

Molecular Modelling Studie zur spezifischen Keratinbindung

Inaugural-Dissertation

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

vorgelegt von

Matthias Volkmar

aus Saalfeld

Düsseldorf 2008

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

- | | |
|--------------|----------------------------------|
| 1. Gutachter | Prof. Dr. Dr. h. c. H.-D. Höltje |
| 2. Gutachter | PD Dr. Klaus-Jürgen Schleifer |

Tag der mündlichen Prüfung: 28.04.2008

Die Farbversion dieser Arbeit ist im Internet unter: www.ulb.uni-duesseldorf.de
unter der Auswahl Elektronische Dissertationen einsehbar.

Inhaltsverzeichnis

| | | |
|----------|--|-----------|
| 1 | EINLEITUNG..... | 7 |
| 1.1 | ALLGEMEINE EINLEITUNG | 8 |
| 1.2 | INTERMEDIÄRFILAMENTE | 9 |
| 1.2.1 | Allgemeines | 9 |
| 1.2.2 | Molekulare Architektur der Intermediärfilamente | 10 |
| 1.2.3 | Die verschiedenen Typen der Intermediärfilamente | 17 |
| 1.3 | DAS C-TERMINALE ENDE VON DESMOPLAKIN..... | 20 |
| 1.3.1 | Allgemeines | 20 |
| 1.3.2 | Struktureller Aufbau von Desmoplakin..... | 20 |
| 1.3.3 | Molekulare Architektur der DPCT Domäne B und C..... | 21 |
| 1.4 | BINDUNGSSTUDIEN ZUR DESMOPLAKIN - IF WECHSELWIRKUNG 23 | |
| 1.4.1 | Bindungsstudien aus der Literatur | 23 |
| 1.4.2 | Bindungsstudien des Kooperationspartners BASF | 24 |
| 1.5 | HYPOTHESE DER AN DER BINDUNG BETEILIGTEN HAAR - KERATINE..... | 24 |
| 1.6 | HYPOTHESE DER AN DER BINDUNG BETEILIGTEN HAUT - KERATINE..... | 25 |
| 2 | AUSGANGSLAGE UND ZIELSETZUNG | 26 |
| 2.1 | AUSGANGSLAGE DIESER ARBEIT | 27 |
| 2.2 | ZIELSETZUNG DIESER ARBEIT | 27 |
| 3 | METHODEN | 28 |
| 3.1 | SEQUENZALIGNMENT..... | 29 |

| | | |
|----------|--|-----------|
| 3.2 | „LOOP SEARCH“ | 29 |
| 3.3 | VORHERSAGE EINER „COILED COIL“-STRUKTUR..... | 30 |
| 3.4 | KONSTRUKTION VON „BACKBONE“-STRUKTURKOORDINATEN DES KERATINMODELLS | 30 |
| 3.5 | ANFÜGEN VON AMINOSÄURESEITENKETTEN | 30 |
| 3.6 | KRAFTFELDMETHODEN | 31 |
| 3.6.1 | Geometrieoptimierung durch Energieminimierung | 31 |
| 3.6.2 | Moleküldynamik-Simulation..... | 32 |
| 3.7 | ANALYSE DER QUALITÄT EINES PROTEINMODELLS | 34 |
| 3.7.1 | Procheck - stereochemische Qualität einer Proteinstruktur..... | 34 |
| 3.8 | DOCKINGMETHODEN FÜR PROTEIN-PROTEIN-INTERAKTIONEN. | 35 |
| 3.8.1 | Hex 4.1 -Software für „Rigid Body Docking“ | 36 |
| 3.8.2 | HADDOCK 1.3 | 37 |
| 4 | DURCHFÜHRUNG UND ERGEBNISSE | 41 |
| 4.1 | PROTEIN-HOMOLOGIE-MODELLING | 42 |
| 4.1.1 | Vervollständigung der Kristallstruktur der Domäne B | 42 |
| 4.1.2 | Vervollständigung der Kristallstruktur der Domäne C..... | 43 |
| 4.1.3 | Konstruktion der Strukturmodelle von Keratinen der Haut und des Haares | 44 |
| 4.2 | AUF DER SUCHE NACH DEN BINDUNGSREGIONEN..... | 56 |
| 4.2.1 | Studium von Proteinkomplexen aus der PDB | 57 |
| 4.2.2 | Die Bindungsregion auf der Domäne B als funktionales Dimer | 60 |
| 4.2.3 | Das erste Experiment resultierend aus der „Monomer-Hypothese“ | 62 |
| 4.2.4 | Die Bindungsregion auf der Domäne B als Monomer - theoretische und experimentelle Verifizierung der „Monomer-Hypothese“ | 65 |
| 4.2.5 | Mutationsvorschläge Teil 1 - Lokalisierung der Bindungsregion | 70 |
| 4.2.6 | Die Suche nach der Bindungsregion auf dem Keratin..... | 72 |

| | | |
|----------|--|------------|
| 4.2.7 | Experimentelle Ergebnisse der Mutations-/Bindungsstudien des Kooperationspartners BASF | 81 |
| 4.2.8 | Nachvollziehbarkeit der Mutationsstudien durch die „LEKG“ | 83 |
| 4.2.9 | Mutationsvorschläge für verbesserte Bindungseigenschaften der Domäne B im Sinne des Hauptzieles..... | 85 |
| 5 | DISKUSSION | 98 |
| 5.1 | ALLGEMEIN | 99 |
| 5.2 | KRAFTFELDMETHODEN UND MUTATIONSVORSCHLÄGE | 101 |
| 5.3 | VERVOLLSTÄNDIGUNG DER KRISTALLSTRUKTUR DER DOMÄNE B 103 | |
| 5.4 | KONSTRUKTION DER STRUKTURMODELLE VON KERATINEN DER HAUT UND DES HAARES | 104 |
| 5.5 | EXPERIMENTELLE ERGEBNISSE DES KOOPERATIONSPARTNERS BASF | 105 |
| 6 | ZUSAMMENFASSUNG | 106 |
| 6.1 | ZUSAMMENFASSUNG | 107 |
| 6.2 | SUMMARY | 108 |
| 7 | ANHANG..... | 109 |
| 7.1 | ABKÜRZUNGEN | 109 |
| 7.2 | EINHEITEN..... | 110 |
| 7.3 | CURRICULUM VITAE | 111 |
| 7.4 | DANKSAGUNG | 111 |

| | | |
|----------|----------------------------------|------------|
| 8 | LITERATURVERZEICHNIS..... | 113 |
|----------|----------------------------------|------------|

1 Einleitung

1.1 Allgemeine Einleitung

Biomoleküle, Wasser und anorganische Ionen, das sind die Grundbausteine einer Zelle. Das Wissen um zelluläre Systeme ist in den vergangenen 100 Jahren geradezu explodiert. Immer besser werdende Analysemethoden erlauben einen noch detaillierteren Einblick in den strukturellen Aufbau einer Zelle, d.h. in welcher Konstellation die einzelnen Elemente zueinander stehen und welche Funktion sie ausüben. Ständig wächst die Zahl der beschriebenen Spezies von Biomolekülen. Doch die Unkenntnis um deren genaue, dynamische Wechselwirkungen und globale Eigenschaften vergrößert sich ebenso.

Physikalisch betrachtet ist eine Zelle weder flüssig noch fest. Viele Bestandteile befinden sich an der Phasengrenze dieser beiden Aggregatzustände. Kleine Moleküle, insbesondere Ionen sind zwar im Zytosol gelöst, die großen funktionalen Strukturelemente aber befinden sich im flüssig-kristallinen Zustand. Keiner vermag heutzutage abzuschätzen, aus wie vielen molekularen Elementen eine Zelle besteht. Doch muss diese Zahl, so viel steht zumindest fest, unglaublich groß sein. Wie aber ist es möglich, dass so viele Strukturen mit höchster Präzision untereinander wechselwirken, um das zu ergeben, was man als belebte Materie im biologischen Sinne bezeichnet? Je intensiver man sich mit dieser Problematik auseinandersetzt, desto staunender wird man vor der Natur stehen. Angesichts dieser ungeheuren Komplexität scheint es auf den ersten Blick schier unmöglich diesem Geheimnis näher zu kommen. Bei genauerer detaillierter Betrachtung lässt sich jedoch feststellen, dass es gewisse Regeln zu geben scheint. Es wird deutlich, dass sich ähnliche Systeme bzw. Strukturelemente unter ähnlichen Bedingungen auch ähnlich verhalten. Ähnlichkeit bezieht sich dabei auf die prinzipiellen räumlich verteilten physikalischen Wechselwirkungsfelder, bedingt durch die jeweilige chemische Struktur. Diese Potentialfeldverteilungen sind aber im zeitlichen Verlauf keineswegs konstant. Vielmehr stellen die in einer Zelle ablaufenden Reaktionen zeitlich hochveränderliche, dynamische Prozesse dar, die man allein auf experimentellem Wege nur unzureichend studieren und

charakterisieren kann. Die theoretische Chemie bietet jedoch die Möglichkeit ein Modell von der Natur zu kreieren. Anhand dieses Modells lassen sich dann die veränderlichen Eigenschaften über einen bestimmten Zeitraum genau beobachten und ableiten. Die Kombination, aus experimentellen Befunden und theoretisch chemischer Betrachtungsweise ist daher bestens geeignet, auf vergleichsweise hohem Niveau zu verstehen, wie bestimmte zelluläre Mechanismen auf molekularer Ebene funktionieren.

Diese Arbeit, eine „Molecular Modelling“ Studie, beleuchtet die bindenden Wechselwirkungen zwischen Proteinen aus der Familie humaner Intermediärfilamente und einem Protein aus der Plectinfamilie. Die daraus resultierende, theoretische Voraussage der Wechselwirkungsgeometrie wurde im Anschluss einer experimentellen Überprüfung in Form einer Mutationsstudie durch den Kooperationspartner BASF unterzogen.

1.2 Intermediärfilamente

1.2.1 Allgemeines

Intermediärfilamente (IF) bilden zusammen mit den Microtubuli und Actinfilamenten das Zytoskelettes in höheren eukariotischen, tierischen Zellen. Der Name Intermediäre Filamente leitet sich aus der Größe ihrer Durchmesser ab. Denn sie liegen mit 10 nm zwischen den Aktinfilamenten mit 7 nm und Microtubuli mit 25 nm [1]. Sie kommen sowohl im Cytoplasma als auch im Nucleus verschiedener Zelltypen vor. Zusammen mit diversen anderen assoziierten Proteinen bilden sie ein großes Netzwerk und verleihen der Zelle auf diese Weise eine gewisse mechanische Integrität. Außerdem sind sie in kritische Prozesse der Zellteilung, Zellmigration sowie der plastischen Verformung involviert. Von einer direkten Bindung zur DNA wird in [2] berichtet. IF sind dynamische Strukturen, die mit anderen Komponenten des Zytoskelettes, insbesondere mit Motor- und Plakin-Proteinen wechselwirken. Die Interaktion mit Motorproteinen gewährleistet einen gerichteten Transport von verschiedenen Stoffen und Zellorganellen. Die Plakinbindung ist nur eine von mehreren Möglichkeiten, IF über bestimmte membranintegrierte Strukturen

wie z.B. Desmosomen in der Zellmembran zu verankern und einen funktionellen Kontakt mit benachbarten Zellen herzustellen [3],[4]. Faktoren, die sich negativ auf die IF Architektur auswirken, haben oft dramatische Folgen für die Funktionalität der entsprechenden Zellen bzw. Zellgewebe. So sind eine ganze Reihe von IF Mutationen bekannt, die sich in verheerenden Erbkrankheiten manifestieren [5]. Weiterhin dürfte die Spaltung von Vimentin, ein spezielles IF mesenchymaler Zellen, durch die HIV-1 Protease ein möglicher Grund für die Zytopathogenität des AIDS-Virus sein [6].

1.2.2 Molekulare Architektur der Intermediärfilamente

Primärstruktur

Proteinsequenzvergleiche haben ergeben, dass alle IF eine hohe Sequenzhomologie untereinander aufweisen und deshalb zu Recht eine eigenständige Proteinfamilie bilden [7]. Diese IF Hauptfamilie wird zur genauen Charakterisierung in sechs verschiedene Proteinklassen weiter unterteilt (siehe 2.2.3 „Die verschiedenen Typen der Intermediärfilamente“). Eine große Gemeinsamkeit aller IF ist eine zentral gelegene, konservierte Region, die 310-355 Aminosäurereste enthält. Dieser Abschnitt beinhaltet ein sich wiederholendes Muster von Aminosäuretypen, ein so genanntes „heptad repeat“ $(a-b-c-d-e-f-g)_n$, welcher an bestimmten Stellen unterbrochen ist. Die einzelnen Buchstaben a-g geben dabei die Position der jeweiligen Aminosäurereste an, wobei a und d hauptsächlich von unpolaren Typen besetzt werden. Darüber hinaus findet man an den übrigen Positionen besonders aber bei e und g häufig basische oder saure Aminosäuren. Diese Struktur beinhaltet also ein hohes Maß an geladenen Aminosäureresten. Im Gegensatz dazu ist die Proteinsequenz des N- bzw. des C-Terminus in flankierendem Bezug zur konservierten „heptad repeat“-Struktur hoch variabel. So gibt es in diesen Bereichen innerhalb der verschiedenen IF hinsichtlich der Größe und der chemischen Eigenschaften große Unterschiede [8]. Wie bereits erwähnt, ist die „heptad repeat“-Region nicht kontinuierlich, sondern an drei Stellen unterbrochen. Die entsprechenden vier „heptad“-Segmente werden mit 1A (35 Aminosäurereste), 1B (101 Aminosäurereste in Vertebraten; 143 in einigen

bezeichnet und ist durch folgende Sequenz (Ein-Buchstabencode der Aminosäuren) gegeben: K/R-X-T/Q-M/I/L-K/Q-L-N-D-R-F-A-S-F/Y-I-D/E-K-V-R-F-L-E. Der andere im Segment 2B liegende Bereich wird als „helix termination motif“ bezeichnet und ist durch folgende Sequenz gegeben: E-Y-Q-X-L-L-D /N-V-K-X-R/A-L-D/E-X-E-I-A-T-Y-R-K/R-L-L-E-G-EE/D-X-R-L/N/I. Darüber hinaus existieren noch eine ganze Reihe weiterer konservierter Sequenzabschnitte, welche in [1] näher beschrieben sind. Zusammenfassend lässt sich sagen, dass die Primärstruktur aller IF Proteine in drei wesentliche Sequenzabschnitte gegliedert werden kann. In den variablen N-Terminus, gefolgt von der zentralen konservierten „heptad repeat“ Struktur und dem abschließenden ebenfalls variablen C-Terminus.

Sekundärstruktur

Die typische Dreiteilung der Primärstruktur bedingt im Fall der IF Proteine eine analoge, grobe Unterteilung in drei entsprechende Sekundärstrukturelemente. Im Folgenden wird der N-Terminus mit „head domain“, die zentrale „heptad repeat“ Struktur mit „rod domain“ und der C-Terminus mit „tail domain“ bezeichnet.

„head domain“:

Das erste Sekundärstrukturelement bildet die „head domain“. Wie bereits erwähnt, variiert die Proteinsequenz in diesen Bereichen sogar innerhalb einer IF Proteinklasse beträchtlich. Bis zum heutigen Tage ist es, bis auf eine Ausnahme, nicht gelungen, atomar aufgelöste Strukturmodelle dieses Proteinabschnitts zu produzieren. Deshalb gibt es hierfür keine gesicherten Strukturinformationen. Lediglich für wenige IF Typen existieren Strukturvoraussagen, die auf der Basis von Kristallstrukturen anderer ähnlicher Proteinabschnitte getroffen wurden [1].

„rod domain“:

Die „rod domain“, flankiert von der „head-“ und „tail domain“ bildet das zentrale Mittelstück der IF Sekundärstruktur. Zentral bezieht sich aber nicht streng auf die Proteinsequenzposition, sondern meint nur, dass es zwischen den beiden anderen Bereichen liegt. Kristallstrukturen kurzer Fragmentstücke dieser Region konnten bestätigen, dass die Segmente 1A, 1B, 2A und 2B, welche alle das typische „heptad repeat“ Muster aufweisen, eine α -helikale Struktur bilden. Zu den „linker“ Regionen liegen keine gesicherten Erkenntnisse vor [1, 9].

„tail domain“:

Ähnlich wie bei der „head domain“, so ist auch hier aufgrund der hohen Proteinsequenzvariabilität mit einer großen Strukturdiversität zu rechnen. Leider gibt es auch für diesen Proteinbereich fast keine atomar aufgelösten Strukturmodelle. Es ist es jedoch im Jahr 2002 gelungen, die Kristallstruktur der „tail domain“ des humanen Lamin A IF Protein aufzulösen. Diese Struktur hat für diese Arbeit aber keine Relevanz, da ihre Proteinsequenz bezüglich der betrachteten IF Proteine zu stark abweicht [1, 9].

Tertiärstruktur und Quartärstruktur

Die Tertiärstruktur und die komplexe Quartärstruktur der IF Proteine resultieren aus der geordneten, nicht kovalenten Polymerisierung vieler einzelner IF Moleküle. Dabei bildet das Dimer, bestehend aus je einem Molekül A und B, die kleinste stabile Grundeinheit. Es kommen je nach IF Typ sowohl Homo- als auch Heterodimere vor. Die Dimerisierung erfolgt durch eine parallele Anordnung von A und B. Dabei gehen also jeweils exakt die gleichen Unterdomänen der Primär- und Sekundärstruktur nicht-bindende Wechselwirkungen ein. Die Bezeichnungen der jeweiligen Sequenzabschnitte, sowie die Einteilung in „head, rod und tail domain“ werden dementsprechend beibehalten. Wie bereits erwähnt, liegen für die „head und tail domain“, bis auf eine Ausnahme und für die verschiedenen „linker“ innerhalb der „rod domain“ keine gesicherten Strukturdaten vor. Die jeweiligen α -helikalen Segmente dimerisieren zu einer linksgängigen Superhelix, einem umeinander gewundenen Doppelstrang in Form einer so genannten „coiled coil“ (siehe

Abbildung 1.2). Dabei umschlingen die Einzelstränge in einer axialen Richtung einander, wobei die Berührungszonen von den zueinander paarweise assoziierten, unpolaren, hydrophoben Aminosäureresten a und d des „heptad repeats“ besetzt sind. Dieser Streifen aus hydrophoben Wechselwirkungen wird auch als „hydrophobic core“ bezeichnet. Ganz ähnlich wie im „leucine zipper“ greifen die apolaren Aminosäureseitenketten dabei ineinander und sorgen somit für die notwendige Stabilisierung im wässrigen Milieu. Die Krümmung beider α -Helices scheint nicht die Konsequenz der „hydrophobic core“-Formation zu sein. Wahrscheinlich ist dies eine intrinsische Eigenschaft einer α -helikalen Struktur, welche viele hintereinander gereihete, apolare Seitenketten auf einer Seite der Helix enthält. Zusätzliche Stabilität wird durch ionische Wechselwirkungen vermittelt. Diese interhelikalen Salzbrücken können in den Positionen g-e', g-a', d-a' und d-e' beobachtet werden. Der Oberstrich kennzeichnet den Dimerisierungspartner [1, 9-11].

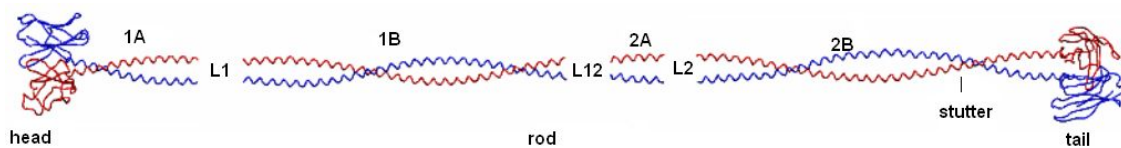


Abbildung 1.2: Quaternärstruktur IF Dimer

Rot und blau kennzeichnen die zwei verschiedenen IF Dimerisierungspartner. In den Segmenten 1A, 1B, 2A, und 2B kommt es zur Ausbildung einer „coiled coil“, welche durch die „linker“ L1, L12 und L2 voneinander getrennt sind. Die Struktur der „head und tail domain“ ist frei erfunden. Im Segment 2B ist die Diskontinuität „stutter“ lokalisiert. Die relativen Längenverhältnisse der „rod“ sowie der „linker“ sind richtig dargestellt.

Intrahelikale ionische Wechselwirkungen stabilisieren die jeweilige Helix. Sie können zwischen den Positionen e-b sowie zwischen c-g beobachtet werden. Oftmals findet man aber in der Kristallstruktur nicht alle theoretisch möglichen Salzbrücken wieder. Um das Arrangement der Aminosäurereste innerhalb einer „coiled coil“ Struktur besser zu verstehen, lohnt sich ein Blick auf die folgende Abbildung 1.3.

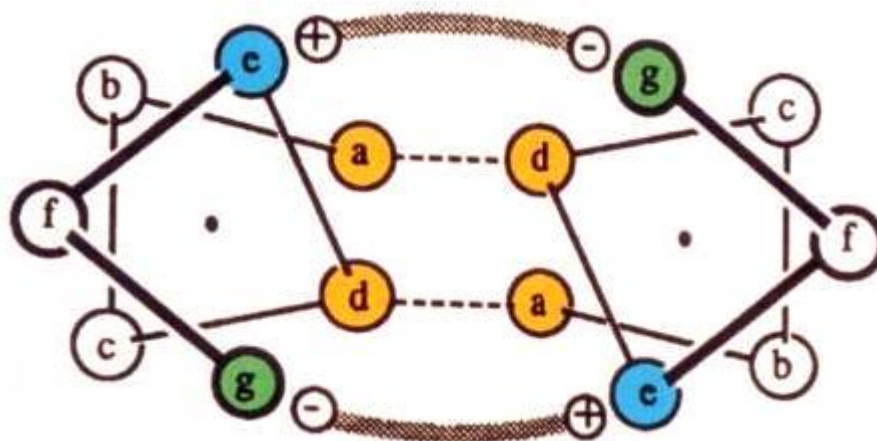


Abbildung 1.3: axiale Projektion einer doppelsträngigen α -helikalen "coiled coil" Struktur
 Dargestellt ist die Wechselwirkungsgeometrie bezogen auf die Positionen im „heptad repeat“. Jeweils ein a und ein d Aminosäurerest haben hydrophoben Kontakt (gelb). Zwischen g und e kann es zur Ausbildung einer Salzbrücke kommen (grün und blau).
 Abbildung wurde übernommen aus [1].

Die geometrischen Parameter von verschiedenen „coiled coil“ Strukturen variieren nur innerhalb kleiner Toleranzbereiche. So findet man in der Literatur folgende Werte [11]:

Radius: 4,9 Å

Aminosäurereste pro Windung: 100

Länge einer Windung: 148 Å

Als Besonderheit beinhaltet das Segment 2B die bereits erwähnte Diskontinuität genannt „stutter“. Hierbei handelt es sich um einen Einschub von vier Aminosäureresten am Ende des achten „heptads“. Um den Fortgang des „hydrophobic core“ zu gewährleisten, müssen sich die α -helikalen Bereiche an dieser Stelle reorganisieren. Dies geschieht durch eine gewisse Entwindung der α -Helix. Danach wird das „heptad repeat“ beginnend bei Position a normal weitergeführt. Kristallstrukturkoordinaten der „stutter“ Region liefert das Cys 2 Fragment [9].

Neuartige Erkenntnisse gibt es über die mögliche molekulare Architektur der Region aus „linker“ L2 und dem sich anschließenden Segment 2A. Das in diesem Sequenzabschnitt neu identifizierte Grundmuster aus elf Aminosäureresten gibt Grund zur Annahme, die bisherige Modellvorstellung

eines flexibeln „linkers“ L2 und einer „coiled coil“ Struktur in 2A zu überdenken. Demnach soll L2 rigide sein und die beiden α -helikalen Bereiche in 2A könnten geradlinig parallel nebeneinander zu liegen kommen, ohne dabei eine linksgängige Superhelix zu bilden. Ein endgültiger Beweis steht jedoch noch aus. Darüber hinaus wurde für das Segment 1A zusätzlich zur „coiled coil“ Struktur eine offene Formation postuliert [12].

Die weitere Zusammenlagerung vieler IF Dimere zum eigentlichen funktionalen IF Netzwerk ist ein überaus komplexer, noch recht wenig verstandener Prozess. Es ist jedoch durch verschiedene Experimente wie z.B. „cross link“ Studien, „cryo-electron tomography“ oder „small angle X-ray scattering“ gelungen, vier verschiedene Modi des axialen Arrangements zweier Dimere zu etablieren.

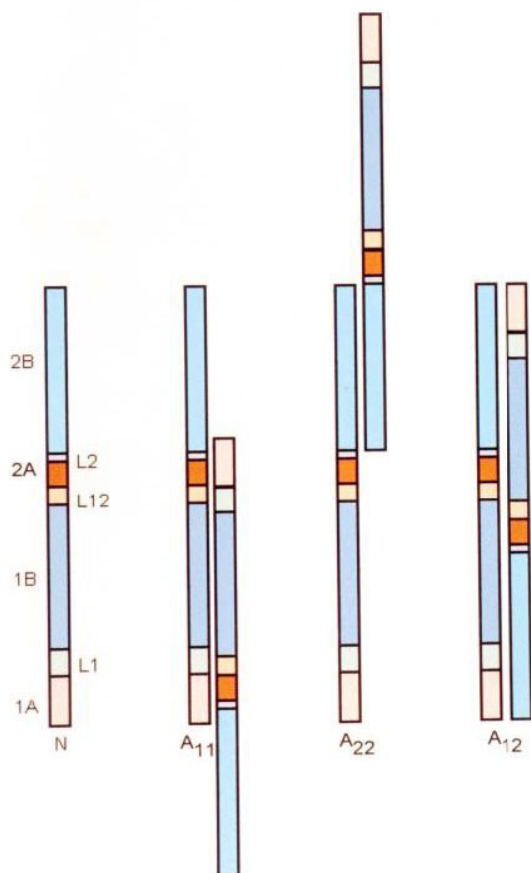


Abbildung 1.4: Zusammenlagerung von IF Dimeren
 IF Dimere (N) kommen im IF Netzwerk in vier unterschiedlichen Arrangements nebeneinander zu liegen. Es wird zwischen drei antiparallelen (A_{11} , A_{22} , A_{12}) und einem parallelen (nicht abgebildet) Modus unterschieden [1].

