

**Computational modeling of function-determining
protein-ligand interactions in adhesion receptors and
ligand-gated ion channels**

Inaugural-Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Michele Bonus

aus Krefeld

Düsseldorf, den 30.05.2021

Aus dem Institut für Pharmazeutische und Medizinische Chemie
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Berichterstattende:

1. Prof. Dr. Holger Gohlke
2. Prof. Dr. Birgit Strodel

Tag der mündlichen Prüfung: 28.01.2022

EIDESSTATTLICHE ERKLÄRUNG

Ich, Michele BONUS, versichere an Eides Statt, dass die Dissertation von mir selbstständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf“ erstellt worden ist.

Diese Dissertation wurde in der vorgelegten oder einer ähnlichen Form noch bei keiner anderen Institution eingereicht, und es wurden bisher keine erfolglosen Promotionsversuche von mir unternommen.

Düsseldorf, im Mai 2021

To my family and friends

“Chemists seek precise answers to well-defined problems, whereas biologists are content with approximate answers to complex problems.”

– Arthur Kornberg

TABLE OF CONTENTS

EIDESSTATTLICHE ERKLÄRUNG	3
TABLE OF CONTENTS	VI
LIST OF PUBLICATIONS	VIII
ABBREVIATIONS	XIV
ZUSAMMENFASSUNG	XVI
ABSTRACT	XVIII
1 INTRODUCTION	1
2 BACKGROUND	4
2.1 Protein-ligand interactions.....	4
2.1.1 Molecular recognition models.....	4
2.1.2 Fundamental quantities to characterize protein-ligand interactions.....	6
2.1.3 Fundamental thermodynamics of protein-ligand interactions.....	8
2.1.4 Enthalpy-entropy compensation.....	11
2.2 Simulation-based characterization of protein-ligand interactions.....	11
2.2.1 MD simulations.....	13
2.2.2 End-point binding free energy calculations.....	15
2.3 Modeling of function-determining protein-ligand interactions.....	20
2.3.1 Functional selectivity of two bile acids for integrin-mediated signaling pathways.....	20
2.3.2 Mechanistically distinct inhibition of NMDA receptor isoforms by tauro-CDC.....	26
2.3.3 Enhanced activation of CNG channels by 8-substituted cyclic nucleotides....	31
2.3.4 Activity determinants of <i>N</i> ⁶ -substituted cAMP analogs on mHCN2 channels...	35
3 SCOPE OF THE THESIS	38
4 PUBLICATION I	40
4.1 Background.....	40
4.2 Results.....	41
4.3 Conclusion and significance.....	48
5 PUBLICATION II	49
5.1 Background.....	49
5.2 Results.....	51
5.3 Conclusion and significance.....	62
6 PUBLICATION III	64
6.1 Background.....	64
6.2 Results.....	65

6.3	Conclusion and significance	72
7	PUBLICATION IV	73
7.1	Background	73
7.2	Results.....	74
7.3	Conclusion and significance	78
8	SUMMARY AND PERSPECTIVE	80
9	ACKNOWLEDGMENT	LXXXII
10	REPRINT PERMISSIONS.....	LXXXIV
11	PUBLICATIONS	LXXXV
	Publication I.....	LXXXV
	Publication II	CXXXIV
	Publication III	CLVII
	Publication IV	CLXXXVIII
12	CURRICULUM VITAE	CXCIX
12.1	Professional and scientific career.....	CXCIX
12.2	Scientific achievements	CCI
13	REFERENCES.....	CCIV

LIST OF PUBLICATIONS

This thesis is based on the following peer-reviewed publications:

Bonus, M., Sommerfeld, A., Qvartskhava, N., Görg, B., Ludwig, B. S., Kessler, H., Gohlke, H., Häussinger, D.

Evidence for functional selectivity in TUDC- and norUDCA-induced signal transduction via $\alpha_5\beta_1$ integrin towards choleresis.

Scientific Reports (2020), 10, 5795.

Impact Factor reported in 2019: 3.998

5-Year Impact Factor reported in 2019: 4.576

Contribution: 35%

Koch, A., **Bonus, M.**, Gohlke, H., Klöcker, N.

Isoform-specific Inhibition of N-methyl-D-aspartate Receptors by Bile Salts.

Scientific Reports (2019), 9, 10068.

Impact Factor reported in 2019: 3.998

5-Year Impact Factor reported in 2019: 4.576

Contribution: 30%

Otte, M., Schweinitz, A., **Bonus, M.**, Enke, U., Schumann, C., Gohlke, H., Benndorf, K.

Hydrophobic alkyl chains substituted to the 8-position of cyclic nucleotides enhance activation of CNG and HCN channels by an intricate enthalpy - entropy compensation.

Scientific Reports (2018), 8, 14960.

Impact Factor reported in 2019: 3.998

5-Year Impact Factor reported in 2019: 4.576

Contribution: 20%

Leybold, T., **Bonus, M.**, Spiegelhalter, F., Schwede, F., Schwabe, T., Gohlke, H., Kusch, J.

N⁶-modified cAMP derivatives that activate protein kinase A also act as full agonists of murine HCN2 channels.

Journal of Biological Chemistry (2019), 294(47), 17978–17987.

Impact Factor reported in 2019: 4.238

5-Year Impact Factor reported in 2019: 4.237

Contribution: 30%

During the work on this thesis, the following publications were additionally authored:

Peer-reviewed articles in scientific journals

Dröge, C., **Bonus, M.**, Baumann, U., Klindt, C., Lainka, E., Kathemann, S., Brinkert, F., Grabhorn, E., Pfister, E.-D., Wenning, D., Fichtner, A., Gotthardt, D. N., Weiss, K. H., McKiernan, P., Puri, R. D., Verma, I. C., Kluge, S., Gohlke, H., Schmitt, L., Kubitz, R., Häussinger, D., Keitel, V.

Sequencing of FIC1, BSEP and MDR3 in a large cohort of patients with cholestasis revealed a high number of different genetic variants.

Journal of Hepatology (2017), 67, 1253–1264.

Impact Factor reported in 2019: 20.582

5-Year Impact Factor reported in 2019: 15.835

Tanabe, K., **Bonus, M.**, Tomiyama, S., Miyoshi, K., Nagi, M., Niimi, K., Chindamporn, A., Gohlke, H., Schmitt, L., Cannon, R. D., Niimi, M., Lamping, E.;

FK506 Resistance of Saccharomyces cerevisiae Pdr5 and Candida albicans Cdr1 Involves Mutations in the Transmembrane Domains and Extracellular Loops.

Antimicrobial Agents and Chemotherapy (2019), 63(1).

Impact Factor reported in 2019: 4.904

5-Year Impact Factor reported in 2019: 4.746

Moussa, M., Ebrahim, W., **Bonus, M.**, Gohlke, H., Mándi, A., Kurtán, T., Hartmann, R., Kalscheuer, R., Lin, W., Liu, Z., Proksch, P.;

Co-culture of the fungus Fusarium tricinctum with Streptomyces lividans induces production of cryptic naphthoquinone dimers.

RSC Advances (2019), 9, 1491–1500.

Impact Factor reported in 2019: 3.119

5-Year Impact Factor reported in 2019: 3.098

Sachs, J., Döhl, K., Weber, A., **Bonus, M.**, Blesse, F., Fleischer, E., Klinger, A., Gohlke, H., Pietruszka, J., Schmitt, L., Teusch, N. E.;

Novel 3,4-Dihydroisocoumarins Inhibit Human P-gp and BCRP in Multidrug Resistant Tumors and Demonstrate Substrate Inhibition of Yeast Pdr5.

Frontiers in Pharmacology (2019), 10, 400.

Impact Factor reported in 2019: 4.225

5-Year Impact Factor reported in 2019: 4.604

Otte, M., Schweinitz, A., Lelle, M., Thon, S., Enke, U., Yüksel, S., Schmauder, R., **Bonus, M.**, Gohlke, H., Benndorf, K.;

Novel Fluorescent Cyclic Nucleotide Derivatives to Study CNG and HCN Channel Function.

Biophysical Journal (2019), 116(12), 2411–2422.

Impact Factor reported in 2019: 3.854

5-Year Impact Factor reported in 2019: 3.845

Preising, M. N., Görg, B., Friedburg, C., Qvartskhava, N., Budde, B. S., **Bonus, M.**, Toliat, M. R., Pflieger, C., Altmüller, J., Herebian, D., Beyer, M., Zöllner, H. J., Wittsack, H.-J., Schaper, J., Klee, D., Zechner, U., Nürnberg, P., Schipper, J., Schnitzler, A., Gohlke, H., Lorenz, B., Häussinger, D., Bolz, H. J.,

Biallelic mutation of human SLC6A6 encoding the taurine transporter TAUT is linked to early retinal degeneration.

The FASEB Journal (2019), 33(10), 11507–11527.

Impact Factor reported in 2019: 4.966

5-Year Impact Factor reported in 2019: 5.394

Form, I, **Bonus, M.**, Gohlke, H., Lin, W., Daletos, G., Proksch, P.

*Xanthone, benzophenone and bianthrone derivatives from the hypersaline lake-derived fungus *Aspergillus wentii*.*

Bioorganic & Medicinal Chemistry (2019), 27, 115005.

Impact Factor reported in 2019: 3.073

5-Year Impact Factor reported in 2019: 2.916

Lelle, M., Otte, M., **Bonus, M.**, Gohlke, H., Benndorf, K.

Fluorophore-labeled cyclic nucleotides as potent agonists of cyclic nucleotide-regulated ion channels.

ChemBioChem (2020), 21(16), 2311–20

Impact Factor reported in 2019: 2.576

5-Year Impact Factor reported in 2019: 2.626

Schirmeyer, J., Hummert, S., Eick, T., Schulz, E., Schwabe, T., Ehrlich, G., Kukaj, T., Wiegand, M., Sattler, C., Schmauder, R., Zimmer, T., Kosmalla, N., Münch, J., **Bonus, M.**, Gohlke, H., Benndorf, K.

Thermodynamic profile of mutual subunit control in a heteromeric receptor.

Proceedings of the National Academy of Sciences of the United States of America (2021), in press.

Impact Factor reported in 2019: 9.412

5-Year Impact Factor reported in 2019: 10.620

Prescher, M., **Bonus, M.**, Stindt, J., Keitel-Anselmino, V., Smits, S. H. J., Gohlke, H., Schmitt, L.

Evidence for a credit-card-swipe mechanism in the human PC-floppase ABCB4.

Structure (2021), in press (aop doi: 10.1016/j.str.2021.05.013).

Impact Factor reported in 2019: 4.862

5-Year Impact Factor reported in 2019: 4.827

Peer-reviewed review articles in scientific journals

Bonus, M., Häussinger, D., Gohlke, H.

Liver cell hydration and integrin signaling.

Biological Chemistry (2021), in press (aop doi: 10.1515/hsz-2021-0193).

Impact Factor reported in 2019: 3.270

5-Year Impact Factor reported in 2019: 3.242

Submitted manuscripts

Tran-Cong, N. M., Hohlfeld, T., **Bonus, M.**, Frank, M., Gohlke, H., Liu, Z., Proksch, P.

The mycotoxin Altenusin is a platelet aggregation inhibitor.

Moe-Lange, J., Gappel, N. M., Machado, M., Wudick, M. M., Sies, C. S. A. Schott-Verdugo, S. N., **Bonus, M.**, Hartwig, T., Bezruczyk, M., Basu, D., Gohlke, H., Malkovskiy, A., Haswell, E. S., Farmer, E. E., Ehrhardt, D. W., Frommer, W. B., Kleist, T. J.

Interdependence of a mechanosensitive anion channel and glutamate receptors in distal wound signaling.

Proceedings and conference contributions

Bonus, M., Sommerfeld, A., Häussinger, D., Gohlke, H.

$\alpha_5\beta_1$ integrins in hepatocytes act as receptors for bile acids with a (nor)ursodeoxycholane scaffold.

European Journal of Medical Research (2014), 19 (Suppl. 1): S13

Bonus, M., Eichler, I., Gohlke, H.

Lipid transport by the ABC transporter MDR3.

InSiDE (2017), 15, 82–86

Yüksel, S., Kondapuram, M., Schwabe, T., **Bonus, M.**, Gohlke, H., Schmauder, R., Kusch, J., Benndorf, K.

Uncoupling the cAMP Binding Domain from the Channel Gate in HCN2 Channels.

Biophysical Journal (2019), 116, 108a

Yüksel, S., Kondapuram, M., Schwabe, T., Zimmer, T., Lelle, M., Schweinitz, A., **Bonus, M.**, Gohlke, H., Schmauder, R., Kusch, J., Benndorf, K.

Uncoupling the cyclic nucleotide-binding domain from the channel gate in HCN2 channels.

Acta Physiologica (2019), 227(719)

ABBREVIATIONS

ABD	<u>A</u> gonist- <u>b</u> inding <u>d</u> omain
ADMIDAS	<u>A</u> djacent to <u>M</u> IDAS
ADP	<u>A</u> nisotropic <u>d</u> isplacement <u>p</u> arameter
AMBER	<u>A</u> ssisted <u>m</u> odel <u>b</u> uilding with <u>e</u> nergy <u>r</u> efinement
AMPA	α - <u>A</u> mino-3-hydroxy-5- <u>m</u> ethyl-4-isoxazolepropionic <u>a</u> cid
ATD	<u>A</u> mino- <u>t</u> erminal <u>d</u> omain
BLAST	<u>B</u> asic <u>L</u> ocal <u>A</u> lignment <u>S</u> earch <u>T</u> ool
cAMP	3',5'- <u>C</u> yclic <u>a</u> denosine <u>m</u> onophosphate
CDC	<u>C</u> henodeoxycholic <u>a</u> cid
cGMP	3',5'- <u>C</u> yclic <u>g</u> uanosine <u>m</u> onophosphate
CHARMM	<u>C</u> hemistry at <u>H</u> arvard <u>M</u> acromolecular <u>M</u> echanics
CL	<u>C</u> - <u>l</u> inker
CLZ	<u>C</u> -terminal <u>l</u> eucine- <u>z</u> ipper
CNBD	<u>C</u> yclic <u>n</u> ucleotide- <u>b</u> inding <u>d</u> omain
CNG	<u>C</u> yclic <u>n</u> ucleotide- <u>g</u> ated
CTD	<u>C</u> -terminal <u>d</u> omain
D-AP5	<u>D</u> -(-)-2- <u>a</u> mino-5- <u>p</u> hospha <u>v</u> aleric <u>a</u> cid
DCKA	5,7- <u>D</u> ichlorokynurenic <u>a</u> cid
ECM	<u>E</u> xtracellular <u>m</u> atrix
EGF	<u>E</u> pidermal growth <u>f</u> actor
EGFR	<u>E</u> GF <u>r</u> eceptor
(cryo-)EM	(<u>c</u> ryo) <u>E</u> lectron <u>m</u> icroscopy
ER	<u>E</u> strogen <u>r</u> eceptor
FAK	<u>F</u> ocal <u>a</u> dhesion <u>k</u> inase
FEP	<u>F</u> ree <u>e</u> nergy <u>p</u> erturbation
FEW	<u>F</u> ree <u>e</u> nergy <u>w</u> orkflow
GABA	<u>G</u> amma- <u>a</u> mino <u>b</u> utyric <u>a</u> cid
GB	<u>G</u> eneralized <u>B</u> orn
GCDC	<u>G</u> lycochenodeoxycholic <u>a</u> cid
GBSA	<u>G</u> eneralized <u>B</u> orn <u>s</u> urface <u>a</u> rea
GPCR	<u>G</u> protein- <u>c</u> oupled <u>r</u> eceptor
GPU	<u>G</u> raphics <u>p</u> rocessing <u>u</u> nit
GROMACS	<u>G</u> roningen <u>M</u> achine for <u>C</u> hemical <u>S</u> imulations
GUDC	<u>G</u> lycoursodeoxycholic <u>a</u> cid
HCN	<u>H</u> yperpolarization-activated <u>c</u> yclic <u>n</u> ucleotide- <u>g</u> ated
ICP	<u>I</u> ntrahepatic <u>c</u> holestasis of <u>p</u> regnancy
LBD	<u>L</u> igand- <u>b</u> inding <u>d</u> omain
LIMBS	<u>L</u> igand-associated <u>m</u> etal- <u>b</u> inding <u>s</u> ite

MAPK	<u>M</u> itogen- <u>a</u> ctivated <u>p</u> rotein <u>k</u> inase
MD	<u>M</u> olecular <u>d</u> ynamics
MIDAS	<u>M</u> etal <u>i</u> on- <u>d</u> eependent <u>a</u> dhesion <u>s</u> ite
MM	<u>M</u> olecular <u>m</u> echanics
NF- κ B	<u>N</u> uclear <u>f</u> actor <u>k</u> appa- <u>l</u> ight- <u>c</u> hain- <u>e</u> nhancer of activated <u>B</u> cells
NMDA	<u>N</u> - <u>m</u> ethyl- <u>D</u> - <u>a</u> spartate
NMDAR	<u>N</u> MDA <u>r</u> eceptor
<i>nor</i> UDCA	24- <u>n</u> or- <u>U</u> DC <u>A</u>
OBC	<u>O</u> nufriev- <u>B</u> ashford- <u>C</u> ase
PBC	<u>P</u> rietary <u>b</u> iliary <u>c</u> holangitis
PBSA	<u>P</u> oisson- <u>B</u> oltzmann <u>s</u> urface <u>a</u> rea
PFIC	<u>P</u> rogressive <u>f</u> amilial <u>i</u> ntrahepatic <u>c</u> holestasis
PH	<u>P</u> ore <u>h</u> elix
PKA	<u>P</u> rotein <u>k</u> inase <u>A</u>
PSC	<u>P</u> rietary <u>s</u> clerosing <u>c</u> holangitis
PSI	<u>P</u> lexin- <u>s</u> emaphorin- <u>i</u> ntegrin
PSI-BLAST	<u>P</u> osition <u>s</u> pecific <u>i</u> terative <u>B</u> LAST
PSIPRED	<u>PSI</u> - <u>B</u> LAST based secondary structure <u>p</u> rediction
RESP	<u>R</u> estrained <u>e</u> lectrostatic <u>p</u> otential
RGD	<u>A</u> rginine- <u>G</u> lycine- <u>A</u> spartic <u>a</u> cid
RMSD	<u>R</u> oot- <u>m</u> ean- <u>s</u> quare <u>d</u> evelopment
SASA	<u>S</u> olvent- <u>a</u> ccessible <u>s</u> urface <u>a</u> rea
SyMBS	<u>S</u> ynergistic <u>m</u> etal ion- <u>b</u> inding <u>s</u> ite
TC	<u>T</u> auro <u>c</u> holic <u>a</u> cid
TI	<u>T</u> hermodynamic <u>i</u> ntegration
TLCAS	<u>T</u> aurolithocholic <u>a</u> cid 3- <u>s</u> ulfate
TM(D)	<u>T</u> ransmembrane (<u>d</u> omain)
<i>Tnor</i> UDCA	Tauro- <u>n</u> orUDCA
TUDC	<u>T</u> auroursodeoxycholic <u>a</u> cid
UDCA	<u>U</u> rsodeoxycholic <u>a</u> cid
XP	<u>E</u> xtra <u>p</u> recision

ZUSAMMENFASSUNG

Proteine empfangen und übersetzen extrazelluläre oder intrazelluläre Signale durch Bindung an kleine Moleküle, Nukleinsäuren oder andere Proteine. Die Langlebigkeit des durch diesen Prozess entstehenden Protein-Ligand-Komplexes kann sowohl die Dauer als auch die Intensität der zellulären Antwort bestimmen und wird maßgeblich durch die Komplementarität der molekularen Interaktionen zwischen den Bindungspartnern bestimmt. Ebendiese Interaktionen bestimmen auch die Art und Weise der Wirkung, die ein Ligand auf ein Protein hat – sie beeinflussen das konformationelle Gleichgewicht des Proteins in bestimmter Weise und legen damit die Rolle des Liganden als Agonist oder Antagonist der Proteinfunktion fest. Oft sind die Interaktionen, welche die Rolle des Liganden determinieren, für ein gegebenes Protein-Ligand-Paar entweder gänzlich unbekannt oder nur teilweise bekannt. Darüber hinaus legen strukturbasierte Überlegungen, wie z.B. die strukturelle Ähnlichkeit mit bekannten Liganden, oft nahe, dass ein Ligand einen Effekt ausüben könnte; ob dieser Effekt die Proteinfunktion jedoch stimuliert oder hemmt, lässt sich selten ableiten. In dieser Arbeit nutze ich eine Reihe computergestützter Methoden, um die Wirkung von Liganden an Adhäsionsrezeptoren vorherzusagen und molekularbiologische, pharmakologische, und elektrophysiologische Daten zur Wirkung von Liganden an ligandgesteuerten Ionenkanälen auf atomistischer Ebene zu interpretieren.

Im speziellen habe ich in Publikation I die Aktivität einer Reihe von Gallensäuren am Adhäsionsrezeptor $\alpha_5\beta_1$ Integrin vorhergesagt und damit einen zweiten Agonisten von $\alpha_5\beta_1$ Integrin identifiziert, der vornehmlich in der Leber wirksam ist. Durch die Interpretation experimenteller Daten, die auf Grundlage meiner Vorhersage erhoben wurden, stellte ich fest, dass dieser Agonist eine funktionelle Selektivität für einen integrinabhängigen Signaltransduktionsweg zeigt, der sich jedoch vom Signaltransduktionsweg unterscheidet, der durch den zuvor identifizierten Agonisten angeschaltet wird. Funktionelle Selektivität für Integrine ohne αI -Domäne wird in dieser Arbeit zum ersten Mal beschrieben.

Weiterhin habe ich im Rahmen von Publikation II eine strukturelle Interpretation für die Beobachtung erarbeitet, dass die Gallensäure tauro-CDC NMDA-Rezeptoren mit GluN2D-Untereinheiten allosterisch, solche mit GluN3B-Untereinheiten jedoch kompetitiv hemmt. Im weiteren Verlauf dieser Studie konnte ich die mögliche allosterische Bindestelle für tauro-CDC im GluN1/GluN2D-Interface identifizieren, deren Lokalisation in exzellentem

Einklang mit strukturellen Daten zu allosteren Bindestellen in GluN2A-tragenden NMDA-Rezeptoren stand.

In Publikation III und Publikation IV habe ich die Struktur-Affinitäts-Beziehungen einer Reihe von Derivaten zyklischer Nukleotide an ligandgesteuerten Ionenkanälen untersucht, deren Aktivität durch cAMP und cGMP reguliert wird. In Publikation III konnte ich durch eine detaillierte Charakterisierung der enthalpischen und entropischen Beiträge zur freien Bindungsenergie eine komplexe Enthalpie-Entropie-Kompensation innerhalb der untersuchten Reihe von Liganden herausarbeiten und unter anderem erklären, warum Derivate mit längeren Alkylketten eine höhere scheinbare Affinität zu den untersuchten CNG-Kanälen zeigen. Diese Arbeit diente als Grundlage für die Synthese und Charakterisierung einer Reihe fluoreszierender Liganden, die zur Untersuchung der Kanalfunktion eingesetzt werden können. Zuletzt lieferte ich in Publikation IV eine strukturbasierte Erklärung für die gemessenen scheinbaren Affinitäten einer weiteren Reihe von Derivaten zyklischer Nucleotide an HCN2-Kanälen. In dieser Arbeit stellten wir fest, dass die eigentlich als PKA-selektiv geltenden Verbindungen auch HCN2-Kanäle aktivieren, und meine strukturbasierten Analysen könnten einen ersten Ansatzpunkt zur Umgehung dieser dualen Aktivierung und für die Entwicklung hochaffiner HCN2-Liganden liefern.

ABSTRACT

Proteins receive and translate extracellular or intracellular signals by binding to small molecules, nucleic acids, or other proteins. The longevity of the protein-ligand complex formed through this process can determine both the duration and the intensity of the cellular response and is largely determined by the complementarity of the molecular interactions between the binding partners. These interactions also determine the way in which a ligand acts on a protein – they influence the conformational equilibrium of the protein in a specific way and thus determine the role of the ligand as an agonist or antagonist of protein function. In many cases, the interactions that determine the role of the ligand are either completely unknown or only partially known for a given protein-ligand pair. Furthermore, structure-based considerations, such as the structural similarity to known ligands, often suggest that a ligand may exert an effect, but whether this effect stimulates or inhibits protein function can rarely be inferred. In this thesis, I employ a range of computational methods to predict the effect of ligands at adhesion receptors and to interpret molecular biological, pharmacological, and electrophysiological data on the effect of ligands at ligand-gated ion channels at the atomistic level.

More specifically, in Publication I, I predicted the activity of a series of bile acids at the adhesion receptor $\alpha_5\beta_1$ integrin and thereby identified a second agonist of $\alpha_5\beta_1$ that predominantly acts in the liver. By interpreting experimental data that was acquired in response to my prediction, I determined that this agonist exhibits a functional selectivity for an integrin-dependent signal transduction pathway different from the pathway triggered by the earlier identified agonist. In this study, functional selectivity is described for the first time for integrins without an αI domain.

Furthermore, in Publication II, I established a structural interpretation for the observation that the bile acid tauro-CDC inhibits NMDA receptors with GluN2D subunits allosterically, but those with GluN3B subunits competitively. Later in this study, I identified the potential allosteric binding site for tauro-CDC in the GluN1/GluN2D interface, which was in excellent agreement with structural data on the location of an allosteric binding site in GluN2A-carrying NMDA receptors.

In Publication III and Publication IV, I studied the structure-apparent affinity relationships of a number of cyclic nucleotide derivatives on cyclic nucleotide-modulated ion channels. In Publication III, I elaborated a complex enthalpy-entropy compensation within the ligand

series by an in-depth characterization of the enthalpic and entropic contributions to the calculated binding free energies. Additionally, I explained why the apparent affinities of the compounds towards the investigated CNG channels increased with the length of the alkyl chain in the attached substituent. This work served as a basis for the synthesis and characterization of a series of fluorescent ligands suitable to study channel function. Lastly, in Publication IV, I provided a structure-based explanation for the experimentally determined apparent affinities of a different set of cyclic nucleotide derivatives towards HCN2 channels. In this work, we determined that the studied compounds, which were considered PKA-selective, also activate HCN2 channels, and my structure-guided analyses might provide a starting point to bypass this dual activation and to develop high-affinity HCN2 ligands.

1 INTRODUCTION

Proteins are the workers, messengers, and guardians of the cell, and hardly any biochemical, let alone physiological process is conceivable without their intervention¹. Enzymes can lower the activation energy for chemical reactions to such an extent that processes that would take almost 80 million years without enzymatic involvement can be carried out in milliseconds^{2,3}. Specialized proteins recognize specific chemical or physical signals and translate these signals into a language whose instructions are understood and passed on by other proteins to ultimately evoke a sufficiently strong cellular response⁴. Obvious errors in such instructions are recognized and corrected by other proteins⁵. However, regardless of the role which a particular protein might play, its ability to communicate only becomes possible through the binding of small molecules, nucleic acids, or other proteins⁵.

Despite the apparent quantitative abundance of proteins in a cell or even in a cellular cluster, only small changes in the absolute or relative populations of certain protein complexes or in the concentration of the products of enzymatic reactions are necessary to substantially alter the fate of the cell. Halving or doubling a certain protein subpopulation may direct a cell from growth to stasis or from replication to apoptosis³. Such population changes are typically the result of only very small changes in the thermodynamics or kinetics of molecular processes^{3,6}. Therefore, a detailed description of the interactions of proteins with their binding partners is of utmost importance for the understanding of biological systems at the molecular level. Furthermore, directly influencing a certain protein-ligand interaction is the prevailing strategy in the development of a pharmacotherapeutic treatment approach⁶. Discovering small-molecular ligands, characterizing their interactions with the protein of interest, and exploiting this information for the development of a lead compound into a drug is an essential element of the drug development pipeline⁷.

Protein-ligand interactions may be described on a qualitative and quantitative level³. On a qualitative level, the first natural and straightforward step would be to describe the types of interactions that stabilize a particular complex. However, this procedure alone is not sufficient for an interpretation regarding the consequence of this interaction, its strength, and its longevity. Without a thorough understanding of whether and how certain interactions influence the conformational equilibrium of the target protein, it is impossible to infer *ad hoc* whether a ligand activates the protein or acts as an inhibitor or antagonist⁸⁻¹⁰. Accordingly, predicting such function-determining interactions is challenging and becomes

increasingly difficult the more the chemotype of the ligand of interest diverges from the chemotype of ligands that have already been characterized. While most experimental approaches provide a relatively quick answer to the question of whether a ligand activates or inhibits protein function, they often cannot provide any information about the molecular determinants that enable the ligand to attain this property. Using computational techniques such as molecular dynamics simulations, protein-ligand interactions can be monitored over a limited time frame—if the simulation is long enough, conformational changes of the protein in response to ligand binding are also accessible¹¹. In this way, a qualitative or semi-quantitative prediction becomes possible as to whether and to what extent a certain ligand activates or inhibits a protein. I made extensive use of this technique in Publication I to successfully predict the activity of a series of bile acids and bile acid derivatives on the adhesion receptor $\alpha_5\beta_1$ integrin. Here, I used structural information from crystal structures of homologous proteins to determine a set of geometric parameters that allowed me to track $\alpha_5\beta_1$ activation during the course of molecular dynamics simulations. These simulations suggested different degrees of integrin activation, depending on the ligand bound; a finding of particular value for the design and interpretation of the subsequent *in vitro* experiments, from which I could infer a functional selectivity of two ligands for overlapping, but distinct signal transduction pathways.

Molecular docking, which is less computationally demanding than a molecular dynamics simulation and often used to predict ligand binding modes, is particularly suitable in cases where structure-activity or structure-function relationships are substantially influenced by energetic contributions of only a few amino acids. In Publication IV, I used molecular docking to identify the determining factors for the apparent affinities of a series of N^6 -substituted cAMP analogs towards the hyperpolarization-activated cyclic nucleotide-gated channel HCN2 and explained why certain substitutions increase or decrease the apparent affinity by one order of magnitude compared to the native ligand. My conclusions from this semi-quantitative approach were validated in subsequent mutagenesis experiments.

Thermodynamic and kinetic quantities, such as the equilibrium association and dissociation constant (K_a and K_d , respectively), the association and dissociation rate constant (k_{on} and k_{off} , respectively), and the binding free energy (ΔG_{bind}), are used to describe protein-ligand interactions on a quantitative level. I provide a detailed description of these quantities in sections 2.1.2 and 2.1.3. Especially ΔG_{bind} is an indispensable measure for estimating the stability of a given protein-ligand complex but is laborious to measure and equally difficult

to predict by computational means. Apart from a number of precise but computationally expensive approaches for the prediction of binding free energies (cf. section 2.2), end-point methods (cf. 2.2.2) offer an attractive compromise between precision and computational costs. I extensively and successfully employed these methods in Publication II and Publication III to provide a structure-based interpretation of data from molecular biological, electrophysical, and pharmacological experiments. In more detail, in Publication II, I used end-point free energy calculations to validate the pharmacologically different mode of action of the bile acid taurochenodeoxycholic acid (tauro-CDC) on *N*-methyl-D-aspartate (NMDA) receptors with different subunit composition. During the course of this study, I also identified an allosteric binding site for tauro-CDC in GluN2D-carrying NMDA receptors. Furthermore, in Publication III, I used end-point approaches to explain the observed structure-apparent affinity relationship within a series of cAMP analogs on a structural and energetic level. These analogs bind to and activate cyclic nucleotide-gated channels with higher apparent affinities than their native ligand cGMP and display an intricate case of enthalpy-entropy compensation within a congeneric series of agonists. The findings from this work led to the synthesis and characterization of a series of fluorescent probes suitable to study the coupling between ligand binding and channel gating in cyclic nucleotide-gated channels¹².

The studies on which this thesis is based show a wide range of applications for computational methods in characterizing the effect of a ligand on its target protein on a qualitative, semi-quantitative, and quantitative level.

2 BACKGROUND

2.1 Protein-ligand interactions

Proteins fulfill a wide range of physiological functions. They reduce the activation energy for vital chemical reactions, transport small molecules and ions across the plasma membrane, and mechanically stabilize the cell³. Proteins further translate the binding of small molecules into a conformational signal and ultimately orchestrate cellular signal transduction by inhibiting or activating other proteins. This very binding of small molecules or other proteins is, therefore, not only a basis for cellular life but also a mainstay of pharmacological interventions^{3,6}. Finding a selective, potent, and efficacious ligand that activates or inhibits a target protein is one of the critical milestones in the drug development pipeline. However, to fully exploit the potential for interactions offered by the biomolecular target, a profound understanding of the mechanisms underlying protein-ligand interactions is necessary. In this section, I provide a general overview of the most popular models that describe the molecular recognition of a ligand by a protein. I introduce fundamental quantities for the description of protein-ligand interactions and later focus on the binding free energy and its dissection into enthalpic and entropic contributions. Finally, I highlight the phenomenon of enthalpy-entropy compensation, which is an essential point of discussion for two of the studies on which this thesis is based (Publication II and Publication III).

2.1.1 Molecular recognition models

Although binding of a ligand to a protein is conceptually simple to describe, it involves an interplay of many overlapping individual processes. It is therefore not surprising that no less than three major models exist that aim to describe individual aspects of molecular recognition. In short, the “lock-and-key” model¹³⁻¹⁵ (**Figure 1A**) was the first model to propose a necessity for shape complementarity between an enzyme and its substrates. The rigidity of protein and ligand implied in this model, however, was not tenable due to numerous anomalies when considering pairs of substrates and non-substrates with clearly complementary shapes¹⁶. The “induced fit” model (**Figure 1B**) accounted for this discrepancy by assuming a flexible catalytic site that responds to substrate binding with a conformational change that brings the catalytic residues in the precise orientation required for catalysis¹⁶. The binding of non-substrates, in turn, would not result in such a conformational response. Both models, however, disregard the natural conformational dynamics of the protein⁶. According to the “conformational selection” model (**Figure 1C**), a suitable ligand selectively binds to the most suitable protein conformational state (or, more

specifically, one of the microstates of the conformational macrostate) and shifts the equilibrium towards this state¹⁷⁻²⁰. The model thereby attractively establishes a direct link between a protein-ligand interaction and an observable decrease or increase in protein function: Activation of a receptor may, for example, be viewed as an increase in the subpopulation of active receptor conformations.

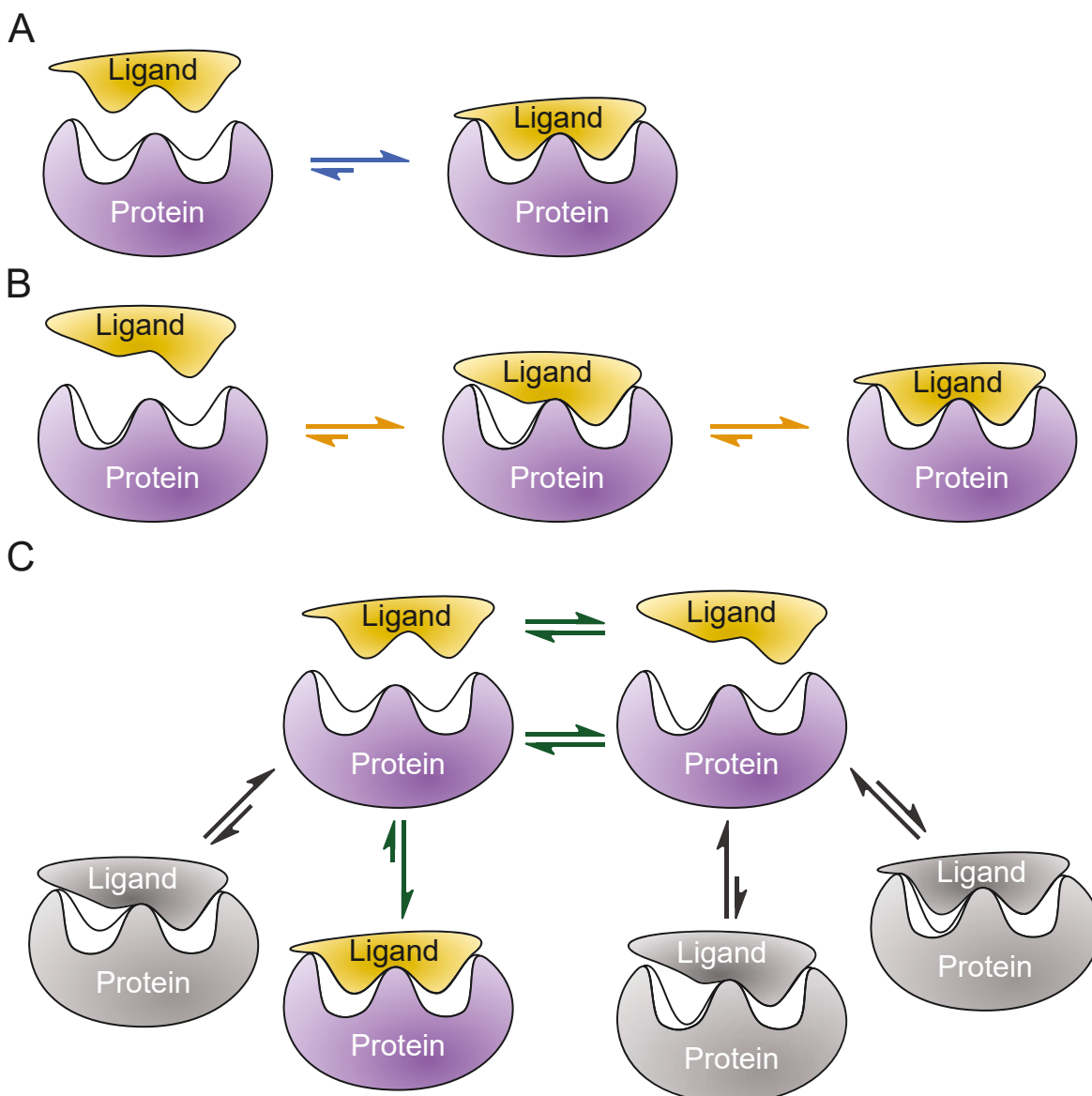


Figure 1: Molecular recognition models.

(A) Illustration of the “lock-and-key” model. Protein and ligand are both assumed to be rigid. Complex formation requires near perfect shape complementarity of both entities. (B) Illustration of the “induced fit” model. After formation of an initial complex with partial shape complementarity (center), local conformational changes lead to shape matching of protein and ligand towards an ideal fit. (C) Illustration of the “conformational selection” model. Within the conformational equilibria of protein and ligand, several matching pairs may exist, only one of which (highlighted in color) is preferentially formed. An interconversion within the possible microstates of the complex in the sense of an induced fit is likely, but not depicted.

The chronological order in which these molecular recognition models were proposed may indicate that the newer models are corrections of the previous model. Indeed, several authors

have already degraded the “lock-and-key” model to a mere historical hypothesis and consider it superseded by its successors^{21,22}. It may be, however, more appropriate to consider them as extensions since experimental evidence for the existence of all three models is available^{6,23}. In many cases, the actual protein-ligand binding process may best be described by an overlap of the mechanisms underlying these models²⁴. This overlap is already evident when considering the “lock-and-key” model as an extreme case of conformational selection²⁵. Assuming that the conformational energy landscape of a very rigid protein can be described with a single, deep energy well, a binding process that takes place in this well may be described by the “lock-and-key” mechanism. Furthermore, the “induced fit” and “conformational selection” models complement each other, as larger backbone motions are better accounted for by the latter, while smaller adjustments in side-chain rotamers are better represented by the former model²⁵⁻²⁷.

As evident from the foregoing, binding processes between proteins and ligands are highly dynamic and complex. Accordingly, computational prediction of protein-ligand interactions is a daunting task, and its success strongly depends on the combination of suitable methods, which must always be chosen in the context of the available experimental evidence on the binding process. The successful computational predictions in two of the studies on which this thesis is based (Publication II) would not have been possible without accounting for an “induced fit” mechanism.

2.1.2 Fundamental quantities to characterize protein-ligand interactions

A bimolecular association-dissociation equilibrium between a protein (P)-ligand (L) pair and the corresponding protein-ligand complex (PL) can be considered as a simple, reversible process³, as illustrated in **Equation 1**.



The strength of the interaction between protein and ligand can be expressed as the affinity of the single species P and L by the equilibrium association constant K_a [l mol^{-1}] (**Equation 2**)^{24,28,29}.

$$K_a = \frac{[PL]}{[P][L]} \quad \text{Equation 2}$$

Here, [P] and [L] refer to the equilibrium concentrations of protein and ligand, and [PL] refers to the equilibrium concentration of the protein-ligand complex (all concentrations

given in mol l⁻¹). In a pharmaceutical context, the reciprocal of K_a , the equilibrium dissociation constant K_d [mol l⁻¹] (**Equation 3**), is encountered more frequently. It corresponds to the ligand concentration at which the binding partner is found bound and unbound with statistically equal probability³⁰ and therefore carries a more intuitive meaning in an experimental context.

$$K_d = \frac{[P][L]}{[PL]} = \frac{1}{K_a} \quad \text{Equation 3}$$

The standard free energy change ΔG° upon complex formation is given by **Equation 4** and **Equation 5** and commonly referred to as binding free energy $\Delta G_{\text{bind}}^\circ$ ^{3,6,31}.

$$\Delta G^\circ = -RT \ln K_a = \Delta G_{\text{bind}}^\circ \quad \text{Equation 4}$$

$$\Delta G^\circ = RT \ln K_d = \Delta G_{\text{bind}}^\circ \quad \text{Equation 5}$$

Here, R is the gas constant (8.3145 J mol⁻¹ K⁻¹), and T is the absolute temperature. The binding free energy, like RT , is expressed in units of energy (e.g., kJ mol⁻¹ or kcal mol⁻¹). Accordingly, both K_a and K_d must be dimensionless but, according to **Equation 3**, are expressed in units of (reciprocal) concentration. This disparity results from the common omission of standard state concentrations in the definition of the equilibrium constants^{3,6,21}. Without this omission, the expression for K_d expands to:

$$K_d = \frac{\frac{[P]}{[P]^\circ} \frac{[L]}{[L]^\circ}}{\frac{[PL]}{[PL]^\circ}} \quad \text{Equation 6}$$

In **Equation 6**, $[P]^\circ$, $[L]^\circ$, and $[PL]^\circ$ are standard state concentrations, i.e., equal to 1 mol l⁻¹. To enable a distinction between both definitions of K_d , the equilibrium dissociation constant specified in concentration units is occasionally referred to as pseudo equilibrium dissociation constant K_d^* , the dimensionless equilibrium dissociation constant as K_d . Although both quantities are numerically identical, this distinction allows for the correct treatment of units. Nevertheless, it is still customary to use K_d^* and K_d interchangeably, and to refer to K_d in units of concentration. Hence, throughout this thesis, I will follow the established practice and imply the correct definition of K_d depending on the context.

While the strength of a protein-ligand interaction may be inferred from K_a , K_d , or ΔG_{bind}^0 , the rate at which the two species associate is given by the second-order association rate constant k_{on} [$\text{l mol}^{-1} \text{s}^{-1}$]⁶. The magnitude of k_{on} depends on the number of productive collision events, i.e., those collisions that result in the formation of a complex³. For a given protein-ligand pair, the number of these events is either directly governed by complementary long-range electrostatic interactions or by strategical arrangements of functional groups on the protein surface, which progressively steer the ligand towards the binding pocket. The dissociation rate constant k_{off} [s^{-1}], on the other hand, is concentration-independent and proportional to the dissociation probability of a single complex. Its reciprocal value, the average complex lifetime $\tau = 1/k_{\text{off}}$ [s], is an intuitive quantity to describe the stability of the protein-ligand complex³.

Association and dissociation rates (v_{on} and v_{off} , respectively) equal the product of the rate constants and the concentrations of the species involved in the reaction (**Equation 7** and **Equation 8**)^{3,30}.

$$v_{\text{on}} = k_{\text{on}} [\text{P}][\text{L}] \quad \text{Equation 7}$$

$$v_{\text{off}} = k_{\text{off}} [\text{PL}] \quad \text{Equation 8}$$

In equilibrium, the rate of association and dissociation are, by definition, the same (**Equation 9**).

$$k_{\text{on}} [\text{P}][\text{L}] = k_{\text{off}} [\text{PL}] \quad \text{Equation 9}$$

Rearranging **Equation 9** recovers the definition of K_a given in **Equation 2**. Therefore:

$$K_a = \frac{k_{\text{on}}}{k_{\text{off}}} = \frac{[\text{PL}]}{[\text{P}][\text{L}]} \quad \text{Equation 10}$$

and

$$K_d = \frac{k_{\text{off}}}{k_{\text{on}}} = \frac{[\text{P}][\text{L}]}{[\text{PL}]} \quad \text{Equation 11}$$

2.1.3 Fundamental thermodynamics of protein-ligand interactions

By a well-established definition, ΔG_{bind}^0 corresponds to the difference between the Gibbs free energy G^0 of the protein-ligand complex and the sum of the free energies of the individual species in solution (**Equation 12**)^{30,32}.

$$\Delta G_{\text{bind}}^{\circ} = G_{\text{solution}}^{\circ, \text{PL}} - \left(G_{\text{solution}}^{\circ, \text{P}} + G_{\text{solution}}^{\circ, \text{L}} \right) \quad \text{Equation 12}$$

The Gibbs free energy itself is composed of an enthalpic (H) and an entropic (S) component.

$$G^{\circ} = H^{\circ} - TS^{\circ} \quad \text{Equation 13}$$

Therefore, one may expand **Equation 12** as follows,

$$\Delta G_{\text{bind}}^{\circ} = H_{\text{solution}}^{\circ, \text{PL}} - TS_{\text{solution}}^{\circ, \text{PL}} - \left(H_{\text{solution}}^{\circ, \text{P}} - TS_{\text{solution}}^{\circ, \text{P}} + H_{\text{solution}}^{\circ, \text{L}} - TS_{\text{solution}}^{\circ, \text{L}} \right) \quad \text{Equation 14}$$

and it becomes evident that the binding free energy of a protein-ligand complex may directly be determined from the changes in enthalpy and entropy of all species participating in the reaction. This equation is equivalent to the typical expression for the change in Gibbs free energy (**Equation 15**)⁶, yet more convenient as a foundation for the discussion of computational methods that calculate enthalpic and entropic components from a set of additional individual contributions (cf. section 2.2.2).

$$\Delta G_{\text{bind}}^{\circ} = \Delta H_{\text{bind}}^{\circ} - T\Delta S_{\text{bind}}^{\circ} \quad \text{Equation 15}$$

From a thermodynamic point of view, the enthalpic component $\Delta H_{\text{bind}}^{\circ}$ in **Equation 15** is a measure for the change in heat associated with protein-ligand binding, and the entropic component $-T\Delta S_{\text{bind}}^{\circ}$ is a measure for the change in the disorder of the overall system³³. When discussing binding processes, $\Delta H_{\text{bind}}^{\circ}$ is often perceived as the sole result of energy changes due to the formation of noncovalent interactions between protein and ligand. In strict terms, however, the heat change during the binding process reflects changes in the entire system under consideration and thus also includes solvent effects^{6,31}. Especially in an aqueous medium, it is hardly imaginable that the formation of new, noncovalent interactions occurs without breaking at least a single preexisting interaction³⁴. Indeed, all processes associated with desolvation and solvation of any species involved in complex formation ultimately manifest in the enthalpy change^{6,34}. The change in entropy during binding $\Delta S_{\text{bind}}^{\circ}$ can be broken down into a configurational part, which treats only the changes in the degrees of freedom in protein and ligand, and a solvent part³⁵ (**Equation 16**):

$$\Delta S_{\text{bind}}^{\circ} = \Delta S_{\text{config}}^{\circ} + \Delta S_{\text{solv}}^{\circ} \quad \text{Equation 16}$$

In particular, the burial of protein and ligand surfaces occupied by water molecules prior to complex formation is reflected in $\Delta S_{\text{solv}}^{\circ}$. Due to the release of these water molecules into the

bulk during complex formation, ΔS_{solv} generally becomes highly positive and thus contributes significantly to the overall entropy difference—in most cases, this term dominates and renders the overall entropy change during the binding reaction favorable. Notable exceptions are binding reactions that are accompanied by a pronounced structuring of disordered protein regions³⁶.

The configurational entropy S_{config} can, in turn, be broken down into a conformational entropy S_{conf} and a vibrational entropy S_{vib}° ^{35,37} (**Equation 17**). While the conformational entropy reflects the number of energy wells occupied by the subsystem, the vibrational entropy reflects their average width³⁸. Hence, S_{vib}° depends on the standard concentration since the margins of the three translational dimensions in each energy well are dictated by the standard volume (i.e., the reciprocal standard concentration).

$$S_{\text{config}}^{\circ} = S_{\text{conf}} + S_{\text{vib}}^{\circ} \quad \text{Equation 17}$$

In a system with M equally stable energy wells, S_{conf} amounts to $-R \ln M$ ³⁵; when the energy wells are not equally stable, S_{conf} can be calculated from

$$S_{\text{conf}} = -R \sum_i^M p_i \ln p_i \quad \text{Equation 18}$$

where p_i is the occupation probability of energy well i . As the difference between conformational and configurational entropy corresponds to the Boltzmann-averaged entropy of the individual energy wells, S_{vib}° can be expressed according to **Equation 19**.

$$S_{\text{vib}}^{\circ} = \langle S_i^{\circ} \rangle = -R \sum_i^M p_i S_i^{\circ} = -R \sum_i^M p_i \left(\frac{\langle E \rangle_i - \mu_i^{\circ}}{T} \right) \quad \text{Equation 19}$$

Here $\langle E \rangle_i$ is the average energy in well i and μ_i° is its standard chemical potential³⁵.

If desired, configurational entropy may also be determined via decomposition into different classes of degrees of freedom³⁵. In this case, however, a simple summation of the component terms would not reproduce the total configurational entropy since inter-term correlations would not be considered. This problem has been addressed in various approaches to the formulation of the molecular partition function, such as the rigid rotor, harmonic oscillator (RRHO) approximation^{39,40} and the flexible molecule (FM) approach^{40,41}, the latter of which has been computationally implemented in the “Binding Entropy Estimation for Rotation and Translation” (BEERT) method⁴¹. Several other approaches to estimate the entropic

contribution to protein-ligand binding have been developed, some of which are briefly outlined in section 2.2.2.

2.1.4 Enthalpy-entropy compensation

During lead structure optimization in drug development, the range of values for $\Delta G_{\text{bind}}^{\circ}$ of newly designed compounds is often substantially narrower than the related changes in enthalpy and entropy would suggest. This discrepancy originates from an inherent correlation between $\Delta H_{\text{bind}}^{\circ}$ and $T\Delta S_{\text{bind}}^{\circ}$ and is popularly termed enthalpy-entropy compensation. First described by Leffler⁴² in 1955, the term was popularized in 1970 by Lumry and Rajender⁴³ and described as a principal phenomenon of processes of small solutes in aqueous solution. More specifically related to protein-ligand binding, enthalpy-entropy compensation refers to the situation where a ligand modification leads to a change in the enthalpic contribution $\Delta\Delta H_{\text{bind}}^{\circ} \equiv \Delta H_{\text{bind},2}^{\circ} - \Delta H_{\text{bind},1}^{\circ}$, which is either partially or entirely canceled out by a change in the entropic contribution $T\Delta\Delta S_{\text{bind}}^{\circ} \equiv T\Delta S_{\text{bind},2}^{\circ} - T\Delta S_{\text{bind},1}^{\circ}$ ³³. For a while, this intriguing phenomenon was considered a thermodynamic stalemate for the optimization of lead compounds. A large-scale analysis of data obtained by isothermal titration calorimetry concluded that the complete cancelation of favorable enthalpic contributions by unfavorable entropic contributions is the rule rather than the exception^{33,44}. However, despite an undeniable ubiquity of the phenomenon, its strong theoretical foundation, and the considerable magnitude of the effect, enthalpy-entropy compensation is not always completely compensatory⁴⁵. Instead, complete compensation observed for a series of ligands is often an artifact brought about by erroneously conducted experiments and incorrect analyses^{33,46}. Nevertheless, such compensation has been observed even after minor changes to a system⁴⁷⁻⁴⁹. Therefore, their consideration is also of particular importance in a computational context when interpreting the effect of such small changes. In two of the studies on which this thesis is based (Publication II and Publication III), accounting for enthalpy-entropy compensation was essential for the interpretation of the acquired data.

2.2 Simulation-based characterization of protein-ligand interactions

Biomolecular simulations are an indispensable tool for the detailed characterization of protein-ligand interactions and are of paramount value in drug design and drug development^{21,50,51}. The underlying methods, whose practical use has earned them the epithet “computational microscope”^{52,53}, provide detailed insights into biological processes on an atomistic scale. These insights may complement existing data and drive new experiments⁵⁴, whose conclusions may be validated by a new set of computations, thus creating an iterative

cycle of mutual complementation between experiment and simulation⁵⁵⁻⁵⁷. In view of the studies on which this thesis is based and the methods and techniques used therein, I will limit the discussion in the following sections to molecular dynamics (MD) simulations, although Monte Carlo simulations can be considered equally valuable.

MD simulations provide both qualitative and quantitative insights on protein-ligand interactions. Qualitatively, they facilitate the identification of key interactions between binding partners and the prediction of ligand-induced conformational changes. These conformational changes can be restricted to the binding pocket in the form of an “induced fit” mechanism but may – in the case of long and large-scale MD simulations – also include movements of entire domains (cf. Publication I). Therefore, short MD simulations may be used to refine ligand poses generated by other methods^{54,58}, while, at the other end of the spectrum, very long and exhaustive MD simulations can capture the complete binding process of a ligand to a protein⁵⁹⁻⁶¹. To reduce the computational resources required for simulating a complete binding process, MD simulations may be coupled with enhanced sampling techniques^{62,63}. However, depending on the method used to enhance sampling, the resulting simulation ensemble may have to be laboriously reweighted to be able to derive kinetic or thermodynamic quantities from it. Regarding their predictive power on a quantitative level, MD simulations are currently the method of choice for predicting binding affinities. In terms of accuracy, MD simulations are vastly superior to other computational techniques such as molecular docking^{54,64}. Although it is currently not possible to simulate binding and unbinding events in sufficient numbers to calculate rate constants or binding affinities directly from unguided MD simulations, MD-based “alchemical” approaches can be considered the gold standard for calculating relative binding free energies between a series of ligands and a protein^{64,65}. In these methods, a ligand is gradually transformed into another ligand through a series of simulations⁶⁶. Two of the most prominent approaches of this type are the free energy perturbation (FEP)^{67,68} and the thermodynamic integration (TI)⁶⁹ approaches. However, despite recent advances in computing power, including GPU support⁷⁰, these techniques are still computationally demanding and time-consuming. Moreover, they are only practical if the ligands compared to each other bear a similar chemical scaffold^{71,72}. An alternative to these methods are so-called end-point methods, such as the Molecular Mechanics/Generalized Born Surface Area (MM-GBSA) and Molecular Mechanics/Poisson–Boltzmann Surface Area (MM-PBSA) approaches^{32,71,73-75}. They use a number of approximations such as continuum solvent models and are therefore considered

less accurate, but are also considerably faster and can therefore be routinely performed in a drug design context³². In contrast to the FEP or TI approaches, however, their predictivity is strongly influenced by the system under investigation and an appropriate choice of parameters^{32,76}, but can also be systematically improved by a variety of modifications, such as the inclusion of an explicit hydration shell or an entropy estimate⁷⁷. Overall, MD simulations coupled with suitable methods for estimating binding free energies are an indispensable tool for characterizing protein-ligand interactions on an atomistic level.

In the following subsections, I briefly describe the theoretical foundations of MD simulations and end-point methods, with special emphasis on the latter in the context of the content of this thesis.

2.2.1 MD simulations

The development of MD simulations dates back to the 1970s⁷⁸ and arose from the desire to circumvent the laborious procedures necessary to study the function of proteins as time-dependent processes in atomic detail¹¹. Due to the inherent complexity of many-body systems like molecular complexes, an exact representation of molecular motion based on quantum mechanical principles was, and is still, impossible. Through simple approximations based on Newtonian physics, however, the computational complexity could be reduced to such an extent that simulations of small proteins and protein-ligand complexes became practicable. Today, with the rapid development of GPU-accelerated computing⁷⁹⁻⁸¹ and better algorithms to support such calculations, even large multi-protein complexes may be simulated within a reasonable time frame. In the course of this development, the great value of MD simulations for a wide range of applications within several disciplines, such as structural bioinformatics, molecular biology, and computational drug discovery, gradually became apparent⁵¹.

The fundamental idea behind MD simulations – the time-dependent monitoring of molecular processes – requires a description of particle motion. Such a description is obtained by solving* the second-order differential equations according to Newton's second law (**Equation 20**).

$$\mathbf{f}_i(t) = m_i \mathbf{a}_i(t) = \frac{\partial V(\mathbf{x}(t))}{\partial \mathbf{x}_i(t)} \quad \text{Equation 20}$$

* An analytical solution is only possible for very small systems; in practice, numerical methods are employed. For details, cf. ref. ¹³.

Here, $\mathbf{f}_i(t)$ is the net force that acts on the i^{th} atom of the system at time t ; this force is given by the mass of atom i , m_i , and its acceleration $\mathbf{a}_i(t)$ at time t . The vector $\mathbf{x}(t)$ contains the locations of all N interacting atoms in cartesian space ($\mathbf{x} = [x_1, y_1, z_1, x_2, y_2, z_2, \dots, x_N, y_N, z_N]$)²¹. This classical mechanical description can only apply to atomic nuclei, requiring the electron motions to be regarded as averaged out. In practice, the potential energy $V(\mathbf{x}(t))$ is calculated from a molecular mechanical model – the “force field” – which may take the form⁸² of

$$V = V_{\text{bond}} + V_{\text{angle}} + V_{\text{dihedral}} + V_{\text{nonbonded}}, \quad \text{Equation 21}$$

where

$$V_{\text{bond}} = \sum_{\text{bonds}} K_r (r - r_{\text{eq}})^2 \quad \text{Equation 22}$$

$$V_{\text{angle}} = \sum_{\text{angles}} K_{\theta} (\theta - \theta_{\text{eq}})^2 \quad \text{Equation 23}$$

$$V_{\text{dihedral}} = \sum_{\text{dihedrals}} \frac{V_n}{2} [1 + \cos(n\phi - \gamma)] \quad \text{Equation 24}$$

$$V_{\text{nonbonded}} = \sum_{i < j} \frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} + \frac{q_i q_j}{\epsilon R_{ij}} \quad \text{Equation 25}$$

Here, the first three terms V_{bond} , V_{angle} , and V_{dihedral} , describe the variation in potential energy as a function of bond stretching and compression, bond angle deviations, and changes in torsion angles, respectively. The fourth term $V_{\text{nonbonded}}$ is a composite of a Lennard-Jones (12-6) potential and a Coulomb potential, describing van der Waals and electrostatic interactions (for further details on the reasoning behind the chosen functional form for each term, cf. ref. ⁸² or ²¹). To emphasize this distinction, it is often given as two separate terms V_{vdw} and V_{elec} . Force fields of this type are implemented in a number of software packages for biomolecular simulations such as AMBER^{83,84}, CHARMM⁸⁵, and GROMACS⁸⁶. The current AMBER version alone (AMBER 2020⁸⁴) contains more than five routinely used amino acid force fields^{82,87-92}, which, together with force fields for nucleic acids^{93,94}, comprise the core of the force field repertoire for biomolecules in this software collection. Further independent but compatible force fields with parameters for carbohydrates⁹⁵, lipids⁹⁶⁻⁹⁸, solvents, ions, and small organic compounds^{99,100} complement these force fields and thus allow MD simulations of highly complex systems.

Although MD simulations provide impressive insights into the nature of molecular interactions already on a visual level, their undeniable value unfolds in the prediction of structural, physical, and thermodynamic quantities. Besides an almost unenumerable range of practical applications, such simulations have successfully been applied to predict structural intermediates along the reaction cycle of enzymes¹⁰¹ and to unravel their mechanistic role in catalysis^{102,103}. Simulation-based predictions aided in narrowing down multimerization patterns of a G protein-coupled receptor¹⁰⁴ and were crucial to determine how its C-terminus regulates membrane trafficking¹⁰⁵. Combined with binding free energy calculations, MD simulations demonstrated superior predictive power in the identification of hotspots of protein-protein interactions¹⁰⁶⁻¹⁰⁸ and the interpretation of thermodynamic profiles of small-molecular binders towards their targets¹⁰⁹. Especially combined with complementary experimental and computational methods, they are an integral tool for the analysis of biomolecular processes and the characterization of structure-function relationships in molecular complexes.

2.2.2 End-point binding free energy calculations

In the end-point approaches MM-GBSA and MM-PBSA, the free energy for binding of complex formation between a ligand and a protein is expressed according to **Equation 12-Equation 15** (Section 2.1.3). However, solvation and desolvation effects are already implied in these equations. Since the solvation free energy (ΔG_{solv}) is calculated as a separate term in both approaches, the binding free energy calculated via the end-point approaches is more appropriately expressed as⁷⁵:

$$\Delta G_{\text{bind}} = \Delta H - T\Delta S = \Delta E_{\text{MM}} + \Delta G_{\text{solv}} - T\Delta S \quad \text{Equation 26}$$

In **Equation 26**, ΔE_{MM} represents the change in the gas-phase molecular mechanics (MM) energy, ΔG_{solv} the change in solvation free energy, and $-T\Delta S$ the change in conformational entropy upon complex formation^{32,76}. The ΔE_{MM} term covers changes in bonded and nonbonded interactions, more precisely the internal energies ΔE_{int} (which are calculated from the bond, angle, and dihedral terms of the force field), the electrostatic energies ΔE_{elec} , and the van der Waals energies ΔE_{vdw} (**Equation 27**).

ΔG_{solv} , in turn, is expressed as a sum of a polar contribution $\Delta G_{\text{GB/PB}}$ and a nonpolar contribution ΔG_{nonpol} . Since ΔG_{nonpol} is, in most cases, calculated from the solvent-accessible surface area (SASA)¹¹⁰, I will refer to this term as ΔG_{SA} from this point on (**Equation 28**).

$$\Delta E_{\text{MM}} = \Delta E_{\text{int}} + \Delta E_{\text{elec}} + \Delta E_{\text{vdw}} \quad \text{Equation 27}$$

$$\Delta G_{\text{solv}} = \Delta G_{\text{GB/PB}} + \Delta G_{\text{SA}} \quad \text{Equation 28}$$

The polar contribution $\Delta G_{\text{GB/PB}}$ is estimated using either a finite-difference method to solve the Poisson–Boltzmann equation – which may be linearized to obtain a faster numerical solution^{32,111} –, or a Generalized Born (GB) pairwise approximation¹¹². For biomolecular systems, the electrostatic potential $\varphi(\mathbf{r})$ can be calculated by solving the following extension of the Poisson–Boltzmann equation, in which the ion concentration is taken into account³²:

$$\nabla \cdot \boldsymbol{\varepsilon}(\mathbf{r}) \nabla \varphi(\mathbf{r}) + 4\pi\lambda(\mathbf{r}) \sum_{i=1}^N z_i e c_i e^{-\frac{z_i e \varphi(\mathbf{r})}{k_B T}} = -4\pi\rho(\mathbf{r}) \quad \text{Equation 29}$$

Here, $\boldsymbol{\varepsilon}(\mathbf{r})$ denotes the dielectric distribution function, $\varphi(\mathbf{r})$ the electrostatic potential distribution function, and $\rho(\mathbf{r})$ the atomic charge density as functions of the atomic coordinates \mathbf{r} . Moreover, $\lambda(\mathbf{r})$ is the ion-exclusion function, which ensures an ion concentration close to zero within solute cavities and includes a Stern-layer correction^{111,113}. For ion types $i = [1..N]$, z_i is the charge, c_i the bulk number density at temperature T , and e the electron charge ($\sim 1.602 \cdot 10^{-19}$ C). Solving **Equation 29** numerically for several thousands of MD snapshots, however, is computationally demanding. GB models provide an approximation of the Poisson–Boltzmann equation at less computational expense. These models assign a spherical volume with a low dielectric constant ($\varepsilon_{\text{in}} = 1$) to each atom, inside which the electrostatics of the surrounding atoms are descreened. The extent to which an atom exerts a descreening effect on its environment is therefore dependent on its ascribed volume, which is calculated from its effective Born radius¹¹⁴. The GB equation without considering mobile ions is given by

$$\Delta G_{\text{GB}} = - \left(\frac{1}{\varepsilon_{\text{in}}} - \frac{1}{\varepsilon_{\text{solv}}} \right) \sum_{i,j} \frac{q_i q_j}{f_{\text{GB}}} \quad \text{Equation 30}$$

where ε_{in} is the dielectric constant of the solute, $\varepsilon_{\text{solv}}$ is the dielectric constant of the solvent, q_i and q_j are the atomic partial charges of atoms i and j , and f_{GB} is a function that interpolates between effective Born radii α at short distances r_{ij} , and r_{ij} itself at large distances^{112,115} like so:

$$f_{\text{GB}} = \sqrt{r_{ij}^2 + \alpha_{ij}^2 e^{\frac{r_{ij}^2}{4\alpha_{ij}^2}}} \quad \text{Equation 31}$$

Here, α_{ij} is the geometric average of the effective Born radii α_i and α_j . Extensions of **Equation 30** that allow for the treatment of mobile ions are available¹¹². These equations show that the calculation of ΔG_{GB} is strongly dependent on the effective Born radii – a dependence reflected in the numerous GB models (GB^{HCT} 116,117, GB^{OBC} 118,119, GBn¹²⁰, GBNSR6¹²¹) implemented in the AMBER software package⁸⁴.

Since the nonpolar contribution to the solvation free energy is mainly attributable to the formation of solvent-filled cavities in the solute and the resulting solute-solvent van der Waals interactions, it is reasonable to estimate this contribution (ΔG_{SA}) via the SASA. In practice, ΔG_{SA} is calculated by

$$\Delta G_{\text{SA}} = \gamma \text{SASA} + b \quad \text{Equation 32}$$

where γ ($\gamma = 0.00542 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ in AMBER⁸⁴) is the surface tension constant, and b ($b = 0.92 \text{ kcal mol}^{-1}$ in AMBER⁸⁴) is a correction constant³². Improved models that include atom-specific parameters for surface tension or split the nonpolar contribution into partial contributions from a cavity and dispersion term (**Equation 33**) are available.

$$\Delta G_{\text{CD}} = \gamma \text{MS} + b + \Delta G_{\text{disp}} \quad \text{Equation 33}$$

Here, MS denotes the molecular surface, and ΔG_{disp} is calculated by integration over the solvent-occupied volume¹²².

Finally, the configurational entropy contribution $-T\Delta S$ is often calculated by estimating S via normal mode analysis^{75,123} or quasi-harmonic analysis. Due to the high computational costs associated with these calculations, the entropic term is often omitted – giving rise to an effective binding free energy ΔG_{eff} –, which seems legitimate when structurally similar ligands are compared for which an equal contribution to conformational entropy is expected. If at all, entropy calculations are often performed on only a fraction of MD snapshots, making it difficult to obtain a converged value. Therefore, in one of the studies on which this thesis is based (Publication III), the differences in the entropic contribution between different ligands were calculated on thousands of snapshots. Alternative methods, such as truncated entropy¹²⁴, interaction entropy^{125,126}, or weighted solvent-accessible surface area¹²⁷, show promising results at reduced computational costs but still have to be extensively evaluated to reveal their full potential.

In practice, first, an explicit-solvent MD of the protein-ligand complex under study is carried out. Subsequently, solvent molecules and ions are removed from the MD trajectory and ΔG_{solv} is calculated using the GBSA or PBSA approach³². To save computing time, this step is often performed on only a fraction of the generated MD snapshots. Finally, the conformational entropy may be estimated from normal mode analysis or quasi-harmonic analysis; a step often performed on an even smaller selection of snapshots due to the computational effort involved. These steps are then repeated twice on the same trajectory: once only for the protein and once only for the ligand. In the sense of a thermodynamic cycle (**Figure 2**), ΔG_{bind} (or ΔG_{eff}) is then calculated as an average of the (effective) free energy differences over all snapshots (**Equation 34**). Since the snapshots for complex, protein, and ligand come from the same trajectory, this procedure is referred to as the “single trajectory approach”.

$$\Delta G_{\text{bind}} = \langle G_{\text{PL}} - G_{\text{P}} - G_{\text{L}} \rangle_{\text{PL}} \quad \text{Equation 34}$$

In the alternative “three trajectory approach”, protein, ligand, and complex are simulated separately, and ΔG_{bind} (or ΔG_{eff}) is calculated as the differences of the average (effective) free energies in the three trajectories (**Equation 35**). This approach is suitable when ligand binding is accompanied by larger conformational changes¹²⁸, but is overall less accurate and yields higher standard errors³².

$$\Delta G_{\text{bind}} = \langle G_{\text{PL}} \rangle_{\text{PL}} - \langle G_{\text{P}} \rangle_{\text{P}} - \langle G_{\text{L}} \rangle_{\text{L}} \quad \text{Equation 35}$$

Many programs were developed to simplify the workflow of setting up simulations, processing the trajectory, and calculating the necessary energy contributions. Within the AMBER software package⁸⁴, MMPBSA.py¹²⁹, FEW^{130,131}, and mm_pbsa.pl are well integrated and particularly noteworthy.

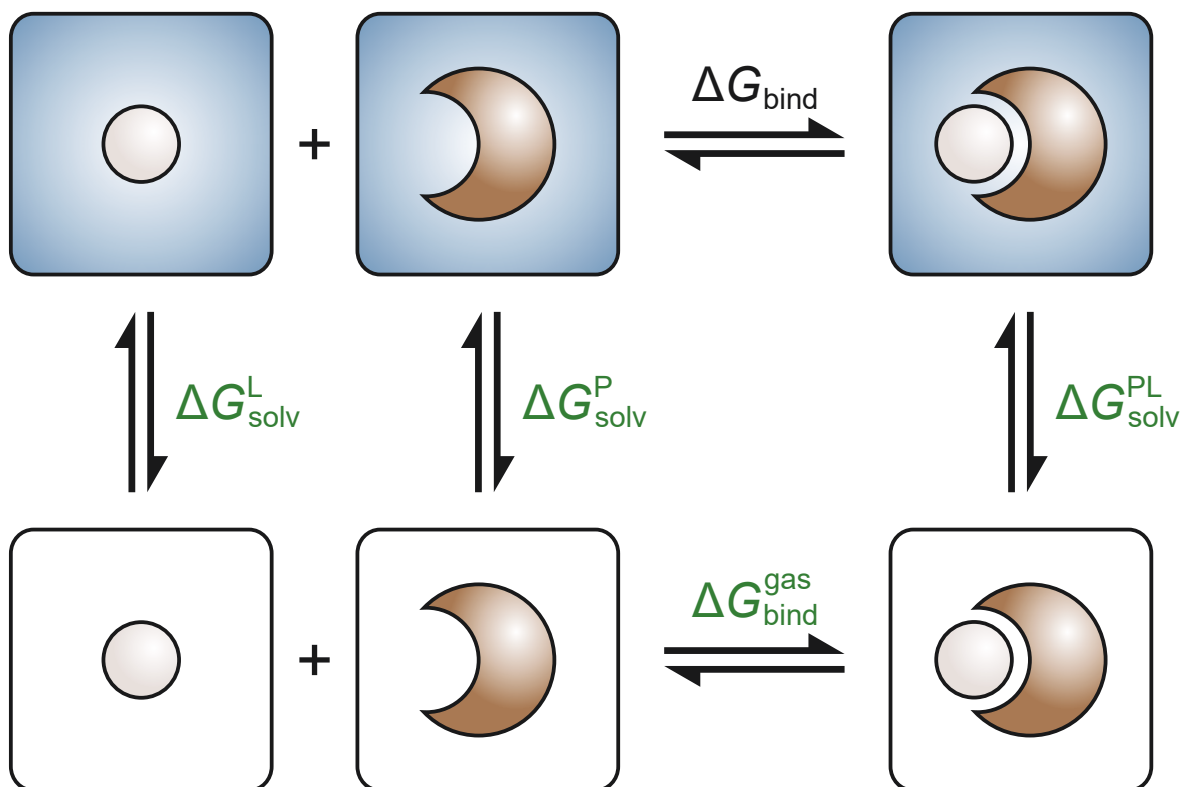


Figure 2: Thermodynamic cycle for binding free energy calculations in the end-point approaches MM-GBSA and MM-PBSA.

Solvated systems are depicted in blue boxes, while systems in the gas phase are depicted in white boxes. Free energies directly calculated as part of the approach are highlighted in green.

The predictive power of these methods and how their performance depends on various parameters has been extensively and systematically investigated in a series of studies^{32,126,132-139}. A detailed description of all results from these studies is beyond the scope of this thesis, yet, in summary, the quality of the binding free energy prediction depends on the force field^{134,140}, the atomic charge model¹³⁴, the continuum solvation method¹⁴¹, the internal dielectric constant¹³⁶, the sampling method¹⁴⁰ and the consideration or disregard of configurational entropy¹²⁶. Particularly noteworthy, however, is the dependence of the results on the simulation length that was discerned in the first of these studies¹³². Comparing simulation lengths between 400 and 4800 ps, it was found that longer simulations often worsen the correlation with experimental values. Unfortunately, this conclusion seems to have had a significant impact on the design of subsequent studies¹³⁴ and has not been revisited yet, leaving end-point methods virtually unevaluated at currently accessible simulation lengths. Own unpublished work on a series of Cyclin-dependent kinase 2-inhibitor complexes suggests that not the simulation length per se, but a greater number of distinct binding poses visited during a simulation may deteriorate the predictions.

Publication II and Publication III rely heavily on end-point methods applied to several simulations with a length in the upper nanosecond range.

Taken together, end-point free energy methods give a reliable and comparatively fast estimate of the binding free energy. Despite their strong dependence on the calculation parameters chosen, they are well suited for a variety of applications. In virtual screening approaches, they are a valuable tool for an efficient rescoring of docking poses, vastly improving the identification of the correct binding pose and the relative ranking of candidate ligands^{133,136,142-148}. They have also proven their value in the identification of critical protein-ligand interactions, which has simplified rational ligand design in a variety of case studies^{57,148-154}. Due to the continuous improvement of end-point methods, presumably supported by machine learning techniques³² in the near future, it can be expected that these methods will continue to play a major role in the quantitative characterization of protein-ligand interactions.

2.3 Modeling of function-determining protein-ligand interactions

So far, I have discussed the basic principles and models of molecular recognition and described the fundamental kinetic and thermodynamic quantities used to characterize protein-ligand interactions. Furthermore, I provided an overview of computational techniques that can be used to predict these quantities. In the following subsections, I will introduce the systems and specific questions I have worked on using these techniques.

2.3.1 Functional selectivity of two bile acids for integrin-mediated signaling pathways

(Parts of this section were published in the review “*Liver cell hydration and integrin signaling*“¹⁵⁵)

Integrins are $\alpha\beta$ heterodimeric transmembrane adhesion receptors and play a major role in the formation and regulation of cell-matrix and cell-cell attachments^{156,157}. By coupling the extracellular environment to the cytoplasmic signal transduction, integrins promote the anchoring of cells and thereby stabilize tissue architecture. Impairment or even loss of integrin function inevitably leads to a disruption of tissue organization and is linked to a range of cancer types, but also musculoskeletal, cardiovascular, and inflammatory diseases^{156,158-161}. Although the role of integrins as anchor proteins, which mediate communication between the intracellular compartment and the extracellular matrix (ECM), is rightfully emphasized in the literature, this focus often overshadows their equally outstanding ability to finely regulate signal transduction¹⁶²⁻¹⁶⁶. The basis for this ability must be an intricate regulation of the conformational equilibrium of these proteins.

Humans express 18 α and 8 β subunits, which can arrange into 24 distinct integrin heterodimers (**Figure 3A**)¹⁶³. The global architecture of these heterodimers is conserved across all multicellular organisms¹⁵⁶. A large extracellular region (ectodomain) that consists of several domains is connected to two generally short cytoplasmic tails via two single-span transmembrane regions (**Figure 3B**). α and β subunits are unrelated with respect to their sequence, but the subunits themselves are globally conserved across almost all species. An exception to this is a group of vertebrate α subunits with an additional inserted domain. This insertion, the α I or α A domain, forms the binding site in collagen-binding integrins and some laminin-binding integrins. A divalent cation (usually Mg^{2+}) within the “metal ion-dependent adhesion site” (MIDAS) of this binding site mediates binding of extracellular ligands either via an aspartate or glutamate residue. Regardless of whether an α I domain is present or absent, the ectodomains of all integrin α subunits carry a β -propeller domain, two calf domains, and a thigh domain with an IgG-like structure. Ectodomains of β subunits all bear a β I (or β A) domain homologous to the α I domain, a hybrid domain, a plexin-semaphorin-integrin (PSI) domain, four epidermal growth factor (EGF)-like domains, and a β -tail domain (**Figure 3B**). In integrin without an α I domain, the β I domain takes over ligand engagement; however, the β I domain carries three divalent cations instead of just one. While the β I MIDAS carries a Mg^{2+} ion, as it does in α I domains, the structural information available to this date suggests that the site “adjacent to MIDAS” (ADMIDAS) and the synergistic metal ion-binding site (SyMBS, formerly referred to as LIMBS for “ligand-associated metal-binding site”) both carry a Ca^{2+} ion.

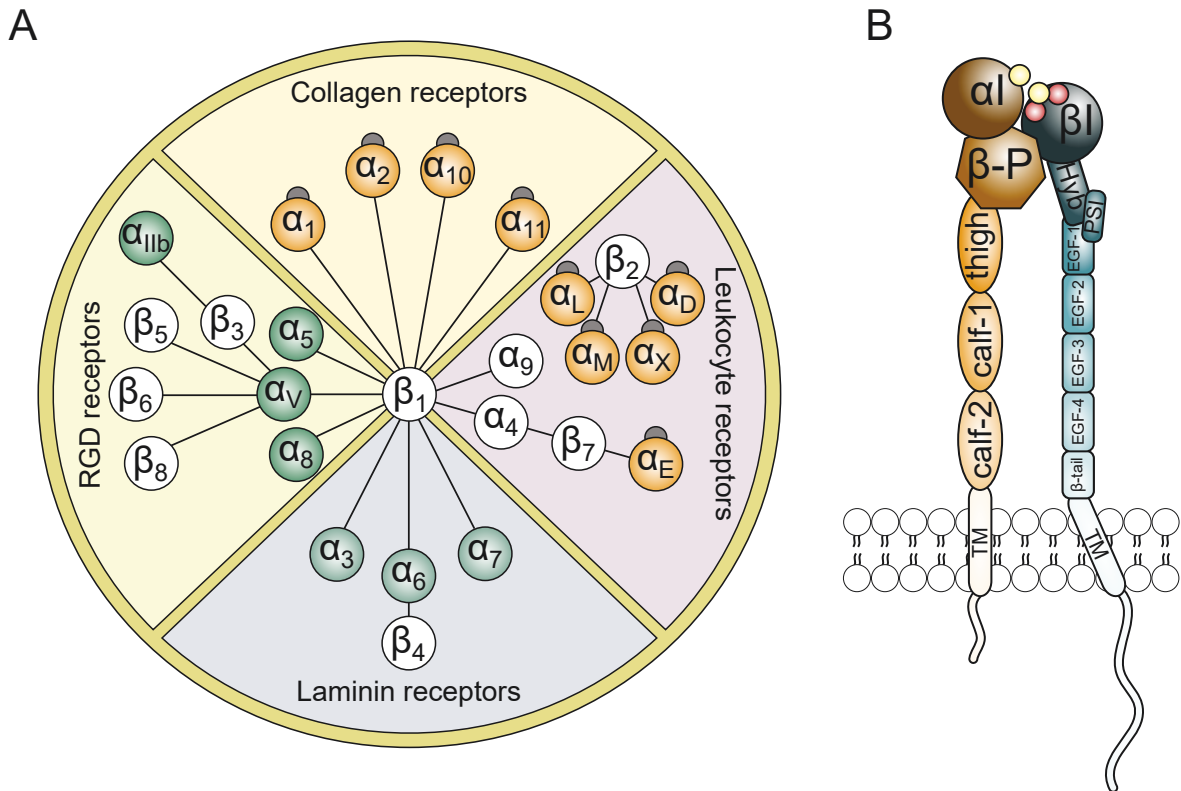


Figure 3: General architecture of integrins.

(A) Representation of known integrin heterodimer pairs in vertebrates and their grouping by ECM ligand. Integrin α subunits are colored according to their evolutionary relationship; α subunits with an αI domain carry a gray attachment (Figure adapted and modified from ref. ¹⁶⁶ and ref. ¹⁶⁷). (B) Domain architecture of an integrin with an αI domain. The MIDAS is depicted as a yellow sphere, the ADMIDAS (distal) and SyMBS (proximal) are depicted as red spheres. Membrane lipids indicate the membrane region in the background of the transmembrane regions. β -P: β -propeller domain; Hyb: hybrid domain; TM: transmembrane domain (Figure adapted and modified from ref. ¹⁵⁶).

According to the currently accepted model, derived from the first crystal structures of a complete integrin ectodomain^{168,169} and electron microscopy studies^{170,171}, integrins adopt at least three stable conformational states throughout their activation cycle¹⁵⁶ (Figure 4). The “bent-closed” (Figure 4, left) conformation with low affinity for ECM ligands represents a resting state of integrins, while the “extended-closed” conformation (Figure 4, center) with medium affinity and the “extended-open” conformation (Figure 4, right) with high affinity are binding-competent^{156,162}.

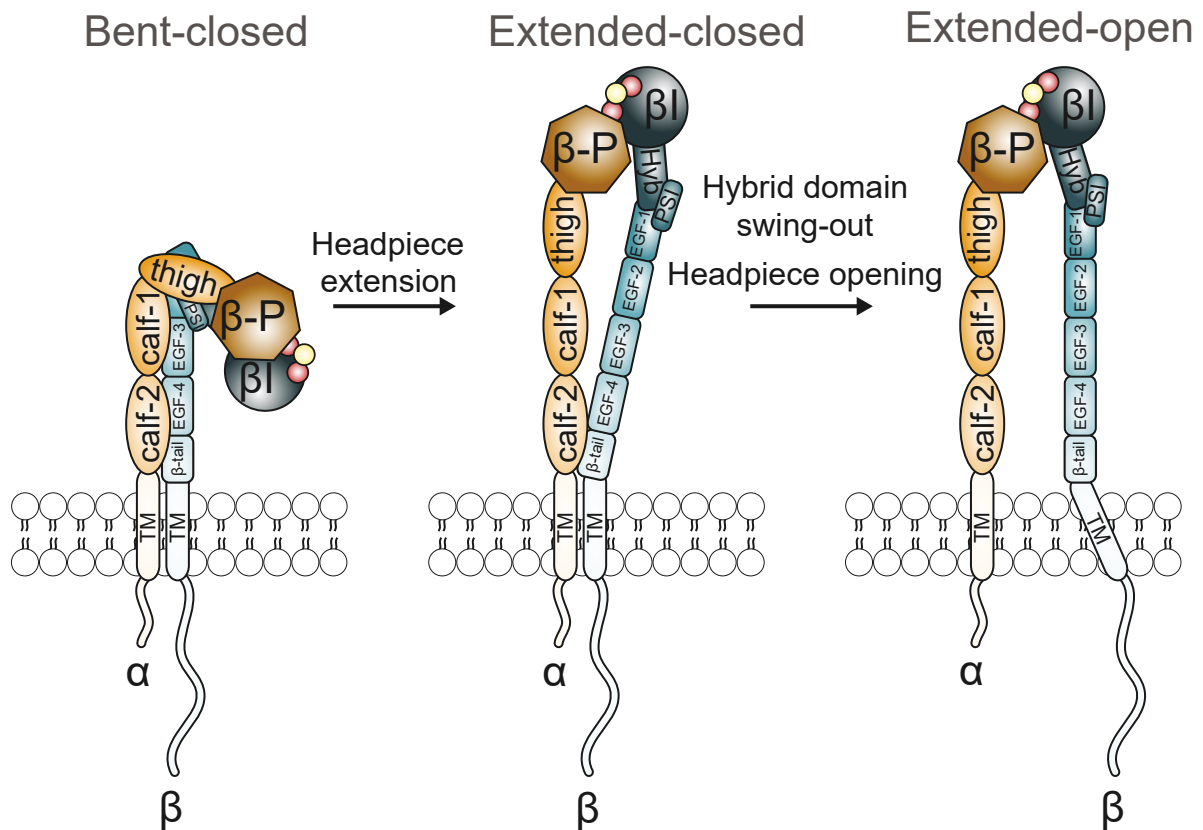


Figure 4: Stable conformational states in the activation cycle of an integrin without an αI domain.

Integrins undergo a conformational transition from a “bent-closed” (left) to an “extended-closed” conformation (center) by extending their ectodomain headpiece via the “joints” between the thigh and calf-1 domain and EGF-1- and EGF-2 domain. The transition between “extended-closed” and “extended-open” conformation is achieved by a swing-out of the hybrid domain and dissociation of the transmembrane helices. Domains are labeled and colored as in **Figure 3** (Figure adapted and modified from ref. ¹⁵⁶).

According to the current model, integrin-dependent ligand binding always follows the same basic multi-step procedure^{156,172-174}. First, intracellular adaptor proteins such as talin and kindlin are recruited. Binding of these adaptor proteins to the integrin cytoplasmic tails induces the long-range conformational changes described in **Figure 4**, which increase the affinity for ECM ligands. Now the integrins engage their respective ECM components with high affinity and simultaneously bind to the actin cytoskeleton. This bilateral connection results in a clustering of integrins and the formation of anchoring cell-matrix connections known as focal adhesions. Numerous intracellular binding partners can subsequently be recruited, which may initiate a variety of downstream signaling pathways^{175,176}.

In the context of this thesis, it is essential to note that this model does neither include nor exclude the possibility of ligand binding to the low-affinity “bent-closed” conformation (**Figure 4**, left). Omitting this scenario is reasonable in the context of binding to voluminous ECM ligands, which are unlikely to reach the binding pocket for steric reasons alone. However, it conveys the impression that the modulation of integrin activity and integrin

signaling by extracellular binding partners of lower molecular weight is not possible. Moreover, the model limits the concept of bidirectionality and reciprocity¹⁶⁶ of integrin signaling to the focal adhesion phase. A large body of experimental evidence suggests that the “bent-closed” conformation may very well bind small- and macromolecular ligands. Among other studies, Adair *et al.* reconstructed structures of bent $\alpha_V\beta_3$ integrin bound to the fibronectin type III domains 9 and 10 from transmission electron microscopy and single-particle image analysis¹⁷⁷, and Zhu *et al.* characterized the complete conformational transition of the $\alpha_{IIb}\beta_3$ integrin headpiece (i.e., β -propeller domain, β I domain, hybrid domain, and PSI domain) after binding of Arg-Gly-Asp (RGD)-containing peptides¹⁷⁸ (**Figure 5**). It is difficult to imagine that the isolated headpiece would be capable of ligand binding while the complete integrin would not. Also, ligand binding to the isolated headpiece induces a conformational change that exactly matches the conformational change following adaptor protein-induced integrin activation^{178,179}. It is, therefore, likely that ligand binding to the “bent-closed” ectodomain influences the integrin conformational equilibrium to a similar extent as binding of intracellular adaptor proteins to the cytoplasmic tail.

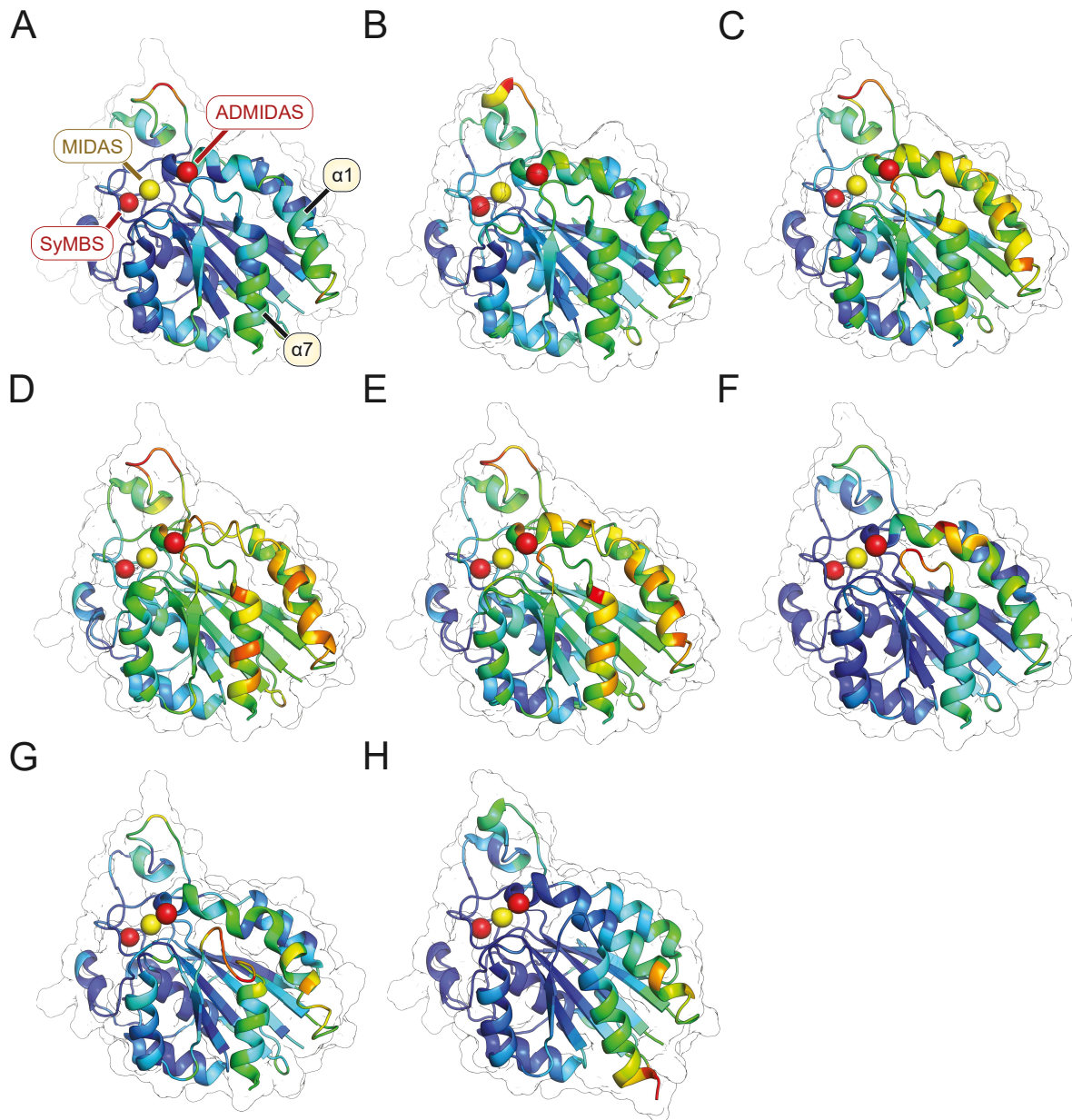


Figure 5: Complete pathway of $\alpha_{11b}\beta_3$ integrin headpiece opening.

Crystal structures of the $\alpha_{11b}\beta_3$ β I domain along the activation pathway from (A) unbound and inactive to (H) bound and fully activated¹⁷⁸. Color coding, according to the rainbow spectrum from blue to red, represents the average residue B factors. The MIDAS ion is depicted as a yellow sphere, the ADMIDAS and SyMBS ions as red spheres. Helices $\alpha 1$ and $\alpha 7$ are labeled accordingly. Ligands are not depicted. (Figure adapted from Publication I).

Binding of the hydrophilic bile acid tauroursodeoxycholic acid (TUDC) to the “bent-closed” conformation of $\alpha_5\beta_1$ integrin serves as another example of integrin activation by extracellular stimuli. $\alpha_5\beta_1$ is one of the predominant integrin isoforms in liver tissue¹⁸⁰ and directly activated by TUDC¹⁸¹⁻¹⁸³. This activation triggers a signal transduction pathway that involves phosphorylation of c-Src, focal adhesion kinase (FAK), epidermal growth factor receptor (EGFR), phosphatidylinositol 3-kinase (PI3-K), mitogen-activated protein kinases (MAPK) Erk-1 and Erk-2, and p38^{MAPK}¹⁸¹. Simultaneous activation of Erk-1/-2 and p38^{MAPK}

leads to a pronounced insertion of the bile salt export pump (Bsep) into the canalicular membrane, which explains the choleric effect of TUDC. Chemical modifications to the 3 α ,7 β -dihydroxycholane (ursodeoxycholane) scaffold of TUDC, however, result in either a complete loss of activity or – in the case of taurocholic acid (TC) – in a compound that even inhibits TUDC-induced activation of $\alpha_5\beta_1$ integrins¹⁸³. Whether chemical modifications of the bile acid's side chain – while preserving the ursodeoxycholane scaffold – retain or even increase activity towards $\alpha_5\beta_1$ integrin has so far not been investigated.

Thus, in Publication I, we first predicted the activity of four side chain-modified TUDC derivatives (24-*nor*-ursodeoxycholic acid (*nor*UDCA), its taurine conjugate *Tnor*UDCA, glyoursodeoxycholic acid (GUDC), and unconjugated ursodeoxycholic acid (UDCA)) by MD simulations and experimentally validated these predictions qualitatively and semi-quantitatively through immunofluorescence staining experiments. For the most active of these four compounds, *nor*UDCA, we characterized the signal transduction events following integrin activation in detail and detected both qualitative and quantitative differences to TUDC-induced signal transduction. After thoroughly challenging our data with integrin binding assays, densitometric fluorescence intensity analyses, and a detailed description of the FAK phosphorylation pattern, we concluded that TUDC and *nor*UDCA exhibit functional selectivity for overlapping but distinct $\alpha_5\beta_1$ -mediated signal transduction pathways. Thus, with this work, we did not only identify a second agonist for $\alpha_5\beta_1$ integrin with an ursodeoxycholane scaffold, but we also described functional selectivity on integrins without an αI domain for the first time. Our findings also provide a rationale for the different therapeutic indications of UDCA and *nor*UDCA in cholestatic liver disease¹⁸⁴.

2.3.2 Mechanistically distinct inhibition of NMDA receptor isoforms by tauro-CDC

Glutamate is the primary excitatory neurotransmitter¹⁸⁵⁻¹⁸⁷. It is involved in learning and memory processes, voluntary motor activity, and the transmission of afferent impulses. The effects of glutamate are mediated by the ionotropic NMDA, α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA), and kainate receptors¹⁸⁵, which are ligand-gated ion channels, and by the metabotropic glutamate receptors (mGluRs), which belong to the family of G protein-coupled receptors. Ionotropic glutamate receptors (iGluRs) are cation channels whose activation directly leads to the depolarization of neurons. They are named based on the respective synthetic ligands that specifically activate these receptor types. To this date, NMDA receptors (NMDARs) are the only pharmacological targets among iGluRs and stand out from other iGluRs due to several unique properties: Activation of most NMDARs

requires binding of glycine (or D-serine) and glutamate, which act as co-agonists so that the occupancy of both binding sites is required for channel opening¹⁸⁵. Moreover, the channel can be blocked by Mg^{2+} ions in a voltage-dependent manner and shows a comparably high permeability for Ca^{2+} ions¹⁸⁷. This high Ca^{2+} permeability is particularly relevant for pathophysiological processes during which substantial amounts of excitatory neurotransmitters are released, such as strokes, epilepsy and neurodegenerative diseases; the excessive stimulation of NMDA receptors leads to cellular Ca^{2+} overload which ultimately results in cell death. In this context, it becomes apparent why NMDAR antagonists appear to exhibit neuroprotective properties¹⁸⁸.

In total, seven genes encode the NMDA receptor subunits. A single *GRIN1* gene encodes the GluN1 subunit, four *GRIN2* genes encode GluN2A-D, and two *GRIN3* genes encode GluN3A-B^{185,187}. These subunits form heterotetramers of two GluN1 and either two GluN2^{189,190} or two GluN3 subunits¹⁹¹. A notable exception is the GluN1/GluN2/GluN3 heterotetramer with a 2:1:1 stoichiometry, but it is still unclear whether such an assembly is biologically relevant¹⁹² (also cf. **Figure 15** in subsection 5.1).

All NMDAR subunits share the same four-domain architecture¹⁸⁷ (**Figure 6A**). The ~350 amino acid long amino-terminal domain (ATD) features a bilobed structure and forms the first N-terminal interface between GluN1 and GluN2 subunits¹⁹⁰ (**Figure 6B**). It plays a guiding role in the assembly of subunits into heterotetramers^{193,194} and governs pharmacological and kinetic properties of GluN2-containing NMDARs^{195,196}. Moreover, ATDs can also bind Zn^{2+} and a number of other negative allosteric modulators¹⁹⁷⁻²⁰⁰ (**Figure 6B**).

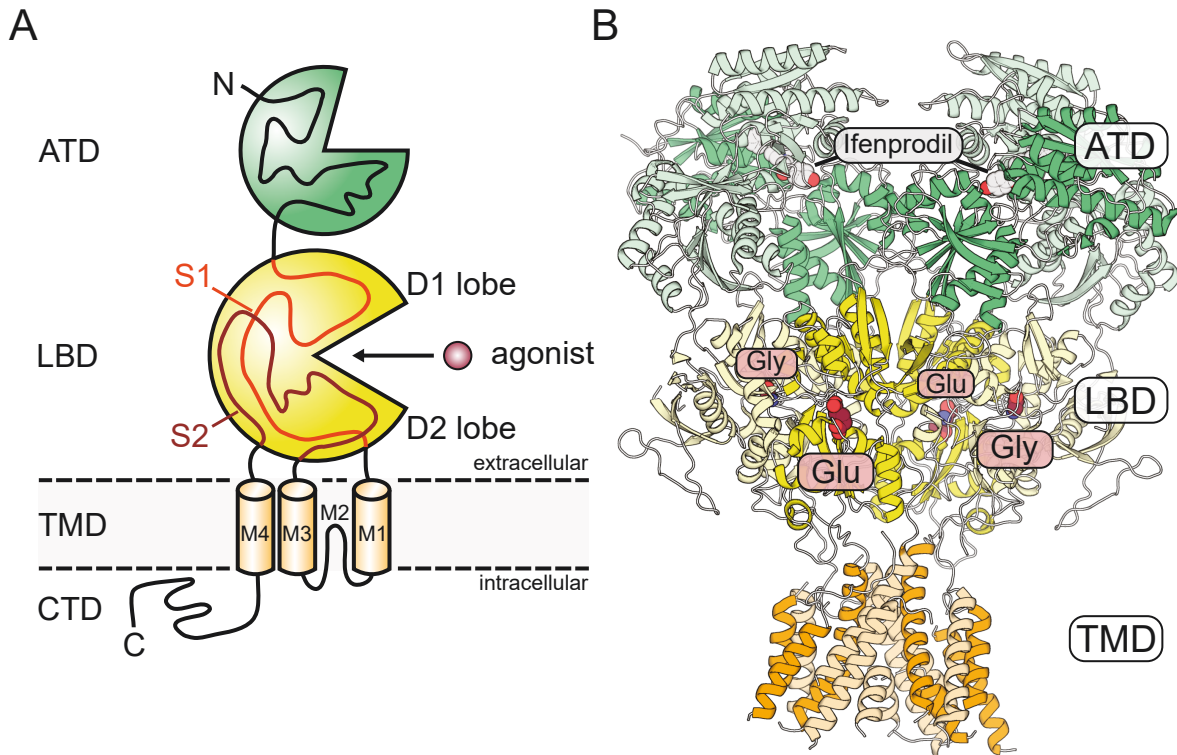


Figure 6: Domain architecture of NMDA receptors.

(A) Schematic representation of an NMDA receptor subunit (modified from ref. ¹⁸⁷). The polypeptide chain is drawn as an undulating line and can be traced from the N- to the C-terminus. The amino-terminal domain (ATD) is colored in green, the ligand-binding domain (LBD) in yellow, and the transmembrane domain (TMD) in orange. The C-terminal domain (CTD) has not been highlighted. NMDA receptor agonists (purple sphere) bind to the cleft between the two lobes (D1 and D2) of the LBD. (B) Cartoon representation of the crystal structure of the heterotetrameric GluN1/GluN2B NMDA receptor. The color scheme corresponds to the one used in panel A. GluN1 subunits are depicted in lighter shades. Ligands are depicted as spheres. The negative allosteric modulator ifenprodil (gray spheres) binds to the GluN1-ATD/GluN2D-ATD interface¹⁹⁰.

Following the ATD in sequence and structure is the ligand-binding domain (LBD), which is sometimes referred to as agonist-binding domain (ABD)¹⁸⁷ to emphasize the existence of other binding sites (i.e., those to which allosteric modulators bind). Similar to the ATD, it forms a bilobed structure (**Figure 7A**), but in which the interlobular cleft constitutes the binding site for glycine and D-Serine (in GluN1 and GluN3 subunits), or glutamate (in GluN2 subunits), respectively²⁰¹⁻²¹⁴. A series of crystal structures of the GluN1/GluN2A LBD heterodimer^{210,211,215} allows for speculations on how agonist binding is coupled to NMDAR activation. LBDs appear to be in an equilibrium between an “open cleft” and a “closed cleft” conformation. While binding of an agonist would shift the equilibrium towards the “closed cleft” conformation, binding of competitive antagonists such as 5,7-dichlorokynurenic acid (DCKA) in glycine binding sites and D(-)-2-amino-5-phosphovaleric acid (D-AP5) in glutamate binding sites would shift the equilibrium towards the “open cleft” conformation¹⁸⁷. LBD closure is accompanied by hydrogen bond formation between the two lobes, which might further stabilize the agonist-bound conformation^{216,217}. The energy

released during these processes induces a series of conformational changes in the NMDAR, which ultimately result in the opening of the channel pore¹⁸⁷. Similar to the ABDs, the GluN1/GluN2 LBD interface plays a crucial role in modulating NMDAR function. It is divided into three contact patches, named sites I-III²⁰². While sites I and III appear to primarily mediate heterodimerization of GluN1 and GluN2 subunits through hydrophobic interactions^{187,202}, site II serves as a binding site for positive and negative allosteric modulators in, at least, GluN2A-containing NMDARs (**Figure 7B**)²¹⁰⁻²¹².

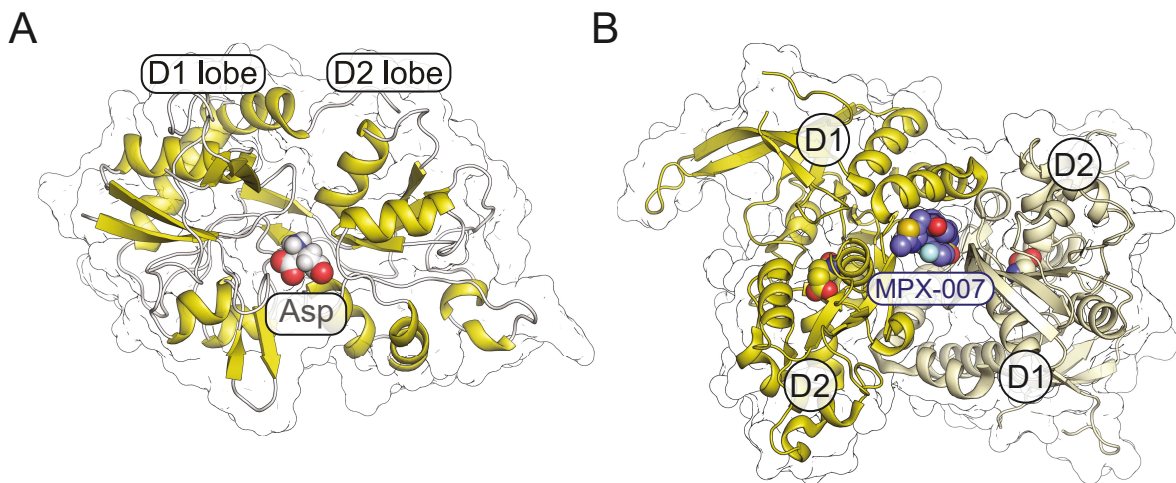


Figure 7: Ligand binding to the LBDs of NMDARs.

(A) Crystal structure of the GluN2D LBD in complex with L-aspartate (PDB ID: 3OEK)²⁰⁶. The two LBD lobes are labeled “D1 lobe” and “D2 lobe”. The binding site lies within the cleft between D1 and D2. (B) Crystal structure of the GluN1/GluN2A LBD interface bound with the negative allosteric modulator MPX-007 (blue spheres)²¹². The GluN1 subunit (right, lighter color shades) is bound to glycine (shown as spheres), the GluN2A subunit (left, darker color shades) to glutamate (shown as spheres). MPX-007 binds to site II²⁰² in the LBD “back-to-back” subunit interface.

The transmembrane domain (TMD) in each NMDAR subunit is composed of three transmembrane helices (M1, M3, and M4) and a reentrant loop (M2)¹⁸⁷ (**Figure 6A**). Within this loop resides the Q/R/N site, which controls permeability for divalent cations and is also responsible for Mg²⁺-dependent channel block^{187,218}.

Note that, while no structure of a GluN1/GluN3 heterotetramer could be determined so far, it is reasonable to assume that its global architecture is identical to that of GluN1/GluN2 heterotetramers. Therefore, the structural insights about the location of GluN1/GluN2 subunit interfaces can most likely be directly transferred to GluN1/GluN3 interfaces. However, the different ligand specificity of GluN2 and GluN3 subunits alone suggests that this is not necessarily applicable to the functions of these interfaces.

Neuronal NMDARs have been thoroughly characterized concerning their localization and their molecular and physiological function¹⁸⁶, but a variety of non-neuronal cells¹⁸⁶, such as

osteoblasts^{219,220} and osteoclasts²²⁰⁻²²², pancreatic beta cells^{223,224}, a wide range of kidney cell types²²⁵⁻²²⁷, lymphocytes^{228,229} and platelets²³⁰ also express NMDARs. Such non-neuronal NMDARs are subject to comparatively high plasma glutamate concentrations, which is why alternative agonists such as L-homocysteic acid and quinolinic acid are discussed¹⁸⁶. Furthermore, a variety of allosteric modulators can fine-tune NMDAR function through positive or negative allosteric modulation. These include bivalent and monovalent ions, protons, polyamines, and neurosteroids¹⁸⁵.

Neurosteroids such as pregnenolone sulfate potentiate or inhibit NMDA receptor activity in an isoform-dependent manner^{185,231,232}, and the close chemical relationship between neurosteroids and bile acids suggests that bile acids may share this property. In fact, chenodeoxycholic acid (CDC), which is also synthesized in the brain, has been identified as a combined antagonist of NMDA and GABA_A receptors²³³, explaining why its dietary supplementation can prevent neurological complications of hereditary diseases that involve defective bile acid synthesis²³⁴. In certain cholestatic liver diseases like progressive familial intrahepatic cholestasis (PFIC) and intrahepatic cholestasis of pregnancy (ICP), or acute hepatitis, plasma concentrations of bile acids and bile salts may exceed 100 μM ²³⁵⁻²³⁷. In such diseases, modulation of non-neuronal NMDARs could contribute to physiological and pathophysiological events, and a deeper understanding of these processes could also pave the way for pharmacotherapeutic interventions.

In Publication II, we combined computational, molecular biological, and electrophysiological approaches to characterize the effects of the most common human bile acids/salts – cholate and (tauro-)CDC – on NMDARs with different subunit compositions. We first found that CDC and tauro-CDC predominantly inhibited NMDA receptors with GluN2D and GluN3B subunits, while GluN2A- and GluN2B-carrying NMDARs were only weakly affected. Intriguingly, the molecular mechanisms underlying the inhibition of NMDARs with GluN2D subunits and GluN3B subunits were different: tauro-CDC acted as a negative allosteric modulator in GluN2D-carrying NMDARs, but as a competitive antagonist in GluN3B-carrying NMDARs. Data from molecular docking, MD simulations, and end-point binding free energy calculations supported the competitive mechanism in GluN3B subunits and helped in isolating Asp447 as a potentially selectivity-determining residue. Using MD simulations, we also predicted the allosteric binding site of tauro-CDC in the GluN1/GluN2D LBD interface, which was in excellent agreement with the known binding site for positive²¹⁵ and negative²¹² allosteric modulators in the GluN1/GluN2A LBD

interface. This binding site has so far not been described for GluN2D-carrying NMDARs and may provide an attractive target for selective modulators within the class of GluN2-carrying NMDARs since tauro-CDC shows little to no effect in NMDARs containing GluN2A or GluN2B subunits.

2.3.3 Enhanced activation of CNG channels by 8-substituted cyclic nucleotides

Cyclic nucleotide-gated (CNG) channels are heterotetrameric ligand-gated ion channels and essential for the transduction of visual^{238,239} and olfactory²⁴⁰⁻²⁴² signals²⁴³. In the corresponding sensory cells, they convert a chemical signal initiated by an increase in the intracellular concentration of guanosine 3',5'-cyclic monophosphate (cGMP) or adenosine 3',5'-cyclic monophosphate (cAMP) into an electrical response by increasing the conductance for cations through the channel pore²⁴³. Each of the four channel subunits carries a cyclic nucleotide-binding domain (CNBD) which converts the binding of a cyclic nucleotide monophosphate (cNMP) into a conformational change that fosters channel opening²⁴³.

A total of six genes – CNGA1-4, CNGB1, and CNGB2 – encode the subunits of CNG channels²⁴⁴⁻²⁵⁰. These subunits assemble in certain combinations to form different homo- or heterotetrameric channels: CNGA1, CNGA2, and CNGA3 subunits can form functional homotetramers in heterologous expression systems while CNGA4, CNGB1, and CNGB3 can only assemble with other subunits to functional channels. Native rod photoreceptor channels consist of three CNGA1 subunits and a single CNGB1b subunit²⁵¹⁻²⁵³—note in this context that CNGB1a and CNGB1b subunits are alternative splicing variants; cone photoreceptor channels consist of two CNGA3 and two CNGB3 subunits²⁵⁴; olfactory channels are composed of two CNGA2, one CNGA4, and one CNGB1b subunit²⁵⁵ (Figure 8).

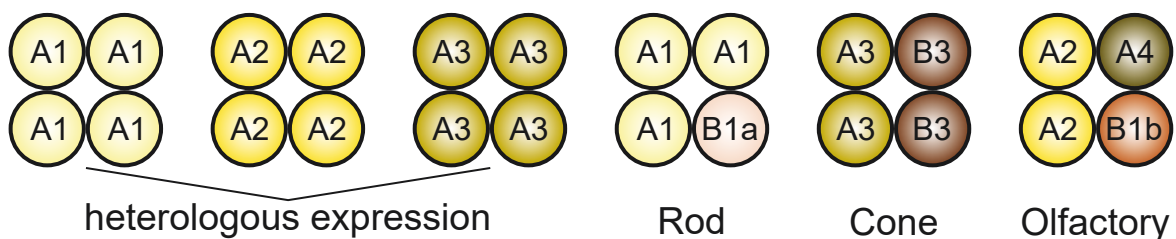


Figure 8: Subunit assemblies in CNG channels.

CNGA1-3 subunits can assemble to functional heterotetramers in heterologous expression systems. Native rod, cone, and olfactory channels are heterotetramers of different subunit composition and stoichiometry (Figure modified from ref. ²⁵⁶).

Both channel and subunit architecture (**Figure 9**) resemble voltage-dependent K^+ (K_v1-9) channels, Ca^{2+} -activated K^+ (K_{Ca}) channels, and transient receptor potential (TRP) channels^{256,257}: the functional channels are tetramers, and each subunit contains six transmembrane helices (S1-S6) divided into a voltage sensor domain (S1-S4) and a pore domain (S5, PH (pore helix), S6) connected by an S4-S5 linker^{258,259}. In CNG channels, a C-linker (CL) domain connects the pore domain to the C-terminally located CNBD and thereby assists the coupling between ligand binding and channel opening. An interesting and unique feature in CNGA subunits is a C-terminal leucine-zipper (CLZ) domain that seems to play an important role in channel assembly and was proposed to control subunit stoichiometry and arrangement in heterotetrameric channels^{252,260}.

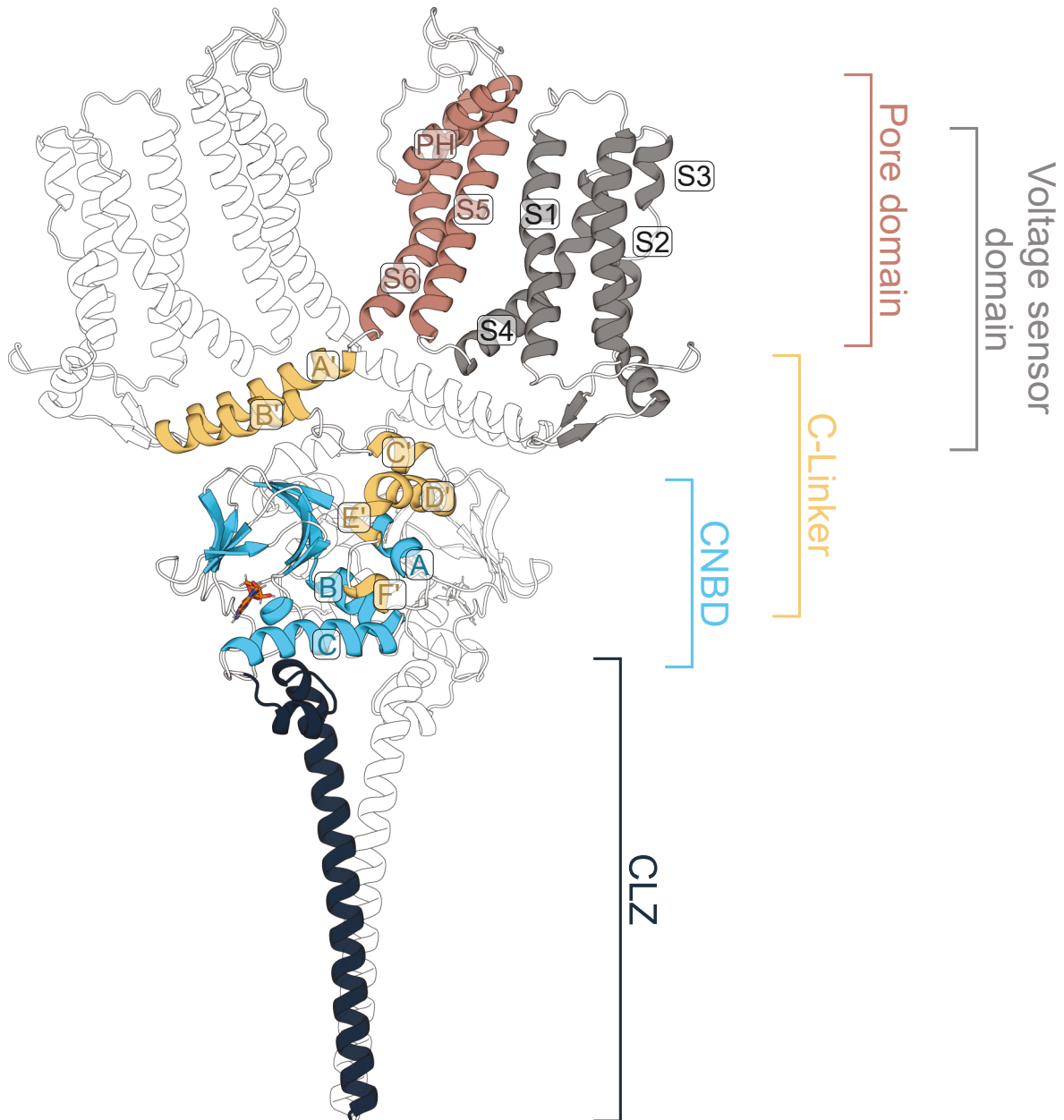


Figure 9: Architecture of two opposite CNGA2 subunits in a homotetrameric CNGA2 channel.

The structure was generated by comparative modeling using RosettaCM^{261,262} based on the structure of TAX-4 (PDB ID: 5H3O) determined via cryo electron microscopy (EM)²⁶³. Transmembrane segments are labeled S1-S6, C-linker helices are labeled A'-F', CNBD helices are labeled A-C. The CLZ domain was modeled based on restraints derived from the crystal structure of the CNGA3 CLZ domain (PDB ID: 3SWY) and a secondary structure prediction with PSIPRED 4.0^{264,265}.

Binding of cyclic nucleotides to the CNBDs (**Figure 10A**) leads to conformational changes, which are transduced into channel opening through the CL, the S6 helix, the PH, and the selectivity filter. Although the apparent affinities towards the different cyclic nucleotides differ between subunits, cGMP is typically superior to cAMP regarding its agonistic potency²⁶⁶. It is believed that the initial events of ligand recognition are dictated by the formation of interactions between the ribose/cyclic phosphate moieties and the β -roll in the CNBD²⁶⁶: A conserved arginine (R538 in rat CNGA2) within the β 7 strand engages in

electrostatic interactions and hydrogen bonds with the negatively charged cyclic phosphate and constitutes the most dominant interaction of the complex (**Figure 10B**). Interestingly, mutation of this arginine to glutamate markedly reduced ligand affinity but had no effect on ligand efficacy, suggesting that this amino acid primarily contributes to complex stabilization but is not involved in channel gating²⁶⁷. The adjacent threonine residue (T539 in rat CNGA2) likely contributes to the selectivity of CNG channels for cGMP versus the selectivity of hyperpolarization-activated cyclic nucleotide-gated (HCN) channels for cAMP^{268,269}. Furthermore, a conserved GE motif within a short α -helix between the β 6 and β 7 strands of the CNBD is suspected to interact with the 2'-OH group of the ribose ring (**Figure 10B**). Only after these initial ligand recognition events have occurred will the C-helix approach the purine ring of the cyclic nucleotide. This triggers a series of conformational changes and ultimately results in the opening of the channel.

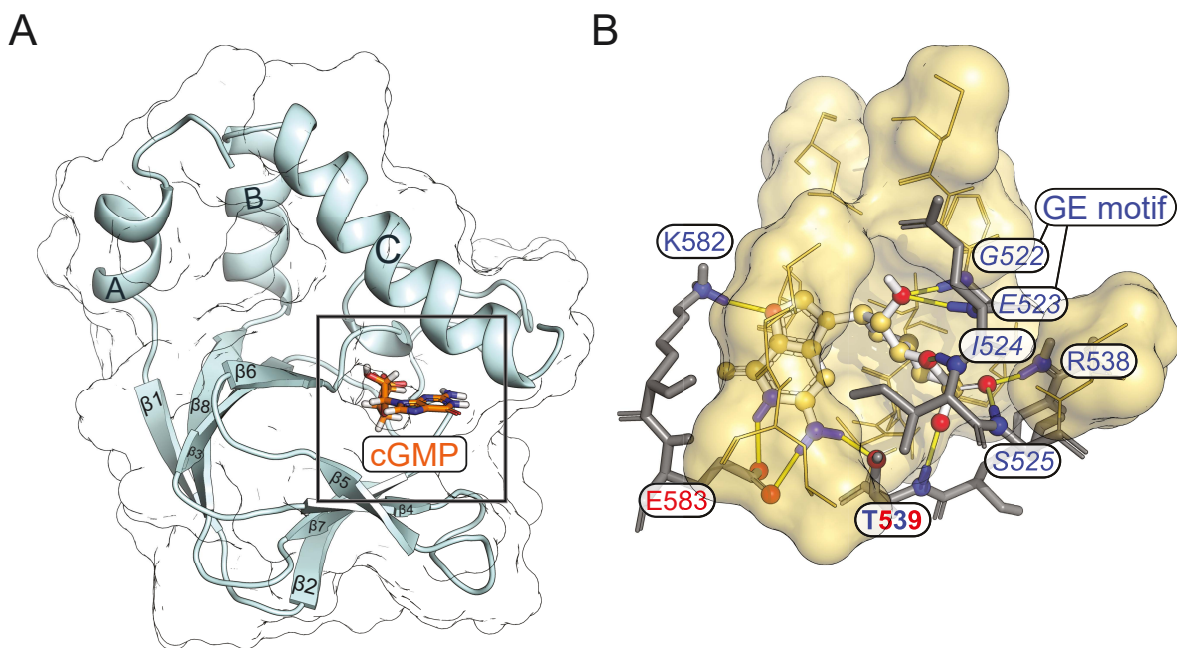


Figure 10: Structure of the CNGA2 CNBD and predicted binding mode of cGMP.

(A) Microarchitecture of the CNBD in CNGA2. The binding site for cyclic nucleotides is highlighted with a black box. cGMP is shown as orange sticks. CNBD helices are labeled A-C; the sheets in the β -roll are labeled β 1- β 8. (B) Predicted binding mode of cGMP (white sticks) in the CNGA2 CNBD binding site (gold surface). A residue is considered part of the binding site if at least one of its atoms lies within 4 Å of any atom of cGMP. Ligand atoms interacting with the protein only through van der Waals contacts are depicted as gold spheres; the corresponding amino acids in the protein are depicted in gold line-and-surface representation. Amino acids that form hydrogen bonds are shown as gray sticks. The atoms involved in hydrogen bonds are colored according to their involvement as donor (blue) or acceptor (red) and are depicted as spheres. Amino acid labels are colored according to the involvement of the respective amino acid as donor (blue), acceptor (red), or mixed donor-acceptor (blue and red). Amino acids in which only the side chain is involved in the hydrogen bond are labeled in normal font style, those in which only the backbone is involved are labeled in italic font style, and those in which both side chain and backbone are involved are labeled in bold font style.

To study the determinants of cyclic nucleotide binding to the CNBD and the coupling between ligand binding and channel opening in CNG channels, a number of modified cNMP

derivatives have been synthesized. While modifications of the ribose or the phosphate moiety were barely tolerated²⁷⁰, modifications of the purine ring emerged as more favorable^{271,272}. Particularly the 8-position turned out to be a promising candidate even for larger substitutions: Compounds like 8-bromo-cGMP and 8-(*p*-chlorophenylthio)-cGMP show higher apparent affinities than the parent nucleotide, are not hydrolyzed by phosphodiesterases, and are membrane permeable²⁷². Since the 8-position tolerates such a wide range of substitutions while preserving – if not increasing – the apparent affinity towards the channel proteins, it is reasonable to exploit these favorable properties for derivatizing cyclic nucleotides into fluorescent probes. Such dye conjugates are valuable tools for studying the relationship between ligand binding and activation gating in CNG and HCN channels.

In Publication III, we systematically studied how a series of 8-alkylthiopurine and 8-heteroalkylthiopurine substitutions of the cAMP and cGMP scaffold influence the apparent affinity towards CNG and HCN channels. Using a combination of molecular docking, MD simulations, and end-point binding free energy calculations, we found that the higher apparent affinities of derivatives with longer alkyl chains can be attributed to higher true affinities, which are the result of a complex enthalpy-entropy compensation. Especially 8-aminohexylthio-cAMP/cGMP and 8-aminodecylthio-cAMP/cGMP were promising starting points for the development of fluorescent cNMP conjugates with a great potential to study ligand binding-coupled activation gating. Such conjugates were synthesized and characterized in a follow-up study¹².

2.3.4 Activity determinants of *N*⁶-substituted cAMP analogs on mHCN2 channels

Shortly after its first description as the “active factor” in hormone-induced glycogen catabolism^{273,274}, cAMP emerged as the prototypical second messenger – a molecule that transmits an extracellular primary signal to the intracellular compartment, where it initiates a cascade of signal transduction events that elicit an appropriate cellular response²⁷⁵. After its synthesis at the cell membrane by adenylate cyclases²⁷⁶, cAMP can activate a variety of effector proteins and thereby initiate a differential intracellular response. Such effector proteins include protein kinase A (PKA)^{277,278}, CNG and HCN channels^{238,279} (cf. section 2.3.3), the exchange protein directly activated by cAMP (Epac)^{280,281} and Popeye domain-containing (POPDC)²⁸²⁻²⁸⁴ protein families. For a detailed discussion of the functions of these effector proteins, cf. the review by Zaccolo *et al.*²⁷⁵.

The evolutionarily ancient structural module that allows binding of cAMP to these proteins is the CNBD (compare section 2.3.3)²⁸⁵. This module, comprising ~120 amino acids, consists of an α -helical subdomain and an eight-stranded β barrel with a highly conserved phosphate-binding cassette (PBC) (**Figure 11A**). A highly conserved arginine and a glutamate or glutamine contained in this submodule (**Figure 11A**) stabilize the cyclic phosphate moiety and the 2'-OH group of the ribose, respectively. The buriedness of the PBC at the base of the β barrel shields the phosphodiester of cyclic nucleotides from solvent molecules and phosphodiesterases (**Figure 11B-D**)^{285,286}. The high degree of similarity between the interaction patterns of the various cAMP-binding proteins and the similar shape of their binding pockets complicates the identification of compounds that selectively activate only one of these proteins and the associated corresponding downstream signaling pathway. Chemical modification of the N^6 -position, however, yielded compounds with various degrees of selectivity for PKA over Epac, such as N^6 -benzoyl-cAMP (N^6 -Bnz-cAMP), N^6 -phenyl-cAMP (N^6 -Phe-cAMP), and N^6 -butyryl-cAMP (N^6 -MB-cAMP)²⁸⁷. Selectivity for Epac over PKA, on the other hand, could be achieved by a 2'-*O*-alkyl substitution, an 8- (*p*-chlorophenylthio) substitution, or a combination thereof²⁸⁷. However, studies on the selectivity profiles of these compounds have so far been devoid of data on their activity towards cyclic nucleotide-regulated.

In Publication IV, we set out to further advance the identification of selective compounds that allow for finer discrimination between the different cAMP-binding proteins. To do so, we studied the effect of N^6 -Phe-cAMP, N^6 -benzyladenosine-3',5'-cyclic monophosphate (N^6 -Bn-cAMP), and N^6 -Bnz-cAMP on functional murine HCN2 (mHCN2). The three compounds showed efficacies similar to cAMP but different apparent affinities. Combining molecular docking and subsequent mutagenesis experiments, we unraveled the structural determinants underlying these differences. Given the close evolutionary relationship between HCN and CNG channels and the structural similarity of their CNBDs, the results from our study may likely be directly transferable to CNG channels.

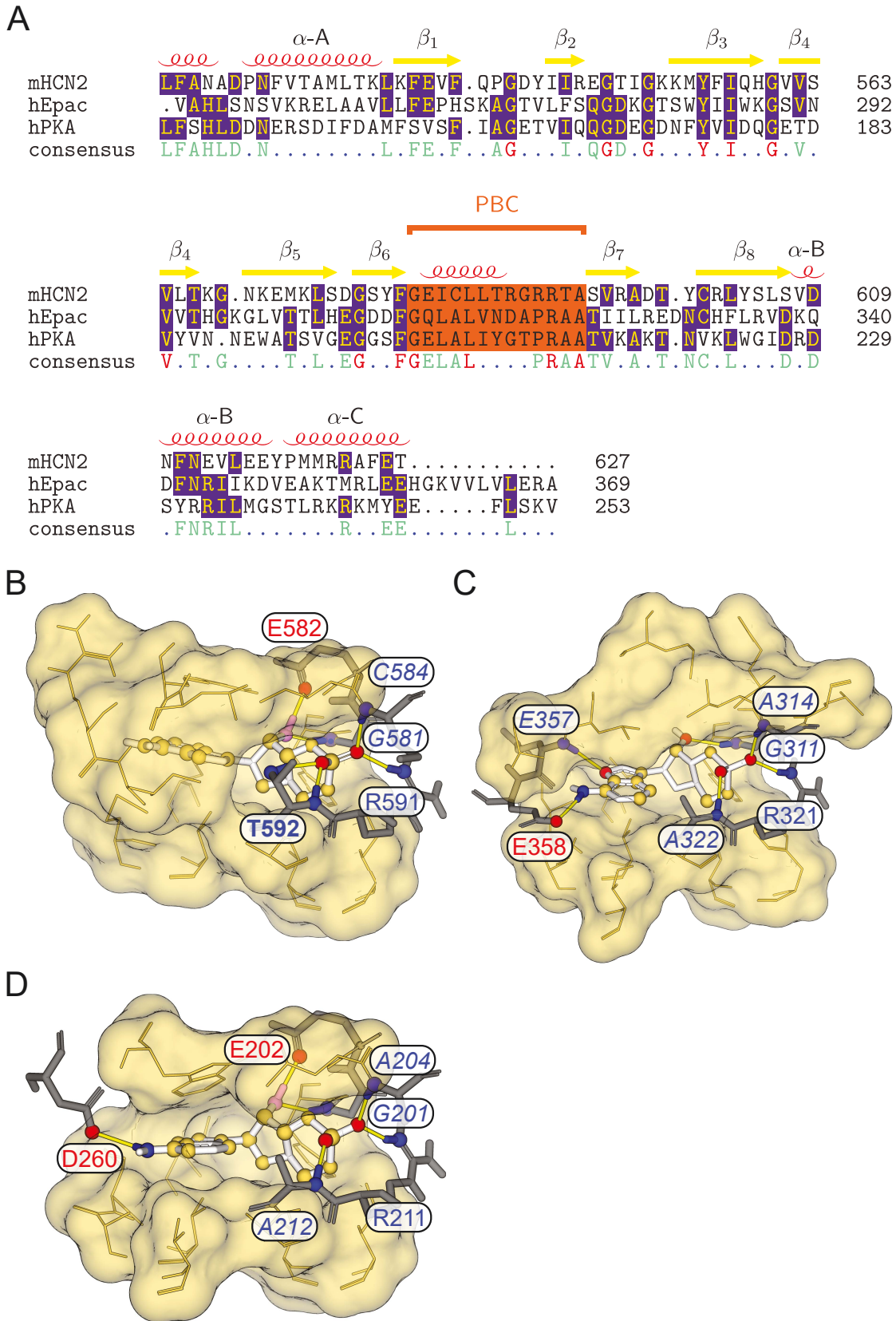


Figure 11: Comparison between the cAMP-binding proteins HCN2, Epac, and PKA.
 (A) CNBD segment from a multiple sequence alignment of murine HCN2, human Epac, and human PKA. Helices and sheets are labeled and indicated by red spirals and yellow arrows, respectively. Amino acids conserved in two of three sequences are highlighted in purple, the PBC is highlighted in orange. (B) Binding mode of cAMP (white sticks) in the mHCN2, (C) Epac, and (D) PKA (site A) binding site (gold surfaces). Color coding corresponds to the one used in Figure 10.

3 SCOPE OF THE THESIS

Protein-ligand interactions lay the foundation for the function of proteins and thus for the fate of the cell (section 1). Describing these interactions at the atomic level as precisely as possible, however, eludes the scope of many experimental techniques. Many descriptions of such interactions on the macroscopic scale are often called into question due to a lack of structural evidence and would therefore hugely benefit from support by structural biological and computational techniques. These techniques can often provide the missing link for experimental observations whose interpretation would otherwise remain speculative. In turn, computational techniques can rarely stand alone and require extensive input from wet lab experiments to provide successful predictions. In this thesis, I demonstrate by means of four examples how my collaborators and I found the answers to biological questions on specific protein-ligand interactions through an interplay of complementary experimental and computational methods.

In all cases, I contributed to answering these biological questions by a qualitative statement on the effect of certain ligands on a target protein (cf. introductory sections 2.1.1, 2.3.1) or a quantitative prediction of the thermodynamic quantities that describe the binding (cf. introductory sections 2.1.2, 2.1.3, and 2.1.4). I achieved this by combining a range of computational techniques, including molecular dynamics simulations (section 2.2.1) and end-point binding free energy calculations (section 2.2.2). In particular, the mutual complementation of experimental and computational results enabled my collaborators and me to answer the following questions:

- I. Besides TUDC, do any other bile acids or bile acid derivatives activate $\alpha_5\beta_1$ integrin in the liver? If so, does this activation initiate the same signal transduction pathways as TUDC? (section 4, Publication I)
- II. Can bile acids modulate the activity of NMDA receptors and could therefore be of physiological and pathophysiological significance in disease states associated with high plasma concentrations of bile acids? If so, which mechanisms underly this modulation (section 5, Publication II)?

In particular, I have contributed to answering these questions:

- i. Why does tauro-CDC act as an allosteric inhibitor of GluN2D-carrying NMDA receptors but as a competitive inhibitor of GluN3B-carrying NMDA receptors?

- ii. Which binding site controls the allosteric modulation of GluN2D-carrying receptors?
- III. Can 8-substituted analogs of cyclic nucleotides activate CNG and HCN channels? What is the structural basis for the measured apparent affinities, and are these molecules suitable for further development into fluorescent probes to study the function of CNG and HCN channels (section 6, Publication III)?

Here, I devoted special attention to the following questions:

- i. Why does the apparent affinity of 8-(acetyl)aminoalkylthio-cNMP derivatives on CNG channels increase with the length of the alkyl chain?
 - ii. Is there a correlation between apparent and true affinities of these compounds towards CNG channels?
- IV. Are N^6 -substituted analogs of cyclic nucleotides selective for PKA, or do they also activate cyclic nucleotide-modulated ion channels? If so, how potent are they compared to the natural ligand, and what is the structural basis for the measured apparent affinities (section 7, Publication IV)?

Here, I was mainly concerned with the question:

- i. What is the structural basis for the increased potency of N^6 -Phe-cAMP and N^6 -Bn-cAMP, and the decreased potency of N^6 -Bnz-cAMP in activating HCN2 channels compared to cAMP?

Apart from their contribution to answering the respective research questions, these studies highlight the importance of cooperation between experimental and computational research groups. Both working approaches – experimental and computational – are constantly enriched by new, innovative techniques, and the onus is on the researchers to fruitfully combine these techniques in the scope of future collaborations.

4 PUBLICATION I

Evidence for functional selectivity in TUDC- and *nor*UDCA-induced signal transduction via $\alpha_5\beta_1$ integrin towards cholestasis

§**Bonus, M.**, §Sommerfeld, A., Qvartskhava, N., Görg, B., Ludwig, B. S., Kessler, H., Gohlke, H., Häussinger, D.

§ Both authors contributed equally to this work

Original publication, see Publication I in section 11; contribution: 35% (details provided below).

(I adapted parts of the following text from the respective publication.)

4.1 Background

Functional selectivity—sometimes referred to as “agonist-directed signaling,” “biased agonism,” or “biased signaling”²⁸⁸—is the phenomenon that biologically active substances can preferentially activate one of several possible signal transduction pathways at pleiotropic receptors²⁸⁸⁻²⁹². While this feature has long been considered a characteristic of G protein-coupled receptors (GPCRs)²⁹², several studies suggest that other receptor types such as nuclear hormone receptors²⁹³⁻²⁹⁵, receptor tyrosine kinases (RTKs)²⁹⁶, and integrins^{297,298} may also engage in ligand-dependent biased signaling. A notable example among the adhesion receptors of the integrin class is the αI domain-containing integrin $\alpha_M\beta_2$ (Macrophage-1 antigen, Mac-1). Integrin $\alpha_M\beta_2$ induces activation of NF- κ B^{299,300} and downregulation of Foxp1^{301,302} when bound to fibronectin, but not when bound to CD40L²⁹⁷. However, for integrins lacking an αI domain, no evidence for an engagement in biased signaling has been found so far.

$\alpha_5\beta_1$ integrin is the predominant integrin isoform in the liver¹⁸⁰ and lacks an αI domain^{166,303}. Upon activation by the conjugated bile acid tauroursodeoxycholic acid (TUDC)¹⁸³, $\alpha_5\beta_1$ integrins stimulate cholestasis via FAK/c-Src/MAPK-dependent signaling pathways towards Bsep und Mrp2 insertion into the canalicular membrane^{181,183,304,305}. Modifying the $3\alpha,7\beta$ -dihydroxycholane (“ursodeoxycholane”) scaffold of TUDC either leads to inactive compounds (tauroolithocholic acid 3-sulfate (TLCAS), taurochenodeoxycholic acid (TCDC), glycochenodeoxycholic acid (GCDC)) or inhibitors of TUDC-induced $\alpha_5\beta_1$ integrin activation (taurocholic acid (TC))¹⁸³. Whether and how a modification of the side chain of TUDC alone, while retaining its ursodeoxycholane scaffold, alters activity towards $\alpha_5\beta_1$ integrin had not been studied so far.

Here we set out to computationally predict and experimentally validate whether and to what extent side chain-modified derivatives of TUDC, in particular 24-*nor*-ursodeoxycholic acid (*nor*UDCA), Tauro-*nor*UDCA (*Tnor*UDCA), glyoursodeoxycholic acid (GUDC), and ursodeoxycholic acid (UDCA) can directly activate $\alpha_5\beta_1$ integrins. To probe for functional selectivity among the active molecules for specific $\alpha_5\beta_1$ integrin-mediated signal transduction pathways, we determined differences in the signaling events following $\alpha_5\beta_1$ integrin activation by TUDC and *nor*UDCA.

4.2 Results

To test whether *nor*UDCA, *Tnor*UDCA, GUDCA, and UDCA can activate $\alpha_5\beta_1$ integrin, I performed three replicate unbiased explicit-solvent MD simulations for each bile acid- $\alpha_5\beta_1$ integrin complex and analyzed these simulations based on three geometric descriptors indicative of conformational changes that occur during integrin activation. Also, I carried out MD simulations of TUDC- and TC-bound $\alpha_5\beta_1$ integrin as reference active and inactive compounds to derive a relative ranking among the bile acids studied. The results of my simulations stimulated subsequent *in vitro* experiments in which my collaborators[†] studied the effect of these bile acids on the activation of $\alpha_5\beta_1$ integrin and the associated downstream signal transduction pathways. I interpreted the results from these experiments regarding the differences between TUDC- and *nor*UDCA-induced signal transduction. Taken together, we demonstrated that *nor*UDCA directly activates $\alpha_5\beta_1$ integrins in hepatocytes and found evidence for functional selectivity of TUDC and *nor*UDCA for overlapping but distinct integrin-mediated signal transduction pathways.

In detail, I assessed the activity of the investigated bile acids on a data set of $3 \times 6 \times 200$ ns of MD simulations based on three geometric descriptors derived from crystal structures of the closed (PDB: 3FCU) and open (PDB: 3FCS) $\alpha_{IIb}\beta_3$ integrin headpiece^{306,307} and previous simulation results¹⁸³: the $\alpha 1$ kink angle, the $\alpha 7$ tilt angle, and the propeller- βA distance (**Figure 12A, B**). To trace changes in the curvature of the $\alpha 1$ helix in the βA domain, I defined the $\alpha 1$ kink angle as the angle between the centers of mass of its three eight-residue segments (**Figure 12B**). Upon ligand-induced activation of $\alpha_{IIb}\beta_3$ integrins, the central segment of the $\alpha 1$ helix moves downward^{306,307}, which leads to its straightening and, therefore, an increase in the $\alpha 1$ kink angle. These conformational changes in the $\alpha 1$ helix are

[†] Dr. Annika Sommerfeld, Dr. Boris Görg, Dr. Natalia Qvartskhava, Prof. Dr. Dieter Häussinger; Clinic for Gastroenterology, Hepatology and Infectious Diseases, Heinrich Heine University Düsseldorf, Düsseldorf, Germany; Beatrice Stefanie Ludwig, Prof. Dr. Dr. Horst Kessler; Institute for Advanced Study and Center for Integrated Protein Science, Department of Chemistry, Technical University of Munich, Garching, Germany.

accompanied by a lateral pivoting movement of the $\alpha 7$ helix in the βA domain³⁰⁶. I traced this movement by the $\alpha 7$ tilt angle, defined as the angle formed by the metal ion in the ADMIDAS (“adjacent to the MIDAS”; a conserved metal ion-binding site in integrins connecting the $\beta 6$ - $\alpha 7$ loop to the $\alpha 1$ helix) and the centers of mass of the first and the last four-residue segments of the $\alpha 7$ helix (**Figure 12B**). Lastly, since previous MD simulations have shown that bile acids that activate $\alpha_5\beta_1$ integrin also lead to an increased distance between the propeller domain in the α_5 subunit and the βA domain in the β_1 subunit¹⁸³, I included the distance between these domains (“propeller- βA distance”) as an additional descriptor for integrin activation (**Figure 12A**). Note that these descriptors represent only minor conformational changes in the βA domain of $\alpha_5\beta_1$ integrin. These precede full activation, characterized by more substantial conformational changes^{178,307-309}, which I, however, did not expect to observe in my sub- μ s long MD simulations when compared to integrin activation times *in vivo*^{157,310}. Although the crystal structure of the activated (open) $\alpha_5\beta_1$ integrin headpiece has not been determined yet, the conformational changes involved in $\alpha_5\beta_1$ integrin activation are likely similar to those observed for other integrin subtypes^{178,311}. Therefore, my selection of the conformational changes represented by the geometric descriptors can be regarded as the most appropriate knowledge-based approximation of a structure-based description of integrin activation in MD simulations.

TUDC- and *nor*UDCA-bound $\alpha_5\beta_1$ showed significantly higher average values for all three geometric descriptors (**Figure 12C**). In particular, TUDC led to an $\alpha 1$ kink angle of $147.3 \pm 0.1^\circ$, an $\alpha 7$ tilt angle of $138.2 \pm 0.1^\circ$, and a propeller- βA distance of $36.68 \pm 0.01 \text{ \AA}$, whereas *nor*UDCA induced $\alpha 1$ kink and $\alpha 7$ tilt angles of similar magnitude ($149.2 \pm 0.1^\circ$ and $138.4 \pm 0.1^\circ$, respectively) but a $\sim 0.6 \text{ \AA}$ smaller propeller- βA distance (**Figure 12C**). In comparison with *nor*UDCA-bound $\alpha_5\beta_1$ integrin ectodomains, *Tnor*UDCA and GUDC-bound ectodomains showed $\alpha 1$ kink angles decreased by $\sim 6^\circ$, while the $\alpha 7$ tilt angle remained in the same range, being $\sim 2^\circ$ smaller. The propeller- βA distance was further reduced by $\sim 0.4 \text{ \AA}$ when *Tnor*UDCA or GUDC were bound (**Figure 12C**). Finally, for the UDCA-bound ectodomain, the $\alpha 1$ kink angle was $139.6 \pm 0.1^\circ$, the $\alpha 7$ tilt angle was $129.6 \pm 0.1^\circ$, and the propeller- βA distance was $35.80 \pm 0.01 \text{ \AA}$ (**Figure 12C**). These values were more similar to those of the inhibitory TC ($\alpha 1_{TC} = 142.0 \pm 0.1^\circ$, $\alpha 7_{TC} = 126.6 \pm 0.1^\circ$, $p\text{-}\beta A_{TC} = 35.77 \pm 0.01 \text{ \AA}$, **Figure 12C**) than to those of any of the other bile acids studied, indicating that UDCA, like TC¹⁸³, cannot activate $\alpha_5\beta_1$ integrins, or might even prevent TUDC-induced integrin activation. Based on this data, I classified TUDC and *nor*UDCA as

highly active, *TnorUDCA* and *GUDC* as weakly active, and *UDCA* and *TC* as inactive or inhibitory ligands, respectively.

To assess the quality of my MD-based activity predictions, I tested for a correlation between the mean values of the three geometric descriptors measured in each triplet of MD simulations and an activity ranking of the bile acids. I derived this ranking from the amount of immunostained, active β_1 integrin determined at fixed time points after liver perfusion with different concentrations of the respective bile acid (**Figure 12D**, cf. Fig. 3 in Publication I for the complete time- and concentration series). Accordingly, *TUDC* exhibited the highest activity towards $\alpha_5\beta_1$ integrin, followed by *norUDCA*, *TnorUDCA*, *GUDC*, *UDCA*, and *TC*. I detected significant linear correlations between the average α_1 kink angle ($R^2 = 0.66$, $p = 0.05$), or the α_7 tilt angle ($R^2 = 0.83$, $p = 0.01$), and the bile acid's position in the ranking (**Figure 12C**). Thus, the set of geometric descriptors I used to predict the activity of the bile acids in my MD simulations was not only able to differentiate between active and inactive bile acids but could also detect more subtle differences in their activities. Similar MD simulations could serve as “computational assays” in further studies to screen candidate molecules for their potential to activate $\alpha_5\beta_1$ integrin.

As predicted by my MD simulations, *norUDCA* caused a dose-dependent activation of $\alpha_5\beta_1$ integrins in hepatocytes (**Figure 12D**). Compared to *TUDC*-induced activation¹⁸³, however, *norUDCA*-induced activation was delayed and less pronounced (**Figure 13A**): While after addition of *TUDC*, active β_1 integrin became clearly visible within 1 min, *norUDCA* reached a similar magnitude of activation only after 15 min (**Figure 13A, B**). A competitive ELISA-based solid-phase assay³¹² confirmed direct binding of *TUDC* and *norUDCA* to the MIDAS in $\alpha_5\beta_1$ integrins and revealed similar binding affinities of both compounds (**Table 1**). Given the latter, I could rule out that the different extent of $\alpha_5\beta_1$ integrin activation by the two bile acids was a consequence of different binding site occupancy.

Table 1: Affinities of *TUDC*, *norUDCA*, and a control peptide towards $\alpha_5\beta_1$ integrin.

Compound	IC ₅₀ [*]
<i>TUDC</i>	4.01 (1.80 ... 8.98) ^a
<i>norUDCA</i>	3.93 (1.34 ... 11.5) ^a
Cilengitide (cyclo(RGD-D-Phe-NMe-V)) ³¹³	15.4 (14.49 ... 16.36) ^b

^{*} IC₅₀ values were obtained from an ELISA-like solid-phase binding assay³¹² after sigmoidal fitting to two independent data rows (serial dilutions). The 95% confidence interval is given in brackets. ^aIn mM. ^bIn nM. (Experiments were performed by Beatrice Stefanie Ludwig under supervision of Prof. Dr. Dr. Horst Kessler, Institute for Advanced Study and Center for Integrated Protein Science, Department of Chemistry, Technical University of Munich, Garching, Germany; Table and caption were modified from Publication I).

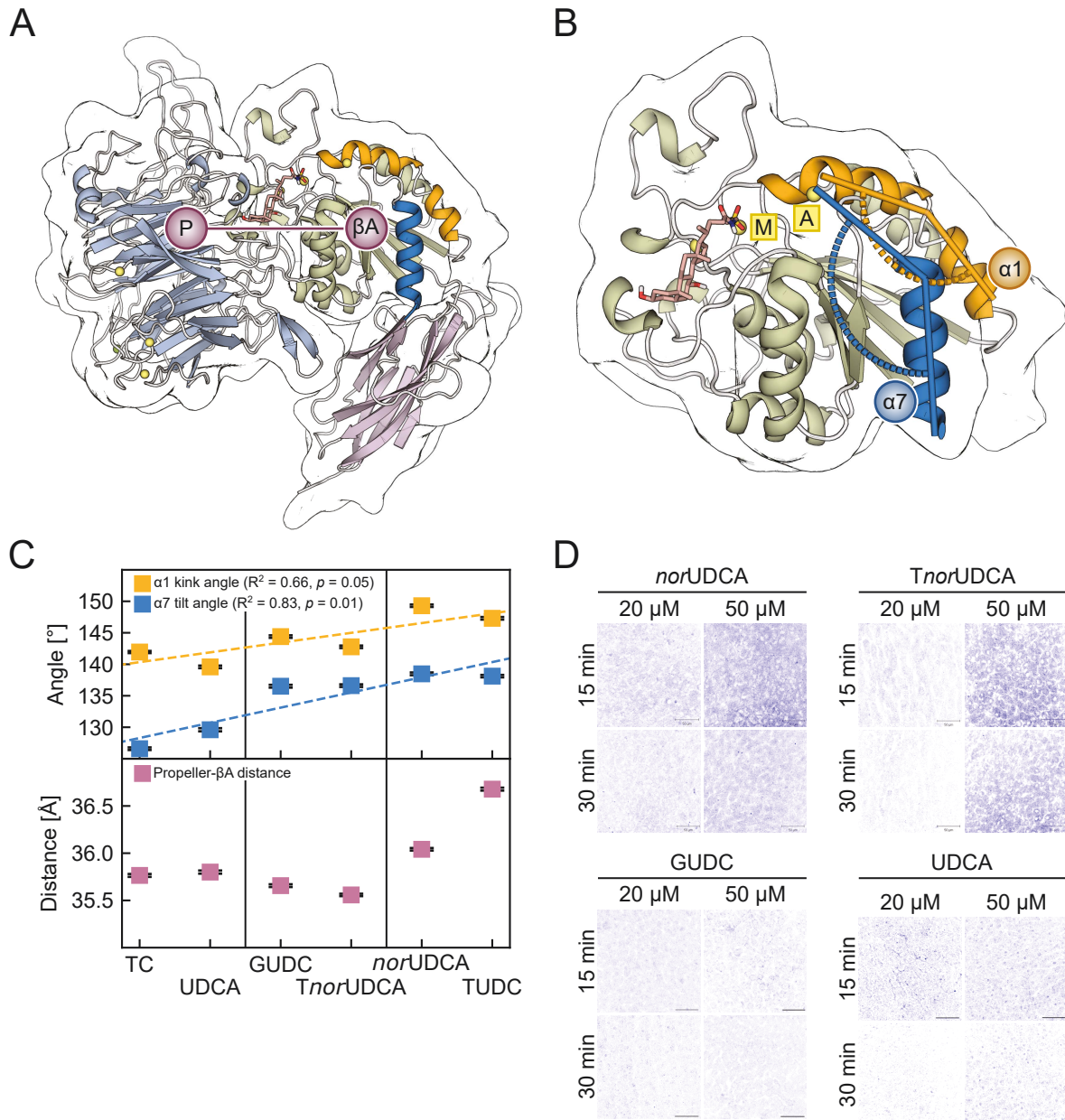


Figure 12: Geometric descriptors and activation of $\alpha_5\beta_1$ integrin in MD simulations and immunofluorescence staining experiments.

(A) Part of the $\alpha_5\beta_1$ integrin headpiece in cartoon representation. Helices $\alpha 1$ and $\alpha 7$ are highlighted in orange and blue, respectively. The propeller- βA distance was measured between the respective centers of mass (violet circles). The propeller domain is colored in light blue, the βA domain in pale green, and the hybrid domain in violet. (B) Close-up view of the βA domain with the docked TUDC structure (stick representation)¹⁸³. This complex structure was used to generate the other starting structures by modifying the bile acid side chain. Angles measured during the MD simulations: orange: $\alpha 1$ kink angle; blue: $\alpha 7$ tilt angle. Mg^{2+} ions are depicted as yellow spheres; the metal ion at the MIDAS is labeled *M*, the metal ion at the ADMIDAS is labeled *A*. (C) Average \pm SEM values of the $\alpha 1$ kink angle (orange), $\alpha 7$ tilt angle (blue), and propeller- βA distance (violet) over three replicate MD simulations versus the rank of the bile acids according to their agonistic activity towards $\alpha_5\beta_1$ integrin as observed in immunofluorescence staining experiments (panel D). Dashed lines depict linear regressions; fit parameters are given in the figures. Vertical lines separate the dataset into inactive (left), weakly active (middle), and highly active (right) bile acids. (D) Effect of *nor*UDCA, *Tnor*UDCA, GUDC, and UDCA on β_1 integrin activation. Rat livers were perfused with the respective bile acids for up to 60 min with the concentrations indicated. Liver samples were immunostained for the active conformation of β_1 integrin (blue, false-color images; for original images cf. Fig. 3 in Publication I). Representative images of at least three independent experiments are depicted for two selected time points and concentrations (for the complete time- and concentration series, cf. Fig. 3 in Publication I). (Figure and caption were modified from Publication I).

As described before^{304,314}, TUDC-induced $\alpha_5\beta_1$ integrin activation led to *sustained* activation of Erk-1/-2 and p38^{MAPK} (**Figure 13C**), which is the critical downstream signal towards choleresis. This kinetic profile was preserved at lower (10 μ M) and higher (50 μ M) concentrations of TUDC. Therefore, I could rule out a concentration effect. In contrast, *norUDCA*-induced $\alpha_5\beta_1$ integrin activation only led to *transient* activation of these MAPKs (**Figure 13C**), which may have been a consequence of weaker $\alpha_5\beta_1$ integrin activation. Additionally, perfusion with TUDC, but not with *norUDCA*, led to a significant EGFR/c-Src association (Publication I – Fig. 7, Supplementary Fig. 11), suggesting a c-Src-dependent trans-activation of EGFR as a determining factor for prolonged MAPK activation. Inhibiting PI3-K abolished *norUDCA*-induced phosphorylation of only Erk-1/-2, but inhibiting c-Src abolished phosphorylation of both Erk-1/-2 and p38^{MAPK}, suggesting that c-Src activation occurs upstream of PI3-K activation. In an earlier study¹⁸¹, which investigated TUDC-induced signaling towards choleresis, Inhibition of c-Src did not prevent Erk-1/-2 phosphorylation, but only delayed it by ~8 min. Hence, inhibition of c-Src activity only prevents Erk-1/-2 phosphorylation after *norUDCA*-induced, but not after TUDC-induced $\alpha_5\beta_1$ integrin activation.

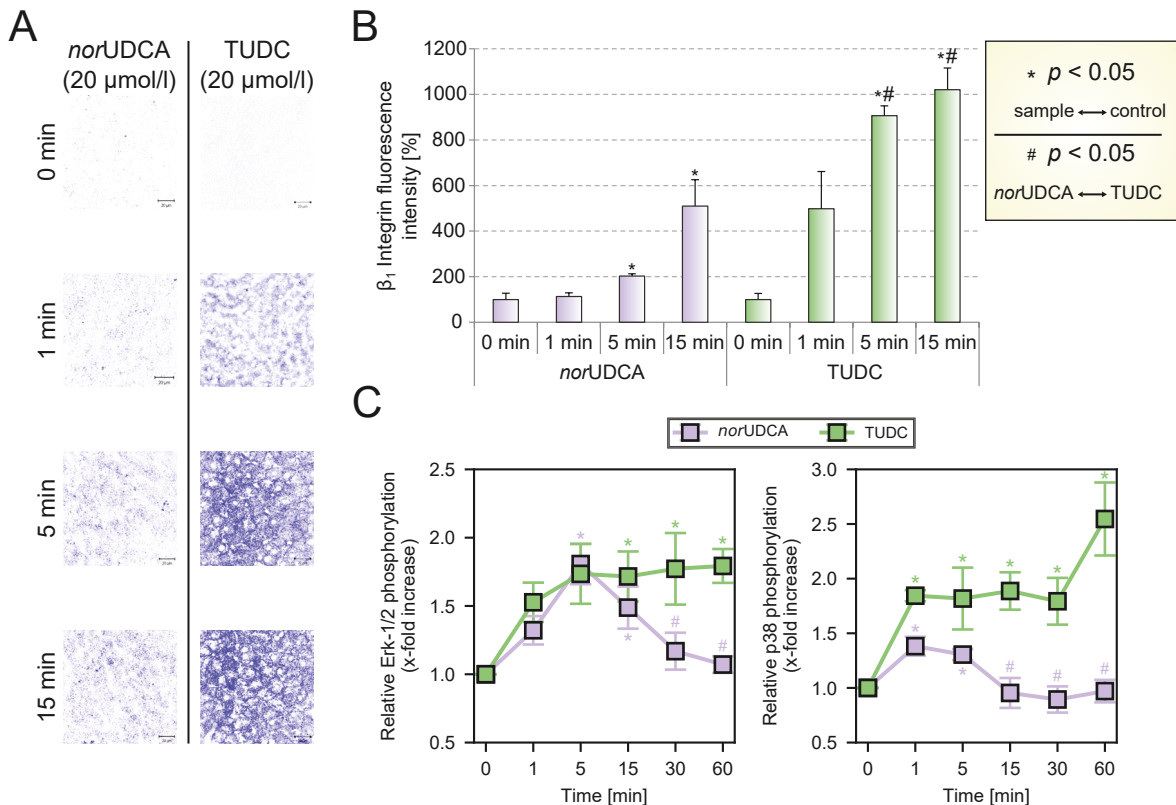


Figure 13: Comparison between TUDC- and *norUDCA*-induced $\alpha_5\beta_1$ integrin activation and MAPK signaling.

(A) Immunofluorescence staining of active β_1 integrin after liver perfusion with either 20 μM *norUDCA* (left) or TUDC (right) at timepoints $t = 0, 1, 5,$ and 15 min. The scale bar corresponds to 20 μm . Representative images of three independent experiments are depicted. (B) Quantification of β_1 integrin fluorescence. (C) Densitometric quantification of Erk-1/-2 (left) and p38^{MAPK} (right) phosphorylation after liver perfusion with either 20 μM *norUDCA* (violet) or TUDC (green). Total Erk-1/-2 and total p38^{MAPK} served as respective loading controls. Phosphorylation at $t = 0$ min was set to 1. Data represents the mean \pm SEM of at least three independent experiments. (Figure and caption were modified from Publication I).

Based on these and further experiments (cf. Publication I)[‡], I proposed a ligand-dependent selectivity for signaling pathways induced by $\alpha_5\beta_1$ integrin (Figure 14): After binding of TUDC¹⁸¹ or *norUDCA*, $\alpha_5\beta_1$ integrins recruit focal adhesion kinase (FAK)¹⁶⁴, which subsequently undergoes autophosphorylation. Since levels of autophosphorylated FAK (FAK^{Y397-P}) increase linearly with the amount of fibronectin-bound (i.e., active and signaling-competent) $\alpha_5\beta_1$ ³¹⁵, a highly efficacious integrin activation as observed with TUDC¹⁸³ (Figure 14A) would result in high FAK^{Y397-P} levels. In contrast, a less efficacious integrin activation—as observed after stimulation with *norUDCA* (Figure 14B)—would result in lower FAK^{Y397-P} levels. FAK^{Y397-P} activates c-Src^{165,316}, which in turn

[‡] These experiments were performed by Dr. Boris Görg and Dr. Natalia Qvartskhava under supervision of Prof. Dr. Dieter Häussinger; Clinic for Gastroenterology, Hepatology and Infectious Diseases, Heinrich Heine University Düsseldorf, Düsseldorf, Germany.

phosphorylates EGFR, and both the activated c-Src and EGFR mediate PI3-K activation^{316,317} and subsequent phosphorylation of Erk-1/-2. However, FAK^{Y397-P} can also *directly* activate PI3-K, independent of c-Src and the EGFR³¹⁶. This direct, FAK-mediated activation of PI3-K may be slower than the c-Src and EGFR-mediated PI3-K activation and may only be triggered by high FAK^{Y397-P} levels. Hence, even when c-Src activity is inhibited by PP-2, a highly efficacious integrin activation by TUDC would lead to a pronounced FAK autophosphorylation and rescue Erk-1/-2 phosphorylation via a direct PI3-K activation, albeit with a time delay, as observed previously¹⁸¹. In contrast, a less efficacious integrin activation by *nor*UDCA would lead to less FAK autophosphorylation and would therefore require activated c-Src to switch on the then necessary PI3-K signal to activate Erk-1/-2, which would occur more rapidly.

Taken together, these results demonstrate that *nor*UDCA directly activates $\alpha_5\beta_1$ integrins in hepatocytes and triggers short-term choleresis via a transient activation of MAPKs, followed by a transient insertion of Bsep into the canalicular membrane. They also provide evidence that TUDC and *nor*UDCA exert a functional selectivity for specific $\alpha_5\beta_1$ -mediated signal transduction pathways. This property has – to the best of my knowledge – not yet been described for ligands interacting with integrins without an αI domain²⁹⁷.

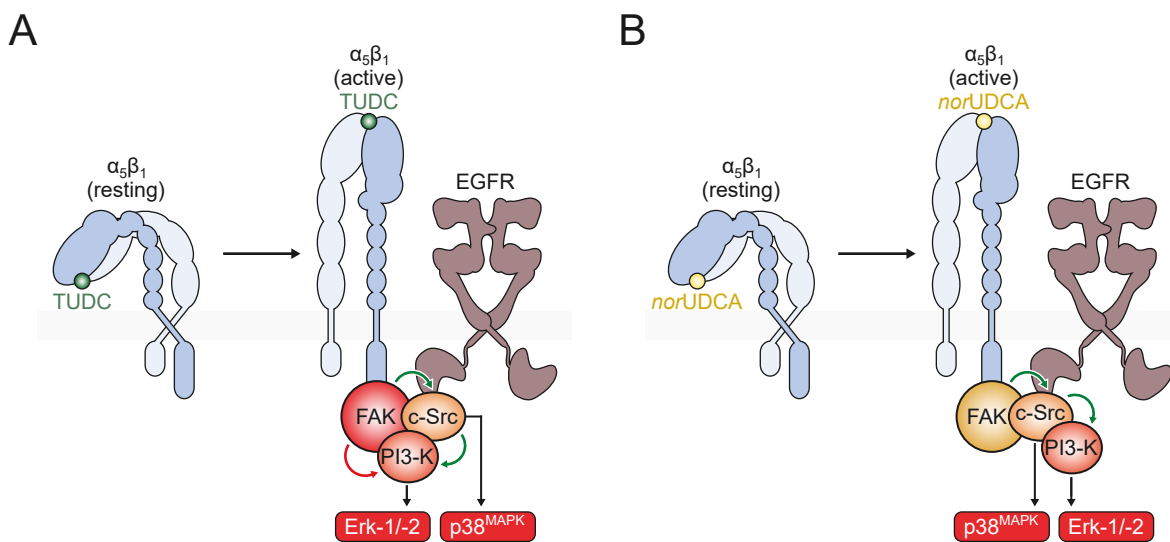


Figure 14: Model of $\alpha_5\beta_1$ integrin activation-dependent differential bile acid signaling.

(A) Activation of $\alpha_5\beta_1$ integrin by TUDC leads to high FAK^{Y397-P} levels, which trigger both a slow, c-Src and EGFR-independent PI3-K activation (red arrow) and a fast, c-Src and EGFR-dependent PI3-K activation (green arrows). The subsequent PI3-K-mediated Erk-1/-2 activation is sustained. (B) Activation of $\alpha_5\beta_1$ integrin by *nor*UDCA leads to lower FAK^{Y397-P} levels, which trigger the fast, c-Src and EGFR-dependent pathway of PI3-K activation (green arrows), but abolish the slower FAK-mediated PI3-K activation. The subsequent PI3-K mediated Erk-1/-2 activation is only transient. (Figure and caption were modified from Publication I).

4.3 Conclusion and significance

With the work described above, I have made a substantial contribution to a collaborative study. I computationally predicted the activity of a set of side chain-modified derivatives of TUDC on $\alpha_5\beta_1$ integrin. My predictions were validated *in vitro* and stimulated further experiments during which we detected differences in TUDC- and *nor*UDCA-induced signaling events downstream of integrin activation. In interpreting these differences, I detected evidence for biased signaling in α I-less integrins, which is – to the best of my knowledge – described for the first time in this study.

The key results and findings of this study are:

- MD-based activity predictions of *nor*UDCA, *Tnor*UDCA, GUDC, and UDCA on $\alpha_5\beta_1$ integrin using knowledge-based geometric descriptors were in excellent agreement with experimentally determined immunofluorescence intensities after liver perfusion with the studied compounds. The approach employed may serve as a basis for further studies to pre-screen a small set of candidate molecules for their potential to activate $\alpha_5\beta_1$ integrin.
- *nor*UDCA is the second identified agonist of $\alpha_5\beta_1$ integrin with a $3\alpha,7\beta$ -dihydroxycholane scaffold. This finding confirms the function of $\alpha_5\beta_1$ integrin as a bile acid sensor¹⁸³ and emphasizes its role as the predominant integrin isoform in the liver¹⁸⁰.
- Dissecting TUDC- and *nor*UDCA-induced integrin signaling revealed evidence for functional selectivity of these two ligands for overlapping but distinct $\alpha_5\beta_1$ integrin-mediated signaling pathways. This feature is described for the first time for integrins without an α I domain.

The insights gained from this study also provide a rationale for the different therapeutic applications of UDCA and *nor*UDCA in primary biliary cholangitis and primary sclerosing cholangitis, respectively¹⁸⁴.

5 PUBLICATION II

Isoform-specific Inhibition of *N*-methyl-D-aspartate Receptors by Bile Salts

§Koch A., §**Bonus, M.**, Gohlke, H., Klöcker, N.

§ Both authors contributed equally to this work

Original publication, see Publication II in section 11; contribution: 30% (details provided below).

(I adapted parts of the following text from the respective publication.)

5.1 Background

Glutamate is the predominant excitatory neurotransmitter in the mammalian central nervous system (CNS)³¹⁸ and exerts its functions via activation of ionotropic glutamate receptors (iGluRs) and metabotropic glutamate receptors (mGluRs), respectively³¹⁹. NMDA receptors (**Figure 15A**) are a subfamily of iGluRs^{187,320} and composed of heterotetrameric subunit assemblies (**Figure 15B**) of two obligatory, glycine- or D-serine-binding³²¹ GluN1 subunits and two regulatory subunits of either GluN2 (GluN2A-D, which bind glutamate³²²) or GluN3 (GluN3A-B, which bind glycine or D-serine³²²) subtype^{192,323}. Neuronal NMDARs are extensively characterized (for an overview, cf. ref. ¹⁸⁶). However, a broad spectrum of non-neuronal cells¹⁸⁶, such as osteoblasts^{219,220} and osteoclasts²²⁰⁻²²², pancreatic beta cells^{223,224}, a variety of kidney cell types²²⁵⁻²²⁷, lymphocytes^{228,229} and platelets²³⁰ also express NMDARs. Since such non-neuronal NMDARs are subject to comparatively high plasma glutamate concentrations, alternative agonists such as L-homocysteic acid and quinolinic acid are discussed¹⁸⁶. Moreover, the literature describes an extensive repertoire of endogenous, allosteric modulators of NMDARs: divalent and monovalent ions, protons, polyamines, and neurosteroids can fine-tune normal receptor function by positive or negative allosteric modulation¹⁸⁵.

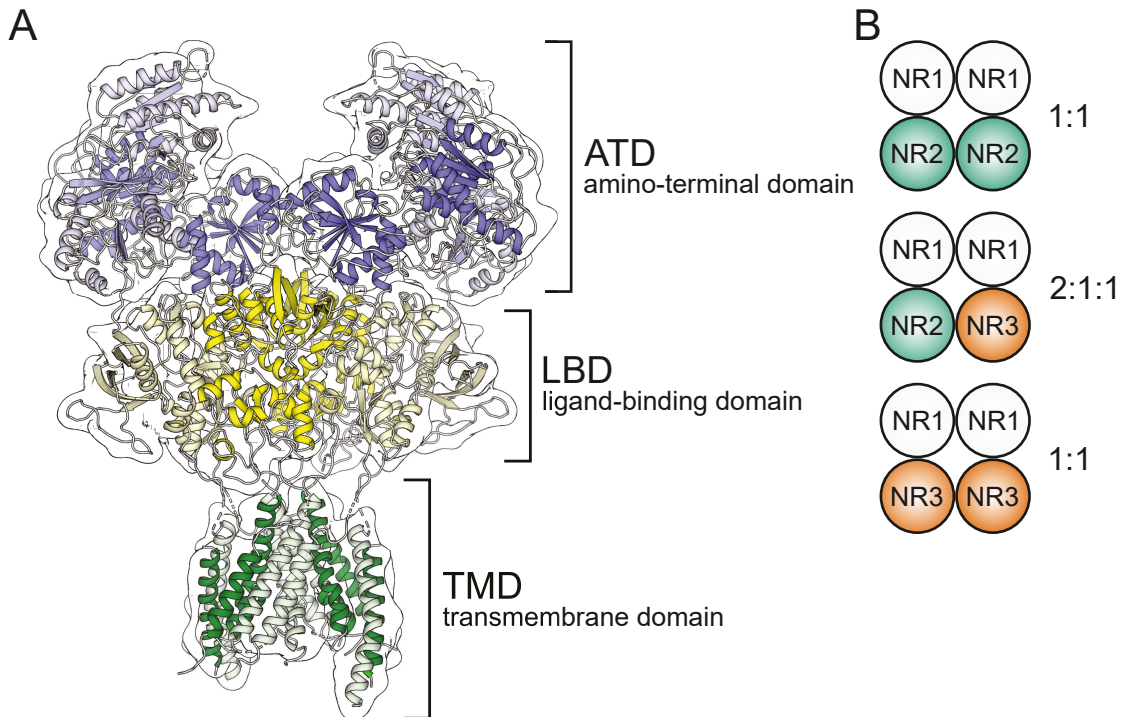


Figure 15: NMDAR structure and subunit assemblies.

(A) Structure of the human GluN1/GluN2A NMDAR in the glutamate/glycine-bound state (PDB ID: 6IRA)³²⁴. Lighter colors represent GluN1 subunits, darker colors GluN2A subunits. The amino-terminal domain (ATD) is colored in blue shades, the ligand-binding domain (LBD) in yellow shades, and the transmembrane domain (TMD) in green shades. (B) Possible NMDAR subunit assemblies: NMDARs comprise two GluN1 (labeled NR1) subunits and either two GluN2 (labeled NR2) or GluN3 (labeled NR3) subunits (1:1 assemblies) or one GluN2 and one GluN3 subunit (2:1:1 assembly). Note that it is still unclear whether the GluN1/GluN2/GluN3 2:1:1 heterotetramer is of biological relevance¹⁹².

Given the chemical similarity between neurosteroids—which can potentiate or inhibit NMDAR activity in an isoform-dependent manner^{185,231,232}—and bile acids, it is reasonable to hypothesize a modulatory function on non-neuronal NMDARs for the latter. Experiments in which cholate and chenodeoxycholate (CDC) blocked NMDARs after being applied to neurons provide evidence for such a function²³³. The observation that early treatment with CDC can prevent neurological complications of cerebrotendinous xanthomatosis – a condition resulting from defective bile acid synthesis^{234,325,326} – further substantiates this hypothesis. In cholestatic liver diseases like PFIC and ICP, or acute hepatitis, plasma concentrations of bile salts may rise to 100 μM and more²³⁵⁻²³⁷. Therefore, modulation of the activity of non-neuronal NMDARs by bile salts could be of physiological and pathophysiological importance.

In this study, combining computational, molecular biological, electrophysiological, and pharmacological techniques[§], we studied the effects of the most common human bile salts – cholate and CDC – on NMDARs with different subunit compositions.

5.2 Results

Electrophysiological and pharmacological experiments showed that both unconjugated and conjugated CDC predominantly inhibit currents of NMDARs that carry GluN subtypes 2D or 3B, rather than 2A or 2B. In addition to this selectivity towards subtypes 2D and 3B, we detected a mechanistic specificity within these subtypes: tauro-CDC inhibited GluN2D-carrying receptors allosterically, but GluN3B-carrying receptors competitively. Tauro-CDC-mediated inhibition of currents was also independent of the agonist used. It did not alter ion selectivity, rendering indirect effects of the bile acid, such as changes in the geometry of the channel pore or the selectivity filter, unlikely. To explain the isoform-specific effect of tauro-CDC on GluN2D- and GluN3B-carrying NMDARs on a structural basis, I conducted docking experiments followed by MD simulations and effective binding energy (ΔG_{eff}) calculations. My results supported the experimental findings and offered a plausible explanation for competitive inhibition of GluN3B-carrying NMDARs by tauro-CDC. Moreover, I identified an allosteric binding site in the GluN1/GluN2D LBD dimer interface that accounted for the allosteric inhibition of GluN2D-carrying NMDARs by tauro-CDC and was in excellent agreement with known binding sites of negative allosteric modulators of GluN1/GluN2A-bearing receptors²¹². In summary, our study provided new molecular insights into the modulation of NMDARs by endogenous steroids and suggested a possible pathophysiological role of these receptors in cholestatic liver disease. From a long-term perspective, it provides a foundation for a pharmacological approach to develop targeted interference strategies that exploit isoform specificity.

Since the results of the electrophysiological and pharmacological experiments indicated that tauro-CDC acts as an allosteric antagonist on GluN2D-carrying NMDARs but as a competitive antagonist on GluN3B-carrying NMDARs, I intended to elucidate the molecular mechanism behind this isoform-specific inhibition using computational methods. To this end, I first conducted induced-fit docking experiments³²⁷ with tauro-CDC on homology models of isolated LBDs of the two NMDAR subunits in question (rat GluN2D_{LBD} and

[§] Molecular biological, electrophysiological, and pharmacological experiments were performed by Dr. Angela Koch under supervision of Prof. Dr. Nikolaj Klöcker; Institute of Neural and Sensory Physiology, Medical Faculty, Heinrich Heine University Düsseldorf, Düsseldorf, Germany.

mouse GluN3B_{LBD}, (**Figure 16A, B**). For detailed protocols regarding the construction of the homology models and the induced fit docking procedure, cf. the “Materials and Methods” section in Publication II. As the experimental findings indicated no or much weaker binding of tauro-CDC to the orthosteric binding site of GluN2D compared to GluN3B, I expected the docking experiments to reveal a similar tendency. Surprisingly, the predicted binding modes were not only geometrically similar (in-place root-mean-square deviation (RMSD) of all tauro-CDC atoms of 2.98 Å after superimposition of the C α atoms of the protein structures (RMSD: 1.45 Å)), but both complexes displayed similar docking scores (IFDScore³²⁷: ~7.4 kcal mol⁻¹) and a similar interaction profile (**Figure 16C, D**). Hence, I concluded that the scoring function of the docking approach was not able to adequately capture the impact of the differences (**Figure 17**) in the amino acid composition between the orthosteric binding pockets of GluN2D and GluN3B on the binding affinity of TCDC.

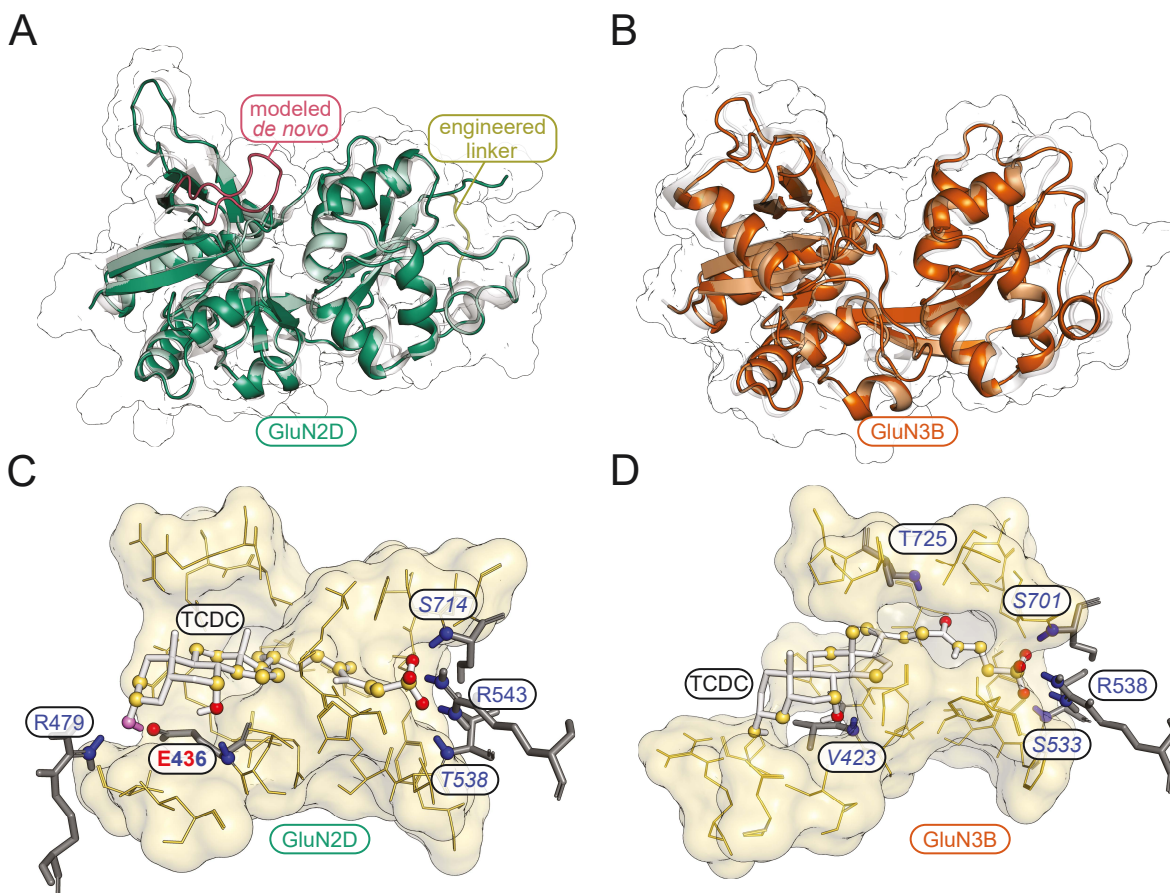


Figure 16: Homology models of GluN2D_{LBD} and GluN3B_{LBD} and docking-derived complex structures with tauro-CDC (TCDC).

(A) Homology model of GluN2D_{LBD} (green). The loop consisting of residues 464–476 (magenta, left) was modeled *de novo* due to its absence in the template structure (gray, PDB ID: 3OEK²⁰⁶). The engineered Gly-Thr linker in the template structure (yellow, right) was removed in the final model. (B) Homology model of GluN3B_{LBD} (orange) superimposed onto the template structure (gray, PDB ID: 2RCA²⁰⁵). (C) Binding mode of tauro-CDC (white sticks) in the GluN2D_{LBD} binding site (gold surface) predicted by induced fit docking. A residue is considered part of the binding site if at least one of its atoms lies within 4 Å of any atom of tauro-CDC. Ligand atoms interacting with the protein only through van der Waals contacts are depicted as gold spheres; the corresponding amino acids in the protein are depicted in gold line-and-surface representation. Amino acids that form hydrogen bonds are shown as gray sticks. The atoms involved in hydrogen bonds are colored according to their involvement as donor (blue), acceptor (red), or mixed donor-acceptor (violet) and are depicted as spheres. Amino acid labels are colored according to the involvement of the respective amino acid as donor (blue), acceptor (red), or mixed donor-acceptor (blue and red). Amino acids in which only the side chain is involved in the hydrogen bond are labeled in normal font style, those in which only the backbone is involved are labeled in italic font style, and those in which both side chain and backbone are involved are labeled in bold font style. (D) Binding mode of tauro-CDC (white sticks) in the GluN3B_{LBD} binding site (gold surface) predicted by induced fit docking. The representation is identical to the one used in panel C. (Figure and caption were modified from Publication II).

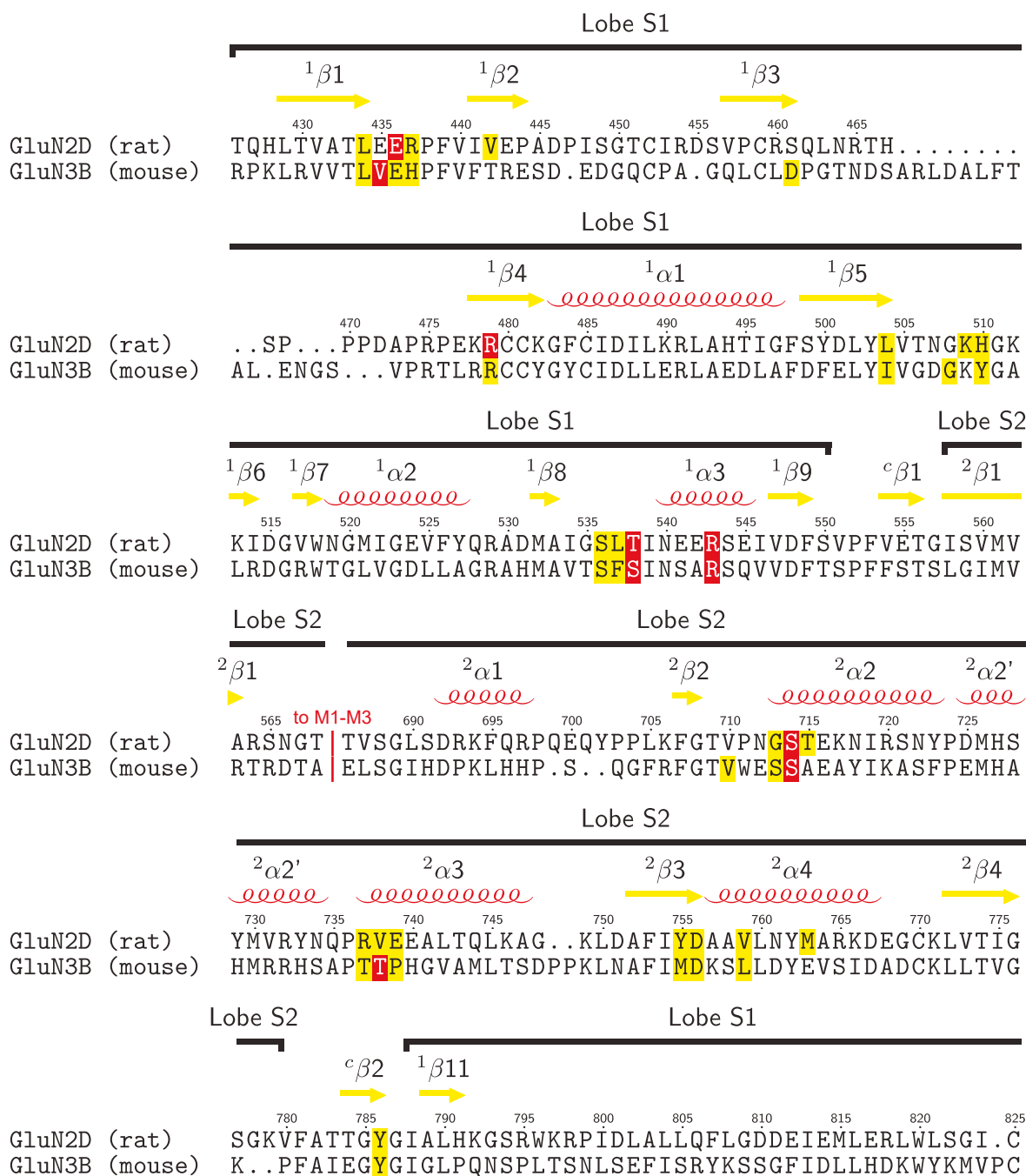


Figure 17: Sequence alignment of rat GluN2D_{LBD} and mouse GluN3B_{LBD}.

Residues that engage in hydrogen bonds with tauro-CDC are colored red, residues that form the binding site of tauro-CDC but do not engage in hydrogen bonds are colored yellow. A residue is considered part of the binding site if at least one of its atoms lies within 4 Å of any atom of tauro-CDC.

In light of the experimental results, the orthosteric binding site of GluN3B, but not the one of GluN2D, should be able to bind tauro-CDC with an affinity in the lower micromolar range. Accordingly, the binding mode of tauro-CDC found by docking should remain stable over time in GluN3B_{LBD} but become unstable in GluN2D_{LBD}. To determine the time-dependent stability of the predicted bile acid-LBD complexes, I carried out ten MD simulations (S₁-S₁₀, for a detailed simulation protocol, cf. the “Materials and Methods”

section in Publication II) of 500 ns length each for both the tauro-CDC/GluN2D_{LBD} and tauro-CDC/GluN3B_{LBD} systems—amounting to a total simulation time of 10.0 μ s—and quantified the structural deviation from the initial tauro-CDC pose within each simulation and between independent simulations. To account for alternate binding modes and metastable states traversed during binding mode transitions, I also quantified the conformational variability of tauro-CDC in and between simulations. While the average structural deviation from the initial pose may be relatively large if a transition to a more stable binding mode occurs already at the very beginning of a simulation, the conformational variability will remain small when only two binding modes are traversed during a simulation, regardless of how different these binding modes are.

Considering all simulations except GluN2D_{LBD} S₉, during which spontaneous complex dissociation occurred after \sim 225 ns, the average RMSD of the tauro-CDC pose in GluN2D was 2.86 ± 0.35 Å higher than in GluN3B. In line with this data was the higher average pairwise RMSD of the tauro-CDC pose of all snapshots of the tauro-CDC/GluN2D simulations (GluN2D_{LBD}: 8.44 ± 0.001 Å (**Figure 18A**), GluN3B_{LBD}: 4.82 ± 0.001 Å (**Figure 18B**); cf. Fig. 7a,b in Publication II). Furthermore, considering the quadratic mean of the pairwise RMSD as a measure for pose diversity³²⁸, the sampled tauro-CDC poses both within and between simulations were considerably more diverse in GluN2D_{LBD} systems than in GluN3B_{LBD} systems (GluN2D_{LBD}: 8.43 ± 0.46 Å, GluN3B_{LBD}: 4.89 ± 0.26 Å; cf. Supplementary Fig. S3 in Publication II).

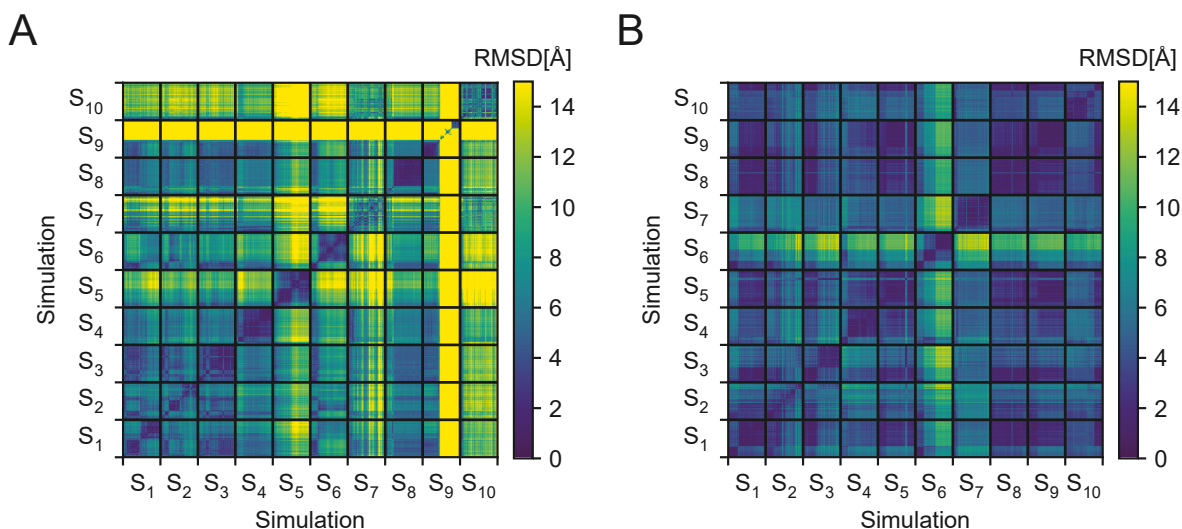


Figure 18: Structural variability and number of hydrogen bonds in the tauro-CDC-GluN2D_{LBD} and tauro-CDC-GluN3B_{LBD} complexes during the MD simulations.

(A) Two-dimensional RMSD of the atomic positions of tauro-CDC after least squares fitting of the C_α atom coordinates of the GluN2D_{LBD} in all 10 MD simulations. (B) Two-dimensional RMSD of the atomic positions of tauro-CDC after least squares fitting of the C_α atom coordinates of the GluN3B_{LBD} in all 10 MD simulations. (Figure and caption were modified from Publication II).

To determine whether the differences in the pose diversity of tauro-CDC between GluN2D and GluN3B were attributable to differences in intermolecular interactions, I analyzed the hydrogen bond interaction patterns of both complexes. In particular, I quantified hydrogen bonds formed between tauro-CDC and Arg543 in GluN2D_{LBD} and Arg538 in GluN3B_{LBD} throughout the MD simulations and found them to be about four times higher in the GluN3B complexes than in the GluN2D complexes (average counts: GluN3B_{LBD}: 1.31 ± 0.01 , GluN2D_{LBD}: 0.31 ± 0.01 , $t(9498) = 58.58$, $p < 10^{-4}$, cf. Figure 7c,d in Publication II). This arginine is crucial for binding of the agonist glutamate to GluN2D²⁰⁶ and of glycine to GluN3B²⁰⁵. Furthermore, the total number of protein-ligand hydrogen bonds was significantly higher in tauro-CDC-GluN3B complexes than in tauro-CDC-GluN2D complexes (average counts: GluN3B_{LBD}: 4.97 ± 0.03 , GluN2D_{LBD}: 3.58 ± 0.03 , $t(9498) = 38.10$, $p < 10^{-4}$, cf. Fig. 7e,f in Publication II). Among the GluN2D_{LBD} residues that form hydrogen bonds with tauro-CDC, Arg543 also exhibited the highest relative hydrogen bonding frequency with 31% of all MD snapshots (cf. Fig. 8a in Publication II). For the corresponding residue in GluN3B_{LBD} (Arg538), this value increased to 131% (cf. Fig. 8b in Publication II). When interpreting these values, note that the relative frequency may exceed values of 100% if the residue forms more than one hydrogen bond at a time. A second dominant hydrogen bond in tauro-CDC-GluN3B_{LBD} complexes (relative hydrogen bonding frequency of 78%) formed between Asp447 and the 3 α -OH group at the cholane scaffold of tauro-CDC. The structural equivalent of this residue in GluN2D_{LBD} is Ser461, which did not

form hydrogen bonds to tauro-CDC and does, therefore, not contribute to the stabilization of the complex. In addition to these residues, Ser701 (16%), Tyr505 (14%), and Ser533 (13%) were among the five residues that most frequently formed hydrogen bonds to tauro-CDC in GluN3B_{LBD} (cf. Fig. 8b in Publication II). In summary, hydrogen bond formation between tauro-CDC and Arg538 in GluN3B_{LBD} was more frequent than between tauro-CDC and Arg543 in GluN2D_{LBD}; The same applies to hydrogen bond formation to the complete LBD. Tauro-CDC/GluN3B_{LBD} complexes are stabilized mainly by Arg538 and, to a lesser extent, Asp447, whereas in tauro-CDC/GluN2D_{LBD} complexes, no amino acid contributes significantly to the stabilization of the complex. In summary, this data supports the interpretation of the pharmacological experiments that tauro-CDC inhibits GluN3B but not GluN2D in a competitive manner.

To validate the interaction of tauro-CDC with the residues of GluN3B_{LBD}, Arg538 would have been the prime residue for mutation studies due to its pronounced involvement in hydrogen bonding with the bile salt (cf. Fig. 8b in Publication II). However, we did not mutate Arg538 as this residue has been reported to be essential for agonist binding²⁰⁵. Instead, we mutated Asp447, Tyr505, and Ser701, which also form hydrogen bonds with the bile salt. Mutation of Tyr505 and Ser701, however, resulted in GluN3B-carrying receptors that could either not or hardly be activated by glycine anymore. Mutation of Asp447 to alanine or arginine changed neither the agonist-induced current amplitudes nor the extent of tauro-CDC block. This outcome, however, does not exclude an interaction of Asp447 with tauro-CDC: abolishing the interaction between Asp447 and the 3 α -OH group of tauro-CDC may adversely affect binding enthalpy but could be compensated by entropic effects³²⁹. I conjectured that the loss in binding enthalpy could be compensated both by an increase in ligand entropy due to an increase in translational and rotational degrees of freedom and by an increase in receptor entropy due to increased conformational freedom of the residues around the mutation site. Particularly the latter effect is likely significant since the Asp447Ala mutation abolishes a stabilizing interaction between Asp447 and Arg474. In the Asp447Arg mutant, the interaction between tauro-CDC and the receptor could be preserved if arginine were to act as a hydrogen bond donor for the 3 α -OH group of tauro-CDC, which in turn would act as a hydrogen bond acceptor. Repulsive forces between Arg447 in the Asp447Arg mutant and Arg474, however, could lead to a gain in receptor entropy.

Next, I complemented my analysis of the structural stability of the tauro-CDC-GluN2D and tauro-CDC-GluN3B complexes and the hydrogen bond-mediated protein-ligand interactions

within these complexes with an energetic analysis. To this end, I determined the effective binding energy (ΔG_{eff}) between tauro-CDC and GluN2D_{LBD} and GluN3B_{LBD}, respectively, using the end-point free energy calculation approaches MM-PBSA and MM-GBSA (data summarized in **Figure 19**, an explanation of the method is provided in section 2.2.2). In line with the results described previously, the effective binding energy in GluN2D_{LBD} (mean: -33.53 ± 0.67 kcal mol⁻¹) was considerably less negative than in GluN3B_{LBD} (mean: -48.90 ± 0.62 kcal mol⁻¹) ($t(17) = 16.84$, $p < 10^{-4}$) when calculated using the MM-GBSA approach. This difference remained when using the MM-PBSA approach (tauro-CDC/GluN2D_{LBD}: mean: -27.14 ± 0.52 kcal mol⁻¹; tauro-CDC/GluN3B_{LBD}: mean: -30.54 ± 0.63 kcal mol⁻¹; $t(17) = 4.07$, $p = 8 \cdot 10^{-4}$), and was robust against changing the solute dielectric constant from $\epsilon_{\text{int}} = 1$ (values above) to $\epsilon_{\text{int}} = 4$ (tauro-CDC/GluN2D_{LBD}: mean: -43.98 ± 0.24 kcal mol⁻¹; tauro-CDC/GluN3B_{LBD}: mean: -45.45 ± 0.22 kcal mol⁻¹; $t(17) = 4.49$, $p = 3 \cdot 10^{-4}$). In summary, both end-point approaches indicated stronger binding of tauro-CDC to GluN3B_{LBD}, thus supporting the experimental results, which suggest competitive inhibition of GluN3B-containing NMDARs by tauro-CDC.

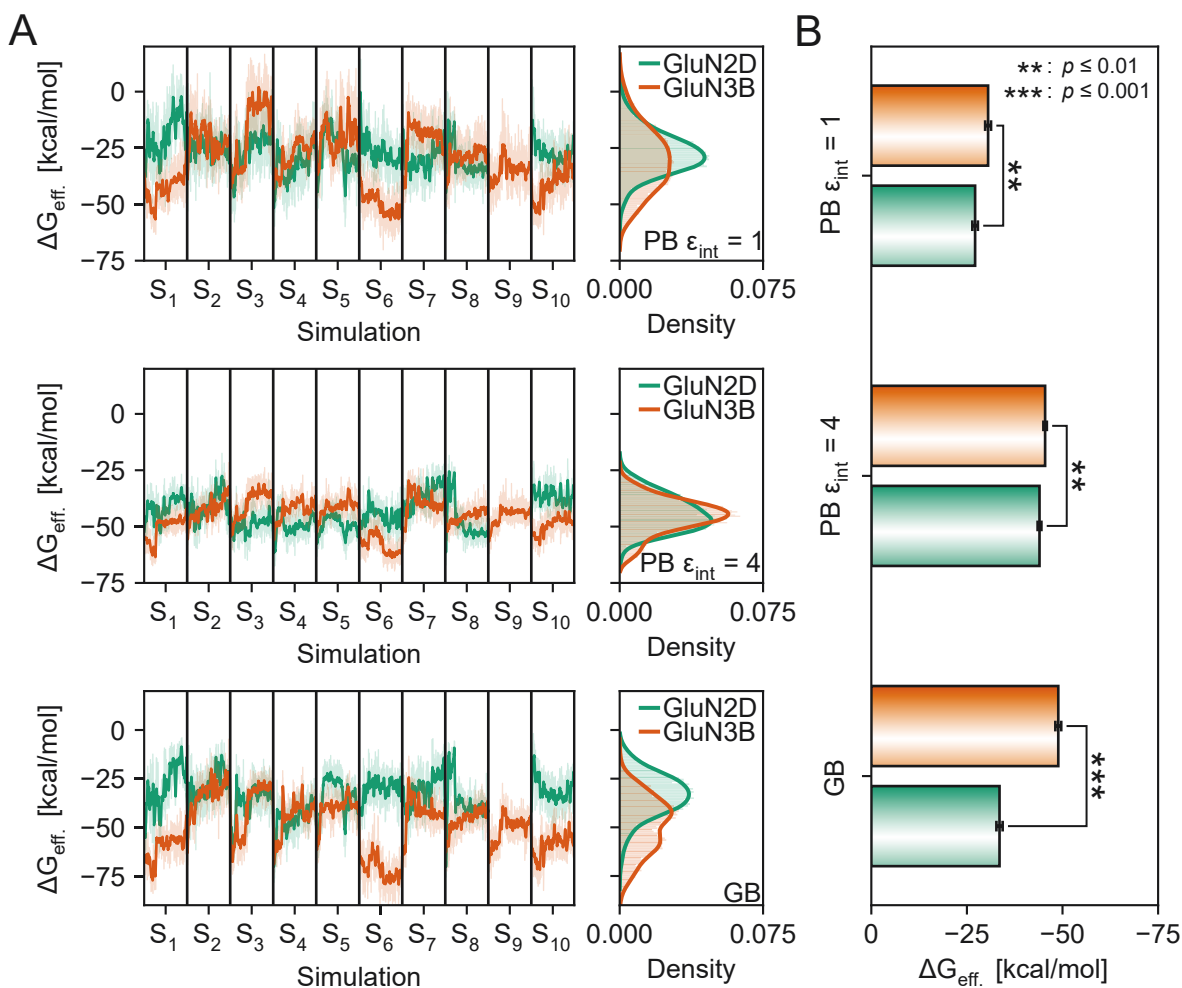


Figure 19: Effective binding energies (ΔG_{eff}) of the tauro-CDC-GluN2D_{LBD} (green) and tauro-CDC-GluN3B_{LBD} (orange) complexes.

(A) Time courses (left) and probability density distributions (right) of ΔG_{eff} , calculated using either the MM-PBSA approach and an internal dielectric constant of $\epsilon_{\text{int}} = 1$ (top row) or $\epsilon_{\text{int}} = 4$ (middle row), or the MM-GBSA approach using the modified OBC model¹¹⁹. Thick, opaque lines in the probability density distributions depict the distribution calculated with a Gaussian kernel density estimator (bandwidth calculated using Scott's rule³³⁰); thin, transparent lines show the underlying data as histograms. The data for MD simulation S₉ of GluN2D was omitted because there, tauro-CDC unbinds after 225 ns (cf. Supplementary Fig. S3b), and the corresponding data was excluded from statistical analysis. (B) Time-averaged mean ΔG_{eff} , calculated using either the MM-PBSA approach and an internal dielectric constant of $\epsilon_{\text{int}} = 1$ (top pair of bars) or $\epsilon_{\text{int}} = 4$ (middle pair of bars), or the MM-GBSA approach using the modified OBC model¹¹⁹. (Figure and caption were modified from Publication II).

After highly likely excluding competitive inhibition of the GluN2D LBD by tauro-CDC, and thereby validating the pharmacological experiments, I intended to identify the binding site that governs allosteric inhibition of GluN2D-carrying NMDARs. To this end, I first constructed a homology model of the GluN1_{LBD}/GluN2D_{LBD} tetramer (**Figure 20A**) and performed a 500 ns long accelerated MD simulation³³¹ (aMD) of this system. For details on the simulation setup, cf. the “Materials and Methods” section in Publication II. Using aMD is uniquely purposeful in this context since an increased number of energetically accessible conformations can be sampled, but an exact representation of the probability distribution of

states is of secondary importance. From the resulting ensemble, I detected potential binding pockets with the MDpocket³³² software. Using this approach, I identified a single binding pocket in the GluN1_{LBD}/GluN2D_{LBD} interface and docked tauro-CDC into the receptor conformation in which this binding pocket exhibited the highest mean local hydrophobic density^{332,333} (**Figure 20B**). The identified allosteric binding pocket (**Figure 20C**) strongly overlaps with the binding pocket of the negative allosteric modulator MPX-007²¹² (**Figure 20D**) and a set of positive allosteric modulators²¹⁵ in the GluN1/GluN2A LBD dimer interface.

During subsequent 500 ns of conventional MD, the predicted binding mode of tauro-CDC in the putative allosteric pocket exhibited low conformational variation (RMSD: 1.90 ± 0.01 Å; 2.17 ± 0.01 Å without fitting, calculated on the last 250 ns of the simulation ensemble). Interactions between the sulfonic acid moiety and Lys531, Lys769, and Ser773 in the GluN1 subunit, as well as Lys797 in the GluN2D subunit, stabilize the tauro-CDC pose (**Figure 21**); Leu808 in the GluN2D subunit (**Figure 21A**) exhibits hydrophobic contacts with the cholane scaffold—this is particularly noteworthy because the corresponding residue in GluN2A, Val783, together with Phe754 in the contacting GluN1 subunit, forms a molecular switch that mediates allosteric inhibition²¹². In summary, the tetramer simulations revealed that the interface between GluN1_{LBD} and GluN2D_{LBD} forms a transient pocket to which tauro-CDC may bind. They thus explain the allosteric inhibition of GluN2D-carrying NMDARs by tauro-CDC found in pharmacological experiments.

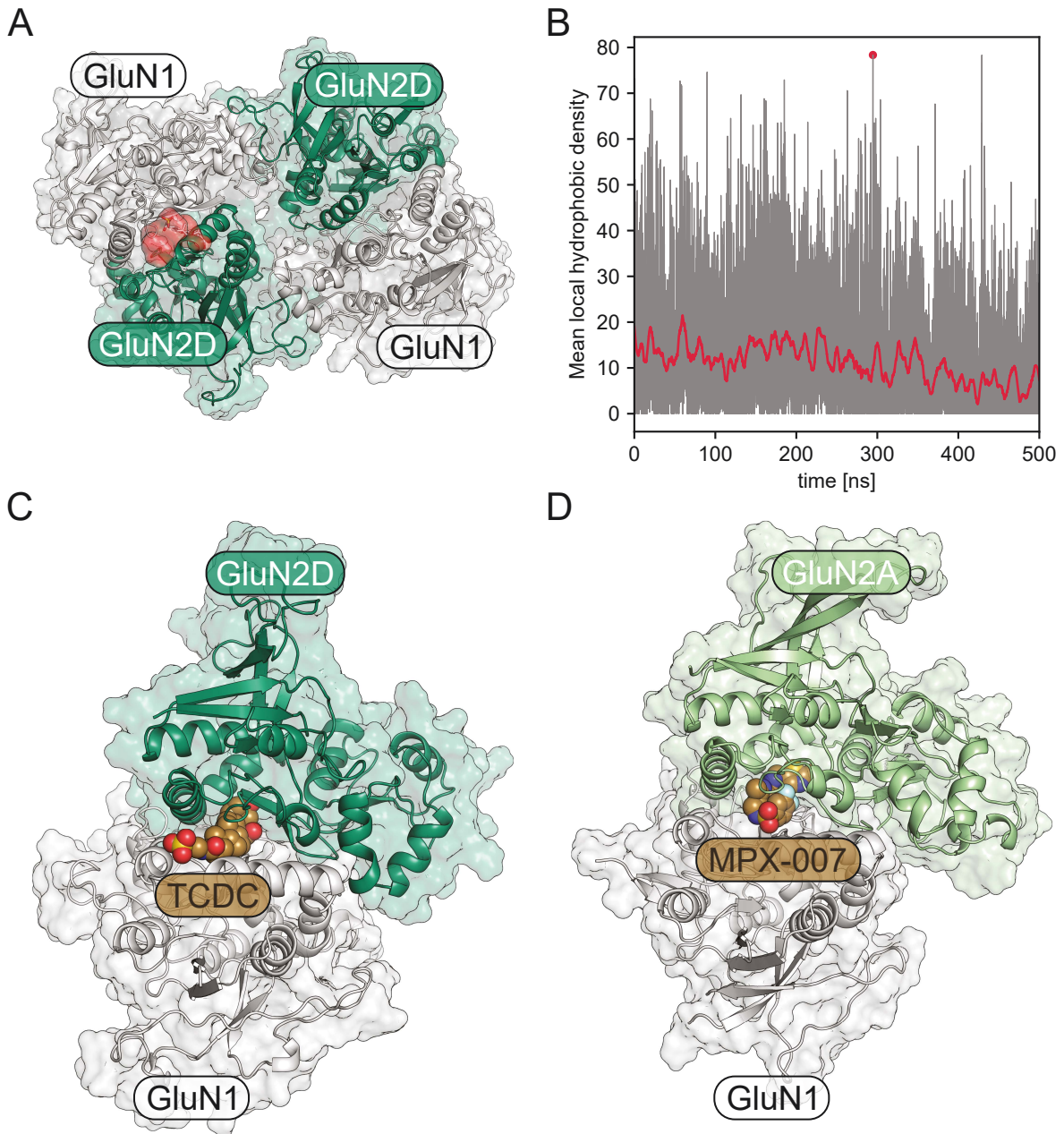


Figure 20: Identification of a potentially allosteric binding site in the GluN1_{LBD}/GluN2D_{LBD} interface. The GluN1 subunit is depicted in gray, the GluN2D subunit in dark green. **(A)** Structure of the GluN1_{LBD}/GluN2D_{LBD} tetramer in complex with tauro-CDC (gray sticks inside red surface). The ligand was docked to the pocket conformation that showed the highest mean local hydrophobic density^{332,333} during 500 ns of accelerated MD. **(B)** Time course of the mean local hydrophobic density^{332,333} for the identified pocket in the GluN1_{LBD}/GluN2D_{LBD} interface. The snapshot in which the pocket displayed the maximum mean local hydrophobic density is marked with a purple dot. **(C)** Binding mode of tauro-CDC (brown spheres) in the alternative binding pocket in the GluN1_{LBD}/GluN2D_{LBD} interface. **(D)** Crystallographically determined binding mode of the negative allosteric modulator MPX-007 (brown spheres) in the GluN1_{LBD}/GluN2A_{LBD} interface²¹². (Figure and caption were modified from Publication II).

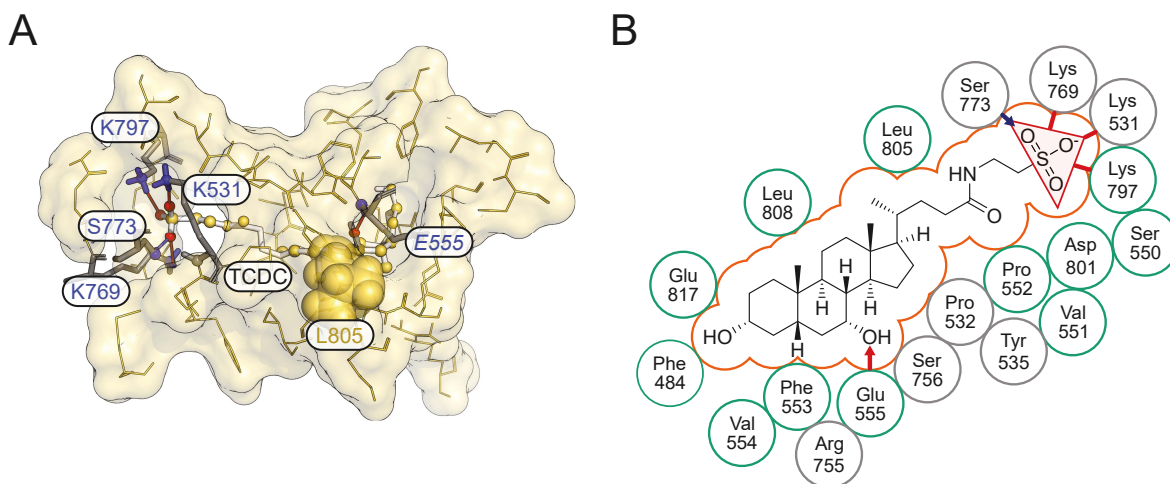


Figure 21: Binding mode of tauro-CDC in the GluN1/GluN2D LBD dimer interface.