Der schrittweise Aufbau zum funktionalen, fertigen IF lässt sich folgendermaßen grob zusammenfassen: Ausgehend vom „coiled coil“ Dimer bilden sich

Tetramere, auch Protofilamente genannt. Man vermutet, dass Protofilamente die löslichen Grundbausteine der IF in vivo sind. Acht dieser Tetramere bilden eine Protofibrille. Dabei besteht die Möglichkeit zur Ausbildung eines Hohlzylinders mit einem Ring aus acht Tetrameren, oder einem Ring aus sieben Tetrameren und einem zentral gelegenen. Viele solcher Protofibrillen lagern sich im weiteren Verlauf in axialer als auch in radialer Richtung zusammen. Im letzten Schritt tempert diese Aggregate und verdichten sich auf ihre endgültige Größe [1, 10, 12].

Längenausdehnung der „linker“

Die statistische Auswertung von „crosslink“ Daten für die Keratine K5/14 und K1/10 (IF Typ I und II - siehe unten) ergab folgende Werte für die axiale Längenausdehnung der „linker“:

$$L1:L12:L2 = 2,52:1,63:0,84 \text{ nm [13]}$$

$$L1:L12:L2 = 2,41:2,0:0,75 \text{ nm [1]}$$

1.2.3 Die verschiedenen Typen der Intermediärfilamente

IF gehören einer Multigenfamilie an, die von etwa 70 verschiedenen Genen codiert werden. Zusätzlich gibt es noch die Möglichkeit alternativer Spleißprozesse. Sie werden aufgrund der Identität ihrer Proteinsequenz (genetische Verwandtschaft) in sechs verschiedene Proteinklassen oder Typen eingeteilt [5]. Bei dieser Einteilung wurden auch Strukturinformationen der jeweiligen Exons und Introns berücksichtigt [8]. Seit dem Jahr 2006 hat man sich auf eine einheitliche Nomenklatur geeinigt [14]. Man sollte deshalb bedenken, dass es zwischen Publikationen die vor bzw. nach 2006 erschienen sind, zu gewissen Diskrepanzen hinsichtlich der Bezeichnung kommen kann.

Typ I und II

Die 28 Typ I und 26 Typ II IF bilden zusammen die Gruppe der Keratine. Diese werden besonders in Epithelzellen expremiert. Sie machen etwa 80 Prozent der gesamten IF Genfamilie aus. Sie bilden nur Heterodimere aus. Ein Dimer besteht immer jeweils aus einem Protein vom Typ I und II [5]. Die Typ I Keratine haben saure Eigenschaften. Die Typ II Keratine haben neutrale bis basische Eigenschaften [1]. In vitro lassen sich verschiedene stabile Dimerkombinationen erzeugen. Hingegen ist die Koexpression eines bestimmten Keratinpaares in vivo ein charakteristisches Merkmal für ein bestimmtes Epithelgewebe und dessen momentanen Differenzierungsstatus', da einige Epithelzellen während ihres Differenzierungsprozesses das spezifische Expressionsmuster verändern [15, 16]. Dies macht man sich in der pathologischen Diagnostik zu nutze, um die Herkunft von Tumormetastasen bestimmen zu können. Dazu bedient man sich spezifischer Keratin Antikörper, mit deren Hilfe sich die entsprechenden Epithelien unterscheiden lassen.

Tabelle 1.1: Beispiele für eine spezifische Koexpression in Epithelzellen

| Typ II | Typ I | Vorkommen |
|--------|--------|---|
| K1;2 | K9;10 | mehrschichtig-verhornendes Epithel (Epidermis) |
| K3 | K12 | Hornhaut (Cornea) |
| K4 | K13 | mehrschichtig-unverhorntes Epithel |
| K5 | K14;15 | Basalzellen komplexer Epithelien sowie Myoepithelzellen |
| K7 | K19 | einschichtiges Epithel, luminale Drüsenzellen |

Tabelle 1.2: Beispiele für Keratinantikörper [17]

| Keratin | Antikörper |
|---------|------------|
| K1 | LL017 |
| K5 | RabK5 |
| K7 | LP1K |
| K8 | LE41 |
| K10 | LHP2 |

Typ III

Die Typ III Klasse wird von vier verschiedenen Genen codiert. Im Gegensatz zu den ersten beiden Typen können diese in vivo sowohl Homo- als auch Heterodimere bilden.

1. Desmin stabilisiert die Sarkomere in Muskelzellen.
2. Vimentin ist weit verbreitet und kommt z.B. in Lymphozyten, Endothelzellen, Fibroblasten und mesenchymalen Geweben vor. Vimentinfilamente assoziieren oft mit Mikrotubuli und erfüllen die wichtige Aufgabe, den Abstand der einzelnen Zellorganellen zueinander zu definieren und konstant zu halten.
3. GFAP (Glial Fibrillary Acidic Protein) kommt in Gliazellen vor. Wahrscheinlich bildet es in vivo nur heterodimere Formen aus.
4. Peripherin findet sich in peripheren Neuronen und ist noch weitgehend unerforscht [1, 18].

Typ IV

Typ IV IF Proteine werden häufig in Nervenzellen exprimiert und sind in das radiale Wachstum der Axone involviert. Die sieben Gene dieser Proteintypen codieren das schwere Neurofilament (NF-H), das mittlere und das leichte Neurofilament (NF-M, NF-L), sowie Nestin und α -Internexin in Nervenzellen, außerdem Synemin α , Synemin β , zwei alternative Spleißvarianten und Syncoilin in Muskelzellen [5, 19, 20].

Typ V

Diese auch als Lamine bezeichnete IF befinden sich im Gegensatz zu allen anderen nicht im Zytoplasma, sondern im Zellkern. Die sechs bekannten Proteine werden von drei verschiedenen Genen codiert. Dabei sind Lamin A, Lamin C1 und Lamin C2 ebenso alternative Spleißvarianten wie die Proteine Lamin B2 und B3. Lamin B1 wird von einem eigenen Gen codiert [5].

Typ VI

Diesem, in der Literatur nicht oft erwähnten Typ gehören die Proteine Phakinin und Filensin an. Es handelt sich dabei um hochspezialisierte Polypeptide in den zellkernlosen Zellen der Augenlinse. Ihre Proteinsequenzhomologie zeigt zum Teil erhebliche Unterschiede zu allen anderen IF [5, 21].

1.3 Das C-terminale Ende von Desmoplakin

1.3.1 Allgemeines

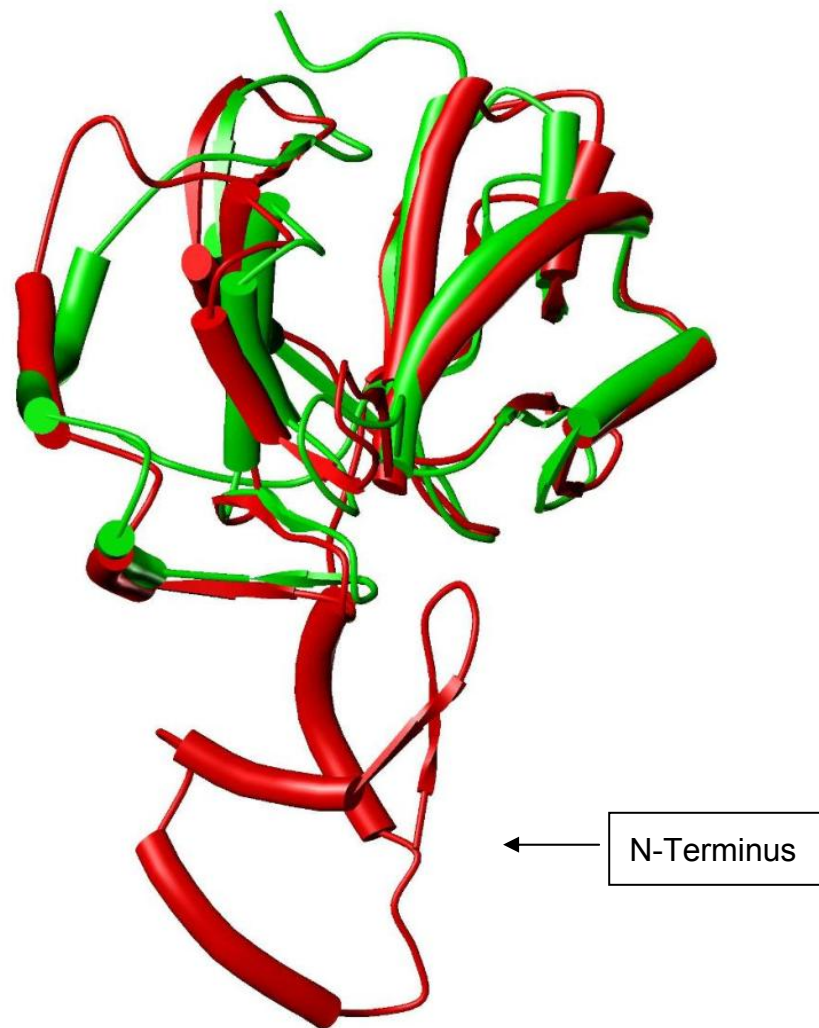
Proteine aus der Plakinfamilie sind involviert in die Organisation des Zytoskelettes. Dabei verbinden sie IF zu anderen Strukturen des Zytoskelettes oder an bestimmte Stellen in der Plasmamembran. Das Plakinprotein Desmoplakin hat u.a. die Aufgabe festbindenden Kontakt zwischen IF und Desmosomen herzustellen. Auf diese Weise entsteht eine funktionale Verbindung zur Zellmembran. Dabei vermitteln drei globuläre Domänen innerhalb des C-terminalen Endes von Desmoplakin (DPCT) die Bindung an IF. Die C-terminalen Bereiche anderer verwandter Plakinproteine enthalten ebenfalls eine oder mehrere dieser Domänen. Der N-Terminus von Desmoplakin interagiert mit desmosomalen Proteinen wie Cadherinen, Desmogleinen und Desmocollinen.

1.3.2 Struktureller Aufbau von Desmoplakin

Desmoplakin besteht aus einem 1056 Aminosäurereste langen N-Terminus, gefolgt von einem 889 Aminosäurereste langen zentralen „rod“-Bereich. Diese zentrale Region besitzt ebenfalls ein „heptad repeat“ Muster und bildet wahrscheinlich analog zu IF eine „coiled coil“ Struktur mit einem weiteren Desmoplakinmolekül aus. Das sich anschließende 926 Aminosäurereste lange DPCT beinhaltet drei globuläre homologe Domänen bezeichnet mit A, B und C. Diese liegen perlschnurartig hintereinander und sind durch Sequenzbereiche variabler Länge untereinander verbunden [22].

1.3.3 Molekulare Architektur der DPCT Domäne B und C

Im Jahr 2002 wurden die Kristallstrukturen der DPCT Domänen B (248 Aminosäurereste) und C (214 Aminosäurereste) veröffentlicht. Sie zeigen eine Sequenzidentität von 29 % und eine Sequenzähnlichkeit von 53 %. Der „rmsd“ Wert beträgt 2,0 Å für 169 C_α Positionen. Jede Domäne besteht aus viereinhalb Kopievarianten eines 38 Aminosäurereste langen Plakinmusters. B enthält einen zusätzlichen N-terminalen Bereich. Eine Plakinkopie enthält einen 11 Reste umfassenden „β-hairpin“ gefolgt von zwei antiparallelen α-Helices, mit acht bzw. 14 Aminosäureresten. Diese Struktur wird von vielen konservierten Aminosäureresten an den relativen Positionen 2, 11, 13, 16, 17, 21, 22, 26 und 30 geprägt. Alle Plakinkopien innerhalb beider Domänen weisen eine ähnlich gefaltete Struktur auf. Aufgrund von vielen apolaren Aminosäureresten auf der Oberfläche, ist diese Faltung jedoch wahrscheinlich nur im Verbund stabil. So kann man in der Kristallstruktur deutlich erkennen, wie sich diese apolaren Bereiche im Inneren gegenseitig zur Deckung bringen. Nach außen hingegen kommen vor allem die hydrophileren Seitenketten zu liegen. Die Gesamtfaltung, also das Arrangement der Plakinkopien innerhalb beider homologen Domänen, ist ebenfalls sehr ähnlich (Abbildung 1.5). Durch den zusätzlich vorhandenen 39 Aminosäurereste langen N-Terminus in B, wird diese Domäne jedoch nicht unwesentlich vergrößert. Er zeigt ähnlich zur Plakinfaltung eine „β-hairpin“ Struktur, die von zwei antiparallelen α-Helices gefolgt wird. Außerdem findet man konservierte Aminosäurereste bezogen auf die anderen Plakinvarianten. Die Faltung ist jedoch unterscheidbar und so wird dieser Bereich nicht als echtes Plakinmuster angesehen [22]. Andererseits vermag vielleicht seine exponierte Lage und die damit verbundene große, mit freiem Zugang zum wässrigen Milieu umgebende Oberfläche, diese Abweichung erklären. Überlagert man die „backbone“ Atome beider Domänen so ergibt sich folgendes Bild:



*Abbildung 1.5: Überlagerung von Domäne B und C
α-helikale Bereiche als Zylinder, „β-hairpins“ als Pfeile mit dünnen Fäden.
Die Domänen B (rot) und C (grün) weisen eine vergleichbar ähnliche Faltung auf. Im unteren Teil der Abbildung erkennt man den zusätzlichen N-terminalen Bereich in B.*

Die Übereinstimmung der Faltung ist im Plakinabschnitt deutlich zu erkennen. Nur in einzelnen „loop“-Bereichen sind Abweichungen zu erkennen. Auch der „shape“ ist abgesehen vom N-Terminus aufgrund der gleichen Faltung in etwa kongruent. Durch unterschiedliche Komposition oberflächenexponierter Aminosäurereste entsteht jedoch eine völlig andere Verteilung molekularer Wechselwirkungsfelder (Abbildung 1.6). Es konnte kein größerer zusammenhängender Bereich mit ähnlichen Eigenschaften identifiziert werden. Viele Bereiche weisen ein entgegengesetztes elektrostatisches Potential auf, wobei saure und basische Aminosäurereste teilweise vertauscht sind.

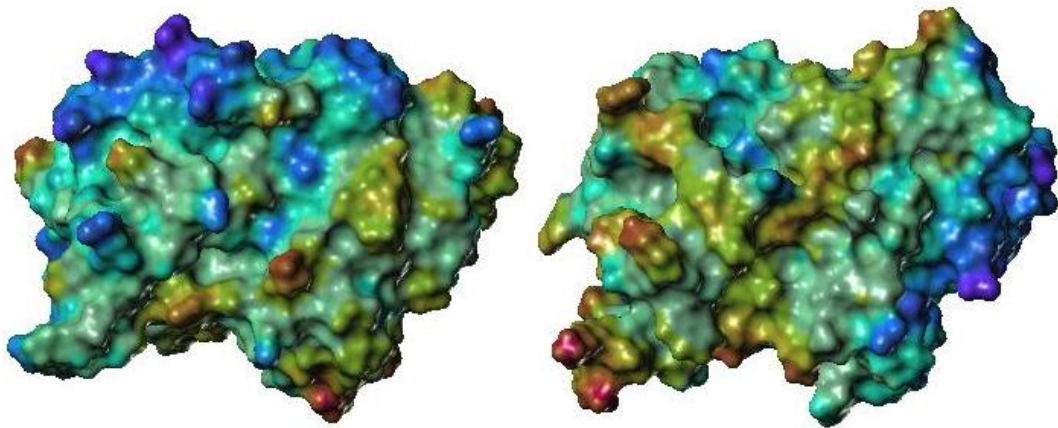


Abbildung 1.6: Elektrostatistisches Potential der Oberflächen von B und C
Aufsicht auf beide Domänen (B links, C rechts) gleicher, relativer Perspektive.
Der N-Terminus von B ist nicht zu erkennen, er befindet sich auf der Rückseite.
Der „shape“ von B und C ist in etwa kongruent. Unterschiedliches Arrangement von sauren (elektrostatistisch negativ - dunkelblaue Töne) und basischen (elektrostatistisch positiv - rotbraune Töne) Aminosäureresten auf der Oberfläche.

1.4 Bindungsstudien zur Desmoplakin - IF Wechselwirkung

1.4.1 Bindungsstudien aus der Literatur

Verschiedene Bindungsstudien haben gezeigt, dass es zwischen dem C-terminalen Ende von Desmoplakin und verschiedenen IF Proteinen zu einer bindenden Proteinwechselwirkung kommt. Die untersuchten IF Proteine waren z.B. Vimentin vom IF Typ III sowie verschiedene Keratine wie K5(IF Typ II)/14(IF Typ I) oder K8(IF Typ II)/18(IF Typ I). Zudem wurde versucht, durch grobe, systematische Eingrenzung der Proteinsequenz, die für die Bindung essentiellen Proteinbereiche zu lokalisieren. Auf Seiten der IF Proteine konnte man eine Bindung innerhalb der „rod domain“, also für den Bereich, für welchen experimentelle Strukturdaten vorliegen, nachweisen. Andererseits wurde die Frage, ob die Desmoplakin - Domänen B und C allein, zusammen oder nur mit den zusätzlichen, dazwischen geschalteten Sequenzregionen essentiell für die Bindung sind, kontrovers bis widersprüchlich diskutiert. Eine Bindung der Domäne B und der Domäne C an die Keratine K8/18 und K5/14, sowie an Vimentin (IF Typ III) ist jedoch nachgewiesen worden [22-25].

1.4.2 Bindungsstudien des Kooperationspartners BASF

Der Kooperationspartner BASF hat zum Bindungsverhalten die folgenden Aussagen getroffen. Die Domäne B ist allein in der Lage, an Keratinstrukturen des Haares und der Hautoberfläche von *Homo sapiens sapiens* zu binden. Die Bindung an Haarkeratine ist dabei etwas stärker als an Hautkeratine. Keine Bindung wurde jedoch am Haar von *Bos taurus* beobachtet. Nicht bekannt ist der genaue Typ von Keratin, welcher an der Bindung beteiligt ist. Weitere Daten zur Bindungsstärke oder andere Einzelheiten liegen nicht vor.

1.5 Hypothese der an der Bindung beteiligten Haar - Keratine

Die gesuchten Keratine müssen sich, das ergibt sich aus der Vorgabe des Kooperationspartners BASF, an der Oberfläche des entsprechenden Zielgewebes befinden. Die Oberfläche des Zielgebietes bildet die Kutikula. In dieser Schicht wurde die Expressierung von zwei Typ I Keratinen (hHa2, hHa5) und zwei Typ II Keratinen (hHb2, hHb5) nachgewiesen („h“ - human; „H“ - Hair; „a“ - acidic; „b“ - basic) [26, 27]. Nach der neuen Nomenklatur aus dem Jahre 2006 werden diese Keratine mit K32 (hHa2), K35 (hHa5), K82 (hHb2) und K85 (hHb5) bezeichnet. Im Jahr 2007 wurde der Katalog der im Haar, speziell in der Kutikula, exprimierten Keratine um zwei weitere vom Typ I (K39, K40) komplettiert [28]. Demnach enthält jede ausdifferenzierte Kutikulazelle des menschlichen Haares sechs Keratinbausteine. Es besteht daher die Möglichkeit zur Ausbildung von acht verschiedenen Keratin - „Dimeren“ (siehe Tab. 1.3).

Tabelle 1.3: „Dimerkombination“ von Keratinen in der Kutikulazelle des menschlichen Haares

| „Dimer“ -Nummer | Typ I | Typ II |
|-----------------|-------|--------|
| 1 | K32 | K82 |
| 2 | K35 | K82 |
| 3 | K39 | K82 |
| 4 | K40 | K82 |
| 5 | K32 | K85 |
| 6 | K35 | K85 |
| 7 | K39 | K85 |
| 8 | K40 | K85 |

1.6 Hypothese der an der Bindung beteiligten Haut - Keratine

Die äußere Schicht der Haut, die Epidermis, wird vor allem aus Keratinozyten aufgebaut. Die äußerste Schicht der Epidermis, das Stratum corneum, wird aus verhornten, abgestorbenen Plattenepithelzellen gebildet. Diese ausdifferenzierten Keratinozyten enthalten hauptsächlich die zwei spezifischen „Keratindimere“ aus K5(IF Typ II)/K14(IF Typ I) sowie K1(IF Typ II)/K10(IF Typ I) [16, 29, 30].

2 Ausgangslage und Zielsetzung

2.1 Ausgangslage dieser Arbeit

Grundlage dieser Arbeit waren die Kristallstrukturen der Domänen B und C (PDB Code 1LM5 und 1LM7), sowie die einiger „coiled coil“ Fragmente (pdb Code 1GK7, 1GK6, 1GK4 und 1D7M). Eine weitere Basis bildeten die (siehe Abschnitt 1.4.2) beschriebenen Bindungsstudien des Kooperationspartners. Die Keratine K8/18 und K5/14 wurden als mögliche Zielstrukturen für Haare und Haut genannt. Zu den Bindungsregionen gab es keine Hinweise.

2.2 Zielsetzung dieser Arbeit

Das Zwischenziel bestand darin, den Proteinbereich auf der Domäne B genau zu charakterisieren, der für die Bindung an Keratine verantwortlich ist. Diese Arbeitshypothese sollte dann im Anschluss durch Mutationsstudien des Kooperationspartners BASF verifiziert werden.

- Zwischenziel: Beschreibung der Bindungsregion auf der Domäne B und deren experimentelle Bestätigung durch den Kooperationspartner BASF

Darauf aufbauend sollte folgendes Hauptziel erreicht werden: Welche Mutationen in der Domäne B führen zu einer stärkeren bzw. selektiveren Bindung an die Keratinstrukturen des Haares im Vergleich zu jenen der Haut.

- Hauptziel: Mutationsvorschläge für die Domäne B, die zu einer stärkeren bzw. selektiveren Bindung an die Keratinstrukturen des Haares im Vergleich zu jenen der Haut führen

Zwei zusätzliche Zielsetzungen waren die Verbesserung der Löslichkeit der Domäne B durch Mutationen sowie die Beantwortung der Frage, ob man die zu exprimierende Sequenz verkürzen kann, ohne dabei die Bindungsaffinität herabzusetzen.

- Nebenziel a: Mutationsvorschläge für die Domäne B, die zu verbesserten Lösungseigenschaften führen
- Nebenziel b: inwiefern man die Domäne B verkleinern kann, ohne deren Bindungsaffinität herabzusetzen

3 Methoden

3.1 Sequenzalignment

Sequenzalignments sind vor allem Methoden der Bioinformatik und der molekularen Phylogenie. Sie werden verwendet, um funktionelle oder evolutionäre Verwandtschaftsverhältnisse von Nukleinsäuresequenzen oder Proteinsäuresequenzen zu untersuchen. Als Teilgebiet des „Pattern Matching“ (engl. für Musterabgleich) versucht dieses Verfahren, ein gegebenes Muster (eine bestimmte Proteinsequenz) in einem Suchbereich (die zu untersuchende Proteinsequenz) wieder zu finden. Das optimale Sequenzalignment von zwei Sequenzen kann dabei exakt berechnet werden. Sind es jedoch mehr als zwei Sequenzen (Multiples Sequenzalignment), ist man aufgrund des hohen Rechenaufwandes auf die Zuhilfenahme heuristischer Denkansätze angewiesen. Dadurch kann zwar das optimale Alignment nicht garantiert werden, man erhält jedoch trotzdem brauchbare Ergebnisse.

Für diese Arbeit hatte das Sequenzalignment große Bedeutung bei der Erstellung des Keratin-Strukturmodells (siehe Abschnitt 4.1). Außerdem hat die übersichtliche Darstellung von verschiedenen, homologen Proteinsequenzen zur besseren Orientierung beigetragen.

Verwendet wurde hierfür das Programm CLUSTAL W [31], mit der BLOSUM62 - Matrix [32].

3.2 „loop search“

Exponierte „loop“-Bereiche besitzen oft eine vergleichsweise hohe Flexibilität. Deswegen können sie in einer Kristallstruktur nicht aufgelöst sein. Im Fall der Domäne B und der Domäne C wurden solche Bereiche durch eine „loop search“-Routine ergänzt. Dabei wurde das Homology Modul des Software-Paketes INSIGHT II in Anspruch genommen [33]. Diese Methode durchsucht eine Datenbank von Kristallstrukturen nach geeigneten Fragmenten. Als Ergebnis bekommt man also Strukturvorschläge von real existierenden „backbone“-Proteingeometrien für den entsprechenden Bereich. Diese können dann automatisch eingefügt werden. Eine Bewertung erfolgt durch den

Anwender, wobei die angegebenen „rmsd“-Werte der Überlagerung Hilfestellung gewähren. Abschließend führt man typischerweise eine Geometrieoptimierung durch.

3.3 Vorhersage einer „coiled coil“-Struktur

Das Programm Pair Coil Scoring Form sowie die neue Version Paircoil2 [34] wurden verwendet, um die „coiled coil“ - Wahrscheinlichkeit einer gegebenen IF Sequenz zu berechnen. Als Ergebnis erhält man eine Sequenz mit unterlegtem „heptad repeat“ und angegebener Wahrscheinlichkeit für die verschiedenen Sequenzbereiche. Diese Informationen sind in die Konstruktion des Keratinmodells eingeflossen (siehe Abschnitt 4.1).

3.4 Konstruktion von „backbone“-Strukturkoordinaten des Keratinmodells

Zur Konstruktion von „backbone“-Strukturkoordinaten des Keratinmodells wurde das Software Paket Sybyl verwendet [35]. Die darin enthaltenen Werkzeuge und Funktionen wurden benutzt, um verschiedene „coiled coil“-Kristallstrukturfragmente in geeigneter Weise abzuschneiden, zu verlängern, zu kombinieren und im Raum zu platzieren. Wichtig war dabei die Erstellung von geometrischen Hilfslinien und Hilfeebenen, an welchen das Modell sukzessive aufgebaut wurde (siehe Abschnitt 4.1).