(A) Binding mode of tauro-CDC (white sticks) in the GluN1_{LBD}/GluN2D_{LBD} interface (gold surface). A residue is considered part of the binding site if at least one of its atoms lies within 4 Å of any atom of tauro-CDC. Ligand atoms interacting with the protein only through van der Waals contacts are depicted as small gold spheres; the corresponding amino acids in the protein are depicted in gold line-and-surface representation. Amino acids that form hydrogen bonds are shown as gray sticks. The atoms involved in hydrogen bonds are colored according to their involvement as donor (blue) or acceptor (red) and are depicted as small spheres. Leu808 is shown as space-filling golden spheres to emphasize its proposed role in the allosteric inhibition mechanism. Amino acid labels are colored according to the involvement of the respective amino acid as donor (blue) or acceptor (red, note that no such donors are present in this complex). Amino acids in which only the side chain is involved in the hydrogen bond are labeled in normal font style and those in which only the backbone is involved are labeled in italic font style. (B) Two-dimensional representation of the interactions between tauro-CDC and the residues in the GluN1/GluN2D interface. Residues that belong to the GluN1 subunit are circled in gray, residues that belong to the GluN2D subunit are circled in green. Hydrogen bonds are indicated by arrows, ionic interactions by lines, regardless of whether the amino acids involved simultaneously form hydrogen bonds.

5.3 Conclusion and significance

With the results described and discussed above, I have significantly contributed to a collaborative, interdisciplinary study. Using the described pipeline of docking experiments, MD simulations, and effective binding energy calculations, I gradually substantiated a set of pharmacological and electrophysiological experiments, suggesting that tauro-CDC inhibits GluN2D-carrying NMDARs allosterically, but GluN3B-carrying NMDARs competitively. Analyzing the interaction partners in tauro-CDC/GluN3B_{LBD} complexes, I could propose specific mutation sites to further validate GluN3B subunit-specific competitive inhibition by tauro-CDC. Lastly, I identified an allosteric binding site in the GluN1/GluN2D LBD dimer interface, which – to the best of my knowledge – has not been described before for this particular pair of subunits.

The key results and findings of this study are:

- Tauro-CDC preferentially inhibits NMDA receptors carrying GluN2D and GluN3B subunits.

- Tauro-CDC acts as a negative allosteric modulator in GluN2D-carrying NMDARs but inhibits GluN3B-carrying NMDARs competitively. In tauro-CDC/GluN3B complexes, Asp447 could, at least in part, play a selectivity-determining role.
- Negative allosteric modulation by tauro-CDC occurs through its occupation of an allosteric binding pocket in the GluN1/GluN2D LBD dimer interface. This binding pocket is in excellent agreement with a known binding site for positive²¹⁵ and negative²¹² allosteric modulators in GluN2A-carrying NMDARs but has not been described for GluN2D-carrying NMDARs so far.

These conclusions substantiate the suspected pathophysiological role of non-neuronal NMDARs in cholestatic liver diseases and provide a basis for treatment strategies that exploit the described isoform specificity. The approach employed to identify the binding pocket and a druggable conformation thereof might be promising for the development of compounds with therapeutic potential.

6 PUBLICATION III

Hydrophobic alkyl chains substituted to the 8-position of cyclic nucleotides enhance activation of CNG and HCN channels by an intricate enthalpy - entropy compensation

Otte, M., Schweinitz, A., **Bonus, M.**, Enke, U., Schumann, C., Gohlke, H., Benndorf, K.

Original publication, see Publication III in section 11; contribution: 20% (details provided below).

(I adapted parts of the following text from the respective publication.)

6.1 Background

CNBD channels, sometimes referred to as cyclic nucleotide-modulated (CNM) channels, are a family of tetrameric cation channels in the voltage-gated K^+ channel superfamily²⁵⁹. Two subfamilies in the family of CNBD channels are the CNG²⁴⁴ and HCN channels^{279,334}. CNG channels generate receptor potentials in retinal rod photoreceptors and olfactory sensory neurons^{335,336}. They are non-selective, voltage-independent, and directly activated by binding of cyclic nucleotide monophosphates (cNMPs). HCN channels, in turn, evoke rhythmic electrical activity in specialized neurons and cardiac cells³³⁷⁻³³⁹. They are weakly K^+ -selective and activated by hyperpolarization. In contrast to CNG channels, binding of cAMP to the CNBD of HCN channels only modulates the response of the channel to activating stimuli: it increases the activation rate, shifts the voltage of half-maximal activation to less hyperpolarizing potentials and increases the maximal current³⁴⁰⁻³⁴².

Due to the inherent complexity of CNG and HCN channel regulation, studying the mechanisms that underly channel function is a challenging task. Investigating the processes of ligand binding and activation gating while considering the mutual influence of subunits at the single-channel level, for example, would provide important functional insights. Such studies, however, would require ligands that can be coupled to fluorescent dyes and, at the same time, exhibit a longer residence time at the channels than cAMP and cGMP. A first step towards such ligands is the identification of high-affinity cNMP derivatives that carry couplable groups.

Several authors have investigated the effects of different chemical modifications of cNMPs and cNMP derivatives on the activation of CNG channels^{343,344}, HCN channels³⁴⁵, or isolated CNBDs of HCN channels³⁴⁶⁻³⁴⁸. In these studies, an 8-alkylthio or 8-arylthio substitution was particularly conducive to increasing the apparent affinity of the compounds. Even large

substituents were readily tolerated, resulting, e.g., in an 8-NBD-thio-cAMP³⁴⁵ or 8-pCP-thio-cGMP³⁴⁸, and indicating a considerable margin of possibilities for substitution.

In this study, we^{**} systematically investigated the effect of a range of 8-alkylthiopurine and 8-heteroalkylthiopurine substitutions within a set of 20 cGMP- and cAMP-derived ligands. We quantified the agonistic potential of these ligands towards functional homotetrameric CNGA2 channels, heterotetrameric CNGA2:CNGA4:CNGB1b channels, and homotetrameric HCN2 channels. For a subset of ligands (**Figure 22**), we combined molecular docking, MD simulations, and end-point free energy calculations to explain the observed structure-apparent affinity relationship. This approach revealed that the higher apparent affinity of the derivatives with longer alkyl chains is a result of an intricate enthalpy-entropy compensation.

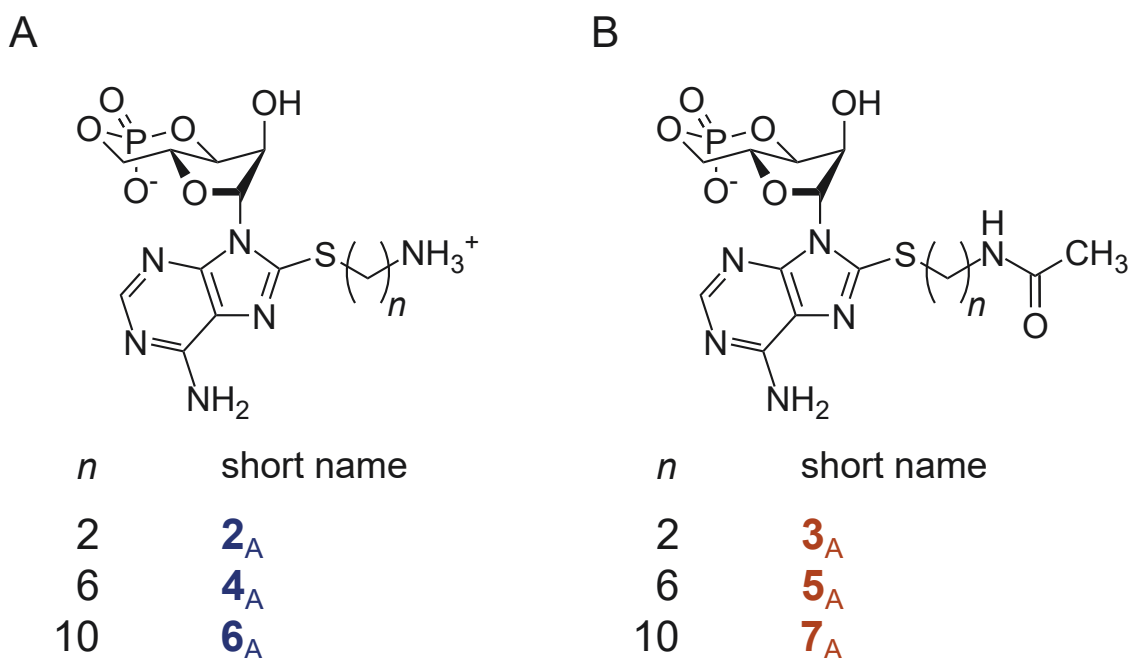


Figure 22: Chemical structure of cAMP derivatives for which the structure-apparent affinity relationship was characterized computationally.

(A) Structure of the studied 8-aminoalkylthio-cAMP derivatives ($n = 2, 6, 10$). (B) Structure of the studied 8-acetylaminoalkylthio-cAMP derivatives ($n = 2, 6, 10$).

6.2 Results

Electrophysiological measurements on homotetrameric CNGA2 (**Figure 23**) and heterotetrameric CNGA2:CNGA4:CNGB1b channels performed by my collaborators

^{**} Synthesis of the compounds, as well as the molecular biological and electrophysiological experiments were performed by Maik Otte, Dr. Andrea Schweinitz, and Uta Enke under supervision of Prof. Dr. Klaus Benndorf (Institute of Physiology II, University Clinics Jena, Friedrich Schiller University Jena, Jena, Germany) and Prof. Dr. Christina Schumann (Department of Medical Didactics and Biotechnology, Ernst Abbe University Jena, Jena, Germany).

indicated a consistent activity pattern within all tested cAMP (**Figure 23A**) and cGMP (**Figure 23B**) derivatives. With increasing length of the alkyl chain of the compounds, their activity also increased regardless of whether the compound carried a free or an acetylated amino group. Furthermore, *N*-acetylation increased the activity in compounds with an ethyl chain but did not change the activity in compounds with hexyl or decyl chains (**Figure 23**; Fig. 2 and Fig. 3 in Publication III). A similar activity pattern was observed in HCN2 channels (Fig. 5c in Publication III).

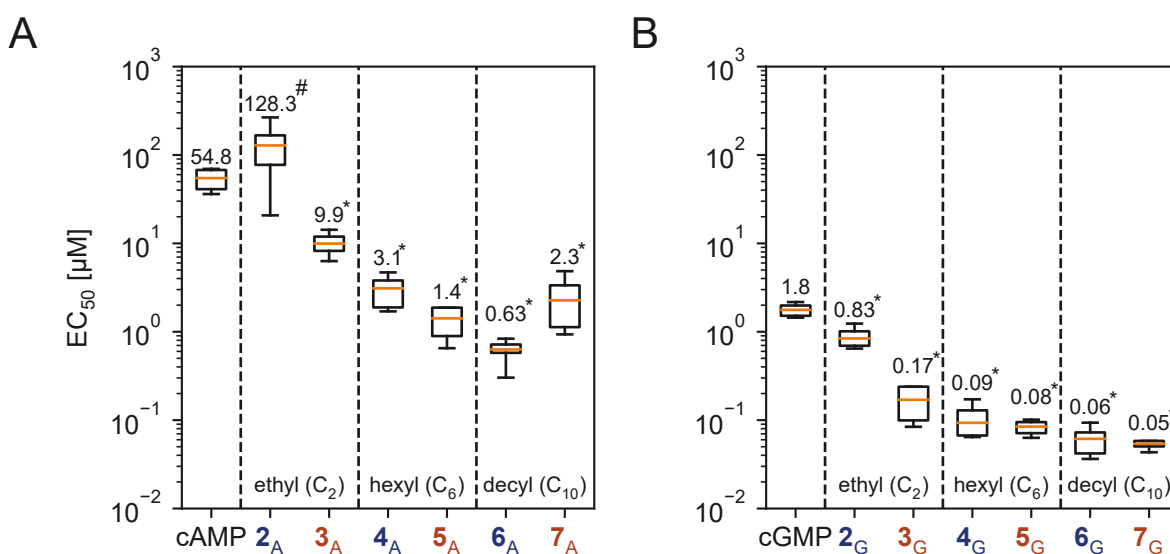


Figure 23: Comparison of the effects between *N*-acetylated and non-acetylated aminoalkylthio-cNMP derivatives in homotetrameric CNGA2 channels.

Box plots of the EC_{50} values for the investigated (A) 8-(acetyl)aminoalkylthio-cAMP and (B) 8-(acetyl)aminoalkylthio-cGMP derivatives at which half-maximum currents were measured in homotetrameric CNGA2 channels. Chain lengths are indicated at the bottom of the plot. For derivatives that contain a charged amino group, the short names (cf. **Figure 22A**) are highlighted in blue; for derivatives that contain an acetylated, uncharged amino group, the short names (cf. **Figure 22B**) are highlighted in brown (Figure and caption were modified from Publication III). The data necessary to recreate this figure was kindly provided by Maik Otte).

To explain the observed relationship between structure and apparent affinity towards heterotetrameric CNGA2 channels within cAMP derivatives 2_A-7_A, I used a combination of molecular docking, MD simulations, and end-point binding free energy calculations. I excluded cAMP derivatives 8_A-11_A (cf. Publication III) as their pEC_{50} range (< 0.3 log units) was similar to the accuracy limit of my intended computational approach³⁴⁹. Similarly, I excluded all cGMP derivatives (2_G-11_G, cf. Publication III) as their pEC_{50} range was almost one log unit smaller than the pEC_{50} range of the cAMP derivatives.

In order to generate binding poses of 2_A-7_A in the heterotetrameric CNGA2 channel as suitable starting conformations for MD simulations, I first constructed a homology model of the heterotetrameric CNGA2 channel based on the cryo-EM structure of TAX-4²⁶³ (overall

sequence identity: ~53%; sequence identity within the cNMP binding pocket: ~77%). Next, I performed molecular docking of compounds **2_A-7_A** to a single CNGA2 subunit using the GlideXP³⁵⁰ methodology (Schrödinger Suite v. 2017-1) and restored C₄ symmetry by chain duplication and superposition onto the respective TAX-4 chains²⁶³. After a thorough reparameterization of the C8-sulfur dihedral in cAMP guided by quantum mechanical calculations (cf. Supplementary Information to Publication III), I subjected each of the six complexes of four cAMP derivatives and a homotetrameric CNGA2 channel to 200 ns of MD simulations in an explicit solvent/explicit membrane environment (for further details, cf. Supplementary Information to Publication III). Subsequently, I used 1,000 snapshots (extraction interval: 0.2 ps) from each of the four simulations to calculate binding free energies with the MM-PBSA approach (cf. section 2.2.2)^{32,75,76,129}. Note that I also used each of the 1,000 snapshots to estimate the entropic contribution $-T\Delta S$ via normal mode analysis. As described in section 2.2.2, this approach is highly computationally demanding, and most – if not all – studies to this date refrained from using such a large number of snapshots to estimate $-T\Delta S$.

During all simulations, the respective ligands remained stably bound to the CNBD (RMSD of the cNMP scaffold ≤ 1.5 Å, **Figure 24A**), although the aminoalkylthio substituents displayed considerable conformational freedom (**Figure 24B**). The computed binding free energies correlated excellently with the experimentally determined pEC₅₀ values, regardless of whether an internal dielectric constant of $\epsilon = 1$ ($R_{\epsilon=1} = 0.89$) or $\epsilon = 4$ ($R_{\epsilon=4} = 0.92$) was chosen^{124,351} (**Figure 24C**). Standard errors of the mean for the ΔG_{bind} estimate over all subunits were < 0.62 kcal mol⁻¹ (**Table 2**), indicating a high precision and neglectable differences between subunits in the ensemble. The results of my calculations mirror the observations from the electrophysiological measurements in that longer-chain alkyl substituents favor the binding to heterotetrameric CNGA2 channels ($\Delta\Delta G_{\text{C2}} = 0.00$ kcal mol⁻¹ (reference), $\Delta\Delta G_{\text{C6}} = -2.33$ kcal mol⁻¹, $\Delta\Delta G_{\text{C10}} = -5.75$ kcal mol⁻¹; values given for $\epsilon = 1$, same trend for $\epsilon = 4$). They further reproduce the trend from the experiment, that derivatives with shorter alkyl chains benefit more from the removal of the N-terminal charge through acetylation than those with longer alkyl chains ($\Delta\Delta G_{\text{C2}} = -3.84$ kcal mol⁻¹, $\Delta\Delta G_{\text{C6}} = -3.51$ kcal mol⁻¹, $\Delta\Delta G_{\text{C10}} = 1.26$ kcal mol⁻¹; values given for $\epsilon = 1$, same trend for $\epsilon = 4$). In summary, they demonstrate that the measured differences in potency between cAMP derivatives **2_A-7_A** are likely a result of affinity differences towards the CNBD and are unrelated to other parts of the channel.

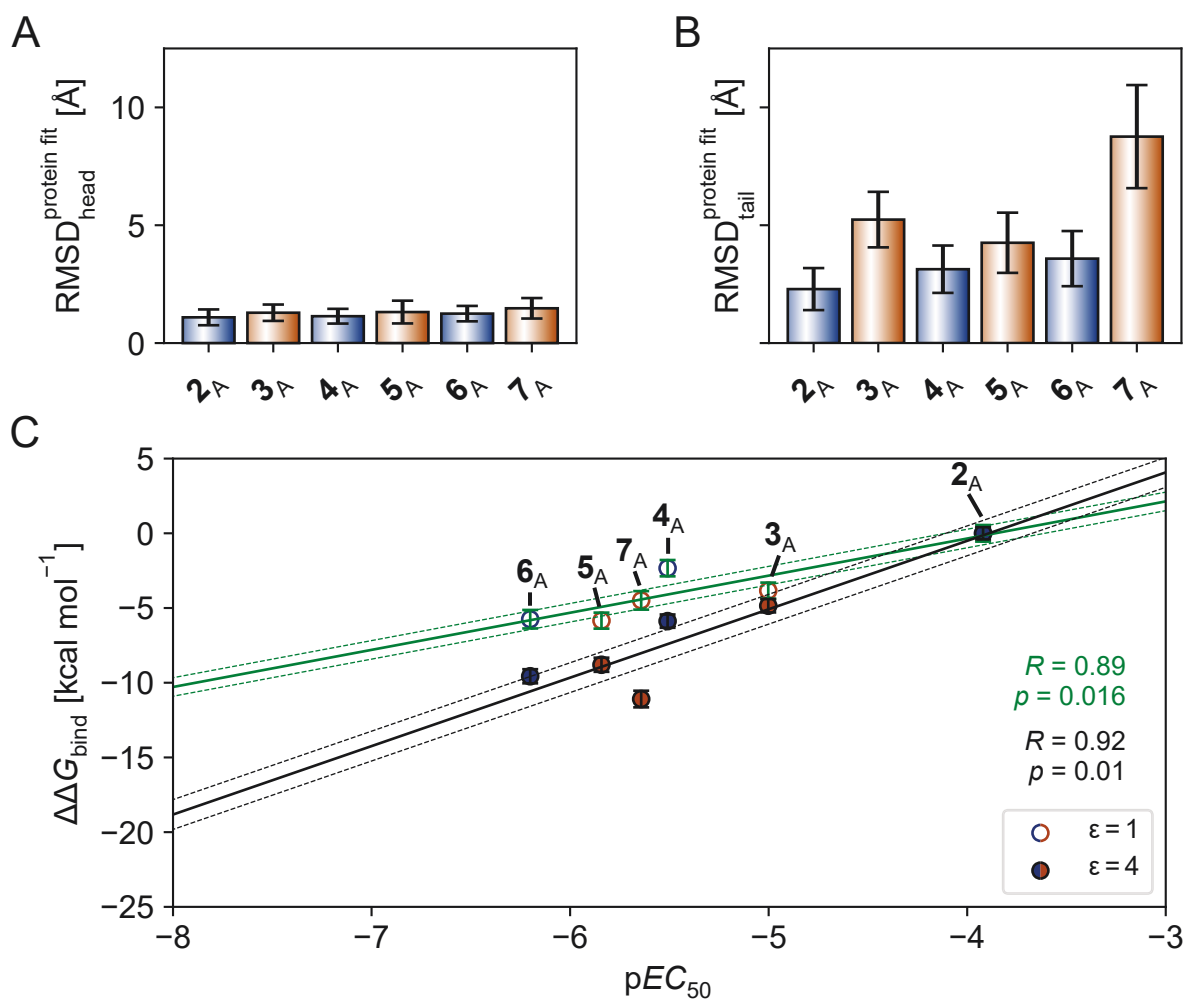


Figure 24: Conformational stability and calculated binding free energies of cNMP derivatives 2_A-7_A. Average RMSD values of the (A) cNMP scaffold (“head”) and the (B) aminoalkylthio substituents (“tail”) over all MD simulations. Derivatives with charged amino groups are highlighted with blue bars, derivatives with acetylated amino groups are highlighted with brown bars. (C) Calculated $\Delta\Delta G_{\text{bind}}$ with respect to 2_A for cAMP derivatives of the congeneric series carrying C₂-, C₆-, and C₁₀-substituents (3_A-7_A) for internal dielectric constants of $\epsilon = 1.0$ (green error bars and regression line) and $\epsilon = 4.0$ (black error bars and regression line). The dashed lines indicate the standard error of the regression estimate. Color coding of the data points corresponds to the one used in panels A and B (Figure and caption were modified from Publication III).

Table 2: Calculated binding free energies for compound 2_A-7_A.

Compound	$\Delta G_{\text{bind}} \pm \text{SEM}$ [kcal mol ⁻¹]*
2 _A	-31.45 ± 0.56 (-27.03 ± 0.41)
3 _A	-35.29 ± 0.54 (-31.88 ± 0.45)
4 _A	-33.78 ± 0.55 (-32.91 ± 0.44)
5 _A	-37.30 ± 0.54 (-35.83 ± 0.47)
6 _A	-37.20 ± 0.62 (-36.60 ± 0.47)
7 _A	-35.94 ± 0.62 (-38.12 ± 0.56)

* Rounded to two decimal places. Values for $\epsilon = 4$ are given in brackets.

Intending to shed more light on the role of the residual tail motions of the alkyl chains after binding, I quantified the tail motions of compounds 2_A-7_A by estimating anisotropic displacement parameters (ADPs) from simulation data. While the ADPs calculated for the

charged derivatives (**2_A**, **4_A**, **6_A**) displayed only a modest, linear dependence on chain length, the ADPs calculated for the acetylated derivatives (**3_A**, **5_A**, **7_A**) were 2.5- to 15-fold higher, and their increase was over-proportional and nonlinear with respect to chain length. Based on these findings, I hypothesized an intricate contribution of the differences in residual tail motions of compounds **2_A**-**7_A** to the differences in binding free energies. To support this hypothesis, I decomposed the calculated binding free energies into the contributions of their individual terms^{352,353} (cf. sections 2.1.3 and 2.2.2).

Regardless of whether a compound was *N*-acetylated or not, enhanced van der Waals interactions and nonpolar contributions to ΔG_{solv} promoted the binding of longer-chain derivatives (**Figure 25A, C**). For derivatives with a charged nitrogen, electrostatic interactions foster the binding of longer-chain derivatives, but their contribution is, in all cases, overcompensated by the penalty resulting from the polar contribution to ΔG_{solv} (**Figure 25A**). The same holds true for the uncharged C₆ derivative (**5_A**), but surprisingly, electrostatics strongly oppose the binding of the uncharged C₁₀ derivative (**Figure 25C**). MD trajectories offered a structural explanation for these observations: three lysine residues at the “gate” of the binding pocket strongly repulse the positive charge in the C₂ derivative. These charge-charge repulsions become weaker in charged C₆ and C₁₀ derivatives since the charge-charge distance increases with chain elongation. In contrast, elongation of *N*-acetylated compounds abrogates charge-assisted hydrogen bonds between the same lysine residues and the acetyl carbonyl group.

Similarly, chain elongation leads to different contributions of the configurational entropy, depending on whether the compounds are charged or acetylated. For charged ligands, the configurational entropy contributions to binding become more disfavored with increasing chain length (**Figure 25B**), while for acetylated ligands, the differences in this contribution with respect to chain length approach zero (**Figure 25D**). This is consistent with the more substantial increase in ADPs for the series of *N*-acetylated ligands and suggests that their residual tail motion after binding is sufficient to minimize the loss of configurational entropy upon binding.

These findings are corroborated when directly comparing derivatives of equal chain length but with either a terminal charge or an acetyl group (**Figure 25E, F**): both the electrostatic and van der Waals contribution to binding is much more favorable for acetylated compounds. Configurational entropy contributions favor the charged C₂ derivative over the acetylated C₂ derivative, presumably due to a more pronounced tail motion of the charged

variant resulting from the charge-charge repulsion described above. This is contrasted by a presumed rigidity of the acetylated C₂ tail, likely mediated by charge-assisted hydrogen bonds to the surrounding lysine residues. For the longer tails, the differences in the configurational entropy contributions favor the acetylated derivatives, which is consistent with the larger calculated ADPs for these compounds.

In summary, I elaborated on an intricate interplay of favorable and unfavorable electrostatic interactions, as well as differential configurational entropy losses that explain the affinity gain for the investigated cAMP derivatives upon chain elongation and the effect of *N*-acetylation. The calculated binding free energies correlated excellently with the experimentally determined apparent affinities, indicating a desirable extended residence time of the C₆ and C₁₀ derivatives. These compounds were, therefore, a promising starting point for fluorophore-coupled ligands that can be used as molecular tools in future functional studies on CNG and HCN channels. Based on the findings of the present study, such fluorophore-coupled ligands were synthesized and characterized in a follow-up study¹².

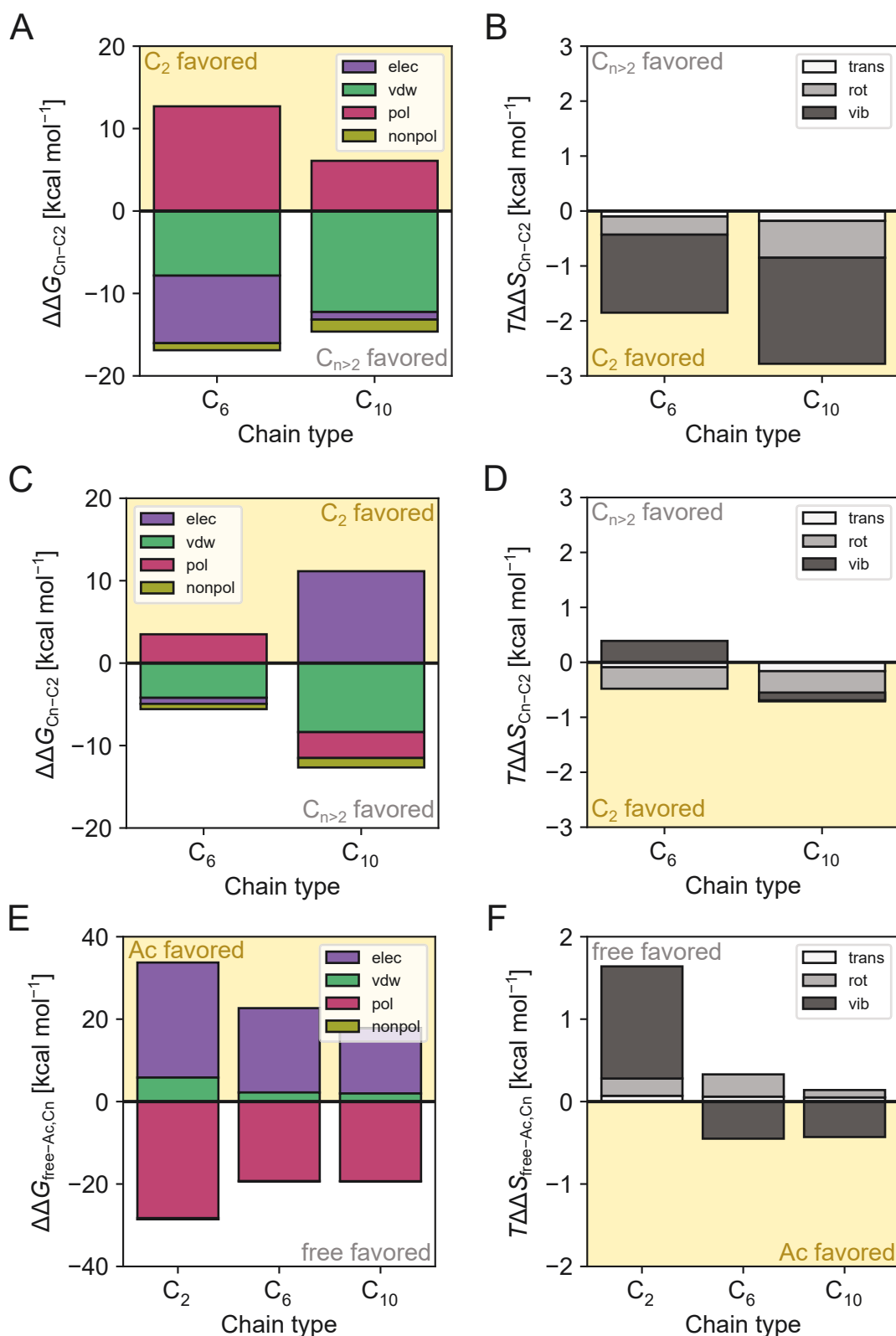


Figure 25: Binding free energy decomposition.

Binding energies were decomposed into their respective energy (A, C, E; electrostatic (elec), van der Waals (vdw), polar solvation (pol) and nonpolar solvation (nonpol)) and configurational entropy (B, D, F; translational (trans), rotational (rot), and vibrational (vib)) terms for $\epsilon = 1$. (A, B) Difference between aminoethyl- (2_A , C_2), aminoethyl- (4_A , C_6), and aminodecyl (6_A , C_{10})-derivatives. (C, D) Difference between acetylaminoethyl- (3_A , C_2), acetylaminoethyl- (5_A , C_6), and acetylaminoethyl- (7_A , C_{10})-derivatives. (E, F) Difference between non-acetylated (free) and acetylated (Ac) derivatives of the same chain length (Figure and caption were modified from Publication III).

6.3 Conclusion and significance

With the approaches and analyses described above, I have made a significant contribution to the results and the design of a collaborative, interdisciplinary study. I explained the observed structure-apparent affinity relationship within the investigated series of ligands on a structural and energetic level and found evidence for an intricate case of enthalpy-entropy compensation. Furthermore, I found evidence that the apparent affinities to CNG and HCN channels can serve as a surrogate for the ligand residence time, a key parameter to be optimized for molecular tools that are used in functional studies on these channels.

In summary, the study yields the following main results:

- Hydrophobic moieties at the 8-position of cyclic nucleotides increase the apparent affinity towards CNG and HCN channels, as does an elongation of the alkyl chain from C₂ over C₆ to C₁₀.
- Apparent affinities and calculated binding free energies for a congeneric series within the investigated cNMP derivatives correlated excellently, suggesting that apparent affinities are suitable surrogates for the ligand residence time.
- An intricate interplay of favorable and unfavorable electrostatic interactions and differential configurational entropy losses explain the affinity gain upon chain elongation and the effect of *N*-acetylation.

The compounds characterized in this study, in particular 8-aminohexylthio-cAMP and 8-aminodecylthio-cAMP, were promising starting points for the development of molecular tools to study CNG and HCN channel function¹².

7 PUBLICATION IV

***N*⁶-modified cAMP derivatives that activate protein kinase A also act as full agonists of murine HCN2 channels**

§Leybold, T., §**Bonus, M.**, Spiegelhalter, F., Schwede, F., Schwabe, T., Gohlke, H., Kusch, J.

§ Both authors contributed equally to this work

Original publication, see Publication IV in section 11; contribution: 30% (details provided below).

(I adapted parts of the following text from the respective publication.)

7.1 Background

cAMP is the first identified second messenger²⁷⁴ and thus one of the most important translators of extracellular signals into a cellular response³⁵⁴. The nature of this response is mediated by binding of cAMP to intracellular effector proteins. Not long ago, protein kinase A (PKA) was considered the only cAMP effector protein³⁵⁵, although this view was already challenged by the discovery of cAMP-mediated activation of CNG channels in photoreceptor cells²³⁸ and olfactory sensory neurons²⁴⁰, as well as cAMP-mediated modulation of HCN channel function in cardiac sino-atrial node cells³⁵⁶. The more recent discoveries of the exchange protein directly activated by cAMP (Epac)^{280,281} and Popeye domain-containing (POPDC)^{282,284} families vastly extended the hitherto mostly linear picture of cAMP-mediated downstream signaling. At the same time, however, these discoveries also highlighted the challenge in distinguishing the different cAMP-mediated signal transduction pathways. Epac, PKA, and cyclic nucleotide-regulated ion channels (i.e., CNG and HCN channels) all rely on a conserved CNBD to sense and bind cAMP³⁵⁷ (cf. section 2.3.4), which greatly impedes the identification of compounds that selectively activate only one of these proteins and the associated downstream signaling pathway. Nevertheless, by modifying the *N*⁶-position, a number of compounds that were selective for PKA over Epac could be obtained. In contrast, modification of the 2'-position led to partial agonists of PKA with conserved activity towards Epac. Combined with an 8-(*p*-chlorophenylthio) (8-pCPT) substitution, these compounds became highly Epac-selective²⁸⁷. However, studies on the selectivity profiles of these compounds typically did not include fully functional cyclic nucleotide-regulated ion channels.

To advance the identification of compounds that allow distinguishing between the different cAMP-sensing proteins, we^{††} studied the effect of three *N*⁶-modified PKA agonists – *N*⁶-phenyladenosine-3',5'-cyclic monophosphate (*N*⁶-Phe-cAMP), *N*⁶-benzyladenosine-3',5'-cyclic monophosphate (*N*⁶-Bn-cAMP), and *N*⁶-benzoyladenosine-3',5'-cyclic monophosphate (*N*⁶-Bnz-cAMP), – on functional murine HCN2 (mHCN2) channels (**Figure 26A**). All three compounds showed efficacies similar to cAMP but different apparent affinities. Using molecular docking and subsequent mutagenesis experiments, we identified the structural determinants underlying these differences.

7.2 Results

The molecular biological and electrophysiological experiments performed by my collaborators revealed that all three compounds regarded as site A-selective PKA agonists^{358,359}, *N*⁶-Phe-cAMP, *N*⁶-Bn-cAMP, and *N*⁶-Bnz-cAMP, increased the current amplitude and accelerated the activation kinetics of mHCN2 channels. The type of substitution, however, greatly impacted the apparent affinity of the compounds. Compared to cAMP, a phenyl or benzyl substituent increased the apparent affinity by one order of magnitude, while a benzoyl substituent decreased it by the same amount (**Figure 26B**). These changes in apparent affinity were not coupled to changes in efficacy as quantified by the maximal agonist-induced voltage shift ($\Delta V_{1/2,max}$) and the maximal fractional increase in tail current amplitude ($I_{max,cpd}/I_{max,cAMP}$, cf. Fig. 3 in Publication IV), and all compounds accelerated the activation kinetics of the channel in a similar way as cAMP (cf. Fig. 4 in Publication IV). Therefore, the differences in apparent affinity likely point to differences in true affinity, which in turn are caused by variations in the interactions stabilizing the complex.

^{††} The molecular biological and electrophysiological experiments and their analysis were performed by Dr. Tim Leypold, Felix Spiegelhalter, Dr. Tina Schwabe, Dr. Jana Kusch (Institute of Physiology II, University Clinics Jena, Friedrich Schiller University Jena, Jena, Germany), and Dr. Frank Schwede (BIOLOG LSI GmbH & Co. KG, Bremen, Germany).

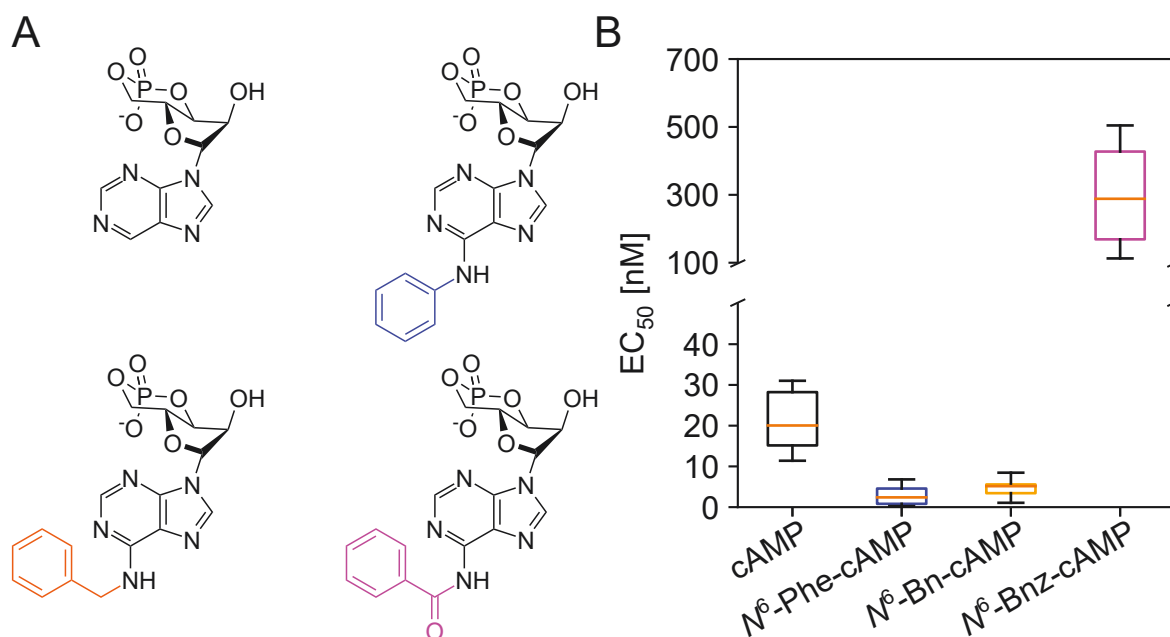


Figure 26: Structure and apparent affinities of the three tested N^6 -modified cAMP derivatives.

(A) Structure of cAMP (top left), N^6 -Phe-cAMP (top right, characteristic structural moiety highlighted in blue), N^6 -Bn-cAMP (bottom left, characteristic structural moiety highlighted in orange), N^6 -Bnz-cAMP (bottom right, characteristic structural moiety highlighted in magenta). (B) Box plot of EC_{50} values obtained from concentration-response relationships with respect to half-maximum activation of murine HCN2 channels. Boxes are drawn from the lower to the upper quartile; whiskers indicate the data range. Color coding corresponds to the one used in panel A (Figure and caption were modified from Publication IV in which individual points of measurement and an alternative representation with standard deviation are also presented).

To unravel the origin of the higher apparent affinities of N^6 -Phe-cAMP and N^6 -Bn-cAMP and the lower apparent affinity of N^6 -Bnz-cAMP at the structural level, I predicted and compared the binding modes of cAMP, N^6 -Phe-cAMP, N^6 -Bn-cAMP, and N^6 -Bnz-cAMP in mHCN2J by molecular docking experiments. The docking protocol I selected reproduced the crystallographically determined binding mode of cAMP with a heavy-atom RMSD of 0.60 Å (**Figure 27A**), strongly indicating that the binding modes of the N^6 -substituted derivatives can also be predicted with a high degree of confidence. In the predicted binding modes of N^6 -Phe-cAMP and N^6 -Bn-cAMP (**Figure 27B, C**), the cAMP moiety of the structures displayed only minor differences to the binding mode of cAMP itself (heavy-atom RMSDs of 0.73 Å and 0.91 Å, respectively). Among these compounds, all docking scores (XP GScores³⁵⁰) were in a similar range (**Table 3**).

Table 3: Docking scores of cAMP and N^6 -substituted derivatives in mHCN2J (PDB ID: 1Q50³⁶⁰)

Compound	Docking score (XP Gscore)*
cAMP	-13.53
N^6 -Phe-cAMP	-10.20
N^6 -Bn-cAMP	-11.54
N^6 -Bnz-cAMP	-6.11

* in kcal mol⁻¹

I visually inspected the binding modes of N^6 -Phe-cAMP and N^6 -Bn-cAMP and detected stabilizing cation- π and/or π - π stacking interactions between the additional phenyl ring and the side chain of Arg-635 (**Figure 27B, C**). These interactions provide a potential explanation for the higher apparent affinities of the two compounds compared to cAMP. Due to the restricted rotation around the N^6 -phenyl bond in N^6 -Phe-cAMP, these interactions might remain more stable over time than for N^6 -Bn-cAMP, which could explain the relative ranking in apparent affinity among these two compounds. In stark contrast to these two compounds, N^6 -Bnz-cAMP displayed an inverted binding mode (RMSD between cAMP and the cAMP moiety in N^6 -Bnz-cAMP: 8.66 Å, **Figure 27D**) and a considerably lower docking score (**Table 3**). To determine why N^6 -Bnz-cAMP cannot adopt the typical binding mode of cyclic nucleotides in HCN2 channels, I replaced the benzyl moiety in the binding mode predicted for N^6 -Bn-cAMP with a benzoyl moiety. In doing so, I found that the orientation of the phenyl ring in N^6 -Bnz-cAMP enforced by the planarity of the amide group inevitably leads to clashes with Arg-635, and thus a strongly unfavorable interaction (**Figure 27E**). In summary, my structural studies thoroughly explain the observed structure-apparent affinity relationship.

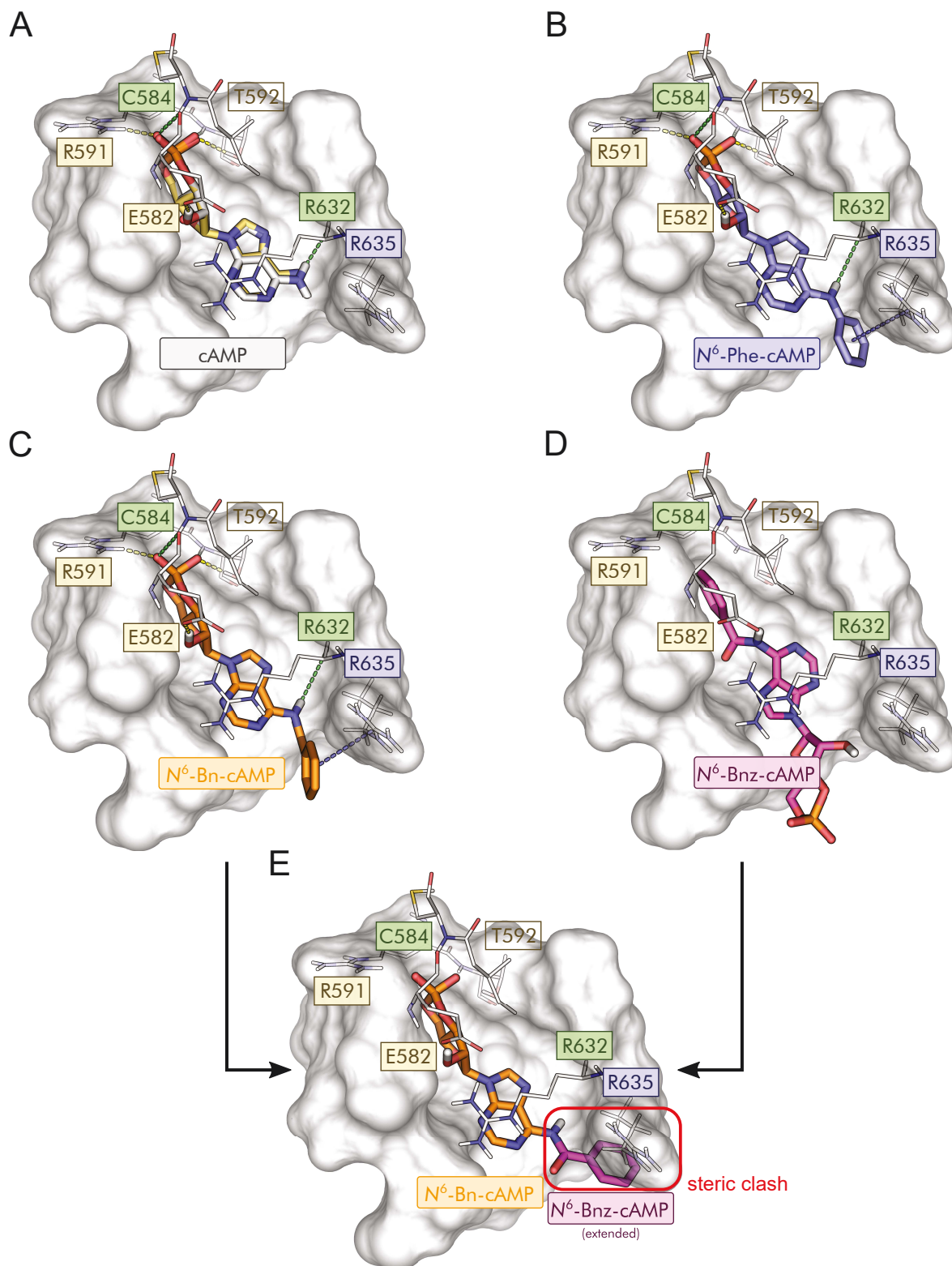


Figure 27: Predicted binding modes of the tested cAMP derivatives.

(A-D) Predicted binding modes of cAMP (A, gray; crystallographic pose (PDB ID 1Q50) shown in gold), *N*⁶-Phe-cAMP (B, blue), *N*⁶-Bn-cAMP (C, orange), and *N*⁶-Bnz-cAMP (D, magenta). In A-C, hydrogen bonds to and from protein side chains are depicted as dashed yellow lines, hydrogen bonds to and from the protein main chain as dashed green lines, and cation- π interactions as dashed blue lines. (E) Predicted binding mode of the cAMP fragment in *N*⁶-Bn-cAMP extended with a benzoyl group as found in *N*⁶-Bnz-cAMP. The steric clash with Arg-635 resulting from this extension is highlighted (red box), explaining the inverted binding mode of *N*⁶-Bnz-cAMP shown in panel D. (Figure taken from Publication IV)

To further validate my structure-based interpretation of the apparent affinity ranking among the N^6 -substituted derivatives, my collaborators performed mutagenesis experiments in which Arg-635 was either neutralized through substitution with alanine (R635A) or charge-inverted through substitution with glutamate (R635E). If Arg-635 indeed conveys the higher apparent affinity of N^6 -Phe-cAMP and N^6 -Bn-cAMP and does not influence cAMP binding itself, such mutations should diminish the apparent affinity to cAMP-like levels. Both the R635A and R635E constructs formed functional channels and displayed a marked current increase and acceleration of activation kinetics in the presence of cAMP. In excellent agreement with my structure-based predictions, the concentration-activity relationship for N^6 -Phe-cAMP determined in both constructs revealed an increase to EC_{50} values indistinguishable from cAMP in wild-type channels (cf. Fig. 6 in Publication IV).

Taken together, our study shows that cross-activation of cyclic nucleotide-modulated channels must be considered when using N^6 -derivatized cyclic nucleotides to study cAMP-dependent cellular processes mediated by PKA. Furthermore, the structure-activity relationships identified in this study provide a starting point for the structure-based design of compounds with higher affinities towards HCN2 channels.

7.3 Conclusion and significance

My structure-based computational analyses and the resulting suggestions for mutagenesis experiments were instrumental in identifying the determining factors for the apparent affinities of the studied compounds. I have significantly contributed to the methodological design, data acquisition, formal analysis, and validation during the course of the interdisciplinary study described above. In summary, the main results of this study are:

- The three N^6 -substituted PKA agonists N^6 -Phe-cAMP, N^6 -Bn-cAMP, and N^6 -Bnz-cAMP also activate HCN2 channels. Although their selectivity profile allows a distinction between PKA- and Epac-mediated signaling, cross-activation of CNG channels could complicate the interpretation of results from experimental settings in which the involvement of these channels cannot be excluded.
- N^6 -Phe-cAMP and N^6 -Bn-cAMP show higher apparent affinities towards HCN2 than the native ligand cAMP. These higher apparent affinities are mediated by π - π and cation- π interactions of the phenyl ring in the compounds with Arg-635. Mutation of this residue resets the apparent affinity of both compounds to that of cAMP towards wild-type channels.

- The lower apparent affinity of N^6 -Bnz-cAMP results from a conformational restriction of the phenyl ring by the adjacent amide group. The enforced coplanarity of these moieties would lead to clashes with Arg-635 if the compound was to adopt the typical binding mode of cyclic nucleotides in HCN2 channels.

When using the investigated N^6 -substituted cyclic nucleotide derivatives to study PKA-mediated cellular processes, at least cross-activation of HCN2 channels needs to be taken into account if the presence of these channels in the experimental setting cannot be excluded. The structure-apparent affinity relationships identified in this study provide a starting point for the structure-based design of potent HCN2 channel agonists, which would ideally also display a favorable selectivity profile.

8 SUMMARY AND PERSPECTIVE

Within the studies on which this thesis is based, I have used computational methods to characterize specific protein-ligand interactions on both a qualitative and quantitative level. Using these approaches, I was able to

- predict – based on known function-determining interactions – whether and to what extent a ligand acts as an agonist or antagonist (Publication I).
- explain differences in the pharmacodynamic mode of action of a single ligand at different receptor subtypes (Publication II).
- predict the thermodynamic binding profile of a series of ligands and use this knowledge to explain the ranking of their relative apparent affinities (Publication III).
- provide structure-guided explanations for increased or decreased potencies within a ligand series, and thus establish starting points for a structure-based optimization of the properties of these ligands (Publication IV).

Through a fruitful combination of experimental and computational approaches, my collaborators and I identified a second bile acid – *norUDCA* – that activates $\alpha_5\beta_1$ integrin in the liver. We also found evidence for a functional selectivity of the, now two, known $\alpha_5\beta_1$ -activating bile acids for overlapping, but distinct $\alpha_5\beta_1$ integrin-mediated signal transduction pathways, and described functional selectivity at integrins without an αI domain for the first time (Publication I). In another study, we found that tauro-CDC acts as a competitive antagonist in NMDA receptors with GluN3B subunits but as a negative allosteric modulator in those with GluN2D subunits, and provide a structure-based explanation for this difference in the pharmacodynamic mode of action (Publication II). In Publication III, we explained the observed structure-apparent affinity relationship within a series of highly potent cAMP analogs towards CNG channels, which underlies a complex enthalpy-entropy compensation. Finally, in Publication IV, we observed that certain cAMP derivatives, previously considered PKA-selective, also activate HCN2 channels. These derivatives displayed an intriguing apparent affinity profile that we could explain on a structural level – the established structure-apparent affinity relationship was validated by mutagenesis experiments.

Structure-based computational approaches often provide the data necessary to put experimental observations into a structural context and can greatly facilitate their interpretation. Conversely, computational techniques can rarely provide the data necessary

to complete the big picture on their own and require extensive input from wet lab experiments to provide successful predictions. Thus, the four studies discussed in this thesis once again underline the importance of cooperation between experimental and computational research groups. Both experimental and computational approaches are continuously evolving through new and innovative methodologies and techniques, and it is up to the researchers to fruitfully combine these developments in future collaborative work.

9 ACKNOWLEDGMENT

First, I would like to thank my supervisor **Prof. Dr. Holger Gohlke** for the opportunity to conduct my research in his group. I am deeply grateful for the seemingly boundless trust he has always placed in me during the course of a great number of projects and for his unremitting optimism that has accompanied me through these projects. I am equally indebted to him for supporting my teaching ambitions and the entrustment with the organization and realization of the “Molecular Modeling” course.

A heartfelt “thanks” also goes to all my colleagues and friends from the CPClab and CBClab. I especially thank **Dr. Nadine Homeyer** for her patient and kind support during the first months of my work. I am convinced that my way of working would never have reached today’s level of maturity without her efforts. My deepest gratitude also goes to **Dr. Bartholomäus Daniel Ciupka**, who always found the right words to rekindle my motivation and to put my perspective back into place during countless sessions of “tümpeln”. I thank **Dr. Benedikt Frieg** and **Dr. Christoph G. W. Gertzen** for the many enjoyable moments and funny memories, even beyond our shared time in the CRC 974. Thanks to my former office mates **Dr. Giulia Pagani** and **Dr. Christian Hanke**, for tolerating my often unpredictable mood, and I direct the same appreciation also towards my current office mates **David Bickel**, **Mohanraj Gopalswamy**, and **Annika Behrendt**. I especially thank **David**, not only because he had to bear me for the longest time so far, but also because he always kept my motivation alive through his humor and his determination. Another special thanks and salutation goes to **Susanne M. A. Hermans**, not only for the odd sense of humor we both seem to share. In addition, I would like to thank **Dr. Christopher Pflieger**, **Dr. Stephan Schott-Verdugo**, **Sabahuddin Ahmad**, and **Dr. Denis Schmidt** for their tireless efforts during the “Molecular Modeling” course and the fruitful and uplifting discussions.

Thanks to all **my collaborators** Prof. Dr. Klaus Benndorf, Uta Enke, Dr. Boris Görg, Prof. Dr. Dieter Häussinger, Prof. Dr. Horst Kessler, Prof. Dr. Nikolaj Klöcker, Dr. Angela Koch, Dr. Jana Kusch, Dr. Tim Leypold, Beatrice Stephanie Ludwig, Maik Otte, Dr. Natalia Qvartskhava, Prof. Dr. Christina Schumann, Dr. Tina Schwabe, Dr. Frank Schwede, Dr. Andrea Schweinitz, Dr. Annika Sommerfeld, and Felix Spiegelhalter, for making this thesis possible.

I also thank **Imke Form**, who sadly passed away in 2019, for all the conversations, her bright positivity, and our work together.

My most sincere appreciation and respect is devoted to **Dr. Alexandra Hamacher**, who has motivated me even beyond my university studies and has always been a role model for me.

In this context, I also thank **Dr. Lars Stank** and **Dr. Thorsten Tuschy**, who were always a great motivation for me during our university studies.

I thank **Susanne Bähz**, not only for everything she has done for me but also for the close friendship that has guided me through many chapters of my life.

I would also like to thank **Dr. Jana Lissy**, **Dr. Friederike Rohn**, **Dr. Helge Jörn Zöllner**, and **all others involved** for their commitment and dedication to the organization of *deLIVER*. Thank you for making this event reality.

I express my sincere gratitude to my WPP and/or diploma students **Daniel Constant**, **Isabelle Eichler**, **Zahra Naderimehr**, **Hela Noori**, and **Shna Saleh** for the excellent work on MDR3, PDR5, and CDK2. It was more than a pleasure to work with you.

I owe my deepest gratitude and appreciation to my **girlfriend**, who has unconditionally supported my efforts in spite of my regular, unpredictable mood eruptions.

Last but not least, I would like to thank **my whole family** and all my **friends** for their support along my journey through life.

I dedicate these last words to my beloved **mother**, who passed away in February 2018.

You left us beautiful memories,

Your love is still our guide,

Although we cannot see you,

You're always at our side

10 REPRINT PERMISSIONS

Publication I (pages LXXXV – CXXXIII)

Reprinted from “*Evidence for functional selectivity in TUDC- and norUDCA-induced signal transduction via $\alpha_5\beta_1$ integrin towards choleresis*”, Bonus, M., Sommerfeld, A., Qvarskhava, N., Görg, B., Ludwig, B. S., Kessler, H., Gohlke, H., Häussinger, D., **Scientific Reports** (2020), 10, 5795.

Copyright © 2020 Bonus *et al.*. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as the original author(s) and source are credited.

Publication II (pages CXXXIV – CLVI)

Reprinted from “*Isoform-specific Inhibition of N-methyl-D-aspartate Receptors by Bile Salts*”, Koch, A., Bonus, M., Gohlke, H., Klöcker, N., **Scientific Reports** (2019), 9, 10068.

Copyright © 2019 Koch *et al.*. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as the original author(s) and source are credited.

Publication III (pages CLVII – CLXXXVII)

Reprinted from “*Hydrophobic alkyl chains substituted to the 8-position of cyclic nucleotides enhance activation of CNG and HCN channels by an intricate enthalpy - entropy compensation*”, Otte, M., Schweinitz, A., Bonus, M., Enke, U., Schumann, C., Gohlke, H., Benndorf, K., **Scientific Reports** (2018), 8, 14960.

Copyright © 2018 Otte *et al.*. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as the original author(s) and source are credited.

Publication IV (pages CLXXXVIII – CXCVIII)

Reprinted from “*N⁶-modified cAMP derivatives that activate protein kinase A also act as full agonists of murine HCN2 channels*”, Leypold, T., Bonus, M., Spiegelhalter, F., Schwede, F., Schwabe, T., Gohlke, H., Kusch, J., **Journal of Biological Chemistry** (2019), 294, 47.

This research was originally published in the Journal of Biological Chemistry. Leypold, T., Bonus, M., Spiegelhalter, F., Schwede, F., Schwabe, T., Gohlke, H., Kusch, J. N⁶-modified cAMP derivatives that activate protein kinase A also act as full agonists of murine HCN2 channels. J. Biol. Chem. 2019; 294:17978–17987. © Leypold, T. *et al.*

11 PUBLICATIONS

Publication I

Evidence for functional selectivity in TUDC- and *nor*UDCA-induced signal transduction via $\alpha_5\beta_1$ integrin towards choleresis

Michele Bonus[§], Annika Sommerfeld[§], Natalia Qvartskhava, Boris Görg,
Beatrice Stefanie Ludwig, Horst Kessler, Holger Gohlke, Dieter Häussinger

Scientific Reports (2020) 10, 5795

Contribution: 35%

[§] Both authors contributed equally to this work

OPEN

Evidence for functional selectivity in TUDC- and *nor*UDCA-induced signal transduction via $\alpha_5\beta_1$ integrin towards choleresis

Michele Bonus^{1,5}, Annika Sommerfeld^{2,5}, Natalia Qvartskhava², Boris Görg², Beatrice Stefanie Ludwig³, Horst Kessler³, Holger Gohlke^{1,4*} & Dieter Häussinger^{2*}

Functional selectivity is the ligand-specific activation of certain signal transduction pathways at a receptor and has been described for G protein-coupled receptors. However, it has not yet been described for ligands interacting with integrins without α I domain. Here, we show by molecular dynamics simulations that four side chain-modified derivatives of tauroursodeoxycholic acid (TUDC), an agonist of $\alpha_5\beta_1$ integrin, differentially shift the conformational equilibrium of $\alpha_5\beta_1$ integrin towards the active state, in line with the extent of β_1 integrin activation from immunostaining. Unlike TUDC, 24-*nor*ursodeoxycholic acid (*nor*UDCA)-induced β_1 integrin activation triggered only transient activation of extracellular signal-regulated kinases and p38 mitogen-activated protein kinase and, consequently, only transient insertion of the bile acid transporter Bsep into the canalicular membrane, and did not involve activation of epidermal growth factor receptor. These results provide evidence that TUDC and *nor*UDCA exert a functional selectivity at $\alpha_5\beta_1$ integrin and may provide a rationale for differential therapeutic use of UDCA and *nor*UDCA.

Functional selectivity is the ligand-specific activation of certain signal transduction pathways at a receptor that can signal through multiple pathways¹. On the molecular level, a ligand likely achieves this type of differential activation by stabilizing only a specific subset of receptor conformations, in particular those that favor interactions with only a specific subset of downstream signaling molecules¹. This phenomenon has so far been described in detail only for G protein-coupled receptors (GPCRs)², but the observation that $\alpha_M\beta_2$ integrins respond differently to fibrinogen- and CD40L-binding has led to the suggestion that this model could be extended to integrins with an α I domain^{3,4}. However, the phenomenon has not yet been described for ligands interacting with integrins lacking an α I domain. Furthermore, a direct connection between differentially ligand-induced integrin conformations and differences in signal transduction pathways downstream of the integrin has not yet been established.

We recently identified the hydrophilic bile acid tauroursodeoxycholic acid (TUDC) as a potent agonist of an $\alpha_5\beta_1$ integrin-mediated signaling pathway towards choleresis^{5,8}. $\alpha_5\beta_1$ integrin is the predominant integrin isoform in the liver and lacks an α I domain⁶. After uptake into hepatocytes through the Na⁺/taurocholate cotransporting polypeptide (Ntcp), TUDC directly activates intracellular $\alpha_5\beta_1$ integrins, followed by an activating phosphorylation of mitogen-activated protein kinases (MAPK) Erk-1/-2 and p38^{MAPK}⁵. These signaling events strongly resemble those initiated in response to hypoosmotic or insulin-induced hepatocyte swelling⁷⁻⁹. There, mechano/swelling-sensitive $\alpha_5\beta_1$ integrins in the plasma membrane become activated and trigger FAK-/c-Src-/MAPK-dependent signaling towards choleresis with Bsep and Mrp2 insertion into the canalicular membrane^{7,10,11}. Chemical modifications of the ursodeoxycholate scaffold in TUDC (Supplementary Fig. 1) either

¹Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, Düsseldorf, Germany. ²Clinic for Gastroenterology, Hepatology and Infectious Diseases, Heinrich Heine University Düsseldorf, Düsseldorf, Germany. ³Institute for Advanced Study and Center for Integrated Protein Science, Department of Chemistry, Technische Universität München, Garching, Germany. ⁴John von Neumann Institute for Computing (NIC), Jülich Supercomputing Centre (JSC), and Institute for Complex Systems - Structural Biochemistry (ICS-6), Forschungszentrum Jülich GmbH, Jülich, Germany. ⁵These author contributed equally: Michele Bonus and Annika Sommerfeld. *email: gohlke@uni-duesseldorf.de; haeussin@uni-duesseldorf.de

completely abolished activity on $\alpha_5\beta_1$ integrin or led to a compound that inhibited the TUDC-induced β_1 integrin activation (taurocholic acid (TC))⁵.

Here, we tested to what extent side chain-modified derivatives of TUDC (24-*nor*-ursodeoxycholic acid (*nor*UDCA), its taurine conjugate (*Tnor*UDCA), glyoursodeoxycholic acid (GUDC), and unconjugated UDCA; Supplementary Fig. 1) can directly activate $\alpha_5\beta_1$ integrins and whether the signaling events downstream of integrin activation differ from those triggered by TUDC. To probe for differences in ligand-induced conformational changes in integrin at the atomistic level, we performed all-atom molecular dynamics (MD) simulations of $\alpha_5\beta_1$ integrin bound to TUDC, *nor*UDCA, *Tnor*UDCA, GUDC, UDCA, or TC of, in total, 3.6 μ s length. In parallel, we studied the extent to which *nor*UDCA-, *Tnor*UDCA-, GUDC-, or UDCA stimulate the activation of β_1 integrins during perfusion of rat liver and compared the signaling events downstream of *nor*UDCA-mediated integrin activation with TUDC-mediated integrin activation. Our results demonstrate that *nor*UDCA directly activates $\alpha_5\beta_1$ integrins in hepatocytes and provide evidence that TUDC and *nor*UDCA exert a functional selectivity for certain signal transduction pathways in $\alpha_5\beta_1$ integrin.

Results

In MD simulations *nor*UDCA induces conformational changes in the $\alpha_5\beta_1$ integrin headpiece that have been linked to integrin activation.

We analyzed all-atom MD simulations of TUDC, *nor*UDCA, *Tnor*UDCA, GUDC, UDCA, and TC bound to the shallow crevice in the subunit interface of the ectodomain of $\alpha_5\beta_1$ integrin for conformational changes in the headpiece region that govern integrin activation. We described these conformational changes by means of three geometric parameters: the kink angle in helix α_1 , the α_7 tilt angle, and the β -propeller – β A domain distance (Figs. 1, 2; details in the Methods subsection “*Analysis of MD trajectories*”). For each complex, three replicates were simulated for 200 ns length to probe for the statistical significance and convergence of the simulation results (Fig. 1c–h). All systems were stable over the course of the simulation time, as demonstrated by the time courses of the root-mean-square deviation (RMSD) of the atomic positions in the β A domain and the full protein (Supplementary Fig. 2), as well as the domain-wise minimum, maximum and average RMSD values (Supplementary Table 1). MD simulations of TUDC- and TC-bound ectodomains served as references as they display the occurrence or absence of conformational changes when the ectodomain is bound to an activating or inhibitory bile acid, respectively⁵. Particularly, TUDC leads to a kink angle of helix α_1 of $147.3 \pm 0.1^\circ$, a tilt angle of helix α_7 of $138.2 \pm 0.1^\circ$, and a distance between the β A-domain and the β -propeller of 36.68 ± 0.01 Å, whereas TC leads to angles of $142.0 \pm 0.1^\circ$ and $126.6 \pm 0.1^\circ$, respectively, and a distance of 35.77 ± 0.01 Å (Figs. 1c,h and 2c, Supplementary Table 2). *nor*UDCA induces α_1 kink and α_7 tilt angles similar in magnitude as in the case of TUDC ($149.2 \pm 0.1^\circ$ and $138.2 \pm 0.1^\circ$, respectively), whereas the distance between β -propeller and β A-domain is ~ 0.6 Å smaller (Figs. 1d and 2a,c, Supplementary Table 2). These findings indicate that *nor*UDCA can activate $\alpha_5\beta_1$ in a similar manner as TUDC but with a lower efficacy. In comparison with *nor*UDCA, for *Tnor*UDCA- and GUDC-bound ectodomains, the α_1 kink angle is decreased by $\sim 6^\circ$, while the α_7 tilt angle remains in the same range, being $\sim 2^\circ$ smaller. Furthermore, the β -propeller – β A domain distance is reduced by ~ 0.4 Å (Figs. 1e,f and 2c, Supplementary Table 2). Finally, for the UDCA-bound ectodomain, the α_1 kink angle is $139.6 \pm 0.1^\circ$, and the α_7 tilt angle is $129.6 \pm 0.1^\circ$; the β -propeller – β A domain distance is 35.80 ± 0.01 Å (Figs. 1g and 2c, Supplementary Table 2). These geometric parameters are more similar to those of TC than to those of any of the previously mentioned bile acids, indicating that UDCA, like TC⁵, cannot activate $\alpha_5\beta_1$ integrins. Furthermore, by comparison, the above results for *Tnor*UDCA- and GUDC indicate that the efficacy of these bile acids in activating $\alpha_5\beta_1$ integrins is, at best, very weak. In all, the computational studies provide evidence that suggests that *nor*UDCA can directly activate $\alpha_5\beta_1$ integrins.

***nor*UDCA activates β_1 integrins in perfused rat liver.** In isolated perfused rat liver, addition of *nor*UDCA at a concentration of ≥ 20 μ mol/l induces the appearance of the active conformation of β_1 integrin after 15 min, whereas in the absence of *nor*UDCA, active β_1 integrin was only scarcely detectable (Fig. 3a). As described for TUDC⁵, β_1 integrin activation was predominantly observed inside the hepatocyte (Fig. 3a, Supplementary Fig. 3). In contrast, active β_1 integrin was only weakly detectable with *Tnor*UDCA (≥ 20 μ mol/l) and GUDC (≥ 20 μ mol/l) (Fig. 3b,c). Perfusion with UDCA (≥ 20 μ mol/l) induces a stronger appearance of the active conformation of β_1 integrin than do *Tnor*UDCA and GUDC (Fig. 3d). None of the bile acids had any effect on the immunostaining for total $\alpha_5\beta_1$ integrins (Fig. 3e).

TUDC induces a stronger appearance of active β_1 integrin than *nor*UDCA. We compared the effect of *nor*UDCA at 20 μ mol/l in inducing the appearance of the active conformation of β_1 integrin in isolated perfused rat liver to that of TUDC at equimolar concentration as a known activator of $\alpha_5\beta_1$ integrin (Fig. 4)⁵. Whereas TUDC induced a pronounced and significant β_1 integrin activation within 5 min, as shown by a β_1 integrin fluorescence intensity of $906 \pm 43\%$ relative to unstimulated control, *nor*UDCA activated β_1 integrins with a lower effect (β_1 integrin fluorescence intensity of $203 \pm 10\%$ relative to unstimulated control). After 5 and 15 min, β_1 integrin fluorescence intensity was significantly higher in the presence of TUDC than in the presence of *nor*UDCA. After 15 min, *nor*UDCA-induced β_1 integrin activation was $510\% \pm 116\%$ of baseline (Fig. 4).

***nor*UDCA induces integrin-dependent signaling cascades similar to TUDC.** Like TUDC⁵, *nor*UDCA (20 μ mol/l) induced within 5 min phosphorylation of extracellular signal-regulated kinases Erk-1/-2, which was abolished in the presence of the RGD motif-containing hexapeptide GRGDSP (10 μ mol/l) but not in the presence of the inactive control hexapeptide GRADSP (10 μ mol/l) (Figs. 5 and 6). Erk-1/-2 phosphorylation due to *nor*UDCA did not increase when phosphatases were inhibited by okadaic acid (5 nmol/l), in contrast to TUDC-induced Erk-1/-2 phosphorylation (Supplementary Fig. 4). *nor*UDCA also increased activation of p38^{MAPK} and the activating Src phosphorylation at tyrosine 418 in an RGD hexapeptide-sensitive way (Fig. 5,

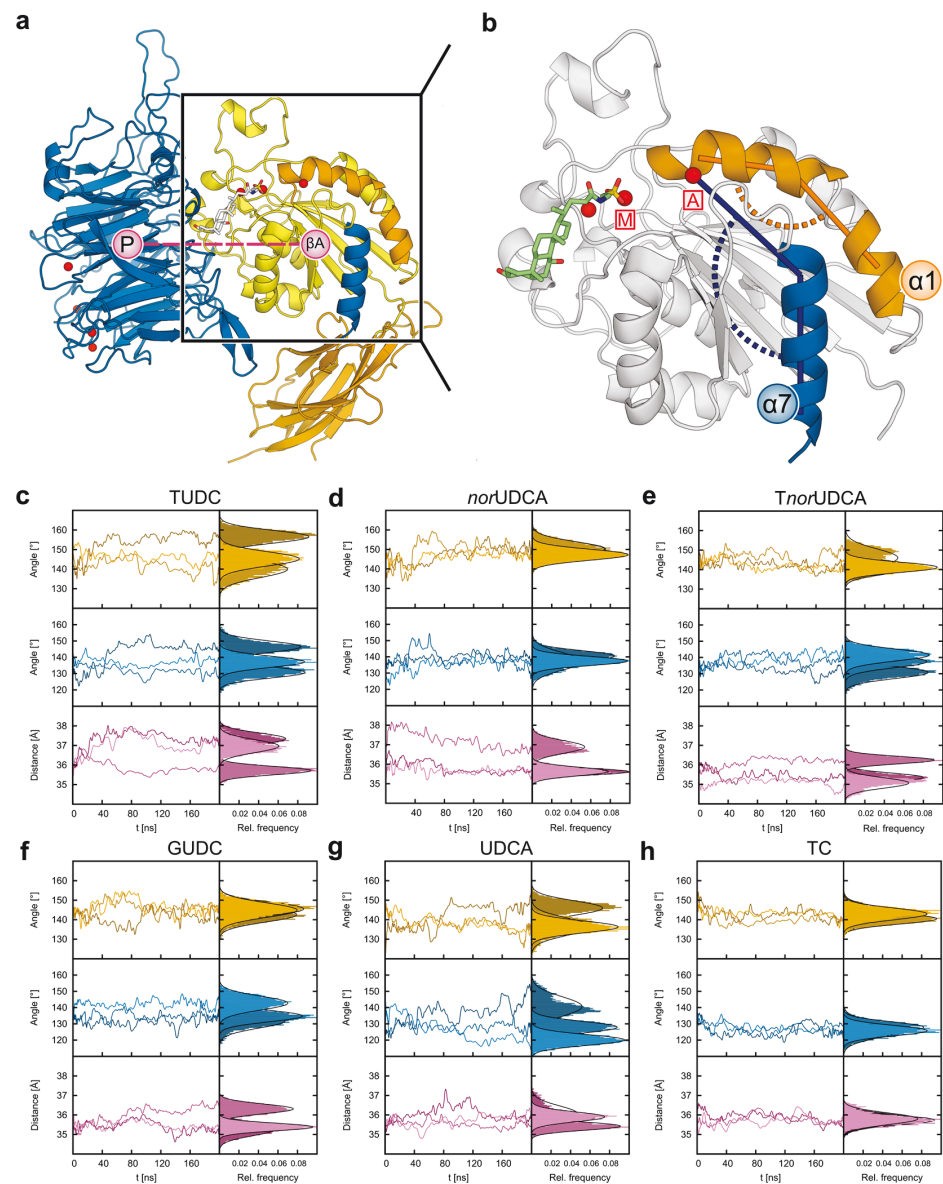


Figure 1. Conformational changes in the $\alpha_5\beta_1$ integrin headpiece. **(a)** Part of the $\alpha_5\beta_1$ integrin headpiece in cartoon representation. Helices α_1 and α_7 are highlighted in orange and blue. The propeller- βA distance is measured between the respective centers of mass (pink circles). Colors of the domains are according to Supplementary Fig. 19B. **(b)** Close-up view of the βA domain with the docked TUDC structure (stick representation)⁵. This complex structure was used to generate other starting structures by modifying the bile acid. Angles measured during the course of the MD simulations: orange: α_1 kink angle; blue: α_7 tilt angle. Mg^{2+} ions are depicted as red spheres; the one at the MIDAS site is labeled M, the one at the ADMIDAS A. **(c–h)** $\alpha_5\beta_1$ integrin and (c) TUDC, (d) *norUDCA*, (e) *TnorUDCA*, (f) GUDC, (g) UDCA, and (h) TC. For clarity, the time course data (left) has been smoothed by Bezier curves. Relative frequencies of the parameters (right) are calculated for the last 100 ns of each simulation. The frequency distributions have been overlaid with Gaussians according to their means and standard deviations (black curves).

Supplementary Fig. 5). PP-2 (250 nmol/l)¹², an inhibitor of Src kinase, largely abolished the *norUDCA*-induced Erk-1/-2 and p38^{MAPK} activation (Fig. 5, Supplementary Fig. 6). Thus, *norUDCA* signaling to both Erk-1/-2 and p38^{MAPK} involves integrins and Src. In order to examine a possible involvement of PI3-K in *norUDCA*-induced

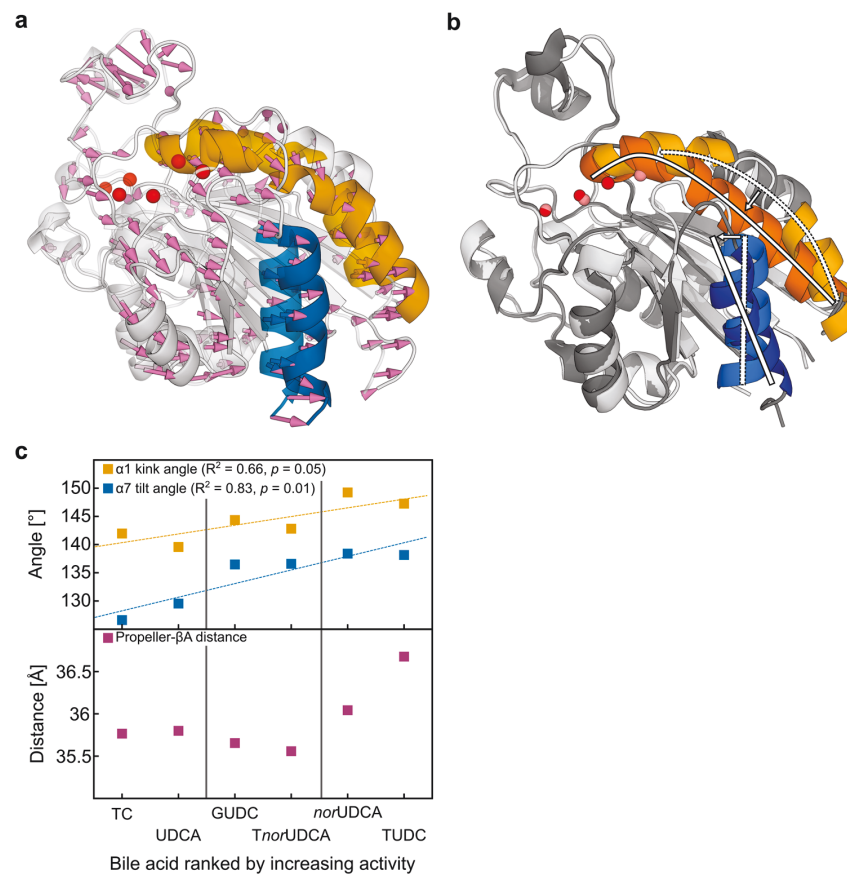


Figure 2. Activation of $\alpha_5\beta_1$ integrins in MD simulations compared to activation of $\alpha_{11b}\beta_3$ integrin in crystal structures. **(a)** Structural overlay of the βA domain (transparent: starting structure; opaque: closest-to-average structure from the last 100 ns) by fitting on the β -propeller domain⁵. Pink arrows denote the positional shift of the βA domain relative to the β -propeller domain, resulting in an increased propeller- βA domain distance. **(b)** Overlay of the closed (lighter colors; PDB ID 3FCS) and open (darker colors; PDB ID 3FCU) conformations of the βA domain in $\alpha_{11b}\beta_3$ integrins. Straightening of the $\alpha 1$ helix (orange) and tilting of the $\alpha 7$ helix (blue) are indicated by white arcs and bars. **(c)** Average of the $\alpha 1$ kink angle (yellow), $\alpha 7$ tilt angle (blue), and β -propeller- βA -domain distance (magenta) over three replicates of MD simulations versus the rank of the bile acids according to their agonist activity towards $\alpha_5\beta_1$ integrin as observed in Fig. 3A-D and ref. ⁵. Dashed lines represent correlation lines; fit parameters are given in the figures. Vertical lines separate the dataset into inactive (left), weakly active (middle), and highly active (right) bile acids.

signaling, the specific inhibitor wortmannin (100 nmol/l) was preperfused. *norUDCA*-induced activation of Erk-1/2 was largely suppressed when wortmannin was present (Fig. 5, Supplementary Fig. 6). In contrast, activation of Src and $p38^{\text{MAPK}}$ was not inhibited by wortmannin (Fig. 5, Supplementary Fig. 6). These findings indicate that Src phosphorylation is upstream of PI3-K activation and that PI3-K is not involved in the signaling towards $p38^{\text{MAPK}}$ activation. In control perfusion experiments without addition of *norUDCA*, no effect on the phosphorylation of Erks, $p38^{\text{MAPK}}$, or Src at tyrosine 418 was found (Supplementary Fig. 7). Next, we examined whether already the initial signaling pathways that usually follow integrin activation are differentially affected by TUDC or *norUDCA*. Perfusion with TUDC (20 $\mu\text{mol/l}$) induced a significant integrin-mediated $\text{FAK}^{\text{Y397-P}}$ autophosphorylation after 10 min (1.87 ± 0.24 -fold amount of $\text{FAK}^{\text{Y397-P}}$) that lasted for up to 30 min compared to livers perfused with normoosmotic medium (Supplementary Fig. 8). In contrast, perfusion with *norUDCA* (20 $\mu\text{mol/l}$) led to an only transient $\text{FAK}^{\text{Y397-P}}$ autophosphorylation that was maximal after 5 min (1.73 ± 0.35 -fold amount of $\text{FAK}^{\text{Y397-P}}$). Similar findings were obtained for other FAK phosphorylation sites, i.e. FAK^{Y407} , $\text{FAK}^{\text{Y576/577}}$, FAK^{Y861} , and FAK^{Y925} (Supplementary Fig. 9).

***norUDCA* does not induce epidermal growth factor receptor (EGFR)-dependent amplification of Erk-1/-2 and $p38^{\text{MAPK}}$ signaling.** Dual activation of Erk-1/-2 and $p38^{\text{MAPK}}$ is involved in the stimulation of canalicular secretion by TUDC. In contrast to TUDC, the effect of *norUDCA* on Erk-1/-2 phosphorylation

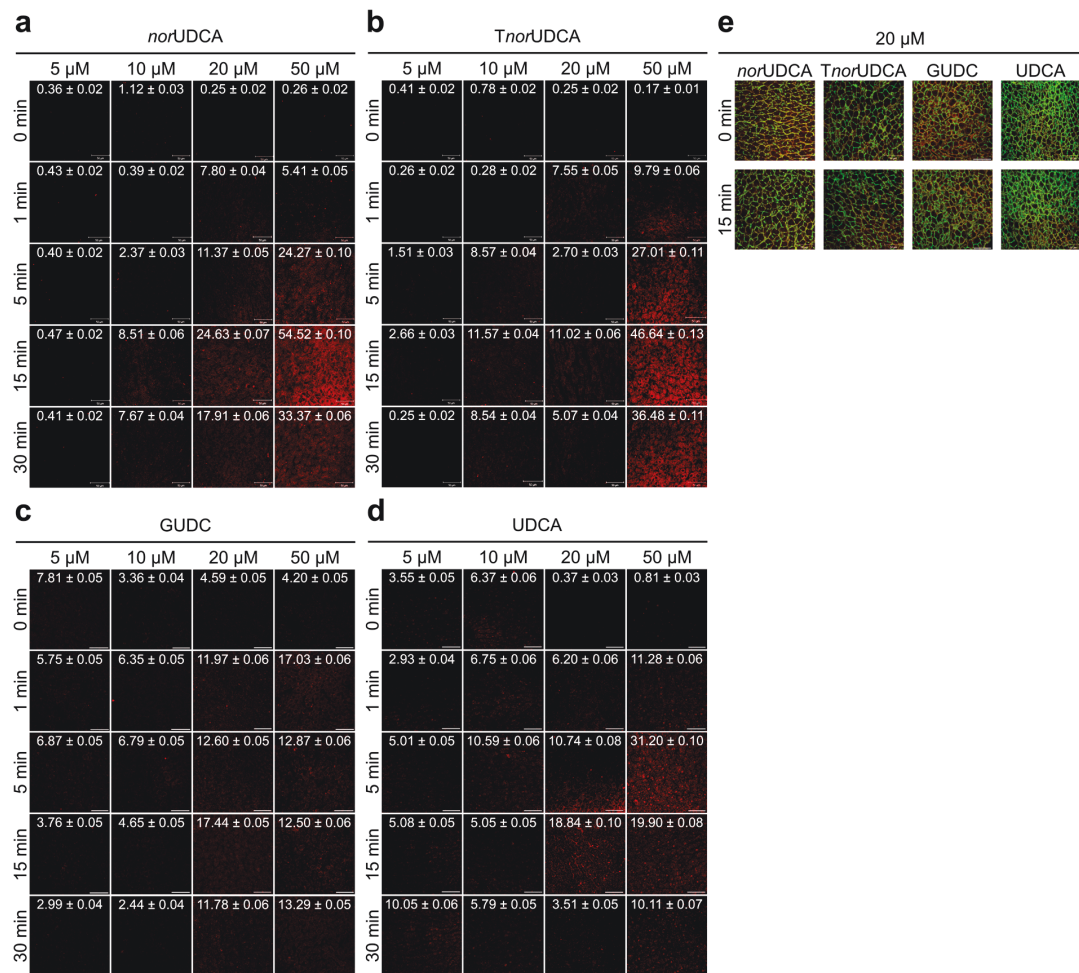


Figure 3. Effect of *norUDCA*, *TnorUDCA*, GUDC, and UDCA on β_1 integrin activation. Rat livers were perfused with (a) *norUDCA*, (b) *TnorUDCA*, (c) GUDC, and (d) UDCA for up to 60 min with the concentrations indicated. Liver samples were immunostained for the active conformation of β_1 integrin (red). The scale bar corresponds to 50 μm . Representative pictures of at least three independent experiments are depicted. To enhance visibility of the images, the white point of all channels in the RGB color space was reduced from the standard value of 255 to a value of 128. For each image, pixel intensities are indicated as average \pm SEM. *NorUDCA* and *TnorUDCA* triggered activation of the β_1 integrin subunit within 15 min, with stronger effects observed with *norUDCA*. In contrast, equimolar concentrations of UDCA and GUDC were ineffective. Like TUDC (Fig. 4)⁵, *norUDCA*-induced β_1 integrin activation occurred primarily in the intracellular compartment of hepatocytes. (e) Staining of total $\alpha_5\beta_1$ integrin (red) and filamentous actin labeled with FITC-coupled phalloidin (green) at $t=0$ min and $t=15$ min after perfusion with *norUDCA*, TUDC, GUDC, and UDCA.

was transient and disappeared largely within 30 min of *norUDCA* exposure (Fig. 6, Supplementary Fig. 10). After 15 min, *norUDCA*-triggered p38^{MAPK} activation was also significantly lower than TUDC-induced phosphorylation of p38^{MAPK} (Fig. 6, Supplementary Fig. 10). Whereas TUDC induced phosphorylation of the EGFR tyrosine residues 845 and 1173, but not of Tyr1045 (Fig. 6, Supplementary Fig. 10), in line with previous data⁵, no activating phosphorylation of EGFR occurred in the presence of *norUDCA* (Fig. 6, Supplementary Fig. 10). TUDC-induced EGFR trans-activation requires an EGFR/c-Src association following Src activation. Compared to the TUDC-triggered Src activation, the *norUDCA*-induced activation of Src was significantly lower (Fig. 7, Supplementary Fig. 11). Immunoprecipitation studies of perfused liver samples revealed that EGFR/c-Src association was absent in *norUDCA*-perfused livers (Fig. 7, Supplementary Fig. 11).

TUDC-induced dual activation of Erk-1/-2 and p38^{MAPK} , and Bsep insertion into the canalicular membrane, are dependent on EGFR phosphorylation. The choleric action of TUDC is largely

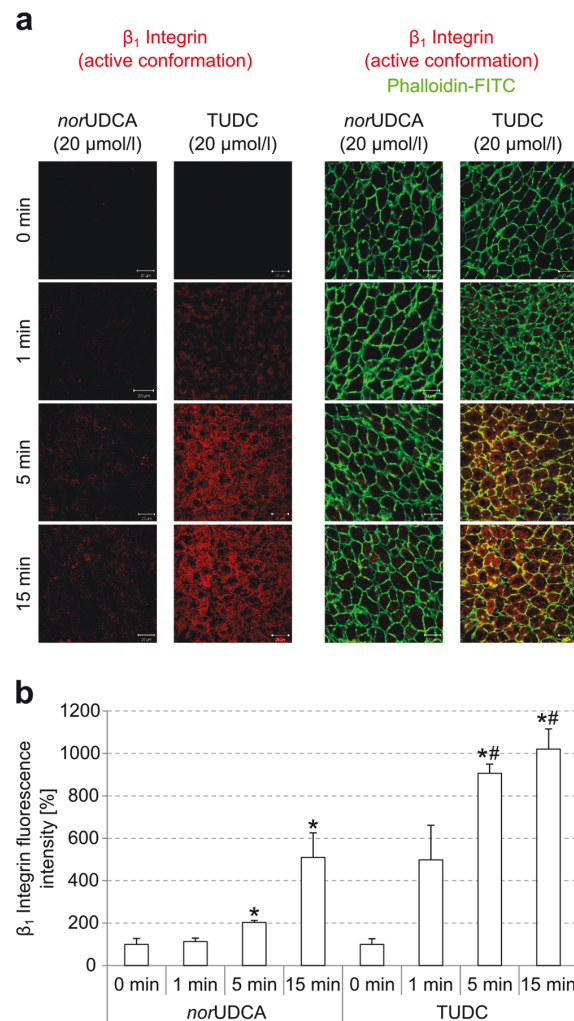


Figure 4. Immunofluorescence staining and quantification of β_1 integrin. (a) Rat livers were perfused with either *norUDCA* or TUDC (20 $\mu\text{mol/l}$ each) for up to 15 min and immunostained for the active β_1 integrin conformation and actin as given under “Experimental Procedures”. The scale bar corresponds to 20 μm . Representative pictures of three independent experiments are depicted. (b) β_1 integrin fluorescence was quantified by using ImageJ analysis software. Whereas TUDC induced β_1 integrin activation within 5 min, *norUDCA* activated β_1 integrins with lower effect. * $p < 0.05$ denotes statistical significance compared with the unstimulated control; # $p < 0.05$ statistical significance between *norUDCA* and TUDC perfusion.

due to an Erk-1/-2- and p38^{MAPK}-dependent insertion of the intracellularly stored canalicular transporters Bsep and Mrp2^{10,13} downstream of EGFR activation. The inhibitor of EGFR tyrosine kinase activity, AG1478, to a large extent abolished the TUDC-induced Erk-1/-2 and p38^{MAPK} activation (Fig. 8, Supplementary Fig. 12). Immunofluorescence stains of the canalicular bile salt transporter Bsep as well as the tight junction complex protein ZO-1, which delineates the bile canaliculi, were analyzed by CLSM and a densitometric analysis procedure^{11,14–16}. In liver tissue, ZO-1 is arranged along two lines, and canalicular transporters within the canalicular membrane are located between these lines (see Supplementary Fig. 13). During control conditions, Bsep was located predominantly in the canalicular membrane (Supplementary Fig. 13). Densitometric analysis after perfusion with TUDC (20 $\mu\text{mol/l}$) revealed significantly different Bsep fluorescence profiles already after 5 min (Fig. 8)¹⁰ ($p < 0.05$; *F*-test for differences in peak heights and variances of Gaussian fits to the data sets) and a narrowing of the fluorescence signal by $0.4 \pm 0.04 \mu\text{m}$, i.e., by ~30%, (determined from the difference in the full width at half maximum (FWHM) values of the fitted Gaussians) after 30 min. Like Erk-1/-2 and p38^{MAPK} activation, Bsep insertion into the canalicular membrane was also inhibited by AG1478 (Fig. 8, $\text{FWHM}_{t=0 \text{ min}}: 1.48 \pm 0.03 \mu\text{m}$ vs. $\text{FWHM}_{t=30 \text{ min}}: 1.51 \pm 0.03 \mu\text{m}$). ZO-1 immunostaining did not change under any condition (see Supplementary Fig. 13).

www.nature.com/scientificreports/

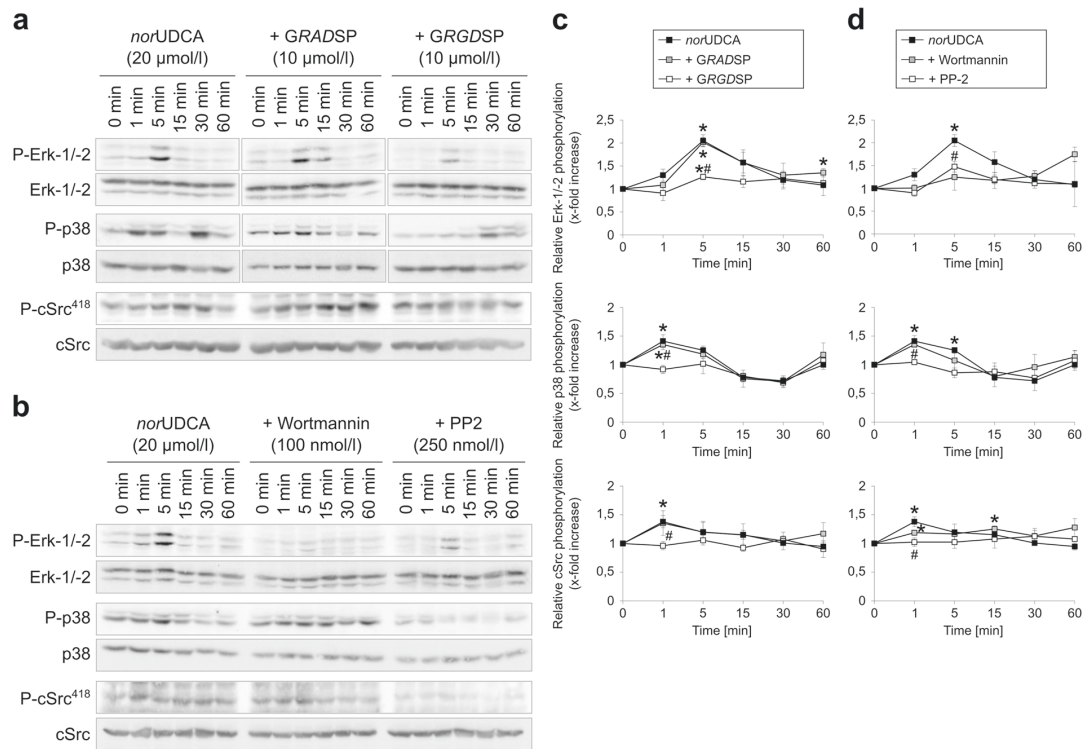


Figure 5. *norUDCA*-induced activation of Erk-1/-2, p38^{MAPK} and Src. Rat livers were perfused with *norUDCA* (20 μmol/l) for up to 60 min. Liver samples were taken at the time points indicated. The integrin antagonistic peptide (*GRGDSP*, 10 μmol/l), the inactive control peptide (*GRADSP*, 10 μmol/l), the PI3-K inhibitor wortmannin (100 nmol/l), and the Src inhibitor PP-2 (250 nmol/l) were added 30 min prior to the addition of *norUDCA*. Activation of Erk-1/-2, p38^{MAPK} and c-Src was analyzed by (a,b) Western blot using specific antibodies and (c,d) subsequent densitometric analysis. Total Erk-1/-2, total p38^{MAPK}, and total c-Src served as respective loading control. Phosphorylation at $t = 0$ min was arbitrarily set as 1. Densitometric analyses (means \pm SEM) and representative blots of at least three independent perfusion experiments are shown.

* $p < 0.05$ statistical significance compared with the unstimulated control. # $p < 0.05$ statistical significance between *norUDCA* in the absence and presence of an inhibitor. *norUDCA* led to a significant activation of Erk-1/-2, p38 MAPK as well as c-Src in the perfused rat liver, which was inhibited by *GRGDSP*, whereas *GRADSP* had no effect. Phosphorylation of Erk-1/-2, p38^{MAPK}, and c-Src was sensitive to PP-2, whereas wortmannin inhibited Erk-1/2 and c-Src activation. Blots were cropped to focus on the area of interest, and full-length blots are presented in Supplementary Figs. 5 and 6.

***norUDCA* induces a transient insertion of Bsep into the canalicular membrane.** TUDC has been shown to increase the capacity for TC excretion into bile^{10,17}. As shown in Fig. 9, *norUDCA* (20 μmol/l) increased bile flow and stimulated a transient TC excretion within the first 10 min of perfusion, whereas the TUDC observed stimulation of TC excretion was prolonged¹⁰. Bsep is responsible for the bile salt-dependent bile flow and transports, among others, conjugates of cholic acid (CA) and chenodeoxycholic acid (CDCA), and the bile acid deoxycholic acid (DCA). In addition, it secretes ursodeoxycholic acid (UDCA) and its conjugates into bile¹⁸. Most of the Bsep immunofluorescence was found between the parallel rows of ZO-1 staining under control conditions, which indicates that Bsep is localized in the canalicular membrane. However, even under control conditions, there was some punctate Bsep staining in the cytosol, mainly in the subcanalicular region, suggestive for the presence of Bsep-containing vesicles inside the cell (Fig. 9). Addition of *norUDCA* (20 μmol/l) results within 5 min in the disappearance of intracellular Bsep, and Bsep staining was almost exclusively found in the canaliculi (Fig. 9). This is reflected in the fluorescence profile, which shows a significant increase in canalicular Bsep fluorescence intensity after 5 min of *norUDCA* addition (Fig. 9); in contrast to TUDC-induced Bsep insertion (Fig. 8), the increase vanished after 30 min, and a punctuated intracellular Bsep staining reappeared (Fig. 9). These findings suggest a *norUDCA*-induced transient translocation and insertion of intracellular Bsep into the canalicular membrane. In contrast, *norUDCA* has no effect on the distribution of the basolateral transporter Ntcp (see Supplementary Fig. 14). Subcellular Ntcp distribution in control and *norUDCA* (20 μmol/l)-perfused livers was analyzed and quantified by CLSM and densitometric fluorescence intensity analysis as described in the Methods section. For labeling of the plasma membrane, liver sections were stained with a specific antibody against the

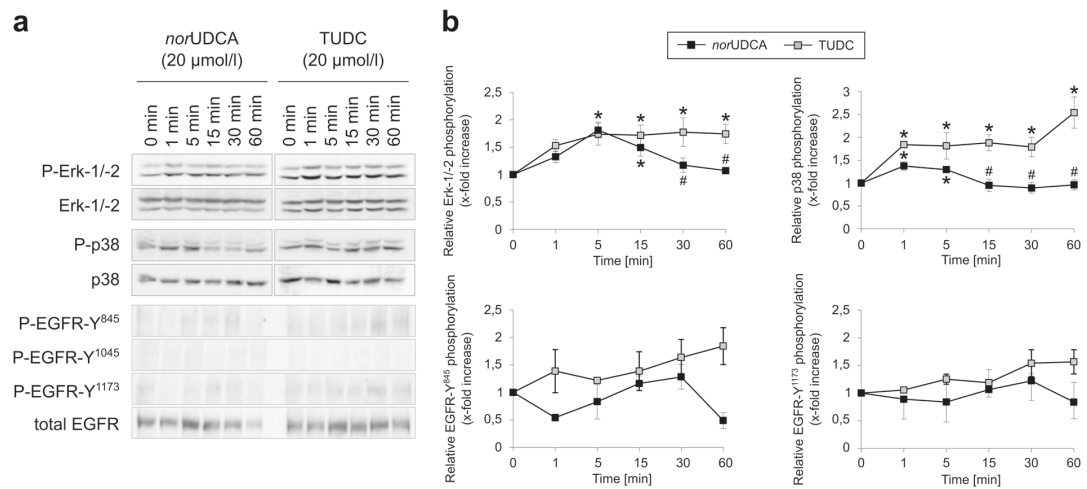


Figure 6. Comparison between *norUDCA*- and TUDC-induced Erk-1/-2, p38^{MAPK} and EGFR activation. Rat livers were perfused with *norUDCA* or TUDC (20 $\mu\text{mol/l}$) each for up to 60 min as described in “Experimental Procedures”. Liver samples were taken at the time points indicated. Phosphorylation of Erk-1/-2, p38^{MAPK}, and EGFR tyrosine residues Tyr⁸⁴⁵, Tyr¹⁰⁴⁵, and Tyr¹¹⁷³ was analyzed by (a) Western blot using specific antibodies and (b) subsequent densitometric analysis (black squares, *norUDCA*; gray squares, TUDC). Total Erk-1/-2, total p38^{MAPK}, and total EGFR served as respective loading controls. Phosphorylation at $t = 0$ was arbitrarily set to 1. Data represent the mean (\pm SEM) of at least three independent experiments; * $p < 0.05$ statistical significance compared with the unstimulated control. # $p < 0.05$ statistical significance between *norUDCA* and TUDC. Blots were cropped to focus on the area of interest, and full-length blots are presented in Supplementary Figure 10. TUDC led to activation of Erk-1/-2, p38^{MAPK}, and EGFR, as indicated by phosphorylation of the EGFR tyrosine residues Tyr⁸⁴⁵ and Tyr¹¹⁷³. *norUDCA* induced a transient Erk-1/-2 phosphorylation and a weak p38^{MAPK} activation. No EGFR activation was observed in *norUDCA*-perfused livers.

plasma membrane marker protein Na⁺/K⁺-ATPase. The immunofluorescence analysis shows that there was no obvious change in Ntcp and Na⁺/K⁺-ATPase distribution at the basolateral membrane within 30 min (see Supplementary Fig. 14).

TC inhibits *norUDCA*-induced $\alpha_5\beta_1$ integrin activation. TC at a concentration of 100 $\mu\text{mol/l}$ had no β_1 integrin-activating activity but interfered with TUDC-induced $\alpha_5\beta_1$ integrin activation⁵. Similarly, when *norUDCA* was added on top of TC (100 $\mu\text{mol/l}$), active β_1 integrin was barely detectable in isolated perfused rat liver (Supplementary Fig. 15). This indicates that TC interferes with *norUDCA*-induced $\alpha_5\beta_1$ integrin activation.

TUDC and *norUDCA* bind directly to $\alpha_5\beta_1$ integrin with similar affinities. Inhibition (i.e., IC₅₀ values) of $\alpha_5\beta_1$ integrin binding to immobilized fibronectin by TUDC and *norUDCA* was determined using a standardized, competitive ELISA-based assay¹⁹ and Cilengitide²⁰ as a control. TUDC and *norUDCA* showed similar IC₅₀ values in the low millimolar range (Table 1, Supplementary Fig. 16), demonstrating a similar binding affinity of both compounds and confirming that the observed activation of $\alpha_5\beta_1$ integrin results from direct binding of the bile acids to the MIDAS site in the integrin head group.

Discussion

In this study, we addressed the question to what extent side chain-modified derivatives of TUDC (*norUDCA*, *TnorUDCA*, GUDC, UDCA) can directly activate $\alpha_5\beta_1$ integrin and whether the signaling events downstream of integrin activation differ from those triggered by TUDC.

Applying all-atom MD simulations, the potential activity of *norUDCA*, *TnorUDCA*, GUDC, and UDCA was assessed on the basis of three geometric parameters, and compared to that of TUDC and TC investigated previously^{5,7}. The geometric parameters were derived from crystal structures of the closed (PDB: 3FCU) and open (PDB: 3FCS) $\alpha_{\text{IIb}}\beta_3$ integrin headpiece (Fig. 2b)²¹, as well as based on previous simulation results⁵: the $\alpha 1$ kink angle, the $\alpha 7$ tilt angle and the propeller- βA distance. Although the crystal structure of the open $\alpha_5\beta_1$ headpiece has remained elusive, it is likely that the conformational changes involved in $\alpha_5\beta_1$ integrin activation are very similar to those observed for other integrin subtypes (Supplementary Fig. 17)^{22,23}. Among the six bile acids tested, TUDC- and *norUDCA*-bound structures displayed on average significantly higher values for all three geometric parameters (Fig. 2, Supplementary Table 3). *TnorUDCA*- and GUDC-bound integrin displayed a larger $\alpha 1$ kink angle and, especially, $\alpha 7$ tilt angle than integrin bound to UDCA and TC but the propeller- βA distance was similar among all four of these bile acids. Hence, we classified TUDC and *norUDCA* as highly activating, *TnorUDCA* and GUDC as weakly activating, and UDCA and TC as inactive or inhibitory ligands, respectively. Note that larger conformational changes in the $\alpha_5\beta_1$ integrin ectodomain, which have been linked to integrin activation^{21,23–25},

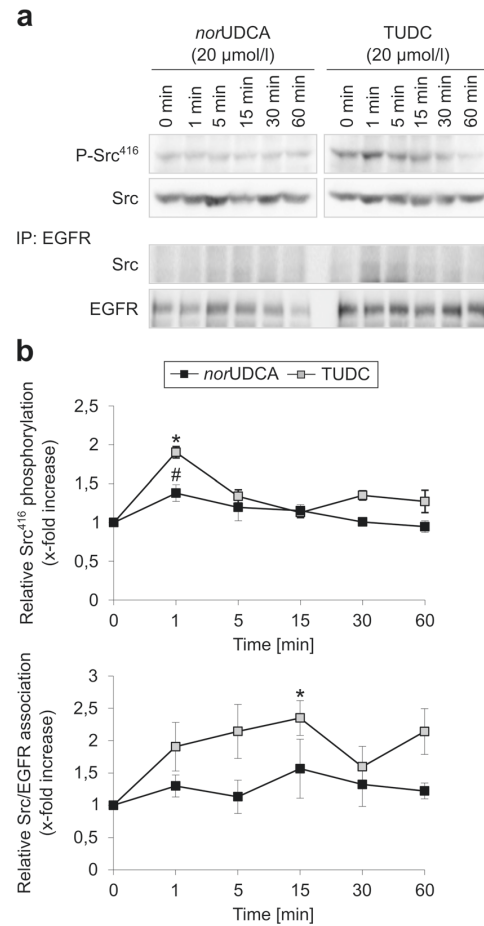


Figure 7. Comparison between *norUDCA*- and TUDC-induced c-Src activation and EGFR/c-Src association. Rat livers were perfused with *norUDCA* or TUDC (20 μmol/l each) for up to 60 min. Liver samples were taken at the time points indicated. Activation of c-Src was analyzed by (a) Western blot using specific antibodies and (b) subsequent densitometric analysis. Total c-Src served as respective loading control. EGFR was immunoprecipitated as described under “Experimental Procedures”. Samples were then analyzed for EGFR/c-Src association by detection of c-Src. Total EGFR served as a loading control. Phosphorylation at $t = 0$ min was set as 1. Densitometric analyses (means \pm SEM) and representative blots of at least three independent perfusion experiments are shown. * $p < 0.05$ statistical significance compared with the unstimulated control. # $p < 0.05$ statistical significance between *norUDCA* and TUDC. Blots were cropped to focus on the area of interest, and full-length blots are presented in Supplementary Figure 11. TUDC led to a significantly more intense phosphorylation of c-Src and EGFR/c-Src association than *norUDCA*.

cannot be expected to be observed during our sub- μ s long MD simulations compared to integrin activation times *in vivo*^{26,27}.

To evaluate the robustness of the predictions from our MD simulations, we correlated the mean values of the three geometric parameters measured in each triplet of MD simulations against the rank of the bile acids in terms of their activity (Fig. 2), as deduced from the amount of immunostained, active β_1 integrin induced by the respective bile acid (Fig. 3). Accordingly, TUDC is the most active bile acid, followed by *norUDCA*, *TnorUDCA*, GUDC, UDCA, and TC. We obtained significant correlations between the average $\alpha 1$ kink angle ($R^2 = 0.66$, $p = 0.05$), or the $\alpha 7$ tilt angle ($R^2 = 0.83$, $p = 0.01$), and the rank (Fig. 2). Thus, the set of geometric parameters used for the analysis of the MD simulations was not only capable to distinguish between active and inactive bile acids but also captured more subtle differences in the activities. Therefore, in future studies, such MD simulations might serve as a “computational assay” to test potential candidate molecules for their ability to activate $\alpha_5\beta_1$ integrin.

As predicted by MD simulations, *norUDCA* caused a dose-dependent activation of $\alpha_5\beta_1$ integrins in hepatocytes (Fig. 3a), and this dose-dependent activation is weaker than the one observed with TUDC (Fig. 4a)⁵. While after addition of TUDC the active conformation of the β_1 integrin subunit becomes markedly visible within

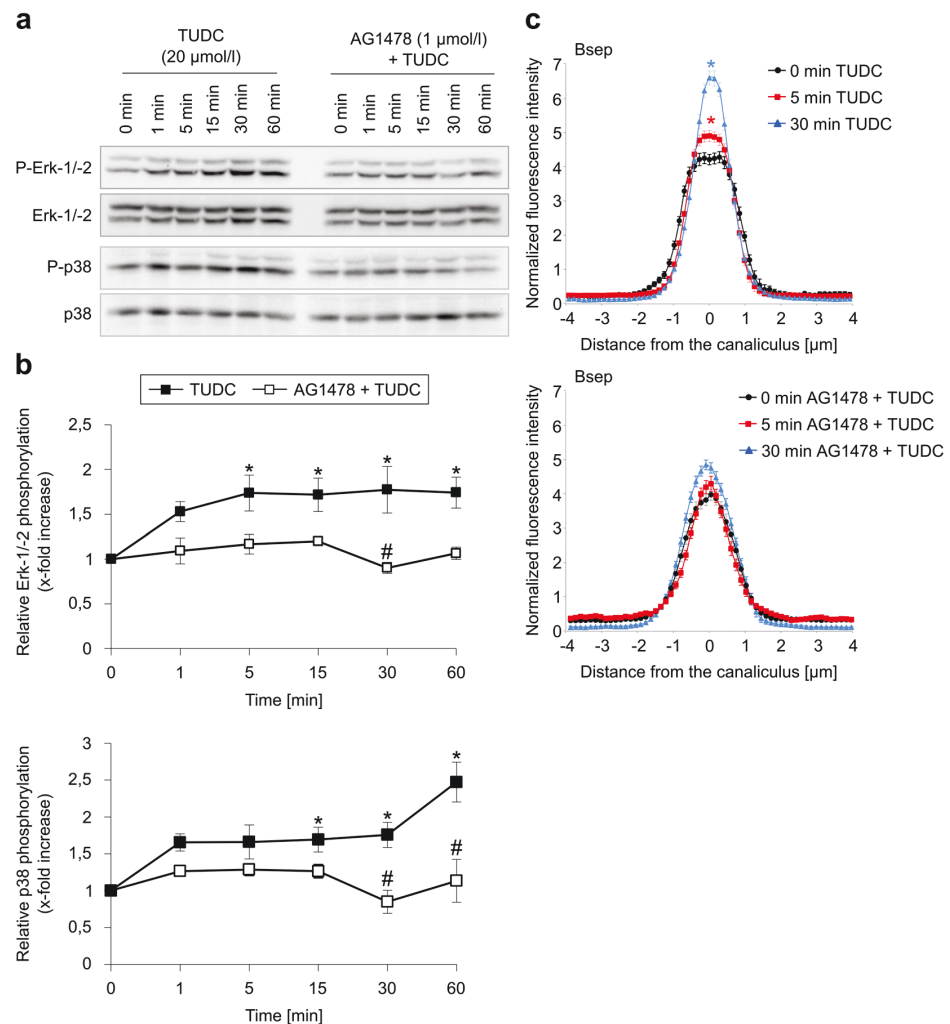


Figure 8. TUDC-induced dual activation of Erk-1/-2 and p38^{MAPK} and Bsep insertion into the canalicular membrane are dependent on EGFR phosphorylation. Rat livers were perfused with TUDC (20 μmol/l) for up to 60 min. When indicated, AG1478 (1 μmol/l) was added 30 min prior to TUDC to the perfusate. (a) Phosphorylation of Erk-1/-2 and p38^{MAPK} was analyzed by use of specific antibodies. Total Erk-1/-2 or p38^{MAPK}, respectively, served as loading controls. (b) Western blots were analyzed densitometrically. Phosphorylation level at $t = 0$ min was set to 1. Representative blots and statistics (mean \pm SEM) of at least three independent perfusion experiments are shown. TUDC induced a significant increase in Erk-1/-2 and p38^{MAPK} phosphorylation ($*p < 0.05$), which was significantly inhibited by AG1478 ($\#p < 0.05$). (c) Cryosections from perfused rat liver were immunostained for Bsep and ZO-1 (see Supplementary Fig. 13), fluorescence images were recorded by confocal LSM (see Supplementary Fig. 13), and analyzed densitometrically. Blots were cropped to focus on the area of interest, and full-length blots are presented in Supplementary Figure 12. Under control conditions (black, $t = 0$ min), Bsep is largely localized between the linear ZO-1, but is also found inside the cells. Addition of TUDC (blue, $t = 5$ min; red, $t = 30$ min) results in the insertion of intracellular Bsep into the canalicular membrane, which was inhibited by AG1478. The fluorescence profiles depicted are statistically significantly ($p < 0.05$) different from each other with respect to variance and peak height.

1 min, *nor*UDCA reaches a similar extent of β_1 integrin activation after 15 min (Figs. 3a and 4a). A standardized, competitive ELISA-based solid-phase assay revealed that TUDC and *nor*UDCA directly bind to the MIDAS site in the integrin head group, confirming that the observed activation of $\alpha_5\beta_1$ integrin results from direct binding of the bile acids, and that the binding affinities of both compounds are similar (Table 1, Supplementary Fig. 16). The latter finding, together with using for TUDC and *nor*UDCA the same concentrations in all experiments, rules out that the different extent of activation of $\alpha_5\beta_1$ integrin by the bile acids is caused by differential occupation of the binding site. The low affinities of both compounds are concordant with the fact that the compounds do not

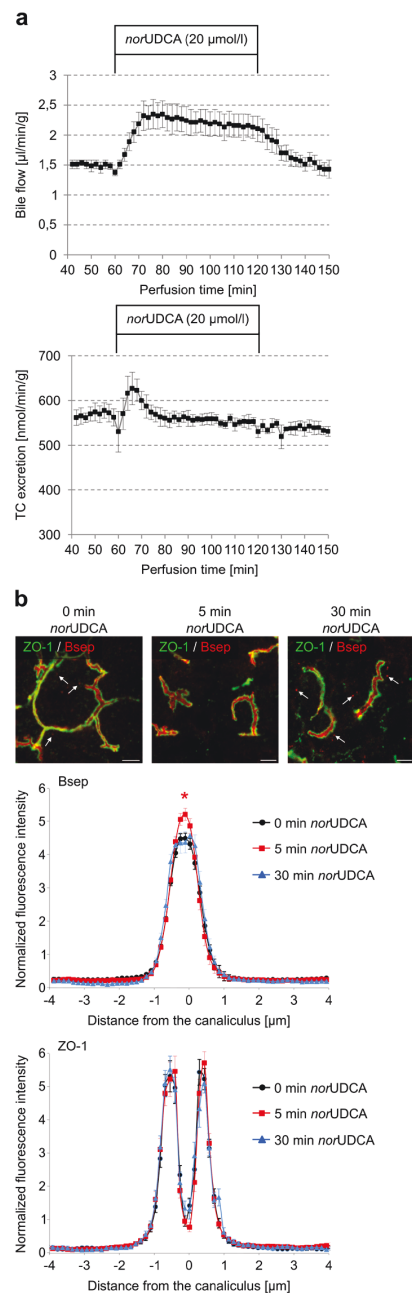


Figure 9. *norUDCA* induced increased bile flow and TC excretion in perfused rat liver. (a) Livers were preperfused in the presence of 10 $\mu\text{mol/l}$ [^3H]TC. Data are given as mean \pm SEM from four different experiments. After a pre-perfusion period of 20 min, *norUDCA* (20 $\mu\text{mol/l}$) was added for 30 min. *norUDCA* increased bile flow over the complete perfusion period and excretion of TC within the first 10 min of perfusion. (b) Cryosections from perfused rat liver were immunostained for Bsep and ZO-1, fluorescence images were recorded by confocal LSM, and analyzed densitometrically. Representative pictures of at least three independent experiments are depicted. The scale bar corresponds to 5 μm . Under control conditions (black, $t = 0$ min), Bsep is largely localized between the linear ZO-1, but is also found inside the cells (white arrows). *norUDCA* (red, $t = 5$ min) led to the insertion of intracellular Bsep into the canalicular membrane. The fluorescence profiles depicted are statistically significantly ($p < 0.05$) different from each other with respect to variance and peak height. Under control conditions, ZO-1 fluorescence profiles show two peaks. Liver perfusion experiments with *norUDCA* resulted in no significant changes of ZO-1 fluorescence profiles with respect to the distance of the peaks and the variance of fluorescence profiles. Means \pm SEM of 30 measurements in each of at least three individual experiments for each condition are shown.

Compound	Sequence	IC ₅₀ ^a
TUDC	—	4.01 (1.80...8.98) ^b
<i>nor</i> UDCA	—	3.93 (1.34...11.5) ^b
Control peptide Cilengitide	c(Arg-Gly-Asp-D-Phe-NMe-Val)	15.4 (14.49...16.36) ^c

Table 1. Affinities of TUDC and *nor*UDCA and the control peptide towards the RGD-recognizing integrin $\alpha_5\beta_1$ obtained from an ELISA-like solid-phase binding assay. ^aThe IC₅₀ values were obtained from a sigmoidal fit to two independent data rows (serial dilutions). The 95% confidence interval is given in brackets. ^bIn mM. ^cIn nM.

activate $\alpha_5\beta_1$ integrin when located in the plasma membrane; extracellular TUDC and *nor*UDCA concentrations in the perfusion experiments were at most 50 μ M. Ntcp-transfected HepG2 cells stimulated with a TUDC concentration of 100 μ M do not show active β_1 integrin in the cell membrane either⁵. In contrast, intracellular bile acid concentrations can reach single digit mM concentrations, as estimated from intracellular bile acid contents for hepatocyte cultures²⁸ or rat hepatoma cells²⁹. The uncertainty in estimating intracellular bile acid concentrations is reflected, however, in that measurements of bile acid concentrations in human liver tissue³⁰ together with those of intracellular water space in rat liver³¹ yielded bile acid concentrations about one order of magnitude smaller than the IC₅₀ values. Finally, with respect to whether the low affinities might be indicative of non-specific binding, note that both RGD peptides and TC inhibit TUDC-induced activation of $\alpha_5\beta_1$ integrin and the signal transduction pathways following integrin activation⁵. Here, we show that this also applies to *nor*UDCA-induced activation of $\alpha_5\beta_1$. We consider particularly the inhibitory effect of TC with respect to TUDC a consequence of competitive antagonism at the MIDAS because we find it difficult to grasp how two bile acids with very similar structures could cause opposing effects via nonspecific mechanisms.

Although the results of our MD simulations indicate that *nor*UDCA is less potent than TUDC with respect to direct $\alpha_5\beta_1$ integrin activation, additional kinetic reasons may contribute as well to this difference. *nor*UDCA, unlike TUDC, is not readily taken up into the hepatocyte via Ntcp or other transport systems³², and the trans-bilayer transport rate of *nor*UDCA, an epimer of *nor*UDCA, is six-fold higher than of CDCA³³, suggesting that *nor*UDCA is passively transported across the sinusoidal membrane. Slow, passive sinusoidal uptake would then be opposed by a fast, active outward transport by a canalicular transporter, presumably Mrp2^{34–37}. Depending on the rates, this situation might prevent concentrating *nor*UDCA inside the hepatocyte. For TUDC, a concentrative uptake into the hepatocyte was proposed as a likely requirement for $\alpha_5\beta_1$ integrin activation⁵.

TUDC-mediated integrin activation is followed by a *sustained* dual activation of Erks and p38^{MAPK}, which is the crucial downstream signaling event towards choleresis^{10,17}. Such a sustained activation of Erks also occurs with lower and higher concentrations (10 μ mol/l and 50 μ mol/l) of TUDC (Supplementary Fig. 18), rendering a concentration effect unlikely. *nor*UDCA also induced a similar but only *transient* dual activation of these MAPKs, which was sensitive to integrin inhibition by an RGD motif-containing hexapeptide (Figs. 5a,b and 6b, Supplementary Figs. 5, 6, 10). This transient MAPK activation might be a consequence of the weaker activation of $\alpha_5\beta_1$ shown above. As *nor*UDCA-induced Erk-1/-2 phosphorylation was not amplified when phosphatases were inhibited with okadaic acid (Supplementary Fig. 4), it is unlikely that the transient MAPK activation by *nor*UDCA is mediated via activation of phosphatases. In this context, note that perfusion with TUDC caused a significant EGFR/c-Src association after 15 min (Fig. 7, Supplementary Fig. 11). By contrast, such an association was not observed following perfusion with *nor*UDCA (Fig. 7, Supplementary Fig. 11). Taken together, our results thus suggest that a c-Src-dependent trans-activation of the EGFR is central for a sustained MAPK activation. At first glance, the suggested sustainer role of EGFR appears contradicted by the observation that AG1478, a selective inhibitor of EGFR tyrosine kinase activity³⁸, abolished the TUDC-induced phosphorylation of Erk and p38^{MAPK} (Fig. 8, Supplementary Fig. 12): If EGFR activation only *sustained* Erk activation, EGFR inhibition should not decrease the extent of TUDC-mediated Erk activation, but only change its time course. However, AG1478 treatment has been shown to compromise basal levels of EGFR phosphorylation³⁹, and such a basal EGFR activity was suggested to be required for proper MAPK signaling⁴⁰.

In line with the transient or sustained character of *nor*UDCA-mediated and TUDC-mediated MAPK activation, respectively, *nor*UDCA induced only a transient insertion of intracellular Bsep into the canalicular membrane, whereas TUDC-induced insertion of Bsep was sustained (Figs. 8 and 9). In line with this, *nor*UDCA only transiently increased TC excretion into bile (Fig. 9a). However, as expected, *nor*UDCA increased bile flow in a sustained way due to *nor*UDCA excretion into bile and induction of a bicarbonate-rich hypercholeresis^{41,42}. Dual activation of Erks and p38^{MAPK} is required for the TUDC-induced stimulation of Bsep insertion into the canalicular membrane¹⁰. In line with this, inhibition of EGFR tyrosine kinase activity by AG1478 prevented Bsep insertion during perfusion with TUDC (Fig. 8c), again suggesting that (at least basal) EGFR activity is an essential requirement for dual MAPK signaling towards choleresis.

Inhibition of PI3-K by wortmannin abolished the *nor*UDCA-induced phosphorylation of Erks, but not of p38^{MAPK}, while inhibition of c-Src by PP-2 abolished phosphorylation of Erks and p38^{MAPK}, suggesting that c-Src activation lies upstream of PI3-K activation. The PI3-K/Ras/Erk pathway has been described as essential for the choleric effect of TUDC⁴³. Whether c-Src directly activates PI3-K⁴⁴, or indirectly via EGFR, was not addressed in this study. An earlier study suggested that genistein-sensitive tyrosine kinases such as EGFR are not involved in the activation of the PI3-K/Ras/Erk pathway by TUDC⁴³. However, whether the Ras/Erk pathway becomes PI3-K-dependent also depends on the extent of EGFR activation⁴⁵.

Notably, in an earlier study with TUDC⁷, inhibition of c-Src did not prevent Erk-1/-2 activation, but only delayed it by ~8 min. Hence, in view of the above results, inhibition of c-Src activity by PP-2 seems to prevent Erk-1/-2 activation only when *nor*UDCA is used as an integrin agonist. Based on our and literature data, we therefore

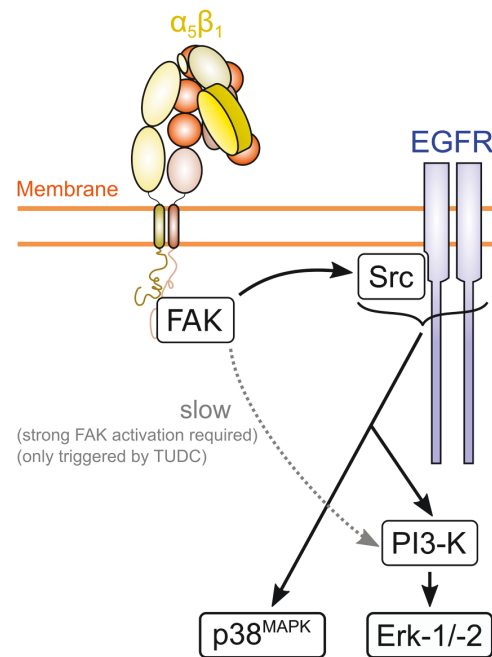


Figure 10. Model of $\alpha_5\beta_1$ integrin activation-dependent differential bile acid signaling. Activation of $\alpha_5\beta_1$ integrin with the less efficacious *norUDCA* results in the formation of $\text{FAK}^{\text{Y397-P}}$, which leads to c-Src- and PI3-K-dependent Erk-1/2 activation. When $\alpha_5\beta_1$ integrin is activated by the more efficacious TUDC, higher levels of $\text{FAK}^{\text{Y397-P}}$ result, which, in addition, trigger a slower activation of Erk-1/-2 via PI3-K in a c-Src-independent manner⁷.

suggest the following ligand-dependent selectivity for signaling pathways induced by $\alpha_5\beta_1$ integrin (Fig. 10): One of the first steps in integrin-mediated signaling is the recruitment of focal adhesion kinase (FAK)⁴⁶ and its subsequent autophosphorylation, an event also observed during TUDC-mediated activation of $\alpha_5\beta_1$ integrin⁷. Levels of autophosphorylated FAK ($\text{FAK}^{\text{Y397-P}}$) were shown to increase linearly with the amount of fibronectin-bound (i.e. active, signaling-competent) $\alpha_5\beta_1$ ⁴⁷. Thus, a highly efficacious integrin activation as observed with TUDC⁵ would result in high $\text{FAK}^{\text{Y397-P}}$ levels, whereas a less efficacious integrin activation as observed with *norUDCA* (this study) would result in lower $\text{FAK}^{\text{Y397-P}}$ levels, as confirmed by densitometric analysis (Supplementary Fig. 8). $\text{FAK}^{\text{Y397-P}}$ activates c-Src^{48,49}, which in turn phosphorylates EGFR, and both the activated c-Src and EGFR mediate PI3-K activation^{48,50} and subsequent phosphorylation of Erk-1/-2. However, $\text{FAK}^{\text{Y397-P}}$ can also *directly* activate PI3-K, independent of c-Src and the EGFR⁴⁸. We now speculate that this direct, FAK-mediated activation of PI3-K is slower than the c-Src and EGFR-mediated PI3-K activation, and that only high $\text{FAK}^{\text{Y397-P}}$ levels trigger this slow pathway. Hence, even when c-Src activity is inhibited by PP-2, a highly efficacious integrin activation by TUDC would lead to a pronounced FAK autophosphorylation and rescue Erk-1/-2 phosphorylation via a direct PI3-K activation, albeit with a time delay, as observed previously⁷. In contrast, a less efficacious integrin activation by *norUDCA* would lead to less FAK autophosphorylation and, thus, require activated c-Src in order to switch on the then necessary PI3-K signal to activate Erk-1/-2, which would occur more rapidly (this study). According to this model, inhibition of EGFR activity by AG1478 should not abolish the Erk response, if TUDC-mediated PI3-K activation occurred via the slow pathway. Regarding the above observation that AG1478 did abolish the TUDC-induced phosphorylation of Erk and p38^{MAPK} , we can only speculate at present that apparently (at least a basal) EGFR activity is required for PI3-K to properly function in this pathway, although the details of this interplay remain elusive.

Taken together, we demonstrated - to our knowledge for the first time - that *norUDCA* directly activates $\alpha_5\beta_1$ integrins in hepatocytes and triggers short-term cholerisis via a transient activation of MAPKs followed by a transient insertion of Bsep into the canalicular membrane in addition to the known bicarbonate-rich hypercholerisis. Furthermore, we provide evidence that TUDC and *norUDCA* exert a functional selectivity for certain signal transduction pathways in $\alpha_5\beta_1$ integrin, a property - to our knowledge - not yet described for ligands interacting with integrins lacking an αI domain³. This functional selectivity may also provide a rationale for the differential therapeutic use of UDCA (which *in vivo* is rapidly conjugated to TUDC) and *norUDCA* (which is resistant to amidation with taurine) in primary biliary cholangitis (PBC) and primary sclerosing cholangitis (PSC), respectively⁴². Although both compounds trigger hypercholerisis, the underlying mechanisms are different. TUDC induces cholerisis by stimulating hepatocellular bile acid secretion, whereas *norUDCA* induces a bicarbonate-rich hypercholerisis by cholehepatic shunting, but has no effect on hepatocellular bile acid

secretion^{41,51}. This may explain why *norUDCA* is superior to UDCA in the treatment of sclerosing cholangitis in *Mdr2* knockout mice⁵².

Materials and Methods

Materials. The materials used were purchased as follows: Ocadaic acid was from Enzo Life Sciences (Lörrach, Germany). PP-2, H-Gly-Arg-Gly-Asp-Ser-Pro-OH (*GRGDSP*), and H-Gly-Arg-Ala-Asp-Ser-Pro-OH (*GRADSP*) were from Merck-Millipore (Darmstadt, Germany), FITC-coupled phalloidin, collagenase, insulin, and TUDC from Sigma Aldrich (Munich, Germany), penicillin/streptomycin and Fluoromount-G from Tocris/Biozol (Eching, Germany), fetal bovine serum (FBS) from Life Technologies GmbH (Darmstadt, Germany), cOmplete™ protease inhibitor cocktail tablets and PhosSTOP™ phosphatase inhibitor cocktail tablets were from Roche Diagnostics (Mannheim, Germany) and William's Medium E from Biochrom (Berlin, Germany). *norUDCA* was kindly provided by Dr. Falk Pharma (Freiburg, Germany). The Bsep antibody (K24)⁵³ and the Ntcp antibody (K4)⁵⁴ were generous gifts from Prof. Dr. B. Steiger (Kantonsspital Zürich, Switzerland). Antibodies recognizing zona occludens-1 (ZO-1, #33–9100), phospho-EGFR Tyr⁸⁴⁵ (#44–784), phospho-EGFR Tyr¹¹⁷³ (#44–794 G), and c-Src (#44–656) were from Life Technologies GmbH (Darmstadt, Germany). The antibodies raised against the $\alpha_5\beta_1$ integrin dimer (AB1950) and the β_1 integrin subunit active conformation (#MAB2079Z), phospho-Erk-1/-2 (#9106), phospho-p38^{MAPK} (#9211), p38^{MAPK} (#9228), phospho-EGFR Tyr¹⁰⁴⁵ (#2237), phospho-Src-Tyr⁴¹⁸ (#2101), phospho-FAK Tyr⁹²⁵ (#3284), and phospho-FAK Tyr^{576/577} (#3281) were from Cell Signaling Technology, Inc. (Danvers, USA), against Erk-1/-2 (#06–182), EGFR (#06–847, Western blot, WB), Na⁺/K⁺-ATPase (#05–369), Cy3-conjugated donkey anti-rabbit IgG (#AP182C), and FITC-conjugated donkey anti-mouse IgG (#AP192C) from Merck-Millipore (Darmstadt, Germany). The antibody against EGFR (sc-03) for immunoprecipitation (IP) studies was from Santa Cruz Biotechnology (Heidelberg, Germany). The polyclonal antibodies against phospho-FAK Tyr³⁹⁷ (#44–650 G), phospho-FAK Tyr⁴⁰⁷ (#44–624 G) and phospho-FAK Tyr⁸⁶¹ (#44–626 G) were purchased from Thermo Fischer Scientific (Waltham, Massachusetts, USA). The monoclonal antibody against glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was from Biodesign International (Saco, Maine, USA). Horseradish peroxidase-conjugated anti-mouse IgG (#1706516) and anti-rabbit IgG (#1721019) were from Bio-Rad Laboratories (Munich, Germany) and Dako (Hamburg, Germany). All other chemicals were from Merck-Millipore (Darmstadt, Germany) at the highest quality available.

Generation of $\alpha_5\beta_1$ integrin-bile acid complex structures and molecular dynamics simulations. A detailed description of how the starting structures for the MD simulations of $\alpha_5\beta_1$ integrin bound to either TUDC, *norUDCA*, *TnorUDCA*, GUDC, UDCA, or TC were generated and how MD simulations of in total 3.6 μ s length of these systems were performed is provided in the Supplementary Text.

Analysis of trajectories from molecular dynamics simulations. MD trajectories were visually inspected for conformational changes in VMD⁵⁵. Conformational changes that may result in integrin activation were evaluated based on three geometric parameters (Fig. 1a,b): Straightening of the $\alpha 1$ helix, tilting of the $\alpha 7$ helix, and the distance between the β -propeller domain in the α -subunit and the βA domain in the β -subunit. Straightening of the $\alpha 1$ helix was monitored through an increase of its kink angle (Fig. 1b). During $\alpha_{\text{ITB}}\beta_3$ integrin activation, this angle increases from $\sim 144^\circ$ to $\sim 166^\circ$, as observed in crystal structures of the closed (PDB: 3FCU) and open (PDB: 3FCS) integrin²¹. Tilting of $\alpha 7$ was measured as the angle between the three points 1) ion at the “Adjacent to MIDAS” (ADMIDAS) site, 2) center of mass of the C_α atoms of the first four residues of the $\alpha 7$ helix, and 3) center of mass of the C_α atoms of the last four residues of the $\alpha 7$ helix (Fig. 1B). Upon activation of $\alpha_{\text{ITB}}\beta_3$ integrins, the $\alpha 7$ helix pivots laterally⁵⁶ (increase of the $\alpha 7$ tilt angle from $\sim 128^\circ$ to $\sim 133^\circ$), accompanied by a marked increase of B-factors in the region of the $\alpha 7$ helix²³ (Supplementary Fig. 17). A larger tilt angle of the $\alpha 7$ helix thus represents a defined, activating conformational change, as does the observation of a higher helix mobility, which is required for subsequent steps in integrin activation. Finally, the distance of the centers of mass of the propeller domain in the α subunit and the βA domain in the β subunit was measured, as it had been shown to increase during TUDC-induced $\alpha_5\beta_1$ integrin activation⁵. All MD trajectory analyses were performed using the programs ptraj from AmberTools 1.5 or cpptraj from AmberTools13⁵⁷.

Liver perfusion. Livers from male Wistar rats (140–160 g) were perfused in a non-recirculating manner as described previously⁵⁸. As a perfusion medium, the bicarbonate-buffered Krebs-Henseleit saline plus L-lactate (2.1 mmol/l) and pyruvate (0.3 mmol/l) gassed with 5% CO₂ and 95% O₂ at 37 °C was used (305 mosmol/l, normoosmotic). Inhibitors and bile acids were added to the influent perfusate by dissolution into the Krebs-Henseleit buffer. Viability of the perfused livers was assessed by measuring lactate dehydrogenase leakage into the perfusate. The portal pressure, the effluent K⁺ concentration, and pH were continuously monitored. In bile formation experiments, livers were perfused with 10 μ mol/l [³H] taurocholate (1 μ Ci/l). Bile was collected at intervals of 2 min. Bile flow was assessed by gravimetry, assuming a specific mass of 1 g/ml. Taurocholate excretion into bile was determined by liquid scintillation counting of the radioactivity present in bile, based on the specific radioactivity of [³H] taurocholate in influent perfusate. To wash out endogenously formed bile acids and obtain a steady-state TC excretion, livers were preperfused for 20 min before experimental maneuvers were started. All experiments were approved by the responsible local authorities of the “Zentrale Einrichtung für Tierforschung und wissenschaftliche Tierschutzaufgaben” (ZETT) of the University of Düsseldorf and the “Landesamt für Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen” (LANUV, NRW) (file number: 84.02-04.2012A214). We confirm that all experiments were performed in accordance with relevant guidelines and regulations.

Immunofluorescence staining. Immunofluorescence staining was performed as described before^{5,16} (see Supplementary Text for a detailed protocol).

Densitometric fluorescence intensity analysis. (see Supplementary Text for a detailed protocol).

Immunoblot analysis. Immunoblot analysis was performed as described before¹⁶ (see Supplementary Text for a detailed protocol).

Immunoprecipitation. Immunoprecipitation was performed as described before¹⁶ (see Supplementary Text for a detailed protocol).

Integrin binding assay. The affinity and selectivity of bile acid derivatives were determined by a solid-phase binding assay applying a previously described protocol¹⁹ that involves coated extracellular matrix proteins and soluble integrins. Cilengitide²⁰ (*c*(f(NMe)VRGD) ($\alpha_5\beta_1$; IC₅₀ = 15.4 nM) was used as internal standard. Flat-bottomed 96-well ELISA plates (BRAND, Wertheim, Germany) were coated overnight at 4 °C with ECM protein (1) (100 μ L per well) in carbonate buffer (15 mM Na₂CO₃, 35 mM NaHCO₃, pH 9.6). Each well was then washed with PBS-T buffer (phosphate-buffered saline/Tween 20, 137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 2 mM KH₂PO₄, 0.01% Tween 20, pH 7.4; 3 \times 200 μ L) and blocked for 1 h at room temperature (RT) with TS-B buffer (Tris-saline/bovine serum albumin (BSA) buffer, 20 mM Tris-HCl, 150 mM NaCl, 1 mM CaCl₂, 1 mM MgCl₂, 1 mM MnCl₂, pH 7.5, 1% BSA; 150 μ L/well). Meanwhile, a dilution series of the compound and internal standard was prepared in an extra plate, ranging from 66 mM to 58 μ M. After washing the assay plate three times with PBS-T (200 μ L), 50 μ L aliquots of the dilution series were transferred to each well from B-G in six appropriate concentrations. Well A was filled with 100 μ L of TS-B buffer (blank), and well H was filled with 50 μ L of TS-B buffer. Then, 50 μ L of a solution of human integrin (2) in TS-B buffer was transferred to wells H-B and incubated for 1 h at RT. The plate was washed three times with PBS-T buffer, and then primary antibody (3) (100 μ L per well) was added to the plate. After incubation for 1 h at RT, the plate was washed three times with PBS-T. Then, secondary peroxidase-conjugated antibody (4) (100 μ L/well) was added to the plate and incubated for 45 min at RT. The plate was then washed three times with PBS-T, developed by the addition of SeramunBlau (50 μ L/well, Seramun Diagnostic GmbH, Heidesee, Germany) and incubated for approx. 1 min at RT in the dark. The reaction was stopped with 3 M H₂SO₄ (50 μ L/well), and the absorbance was measured at 450 nm with a plate reader (infinite M200 Pro, TECAN). The IC₅₀ value (with 95% confidence interval) of each compound resulted from a sigmoidal fit to 32 data points, obtained from two serial dilution rows, by using the GraphPad Prism software package. All IC₅₀ values determined were referenced to the affinity of the internal standard.

- (1) 0.5 μ g mL⁻¹, human fibronectin, Sigma-Aldrich.
- (2) 2.0 μ g mL⁻¹, human $\alpha_5\beta_1$ -integrin, R&D.
- (3) 1.0 μ g mL⁻¹, mouse anti-human CD49e, BD Biosciences.
- (4) 2.0 μ g mL⁻¹, anti-mouse IgG-POD, Sigma-Aldrich.

Statistical analysis. Statistical analysis of the data from MD simulations was performed in R⁵⁹. Mean values and their respective standard errors were computed using the last 100 ns of each simulation. The statistical significance of differences in simulation means was assessed by Student's *t*-test. $p < 0.05$ was considered statistically significant.

As to experimental work, unless stated otherwise in the respective subsections of the Materials and Methods section, results from at least three independent experiments are expressed as mean values \pm SEM. *n* refers to the number of independent experiments. Differences between experimental groups were analyzed by Student's *t*-test, one-way analysis of variance following Dunnett's multiple comparison post hoc test, or two-way analysis of variance following Bonferroni's multiple comparison post hoc test where appropriate (GraphPad Prism; GraphPad, La Jolla, USA; Microsoft Excel for Windows). $p < 0.05$ was considered statistically significant.

Data availability

All data generated or analyzed during this study are included in this published article (and its Supplementary Information file).

Received: 26 July 2018; Accepted: 2 March 2020;

Published online: 02 April 2020

References

1. Urban, J. D. *et al.* Functional selectivity and classical concepts of quantitative pharmacology. *J. Pharmacol. Exp. Ther.* **320**, 1–13, <https://doi.org/10.1124/jpet.106.104463> (2007).
2. Violin, J. D. & Lefkowitz, R. J. Beta-arrestin-biased ligands at seven-transmembrane receptors. *Trends Pharmacol. Sci.* **28**, 416–422, <https://doi.org/10.1016/j.tips.2007.06.006> (2007).
3. Simon, D. I. Opening the field of integrin biology to “biased agonism”. *Circ. Res.* **109**, 1199–1201, <https://doi.org/10.1161/CIRCRESAHA.111.257980> (2011).
4. Wolf, D. *et al.* Binding of CD40L to Mac-1's I-domain involves the EQLKSKTL motif and mediates leukocyte recruitment and atherosclerosis—but does not affect immunity and thrombosis in mice. *Circ. Res.* **109**, 1269–1279, <https://doi.org/10.1161/CIRCRESAHA.111.247684> (2011).
5. Gohlke, H., Schmitz, B., Sommerfeld, A., Reinehr, R. & Häussinger, D. $\alpha_5\beta_1$ -integrins are sensors for tauroursodeoxycholic acid in hepatocytes. *Hepatology* **57**, 1117–1129, <https://doi.org/10.1002/hep.25992> (2013).
6. Volpes, R., van den Oord, J. J. & Desmet, V. J. Integrins as differential cell lineage markers of primary liver tumors. *Am. J. Pathol.* **142**, 1483–1492 (1993).
7. Häussinger, D. *et al.* Involvement of Integrins and Src in Tauroursodeoxycholate-Induced and Swelling-Induced Cholelithiasis. *Gastroenterology* **124**, 1476–1487, [https://doi.org/10.1016/S0016-5085\(03\)00274-9](https://doi.org/10.1016/S0016-5085(03)00274-9) (2003).

8. Schliess, F., Reissmann, R., Reinehr, R., vom Dahl, S. & Häussinger, D. Involvement of Integrins and Src in Insulin Signaling toward Autophagic Proteolysis in Rat Liver. *J. Biol. Chem.* **279**, 21294–21301, <https://doi.org/10.1074/jbc.M313901200> (2004).
9. vom Dahl, S. *et al.* Involvement of Integrins in Osmosensing and Signaling toward Autophagic Proteolysis in Rat Liver. *J. Biol. Chem.* **278**, 27088–27095, <https://doi.org/10.1074/jbc.M210699200> (2003).
10. Kurz, A. K., Graf, D., Schmitt, M., vom Dahl, S. & Häussinger, D. Tauroursodesoxycholate-Induced Cholestasis Involves p38(MAPK) Activation and Translocation of the Bile Salt Export Pump in Rats. *Gastroenterology* **121**, 407–419, [gast.2001.26262](https://doi.org/10.1053/j.gastro.2001.26262) (2001).
11. Schmitt, M., Kubitz, R., Lizun, S., Wettstein, M. & Häussinger, D. Regulation of the Dynamic Localization of the Rat Bsep Gene-Encoded Bile Salt Export Pump by Anisoosmolarity. *Hepatology* **33**, 509–518, <https://doi.org/10.1053/jhep.2001.22648> (2001).
12. Hanke, J. H. *et al.* Discovery of a novel, potent, and Src family-selective tyrosine kinase inhibitor. Study of Lck- and FynT-dependent T cell activation. *J. Biol. Chem.* **271**, 695–701 (1996).
13. Beuers, U. *et al.* Tauroursodeoxycholic acid inserts the apical conjugate export pump, Mrp2, into canalicular membranes and stimulates organic anion secretion by protein kinase C-dependent mechanisms in cholestatic rat liver. *Hepatology* **33**, 1206–1216, <https://doi.org/10.1053/jhep.2001.24034> (2001).
14. Kubitz, R., D'Urso, D., Keppler, D. & Häussinger, D. Osmodependent Dynamic Localization of the Multidrug Resistance Protein 2 in the Rat Hepatocyte Canalicular Membrane. *Gastroenterology* **113**, 1438–1442, [gast.1997.v113.pm9352844](https://doi.org/10.1053/j.gastro.1997.v113.pm9352844) (1997).
15. Cantore, M., Reinehr, R., Sommerfeld, A., Becker, M. & Häussinger, D. The Src Family Kinase Fyn Mediates Hyperosmolarity-Induced Mrp2 and Bsep Retrieval from Canalicular Membrane. *J. Biol. Chem.* **286**, 45014–45029, <https://doi.org/10.1074/jbc.M111.292896> (2011).
16. Sommerfeld, A., Mayer, P. G. K., Cantore, M. & Häussinger, D. Regulation of Plasma Membrane Localization of the Na⁺-Taurocholate Cotransporting Polypeptide (Ntcp) by Hyperosmolarity and Tauroursodeoxycholate. *J. Biol. Chem.* **290**, 24237–24254, <https://doi.org/10.1074/jbc.M115.666883> (2015).
17. Schliess, F., Kurz, A. K., vom Dahl, S. & Häussinger, D. Mitogen-Activated Protein Kinases Mediate the Stimulation of Bile Acid Secretion by Tauroursodeoxycholate in Rat Liver. *Gastroenterology* **113**, 1306–1314, <https://doi.org/10.1053/gast.1997.v113.pm9322526> (1997).
18. Steiger, B. Recent insights into the function and regulation of the bile salt export pump (ABCB11). *Curr. Opin. Lipidol.* **20**, 176–181, <https://doi.org/10.1097/MOL.0b013e32832b677c> (2009).
19. Bochen, A. *et al.* Biselectivity of isoDGR peptides for fibronectin binding integrin subtypes alpha5beta1 and alphavbeta6: conformational control through flanking amino acids. *J. Med. Chem.* **56**, 1509–1519, <https://doi.org/10.1021/jm301221x> (2013).
20. Mas-Moruno, C., Rechenmacher, F. & Kessler, H. Cilengitide: the first anti-angiogenic small molecule drug candidate design, synthesis and clinical evaluation. *Anticancer Agents Med. Chem.* **10**, 753–768 (2010).
21. Zhu, J. *et al.* Structure of a Complete Integrin Ectodomain in a Physiologic Resting State and Activation and Deactivation by Applied Forces. *Mol. Cell* **32**, 849–861, <https://doi.org/10.1016/j.molcel.2008.11.018> (2008).
22. Van Agthoven, J. F. *et al.* Structural basis for pure antagonism of integrin α V β 3 by a high-affinity form of fibronectin. *Nat. Struct. Mol. Biol.* **21**, 383–388, <https://doi.org/10.1038/nsmb.2797> (2014).
23. Zhu, J., Zhu, J. & Springer, T. A. Complete integrin headpiece opening in eight steps. *J. Cell Biol.* **201**, 1053–1068, <https://doi.org/10.1083/jcb.201212037> (2013).
24. Puklin-Faucher, E. & Vogel, V. Integrin activation dynamics between the RGD-binding site and the headpiece hinge. *J. Biol. Chem.* **284**, 36557–36568, <https://doi.org/10.1074/jbc.M109.041194> (2009).
25. Puklin-Faucher, E., Gao, M., Schulten, K. & Vogel, V. How the headpiece hinge angle is opened: New insights into the dynamics of integrin activation. *J. Cell Biol.* **175**, 349–360, <https://doi.org/10.1083/jcb.200602071> (2006).
26. Puklin-Faucher, E. & Sheetz, M. P. The mechanical integrin cycle. *J. Cell Sci.* **122**, 179–186, <https://doi.org/10.1242/jcs.042127> (2009).
27. Craig, D., Gao, M., Schulten, K. & Vogel, V. Structural insights into how the MIDAS ion stabilizes integrin binding to an RGD peptide under force. *Structure* **12**, 2049–2058, <https://doi.org/10.1016/j.str.2004.09.009> (2004).
28. Nakazawa, T., Hoshino, M., Hayakawa, T., Tanaka, A. & Ohiwa, T. Vasopressin reduces taurochenodeoxycholate-induced hepatotoxicity by lowering the hepatocyte taurochenodeoxycholate content. *J. Hepatol.* **25**, 739–747, [https://doi.org/10.1016/s0168-8278\(96\)80247-9](https://doi.org/10.1016/s0168-8278(96)80247-9) (1996).
29. Agellon, L. B. & Torchia, E. C. Intracellular transport of bile acids. *Biochim. Biophys. Acta* **1486**, 198–209, [https://doi.org/10.1016/s1388-1981\(00\)00057-3](https://doi.org/10.1016/s1388-1981(00)00057-3) (2000).
30. Setchell, K. D. *et al.* Bile acid concentrations in human and rat liver tissue and in hepatocyte nuclei. *Gastroenterology* **112**, 226–235, [https://doi.org/10.1016/s0016-5085\(97\)70239-7](https://doi.org/10.1016/s0016-5085(97)70239-7) (1997).
31. vom Dahl, S., Hallbrucker, C., Lang, F. & Häussinger, D. Regulation of cell volume in the perfused rat liver by hormones. *Biochem. J.* **280**(Pt 1), 105–109 (1991).
32. König, J., Klatt, S., Dilger, K. & Fromm, M. F. Characterization of Ursodeoxycholic and Norursodeoxycholic Acid as Substrates of the Hepatic Uptake Transporters OATP1B1, OATP1B3, OATP2B1 and NTCP. *Basic Clin. Pharmacol. Toxicol.* **111**, 81–86, <https://doi.org/10.1111/j.1742-7843.2012.00865.x> (2012).
33. Ko, J. *et al.* Effects of side chain length on ionization behavior and transbilayer transport of unconjugated dihydroxy bile acids: a comparison of nor-chenodeoxycholic acid and chenodeoxycholic acid. *J. Lipid Res.* **35**, 883–892 (1994).
34. Jedlitschky, G. *et al.* ATP-dependent transport of bilirubin glucuronides by the multidrug resistance protein MRP1 and its hepatocyte canalicular isoform MRP2. *Biochem. J.* **327**(Pt 1), 305–310 (1997).
35. Kamisako, T. *et al.* Transport of monoglucuronosyl and bisglucuronosyl bilirubin by recombinant human and rat multidrug resistance protein 2. *Hepatology* **30**, 485–490, <https://doi.org/10.1002/hep.510300220> (1999).
36. Cui, Y. *et al.* Drug resistance and ATP-dependent conjugate transport mediated by the apical multidrug resistance protein, MRP2, permanently expressed in human and canine cells. *Mol. Pharmacol.* **55**, 929–937 (1999).
37. Denk, G. U. *et al.* Conjugation is essential for the anticholestatic effect of Norursodeoxycholic acid in tauroolithocholic acid-induced cholestasis in rat liver. *Hepatology* **52**, 1758–1768, <https://doi.org/10.1002/hep.23911> (2010).
38. Norris, R. P., Freudzon, M., Nikolaev, V. O. & Jaffe, L. A. Epidermal growth factor receptor kinase activity is required for gap junction closure and for part of the decrease in ovarian follicle cGMP in response to LH. *Reproduction* **140**, 655–662, <https://doi.org/10.1530/REP-10-0288> (2010).
39. Rubio, I. *et al.* Ras activation in response to phorbol ester proceeds independently of the EGFR via an unconventional nucleotide-exchange factor system in COS-7 cells. *Biochem. J.* **398**, 243–256, <https://doi.org/10.1042/BJ20060160> (2006).
40. Berger, C., Krengel, U., Stang, E., Moreno, E. & Madshus, I. H. Nimotuzumab and cetuximab block ligand-independent EGF receptor signaling efficiently at different concentrations. *J. Immunother.* **34**, 550–555, <https://doi.org/10.1097/CJI.0b013e31822a5ca6> (2011).
41. Hofmann, A. F. *et al.* Novel biotransformation and physiological properties of norursodeoxycholic acid in humans. *Hepatology* **42**, 1391–1398, <https://doi.org/10.1002/hep.20943> (2005).
42. Trauner, M. *et al.* Potential of nor-Ursodeoxycholic Acid in Cholestatic and Metabolic Disorders. *Dig. Dis.* **33**, 433–439, <https://doi.org/10.1159/000371904> (2015).
43. Kurz, A. K. *et al.* Phosphoinositide 3-kinase-dependent Ras activation by tauroursodesoxycholate in rat liver. *Biochem. J.* **350**(Pt 1), 207–213, <https://doi.org/10.1042/bj3500207> (2000).

44. Kim, N.-G. & Gumbiner, B. M. Adhesion to fibronectin regulates Hippo signaling via the FAK-Src-PI3K pathway. *J. Cell Biol.* **210**, 503–515, <https://doi.org/10.1083/jcb.201501025> (2015).
45. Sampaio, C. *et al.* Signal Strength Dictates Phosphoinositide 3-Kinase Contribution to Ras/Extracellular Signal-Regulated Kinase 1 and 2 Activation via Differential Gab1/Shp2 Recruitment: Consequences for Resistance to Epidermal Growth Factor Receptor Inhibition. *Mol. Cell. Biol.* **28**, 587–600, <https://doi.org/10.1128/MCB.01318-07> (2008).
46. Clark, E. A. & Brugge, J. S. Integrins and signal transduction pathways: the road taken. *Science* **268**, 233–239, <https://doi.org/10.1126/science.7716514> (1995).
47. Garcia, A. J. & Boettiger, D. Integrin-fibronectin interactions at the cell-material interface: initial integrin binding and signaling. *Biomaterials* **20**, 2427–2433, [https://doi.org/10.1016/S0142-9612\(99\)00170-2](https://doi.org/10.1016/S0142-9612(99)00170-2) (1999).
48. Mitra, S. K., Hanson, D. A. & Schlaepfer, D. D. Focal adhesion kinase: in command and control of cell motility. *Nature Reviews Molecular Cell Biology* **6**, 56–68, <https://doi.org/10.1038/nrm1549> (2005).
49. Huveneers, S. & Danen, E. H. J. Adhesion signaling - crosstalk between integrins, Src and Rho. *J. Cell Sci.* **122**, 1059–1069, <https://doi.org/10.1242/jcs.039446> (2009).
50. Shah, B. H., Neithardt, A., Chu, D. B., Shah, F. B. & Catt, K. J. Role of EGF receptor transactivation in phosphoinositide 3-kinase-dependent activation of MAP kinase by GPCRs. *J. Cell. Physiol.* **206**, 47–57, <https://doi.org/10.1002/jcp.20423> (2006).
51. Yoon, Y. B. *et al.* Effect of side-chain shortening on the physiologic properties of bile acids: hepatic transport and effect on biliary secretion of 23-nor-ursodeoxycholate in rodents. *Gastroenterology* **90**, 837–852 (1986).
52. Fickert, P. *et al.* 24-norUrsodeoxycholic Acid is Superior to Ursodeoxycholic Acid in the Treatment of Sclerosing Cholangitis in Mdr2 (Abcb4) Knockout Mice. *Gastroenterology* **130**, 465–481, <https://doi.org/10.1053/j.gastro.2005.10.018> (2006).
53. Noe, J., Stieger, B. & Meier, P. J. Functional expression of the canalicular bile salt export pump of human liver. *Gastroenterology* **123**, 1659–1666, <https://doi.org/10.1053/gast.2002.36587> (2002).
54. Stieger, B. *et al.* *In situ* localization of the hepatocytic Na⁺/Taurocholate cotransporting polypeptide in rat liver. *Gastroenterology* **107**, 1781–1787, [https://doi.org/10.1016/0016-5085\(94\)90821-4](https://doi.org/10.1016/0016-5085(94)90821-4) (1994).
55. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graphics* **14**(33–38), 27–38, [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5) (1996).
56. Xiao, T., Takagi, J., Collier, B. S., Wang, J.-H. & Springer, T. A. Structural basis for allostery in integrins and binding to fibrinogen-mimetic therapeutics. *Nature* **432**, 59–67, <https://doi.org/10.1038/nature02976> (2004).
57. Roe, D. R. & Cheatham, T. E. III. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **9**, 3084–3095, <https://doi.org/10.1021/ct400341p> (2013).
58. Sies, H. The use of perfusion of liver and other organs for the study of microsomal electron-transport and cytochrome P-450 systems. *Methods Enzymol.* **52**, 48–59, [https://doi.org/10.1016/S0076-6879\(78\)52005-3](https://doi.org/10.1016/S0076-6879(78)52005-3) (1978).
59. R: A language and environment for statistical computing (R Foundation for Statistical Computing, Vienna, Austria, 2010).

Acknowledgements

We are grateful for computational support by the “Zentrum für Informations und Medientechnologie” at the Heinrich Heine University and the computing time provided by the John von Neumann Institute for Computing (NIC) to H.G. on the supercomputer JURECA at Jülich Supercomputing Centre (JSC) (user ID: HKF7). The authors thank Nicole Eichhorst and Janina Thies for their expert technical assistance. This study was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) through the Collaborative Research Center SFB 974 (‘Communication and Systems Relevance during Liver Damage and Regeneration’, Düsseldorf) and the Clinical Research Group KFO 217 (‘Hepatobiliary Transport in Health and Disease’, Düsseldorf). Funding by Deutsche Forschungsgemeinschaft (DFG) (INST 208/704–1 FUGG) to purchase the hybrid computer cluster used in this study is gratefully acknowledged.

Author contributions

M.B. performed MD simulations, analyzed data, wrote manuscript; A.S. performed experiments, analyzed data, wrote manuscript; N.Q. performed experiments, analyzed data; B.G. performed experiments, analyzed data; B.S.L. performed affinity determination, analyzed data; H.K. analyzed affinity data; H.G. designed study, analyzed data, wrote manuscript; D.H. designed study, analyzed data, wrote manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-62326-y>.

Correspondence and requests for materials should be addressed to H.G. or D.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020

Supplementary Information

Evidence for functional selectivity in TUDC- and *nor*UDCA-induced signal transduction via $\alpha_5\beta_1$ integrin towards choleresis

*Michele Bonus*¹, *Annika Sommerfeld*², *Natalia Qvartskhava*², *Boris Görg*², *Beatrice Stefanie Ludwig*³, *Horst Kessler*³, *Holger Gohlke*^{1,4*}, *Dieter Häussinger*^{2*}

¹Institute for Pharmaceutical and Medicinal Chemistry, Department of Mathematics and Natural Sciences, Heinrich Heine University Düsseldorf, Universitätsstr. 1, 40225 Düsseldorf, Germany

²Clinic for Gastroenterology, Hepatology and Infectious Diseases, Heinrich Heine University Düsseldorf, Moorenstr. 5, 40225 Düsseldorf, Germany

³Institute for Advanced Study and Center for Integrated Protein Science, Department of Chemistry, Technische Universität München, Lichtenbergstr. 4, 85747 Garching, Germany

⁴John von Neumann Institute for Computing (NIC), Jülich Supercomputing Centre (JSC), and Institute for Complex Systems - Structural Biochemistry (ICS-6), Forschungszentrum Jülich GmbH, Wilhelm-Johnen-Straße, 52425 Jülich, Germany

Table of Contents

Suppl. Text	3
Suppl. Figure 1: Primary and conjugated bile acids	7
Suppl. Figure 2: Structural stability during the MD simulations	8
Suppl. Figure 3: Effects of <i>nor</i> UDCA, <i>Tnor</i> UDCA, GUDC and UDCA on β_1 integrin activation	9
Suppl. Figure 4: Effects of OA on <i>nor</i> UDCA- and TUDC-induced Erk-1/-2 activation	10
Suppl. Figure 5: <i>nor</i> UDCA-induced activation of Erk-1/-2, p38 ^{MAPK} and Src	11
Suppl. Figure 6: <i>nor</i> UDCA-induced activation of Erk-1/-2, p38 ^{MAPK} and Src	12
Suppl. Figure 7: Control perfusion experiments	13
Suppl. Figure 8: Phosphorylation of FAK ^{Y397} in <i>nor</i> UDCA- and TUDC-perfused rat livers	14
Suppl. Figure 9: Phosphorylation of FAK ^{Y925,861,576/577,407} in perfused rat livers	15
Suppl. Figure 10: <i>nor</i> UDCA- and TUDC-induced Erk-1/-2, p38 ^{MAPK} and EGFR activation	17
Suppl. Figure 11: <i>nor</i> UDCA- and TUDC-induced c-Src activ. and EGFR/c-Src association	18
Suppl. Figure 12: TUDC-induced dual activation of Erk-1/-2 and p38 ^{MAPK}	19
Suppl. Figure 13: Immunohistochemical determination of Bsep and ZO-1 distribution	20
Suppl. Figure 14: Regulation of Ntcp in <i>nor</i> UDCA-perfused rat liver	21
Suppl. Figure 15: Inhibition of <i>nor</i> UDCA-induced β_1 integrin activation by TC	22
Suppl. Figure 16: Binding of TUDC and <i>nor</i> UDCA to $\alpha_5\beta_1$ integrin	23
Suppl. Figure 17: Flexibility changes in the βA domain upon activation	24
Suppl. Figure 18: Comparison between <i>nor</i> UDCA- and TUDC-induced Erk-1/-2 activation	25
Suppl. Figure 19: Subunit and domain organization of the $\alpha_5\beta_1$ integrin ectodomain	26
Suppl. Table 1: Domain-wise minimum, maximum and average RMSD values	27
Suppl. Table 2: Angles and distances computed from MD trajectories	28
Suppl. Table 3: Statistical testing for differences in mean values	29
Suppl. References	30

Note: some figure titles in the table of contents were shortened for reasons of clarity.

Supplementary Text

Materials and Methods

Generation of $\alpha_5\beta_1$ integrin-bile acid complex structures

To generate starting structures for the MD simulations, a 3D model of the $\alpha_5\beta_1$ integrin ectodomain was used as the receptor. This model had been generated by us by homology modeling in a previous study¹. The same model had been used in the study identifying TUDC as a potent activator of $\alpha_5\beta_1$ integrin-mediated signaling².

3D structures of the bile acids *nor*UDCA, *Tnor*UDCA, GUDC, and UDCA were generated with Maestro³ and prepared with the LigPrep suite of programs⁴: I) Protonation states for each bile acid were calculated at pH 7.4 ± 0.2 using Epik⁵. II) Energy-minimized ring conformations were obtained via ring sampling from the template collection embedded in the ring_conf utility. III) Each bile acid was energy-minimized using the premin tool in LigPrep and bmin in MacroModel⁶.

To generate bound conformations of each bile acid in the binding site of $\alpha_5\beta_1$ integrin, the conformation of TUDC in the ligand-bound homology model was used as a template for structural alignment of the other bile acids with the program Moloc. To maintain coordination to the Mg^{2+} ion in the binding site, the acidic moieties and the side chains of each bile acid were first superimposed on the sulfonate group and the side chain of TUDC. Second, the 5 β -androstane core of *Tnor*UDCA and GUDC was superimposed onto the 5 β -androstane core of TUDC. This procedure was omitted for *nor*UDCA and UDCA, because the side chains in these bile acids are not sufficiently long to allow superposition of the core. Instead, the orientation of the 5 β -androstane core of these bile acid was manually adjusted to resemble its orientation in TUDC. Finally, the ligand poses were energy-minimized in the receptor with the MAB force field⁷.

MD simulations

All MD simulations were set up and performed with the AMBER 11 suite of programs⁸ using the ff99SB force field⁹ for the protein and the general AMBER force field (GAFF)¹⁰ for the bile acids. Parameters from Åquist *et al.*¹¹ were used to treat the Mg^{2+} ions. Partial charges for the ligand atoms were derived according to the RESP procedure¹². Each integrin-bile acid complex

was neutralized by addition of Na^+ ions and solvated in an octahedral periodic box of TIP3P water molecules¹³. The closest allowed distance between the outermost atom of the protein and the border of the periodic box was 11 Å. The Particle Mesh Ewald method¹⁴ was used to calculate long-range electrostatic interactions. For each simulation, a 2 fs integration time step was used, and the cutoff distance for direct space non-bonded interactions was set to 8 Å.

For thermalizing the systems, solute atoms were restrained by harmonic positional restraints with a force constant of $25 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{Å}^{-2}$, and the systems were energy-minimized for 50 steps with the conjugate gradient method, followed by 200 steps with the steepest descent method. This procedure was repeated with a lower force constant of $5 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{Å}^{-2}$. While keeping the restraints on the solute atoms, gradual heating from 100 K to 300 K was performed for 40 ps in the canonical (NVT) ensemble. Target temperature and restraints were kept for another 10 ps. To adjust the solvent density, the systems were subjected to three consecutive restrained simulations of 50 ps length in the isothermal-isobaric (NPT) ensemble. Gradual removal of the restraints was performed during 100 ps of NVT-MD. The systems were then subjected to 200 ns of NVT-MD, with coordinates extracted every 20 ps. We performed three independent simulations of 200 ns length of the $\alpha_5\beta_1$ -bile acid complexes containing *norUDCA*, *TnorUDCA*, *GUDC*, and *UDCA*, and two additional independent simulations of $\alpha_5\beta_1$ integrin bound to *TUDC* and *TC*. In addition, for *TUDC* and *TC*, data of MD simulations of 200 ns length was reused from ref.².

Immunofluorescence staining

Cryosections of liver samples (7 μm) were either fixed in methanol (2 min, -20°C) or fixed and permeabilized with Zamboni (20 min, room temperature) and Triton X-100 (2 min, 4°C). Then, liver sections were blocked with FBS (5% (w/v), 30 min, room temperature) and incubated with the primary antibodies against Bsep (1:200) and ZO-1 (1:500) (overnight, 4°C), Ntcp (1:200) and Na^+/K^+ -ATPase (1:100), or β_1 integrin subunit active conformation (1:100) and $\alpha_5\beta_1$ integrin dimer (1:100), washed thrice, and stained with an anti-mouse-FITC and an anti-rabbit Cy3-conjugated antibody (1:500 in PBS, 2 h, room temperature, respectively). To visualize filamentous actin in order to document cell shape, FITC-coupled phalloidin (1 $\mu\text{g}/\text{ml}$) was applied. Following immunofluorescence staining, samples were covered with Fluoromount-G reagent and visualized by confocal laser scanning microscopy (CLSM) using LSM510 META (Zeiss, Oberkochen, Germany). Average pixel intensities for the red channel were calculated in

ImageJ by separating the red channel from each image and extracting the data with the Histogram function.

Densitometric fluorescence intensity analysis

Cryosections of perfused rat liver for the analysis at the canalicular membrane were stained for Bsep and for the tight junction protein ZO-1, which forms the sealing border between canalicular and sinusoidal membrane. Apparent integrity and comparability of the canaliculi was assumed when the bordering tight junction lines (detected by the immunostained ZO-1) were intact and run in parallel. Cryosections of perfused rat liver for the analysis at the basolateral membrane were stained for Ntcp and for Na^+/K^+ -ATPase (plasma membrane marker protein). Na^+/K^+ -ATPase profiles were selected according apparent integrity and comparability. Acceptable Na^+/K^+ -ATPase intensity profiles have a sufficiently high peak fluorescence in the central part (corresponding to the basolateral membrane) and low intracellular fluorescence. For negative controls, primary antibodies were omitted in each experiment. Densitometric analysis was performed as described previously^{15,16}. For densitometric analysis using digitalized microscopic pictures of the membranes, the software Image-Pro Plus (Media Cybernetics, Rockville, USA) was used. The profile of the fluorescence intensity was measured over a thick line at a right angle to the membrane. The length of the line was always 8 μm . The mean fluorescence intensity to each pixel over the line perpendicular to the length was calculated by Image-Pro Plus. Each measurement was normalized to the sum of all intensities of the respective measurement. Values are given as means \pm SEM. Densitometric analysis of protein distribution in immunofluorescence images were performed using Wilcoxon rank sum test. $p < 0.05$ was considered statistically significant. Data from at least 10 different areas per tissue sample and from at least three independent liver preparations were processed.

Immunoblot analysis

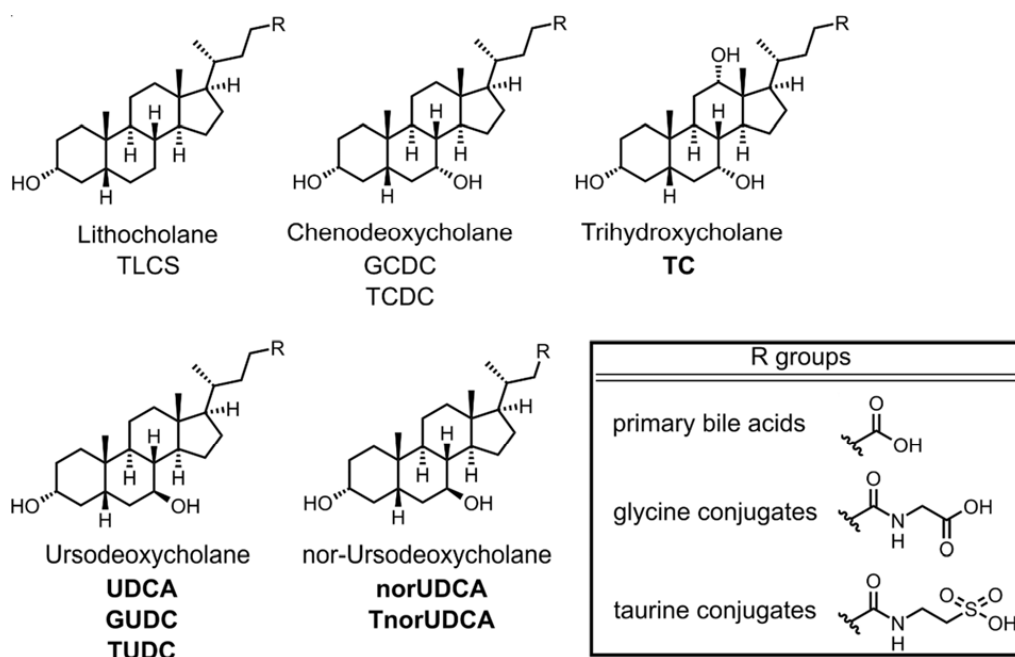
Liver samples were immediately lysed at 4°C by using a lysis buffer containing 20 mmol/l Tris-HCl (pH 7.4), 140 mmol/l NaCl, 10 mmol/l NaF, 10 mmol/l sodium pyrophosphate, 1% (v/v) Triton X-100, 1 mmol/l EDTA, 1 mmol/l EGTA, 1 mmol/l sodium vanadate, 20 mmol/l β -glycerophosphate, protease inhibitor (cOmplete™, Roche Diagnostics, Mannheim, Germany) and phosphatase inhibitor cocktail (PhosSTOP™, Roche Diagnostics, Mannheim, Germany). The

lysates were kept on ice for 10 min and then centrifuged at 8000 rpm for 8 min at 4°C, and aliquots of the supernatant were taken for protein determination using the Bio-Rad protein assay (Bio-Rad Laboratories, Munich, Germany). Equal amounts of protein were subjected to sodium dodecyl sulfate / polyacrylamide gel electrophoresis and transferred onto nitrocellulose membranes using a semidry transfer apparatus (GE Healthcare, Freiburg, Germany). Membranes were blocked for 60 min in 5% (w/v) bovine serum albumin or 5% (w/v) milk powder containing 20 mmol/l Tris (pH 7.5), 150 mmol/l NaCl, and 0.1% Tween 20 (TBS-T) and exposed to primary antibodies overnight at 4°C. After washing with TBS-T and incubation at room temperature for 2 h with horseradish peroxidase-coupled anti-mouse or anti-rabbit IgG antibody, respectively (all diluted 1:10,000), the immunoblots were washed extensively, and bands were visualized using the ChemiDoc™ Touch Imaging System from Bio-Rad (Munich, Germany). Semi-quantitative evaluation was carried out by densitometry using the Image Lab Touch Software from Bio-Rad. Protein phosphorylation is given as the ratio of detected phospho-protein/total protein.

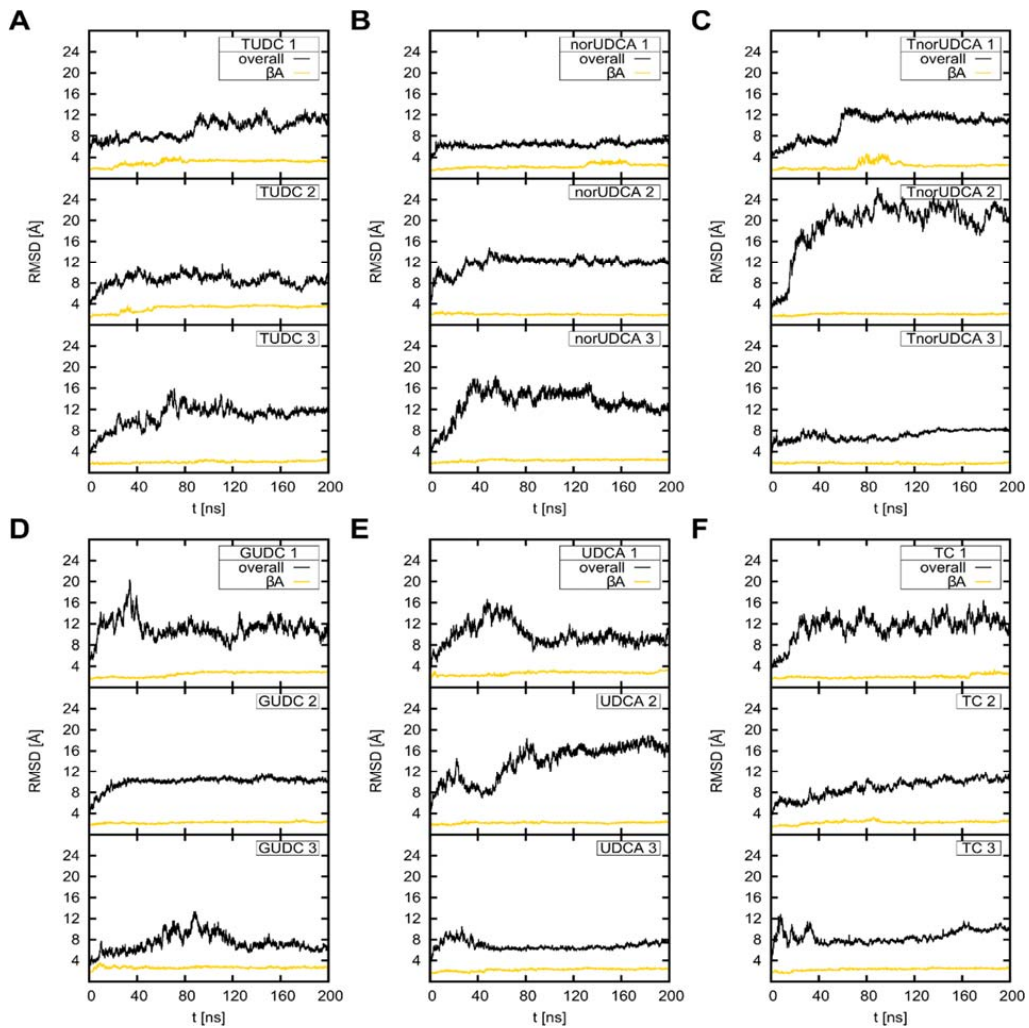
Immunoprecipitation

Liver samples were harvested in lysis buffer containing 136 mmol/l NaCl, 20 mmol/l Tris-HCl, 10% (v/v) glycerol, 2 mmol/l EDTA, 50 mmol/l β -glycerophosphate, 20 mmol/l sodium pyrophosphate, 0.2 mmol/l Pefablock, 5 mg/l aprotinin, 5 mg/l leupeptin, 4 mmol/l benzamidine, 1 mmol/l sodium vanadate, supplemented with 1% (v/v) Triton X-100. The protein amount was determined as described above. Samples containing equal protein amounts were incubated for 2 h at 4°C with the respective antibody in order to immunoprecipitate the required protein. Then protein A-/G-agarose (Santa Cruz Biotechnology, Heidelberg, Germany) was added and incubated at 4°C overnight. Immunoprecipitates were washed thrice with lysis buffer supplemented with 0.1% (v/v) Triton X-100 and then transferred to Western blot analysis as described above.

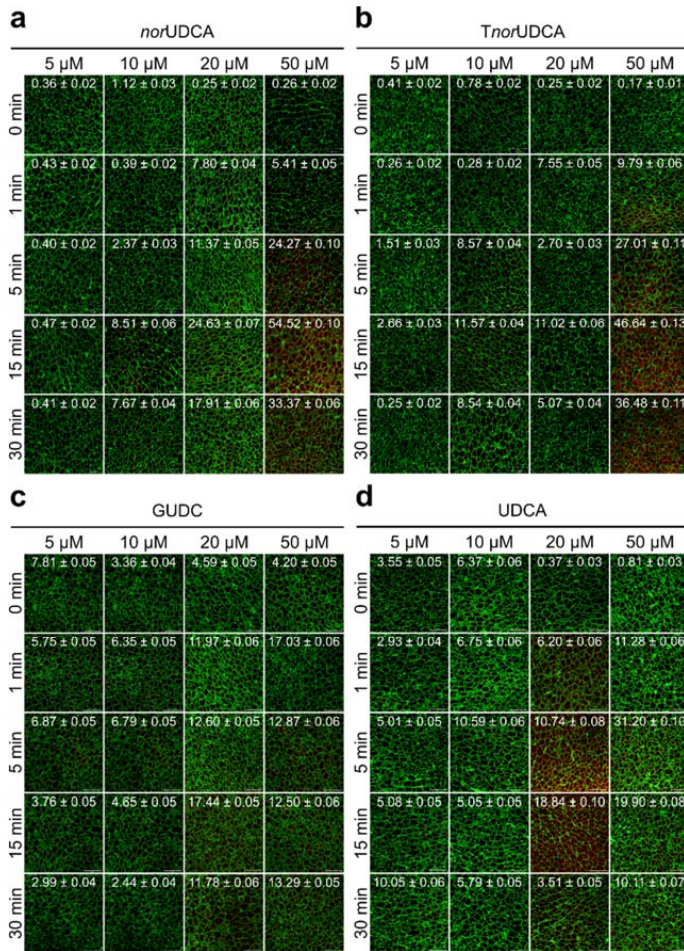
Supplementary Figures



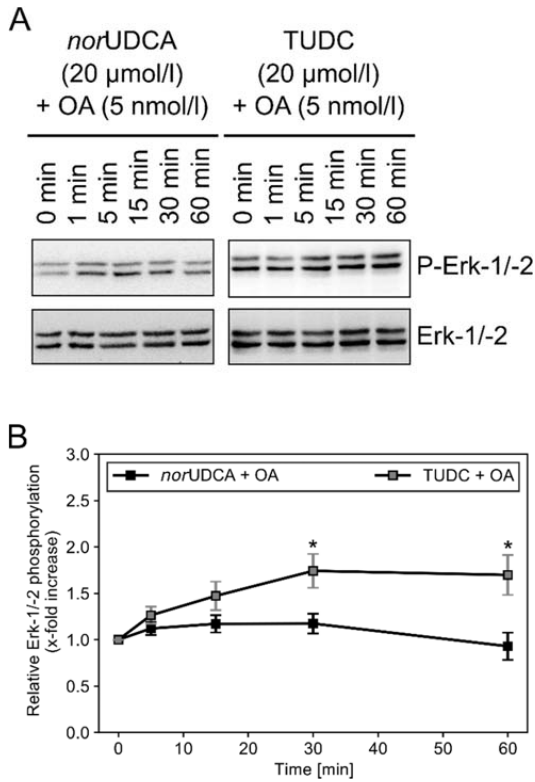
Supplementary Figure 1. Primary and conjugated bile acids. Chemical structures of primary and conjugated bile acids. Bile acids for which agonistic activity towards $\alpha_5\beta_1$ integrins is known or has been determined in this study and bile acids for which antagonistic activity is known are marked in bold.



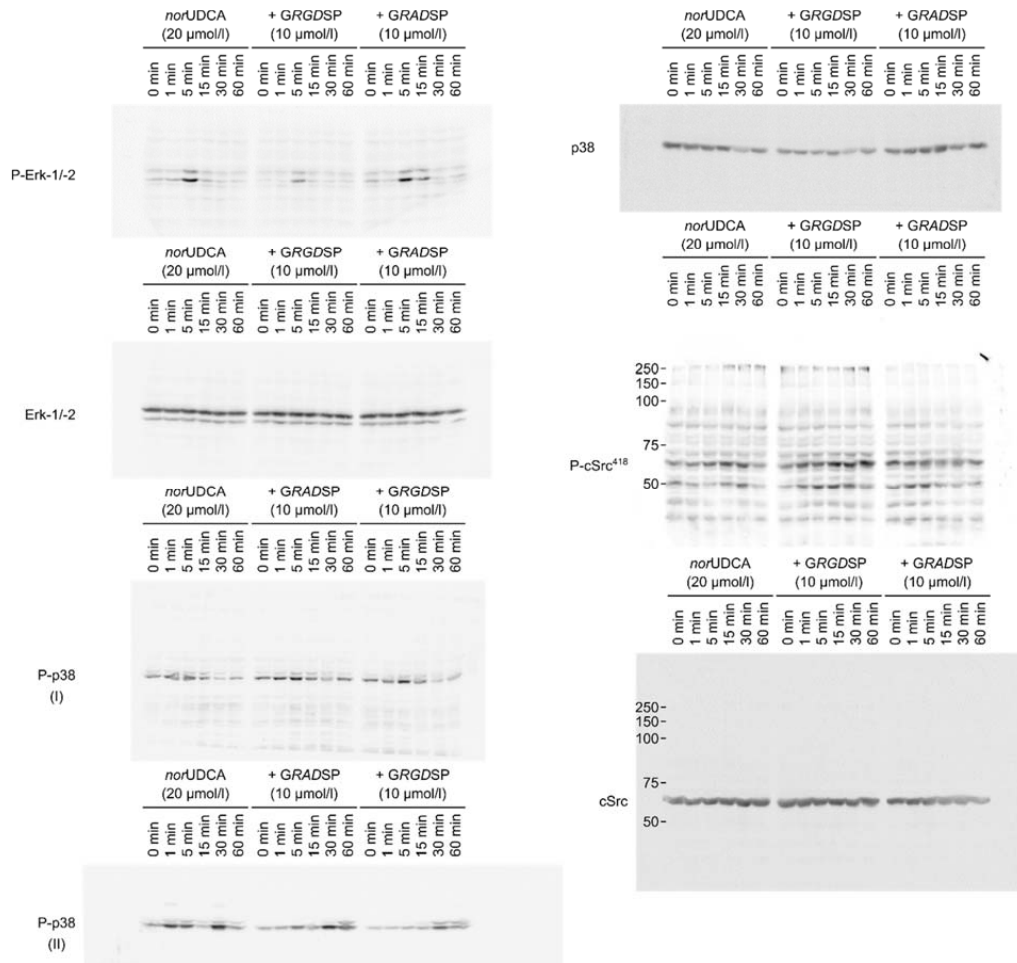
Supplementary Figure 2. Structural stability during the MD simulations. Root mean square deviation (RMSD) of the overall protein structure (black lines) and the βA domain (yellow lines) during three (rows) MD simulations of the complexes between $\alpha_5\beta_1$ integrin and (A) TUDC, (B) *nor*UDCA, (C) *Tnor*UDCA, (D) GUDC, (E) UDCA, and (F) TC.



Supplementary Figure 3. Effects of norUDCA, TnorUDCA, GUDC and UDCA on β_1 integrin activation. Rat livers were perfused with (a) norUDCA, (b) TnorUDCA, (c) GUDC, and (d) UDCA for up to 60 min with the concentrations indicated. Liver samples were immunostained for the active conformation of β_1 integrin (red) and $\text{Na}^+\text{-K}^+\text{-ATPase}$ (green).



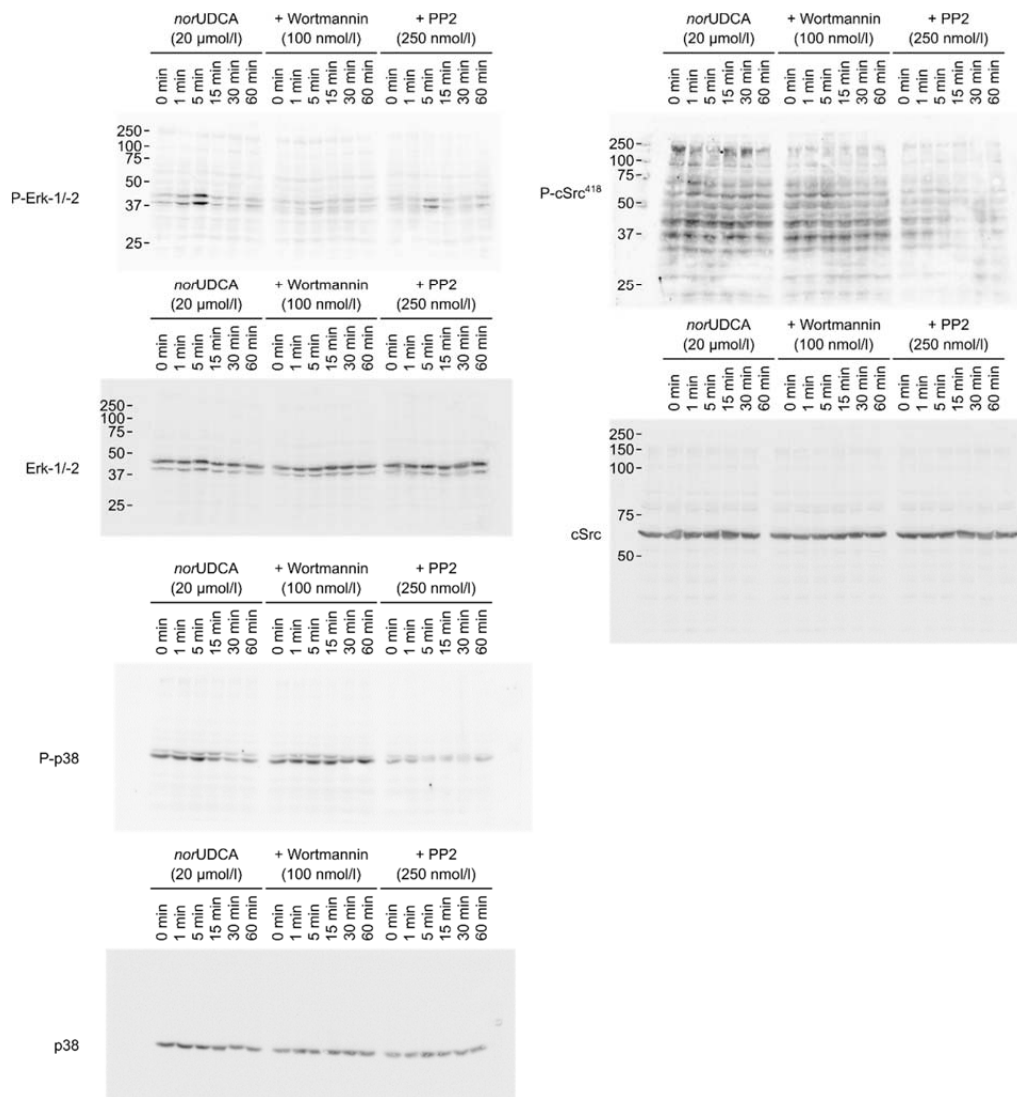
Supplementary Figure 4. Effects of okadaic acid on *nor*UDCA- and TUDC-induced Erk-1/-2 activation. Rat livers were perfused with *nor*UDCA or TUDC (20 μ mol/l each) in the presence of okadaic acid (OA) for up to 60 min. Liver samples were taken at the time points indicated. Phosphorylation of Erk-1/-2 was analyzed by (A) Western blot using specific antibodies and (B) subsequent densitometric analysis. Levels of phosphorylated Erk-1/-2 were normalized to total Erk-1/-2. Phosphorylation at $t = 0$ was arbitrarily set to 1. Data represent the mean (mean \pm SEM) of at least three independent experiments; *, Statistically significant compared to the unstimulated control ($t = 0$) (one-way ANOVA, Dunnett's *post hoc* test).



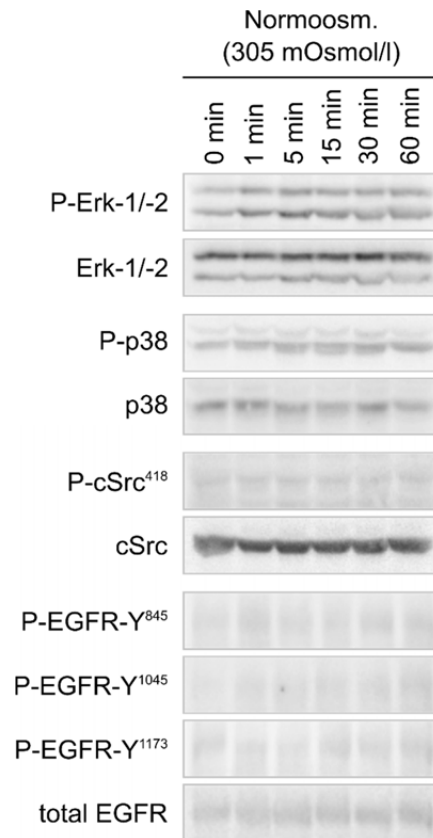
Supplementary Figure 5. *norUDCA*-induced activation of Erk-1/-2, p38^{MAPK} and Src. Full-length blots of cropped blots from Figure 5a in the main text.

SI for "Evidence for functional selectivity in signal transduction via $\alpha_5\beta_1$ integrin"

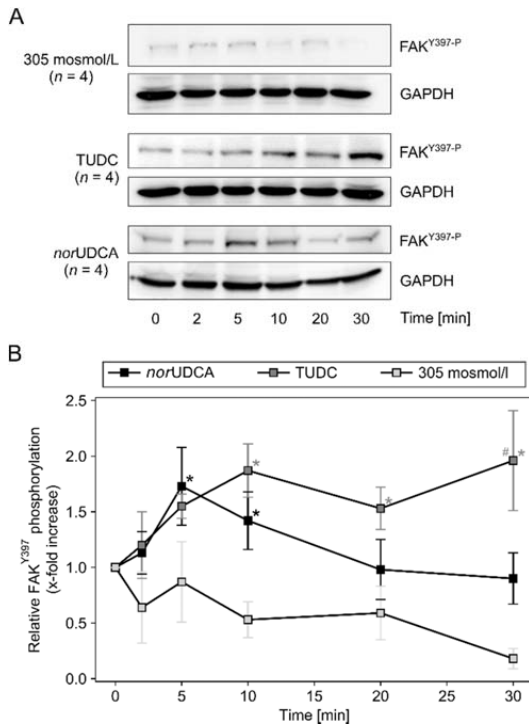
12



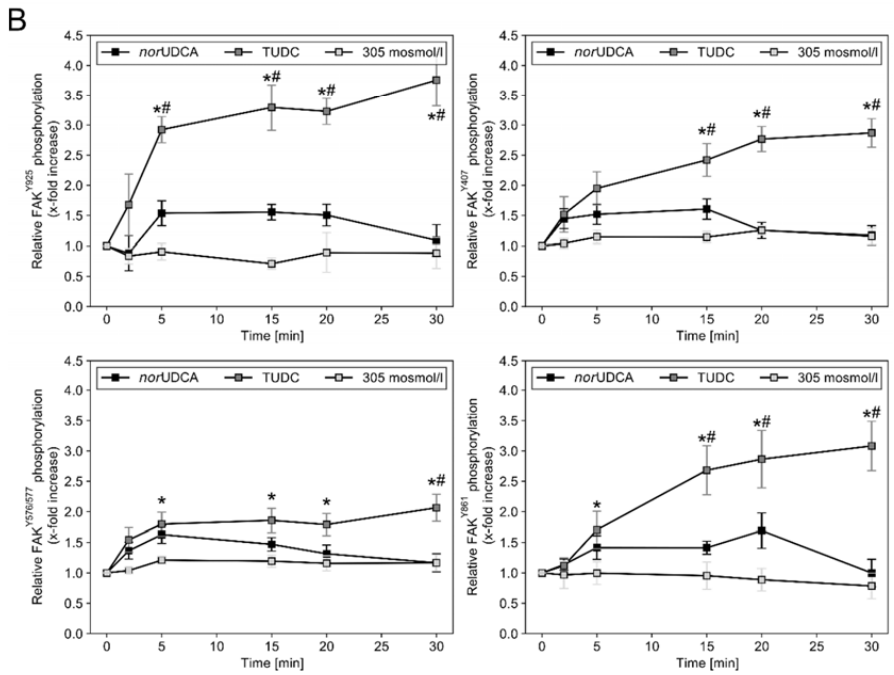
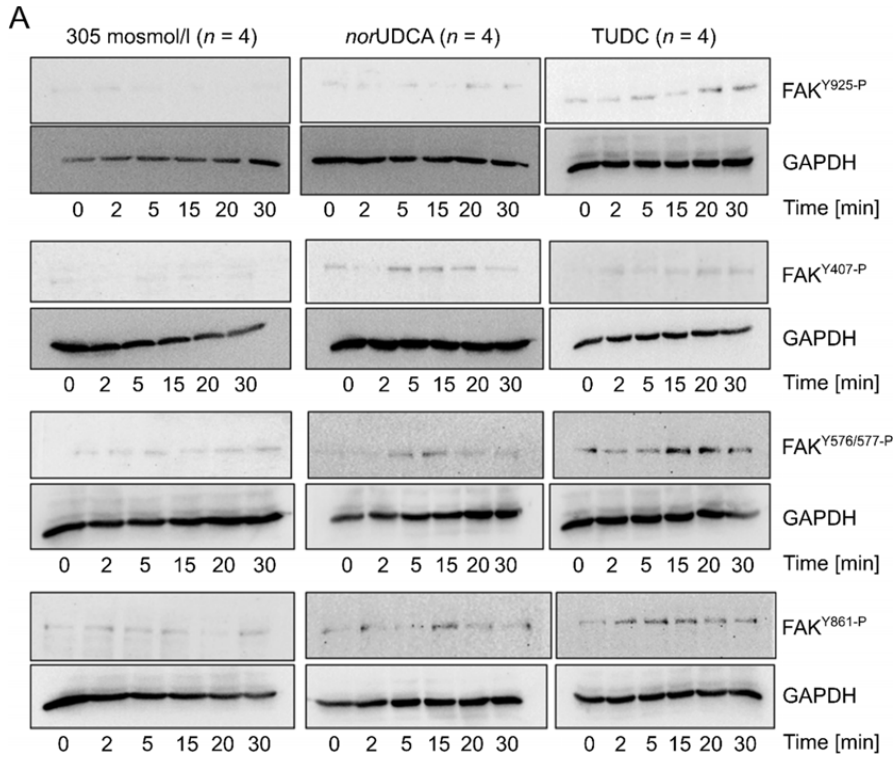
Supplementary Figure 6. *norUDCA*-induced activation of Erk-1/-2, p38^{MAPK} and Src. Full-length blots of cropped blots from Figure 5b in the main text.



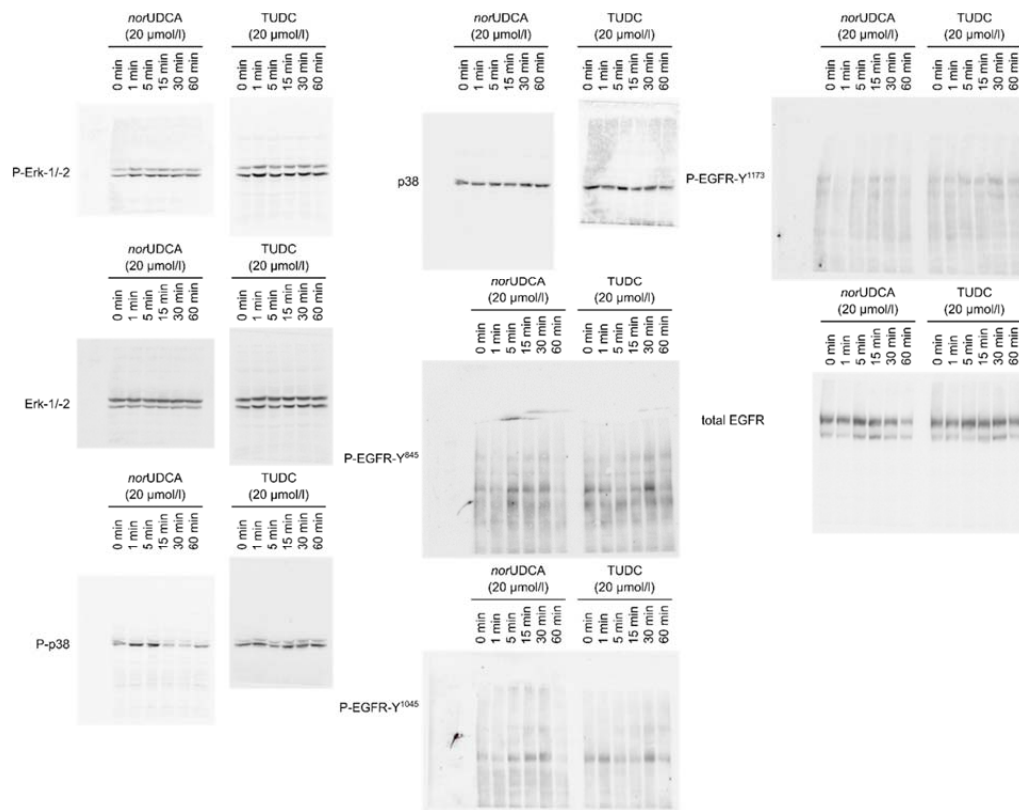
Supplementary Figure 7. Control perfusion experiments. Rat livers were perfused with control buffer for up to 60 min. Liver samples were taken at the time points indicated. Activation of Erk-1/-2, p38^{MAPK}, cSrc, EGFR Tyr⁸⁴⁵, Tyr¹⁰⁴⁵, and Tyr¹¹⁷³ were analyzed by Western blot using specific antibodies. Total Erk-1/-2, p38^{MAPK}, cSrc, and EGFR served as respective loading controls. Representative pictures of at least three independent experiments are depicted. Blots were cropped to focus on the area of interest. No effect on the phosphorylation of kinases and the EGFR were observed under control conditions.



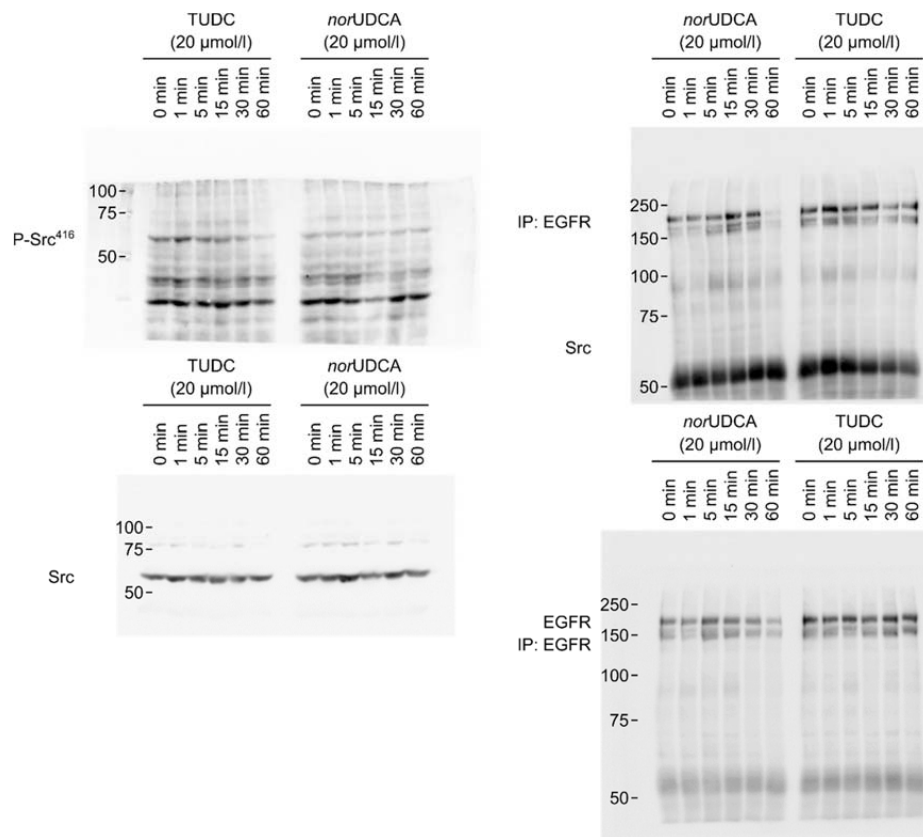
Supplementary Figure 8. Phosphorylation of FAK^{Y397} in *norUDCA*- and TUDC-perfused rat livers. The livers were perfused either with normoosmotic medium (305 mosmol/l) or with TUDC (20 μ mol/l) and *norUDCA* (20 μ mol/l), and the protein samples were gathered at the indicated time points (0-30 min). Thereafter, phosphorylation of FAK^{Y397} was analyzed by Western blotting using specific antibodies. Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) served as a loading control. Blots were cropped to focus on the area of interest. FAK^{Y397} phosphorylation was significantly increased in the TUDC-perfused livers within 10 min and lasted for up to 30 min compared to livers perfused with normoosmotic medium. However, FAK^{Y397} phosphorylation was not significantly increased in *norUDCA* perfused livers compared to controls. (A) Representative blots are shown from four independent experiments. (B) Phosphorylation of focal adhesion kinase (FAK^{Y397}) was analyzed densitometrically, and the means of measurements in each of the four individual experiments for each condition are shown (means \pm SEM). *: Statistically significant against “305 mosmol/l” ($p < 0.05$, $n = 4$, two-way ANOVA, Bonferroni *post hoc* test).



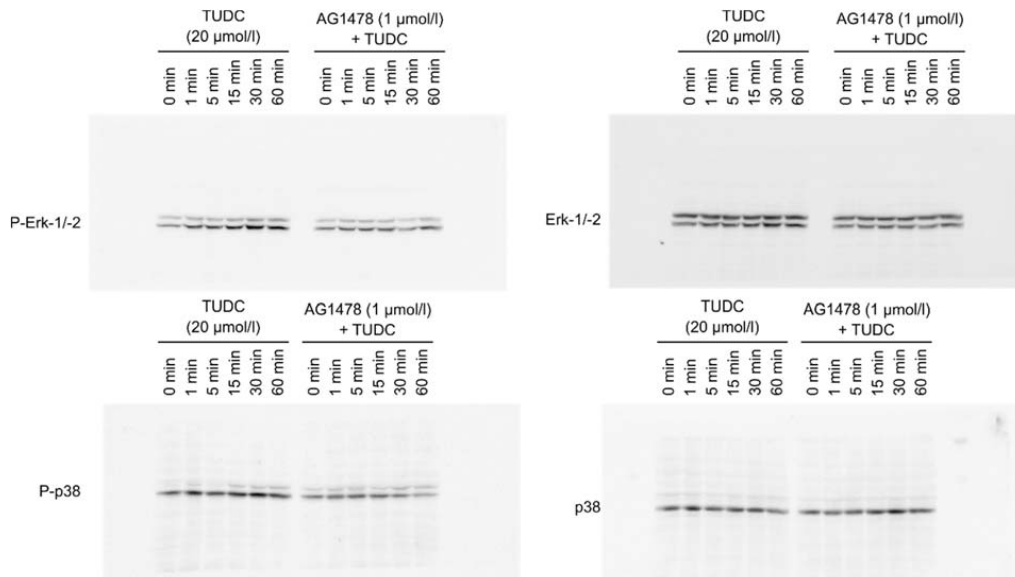
Supplementary Figure 9. Phosphorylation of FAK^{Y925,861,576/577,407} in *nor*UDCA- and TUDC-perfused rat livers. The livers were perfused either with normoosmotic medium (305 mosmol/l) or with TUDC (20 μ mol/l) and *nor*UDCA (20 μ mol/l), and the protein samples were gathered at the indicated time points (0-30 min). Thereafter, phosphorylation of FAK^{Y925}, FAK^{Y407}, FAK^{Y576/577}, and FAK^{Y861} was analyzed by Western blotting using specific antibodies. Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) served as a loading control. Blots were cropped to focus on the area of interest. FAK^{Y925,861,576/577,407} phosphorylation was significantly increased in the TUDC-perfused livers within 5-15 min and lasted for up to 30 min compared to livers perfused with normoosmotic medium. However, phosphorylation was not significantly increased in *nor*UDCA perfused livers compared to controls. (A) Representative blots are shown from four independent experiments. (B) Phosphorylation of focal adhesion kinase (FAK^{Y925,861,576/577,407}) was analyzed densitometrically, and the means of measurements in each of the four individual experiments for each different condition are shown (means \pm SEM). *: Statistically significant against "305 mosmol/l"; #: Statistically significant against "*nor*UDCA" ($p < 0.05$, $n = 4$, two-way ANOVA, Bonferroni *post hoc* test).



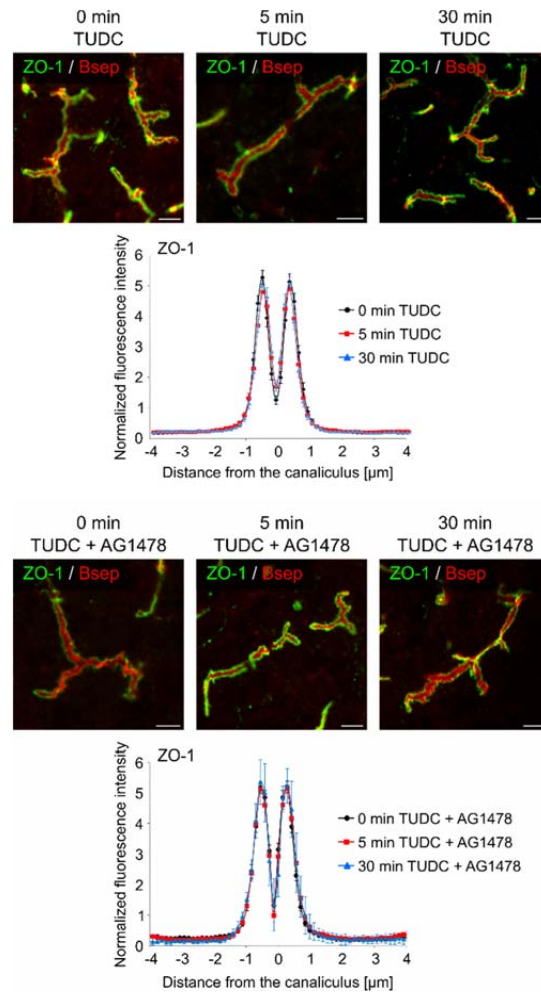
Supplementary Figure 10. Comparison between *norUDCA*- and TUDC-induced Erk-1/-2, p38^{MAPK} and EGFR activation. Full-length blots of cropped blots from Figure 6a in the main text.



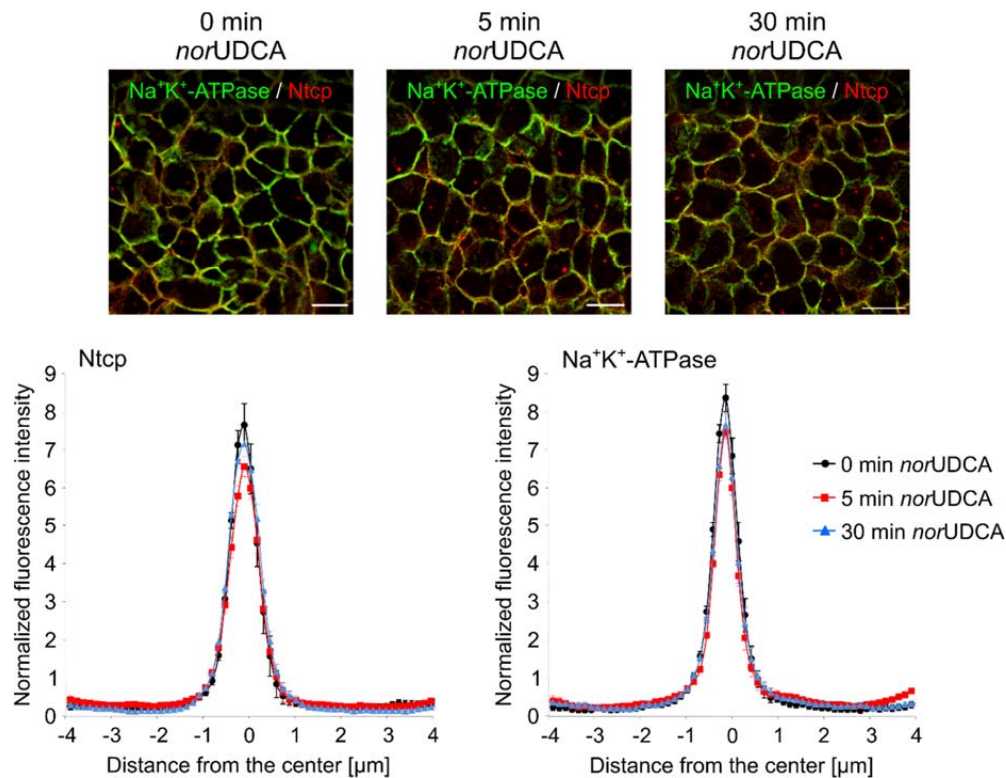
Supplementary Figure 11. Comparison between *norUDCA*- and TUDC-induced c-Src activation and EGFR/c-Src association. Full-length blots of cropped blots from Figure 7a in the main text.