3.5 Anfügen von Aminosäureseitenketten

Das Programm SCWRL („sidechain placement with a rotamer library“) wurde dazu verwendet, die entsprechenden Aminosäureseitenketten auf die „backbone“-Strukturkoordinaten des Keratinmodells zu übertragen [36]. Auf der Basis einer Rotamer-Bibliothek werden hierbei nur energetisch begünstigte Konformationen berücksichtigt. Somit ist ein wahrscheinliches Arrangement der Seitenkettenkonformation gewährleistet (siehe Abschnitt 4.1).

3.6 Kraftfeldmethoden

Bei Kraftfeldmethoden greift man auf eine klassische Vorstellungsweise zurück, gemäß welcher die Atome als Massepunkte in Molekülen durch kleine Federn mit bestimmter Federkonstante untereinander verbunden sind. Die Änderung von Bindungs- und Torsionswinkeln wird berücksichtigt. Für eine Vielzahl von möglichen Atomkombinationen müssen hierzu die entsprechenden Federkonstanten, sowie die Partialladungen der Atome und deren gegenseitige Wechselwirkung ermittelt werden. Durch diese starke Vereinfachung wird es möglich, vergleichsweise große Systeme wie Proteine in einer wässrigen Umgebung zu betrachten. Verschiedene Kraftfelder, mit teilweise spezieller Ausrichtung auf bestimmte Molekülarten, stehen dem Anwender zur Verfügung. In dieser Arbeit wurden das in Sybyl integrierte Tripos - Kraftfeld und das GROMOS96 - Kraftfeld unverändert verwendet [37]. Hauptanwendungsgebiete sind Geometrieoptimierungen und Moleküldynamiksimulationen.

3.6.1 Geometrieoptimierung durch Energieminimierung

Energetisch ungünstige Molekülkonformationen können mit Hilfe von Kraftfeldmethoden schnell und zuverlässig verbessert werden. So treten beim manuellen Erstellen von Strukturkoordinaten oder bei einer Momentaufnahme einer Moleküldynamiksimulation solche ungünstigen Konformationen auf. Man muss jedoch berücksichtigen, dass hierbei nur lokale Minima gefunden werden. Das Auffinden des globalen Minimums aller Konformationen kann somit nicht garantiert werden. Folgende zwei Methoden wurden mit den Programmen Sybyl und GROMACS angewendet [38]:

„Steepest Descent“ - Methode des steilsten Abstiegs

Ausgehend von einer festgelegten Schrittweite werden die Koordinaten der Atome in Richtung der negativen Steigung (in Richtung lokales Minimum) der Energiefunktion verschoben. Dieser Vorgang wird wiederholt, solange sich die potentielle Energie der betrachteten Konformation verringert. Die Schrittweite wird dabei stetig erhöht. Steigt die Energie wieder an, so wird dieser Schritt

rückgängig gemacht und die Schrittweite verringert. In der Nähe eines Minimums konvergiert diese Methode oft nur sehr langsam. Es kommt dabei zu einer Oszillation, die dann durch ein Abbruchkriterium beendet wird. Praktisch gesehen kommt man mit dieser Methode schnell in die Nähe eines lokalen Minimums, so dass sie sich als erste Anwendung für eine Geometrieoptimierung eignet. Eine weitere Optimierung lässt sich im Anschluss mit der „Conjugate Gradient“ Methode erzielen.

„Conjugate Gradient“

Im Unterschied zu „Steepest Decent“ wird hier zur Festlegung der Schrittweite die Steigung der zuvor berechneten Tangente an die Energiefunktion in Betracht gezogen. Dadurch konvergiert diese Methode schnell gegen das Minimum.

3.6.2 Moleküldynamik-Simulation

Die Moleküldynamik-Simulation ist eine wichtige, theoretische Methode zur Erforschung der Struktur, Dynamik und Thermodynamik biologischer Moleküle wie Proteine und Nucleinsäuren. Dabei wird das zeitliche Verhalten des zu betrachtenden Systems berechnet. Auf diese Weise erhält man detaillierte Einblicke in die Wechselwirkungen auf molekularer Ebene. So lassen sich beispielsweise Konformationsänderungen von Aminosäureseitenketten, Bewegungen von „Loopregionen“ und Unterdomänen, Faltungswege von ganzen Proteinen sowie Assoziations- und Dissoziationsvorgänge in hoher zeitlicher Auflösung reproduzierbar studieren. Eine weitere wichtige Anwendung ist auch die Betrachtung von kleinen Liganden in ihrer Proteinbindetasche. Theoretisch gesehen, müsste man zur Durchführung einer Moleküldynamik-Simulation, die simultane Integration der Newton'schen Bewegungsgleichungen aller Atome des Systems ermitteln. Dieses Gleichungssystem ist aber für mehr als zwei Komponenten mathematisch nicht mehr exakt lösbar. Ein numerisches Näherungsverfahren, das häufig dafür in der Praxis verwendet wird, ist die Leapfrog-Methode. Dabei werden die Geschwindigkeiten v zum Zeitpunkt $v(t+1/2\Delta t)$ und die Ortskoordinaten r zum Zeitpunkt $r(t+\Delta t)$ als Funktion der Zeit

unter Verwendung des Kraftfeldes aus den vorherigen Geschwindigkeiten $v(t-1/2\Delta t)$ und den Ortskoordinaten $r(t)$ ermittelt. Zu Beginn einer Moleküldynamik-Simulation müssen daher die Startgeometrie, die Starttemperatur und die Größe des Zeitintervalls festgelegt werden. Als Startgeometrie (Startkoordinaten) verwendet man experimentelle Strukturkoordinaten wie Kristall- oder NMR-Strukturen oder aus theoretischen Berechnungen resultierende, wie sie z.B. aus dem Bau eines Homologiemodells oder einer Geometrieoptimierung erhalten werden können. Eine Kombination aus beiden ist in der Praxis häufig anzutreffen. Aus der festgelegten Temperatur werden die Startgeschwindigkeiten für den Leapfrog-Algorithmus berechnet. Dabei verwendet man die Randbedingung, dass auch die Geschwindigkeiten der einzelnen Atome in Molekülen nach der kinetischen Gastheorie durch die Maxwell-Boltzmann-Gleichung beschrieben werden können. Per Zufallsgenerator werden Geschwindigkeiten erzeugt und anschließend an diese Gauß-Verteilung für die vorgegebene Temperatur angepasst. Aus der Ableitung der potentiellen Energie des Kraftfeldes, nach der jeweiligen Ortskoordinate der Atome, erhält man die auf sie einwirkenden Kräfte bzw. Beschleunigungen. Diese werden dann zur Berechnung der folgenden Geschwindigkeiten und Ortskoordinaten herangezogen. Durch diese zeitdiskrete Integration (Diskretisierung der Geschwindigkeit) macht man zwangsläufig Fehler. Somit ist es wichtig, die berechneten Zeitschritte kleiner zu wählen, als die schnellsten Bewegungen im Molekül ablaufen können (Streckschwingungen, 10^{-14} s). Man verwendet üblicherweise als Schrittweite eine Femtosekunde (10^{-15} s). Die Verwendung des 'Settle-Algorithmus' für das SPC Wassermolekülmodell und des 'Lincs-Algorithmus' für den Rest erlaubt jedoch eine Schrittweite von zwei Femtosekunden. Bei diesen Methoden werden die schnellen Streckschwingungen ausgekoppelt und der Abstand betreffender Atome auf die Gleichgewichtslage „eingefroren“. Der Lincs-Algorithmus wurde auch auf die Bindungswinkel von essentiellen Wasserstoffatomen angewendet.

Aufgrund numerischer Fehler, Unzulänglichkeiten des Kraftfeldes, etc., kann es im Laufe einer MD-Simulation zu einem unphysikalischen Aufheizen oder Abkühlen des Systems kommen. Um dem entgegenzuwirken ist es notwendig,

die Temperatur des Systems zu kontrollieren. Die Simulation wurde als NPT-Ensemble durchgeführt (konstante Anzahl an Teilchen N , konstanter Druck und konstante Temperatur). Sowohl die Temperatur als auch der Druck wurden durch Skalierung der Geschwindigkeiten bzw. der Simulationsbox (inklusive aller Ortskoordinaten) unter Verwendung des Berendsen-Algorithmus' konstant gehalten. Die Randbedingungen der Simulationsbox wurden mit PBC (periodic boundary conditions) durchgeführt. Eine genaue Beschreibung der erwähnten Methoden und weiterführende Informationen finden sich unter [39].

3.7 Analyse der Qualität eines Proteinmodells

Die aussagekräftigste Methode zur Beurteilung der Qualität eines Proteinmodells, wäre die exakte Vorhersage oder die Nachvollziehbarkeit der Realität in Form von Experimenten. Das ist jedoch aus verschiedenen Gründen nicht immer möglich. Allein die Frage, wie man bestimmte auf experimentellem Wege erhaltene Daten mit dem künstlichen Modell korreliert, birgt oft Schwierigkeiten in sich. Denn die Erklärung eines experimentellen Befundes kann häufig multiple Ursachen haben. In vielen Fällen stehen anfänglich nur wenige experimentelle Ergebnisse zur Verfügung, vielmehr bilden diese dann das eigentliche Arbeitsziel der theoretischen Betrachtungsweise. Theoretische Beurteilungssysteme, die sich auf experimentelle Daten beziehen, stellen in diesem Zusammenhang eine gute Alternative oder Ergänzung dar.

3.7.1 Procheck - stereochemische Qualität einer Proteinstruktur

Das Programm PROCHECK beurteilt unter anderem die Wahrscheinlichkeit der Geometrie von einzelnen Aminosäureresten in einem zu untersuchenden Proteinmodell [40]. Dabei bezieht es sich auf eine Referenzverteilung stereochemischer Parameter wie Torsionswinkel, Chiralität und Bindungslängen, die aus einer großen Anzahl hochaufgelöster Kristallstrukturen erhalten wurden [41]. Speziell der Ramachandran Plot stellt die Phi- und Psi-Torsionswinkelverteilung graphisch dar und ist somit ein wichtiges und übersichtliches Gütekriterium für die Qualität eines Proteinmodells. Es muss jedoch darauf hingewiesen werden, dass Proteinstrukturen durchaus

unwahrscheinliche Konformationen wie beispielsweise cis-Aminosäuren enthalten können. Diese werden relativ häufig in Tyrosin- oder Serin-Resten beobachtet, wenn sie direkt vor einem Polin-Rest stehen. Die Kristallstruktur der humanen Arylsulfatase A (PDB-Code: 1AUK) beinhaltet beispielsweise acht cis-Aminosäuren. Procheck wurde zur stereochemischen Beurteilung sowohl der eingefügten „loop“-Bereiche in die Kristallstruktur der Domäne B als auch des Keratinmodells verwendet.

3.8 Dockingmethoden für Protein-Protein-Interaktionen

Seien die räumlichen Strukturen der Proteine A und B gegeben, von denen bekannt ist, dass sie einen Komplex AB bilden, dann versteht man unter dem Protein-Protein-Docking die korrekte Vorhersage der Wechselwirkungsgeometrie AB. Es gibt verschiedene Ansätze für das Protein-Protein-Docking. Der einfachste Ansatz betrachtet Proteine als starre Gebilde. Hier spricht man vom „Rigid Body Docking“. Dieser Ansatz geht auf das bekannte Schlüssel-Schloss-Prinzip zurück, das Emil Fischer 1894 vorschlug. Es besagt, dass die zu interagierenden Proteine, in seinem Fall Enzym und passendes Substrat, starre Körper sind, die geometrisch komplementäre Oberflächenregionen besitzen müssen. Es gibt Proteine, die während der Komplexbildung Teilbereiche ihres „backbones“ verändern bzw. anpassen. Andere wiederum bleiben relativ inert gegenüber konformatorischen Änderungen im Proteinerückgrat. Die Seitenketten bewegen sich aber in jedem Fall. Sie adaptieren dabei in die für den Komplex günstigste Konformation. Deshalb erscheinen Methoden als sinnvoll, die versuchen, verschiedene Seitenkettenkonformationen in den Docking-Prozess mit einzubeziehen. Eine systematische Durchmusterung des gesamten Konformationsbereichs ist jedoch aufgrund der enorm großen Anzahl an Freiheitsgraden und der damit verbundenen langen Rechenzeit nur für kleine Peptide möglich. Für große Proteine bleibt bisher nur die Möglichkeit, ein kleines Ensemble an verschiedenen, zufällig ausgewählten Startgeometrien für den bevorstehenden Dockingprozess zu verwenden. Innovative Ansätze zielen zusätzlich darauf ab, biochemische sowie biophysikalische Daten im Dockingprozess zu

berücksichtigen. Je nach Datenlage kann auf diese Weise die Vielfalt an möglichen Komplexgeometrien auf ein Minimum reduziert werden. Durch diese Konzentration auf wenige Wechselwirkungsmöglichkeiten sollte die Wahrscheinlichkeit für eine korrekte Vorhersage stark zunehmen. Folgende Methoden bzw. Programme wurden in dieser Arbeit verwendet.

3.8.1 Hex 4.1 -Software für „Rigid Body Docking“

Hex steht für ein interaktives, molekulares Graphikprogramm zur Berechnung und Visualisierung von möglichen Dockingvarianten von Paaren von Protein- und oder DNA-Molekülen. Es basiert auf der Kombination von Energie- und Oberflächenkomplementarität.

Zu Beginn einer jeden Docking-Prozedur werden die Strukturkoordinaten der entsprechenden Moleküle (Proteine) auf eine dreidimensionale Funktion reduziert. Diese aus verschiedenen Parametern zusammengesetzte „Shape“-Funktion reduziert die Daten der Strukturkoordinaten auf die Verkörperung der räumlichen Gestalt der Oberflächen beider Proteine. Ein Protein wird als Rezeptor, das andere als Ligand bezeichnet. Jedem Molekül wird außerdem ein Schwerpunkt zugewiesen. Die Verbindung zwischen diesen beiden Massezentren ist als intermolekulare Achse definiert. Beide Proteine werden vom Benutzer in geeigneter Weise auf einen bestimmten Abstand entlang der intermolekularen Achse separiert. Während des Dockingprozesses rotieren der Rezeptor und der Ligand in einer vom Benutzer festgelegten Schrittweite um ihre Schwerpunkte. Der Ligand wird nach jedem Rotationsschritt vom Ausgangsabstand wiederum entlang der intermolekularen Achse an den Rezeptor geführt und auch wieder entfernt, wobei die intermolekulare Wechselwirkungsenergie im energetisch günstigsten Abstand abgespeichert wird. Die Berechnung von Wechselwirkungsenergien basiert auf der speziellen „Shape“-Funktion sowie auf Kraftfeldmethoden (OPLS Kraftfeld). Die Bewertung der Proteinkomplexe erfolgt primär durch das (Hex)-„Shape“ Potential. Optional kann das Elektrostatische Potential (entspricht dem Coulomb-Potential) zusätzlich berücksichtigt werden. Die Schrittweite der Annäherung ist wählbar. Zusätzlich können weitere Filter und Parameter verändert werden, um einen

groben, aber sehr schnellen Vorausscheid von sterisch ungünstigen Konstellationen zu bewirken. Anschließend kann jede Proteinkomplexkonstellation einer Energieminimierung unterworfen werden. Abschließend werden alle Dockingresultate in Cluster zugeordnet bzw. eingeteilt und graphisch visualisiert. Das Programm bietet auch die Möglichkeit, ein Ensemble von verschiedenen Startkoordinaten der jeweiligen Proteine zu untersuchen [42].

Hex wurde zum Docken von Proteinen verwendet. Sehr praktisch hat sich dabei die Programmfunktion erwiesen, den Suchraum an Proteinkomplexkonstellationen beliebig einzuschränken, gleichzeitig aber die Genauigkeit zu steigern. So kann man u.a. die Wechselwirkungsenergie von Proteinkomplexen aus der PDB einfach und schnell berechnen. Die Energien sind natürlich nur relativ zu betrachten und stellen keine realen Wechselwirkungsenergien dar.

3.8.2 HADDOCK 1.3

Das Programm HADDOCK steht für „**H**igh **A**mbiguity **D**riven biomolecular **DOCK**ing“. Es basiert ebenfalls auf der Kombination von Energie- und Oberflächenkomplementarität und betrachtet Proteine und andere Biomoleküle variabler Größe. Wie der Name vielleicht erahnen lässt, versucht dieses Programm jedoch zusätzlich den Dockingprozess zu steuern bzw. auf gewisse Zielgebiete zu konzentrieren. Dazu bedient es sich biophysikalischer oder biochemischer Daten, welche beispielsweise die Bindungsregionen näher eingrenzen können. Diese Informationen werden als so genannte „Ambiguous Interaction Restraints“ in den Dockingprozess involviert. Zusätzlich können auch direkt orientierende NMR-Daten wie beispielsweise „residual dipolar couplings“ verwendet werden. HADDOCK besteht aus einer Sammlung von Python-Skripten, erhalten von ARIA und macht Gebrauch von CNS zur Strukturberechnung. Die Kraftfeldrechnungen zur Bestimmung von inter- und intramolekularen Wechselwirkungsenergien basieren auf dem OPLS-Kraftfeld mit einem modifizierten Parametersatz. Der Dockingprozess selbst umfasst drei aufeinander folgende sowie aufbauende Phasen, wobei entsprechende

Parameter vom Benutzer verändert werden können [43]. (In dieser Arbeit wurden die Standardeinstellungen beibehalten.)

Phase 1: Zufällige Strukturgeneration und „rigid body“-Energieminimierung

Phase 2: „Semi-flexible simulated annealing“

Phase 3: „Refinement“ mittels Moleküldynamik-Simulation

Ambiguous Interaction Restraint

Ein „Ambiguous Interaction Restraint“ ist definiert als eine mehrdeutige, intermolekulare van der Waals ähnliche Energieterm-Distanz (mit einem maximalen Wert von „3 Å“) zwischen den Atomen von „aktiven Resten“ des einen und den Atomen von „aktiven Resten“ sowie „passiven Resten“ des anderen Komplexpartners. Dieses Verfahren wird für beide Komplexpartner durchgeführt

„Aktive Reste“ sind im Fall von Proteinen die Aminosäurereste, von denen man annimmt sie seien an der Bindung beteiligt. „Passive Reste“ sind dann die unmittelbaren Nachbarn der „aktiven Reste“. Für beide gilt die Bedingung, dass sie für Wassermoleküle frei zugänglich sein müssen.

Ist der maximale Abstand also unterschritten, so wird dieser Energieterm beim Docking berücksichtigt und die steuernde Funktion ist somit aktiviert.

Phase 1: Zufällige Strukturgeneration und „rigid body“-Energieminimierung

Die Proteine werden auf einen bestimmten Abstand separiert und zufällig um ihre Massenpunkte rotiert. Jede Konstellation wird dann im Anschluss einer „rigid body“-Energieminimierung unterzogen, wobei Rotations- und Translationsbewegungen erlaubt sind. „Rigid body“ bedeutet, dass die Bindungswinkel und Bindungslängen in beiden Molekülen als starr und unflexibel behandelt werden. Die besten Strukturen bezüglich ihrer

intermolekulare Wechselwirkungsenergie werden in der anschließenden Phase weiterverwendet.

Phase 2: „semi-flexible simulated annealing“

„Simulated annealing“ steht für die simulierte Abkühlung eines molekularen Systems bis in die Nähe des absoluten Nullpunktes. Durch die Anwendung dieses Verfahrens wird die Wahrscheinlichkeit erhöht, dass sich die molekularen Elemente in energetisch günstigen Konformationen arrangieren. „Semi flexible“ bedeutet hierbei, dass sich im Fall von Proteinen nur die Seitenketten frei bewegen können. Die „backbone“-Atome werden dagegen durch Rückhaltekräfte auf ihren Positionen festgehalten. HADDOCK verwendet ein auf Moleküldynamik-Simulation beruhendes Verfahren. Ausgehend von einer ausreichend hohen Temperatur, bei welcher die Seitenketten den möglichen Konformationsbereich über bestehende Energiebarrieren hinweg beschreiten können, wird das System sukzessive abgekühlt. Als Resultat sollte sich die Optimierung der Seitenkettenkonformation in einer erhöhten Wechselwirkungsenergie widerspiegeln. Es besteht zusätzlich die Möglichkeit, bestimmte Proteinbereiche in dieser Phase völlig flexibel zu betrachten. Bedingt durch die hohe Temperatur, muss man jedoch mit stärkeren, strukturellen Veränderungen rechnen. Dies kann man z.B. beim Docking von Liganden oder Peptiden sinnvoll einsetzen, um deren Konformationsraum besser auszuloten.

Die zweite Phase besteht aus drei hintereinander geschalteten „simulated annealing“ Protokollen.

Protokoll 1: 1000 Zeitschritte; von 5000 K auf 50 K; 8 fs Schrittweite; jeder Komplexpartner wird als „rigid body“ behandelt

Protokoll 2: 4000 Zeitschritte; von 2000 K auf 50 K; 4 fs Schrittweite; die definierten Seitenketten der Bindungsregionen werden als „semi-flexible“ betrachtet

Protokoll 3: 1000 Zeitschritte; von 500 K auf 50 K; 2 fs Schrittweite; die definierten Seitenketten und ggf. die „backbone“-Bereiche sind freibeweglich; Energieminimierung mit „steepest descent“

Alle so erhaltenen Proteinkomplexe werden in der nächsten Phase weiter verfeinert.

Phase 3: “Refinement” mittels Moleküldynamik-Simulation

In dieser Phase wird eine Moleküldynamik-Simulation in Gegenwart von Wassermolekülen durchgeführt. Dadurch soll die energetische Bewertung der Proteinkomplexe verbessert werden.

Protokoll der Dynamik: Schrittweite 2 fs; aufheizen auf 500 K in 3 Stufen (500 Zeitschritte bei 100 K; 200 K; 300 K; Rückhaltekräfte auf alle Atome außer den definierten Seitenketten der Bindungsregion); 5000 Zeitschritte bei 300 K (Rückhaltekräfte auf alle „backbone“-Atome außer denen der Bindungsregion); Abkühlung in jeweils 1000 Zeitschritten bei 300 K ,200 K , 100 K (Rückhaltekräfte auf alle „backbone“-Atome außer denen der Bindungsregion)

Abschließend erfolgt eine Sortierung/Bewertung der Proteinkomplexe entsprechend der Kraftfeldenergien aus elektrostatischen und van der Waals Energietermen. Diese Energien sind relativ zu betrachten und stellen keine realen Wechselwirkungsenergien dar. Eine weitere Bewertung erfolgt aufgrund der bedeckten bzw. begrabenen Oberfläche der einzelnen Komplexpartner bedingt durch die jeweilige Komplexgeometrie. Weitere automatische Strukturvalidierungen können beispielsweise durch Verknüpfung mit dem Programm Procheck eingestellt werden.

Um das Programm benutzen zu können, müssen „aktive Reste“ definiert werden. Die „aktiven Reste“ wurden aus der Arbeitshypothese abgeleitet. HADDOCK wurde verwendet, um Vorschläge für Komplexgeometrien zu erhalten.

4 Durchführung und Ergebnisse

4.1 Protein-Homologie-Modelling

Das Homologie-Modelling ist ein theoretisches Verfahren, dreidimensionale Strukturkoordinaten von Proteinen zu generieren, deren atomar aufgelöste Strukturen nicht vorhanden sind oder nicht zur Verfügung stehen. Man macht sich dabei die Tatsache zu nutze, dass die räumliche Struktur (die Faltung der Proteine) höher konserviert sein kann als die Aminosäuresequenz. Deshalb erscheint es angemessen, vorhandene Strukturdaten auf homologe Proteine/Proteinbereiche zu übertragen.

4.1.1 Vervollständigung der Kristallstruktur der Domäne B

Die Kristallstruktur der Domäne B ist ein Homodimer. Die einzelnen Monomere werden mit A bzw. B bezeichnet. Sowohl bei A als auch bei B konnten die Koordinaten eines „loop“-Bereichs nicht angegeben werden. Es handelt sich dabei um den gleichen „loop“ in beiden Kristallstruktur-Monomeren. Die Größe der fehlenden Bereiche ist verschieden. Entsprechende Proteinbereiche unmittelbar vor und nach diesem „loop“ haben zudem etwas verschiedene, relative Koordinaten. Deshalb wurde die „loop search“-Routine (siehe Abschnitt 3.2) für jedes Monomer separat durchgeführt und ein entsprechender „loop“ eingefügt.

Dem Monomer A fehlen die drei Aminosäurereste E2239, E2240 und I2241. Eingefügt wurde ein Strukturbereich aus der Kristallstruktur mit dem PDB-Code 1ABZ. In die Suche wurden jeweils fünf Aminosäurereste vor und nach der Fehlstelle mit berücksichtigt. Die Sequenz der eingefügten „backbone“-Koordinaten lautet ALEARGTDSNAEL. GTD entspricht dabei den fehlenden Aminosäureresten (E2239, E2240 und I2241). Der „rmsd“-Wert beträgt 0,42 Å.

Dem Monomer B fehlen die zwei Aminosäurereste E2240 und I2241. Eingefügt wurde ein Strukturbereich aus der Kristallstruktur mit dem PDB-Code 3CEQ. In die Suche wurden jeweils fünf Aminosäurereste vor und nach der Fehlstelle mit berücksichtigt. Die Sequenz der eingefügten „backbone“-Koordinaten lautet

FINNKLISDAEL. LI entspricht dabei den fehlenden Aminosäureresten (E2240 und I2241). Der „rmsd“-Wert beträgt 0,8363 Å.

Die stereochemische Beurteilung der eingefügten Bereiche wurde mit dem Programm Procheck kontrolliert und sichergestellt.