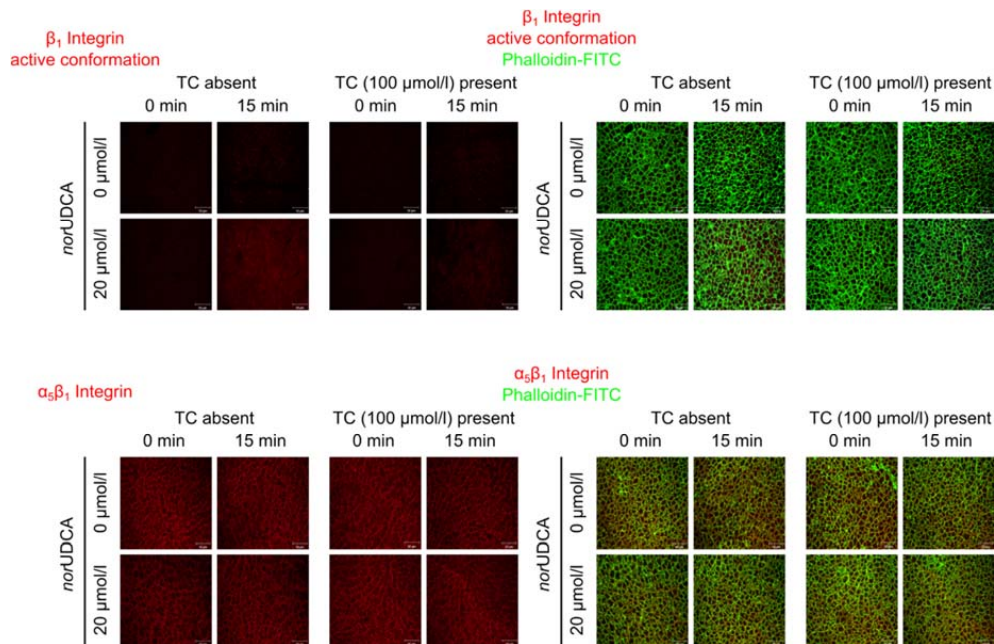
Supplementary Figure 12. TUDC-induced dual activation of Erk-1/2 and p38^{MAPK} and Bsep insertion into the canalicular membrane are dependent on EGFR phosphorylation. Full-length blots of cropped blots from Figure 8a in the main text.



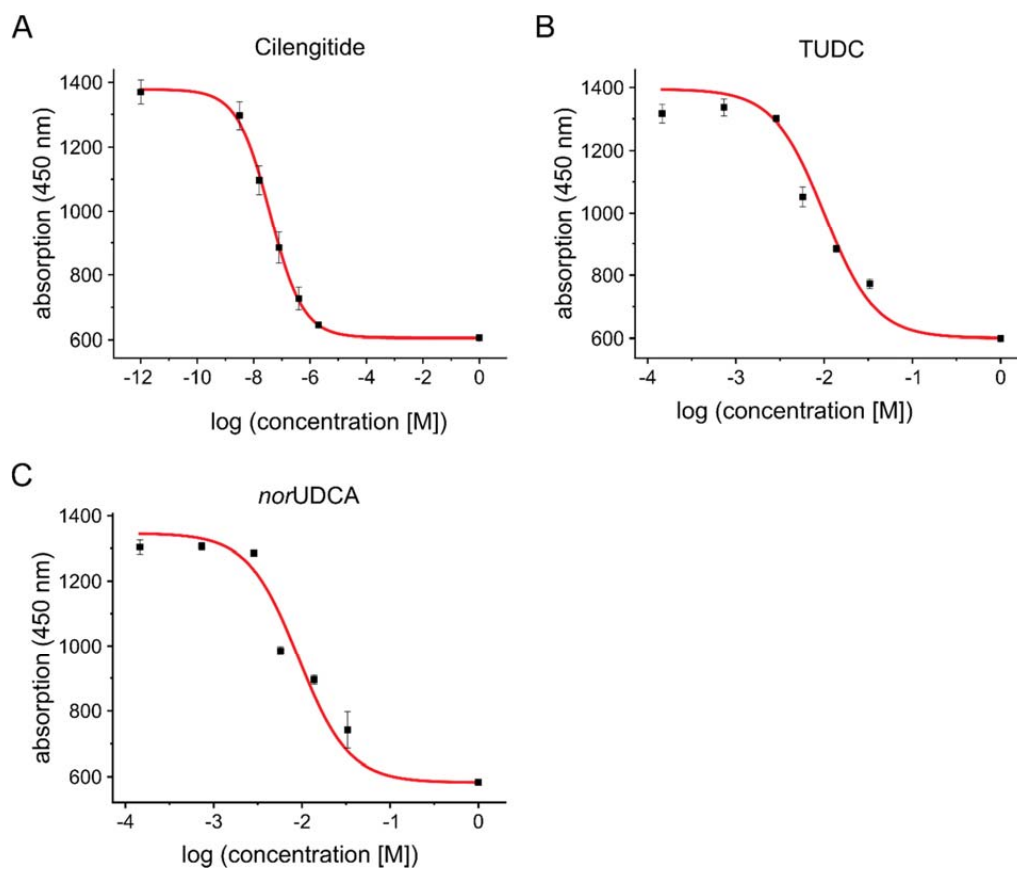
Supplementary Figure 13. Immunohistochemical determination of Bsep and ZO-1 distribution in TUDC perfused rat livers. For experimental details see caption of Fig. 8. Cryosections from perfused rat liver were immunostained for Bsep and ZO-1, fluorescence images were recorded by confocal LSM, and analyzed densitometrically. Representative pictures of at least three independent experiments are depicted. The scale bar corresponds to 5 μm . TUDC led to insertion of intracellular Bsep into the canalicular membrane, which was inhibited by AG1478 (see also Fig. 8). Liver perfusion experiments with *nor*UDCA resulted in no significant changes of ZO-1 fluorescence profiles. Means \pm SEM of 30 measurements in each of at least three individual experiments for each condition are shown.



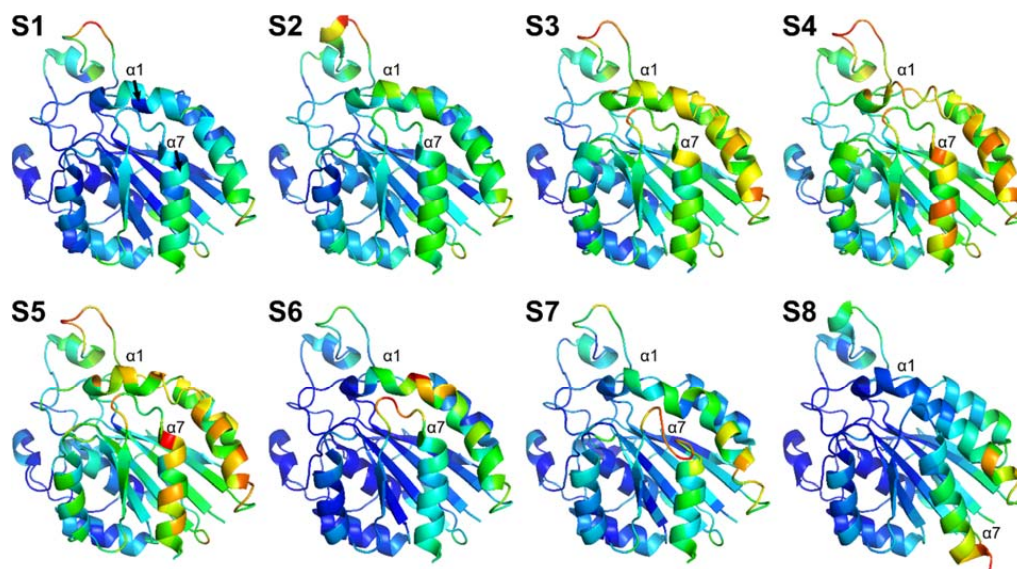
Supplementary Figure 14. Regulation of Ntcp in *norUDCA*-perfused rat liver. Rat livers were perfused with *norUDCA* (20 μmol/l) for up to 60 min, immunostained for Ntcp and Na⁺/K⁺-ATPase. The densitometric analysis of fluorescence profiles and intensity of Ntcp and Na⁺/K⁺-ATPase distribution are shown. Representative pictures of at least three independent experiments are depicted. The scale bars correspond to 10 μm. Under control conditions (black, $t=0$ min), Ntcp-bound fluorescence was largely localized in the center of the basolateral membrane. No significant changes of Ntcp and Na⁺/K⁺-ATPase fluorescence profiles were observed after perfusion with *norUDCA* (blue, $t=5$ min; red, $t=30$ min). Means \pm SEM of 30 measurements in each of at least three individual experiments for each condition are shown.



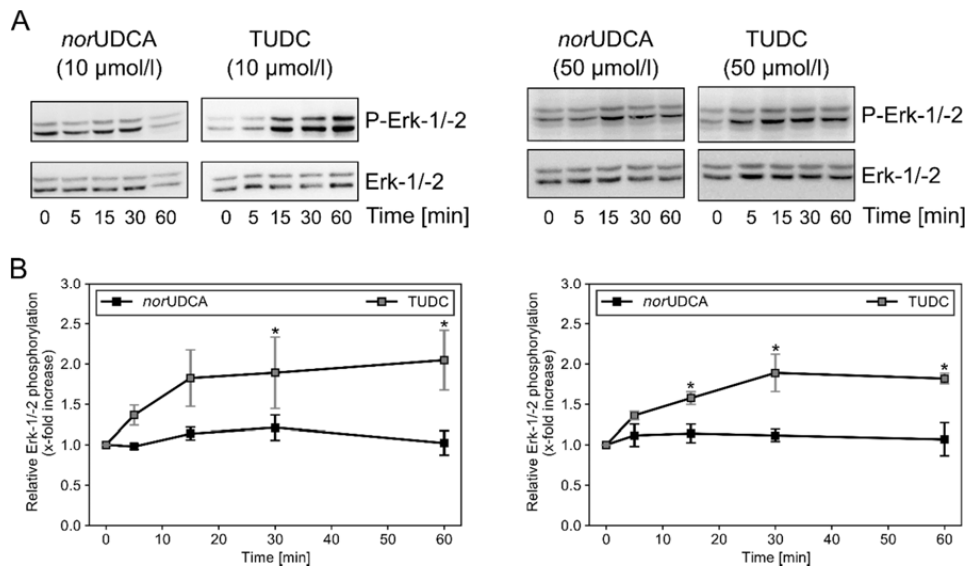
Supplementary Figure 15. Inhibition of *norUDCA*-induced β_1 integrin activation by TC. Rat livers were perfused with *norUDCA* (20 $\mu\text{mol/l}$) for 15 min. Experiments were carried out in the absence or presence of TC (100 $\mu\text{mol/l}$), which was added 30 min prior to *norUDCA* addition. Liver samples were taken at the time points indicated. Cryosections were stained for the β_1 integrin subunit or for total $\alpha_5\beta_1$ integrin, and fluorescence images were recorded by LSM. In order to visualize cell shapes and integrity of the cytoskeleton, filamentous actin was labeled by use of FITC-coupled phalloidin as a control. The scale bar corresponds to 50 μm . In the absence of TC, *norUDCA* induced within 15 min an activation of the β_1 integrin subunit, whereas in the presence of TC the effect was abolished. Immunoreactivity of the total $\alpha_5\beta_1$ integrin heterodimer as well as the integrity of the cytoskeleton remained unchanged under these conditions.



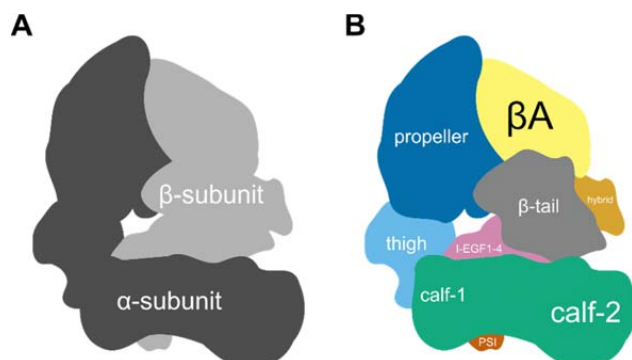
Supplementary Figure 16. Binding of TUDC and *nor*UDCA towards $\alpha_5\beta_1$. Absorption-concentration curves obtained from the competitive ELISA-based assay for (A) Cilengitide (reference compound), (B) TUDC, and (C) *nor*UDCA binding to $\alpha_5\beta_1$ integrin. The red curves represent a sigmoidal fit to 32 data points, obtained from two serial dilution rows.



Supplementary Figure 17. Flexibility changes in the β A domain upon activation. Structural time course (S1 \rightarrow S8) of integrin activation in the β A domain of $\alpha_{11b}\beta_3$ integrins derived from crystal structures (PDB entries: 3ZDX, 3ZDY, 3ZDZ, 3ZE0, 3ZE1, 3ZE2)¹⁷. Colors denote the average B factor of the amino acids, ranging from 0 (blue) to the maximum in the β A domain of the respective structure (red).



Supplementary Figure 18. Comparison between *nor*UDCA- and TUDC-induced Erk-1/2 activation. Rat livers were perfused with *nor*UDCA or TUDC (10 or 50 $\mu\text{mol/l}$ each) for up to 60 min. Liver samples were taken at the time points indicated. Phosphorylation of Erk-1/2 was analyzed by (A) Western blot using specific antibodies and (B) subsequent densitometric analysis. Levels of phosphorylated Erk-1/2 were normalized to total Erk-1/2. Phosphorylation at $t = 0$ was arbitrarily set to 1. Data represent the mean (mean \pm SEM) of at least three independent experiments; *: Statistically significant compared to the unstimulated control ($t = 0$) (one-way ANOVA, Dunnett's *post hoc* test).



Supplementary Figure 19. Subunit (A) and domain organization (B) of the $\alpha_5\beta_1$ integrin ectodomain. (A) Illustration of the subunit organization in $\alpha_5\beta_1$ integrin. The α -subunit is depicted in dark gray, the β -subunit is depicted in light gray. (B) Illustration of the domain organization in $\alpha_5\beta_1$ integrin. β -Propeller domain: dark blue, thigh domain: light blue, calf-1/2 domains: green; β A domain: yellow, hybrid domain: ochre, PSI domain: dark orange, I-EGF 1-4 domains: pink, β -tail domain: grey.

Supplementary Tables

Supplementary Table 1. Domain-wise minimum, maximum and average RMSD values.^a

α -subunit RMSD	propeller			thigh			calf-1			calf-2											
	min	mean	max	min	mean	max	min	mean	max	min	mean	max		min	mean	max		min	mean	max	
TUDC	1	1.53	1.99	2.71	1.87	2.47	3.37	1.49	2.04	2.58	2.13	4.51	5.86								
	2	1.58	2.08	2.60	1.63	3.05	3.83	1.68	2.21	2.73	2.24	4.02	5.30								
	3	1.48	2.14	2.82	1.59	2.60	3.41	1.81	2.31	2.76	2.11	5.13	6.61								
norUDCA	1	1.61	2.33	3.18	1.41	2.59	3.51	1.55	2.07	2.53	2.04	4.86	6.04								
	2	1.34	2.24	3.02	1.69	2.94	4.04	1.59	2.57	3.19	2.25	3.33	4.58								
	3	1.38	2.06	2.56	1.62	2.62	3.37	1.67	2.27	2.82	2.19	4.49	5.25								
TnorUDCA	1	1.69	2.64	3.28	1.57	2.68	3.18	1.63	2.22	2.74	2.14	4.62	5.36								
	2	1.28	2.18	2.74	1.66	2.67	3.54	1.60	2.52	3.08	2.07	4.20	5.16								
	3	1.56	1.90	2.43	1.48	2.46	3.38	1.56	2.16	2.87	2.26	3.18	4.25								
GUDC	1	1.40	2.17	3.22	1.81	2.42	3.10	1.44	2.21	2.92	2.11	4.76	6.74								
	2	1.58	2.51	3.36	1.70	2.82	3.54	1.57	2.48	2.94	2.22	4.32	5.54								
	3	1.44	2.02	2.77	1.65	2.38	3.32	1.86	2.47	3.10	2.14	4.44	5.60								
UDCA	1	1.55	2.43	3.17	1.73	2.90	3.48	1.54	1.91	2.36	2.36	3.54	4.50								
	2	1.42	2.24	2.77	1.75	2.86	3.63	1.71	2.23	2.79	2.12	4.59	5.55								
	3	1.42	2.14	3.03	1.65	2.44	3.02	1.56	1.99	2.48	2.38	5.98	6.91								
TC	1	1.63	2.42	3.15	1.65	2.58	3.32	1.53	2.17	2.76	2.06	3.37	4.13								
	2	1.49	2.26	3.20	1.68	2.53	3.87	1.63	2.29	2.77	2.07	4.16	5.92								
	3	1.56	2.10	3.11	1.83	2.74	3.57	1.72	2.78	3.47	2.31	5.58	6.55								
β -subunit RMSD	β A			hybrid			PSI			I-EGF3			I-EGF4			β -tail					
	min	mean	max	min	mean	max	min	mean	max	min	mean	max	min	mean	max	min	mean	max			
TUDC	1	1.38	3.07	4.23	1.86	3.79	4.50	2.37	5.09	7.16	0.88	1.64	2.64	1.23	3.99	6.02	1.99	4.25	5.59		
	2	1.35	3.21	3.89	1.88	3.43	4.53	2.21	5.60	7.26	0.98	2.47	3.35	1.12	1.84	3.25	2.02	4.74	6.16		
	3	1.29	2.02	2.66	1.72	3.84	5.06	2.25	5.24	6.77	1.12	3.41	5.86	1.29	2.18	2.86	1.82	4.01	5.27		
norUDCA	1	1.26	2.30	3.84	1.69	4.08	4.77	2.34	4.55	6.85	1.00	1.60	2.82	1.17	1.92	2.77	2.14	5.47	6.22		
	2	1.33	2.01	2.78	1.85	3.94	4.59	2.56	4.93	6.23	1.15	2.13	3.26	0.96	1.86	2.92	1.85	3.76	4.86		
	3	1.37	2.33	2.77	1.79	3.36	4.86	2.95	6.86	8.71	1.03	2.68	4.15	1.18	2.91	4.43	1.83	3.63	4.96		
TnorUDCA	1	1.40	2.39	4.80	1.50	3.70	4.86	2.09	4.41	6.57	1.01	2.82	5.01	1.19	1.90	2.86	1.86	6.61	7.94		
	2	1.39	2.03	2.42	1.55	4.11	4.89	2.37	5.66	7.11	0.97	2.11	3.77	1.03	2.30	3.40	2.10	4.06	5.31		
	3	1.37	1.77	2.22	1.92	4.2	5.39	2.90	5.82	7.56	0.99	3.13	5.01	1.09	1.64	2.76	2.10	5.36	8.95		
GUDC	1	1.44	2.55	3.17	1.48	3.71	5.01	2.87	4.41	6.00	1.19	3.46	4.93	1.09	2.15	4.01	2.01	3.66	4.92		
	2	1.31	2.26	2.89	1.66	3.16	4.66	1.96	5.28	6.75	1.02	2.53	3.38	1.12	1.80	2.44	1.79	3.75	5.13		
	3	1.54	2.67	3.65	1.77	3.25	4.55	2.75	5.21	6.73	1.04	2.12	3.19	0.97	1.83	3.15	1.81	3.82	4.94		
UDCA	1	1.46	2.71	3.38	1.53	3.70	5.32	2.30	5.99	7.54	0.92	1.65	2.38	1.10	1.78	2.57	1.89	3.69	5.21		
	2	1.40	2.20	2.82	1.75	4.26	5.18	2.10	4.06	5.18	1.02	3.44	4.38	1.01	1.84	2.95	1.92	4.54	5.77		
	3	1.47	2.25	2.72	1.62	3.41	4.88	2.68	5.20	6.76	1.09	1.84	2.30	1.02	1.62	2.58	1.78	4.47	5.72		
TC	1	1.48	2.06	3.36	1.72	4.97	6.89	2.38	4.04	5.39	1.09	2.92	4.44	1.07	1.97	4.01	1.79	4.99	6.94		
	2	1.33	2.29	3.28	1.49	4.15	5.40	2.32	4.58	5.91	0.97	2.42	3.09	1.05	1.65	2.41	2.07	3.99	5.47		
	3	1.49	2.35	2.78	1.50	3.70	4.74	2.53	5.46	6.81	1.27	2.61	3.78	1.06	1.55	2.63	1.84	4.23	7.11		

^a RMSD values larger than 4 Å and larger than 6 Å are highlighted in orange and red, respectively.

Supplementary Table 2. Angles and distances computed from MD trajectories.^a

Simulation (number)	$\alpha 1$ kink angle^b	$\alpha 7$ tilt angle^b	Propeller-βA distance^c
TUDC (1)	156.54 \pm 0.05	146.48 \pm 0.05	37.35 \pm 0.004
TUDC (2)	139.98 \pm 0.10	130.74 \pm 0.08	35.72 \pm 0.004
TUDC (3)	145.25 \pm 0.08	137.27 \pm 0.08	36.97 \pm 0.007
TUDC ^d	147.26 \pm 0.13	138.16 \pm 0.12	36.68 \pm 0.010
<i>nor</i> UDCA (1)	150.23 \pm 0.05	140.53 \pm 0.05	35.61 \pm 0.002
<i>nor</i> UDCA (2)	150.22 \pm 0.05	137.21 \pm 0.08	36.89 \pm 0.005
<i>nor</i> UDCA (3)	147.29 \pm 0.04	137.44 \pm 0.04	35.63 \pm 0.003
<i>nor</i> UDCA ^d	149.25 \pm 0.08	138.39 \pm 0.10	36.04 \pm 0.006
<i>Tnor</i> UDCA (1)	141.21 \pm 0.05	130.77 \pm 0.05	35.36 \pm 0.003
<i>Tnor</i> UDCA (2)	146.04 \pm 0.06	137.48 \pm 0.05	36.24 \pm 0.003
<i>Tnor</i> UDCA (3)	141.17 \pm 0.04	141.50 \pm 0.05	35.07 \pm 0.004
<i>Tnor</i> UDCA ^d	142.81 \pm 0.09	136.58 \pm 0.09	35.56 \pm 0.006
GUDC (1)	142.60 \pm 0.05	131.86 \pm 0.06	35.27 \pm 0.005
GUDC (2)	145.54 \pm 0.05	134.80 \pm 0.05	36.31 \pm 0.003
GUDC (3)	144.91 \pm 0.05	142.76 \pm 0.06	35.38 \pm 0.002
GUDC ^d	144.35 \pm 0.09	136.47 \pm 0.10	35.65 \pm 0.006
UDCA (1)	145.98 \pm 0.05	140.64 \pm 0.09	36.07 \pm 0.006
UDCA (2)	136.43 \pm 0.06	127.92 \pm 0.05	35.43 \pm 0.003
UDCA (3)	136.31 \pm 0.07	120.08 \pm 0.07	35.90 \pm 0.006
UDCA ^d	139.58 \pm 0.11	129.55 \pm 0.12	35.80 \pm 0.006
TC (1)	142.46 \pm 0.05	126.11 \pm 0.07	35.76 \pm 0.004
TC (2)	140.49 \pm 0.06	127.91 \pm 0.07	35.74 \pm 0.005
TC (3)	142.96 \pm 0.07	125.81 \pm 0.07	35.80 \pm 0.007
TC ^{d, e}	141.97 \pm 0.10	126.61 \pm 0.12	35.77 \pm 0.010
Overall mean ^f	144.20 \pm 0.25	134.30 \pm 0.27	35.92 \pm 0.019

^a For the geometric parameters, the average over the last 100 ns \pm SEM is shown. See Figure 2 in the main text for a definition of these parameters.

^b In $^\circ$.

^c In Å

^d Data averaged over the three replicate MD simulations performed for each bile acid.

^e MD simulations with TC served as negative control.

^f Data averaged over all 18 MD simulations performed for the bile acids.

Supplementary Table 3. Statistical testing for differences in mean values.^a

$\alpha 1$ tilt angle ^b	TUDC			norUDCA			TnorUDCA			GUDC			UDCA			TC			
	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	
TUDC	1	2	3	****															
norUDCA	1	2	3	****	****	****	****	****	****										
	2	3	****	****	****	****	****	****	****										
	3	****	****	****	****	****	****	****	****										
TnorUDCA	1	2	3	****	****	****	****	****	****	****	****	****							
	2	3	****	****	****	****	****	****	****	****	****	****	****						
	3	****	****	****	****	****	****	****	****	****	****	****	****	****					
GUDC	1	2	3	****	****	****	****	****	****	****	****	****	****	****	****				
	2	3	****	****	****	****	****	****	****	****	****	****	****	****	****				
	3	****	****	****	****	****	****	****	****	****	****	****	****	****	****				
UDCA	1	2	3	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****
	2	3	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****
	3	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****
TC	1	2	3	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****
	2	3	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****
	3	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****
$\alpha 7$ kink angle ^b	TUDC			norUDCA			TnorUDCA			GUDC			UDCA			TC			
	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	
TUDC	1	2	3	****															
norUDCA	1	2	3	****	****	****	****	****	****										
	2	3	****	****	****	****	****	****	****										
	3	****	****	****	****	****	****	****	****										
TnorUDCA	1	2	3	****	****	****	****	****	****	****	****	****							
	2	3	****	****	****	****	****	****	****	****	****	****							
	3	****	****	****	****	****	****	****	****	****	****	****							
GUDC	1	2	3	****	****	****	****	****	****	****	****	****	****	****	****				
	2	3	****	****	****	****	****	****	****	****	****	****	****	****	****				
	3	****	****	****	****	****	****	****	****	****	****	****	****	****	****				
UDCA	1	2	3	****	ns	****	****	****	****	****	****	****	****	****	****	****	****	****	****
	2	3	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****
	3	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****
TC	1	2	3	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****
	2	3	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****
	3	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****
Propeller- βA distance ^b	TUDC			norUDCA			TnorUDCA			GUDC			UDCA			TC			
	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	
TUDC	1	2	3	****															
norUDCA	1	2	3	****	****	****	****	****	****										
	2	3	****	****	****	****	****	****	****										
	3	****	****	****	****	****	****	****	****										
TnorUDCA	1	2	3	****	****	****	****	****	****	****	****	****							
	2	3	****	****	****	****	****	****	****	****	****	****							
	3	****	****	****	****	****	****	****	****	****	****	****							
GUDC	1	2	3	****	****	****	****	****	****	****	****	****	****	****	****				
	2	3	****	****	****	****	****	****	****	****	****	****	****	****	****				
	3	****	****	****	****	****	****	****	****	****	****	****	****	****	****				
UDCA	1	2	3	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****
	2	3	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****
	3	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****
TC	1	2	3	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****
	2	3	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****
	3	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****	****

^a ****: $p < 0.0001$, ***: $p < 0.001$, **: $p < 0.01$, *: $p < 0.05$, ns: $p \geq 0.05$. Results from redundant tests are omitted.

^b See Figure 2 in the main text for a definition of the geometric parameters.

Supplementary References

- 1 Reinehr, R., Gohlke, H., Sommerfeld, A., vom Dahl, S. & Häussinger, D. Activation of Integrins by Urea in Perfused Rat Liver. *J. Biol. Chem.* **285**, 29348-29356, doi:10.1074/jbc.M110.155135 (2010).
- 2 Gohlke, H., Schmitz, B., Sommerfeld, A., Reinehr, R. & Häussinger, D. $\alpha_5\beta_1$ -integrins are sensors for tauroursodeoxycholic acid in hepatocytes. *Hepatology* **57**, 1117-1129, doi:10.1002/hep.25992 (2013).
- 3 Suite 2012: Maestro, version 9.3 (Schrödinger LLC, New York, NY, 2012).
- 4 Suite 2012: LigPrep, version 2.5 (Schrödinger LLC, New York, NY, 2012).
- 5 Shelley, J. C. *et al.* Epik: a software program for pK(a) prediction and protonation state generation for drug-like molecules. *J. Comput.-Aided Mol. Des.* **21**, 681-691, doi:10.1007/s10822-007-9133-z (2007).
- 6 MacroModel, version 9.9 (Schrödinger LLC, New York, NY, 2012).
- 7 Gerber, P. R. & Müller, K. MAB, a generally applicable molecular force field for structure modelling in medicinal chemistry. *J. Comput.-Aided Mol. Des.* **9**, 251-268, doi:10.1007/BF00124456 (1995).
- 8 AMBER 11 (University of California, San Francisco, 2010).
- 9 Hornak, V. *et al.* Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins* **65**, 712-725, doi:10.1002/prot.21123 (2006).
- 10 Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **25**, 1157-1174, doi:10.1002/jcc.20035 (2004).
- 11 Åqvist, J. Ion-Water Interaction Potentials Derived from Free-Energy Perturbation Simulations. *J. Phys. Chem.* **94**, 8021-8024, doi:10.1021/j100384a009 (1990).
- 12 Bayly, C. I., Cieplak, P., Cornell, W. D. & Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.* **97**, 10269-10280, doi:10.1021/j100142a004 (1993).
- 13 Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926-935, doi:10.1063/1.445869 (1983).
- 14 Darden, T. A., York, D. & Pedersen, L. Particle mesh Ewald - an N·Log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089-10092, doi:10.1063/1.464397 (1993).

-
- 15 Schmitt, M., Kubitz, R., Lizun, S., Wettstein, M. & Häussinger, D. Regulation of the Dynamic Localization of the Rat Bsep Gene-Encoded Bile Salt Export Pump by Anisoosmolarity. *Hepatology* **33**, 509-518, doi:10.1053/jhep.2001.22648 (2001).
 - 16 Kubitz, R., D'Urso, D., Keppler, D. & Häussinger, D. Osmodependent Dynamic Localization of the Multidrug Resistance Protein 2 in the Rat Hepatocyte Canalicular Membrane. *Gastroenterology* **113**, 1438-1442, doi:gast.1997.v113.pm9352844 (1997).
 - 17 Zhu, J., Zhu, J. & Springer, T. A. Complete integrin headpiece opening in eight steps. *J. Cell Biol.* **201**, 1053-1068, doi:10.1083/jcb.201212037 (2013).

Publication II**Isoform-specific Inhibition of *N*-methyl-D-aspartate Receptors
by Bile Salts**

Angela Koch[§], **Michele Bonus**[§], Holger Gohlke, Nikolaj Klöcker

Scientific Reports (2019), 9, 10068

Contribution: 30%

[§] Both authors contributed equally to this work

SCIENTIFIC REPORTS

OPEN

Isoform-specific Inhibition of *N*-methyl-D-aspartate Receptors by Bile Salts

Angela Koch¹, Michele Bonus^{1,2}, Holger Gohlke^{1,2,3} & Nikolaj Klöcker¹

Received: 28 January 2019

Accepted: 18 June 2019

Published online: 11 July 2019

The *N*-methyl-D-aspartate subfamily of ionotropic glutamate receptors (NMDARs) is well known for its important roles in the central nervous system (CNS), e.g. learning and memory formation. Besides the CNS, NMDARs are also expressed in numerous peripheral tissues including the pancreas, kidney, stomach, and blood cells, where an understanding of their physiological and pathophysiological roles is only evolving. Whereas subunit composition increases functional diversity of NMDARs, a great number of endogenous cues tune receptor signaling. Here, we characterized the effects of the steroid bile salts cholate and chenodeoxycholate (CDC) on recombinantly expressed NMDARs of defined molecular composition. CDC inhibited NMDARs in an isoform-dependent manner, preferring GluN2D and GluN3B over GluN2A and GluN2B receptors. Determined IC_{50} values were in the range of bile salt serum concentrations in severe cholestatic disease states, pointing at a putative pathophysiological significance of the identified receptor modulation. Both pharmacological and molecular simulation analyses indicate that CDC acts allosterically on GluN2D, whereas it competes with agonist binding on GluN3B receptors. Such differential modes of inhibition may allow isoform-specific targeted interference with the NMDAR/bile salt interaction. In summary, our study provides further molecular insight into the modulation of NMDARs by endogenous steroids and points at a putative pathophysiological role of the receptors in cholestatic disease.

N-methyl-D-aspartate receptors (NMDARs) comprise a subfamily of ionotropic glutamate receptors (iGluRs). NMDARs assemble from combinations of GluN1, GluN2 and GluN3 subunits, whereby they build tetrameric ion channels, with GluN1 being obligatory for channel function. Distinct from other iGluR subfamilies, most NMDARs show significant Ca^{2+} permeability and exhibit an apparent voltage-dependence due to voltage-dependent block by extracellular Mg^{2+} ions^{1,2}. Activation of most NMDAR subunit combinations requires binding of the agonist glutamate to GluN2 and the co-agonists glycine or D-serine to GluN1 subunits³. The exceptions are channels assembled from GluN1 and GluN3, which lack a glutamate-binding site and hence serve as glycine-activated cation channels⁴. Aside from their well-characterized expression in the brain, both in neuronal and non-neuronal cells, NMDARs have been identified in a number of peripheral tissues and systems⁵. These include circulating blood cells such as lymphocytes and platelets, which may employ NMDARs for cell activation. As non-neuronal NMDARs are presumably directly exposed to high levels of plasma glutamate, alternative agonists such as L-homocysteic acid (L-HCA) and quinolinic acid (QN) are discussed as receptor ligands⁵.

NMDAR activity is modulated by numerous endogenous cues including pH, metal ions, polyamines, and lipids⁶. Neurosteroids derived from the cholesterol backbone may allosterically either potentiate or inhibit NMDAR function in an isoform-dependent manner^{6–9}. The main steroid salts contained in human bile, cholate and chenodeoxycholate (CDC), are also derived from cholesterol. These primary bile salts are synthesized by hepatocytes, conjugated to glycine or taurine to increase water solubility, and finally secreted into the intestine to help emulsifying food lipids. Bile salts are efficiently reabsorbed from the intestine and transported back to the liver for reuse by the hepatic portal system (enterohepatic cycle). Under physiological conditions, only insignificant amounts of bile salts enter systemic circulation. In certain cholestatic disease states like progressive

¹Institute of Neural and Sensory Physiology, Medical Faculty, Heinrich Heine University Düsseldorf, 40225, Düsseldorf, Germany. ²Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, 40225, Düsseldorf, Germany. ³John von Neumann Institute for Computing (NIC), Jülich Supercomputing Centre (JSC) & Institute for Complex Systems - Structural Biochemistry (ICS 6), Forschungszentrum Jülich GmbH, 52425, Jülich, Germany. Angela Koch and Michele Bonus contributed equally. Correspondence and requests for materials should be addressed to N.K. (email: nikolaj.kloecker@uni-duesseldorf.de)

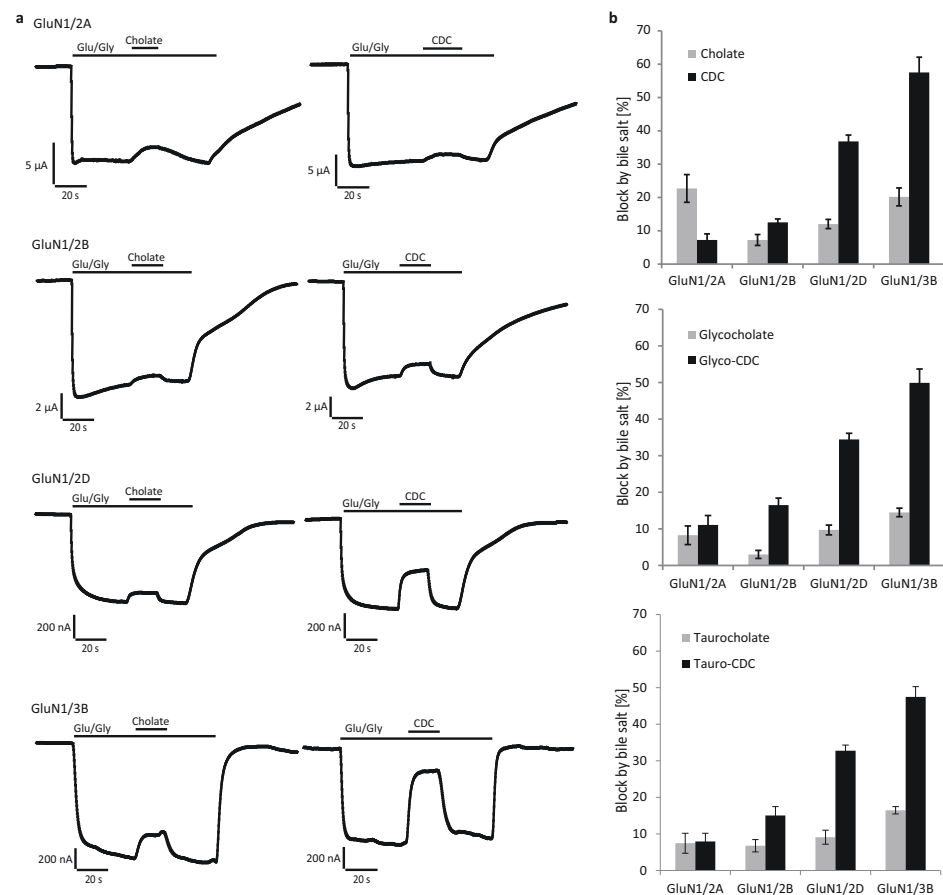


Figure 1. Inhibition of NMDARs by the bile salts cholate and CDC. **(a)** Representative current traces recorded from *Xenopus laevis* oocytes expressing NMDARs of indicated molecular compositions. Note the reversible reduction of agonist-induced currents by cholate and CDC. **(b)** Quantification of the experiments in **(a)**. Given is the mean reduction (\pm SEM) of indicated NMDAR currents by application of 100 μ M bile salts (in % of steady-state agonist-induced current, $n = 6$ –14 oocytes). Glutamate (Glu): 150 μ M, Glycine (Gly): 10 μ M.

familial forms of intrahepatic cholestasis or severe forms of it in pregnancy or acute hepatitis, however, plasma concentrations of bile salts may rise to hundred μ M and more^{10–12}. When applied to neurons, cholate and CDC block NMDA receptors¹³ and it is known that CDC can improve neurological symptoms of cerebrotendinous xanthomatosis, a disorder of bile salt metabolism¹⁴. Another bile salt, tauroursodeoxycholate, is neuroprotective on NMDA-induced retinal ganglion cell degeneration¹⁵. Furthermore, NMDA receptors are reported to be involved in modulation of cholestasis-induced antinociception in rat^{16,17} and there is a potential involvement of the dorsal hippocampal (CA1) glutamatergic system in cholestasis-induced decrease in rats' memory function¹⁸. Thus, bile salt effects on NMDA receptors can have a physiological and pathophysiological relevance in nervous system or peripheral tissues.

Here, we have studied the effects of the most common human bile salts, cholate and CDC, on NMDARs of different subunit composition. We demonstrate that bile salts inhibit NMDARs in an isoform-specific manner, both with respect to efficacy and to the molecular mechanism of action. Our data shed new light on the effects of bile salts in cholestatic diseases as well as when used in therapy.

Results

CDC inhibits preferentially the GluN subtypes 2D and 3B.

The effects of the bile salts cholate and CDC on GluNs were probed on recombinant receptors of defined subunit composition expressed in *Xenopus laevis* oocytes. For initial experiments, agonists were applied at saturating concentration to ensure maximum activation of GluN2A, GluN2B, GluN2D, and GluN3B, respectively^{4,19}. Cholate inhibited steady-state currents only moderately irrespective of receptor subtype, while CDC was more potent and reduced preferentially GluN2D and GluN3B currents (Fig. 1, Table S1). Inhibition by 100 μ M CDC ranged between $7 \pm 2\%$ of GluN2A and $13 \pm 1\%$ of GluN2B steady-state currents, whereas it amounted to $37 \pm 2\%$ of GluN2D and even to $58 \pm 5\%$ of GluN3B currents. The bile salt effects were specific for GluNs, as no inhibition of glutamate-induced steady-state kainate

receptor GluK2 current ($n = 3$, after block of desensitization with ConA) and only a minor reduction of α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) receptor GluA1/GluA2 currents ($7 \pm 1\%$, $n = 6$) were observed after application of $100 \mu\text{M}$ CDC. Somewhat unexpected, taurine or glycine conjugation of bile salts did neither change their overall potencies in inhibiting GluNs nor their subunit preferences.

Allosteric inhibition of GluN2D by tauro-CDC. To dissect the mechanism of GluN2D inhibition by bile salts, we recorded dose-response relations for CDC at both saturating and EC_{50} NMDAR agonist concentrations^{20,21}. We used routinely taurine-conjugated CDC in these experiments, because of higher water solubility and hence more feasible handling.

As depicted in Fig. 2a, maximum inhibition of glutamate/glycine-induced currents by tauro-CDC was significantly smaller at saturating than at EC_{50} agonist concentrations ($n = 8-13$; $p = 0.01$ and 0.03 for glutamate and glycine, respectively).

Conversely, dose-response curves for glutamate and glycine in the absence or presence of $180 \mu\text{M}$ tauro-CDC, which was calculated to be the mean IC_{50} ($179 \pm 11 \mu\text{M}$, $n = 13$, saturating agonist concentration), showed that the maximum agonist-induced currents were significantly reduced by the bile salt (Fig. 2b,c; $n = 7-13$, $p = 0.0002$ and 0.003 for glycine and glutamate, respectively). There were neither significant differences between the EC_{50} values for glutamate or glycine at varying concentrations of tauro-CDC ($n = 7-13$, $p = 0.3$ and 0.2 for glutamate and glycine respectively) nor were there different IC_{50} values for tauro-CDC at varying concentrations of glycine ($n = 8-13$; $p = 0.1$; Table S2). Only the IC_{50} value for tauro-CDC at the glutamate EC_{50} concentration was significantly smaller compared to inhibition at a saturating glutamate concentration ($n = 11-13$, $p = 0.009$).

Application of $100 \mu\text{M}$ tauro-CDC in the absence of NMDAR agonists showed no effects on GluN2D or GluN3B receptors ($n = 4$) excluding partial agonist action of bile salts on NMDA receptors (data not shown). Taken together, these results strongly suggest that tauro-CDC inhibits GluN2D receptors by an allosteric mechanism.

Competitive inhibition of GluN3B by tauro-CDC. Next, we sought to investigate the mechanism of action of tauro-CDC-mediated inhibition of GluN3B receptors. As depicted in Fig. 3a, the mean IC_{50} for tauro-CDC was significantly higher in the presence of $10 \mu\text{M}$ glycine than in the presence of $5 \mu\text{M}$ glycine ($83 \pm 8 \mu\text{M}$ vs. $44 \pm 7 \mu\text{M}$, $n = 6$, $p = 0.007$). The maximum block by tauro-CDC, however, did not depend on the glycine concentration ($95 \pm 3\%$ vs. $92 \pm 3\%$, $n = 6$, $p = 0.5$). These data indicate that tauro-CDC inhibits GluN3B receptors – other than GluN2D receptors – in a competitive manner. We also recorded glycine dose-response curves in the presence or absence of $100 \mu\text{M}$ tauro-CDC ($n = 6-11$, Fig. 3b). Known receptor desensitization at higher glycine concentrations prevented calculation of precise EC_{50} values^{4,22}. However, as depicted in Fig. 3b, the bell-shaped concentration-response curve shifted to higher glycine concentrations in the presence of tauro-CDC.

Tauro-CDC has no effect on ion selectivity. Bile salts are known to integrate into the plasma membrane of cells²³, which might impact ion channel pores including their selectivity filters. Therefore, we checked for potential bile salt effects on GluN receptor ion selectivity. Current-voltage (I/V) relationships of agonist-induced currents were recorded for GluN2D and GluN3B in the absence and presence of $100 \mu\text{M}$ tauro-CDC ($n = 10-11$). For better comparison of I/V curves, individual oocyte data were normalized to their maximum inward current at -140 mV. As shown in Fig. 4, tauro-CDC did neither change the shape of I/V relationships nor did it shift the reversal potentials of the receptor currents. These results render a change in ion selectivity of the two GluNs by bile salts unlikely.

Bile salt inhibition of GluNs when activated by alternative agonists. As the inhibition of GluNs by bile salts required concentrations that are hardly reached in the CNS, we tested whether tauro-CDC also reduced currents elicited by the alternative non-neuronal NMDAR agonists L-homocysteic acid (L-HCA) and quinolinic acid (QN)⁵.

As shown in Fig. 5, GluN2D currents induced by application of L-HCA at its reported EC_{50} of $13 \mu\text{M}$ ²⁴ were blocked by $30 \pm 2\%$ ($n = 12$), whereas currents elicited by QN at its reported EC_{50} of 7.2 mM ²⁴ were reduced by $26 \pm 2\%$ ($n = 7$) when co-applying tauro-CDC at $100 \mu\text{M}$. Lowering the agonist concentration of QN to 1 mM did not change inhibition by tauro-CDC, which was still $29 \pm 2\%$ ($n = 7$). For all these experiments involving the non-neuronal NMDAR agonists L-HCA and QN on GluN2D receptors, $10 \mu\text{M}$ glycine was used as co-agonist. GluN3B could not be activated by L-HCA, even at high concentrations of up to 1 mM ($n = 4$). In contrast, currents activated by 7.2 mM QN were reduced by about $68 \pm 4\%$ ($n = 5$), with QN showing no effect on uninjected oocytes ($n = 3$).

Binding modes of tauro-CDC in GluN2D_{LBD} and GluN3B_{LBD}. The pharmacological experiments indicated that tauro-CDC acts as a non-competitive allosteric antagonist of GluN2D, but as a competitive antagonist of GluN3B. In order to assess the molecular origin of isoform-specific inhibition of NMDARs by tauro-CDC, we performed induced fit docking experiments of tauro-CDC in respective ligand binding domains (LBDs) of the NMDAR isoforms (Fig. 6a,b), which allows to consider receptor flexibility in terms of side chain rearrangements.

For both GluN2D_{LBD} and GluN3B_{LBD}, binding modes of tauro-CDC (Fig. 6c,d) were predicted with similar docking scores (IFDScore²⁵) of $\sim 7.4 \text{ kcal}\cdot\text{mol}^{-1}$. In agreement, both binding modes were of similar geometry with an in-place root-mean-square deviation (RMSD) of the ligand atoms of 2.98 \AA after superimposition of the C_{α} atoms of the protein structures (RMSD: 1.45 \AA). The similarity in binding mode geometry was paralleled by a similar interaction profile in both complexes (Fig. 6c,d).

Hence, the static structures obtained from docking experiments cannot explain the observed isoform specificity of NMDAR inhibition by tauro-CDC.

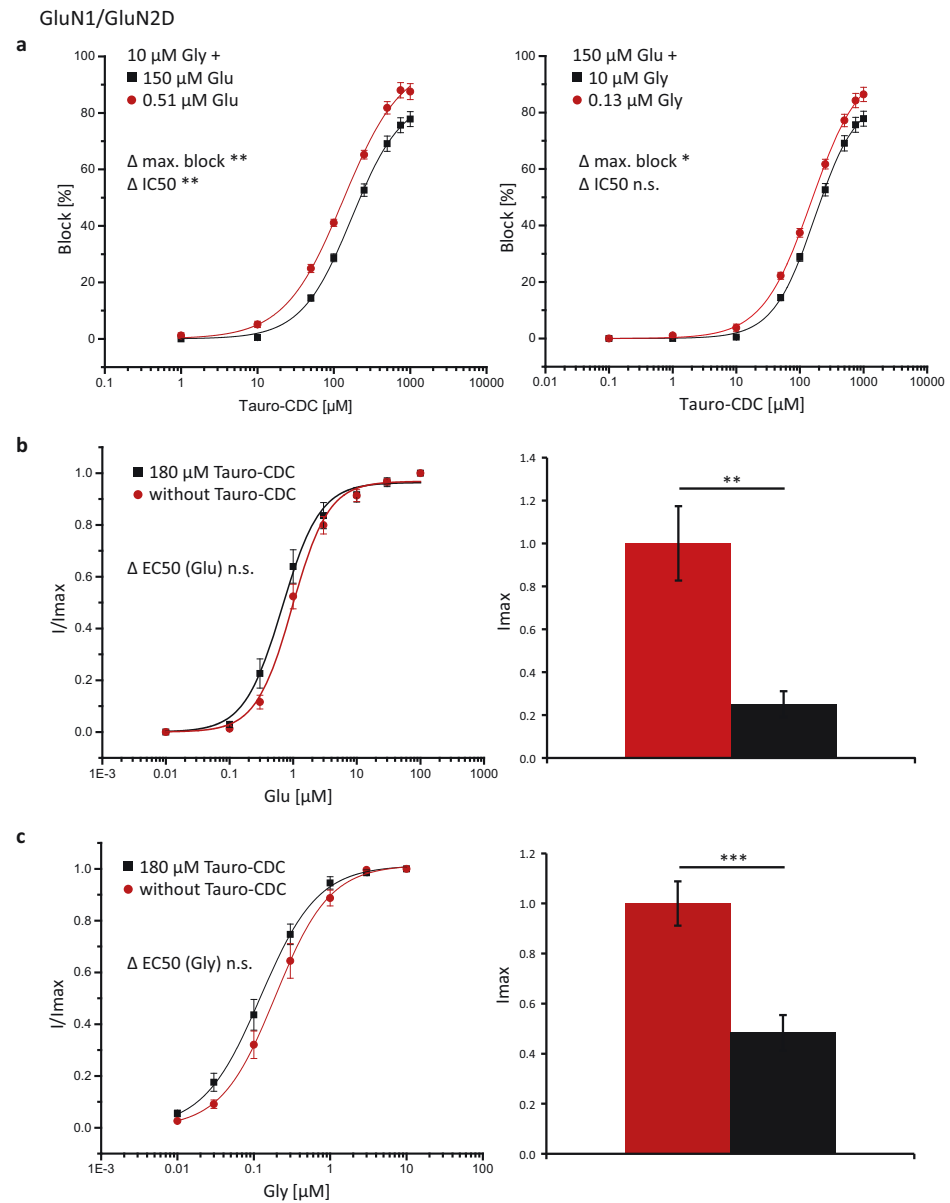


Figure 2. Tauro-CDC inhibits GluN2D receptors in an allosteric manner. (a) Concentration-response curves for inhibition of GluN1/GluN2D receptors by tauro-CDC at saturating (squares) and half-maximum (dots) agonist concentrations (glutamate/glycine, Glu/Gly) as indicated^{20,21}. Displayed is mean \pm SEM of recordings from $n = 8$ –13 oocytes. (b) Concentration-response curves (left) and maximum currents (right) induced by the agonist Glu in the presence (squares) and absence (dots) of 180 μM tauro-CDC. The concentration of the co-agonist Gly was kept constant at 10 μM . Displayed is the mean \pm SEM of recordings from $n = 7$ oocytes. (c) Concentration-response curves and maximum currents induced by the co-agonist Gly in the presence (squares) and absence (dots) of 180 μM tauro-CDC. The concentration of the agonist Glu was kept constant at 150 μM . Displayed is the mean \pm SEM of recordings from $n = 12$ –13 oocytes. The results of t-tests are reported as follows: n.s. $p > 0.05$, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p < 0.001$.

Structural variability of tauro-CDC/GluN2D_{LBD} and tauro-CDC/GluN3B_{LBD} binding modes. We next asked whether the predicted binding modes of tauro-CDC in GluN2D_{LBD} or GluN3B_{LBD} structurally vary over time scales accessible by molecular dynamics (MD) simulations. We therefore carried out ten MD

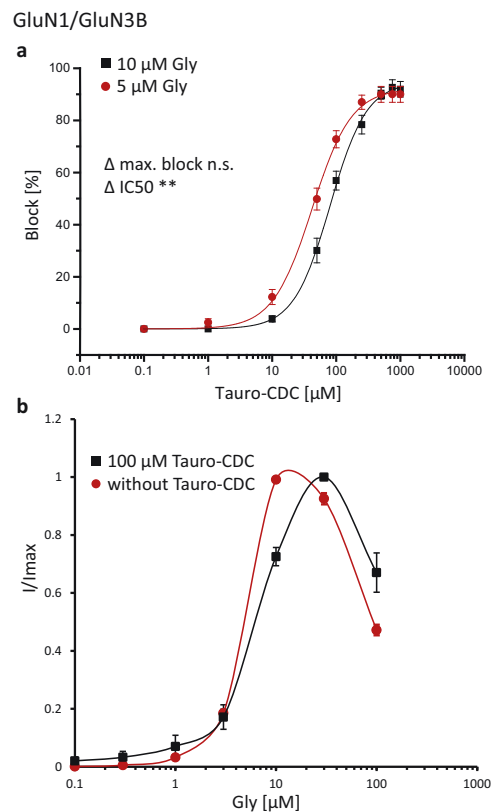


Figure 3. Tauro-CDC inhibits GluN3B receptors in a competitive manner. **(a)** Concentration-response curves for inhibition of GluN1/GluN3B receptors by tauro-CDC at saturating (squares) and estimated half-maximum (dots) glycine (Gly) concentrations as indicated⁴. Displayed is mean \pm SEM of recordings from $n = 6$ oocytes. **(b)** Concentration-response curves for Gly in the presence or absence of 100 μ M Tauro-CDC ($n = 6-11$ oocytes). The results of t-tests are reported as follows: n.s. $p > 0.05$, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p < 0.001$.

simulations (S_1-S_{10}) of 500 ns length each for both the tauro-CDC/GluN2D_{LBD} and tauro-CDC/GluN3B_{LBD} systems and quantified the structural deviation from the initial tauro-CDC pose within each simulation and between independent simulations. Similarly, we quantified the conformational variability of tauro-CDC to account for potential alternative binding modes.

Taking all simulations but S_9 into account, the average RMSD of the tauro-CDC pose in GluN2D_{LBD} was 2.86 ± 0.35 Å higher than in GluN3B_{LBD} (Supplementary Fig. 3a). Simulation S_9 of GluN2D_{LBD} was omitted from this and all subsequent analyses when tauro-CDC unbound after ~ 225 ns (Supplementary Fig. 3b). No unbinding event was observed in the simulations of tauro-CDC/GluN3B_{LBD}.

The average pairwise RMSD of the tauro-CDC pose of all snapshots extracted from the MD simulations was 8.44 ± 0.001 Å for tauro-CDC/GluN2D_{LBD} (Figs. 7a) and 4.82 ± 0.001 Å for tauro-CDC/GluN3B_{LBD} (Fig. 7b). Taking the quadratic mean of the pairwise RMSD as a measure for pose diversity²⁶ revealed that the tauro-CDC poses observed both within and between simulations were substantially more diverse in GluN2D_{LBD} systems (mean: 8.43 ± 0.46 Å, Supplementary Fig. 3c) than in GluN3B_{LBD} systems (mean: 4.89 ± 0.26 Å, Supplementary Fig. 3d) ($t(98) = 7.01$, $p < 10^{-4}$).

To conclude, these results reveal that the predicted binding mode of tauro-CDC in GluN2D_{LBD} is structurally less stable than in GluN3B_{LBD}. In one simulation of tauro-CDC in GluN2D_{LBD}, tauro-CDC even unbound. Note that simulating ligand unbinding events is challenging also for ligands with moderate affinities due to low off-rates²⁷, which may explain why the unbinding occurred only in one simulation. As to GluN3B_{LBD}, the results are in line with our experimental data, according to which tauro-CDC must bind to the glycine binding site in order to competitively inhibit GluN3B-containing NMDARs.

Contributions of hydrogen bonds to tauro-CDC/GluN2D_{LBD} and tauro-CDC/GluN3B_{LBD} complex stability. To assess whether a difference in structural variability between tauro-CDC/GluN2D_{LBD} and tauro-CDC/GluN3B_{LBD} may be attributed to differences in intermolecular interactions within the complexes, we quantified the amount of hydrogen bonds formed between tauro-CDC and an arginine residue that is crucial for binding of the agonist glutamate to GluN2D_{LBD} (Arg543) and of glycine to GluN3B_{LBD} (Arg538) over the course

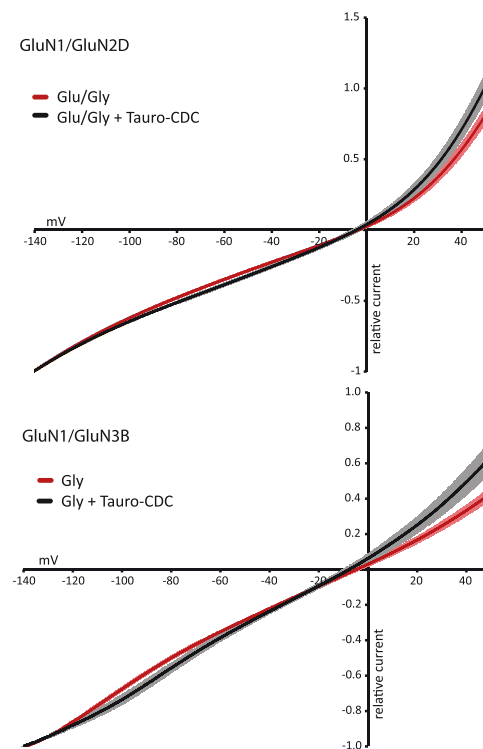


Figure 4. Tauro-CDC does not affect the ion selectivity of NMDARs. I/V relationships ($n = 10-11$, mean \pm SEM) for indicated NMDARs of agonist-induced currents in presence (black) and absence (red) of $100\mu\text{M}$ tauro-CDC. Data were normalized to maximum inward current at -140 mV . Note that the reversal potentials do not change in the presence of tauro-CDC.

of all MD simulations. Additionally, we quantified the total amount of hydrogen bonds between tauro-CDC and GluN2D_{LBD} or GluN3B_{LBD}. We focused here on identifying hydrogen bonds as interactions that contribute to the stability of the tauro-CDC/GluN3B_{LBD} complex in order to predict mutations for *ex vivo* validation of a suggested tauro-CDC binding mode.

The average hydrogen bond count between Arg543 in GluN2D_{LBD} and tauro-CDC was 0.31 ± 0.01 (Fig. 7c). In contrast, Arg538 in GluN3B_{LBD} formed hydrogen bonds to tauro-CDC about four times more frequently (mean: 1.31 ± 0.01) ($t(9498) = 58.58$, $p < 10^{-4}$) (Fig. 7d). Similarly, the average number of hydrogen bonds formed between the complete GluN2D_{LBD} and tauro-CDC (mean: 3.58 ± 0.03) (Fig. 7e) was significantly lower than the average number of hydrogen bonds formed between the complete GluN3B_{LBD} and tauro-CDC (mean: 4.97 ± 0.03) ($t(9498) = 38.10$, $p < 10^{-4}$) (Fig. 7f).

At a per-residue level, Arg543 showed the largest hydrogen bond occupancy (31%) for tauro-CDC/GluN2D_{LBD} (Fig. 8a). The corresponding residue in GluN3B_{LBD} (Arg538) displayed an occupancy of 131% (occupancy may exceed 100% as an arginine side chain can form more than one hydrogen bond) (Fig. 8b). Tauro-CDC in GluN3B_{LBD} also formed a hydrogen bond with the $3\alpha\text{-OH}$ group of the cholane scaffold to Asp447 (78% occupancy). This residue has no counterpart in the GluN2D sequence (Supplementary Fig. 1). In addition to Arg538 and Asp447, Ser701 (16%), Tyr505 (14%), and Ser533 (13%) were among the five residues that most frequently formed hydrogen bonds to tauro-CDC during the GluN3B_{LBD} simulations (Fig. 8b). Lys509 (29%), Arg437 (21%), Ser714 (15%) and Tyr755 (14%) additionally stabilized the GluN2D_{LBD}-tauro-CDC complex.

To conclude, hydrogen bond formation between tauro-CDC and Arg543 in GluN2D_{LBD} occurred less frequently than between tauro-CDC and Arg538 in GluN3B_{LBD}; the same held true when considering hydrogen bond formation to the complete LBD. The tauro-CDC/GluN3B_{LBD} complex was particularly stabilized by Arg538 and, to a lesser extent, Asp447, whereas no particular amino acid contributed to stabilizing tauro-CDC/GluN2D_{LBD}.

Effective binding energies of tauro-CDC/GluN2D_{LBD} and tauro-CDC/GluN3B_{LBD} complexes. To complement our findings from the structural analysis of tauro-CDC complexes, we quantified the effective binding energy of tauro-CDC to GluN2D_{LBD} or GluN3B_{LBD} with the end-point effective energy approaches MM/PBSA and MM/GBSA.

The effective binding energy computed by MM/GBSA for tauro-CDC/GluN2D_{LBD} (mean: $-33.53 \pm 0.67\text{ kcal}\cdot\text{mol}^{-1}$) was significantly less negative than for tauro-CDC/GluN3B_{LBD} (mean:

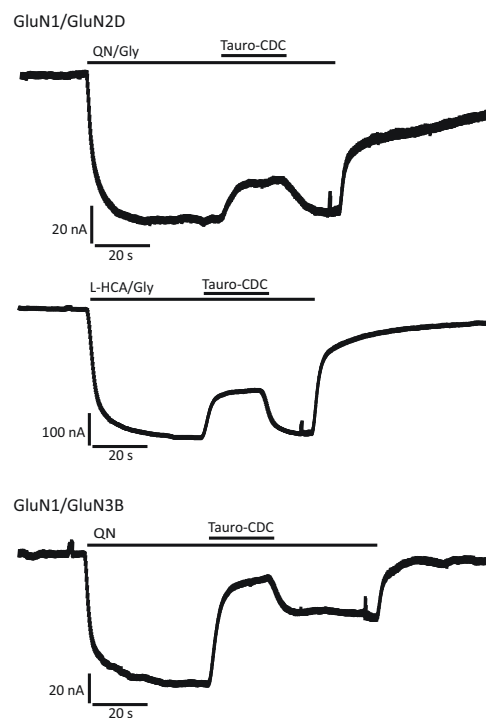


Figure 5. Bile salt inhibition of NMDARs activated by the alternative agonists quinolinic acid (QN) and L-homocysteic acid (L-HCA). Representative current traces recorded from *Xenopus laevis* oocytes expressing NMDARs of indicated molecular compositions. Note the reversible reduction of currents induced by QN (7.2 mM) and L-HCA (13 μ M) by tauro-CDC (100 μ M). Concentrations of QN and L-HCA are reported EC₅₀²⁴.

-48.90 ± 0.62 kcal·mol⁻¹) ($t(17) = 16.84$, $p < 10^{-4}$) (Fig. 9d). The same trend was obtained with the MM/PBSA approach (tauro-CDC/GluN2D_{LBD}: mean: -27.14 ± 0.52 kcal·mol⁻¹; tauro-CDC/GluN3B_{LBD}: mean: -30.54 ± 0.63 kcal·mol⁻¹; $t(17) = 4.07$, $p = 8 \cdot 10^{-4}$). The trend was stable irrespective of using a dielectric constant of the solute of $\epsilon = 1$ (values above; Fig. 9a,d) or 4 (tauro-CDC/GluN2D_{LBD}: mean: -43.98 ± 0.24 kcal·mol⁻¹; tauro-CDC/GluN3B_{LBD}: mean: -45.45 ± 0.22 kcal·mol⁻¹; $t(17) = 4.49$, $p = 3 \cdot 10^{-4}$) (Fig. 9d); higher dielectric constants have been used before to account for highly charged binding sites²⁸, as given in the case of the LBDs.

To conclude, both approaches revealed more negative effective energies of tauro-CDC binding to GluN3B_{LBD} than to GluN2D_{LBD}. Thus, they indicate a stronger binding of tauro-CDC to GluN3B_{LBD}.

Identification of an alternative binding site in the GluN1_{LBD}/GluN2D_{LBD} interface. We further aimed to identify an alternative binding mode of tauro-CDC to a potentially allosteric site in the GluN1_{LBD}/GluN2D_{LBD} tetramer. To do so, we first detected potential binding pockets in a conformational ensemble obtained from 500 ns of accelerated MD simulations of the tetrameric GluN1_{LBD}/GluN2D_{LBD} interface (Fig. 10a) using the MDpocket software²⁹. A single binding pocket at the GluN1/GluN2D_{LBD} interface was identified, and tauro-CDC was docked into the receptor conformation in which the pocket showed the highest mean local hydrophobic density^{29,30} (Fig. 10b). The identified pocket largely overlaps with the allosteric binding pocket of the negative allosteric modulator MPX-007 in the GluN1/GluN2A interface³¹.

During 500 ns of conventional MD simulations, tauro-CDC displayed an average RMSD value of 1.90 ± 0.01 Å (2.17 ± 0.01 Å without fitting, calculated on the last 250 ns of the simulation), indicating a low structural variability of the predicted binding mode. The binding mode is stabilized by interactions of the acid moiety with the positively charged residues Lys531 and Lys769 in the GluN1 subunit and Lys797 and Arg798 in the GluN2D subunit (Fig. 10d); Leu808 in the GluN2D subunit makes a hydrophobic contact with the cholane scaffold (Fig. 10d). Functional data on GluN1/GluN2A NMDARs indicates that Phe754 in GluN1 and Val783 in GluN2A form a molecular switch that mediates allosteric inhibition³¹. The corresponding Leu808 in GluN2D may assume a similar role as Val783 in GluN2A.

These tetramer simulations suggest that the interface between the GluN1 and GluN2D subunit forms a transient pocket, to which tauro-CDC can bind.

Point mutants. In order to verify the interaction of tauro-CDC with residues of the GluN3B LBD, Arg538 would be the primary residue for mutation studies due to its pronounced involvement in hydrogen bonds with the bile salt (Fig. 8b). However, we did not mutate Arg538 because this residue is reported to be important for

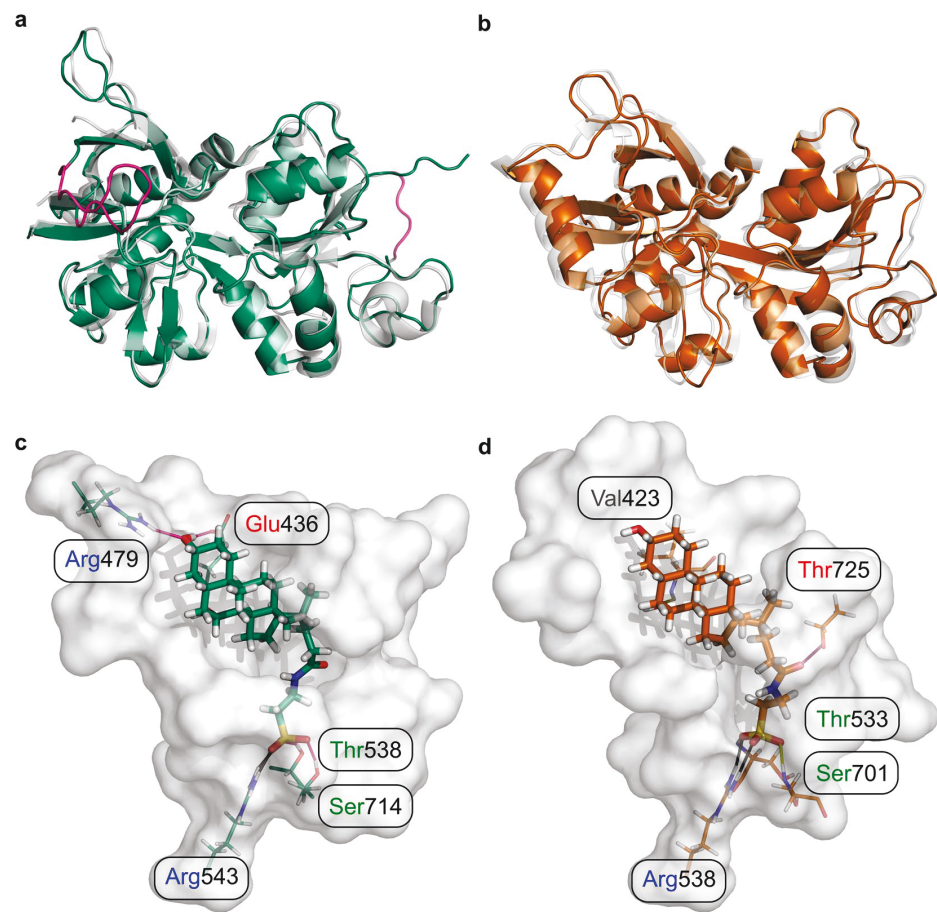


Figure 6. Homology models of GluN2D_{LBD} and GluN3B_{LBD} and docking-derived complex structures with tauro-CDC. **(a)** Homology model of GluN2D_{LBD} (green). The loop consisting of residues 464–476 (magenta, left) was modeled *de novo* due to its absence in the template structure (grey, PDB ID: 3OEK⁴⁵). The engineered Gly-Thr linker in the template structure (magenta, right) was removed in the final model. **(b)** Homology model of GluN2D_{LBD} (orange) superimposed onto the template structure (grey, PDB ID: 2RCA³²). **(c)** Binding mode of tauro-CDC (green sticks) in GluN2D_{LBD} (white surface) predicted by docking. Residues involved in hydrogen bonds and salt bridges are depicted as lines. Positively and negatively charged residues are labeled in blue and red, respectively. Polar residues are labeled in green. **(d)** Binding mode of tauro-CDC (orange sticks) in GluN3B_{LBD} (white surface) predicted by docking. Residues involved in hydrogen bonds and salt bridges are depicted as lines. Positively and negatively charged residues are labeled in blue and red, respectively. Polar residues are labeled in green and hydrophobic residues are labeled in grey.

agonist binding³². Rather, we mutated Asp447, Tyr505, and Ser701 (Fig. 8b) that also engage in hydrogen bonds with the bile salt. Yet, as these interactions are transient and, hence, likely rather weak, a decrease in the extent of tauro-CDC block would indicate an interaction of the bile salt with the mutated residue, while no effect would not exclude such an interaction due to potential enthalpy-entropy compensation effects³³.

Mutation of Asp447 to Ala or Arg had neither an effect on agonist-induced current amplitudes ($n = 2-3$ dose-response curves) nor on the extent of tauro-CDC block in comparison to wild type receptor block ($100 \mu\text{M}$ tauro-CDC, $n = 9-12$ at $10 \mu\text{M}$ Gly, $n = 9-10$ at $5 \mu\text{M}$ Gly). Mutation of Tyr505 to an Ala resulted in GluN3B receptors that could not be activated by its agonist glycine anymore ($n = 9$). Mutation of Ser701 to Ala strongly reduced glycine-induced currents from $835 \pm 85 \text{ nA}$ ($n = 30$) to $106 \pm 12 \text{ nA}$ ($n = 29$). While we observed an increase of 15% in CDC and tauro-CDC block (both $100 \mu\text{M}$) for this mutant ($n = 11-13$), which might indicate an influence on CDC and tauro-CDC binding, it may also well be due to reduced glycine binding.

To conclude, none of the mutants showed a significant influence on tauro-CDC block such that no further information about binding (or non-binding) of tauro-CDC to GluN3B LBD may be inferred.

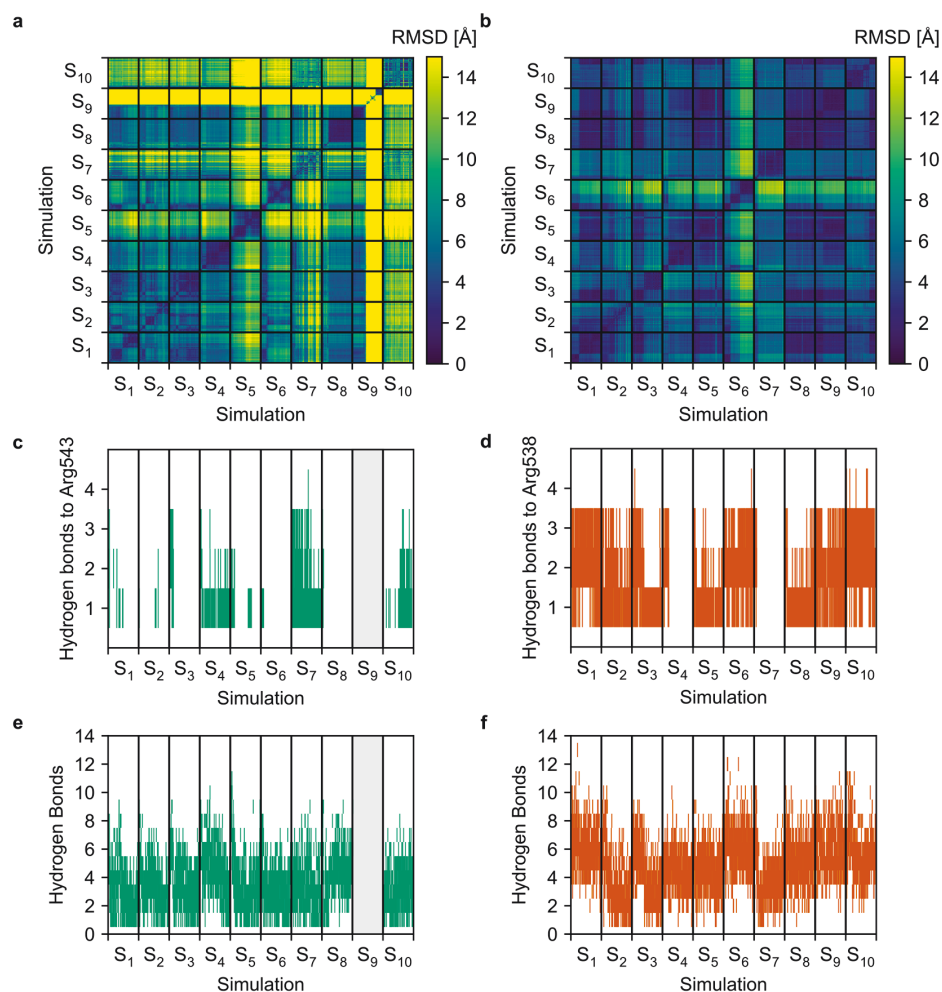


Figure 7. Structural variability and number of hydrogen bonds in the tauro-CDC-GluN2D_{LBD} and tauro-CDC-GluN3B_{LBD} complexes during the MD simulations. **(a)** Two-dimensional root-mean-square deviation (RMSD) of the atomic positions of tauro-CDC after least squares fitting of the C_α atom coordinates of the GluN2D_{LBD} in all 10 MD simulations. **(b)** Two-dimensional root-mean-square deviation (RMSD) of the atomic positions of tauro-CDC after least squares fitting of the C_α atom coordinates of the GluN3B_{LBD} in all 10 MD simulations. **(c)** Number of hydrogen bonds between tauro-CDC and Arg543 in GluN2D_{LBD} over the course of 10 MD simulations. **(d)** Number of hydrogen bonds between tauro-CDC and Arg538 in GluN3B_{LBD} over the course of 10 MD simulations. **(e)** Total number of hydrogen bonds between tauro-CDC and GluN2D_{LBD} over the course of 10 MD simulations. **(f)** Total number of hydrogen bonds between tauro-CDC and GluN3B_{LBD} over the course of 10 MD simulations. In panels c and e, data for MD simulation S₉ is not shown because here tauro-CDC unbinds after 225 ns (c.f. Supplementary Fig. 3b), and the corresponding data was excluded from statistical analysis.

Discussion

Here, we have shown that the bile salt CDC effectively inhibits NMDARs with a preference on GluN2D and GluN3B containing receptors. Neither did conjugation of CDC *per se* nor the type of conjugated amino salt – taurine or glycine – change the overall inhibitory potency of the bile salt or shift its subunit preference. Based on detailed pharmacological analyses and supported by MD experiments, we infer from our data that CDC inhibits NMDARs by different mechanisms that are allosteric inhibition of GluN2D and competitive block of GluN3B containing receptors.

The observed subtype preference of NMDAR inhibition by CDC is in good agreement with the one reported for structurally similar neurosteroids. Thus, pregnenolone sulfate has been shown to significantly potentiate GluN2A and GluN2B containing NMDARs but has much lower efficacy on GluN2C and GluN2D receptors^{6–9}. Also, the apparent insignificance of amino acid conjugates in the interaction of bile salts with receptors is

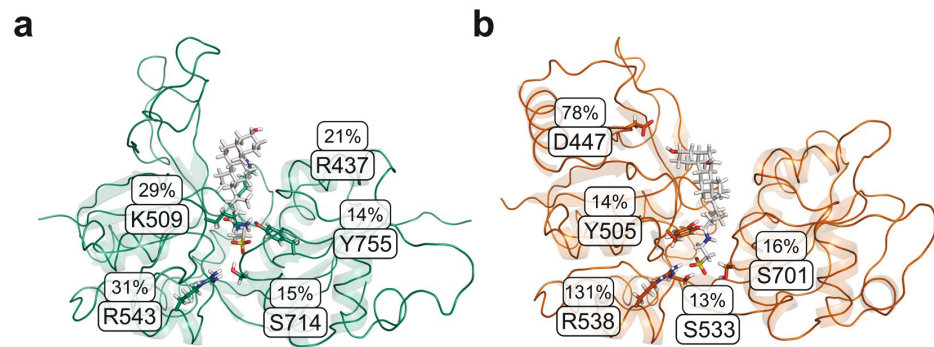


Figure 8. Formation of protein-ligand hydrogen bonds during the MD simulations. Average hydrogen bond occupancy of the five amino acids most frequently involved in hydrogen bonding to tauro-CDC in (a) the 9 tauro-CDC-GluN2D_{LBD} simulations and (b) the 10 tauro-CDC-GluN3B_{LBD} simulations.

paralleled in CDC-mediated activation of the bile acid receptor TGR5, which does not discriminate between conjugation substrates³⁴. Intriguingly, however, our pharmacological experiments indicated different modes of inhibition for different GluN subtypes by CDC: allosteric inhibition of GluN2D and competitive block of GluN3B. In line with this suggestion, MD simulations predicted less structural stability for tauro-CDC binding in the orthosteric pocket of GluN2D_{LBD} than in the one of GluN3B_{LBD}. Effective binding energy computations comparing tauro-CDC/GluN2D_{LBD} with tauro-CDC/GluN3B_{LBD} complexes performed by two different approaches consistently revealed more negative effective binding energies of tauro-CDC binding to GluN3B_{LBD} than to GluN2D_{LBD}, thus supporting stronger binding of tauro-CDC to GluN3B_{LBD}.

Previous studies have shown that inclusion of configurational entropy is crucial for calculating *absolute* binding free energies^{35,36}. In the present study, however, we were rather interested in *relative* binding free energies and, thus, decided to neglect contributions due to changes in the configurational entropy of the ligand or the receptor upon complex formation in order to avoid introducing additional uncertainty in the computations^{35,37,38}. Finally, MD simulations of the GluN1_{LBD}/GluN2D_{LBD} tetramer suggested that the interface between the GluN1 and GluN2D subunit forms a transient pocket, to which tauro-CDC can bind, which may explain the allosteric mechanism of action in this GluN subtype. We selected this interfacial pocket based on a combined approach of geometry-based cavity detection and druggability prediction that has been shown to be able to successfully identify binding-competent pocket conformations from an MD ensemble²⁹. Notably, the detected pocket largely overlaps with the binding site of the negative allosteric modulator MPX-007 in the GluN1/GluN2A interface²⁹. In order to verify binding of tauro-CDC to the orthosteric pocket of GluN3B_{LBD} but not GluN2D_{LBD}, we quantified hydrogen bond formation between tauro-CDC and neighboring residues based on geometric criteria over the course of the MD simulations. The results show that hydrogen bonding occurred less frequently between tauro-CDC and the GluN2D_{LBD} than between tauro-CDC and the GluN3B_{LBD}, in agreement with the above structural and energetic analyses of MD simulations. The tauro-CDC/GluN3B_{LBD} complex is predominantly stabilized by a hydrogen bond to Arg538 and, to a lesser extent, Asp447. Of the residues mutated to alanine in the GluN3B_{LBD}, positions 505 and 701 resulted in GluN3B receptors that could either not or hardly be activated by glycine anymore. For Ser701, it was reported that it forms hydrogen bonds with the glycine carboxylate group³², so it is not surprising that mutation of this site affected channel function. Mutation of Asp447 to an Ala or Arg had neither an effect on agonist-induced current amplitudes nor on the extent of tauro-CDC block, which does not exclude that Asp447 interacts with tauro-CDC, as enthalpy-entropy compensation effects upon mutation may account for the loss of interaction in the mutant³³. In particular, we speculate that the Asp447Ala mutation leads to a loss in binding enthalpy upon disruption of the tauro-CDC-4Asp47 hydrogen bond, but this loss in binding enthalpy could be compensated by both a gain in entropy of the ligand due to an increase in translational and rotational degrees of freedom and a gain in entropy of the receptor due to increased conformational freedom of the residues surrounding the mutation site. The latter effect might especially be relevant due to the stabilizing interaction between Asp447 and Arg474, which would break upon mutation of Asp447 to alanine. Similarly, in the Asp447Arg mutant, the interaction between tauro-CDC and receptor could be preserved if arginine acted as a hydrogen bond donor for the 3-OH group of tauro-CDC, which would in turn act as a hydrogen bond acceptor. Moreover, repulsive forces between Arg447 in the Arg447Arg mutant and Arg474 could lead to a gain in receptor entropy.

The inhibition of GluNs by bile salts required concentrations that are hardly reached in the CNS, where CDC and cholate exists in the nmol/g range^{39,40}. However, NMDARs are also widely expressed in non-neural peripheral tissues. These receptors have many distinct physiological and pathophysiological roles, and there is evidence that peripheral NMDARs may use alternative agonists such as L-HCA and QN⁵. We could show that tauro-CDC also reduced currents elicited by these alternative NMDAR agonists, indicating potential modulation of peripheral NMDAR function by bile salts. Whereas under physiological conditions only low μ M amounts of bile salts are present in serum and urine, in cholestatic disease states like inherited progressive cholestasis, severe forms

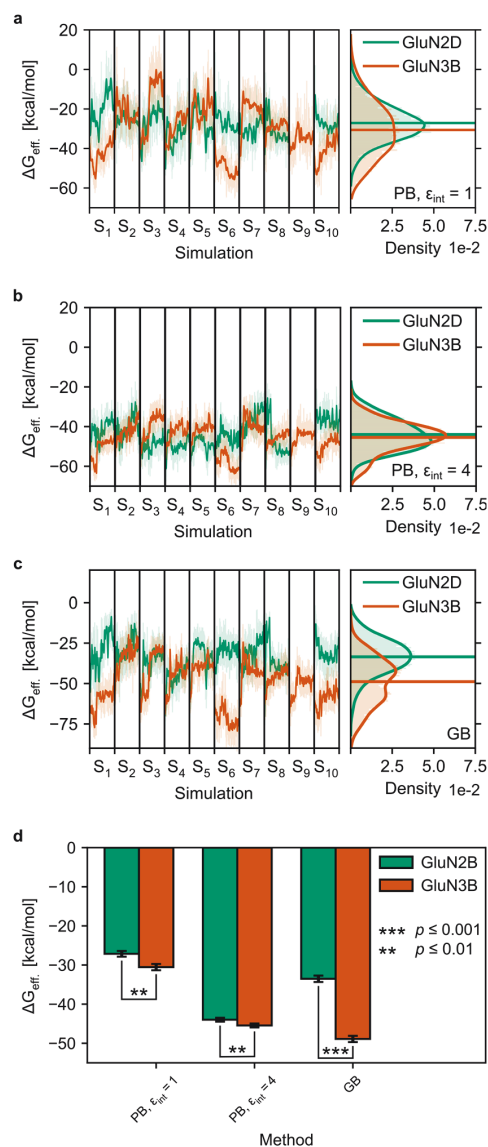


Figure 9. Effective binding energies (ΔG_{eff}) of the tauro-CDC-GluN2D_{LBD} (green) and tauro-CDC-GluN3B_{LBD} (orange) complexes. (a–c) Time courses (left) and probability distributions (right) of ΔG_{eff} calculated using either the MM-PBSA method and an internal dielectric constant of (a) $\epsilon_{\text{int}} = 1$ or (b) $\epsilon_{\text{int}} = 4$, or (c) the MM-GBSA method using the modified OBC model⁶⁷. Thick, opaque lines in the time courses represent the data smoothed with a Savitzky-Golay filter (window length = 51, degree of the smoothing polynomial = 2); thin, transparent lines show the unsmoothed data. Thick, opaque lines in the probability distributions show the distribution calculated with a Gaussian core density estimator (bandwidth calculated using Scott's rule⁷²); thin, transparent lines show the underlying data as a histogram. (d) Time-averaged mean ΔG_{eff} calculated using either the MM-PBSA method and an internal dielectric constant of $\epsilon_{\text{int}} = 1$ or $\epsilon_{\text{int}} = 4$, or the MM-GBSA method using the modified OBC model⁶⁷. In panels a–c, data for MD simulation S₉ is not shown because here, tauro-CDC unbinds after 225 ns (c.f. Supplementary Fig. 3b), and the corresponding data was excluded from statistical analysis.

of intrahepatic cholestasis of pregnancy or acute hepatitis, plasma concentrations may rise to hundred μM and more^{10–12}. As blood cells including platelets express NMDARs, it may well be hypothesized that bile salt mediated receptor inhibition contributes to the reported platelet activation defects and impaired thrombus formation in cholestatic liver disease⁴¹. Despite the observed preference of bile salts for specific receptor isoforms, it remains difficult to predict the exact physiological or pathophysiological consequences of their action on NMDAR

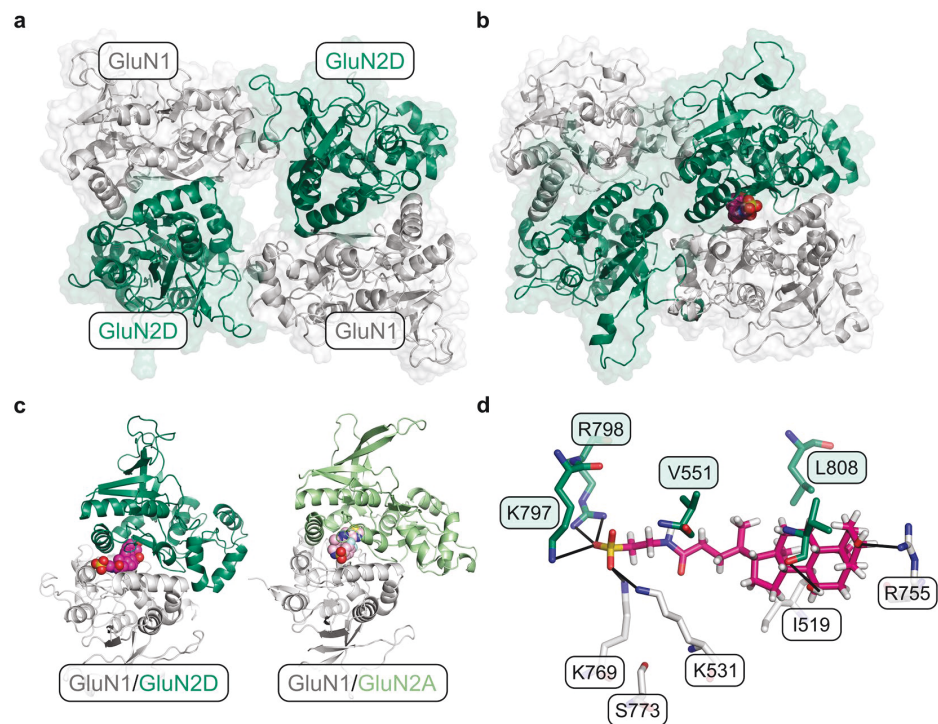


Figure 10. Identification of a potentially allosteric binding site in the GluN1_{LBD}/GluN2_{LBD} interface. The GluN1 subunit is depicted in grey, the GluN2D subunit in dark green. (a) Initial model of the GluN1_{LBD}/GluN2_{LBD} tetramer. (b) Structure of the GluN1_{LBD}/GluN2_{LBD} tetramer in complex with tauro-CDC (magenta spheres). The ligand was docked to the pocket conformation that showed the highest mean local hydrophobic density^{29,30} during 500 ns of accelerated MD. (c) Structural comparison of the tauro-CDC-bound GluN1_{LBD}/GluN2_{LBD} interface (left) and the crystal structure of the MPX-007-bound GluN1_{LBD}/GluN2A_{LBD} interface (right, PDB-ID: 5JTY³¹). The ligands are depicted as magenta and pink spheres, respectively. (d) Binding mode of tauro-CDC (magenta sticks) in the alternative binding pocket. Amino acids that belong to the GluN1 or GluN2D subunit are shown as white or green sticks, respectively. Possible hydrogen bonds and salt bridges are depicted as black lines.

signaling. The net effect of bile salts on NMDARs will eventually depend on the overall expression, stoichiometric assembly, and posttranscriptional and posttranslational modifications of individual GluN isoforms in a given cell type⁴². On a pharmacological perspective, however, the distinct modes of NMDAR inhibition by bile salts, i.e. allosteric versus competitive, may be exploited for developing targeted interference strategies holding isoform specificity.

Materials and Methods

Molecular biology. The cDNAs coding for NMDA and kainate receptors, subcloned into pSGEM for expression in *Xenopus laevis* oocytes, were kindly provided by Michael Hollmann (Ruhr University Bochum). AMPA receptor cDNAs were inserted in pGEM-HE for cRNA synthesis. The following cDNAs were used: GluN1-1a (U08261), GluN2A (NM_012573), GluN2B (NM_012574), GluN2D (U08260), GluN3B (NM_130455), GluK2 (NM_019309), GluA1 (NM_031608) and GluA2 (M38061). These cDNAs are from rat except for GluN3B, which is from mouse. Point mutants were generated by polymerase chain reaction (PCR)-directed mutagenesis with specific mismatch primers (biomers.net). cRNA was synthesized from 1 µg of linearized plasmid DNA using an *in vitro* transcription kit (mMESSAGE mMACHINE[®] T7 Transcription Kit, Thermo Fisher Scientific).

Heterologous expression of iGluRs in *Xenopus laevis* oocytes. Oocytes were obtained from parts of the ovaries surgically removed from *Xenopus laevis* in accordance with German law and approved by the local authorities of the Heinrich Heine University of Düsseldorf. Lumps of around 20 oocytes were incubated with 10 mg/ml collagenase type 2 (Worthington Biochemical Corporation) for 1–1.5 h in Ca²⁺-free Barth's solution (in mM: 88 NaCl, 1 KCl, 2.4 NaHCO₃, 0.82 MgSO₄, 20 HEPES, pH adjusted to 7.5 with NaOH) with slow agitation to remove the follicular cell layer and then washed extensively with Barth's solution containing Ca²⁺ (in mM: 88 NaCl, 1 KCl, 2.4 NaHCO₃, 0.82 MgSO₄, 0.41 CaCl₂, 0.33 Ca(NO₃)₂, 20 HEPES, pH adjusted to 7.5 with NaOH) to inactivate the collagenase. Oocytes were maintained in Barth's solution charged with antibiotic and antimycotic solution (containing penicillin, streptomycin and amphotericin B, Sigma-Aldrich) at 18 °C. Alternatively,

Xenopus laevis oocytes were purchased from EcoCyte Bioscience (Castrop-Rauxel, Germany). Within 24 h after surgery, oocytes of stages V–VI were injected with 3.8 ng GluN1 and 5 ng GluN2A or GluN2B cRNAs, 7.6 ng GluN1 and 10 ng GluN2D cRNA or 7.6 ng GluN1 and 6.4 ng GluN3B-cRNA per oocyte (molecular GluN1 to GluN2 ratio: 1:1; GluN1 to GluN3 ratio: 1:1) using a Micro4 nanoliter injector (World Precision Instruments, Sarasota, FL, USA).

Electrophysiological recordings. Two-electrode voltage clamp recordings of oocyte current responses were performed 3–5 days after cRNA injection at -70 mV holding potential with a Turbo Tec-03X amplifier (npi electronic, Germany) controlled by Pulse software (HEKA, Germany). Electrodes were filled with 3 M KCl and had resistances of 0.5–1.5 M Ω . Oocytes were continuously superfused with calcium-free Ba²⁺-Ringer's solution (in mM: 115 NaCl, 2.5 KCl, 1.8 BaCl₂ and 10 HEPES, pH 7.2) to prevent the activation of endogenous Ca²⁺-gated chloride channels. Agonists and bile salts were applied for 20–40 s by superfusion. Agonist and bile salt blocker potencies were determined by recording current responses induced by the application of increasing agonist concentrations or increasing bile salt concentration to the same oocyte. Agonist-induced currents were normalized to the maximum response under saturating agonist concentration. In case of inhibition by bile salts the relative reduction in agonist-induced current was determined. Data from each oocyte was fitted separately to the Hill equation using Origin 9 software. The resulting EC₅₀ or IC₅₀ values were averaged. Current-voltage relationships were determined by ramping the holding potential from -140 mV to $+50$ mV corrected for background conductivities. Data are reported as mean \pm standard error of the mean (SEM). Statistical significance was determined using the unpaired, two-tailed Student's t-test.

Molecular modeling and molecular dynamics simulations. *Preparation of tauro-CDC/GluN2D_{LBD} and tauro-CDC/GluN3B_{LBD} complex structures.* In order to generate full-length structures of rat GluN2D and mouse GluN3B LBDs (from now on referred to as GluN2D_{LBD} and GluN3B_{LBD}), we first generated a multiple sequence alignment of all rat and mouse GluN1, GluN2, and GluN3 subunit sequences using the MAFFT⁴³ server with default settings applied. The alignment was further refined using the GLProbs software⁴⁴ with two passes of consistency transformation and 100 passes of iterative refinement. The resulting alignment was then used as a template to manually align the sequence portions resolved in the selected template structures for GluN2D_{LBD} (PDB ID: 3OEK⁴⁵) and GluN3B_{LBD} (PDB ID: 2RCA³²) to the target sequences using the BioLuminate package (Schrödinger Release 2018-3: BioLuminate, Schrödinger, LLC, New York, NY, 2018) of the Schrödinger Suite (Small-Molecule Drug Discovery Suite 2018-3, Schrödinger, LLC, New York, NY, 2018). This procedure was necessary to account for insertions and deletions present in the template structures (Supplementary Fig. 1). Structural modeling was also performed with the BioLuminate package using the energy-based model building method (Schrödinger Release 2018-3: Prime, Schrödinger, LLC, New York, NY, 2018). For identical residues in the alignment, side chain rotamers were retained, whereas for non-identical residues, side chain and main chain atoms were energy-minimized using the Prime module (Schrödinger Release 2018-3: Prime, Schrödinger, LLC, New York, NY, 2018) employing the VSGB solvation model⁴⁶ and the OPLS3 force field⁴⁷; all other settings were kept at their default values.

To generate atomic charges for tauro-CDC, we first generated a three-dimensional molecular structure of tauro-CDC in the Maestro interface (Schrödinger Release 2018-3: Maestro, Schrödinger, LLC, New York, NY, 2018) of the Schrödinger Suite. Further preparation of the structure in LigPrep (Schrödinger Release 2018-3: LigPrep, Schrödinger, LLC, New York, NY, 2018) (default settings) returned a single state with a negatively charged sulfonate moiety. Conformers of this structure were then generated in the MacroModel module (Schrödinger Release 2018-3: MacroModel, Schrödinger, LLC, New York, NY, 2018) using the Monte-Carlo Multiple Minimum search algorithm⁴⁸, keeping all other settings at their default values. The conformer with the lowest potential energy that did not display intramolecular hydrogen bonding was subjected to quantum mechanical geometry optimization at the HF/6-31 G(d) level of theory in the GAUSSIAN 09 software (Revision A.02)⁴⁹. A tight convergence criterion for the self-consistent field iteration process was set, and computation of electrostatic potential points with a density of ~ 1.68 pt/au² (6 pt/Å²) was invoked. The geometry-optimized structure was visually inspected for the absence of intramolecular hydrogen bonds and subjected to the standard RESP procedure^{50,51} as implemented in the Antechamber set of programs to obtain the atomic point charges.

Binding modes for tauro-CDC in GluN2D_{LBD} and GluN3B_{LBD} were generated using the standard Induced Fit protocol implemented in the Schrödinger Suite²⁵. The docking grid was centered on the coordinates of the bound ligand in the GluN2D and GluN3B template structure, respectively. In the initial Glide docking stage, the van der Waals radii of the receptor and ligand atoms were reduced to half their initial value, and a maximum of 20 poses were allowed to be carried over to the Prime (Schrödinger Release 2018-3: Prime, Schrödinger, LLC, New York, NY, 2018) refinement stage. For each generated pose, the receptor residues within 5 Å were energy-minimized and their side chain rotamers optimized. During the final Glide (Schrödinger Release 2018-3: Glide, Schrödinger, LLC, New York, NY, 2018)^{52,53} docking stage, tauro-CDC was redocked into all receptor conformers using XP precision⁵⁴. The tauro-CDC/GluN2D_{LBD} and tauro-CDC/GluN3B_{LBD} complexes with the lowest IFD score²⁵ were considered for all subsequent calculations.

Molecular dynamics simulations of tauro-CDC/GluN2D_{LBD} and tauro-CDC/GluN3B_{LBD} complexes. All molecular dynamics (MD) simulations were performed using the mixed single precision/precision GPU (CUDA) version of PMEMD⁵⁵ in the Amber14 suite of molecular simulation programs⁵⁶. Hydrogen mass repartitioning⁵⁷ was employed to enable a time step of 4 fs for integration. The Langevin thermostat⁵⁸ with a collision frequency of 0.01 ps⁻¹ and a target temperature of $T = 300$ K was used for temperature control. Covalent bonds involving hydrogen atoms were constrained using the SHAKE algorithm⁵⁹. Long-range electrostatic interactions were

estimated using the Particle Mesh Ewald method⁶⁰, and a cutoff of 8 Å was used for short-range electrostatics and van der Waals forces.

All following steps were performed 10 times for each of the two complexes (tauro-CDC/GluN2D_{LBD} and tauro-CDC/GluN3B_{LBD}) so as to obtain 10 independent simulations per complex. The initial structures were minimized using a three-step procedure. First, the coordinates of all solute molecules were restrained by a harmonic potential with a force constant of 2.0 kcal mol⁻¹ Å⁻² while 2,000 steps of steepest descent minimization followed by 3,000 steps of conjugate gradient minimization were carried out. This step was repeated with the restraints switched from the solute to the solvent molecules. During the following final minimization step, the restraints were removed, and 3,000 steps of steepest descent minimization followed by 7,000 steps of conjugate gradient minimization were performed. 20 ps of NVT-MD (the solute was restrained with a force constant of 2.0 kcal mol⁻¹ Å⁻²) were carried out while heating the system from 0 K to 300 K, followed by additional 5 ps of NVT-MD at 300 K. Density adaptation of the system was achieved by 75 ps of NPT-MD (solute restrained, force constant: 2.0 kcal mol⁻¹ Å⁻²). An additional 1.7 ns of restrained NPT-MD were carried out before switching to the NVT ensemble. Here, 3.2 ns of restrained MD were performed prior to the start of the production phase, with harmonic restraints (force constant: 2 kcal mol⁻¹ Å⁻²) applied to only those C_α atoms that are closest to the center of mass of a secondary structure element (α-helix or β-sheet). The subsequent production phase consisted of 500 ns of NVT-MD (restraints as in the final NVT step of the equilibration phase), resulting in an aggregate simulation time of 5 μs per complex (10 μs in total). Coordinates for analysis and post-processing were saved every 20 ps. Post-processing and analysis of the MD trajectories was performed in CPPTRAJ⁶¹ as implemented in AmberTools15. Trajectories were visually inspected with VMD⁶². The two-tailed Student's *t*-test was used to determine statistically significant differences in time-averaged quantities between GluN2D and GluN3B.

Calculation of effective binding energies. Effective binding energies (ΔG_{eff}) of tauro-CDC to GluN2D_{LBD} and GluN3B_{LBD} were calculated using the MM/PBSA and MM/GBSA approaches as implemented in the MMPBSA.py module⁶³ in AmberTools15⁶⁶. All calculations were performed on 500 snapshots per trajectory extracted at regular intervals of 1 ns, corresponding to a total of 5,000 snapshots per approach for both the tauro-CDC/GluN2D_{LBD} complex and the tauro-CDC/GluN3B_{LBD} complex. The MM/PBSA calculations were performed with values of 1 and 4 for the internal dielectric constant (ϵ_{int}), to test for robustness of the relative ΔG_{eff} with respect to this parameter²⁸. The polar part of the contribution to the solvation free energy was calculated by the linear Poisson-Boltzmann equation using an ionic strength of 100 mM and an external dielectric constant (ϵ_{ext}) of 80. Solutions for the linear PB equation were computed with 1,000 iterations on a cubic grid with 0.5 Å spacing between grid points. The nonpolar part of the contribution to the solvation free energy was considered proportional to the solvent accessible surface area (SASA) and computed with the equation $\Delta G_{\text{nonpol,solv}} = \gamma(\text{SASA}) + \beta$, where γ and β are parameterized constants⁶⁴. The SASA was calculated with a solvent probe radius of 1.4 Å, Tan & Luo radii⁶⁵ for the protein, and mbondi radii⁶⁶ for the ligands. The MM/GBSA calculations were performed using the modified OBC model (GB^{OBC}II)⁶⁷, an ionic strength of 100 mM and default values for γ (0.005 kcal mol⁻¹ Å⁻²) and β (0.0 kcal mol⁻¹).

Results are expressed as $\Delta G_{\text{eff}} \pm$ standard error of the mean (SEM), where ΔG_{eff} is the average effective binding energy over all snapshots and the overall SEM is estimated as the SEM of the single trajectories propagated over all trajectories. The two-tailed Student's *t*-test was used to determine statistically significant differences in ΔG_{eff} between GluN2D and GluN3B.

Preparation of GluN1_{LBD}/GluN2D_{LBD} tetramer structures. The structure of the rat GluN1 LBD (GluN1_{LBD}) was prepared analogously to the structure of GluN2D_{LBD} (see above for further details), using the crystal structure of the rat GluN1 LBD (PDB ID: 1PB7⁶⁸) as a template to which only missing residues and loops were added during the modeling step. To construct the LBD tetramer from the individual LBD structures, the coordinates of the respective LBD monomers were superimposed onto their corresponding monomer units in the structure of the GluN1/GluN2D NMDA receptor (PDB ID: 5FXH⁶⁹) (Fig. 10a).

Accelerated molecular dynamics simulation of the GluN1_{LBD}/GluN2D_{LBD} tetramer. A single, 500 ns long accelerated MD (aMD) simulation⁷⁰ of the GluN1_{LBD}/GluN2D_{LBD} tetramer was prepared, heated, and equilibrated analogously to the classical MD simulations of the tauro-CDC/GluN2D_{LBD} and tauro-CDC/GluN3B_{LBD} complex (see above). During the production phase, bias parameters of $\alpha_{\text{D}} = 781.0$ kcal mol⁻¹ and $\alpha_{\text{P}} = 50,035.0$ kcal mol⁻¹, a total dihedral energy threshold of $E_{\text{threshD}} = 18,950.0$ kcal mol⁻¹, and a total potential energy threshold of $E_{\text{threshP}} = -341,718.0$ kcal mol⁻¹ were set on the basis of empirical formulas that were previously used to estimate these parameters^{70,71}.

Preparation of tauro-CDC//GluN1_{LBD}/GluN2D_{LBD} complex structures. To identify a possible allosteric binding site in the GluN1_{LBD}/GluN2D_{LBD} tetramer, the MDpocket software²⁹ was applied on all 25,000 snapshots of the 500 ns long aMD trajectory (see above). A binding pocket at the GluN1/GluN2D_{LBD} interface was identified into which tauro-CDC was docked, using the snapshot in which the selected pocket displayed the highest mean local hydrophobic density^{29,30} (Supplementary Fig. 2).

Molecular dynamics simulations of the tauro-CDC//GluN1_{LBD}/GluN2D_{LBD} complex. A single 500 ns long conventional MD simulation of the tauro-CDC//GluN1_{LBD}/GluN2D_{LBD} tetramer complex was generated and analyzed analogously to the classical MD simulations of the tauro-CDC/GluN2D_{LBD} and tauro-CDC/GluN3B_{LBD} complex (see above).

References

- Mayer, M. L., Westbrook, G. L. & Guthrie, P. B. Voltage-dependent block by Mg²⁺ of NMDA responses in spinal cord neurones. *Nature* **309**, 261–263, <https://doi.org/10.1038/309261a0> (1984).
- MacDermott, A. B., Mayer, M. L., Westbrook, G. L., Smith, S. J. & Barker, J. L. NMDA-receptor activation increases cytoplasmic calcium concentration in cultured spinal cord neurones. *Nature* **321**, 519–522, <https://doi.org/10.1038/321519a0> (1986).
- Kleckner, N. W. & Dingledine, R. Requirement for glycine in activation of NMDA-receptors expressed in *Xenopus* oocytes. *Science* **241**, 835–837 (1988).
- Chatterton, J. E. *et al.* Excitatory glycine receptors containing the NR3 family of NMDA receptor subunits. *Nature* **415**, 793–798, <https://doi.org/10.1038/nature715> (2002).
- Hogan-Cann, A. D. & Anderson, C. M. Physiological Roles of Non-Neuronal NMDA Receptors. *Trends in pharmacological sciences* **37**, 750–767, <https://doi.org/10.1016/j.tips.2016.05.012> (2016).
- Traynelis, S. F. *et al.* Glutamate receptor ion channels: structure, regulation, and function. *Pharmacological Reviews* **62**, 405–496, <https://doi.org/10.1124/pr.109.002451> (2010).
- Malayev, A., Gibbs, T. T. & Farb, D. H. Inhibition of the NMDA response by pregnenolone sulphate reveals subtype selective modulation of NMDA receptors by sulphated steroids. *British journal of pharmacology* **135**, 901–909, <https://doi.org/10.1038/sj.bjp.0704543> (2002).
- Jang, M.-K., Mierke, D. F., Russek, S. J. & Farb, D. H. A steroid modulatory domain on NR2B controls N-methyl-D-aspartate receptor proton sensitivity. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 8198–8203, <https://doi.org/10.1073/pnas.0401838101> (2004).
- Horák, M., Vlček, K., Chodounská, H. & Vyklícký, L. Jr. Subtype-dependence of N-methyl-D-aspartate receptor modulation by pregnenolone sulfate. *Neuroscience* **137**, 93–102, <https://doi.org/10.1016/j.neuroscience.2005.08.058> (2006).
- Keitel, V. *et al.* Expression and localization of hepatobiliary transport proteins in progressive familial intrahepatic cholestasis. *Hepatology* **41**, 1160–1172, <https://doi.org/10.1002/hep.20682> (2005).
- Glantz, A., Marschall, H.-U. & Mattsson, L.-Å. Intrahepatic cholestasis of pregnancy: Relationships between bile acid levels and fetal complication rates. *Hepatology* **40**, 467–474, <https://doi.org/10.1002/hep.20336> (2004).
- Makino, I., Hashimoto, H., Shinozaki, K., Yoshino, K. & Nakagawa, S. Sulfated and nonsulfated bile acids in urine, serum, and bile of patients with hepatobiliary diseases. *Gastroenterology* **68**, 545–553 (1975).
- Schubring, S. R., Fleischer, W., Lin, J.-S., Haas, H. L. & Sergeeva, O. A. The bile steroid chenodeoxycholate is a potent antagonist at NMDA and GABA(A) receptors. *Neuroscience Letters* **506**, 322–326, <https://doi.org/10.1016/j.neulet.2011.11.036> (2012).
- Pilo de la Fuente, B., Ruiz, I., Lopez-de-Munain, A. & Jiménez-Escrig, A. Cerebrotendinous xanthomatosis: neuropathological findings. *J Neurol* **255**, 839–842, <https://doi.org/10.1007/s00415-008-0729-6> (2008).
- Gómez-Vicente, V. *et al.* Neuroprotective Effect of Tauroursodeoxycholic Acid on N-Methyl-D-Aspartate-Induced Retinal Ganglion Cell Degeneration. *PLoS one* **10**, e0137826, <https://doi.org/10.1371/journal.pone.0137826> (2015).
- Hasanein, P. *et al.* Modulation of cholestasis-induced antinociception in rats by two NMDA receptor antagonists: MK-801 and magnesium sulfate. *European journal of pharmacology* **554**, 123–127, <https://doi.org/10.1016/j.ejphar.2006.10.026> (2007).
- Dehpour, A. R. *et al.* Effect of NMDA receptor antagonist on naloxone-precipitated withdrawal signs in cholestatic mice. *Hum Psychopharmacol* **15**, 213–218, [https://doi.org/10.1002/\(SICI\)1099-1077\(200004\)15:3<213::AID-HUP159>3.0.CO;2-Z](https://doi.org/10.1002/(SICI)1099-1077(200004)15:3<213::AID-HUP159>3.0.CO;2-Z) (2000).
- Hosseini, N., Nasehi, M., Radahmadi, M. & Zarrindast, M.-R. Effects of CA1 glutamatergic systems upon memory impairments in cholestatic rats. *Behavioural brain research* **256**, 636–645, <https://doi.org/10.1016/j.bbr.2013.08.018> (2013).
- Hollmann, M. *et al.* Zinc potentiates agonist-induced currents at certain splice variants of the NMDA receptor. *Neuron* **10**, 943–954, [https://doi.org/10.1016/0896-6273\(93\)90209-A](https://doi.org/10.1016/0896-6273(93)90209-A) (1993).
- Erreger, K. *et al.* Subunit-specific agonist activity at NR2A-, NR2B-, NR2C-, and NR2D-containing N-methyl-D-aspartate glutamate receptors. *Molecular Pharmacology* **72**, 907–920, <https://doi.org/10.1124/mol.107.037333> (2007).
- Chen, P. E. *et al.* Modulation of glycine potency in rat recombinant NMDA receptors containing chimeric NR2A/2D subunits expressed in *Xenopus laevis* oocytes. *Journal of Physiology* **586**, 227–245, <https://doi.org/10.1113/jphysiol.2007.143172> (2008).
- Awobuluyi, M. *et al.* Subunit-specific roles of glycine-binding domains in activation of NR1/NR3 N-methyl-D-aspartate receptors. *Molecular Pharmacology* **71**, 112–122, <https://doi.org/10.1124/mol.106.030700> (2007).
- Schmidt, A. *et al.* The bile acid-sensitive ion channel (BASIC) is activated by alterations of its membrane environment. *PLoS one* **9**, e111549, <https://doi.org/10.1371/journal.pone.0111549> (2014).
- Curras, M. C. & Dingledine, R. Selectivity of amino acid transmitters acting at N-methyl-D-aspartate and amino-3-hydroxy-5-methyl-4-isoxazolepropionate receptors. *Molecular Pharmacology* **41**, 520–526 (1992).
- Sherman, W., Day, T., Jacobson, M. P., Friesner, R. A. & Farid, R. Novel procedure for modeling ligand/receptor induced fit effects. *Journal of Medicinal Chemistry* **49**, 534–553, <https://doi.org/10.1021/jm050540c> (2006).
- Kuzmanic, A. & Zagrovic, B. Determination of ensemble-average pairwise root mean-square deviation from experimental B-factors. *Biophysical Journal* **98**, 861–871, <https://doi.org/10.1016/j.bpj.2009.11.011> (2010).
- Plattner, N. & Noé, F. Protein conformational plasticity and complex ligand-binding kinetics explored by atomistic simulations and Markov models. *Nat. Commun.* **6**, 7653, <https://doi.org/10.1038/ncomms8653> (2015).
- Homeyer, N., Stoll, F., Hillisch, A. & Gohlke, H. Binding Free Energy Calculations for Lead Optimization: Assessment of Their Accuracy in an Industrial Drug Design Context. *Journal of Chemical Theory and Computation* **10**, 3331–3344, <https://doi.org/10.1021/ct500029e> (2014).
- Schmidtke, P., Bidon-Chanal, A., Luque, F. J. & Barril, X. MDpocket: open-source cavity detection and characterization on molecular dynamics trajectories. *Bioinformatics* **27**, 3276–3285, <https://doi.org/10.1093/bioinformatics/btr550> (2011).
- Schmidtke, P. & Barril, X. Understanding and predicting druggability. A high-throughput method for detection of drug binding sites. *Journal of Medicinal Chemistry* **53**, 5858–5867, <https://doi.org/10.1021/jm100574m> (2010).
- Yi, F. *et al.* Structural Basis for Negative Allosteric Modulation of GluN2A-Containing NMDA Receptors. *Neuron* **91**, 1316–1329, <https://doi.org/10.1016/j.neuron.2016.08.014> (2016).
- Yao, Y., Harrison, C. B., Freddolino, P. L., Schulten, K. & Mayer, M. L. Molecular mechanism of ligand recognition by NR3 subtype glutamate receptors. *EMBO Journal* **27**, 2158–2170, <https://doi.org/10.1038/emboj.2008.140> (2008).
- Richieri, G. V., Low, P. J., Ogata, R. T. & Kleinfeld, A. M. Mutants of rat intestinal fatty acid-binding protein illustrate the critical role played by enthalpy-entropy compensation in ligand binding. *Journal of Biological Chemistry* **272**, 16737–16740, <https://doi.org/10.1074/jbc.272.27.16737> (1997).
- Gertzen, C. G. W. *et al.* Mutational mapping of the transmembrane binding site of the G-protein coupled receptor TGR5 and binding mode prediction of TGR5 agonists. *Eur J Med Chem* **104**, 57–72, <https://doi.org/10.1016/j.ejmech.2015.09.024> (2015).
- Hou, T., Wang, J., Li, Y. & Wang, W. Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *Journal of Chemical Information and Modeling* **51**, 69–82, <https://doi.org/10.1021/ci100275a> (2011).
- Genheden, S. & Ryde, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin Drug Discov* **10**, 449–461, <https://doi.org/10.1517/17460441.2015.1032936> (2015).
- Gohlke, H. & Case, D. A. Converging free energy estimates: MM-PB(GB)SA studies on the protein-protein complex Ras-Raf. *Journal of Computational Chemistry* **25**, 238–250, <https://doi.org/10.1002/jcc.10379> (2004).

38. Weis, A., Katebzadeh, K., Söderhjelm, P., Nilsson, I. & Ryde, U. Ligand affinities predicted with the MM/PBSA method: dependence on the simulation method and the force field. *Journal of Medicinal Chemistry* **49**, 6596–6606, <https://doi.org/10.1021/jm0608210> (2006).
39. Mano, N. *et al.* Presence of protein-bound unconjugated bile acids in the cytoplasmic fraction of rat brain. *Journal of Lipid Research* **45**, 295–300, <https://doi.org/10.1194/jlr.M300369-JLR200> (2004).
40. Tripodi, V., Contini, M., Fernández, M. A. & Lemberg, A. Bile acids content in brain of common duct ligated rats. *Ann Hepatol* **11**, 930–934 (2012).
41. Gowert, N. S. *et al.* Defective Platelet Activation and Bleeding Complications upon Cholestasis in Mice. *Cell. Physiol. Biochem* **41**, 2133–2149, <https://doi.org/10.1159/000475566> (2017).
42. Paoletti, P., Bellone, C. & Zhou, Q. NMDA receptor subunit diversity: impact on receptor properties, synaptic plasticity and disease. *Nature Reviews Neuroscience* **14**, 383–400, <https://doi.org/10.1038/nrn3504> (2013).
43. Katoh, K., Rozewicki, J. & Yamada, K. D. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform*, <https://doi.org/10.1093/bib/bbx108> (2017).
44. Ye, Y. *et al.* GLProbs: Aligning Multiple Sequences Adaptively. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* **12**, 67–78, <https://doi.org/10.1109/TCBB.2014.2316820> (2015).
45. Vance, K. M., Simorowski, N., Traynelis, S. F. & Furukawa, H. Ligand-specific deactivation time course of GluN1/GluN2D NMDA receptors. *Nat. Commun.* **2**, 294, <https://doi.org/10.1038/ncomms1295> (2011).
46. Li, J. *et al.* The VSGB 2.0 model: a next generation energy model for high resolution protein structure modeling. *Proteins* **79**, 2794–2812, <https://doi.org/10.1002/prot.23106> (2011).
47. Harder, E. *et al.* OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. *Journal of Chemical Theory and Computation* **12**, 281–296, <https://doi.org/10.1021/acs.jctc.5b00864> (2016).
48. Chang, G., Guida, W. C. & Still, W. C. An Internal Coordinate Monte-Carlo Method for Searching Conformational Space. *J Am Chem Soc* **111**, 4379–4386, <https://doi.org/10.1021/ja00194a035> (1989).
49. Gaussian 09, Revision A.02 (Wallingford CT, 2009).
50. Bayly, C. I., Cieplak, P., Cornell, W. D. & Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J Phys Chem* **97**, 10269–10280, <https://doi.org/10.1021/j100142a004> (1993).
51. Wang, J., Cieplak, P. & Kollman, P. A. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *Journal of Computational Chemistry* **21**, 1049–1074, [https://doi.org/10.1002/1096-987X\(200009\)21:12<1049::AID-JCC3>3.0.CO;2-F](https://doi.org/10.1002/1096-987X(200009)21:12<1049::AID-JCC3>3.0.CO;2-F) (2000).
52. Friesner, R. A. *et al.* Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *Journal of Medicinal Chemistry* **47**, 1739–1749, <https://doi.org/10.1021/jm0306430> (2004).
53. Halgren, T. A. *et al.* Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *Journal of Medicinal Chemistry* **47**, 1750–1759, <https://doi.org/10.1021/jm030644s> (2004).
54. Friesner, R. A. *et al.* Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *Journal of Medicinal Chemistry* **49**, 6177–6196, <https://doi.org/10.1021/jm051256o> (2006).
55. Le Grand, S., Götz, A. W. & Walker, R. C. SPFP: Speed without compromise—A mixed precision model for GPU accelerated molecular dynamics simulations. *Comput Phys Commun* **184**, 374–380, <https://doi.org/10.1016/j.cpc.2012.09.022> (2013).
56. AMBER 14 (University of California, San Francisco, 2014).
57. Hopkins, C. W., Le Grand, S., Walker, R. C. & Roitberg, A. Long-Time-Step Molecular Dynamics through Hydrogen Mass Repartitioning. *Journal of Chemical Theory and Computation* **11**, 1864–1874, <https://doi.org/10.1021/ct5010406> (2015).
58. Pastor, R. W., Brooks, B. R. & Szabo, A. An Analysis of the Accuracy of Langevin and Molecular-Dynamics Algorithms. *Mol Phys* **65**, 1409–1419, <https://doi.org/10.1080/0026878800101881> (1988).
59. Ryckaert, J.-P., Ciccolini, G. & Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics* **23**, 327–341, [https://doi.org/10.1016/0021-9991\(77\)90098-5](https://doi.org/10.1016/0021-9991(77)90098-5) (1977).
60. Darden, T. A., York, D. M. & Pedersen, L. Particle mesh Ewald - an N-Log(N) method for Ewald sums in large systems. *J Chem Phys* **98**, 10089–10092, <https://doi.org/10.1063/1.464397> (1993).
61. Roe, D. R. & Cheatham, T. E. III. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *Journal of Chemical Theory and Computation* **9**, 3084–3095, <https://doi.org/10.1021/ct400341p> (2013).
62. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *Journal of molecular graphics* **14**(33–38), 27–38, [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5) (1996).
63. Miller, B. R. III *et al.* MMPBSA.py: An Efficient Program for End-State Free Energy Calculations. *Journal of Chemical Theory and Computation* **8**, 3314–3321, <https://doi.org/10.1021/ct300418h> (2012).
64. Sitkoff, D., Sharp, K. A. & Honig, B. Accurate Calculation of Hydration Free-Energies Using Macroscopic Solvent Models. *J Phys Chem* **98**, 1978–1988, <https://doi.org/10.1021/j100058a043> (1994).
65. Tan, C., Yang, L. & Luo, R. How well does Poisson-Boltzmann implicit solvent agree with explicit solvent? A quantitative analysis. *The Journal of Physical Chemistry B* **110**, 18680–18687, <https://doi.org/10.1021/jp063479b> (2006).
66. Tsui, V. & Case, D. A. Theory and applications of the generalized Born solvation model in macromolecular simulations. *Biopolymers* **56**, 275–291, [https://doi.org/10.1002/1097-0282\(2000\)56:4<275::AID-BIP10024>3.0.CO;2-E](https://doi.org/10.1002/1097-0282(2000)56:4<275::AID-BIP10024>3.0.CO;2-E) (2000).
67. Onufriev, A., Bashford, D. & Case, D. A. Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins* **55**, 383–394, <https://doi.org/10.1002/prot.20033> (2004).
68. Furukawa, H. & Gouaux, E. Mechanisms of activation, inhibition and specificity: crystal structures of the NMDA receptor NR1 ligand-binding core. *EMBO Journal* **22**, 2873–2885, <https://doi.org/10.1093/emboj/cdg303> (2003).
69. Tajima, N. *et al.* Activation of NMDA receptors and the mechanism of inhibition by ifenprodil. *Nature* **534**, 63–68, <https://doi.org/10.1038/nature17679> (2016).
70. Hamelberg, D., Mongan, J. & McCammon, J. A. Accelerated molecular dynamics: a promising and efficient simulation method for biomolecules. *J Chem Phys* **120**, 11919–11929, <https://doi.org/10.1063/1.1755656> (2004).
71. Pierce, L. C. T., Salomon-Ferrer, R., de Oliveira, C. A. F., McCammon, J. A. & Walker, R. C. Routine Access to Millisecond Time Scale Events with Accelerated Molecular Dynamics. *Journal of Chemical Theory and Computation* **8**, 2997–3002, <https://doi.org/10.1021/ct300284c> (2012).
72. Scott, D. W. *Multivariate density estimation: theory, practice, and visualization*. 2. edn, (John Wiley & Sons, Inc., 2015).

Acknowledgements

The study was supported by the Deutsche Forschungsgemeinschaft SFB974, TP B05 to N.K. and financial support for GPU server to H.G. We thank Prof. Michael Hollmann (Ruhr-Universität Bochum, Germany) for iGluR cDNAs. H.G. is grateful to the Jülich Supercomputing Centre at the Forschungszentrum Jülich for computing time on the supercomputer JURECA and JUWELS (NIC project ID: HKF7 and HDD17) and to the “Zentrum für Informations- und Medientechnologie” (ZIM) at the Heinrich Heine University Düsseldorf for computational support.

www.nature.com/scientificreports/

Author Contributions

N.K. and H.G. designed the research, A.K. conceptualized and performed TEVC experiments, M.B. performed computations, all authors analyzed the data and contributed to manuscript writing. N.K. and H.G. acquired funding.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-019-46496-y>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019

Supplementary Information

Isoform-specific Inhibition of N-methyl-D-aspartate Receptors by Bile Salts

Angela Koch^{a,§}, Michele Bonus^{b,§}, Holger Gohlke^{b,c}, Nikolaj Klöcker^{a,*}

^a *Institute of Neural and Sensory Physiology, Medical Faculty, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany*

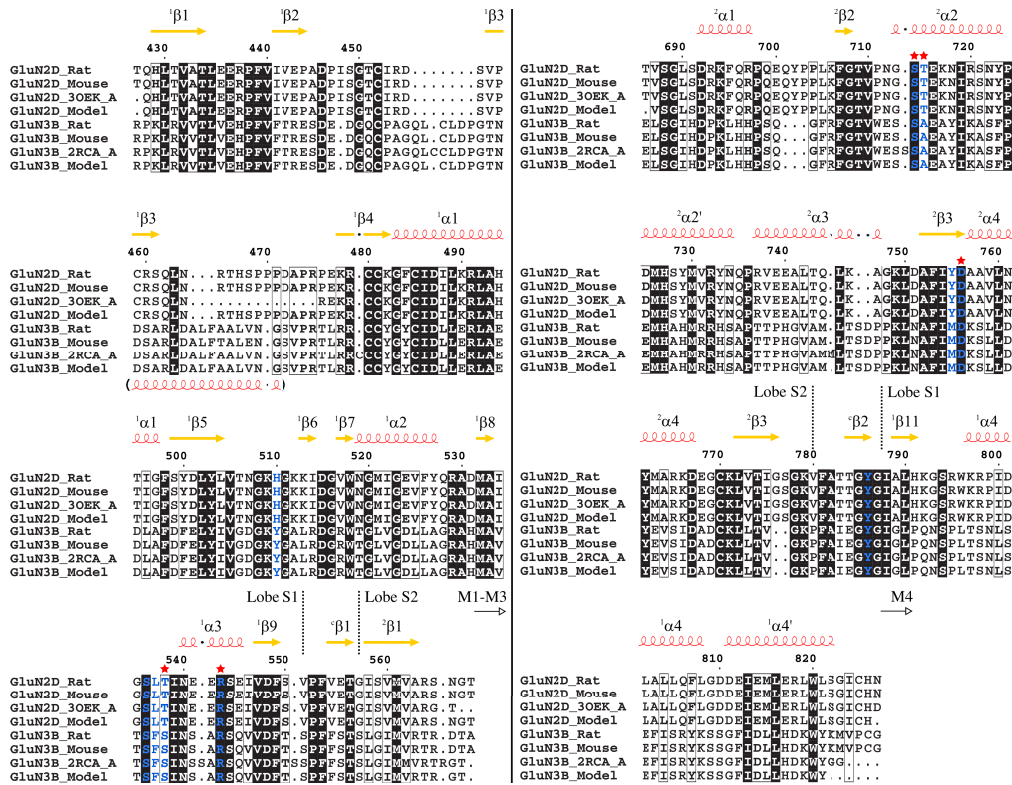
^b *Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany*

^c *John von Neumann Institute for Computing (NIC), Jülich Supercomputing Centre (JSC) & Institute for Complex Systems - Structural Biochemistry (ICS 6), Forschungszentrum Jülich GmbH, 52425 Jülich, Germany*

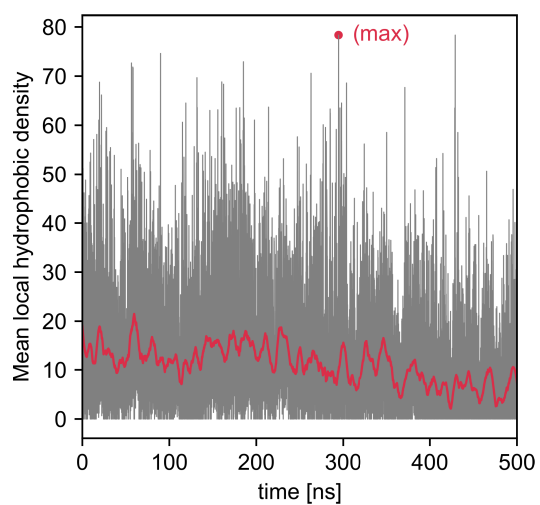
§ these authors contributed equally

**Correspondence to: Nikolaj Klöcker, Institute of Neural and Sensory Physiology, Medical Faculty, Heinrich Heine University, Universitätsstraße 1, 40225 Düsseldorf, Germany, Tel.: 0049-211-81-12687; Fax: 0049-211-81-14231; E-mail: nikolaj.kloecker@uni-duesseldorf.de*

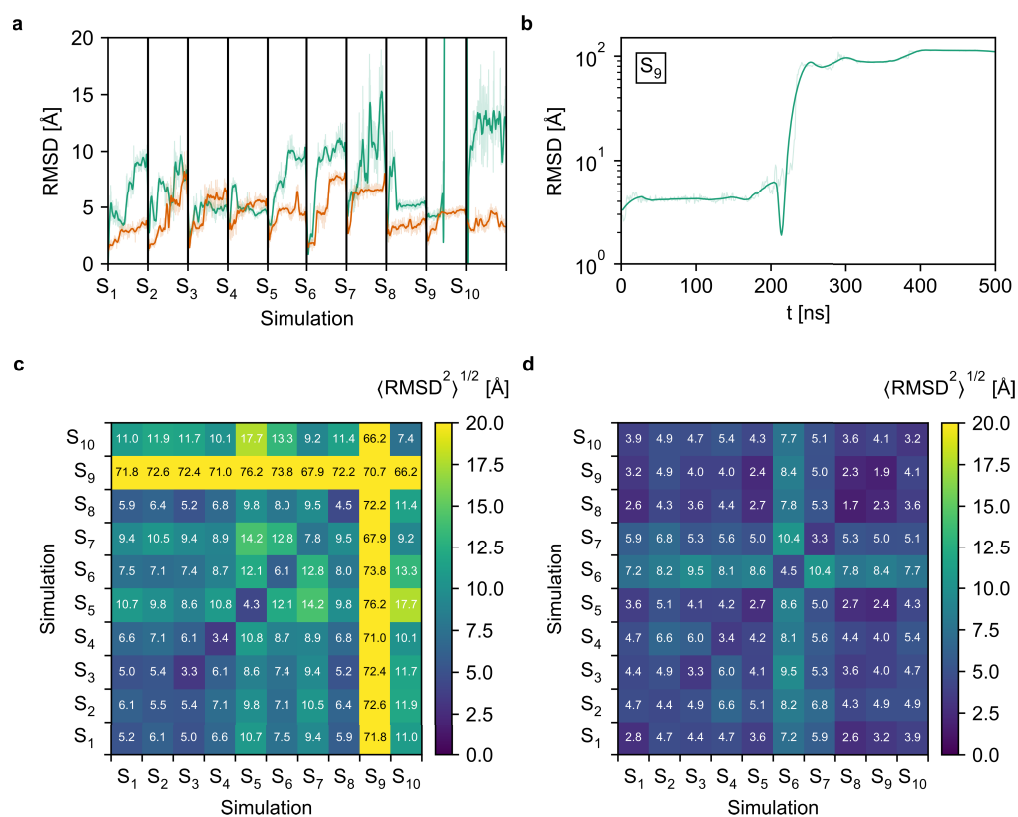
Supplementary Figures



Supplementary Figure S1. Sequence alignment of the rat and mouse GluN2D_{LBD} and GluN3B_{LBD} sequences and the sequences of the respective template structures and the final models. α helices and β sheets are shown in red and yellow, respectively, and numbered sequentially in each of the two structural lobes S1 and S2. The left part of the figure shows the portion of the LBD sequence that precedes transmembrane segments M1-M3, the right part the portion of the LBD sequence that follows these segments.



Supplementary Figure S2. Time course of the mean local hydrophobic density (MLHD)^{29,30} for the identified pocket in the GluN1_{LBD}/GluN2_{LBD} interface. The snapshot in which the pocket displayed the maximum MLHD (denoted as "max") was selected for docking of tauro-CDC.



Supplementary Figure S3. (a) RMSD of tauro-CDC after least squares fitting of the C_α atom coordinates of the GluN2DLDB (green) or GluN3BLDB (orange). (b) RMSD of tauro-CDC during MD simulation 9 (S₉) after least squares fitting of the C_α atom coordinates of the GluN2DLDB. (c) Quadratic means of the pairwise RMSDs calculated for pairs of individual simulations of the tauro-CDC/GluN2DLDB complex and (d) the tauro-CDC/GluN3BLDB complex. Coordinate fitting was performed as described in panel a and Fig. 7a, b, respectively.

Supplementary Tables

Table S1. Mean (\pm SEM) reduction of agonist-induced NMDA receptor currents

Bile salt	Reduction of $I_{\text{Glu/Gly}}$ by bile salt [%] (n)			
	GluN2A	GluN2B	GluN2D	GluN3B
Cholate	23 \pm 4 (10)	7 \pm 2 (11)	12 \pm 1 (11)	20 \pm 3 (6)
Glycocholate	8 \pm 3 (9)	3 \pm 1 (8)	10 \pm 1 (8)	14 \pm 1 (6)
Taurocholate	7 \pm 3 (9)	7 \pm 2 (7)	9 \pm 2 (8)	16 \pm 1 (6)
CDC	7 \pm 2 (10)	13 \pm 1 (13)	37 \pm 2 (14)	58 \pm 5 (12)
Glyco-CDC	11 \pm 3 (9)	17 \pm 2 (9)	34 \pm 2 (9)	50 \pm 4 (8)
Tauro-CDC	8 \pm 2 (8)	15 \pm 2 (9)	33 \pm 2 (14)	47 \pm 3 (12)

Mean block by 100 μM bile salt of currents induced by 150 μM glutamate and 10 μM glycine.
 n , number of oocytes.

Table S2. Mean IC_{50} or EC_{50} values for Tauro-CDC and agonists (\pm SEM) on GluN2D

Condition	IC_{50} Tauro-CDC [μM] (n)	max. block by Tauro-CDC [%] (n)
10 μM Gly, 150 μM Glu	179 \pm 11 (13)	87 \pm 3 (13)
10 μM Gly, 0.51 μM Glu	137 \pm 8 (11)	100 \pm 4 (11)
0.13 μM Gly, 150 μM Glu	153 \pm 8 (8)	98 \pm 3 (8)
	EC_{50} Glu [μM] (n)	max. Glu-induced current [normalized] (n)
10 μM Gly	1.03 \pm 0.12 (7)	1.00 \pm 0.17 (7)
10 μM Gly + 180 μM Tauro-CDC	0.82 \pm 0.15 (7)	0.25 \pm 0.06 (7)
	EC_{50} Gly [μM] (n)	max. Gly-induced current [normalized] (n)
150 μM Glu	0.25 \pm 0.05 (12)	1.00 \pm 0.09 (12)
150 μM Glu + 180 μM Tauro-CDC	0.15 \pm 0.03 (13)	0.48 \pm 0.07 (13)

Publication III**Hydrophobic alkyl chains substituted to the 8-position of cyclic nucleotides enhance activation of CNG and HCN channels by an intricate enthalpy – entropy compensation**

Maik Otte, Andrea Schweinitz, **Michele Bonus**, Uta Enke, Christina Schumann,
Holger Gohlke, Klaus Benndorf

Scientific Reports (2018), 8, 14960

Contribution: 20%

SCIENTIFIC REPORTS

OPEN

Hydrophobic alkyl chains substituted to the 8-position of cyclic nucleotides enhance activation of CNG and HCN channels by an intricate enthalpy - entropy compensation

Received: 20 June 2018
Accepted: 19 September 2018
Published online: 08 October 2018

Maik Otte¹, Andrea Schweinitz¹, Michele Bonus², Uta Enke¹, Christina Schumann³, Holger Gohlke^{2,4} & Klaus Benndorf¹

Cyclic nucleotide-gated (CNG) and hyperpolarization-activated cyclic nucleotide-gated (HCN) channels are tetrameric non-specific cation channels in the plasma membrane that are activated by either cAMP or cGMP binding to specific binding domains incorporated in each subunit. Typical apparent affinities of these channels for these cyclic nucleotides range from several hundred nanomolar to tens of micromolar. Here we synthesized and characterized novel cAMP and cGMP derivatives by substituting either hydrophobic alkyl chains or similar-sized more hydrophilic heteroalkyl chains to the 8-position of the purine ring with the aim to obtain full agonists of higher potency. The compounds were tested in homotetrameric CNGA2, heterotetrameric CNGA2:CNGA4:CNGB1b and homotetrameric HCN2 channels. We show that nearly all compounds are full agonists and that longer alkyl chains systematically increase the apparent affinity, at the best more than 30 times. The effects are stronger in CNG than HCN2 channels which, however, are constitutively more sensitive to cAMP. Kinetic analyses reveal that the off-rate is significantly slowed by the hydrophobic alkyl chains. Molecular dynamics simulations and free energy calculations suggest that an intricate enthalpy - entropy compensation underlies the higher apparent affinity of the derivatives with the longer alkyl chains, which is shown to result from a reduced loss of configurational entropy upon binding.

Cyclic nucleotide-gated (CNG) and hyperpolarization-activated cyclic nucleotide-gated (HCN) channels are activated by the binding of cyclic nucleotide monophosphates (cNMPs) to specific intracellular cyclic nucleotide binding domains (CNBDs)¹⁻³. Despite this functional similarity, the physiological role and activation mechanism of both channel types are markedly different: CNG channels generate receptor potentials in photoreceptors and olfactory sensory neurons^{4,5} and are solely activated by the binding of cNMPs. In contrast, HCN channels are primarily activated by hyperpolarizing voltage and activation is secondarily enhanced by the binding of cAMP^{2,3,6-10}. HCN channels generate electrical rhythmicity in specialized cells of the heart and brain. In brain neurons they also contribute to setting the membrane potential and to dampen arriving inhibitory and excitatory stimuli (for review see Biel and coworkers¹¹). Both channel types are nonspecific cation channels and, concerning their topology, they belong to the superfamily of six-transmembrane domain voltage-gated channels¹². An intracellular

¹Institut für Physiologie II, Universitätsklinikum Jena, Friedrich-Schiller-Universität Jena, 07740, Jena, Germany. ²Institut für Pharmazeutische und Medizinische Chemie, Heinrich-Heine-Universität Düsseldorf, 40225, Düsseldorf, Germany. ³Fachbereich Medizintechnik und Biotechnologie, Ernst-Abbe-Hochschule Jena, 07745, Jena, Germany. ⁴John von Neumann Institute for Computing (NIC), Jülich Supercomputing Centre (JSC) & Institute for Complex Systems - Structural Biochemistry (ICS 6), Forschungszentrum Jülich GmbH, Jülich, Germany. Maik Otte and Andrea Schweinitz contributed equally. Correspondence and requests for materials should be addressed to K.B. (email: Klaus.Benndorf@med.uni-jena.de)

N-terminus is followed by the α -helical transmembrane domains S1 to S6, with S4 being the voltage sensor, and a hairpin loop between S5 and S6 dipping into the membrane, thereby contributing to the wall of the common pore. In contrast to solely voltage-gated channels, the C-terminus in each subunit contains the CNBD that is connected to the S6 helix via the C-linker.

Native mammalian olfactory CNG channels are composed of three different subunit isoforms, 2xCNGA2, CNGB4 and CNGB1b, of which only CNGA2 subunits can form functional homotetrameric channels when expressed heterologously⁴. For homologous TAX-4 channels of *Caenorhabditis elegans*, a full channel structure at 3.5 Å resolution has been reported recently¹³. Mammalian HCN channels are either homo- or heterotetrameric channels composed of the four subunit isoforms HCN1 to HCN4^{14,15}. All four isoforms can form functional homotetrameric channels. Among the different homotetrameric channels HCN2 channels are strongly modulated by cAMP¹⁶. The binding of cAMP shifts voltage-dependent activation to more depolarized potentials and increases both the rate of channel opening and the maximal current at saturating hyperpolarizing voltages¹⁷. Recently, the structure of the full length HCN1 isoform has been published by means of cryo-electron microscopy at 3.5 Å resolution both in the absence and presence of cAMP¹⁸. Notably, the structure of the tetrameric CNBD of the full-length HCN1 channel is closely similar to respective structures of the isolated CNBDs from three mammalian HCN isoforms^{19–21}.

Based on this structure a scenario for the duality of voltage- and cAMP-induced activation has been proposed. The cAMP-effect to activate the channels has been explained by inducing a concerted rotation of the cytoplasmic domains enhancing opening of the gate formed by the S6-helices in the sense of a relief of auto-inhibition^{22,23}. In more detail, the molecular mechanism of cNMP binding to the CNBD is complex. The CNBD consists of the helices A to C and a β -roll, in which a phosphate binding cassette is embedded, which in turn contains a further short α -helix termed P-helix. For the CNBD of HCN channels, cAMP has been reported to bind in its *anti* conformation¹⁹. The purine ring interacts with the C-helix by a hydrophobic interaction (I636) and by the N6 amine forming a hydrogen bond with backbone carbonyl oxygen of R632. The purine ring interacts via hydrophobic interactions with V564, M572, L574 in the β -roll. Furthermore, the phosphate is stabilized by three hydrogen bonds and an ionic interaction at the phosphate binding cassette. In contrast to cAMP, cGMP binds in the *syn* conformation, enabling its purine ring N2 atom to form a hydrogen bond with T592 of the β -roll¹⁹.

In the past decades, several studies addressed the effects of cNMP modifications on the activation of CNG channels^{24,25}, HCN channels²⁶ and isolated CNBDs of HCN channels^{27–29}. A major result is that an enhanced apparent affinity can be reached by 8-thio-substituted compounds³⁰, suggesting sizeable space around the 8-position of the purine ring of a bound cNMP. Even bulky substituents such as 8-pCPT-, 8-Fluo- or 8-NBD-cNMP can be accommodated^{26,31–33}.

Herein we systematically investigated the impact of modifications at the 8-position of the purine base by thioalkylamines or thioheteroalkylamines within a set of 20 cGMP- and cAMP-derived ligands and quantified their agonistic effects on functional channels. This approach provides the advantage that all cooperative effects among the four CNBDs^{34,35} and reciprocal interactions with the channel core³⁶ are included. To enable more general conclusions, we analyzed three types of channels: homotetrameric CNGA2 channels, heterotetrameric CNGA2:CNGA4:CNGB1b channels and homotetrameric HCN2 channels. For a subset of cNMP derivatives, molecular docking and simulations as well as end-point free energy computations were performed to reveal the underlying reasons for the observed structure-apparent affinity relationships.

Results

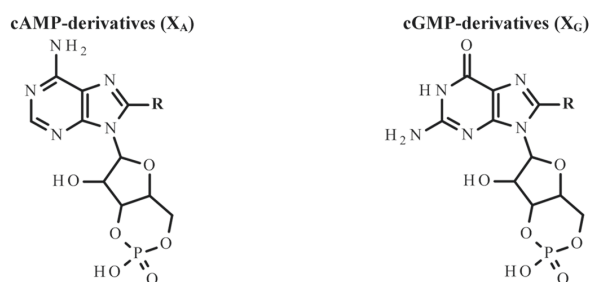
Cyclic nucleotide derivatives. We systematically modified the 8-position of the cNMPs by thioalkylamine residues of increasing length including an ethyl, hexyl or decyl chain. In general an increasing length of the alkyl chain introduces stronger hydrophobicity. To probe the effects of charged versus uncharged residues, N-terminal acetylated and non-acetylated compounds were synthesized. In addition, similar-sized more hydrophilic thioheteroalkylamine residues with different hetero atoms were attached to the 8-position. These include a short polyethylene glycol, and glycine. Names, structures and short names, as used in the text, are summarized in Table 1. The synthesis strategies of the cNMP derivatives are described in the Methods section.

Hydrophobic moieties in 8-position of cyclic nucleotides enhance the apparent affinity to CNG channels.

Ionic currents of both homotetrameric CNGA2 and heterotetrameric CNGA2:CNGA4:CNGB1b channels incorporated in inside-out patches were elicited by applying voltage pulses to -10 and $+10$ mV, starting from a holding potential of 0 mV (Fig. 1a). The current amplitude was generally measured as the late current amplitude at $+10$ mV with respect to the current amplitude in the absence of the cNMP, which was only negligibly small. For each cNMP derivative a complete concentration-activation relationship was recorded, and the data points were fitted with the Hill equation (equation 1) yielding the concentration of half maximum activation (EC_{50}) and the Hill coefficient H (Fig. 1b,c).

We first determined the effect of alkyl chains with increasing lengths on CNGA2 and CNGA2:CNGA4:CNGB1b channels (Fig. 2). First consider the effects of the cGMP derivatives (Fig. 2a,c). Compared to cGMP (1_C), 8-(Ac)AET-cGMP (3_C) reduced the EC_{50} value significantly, and this effect was further enhanced by 8-(Ac)AHT-cGMP (5_C). Prolongation with a decyl chain in 8-(Ac)ADT-cGMP (7_C) did not further reduce the EC_{50} value. In contrast, respectively sized more hydrophilic residues (compounds 8_C – 11_C) left the EC_{50} value approximately unaffected compared to 1_C .

Next consider the effects of the respective cAMP derivatives (Fig. 2b,d). Notably, the pattern produced by the cGMP derivatives was largely reproduced by the respective cAMP derivatives with the only difference that all EC_{50} values were shifted towards higher values. When considering the Hill coefficient H for the respective cGMP and cAMP derivatives, a systematic change at prolonged hydrophobic chains could not be observed (Supplementary Tables 1a–e).



No. of compound	Chain type	Structures at 8-Position (R)	R'
1 _{A/G}	–	–H	–
2 _{A/G}	8-AET-		–H
3 _{A/G}	8-(Ac)AET-		–COCH ₃
4 _{A/G}	8-AHT-		–H
5 _{A/G}	8-(Ac)AHT-		–COCH ₃
6 _{A/G}	8-ADT-		–H
7 _{A/G}	8-(Ac)ADT-		–COCH ₃
8 _{A/G}	8-(Ac)APT-		–COCH ₃
9 _{A/G}	8-(Ac)APET-		–COCH ₃
10 _{A/G}	8-(Ac)AGET-		–COCH ₃
11 _{A/G}	8-(Ac)AGGET-		–COCH ₃

Table 1. Names and structures of cyclic nucleotide derivatives used in the experiments. cNMP is either cGMP (X_G) or cAMP (X_A).

Together, the tested cGMP and cAMP derivatives produce a consistent activity pattern on both homo- and heterotetrameric CNG channels. Moreover, the results allowed us to identify highly potent agonistic cGMP derivatives for homotetrameric CNGA2 channels with EC_{50} values in the range of tens of nanomolar.

The effect of the length of the alkyl chain on the apparent affinity dominates over the effect of a terminal charge. So far, we showed that a thioalkylamino chain in 8-position of the cNMPs enhanced the apparent affinity for cNMPs and that six C-atoms apparently suffice to generate the maximum effect. To test whether the observed effects are influenced by a terminal charge at the chain, we compared the effects of N-acetylated with non-acetylated chains, i.e., neutral with positively charged chains. For the cGMP derivatives with an ethyl chain the acetyl group reduced the EC_{50} value about five times (0.83 μ M for 2_G versus 0.17 μ M for 3_G) in homotetrameric CNGA2 channels. However, with a hexyl or decyl chain, this effect vanished, suggesting that the total length of the chain is the dominating influence on the EC_{50} value (Fig. 3a), while the charge at the terminus of the chain is rather irrelevant. For the cAMP derivatives the pattern is approximately preserved (Fig. 3b). Concerning the ethyl chain, the acetyl group reduced the EC_{50} value about 12 times compared to the non-acetylated chain (128.3 μ M for 2_A versus 9.9 μ M for 3_A), though it should be noted that the derivative 2_A is exceptionally only a partial agonist. Prolongation of the chain increases the apparent affinity further, independent of the charge (derivatives 4_A–7_A in Fig. 3b).

In CNG channels the hydrophobic residues in 8-position predominantly slow down the unbinding rate. To learn more about the molecular mechanism of the enhanced apparent activity by alkyl chains in 8-position of cGMP and cAMP, we analyzed for a representative highly potent cGMP derivative the activation and deactivation time course. We applied concentration jumps from zero to a defined level and back to zero, yielding the time constants τ_a and τ_d , respectively (c.f. Methods section). If the enhancement of the apparent affinity is

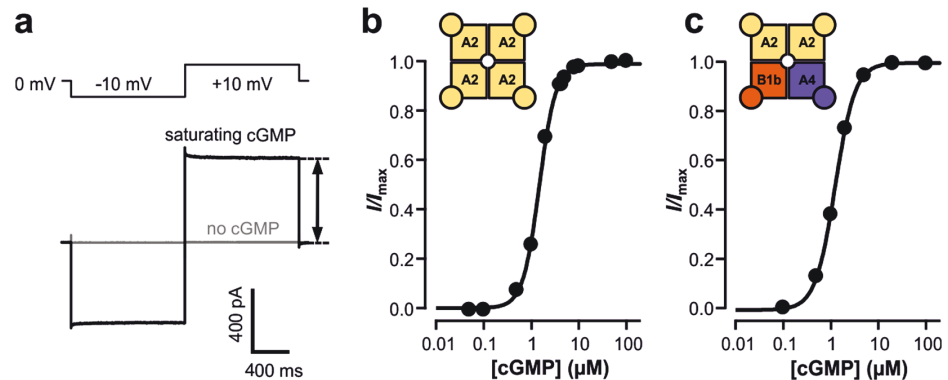


Figure 1. Current measurement and concentration-activation relationships in olfactory CNG channels. (a) Representative current recording from CNGA2 channels activated by saturating cGMP (I_G , 100 μM) according to the indicated voltage protocol. The amplitude of the late current at +10 mV was evaluated for the concentration-activation relationships (black arrow). (b) Concentration-activation relationship for CNGA2 channels. The continuous curve was obtained by fitting equation (1) yielding $EC_{50} = 1.47 \mu\text{M}$ and $H = 2.52$. (c) Concentration-activation relationship for heterotetrameric CNGA2:CNGA4:CNGB1b channels. $EC_{50} = 1.24 \mu\text{M}$ and $H = 2.09$.

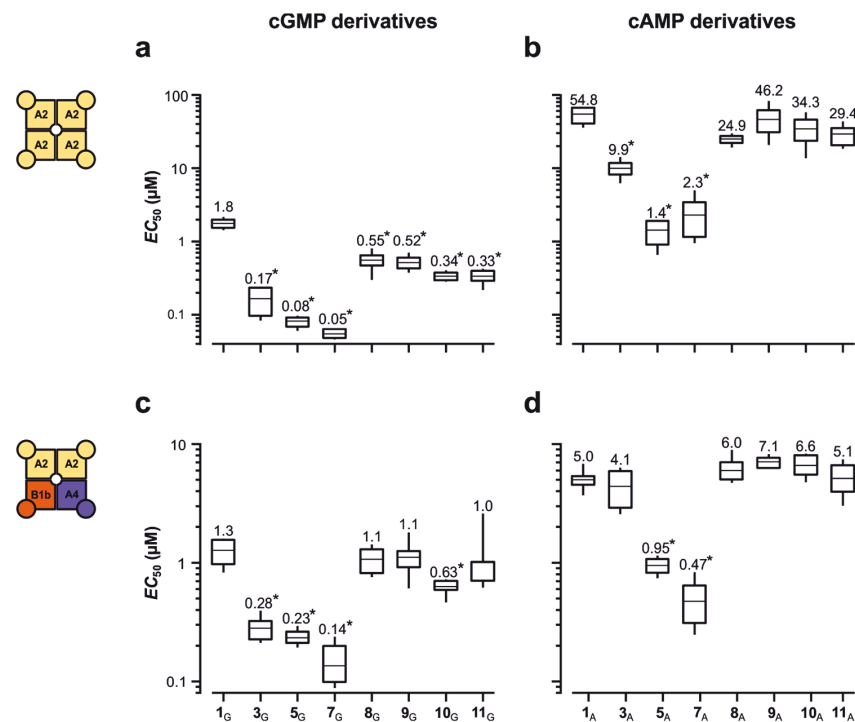


Figure 2. Box plot of the effect of cNMP derivatives on the EC_{50} value for CNG channels. Shown are mean as a horizontal line within each box, the boxes as 25th and 75th percentiles and the whiskers as 10th and 90th percentiles of the data. The numeric mean values are indicated above each box. The asterisks indicate EC_{50} values that were significantly smaller than the respective natural cyclic nucleotide $I_{G/A}$ ($p < 0.01$). (a) CNGA2 channels with cGMP derivatives. (b) CNGA2 channels with cAMP derivatives. (c) CNGA2:CNGA4:CNGB1b channels with cGMP derivatives. (d) CNGA2:CNGA4:CNGB1b channels with cAMP derivatives. Independent of the channel type and the cNMP, the most hydrophobic residues produce the highest apparent affinity whereas the hydrophilic residues, though similar by size, leave the apparent affinity approximately unaffected.

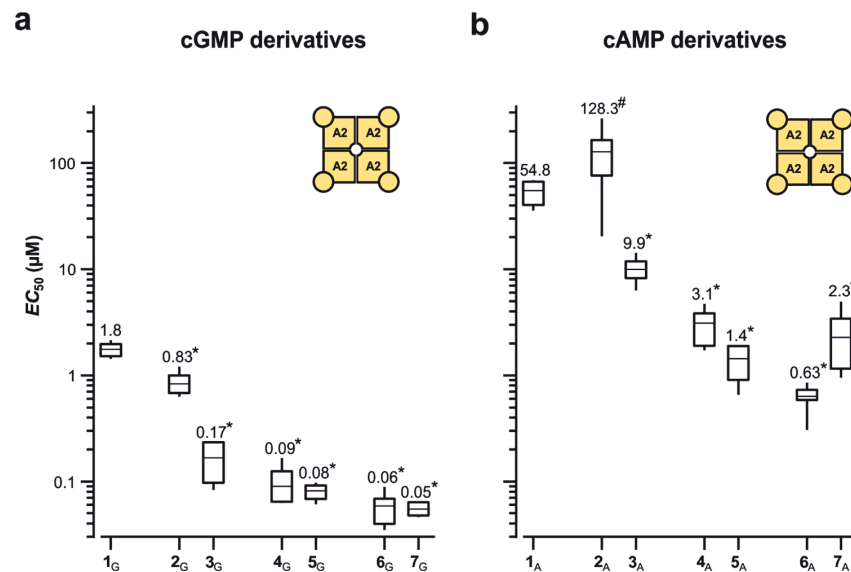


Figure 3. Comparison of the effects of N-acetylated with non-acetylated thioalkylamine residues in CNGA2 channels. Box plot of the EC_{50} values. Generally, the longer residues cause a lower EC_{50} value (c.f. Supplementary Table 1a,b). The EC_{50} values of all cGMP and cAMP derivatives were significantly smaller (asterisks) than those of the natural cyclic nucleotides 1_G and 1_A, respectively ($p < 0.01$). The “#” indicates that compound 2_A is a partial agonist.

caused by solely accelerating the binding rate, τ_a would be decreased and τ_d would be unaffected. Conversely, if the hydrophobic alkyl chain enhances the apparent affinity by solely decelerating the unbinding rate, τ_d would be increased and τ_a could be increased preferentially at lower concentrations.

We chose 8-(Ac)AHT-cGMP (5_G) as representative cGMP derivative and tested it with both CNGA2 and CNGA2:CNGA4:CNGB1b channels. With CNGA2 channels (10 μM cGMP) the activation time course was rather unaffected (Fig. 4a), whereas deactivation was markedly slowed. The plot of the activation time constant τ_a as function of the ligand concentration showed that at 1 μM 8-(Ac)AHT-cGMP (5_G) and higher concentrations, τ_a was similar to that for cGMP (1_G), whereas it was larger at lower concentrations (Fig. 4b). In contrast, the deactivation time constant τ_d was approximately an order of magnitude increased at all tested concentrations of 5_G (Fig. 4c).

With CNGA2:CNGA4:CNGB1b channels, the effects of 10 μM 5_G on activation and deactivation kinetics were qualitatively similar to those with CNGA2 channels, with the difference that deactivation was slowed to a much lower degree (Fig. 4d–f). This reduced effect on deactivation also caused that τ_a was not affected at low concentrations of 5_G (Fig. 4e). Together these results suggest that the hydrophobic chain in 8-position specifically decreases the unbinding rate in both types of channels, which would increase the dwell time in the CNBD. Moreover, the smaller effect of 5_G on CNGA2:CNGA4:CNGB1b channels suggests that the dwell time in the CNGA4 and/or the CNGB1b subunit is shorter than that in the CNGA2 subunits.

A hydrophobic residue in 8-position of cAMP (1_A) also enhances the apparent affinity to HCN2 pacemaker channels.

We next tested whether related HCN2 channels are similarly affected by a hydrophobic chain in 8-position of the purine ring using the natural ligand cAMP (1_A). From a holding potential of 0 mV the channels were pre-activated by hyperpolarizing pulses to -130 mV, and the current amplitude was measured as instantaneous current at -100 mV in both the absence and presence of 1_A (Fig. 5a) and a series of cAMP derivatives. The amplitude of the current differences evoked by 1_A and its derivatives were plotted as function of the concentration, and the data points were fitted by equation 1, yielding again values for EC_{50} and H (Fig. 5b). Comparing the EC_{50} values reveals the following: First, as known for HCN2 channels, a lower 1_A concentration (28.8 nM) was required for activation than for either 1_A or 1_G on CNG channels (c.f. Fig. 2). Second, 3_A did not decrease but slightly increase the EC_{50} value. Third, 5_A and 7_A further decreased the EC_{50} value but the effects were only moderate. Fourth, the more hydrophilic thioheteroalkyl chains with a PEG (8_A) or a glycine (10_A) increased the EC_{50} value about threefold. Nevertheless, the overall pattern for the five tested cAMP derivatives was similar to that for CNG channels: A longer hydrophobic alkyl chain increased the apparent affinity of the derivatives whereas the more hydrophilic heteroalkyl derivatives were ineffective in this respect. The Hill coefficients did not show a systematic dependence on the 8-substitutions (c.f. Supplementary Table 1e).

Different potencies of cAMP derivatives in CNGA2 channels likely result from differential binding affinities.

We intended to provide an explanation at the atomistic level as to the observed

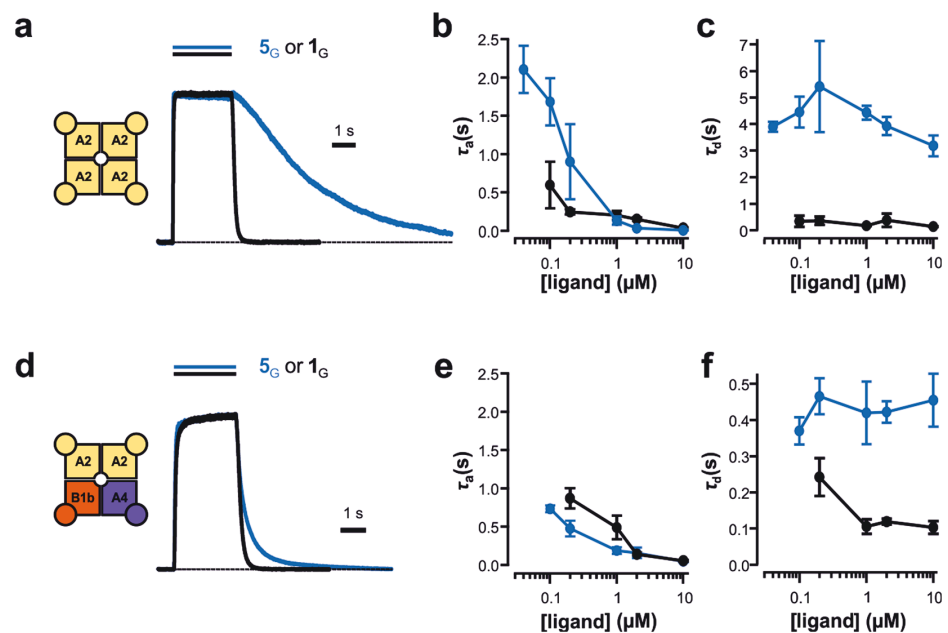


Figure 4. The hydrophobic residue in 8-position slows down the unbinding of 8-(Ac)AHT-cGMP (5_G) compared to cGMP (1_G). The channels were activated and deactivated by concentration jumps evoked by a piezo actuator switching between a control solution and a test solution containing a defined ligand concentration. **(a)** Superimposition of a current time course at $10\ \mu\text{M}$ of 5_G and 1_G in CNGA2 channels. **(b)** Activation time constant τ_a as function of the 5_G and 1_G concentration in CNGA2 channels. **(c)** Deactivation time constant τ_d as function of the 5_G and 1_G concentration in CNGA2 channels. **(d,e,f)** Analog to a, b, c for CNGA2:CNGA4:CNGB1b channels. Error bars indicate SEM.

structure-apparent affinity relationships of cAMP derivatives (2_A-7_A) shown in Supplementary Fig. 1 with respect to binding to CNGA2 channels. We did not consider the heteroalkyl cAMP derivatives (8_A-11_A), because in these cases the range of pEC_{50} values (<0.3 log units) is similar to the accuracy limit of our computations³⁷. By molecular docking³⁸, we initially predicted binding poses (orientations and conformations) of the respective derivatives in a homology model of the homotetrameric CNGA2 channel, which was based on the cryo-EM structure of TAX-4¹³ as a template (overall sequence identity: $\sim 53\%$; sequence identity in the binding pocket of cGMP: $\sim 77\%$).

To further confirm the obtained binding poses, we performed 200 ns of MD simulations for each complex of four cAMP derivatives and a homotetrameric CNGA2 channel in an explicit solvent/explicit membrane environment, followed by MM-PBSA calculations of binding free energies³⁹. Prior to that, we derived dihedral parameters for the torsions involving the bond between the carbon at C8 (GAFF atom type *cc*) and the sulfur of the alkylsulfanyl moiety (GAFF atom type *ss*), which show very good agreement with QM-derived torsion energies (see section *Validation of force field parameters* in the Supplementary Results; Supplementary Fig. 2). Although an implementation of MM-PBSA for implicit solvent/implicit membrane calculations of binding free energies is available⁴⁰, the ligand binding site in the intracellular CNBD is far enough away ($\sim 45\ \text{\AA}$) from the inner leaflet of the membrane to justify omitting an implicit membrane representation in our MM-PBSA computations. All ligands remained stably bound to the cAMP binding site throughout the MD simulations, as evident from an RMSD of the cAMP head group $\leq 1.5\ \text{\AA}$ (Supplementary Fig. 3a), though the tails move considerably (Supplementary Fig. 3b), resulting in overall moderate to large chain motions relative to the protein (Supplementary Fig. 3c). The computed binding free energies display a significant and very good correlation with experimentally determined pEC_{50} values (Fig. 6a); $R_{\varepsilon=1} = 0.89$, $R_{\varepsilon=4} = 0.92$), which indirectly supports the quality of the initial binding poses. The standard error of the mean for ΔG over all subunits was $<0.62\ \text{kcal mol}^{-1}$ for all ligands (Fig. 6a), suggesting a precise ΔG estimate and only small differences between subunits if all four subunits are ligand-bound. Furthermore, very good correlations were found irrespective of whether the dielectric constant of the protein was set to $\varepsilon = 1$ or 4 ⁴¹, indicating that the results are robust with respect to this parameter⁴². Our computations demonstrate that binding of the 8-substituted cAMP derivatives 2_A-7_A to CNGA2 channels becomes increasingly favorable with increasing length of the alkyl chain (for $\varepsilon = 1$: $\Delta\Delta G_{C_2} = 0.00\ \text{kcal mol}^{-1}$ (reference), $\Delta\Delta G_{C_6} = -2.33\ \text{kcal mol}^{-1}$, $\Delta\Delta G_{C_{10}} = -5.75\ \text{kcal mol}^{-1}$; the same trend is observed for $\varepsilon = 4$) and that removal of the terminal charge via N-acetylation provides the largest gain in affinity for those derivatives carrying an ethyl side chain (for $\varepsilon = 1$: $\Delta\Delta G_{C_2} = -3.84\ \text{kcal mol}^{-1}$, $\Delta\Delta G_{C_6} = -3.51\ \text{kcal mol}^{-1}$, $\Delta\Delta G_{C_{10}} = 1.26\ \text{kcal mol}^{-1}$; same trend for $\varepsilon = 4$), which parallels experimental findings on apparent affinities (c.f. Fig. 3b). Taken together, these results suggest that the measured differences in potency of the cAMP derivatives 2_A-7_A predominantly result from affinity differences at the CNBD itself rather than from differences associated with other parts of the channel.

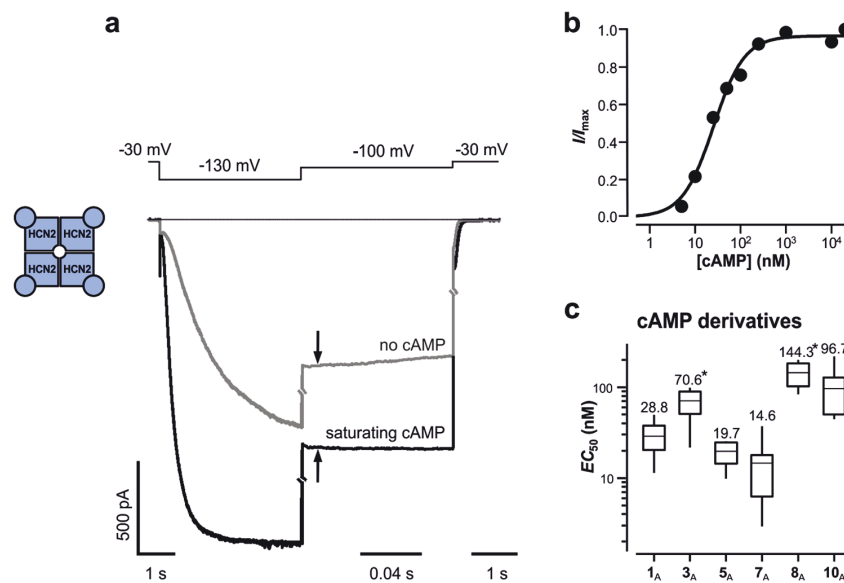


Figure 5. Effects of cAMP derivatives on HCN2 channels. (a) Current traces in the absence and presence of saturating cAMP (20 μM) evoked by the indicated pulse scheme. The increase of the instantaneous current at 130 mV by cAMP (arrows) was evaluated. (b) Concentration-activation relationship for the cAMP-induced current increase of an individual experiment. The continuous curve was obtained by fitting the Hill equation (equation (1)) yielding $EC_{50} = 25.42$ nM and $H = 1.28$. (c) Box plot of the EC_{50} values for the indicated cAMP derivatives. Shown are mean as a horizontal line within each box, the boxes as 25th and 75th percentiles and the whiskers as 10th and 90th percentiles of the data. The numeric mean values are indicated above each box. The EC_{50} values were not significantly different from 1_A apart from 3_A and 8_A which were significantly larger (asterisks; $p < 0.01$).

Interplay of specific electrostatic interactions and differential configurational entropy losses can explain affinity gains of cAMP derivatives with longer alkyl chains and/or N-acetylation. Analysis of tail motions of cAMP derivatives 2_A–7_A bound to CNGA2 channels in terms of anisotropic displacement parameters (ADP; Fig. 6b,c) reveals rather small ADPs for the derivatives with a terminal charge (2_A, 4_A, 6_A), which increase only moderately with respect to an increasing length of the alkyl linker. In contrast, for N-acetylated cAMP derivatives (3_A, 5_A, 7_A), the ADPs are 2.5- to 15-fold larger, and they markedly increase with increasing length of the alkyl linker. Characterizing the conformational heterogeneity of the tail region of the cAMP derivatives in terms of the quadratic means of the pairwise RMSD ($\langle \text{RMSD}^2 \rangle^{1/2}$; Fig. 6d, Supplementary Fig. 4)⁴³ reveals a similar picture in that an N-acetylated cAMP derivative shows a larger $\langle \text{RMSD}^2 \rangle^{1/2}$ than the related one with terminal charge, and the $\langle \text{RMSD}^2 \rangle^{1/2}$ values increase more strongly with the length of the alkyl chain for the former than the latter derivatives. These findings point to an intricate contribution of differences in residual tail motions of bound cAMP derivatives 2_A–7_A to differences in the computed binding affinities.

To further scrutinize the role of the tail, we decomposed the computed binding free energies into respective energy and entropy terms (Supplementary eqs 1, 7), as done previously⁴⁴. For both types of cAMP derivatives, either with a terminal charge or N-acetylation, improved van der Waals interactions and non-polar contributions to the solvation free energy foster binding of derivatives with longer alkyl chains (Fig. 7a,c). Electrostatic interactions, counteracted by polar contributions to the solvation free energy, favor binding of longer cAMP derivatives with a terminal charge (Fig. 7a), whereas they disfavor the C₁₀ derivative with N-acetylation (Fig. 7c). On a structural level, MD trajectories suggest charge-charge repulsions between short-chain cAMP derivatives with a terminal charge and arginine/lysine residues surrounding the binding site (R635 in HCN2, K486, K533, K582 in rCNGA2; Fig. 6b), which are reduced with increasing chain length. In contrast, in the case of N-acetylated derivatives, a longer alkyl chain abrogates charge-assisted hydrogen bonds between the arginine/lysine residues and the acetyl carbonyl group. Similarly, an increasing length of the alkyl linker leads to different configurational entropy contributions of the cAMP derivatives depending on whether a terminal charge is present or N-acetylation: in the former case, configurational entropy contributions to binding become the more disfavored the longer the alkyl linker is (Fig. 7b), in agreement with the fact that ADPs and $\langle \text{RMSD}^2 \rangle^{1/2}$ of bound ligands increase only moderately in this series. In the latter case, differences in the configurational entropy contributions with respect to the length of the alkyl chain are close to zero (Fig. 7d). This is in agreement with the fact that ADPs and $\langle \text{RMSD}^2 \rangle^{1/2}$ of bound ligands increase more strongly in this series, thereby apparently keeping the loss of configurational entropy upon binding constant and small.

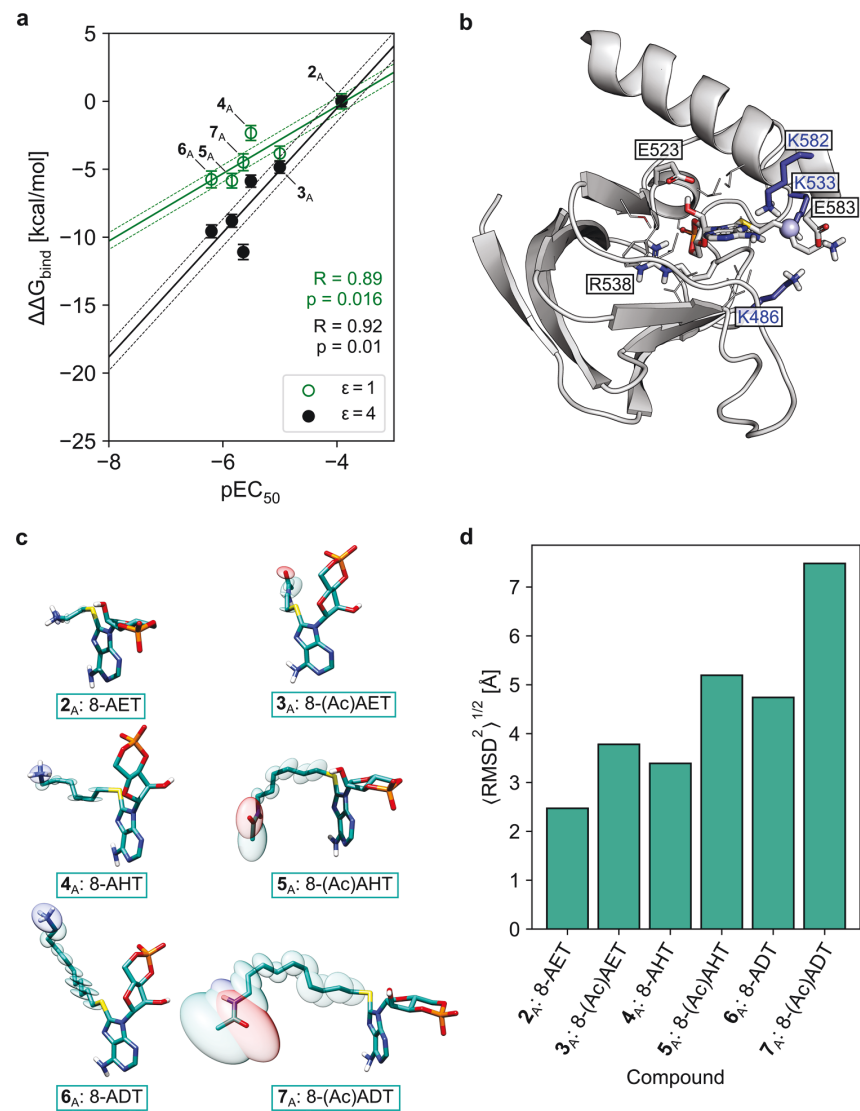


Figure 6. Calculated relative binding free energies ($\Delta\Delta G_{\text{bind}}$) and conformational heterogeneity of cAMP derivatives in CNGA2 channels. **(a)** $\Delta\Delta G_{\text{bind}}$ with respect to 8-AET-cAMP (2_A) for cAMP derivatives of the congeneric series carrying C₂-, C₆-, and C₁₀-substituents (3_A-7_A) for internal dielectric constants of $\epsilon = 1.0$ (green circles) and $\epsilon = 4.0$ (black, filled circles). Linear regression lines are drawn as solid lines in the respective colors, and the standard error of the estimate is indicated by the surrounding dashed lines. **(b)** Binding pose of 8-AHT-cAMP (4_A; grey sticks, center) in a single CNBD. Residues within 5 Å of the ligand are depicted explicitly. Residues forming hydrogen bonds and/or electrostatic interactions to the ligand are depicted as sticks, other residues are depicted as lines. Lysines K486, K533 and K582 are colored in blue. A representative position in which the positively charged nitrogen of a C₂ ligand would be is depicted as blue sphere; the spatial proximity to the surrounding lysines indicates a potential electrostatic repulsion in C₂ derivatives that opposes binding. **(c)** Visualization of the calculated ADPs (ellipsoid scaling factor: 0.5) for the six cAMP derivatives 2_A-7_A. **(d)** Conformational heterogeneity of the tail in the respective cAMP derivatives in the complex with a CNGA2 channel. Bar heights represent the quadratic means of the pairwise root-mean-square deviation (2D-RMSD) of the atomic coordinates of the ligand tail region after root-mean-square fitting of the core/head region.

This view is corroborated if cAMP derivatives of the same chain length but with a terminal charge or N-acetylation are compared (Fig. 7e,f): the latter make favorable electrostatic and van der Waals interactions compared to the first. Configurational entropy contributions favor the short-chain cAMP derivative with a terminal charge, likely because its tail motion is less restricted due to the charge-charge repulsion in contrast to the

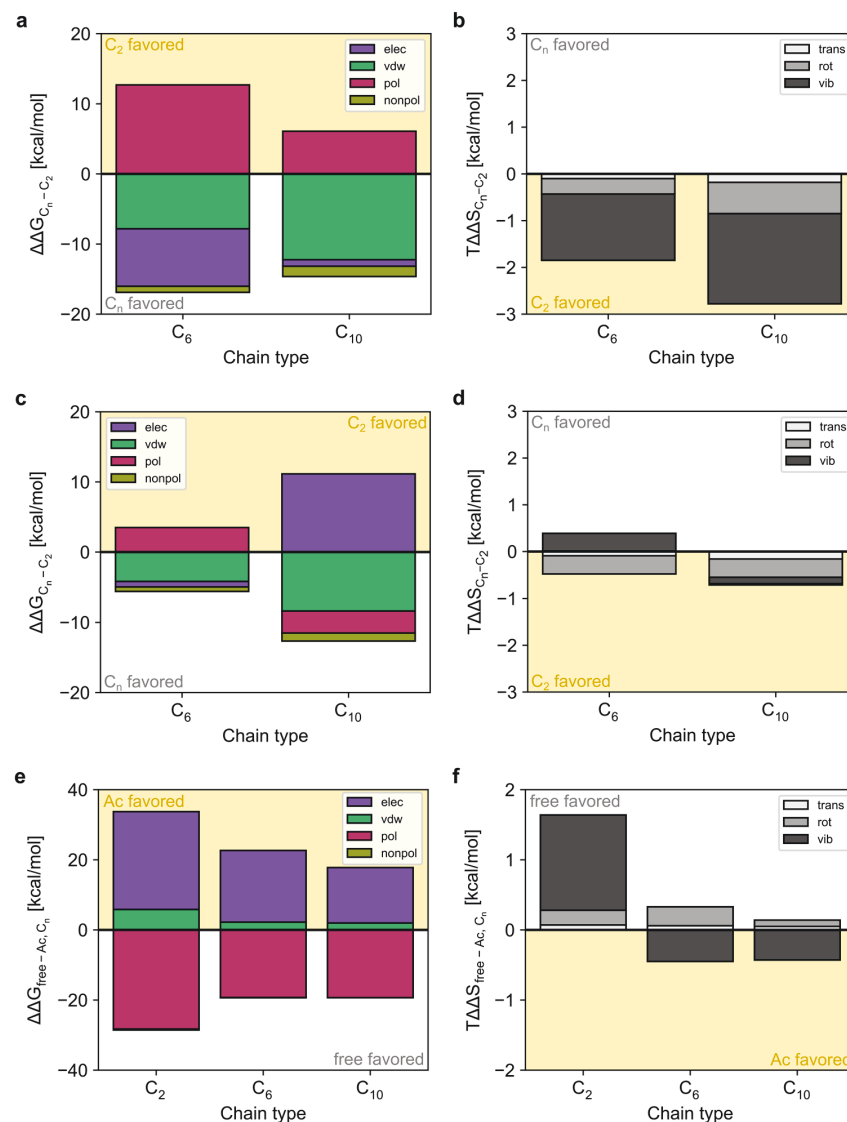


Figure 7. Decomposition of the binding free energy of cAMP derivatives. Binding free energies were decomposed into their respective energy (a,c,e; electrostatic (elec), van der Waals (vdw), polar solvation (pol), and nonpolar solvation (nonpol)) and configurational entropy (b,d,f; translational (trans), rotational (rot), and vibrational (vib)) terms for $\epsilon = 1$. (a,b) Difference between 8-AET-cAMP (2_A) and 8-AHT-cAMP (C₆; 4_A) or 8-ADT-cAMP (C₁₀; 6_A). (c,d) Difference between 8-(Ac)AET-cAMP (3_A) and 8-(Ac)AHT-cAMP (C₆; 5_A) or 8-(Ac)ADT-cAMP (C₁₀; 7_A). (e,f) Difference between 8-AET-cAMP (2_A) and 8-(Ac)AET-cAMP (C₂; 3_A), 8-AHT-cAMP (4_A) and 8-(Ac)AHT-cAMP (C₆; 5_A), and 8-ADT-cAMP (6_A) and 8-(Ac)ADT-cAMP (C₁₀; 7_A).

charge-assisted hydrogen bonds formed by the N-acetyl group; for the longer tails, the differences in the configurational entropy contributions favor cAMP derivatives with N-acetylation, in agreement with the respective larger ADPs and $\langle \text{RMSD}^2 \rangle^{1/2}$ found for those derivatives. Taken together, an intricate interplay of favorable and unfavorable electrostatic interactions as well as differential configurational entropy losses can explain affinity gains of cAMP derivatives with longer alkyl chains and/or N-acetylation.

Discussion

Herein we synthesized 20 mostly new cNMP derivatives modified at the 8-position of the purine moiety and studied the effects of these compounds on the activation gating of homotetrameric CNGA2, heterotetrameric CNGA2:CNGA4:CNGB1b and homotetrameric HCN2 channels. Notably, almost all of these cNMP derivatives are full agonists with potency at the best more than thirty times higher than the respective natural ligands cGMP

(I_G) or cAMP (I_A). We show that alkyl chains of increasing lengths substantially enhance the apparent affinity of both CNGA2 and CNGA2:CNGA4:CNGB1b channels whereas their effects are much weaker in HCN2 channels, which have a high apparent affinity already for I_A . Kinetic analyses of activation and deactivation time courses revealed that the effect of the alkyl chain is caused by a markedly slowed deactivation rate, suggesting that the dwell time of the cNMP derivatives at the CNBD is significantly prolonged when a hydrophobic residue is present in 8-position. Molecular docking and simulations in connection with end-point free energy calculations suggest that the measured differences in potency of the cAMP derivatives predominantly result from affinity differences.

In a previous report it has been shown that I_A and I_G bind in a different conformation to the CNBD: cAMP in *anti* and cGMP in *syn*¹⁹. For TAX-4 channels the *syn* conformation of cGMP has been verified recently¹³. The observed similarity of the effectivity pattern with cGMP and cAMP derivatives may support an earlier result²⁸, suggesting that 8-substitution at the purine ring forces also cAMP derivatives to the *syn* conformation. This was, however, not confirmed by the present MD simulations of cAMP derivatives that were started from the *anti* conformation and displayed only moderate structural deviations of the cAMP head group, indicating the absence of steric strain for the 8-substituted adenine ring. Notably, for several analogs intermediate to cAMP and cGMP a consistent binding to isolated CNBDs of HCN2 in the *anti* conformation was shown by X-ray crystallography⁴⁵. Though these cyclic nucleotides were not substituted as herein, this further supports the idea that the steric strains are only moderate.

Our most potent cNMP derivatives were those that contained a hexyl or decyl chain in 8-position of the purine ring. The EC_{50} values were 50, 140 and 15 nM for functional homotetrameric CNGA2, heterotetrameric CNGA2:CNGA4:CNGB1b and homotetrameric HCN2 channels, respectively (Figs 2 and 5c). To the best of our knowledge these apparent affinities are the highest for CNGA2 and HCN2 channels. For CNGA2:CNGA4:CNGB1b channels an even twofold lower EC_{50} value of 73 nM has been reported for 8-pCPT-cGMP²⁵.

In a recent study the binding affinity of a large number of cAMP derivatives on isolated monomeric CNBDs of HCN1, HCN2 and HCN4 channels has been analyzed²⁹ yielding for cAMP BC_{50} values between 2.0 and 3.7 μ M. These values are much higher than typical EC_{50} values in functional channels in the order of 100 nM and below³⁶. This suggests that isolated CNBDs alone are not able to generate these high affinities typical for functional channels with interacting subunits. This result also fits to previous results obtained by isothermal titration calorimetry⁴⁶ showing that the cAMP binding to an isolated tetrameric CNBD of HCN2 has a high and a low affinity component and that in the monomeric CNBD the high affinity component is lost. A related result has also been reported recently for a single functional subunit in concatameric CNGA2 channels with only one functional CNBD: The EC_{50} value of its activating effect was approximately 22 times larger than that of the channel with four functional subunits⁴⁷. When considering in the report of Moller and coworkers the effects of the best 8-substituted cAMP derivatives, 8-[Fluo]-cAMP and 8-Br-cAMP, on isolated monomeric CNBDs, the BC_{50} values ranged between 100 and 300 nM²⁹. It remains to be tested whether these compounds generate significantly lower EC_{50} values in functional channels than in isolated CNBDs or whether these BC_{50} values for monomers are a lower limit. The latter idea is supported by the result that the EC_{50} value for cAMP binding to HCN2 channels with only one functional CNBD is similar to HCN2 channels with four functional CNBDs⁴⁸, suggesting that, in contrast to CNGA2 channels, the subunits in HCN2 channels form high affinity CNBDs independent of the liganding of the other subunits.

The end-point free energy calculations yielded a very good correlation between computed binding free energies and experimentally determined pEC_{50} values, which is in contrast to numerous earlier results^{42,49,50}. One reason may be the use of a high number of statistically independent conformations (4,000 snapshots extracted every 200 ps), resulting in precise ($SEM < 0.7$ kcal mol⁻¹) computed binding free energies. Furthermore, here, the influence of charge differences in the cAMP derivatives may be less relevant than in previous studies⁴² because the tails remain largely accessible to the solvent. This view is also suggested by very similar trends in computed binding free energies irrespective of using $\epsilon = 1$ or 4. For HCN2 channels, it has been observed that several conformational states in CNBDs regulate the binding of cNMPs⁵¹. It is therefore attractive to speculate that a similar mechanism regulates cNMP binding to CNG channels. Since our experimental results indicate that hydrophobic substituents in 8-position of cNMPs decrease the unbinding rate, we only considered the cNMP “trapped” state in our simulations, which should be the state that predominantly determines the affinities of the tested cNMP derivatives. A detailed investigation of energetic contributions with respect to tail length and presence or absence of N-acetylation revealed an intricate interplay of favorable and unfavorable electrostatic interactions as well as differential configurational entropy losses as underlying reasons for affinity differences. Notably, an increase in residual mobility of the tails of bound, N-acetylated cAMP derivatives results in configurational entropy contributions that become less unfavorable with increasing alkyl chain length compared to non-acetylated derivatives, while at the same time favorable electrostatic (and van der Waals) contributions decrease. This compensation is reflected in moderate decreases in EC_{50} values of these compounds by at most a factor of 15 (excluding 2A) (Fig. 3b) despite adding four methylene units on going from 8-AET to 8-AHT and 8-AHT to 8-ADT; for comparison, for a single buried methyl group, increases of association constants at 298 K by a factor of 3–11 have been estimated^{52–55}. The cAMP derivatives investigated here computationally thus reflect an example of enthalpy/entropy compensation similarly observed for *para*-substituted benzenesulfonamides with “greasy tails” that bind to bovine carbonic anhydrase II, for which a model of decreasing “tightness” of the protein–ligand interface as the chain length of the ligand increases was proposed⁵⁶. Note that in our context “enthalpy” relates to changes in electrostatic and van der Waals interactions and “entropy” to changes in the configurational entropy, as the polar and non-polar parts of solvation free energies contain both enthalpic and entropic components (see Homeyer and coworkers³⁷ for further discussion). A sequence alignment of human, murine and rat HCN1/2 and CNGA1/2 (Supplementary Fig. 5) revealed a high similarity between the residues of rCNGA2 that contact the investigated cNMP derivatives and the corresponding residues in HCN channels, which suggests a similar mechanism of

enthalpy - entropy compensation in HCN channels. This may provide an explanation for the highly similar EC_{50} values of the ligands in CNG and HCN channels, respectively (Fig. 5c). This sequence alignment also reveals distinct (R632) and homologous (R635, which corresponds to K582 in rCNGA2) residues in HCN channels that are suggested to be involved in potential charge-charge repulsions between *N*-terminally charged C_2 derivatives and CNG channels. It is of note that, to the best of our knowledge, the other residues involved in these charge-charge repulsions in CNG channels (Fig. 6b), have not yet been characterized in other studies^{19,58,59}.

It should also be noted that the structure of the CNBDs in CNG and HCN channels is similar to that of other proteins with entirely different function including protein kinase A, exchange protein directly activated by cAMP (Epac) and the prokaryotic cAMP receptor CAP^{60,61}. The interactions of cyclic nucleotides with the CNBDs of these proteins have been intensively studied^{62–66}. It would be of great interest to test the effectivity of the highly affine cNMP derivatives described herein also on these proteins.

Together, our results provide a strategy to generate cNMPs with higher affinity to CNG and HCN channels. On the one hand, this strategy has potential for studying the process of ligand binding, in particular at the single-channel level, because the dwell time of these cNMP derivatives is presumably considerably prolonged. Hence, an appropriate coupling of dyes to these cNMP derivatives is promising to study in single channels individual steps of ligand binding and activation gating in parallel, analogous to previous studies in macropatches^{34,36}. For one fluorescent cAMP derivative, containing the hexyl chain studied herein, its usefulness has been demonstrated: In combination with mutagenesis strategies the binding to defined subunits in CNGA2:CNGA4:CNGB1b channels could be quantified⁶⁷. Furthermore, our high-affinity cNMP derivatives should also be of interest for the development of specific drugs targeting HCN or CNG channels to treat such severe diseases as bradycardia⁶⁸ or achromatopsia⁶⁹.

Methods

Synthesis of cNMP derivatives. The cyclic nucleotides cGMP and cAMP were obtained from Sigma. All synthesized derivatives were monitored and analyzed by reversed phase HPLC and thin layer chromatography (TLC), visualization was effected by UV (254 nm) and confirmed by mass spectroscopy. The structures of all C8-modified cNMP derivatives are given in Table 1. Three of the analogs (**2**_C, **4**_{A/G}) were previously studied in other channel types or isolated CNBDs^{29,30}.

8-Br-cAMP and 8-Br-cGMP (Biolog Life Science Institute, Bremen) are appropriate compounds for the synthesis of the C8 substituted cAMP and cGMP derivatives, respectively. Following the synthesis route described by Brown *et al.*³⁰ with minor modifications, the 8-Br-cNMP was first substituted to the thiol by thiourea. The desired thioether was then obtained with an appropriate linker derivative. The respective linker derivative comprised an *N*-terminal protected bromoalkylamine or bromoheteroalkylamine. The compounds described herein were obtained by deprotection of the terminal amino group and part of them was further acetylated. A detailed description of the synthesis is provided in Supplementary Methods.

Molecular Biology and functional expression. The subunits CNGA2 (accession No. AF126808), CNGA4 (accession No. U12623) and CNGB1b (accession No. AF068572) of rat olfactory channels as well as mouse HCN2 channels (NM008226) were subcloned in front of the T7 promoter of pGEMHEnew. The respective cRNAs were produced by using the mMMESSAGE mMACHINE T7 Kit (Ambion, Austin, TX).

Oocytes of *Xenopus laevis* were obtained either from Ecocyte[®] (Castrop-Rauxel, Germany) or surgically under anesthesia (0.3% 3-aminobenzoic acid ethyl ester) from female adults. The procedures had approval from the authorized animal ethical committee of the Friedrich Schiller University Jena. The methods were carried out in accordance with the approved guidelines.

The oocytes were exposed to collagenase A (3 mg/ml; Roche, Grenzach-Wyhlen, Germany) for 105 min in Ca^{2+} -free Barth's solution containing (in mM) 82.5 NaCl, 2 KCl, 1 MgCl₂, 2, and 5 Hepes, pH 7.4. Oocytes at stages IV and V were injected with 50–130 ng of cRNA encoding either CNGA2, CNGA2:CNGA4:CNGB1b (2:1:1 ratio) or HCN2 channels either manually or by means of an injection robot (RoboInject[®]). The injected oocytes were incubated at 18 °C for up to 6 days in Barth's solution containing (in mM) 84 NaCl, 1 KCl, 2.4 NaHCO₃, 0.82 MgSO₄, 0.41 CaCl₂, 0.33 Ca(NO₃)₂, 7.5 TRIS, cefuroxime (4.0 μg × ml⁻¹), and penicillin/streptomycin (100 μg × ml⁻¹), pH 7.4.

Electrophysiology. Macroscopic currents were recorded from inside-out patches of the oocytes by using standard patch-clamp techniques. The patch pipettes were pulled from quartz tubing (P-2000, Sutter Instrument, Novato, USA) with an outer and inner diameter of 1.0 and 0.7 mm (VITROCOM, New Jersey, USA). The corresponding pipette resistance was 0.9–2.3 MΩ. The bath and pipette solution contained (in mM): 150 KCl, 1 EGTA, 5 Hepes (pH 7.4 with KOH) for CNG measurements. For HCN measurements the bath solution contained 100 mM KCl, 10 mM EGTA, 10 mM HEPES (pH 7.2) and 120 mM KCl, 10 mM HEPES, and 1.0 mM CaCl₂ (pH 7.2) in the pipette. Recording was carried out at room temperature using an Axopatch 200B amplifier (Axon Instruments, Foster City, CA). Electrophysiology was controlled by the Patchmaster-software (HEKA Elektronik Dr. Schulze GmbH, Lambrecht, Germany). The sampling rate was 5 kHz and the filter implemented in the amplifier (4-pole Bessel) was set to 2 kHz. Measurements in HCN2 channels were generally started 3.5 min after patch excision to minimize channel run down during the measurement^{14,27,70}. The solutions with the ligand concentrations to be tested were applied via a multi-barrel device to the patches with a flow rate of 0.8 to 1.2 ml/min.

In experiments designed to record activation and deactivation kinetics in CNG channels, theta-glass pipettes were used in which one barrel contained control solution and the other barrel the test solution supplemented with the ligand. The theta-glass pipette was mounted on a piezo actuator which was controlled by the computer controlling the voltage protocols. The effective switch time of the solution exchange, t_{10-90} , determined by an open

patch pipette and different solutions in the barrels⁷¹, was negligible compared to the time courses of activation and deactivation to be considered. The concentration of all ligands was verified by UV spectroscopy.

Fitting steady-state concentration-activation relationships by Hill functions. Concentration-activation relationships were fitted with the Igor software[®] by

$$I/I_{\max} = 1/(1 + EC_{50}/[CN]^H). \quad (1)$$

I is the actual current amplitude and I_{\max} the maximum current amplitude at a saturating cNMP concentration specified for each cNMP and channel. EC_{50} is the cNMP concentration generating the half maximum current and H the Hill coefficient.

The speed of activation and deactivation was determined by fitting the respective time courses with single exponentials yielding the time constant τ . Part of the time courses required the sum of two exponentials for adequate description, yielding the time constants τ_1 and τ_2 and the contributions A_1 and A_2 , respectively. From these a weighted mean time constant was computed according to

$$\tau_{\text{wm}} = (A_1\tau_1 + A_2\tau_2)/(A_1 + A_2). \quad (2)$$

To quantify the time course at a given cNMP concentration for a number of measurements τ_{wm} and τ were treated equally. Errors are given as mean \pm s.e.m. Statistical tests were performed by the Igor software[®]. The t -test level of significance was set to 0.01 with respect to cGMP or cAMP.

Molecular modelling and simulations. In order to explain the observed structure/apparent affinity relationships of the non-acetylated and N-acetylated cAMP derivatives 2_A–7_A, we first generated a homology model of homotetrameric rat CNGA2 using the cryo-EM structure of the cyclic nucleotide-gated cation channel TAX-4 from *C. elegans* in the cGMP-bound open state (PDB-ID: 5H3O)¹³ as a template. Molecular docking was then used to generate complex structures of derivatives of the congeneric series 8-AET-cAMP (2_A), 8-(Ac)AET-cAMP (3_A), 8-AHT-cAMP (4_A), 8-(Ac)AHT-cAMP (5_A), 8-ADT-cAMP (6_A), and 8-(Ac)ADT-cAMP (7_A) and the homology model of CNGA2. The resulting six complex structures were subjected to molecular dynamics (MD) simulations in conjunction with binding free energy calculations. Details are given in Supplementary Methods.

References

- Kaupp, U. B. *et al.* Primary structure and functional expression from complementary DNA of the rod photoreceptor cyclic GMP-gated channel. *Nature* **342**, 762–766 (1989).
- Santoro, B. *et al.* Identification of a gene encoding a hyperpolarization-activated pacemaker channel of brain. *Cell* **93**, 717–729 (1998).
- Craven, K. & Zagotta, W. CNG and HCN channels: two peas, one pod. *Annu Rev Physiol* **68**, 375–401 (2006).
- Kaupp, U. B. & Seifert, R. Cyclic Nucleotide-Gated Ion Channels. *Physiol Rev* **82**, 769–824 (2002).
- Matulef, K. & Zagotta, W. N. Cyclic nucleotide-gated ion channels. *Annu Rev Cell Dev Biol* **19**, 23–44 (2003).
- DiFrancesco, D. Dual allosteric modulation of pacemaker (f) channels by cAMP and voltage in rabbit SA node. *J Physiol* **515**(Pt 2), 367–376 (1999).
- Santoro, B. & Tibbs, G. R. The HCN gene family: molecular basis of the hyperpolarization-activated pacemaker channels. *Ann NY Acad Sci* **868**, 741–764 (1999).
- Kaupp, U. B. & Seifert, G. Molecular Diversity of Pacemaker Ion Channels. *Annu Rev Physiol* **63**, 235–257 (2001).
- Wang, Z. & Fedida, D. Gating Charge Immobilization Caused by the Transition between Inactivated States in the Kv1.5 Channel. *Biophys J* **81**, 2614–2627 (2001).
- Robinson, R. B. & Siegelbaum, S. A. Hyperpolarization-activated cation currents: from molecules to physiological function. *Annu Rev Physiol* **65**, 453–480, <https://doi.org/10.1146/annurev.physiol.65.092101.142734> (2003).
- Biel, M., Wahl-Schott, C., Michalak, S. & Zong, X. Hyperpolarization-activated cation channels: from genes to function. *Physiol Rev* **89**, 847–885, <https://doi.org/10.1152/physrev.00029.2008> (2009).
- Yu, F. H., Yarov-Yarovoy, V., Gutman, G. A. & Catterall, W. A. Overview of molecular relationships in the voltage-gated ion channel superfamily. *Pharmacol Rev* **57**, 387–395, <https://doi.org/10.1124/pr.57.4.13> (2005).
- Li, M. *et al.* Structure of a eukaryotic cyclic-nucleotide-gated channel. *Nature* **542**, 60–65, <https://doi.org/10.1038/nature20819> (2017).
- Ludwig, A., Zong, X., Jeglitsch, M., Hofmann, F. & Biel, M. A family of hyperpolarization-activated mammalian cation channels. *Nature* **393**, 587–591 (1998).
- Ludwig, A. *et al.* Two pacemaker channels from human heart with profoundly different activation kinetics. *EMBO J* **18**, 2323–2329, <https://doi.org/10.1093/emboj/18.9.2323> (1999).
- Stieber, J., Stockl, G., Herrmann, S., Hassfurth, B. & Hofmann, F. Functional expression of the human HCN3 channel. *J Biol Chem* **280**, 34635–34643, <https://doi.org/10.1074/jbc.M502508200> (2005).
- DiFrancesco, D. & Tortora, P. Direct activation of cardiac pacemaker channels by intracellular cyclic AMP. *Nature* **351**, 145–147, <https://doi.org/10.1038/351145a0> (1991).
- Lee, C. H. & MacKinnon, R. Structures of the Human HCN1 Hyperpolarization-Activated Channel. *Cell* **168**, 111–120 e111, <https://doi.org/10.1016/j.cell.2016.12.023> (2017).
- Zagotta, W. N. *et al.* Structural basis for modulation and agonist specificity of HCN pacemaker channels. *Nature* **425**, 200–205, <https://doi.org/10.1038/nature01922> (2003).
- Xu, X., Vysotskaya, Z. V., Liu, Q. & Zhou, L. Structural basis for the cAMP-dependent gating in the human HCN4 channel. *J Biol Chem* **285**, 37082–37091, <https://doi.org/10.1074/jbc.M111.152033> (2010).
- Lolicato, M. *et al.* Tetramerization dynamics of C-terminal domain underlies isoform-specific cAMP gating in hyperpolarization-activated cyclic nucleotide-gated channels. *J Biol Chem* **286**, 44811–44820, <https://doi.org/10.1074/jbc.M111.297606> (2011).
- Wainger, B. J., DeGennaro, M., Santoro, B., Siegelbaum, S. A. & Tibbs, G. R. Molecular mechanism of cAMP modulation of HCN pacemaker channels. *Nature* **411**, 805–810, <https://doi.org/10.1038/35081088> (2001).
- Wicks, N. L., Wong, T., Sun, J., Madden, Z. & Young, E. C. Cytoplasmic cAMP-sensing domain of hyperpolarization-activated cation (HCN) channels uses two structurally distinct mechanisms to regulate voltage gating. *Proc Natl Acad Sci USA* **108**, 609–614, <https://doi.org/10.1073/pnas.1012750108> (2011).

24. Brown, R. L., Strassmaier, T., Brady, J. D. & Karpen, J. W. The pharmacology of cyclic nucleotide-gated channels: emerging from the darkness. *Curr Pharm Des* **12**, 3597–3613 (2006).
25. Strassmaier, T. & Karpen, J. W. Novel N7- and N1-substituted cGMP derivatives are potent activators of cyclic nucleotide-gated channels. *J Med Chem* **50**, 4186–4194, <https://doi.org/10.1021/jm0702581> (2007).
26. Wu, S. *et al.* State-dependent cAMP binding to functioning HCN channels studied by patch-clamp fluorometry. *Biophys J* **100**, 1226–1232, <https://doi.org/10.1016/j.bpj.2011.01.034> (2011).
27. Bois, P., Renaudon, B., Baruscotti, M., Lénfant, J. & DiFrancesco, D. Activation of f-channels by cAMP analogues in macropatches from rabbit sino-atrial node myocytes. *J Physiol* **501**(Pt 3), 565–571 (1997).
28. Scott, S. P., Shea, P. W. & Dryer, S. E. Mapping ligand interactions with the hyperpolarization activated cyclic nucleotide modulated (HCN) ion channel binding domain using a soluble construct. *Biochemistry* **46**, 9417–9431, <https://doi.org/10.1021/bi6026049> (2007).
29. Moller, S. *et al.* Cyclic nucleotide mapping of hyperpolarization-activated cyclic nucleotide-gated (HCN) channels. *ACS Chem Biol* **9**, 1128–1137, <https://doi.org/10.1021/cb400904s> (2014).
30. Brown, R. L., Bert, R. J., Evans, F. E. & Karpen, J. W. Activation of retinal rod cGMP-gated channels: what makes for an effective 8-substituted derivative of cGMP? *Biochemistry* **32**, 10089–10095 (1993).
31. Caretta, A., Cavaggoni, A. & Sorbi, R. T. Binding stoichiometry of a fluorescent cGMP analogue to membranes of retinal rod outer segments. *Eur J Biochem* **153**, 49–53 (1985).
32. Tanaka, J. C., Eccleston, J. F. & Furman, R. E. Photoreceptor channel activation by nucleotide derivatives. *Biochemistry* **28**, 2776–2784 (1989).
33. Wei, J. Y., Cohen, E. D., Genieser, H. G. & Barnstable, C. J. Substituted cGMP analogs can act as selective agonists of the rod photoreceptor cGMP-gated cation channel. *J Mol Neurosci* **10**, 53–64, <https://doi.org/10.1007/BF02737085> (1998).
34. Biskup, C. *et al.* Relating ligand binding to activation gating in CNGA2 channels. *Nature* **446**, 440–443, <https://doi.org/10.1038/nature05596> (2007).
35. Kusch, J. *et al.* How subunits cooperate in cAMP-induced activation of homotetrameric HCN2 channels. *Nat Chem Biol* **8**, 162–169, <https://doi.org/10.1038/nchembio.747> (2012).
36. Kusch, J. *et al.* Interdependence of receptor activation and ligand binding in HCN2 pacemaker channels. *Neuron* **67**, 75–85, <https://doi.org/10.1016/j.neuron.2010.05.022> (2010).
37. Bash, P. A., Ho, L. L., MacKerell, A. D., Levine, D. & Hallstrom, P. Progress toward chemical accuracy in the computer simulation of condensed phase reactions. *Proc Natl Acad Sci USA* **93**, 3698–3703 (1996).
38. Friesner, R. A. *et al.* Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J Med Chem* **49**, 6177–6196, <https://doi.org/10.1021/jm051256o> (2006).
39. Homeyer, N. & Gohlke, H. Free Energy Calculations by the Molecular Mechanics Poisson-Boltzmann Surface Area Method. *Mol Inform* **31**, 114–122, <https://doi.org/10.1002/minf.201100135> (2012).
40. Homeyer, N. & Gohlke, H. Extension of the free energy workflow FEW towards implicit solvent/implicit membrane MM-PBSA calculations. *Biochim Biophys Acta* **1850**, 972–982, <https://doi.org/10.1016/j.bbagen.2014.10.013> (2015).
41. Genheden, S., Kuhn, O., Mikulskis, P., Hoffmann, D. & Ryde, U. The normal-mode entropy in the MM/GBSA method: effect of system truncation, buffer region, and dielectric constant. *J Chem Inf Model* **52**, 2079–2088, <https://doi.org/10.1021/ci3001919> (2012).
42. Homeyer, N., Stoll, F., Hillisch, A. & Gohlke, H. Binding Free Energy Calculations for Lead Optimization: Assessment of Their Accuracy in an Industrial Drug Design Context. *J Chem Theory Comput* **10**, 3331–3344, <https://doi.org/10.1021/ct5000296> (2014).
43. Kuzmanic, A. & Zagrovic, B. Determination of ensemble-average pairwise root mean-square deviation from experimental B-factors. *Biophys J* **98**, 861–871, <https://doi.org/10.1016/j.bpj.2009.11.011> (2010).
44. Gohlke, H. & Case, D. A. Converging free energy estimates: MM-PB(GB)SA studies on the protein-protein complex Ras-Raf. *J Comput Chem* **25**, 238–250, <https://doi.org/10.1002/jcc.10379> (2004).
45. Ng, L. C. T., Putrenko, I., Baronas, V., Van Petegem, F. & Accili, E. A. Cyclic Purine and Pyrimidine Nucleotides Bind to the HCN2 Ion Channel and Variably Promote C-Terminal Domain Interactions and Opening. *Structure* **24**, 1629–1642, <https://doi.org/10.1016/j.str.2016.06.024> (2016).
46. Chow, S. S., Van Petegem, F. & Accili, E. A. Energetics of cyclic AMP binding to HCN channel C terminus reveal negative cooperativity. *J Biol Chem* **287**, 600–606, <https://doi.org/10.1074/jbc.M111.269563> (2012).
47. Wongsamitkul, N. *et al.* Quantifying the cooperative subunit action in a multimeric membrane receptor. *Sci Rep* **6**, 20974, <https://doi.org/10.1038/srep20974> (2016).
48. Ulens, C. & Siegelbaum, S. A. Regulation of hyperpolarization-activated HCN channels by cAMP through a gating switch in binding domain symmetry. *Neuron* **40**, 959–970 (2003).
49. Hou, T., Wang, J., Li, Y. & Wang, W. Assessing the performance of the molecular mechanics/Poisson Boltzmann surface area and molecular mechanics/generalized Born surface area methods. II. The accuracy of ranking poses generated from docking. *J Comput Chem* **32**, 866–877, <https://doi.org/10.1002/jcc.21666> (2011).
50. Yang, T. *et al.* Virtual screening using molecular simulations. *Proteins* **79**, 1940–1951, <https://doi.org/10.1002/prot.23018> (2011).
51. Goldschen-Ohm, M. P. *et al.* Structure and dynamics underlying elementary ligand binding events in human pacemaking channels. *eLife* **5**, <https://doi.org/10.7554/eLife.20797> (2016).
52. Nicholls, A., Sharp, K. A. & Honig, B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**, 281–296, <https://doi.org/10.1002/prot.340110407> (1991).
53. Williams, D. H., Searle, M. S., Mackay, J. P., Gerhard, U. & Maplestone, R. A. Toward an estimation of binding constants in aqueous solution: studies of associations of vancomycin group antibiotics. *Proc Natl Acad Sci USA* **90**, 1172–1178 (1993).
54. Ross, P. D. & Rekharsky, M. V. Thermodynamics of hydrogen bond and hydrophobic interactions in cyclodextrin complexes. *Biophys J* **71**, 2144–2154, [https://doi.org/10.1016/S0006-3495\(96\)79415-8](https://doi.org/10.1016/S0006-3495(96)79415-8) (1996).
55. Gohlke, H. & Klebe, G. Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors. *Angew Chem Int Ed Engl* **41**, 2644–2676, doi:10.1002/1521-3773(20020802)41:15<2644::AID-ANIE2644>3.0.CO;2-O (2002).
56. Krishnamurthy, V. M., Bohall, B. R., Semetey, V. & Whitesides, G. M. The paradoxical thermodynamic basis for the interaction of ethylene glycol, glycine, and sarcosine chains with bovine carbonic anhydrase II: an unexpected manifestation of enthalpy/entropy compensation. *J Am Chem Soc* **128**, 5802–5812, <https://doi.org/10.1021/ja060070r> (2006).
57. Homeyer, N. *et al.* Interpreting Thermodynamic Profiles of Aminoadamantane Compounds Inhibiting the M2 Proton Channel of Influenza A by Free Energy Calculations. *J Chem Inf Model* **56**, 110–126, <https://doi.org/10.1021/acs.jcim.5b00467> (2016).
58. Zhou, L. & Siegelbaum, S. A. Gating of HCN channels by cyclic nucleotides: residue contacts that underlie ligand binding, selectivity, and efficacy. *Structure* **15**, 655–670, <https://doi.org/10.1016/j.str.2007.04.012> (2007).
59. Saponaro, A. *et al.* A synthetic peptide that prevents cAMP regulation in mammalian hyperpolarization-activated cyclic nucleotide-gated (HCN) channels. *eLife* **7**, <https://doi.org/10.7554/eLife.35753> (2018).
60. Dremier, S., Kopperud, R., Doskeland, S. O., Dumont, J. E. & Maenhaut, C. Search for new cyclic AMP-binding proteins. *FEBS Lett* **546**, 103–107 (2003).
61. Bos, J. L. Epac proteins: multi-purpose cAMP targets. *Trends Biochem Sci* **31**, 680–686, <https://doi.org/10.1016/j.tibs.2006.10.002> (2006).