4.1.2 Vervollständigung der Kristallstruktur der Domäne C

Die Kristallstruktur der Domäne C ist ein Homodimer. Die einzelnen Monomere werden mit A bzw. B bezeichnet. Sowohl bei A als auch bei B konnten die Koordinaten eines „loop“-Bereichs nicht angegeben werden. Es handelt sich dabei um den gleichen „loop“ in beiden Kristallstruktur-Monomeren. Die Größe der fehlenden Bereiche ist verschieden. Entsprechende Proteinbereiche unmittelbar vor und nach diesem „loop“ haben zudem etwas verschiedene, relative Koordinaten. Deshalb wurde die „loop search“-Routine (siehe Abschnitt 3.2) für jedes Monomer separat durchgeführt und ein entsprechender „loop“ eingefügt.

Dem Monomer A fehlen die sieben Aminosäurereste E2699, G2700, V2701, K2702, G2703, K2704 und K2705. Eingefügt wurde ein Strukturbereich aus der Kristallstruktur mit dem PDB-Code 1FAG. In die Suche wurden jeweils fünf Aminosäurereste vor und nach der Fehlstelle mit berücksichtigt. Die Sequenz der eingefügten „backbone“-Koordinaten lautet FLKLAFCDICGKFLNG. FCDICGK entspricht dabei den fehlenden Aminosäureresten. Der „rmsd“-Wert beträgt 0,6102 Å.

Dem Monomer B fehlen die drei Aminosäurereste K2702, G2703 und K2704. Eingefügt wurde ein Strukturbereich aus der Kristallstruktur mit dem PDB-Code 1GPE. In die Suche wurden jeweils fünf Aminosäurereste vor und nach der Fehlstelle mit berücksichtigt. Die Sequenz der eingefügten „backbone“-Koordinaten lautet LTVPLINNRTNNI. INN entspricht dabei den fehlenden Aminosäureresten. Der „rmsd“-Wert beträgt 1,66 Å.

Die stereochemische Beurteilung der eingefügten Bereiche wurde mit dem Programm Procheck kontrolliert und sichergestellt.

4.1.3 Konstruktion der Strukturmodelle von Keratinen der Haut und des Haares

Vorüberlegungen

Im Fall der Keratine liegen keine homologen Strukturdaten für die globulären Bereiche aus „head“ und „tail“ vor. Es ist deshalb nicht möglich, ein vollständiges Homologiemodell zu erstellen. Für die Domäne B wurde eine Bindung an Keratine innerhalb der „rod“ nachgewiesen (siehe Abschnitt 1.4.1 Bindungsstudien aus der Literatur). Diese wichtige Information gestattet es im Sinne der Zielsetzung, sowohl die „head“- als auch die „tail“-Domäne bei der Konstruktion des Modells nicht zu berücksichtigen und nur die „rod“ zu modellieren.

Für die „rod“ gibt es keine vollständigen Strukturdaten. Die „linker“ konnten deshalb ebenfalls nicht in das Modell einbezogen werden. Der Großteil der „rod“ besteht aus den vier „coiled coil“-Segmenten 1A, 1B, 2A und 2B. Diese Segmente wurden aus publizierten Kristallstrukturen sukzessive aufgebaut.

Konstruktion der „coiled coil - rod“-Segmente 1A, 1B, 2A und 2B von Keratin K8/18

Eine „coiled coil“-Struktur ist das Ergebnis einer entsprechenden Dimerisierung zweier α -Helices, welche beide ein „heptad repeat“-Muster aufweisen (siehe Abschnitt 1.2.2 Molekulare Architektur der Intermediärfilamente). Es handelt sich dabei um ein axial homogen fortlaufendes Gebilde. Jede „coiled coil“ kann man deswegen als Summe von hintereinander gereihten „heptad repeat“-Einheiten betrachten. Mit anderen Worten, aus einem „heptad repeat“-Dimer kann man durch dessen Vervielfältigung eine beliebig lange „coiled coil“-Struktur erzeugen. Es stellt sich dabei die Herausforderung, die einzelnen identischen Dimerkopien in geeigneter Art und Weise auszurichten und miteinander zu verbinden. Diese Fehlerquelle kann durch Verwendung möglichst langer „coiled coil“-Fragmente minimiert werden. Des Weiteren muss die Sequenz richtig auf das „backbone“ übertragen werden.

Cortexillin I PDB:1D7M Position 243-343

12345678901234567890123456789012345678901234567890123456
 EMANRLAGLENSLESEKVSREQLIKQKQDQLNSLLASLESEGAEREKRLRELEAKLD
 cdefgabcdefgabcdefgabcdefgabcdefgabcdefgabcdefgabcdefgabc

789012345678901234567890123456789012345678901
 ETLKNLELEKRLARMELEARLAKTEKDRAILELKLAEAIDEKSKLE
 cdefgabcdefgabcdefgabcdefgabcdefgabcdefgabcdefgabcde

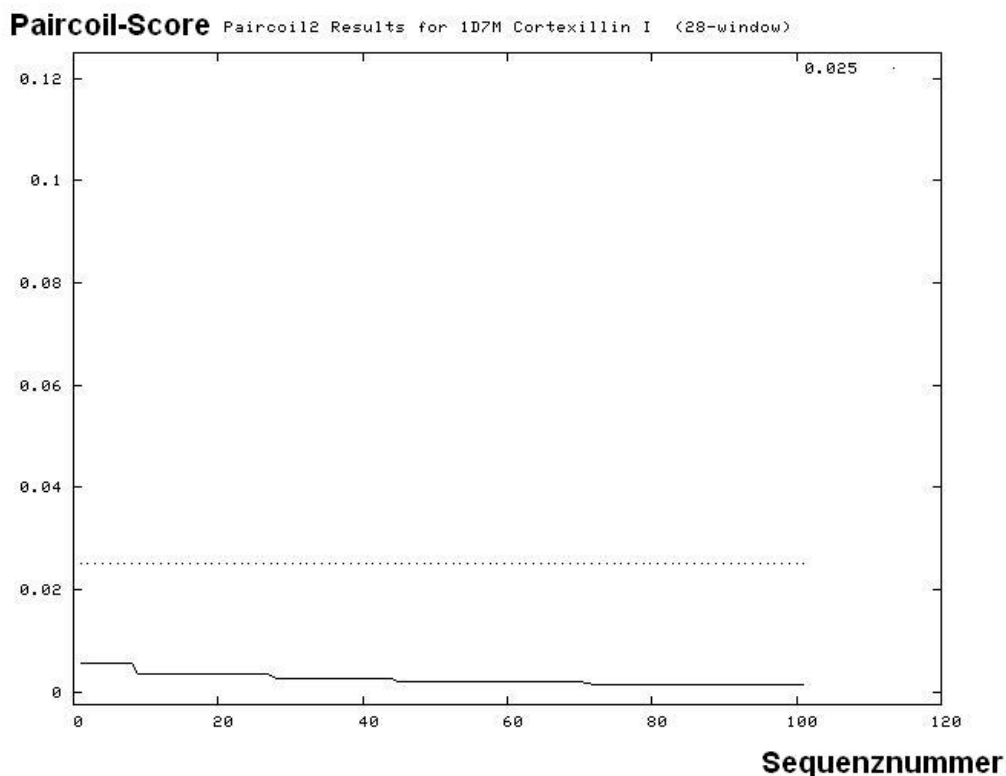


Abbildung 4.1: Cortexillin I „coiled coil“-Wahrscheinlichkeit

Die Sequenz der Kristallstruktur wurde mit dem Programm Paircoil2 ausgewertet.
 Oben: Unterhalb der Cortexillin I-Sequenz ist das zugehörige „heptad repeat“
 wiedergegeben. Unten: Der Paircoil2 Graph indiziert hohe „coiled coil“-
 Wahrscheinlichkeit für die komplette Sequenz unterhalb der gepunkteten Linie.

Die Kristallstruktur des „coiled coil“-Abschnittes von Cortexillin I (PDB-Code: 1D7M) wurde diesbezüglich als Haupttemplat für alle vier Segmente benutzt. Dieses Homodimer beinhaltet 14 „heptad repeat“-Einheiten. Es setzt sich aus den Aminosäureresten 243 bis 343 zusammen. Mit „Paircoil2“ wurde das „heptad repeat“ auf die Sequenz übertragen und die „coiled coil“-

Wahrscheinlichkeit berechnet (Abbildung 4.1). Eine Kontrolle erfolgte visuell und durch Vergleich mit den Literaturangaben [11, 44].

Bevor mit der Konstruktion begonnen werden konnte, mussten die genauen Längenverhältnisse der vier „coiled coil“-Segmente mit dem richtigen „heptad repeat“-Muster bestimmt werden.

Am Beispiel des Keratins K8(IF Typ II)/18(IF Typ I) wird dazu im Folgenden die genaue Vorgehensweise erklärt. Für alle anderen Keratinmodelle wurde analog verfahren. (Das erste Modell wurde von K8/18 gebaut, weil eine Bindung mit der Domäne B nachgewiesen ist. Außerdem wurde anfangs angenommen, dass es auch die Zielstruktur im Haar sei.) Im Gegensatz zum Cortexillin I handelt es sich beim Keratin 8/18 um ein Heterodimer. Die genaue Analyse und Aufarbeitung beider Sequenzen erfolgte mit Hilfe von Literaturangaben, Sequenzalignments und Pair Coil Scoring Form [1, 45]. Über die Angaben der „linker“-Positionen konnte beispielsweise auf Start- bzw. Endpositionen der angrenzenden Segmente geschlossen werden. Das von Pair Coil Scoring Form unterlegte „heptad repeat“-Muster wurde dann mit Literaturangaben verglichen und visuell kontrolliert. Für die „linker“ L12 und L2 waren die Literaturangaben eindeutig und konnten so problemlos übernommen werden. Die Informationen zum „linker“ L1 waren dagegen nicht einheitlich. Er wurde deshalb zugunsten eines um zwei plus zwei Aminosäurereste längeren Segmentes 1B um dementsprechend zwei plus zwei Aminosäurereste verkürzt. Die möglichst vollständige Konstruktion der Segmente hatte diesbezüglich Vorrang. Auch die Grenzen der gesamten „rod“ waren nicht ganz eindeutig zu entnehmen. Sie wurden deshalb auch eher großzügig interpretiert. Sollte sich später herausstellen, dass diese unsicheren Positionen in die Bindungsregion involviert sind, so hätte man diesen Teil neu bedenken müssen. Alle anderen Strukturkoordinaten wurden jedoch dadurch nicht beeinflusst. Im Folgenden wird das Ergebnis dieser Prozedur dargestellt (siehe auch Abbildung 4.2).

Die gesamte „rod“ beinhaltet 315 Aminosäurereste jedes Keratins. Das erste Segment 1A beginnt und endet mit einer d-Position im „heptad repeat“. Es wird aus den Aminosäureresten 83-125 des K8 und 73-115 des K18 aufgebaut.

Daran schließt sich der „linker“ L1 aus je 10 Resten an. Das nächste Segment 1B beginnt bzw. endet auf einer c-Position im „heptad repeat“ und beinhaltet die Sequenzpositionen 136-236 (K8) und 126-226 (K18). Der folgende „linker“ L12, das Verbindungselement der 1-Segmente und 2-Segmente, enthält je 17 Aminosäurereste. Das Segment 2A, Positionen 254-272 (K8) und 244-262, (K18) beginnt auf einer c-Position und endet auf einer g-Position im „heptad repeat“. Es ist mit je 19 Aminosäureresten das mit Abstand kleinste Segment. Der „linker“ L2 wird von je acht Resten aufgebaut. Das größte und letzte Segment 2B beginnt und endet auf einer d-Position und baut sich aus den Aminosäureresten 281-397 (K8) sowie 271-387 (K18) auf. Als Besonderheit weist es im Gegensatz zu allen anderen Segmenten kein durchgängiges „heptad repeat“-Muster auf. Es ist an einer Stelle nach einer d-Position durch die Diskontinuität „stutter“ der Reste 341-344 (K8) und 331-334 (K18) unterbrochen. Im Anschluss daran wird es beginnend auf einer a-Position bis zum Schluss kontinuierlich fortgesetzt.

Zum Vergleich sind die „coiled coil“-Wahrscheinlichkeiten für beide Sequenzen von K8/18, bereinigt von „head-“ und „tail“-Bereichen graphisch, aufgetragen (Abbildung 4.3). Alle Aminosäurereste der Segmente weisen in ihrer Konstellation zueinander eine hohe Wahrscheinlichkeit zur Ausbildung einer „coiled coil“-Struktur auf. In den Bereichen stark abnehmender Wahrscheinlichkeiten kommen „linker“-Bereiche zu liegen.

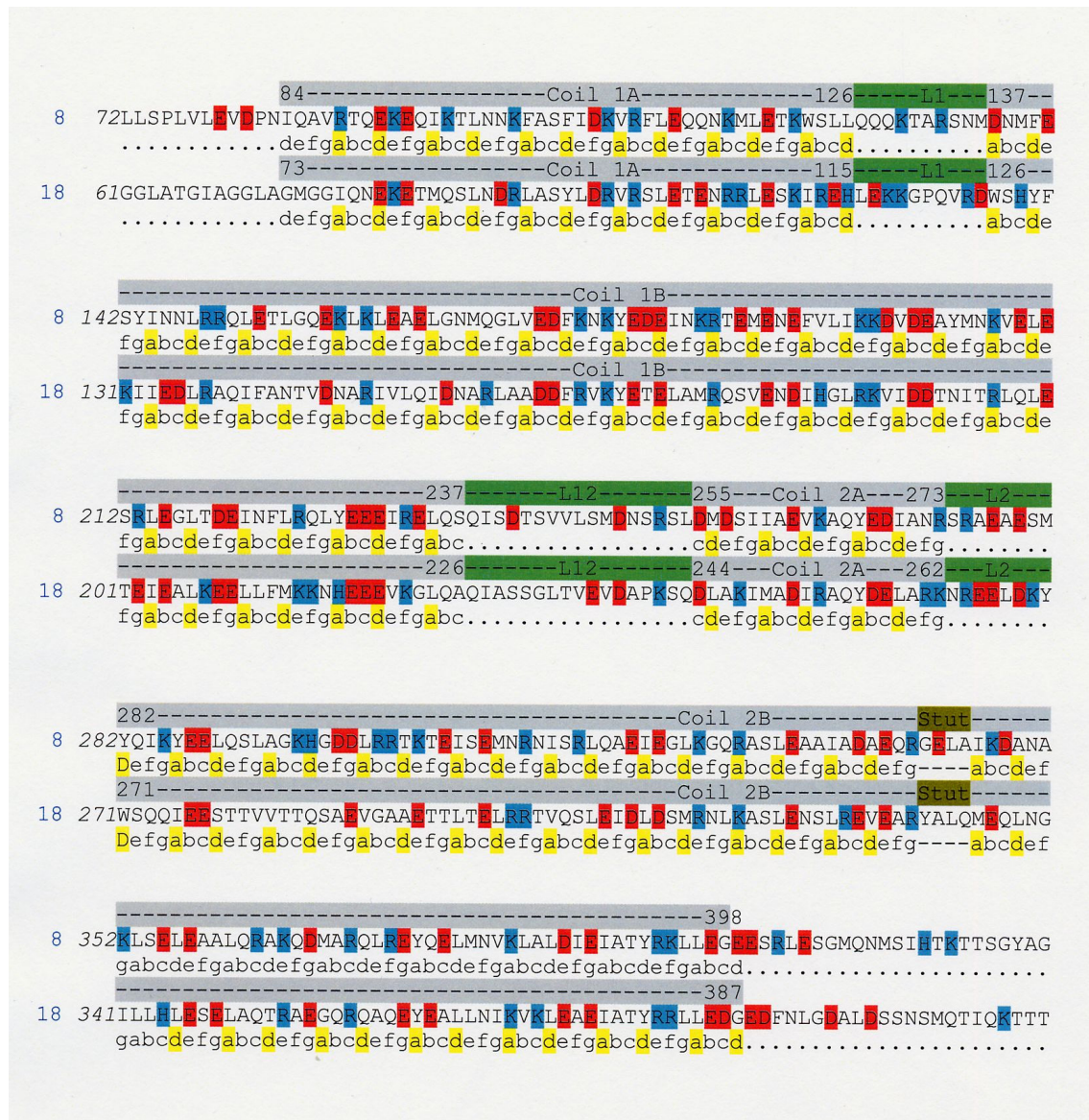


Abbildung 4.2: Die "rod" von K8/18

Über der Sequenz sind die „coiled coil“ Regionen (grau), die „linker“ (grün) sowie der „stutter“ (braun) angegeben. Farbiger unterlegt ist die Sequenz für saure (E, D rot) und basische (R, K und H hellblau) Aminosäurereste. Unterhalb der Sequenz ist das „heptad repeat“ angegeben. Dabei sind die Positionen der apolaren Reste des „hydrophobic core“ hervorgehoben (gelb).

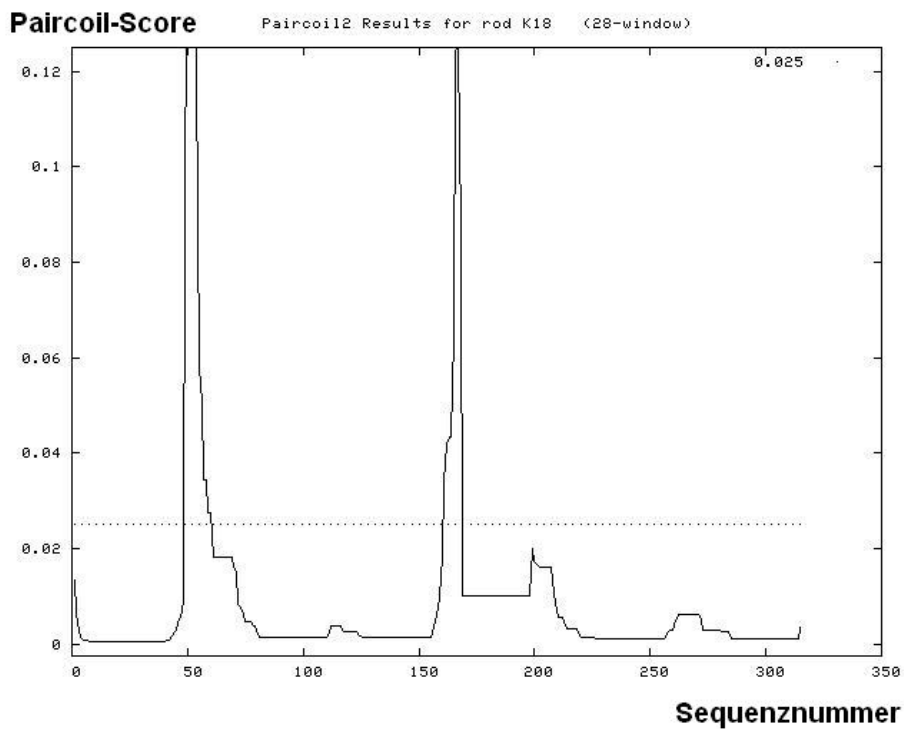
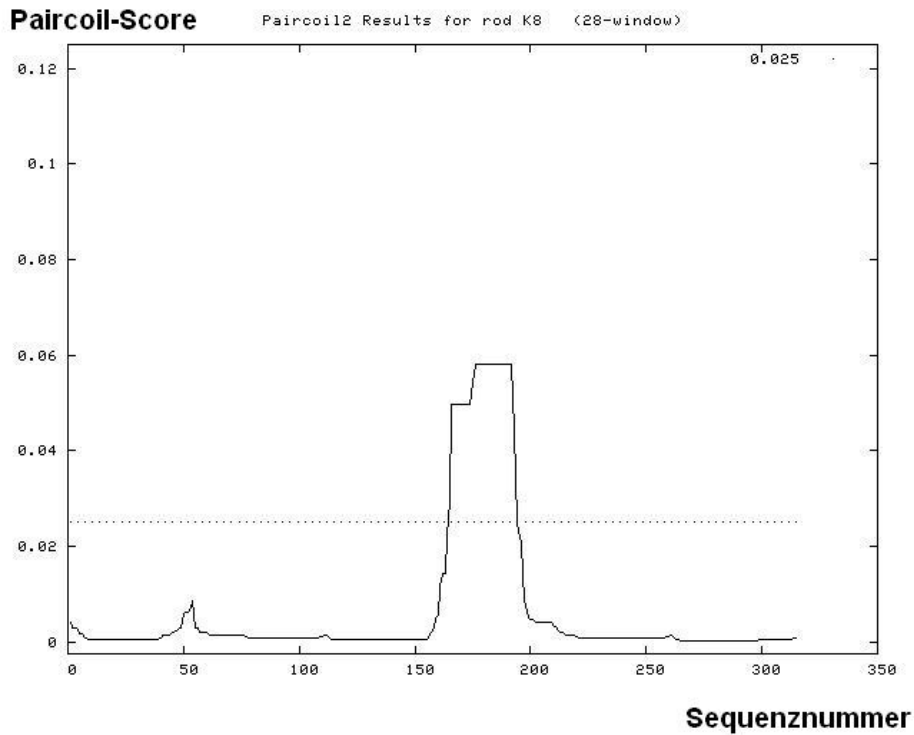


Abbildung 4.3: „rod“-K8/18 „coiled coil“-Wahrscheinlichkeit

Die Sequenz der „rod“ wurde mit dem Programm Paircoil2 ausgewertet. oben: K8; unten:K18;. Die „rod“ beginnt bei 0. Der Paircoil2 Graph indiziert hohe „coiled coil“-Wahrscheinlichkeit für die Segmente(unterhalb der gepunkteten Linie). In den „linker“-Bereichen fällt sie teilweise stark ab.

Der nächste Schritt zur Konstruktion des Modells bestand in der entsprechenden Modifizierung der Kristallstruktur von Cortexilin I, um die Strukturkoordinaten der „rod“ von Keratinen ohne „linker“ zu erhalten. Diese Arbeiten wurden mit dem Programm Sybyl durchgeführt. Zuerst wurden alle Aminosäureseitenketten entfernt, es wurde also nur das „backbone“ der Kristallstruktur als Templat verwendet bzw. verändert. Die beiden kürzeren Segmente 1A und 2A wurden erhalten, indem man das Templat an den richtigen Stellen verkleinert. Das Segment 1B besteht analog zum Templat aus 101 Aminosäureresten. Dennoch musste das Templat verändert werden, da sein „heptad repeat“ nicht mit einer α -Position beginnt. Dazu wurde es entsprechend verkürzt. Die fehlenden Bereiche wurden durch Überlagerung eines zweiten Templates, dessen Verkürzung und anschließender kovalenter Verbindung, erhalten. Das Segment 2B wurde zu dessen Konstruktion in drei Abschnitte unterteilt. Der erste und der letzte Abschnitt waren die Bereiche vor bzw. nach dem „stutter“. Sie wurden auf ähnliche Weise wie das Segment 1B erhalten. Die „stutter“-Region („stutter“ und angrenzende Bereiche, welche zum Überlagern dienten) wurde aus der Kristallstruktur des „Cys2“-Fragmentes von Vimentin (PDB-Code: 1GK4) herausgeschnitten. Anschließend wurden alle drei Abschnitte zum vollständigen Segment 2B zusammengefügt.

Die nächste Herausforderung beim Modellbau bestand darin, die einzelnen Segmente so gut wie möglich im Raum zu arrangieren. Unter Verwendung zahlreicher geometrischer Hilfspunkte, Hilfslinien und Hilfsebenen wurden die einzelnen Segmente den Vorstellungen entsprechend axial zueinander im Raum angeordnet. Die axialen Abstände benachbarter Segmente, bedingt durch die „linker“, wurden dabei berücksichtigt.

Das so erhaltene „backbone“-Modell der „rod“ wurde im letzten Schritt um die Aminosäureseitenketten (entsprechend der Sequenz) erweitert. Dies geschah mit dem Programm SCWRL. Die stereochemische Qualität wurde abschließend mit dem Programm Procheck kontrolliert und sichergestellt (Abbildungen 4.4-4.5).

PROCHECK

Ramachandran Plot

K8_18

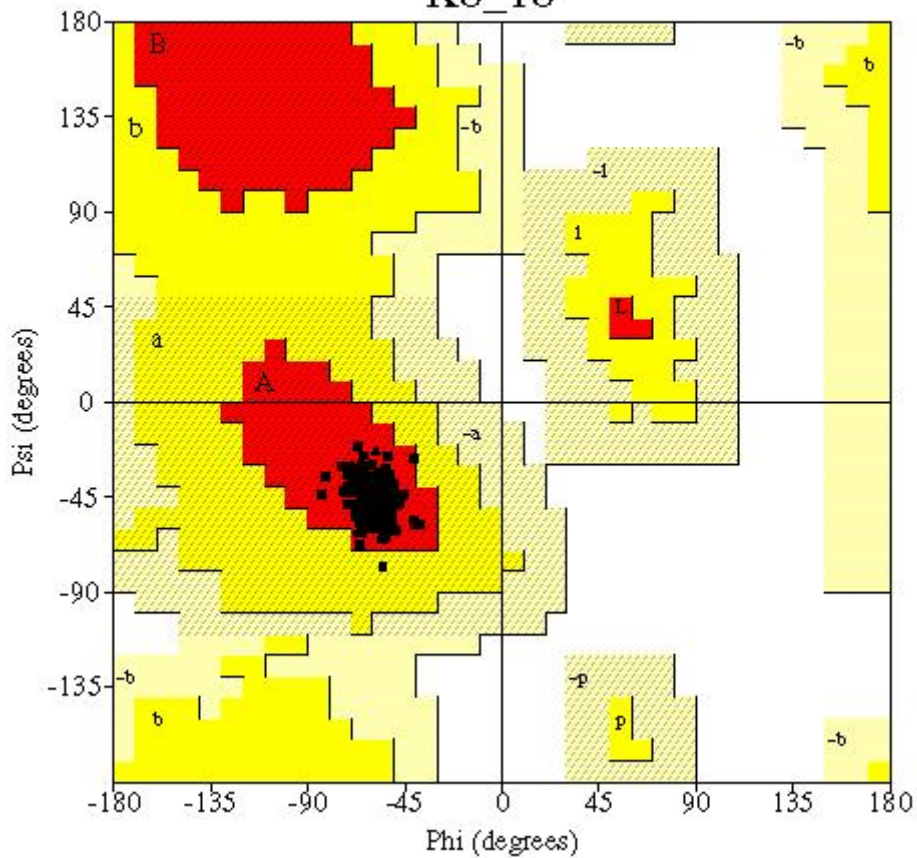


Abbildung 4.4: Ramachandran Plot von K8/18-"rod"

Alle Psi- und Phiwinke (schwarze Quadrate) liegen in erlaubten Bereichen (rot und gelb). Die enge Verteilung leitet sich aus der Gleichmäßigkeit der „coiled coil“ ab.