62. Berman, H. M. *et al.* The cAMP binding domain: an ancient signaling module. *Proc Natl Acad Sci USA* **102**, 45–50, <https://doi.org/10.1073/pnas.0408579102> (2005).
63. Popovych, N., Tzeng, S. R., Tonelli, M., Ebright, R. H. & Kalodimos, C. G. Structural basis for cAMP-mediated allosteric control of the catabolite activator protein. *Proc Natl Acad Sci USA* **106**, 6927–6932, <https://doi.org/10.1073/pnas.0900595106> (2009).
64. Selvaratnam, R., Mazhab-Jafari, M. T., Das, R. & Melacini, G. The auto-inhibitory role of the EPAC hinge helix as mapped by NMR. *PLoS One* **7**, e48707, <https://doi.org/10.1371/journal.pone.0048707> (2012).
65. Selvaratnam, R. *et al.* The projection analysis of NMR chemical shifts reveals extended EPAC autoinhibition determinants. *Biophys J* **102**, 630–639, <https://doi.org/10.1016/j.bpj.2011.12.030> (2012).
66. Akimoto, M. *et al.* Signaling through dynamic linkers as revealed by PKA. *Proc Natl Acad Sci USA* **110**, 14231–14236, <https://doi.org/10.1073/pnas.1312644110> (2013).
67. Nache, V. *et al.* Deciphering the function of the CNGB1b subunit in olfactory CNG channels. *Sci Rep* **6**, 29378, <https://doi.org/10.1038/srep29378> (2016).
68. Baruscotti, M., Bottelli, G., Milanesi, R., DiFrancesco, J. C. & DiFrancesco, D. HCN-related channelopathies. *Pflugers Arch* **460**, 405–415, <https://doi.org/10.1007/s00424-010-0810-8> (2010).
69. Giblin, J. P., Comes, N., Strauss, O. & Gasull, X. Ion Channels in the Eye: Involvement in Ocular Pathologies. *Adv Protein Chem Struct Biol* **104**, 157–231, <https://doi.org/10.1016/bs.apcsb.2015.11.006> (2016).
70. DiFrancesco, D. Characterization of single pacemaker channels in cardiac sino-atrial node cells. *Nature* **324**, 470–473, <https://doi.org/10.1038/324470a0> (1986).
71. Jonas, P. In *Single-Channel Recording* (eds Sakmann, B. & Neher, E.) 231–243 (Plenum Press, New York, 1995).

Acknowledgements

We thank K. Schoknecht, S. Bernhardt, A. Kolchmeier and C. Ranke for excellent technical assistance, to A. Perner, Hans-Knöll Institut Jena, for performing mass spectroscopic measurements. HG is grateful for computational support and infrastructure provided by the “Zentrum für Informations- und Medientechnologie” (ZIM) at the Heinrich Heine University Düsseldorf and the computing time provided by the John von Neumann Institute for Computing (NIC) on the supercomputer JURECA at Jülich Supercomputing Centre (JSC) (user ID: HDD17). This project was partly funded by the German Research Foundation (DFG) within the CRC/TR 166 ReceptorLight (project A05) to KB, Research Group FOR 2518 DynIon (project P2 to KB and project P7 to HG), and by the Thüringer Ministerium für Wirtschaft Wissenschaft und Digitale Gesellschaft, Dag-ION (project 2025-514 to KB).

Author Contributions

A.S., C.S., K.B. participated in research design; M.O., A.S., U.E. conducted the experiments; M.B. performed modeling and simulations; A.S., M.O., U.E., M.B., H.G. performed data analysis; M.O. and M.B. designed the figures; K.B., A.S., M.B. and H.G. wrote the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-33050-5>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018

Supplementary Information

to

Hydrophobic alkyl chains substituted to the 8-position of cyclic nucleotides enhance activation of CNG and HCN channels by an intricate enthalpy - entropy compensation

by Maik Otte, Andrea Schweinitz, Michele Bonus, Uta Enke, Christina Schumann, Holger Gohlke, Klaus Benndorf

Supplementary Methods

Chemical Syntheses

General

All reagents were of analytical or HPLC grade and were purchased from commercial suppliers.

Reactions were followed by analytical reversed phase HPLC at 30°C with a JASCO PU-2080 Plus system (JASCO, Labor- und Datentechnik GmbH, Gross-Umstadt, Germany) consisting of a DG-2080-53 In-Line Degasser, a PU-2080 pump, a column oven (CO-2060 Plus) and a UV-2075 detector using a linear water (A)-methanol (B) gradient (increase of 1 % B/min) both containing 0.1 % formic acid at a flow rate of 1 ml/min. A 250 x 4.6 mm column filled with OTU LipoMare C₁₈ (5 µm, 105 Å) as stationary phase including an OTU LipoMare C₁₈ 5 µm (10 x 4.6 mm, 105 Å) precolumn cartridge (all AppliChrom, Oranienburg, Germany) were used.

Semipreparative HPLC purification was performed on a JASCO chromatograph PU-2087 Plus system equipped with a DG-2080-54 In-Line Degasser, a PU-2087 pump and a UV-975 UV detector using an OTU LipoMare C₁₈ column (5 µm 250 x 8 mm, flow rate 2 ml/min) including a 30 x 8 mm precolumn. Elution was achieved running a linear water (A)-methanol (B) gradient (increase of 45 % B in 120 min) with 0.1 % formic acid, in some cases with 5 mM ammonium acetate (pH 5) was used. Column effluents were monitored at 254 nm. Both HPLC systems were managed by ChromPass Chromatography DataSystem software version 1.8.6.1.

MS characterization was performed by High-resolution UHPLC-MS on an Accela UHPLC system connected to a Q Exactive Orbitrap (both Thermo Scientific). The system was operated in the ESI-MS mode and full scan mass spectra (m/z 100 to 1550) were measured. The UHPLC was run on an RP C₁₈ column (Accucore C₁₈ 100 x 2.1 mm), developed by a flow rate of 0.2 ml/min with a water-acetonitrile gradient containing 0.1 % formic acid as mobile phases. A linear gradient was used starting from 5 % to 98 % acetonitrile in 14 min. Data acquisition and processing was controlled by Thermo Scientific Xcalibur software.

Thin layer chromatography (TLC) was performed on silica gel plates (silica gel 60 F254, Merck, Darmstadt, Germany) in n-butanol/acetic acid/water 5/3/2 (v/v/v) visualization was effected by UV (250 nm).

Synthesis of Precursor 8-thio-cGMP and 8-thio-cAMP

The starting material 8-bromo-cGMP or 8-bromo-cAMP (sodium salt) was purchased from Biolog Life Science Institute (Bremen, Germany) and was converted into the corresponding thiol in accordance with the reported procedure by Brown et al.¹ with minor changes.

Briefly, 50 mg (0.12 mmol) of 8-bromo-cGMP (sodium salt) and a tenfold excess of thiourea (85 mg, 1.2 mmol) were dissolved in 1 ml DMSO and heated on a heating block at 110 °C. The progress of the reactions was monitored by analytical reversed phase HPLC. After completion of the reactions the volume of the raw mixtures was reduced to ~100 µL in a vacuum concentrator RVC 2-18 CDplus (Christ, Osterode, Germany). The crude product was precipitated first in acetone to remove excess of thiourea and then twice with tert-Butyl methyl ether (TBME) to give a pale yellow solid with > 95 % purity (area % HPLC) based on UV absorbance to be used without further purification. Successful

synthesis was confirmed by mass spectrometry (8-thio-cGMP: HPLC R_t = 13.9 min; TLC R_f = 0.55; MS calc. 377.02, MS found m/z : 376.01 (M-H)⁻, 400.0 (M+Na)⁺). The same procedure was used in parallel for 8-thio-cAMP (8-thio-cAMP: HPLC R_t = 12.5 min; TLC R_f = 0.65; MS calc. 361.02, MS found m/z : 360.02 (M-H)⁻, 362.03 (M+H)⁺).

Synthesis of 8-AHT-cGMP (4_G) and 8-(Ac)AHT-cGMP (5_G)

8-thio-cGMP (10 mg, 25 μ M) and N-(6-Bromohexyl)-phthalimide (16 mg, 50 μ M, Alfa Aesar, Karlsruhe, Germany) were dissolved in 1 ml methanol and brought to pH ~ 10 by addition of sodium methoxide in accordance to Caretta et. al.². The mixture was held on 40 °C for 2 h. When the HPLC analysis indicated that the reaction was complete, the mixture was centrifuged the resulting supernatant was concentrated and precipitated in TBME. The N-terminal phthaloyl group was then cleaved with hydrazine hydrate in ethanol containing 5 % water at 80 °C, followed by 6M HCl³. The progress of the reaction was monitored by analytical HPLC. Finally, the mixture was adjusted to neutral pH with NaOH, centrifuged and the resulting deposit precipitated with TBME. The crude product was purified by semipreparative HPLC, followed by lyophilization. The overall yield after purification was 42 % with a purity of > 99 % (HPLC area %) based on UV absorbance (8-AHT-cGMP (4_G): HPLC R_t = 21.9 min; TLC R_f = 0.34; MS calc. 476.12, MS found m/z : 475.12 (M-H)⁻, 499.11 (M+Na)⁺). Part of the product 4_G (3 mg, 6 μ M) was acetylated with a 5 fold excess of acetic anhydride in water/MeOH 2/1 (v/v). The reaction was maintained at pH ~ 8 by the addition of NH₄OH (30 %). The progress was monitored by analytical HPLC. After completion the solvent was removed under reduced pressure. The residue was purified by semipreparative HPLC. After lyophilisation the product was obtained as white powder. Successful synthesis was confirmed by mass spectrometry (8-(Ac)AHT-cGMP (5_G): HPLC R_t = 32,1 min; TLC R_f = 0.55; MS calc. 518.13, MS found m/z : 517.13 (M-H)⁻).

The other compounds used herein were synthesized in an analogue manner. Successful synthesis for all compounds was confirmed by mass spectrometry. Analytical details are summarized in Supplementary Table 3.

Synthesis of 8-AET-cGMP (2_G) and 8-AET-cAMP (2_A)

For 2_G and 2_A the alkylation was done without protecting group und further deprotection. Therefore, 8-thio-cGMP and 8-thio-cAMP were alkylated with 2-Bromoethan-1-amine hydrobromide (Sigma, Darmstadt, Germany), respectively. The compounds were obtained as white powder after lyophilisation. Successful synthesis was confirmed by mass spectrometry. For analytical details see Supplementary Table 3.

Synthesis of 8-AHT-cAMP (4_A), 8-ADT-cGMP (6_G) and 8-ADT-cAMP (6_A)

N-(6-Bromohexyl)-phthalimide (Alfa Aesar, Karlsruhe, Germany) for compound 4_A and N-(10-Bromodecyl)-phthalimide (abcr, Karlsruhe, Germany) for compounds 6_G and 6_A were added to the corresponding 8-thio-cAMP or 8-thio-cGMP as described for 4_G. The compounds were obtained as white powder after lyophilisation. Successful synthesis was confirmed by mass spectrometry. For analytical details see Supplementary Table 3.

Synthesis of 8-(Ac)AET-cGMP (3_G), 8-(Ac)ADT-cGMP (7_G), 8-(Ac)AET-cAMP (3_A), 8-(Ac)AHT-cAMP (5_A), 8-(Ac)ADT-cAMP (7_A)

Acetylation was conducted similarly to the synthesis of 5_G using 2_G, 6_G, 2_A, 4_A, and 6_A for the respective synthesis of 3_G, 7_G, 3_A, 5_A, and 7_A. The compounds were obtained as white powder after lyophilisation. Successful synthesis was confirmed by mass spectrometry. For analytical details see Supplementary Table 3.

Synthesis of 8-(Ac)APT-cGMP (**8_G**), 8-(Ac)APT-cAMP (**8_A**)

BOC-Amino-(PEG)₂-Br (BIOZOL, Eching, Germany) was coupled to the corresponding 8-thio-cGMP or 8-thio-cAMP for the respective synthesis of **8_G** and **8_A**, as described for the synthesis of **4_G**. Deprotection was done with 1M HCl in acetic acid using a standard procedure. Further acetylation was conducted similarly to the synthesis of **5_G**. The compounds were obtained as white powder after lyophilisation. Successful synthesis was confirmed by mass spectrometry. For analytical details see Supplementary Table 3.

Synthesis of 8-(Ac)APET-cGMP (**9_G**), 8-(Ac)APET-cAMP (**9_A**) and 8-(Ac)AGET-cGMP (**10_G**), 8-(Ac)AGGET-cGMP (**11_G**), 8-(Ac)AGET-cAMP (**10_A**), 8-(Ac)AGGET-cAMP (**11_A**)

2_G and **2_A** were coupled to Fmoc-Adoa-OH (Iris Biotech, Marktredwitz, Germany) using a standard PyBOP coupling method for the respective synthesis of **9_G** and **9_A** and to Boc-Gly-OSu (Aldrich, Darmstadt, Germany) for the respective synthesis of **10_G**, **10_A**. Additionally, Boc-Gly-OSu was coupled to **10_G** and **10_A** for the synthesis of **11_G**, **11_A**.

Deprotection was done with 20 % piperidin in dimethylformamide for the Fmoc- and 1M HCl in acetic acid for the Boc-protecting group using common procedures. Further acetylation was conducted similarly to the synthesis of **5_G**. The compounds were obtained as white powder after lyophilisation. Successful synthesis was confirmed by mass spectrometry. For analytical details see Supplementary Table 3.

Homology modeling of the homotetrameric CNGA2 channel structure

A homology model of homotetrameric rat CNGA2 (UniProt accession ID: Q00195) was generated using the 3.5 Å cryo-EM structure of the cyclic nucleotide-gated cation channel TAX-4 from *C. elegans* in the cGMP-bound open state (PDB ID: 5H3O)⁴ as a template. Target-template alignment (sequence identity: ~53 %, coverage: ~77 %, target residues: 133-583 of each monomer) and structural modeling were carried out using the SWISS-MODEL workspace⁵. The C₄ symmetry of the template structure was preserved in the final model. Model quality was assessed with the server implementation of the MolProbity software⁶, version 4.4 and the visualization software KiNG⁷, version 2.23: a Clashescore of 1.72 and a MolProbity score of 1.55 (99th and 94th percentile, respectively) indicated a well-defined structure. No clashes or geometry outliers were detected in or near the cyclic nucleotide-binding domains.

Parameterization of cAMP derivatives

To derive a consistent set of point charges and force field parameters for the cAMP derivatives, the ligand set was split into a core fragment and ten side chain fragments (Supplementary Fig. 1). In particular, 8-methylthio-cAMP was involved in the charge derivation for the core fragment, and the respective *S*-substituted 1-methyl-1*H*-imidazole-2-thiol derivatives were involved in the charge derivation for the side chain fragments.

Construction of molecular fragments for charge derivation

To generate three-dimensional structures of 8-methylthio-cAMP, the structure of cAMP was obtained from the crystal structure of the hyperpolarized-activated cyclic nucleotide-gated ion channel 2 (HCN2, PDB ID: 1Q5O)⁸. A methylthio substituent was added to the nucleotide using the 3D builder of the Maestro (Schrödinger, LLC)⁹ GUI, and free valences were filled with hydrogens while preserving the negative charge at the cyclic phosphate. Three-dimensional structures of the *S*-substituted 1-methyl-1*H*-imidazole-2-thiol derivatives were prepared accordingly. To prevent the formation of intramolecular interactions in the following geometry optimization step, the alkyl chains in these fragments were constructed to adopt an all-anti conformation.

Derivation of atom-centered point charges for the molecular fragments

To obtain charges for the molecular fragments, a gas-phase geometry optimization of each fragment was carried out at the HF/6-31G(d) level using GAUSSIAN 09, Revision B.01¹⁰. To ensure invariance of the molecular electrostatic potential (MEP) with respect to molecular orientation, the MEP computation (level of theory: HF/6-31G(d)) and charge fitting procedures were carried out on the

R.E.D. server¹¹, which uses a rigid-body reorientation algorithm¹² before computing the electrostatic potential. The RESP procedure¹³ with two fitting stages (hyperbolic constraint values: 0.0005/0.001) was used for charge derivation^{14,15}. The MEP was calculated on four layers defined by scaling the atomic van der Waals radii by a factor of 1.4, 1.8, 2.0 and 2.2, respectively, and a point density of 0.28 points au⁻² (1 pt Å⁻²). During charge fitting, additional inter-molecular charge constraints between the methyl group of the 8-methylthio-cAMP and the (1-methyl-1*H*-imidazolyl-2-yl)sulfanyl portion of the *S*-substituted 1-methyl-1*H*-imidazole-2-thiol derivatives with a target value of zero were employed.

Derivation of atom types and missing force field parameters for the cAMP derivatives

Three-dimensional structures of the cAMP derivatives were constructed from the molecular fragments in the xLEaP module of AmberTools17¹⁶. Atom types from the GAFF force field¹⁷ were assigned to each structure using the antechamber program (AmberTools17), and missing force field parameters were determined using the parmchk module¹⁸ (AmberTools17), respectively. The missing dihedral parameters for the na-cc-ss-c3 and the nd-cc-ss-c3 torsions were derived from ab initio calculations. In detail, using 2-(methylthio)-1*H*-imidazole as a model compound, a quantum-mechanically (QM)-derived energy profile of the 360° rotation about the C-S bond in the gas phase was computed in increments of 8° at the HF/6-31G(d) level using GAUSSIAN 09, Revision B.01¹⁰. For each of the 46 output structures, the molecular mechanics (MM) energy was evaluated using the GAFF force field¹⁷, while setting the values of the missing torsion parameters to zero. The values for the potential energy function that need to be described by the missing dihedral parameters to match the QM energies were then calculated as the difference between the QM-derived energy profile and the MM-derived energy profile. The missing parameters were derived using a global optimization strategy.

Generation of cAMP derivatives/CNGA2 complex structures

Structures for the protein-ligand complexes of the cAMP derivatives and the CNGA2 tetramer were generated by molecular docking using the Maestro (Schrödinger, LLC)⁹ GUI and the Glide XP methodology¹⁹. The structure of the CNGA2 homology model was prepared for docking using the default settings of the Protein Preparation Wizard built into Maestro (Schrödinger, LLC)⁹. Protein protonation states were determined using the built-in implementation of PROPKA^{20,21}. No minimization was performed to preserve C₄ symmetry. To determine the optimal location of the docking grid in a CNGA2 monomer, chain A of the cGMP-bound TAX-4 cryo-EM structure⁴ was superimposed onto chain A of the CNGA2 homology model. After this superimposition, the coordinates of the center of mass of cGMP in the TAX-4 structure were used as coordinates for the center of the docking grid in CNGA2. The volume of the inner box of the docking grid was set to 15 Å³, and the dimensions of the outer box were set such that ligands with a length of ~20 Å could be docked. No constraints, rotatable groups, or excluded volumes were defined. Docking was performed using the default settings of the Glide XP methodology as implemented in the Schrödinger Suite v. 2017-1. After completion of the docking runs in chain A, the coordinates of the ligands were duplicated into the other three chains.

Molecular dynamics simulations of cAMP/CNGA2 complexes

System setup

The orientation of the CNGA2 tetramer in the membrane was predicted using the PPM web server²². The calculated transfer free energy of 112.1 kcal/mol and the predicted tilt angle of 0.0 ± 0.0° indicated a high reliability of the predicted orientation. This orientation was then used to embed the simulation systems into a lipid bilayer, consisting of ~80 % DOPC (upper leaflet: 168, lower leaflet: 206) and ~20 % POPC (upper leaflet: 44, lower leaflet: 40) molecules in both membrane leaflets using CHARMM-GUI²³. The minimum water height on the top and bottom of the system was chosen to be 25 Å. The program was then invoked to add 154 mM sodium chloride. The final system comprised ~264,000 atoms. To convert the pdb file created with CHARMM-GUI²³ into a pdb file recognized by Amber and the AmberTools programs, the charmm2amber.py script of AmberTools17 was used. Topology files were built using the tLEaP module in AmberTools17. The ff14SB force field²⁴ was used for the protein, the Lipid14 force field²⁵ was used for the lipids, and the GAFF force field¹⁷ with the modified parameters was used for the ligands.

Equilibration protocol

All MD simulations were carried out using the mixed precision SPFP GPU implementation of the Amber 17 package of molecular simulation programs^{16,26}. A time step of 2 fs was used for integration, and the Langevin thermostat²⁷ was used for temperature control with a collision frequency of $\gamma = 0.01$ ps⁻¹ and a target temperature of $T = 300$ K. Covalent bonds involving hydrogen atoms were constrained using the SHAKE algorithm²⁸. The Particle Mesh Ewald²⁹ method was used to estimate long-range electrostatic interactions, and a cutoff of 10 Å was used for short-range electrostatics and van der Waals forces.

The initial structures were energy minimized for 5,000 steps using the steepest descent algorithm, followed by 5,000 steps of minimization with the conjugate gradient algorithm; during this procedure, all protein and membrane atoms were restrained to their initial positions by harmonic restraints with a force constant of 25.0 kcal mol⁻¹ Å⁻². This step was first repeated with the force constant of the harmonic restraints lowered to 5.0 kcal mol⁻¹ Å⁻², then repeated with the restraints removed from the lipid atoms, and, finally, repeated without restraints. After the final minimization step, the root-mean-square deviation (RMSD) of the atomic positions of the backbone atoms was ≤ 0.73 Å for all systems with respect to the initial structure. 5 ps of NVT-MD (protein restrained; force constant: 5.0 kcal mol⁻¹ Å⁻²) were performed while heating the system from 0 K to 100 K. 115 ps of NPT-MD (protein restrained; force constant: 5.0 kcal mol⁻¹ Å⁻²) were performed for density adaptation while heating the system from 100 K to 300 K. After removing the restraints, additional 4,880 ps of NPT-MD were performed.

Production protocol

For each of the 6 complex structures, 200 ns of NPT-MD were carried out, resulting in an aggregate simulation time of 1.2 μs. Coordinates for analysis and post-processing were saved every 20 ps.

Post-processing and data analysis

Post-processing and analysis of the MD trajectories was performed in CPPTRAJ³⁰ as implemented in AmberTools17.

Binding free energy calculations

All calculations described in the following subsections were performed using the MD trajectories of those CNGA2 complexes that contain ligands of the congeneric series 8-AET-, 8-(Ac)AET-, 8-AHT-, 8-(Ac)AHT-, 8-ADT-, and 8-(Ac)ADT-cAMP. For 8-(Ac)APT-, 8-(Ac)APET-, 8-(Ac)AGET-, and 8-(Ac)AGGET-cAMP, no such calculations were performed, because their pEC₅₀ values span < 0.3 log units.

Calculation of effective binding energies

For each simulation system, the change in effective energy due to ligand binding ($\Delta G_{\text{eff}}^{\text{MM-PBSA}}$) was calculated by the single-trajectory MM-PBSA approach³¹⁻³³ using the MMPBSA.py module³⁴ as implemented in AmberTools17¹⁶. In short, for a set of simulation snapshots $i = [1, N]$, where N is the total number of snapshots selected for analysis, $\Delta G_{\text{eff}, i}^{\text{MM-PBSA}}$ was calculated as the sum of the contributions to the molecular mechanics free energy difference (ΔG_i^{MM}) and the solvation free energy difference (ΔG_i^{solv}) (eq. 1):

$$\Delta G_{\text{eff}, i}^{\text{MM-PBSA}} = \Delta G_{\text{elec}, i}^{\text{MM}} + \Delta G_{\text{vdW}, i}^{\text{MM}} + \Delta G_{\text{int}, i}^{\text{MM}} + \Delta G_{\text{nonpol}, i}^{\text{solv}} + \Delta G_{\text{pol}, i}^{\text{solv}} \quad (\text{eq. 1})$$

where $\Delta G_{\text{elec}}^{\text{MM}}$, $\Delta G_{\text{vdW}}^{\text{MM}}$, and $\Delta G_{\text{int}}^{\text{MM}}$ are contributions of the electrostatic, van der Waals, and internal energies, and $\Delta G_{\text{nonpol}}^{\text{solv}}$ and $\Delta G_{\text{pol}}^{\text{solv}}$ are the nonpolar and polar parts of the solvation free energy. The effective binding energy of each ligand to CNGA2 was then expressed as an average over all snapshots (eq. 2):

$$\Delta G_{\text{eff}}^{\text{MM-PBSA}} = \frac{\sum_{i=1}^N \Delta G_{\text{eff}, i}^{\text{MM-PBSA}}}{N} \quad (\text{eq. 2})$$

with standard error of the mean (SEM; eq. 3)

$$\text{SEM}(\Delta G_{\text{eff}}^{\text{MM-PBSA}}) = \frac{s(\Delta G_{\text{eff}}^{\text{MM-PBSA}})}{\sqrt{N}} \quad (\text{eq. 3})$$

where s denotes the sample standard deviation.

In this study, $\Delta G_{\text{pol}}^{\text{solv}}$ was calculated by the linear Poisson-Boltzmann (PB) equation using an ionic strength of 100 mM, internal dielectric constants (ϵ_{int}) of 1 and 4, and an external dielectric constant (ϵ_{ext}) of 80. Solutions for the linear PB equation were computed with 1,000 iterations on a cubic (lattice) grid with 0.5 Å spacing between grid points. $\Delta G_{\text{nonpol}}^{\text{solv}}$ was considered proportional to the solvent accessible surface area (SASA) according to eq. 4:

$$\Delta G_{\text{nonpol}}^{\text{solv}} = \gamma(\text{SASA}) + \beta + E_{\text{dispersion}} \quad (\text{eq. 4})$$

with $\gamma = 0.0378 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ and $\beta = -0.5692 \text{ kcal mol}^{-1}$. The SASA was calculated with a solvent probe radius of 1.4 Å, Tan & Luo radii³⁵ for the protein, and mbondi radii³⁶ for the ligands. The dispersion term was calculated by a surface-integration approach³⁷.

For each simulation, $\Delta G_{\text{eff}}^{\text{MM-PBSA}}$ was calculated from 1,000 simulation snapshots (time between snapshots: 200 ps). Since the simulated CNGA2 tetramer systems contained a ligand bound to each CNBD, the effective binding energies were calculated as averages over all CNBDs (eq. 5):

$$\Delta G_{\text{eff}}^{\text{MM-PBSA}} = \frac{\Delta G_{\text{eff, CNBD1}}^{\text{MM-PBSA}} + \Delta G_{\text{eff, CNBD2}}^{\text{MM-PBSA}} + \Delta G_{\text{eff, CNBD3}}^{\text{MM-PBSA}} + \Delta G_{\text{eff, CNBD4}}^{\text{MM-PBSA}}}{4} \quad (\text{eq. 5})$$

with standard error of the mean (eq. 6):

$$\text{SEM}(\Delta G_{\text{eff}}^{\text{MM-PBSA}}) = \sqrt{\text{SEM}(\Delta G_{\text{eff, CNBD1}}^{\text{MM-PBSA}})^2 + \text{SEM}(\Delta G_{\text{eff, CNBD2}}^{\text{MM-PBSA}})^2 + \text{SEM}(\Delta G_{\text{eff, CNBD3}}^{\text{MM-PBSA}})^2 + \text{SEM}(\Delta G_{\text{eff, CNBD4}}^{\text{MM-PBSA}})^2} \quad (\text{eq. 6})$$

Calculation of the entropic contribution to binding

Similarly, for each simulation system, the changes in configurational entropy due to ligand binding (ΔS_{NM}) was calculated by normal mode analysis (NMA) as implemented in MMPBSA.py. In short, for a subset of simulation snapshots $i = [1, N]$, where N is the total number of snapshots selected for analysis, $\Delta S_{\text{NM}, i}$ was calculated as the sum of the contributions that arise from changes in translational ($\Delta S_{\text{trans}, i}$), rotational ($\Delta S_{\text{rot}, i}$), and vibrational ($\Delta S_{\text{vib}, i}$) entropy:

$$\Delta S_{\text{NM}, i} = \Delta S_{\text{trans}, i} + \Delta S_{\text{rot}, i} + \Delta S_{\text{vib}, i} \quad (\text{eq. 7})$$

The change in configurational entropy after ligand binding was then estimated as the average over all N snapshots:

$$\Delta S_{\text{NM}} = \frac{\sum_{i=1}^M \Delta S_{\text{NM}, i}}{N} \quad (\text{eq. 8})$$

with SEM:

$$\text{SEM}(\Delta S_{\text{NM}}) = \frac{s(\Delta S_{\text{NM}})}{\sqrt{N}} \quad (\text{eq. 9})$$

A convergence criterion of $0.001 \text{ kcal mol}^{-1} \text{ \AA}^{-1}$ for minimization and the HCT Generalized Born model for solvation³⁸. Due to the computational expense, all calculations were performed on truncated systems that only contain the ligand and the first and second shell of the binding site residues. It has been suggested that such an approach gives rise to a large statistical uncertainty of the calculated entropies^{39,40}, which probably results from larger structural changes in the truncated systems during minimization compared to the complete systems. Solutions to this problem, such as the consideration of a buffer region around the truncated systems that is kept fixed during minimization⁴¹, have been brought forward. However, in these studies, comparably short MD trajectories (~1-2 ns) were used for

the MM-GBSA/PBSA calculations. Since we evaluate here 1,000 snapshots from a 200 ns trajectory for four binding sites, we do not expect the uncertainty in our estimate of the entropy term to become a limiting factor. This expectation is confirmed by a maximal $\text{SEM}(\Delta S_{\text{NM}}) = 0.41 \text{ kcal mol}^{-1}$ found for 8-ADT-cAMP (data not shown).

Calculation of binding free energies

After effective binding energies and contributions from changes in configurational entropy were estimated for each complex, binding free energies (ΔG_{bind}) were calculated according to (eq. 10):

$$\Delta G_{\text{bind}} = \Delta G_{\text{eff}}^{\text{MM-PBSA}} - T\Delta S_{\text{NM}} \quad (\text{eq. 10})$$

with $T = 298.15 \text{ K}$ and SEM (eq. 11):

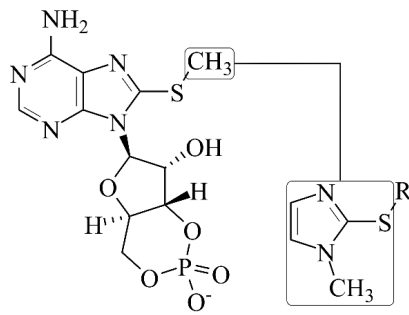
$$\text{SEM}(\Delta G_{\text{bind}}) = \sqrt{\text{SEM}(\Delta G_{\text{eff}}^{\text{MM-PBSA}})^2 + T^2 \text{SEM}(\Delta S_{\text{NM}})^2} \quad (\text{eq. 11})$$

Supplementary Results

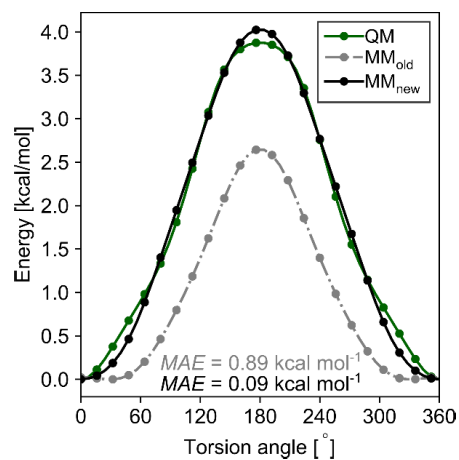
Validation of force field parameters

To assess whether the newly derived na-cc-ss-c3 and nd-cc-ss-c3 dihedral parameters improve the agreement with QM-derived energies, the QM energies of 46 conformers of 2-(methylthio)-1*H*-imidazole, generated by rotation about the na-cc-ss-c3 torsion in steps of 8° , were compared with their respective MM energies obtained with either the parmchk2-derived parameters or the new parameters. While the dihedral parameters obtained from the parmchk2 tool underestimate the torsional energy profile by, on average, $0.89 \text{ kcal mol}^{-1}$ (MAE), the newly derived parameters reduce this error to, on average, $0.09 \text{ kcal mol}^{-1}$ (MAE) (Supplementary Fig. 2). These results demonstrate that the newly derived dihedral parameters improve the agreement with QM-derived energies and, thus, should be better suited to reproduce the conformational energies of the investigated cAMP derivatives.

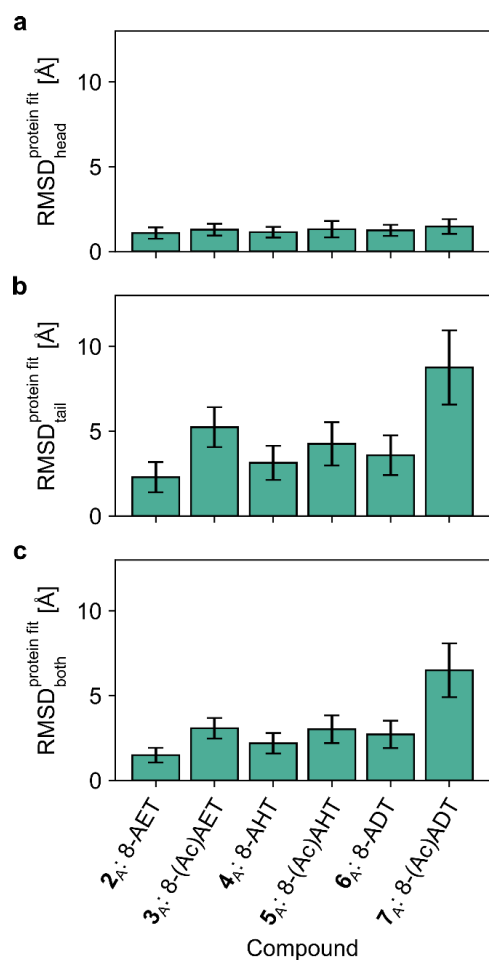
Supplementary Figures

cAMP derivatives	Compound	R-group
	8-AET-cAMP (2 _A)	-C ₂ H ₄ -NH ₃ ⁺
	8-(Ac)AET-cAMP (3 _A)	-C ₂ H ₄ -NHAc
	8-AHT-cAMP (4 _A)	-C ₆ H ₁₂ -NH ₃ ⁺
	8-(Ac)AHT-cAMP (5 _A)	-C ₆ H ₁₂ -NHAc
	8-ADT-cAMP (6 _A)	-C ₁₀ H ₂₀ -NH ₃ ⁺
	8-(Ac)ADT-cAMP (7 _A)	-C ₁₀ H ₂₀ -NHAc

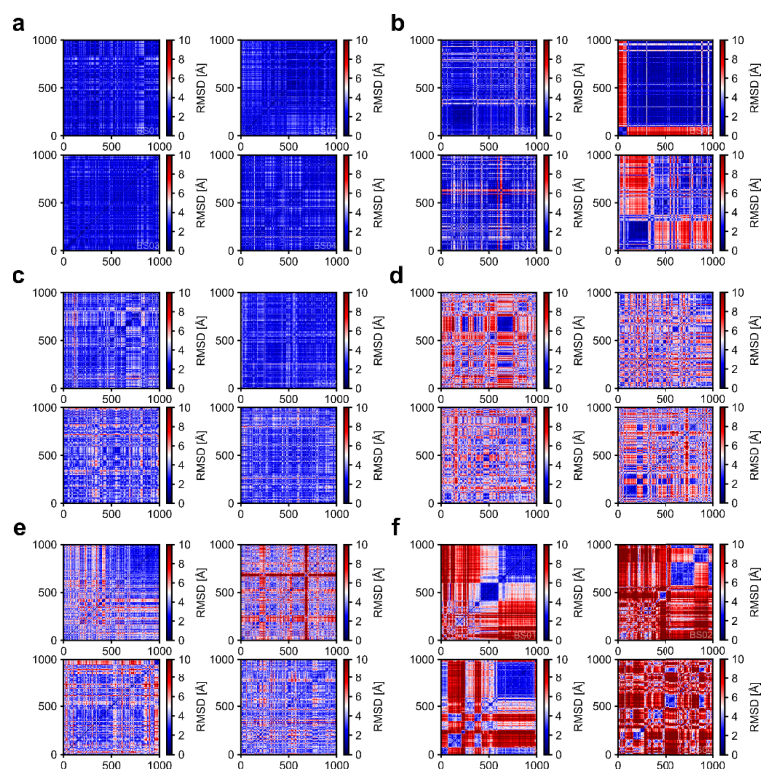
Supplementary Figure 1: Ligand parameterization scheme. Groups for which inter-molecular charge constraints were applied are indicated by rounded boxes and a connecting line.



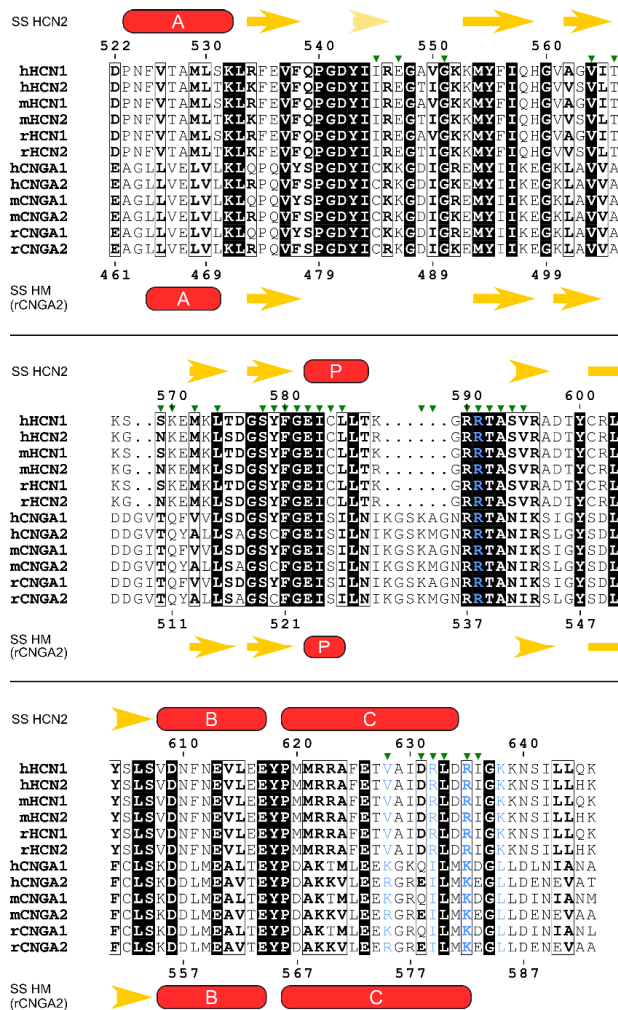
Supplementary Figure 2: Energy profile of the rotation about the na-cc-ss-c3 torsion in 2-(methylthio)-1H-imidazole. The new parameters (black) improve the agreement with the QM-derived energies (green) by, on average, 0.80 kcal mol⁻¹ with respect to the parameters derived from the parmchk2 tool (grey).



Supplementary Figure 3: Root-mean-square deviation (RMSD) of the atomic coordinates of the (a) head region, (b) tail region, and (c) complete ligand of the investigated cAMP derivatives after root-mean-square fitting of the protein backbone. RMSD values were calculated for all four ligands separately with respect to the starting structure and were averaged over the whole MD trajectory. Error bars display sample standard deviations.



Supplementary Figure 4: Conformational heterogeneity of the ligand in complexes of rCNGA2 and cAMP-derivatives. Pairwise root-mean-square deviations (2D-RMSD) of the atomic coordinates of the tail region of the investigated cAMP-derivatives after root-mean-square fitting of the core/head region are color-coded from blue (RMSD = 0 Å) over white (RMSD = 5 Å) to red (RMSD = 10 Å). (a) 8-AET-cAMP (**2_A**), (b) 8-(Ac)AET-cAMP (**3_A**), (c) 8-AHT-cAMP (**4_A**), (d) 8-(Ac)AHT-cAMP (**5_A**), (e) 8-ADT-cAMP (**6_A**), (f) 8-(Ac)ADT-cAMP (**7_A**). Each of the four subpanels represents one binding site (BS01-BS04).



Supplementary Figure 5: Sequence alignment of human, mouse, and rat HCN1/2 and CNGA1/2. Top residue numbers correspond to HCN1/2, bottom residue numbers to rCNGA2. Secondary structure elements are indicated as red rods (helices) and orange arrows (sheets). Residues within 5 Å of the cNMPs are highlighted with green triangles. Identical regions are highlighted with white letters on black background, similar regions are highlighted with a surrounding box. Residues highlighted in blue have been described in other studies as important for cNMP binding in HCN2 channels⁸.

Supplementary Tables

Supplementary Table 1. Effects of cNMP derivatives on homotetrameric CNGA2 and heterotetrameric CNGA2:CNGA4:CNGB1b channels (a-d) and on homotetrameric HCN2 channels (e). The table summarizes the EC_{50} values, Hill coefficients H (mean \pm SEM), and number of experiments, n . * indicates that 2_A acted as a partial agonist on CNGA2 channels. ** indicates that 6_A evoked an unusual inhibition at concentrations $> 2 \mu\text{M}$ in CNGA2:CNGA4:CNGB1b channels.

a) Effects of cGMP derivatives on homotetrameric CNGA2 channels

No. of compound	Chain type	EC_{50}	H	n
1 _G	cGMP	1.8 \pm 0.1	2.6 \pm 0.04	16
2 _G	8-AET-	0.83 \pm 0.05	2.4 \pm 0.29	14
3 _G	8-(Ac)AET-	0.17 \pm 0.03	1.8 \pm 0.24	5
4 _G	8-AHT-	0.09 \pm 0.02	2.5 \pm 0.54	5
5 _G	8-(Ac)AHT-	0.08 \pm 0.004	2.7 \pm 0.14	10
6 _G	8-ADT-	0.06 \pm 0.006	1.8 \pm 0.09	7
7 _G	8-(Ac)ADT-	0.05 \pm 0.003	2.2 \pm 0.11	8
8 _G	8-(Ac)APT-	0.55 \pm 0.04	2.7 \pm 0.25	12
9 _G	8-(Ac)APET-	0.52 \pm 0.03	2.3 \pm 0.13	11
10 _G	8-(Ac)AGET-	0.34 \pm 0.01	2.6 \pm 0.08	12
11 _G	8-(Ac)AGGET-	0.33 \pm 0.02	2.4 \pm 0.11	11

b) Effects of cAMP derivatives on homotetrameric CNGA2 channels

No. of compound	Chain type	EC_{50}	H	n
1 _A	cAMP	54.8 \pm 6.0	2.2 \pm 0.04	4
2 _A	8-AET-	128.3 \pm 20.4*	1.8 \pm 0.17	14
3 _A	8-(Ac)AET-	9.9 \pm 0.8	2.3 \pm 0.11	9
4 _A	8-AHT-	3.1 \pm 0.3	2.3 \pm 0.12	10
5 _A	8-(Ac)AHT-	1.4 \pm 0.2	3.1 \pm 0.28	8
6 _A	8-ADT-	0.63 \pm 0.06	2.3 \pm 0.18	7
7 _A	8-(Ac)ADT-	2.3 \pm 0.4	2.4 \pm 0.23	14
8 _A	8-(Ac)APT-	24.9 \pm 1.3	2.4 \pm 0.11	7
9 _A	8-(Ac)APET-	46.2 \pm 6.9	2.3 \pm 0.09	8
10 _A	8-(Ac)AGET-	34.2 \pm 4.5	2.1 \pm 0.22	10
11 _A	8-(Ac)AGGET-	29.37 \pm 3.1	1.9 \pm 0.11	7

c) Effects of cGMP derivatives on heterotetrameric CNGA2:CNGA4:CNGB1b channels

No. of compound	Chain type	EC_{50}	H	n
1 _G	cGMP	1.3 ± 0.1	2.5 ± 0.28	8
2 _G	8-AET-	1.5 ± 0.2	1.8 ± 0.21	5
3 _G	8-(Ac)AET-	0.28 ± 0.02	2.0 ± 0.10	6
4 _G	8-AHT-	0.62 ± 0.17	1.3 ± 0.14	5
5 _G	8-(Ac)AHT-	0.23 ± 0.01	2.0 ± 0.09	10
6 _G	8-ADT-	0.30 ± 0.02	1.6 ± 0.10	6
7 _G	8-(Ac)ADT-	0.14 ± 0.02	1.9 ± 0.12	7
8 _G	8-(Ac)APT-	1.1 ± 0.1	1.6 ± 0.07	8
9 _G	8-(Ac)APET-	1.1 ± 0.1	2.0 ± 0.18	9
10 _G	8-(Ac)AGET-	0.63 ± 0.03	1.9 ± 0.07	9
11 _G	8-(Ac)AGGET-	1.0 ± 0.2	1.7 ± 0.12	10

d) Effects of cAMP derivatives on heterotetrameric CNGA2:CNGA4:CNGB1b channels

No. of compound	Chain type	EC_{50}	H	n
1 _A	cAMP	5.0 ± 0.3	2.1 ± 0.08	7
2 _A	8-AET-	13.3 ± 0.7	1.0 ± 0.02	8
3 _A	8-(Ac)AET-	4.1 ± 0.6	1.8 ± 0.37	5
4 _A	8-AHT-	1.4 ± 0.1	2.5 ± 0.06	6
5 _A	8-(Ac)AHT-	0.95 ± 0.06	2.5 ± 0.19	5
6 _A	8-ADT-	0.29 ± 0.04**	3.9 ± 1.04	9
7 _A	8-(Ac)ADT-	0.47 ± 0.06	2.7 ± 0.20	10
8 _A	8-(Ac)APT-	6.0 ± 0.4	2.2 ± 0.10	10
9 _A	8-(Ac)APET-	7.1 ± 0.3	2.2 ± 0.13	7
10 _A	8-(Ac)AGET-	6.6 ± 0.4	2.2 ± 0.14	8
11 _A	8-(Ac)AGGET-	5.1 ± 0.6	2.5 ± 0.23	6

e) Effects of cAMP derivatives on homotetrameric HCN2 channels

No. of compound	Chain type	EC_{50}	H	n
1 _A	cAMP	28.8 ± 3.7	1.3 ± 0.13	9
3 _A	8-(Ac)AET-	70.6 ± 9.2	1.7 ± 0.61	7
5 _A	8-(Ac)AHT-	19.7 ± 1.9	0.8 ± 0.07	8
7 _A	8-(Ac)ADT-	14.6 ± 3.4	1.5 ± 0.27	9
8 _A	8-(Ac)APT-	144.3 ± 15.9	0.9 ± 0.11	6
10 _A	8-(Ac)AGET-	96.7 ± 23.7	1.5 ± 0.45	6

Supplementary Table 2. Analytical details of the synthesized compounds.

No.	Short Name	Purity (area %)	HPLC ^a (Rt in min)	MS found (m/z)	TLC ^b (R _f -values)
2 _A	8-AET-cAMP	>99	11.2 ^c	(M-H) ⁻ 403.06 (M+H) ⁺ 405.07	0.18
2 _G	8-AET-cGMP	>99	12.4 ^c	(M-H) ⁻ 419.05 (M+H) ⁺ 421.07	0.28
3 _A	8-(Ac)AET-cAMP	>99	19.5 ^c	(M-H) ⁻ 445.07	0.40
3 _G	8-(Ac)AET-cGMP	>99	19.9 ^c	(M-H) ⁻ 461.07	0.48
4 _A	8-AHT-cAMP	>99	19.8 ^c	(M-H) ⁻ 459.12 (M+Na) ⁺ 483.12	0.24
4 _G	8-AHT-cGMP	>99	21.9	(M-H) ⁻ 475.12 (M+Na) ⁺ 499.11	0.34
5 _A	8-(Ac)AHT-cAMP	>99	34.8	(M-H) ⁻ 501.13 (M+H) ⁺ 503.15	0.65
5 _G	8-(Ac)AHT-cGMP	97	32.1	(M-H) ⁻ 517.13	0.55
6 _A	8-ADT-cAMP	97	15.0 ^{d,e}	(M-H) ⁻ 515.19 (M+Na) ⁺ 539.18	0.60
6 _G	8-ADT-cGMP	>99	18.1 ^d	(M-H) ⁻ 531.18 (M+Na) ⁺ 555.18	0.60
7 _A	8-(Ac)ADT-cAMP	>99	25.5 ^d	(M-H) ⁻ 557.19 (M+H) ⁺ 559.21	0.64
7 _G	8-(Ac)ADT-cGMP	95	28.4 ^{d,e}	(M-H) ⁻ 573.19 (M+H) ⁺ 575.2	0.68
8 _A	8-(Ac)APT-cAMP	>99	24.6	(M-H) ⁻ 553.13 (M+H) ⁺ 535.14	0.29
8 _G	8-(Ac)APT-cGMP	>99	23.5	(M-H) ⁻ 549.12 (M+H) ⁺ 551.13	0.46
9 _A	8-(Ac)APET-cAMP	98	22.9	(M-H) ⁻ 590.14 (M+H) ⁺ 592.16	0.21
9 _G	8-(Ac)APET-cGMP	>95	22.0	(M-H) ⁻ 606.14 (M+H) ⁺ 608.15	0.32
10 _A	8-(Ac)AGET-cAMP	>99	16.0	(M-H) ⁻ 502.09 (M+H) ⁺ 504.11	0.30
10 _G	8-(Ac)AGET-cGMP	>99	15.7	(M-H) ⁻ 518.09 (M+H) ⁺ 520.10	0.35
11 _A	8-(Ac)AGGET-cAMP	>99	16.7	(M-H) ⁻ 559.11 (M+H) ⁺ 561.13	0.20
11 _G	8-(Ac)AGGET-cGMP	>99	16.6	(M-H) ⁻ 575.11 (M+H) ⁺ 577.19	0.26

Solvent systems:

^a standard conditions for analytical HPLC: linear water (A)-methanol (B) gradient with 0.1 % formic acid at a flow rate of 1 ml/min (increase of 1 % B/min), 30°C, monitored at 254 nm, start at 1 % B (^c start at 0 % B, 25°C; ^d start at 30 % B). All compounds were purified by semipreparative HPLC with a water-methanol gradient containing 0.1 % formic acid; ^e in these rare cases the eluent for semipreparative HPLC contained 5 mM ammonium acetate (pH 5) in the aqueous phase.

^b TLC n-butanol/acetic acid/H₂O: 5/3/2(v/v/v).

References

- 1 Brown, R. L., Bert, R. J., Evans, F. E. & Karpen, J. W. Activation of retinal rod cGMP-gated channels: what makes for an effective 8-substituted derivative of cGMP? *Biochemistry* **32**, 10089-10095 (1993).
- 2 Caretta, A., Cavaggioni, A. & Sorbi, R. T. Binding Stoichiometry of a Fluorescent Cgmp Analog to Membranes of Retinal Rod Outer Segments. *Eur J Biochem* **153**, 49-53, doi:DOI 10.1111/j.1432-1033.1985.tb09265.x (1985).
- 3 Bergmann, E. D. & Migron, Y. Preparation of Olefinic Derivatives of Phthalimide and Their Use as Precursors for Homoallylic Amines. *Org Prep Proced Int* **8**, 75-80, doi:DOI 10.1080/00304947609355592 (1976).
- 4 Li, M. *et al.* Structure of a eukaryotic cyclic-nucleotide-gated channel. *Nature* **542**, 60-65, doi:10.1038/nature20819 (2017).
- 5 Biasini, M. *et al.* SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res* **42**, W252-258, doi:10.1093/nar/gku340 (2014).
- 6 Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12-21, doi:10.1107/S0907444909042073 (2010).
- 7 Chen, V. B., Davis, I. W. & Richardson, D. C. KING (Kinemage, Next Generation): a versatile interactive molecular and scientific visualization program. *Protein Sci* **18**, 2403-2409, doi:10.1002/pro.250 (2009).
- 8 Zagotta, W. N. *et al.* Structural basis for modulation and agonist specificity of HCN pacemaker channels. *Nature* **425**, 200-205, doi:10.1038/nature01922 (2003).
- 9 Schrödinger, LLC. *Schrödinger Release 2017-1: Maestro* (2017).
- 10 Frisch, M. J. *et al.* (Wallingford CT, 2009).
- 11 Vanqualef, E. *et al.* R.E.D. Server: a web service for deriving RESP and ESP charges and building force field libraries for new molecules and molecular fragments. *Nucleic Acids Res* **39**, W511-517, doi:10.1093/nar/gkr288 (2011).
- 12 Dupradeau, F.-Y. *et al.* The R.E.D. tools: advances in RESP and ESP charge derivation and force field library building. *Phys Chem Chem Phys* **12**, 7821-7839, doi:10.1039/c0cp00111b (2010).
- 13 Wang, J., Cieplak, P. & Kollman, P. A. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J Comput Chem* **21**, 1049-1074, doi:Doi 10.1002/1096-987x(200009)21:12<1049::Aid-Jcc3>3.3.Co;2-6 (2000).
- 14 Cornell, W. D. *et al.* A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J Am Chem Soc* **117**, 5179-5197, doi:10.1021/ja00124a002 (1995).
- 15 Hornak, V. *et al.* Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins* **65**, 712-725, doi:10.1002/prot.21123 (2006).
- 16 Case, D. A. *et al.* *AMBER 2017* (University of California, San Francisco, 2017).
- 17 Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and testing of a general amber force field. *J Comput Chem* **25**, 1157-1174, doi:10.1002/jcc.20035 (2004).
- 18 Wang, J., Wang, W., Kollman, P. A. & Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *J Mol Graph Model* **25**, 247-260, doi:10.1016/j.jmglm.2005.12.005 (2006).
- 19 Friesner, R. A. *et al.* Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J Med Chem* **49**, 6177-6196, doi:10.1021/jm051256o (2006).
- 20 Rostkowski, M., Olsson, M. H., Sondergaard, C. R. & Jensen, J. H. Graphical analysis of pH-dependent properties of proteins predicted using PROPKA. *BMC Struct Biol* **11**, 6, doi:10.1186/1472-6807-11-6 (2011).
- 21 Olsson, M. H. M., Sondergaard, C. R., Rostkowski, M. & Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. *J Chem Theory Comput* **7**, 525-537, doi:10.1021/ct100578z (2011).
- 22 Lomize, M. A., Pogozheva, I. D., Joo, H., Mosberg, H. I. & Lomize, A. L. OPM database and PPM web server: resources for positioning of proteins in membranes. *Nucleic Acids Res* **40**, D370-376, doi:10.1093/nar/gkr703 (2012).

- 23 Jo, S., Kim, T., Iyer, V. G. & Im, W. CHARMM-GUI: a web-based graphical user interface for CHARMM. *J Comput Chem* **29**, 1859-1865, doi:10.1002/jcc.20945 (2008).
- 24 Maier, J. A. *et al.* ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J Chem Theory Comput* **11**, 3696-3713, doi:10.1021/acs.jctc.5b00255 (2015).
- 25 Dickson, C. J. *et al.* Lipid14: The Amber Lipid Force Field. *J Chem Theory Comput* **10**, 865-879, doi:10.1021/ct4010307 (2014).
- 26 Le Grand, S., Götz, A. W. & Walker, R. C. SPFP: Speed without compromise-A mixed precision model for GPU accelerated molecular dynamics simulations. *Comput Phys Commun* **184**, 374-380, doi:10.1016/j.cpc.2012.09.022 (2013).
- 27 Pastor, R. W., Brooks, B. R. & Szabo, A. An Analysis of the Accuracy of Langevin and Molecular-Dynamics Algorithms. *Mol Phys* **65**, 1409-1419, doi:Doi 10.1080/00268978800101881 (1988).
- 28 Ryckaert, J.-P., Ciccotti, G. & Berendsen, H. J. C. Numerical-Integration of Cartesian Equations of Motion of a System with Constraints - Molecular-Dynamics of N-Alkanes. *J Comput Phys* **23**, 327-341, doi:Doi 10.1016/0021-9991(77)90098-5 (1977).
- 29 Darden, T., York, D. & Pedersen, L. Particle Mesh Ewald - an N-log(N) Method for Ewald Sums in Large Systems. *J Chem Phys* **98**, 10089-10092, doi:Doi 10.1063/1.464397 (1993).
- 30 Roe, D. R. & Cheatham, T. E., III. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *Journal of Chemical Theory and Computation* **9**, 3084-3095, doi:10.1021/Ct400341p (2013).
- 31 Wang, C., Greene, D. A., Xiao, L., Qi, R. & Luo, R. Recent Developments and Applications of the MMPBSA Method. *Front Mol Biosci* **4**, 87, doi:10.3389/fmolb.2017.00087 (2017).
- 32 Gohlke, H. & Case, D. A. Converging free energy estimates: MM-PB(GB)SA studies on the protein-protein complex Ras-Raf. *J Comput Chem* **25**, 238-250, doi:10.1002/jcc.10379 (2004).
- 33 Srinivasan, J., Cheatham, T. E., III, Cieplak, P., Kollman, P. A. & Case, D. A. Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate - DNA helices. *J Am Chem Soc* **120**, 9401-9409, doi:DOI 10.1021/ja981844+ (1998).
- 34 Miller, B. R., III *et al.* MMPBSA.py: An Efficient Program for End-State Free Energy Calculations. *J Chem Theory Comput* **8**, 3314-3321, doi:10.1021/ct300418h (2012).
- 35 Tan, C., Yang, L. & Luo, R. How well does Poisson-Boltzmann implicit solvent agree with explicit solvent? A quantitative analysis. *J Phys Chem B* **110**, 18680-18687, doi:10.1021/jp063479b (2006).
- 36 Tsui, V. & Case, D. A. Theory and applications of the generalized Born solvation model in macromolecular simulations. *Biopolymers* **56**, 275-291, doi:10.1002/1097-0282(2000)56:4<275::AID-BIP10024>3.0.CO;2-E (2000).
- 37 Tan, C., Tan, Y.-H. & Luo, R. Implicit nonpolar solvent models. *J Phys Chem B* **111**, 12263-12274, doi:10.1021/jp073399n (2007).
- 38 Hawkins, G. D., Cramer, C. J. & Truhlar, D. G. Pairwise Solute Descreening of Solute Charges from a Dielectric Medium. *Chem Phys Lett* **246**, 122-129, doi:Doi 10.1016/0009-2614(95)01082-K (1995).
- 39 Weis, A., Katebzadeh, K., Söderhjelm, P., Nilsson, I. & Ryde, U. Ligand affinities predicted with the MM/PBSA method: dependence on the simulation method and the force field. *J Med Chem* **49**, 6596-6606, doi:10.1021/jm0608210 (2006).
- 40 Page, C. S. & Bates, P. A. Can MM-PBSA calculations predict the specificities of protein kinase inhibitors? *J Comput Chem* **27**, 1990-2007, doi:10.1002/jcc.20534 (2006).
- 41 Genheden, S., Kuhn, O., Mikulskis, P., Hoffmann, D. & Ryde, U. The normal-mode entropy in the MM/GBSA method: effect of system truncation, buffer region, and dielectric constant. *J Chem Inf Model* **52**, 2079-2088, doi:10.1021/ci3001919 (2012).

Publication IV***N*⁶-modified cAMP derivatives that activate protein kinase A
also act as full agonists of murine HCN2 channels**

Tim Leypold[§], **Michele Bonus**[§], Felix Spiegelhalter, Frank Schwede, Tina Schwabe,
Holger Gohlke, Jana Kusch

Journal of Biological Chemistry (2019), 294(47), 17978-17987



Contribution: 30%

[§] Both authors contributed equally to this work



N⁶-modified cAMP derivatives that activate protein kinase A also act as full agonists of murine HCN2 channels

Received for publication, July 17, 2019, and in revised form, October 15, 2019. Published, Papers in Press, October 15, 2019. DOI 10.1074/jbc.RA119.010246

Tim Leypold^{#1}, Michele Bonus^{S1}, Felix Spiegelhalter[‡], Frank Schwede[¶], Tina Schwabe[‡],  Holger Gohlke^{SII}, and  Jana Kusch^{#2}

From the [‡]Friedrich Schiller University, University Hospital Jena, Institute of Physiology II, Kollegiengasse 9, 07743 Jena, Germany, the ^SInstitute for Pharmaceutical and Medical Chemistry, Heinrich Heine University, Universitätsstraße 1, 40225 Düsseldorf, Germany, [¶]BIOLOG LSI GmbH & Co. KG, 28199 Bremen, Germany, and the ^{II}John von Neumann Institute for Computing, Jülich Supercomputing Centre and Institute for Complex Systems - Structural Biochemistry (ICS 6), Forschungszentrum Jülich GmbH, 52425 Jülich, Germany

Edited by Mike Shipston

cAMP acts as a second messenger in many cellular processes. Three protein types mainly mediate cAMP-induced effects: PKA, exchange protein directly activated by cAMP (Epac), and cyclic nucleotide–modulated channels (cyclic nucleotide–gated or hyperpolarization-activated and cyclic nucleotide–modulated (HCN) channels). Discrimination among these cAMP signaling pathways requires specific targeting of only one protein. Previously, cAMP modifications at position N⁶ of the adenine ring (PKA) and position 2′-OH of the ribose (Epac) have been used to produce target-selective compounds. However, cyclic nucleotide–modulated ion channels were usually outside of the scope of these previous studies. These channels are widely distributed, so possible channel cross-activation by PKA- or Epac-selective agonists warrants serious consideration. Here we demonstrate the agonistic effects of three PKA-selective cAMP derivatives, N⁶-phenyladenosine-3′,5′-cyclic monophosphate (N⁶-Phe-cAMP), N⁶-benzyladenosine-3′,5′-cyclic monophosphate (N⁶-Bn-cAMP), and N⁶-benzoyl-adenosine-3′,5′-cyclic monophosphate (N⁶-Bnz-cAMP), on murine HCN2 pacemaker channels. Electrophysiological characterization in *Xenopus* oocytes revealed that these derivatives differ in apparent affinities depending on the modification type but that their efficacy and effects on HCN2 activation kinetics are similar to those of cAMP. Docking experiments suggested a pivotal role of Arg-635 at the entrance of the binding pocket in HCN2, either causing stabilizing cation– π interactions with the aromatic ring in N⁶-Phe-cAMP or N⁶-Bn-cAMP or a steric clash with the aromatic ring in N⁶-Bnz-cAMP. A reduced apparent affinity of N⁶-Phe-cAMP toward the variants R635A and R635E strengthened that notion. We conclude that some PKA activators also effectively activate HCN2 channels. Hence, when studying PKA-mediated cAMP signaling with cAMP derivatives in a native environment, activation of HCN channels should be considered.

cAMP, the first identified second messenger (1), plays a key role in living organisms ranging from *Dictyostelium* to *Homo sapiens*. It is involved in a wide variety of cellular processes, including proliferation, differentiation, secretion, migration, pacemaking, sensation, and apoptosis. The concentration of cytosolic cAMP is controlled by the action of adenylyl cyclases, catalyzing the conversion of ATP to cAMP, and cyclic nucleotide phosphodiesterases, catalyzing the conversion of cAMP to 5′-AMP. In mammalian cells, four main types of proteins respond to cAMP: PKA, exchange protein directly activated by cAMP (Epac1 and Epac2),³ CN-modulated ion channels (cyclic nucleotide–gated channels and hyperpolarization-activated and cyclic-nucleotide modulated (HCN) channels) (2) and the relatively new class of Popeye domain–containing proteins (3).

HCN channels, as one type of CN-modulated ion channels, are activated by hyperpolarizing membrane potentials as a primary stimulus but are secondarily stimulated by direct binding of cyclic nucleotides, mainly cAMP, to intracellular binding sites (4–10). Structurally, HCN channels belong to the superfamily of tetrameric voltage-gated ion channels (11). In these channels, each subunit contains a voltage sensor domain and a pore domain. In contrast to most other members of this superfamily, HCN channel subunits additionally contain a cyclic nucleotide–binding domain (CNBD) in the C terminus, which is connected to the membrane portion by a so-called C-linker (12). It has been shown that the unoccupied CNBD together with the C-linker has an inhibitory effect on channel gating, which is relieved by CN binding (13). Such a relieving effect is reflected by a shift of the steady-state activation curve to more depolarized voltages, thereby increasing the maximum current amplitude, accelerating activation and decelerating deactivation (14, 15).

The CNBD is a highly conserved structure (16). Combined data from PKA, Epac, and ion channel studies illustrate that CNBDs use a common mechanism to bind and sense cAMP

This work was supported by grant DFG Research Unit 2518 Dynlon of the Deutsche Forschungsgemeinschaft (projects P2 (KU 3092/2–1) and P7 (GO 1367/2–1)). F. S. is Head of Research and Development of the BIOLOG Life Science Institute, which sells the cAMP analogs that were used in this study.

¹ These authors contributed equally to this work.

² To whom correspondence should be addressed. E-mail: jana.kusch@med.uni-jena.de.

³ The abbreviations used are: Epac, exchange protein directly activated by cAMP; CN, cyclic nucleotide; HCN, hyperpolarization-activated and cyclic nucleotide–modulated; CNBD, cyclic nucleotide–binding domain; N⁶-Phe-cAMP, N⁶-phenyladenosine-3′,5′-cyclic monophosphate; N⁶-Bn-cAMP, N⁶-benzyladenosine-3′,5′-cyclic monophosphate; N⁶-Bnz-cAMP, N⁶-benzoyl-adenosine-3′,5′-cyclic monophosphate; RMSD, root mean square deviation.

HCN2 channel activation by N^6 -modified cAMP derivatives

(17). This similarity in the cAMP-binding mechanism of all cAMP-responsive proteins makes it challenging to target only one type of protein to discriminate between the different cAMP signaling pathways. However, discrimination between two proteins, PKA and Epac, could be realized and has been used widely (e.g. (18–22)). Although Epac tolerates 2'-OH modifications, modifications at the N^6 position of the purine ring are not accepted by this protein. However, such N^6 -modified derivatives are often potent agonists for PKA and can be used to exclude Epac activation (23). Notably, fully functional ion channels were often outside the scope when studying the selectivity of cAMP derivatives.

Besides a systematic study performed by Ng et al. (24) comparing binding and gating parameters for different cyclic nucleotides in whole HCN channels, the effect of purine ring modifications at cAMP on functional channels have so far not been tested systematically. To contribute to the urgent problem of developing cAMP analogs that can discriminate between the different cAMP-binding proteins, here we tested three N^6 -modified cAMP derivatives: N^6 -phenyladenosine-3',5'-cyclic monophosphate (N^6 -Phe-cAMP), N^6 -benzyladenosine-3',5'-cyclic monophosphate (N^6 -Bn-cAMP), and N^6 -benzoyladenosine-3',5'-cyclic monophosphate (N^6 -Bnz-cAMP), known to be activators of PKA, on functional HCN2 channels. We show that all of them are HCN2 channel activators with efficacies similar to cAMP but with different apparent affinities. Docking and mutagenesis experiments revealed the specific interactions between molecule and binding pocket that underlie those differences.

Results

Effect of the native ligand cAMP on HCN2 channels in oocyte macropatches

HCN2 channels produce slowly activating inward currents in response to hyperpolarizing voltage jumps. It has been shown that the cyclic nucleotide-binding domain has an inhibitory effect on channel gating that is relieved by cAMP binding (13). Such a relieving effect is reflected by an increase in current amplitude at a given voltage, by a shift of the voltage of half-maximum activation, $V_{1/2}$, to more depolarized values, and by an acceleration of activation kinetics and a deceleration of deactivation kinetics (14). This could also be observed here for mHCN2 channels expressed in *Xenopus laevis* oocytes (Fig. 1).

Fig. 1, A and B, shows representative current traces and the respective protocols used to determine channel activation at varying agonist concentrations at a given command voltage of -130 mV (Fig. 1A) and to determine channel activation at varying voltages at zero or saturating agonist concentrations (Fig. 1B), respectively. The cAMP concentration required to cause EC_{50} , a measure of the apparent affinity of the agonist to the receptor, was determined by approximating the Hill equation (Equation 1) to relative current amplitudes plotted against the cAMP concentration (Fig. 1C). Current amplitudes were obtained from tail currents at -100 mV following an activating pulse of -130 mV. The EC_{50} value was found to be 21.3 ± 3.3 nM, and the Hill coefficient was 1.1 ± 0.2 . From this, $10 \mu\text{M}$ was defined as the saturating concentration for all experiments.

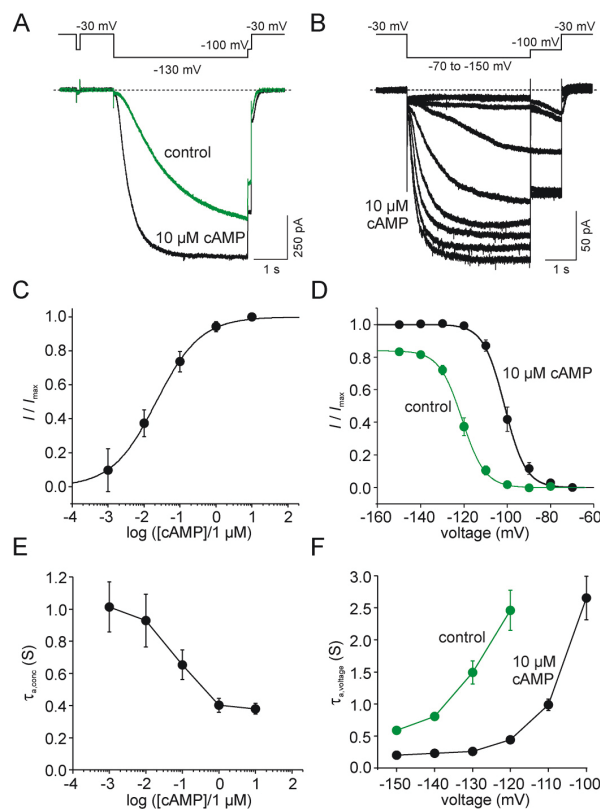


Figure 1. cAMP effects on steady-state and nonsteady state parameters of mHCN2 channel activation. A, protocol and representative current traces to study concentration-dependent gating at a fixed command voltage. Tail currents were obtained from a -100 mV pulse following an activating -130 mV pulse. B, protocol and representative current traces (exemplary of a saturating concentration of $10 \mu\text{M}$ cAMP) to study voltage-dependent gating at a fixed agonist concentration. A voltage family from -70 mV to -150 mV was applied with 10 -mV increments. Tail currents were obtained from a -100 mV pulse following the variable test pulse. C, concentration-response relationship for cAMP. Mean values for I/I_{max} were obtained from six to 13 recordings and plotted against the cAMP concentration. The Hill equation (Equation 1) was approximated to the data, yielding EC_{50} (21.3 ± 3.3 nM) and H (1.1 ± 0.2), respectively. D, steady-state activation relationship at zero and saturating [cAMP]. Mean values for relative current amplitudes (I/I_{max}) for zero cAMP ($n = 14$) and for saturating cAMP of $10 \mu\text{M}$ ($n = 13$) were plotted against the command voltage. The Boltzmann equation was approximated to the data, yielding $V_{1/2} = -117.8 \pm 1.5$ mV and a slope of $z\delta = 4.2 \pm 0.3$ for zero and $V_{1/2} = -97.9 \pm 1.9$ mV and a slope of $z\delta = 4.2 \pm 0.3$ for $10 \mu\text{M}$ cAMP. E, activation kinetics at different cAMP concentrations. Activation time constants were obtained from approximating a monoexponential function (Equation 3) to the current time courses, yielding $\tau_{a,\text{conc}}$. Mean values were obtained from three to seven recordings. The protocol used is shown in A. F, activation kinetics at different command voltages at zero and saturating [cAMP]. Activation time constants were obtained from approximating a monoexponential function (Equation 3) to the current time courses, yielding $\tau_{a,\text{voltage}}$. Mean values were obtained from 14 recordings for zero and 13 recordings for $10 \mu\text{M}$ cAMP. The protocol used is shown in B.

To estimate $V_{1/2}$, relative current amplitudes were plotted versus the command voltage. The Boltzmann equation was fitted to the data points of individual recordings, yielding $V_{1/2}$ under control conditions of -117.8 ± 1.5 mV and after application of $10 \mu\text{M}$ cAMP of -97.9 ± 1.9 mV. Thus, the steady-state activation was shifted by 19.9 ± 1.2 mV to more depolarized values because of cAMP binding (Fig. 1D). The slopes of

HCN2 channel activation by *N*⁶-modified cAMP derivatives

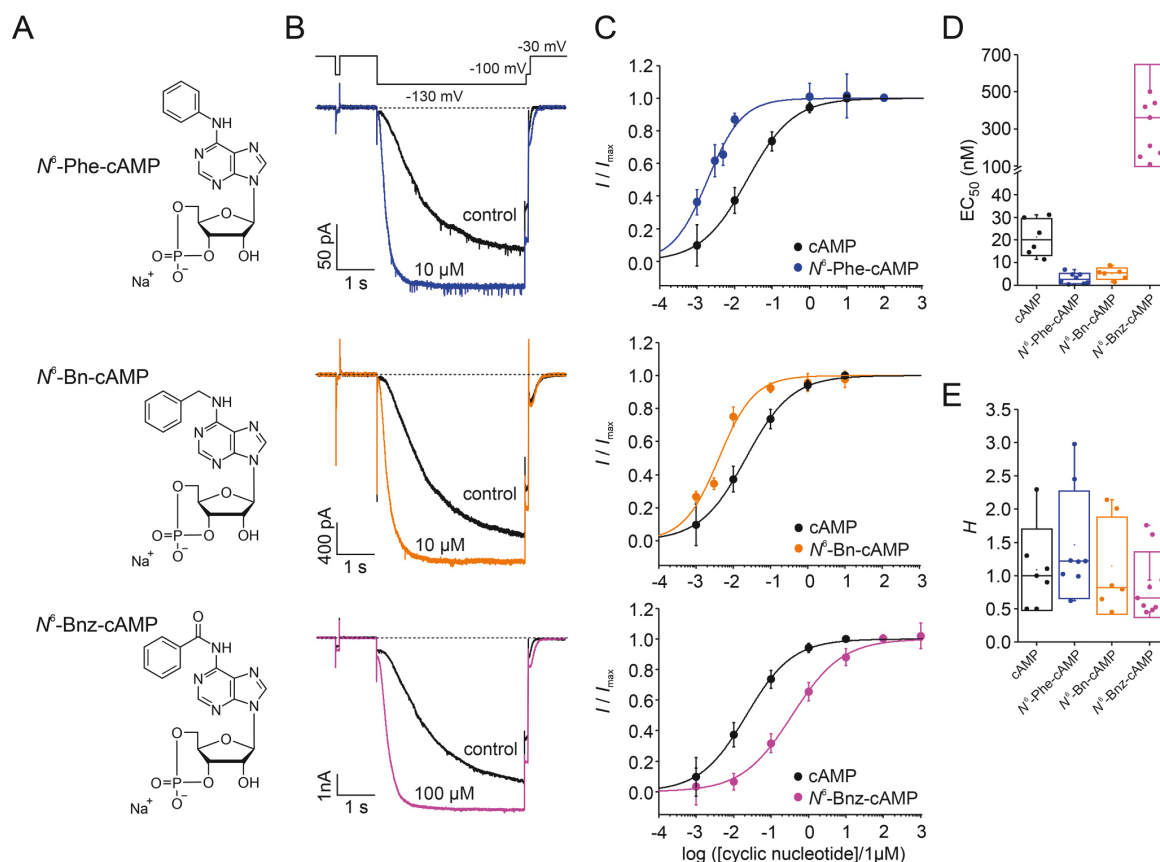


Figure 2. Structure and apparent affinities of the three tested derivatives. *A*, molecular formulas for three *N*⁶-modified derivatives known to activate PKA. *B*, representative current responses before and after application of the respective ligand. *Black traces* represent recordings in the absence of the ligand, and *colored traces* represent recordings during application of a saturating concentration of the respective ligand. Each ligand turned out to be an activator of heterologously expressed HCN2 channels. *C*, concentration–response relationships for the three derivatives (*colored symbols and fits*) compared with mHCN2 WT data (*black symbols and fits*). The Hill equation (Equation 1) was approximated to the data to obtain EC_{50} and *H*. *Error bars* indicate S.E. *D*, box plot of EC_{50} values obtained from *C*. *Filled circles* indicate individual recordings. *Error bars* indicate S.D. *E*, box plot of Hill coefficients obtained from *C*. *Filled circles* indicate individual recordings. *Error bars* indicate S.D.

the curves were similar with and without cAMP (4.2 ± 0.2 and 4.2 ± 0.3 mV/e⁻ fold change, respectively).

The time constant of activation, given as $\tau_{a,conc}$, decreased with increasing cAMP concentrations (Fig. 1E). Furthermore, the time constant of activation decreased with increasing hyperpolarizing voltages under control conditions without cAMP as well as after application of 10 μ M cAMP (Fig. 1F).

All three tested cAMP derivatives were able to promote HCN2 channel activation

We tested three cAMP derivatives, used as PKA activators, for their ability to activate HCN2 channels: *N*⁶-Phe-cAMP, *N*⁶-Bn-cAMP, and *N*⁶-Bnz-cAMP (Fig. 2A). In these derivatives, a phenyl group, a benzyl group, or a benzoyl group, respectively, is attached to the amino group in position 6 of the adenine moiety. All three *N*⁶-modified derivatives have been shown to be site-selective for site A of both PKA type I and II (27–29).

To test whether the three cAMP derivatives are able to modulate HCN2 channels, channel activation was monitored at

zero cAMP and in the presence of different agonist concentrations. Representative current recordings for zero cAMP and saturation are shown in Fig. 2B. All three derivatives caused an increase in current amplitude and acceleration of the activation kinetics, as known for cAMP (Fig. 1).

cAMP modifications at position *N*⁶ affected the apparent affinity of the agonist for HCN2

To study the concentration dependence in more detail, concentration–response relationships were recorded for each derivative (Fig. 2C) and compared with the relationship for cAMP. For all cases, the Hill equation (Equation 1) was approximated to the relative currents of each individual recording. The results for EC_{50} and the Hill coefficient, *H*, are summarized in Fig. 2, D and E, respectively. The nature of the *N*⁶ modification had a major influence on the apparent affinity ($EC_{50,Phe} = 3.02 \pm 0.83$ nM, $EC_{50,Bn} = 5.12 \pm 1.03$ nM, $EC_{50,Bnz} = 374 \pm 91.3$ nM) (Fig. 2D). Adding a phenyl or benzyl group to position *N*⁶ shifted the EC_{50} value to lower concentrations compared with native cAMP, causing an apparent affinity one order of magnitude higher than for

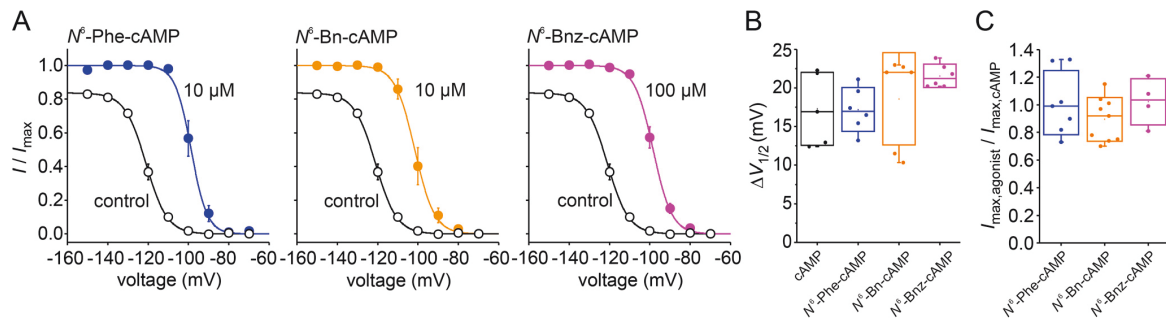
HCN2 channel activation by N^6 -modified cAMP derivatives

Figure 3. Efficacy of the three tested derivatives. *A*, steady-state activation relationships at zero and saturating agonist concentrations. The graphs show the relationships of the tested derivatives at saturating concentrations and in the absence of any ligand. Except for N^6 -Bnz-cAMP, for which $100 \mu\text{M}$ was used as a saturating concentration, the concentration was $10 \mu\text{M}$ for saturation. I/I_{max} are mean values obtained from three to six recordings. The Boltzmann equation was approximated to the relative current values to yield $V_{1/2}$. The results of the fits were as follows: $-97.5 \pm 2.0 \text{ mV}$ for N^6 -Phe-cAMP, $-104.9 \pm 2.7 \text{ mV}$ for N^6 -Bn-cAMP, and $-98.3 \pm 1.4 \text{ mV}$ for N^6 -Bnz-cAMP. The slopes were 3.4 ± 0.2 , 4.1 ± 0.2 , and 4.4 ± 0.3 , respectively. *B* and *C*, comparison of the efficiency of the three derivatives compared with cAMP in two measures. *B*, box plots of $I_{\text{max,agonist}}/I_{\text{max,cAMP}}$. *Filled circles* indicate individual recordings. *Error bars* indicate S.D. *C*, box plots of $\Delta V_{1/2}$. *Filled circles* indicate individual recordings. *Error bars* indicate S.D.

cAMP. However, adding a benzoyl group shifted the EC_{50} value to higher concentrations, indicating an apparent affinity one order of magnitude lower than for cAMP (Fig. 2*D*). The Hill coefficients for the derivatives were not significantly different from that for cAMP: $H_{\text{Phe}} = 1.4 \pm 0.3$, $H_{\text{Bn}} = 1.2 \pm 0.3$, and $H_{\text{Bnz}} = 0.9 \pm 0.2$ (Student's *t* test, $p = 0.05$) (Fig. 2*E*).

None of the cAMP modifications tested changed the efficacy

We used two measures to compare the efficacy of the cAMP derivatives with that of native cAMP: the maximal agonist-induced voltage shift ($\Delta V_{1/2,\text{max}}$) and the maximal fractional increase in tail current amplitude ($I_{\text{max,agonist}}/I_{\text{max,cAMP}}$) observed with saturating agonist concentrations (Fig. 3) (30).

To yield $\Delta V_{1/2,\text{max}}$, $V_{1/2}$ values before and after application of saturating agonist concentrations were estimated by fitting the Boltzmann equation to steady-state activation relationships (Fig. 3*A*). There was no difference between $\Delta V_{1/2,\text{max}}$ of cAMP and either one of the tested derivatives ($16.1 \pm 2.3 \text{ mV}$ for cAMP, $18.3 \pm 1.4 \text{ mV}$ for N^6 -Phe-cAMP, $18.6 \pm 2.4 \text{ mV}$ for N^6 -Bn-cAMP, and $22.3 \pm 1.1 \text{ mV}$ for N^6 -Bnz-cAMP) (Fig. 3*B*).

To yield the maximal fractional increase in tail current amplitude, we determined the tail current amplitudes at a test pulse of -100 mV , following a hyperpolarizing pulse of -130 mV , for each cAMP derivative at saturating concentrations and related that to the maximum tail current amplitude of $10 \mu\text{M}$ cAMP in the same patch. All derivatives caused a fractional current of around 1.0, suggesting that they all cause a similar current increase as cAMP (Fig. 3*C*). The values were 1.02 ± 0.14 for N^6 -Phe-cAMP, 0.98 ± 0.05 for N^6 -Bn-cAMP, and 1.00 ± 0.02 for N^6 -Bnz-cAMP (Fig. 3*B*). Thus, both measures led us to conclude that the modifications performed in the tested cAMP derivatives did not affect the efficacy of the agonists.

None of the cAMP modifications tested changed activation kinetics

Fig. 4 summarizes the results regarding activation kinetics upon hyperpolarizing voltage jumps. First, we studied the concentration dependence of these kinetics. For this purpose, we applied the agonists at different concentrations, covering a wide range, and recorded the currents at a nearly saturating

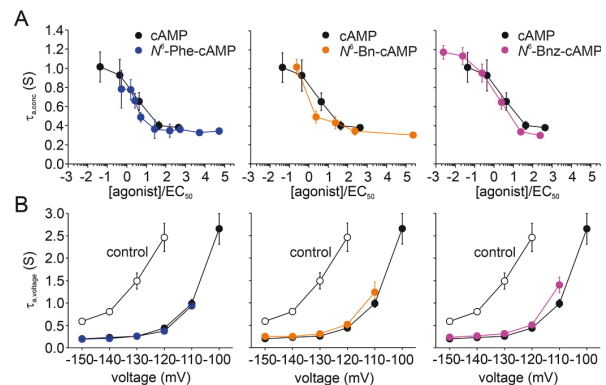


Figure 4. Activation kinetics for the tested cAMP derivatives. *A*, activation kinetics in dependence on agonist concentration at a given command voltage of -130 mV . $\tau_{\text{a,conc}}$ values, obtained from approximating a monoexponential equation (Equation 3) to the current time courses, are plotted against the cAMP concentration. Mean values were obtained from three to nine recordings. *Black symbols* in each plot illustrate the case for cAMP. The protocol used is shown in Fig. 1*A*. *B*, activation kinetics in dependence on command voltage at a saturating agonist concentration. $\tau_{\text{a,voltage}}$ values were plotted against the command voltage. *Open symbols* represent recordings in the absence of ligands and *filled symbols* in the presence of saturating agonist concentrations. Mean values were obtained from three to 19 recordings. The protocol used is shown in Fig. 1*B*.

voltage of -130 mV . Kinetics were quantified by $\tau_{\text{a,conc}}$ obtained from approximating a monoexponential equation (Equation 3) to the current time courses. The values were plotted against the normalized concentration, which is the ratio of applied concentration and apparent affinity $[\text{agonist}]/\text{EC}_{50}$ (Fig. 4*A*). Such normalization was required because large differences in the EC_{50} values for the derivatives and cAMP did not allow comparison at absolute concentrations. As a result, the concentration dependence of $\tau_{\text{a,conc}}$ for derivatives could be superimposed with that of cAMP.

Second, we studied the voltage dependence of activation kinetics. For this purpose, we applied a family of hyperpolarizing voltages at saturating agonist concentrations and in the absence of agonists, respectively. The results are shown in Fig. 4*B*. Analogous to the concentration dependence, the voltage dependence of $\tau_{\text{a,voltage}}$ could be superimposed with that of

HCN2 channel activation by N^6 -modified cAMP derivatives

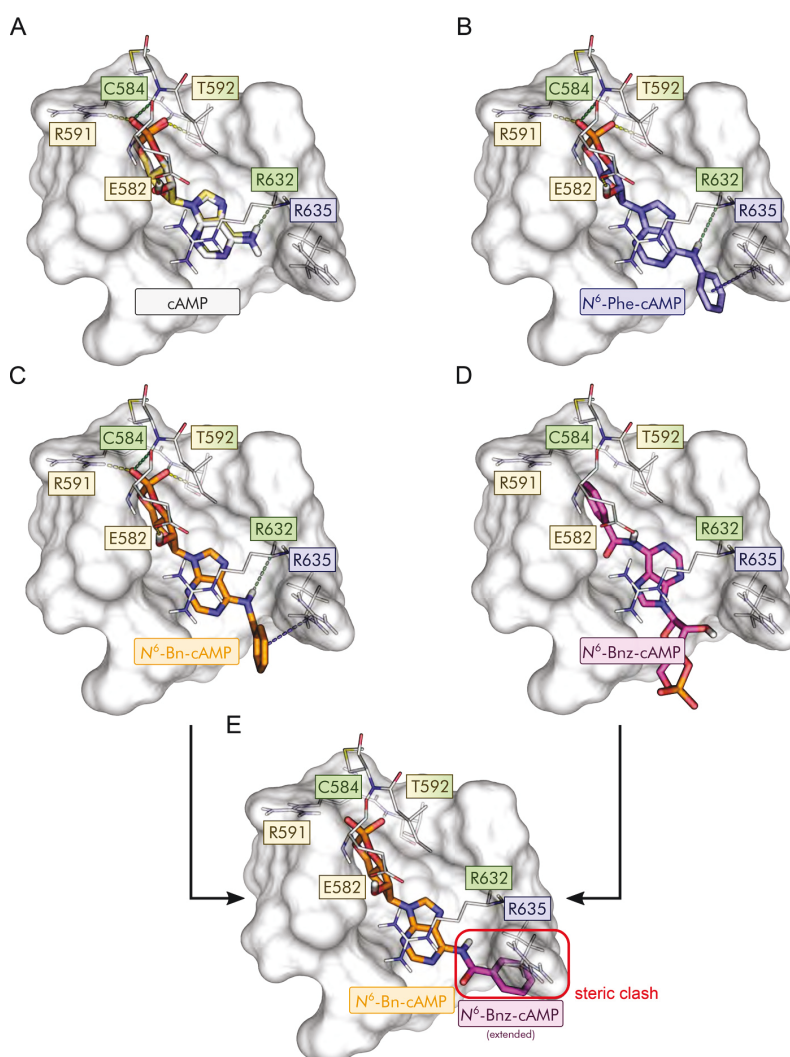


Figure 5. Predicted binding modes of the tested cAMP derivatives. A–D, predicted binding modes of cAMP (A, gray; crystallographic pose (PDB code 1Q50 (25) shown in gold), N^6 -Phe-cAMP (B, blue), N^6 -Bn-cAMP (C, orange), and N^6 -Bnz-cAMP (D, magenta). In A–C, hydrogen bonds to and from protein side chains are depicted as dashed yellow lines, hydrogen bonds to and from the protein main chain as dashed green lines, and cation– π interactions as dashed blue lines. E, predicted binding mode of the cAMP fragment in N^6 -Bn-cAMP extended with a benzoyl group as found in N^6 -Bnz-cAMP. The steric clash with Arg-635 resulting from this extension is highlighted (red box), explaining the inverted binding mode of N^6 -Bnz-cAMP shown in D.

cAMP. From these results, we conclude that all derivatives had a similar accelerating effect as the native ligand.

Molecular docking suggests a pivotal role of Arg-635 for determining the binding behavior of N^6 -modified derivatives

To determine the structural basis of the higher apparent affinities of N^6 -Phe-cAMP and N^6 -Bn-cAMP and the lower apparent affinity of N^6 -Bnz-cAMP, we predicted and compared the binding modes of cAMP, N^6 -Phe-cAMP, N^6 -Bn-cAMP, and N^6 -Bnz-cAMP in mHCN2] by molecular docking (Fig. 5). The selected docking protocol was able to reproduce the crystallographically determined binding mode of cAMP with a heavy-atom RMSD of 0.60 Å (Fig. 5A), indicating that the binding modes of the N^6 -substituted derivatives can be reliably predicted. The predicted binding modes of the cAMP moiety of

N^6 -Phe-cAMP and N^6 -Bn-cAMP deviate only marginally from the native binding mode of cAMP (heavy-atom RMSDs of 0.73 Å and 0.91 Å, respectively; Fig. 5, B and C). Docking scores (XP GScores) for all three compounds were in a similar range (cAMP, -13.53 kcal mol $^{-1}$; N^6 -Phe-cAMP, -10.20 kcal mol $^{-1}$; N^6 -Bn-cAMP, -11.54 kcal mol $^{-1}$).

Compared with cAMP, the predicted binding modes of N^6 -Phe-cAMP and N^6 -Bn-cAMP reveal stabilizing cation– π and/or π – π stacking interactions between the additional phenyl ring and the side chain of Arg-635 (Fig. 5, A–C), which can explain the higher apparent affinities of these two derivatives. This interaction may be more stable in the case of N^6 -Phe-cAMP because of the more restricted conformational freedom of the N^6 -ring bonds compared with the less restricted N^6 -benzyl carbon bond in N^6 -Bn-cAMP, which

HCN2 channel activation by N^6 -modified cAMP derivatives

can explain the higher apparent affinity of N^6 -Phe-cAMP compared with N^6 -Bn-cAMP.

In contrast to the N^6 -Phe and N^6 -Bn derivatives, the orientation of the cAMP moiety in the predicted binding mode of N^6 -Bnz-cAMP did not match the crystallographic pose (RMSD of 8.66 Å) but was instead inverted (Fig. 5D) and associated with a markedly lower docking score (-6.11 kcal mol $^{-1}$). To determine why N^6 -Bnz-cAMP cannot adopt the binding mode favorable for cAMP and the other N^6 -substituted derivatives, we used the predicted binding mode of N^6 -Bn-cAMP as a template and replaced the benzyl moiety with a benzoyl moiety. Because of the planarity of the amide group in N^6 -Bnz-cAMP, the phenyl ring is then forced into an orientation in which it inevitably clashes with Arg-635 (Fig. 5E), resulting in a strongly unfavorable interaction.

Substituting Arg-635 at the entrance of the binding pocket prevented the higher apparent affinity of N^6 -Phe-cAMP

The arginine at position 635 was substituted with either alanine for neutralization (R635A) or glutamate for charge reversal (R635E). If Arg-635 is indeed involved in the higher apparent affinity of N^6 -Phe-cAMP and N^6 -Bn-cAMP, then the mutations should cause an apparent affinity that resembles that of cAMP to WT channels. To test this, the concentration-activation relationship was tested for N^6 -Phe-cAMP in both R635A and R635E. Both constructs, R635A and R635E, formed functional channels and expressed in high densities in *Xenopus* oocytes. First, we tested whether those constructs are still reactive to 10 μ M cAMP. As for WT channels, we found a substantial current increase and an acceleration of the activation speed (Fig. 6A). When studying the concentration-response relationships for N^6 -Phe-cAMP, we found, as expected, that the EC_{50} values for both constructs were increased (Fig. 6, B and C), 27.7 ± 5.5 nM for R635A and 33.9 ± 5.3 nM for R635E. They are not different from the EC_{50} values found for cAMP (21.3 ± 3.3 nM) in WT channels.

Discussion

PKA agonists activate HCN2 channels

Here we performed a comparative study of three cAMP derivatives, N^6 -Phe-cAMP, N^6 -Bn-cAMP, and N^6 -Bnz-cAMP, earlier described as activators of PKA, to investigate their effects on HCN2 channels. In the last decades, these derivatives have been shown to be useful tools when discrimination between activation of PKA and Epac was required (18–22). Interestingly, all three tested derivatives turned out to also be activators for HCN2 channels heterologously expressed in *Xenopus laevis* oocytes by promoting HCN2 channel gating in the presence of a primary hyperpolarizing voltage stimulus.

To study the agonistic effect in more detail, we determined the concentration-activation relationship and quantified EC_{50} as a measure of the apparent affinity and the Hill coefficient of activation. The apparent affinity was considerably affected by the type of N^6 modification: Substituting one of the N^6 -bound hydrogens with a phenyl or benzyl group dramatically increased the apparent affinity (seven times and four times higher than cAMP, respectively), whereas a benzoyl ring decreased it (18 times lower than cAMP). However, because EC_{50} is a function of both

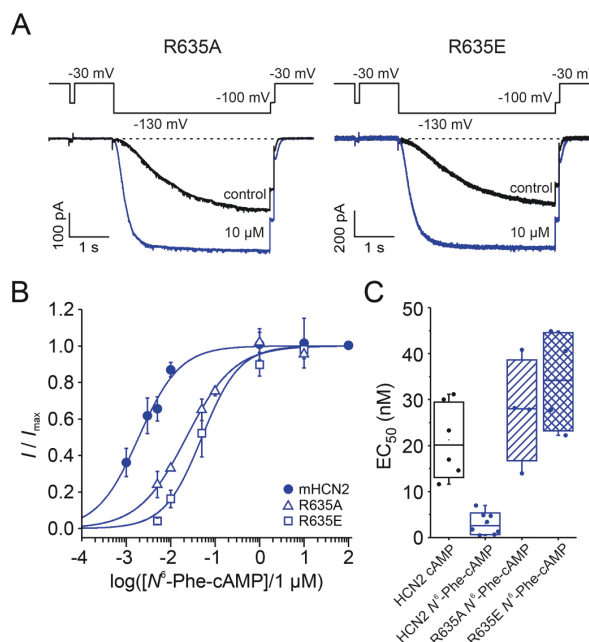


Figure 6. Response to N^6 -Phe-cAMP after charge neutralization or reversal at position Arg-635 in the CNBD. A, representative current traces for R635A and R635E at zero and saturating cAMP. B, concentration-activation relationships for mHCN2, R635A, and R635E. The Hill equation (Equation 1) was approximated to the relative currents to obtain EC_{50} and the Hill coefficient. C, box plot of EC_{50} values. Filled circles indicate individual recordings. Error bars indicate S.D. EC_{50} values obtained with N^6 -Phe-cAMP for R635A and R635E, respectively, are not different from the EC_{50} value obtained for cAMP and mHCN2 (Student's *t* test).

binding affinity and efficacy, this parameter alone does not allow any interpretation of the mechanisms behind the agonist-dependent differences. The Hill coefficients for all agonists were similar to the Hill coefficient for cAMP. Even though this parameter is often used to describe the magnitude of cooperativity in ligand binding to allosteric systems, we hesitate to interpret those numbers in detail because, when derived from electrophysiological measurements, the Hill coefficient is just an empirical description without physical meaning because it depends on both binding and gating (31).

To address the question of why the tested agonists show those clear differences in their apparent affinities, we determined the efficacy, the ability of a ligand to elicit a response upon binding to a receptor (32), compared with unmodified cAMP. Two parameters were quantified: the maximum extent of shifting the steady-state activation to more depolarized voltages, $\Delta V_{1/2}$, and the ability of increasing the current in channels maximally activated by voltage. For both parameters, none of the three agonists differed from cAMP, indicating a similar efficacy of the native ligand and the three derivatives. Following a definition formulated by Colquhoun (31), according to which efficacy is determined by the sum of transduction events that follow the initial binding reaction, in HCN2 channels, those transduction events are similar whether matter they bind N^6 -Phe-cAMP, N^6 -Bn-cAMP, N^6 -Bnz-cAMP, or the native ligand cAMP. Provided that cAMP is a full agonist in HCN2

HCN2 channel activation by N^6 -modified cAMP derivatives

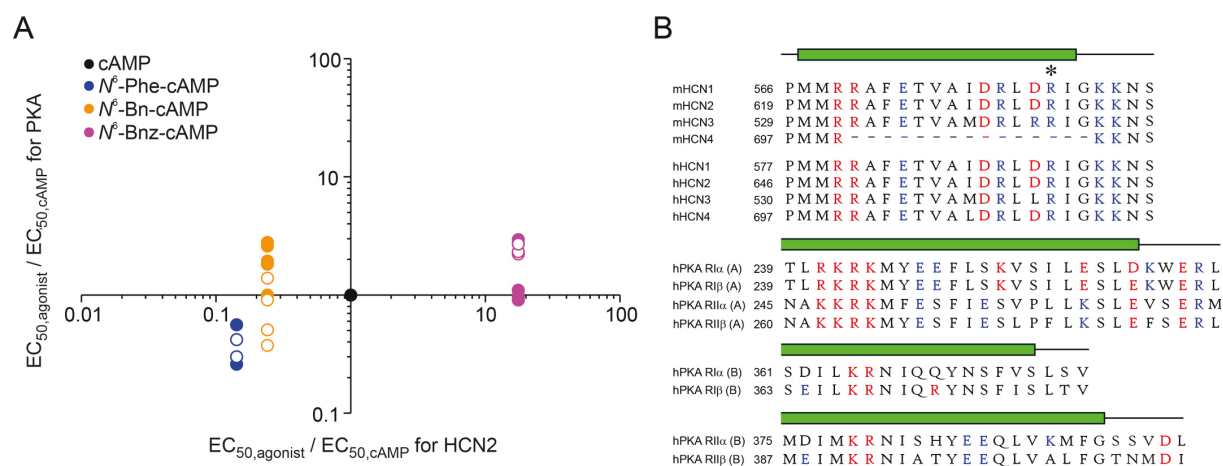


Figure 7. Comparison of relative apparent affinity for PKA and HCN2. *A*, apparent affinities for cAMP derivatives were related to apparent affinities for unmodified cAMP. Among the tested derivatives, the best discrimination can be realized with N^6 -Bnz-cAMP. Data for PKA were obtained from Refs. 29, 34–36. Filled symbols represent PKAI and empty symbols PKAII. *B*, sequence alignment comparing the distal parts of the C-helix sequences of mouse and human HCN isoforms with those from A and B sites in regulatory subunits of human PKAI and II. Alignments and structure predictions (α -helix C marked in green) are shown as proposed by Berman *et al.* (39), who used a structure and transformation method (SAT). Prediction of the HCN C-helix follows a suggestion from Lee and MacKinnon (40). The asterisks mark Arg-635 mutated herein. Positively charged residues are shown in red and negatively charged in blue.

channels (33), those data further suggest consideration of all three N^6 -modified derivatives also as full agonists.

In addition to those steady-state parameters, we quantified the activation kinetics of the macroscopic current time course following a hyperpolarizing voltage jump. The kinetics of a macroscopic current reflect the kinetics of the underlying channels and are determined by the rate at which a new equilibrium occurs after a perturbation. We compared the effect of the N^6 -modified derivatives under two conditions: at varying voltages in the presence of a saturating agonist concentration and at varying concentrations at a saturating voltage. For both conditions, the activation time constants for each of the three derivatives were not different from the activation time constants for cAMP. This supports the interpretation of efficacy that the transduction events that follow the initial binding reaction are similar for all tested agonists. Together, those data led us to suggest that the N^6 -modified derivatives tested here are differently potent in HCN2 channels because they differ in their ability to bind to the HCN2 binding sites.

Arginine 635 in the CNBD affects apparent affinities

We performed docking experiments to identify molecular interactions between N^6 derivatives and binding site residues that might underlie the observed differences in apparent affinities. The binding modes of N^6 -Phe-cAMP and N^6 -Bn-cAMP are stabilized by a cation– π and/or π – π stacking interaction between the additional phenyl moiety and the side chain of Arg-635. The higher apparent affinity of N^6 -Phe-cAMP compared with N^6 -Bn-cAMP is likely explained by the more restricted conformational freedom of the amino-phenyl side chain in the former case, which leads to a reduced loss in configurational entropy upon binding. In contrast, the benzoyl moiety in N^6 -Bnz-cAMP enforces coplanarity of the two aryl systems, which would lead to a clash between the phenyl ring and Arg-635 if the cAMP moiety of N^6 -Bnz-cAMP adopted the usual binding mode in the cAMP binding site. Based on these

results, we hypothesized that the cation– π and/or π – π stacking interaction between the phenyl ring and Arg-635 is the main cause of the differential apparent affinities of N^6 -Phe-cAMP and N^6 -Bn-cAMP versus N^6 -Bnz-cAMP.