PROCHECK

Page 1

Chi1-Chi2 plots

K8_18

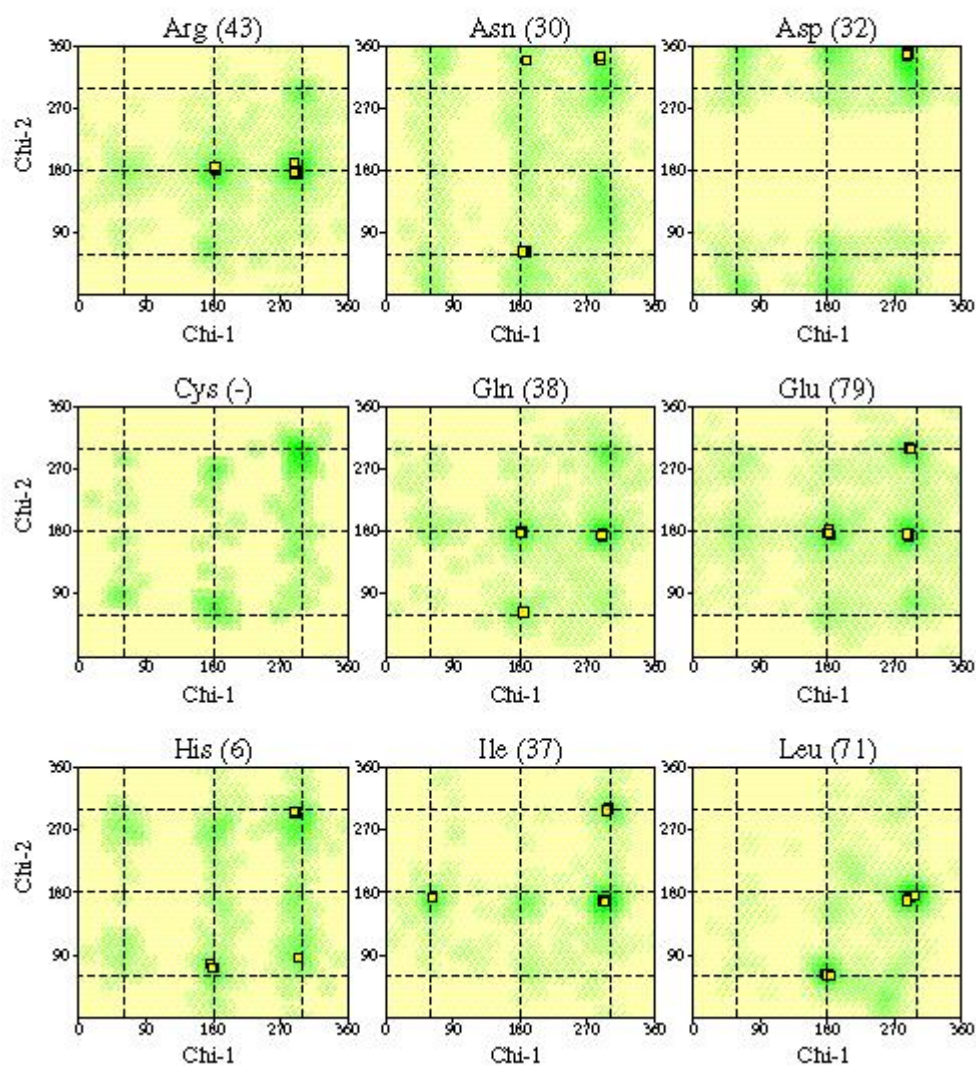


Abbildung 4.5 Teil1: Chi1- und Chi2-Verteilung von K8/18-"rod"

Alle Chi1- und Chi2-Winkel(gelbe Quadrate) liegen in erlaubten Bereichen(grün).

PROCHECK

Page 2

Chi1-Chi2 plots

K8_18

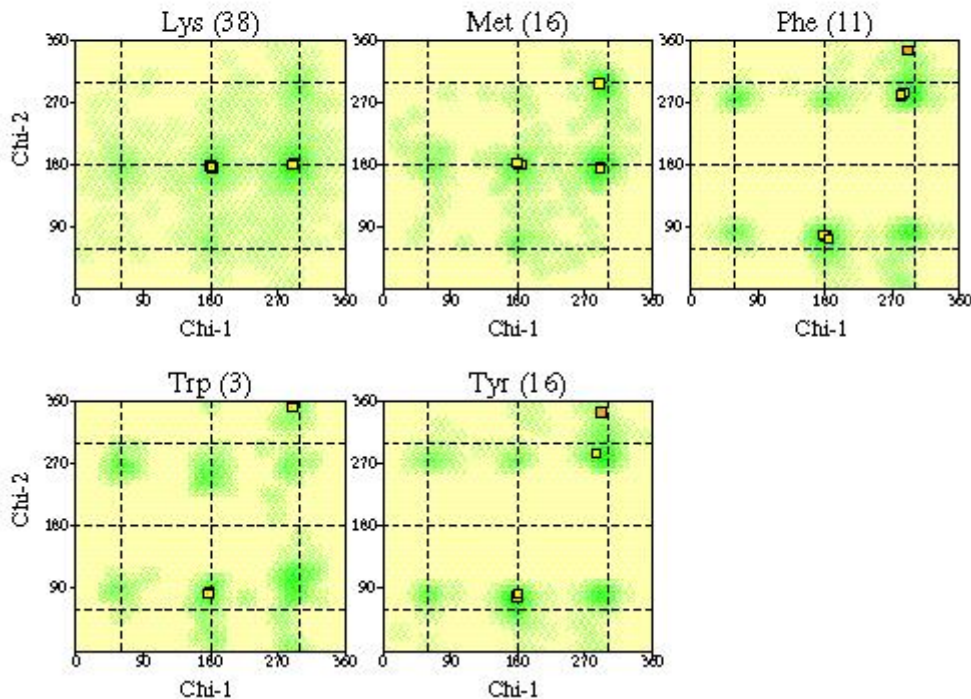


Abbildung 4.5: Teil 2: Chi1- und Chi2-Verteilung von K8/18-"rod"

Alle Chi1- und Chi2-Winkel(gelbe Quadrate) liegen in erlaubten Bereichen(grün).

Die van der Waals-Oberfläche des fertigen Modells könnte man am besten mit einem viergeteilten, stabförmigen und gleichmäßigen Gebilde beschreiben (Abbildung 4.6). Bedingt durch die periodische Superspiralisierung der „coiled coil“ und die damit hervorgerufene Furchenbildung, ähnlich der DNA, gewinnt man im Zweidimensionalen zusätzlich den Eindruck, als würde der Stab an seiner Oberfläche sinuswellenförmige Berge und Täler beschreiten. Das zusätzlich abgebildete elektrostatische und das lipophile Potential verdeutlichen außerdem, dass die Verteilung der molekularen Wechselwirkungsfelder keine Periodizität aufweist. Besonders beim Elektrostatischen Potential werden die Unterschiede zwischen den Segmenten und innerhalb eines Segmentes deutlich. Eine Ausnahme bildet natürlich der „hydrophobic core“ im Proteininterface, der als globale Konstante die „coiled coil“ in gestapelter Packung stabilisiert (Abbildung 4.7).

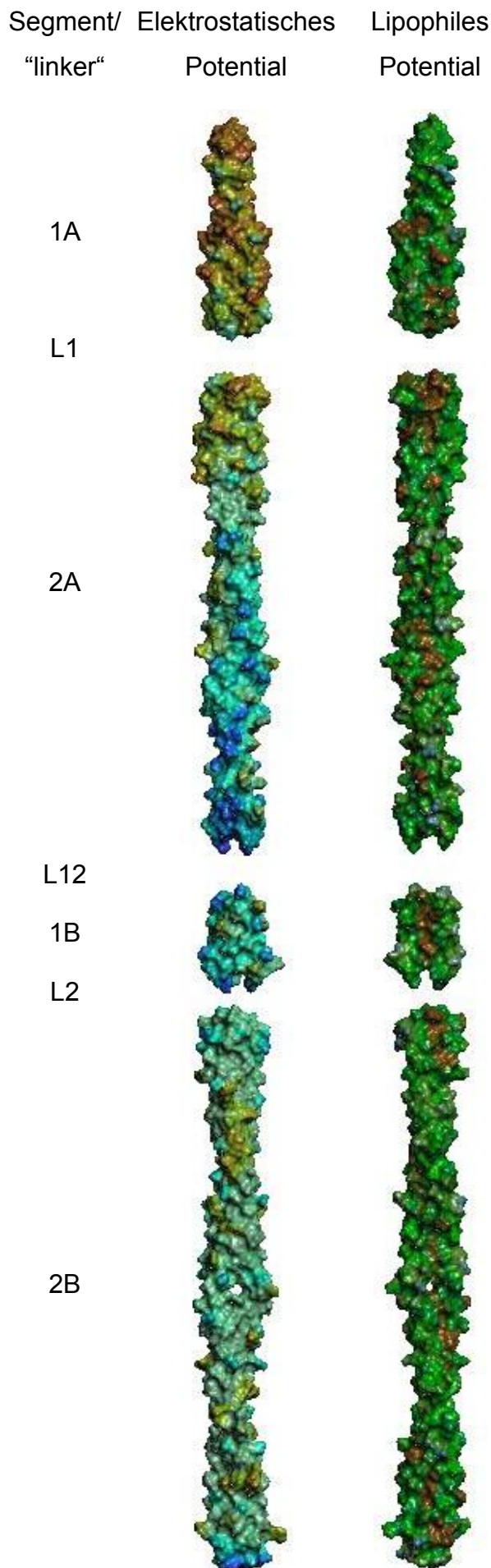


Abbildung 4.6: Oberflächendarstellung der K8/18-"rod"
links: Elektrostatisches Potential mit positiv geladenen (rote und braune Farbtöne) und negativ geladenen (blaue Farbtöne) Bereichen.
rechts: Lipophiles Potential mit hydrophoben (braune Farbtöne), gemischten (grüne Farbtöne) und hydrophilen (blaue Farbtöne) Bereichen.

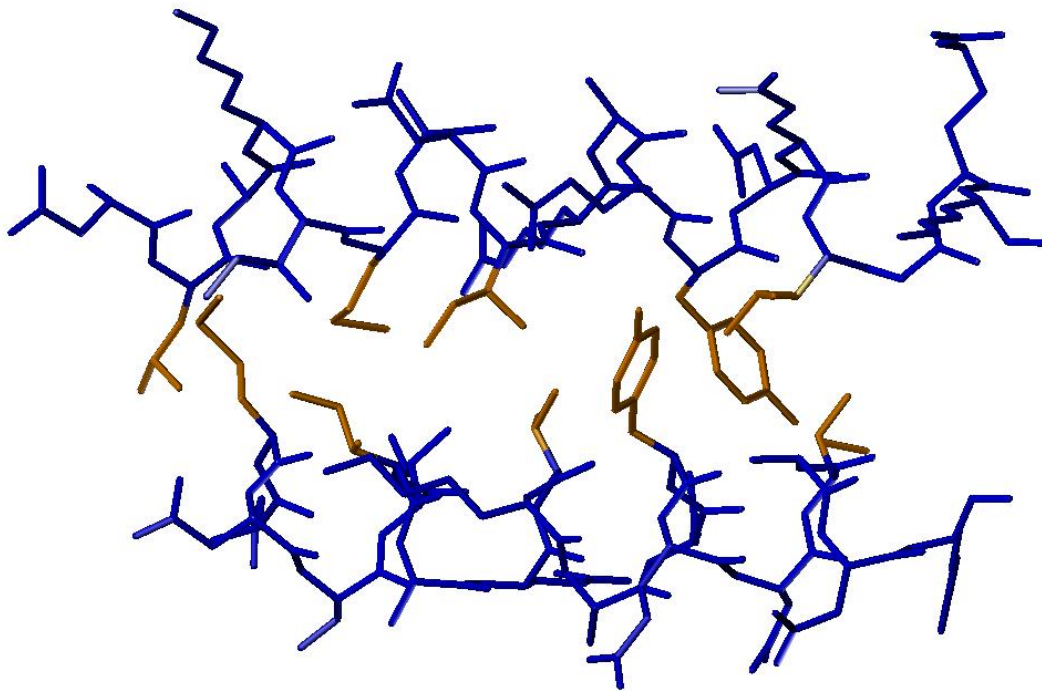


Abbildung 4.7: K8/18-"hydrophobic core" im Segment 1B
Dargestellt ist das komplette Segment 1B von K8/18. Die beiden Helices der „coiled coil“ (blau), die Seitenketten des „hydrophobic core“ auf den a- und d-Positionen (braun).

Konstruktion der Keratinmodelle von Haut und Haar

Die Länge und die jeweilige „heptad repeat“-Startposition der vier Segmente der „rod“ sind höchstwahrscheinlich identisch. Diese Schlussfolgerung ergab sich aufgrund der Analyse von multiplen Sequenzalignments mit dem Programm Clustal W. Insofern konnte das „backbone“-Modell vom K8/18 unverändert als Templat genutzt werden. Das weitere Vorgehen (das Einfügen der Seitenketten und Kontrolle des Modells) verlief dann analog zur Modellkonstruktion der K8/18-„rod“. Folgende Hautkeratin-„rod“-Strukturmodelle wurden generiert:

- K1/10
- K5/14

Folgende Haarkeratin-„rod“-Strukturmodelle wurden generiert:

- K32/82
- K35/82
- K32/85
- K35/K85

Aufgrund der großen Ähnlichkeit von K39 und K40 zu K35 und K32 wurden die restlichen Strukturmodelle nicht konstruiert (siehe Abschnitt 1.6). Bei Bedarf konnten durch Mutationen einzelne Abschnitte von K35 bzw. K32 schnell zu K39 oder K40 transformiert werden.

4.2 Auf der Suche nach den Bindungsregionen

Das allgemeine Vorgehen, um dieses Ziel zu erreichen, stellt sich folgendermaßen dar: Am Anfang stand die Formulierung einer Arbeitshypothese. Im weiteren Verlauf wurde versucht, diese mit verschiedenen Ansätzen und Methoden zu bestätigen. Gelingt dies nicht, so wurde in Folge dessen eine neue Arbeitshypothese erstellt und gleichermaßen überprüft. Die visuelle Betrachtung und Analyse der Strukturmodelle im Stereomodus erlangte in diesem Zusammenhang verhältnismäßig große Bedeutung.

Arbeitsplan für das weitere Vorgehen

Zunächst sollte die Bindungsregion auf der Domäne B näher charakterisiert werden.

- Welche Region der Domäne B ist prinzipiell befähigt, mit einer „coiled coil“-Struktur in Wechselwirkung zu treten?

Diese Arbeitshypothese hätte dann durch Mutationsstudien des Kooperationspartners BASF überprüft werden können. Das Zwischenziel wäre damit erreicht (siehe Abschnitt 2.2).

- Experimentelle Überprüfung der Arbeitshypothese

Nach einer eventuellen Bestätigung könnte dann der Versuch unternommen werden, die Bindungsregion am Keratin und somit die Proteinkomplex-Wechselwirkungsgeometrie zu bestimmen.

- Bestimmung der Protein-Protein-Wechselwirkungsgeometrie

Aufgrund dessen wäre man dann in der Lage, Mutationsvorschläge für verbesserte Bindungseigenschaften abzuleiten.

- Mutationsvorschläge für verbesserte Bindungseigenschaften

Diese hätten dann noch einer experimentellen Bestätigung durch den Kooperationspartner BASF bedurft. Das Hauptziel wäre erreicht (siehe Abschnitt 2.2).

Im Folgenden stelle ich das Vorgehen und die entstandenen theoretischen und experimentellen Ergebnisse dar.

4.2.1 Studium von Proteinkomplexen aus der PDB

Strukturkoordinaten von Proteinkomplexen in der PDB wurden im Stereomodus des Programms Sybyl visuell betrachtet und analysiert. Zusätzlich wurden zum Vergleich Protein-Protein-Wechselwirkungsenergien mit dem Dockingprogramm Hex berechnet. Damit sollte die Problematik beleuchtet werden, inwiefern bestimmte Proteine miteinander in Wechselwirkung treten bzw. stabile Komplexe bilden können. Besonders interessant waren Komplexe, deren freie, ungebundene Partner ebenfalls in der PDB veröffentlicht sind. Durch Vergleich konnte so das Ausmaß an konformatorischen Veränderungen während des Komplexbildungsprozesses abgeschätzt werden.

Es fiel auf, dass die Oberflächen der Bindungsregionen schon vor der Komplexbildung meist komplementär zueinander gestaltet sind. Es gibt aber auch Ausnahmen, wobei zumindest ein Teilbereich der Protein-Interfacekomplementarität erst während der Komplexbildung induziert wird. Vermutlich gibt es zwischen beiden Situationen fließende Übergänge. Eine Veränderung von Seitenkettenkonformationen im Proteinkomplex, verglichen mit den freien Proteinen, konnten jedoch beobachtet werden. Überaus wichtig scheint außerdem zu sein, dass die Komplementarität einen gewissen Grad an Spezifität aufweist. Diese Spezifität kommt durch eine „zerklüftete Topographie“ zustande. Die Proteine dürfen also im Interface nicht einfach nur planar sein. Somit gibt es folglich nur eine Wechselwirkungsgeometrie, in der beide

Komplexparker komplementär ineinander greifen können. Dies bedingt eine Art „mechanische Rigidität“, die für einen stabilen Proteinkomplex unabdingbar ist. In großen funktionalen Oligomeren wie Enzymen, findet sich eine exakte Übereinstimmung der Komplementarität bis ins kleinste Detail. Dies mag in der Evolution begründet liegen, in welcher die Natur die Strukturen über lange Zeiträume hinweg präzise aneinander anpassen konnte. Im Gegensatz dazu kann man in Antikörper-Antigenkomplexen auch gewisse Freiräume beobachten, in welchen keine Proteinatome zu liegen kommen. Diese Freiräume bleiben aber zumeist nicht unbesetzt, sondern werden häufig (besonders dann, wenn diese polare Eigenschaften aufweisen) mit Wassermolekülen ausgefüllt. Wasser spielt die Rolle eines „evolutionären Kittes“, der es diesen Proteinen erlaubt, trotz kleinerer Interface-Komplementaritätsdifferenzen, stabile und hoch spezifische Proteinkomplexe zu bilden. Weiterhin zeigte sich, dass apolare Bereiche an der Oberfläche nach Möglichkeit im Interface zu liegen kommen. Sie können sich so einer energetisch ungünstigen, wässrigen Umgebung entziehen. In dieser Konstellation würde die Grenzflächenenergie in wässriger Umgebung minimiert und die Komplexstabilität vergrößert werden. Auffällig war außerdem die hohe Anzahl an geladenen Aminosäuren an der Proteinoberfläche. Ich vermute, dass dadurch eine unspezifische Komplexbildung unterbunden wird. Auf der anderen Seite trägt eine ganz bestimmte Verteilung von Elektrostatischen Potentialen auf der Oberfläche zur Spezifität und Erkennung von bindenden Protein-Protein Wechselwirkungen bei. Die Analyse der gegenseitigen Wechselwirkungsarten im Interface ergab allerdings kein einheitliches Bild. So können mitunter Salzbrücken beobachtet werden, dieses entspräche einer elektrostatischen Komplementarität. Des Weiteren kommen viele verschiedene Kombinationen von intermolekular wechselwirkenden Aminosäureseitenketten vor, solange die Oberflächenkomplementarität gewahrt bleibt. Sogar die räumliche Nähe von gleichgeladenen Aminosäureresten und die damit verbundene starke Abstoßung kann toleriert bzw. kompensiert werden. Eine Erklärung hierfür könnte in der Anwesenheit von Wasser und Ionen begründet sein. Diesen

Sachverhalt kann man z.B. an dem Dimer der Kristallstruktur der Domäne B eindrucksvoll verdeutlichen:

Die Domäne B liegt sowohl unter den Bedingungen des Aufreinigungsprozesses (also in wässriger Lösung) als auch unter Kristallisationsbedingungen als Dimer vor [22]. Diese Eigenschaft der Domäne B, als stabiles Dimer in wässriger Lösung, wurde auch durch den Kooperationspartner BASF bestätigt. Das Programm Hex berechnete die Werte der Monomer-Monomer-Wechselwirkungsenergien im Vakuum. Dazu wurden die beiden Monomere der Kristallstruktur entlang ihrer intermolekularen Achse separiert. Die Geometrie der Kristallstruktur konnte von dem Programm nur gefunden werden, indem eine Rotation während des Docking-Prozesses verhindert wurde (totale Einschränkung des Suchraumes). Der Gesamtwert der Wechselwirkungsenergie beträgt +44 kJ/mol. Er setzt sich aus -444 kJ/mol für das (Hex)-„Shape“ - Potential und +488 kJ/mol für das Elektrostatische Potential zusammen. Die Interface-Kontaktfläche (nicht von Wassermolekülen frei zugängliches Proteininterface) besitzt mit 400 \AA^2 zudem einen vergleichsweise durchschnittlichen Betrag. Diesen Werten nach zu urteilen, wäre die Dimerisierung der Domäne B folglich gar nicht möglich. Der hohe, abstoßend wirkende, elektrostatische Energieanteil kann durch die räumliche Nähe von negativ geladenen Aminosäureresten in der Kristallstruktur erklärt werden. Er muss also in wässriger, ionenhaltiger Lösung zwangsläufig stark reduziert sein, um eine Dimerisierung zu gewährleisten. Dieses Beispiel zeigt allgemein, dass man die im Vakuum berechneten Kraftfeld-Wechselwirkungsenergien differenziert und kritisch betrachten muss. Im Vergleich zur Domäne B betragen die gleichermaßen berechneten Wechselwirkungsenergien des Dimers der Domäne C -387 kJ/mol für das Hex-„Shape Potential und -136 kJ/mol für das Elektrostatische Potential. Die Kontaktfläche beträgt 380 \AA^2 . Die Dimerisierung ist in diesem Fall also nachvollziehbar.

Aus diesen Überlegungen heraus ergaben sich folgende Kriterien, die in der Suche nach den Bindungsregionen meiner zu betrachtenden Proteine berücksichtigt worden sind:

- spezifisch-komplementäre Oberflächen
- lipophile Komplementarität
- elektrostatische Komplementarität
- Phenylalanin, Tryptophan und Tyrosin sind nach Möglichkeit im Interface zu platzieren
- vertretbare konformative Veränderungen können zu komplementäreren Oberflächen führen (Abschätzung)
- mindestens 400 kJ/mol (Hex-„Shape“ Potential; Protein-Protein-Wechselwirkungsenergie)

4.2.2 Die Bindungsregion auf der Domäne B als funktionales Dimer

Die erste Arbeitshypothese bezüglich der Bindungsregion auf der Domäne B war die Vorstellung, dass das Homodimer der Kristallstruktur ein funktionales Dimer ist und kein unter Kristallisationsbedingungen entstandenes Artefakt. Das würde bedeuten, dass die Dimerisierung eine Voraussetzung für die Bindung zum Keratin ist. Dies stand in Einklang mit der Auffassung des Kooperationspartners BASF.

Die Untersuchungen (bestehend aus Docking mit Hex, Moleküldynamik-Simulationen und visueller Stereobetrachtung) gaben außerdem keine Hinweise darauf, dass in die Bindungsregion Bereiche aus beiden Monomeren involviert sein könnten. Der Vergleich mit der Kristallstruktur der homologen Domäne C, welche ebenfalls ein Homodimer ist, unterstützte dahingehend meine Auffassung. Denn die Konstellation der beiden Monomere (Kristallstruktur der Domäne C) zueinander unterscheidet sich stark von jener der Kristallstruktur der Domäne B. Beide Dimere besitzen also kein, durch Dimerisierung entstandenes, „homologes Interface“ (was man bei homologen, funktionalen Dimeren erwarten kann). Meiner Auffassung und Einschätzung nach, könnte

durch diese Arbeitshypothese kein stabiler, von Wassermolekülen nicht frei zugänglicher Proteinkomplex gewährleistet werden. Das Programm Hex lieferte Geometrien mit relativ niedrigen Wechselwirkungsenergien:

- die Werte lagen unter 300 kJ/mol Hex-„Shape“ Potential

Diese Werte wurden aber auch von vielen, völlig unterschiedlichen Komplexgeometrien erreicht. Sie kommen zustande, indem das Programm die größten Oberflächen kombiniert. Die genaue Betrachtung der Komplexe im Stereomodus offenbarte aber kein Potential für konformative Veränderungen, die zu einem spezifisch-komplementären Interface führen würden. Meiner Einschätzung nach ist das Dimer in der Kristallstruktur ein Artefakt und somit in vivo nicht als funktionales Dimer in Hinsicht einer IF Protein-Bindedomäne existent. Die Dimer-Hypothese wurde folglich von mir verworfen. Stattdessen konzentrierte ich mich auf das Monomer der Kristallstruktur. Ich vermutete, dass ein Monomer für die Bindung ausreichend ist (Monomer-Hypothese).

- Ist ein Monomer der Domäne B für die Bindung zum Keratin ausreichend?

Wenige Zeit später kontaktierte ich in diesem Zusammenhang die Arbeitsgruppe, welche die Kristallstrukturen der Domäne B und C publizierte (siehe Abschnitt 1.3.3) [22]. Sie erwähnten mir gegenüber, dass sie noch eine weitere unveröffentlichte Kristallstruktur, bestehend aus der Sequenz der Domäne A und der Domäne B des Desmoplakin C-terminal, besitzen. Diese weist keine Dimerisierung auf. Außerdem liegt dieses Proteinkonstrukt auch in wässriger Lösung als Monomer vor. Dadurch wird die Hypothese eines funktionalen Dimers ebenfalls entkräftet. Außerdem hätten sie die unbewiesene Auffassung, dass das Dimer in der Kristallstruktur der Domäne B ein Artefakt ist. Im Nachhinein steht mir eine neue, nicht veröffentlichte Kristallstruktur der Domäne B des Kooperationspartners BASF zur Verfügung. Diese zeigt ebenfalls eine Dimerisierung. Es handelt sich dabei jedoch um ein anders aufgebautes Dimer, wodurch die Artefakttheorie ebenfalls unterstützt wird.

- Die Dimerisierung der Domäne B ist wahrscheinlich ein Artefakt und somit für die Bindung zum Keratin nicht notwendig.

4.2.3 Das erste Experiment resultierend aus der „Monomer-Hypothese“

Der Kooperationspartner BASF hat aufgrund der „Monomer-Hypothese“ und auf der Basis der publizierten Kristallstruktur (die „neue“ BASF-Kristallstruktur gab es zum damaligen Zeitpunkt noch nicht) die Mutation D2407A (Aspartat2407 nach Alanin) durchgeführt. Diese Mutation sollte die Dimerisierungswahrscheinlichkeit herabsetzen, ohne die Bindungsaktivität zum Keratin zu beeinträchtigen.

- Durchführung der Mutation D2407A durch den Kooperationspartner BASF

Die Domäne B liegt sowohl unter den Bedingungen des Aufreinigungsprozesses als auch unter Kristallisationsbedingungen als Homodimer vor [22]. Auch unter den Bedingungen einer Gelelektrophorese (mit SDS als Detergenz) liegt sie im dimerisierten Stadium vor [Aussage der BASF]. Es muss sich also um ein außerordentlich stabiles, so genanntes „SDS-resistentes“ Dimer handeln [46-50]. Vergleicht man im Nachhinein die beiden Dimerkonstellationen der publizierten und der BASF-Kristallstruktur, so erscheint das Dimer der publizierten wesentlich stabiler (vor allem in Gegenwart von SDS). In dieser ist der Aminosäurerest D2407 des einen Monomers A von zwei Argininresten und einem Threoninrest des anderen Monomers B räumlich nahe umgeben. Dadurch verzahnen sich beide Monomere, wobei sich die dabei ausgebildete Salzbrücke in Proteinumgebung (frei von Lösungsmiteleinflüssen) befindet. Umgekehrt gilt das auch für D2407 des Monomers B, welches von zwei Argininresten und einem Threoninrest des anderen Monomers A umgeben ist (siehe Abbildung 4.8). (Die Konstellation beider Monomere in beiden jeweiligen Kristallstrukturen zueinander besitzt als Symmetrieelement eine Drehachse mit dem Drehwinkel 180 Grad). Das Dimer der BASF-Kristallstruktur weist keine stabile Verzahnung auf. Das Dimer-Interface ist größtenteils mit Wassermolekülen besetzt. Zwei vorkommende Salzbrücken (zwischen Arginin- und Asparaginsäureresten) sind lösungsmittel exponiert und somit stark abgeschwächt. SDS-Moleküle könnten sie außerdem leicht zerstören, im Gegensatz zu den starken, weil in Proteinumgebung „verborgenen“ Salzbrücken in der publizierten Kristallstruktur. Ich halte es deshalb für sehr

wahrscheinlich, dass die Dimerisierung in der publizierten Kristallstruktur in Lösung, also auch nach der Bindung zum Keratin, vorliegt. Die doppelt vorhandenen, starken elektrostatischen Wechselwirkungen (Salzbrücken) zwischen den beiden Monomeren, kommen in der hergestellten Mutante D2407A nicht mehr zustande. In Folge dessen sollte diese Dimerisierung weniger wahrscheinlich bis unwahrscheinlich sein. Würde die Bindungsaffinität dieser Mutante zum Keratin nicht geringer sein als die des Wildtyps, so wäre das ein weiteres Indiz für die „Monomer-Hypothese“. (Vorausgesetzt, dieses Dimer liegt auch wirklich in Lösung vor.)

- Haben Mutante und Wildtyp vergleichbare Bindungseigenschaften?

Ein weiterer Nebeneffekt dieser Mutante könnte eine verbesserte Löslichkeit sein. Wenn die Dimerisierungswahrscheinlichkeit herabgesetzt ist, so könnte dies bedeuten, dass die Neigung zur Ausbildung von Oligomeren ebenfalls vermindert ist. Das Nebenziel b wäre somit erreicht (siehe Abschnitt 2.2).

- Besitzt die Mutante verbesserte Lösungseigenschaften?

Besäße die Mutante verbesserte Lösungseigenschaften, so sollten sich im Allgemeinen auch die Bindungseigenschaften verbessern, zumindest relativ gesehen (höhere Proteinkonzentrationen wären möglich, wodurch die Absättigung am Haar wesentlich schneller ablaufen könnte). Das Hauptziel wäre damit zum Teil erfüllt (siehe Abschnitt 2.2).

- Besitzt die Mutante verbesserte Bindungseigenschaften?

Wenn für die Bindung zum Keratin ein Monomer ausreichend ist, dann sollte die Mutante (die möglicherweise verstärkt als Monomer vorliegt) auch eine höhere Bindungskonstante bzw. Assoziationskonstante haben. Denn eine Verringerung der Massenträgheit um 50 Prozent (Dimer zu Monomer) sowie die effektiv höhere Konzentration an bindenden Teilchen bei gleicher Proteineinwaage, könnte sich bei der Mutante gegenüber dem Wildtyp positiv bemerkbar machen.

- Besitzt die Mutante eine höhere Assoziationskonstante und somit auch eine höhere Bindungsaffinität?

Man würde also zur Besetzung der jeweils gleichen Anzahl an Keratinbindungsstellen durch die Mutante mit entsprechend weniger Proteinmenge verglichen mit dem Wildtyp (Dimer der Domäne B), auskommen.

Der Kooperationspartner BASF hat mir mitgeteilt, dass die Bindungsaktivität (bezogen auf die Proteineinwaage bzw Konzentration) der Mutante und des Wildtyps in etwa gleich sind. Die Bestimmung der Aktivität läuft über die photometrische Bestimmung des Substratumsatzes der pro Monomereinheit über (His-Tag-) gekoppelten Peroxidase [51]. Es wurde jedoch nicht untersucht, ob die Mutante in Lösung verstärkt als Monomer vorliegt. Wäre das so, dann hätte diese Mutante eine höhere Bindungsaffinität, da von den entsprechenden Monomeren mehr Bindungsstellen besetzt sein müssen (eine Peroxidase pro Monomer und Bindungsstelle), als von den Wildtyp-Dimeren (zwei Peroxidasen pro Dimer und Bindungsstelle).

Die Mutation D2407A hat also nicht zu einer Beeinträchtigung der Bindungsaktivität am Keratin geführt. Dies kann als weitere Unterstützung der „Monomerhypothese“ bewertet werden.

- Die Mutante D2407A unterstützt die „Monomerhypothese“.

Da der Kooperationspartner BASF nicht untersucht hat, ob diese Mutante in Lösung verstärkt als Monomer vorliegt (was die ursprüngliche Intension dieser Mutationsstudie war), besteht die Möglichkeit, dass die Mutation D2407A eine höhere Bindungsaffinität im Vergleich zum Wildtyp besitzt. Die Bindungsaktivität der löslichen Proteinfraction der Mutante ist, anders als beim Wildtyp, an eine Renaturierung mit Harnstoff gekoppelt. Das könnte bedeuten, dass es unter den Bedingungen der Proteinexpression zu einer Fehlfaltung kommt. Es ist jedoch für den Kooperationspartner BASF wichtig, dass die native, lösliche Proteinfraction an Keratinstrukturen binden kann. Mit dem Austausch von D2407 durch eine andere Aminosäure als Alanin könnte dieses Problem umgangen werden.

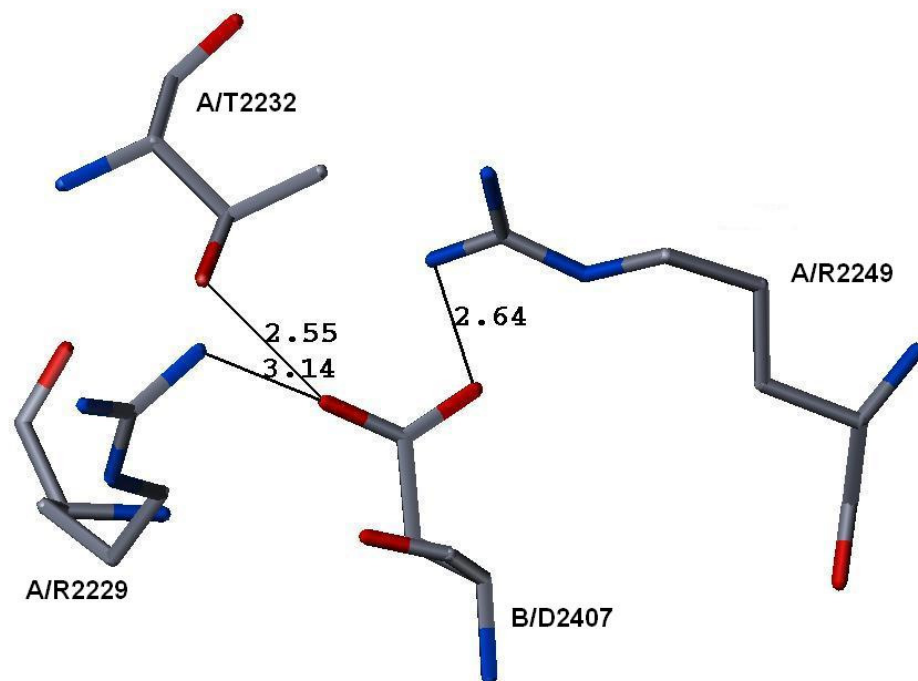


Abbildung 4.8: Salzbrücken im Dimerinterface der Kristallstruktur der Domäne B
 Das Dimer der Kristallstruktur der Domäne B (PDB Code: 1LM7) wird durch die dargestellte Salzbrückenkonstellation im Interface beider Monomere stabilisiert. Der Aminosäurerest D2407 des einen Monomers B ist von zwei Argininresten und einem Threoninrest des anderen Monomers A räumlich nahe umgeben (Abstände in Å). Diese Salzbrückenkonstellation kommt zweimal in dem Dimer vor, da es eine Drehachse als Symmetrieelement besitzt. Eine Mutation von D2407A unterbindet somit zwei starke elektrostatische Wechselwirkungen sowie zwei Wasserstoffbrücken zum Threonin.

4.2.4 Die Bindungsregion auf der Domäne B als Monomer - theoretische und experimentelle Verifizierung der „Monomer-Hypothese“

Die neue Arbeitshypothese besagt, dass die Domäne B als Monomer in der Lage, mit „coiled coil“-Strukturen in bindende Wechselwirkungen zu treten. Zu Beginn dieser Untersuchungen wurden keine Einschränkungen bezüglich der potentiell möglichen Bindungsregionen auf der Oberfläche des Monomers der Domäne B gemacht. Meine Untersuchungen (bestehend aus Docking mit Hex und visueller Stereobetrachtung) brachten zunächst keine brauchbaren Ergebnisse, welche die „Monomer-Hypothese“ unterstützen würde. Später führte ich folgende notwendige Randbedingungen ein:

- die Bindungsregion darf nicht den N-Terminus involvieren (Sequenzposition 2209-2249)
- die Domänen B und C haben denselben Bindungsmodus bzw. dieselbe Bindungsregion
- im Umkehrschluss darf die Bindungsregion auf der Domäne C nicht in den Bereichen des N-Terminus der Domäne B lokalisiert sein (Überlagerung beider Domänen)

Ich begründete diese Entscheidung durch erneuten Vergleich mit der homologen Domäne C. Diese vermag auch ohne entsprechenden N-Terminus an IF Proteine zu binden (siehe Abschnitt 1.3.3 und 1.4.1).

- Die Bindungsregion darf bei beiden Domänen nicht den N-Terminus und den C-Terminus total bedecken

An diese Endstellen schließen sich in vivo weitere Sequenzabschnitte an, eine Bindung in diesen Bereichen erscheint deshalb aus sterischen Gründen unmöglich.

Mit diesen Randbedingungen und den Kriterien (siehe Abschnitt 4.2.1) führte ich also meine Untersuchungen fort. Eine mögliche Bindungsregion konnte ich so nicht näher eingrenzen. Die besten Komplexe lieferten für das Hex-„Shape“ Potential zwar hohe Werte bei 500 kJ/mol. Die Randbedingungen und Kriterien führten jedoch zum Ausschluss dieser Geometrien. In der Zwischenzeit führte ich eine weitere Randbedingung ein:

- Da, wie vermutet, die Domäne B in vitro als Dimer bindet, wäre es somit notwendig (je nach Stabilität des Dimers), dass die Bindungsregion auch am Dimer freizugänglich ist. Gleichmaßen sollte dies nach Möglichkeit auch für das Kristallstruktur-Dimer der Domäne C gelten.

Nachdem ich diesen Sachverhalt mit den mir bekannten bzw. zur Verfügung stehenden Mitteln bearbeitet hatte, konnte ich nur unsichere Vorstellungen von einer Bindungsregion für „coiled coil“-Strukturen präsentieren. Ich gelangte immer mehr zu der Überzeugung, dass es zu gewissen konformativen Veränderungen während des Komplexbildungsprozesses kommen musste. Das

„backbone“ der „coiled coil“ betrachtete ich dabei als unveränderlich, da es in vivo zu höheren, sich gegenseitig stabilisierenden Aggregaten kommt (siehe Abschnitt 1.2.2). Während der Betrachtung im Stereomodus schätzte ich mögliche Konformationsänderungen ab. Ich wechselte dabei zwischen verschiedenen Ansichten der „backbone“-Darstellungen, mit und ohne Seitenketten, und zwischen verschiedenen schematischen Darstellungen der Sekundärstrukturelemente. Nach dieser Prozedur konnte ich eine mögliche Wechselwirkungsgeometrie ausfindig machen. Deren genaue Analyse ergab, dass sie alle Kriterien und alle Randbedingungen erfüllt. Das Programm Hex gab für diesen Komplex 440 kJ/mol an. Nachdem dieser mit „Steepest Descent“ im Programm Sybyl minimiert wurde, folgte ein erneuter gezielter Dockingversuch mit Hex. Dieser Kreislauf wurde mehrmals wiederholt, wobei die Wechselwirkungsenergien auf 660 kJ/mol anstiegen. Es kam dabei zu geringen Veränderungen der Seitenkettenkonformation, wodurch beide Komplexpartner stärker angenähert werden konnten. Es muss an dieser Stelle angemerkt werden, dass dieses Ergebnis mit vielen verschiedenen „coiled coil“-Struktur-Abschnitten möglich ist. Auf Seiten der Domäne B ist dieser Bereich jedoch insofern außergewöhnlich, als er es erlaubt, durch geringe konformative Veränderungen eine „coiled coil“-Struktur aufzunehmen. Die „coiled coil“ kommt dabei in einem V-förmigen Tal zu liegen (siehe Abbildung 4.9). In dem Bereich, wo die beiden Helices der „coiled coil“ senkrecht übereinander stehen (relativ zur Oberfläche der Domäne B), kann die untere etwas in den V-förmigen Innenbereich der Domäne B eindringen. In dem Bereich horizontaler Konstellation beider „coiled coil“-Helices kommen diese auf einem ebenfalls waagrecht angeordneten „loop“ der Domäne B zu liegen.

- eine Wechselwirkungsgeometrie, die alle Kriterien und Randbedingungen erfüllt

Der Oberflächenbereich der Domäne B, welcher durch die Komplexbildung von Wassermolekülen nicht mehr frei zugänglich ist, kann als Komplexinterface der Domäne B definiert bzw. dargestellt werden (siehe Abbildung 4.10).

Die Fläche des Interfaces beträgt etwa 670 \AA^2 (zum Vergleich: die Dimere der Domänen B und C haben nicht mehr als 400 \AA^2 ; siehe Abschnitt 4.2.1) und eine Länge von etwa 45 \AA . Das Interface der Domäne B wird durch die folgenden Aminosäurereste definiert: R2284, P2285, G2286, T2287, L2289, E2290, E2293, E2327, K2328, S2331, R2334, Y2339, N2340, P2342, E2343, G2345, E2357, L2358, I2359, E2360, G2362, H2363, R2366, D2410, D2411, K2413, G2414, F2416, E2421, E2422, N2423

Diese Hypothese der Keratin-Bindungsregion konnte also ausreichend begründet werden. Eine experimentelle Überprüfung war deshalb angezeigt. Im folgenden Abschnitt werden die aus dieser Hypothese abgeleiteten Mutationsvorschläge näher erörtert.

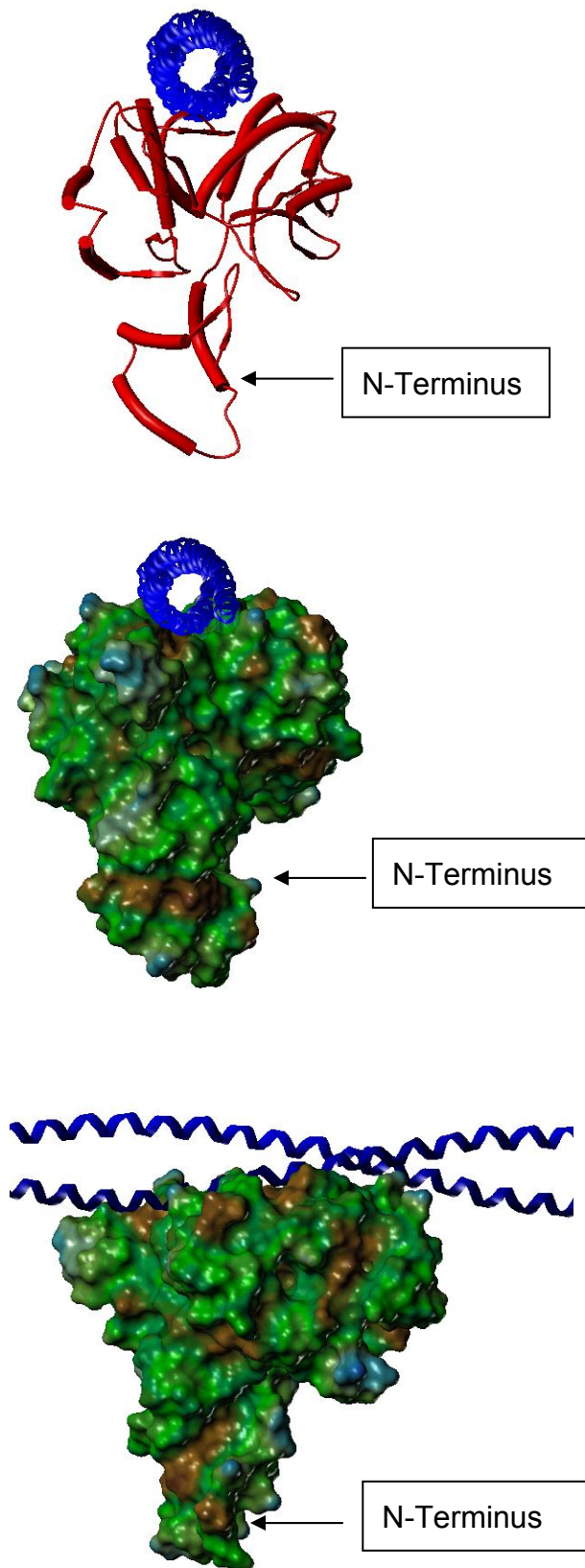


Abbildung 4.9: "coiled coil"-Domäne B-Wechselwirkungsgeometrie
 oben: Schematische Darstellung der Komplexgeometrie aus „coiled coil“-Struktur (blaue Helices) und Domäne B (rote, Helix als Zylinder, Pfeile als Faltblätter)
 mitte: Gleiche Perspektive wie oben, die Domäne B als Oberfläche mit lipophilem Potential - siehe Abbildung 4.6)
 unten: Seitenansicht der mittleren Perspektive.

Die „coiled coil“ kommt in einem V-förmigen Tal der Domäne B zu liegen. An der Stelle, wo die beiden Helices der „coiled coil“ senkrecht übereinander stehen (relativ zur Oberfläche der Domäne B), kann die untere weit in den V-förmigen Innenbereich der Domäne B eindringen (unten). In dem Bereich horizontaler Konstellation beider „coiled coil“-Helices kommen diese auf einem ebenfalls waagrecht angeordneten „loop“ der Domäne B zu liegen (siehe oben)

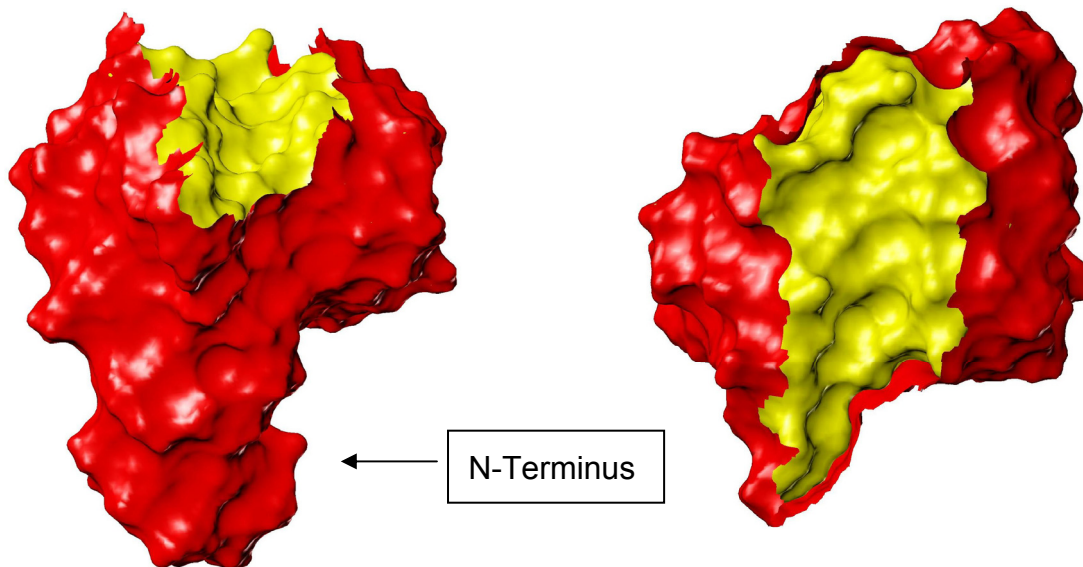


Abbildung 4.10: *Interface der Domäne B*
 Darstellung der Oberfläche des Interfaces (gelb) der Domäne B (rot und gelb),
 welches im Komplex durch die „coiled coil“ bedeckt wird.
 links: Gleiche Perspektive wie Abbildung 4.9 oben und mitte.
 rechts: Aufsicht der linken Perspektive

4.2.5 Mutationsvorschläge Teil 1 - Lokalisierung der Bindungsregion

Zusammenfassend stellt sich folgendes Bild dar: Die Domäne B besitzt nach theoretischer Betrachtung eine Region, die befähigt scheint, mit „coiled coil“-Strukturen einen Protein-Protein Komplex einzugehen. Unbekannt ist jedoch der genaue „coiled coil“-Abschnitt. Daraus sollen Mutationsvorschläge abgeleitet werden, um deren experimentelle Veränderung der Bindungseigenschaften als Indiz für die Richtigkeit der Hypothese der „coiled coil“-Binderegion der Domäne B zu sehen. Eine Veränderung im Sinne einer Verschlechterung erschien mir möglich. Auch wenn die genaue Bindestelle am Keratin nicht bekannt war, so sollte der gezielte Austausch von bestimmten Aminosäuren zu einer unspezifischen Abnahme der Bindungsaffinität führen. Dem gegenüber ist das Ableiten von Mutationsvorschlägen, die zu einer Verbesserung der Bindungseigenschaften führen, nur mit Kenntnis der genauen Wechselwirkungsgeometrie möglich. Folgende Randbedingungen habe ich für die Auswahl der Mutationsvorschläge eingeführt:

- Lokalisierung in der Binderegion

So können sie einen direkten, nachvollziehbaren Effekt ausüben.

- Lokalisierung möglichst in „loopbereichen“ an der Proteinoberfläche; weniger in lipophilen Clustern

Dieses soll die Wahrscheinlichkeit einer Fehlfaltung minimieren.

- 1-2 Mutationen pro Mutante, die nicht weiter als 3 Aminosäurereste voneinander entfernt sein dürfen

Das war eine Vorgabe des Kooperationspartners BASF, da man innerhalb dieses Abstandes die Mutationen in einem Mutationsschritt durchführen konnte.

- Die Mutationen sollten hinreichend starke Veränderung der molekularen Wechselwirkungsfelder im Proteininterface bewirken.

Dadurch soll eine messbare Veränderung der Bindungseigenschaften gewährleistet werden. Aus der Überlagerung von homologer Domäne B und C konnten Mutationen abgeleitet werden, die weniger wahrscheinlich zu einer Veränderung der Faltung führen.

Folgende Mutationsvorschläge wurden in Absprache mit dem Kooperationspartner BASF ausgewählt und experimentell umgesetzt.

F2416A

Dies sollte eine Abnahme der molekularen Wechselwirkungen bewirken.

L2358E

Diese Mutation führt zur Umwandlung eines lipophilen in ein elektrostatisch negativ geladenes molekulares Wechselwirkungsfeld. Hier besteht ein hohes Risiko einer Fehlfaltung, da Leucin 2358 Teil eines lipophilen Clusters ist. Sie wurde dennoch bewusst ausgewählt, da sie eine starke Veränderung darstellt (siehe nächsten Abschnitt. 4.2.6).

Y2339A

Dieses bewirkt eine Abnahme der molekularen Wechselwirkungen. Auch hier besteht ein hohes Risiko bei gleichzeitiger Aussicht auf einen messbaren Effekt,

da Tyrosin 2339 Teil des gleichen lipophilen Clusters wie Leucin 2358 ist (siehe nächsten Abschnitt. 4.2.6).