To test this hypothesis, we either neutralized or reversed the charge at position 635 by replacing the arginine with either alanine (R635A) or glutamate (R635E). Neutralization resulted in an apparent affinity of N^6 -Phe-cAMP similar to the apparent affinity of cAMP for mHCN2 WT channels. From this result, we suggest that now N^6 -Phe-cAMP interacts with R635A in a similar manner as cAMP with the mHCN2 WT. Reversing the charge by replacing the arginine with a glutamate led to an apparent affinity of N^6 -Phe-cAMP slightly lower than that of cAMP for mHCN2 WT channels. This might be caused by a repulsive force between the aromatic ring of N^6 -Phe-cAMP and the negative side chain of the newly introduced glutamate. Together, these data support the hypothesis resulting from the docking experiments, proposing a stabilizing or destabilizing interaction between N^6 modifications and Arg-635. A role of Arg-635 for binding and selectivity was discussed earlier by Zhou and Siegelbaum (30). They described this residue as being involved in discriminating between cAMP and cGMP without contributing to efficacy.

Usability of N^6 modifications in cAMP to discriminate between PKA and HCN channels

What can we learn from the results presented here regarding discrimination between PKA and HCN channel activation? To compare the apparent affinities of the derivatives at HCN2 channels with that at PKA, we plotted relative EC_{50} values for PKAI and PKAII obtained from the literature ($EC_{50} / EC_{50,cAMP}$) versus relative EC_{50} values for HCN2 obtained here (Fig. 7) (29, 34–36).

The very high apparent affinity of N^6 -Phe-cAMP as well as the very low one of N^6 -Bnz-cAMP for HCN2 channels brings the N^6 position into focus for the search for discriminating ago-

HCN2 channel activation by N^6 -modified cAMP derivatives

nists. However, in the case of N^6 -Phe-cAMP, the apparent affinity for PKA I and PKAII has also been shown to be relatively high, with a factor of ~ 2.4 and ~ 2.8 higher for PKA I and PKAII, respectively, compared with cAMP (29). Thus, the discrimination between HCN and PKA is not very strong.

N^6 -Bn-cAMP was also very potent for HCN2 channels, whereas it was in most cases less potent than cAMP in PKA I (29, 34, 36). The differences are not very high; however, a benzyl ring in the N^6 position could be at least one of several modifications in the cAMP molecule to yield discrimination between HCN2 and PKA I. A discrimination to PKAII is questionable because the apparent affinity has been shown to be either unchanged (29) or two times better than for cAMP (36).

The derivative with the best discrimination ability found here is N^6 -Bnz-cAMP. It is a potent agonist of PKA I, with a similar apparent affinity as cAMP (29, 35), but a relatively poor agonist of HCN2, with an apparent affinity 18 times worse than cAMP. This is a very important finding because N^6 -Bnz-cAMP is a well-established tool for studying cAMP-mediated signal transduction in a wide range of cellular processes (e.g. 23, 37, 38). Discrimination between PKAII and HCN2 is expected to be less strong because N^6 -Bnz-cAMP is two times less potent to PKAII than cAMP (29, 35).

The C-helix of CNBDs (green bars in Fig. 7) acts as a hydrophobic lid covering the base moiety of the cyclic nucleotide and thereby capping the binding pocket (39). Is the positively charged Arg-635 at the C-terminal end of this C-helix a unique feature of HCN channels and can it be used for further discrimination approaches? The alignment in Fig. 7 shows that, with the exception of mHCN4, Arg-635 is highly conserved in mouse and human HCN channel isoforms, which allows speculation regarding an important role in ligand binding. A comparison with A and B sites of PKA I and PKAII shows that there is no such arginine at the C-terminal ends of the PKA C-helices. However, there is an interesting pattern of charged amino acids right before R635 that reads negative–positive–X–negative–positive and can be found in most HCN channels. This pattern is also present in A sites of PKA RI and, slightly modified, in PKA RII, although not as part of the C-helix sequence but in the consecutive loop directly behind. Such a pattern could not be found in any B sites of PKA presented here. Another pattern in HCN CNBDs is a strong positive spot at the N-terminal part of the C-helix followed by negative residues two positions C-terminal from that. Again this pattern can be found in A sites of PKA RI and additionally in RII but not in B sites. Those differences might be responsible for N^6 -Phe-cAMP, N^6 -Bn-cAMP, and N^6 -Bnz-cAMP preferring A sites over B sites (34, 41).

Because the C-helix is the most variable structure of CNBDs (39), residues in that helix are potential key players in determining protein-specific ligand selectivity. After comparing sequences of HCN CNBDs with A sites and B sites, we speculate that there is a greater similarity between binding to HCN CNBDs and A sites than between HCN CNBDs and B sites. Thus, to identify further cAMP derivatives with a higher degree of discrimination between HCN and PKA than shown here, we propose to focus on the unique characteristics of binding to B sites.

In summary, our analysis shows that possible cross-activation of CN-modulated ion channels should be taken into account when CN-dependent cellular processes are investigated with N^6 -modified derivatives. To overcome this problem, we further suggest that, even with a missing all-or-nothing principle in protein activation by cyclic nucleotides, carefully performed concentration–activation relationships can be used to identify a concentration range in which selective protein activation is possible.

Experimental procedures

X. laevis oocytes as a heterologous expression system

Surgical removal of oocytes was performed under anesthesia (0.3% tricaine, MS-222, Pharmaq Ltd., Fordingbridge, UK) from adult females of the South African claw frog *X. laevis*. The oocytes were treated with collagenase A (3 mg/ml, Roche) for 105 min in Ca^{2+} -free Barth's solution containing 82.5 mM NaCl, 2 mM KCl, 1 mM $MgCl_2$, and 5 mM Hepes (pH 7.5). After this procedure, oocytes of stages IV and V were manually dissected and injected with cRNA encoding either mHCN2 channels of *Mus musculus* (NM_008226) or the mHCN2 mutants R635A and R635E, respectively. After injection with cRNA, the oocytes were cultured at 18 °C for 2–6 days in Barth's solution containing 84 mM NaCl, 1 mM KCl, 2.4 mM $NaHCO_3$, 0.82 mM $MgSO_4$, 0.41 mM $CaCl_2$, 0.33 mM $Ca(NO_3)_2$, and 7.5 mM Tris (pH 7.4). The procedures had approval from the authorized animal ethical committee of the Friedrich Schiller University Jena. The methods were carried out in accordance with the approved guidelines. Oocytes harvested in our own laboratory were complemented with ready-to-use oocytes purchased from Ecocyte Bioscience (Dortmund, Germany).

Molecular biology

The mouse HCN2 (accession number NM_008226) channel and the modified subunit variants were subcloned in front of the T7 promoter of pGEM-HCN2. The point mutations R635E and R635A were introduced by overlapping PCR using the following primers: PCR 1 forward, 5'-TCATACTCGCGCGGG-CCCCAAGGTTTC-3' (restriction site *MauBI*, outer primer); PCR 1 reverse, 5'-CTTCTTGCCTATCATCTAGCCGGTCAATAGC-3' (R635E) or 5'-CTTCTTGCCTATGGCATCTAGCCGGTCAATAGC-3' (R635A); PCR 2 forward, 5'-GCTATTGACCGGCTAGATGAGATAGGCAAGAAG-3' (R635E) or 5'-GCTATTGACCGGCTAGATGCCATAGGCAAGAAG-3' (R635A); PCR 2 reverse, 5'-AGCAGGGTTGGTCTAGAGTCAAGTTGGAAGAG-3' (restriction site *XbaI*, outer primer). In a third PCR, the PCR products of PCR 1 and 2 were used as a template using the outer primers. The resulting fragment was subcloned into the pGEM-HCN2 subunit. Correctness of the sequences was confirmed by restriction analysis and sequencing (Microsynth SEQLAB). cRNAs were prepared using the mMMESSAGE mMACHINE T7 Kit (Ambion).

Electrophysiological experiments

Macroscopic currents were recorded using the patch clamp technique in the inside-out configuration. All measurements were started after a delay of 3.5 min to minimize rundown phe-

HCN2 channel activation by N^6 -modified cAMP derivatives

nomena. Patch pipettes were pulled from quartz tubing whose outer and inner diameters were 1.0 and 0.7 mm (Vitrocom), respectively, using a laser puller (P-2000, Sutter Instrument). The pipette resistance was 1.2–2.1 megaohm. The bath solution contained 100 mM KCl, 10 mM EGTA, and 10 mM Hepes (pH 7.2), and the pipette solution contained 120 mM KCl, 10 mM Hepes, and 1.0 mM CaCl_2 (pH 7.2). For parts of the experiments, different concentrations of cAMP, N^6 -Phe-cAMP, N^6 -Bn-cAMP, or N^6 -Bnz-cAMP (BIOLOG LSI, Bremen, Germany) were applied with the bath solution. A saturating concentration of 10 μM cAMP was applied to each patch to define the maximum current amplitude. An HEKA EPC 10 USB amplifier (Harvard Apparatus) was used for current recording. Pulsing and data recording were controlled by the Patchmaster software (Harvard Apparatus). The sampling rate was 5 kHz. The holding potential was generally -30 mV.

Quantification and statistical analysis

Concentration–activation relationships were analyzed by approximating the Hill equation to each individual recording using the OriginPro 9.0G software.

$$I/I_{\max} = 1/[1 + (\text{EC}_{50}/[\text{agonist}])^H] \quad (\text{Eq. 1})$$

I is the actual current amplitude at a given agonist concentration, I_{\max} the maximal current amplitude at a saturating concentration of 10 μM cAMP, EC_{50} the concentration of half-maximum activation, and H the Hill coefficient. Current amplitudes were generally obtained from tail currents at a -100 mV pulse following an activating -130 mV pulse and corrected for leak currents obtained from a short -100 mV pulse at the beginning of each trace. Values for EC_{50} and H were yielded for each individual recording and averaged. Steady-state activation curves were analyzed by fitting the Boltzmann equation to each individual recording using the OriginPro 9.0G software.

$$I/I_{\max} = 1/[1 + \exp(z\delta F(V - V_{1/2})/RT)] \quad (\text{Eq. 2})$$

$V_{1/2}$ is the voltage of half-maximum activation and $z\delta$ the effective gating charge. F , R , and T are the Faraday constant, the molar gas constant, and the temperature in Kelvin, respectively. I is the actual current amplitude and I_{\max} the maximum current amplitude at the saturating hyperpolarizing voltage of -150 mV specified for each patch.

The time courses of current activation were fitted with a single exponential starting after an initial delay.

$$I(t) = A \times \exp[-t/\tau] \quad (\text{Eq. 3})$$

A is the amplitude, t the time, and τ the time constant for activation.

Experimental data are given as mean \pm S.E. Statistical analysis was performed by an unpaired Student's t test. A value of $p < 0.05$ was accepted as statistically significant.

Molecular modeling and docking experiments

Three-dimensional structures of the anionic species of cAMP, N^6 -Phe-cAMP, N^6 -Bn-cAMP, and N^6 -Bnz-cAMP were generated in Maestro (Schrödinger, LLC) and prepared using

the LigPrep workflow (Schrödinger, LLC). One monomer (chain A) of the crystal structure of cAMP-bound mHCN2J (25) was prepared using the Protein Preparation Wizard in Maestro; the program was evoked to cap the termini with acetyl and N -methyl amide moieties and to convert selenomethionine residues back to methionine residues. The potential grid for docking was centered on the cocrystallized cAMP, and the ligand length was set to 16 Å or less. The dimensions of the inner box that restricts the region in which the diameter midpoint of each ligand can be located were set to $10 \times 10 \times 10$ Å. No constraints, rotatable OH- and SH-groups, or excluded volumes were defined. Molecular docking was performed with Glide (Schrödinger, LLC) using the Extra Precision (26) mode with default options.

Author contributions—T. L., M. B., F. Spiegelhalter, and J. K. data curation; T. L., M. B., F. Spiegelhalter, and J. K. formal analysis; T. L., M. B., F. Spiegelhalter, F. Schwede, T. S., H. G., and J. K. methodology; T. L., M. B., F. Spiegelhalter, F. Schwede, T. S., and H. G. writing-review and editing; M. B. and J. K. validation; F. Schwede, H. G., and J. K. conceptualization; F. Schwede and T. S. resources; H. G. and J. K. supervision; H. G. and J. K. funding acquisition; J. K. writing-original draft; J. K. project administration.

Acknowledgments—We thank Sandra Bernhardt, Uta Enke, Andrea Kolchmeier, Claudia Ranke, and Karin Schoknecht for excellent technical assistance. We are grateful for computational support and infrastructure provided by the Zentrum für Informations- und Medientechnologie at the Heinrich Heine University Düsseldorf and the computing time provided by the John von Neumann Institute for Computing (to H. G.) on the supercomputer JUWELS at the Jülich Supercomputing Center (user IDs HKF7 and HDD17).

References

- Sutherland, E. W., and Rall, T. W. (1958) Fractionation and characterization of a cyclic adenine ribonucleotide formed by tissue particles. *J. Biol. Chem.* **232**, 1077–1091 Medline
- Beavo, J. A., and Brunton, L. L. (2002) Cyclic nucleotide research: still expanding after half a century. *Nat. Rev. Mol. Cell Biol.* **3**, 710–718 CrossRef Medline
- Schindler, R. F., and Brand, T. (2016) The Popeye domain containing protein family: a novel class of cAMP effectors with important functions in multiple tissues. *Prog. Biophys. Mol. Biol.* **120**, 28–36 CrossRef Medline
- Craven, K. B., and Zagotta, W. N. (2006) CNG and HCN channels: two peas, one pod. *Annu. Rev. Physiol.* **68**, 375–401 CrossRef Medline
- DiFrancesco, D. (1986) Characterization of single pacemaker channels in cardiac sino-atrial node cells. *Nature* **324**, 470–473 CrossRef Medline
- DiFrancesco, D. (1999) Dual allosteric modulation of pacemaker (f) channels by cAMP and voltage in rabbit SA node. *J. Physiol.* **515**, 367–376
- Robinson, R. B., and Siegelbaum, S. A. (2003) Hyperpolarization-activated cation currents: from molecules to physiological function. *Annu. Rev. Physiol.* **65**, 453–480 CrossRef Medline
- Santoro, B., and Tibbs, G. R. (1999) The HCN gene family: molecular basis of the hyperpolarization-activated pacemaker channels. *Ann. N.Y. Acad. Sci.* **868**, 741–764 CrossRef Medline
- Wang, J., Chen, S., Nolan, M. F., and Siegelbaum, S. A. (2002) Activity-dependent regulation of HCN pacemaker channels by cyclic AMP: signaling through dynamic allosteric coupling. *Neuron* **36**, 451–461 CrossRef Medline
- Wang, J., Chen, S., and Siegelbaum, S. A. (2001) Regulation of hyperpolarization-activated HCN channel gating and cAMP modulation due to interactions of COOH terminus and core transmembrane regions. *J. Gen. Physiol.* **118**, 237–250 CrossRef Medline

HCN2 channel activation by N⁶-modified cAMP derivatives

11. Clapham, D. E. (1998) Not so funny anymore: pacing channels are cloned. *Neuron* **21**, 5–7 [CrossRef Medline](#)
12. Santoro, B., Grant, S. G., Bartsch, D., and Kandel, E. R. (1997) Interactive cloning with the SH3 domain of N-src identifies a new brain specific ion channel protein, with homology to eag and cyclic nucleotide-gated channels. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 14815–14820 [CrossRef Medline](#)
13. Wainger, B. J., DeGennaro, M., Santoro, B., Siegelbaum, S. A., and Tibbs, G. R. (2001) Molecular mechanism of cAMP modulation of HCN pacemaker channels. *Nature* **411**, 805–810 [CrossRef Medline](#)
14. DiFrancesco, D., and Tortora, P. (1991) Direct activation of cardiac pacemaker channels by intracellular cyclic AMP. *Nature* **351**, 145–147 [CrossRef Medline](#)
15. Wicks, N. L., Wong, T., Sun, J., Madden, Z., and Young, E. C. (2011) Cytoplasmic cAMP-sensing domain of hyperpolarization-activated cation (HCN) channels uses two structurally distinct mechanisms to regulate voltage gating. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 609–614 [CrossRef Medline](#)
16. Shabb, J. B., and Corbin, J. D. (1992) Cyclic nucleotide-binding domains in proteins having diverse functions. *J. Biol. Chem.* **267**, 5723–5726 [Medline](#)
17. Rehmann, H., Wittinghofer, A., and Bos, J. L. (2007) Capturing cyclic nucleotides in action: snapshots from crystallographic studies. *Nat. Rev. Mol. Cell Biol.* **8**, 63–73 [CrossRef Medline](#)
18. Chen, L., Wang, P., Andrade, C. F., Zhao, I. Y., Dubé, P. E., Brubaker, P. L., Liu, M., and Jin, T. (2005) PKA independent and cell type specific activation of the expression of caudal homeobox gene Cdx-2 by cyclic AMP. *FEBS J.* **272**, 2746–2759 [CrossRef Medline](#)
19. Hewer, R. C., Sala-Newby, G. B., Wu, Y. J., Newby, A. C., and Bond, M. (2011) PKA and Epac synergistically inhibit smooth muscle cell proliferation. *J. Mol. Cell Cardiol.* **50**, 87–98 [CrossRef Medline](#)
20. Kang, G., Joseph, J. W., Chepurny, O. G., Monaco, M., Wheeler, M. B., Bos, J. L., Schwede, F., Genieser, H. G., and Holz, G. G. (2003) Epac-selective cAMP analog 8-pCPT-2'-O-Me-cAMP as a stimulus for Ca²⁺-induced Ca²⁺ release and exocytosis in pancreatic β -cells. *J. Biol. Chem.* **278**, 8279–8285 [CrossRef Medline](#)
21. Rangarajan, S., Enserink, J. M., Kuiperij, H. B., de Rooij, J., Price, L. S., Schwede, F., and Bos, J. L. (2003) Cyclic AMP induces integrin-mediated cell adhesion through Epac and Rap1 upon stimulation of the β 2-adrenergic receptor. *J. Cell Biol.* **160**, 487–493 [CrossRef Medline](#)
22. Shimomura, H., Imai, A., and Nashida, T. (2004) Evidence for the involvement of cAMP-GEF (Epac) pathway in amylase release from the rat parotid gland. *Arch. Biochem. Biophys.* **431**, 124–128 [CrossRef Medline](#)
23. Christensen, A. E., Selheim, F., de Rooij, J., Dremier, S., Schwede, F., Dao, K. K., Martinez, A., Maenhaut, C., Bos, J. L., Genieser, H. G., and Døskeland, S. O. (2003) cAMP analog mapping of Epac1 and cAMP kinase. Discriminating analogs demonstrate that Epac and cAMP kinase act synergistically to promote PC-12 cell neurite extension. *J. Biol. Chem.* **278**, 35394–35402 [CrossRef Medline](#)
24. Ng, L. C. T., Putrenko, I., Baronas, V., Van Petegem, F., and Accili, E. A. (2016) Cyclic purine and pyrimidine nucleotides bind to the HCN2 ion channel and variably promote C-terminal domain interactions and opening. *Structure* **24**, 1629–1642 [CrossRef Medline](#)
25. Zagotta, W. N., Olivier, N. B., Black, K. D., Young, E. C., Olson, R., and Gouaux, E. (2003) Structural basis for modulation and agonist specificity of HCN pacemaker channels. *Nature* **425**, 200–205 [CrossRef Medline](#)
26. Friesner, R. A., Murphy, R. B., Repasky, M. P., Frye, L. L., Greenwood, J. R., Halgren, T. A., Sanschagrin, P. C., and Mainz, D. T. (2006) Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J. Med. Chem.* **49**, 6177–6196 [CrossRef Medline](#)
27. Døskeland, S. O., OGREID, D., Ekanger, R., Sturm, P. A., Miller, J. P., and Suva, R. H. (1983) Mapping of the two intrachain cyclic nucleotide binding sites of adenosine cyclic 3',5'-phosphate dependent protein kinase I. *Biochemistry* **22**, 1094–1101 [CrossRef Medline](#)
28. OGREID, D., Ekanger, R., Suva, R. H., Miller, J. P., and Døskeland, S. O. (1989) Comparison of the two classes of binding sites (A and B) of type I and type II cyclic-AMP-dependent protein kinases by using cyclic nucleotide analogs. *Eur. J. Biochem.* **181**, 19–31 [CrossRef Medline](#)
29. OGREID, D., Ekanger, R., Suva, R. H., Miller, J. P., Sturm, P., Corbin, J. D., and Døskeland, S. O. (1985) Activation of protein kinase isozymes by cyclic nucleotide analogs used singly or in combination: principles for optimizing the isozyme specificity of analog combinations. *Eur. J. Biochem.* **150**, 219–227 [CrossRef Medline](#)
30. Zhou, L., and Siegelbaum, S. A. (2007) Gating of HCN channels by cyclic nucleotides: residue contacts that underlie ligand binding, selectivity, and efficacy. *Structure* **15**, 655–670 [CrossRef Medline](#)
31. Colquhoun, D. (1998) Binding, gating, affinity and efficacy: the interpretation of structure-activity relationships for agonists and of the effects of mutating receptors. *Br. J. Pharmacol.* **125**, 924–947 [Medline](#)
32. Stephenson, R. P. (1956) A modification of receptor theory. *Br. J. Pharmacol. Chemother.* **11**, 379–393 [CrossRef Medline](#)
33. Flynn, G. E., Black, K. D., Islas, L. D., Sankaran, B., and Zagotta, W. N. (2007) Structure and rearrangements in the carboxy-terminal region of SpIH channels. *Structure* **15**, 671–682 [CrossRef Medline](#)
34. Zorn, M., Fladmark, K. E., OGREID, D., Jastorff, B., Døskeland, S. O., and Dostmann, W. R. (1995) Ala335 is essential for high-affinity cAMP-binding of both sites A and B of cAMP-dependent protein kinase type I. *FEBS Lett.* **362**, 291–294 [CrossRef Medline](#)
35. Poppe, H., Rybalkin, S. D., Rehmann, H., Hinds, T. R., Tang, X. B., Christensen, A. E., Schwede, F., Genieser, H. G., Bos, J. L., Døskeland, S. O., Beavo, J. A., and Butt, E. (2008) Cyclic nucleotide analogs as probes of signaling pathways. *Nat. Methods* **5**, 277–278 [CrossRef Medline](#)
36. Chepurny, O. G., Bertinetti, D., Diskar, M., Leech, C. A., Afshari, P., Tsalkova, T., Cheng, X., Schwede, F., Genieser, H. G., Herberg, F. W., and Holz, G. G. (2013) Stimulation of proglucagon gene expression by human GPR119 in enteroendocrine L-cell line GLUTag. *Mol. Endocrinol.* **27**, 1267–1282 [CrossRef Medline](#)
37. Lo, K. W., Ashe, K. M., Kan, H. M., Lee, D. A., and Laurencin, C. T. (2011) Activation of cyclic amp/protein kinase: a signaling pathway enhances osteoblast cell adhesion on biomaterials for regenerative engineering. *J. Orthop. Res.* **29**, 602–608 [CrossRef Medline](#)
38. Monje, P. V., Athauda, G., and Wood, P. M. (2008) Protein kinase A-mediated gating of neuregulin-dependent ErbB2-ErbB3 activation underlies the synergistic action of cAMP on Schwann cell proliferation. *J. Biol. Chem.* **283**, 34087–34100 [CrossRef Medline](#)
39. Berman, H. M., Ten Eyck, L. F., Goodsell, D. S., Haste, N. M., Kornev, A., and Taylor, S. S. (2005) The cAMP binding domain: an ancient signaling module. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 45–50 [CrossRef Medline](#)
40. Lee, C. H., and MacKinnon, R. (2017) Structures of the human HCN1 hyperpolarization-activated channel. *Cell* **168**, 111–120.e11 [CrossRef Medline](#)
41. Huq, I., Dostmann, W. R., and OGREID, D. (1996) Isoleucine 368 is involved in low-affinity binding of N⁶-modified cAMP analogues to site B of the regulatory subunit of cAMP-dependent protein kinase I. *Biochem. J.* **316**, 337–343 [CrossRef Medline](#)

12 CURRICULUM VITAE

12.1 Professional and scientific career

06.14 – present **Research associate**

Institute for Pharmaceutical and Medicinal Chemistry
Heinrich Heine University Düsseldorf
Düsseldorf, Germany

Key tasks:

- Head assistant
“Molecular Modeling” (Pharmacy studies, 7th semester)
- Course assistant
“Drug Analysis” (Pharmacy studies, 7th semester)
- Lecture substitute
“Medicinal Chemistry” (Pharmacy studies, 6th – 8th semester)
“Stereochemistry” (Pharmacy studies, 3rd semester)

06.14 – present **Ph. D. student**

Institute for Pharmaceutical and Medicinal Chemistry
Heinrich Heine University Düsseldorf
Research group of Prof. Dr. Holger Gohlke
Düsseldorf, Germany

Key research areas:

- Investigation of dynamic protein-ligand interactions and protein-ligand binding free energies
- Computational identification and design of novel protein ligands
- Computational elucidation of the stereochemistry of natural products
- Prediction of allosteric modulation of protein-ligand binding

Key tasks:

- Advanced internal training
- Mentoring of diploma, master, and junior Ph. D. students
- Writing of project proposals for research grants
- Newcomer care

06.13 – 06.14 **Scholarship holder**

iRTG of the CRC 974
Düsseldorf, Germany

06.13 – present **Member of the integrated Research Training Group (iRTG) of the Collaborative Research Centre 974 (CRC 974) – “Communication and System Relevance in Liver Injury and Regeneration”**

Düsseldorf, Germany

Key tasks:

- Head of the organizing committee of the doctoral symposium *deLIVER – Technology in Hepatology* (2017)
- Deputy spokesperson of the graduate students in the iRTG (2019)

03.16 **License to practice pharmacy (Approbation)**

05.13 – 11.13 **Trainee pharmacist**

Institute for Pharmaceutical and Medicinal Chemistry
Heinrich Heine University Düsseldorf
Research group of Prof. Dr. Holger Gohlke
Düsseldorf, Germany

11.12 – 05.13 **Trainee pharmacist**

Linner Apotheke
Krefeld, Germany

04.08 – 10.12 **Pharmacy studies**

Heinrich Heine University Düsseldorf
Düsseldorf, Germany

12.2 Scientific achievements

Recent publications

Bonus, M., et al.

Liver cell hydration and integrin signaling.
Biological Chemistry (2021), in press.
(aop doi: 10.1515/hsz-2021-0193)

Prescher, M., Bonus, M., et al.

Evidence for a credit-card-swipe mechanism in the human PC-floppase ABCB4.
Structure (2021), in press.
(aop doi: 10.1016/j.str.2021.05.013)

Moe-Lange, J., Gappel, N. M., Machado, M., Wudick, M. M., Sies, C. S. A. Schott-Verdugo, S. N., Bonus, M., et al.

Interdependence of a mechanosensitive anion channel and glutamate receptors in distal wound signaling.
Submitted (2021).

Schirmeyer, J., Hummert, S., Eick, T., Schulz, E., Schwabe, T., Ehrlich, G., Kukaj, T., Wiegand, M., Sattler, C., Schmauder, R., Zimmer, T., Kosmalla, N., Bonus, M., et al.

Thermodynamic profile of mutual subunit control in a heteromeric receptor.
Proceedings of the National Academy of Sciences of the United States of America (2021), in press.

Tran-Cong, N. M., Hohlfeld, T., Bonus, M., et al.

The mycotoxin Altenusin is a platelet aggregation inhibitor.
Submitted (2021).

Lelle, M., Otte, M., Bonus, M., et al.

Fluorophore-labeled cyclic nucleotides as potent agonists of cyclic nucleotide-regulated ion channels.
ChemBioChem 21(16), 2311–20 (2020).

Bonus, M., et al.

Evidence for functional selectivity in TUDC- and *nor*UDCA-induced signal transduction via $\alpha_5\beta_1$ integrin towards choleresis.
Scientific Reports 10, 5795 (2020).

Leypold, T., Bonus, M., et al.

*N*⁶-modified cAMP derivatives that activate protein kinase A also act as full agonists of murine HCN2 channels.
Journal of Biological Chemistry 294, 17978-17987 (2019).

Form, I., Bonus, M., et al.

Xanthone, benzophenone and bianthrone derivatives from the hypersaline lake-derived fungus *Aspergillus wentii*.
Bioorganic & Medicinal Chemistry 27, 115005 (2019).

Preising, M. N., Görg, B., Friedburg, C., Qvartskhava, N., Budde, B. S., Bonus, M., *et al.*

Biallelic mutation of human SLC6A6 encoding the taurine transporter TAUT is linked to early retinal degeneration.

The FASEB Journal 33(10), 11507–11527 (2019).

Otte, M., Schweinitz, A., Lelle, M., Thon, S., Enke, U., Yüksel, S., Schmauder, R., Bonus, M., *et al.*

Novel Fluorescent Cyclic Nucleotide Derivatives to Study CNG and HCN Channel Function.

Biophysical Journal 116(12), 2411–2422 (2019)

Oral presentations

Bonus, M.

Elucidating the Mechanism of Lipid Translocation by the Canalicular ABC Transporter MDR3.

deLIVER – Technology in Hepatology, Düsseldorf, Germany
September 28th – September 29th, 2017

Prescher, M., Bonus, M.

MDR3 – State of the Art.

CRC 974 Retreat, Trier, Germany
February 4th – February 5th, 2017

Poster presentations

Pfleger, C., Bonus, M., *et al.*

Disinhibition and inhibition of HCN2 channel function by ligand binding to the cyclic nucleotide binding domain.

NIC Symposium, Jülich, Germany
February 27th – February 28th, 2020

Bonus, M., *et al.*

Lipid Transport by the ABC Transporter MDR3.

NIC Symposium, Jülich, Germany
February 22nd – February 23rd, 2018

Pfleger, C., Bonus, M., *et al.*

Disinhibition and inhibition of HCN2 channel function by ligand binding to the cyclic nucleotide binding domain.

NIC Symposium, Jülich, Germany
February 22nd – February 23rd, 2018

Bonus, M., *et al.*

$\alpha_5\beta_1$ integrins are receptors for bile acids with a (nor-)ursodeoxycholate scaffold

XXIII International Bile Acid Meeting, Düsseldorf, Germany
June 17th – June 18th, 2016

Awarded the 2nd best poster award

Bonus, M., et al.

Lipid Transport by the ABC Transporter MDR3.

NIC Symposium, Jülich, Germany

February 11th – February 12th, 2016

Bonus, M., et al.

Conformational Response of $\alpha_5\beta_1$ Integrin to Bile Acid Binding.

Workshop on Computer Simulations and Theory of Macromolecules,

Hünfeld, Germany

April 11th – April 12th, 2014

Grants and funding	05.17 – present
	Computing time grant on JUWELS/JURECA, Jülich Supercomputing Centre (Project ID: HDD17)
	11.15 – 10.17
	Computing time grant on JURECA/JUROPA, Jülich Supercomputing Centre (Project ID: HDU19)
	<i>“Energetics of Lipid and Drug Transport by the ABC Transporters MDR3 and MDRI”</i>
	06.13 – 06.14
	Research fellowship in the integrated Research Training Group (iRTG) of the Collaborative Research Centre 974 – “Communication and System Relevance in Liver Injury and Regeneration”
	<i>“Mode of action of bile acid derivatives on $\alpha_5\beta_1$ integrin.”</i>
<hr/>	
Membership in scientific consortia	04.17 – present
	Research associate of the Research Unit 2518
	<i>“Functional dynamics of ion channels and transporters – DynIon”</i>
	07.14 – present
	Research associate of the Collaborative Research Centre 974
	<i>“Communication and System Relevance in Liver Injury and Regeneration”</i>
	06.13 – 06.14
	Research fellow of the Collaborative Research Centre 974
	<i>“Communication and System Relevance in Liver Injury and Regeneration”</i>

13 REFERENCES

1. Carlberg, C. & Molnár, F. Mechanisms of Gene Regulation. 2nd edn 1 online resource (XVI, 211 pages 82 illustrations in color (Springer Netherlands : Imprint: Springer., Dordrecht, 2016).
2. Radzicka, A. & Wolfenden, R. A proficient enzyme. *Science* **267**, 90-3 (1995).
3. Williams, M.A. & Daviter, T. *Protein-Ligand Interactions - Methods and Applications*, X, 530 (Humana Press; Springer Science+Business Media, New York, 2013).
4. Bradshaw, R.A. & Dennis, E.A. *Handbook of cell signaling*, (Academic Press, Amsterdam San Diego, Calif., 2004).
5. Krauss, G. *Biochemistry of Signal Transduction and Regulation*, XVI, 541 (Wiley-VCH, Weinheim Great Britain, 2003).
6. Du, X. et al. Insights into Protein-Ligand Interactions: Mechanisms, Models, and Methods. *International Journal of Molecular Sciences* **17**(2016).
7. Renaud, J.-P. *Structural Biology in Drug Discovery - Methods, Techniques, and Practices*, 1 online resource (Wiley, Hoboken, NJ, 2020).
8. Gao, Z.-G. et al. Structural determinants of A(3) adenosine receptor activation: nucleoside ligands at the agonist/antagonist boundary. *Journal of Medicinal Chemistry* **45**, 4471-84 (2002).
9. Wifling, D. et al. Basal Histamine H4 Receptor Activation: Agonist Mimicry by the Diphenylalanine Motif. *Chemistry* **25**, 14613-14624 (2019).
10. Klein Herenbrink, C. et al. Molecular Determinants of the Intrinsic Efficacy of the Antipsychotic Aripiprazole. *ACS Chemical Biology* **14**, 1780-1792 (2019).
11. Durrant, J.D. & McCammon, J.A. Molecular dynamics simulations and drug discovery. *BMC Biology* **9**, 71 (2011).
12. Lelle, M., Otte, M., Bonus, M., Gohlke, H. & Benndorf, K. Fluorophore-labeled cyclic nucleotides as potent agonists of cyclic nucleotide-regulated ion channels. *ChemBioChem* **21**, 2311-2320 (2020).
13. Fischer, E. Einfluss der Configuration auf die Wirkung der Enzyme. *Berichte der deutschen chemischen Gesellschaft* **27**, 2985-2993 (1894).
14. Fischer, E. Einfluss der Configuration auf die Wirkung der Enzyme. II. *Berichte der deutschen chemischen Gesellschaft* **27**, 3479-3483 (1894).
15. Fischer, E. Ueber den Einfluss der Configuration auf die Wirkung der Enzyme III. *Berichte der deutschen chemischen Gesellschaft* **28**, 1429-1438 (1895).
16. Koshland, D.E. Application of a Theory of Enzyme Specificity to Protein Synthesis. *Proceedings of the National Academy of Sciences of the United States of America* **44**, 98-104 (1958).
17. Frauenfelder, H., Sligar, S.G. & Wolynes, P.G. The energy landscapes and motions of proteins. *Science* **254**, 1598-603 (1991).
18. Henzler-Wildman, K. & Kern, D. Dynamic personalities of proteins. *Nature* **450**, 964-72 (2007).
19. Bryngelson, J.D., Onuchic, J.N., Socci, N.D. & Wolynes, P.G. Funnels, pathways, and the energy landscape of protein folding: a synthesis. *Proteins* **21**, 167-95 (1995).
20. Miller, D.W. & Dill, K.A. Ligand binding to proteins: the binding landscape model. *Protein Science* **6**, 2166-79 (1997).
21. De Vivo, M., Masetti, M., Bottegoni, G. & Cavalli, A. Role of Molecular Dynamics and Related Methods in Drug Discovery. *Journal of Medicinal Chemistry* **59**, 4035-61 (2016).
22. Changeux, J.P. & Edelstein, S. Conformational selection or induced fit? 50 years of debate resolved. *F1000 Biol Rep* **3**, 19 (2011).
23. Csermely, P., Palotai, R. & Nussinov, R. Induced fit, conformational selection and independent dynamic segments: an extended view of binding events. *Trends in Biochemical Sciences* **35**, 539-46 (2010).
24. Kastriitis, P.L. & Bonvin, A.M.J.J. On the binding affinity of macromolecular interactions: daring to ask why proteins interact. *Journal of the Royal Society Interface* **10**, 20120835 (2013).
25. Liu, S.-Q. et al. Protein Folding, Binding and Energy Landscape: A Synthesis. in *Protein Engineering* (ed. Kaumaya, P.) 207-244 (IntechOpen, 2012).
26. Perica, T. & Chothia, C. Ubiquitin--molecular mechanisms for recognition of different structures. *Current Opinion in Structural Biology* **20**, 367-76 (2010).
27. Boehr, D.D., Nussinov, R. & Wright, P.E. The role of dynamic conformational ensembles in biomolecular recognition. *Nature Chemical Biology* **5**, 789-96 (2009).
28. Martin, S.F. & Clements, J.H. Correlating structure and energetics in protein-ligand interactions: paradigms and paradoxes. *Annual Review of Biochemistry* **82**, 267-93 (2013).

29. Gohlke, H. & Klebe, G. Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors. *Angewandte Chemie, International Edition in English* **41**, 2644-76 (2002).
30. Phillips, R., Kondev, J., Theriot, J. & Garcia, H.G. *Physical Biology of the Cell*, 1057 (Garland Science, Taylor & Francis Group, LLC, New York, NY; Abingdon, UK, 2013).
31. Cooper, A. & Johnson, C.M. Introduction to microcalorimetry and biomolecular energetics. *Methods in Molecular Biology* **22**, 109-24 (1994).
32. Wang, E. et al. End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: Strategies and Applications in Drug Design. *Chemical Reviews* **119**, 9478-9508 (2019).
33. Chodera, J.D. & Mobley, D.L. Entropy-enthalpy compensation: role and ramifications in biomolecular ligand recognition and design. *Annual Review of Biophysics* **42**, 121-42 (2013).
34. Perozzo, R., Folkers, G. & Scapozza, L. Thermodynamics of protein-ligand interactions: history, presence, and future aspects. *Journal of Receptors and Signal Transduction* **24**, 1-52 (2004).
35. Chang, C.-e.A., Chen, W. & Gilson, M.K. Ligand configurational entropy and protein binding. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 1534-9 (2007).
36. Schön, A. et al. Thermodynamics of binding of a low-molecular-weight CD4 mimetic to HIV-1 gp120. *Biochemistry* **45**, 10973-80 (2006).
37. Chong, S.-H. & Ham, S. Dissecting Protein Configurational Entropy into Conformational and Vibrational Contributions. *Journal of Physical Chemistry B* **119**, 12623-31 (2015).
38. Karplus, M., Ichiye, T. & Pettitt, B.M. Configurational entropy of native proteins. *Biophysical Journal* **52**, 1083-5 (1987).
39. Hill, T.L. *An introduction to statistical thermodynamics*, 508 (Dover Publications, New York, 1986).
40. Zhou, H.-X. & Gilson, M.K. Theory of free energy and entropy in noncovalent binding. *Chemical Reviews* **109**, 4092-107 (2009).
41. Ben-Shalom, I.Y., Pfeiffer-Marek, S., Baringhaus, K.-H. & Gohlke, H. Efficient Approximation of Ligand Rotational and Translational Entropy Changes upon Binding for Use in MM-PBSA Calculations. *Journal of Chemical Information and Modeling* **57**, 170-189 (2017).
42. Leffler, J.E. The Enthalpy-Entropy Relationship and Its Implications for Organic Chemistry. *Journal of Organic Chemistry* **20**, 1202-1231 (1955).
43. Lumry, R. & Rajender, S. Enthalpy-entropy compensation phenomena in water solutions of proteins and small molecules: a ubiquitous property of water. *Biopolymers* **9**, 1125-227 (1970).
44. Reynolds, C.H. & Holloway, M.K. Thermodynamics of ligand binding and efficiency. *ACS Medicinal Chemistry Letters* **2**, 433-7 (2011).
45. Olsson, T.S., Ladbury, J.E., Pitt, W.R. & Williams, M.A. Extent of enthalpy-entropy compensation in protein-ligand interactions. *Protein Science* **20**, 1607-18 (2011).
46. Myszka, D.G. et al. The ABRF-MIRG'02 study: assembly state, thermodynamic, and kinetic analysis of an enzyme/inhibitor interaction. *Journal of Biomolecular Techniques* **14**, 247-69 (2003).
47. Blasie, C.A. & Berg, J.M. Entropy-enthalpy compensation in ionic interactions probed in a zinc finger peptide. *Biochemistry* **43**, 10600-4 (2004).
48. Davidson, J.P., Lubman, O., Rose, T., Waksman, G. & Martin, S.F. Calorimetric and structural studies of 1,2,3-trisubstituted cyclopropanes as conformationally constrained peptide inhibitors of Src SH2 domain binding. *Journal of the American Chemical Society* **124**, 205-15 (2002).
49. Houk, K.N., Leach, A.G., Kim, S.P. & Zhang, X. Binding affinities of host-guest, protein-ligand, and protein-transition-state complexes. *Angewandte Chemie International Edition* **42**, 4872-97 (2003).
50. Amaro, R.E. & Mulholland, A.J. Multiscale Methods in Drug Design Bridge Chemical and Biological Complexity in the Search for Cures. *Nature Reviews Chemistry* **2**(2018).
51. Jorgensen, W.L. The many roles of computation in drug discovery. *Science* **303**, 1813-8 (2004).
52. Lee, E.H., Hsin, J., Sotomayor, M., Comellas, G. & Schulten, K. Discovery through the computational microscope. *Structure* **17**, 1295-306 (2009).
53. Dror, R.O., Dirks, R.M., Grossman, J.P., Xu, H. & Shaw, D.E. Biomolecular simulation: a computational microscope for molecular biology. *Annual Review of Biophysics* **41**, 429-52 (2012).
54. Hollingsworth, S.A. & Dror, R.O. Molecular Dynamics Simulation for All. *Neuron* **99**, 1129-1143 (2018).
55. Borhani, D.W. & Shaw, D.E. The future of molecular dynamics simulations in drug discovery. *Journal of Computer-Aided Molecular Design* **26**, 15-26 (2012).
56. Bonus, M. et al. Evidence for functional selectivity in TUDC- and norUDCA-induced signal transduction via alpha5beta1 integrin towards choleresis. *Scientific Reports* **10**, 5795 (2020).
57. Koch, A., Bonus, M., Gohlke, H. & Klöcker, N. Isoform-specific Inhibition of N-methyl-D-aspartate Receptors by Bile Salts. *Scientific Reports* **9**, 10068 (2019).

58. Spahn, V. et al. A nontoxic pain killer designed by modeling of pathological receptor conformations. *Science* **355**, 966-969 (2017).
59. Shan, Y. et al. How does a drug molecule find its target binding site? *Journal of the American Chemical Society* **133**, 9181-3 (2011).
60. Dror, R.O. et al. Pathway and mechanism of drug binding to G-protein-coupled receptors. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 13118-23 (2011).
61. Dror, R.O. et al. Structural basis for modulation of a G-protein-coupled receptor by allosteric drugs. *Nature* **503**, 295-9 (2013).
62. Kappel, K., Miao, Y. & McCammon, J.A. Accelerated molecular dynamics simulations of ligand binding to a muscarinic G-protein-coupled receptor. *Quarterly Reviews of Biophysics* **48**, 479-87 (2015).
63. Miao, Y., Bhattarai, A. & Wang, J. Ligand Gaussian accelerated molecular dynamics (LiGaMD): Characterization of ligand binding thermodynamics and kinetics. *bioRxiv*, 2020.04.20.051979 (2020).
64. Perez, A., Morrone, J.A., Simmerling, C.L. & Dill, K.A. Advances in free-energy-based simulations of protein folding and ligand binding. *Current Opinion in Structural Biology* **36**, 25-31 (2016).
65. Chodera, J.D. et al. Alchemical free energy methods for drug discovery: progress and challenges. *Current Opinion in Structural Biology* **21**, 150-60 (2011).
66. Aleksandrov, A., Thompson, D. & Simonson, T. Alchemical free energy simulations for biological complexes: powerful but temperamental. *Journal of Molecular Recognition* **23**, 117-27 (2010).
67. Jorgensen, W.L. & Thomas, L.L. Perspective on Free-Energy Perturbation Calculations for Chemical Equilibria. *Journal of Chemical Theory and Computation* **4**, 869-876 (2008).
68. Zwanzig, R.W. High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *Journal of Chemical Physics* **22**, 1420-1426 (1954).
69. Kirkwood, J.G. Statistical Mechanics of Fluid Mixtures. *The Journal of Chemical Physics* **3**, 300-313 (1935).
70. He, X. et al. Fast, Accurate, and Reliable Protocols for Routine Calculations of Protein-Ligand Binding Affinities in Drug Design Projects Using AMBER GPU-TI with ff14SB/GAFF. *ACS Omega* **5**, 4611-4619 (2020).
71. Wang, L. et al. Accurate and reliable prediction of relative ligand binding potency in prospective drug discovery by way of a modern free-energy calculation protocol and force field. *Journal of the American Chemical Society* **137**, 2695-703 (2015).
72. Mobley, D.L. & Dill, K.A. Binding of small-molecule ligands to proteins: "what you see" is not always "what you get". *Structure* **17**, 489-98 (2009).
73. Kollman, P.A. et al. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Accounts of Chemical Research* **33**, 889-97 (2000).
74. Wang, W., Donini, O., Reyes, C.M. & Kollman, P.A. Biomolecular simulations: recent developments in force fields, simulations of enzyme catalysis, protein-ligand, protein-protein, and protein-nucleic acid noncovalent interactions. *Annual Review of Biophysics and Biomolecular Structure* **30**, 211-43 (2001).
75. Wang, C., Greene, D.A., Xiao, L., Qi, R. & Luo, R. Recent Developments and Applications of the MMPBSA Method. *Frontiers in Molecular Biosciences* **4**, 87 (2017).
76. Genheden, S. & Ryde, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opinion on Drug Discovery* **10**, 449-61 (2015).
77. Aldeghi, M., Bodkin, M.J., Knapp, S. & Biggin, P.C. Statistical Analysis on the Performance of Molecular Mechanics Poisson-Boltzmann Surface Area versus Absolute Binding Free Energy Calculations: Bromodomains as a Case Study. *Journal of Chemical Information and Modeling* **57**, 2203-2221 (2017).
78. McCammon, J.A., Gelin, B.R. & Karplus, M. Dynamics of folded proteins. *Nature* **267**, 585-90 (1977).
79. Le Grand, S., Götz, A.W. & Walker, R.C. SPFP: Speed without compromise-A mixed precision model for GPU accelerated molecular dynamics simulations. *Computer Physics Communications* **184**, 374-380 (2013).
80. Götz, A.W. et al. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born. *Journal of Chemical Theory and Computation* **8**, 1542-1555 (2012).
81. Salomon-Ferrer, R., Götz, A.W., Poole, D., Le Grand, S. & Walker, R.C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *Journal of Chemical Theory and Computation* **9**, 3878-88 (2013).
82. Cornell, W.D. et al. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *Journal of the American Chemical Society* **117**, 5179-5197 (1995).

83. Case, D.A. et al. The Amber biomolecular simulation programs. *Journal of Computational Chemistry* **26**, 1668-88 (2005).
84. Case, D.A. et al. AMBER 2020. (University of California, San Francisco, 2020).
85. Brooks, B.R. et al. CHARMM: the biomolecular simulation program. *Journal of Computational Chemistry* **30**, 1545-614 (2009).
86. Van Der Spoel, D. et al. GROMACS: fast, flexible, and free. *Journal of Computational Chemistry* **26**, 1701-18 (2005).
87. Tian, C. et al. ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution. *Journal of Chemical Theory and Computation* **16**, 528-552 (2020).
88. Maier, J.A. et al. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *Journal of Chemical Theory and Computation* **11**, 3696-713 (2015).
89. Debiec, K.T. et al. Further along the Road Less Traveled: AMBER ff15ipq, an Original Protein Force Field Built on a Self-Consistent Physical Model. *Journal of Chemical Theory and Computation* **12**, 3926-47 (2016).
90. Lindorff-Larsen, K. et al. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **78**, 1950-8 (2010).
91. Hornak, V. et al. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins* **65**, 712-25 (2006).
92. Wang, J., Cieplak, P. & Kollman, P.A. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *Journal of Computational Chemistry* **21**, 1049-1074 (2000).
93. Pérez, A. et al. Refinement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers. *Biophysical Journal* **92**, 3817-29 (2007).
94. Zgarbová, M. et al. Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles. *Journal of Chemical Theory and Computation* **7**, 2886-2902 (2011).
95. Kirschner, K.N. et al. GLYCAM06: a generalizable biomolecular force field. Carbohydrates. *Journal of Computational Chemistry* **29**, 622-55 (2008).
96. Skjevik, Å.A., Madej, B.D., Walker, R.C. & Teigen, K. LIPID11: a modular framework for lipid simulations using amber. *Journal of Physical Chemistry B* **116**, 11124-36 (2012).
97. Dickson, C.J. et al. Lipid14: The Amber Lipid Force Field. *Journal of Chemical Theory and Computation* **10**, 865-879 (2014).
98. Gould, I.R., Skjevik, Å.A., Dickson, C.J., Madej, B.D. & Walker, R.C. Lipid17: A Comprehensive AMBER Force Field for the Simulation of Zwitterionic and Anionic Lipids. *Manuscript in preparation* (2018).
99. Wang, J., Wolf, R.M., Caldwell, J.W., Kollman, P.A. & Case, D.A. Development and testing of a general amber force field. *Journal of Computational Chemistry* **25**, 1157-74 (2004).
100. Vasseti, D., Pagliai, M. & Procacci, P. Assessment of GAFF2 and OPLS-AA General Force Fields in Combination with the Water Models TIP3P, SPCE, and OPC3 for the Solvation Free Energy of Druglike Organic Molecules. *Journal of Chemical Theory and Computation* **15**, 1983-1995 (2019).
101. Minges, A. et al. Structural intermediates and directionality of the swiveling motion of Pyruvate Phosphate Dikinase. *Scientific Reports* **7**, 45389 (2017).
102. Ciupka, D. & Gohlke, H. On the potential alternate binding change mechanism in a dimeric structure of Pyruvate Phosphate Dikinase. *Scientific Reports* **7**, 8020 (2017).
103. Frieg, B. et al. Mechanism of Fully Reversible, pH-Sensitive Inhibition of Human Glutamine Synthetase by Tyrosine Nitration. *Journal of Chemical Theory and Computation* **16**, 4694-4705 (2020).
104. Greife, A. et al. Structural assemblies of the di- and oligomeric G-protein coupled receptor TGR5 in live cells: an MFIS-FRET and integrative modelling study. *Scientific Reports* **6**, 36792 (2016).
105. Spomer, L. et al. A membrane-proximal, C-terminal alpha-helix is required for plasma membrane localization and function of the G Protein-coupled receptor (GPCR) TGR5. *Journal of Biological Chemistry* **289**, 3689-702 (2014).
106. Ciglia, E. et al. Resolving hot spots in the C-terminal dimerization domain that determine the stability of the molecular chaperone Hsp90. *PLOS ONE* **9**, e96031 (2014).
107. Bopp, B. et al. Design and biological testing of peptidic dimerization inhibitors of human Hsp90 that target the C-terminal domain. *Biochimica et Biophysica Acta (BBA)* **1860**, 1043-55 (2016).
108. Bhatia, S. et al. Targeting HSP90 dimerization via the C terminus is effective in imatinib-resistant CML and lacks the heat shock response. *Blood* **132**, 307-320 (2018).

109. Homeyer, N. et al. Interpreting Thermodynamic Profiles of Aminoadamantane Compounds Inhibiting the M2 Proton Channel of Influenza A by Free Energy Calculations. *Journal of Chemical Information and Modeling* **56**, 110-26 (2016).
110. Gilson, M.K. & Honig, B. Calculation of the total electrostatic energy of a macromolecular system: solvation energies, binding energies, and conformational analysis. *Proteins* **4**, 7-18 (1988).
111. Sharp, K.A. & Honig, B. Calculating Total Electrostatic Energies with the Nonlinear Poisson-Boltzmann Equation. *Journal of Physical Chemistry* **94**, 7684-7692 (1990).
112. Srinivasan, J., Trevathan, M.W., Beroza, P. & Case, D.A. Application of a pairwise generalized Born model to proteins and nucleic acids: inclusion of salt effects. *Theoretical Chemistry Accounts* **101**, 426-434 (1999).
113. Stein, C.J., Herbert, J.M. & Head-Gordon, M. The Poisson-Boltzmann model for implicit solvation of electrolyte solutions: Quantum chemical implementation and assessment via Sechenov coefficients. *Journal of Chemical Physics* **151**, 224111 (2019).
114. Onufriev, A.V., Case, D.A. & Bashford, D. Effective Born radii in the generalized Born approximation: The importance of being perfect. *Journal of Computational Chemistry* **23**, 1297-1304 (2002).
115. Still, W.C., Tempczyk, A., Hawley, R.C. & Hendrickson, T. Semianalytical Treatment of Solvation for Molecular Mechanics and Dynamics. *Journal of the American Chemical Society* **112**, 6127-6129 (1990).
116. Hawkins, G.D., Cramer, C.J. & Truhlar, D.G. Pairwise Solute Descreening of Solute Charges from a Dielectric Medium. *Chemical Physics Letters* **246**, 122-129 (1995).
117. Hawkins, G.D., Cramer, C.J. & Truhlar, D.G. Parametrized models of aqueous free energies of solvation based on pairwise descreening of solute atomic charges from a dielectric medium. *Journal of Physical Chemistry* **100**, 19824-19839 (1996).
118. Onufriev, A.V., Bashford, D. & Case, D.A. Modification of the generalized Born model suitable for macromolecules. *Journal of Physical Chemistry B* **104**, 3712-3720 (2000).
119. Onufriev, A.V., Bashford, D. & Case, D.A. Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins* **55**, 383-94 (2004).
120. Mongan, J., Simmerling, C.L., McCammon, J.A., Case, D.A. & Onufriev, A.V. Generalized Born model with a simple, robust molecular volume correction. *Journal of Chemical Theory and Computation* **3**, 156-169 (2007).
121. Forouzesh, N., Izadi, S. & Onufriev, A.V. Grid-Based Surface Generalized Born Model for Calculation of Electrostatic Binding Free Energies. *Journal of Chemical Information and Modeling* **57**, 2505-2513 (2017).
122. Tan, C., Tan, Y.-H. & Luo, R. Implicit nonpolar solvent models. *The Journal of Physical Chemistry B* **111**, 12263-74 (2007).
123. Ma, J. Usefulness and limitations of normal mode analysis in modeling dynamics of biomolecular complexes. *Structure* **13**, 373-80 (2005).
124. Genheden, S., Kuhn, O., Mikulskis, P., Hoffmann, D. & Ryde, U. The normal-mode entropy in the MM/GBSA method: effect of system truncation, buffer region, and dielectric constant. *Journal of Chemical Information and Modeling* **52**, 2079-88 (2012).
125. Duan, L., Liu, X. & Zhang, J.Z.H. Interaction Entropy: A New Paradigm for Highly Efficient and Reliable Computation of Protein-Ligand Binding Free Energy. *Journal of the American Chemical Society* **138**, 5722-8 (2016).
126. Sun, H. et al. Assessing the performance of MM/PBSA and MM/GBSA methods. 7. Entropy effects on the performance of end-point binding free energy calculation approaches. *Physical Chemistry Chemical Physics* **20**, 14450-14460 (2018).
127. Wang, J. & Hou, T. Develop and test a solvent accessible surface area-based model in conformational entropy calculations. *Journal of Chemical Information and Modeling* **52**, 1199-212 (2012).
128. Swanson, J.M., Henchman, R.H. & McCammon, J.A. Revisiting free energy calculations: a theoretical connection to MM/PBSA and direct calculation of the association free energy. *Biophysical Journal* **86**, 67-74 (2004).
129. Miller, B.R., III et al. MMPBSA.py: An Efficient Program for End-State Free Energy Calculations. *Journal of Chemical Theory and Computation* **8**, 3314-21 (2012).
130. Homeyer, N. & Gohlke, H. FEW: a workflow tool for free energy calculations of ligand binding. *Journal of Computational Chemistry* **34**, 965-73 (2013).
131. Homeyer, N. & Gohlke, H. Extension of the free energy workflow FEW towards implicit solvent/implicit membrane MM-PBSA calculations. *Biochimica et Biophysica Acta (BBA)* **1850**, 972-982 (2015).

132. Hou, T., Wang, J., Li, Y. & Wang, W. Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *Journal of Chemical Information and Modeling* **51**, 69-82 (2011).
133. Hou, T., Wang, J., Li, Y. & Wang, W. Assessing the performance of the molecular mechanics/Poisson Boltzmann surface area and molecular mechanics/generalized Born surface area methods. II. The accuracy of ranking poses generated from docking. *Journal of Computational Chemistry* **32**, 866-77 (2011).
134. Xu, L., Sun, H., Li, Y., Wang, J. & Hou, T. Assessing the performance of MM/PBSA and MM/GBSA methods. 3. The impact of force fields and ligand charge models. *Journal of Physical Chemistry B* **117**, 8408-21 (2013).
135. Sun, H., Li, Y., Tian, S., Xu, L. & Hou, T. Assessing the performance of MM/PBSA and MM/GBSA methods. 4. Accuracies of MM/PBSA and MM/GBSA methodologies evaluated by various simulation protocols using PDBbind data set. *Physical Chemistry Chemical Physics* **16**, 16719-29 (2014).
136. Sun, H. et al. Assessing the performance of MM/PBSA and MM/GBSA methods. 5. Improved docking performance using high solute dielectric constant MM/GBSA and MM/PBSA rescoring. *Physical Chemistry Chemical Physics* **16**, 22035-45 (2014).
137. Chen, F. et al. Assessing the performance of the MM/PBSA and MM/GBSA methods. 6. Capability to predict protein-protein binding free energies and re-rank binding poses generated by protein-protein docking. *Physical Chemistry Chemical Physics* **18**, 22129-39 (2016).
138. Chen, F. et al. Assessing the performance of MM/PBSA and MM/GBSA methods. 8. Predicting binding free energies and poses of protein-RNA complexes. *RNA* **24**, 1183-1194 (2018).
139. Mishra, S.K. & Koča, J. Assessing the Performance of MM/PBSA, MM/GBSA, and QM-MM/GBSA Approaches on Protein/Carbohydrate Complexes: Effect of Implicit Solvent Models, QM Methods, and Entropic Contributions. *Journal of Physical Chemistry B* **122**, 8113-8121 (2018).
140. Weis, A., Katebzadeh, K., Söderhjelm, P., Nilsson, I. & Ryde, U. Ligand affinities predicted with the MM/PBSA method: dependence on the simulation method and the force field. *Journal of Medicinal Chemistry* **49**, 6596-606 (2006).
141. Genheden, S., Luchko, T., Gusarov, S., Kovalenko, A. & Ryde, U. An MM/3D-RISM approach for ligand binding affinities. *Journal of Physical Chemistry B* **114**, 8505-16 (2010).
142. Rastelli, G., Del Rio, A., Degliesposti, G. & Sgobba, M. Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA. *Journal of Computational Chemistry* **31**, 797-810 (2010).
143. Lindström, A. et al. Postprocessing of docked protein-ligand complexes using implicit solvation models. *Journal of Chemical Information and Modeling* **51**, 267-82 (2011).
144. Steinbrecher, T., Case, D.A. & Labahn, A. A multistep approach to structure-based drug design: studying ligand binding at the human neutrophil elastase. *Journal of Medicinal Chemistry* **49**, 1837-44 (2006).
145. Thompson, D.C., Humblet, C. & Joseph-McCarthy, D. Investigation of MM-PBSA rescoring of docking poses. *Journal of Chemical Information and Modeling* **48**, 1081-91 (2008).
146. Zhang, X., Wong, S.E. & Lightstone, F.C. Toward fully automated high performance computing drug discovery: a massively parallel virtual screening pipeline for docking and molecular mechanics/generalized Born surface area rescoring to improve enrichment. *Journal of Chemical Information and Modeling* **54**, 324-37 (2014).
147. Greenidge, P.A., Kramer, C., Mozziconacci, J.-C. & Sherman, W. Improving docking results via reranking of ensembles of ligand poses in multiple X-ray protein conformations with MM-GBSA. *Journal of Chemical Information and Modeling* **54**, 2697-717 (2014).
148. Otte, M. et al. Hydrophobic alkyl chains substituted to the 8-position of cyclic nucleotides enhance activation of CNG and HCN channels by an intricate enthalpy - entropy compensation. *Scientific Reports* **8**, 14960 (2018).
149. Pan, P. et al. Combating Drug-Resistant Mutants of Anaplastic Lymphoma Kinase with Potent and Selective Type-I(1/2) Inhibitors by Stabilizing Unique DFG-Shifted Loop Conformation. *ACS Central Science* **3**, 1208-1220 (2017).
150. Xu, W. et al. Macrocyclized Extended Peptides: Inhibiting the Substrate-Recognition Domain of Tankyrase. *Journal of the American Chemical Society* **139**, 2245-2256 (2017).
151. Roca, C. et al. Deciphering the Inhibition of the Neuronal Calcium Sensor 1 and the Guanine Exchange Factor Ric8a with a Small Phenothiazine Molecule for the Rational Generation of Therapeutic Synapse Function Regulators. *Journal of Medicinal Chemistry* **61**, 5910-5921 (2018).
152. Jiang, Z.-Y. et al. Discovery of potent Keap1-Nrf2 protein-protein interaction inhibitor based on molecular binding determinants analysis. *Journal of Medicinal Chemistry* **57**, 2736-45 (2014).

153. Guan, Y. et al. Exploring resistance mechanisms of HCV NS3/4A protease mutations to MK5172: insight from molecular dynamics simulations and free energy calculations. *Molecular BioSystems* **11**, 2568-78 (2015).
154. Pan, P., Li, Y., Yu, H., Sun, H. & Hou, T. Molecular principle of topotecan resistance by topoisomerase I mutations through molecular modeling approaches. *Journal of Chemical Information and Modeling* **53**, 997-1006 (2013).
155. Bonus, M., Häussinger, D. & Gohlke, H. Liver cell hydration and integrin signaling. *Biological Chemistry* (2021).
156. Kadry, Y.A. & Calderwood, D.A. Chapter 22: Structural and signaling functions of integrins. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1862**, 183206 (2020).
157. Puklin-Faucher, E. & Sheetz, M.P. The mechanical integrin cycle. *Journal of Cell Science* **122**, 179-86 (2009).
158. Wickström, S.A., Radovanac, K. & Fässler, R. Genetic analyses of integrin signaling. *Cold Spring Harbor Perspectives in Biology* **3**(2011).
159. Hegde, S. & Raghavan, S. A skin-depth analysis of integrins: role of the integrin network in health and disease. *Cell Communication & Adhesion* **20**, 155-69 (2013).
160. Plow, E.F., Meller, J. & Byzova, T.V. Integrin function in vascular biology: a view from 2013. *Current Opinion in Hematology* **21**, 241-7 (2014).
161. Winograd-Katz, S.E., Fässler, R., Geiger, B. & Legate, K.R. The integrin adhesome: from genes and proteins to human disease. *Nature Reviews Molecular Cell Biology* **15**, 273-88 (2014).
162. Arnaout, M.A., Mahalingam, B. & Xiong, J.-P. Integrin structure, allostery, and bidirectional signaling. *Annual Review of Cell and Developmental Biology* **21**, 381-410 (2005).
163. Campbell, I.D. & Humphries, M.J. Integrin structure, activation, and interactions. *Cold Spring Harbor Perspectives in Biology* **3**(2011).
164. Clark, E.A. & Brugge, J.S. Integrins and signal transduction pathways: the road taken. *Science* **268**, 233-9 (1995).
165. Huvencers, S. & Danen, E.H.J. Adhesion signaling - crosstalk between integrins, Src and Rho. *Journal of Cell Science* **122**, 1059-69 (2009).
166. Hynes, R.O. Integrins: bidirectional, allosteric signaling machines. *Cell* **110**, 673-87 (2002).
167. Barczyk, M., Carracedo, S. & Gullberg, D. Integrins. *Cell and Tissue Research* **339**, 269-80 (2010).
168. Xiong, J.-P. et al. Crystal structure of the extracellular segment of integrin alpha Vbeta3. *Science* **294**, 339-45 (2001).
169. Xiong, J.-P. et al. Crystal structure of the extracellular segment of integrin alpha Vbeta3 in complex with an Arg-Gly-Asp ligand. *Science* **296**, 151-5 (2002).
170. Takagi, J., Petre, B.M., Walz, T. & Springer, T.A. Global conformational rearrangements in integrin extracellular domains in outside-in and inside-out signaling. *Cell* **110**, 599-11 (2002).
171. Nishida, N. et al. Activation of leukocyte beta2 integrins by conversion from bent to extended conformations. *Immunity* **25**, 583-94 (2006).
172. Kechagia, J.Z., Ivaska, J. & Roca-Cusachs, P. Integrins as biomechanical sensors of the microenvironment. *Nature Reviews Molecular Cell Biology* **20**, 457-473 (2019).
173. Sun, Z., Costell, M. & Fässler, R. Integrin activation by talin, kindlin and mechanical forces. *Nature Cell Biology* **21**, 25-31 (2019).
174. Bachmann, M., Kukkurainen, S., Hytönen, V.P. & Wehrle-Haller, B. Cell Adhesion by Integrins. *Physiol Rev* **99**, 1655-1699 (2019).
175. Morse, E.M., Brahme, N.N. & Calderwood, D.A. Integrin cytoplasmic tail interactions. *Biochemistry* **53**, 810-20 (2014).
176. Kim, C., Ye, F. & Ginsberg, M.H. Regulation of integrin activation. *Annual Reviews of Cell and Developmental Biology* **27**, 321-45 (2011).
177. Adair, B.D. et al. Three-dimensional EM structure of the ectodomain of integrin {alpha}V{beta}3 in a complex with fibronectin. *Journal of Cell Biology* **168**, 1109-18 (2005).
178. Zhu, J., Zhu, J. & Springer, T.A. Complete integrin headpiece opening in eight steps. *Journal of Cell Biology* **201**, 1053-68 (2013).
179. Springer, T.A., Zhu, J. & Xiao, T. Structural basis for distinctive recognition of fibrinogen gammaC peptide by the platelet integrin alphaIIbbeta3. *Journal of Cell Biology* **182**, 791-800 (2008).
180. Volpes, R., van den Oord, J.J. & Desmet, V.J. Integrins as differential cell lineage markers of primary liver tumors. *American Journal of Pathology* **142**, 1483-92 (1993).
181. Häussinger, D. et al. Involvement of Integrins and Src in Tauroursodeoxycholate-Induced and Swelling-Induced Cholereresis. *Gastroenterology* **124**, 1476-87 (2003).

182. Schliess, F., Reissmann, R., Reinehr, R., vom Dahl, S. & Häussinger, D. Involvement of Integrins and Src in Insulin Signaling toward Autophagic Proteolysis in Rat Liver. *Journal of Biological Chemistry* **279**, 21294-301 (2004).
183. Gohlke, H., Schmitz, B., Sommerfeld, A., Reinehr, R. & Häussinger, D. $\alpha 5\beta 1$ -integrins are sensors for tauroursodeoxycholic acid in hepatocytes. *Hepatology* **57**, 1117-29 (2013).
184. Trauner, M. et al. Potential of nor-Ursodeoxycholic Acid in Cholestatic and Metabolic Disorders. *Digestive Diseases* **33**, 433-439 (2015).
185. Traynelis, S.F. et al. Glutamate receptor ion channels: structure, regulation, and function. *Pharmacological Reviews* **62**, 405-96 (2010).
186. Hogan-Cann, A.D. & Anderson, C.M. Physiological Roles of Non-Neuronal NMDA Receptors. *Trends in Pharmacological Sciences* **37**, 750-767 (2016).
187. Hansen, K.B. et al. Structure, function, and allosteric modulation of NMDA receptors. *Journal of General Physiology* **150**, 1081-1105 (2018).
188. Hardingham, G.E. Coupling of the NMDA receptor to neuroprotective and neurodestructive events. *Biochemical Society Transactions* **37**, 1147-60 (2009).
189. Lee, C.-H. et al. NMDA receptor structures reveal subunit arrangement and pore architecture. *Nature* **511**, 191-7 (2014).
190. Karakas, E. & Furukawa, H. Crystal structure of a heterotetrameric NMDA receptor ion channel. *Science* **344**, 992-7 (2014).
191. Ulbrich, M.H. & Isacoff, E.Y. Subunit counting in membrane-bound proteins. *Nature Methods* **4**, 319-21 (2007).
192. Ulbrich, M.H. & Isacoff, E.Y. Rules of engagement for NMDA receptor subunits. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 14163-8 (2008).
193. Atlason, P.T., Garside, M.L., Meddows, E., Whiting, P. & McIlhinney, R.A.J. N-Methyl-D-aspartate (NMDA) receptor subunit NR1 forms the substrate for oligomeric assembly of the NMDA receptor. *Journal of Biological Chemistry* **282**, 25299-307 (2007).
194. Farina, A.N. et al. Separation of domain contacts is required for heterotetrameric assembly of functional NMDA receptors. *Journal of Neuroscience* **31**, 3565-79 (2011).
195. Gielen, M., Siegler Retchless, B., Mony, L., Johnson, J.W. & Paoletti, P. Mechanism of differential control of NMDA receptor activity by NR2 subunits. *Nature* **459**, 703-7 (2009).
196. Yuan, H., Hansen, K.B., Vance, K.M., Ogden, K.K. & Traynelis, S.F. Control of NMDA receptor function by the NR2 subunit amino-terminal domain. *Journal of Neuroscience* **29**, 12045-58 (2009).
197. Karakas, E., Simorowski, N. & Furukawa, H. Structure of the zinc-bound amino-terminal domain of the NMDA receptor NR2B subunit. *EMBO Journal* **28**, 3910-20 (2009).
198. Romero-Hernandez, A., Simorowski, N., Karakas, E. & Furukawa, H. Molecular Basis for Subtype Specificity and High-Affinity Zinc Inhibition in the GluN1-GluN2A NMDA Receptor Amino-Terminal Domain. *Neuron* **92**, 1324-1336 (2016).
199. Karakas, E., Simorowski, N. & Furukawa, H. Subunit arrangement and phenylethanolamine binding in GluN1/GluN2B NMDA receptors. *Nature* **475**, 249-53 (2011).
200. Tajima, N. et al. Activation of NMDA receptors and the mechanism of inhibition by ifenprodil. *Nature* **534**, 63-8 (2016).
201. Furukawa, H. & Gouaux, E. Mechanisms of activation, inhibition and specificity: crystal structures of the NMDA receptor NR1 ligand-binding core. *EMBO Journal* **22**, 2873-85 (2003).
202. Furukawa, H., Singh, S.K., Mancusso, R. & Gouaux, E. Subunit arrangement and function in NMDA receptors. *Nature* **438**, 185-92 (2005).
203. Inanobe, A., Furukawa, H. & Gouaux, E. Mechanism of partial agonist action at the NR1 subunit of NMDA receptors. *Neuron* **47**, 71-84 (2005).
204. Yao, Y. & Mayer, M.L. Characterization of a soluble ligand binding domain of the NMDA receptor regulatory subunit NR3A. *Journal of Neuroscience* **26**, 4559-66 (2006).
205. Yao, Y., Harrison, C.B., Freddolino, P.L., Schulten, K. & Mayer, M.L. Molecular mechanism of ligand recognition by NR3 subtype glutamate receptors. *EMBO Journal* **27**, 2158-70 (2008).
206. Vance, K.M., Simorowski, N., Traynelis, S.F. & Furukawa, H. Ligand-specific deactivation time course of GluN1/GluN2D NMDA receptors. *Nature Communications* **2**, 294 (2011).
207. Yao, Y., Belcher, J., Berger, A.J., Mayer, M.L. & Lau, A.Y. Conformational analysis of NMDA receptor GluN1, GluN2, and GluN3 ligand-binding domains reveals subtype-specific characteristics. *Structure* **21**, 1788-99 (2013).
208. Kvist, T. et al. Crystal structure and pharmacological characterization of a novel N-methyl-D-aspartate (NMDA) receptor antagonist at the GluN1 glycine binding site. *Journal of Biological Chemistry* **288**, 33124-35 (2013).

209. Jespersen, A., Tajima, N., Fernandez-Cuervo, G., Garnier-Amblard, E.C. & Furukawa, H. Structural insights into competitive antagonism in NMDA receptors. *Neuron* **81**, 366-78 (2014).
210. Hackos, D.H. et al. Positive Allosteric Modulators of GluN2A-Containing NMDARs with Distinct Modes of Action and Impacts on Circuit Function. *Neuron* **89**, 983-99 (2016).
211. Volgraf, M. et al. Discovery of GluN2A-Selective NMDA Receptor Positive Allosteric Modulators (PAMs): Tuning Deactivation Kinetics via Structure-Based Design. *Journal of Medicinal Chemistry* **59**, 2760-79 (2016).
212. Yi, F. et al. Structural Basis for Negative Allosteric Modulation of GluN2A-Containing NMDA Receptors. *Neuron* **91**, 1316-1329 (2016).
213. Lind, G.E. et al. Structural basis of subunit selectivity for competitive NMDA receptor antagonists with preference for GluN2A over GluN2B subunits. *Proceedings of the National Academy of Sciences of the United States of America* **114**, E6942-E6951 (2017).
214. Romero-Hernandez, A. & Furukawa, H. Novel Mode of Antagonist Binding in NMDA Receptors Revealed by the Crystal Structure of the GluN1-GluN2A Ligand-Binding Domain Complexed to NVP-AAM077. *Molecular Pharmacology* **92**, 22-29 (2017).
215. Villemure, E. et al. GluN2A-Selective Pyridopyrimidinone Series of NMDAR Positive Allosteric Modulators with an Improved in Vivo Profile. *ACS Medicinal Chemistry Letters* **8**, 84-89 (2017).
216. Kalbaugh, T.L., VanDongen, H.M.A. & VanDongen, A.M.J. Ligand-binding residues integrate affinity and efficacy in the NMDA receptor. *Molecular Pharmacology* **66**, 209-19 (2004).
217. Paganelli, M.A., Kussius, C.L. & Popescu, G.K. Role of cross-cleft contacts in NMDA receptor gating. *PLOS ONE* **8**, e80953 (2013).
218. Wollmuth, L.P. Ion permeation in ionotropic glutamate receptors: Still dynamic after all these years. *Current Opinion in Physiology* **2**, 36-41 (2018).
219. Hinoi, E., Fujimori, S. & Yoneda, Y. Modulation of cellular differentiation by N-methyl-D-aspartate receptors in osteoblasts. *FASEB Journal* **17**, 1532-4 (2003).
220. Genever, P.G. et al. Expression of a functional N-methyl-D-aspartate-type glutamate receptor by bone marrow megakaryocytes. *Blood* **93**, 2876-83 (1999).
221. Mentaverri, R. et al. Regulation of bone resorption and osteoclast survival by nitric oxide: possible involvement of NMDA-receptor. *Journal of Cellular Biochemistry* **88**, 1145-56 (2003).
222. Merle, B., Itzstein, C., Delmas, P.D. & Chenu, C. NMDA glutamate receptors are expressed by osteoclast precursors and involved in the regulation of osteoclastogenesis. *Journal of Cellular Biochemistry* **90**, 424-36 (2003).
223. Inagaki, N. et al. Expression and role of ionotropic glutamate receptors in pancreatic islet cells. *FASEB Journal* **9**, 686-91 (1995).
224. Marquard, J. et al. Characterization of pancreatic NMDA receptors as possible drug targets for diabetes treatment. *Nature Medicine* **21**, 363-72 (2015).
225. Anderson, M., Suh, J.M., Kim, E.Y. & Dryer, S.E. Functional NMDA receptors with atypical properties are expressed in podocytes. *American Journal of Physiology: Cell Physiology* **300**, C22-32 (2011).
226. Sproul, A. et al. N-methyl-D-aspartate receptor subunit NR3a expression and function in principal cells of the collecting duct. *American Journal of Physiology: Renal Physiology* **301**, F44-F54 (2011).
227. Deng, A. & Thomson, S.C. Renal NMDA receptors independently stimulate proximal reabsorption and glomerular filtration. *American Journal of Physiology: Renal Physiology* **296**, F976-82 (2009).
228. Boldyrev, A.A. Homocysteinic acid causes oxidative stress in lymphocytes by potentiating toxic effect of NMDA. *Bulletin of Experimental Biology and Medicine* **140**, 33-7 (2005).
229. Gill, S.S. & Pulido, O.M. Glutamate receptors in peripheral tissues: current knowledge, future research, and implications for toxicology. *Toxicologic Pathology* **29**, 208-23 (2001).
230. Hitchcock, I.S., Skerry, T.M., Howard, M.R. & Genever, P.G. NMDA receptor-mediated regulation of human megakaryocytopoiesis. *Blood* **102**, 1254-9 (2003).
231. Malayev, A., Gibbs, T.T. & Farb, D.H. Inhibition of the NMDA response by pregnenolone sulphate reveals subtype selective modulation of NMDA receptors by sulphated steroids. *British Journal of Pharmacology* **135**, 901-9 (2002).
232. Horák, M., Vlček, K., Chodounská, H. & Vyklický, L., Jr. Subtype-dependence of N-methyl-D-aspartate receptor modulation by pregnenolone sulfate. *Neuroscience* **137**, 93-102 (2006).
233. Schubring, S.R., Fleischer, W., Lin, J.-S., Haas, H.L. & Sergeeva, O.A. The bile steroid chenodeoxycholate is a potent antagonist at NMDA and GABA(A) receptors. *Neuroscience Letters* **506**, 322-6 (2012).
234. Pilo de la Fuente, B., Ruiz, I., Lopez-de-Munain, A. & Jiménez-Escrig, A. Cerebrotendinous xanthomatosis: neuropathological findings. *Journal of Neurology* **255**, 839-42 (2008).

235. Keitel, V. et al. Expression and localization of hepatobiliary transport proteins in progressive familial intrahepatic cholestasis. *Hepatology* **41**, 1160-72 (2005).
236. Glantz, A., Marschall, H.-U. & Mattsson, L.-Å. Intrahepatic cholestasis of pregnancy: Relationships between bile acid levels and fetal complication rates. *Hepatology* **40**, 467-74 (2004).
237. Makino, I., Hashimoto, H., Shinozaki, K., Yoshino, K. & Nakagawa, S. Sulfated and nonsulfated bile acids in urine, serum, and bile of patients with hepatobiliary diseases. *Gastroenterology* **68**, 545-53 (1975).
238. Fesenko, E.E., Kolesnikov, S.S. & Lyubarsky, A.L. Induction by cyclic GMP of cationic conductance in plasma membrane of retinal rod outer segment. *Nature* **313**, 310-3 (1985).
239. Burns, M.E. & Arshavsky, V.Y. Beyond counting photons: trials and trends in vertebrate visual transduction. *Neuron* **48**, 387-401 (2005).
240. Lancet, D. Vertebrate olfactory reception. *Annual Reviews Neuroscience* **9**, 329-55 (1986).
241. Nakamura, T. & Gold, G.H. A cyclic nucleotide-gated conductance in olfactory receptor cilia. *Nature* **325**, 442-4 (1987).
242. Menini, A. Cyclic nucleotide-gated channels in visual and olfactory transduction. *Biophysical Chemistry* **55**, 185-96 (1995).
243. Varnum, M.D. & Dai, G. Cyclic nucleotide-gated channels. in *Handbook of ion-channels* (eds. Zheng, J. & Trudeau, M.C.) 361-382 (Taylor & Francis Group, LLC, Boca Raton, Florida, 2015).
244. Kaupp, U.B. et al. Primary structure and functional expression from complementary DNA of the rod photoreceptor cyclic GMP-gated channel. *Nature* **342**, 762-6 (1989).
245. Dhallan, R.S., Yau, K.-W., Schrader, K.A. & Reed, R.R. Primary structure and functional expression of a cyclic nucleotide-activated channel from olfactory neurons. *Nature* **347**, 184-7 (1990).
246. Bönigk, W. et al. Rod and cone photoreceptor cells express distinct genes for cGMP-gated channels. *Neuron* **10**, 865-77 (1993).
247. Chen, T.-Y. et al. A new subunit of the cyclic nucleotide-gated cation channel in retinal rods. *Nature* **362**, 764-7 (1993).
248. Bradley, J., Li, J., Davidson, N., Lester, H.A. & Zinn, K. Heteromeric olfactory cyclic nucleotide-gated channels: a subunit that confers increased sensitivity to cAMP. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 8890-4 (1994).
249. Liman, E.R. & Buck, L.B. A second subunit of the olfactory cyclic nucleotide-gated channel confers high sensitivity to cAMP. *Neuron* **13**, 611-21 (1994).
250. Gerstner, A., Zong, X., Hofmann, F. & Biel, M. Molecular cloning and functional characterization of a new modulatory cyclic nucleotide-gated channel subunit from mouse retina. *Journal of Neuroscience* **20**, 1324-32 (2000).
251. Weitz, D., Ficek, N., Kremmer, E., Bauer, P.J. & Kaupp, U.B. Subunit stoichiometry of the CNG channel of rod photoreceptors. *Neuron* **36**, 881-9 (2002).
252. Zhong, H., Molday, L.L., Molday, R.S. & Yau, K.-W. The heteromeric cyclic nucleotide-gated channel adopts a 3A:1B stoichiometry. *Nature* **420**, 193-8 (2002).
253. Zheng, J., Trudeau, M.C. & Zagotta, W.N. Rod cyclic nucleotide-gated channels have a stoichiometry of three CNGB1 subunits and one CNGB1 subunit. *Neuron* **36**, 891-6 (2002).
254. Peng, C., Rich, E.D. & Varnum, M.D. Subunit configuration of heteromeric cone cyclic nucleotide-gated channels. *Neuron* **42**, 401-10 (2004).
255. Zheng, J. & Zagotta, W.N. Stoichiometry and assembly of olfactory cyclic nucleotide-gated channels. *Neuron* **42**, 411-21 (2004).
256. Zheng, J. & Trudeau, M.C. *Handbook of ion channels*, xx, 671 pages (CRC Press, Boca Raton, 2015).
257. Gao, Y., Cao, E., Julius, D. & Cheng, Y. TRPV1 structures in nanodiscs reveal mechanisms of ligand and lipid action. *Nature* **534**, 347-51 (2016).
258. Craven, K.B. & Zagotta, W.N. CNG and HCN channels: two peas, one pod. *Annual Review of Physiology* **68**, 375-401 (2006).
259. James, Z.M. & Zagotta, W.N. Structural insights into the mechanisms of CNBD channel function. *Journal of General Physiology* **150**, 225-244 (2018).
260. Shuart, N.G., Haitin, Y., Camp, S.S., Black, K.D. & Zagotta, W.N. Molecular mechanism for 3:1 subunit stoichiometry of rod cyclic nucleotide-gated ion channels. *Nature Communications* **2**, 457 (2011).
261. Song, Y. et al. High-resolution comparative modeling with RosettaCM. *Structure* **21**, 1735-42 (2013).
262. Alford, R.F. et al. An Integrated Framework Advancing Membrane Protein Modeling and Design. *PLOS Computational Biology* **11**, e1004398 (2015).
263. Li, M. et al. Structure of a eukaryotic cyclic-nucleotide-gated channel. *Nature* **542**, 60-65 (2017).
264. Buchan, D.W.A., Minneci, F., Nugent, T.C.O., Bryson, K. & Jones, D.T. Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Research* **41**, W349-57 (2013).

265. Buchan, D.W.A. & Jones, D.T. The PSIPRED Protein Analysis Workbench: 20 years on. *Nucleic Acids Research* **47**, W402-W407 (2019).
266. Varnum, M.D., Black, K.D. & Zagotta, W.N. Molecular mechanism for ligand discrimination of cyclic nucleotide-gated channels. *Neuron* **15**, 619-25 (1995).
267. Tibbs, G.R., Liu, D.T., Leybold, B.G. & Siegelbaum, S.A. A state-independent interaction between ligand and a conserved arginine residue in cyclic nucleotide-gated channels reveals a functional polarity of the cyclic nucleotide binding site. *Journal of Biological Chemistry* **273**, 4497-505 (1998).
268. Altenhofen, W. et al. Control of ligand specificity in cyclic nucleotide-gated channels from rod photoreceptors and olfactory epithelium. *Proceedings of the National Academy of Sciences of the United States of America* **88**, 9868-72 (1991).
269. Flynn, G.E., Black, K.D., Islas, L.D., Sankaran, B. & Zagotta, W.N. Structure and rearrangements in the carboxy-terminal region of SpIH channels. *Structure* **15**, 671-82 (2007).
270. Zimmerman, A.L., Yamanaka, G., Eckstein, F., Baylor, D.A. & Stryer, L. Interaction of hydrolysis-resistant analogs of cyclic GMP with the phosphodiesterase and light-sensitive channel of retinal rod outer segments. *Proceedings of the National Academy of Sciences of the United States of America* **82**, 8813-7 (1985).
271. Tanaka, J.C., Eccleston, J.F. & Furman, R.E. Photoreceptor channel activation by nucleotide derivatives. *Biochemistry* **28**, 2776-84 (1989).
272. Wei, J.-Y., Cohen, E.D., Genieser, H.-G. & Barnstable, C.J. Substituted cGMP analogs can act as selective agonists of the rod photoreceptor cGMP-gated cation channel. *Journal of Molecular Neuroscience* **10**, 53-64 (1998).
273. Rall, T.W., Sutherland, E.W. & Berthet, J. The relationship of epinephrine and glucagon to liver phosphorylase. IV. Effect of epinephrine and glucagon on the reactivation of phosphorylase in liver homogenates. *Journal of Biological Chemistry* **224**, 463-75 (1957).
274. Sutherland, E.W. & Rall, T.W. Fractionation and characterization of a cyclic adenine ribonucleotide formed by tissue particles. *Journal of Biological Chemistry* **232**, 1077-91 (1958).
275. Zaccolo, M., Zerio, A. & Lobo, M.J. Subcellular Organization of the cAMP Signaling Pathway. *Pharmacological Reviews* **73**, 278-309 (2021).
276. Levitzki, A. From epinephrine to cyclic AMP. *Science* **241**, 800-6 (1988).
277. Walsh, D.A., Perkins, J.P. & Krebs, E.G. An adenosine 3',5'-monophosphate-dependant protein kinase from rabbit skeletal muscle. *J Biol Chem* **243**, 3763-5 (1968).
278. Krebs, E.G. & Beavo, J.A. Phosphorylation-dephosphorylation of enzymes. *Annual Review of Biochemistry* **48**, 923-59 (1979).
279. Santoro, B. & Tibbs, G.R. The HCN gene family: molecular basis of the hyperpolarization-activated pacemaker channels. *Annals of the New York Academy of Sciences* **868**, 741-64 (1999).
280. de Rooij, J. et al. Epac is a Rap1 guanine-nucleotide-exchange factor directly activated by cyclic AMP. *Nature* **396**, 474-7 (1998).
281. Kawasaki, H. et al. A family of cAMP-binding proteins that directly activate Rap1. *Science* **282**, 2275-9 (1998).
282. Andrée, B. et al. Isolation and characterization of the novel popeye gene family expressed in skeletal muscle and heart. *Developmental Biology* **223**, 371-82 (2000).
283. Brand, T. The Popeye domain-containing gene family. *Cell Biochemistry and Biophysics* **43**, 95-103 (2005).
284. Schindler, R.F.R. & Brand, T. The Popeye domain containing protein family--A novel class of cAMP effectors with important functions in multiple tissues. *Progress in Biophysics and Molecular Biology* **120**, 28-36 (2016).
285. Berman, H.M. et al. The cAMP binding domain: an ancient signaling module. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 45-50 (2005).
286. Diller, T.C., Madhusudan, Xuong, N.-H. & Taylor, S.S. Molecular basis for regulatory subunit diversity in cAMP-dependent protein kinase: crystal structure of the type II beta regulatory subunit. *Structure* **9**, 73-82 (2001).
287. Christensen, A.E. et al. cAMP analog mapping of Epac1 and cAMP kinase. Discriminating analogs demonstrate that Epac and cAMP kinase act synergistically to promote PC-12 cell neurite extension. *Journal of Biological Chemistry* **278**, 35394-402 (2003).
288. Simmons, M.A. Functional selectivity, ligand-directed trafficking, conformation-specific agonism: what's in a name? *Molecular Interventions* **5**, 154-7 (2005).
289. Urban, J.D. et al. Functional selectivity and classical concepts of quantitative pharmacology. *Journal of Pharmacology and Experimental Therapeutics* **320**, 1-13 (2007).

290. Laprairie, R.B., Bagher, A.M., Kelly, M.E.M., Dupré, D.J. & Denovan-Wright, E.M. Type 1 cannabinoid receptor ligands display functional selectivity in a cell culture model of striatal medium spiny projection neurons. *Journal of Biological Chemistry* **289**, 24845-62 (2014).
291. Kenakin, T. Biased Receptor Signaling in Drug Discovery. *Pharmacological Reviews* **71**, 267-315 (2019).
292. Violin, J.D. & Lefkowitz, R.J. Beta-arrestin-biased ligands at seven-transmembrane receptors. *Trends in Pharmacological Sciences* **28**, 416-22 (2007).
293. Wardell, S.E., Marks, J.R. & McDonnell, D.P. The turnover of estrogen receptor alpha by the selective estrogen receptor degrader (SERD) fulvestrant is a saturable process that is not required for antagonist efficacy. *Biochemical Pharmacology* **82**, 122-30 (2011).
294. Norris, J.D. et al. Peptide antagonists of the human estrogen receptor. *Science* **285**, 744-6 (1999).
295. Paige, L.A. et al. Estrogen receptor (ER) modulators each induce distinct conformational changes in ER alpha and ER beta. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 3999-4004 (1999).
296. Kovalenko, M. et al. Phosphorylation site-specific inhibition of platelet-derived growth factor beta-receptor autophosphorylation by the receptor blocking typhostin AG1296. *Biochemistry* **36**, 6260-9 (1997).
297. Simon, D.I. Opening the field of integrin biology to "biased agonism". *Circulation Research* **109**, 1199-201 (2011).
298. Wolf, D. et al. Binding of CD40L to Mac-1's I-domain involves the EQLKKSRTL motif and mediates leukocyte recruitment and atherosclerosis--but does not affect immunity and thrombosis in mice. *Circulation Research* **109**, 1269-79 (2011).
299. Sitrin, R.G., Pan, P.M., Srikanth, S. & Todd, R.F., III. Fibrinogen activates NF-kappa B transcription factors in mononuclear phagocytes. *Journal of Immunology* **161**, 1462-70 (1998).
300. Shi, C., Zhang, X., Chen, Z., Robinson, M.K. & Simon, D.I. Leukocyte integrin Mac-1 recruits toll/interleukin-1 receptor superfamily signaling intermediates to modulate NF-kappaB activity. *Circulation Research* **89**, 859-65 (2001).
301. Shi, C. et al. Integrin engagement regulates monocyte differentiation through the forkhead transcription factor Foxp1. *Journal of Clinical Investigation* **114**, 408-18 (2004).
302. Shi, C. et al. Down-regulation of the forkhead transcription factor Foxp1 is required for monocyte differentiation and macrophage function. *Blood* **112**, 4699-711 (2008).
303. Springer, T.A. & Wang, J.-H. The three-dimensional structure of integrins and their ligands, and conformational regulation of cell adhesion. *Advances in Protein Chemistry* **68**, 29-63 (2004).
304. Kurz, A.K., Graf, D., Schmitt, M., vom Dahl, S. & Häussinger, D. Tauroursodesoxycholate-Induced Cholestasis Involves p38(MAPK) Activation and Translocation of the Bile Salt Export Pump in Rats. *Gastroenterology* **121**, 407-19 (2001).
305. Schmitt, M., Kubitz, R., Lizun, S., Wettstein, M. & Häussinger, D. Regulation of the Dynamic Localization of the Rat Bsep Gene-Encoded Bile Salt Export Pump by Anisoosmolarity. *Hepatology* **33**, 509-18 (2001).
306. Xiao, T., Takagi, J., Collier, B.S., Wang, J.-H. & Springer, T.A. Structural basis for allostery in integrins and binding to fibrinogen-mimetic therapeutics. *Nature* **432**, 59-67 (2004).
307. Zhu, J. et al. Structure of a Complete Integrin Ectodomain in a Physiologic Resting State and Activation and Deactivation by Applied Forces. *Molecular Cell* **32**, 849-61 (2008).
308. Puklin-Faucher, E. & Vogel, V. Integrin activation dynamics between the RGD-binding site and the headpiece hinge. *Journal of Biological Chemistry* **284**, 36557-68 (2009).
309. Puklin-Faucher, E., Gao, M., Schulten, K. & Vogel, V. How the headpiece hinge angle is opened: New insights into the dynamics of integrin activation. *Journal of Cell Biology* **175**, 349-60 (2006).
310. Craig, D., Gao, M., Schulten, K. & Vogel, V. Structural insights into how the MIDAS ion stabilizes integrin binding to an RGD peptide under force. *Structure* **12**, 2049-58 (2004).
311. Van Aghthoven, J.F. et al. Structural basis for pure antagonism of integrin $\alpha V\beta 3$ by a high-affinity form of fibronectin. *Nature Structural & Molecular Biology* **21**, 383-8 (2014).
312. Bochen, A. et al. Biselectivity of isoDGR peptides for fibronectin binding integrin subtypes $\alpha 5\beta 1$ and $\alpha v\beta 6$: conformational control through flanking amino acids. *Journal of Medicinal Chemistry* **56**, 1509-19 (2013).
313. Mas-Moruno, C., Rechenmacher, F. & Kessler, H. Cilengitide: the first anti-angiogenic small molecule drug candidate design, synthesis and clinical evaluation. *Anti-Cancer Agents in Medicinal Chemistry* **10**, 753-68 (2010).
314. Schliess, F., Kurz, A.K., vom Dahl, S. & Häussinger, D. Mitogen-Activated Protein Kinases Mediate the Stimulation of Bile Acid Secretion by Tauroursodesoxycholate in Rat Liver. *Gastroenterology* **113**, 1306-14 (1997).

315. García, A.J. & Boettiger, D. Integrin-fibronectin interactions at the cell-material interface: initial integrin binding and signaling. *Biomaterials* **20**, 2427-33 (1999).
316. Mitra, S.K., Hanson, D.A. & Schlaepfer, D.D. Focal adhesion kinase: in command and control of cell motility. *Nature Reviews Molecular Cell Biology* **6**, 56-68 (2005).
317. Shah, B.H., Neithardt, A., Chu, D.B., Shah, F.B. & Catt, K.J. Role of EGF receptor transactivation in phosphoinositide 3-kinase-dependent activation of MAP kinase by GPCRs. *Journal of Cellular Physiology* **206**, 47-57 (2006).
318. Danbolt, N.C. Glutamate uptake. *Progress in Neurobiology* **65**, 1-105 (2001).
319. Kew, J.N.C. & Kemp, J.A. Ionotropic and metabotropic glutamate receptor structure and pharmacology. *Psychopharmacology* **179**, 4-29 (2005).
320. Moriyoshi, K. et al. Molecular cloning and characterization of the rat NMDA receptor. *Nature* **354**, 31-7 (1991).
321. Kleckner, N.W. & Dingledine, R. Requirement for glycine in activation of NMDA-receptors expressed in *Xenopus* oocytes. *Science* **241**, 835-7 (1988).
322. Bard, L. & Groc, L. Glutamate receptor dynamics and protein interaction: lessons from the NMDA receptor. *Molecular and Cellular Neuroscience* **48**, 298-307 (2011).
323. Kokane, S.S., Gong, K., Jin, J. & Lin, Q. Prolonged ketamine exposure induces increased activity of the GluN2B-containing N-methyl-d-aspartate receptor in the anterior cingulate cortex of neonatal rats. *Neurotoxicology and Teratology* **63**, 1-8 (2017).
324. Zhang, J.-B. et al. Structural Basis of the Proton Sensitivity of Human GluN1-GluN2A NMDA Receptors. *Cell Reports* **25**, 3582-3590 e4 (2018).
325. Moghadasian, M.H., Salen, G., Fröhlich, J.J. & Scudamore, C.H. Cerebrotendinous xanthomatosis: a rare disease with diverse manifestations. *Archives of Neurology* **59**, 527-9 (2002).
326. Salen, G. et al. Increased concentrations of cholestanol and apolipoprotein B in the cerebrospinal fluid of patients with cerebrotendinous xanthomatosis. Effect of chenodeoxycholic acid. *New England Journal of Medicine* **316**, 1233-8 (1987).
327. Sherman, W., Day, T., Jacobson, M.P., Friesner, R.A. & Farid, R. Novel procedure for modeling ligand/receptor induced fit effects. *Journal of Medicinal Chemistry* **49**, 534-53 (2006).
328. Kuzmanic, A. & Zagrovic, B. Determination of ensemble-average pairwise root mean-square deviation from experimental B-factors. *Biophysical Journal* **98**, 861-71 (2010).
329. Richieri, G.V., Low, P.J., Ogata, R.T. & Kleinfeld, A.M. Mutants of rat intestinal fatty acid-binding protein illustrate the critical role played by enthalpy-entropy compensation in ligand binding. *Journal of Biological Chemistry* **272**, 16737-40 (1997).
330. Scott, D.W. *Multivariate density estimation : theory, practice, and visualization*, xviii, 350 pages (John Wiley & Sons, Inc., Hoboken, New Jersey, 2015).
331. Hamelberg, D., Mongan, J. & McCammon, J.A. Accelerated molecular dynamics: a promising and efficient simulation method for biomolecules. *Journal of Chemical Physics* **120**, 11919-29 (2004).
332. Schmidtke, P., Bidon-Chanal, A., Luque, F.J. & Barril, X. MDpocket: open-source cavity detection and characterization on molecular dynamics trajectories. *Bioinformatics* **27**, 3276-85 (2011).
333. Schmidtke, P. & Barril, X. Understanding and predicting druggability. A high-throughput method for detection of drug binding sites. *Journal of Medicinal Chemistry* **53**, 5858-67 (2010).
334. Santoro, B. et al. Identification of a gene encoding a hyperpolarization-activated pacemaker channel of brain. *Cell* **93**, 717-29 (1998).
335. Kaupp, U.B. & Seifert, R. Cyclic nucleotide-gated ion channels. *Physiological Reviews* **82**, 769-824 (2002).
336. Matulef, K. & Zagotta, W.N. Cyclic nucleotide-gated ion channels. *Annual Review of Cell and Developmental Biology* **19**, 23-44 (2003).
337. Moosmang, S. et al. Cellular expression and functional characterization of four hyperpolarization-activated pacemaker channels in cardiac and neuronal tissues. *European Journal of Biochemistry* **268**, 1646-52 (2001).
338. Brown, H.F., Difrancesco, D. & Noble, S.J. How Does Adrenaline Accelerate the Heart. *Nature* **280**, 235-236 (1979).
339. DiFrancesco, D. Characterization of single pacemaker channels in cardiac sino-atrial node cells. *Nature* **324**, 470-3 (1986).
340. Wang, J., Chen, S. & Siegelbaum, S.A. Regulation of hyperpolarization-activated HCN channel gating and cAMP modulation due to interactions of COOH terminus and core transmembrane regions. *Journal of General Physiology* **118**, 237-50 (2001).
341. Craven, K.B. & Zagotta, W.N. Salt bridges and gating in the COOH-terminal region of HCN2 and CNGA1 channels. *Journal of General Physiology* **124**, 663-77 (2004).

342. Zhou, L., Olivier, N.B., Yao, H., Young, E.C. & Siegelbaum, S.A. A conserved tripeptide in CNG and HCN channels regulates ligand gating by controlling C-terminal oligomerization. *Neuron* **44**, 823-34 (2004).
343. Brown, R.L., Strassmaier, T., Brady, J.D. & Karpen, J.W. The pharmacology of cyclic nucleotide-gated channels: emerging from the darkness. *Current Pharmaceutical Design* **12**, 3597-613 (2006).
344. Strassmaier, T. & Karpen, J.W. Novel N7- and N1-substituted cGMP derivatives are potent activators of cyclic nucleotide-gated channels. *Journal of Medicinal Chemistry* **50**, 4186-94 (2007).
345. Wu, S. et al. State-dependent cAMP binding to functioning HCN channels studied by patch-clamp fluorometry. *Biophysical Journal* **100**, 1226-32 (2011).
346. Bois, P., Renaudon, B., Baruscotti, M., Lenfant, J. & DiFrancesco, D. Activation of f-channels by cAMP analogues in macropatches from rabbit sino-atrial node myocytes. *Journal of Physiology* **501** (Pt 3), 565-71 (1997).
347. Scott, S.-P., Shea, P.W. & Dryer, S.E. Mapping ligand interactions with the hyperpolarization activated cyclic nucleotide modulated (HCN) ion channel binding domain using a soluble construct. *Biochemistry* **46**, 9417-31 (2007).
348. Möller, S. et al. Cyclic nucleotide mapping of hyperpolarization-activated cyclic nucleotide-gated (HCN) channels. *ACS Chemical Biology* **9**, 1128-37 (2014).
349. Bash, P.A., Ho, L.L., MacKerell, A.D., Jr., Levine, D. & Hallstrom, P. Progress toward chemical accuracy in the computer simulation of condensed phase reactions. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 3698-703 (1996).
350. Friesner, R.A. et al. Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *Journal of Medicinal Chemistry* **49**, 6177-96 (2006).
351. Homeyer, N., Stoll, F., Hillisch, A. & Gohlke, H. Binding Free Energy Calculations for Lead Optimization: Assessment of Their Accuracy in an Industrial Drug Design Context. *Journal of Chemical Theory and Computation* **10**, 3331-44 (2014).
352. Gohlke, H. & Case, D.A. Converging free energy estimates: MM-PB(GB)SA studies on the protein-protein complex Ras-Raf. *Journal of Computational Chemistry* **25**, 238-50 (2004).
353. Gohlke, H., Kiel, C. & Case, D.A. Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-RalGDS complexes. *Journal of Molecular Biology* **330**, 891-913 (2003).
354. Fujita, T., Umemura, M., Yokoyama, U., Okumura, S. & Ishikawa, Y. The role of Epac in the heart. *Cellular and Molecular Life Sciences* **74**, 591-606 (2017).
355. Cheng, X., Ji, Z., Tsalkova, T. & Mei, F. Epac and PKA: a tale of two intracellular cAMP receptors. *Acta Biochimica et Biophysica Sinica* **40**, 651-62 (2008).
356. Kaupp, U.B. & Seifert, R. Molecular diversity of pacemaker ion channels. *Annual Review of Physiology* **63**, 235-57 (2001).
357. Rehmann, H., Wittinghofer, A. & Bos, J.L. Capturing cyclic nucleotides in action: snapshots from crystallographic studies. *Nature Reviews Molecular Cell Biology* **8**, 63-73 (2007).
358. Døskeland, S.O. et al. Mapping of the two intrachain cyclic nucleotide binding sites of adenosine cyclic 3',5'-phosphate dependent protein kinase I. *Biochemistry* **22**, 1094-101 (1983).
359. Øgreid, D., Ekanger, R., Suva, R.H., Miller, J.P. & Døskeland, S.O. Comparison of the two classes of binding sites (A and B) of type I and type II cyclic-AMP-dependent protein kinases by using cyclic nucleotide analogs. *European Journal of Biochemistry* **181**, 19-31 (1989).
360. Zagotta, W.N. et al. Structural basis for modulation and agonist specificity of HCN pacemaker channels. *Nature* **425**, 200-5 (2003).