K2328A/S2331A

Diese Mutation bedeutet eine Veränderung des elektrostatisch positiv geladenen zu einem kleinen lipophilen molekularen Wechselwirkungsfeld. Zusätzlich wird eine mögliche Wasserstoffbrücke unterbunden.

S2331A/R2334A

Dies verändert das molekulare Wechselwirkungsfeld eines elektrostatisch positiv geladenen Bereiches zu einem kleinen lipophilen Areal. Zusätzlich wird eine mögliche Wasserstoffbrücke unterbunden.

KBD R2284D/G2286D

Diese Mutation bedeutet eine Veränderung eines elektrostatisch positiv geladenen zu einem elektrostatisch negativ geladenen molekularen Wechselwirkungsfeld. Die Domäne C besitzt an derselben homologen Position auch ein Aspartat. Zusätzlich wird mit G2286D ein elektrostatisch negativ geladenes molekulares Wechselwirkungsfeld bzw. eine sterische Barriere eingeführt.

4.2.6 Die Suche nach der Bindungsregion auf dem Keratin

Die Suche nach der Bindungsregion konzentrierte sich auf das Keratin K8/18. Denn an diesem Keratin ist eine Bindung nachgewiesen worden (siehe Abschnitt 1.4.1). In der Literatur konnten keine Einschränkungen für die Bindungsstelle auf der IF „rod“ gefunden werden. Ihre Lokalisation ist also in allen vier Segmenten und den „link“-Bereichen möglich. Letztere konnten allerdings nicht berücksichtigt werden, da entsprechende Strukturkoordinaten fehlen (siehe Abschnitt 1.2.2). Durch die (wenig verstandene) nicht kovalente Polymerisierung der IF Dimere zum funktionalen IF Netzwerk (siehe Abschnitt 1.2.2) kommen aber nur wenige Regionen der IF „rod“ als Bindungsstelle in Frage. Auf der anderen Seite sollte zumindest für eng verwandte IF diese Polymerisierung ähnlich strukturiert sein (persönliche Vermutung) und die Bindungsregion somit denselben homologen Sequenzbereich umfassen. Ich

konnte auf der Oberfläche IF „rod“ keine sich wiederholenden Regionen bezüglich der Verteilung der molekularen Wechselwirkungsfelder finden (in einer Größenordnung vergleichbar zum Interface der Domäne B). Außerdem handelt es sich um eine spezifische Wechselwirkung zwischen der Domäne B und dem Keratin. Aus diesen Gründen denke ich, dass jede IF „rod“ auch nur eine Bindestelle bereitstellt.

Aufgrund der engen Verwandtschaft von K8/18 zu den Haarkeratinen (alle Keratine sind eng verwandt - siehe Abschnitt 1.2.3) kann davon ausgegangen werden, dass die Bindungsregion für alle Keratine homolog ist. Über multiple Sequenzalignments konnte ich schnell den Bezug zu den Haarkeratinen, Hautkeratinen und zu Vimentin (IF Typ III - weniger eng verwandt mit Keratinen - siehe Abschnitt 1.2.3) herstellen, an welchem ebenfalls eine Bindung nachgewiesen wurde (siehe Abschnitt 1.4.1).

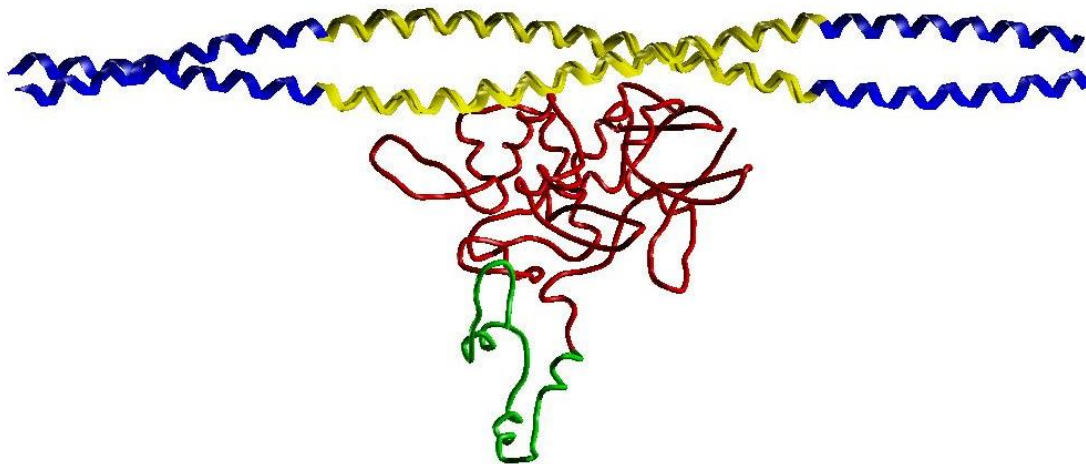
Das Interface der Domäne B wird vornehmlich von polaren bzw. geladenen Aminosäureresten besetzt. In diesem Zusammenhang stellt das Vorhandensein von zwei oberflächenexponierten, lipophilen Bereichen im Abstand von 27 Å eine Besonderheit dieses Interfaces dar.

Die verschiedenen Strukturmodelle nutzte ich, um mit den Programmen Sybyl und Grid die verschiedenen molekularen Wechselwirkungsfelder im Stereomodus sichtbar zu machen.

Folgende einander ergänzende Randbedingungen wurden eingeführt:

- elektrostatische Komplementarität
- möglichst wenig gleiche Ladungen auf engem Raum
- lipophile Komplementarität
- Phenylalanin, Tryptophan und Tyrosin sind nach Möglichkeit im Interface zu platzieren
- die Bindungsregion auf K8/18 muss auf andere Keratine übertragbar sein
- die Bindungsregion auf K8/18 kann bzw. sollte nach Möglichkeit auf Vimentin übertragbar sein

Aufgrund der Oberflächenform sind verschiedene „coiled coil“-Abschnitte praktisch nicht unterscheidbar (siehe Abschnitt 4.1.3 und Abbildung 4.6). Dieses erschwerte die Suche. Ein spezieller Keratinbereich, der eine auffällige lipophile Komplementarität zu den oben erwähnten lipophilen Bereichen der Domäne B besitzt, wurde ausgewählt, um die Mutation F2416A (siehe Abschnitt 4.2.5) in einer vergleichenden Molekül-Dynamiksimulation nachzuvollziehen. Dazu wurde für das Wildtypmodell zunächst die entsprechende Komplexgeometrie aus Domäne B und des lipophil-komplementären Keratinbereichs mit den Programmen Sybyl und Hex erzeugt. Dabei wurden die lipophil-komplementären Bereiche gegenseitig zur Deckung gebracht. Der Austausch von F2416A in dieser Geometrie ergab das Mutationsmodell. Das zeitliche Verhalten beider Modelle wurde über vier Nanosekunden betrachtet. Der für die Bindung nicht notwendige N-Terminus der Domäne B (siehe Abschnitt 4.2.4) wurde zur Verkleinerung des Systems und somit zur Verkürzung der Rechenzeit entfernt (Beginn bei I2250). Der ausgewählte Keratinabschnitt, lokalisiert im „rod“-Segment 1B, umfasst die Aminosäurereste M168 (d-Position im „heptad repeat“) bis S212 (f-Position im „heptad repeat“) vom Keratin 8 und A157 (d-Position im „heptad repeat“) bis T201 (f-Position im „heptad repeat“) vom Keratin 18 (siehe Abbildung 4.11).



*Abbildung 4.11: Der ausgewählte Keratinbereich und die Domäne B
Dargestellt ist das IF Segment 1B (gelbes und blaues „backbone“) und die Domäne B (rotes und grünes „backbone“) im richtigen Größenverhältnis zueinander.
Gekennzeichnet sind der nicht berücksichtigte N-Terminus der Domäne B (grün) und der ausgewählte Keratinbereich (gelb).*

In dieser Wechselwirkungsgeometrie ist der Hauptwechselwirkungspartner der Domäne B das Keratin 18. Die lipophile Komplementarität des Keratins ist besonders an zwei (in die Mutationsstudien involvierte) Stellen zur Domäne B gegeben (siehe Abschnitt 4.2.5). Im lipophilen Cluster 1 befinden sich in unmittelbarer Umgebung von F2416 (Domäne B) drei lipophile Aminosäurereste des Keratins 8/18 (siehe Abbildung 4.12): V188 (Keratin 18; g-Position im „heptad repeat“) hat einen Abstand (kürzester C-C Abstand) von 3,2 Å vor und 3,8 Å nach der Dynamik. L185 (Keratin 18; d-Position im „heptad repeat“) hat einen Abstand (kürzester C-C Abstand) von 5,3 Å vor und 4,2 Å nach der Dynamik. V200 (Keratin 8; a-Position im „heptad repeat“) hat einen Abstand (kürzester C-C Abstand) von 4,1 Å vor und 4,3 Å nach der Dynamik. Dieses lipophile Cluster ist somit während der Dynamik stabil geblieben.

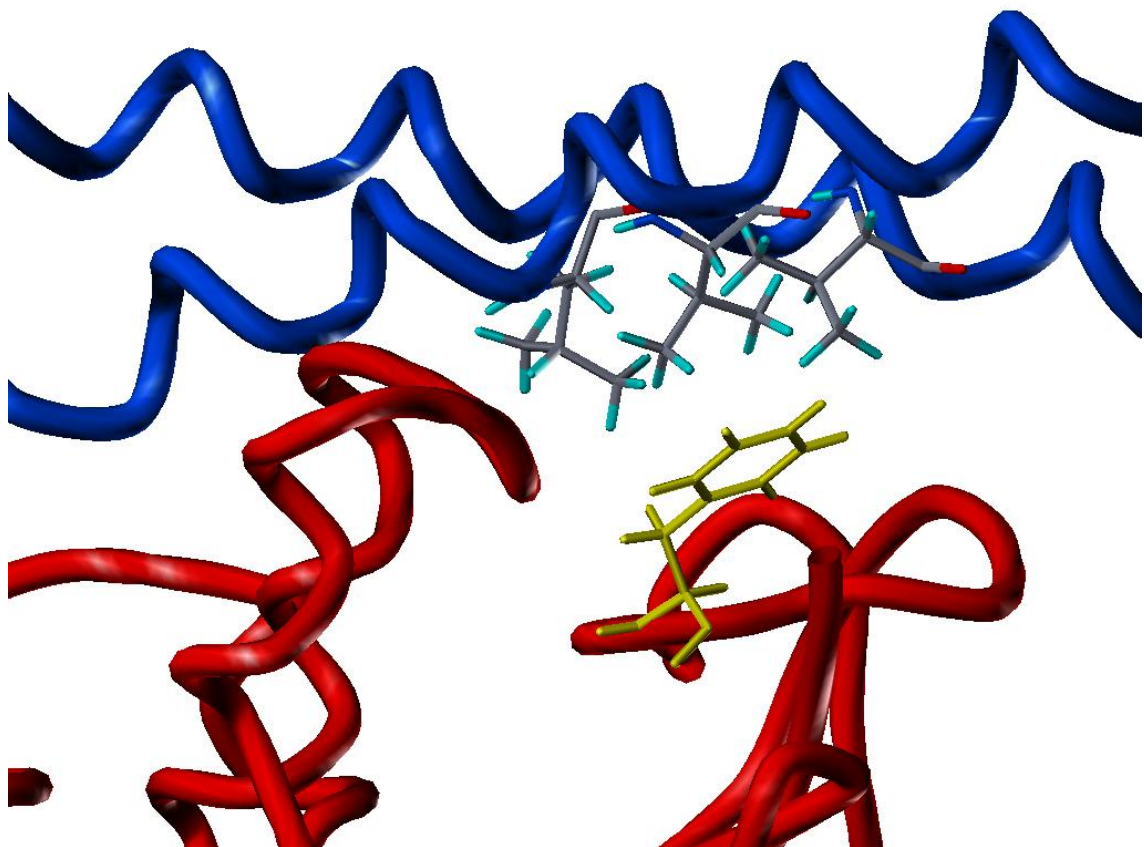


Abbildung 4.12: Lipophiles Cluster 1
Dargestellt ist das lipophile Cluster 1 zwischen Keratin 8/18 (blaues „backbone“) und der Domäne B (rotes „backbone“). Die drei oberen Aminosäurereste (L185, V188 vom K18 und V200 vom K8 - in Atomtyp-Färbung) können in Wechselwirkung mit dem unteren Phenylalaninrest (gelb) der Domäne B treten. Aus Gründen der Übersichtlichkeit wurde auf eine genaue Bezeichnung verzichtet.

Im Abstand von etwa 27 Å davon entfernt, befindet sich die zweite lipophil-komplementäre Stelle (lipophiles Cluster 2) in der nahen Umgebung der drei Aminosäuren L2358, I2359 und Y2339 (Domäne B). In dieser Region stehen die beiden Helices der „coiled coil“ senkrecht übereinander (siehe Abbildung 4.9). Die untere Helix (Keratin 18) kann in dieser Geometrie über drei Aminosäuren mit dem Lipophilen Cluster der Domäne B (aus P2342, L2358, I2359 und Y2339) wechselwirken (siehe Abbildung 4.13): V166 (f-Position im „heptad repeat“), T170 (c-Position im „heptad repeat“) und M174 (g-Position im „heptad repeat“). V166 hat einen Abstand (kürzester C-C Abstand) zum L2358 von 4,8 Å bzw. zum P2342 von 5,1 Å vor und 4,0 Å bzw. 4,1 Å nach der Dynamik.

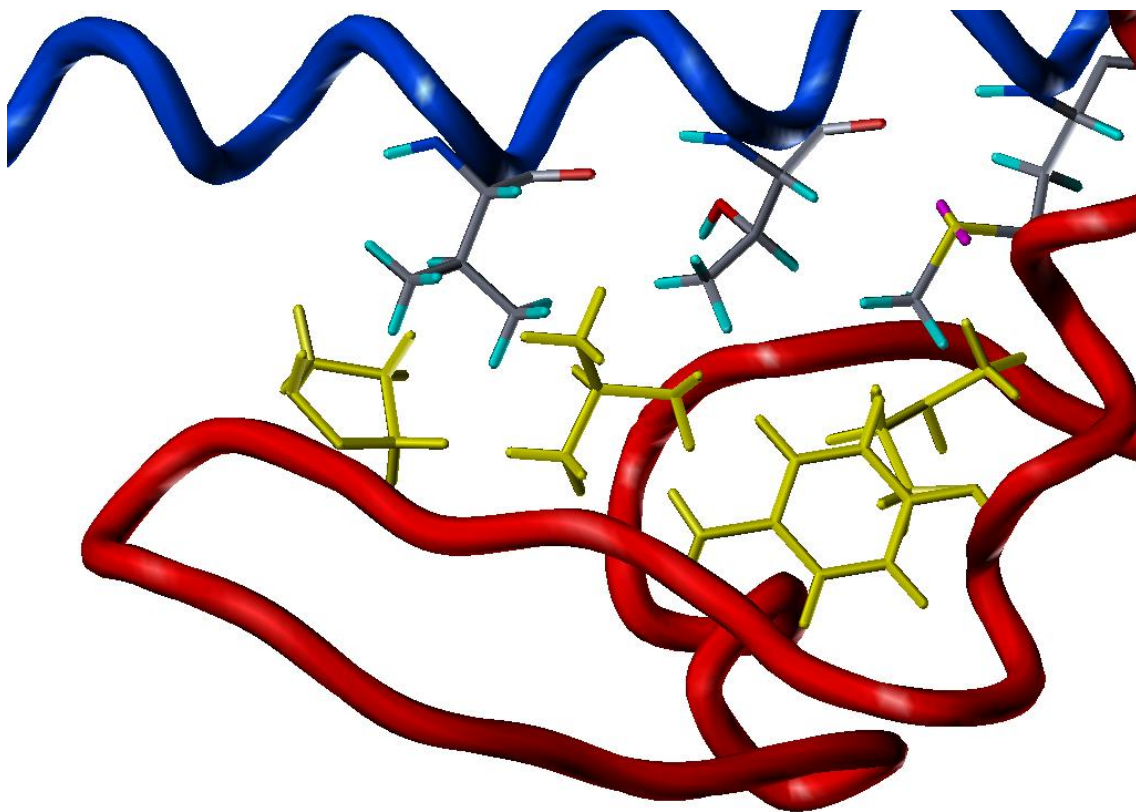


Abbildung 4.13: *Lipophiles Cluster 2*
Dargestellt ist das lipophile Cluster 2 zwischen Keratin 18 (blaues „backbone“) und der Domäne B (rotes „backbone“). Die drei oberen Aminosäurereste (V166, T170 und M174 in Atomtyp-Färbung) vom Keratin 18 können in Wechselwirkung mit den vier unteren Aminosäureresten (P2342, L2358, I2359 und Y2339 - gelbe Seitenketten) der Domäne B treten.

T170 hat einen Abstand (kürzester C-C Abstand) zum Y2339 von 3,7 Å vor und 3,7 Å nach der Dynamik. M174 hat einen Abstand (kürzester C-C Abstand) zum I2359 von 4,3 Å vor und 4,2 Å nach der Dynamik. Somit ist auch diese in wässriger Umgebung besonders günstige Konstellation, stabil geblieben.

- zwei lipophile Cluster im Abstand von 27 Å

Die folgenden konservierten Keratin-Aminosäurereste sind ebenfalls als besonders gute Wechselwirkungspartner in dieser Wechselwirkungsgeometrie hervorzuheben: F193 (Keratin 8; a-Position im „heptad repeat“) füllt an dieser Stelle als großvolumiger Rest den Raum zur Domäne B aus. D181 (Keratin 18; g-Position im „heptad repeat“) kann eine Salzbrücke zu R2366 (hoch konservierte Aminosäure auf der Domäne B und C) ausbilden. D201 (Keratin 8; b-Position im „heptad repeat“) kann eine Salzbrücke zu K2413 (ebenfalls hoch konservierte Aminosäure auf der Domäne B und C) ausbilden. Y204 (Keratin 8; e-Position im „heptad repeat“) hat guten Kontakt zur Domäne B. R186 (Keratin 18; e-Position im „heptad repeat“) kann mit E2360 (Domäne B) elektrostatisch wechselwirken. Eine echte Salzbrücke ist jedoch geometrisch nicht möglich. Die räumliche Nähe von E2360 und E190 (Keratin 8; e-Position im „heptad repeat“) kann durch R186 gut kompensiert werden. D191 (Keratin 18; c-Position im „heptad repeat“) kann eine Salzbrücke zu R2284 (Domäne B) ausbilden. E179 (Keratin 18; e-Position im „heptad repeat“) kann ebenfalls eine Salzbrücke zu K2328 (Domäne B) ausbilden. Eine weitere Salzbrücke kann zwischen K207 (Keratin 8; a-Position im „heptad repeat“) und E2421 (Domäne B) ausgebildet werden. Über diese speziellen Protein-Protein-Kontakte hinaus besteht noch eine Vielzahl weiterer Wechselwirkungsmöglichkeiten. Aufgrund der sechs elektrostatischen Wechselwirkungen zwischen Keratin und der Domäne B, besitzt diese Wechselwirkungsgeometrie auch elektrostatische Komplementarität (siehe Abbildung 4.14). Wichtig in diesem Zusammenhang ist auch, dass eine energetisch ungünstige Konstellation von zwei gleichgeladenen Aminosäureresten in engem Kontakt durch diese Geometrie nicht erzwungen wird.

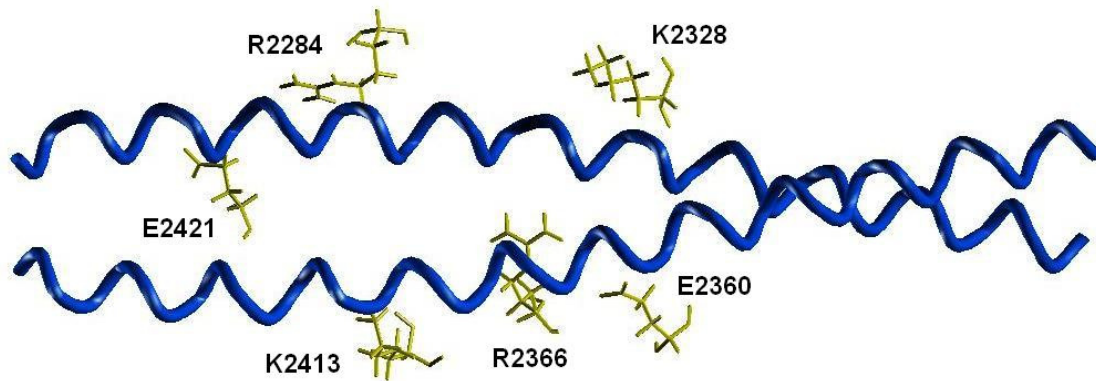


Abbildung 4.14: Verteilung der elektrostatischen Komplementarität
 Dargestellt ist die breite, räumliche Verteilung der elektrostatischen Komplementarität durch die jeweiligen Reste der Domäne B (gelb) in Bezug zur „coiled coil“ (blau).

Eine genauere Analyse der beiden vergleichenden Dynamiken ergab, dass die Mutante F2416A eine geringere Wechselwirkungsenergie über die Zeit produziert als der Wildtyp (siehe Abbildung 4.12). Zudem blieb die Wechselwirkungsgeometrie stabil. Dieses Ergebnis unterstützt die Vermutung, dass die Mutante F2416A verschlechterte Bindungseigenschaften aufweist.

Die gefundene Wechselwirkungsgeometrie bzw. die Bindungsregion auf dem Keratin soll im Weiteren mit „LEKG“ (von: lipophil und elektrostatisch komplementäre Geometrie) bezeichnet werden.

- „LEKG“ als mögliche Wechselwirkungsgeometrie

Eine Übertragung dieser Wechselwirkungsgeometrie auf die Keratine der Zielgebiete (siehe Abschnitt 1.5 und 1.6) ist gewährleistet (siehe Abbildung 4.16). Eine Übertragung auf Vimentin ist ebenfalls gut möglich, jedoch mit etwas geringerer Übereinstimmung. Zusammenfassend betrachtet, erfüllt die „LEKG“ alle Randbedingungen.

Die Ergebnisse der durchgeführten Mutationsstudien werden im folgenden Abschnitt dargestellt.

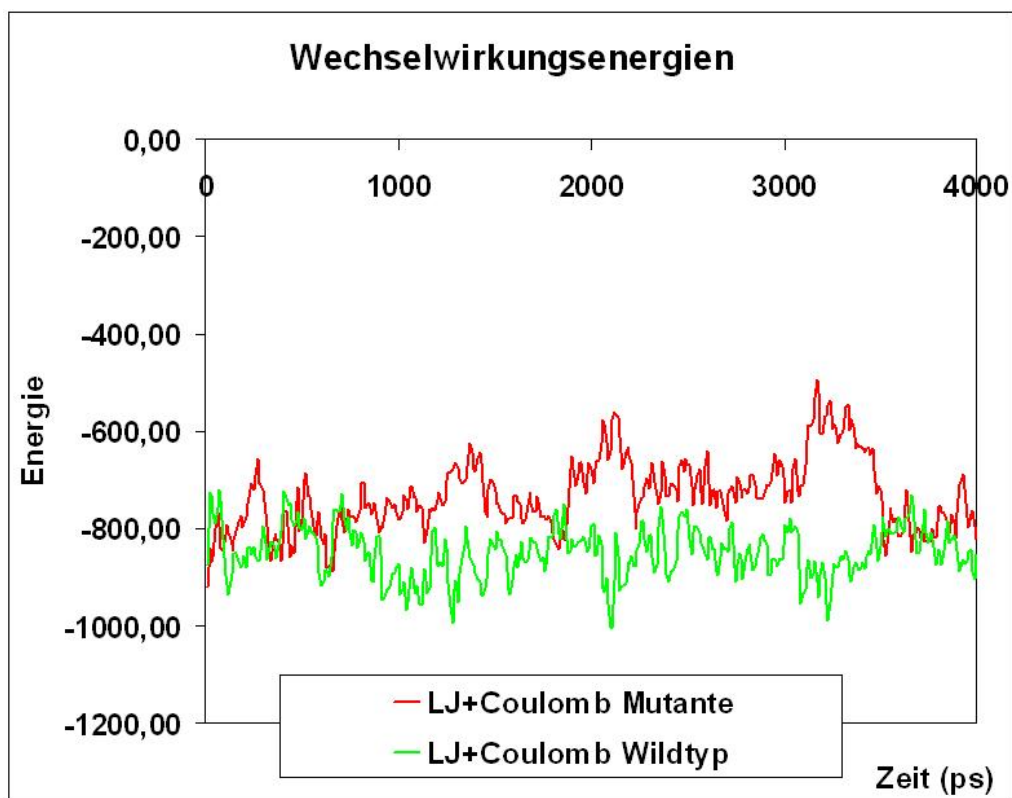
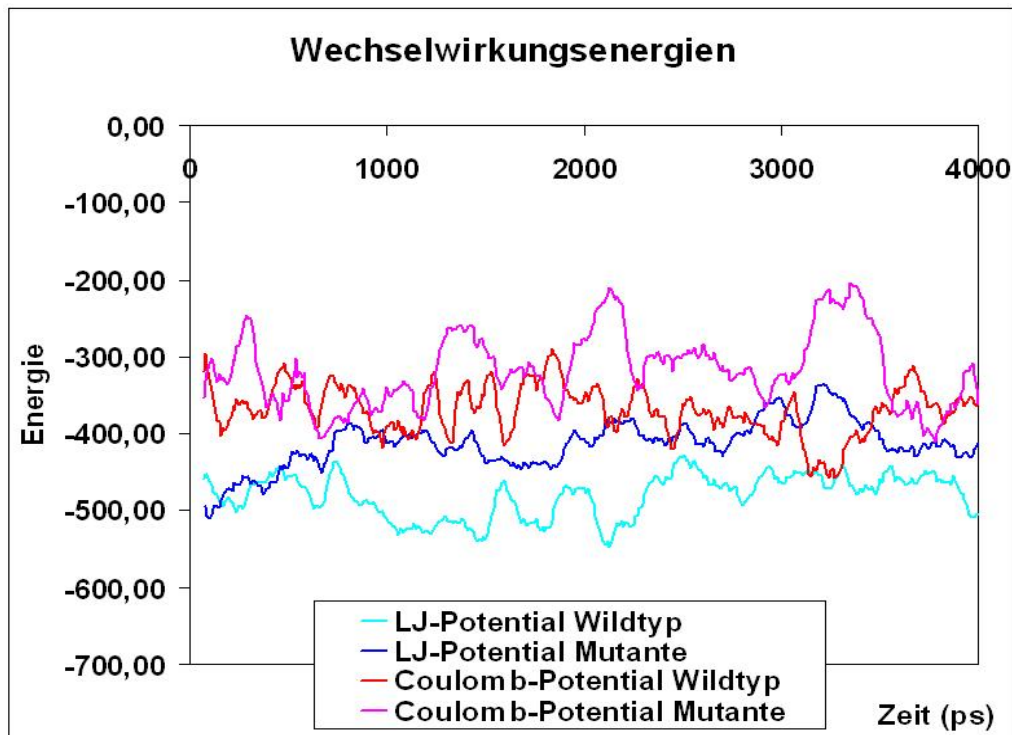


Abbildung 4.15: Oben: Moleküldynamik-Simulation Domäne B/MutanteF2416A - K8/18 Dargestellt sind die Wechselwirkungsenergien aus Coulomb-Potential und Lennard-Jones-Potential (LJ-Potential). Die Mutante produziert eine geringere Wechselwirkungsenergie über die Zeit, als der Wildtyp. Unten: Summe der beiden Potentiale für Mutante und Wildtyp.

4.2.7 Experimentelle Ergebnisse der Mutations-/Bindungsstudien des Kooperationspartners BASF

Der Kooperationspartner BASF führte den Austausch der unter Abschnitt 4.2.5 erläuterten Mutationsvorschläge für die Domäne B durch. Anschließend wurde die Bindungsaktivität am Haarkeratin der jeweiligen Mutante bestimmt, wobei dem Wildtyp eine Referenzaktivität von 100 Prozent zugeordnet wurde (siehe Tab. 4.1). Diesem Experiment zufolge, besitzt die Mutante F2416A eine Bindungsaktivität von 37 Prozent. Die Bindungsaktivität der beiden Mutanten L2358E und Y2339A konnte nicht bestimmt werden. Sie wurden bereits im Expressionssystem proteolytisch verdaut. Die Doppelmutante K2328A/S2331A

Tabelle 4.1: Mutationsstudien und die Bestimmung der Bindungsaktivität am Haarkeratin

| Protein | Aktivität [%] |
|--------------------------|-----------------------|
| 1. Wildtyp Domäne B | 100 |
| 2. Mutante F2416A | 37 |
| 3. Mutante L2358E | proteolytischer Abbau |
| 4. Mutante Y2339A | proteolytischer Abbau |
| 5. Mutante K2328A/S2331A | 81 |
| 6. Mutante S2331A/R2334A | 42 |
| 7. Mutante R2284D/G2286D | 0 |

besitzt eine Aktivität von 81 Prozent und die Doppelmutante S2331A/R2334A besitzt eine von 42 Prozent. Bei der Doppelmutante R2284D/G2286D wurde keine Bindungsaktivität detektiert. Alle Mutanten, die einer Bestimmung der Bindungsaktivität unterzogen werden konnten (die nicht proteolytisch verdaut wurden), besitzen demzufolge verschlechterte Bindungseigenschaften. Die theoretische Voraussage ist somit bestätigt wurden. Betrachtet man die Lokalisation der Mutationen auf der Oberfläche der Domäne B, so kann eine gute Übereinstimmung mit der vorgeschlagenen Wechselwirkungsgeometrie gefunden werden (siehe Abbildung 4.17).

- Das Zwischenziel wurde erreicht (siehe Abschnitt 2.2).

Die proteolytische Verdauung der Mutanten L2358E und Y2339A deutet meiner Meinung nach auf eine Fehlfaltung hin. Denn diese Aminosäuren sind Teil eines lipophilen Clusters (siehe Abschnitt 4.2.6). Für diese Mutationen wurde ein hohes Fehlfaltungsrisiko vorhergesagt, da proteininterne, apolare Wechselwirkungen für eine korrekte Faltung sehr wichtig sind

Bei einer Wiederholung des Mutations- und Bindungsaktivitätstestes der Mutanten F2416A, R2284D/G2286D und S2331A/R2334A konnten die Ergebnisse für die ersten beiden bestätigt werden. Die genauen Daten der Messwerte sind mir nicht bekannt.

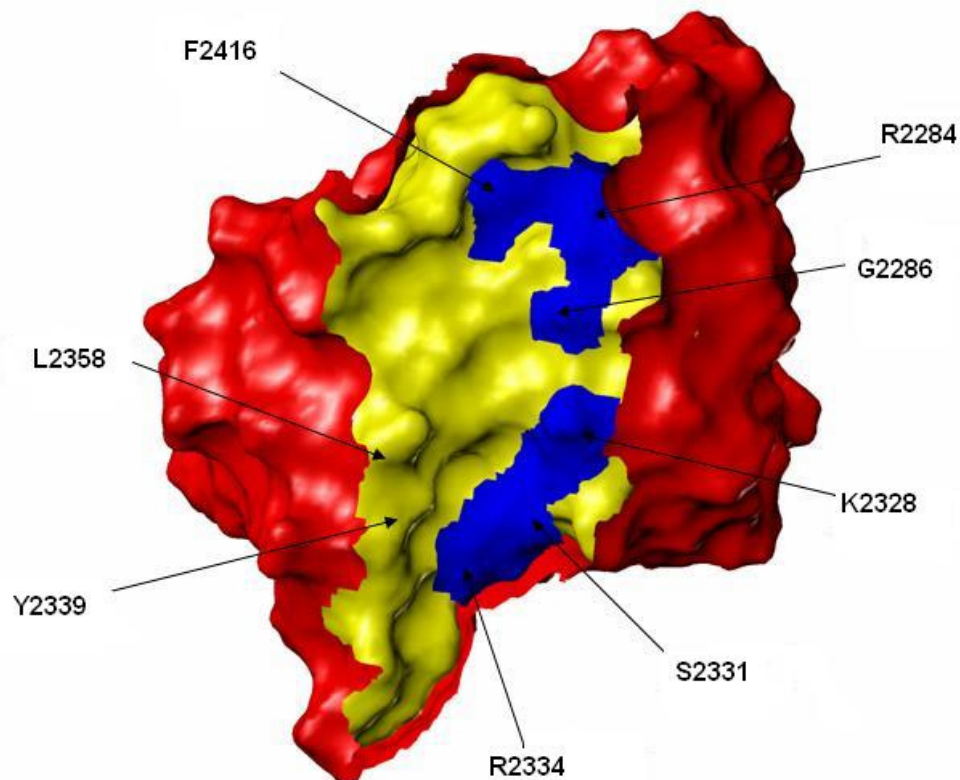


Abbildung 4.17: Lokalisation der mutierten Aminosäurereste auf der Domäne B
 Dargestellt ist die Oberfläche der Domäne B (rot, gelb und blau), das vorgeschlagene Interface (gelb und blau). Die Pfeile kennzeichnen den Ort der mutierten Aminosäuren. Die Oberflächen der mutierten Aminosäuren, die zu einer Abnahme der Bindungsaktivität geführt haben, sind hervorgehoben (blau). Die Mutationen von L2358 und Y2339 verursachten eine proteolytische Verdauung.

4.2.8 Nachvollziehbarkeit der Mutationsstudien durch die „LEKG“

Ich möchte im Folgenden die Frage erörtern, ob die bisherigen experimentellen Ergebnisse des Kooperationspartners BASF anhand der „LEKG“ nachvollziehbar sind (siehe Abschnitt 4.2.6 und 4.2.7).

Die Übereinstimmung der „LEKG“ in Bezug auf die Mutation F2416A und die Übertragbarkeit auf die betrachteten Keratine wurde bereits gezeigt (siehe Abschnitt 4.2.6).

Der vollständige Bindungs-/Aktivitätsverlust der Doppelmutante R2284D/G2286D kann ebenfalls sehr gut erklärt werden: Die zusätzliche Seitenkette in G2286D kann aus sterischen Gründen nicht in den Komplex eingepasst werden (siehe Abbildung 4.18). Wichtig in diesem Zusammenhang ist natürlich, dass sich an gleicher homologer Stelle der Zielkeratine ebenfalls kein sperriger Rest befindet. Alle in Frage kommenden Keratine besitzen an dieser Position ein konserviertes Glycin. Nur K40 und Vimentin besitzen ein Serin. Durch geringe Konformationsänderung der Domäne B erscheint ein Serinrest in dieser Konstellation möglich. Die nachgewiesene Bindung von Vimentin könnte also auch in dieser Geometrie erklärt werden (siehe Abschnitt 1.4.1).

Eine Verschlechterung der Bindungsaktivität der Mutante S2331A/R2334A erscheint ebenfalls sehr wahrscheinlich. Arginin2334 besitzt gute van der Waals-Kontakte zu Threonin170 und Methionin174 des Keratins 18. Außerdem befindet sich in 6Å Entfernung eine Carboxylatgruppe von Glutaminsäure169 (K18). Diese Mutation kann am besten auf das Haarkeratin K40 übertragen werden (siehe Abbildung 4.16).

Auch die letzte Mutante K2328A/S2331A kann aufgrund der möglichen Salzbrücke zwischen Lysin2328 und Glutaminsäure179 (absolut konserviert) verschlechterte Bindungseigenschaften aufweisen. Alle Mutationen sind mit der „LEKG“ nachvollziehbar und auf die anderen betrachteten Keratine übertragbar.

- Die Nachvollziehbarkeit und die Übertragbarkeit, auf andere Keratine, der Mutationsstudien durch die „LEKG“ sind gewährleistet.

Die „LEKG“ ist von allen betrachteten Wechselwirkungsgeometrien die einzige, die als Bindungsregion der Domäne B an Keratine wahrscheinlich ist. Sie zeichnet sich durch drei wesentliche Eigenschaften aus, die bei allen anderen analysierten Komplexen nicht gegeben waren: Die Geometrie der „coiled coil“ ist optimal zur Domäne B positioniert. Dadurch können beide Proteine über das gesamte Interface hinweg, bindende van der Waals-Kontakte ausbilden. Weiterhin besitzt die „LEKG“ eine sehr hohe Präzision bezüglich der Komplementarität in der Verteilung molekularer Wechselwirkungsfelder. Vor allem die besondere, lipophile Komplementarität im mittleren Abstand von 27 Å konnte in keiner anderen Wechselwirkungsgeometrie beobachtet werden. Zudem sind die experimentellen Daten nachvollziehbar.

Aus der „LEKG“ wurden Mutationsvorschläge abgeleitet, welche dem Hauptziel entsprechen würden (siehe Abschnitt 2.2). Diese werden im nächsten Abschnitt näher erläutert.

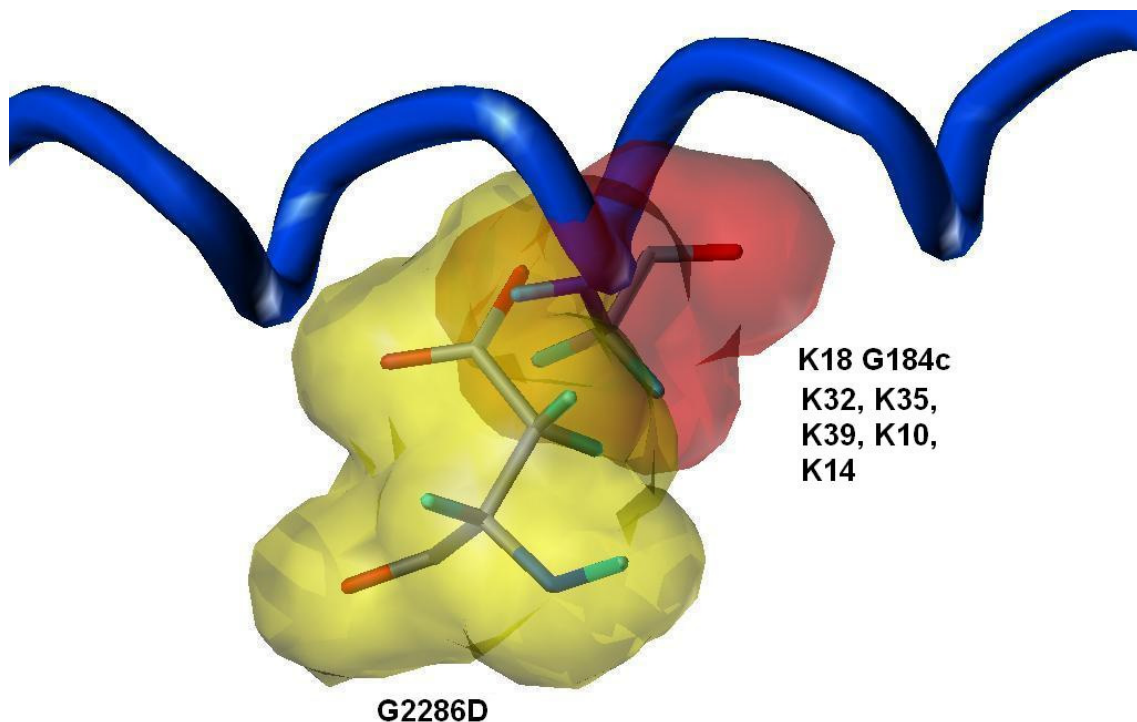


Abbildung 4.18: Die Mutation G2286D in Bezug zur „LEKG“
 Die Wechselwirkungsgeometrie „LEKG“ wäre mit der Mutante G2286D nicht möglich. Die van der Waals Radien von Asparaginsäure (gelb) und Glycin 184 vom Keratin 18 auf einer c-Position im „heptad repeat“ (rot) würden stark überlappen. Das ist aufgrund des Pauli-Prinzips nicht möglich. Außer K40 und Vimentin(Serin) besitzen alle anderen betrachteten Keratine an dieser homologen Position ebenfalls ein Glycin.

4.2.9 Mutationsvorschläge für verbesserte Bindungseigenschaften der Domäne B im Sinne des Hauptzieles

Mutation 1: D2407A bzw. D2407X

Der erste Mutationsvorschlag für verbesserte Bindungseigenschaften im Sinne des Hauptzieles dieser Arbeit (siehe Abschnitt 2.2), ist die Mutation D2407A bzw. D2407X. Eine ausführliche Begründung hierfür wurde bereits im Abschnitt 4.2.3 gegeben. „X“ steht für eine der restlichen Aminosäuren.

Die folgenden Mutationsvorschläge wurden aus der „LEGK“ abgeleitet (siehe Abschnitt 4.2.6). Durch Überlagerung der verschiedenen Keratin-Strukturmodelle konnten die Wechselwirkungen mit der Domäne B in der „LEKG“ studiert werden. Das Sequenzalignment der Keratine diente zur groben Orientierung und zur Bestimmung von Sequenzunterschieden (siehe Abbildung 4.16). Bei der Auswahl der Mutationen wurde ein Sequenzalignment der drei homologen Domänen A, B und C (siehe Abschnitt 1.3.2) (und weiterer verwandter Proteine) berücksichtigt, um einen Überblick über die erlaubten Aminosäurediversitäten an den jeweiligen homologen Sequenzpositionen zu erhalten. Mit dem Programm Sybyl (Tripos-Kraffteld) wurden die Unterschiede der totalen Energien von Mutante und Wildtyp (in der analogen Komplexgeometrie mit gleicher energieminimierter Konformation) berechnet, um eine grobe Aussage über die Veränderung der Wechselwirkungsenergie zu treffen. Dabei wurden van der Waals und elektrostatische Anteile getrennt betrachtet. Dadurch gelang eine bessere Bewertung von Mutationen, welche eine Veränderung von geladenen Aminosäuren herbeiführen. Innerhalb des Proteinkomplexes kann man eine verringerte relative Permittivität (im Vergleich zu einer wässrigen Umgebung mit gelösten Ionen) annehmen. Deshalb müssen hier die langreichenden elektrostatischen Kräfte besonders berücksichtigt werden. Eine Bewertung von Mutationen von geladenen Aminosäureresten, welche Kontakt zum wässrigen Milieu haben, erscheint mit dieser Berechnung jedoch nicht sinnvoll. Weiterhin muss bei diesen Kraffteldrechnungen beachtet werden, dass die überaus wichtige Veränderung (Minimierung) der

Grenzflächenenergie von apolaren Bereichen in Wasser (die im Proteinkomplex aneinander zu liegen kommen) nicht berücksichtigt wird.

(Das Keratin 8/18 dient im Folgenden als Referenz zur Angabe der jeweiligen homologen Sequenzposition. Die Abbildungen von Wildtyp und Mutante sind aus der gleichen Perspektive aufgenommen und deshalb zur besseren Orientierung im Vergleich zu betrachten. Die dargestellten Kohlenstoffatome der Domäne B sind schwarz, die des Keratins hellgrau.)

Mutation 2: E2293I

Der Aminosäurerest E2293 (Domäne B) kann vor allem mit S177 (K18) wechselwirken (siehe Abbildung 4.19 A). Bei den Hautkeratinen befindet sich ebenfalls ein Serin, bei allen Haarkeratinen hingegen ein Leucin, an dieser Stelle. Die Mutation E2293I führt also dazu, dass im Komplex mit Haarkeratin (im Gegensatz zum Hautkeratin) zwei größere apolare Bereiche zur Deckung gebracht (und die Grenzflächenenergie somit minimiert) werden (siehe Abbildung 4.19 B). Zudem ist die im Komplex beobachtete Wasserstoffbrücke vom Glutamat zum Serin nicht mehr möglich. Diese Mutation erscheint deshalb besonders geeignet zu sein, die Domäne B verstärkt auf Haarkeratine zu orientieren.

Diese Mutation bedingt am Haarkeratin eine Verbesserung von 2,7 kJ/mol von der Waals Energie und eine Verbesserung von 0,5 kJ/mol im elektrostatischen Anteil, was wohl auf die räumliche Nähe von E2293 beim Wildtyp zu einem Glutamat zurückzuführen ist.

Ein Risiko (Fehlfaltung) bei dieser Mutation könnte insofern bestehen, als es sich bei E2293 um eine hoch konservierte Aminosäure in einem ebenfalls hoch konservierten Umfeld (zwei Aminosäurereste vorher bis einer danach) handelt. Diese Mutation ist aber dennoch sehr interessant und viel versprechend, da man bei ihr besonders gut zwischen Haut- und Haarkeratinen differenzieren kann.

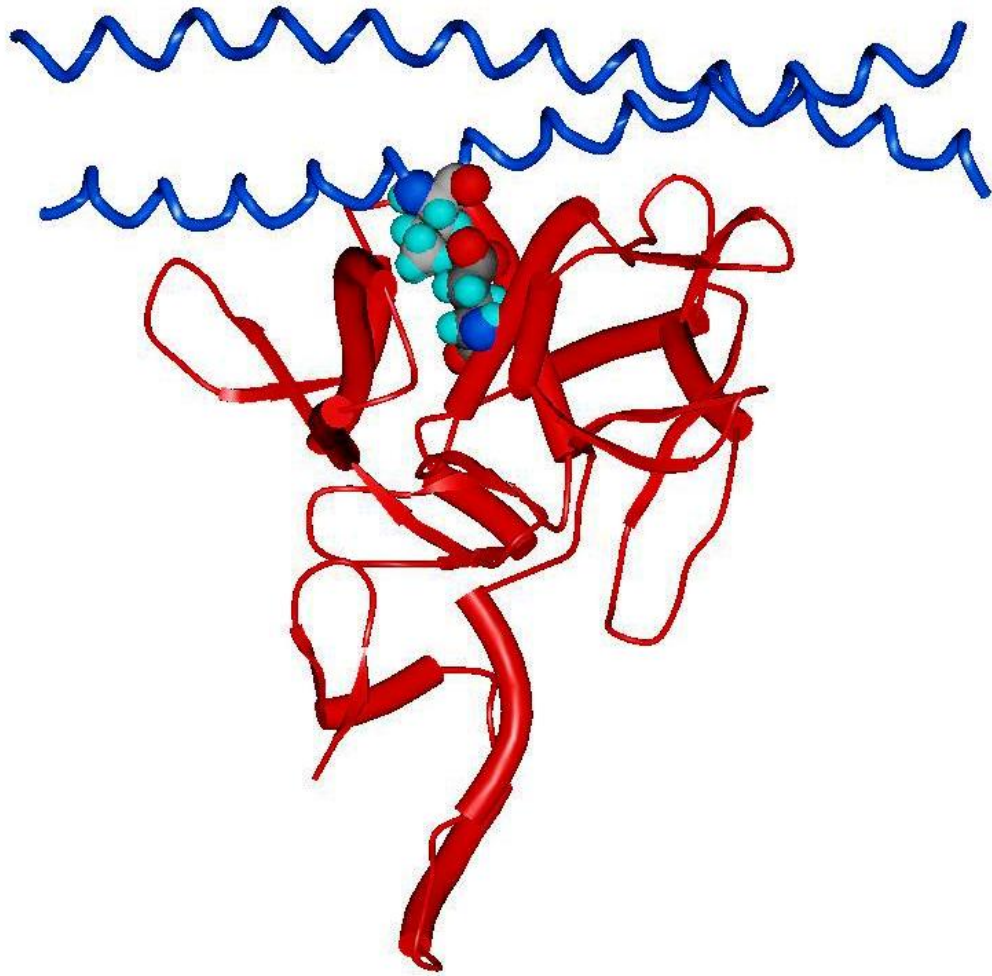


Abbildung 4.19 A: Mutation 2
Dargestellt ist der Proteinkomplex „LEGK“ aus Domäne B (rote Sekundärstruktur) und Haar-Keratin (blaue Helices). Beim Wildtyp befindet sich E2293 (Domäne B) in Nachbarschaft zu einem Leucin am Haarkeratin.

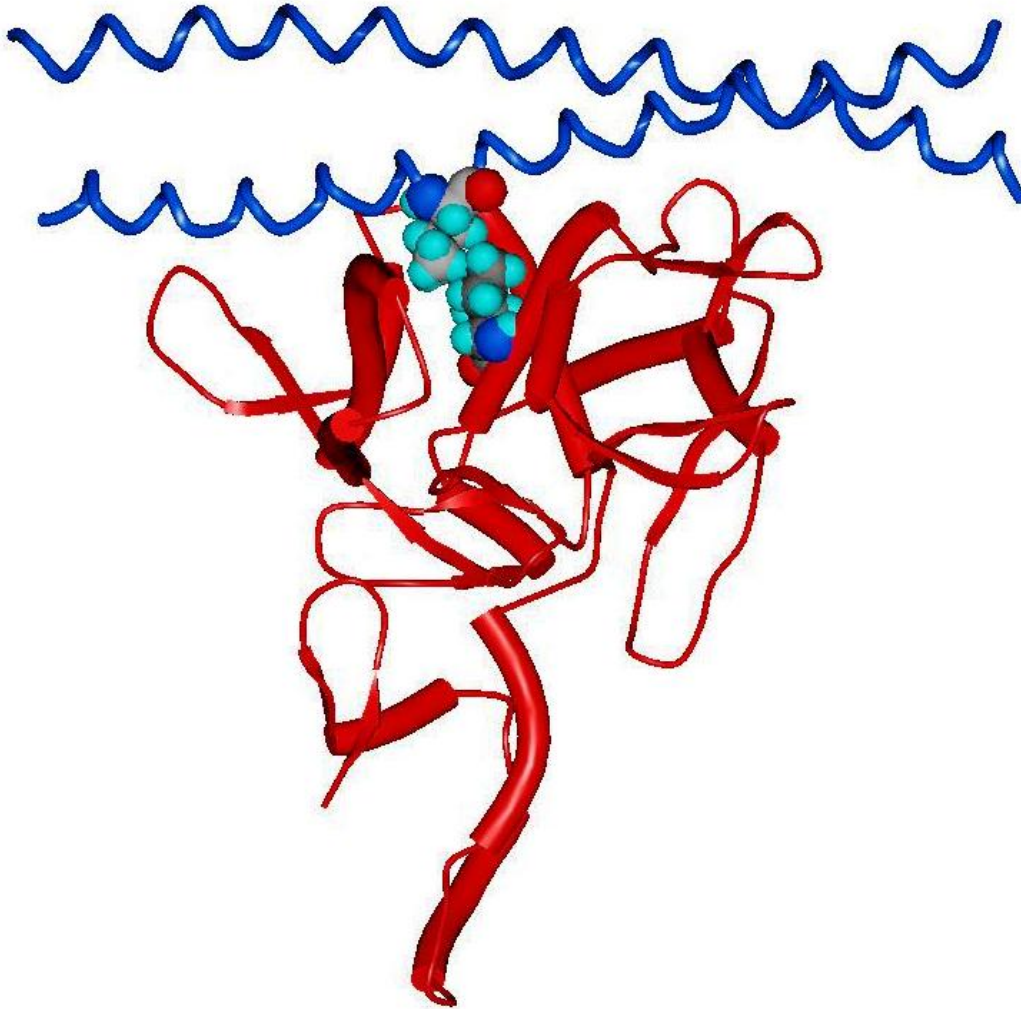


Abbildung 4.19:B: *Mutation 2*
Dargestellt ist der Proteinkomplex „LEGK“ aus Domäne B (rote Sekundärstruktur) und Keratin (blaue Helices). In der Mutante ist das E2293 durch ein I ersetzt.

Mutation 3 (zwei Varianten): E2422I oder E2422F

Der dem wässrigen Milieu ausgesetzte Aminosäurerest E2422 (Domäne B), befindet sich in nächster Nachbarschaft zu einem hoch konservierten Y204 am Keratin 8 (siehe Abbildung 4.20 A). In der Domäne C befindet sich an dieser Stelle ein Leucin. Ein Isoleucin an dieser Position könnte gut mit dem aromatischen Ringsystem in Deckung gebracht werden (siehe Abbildung

4.20 B). Diese Konstellation hat in wässriger Umgebung stabilisierende Eigenschaften. Die Einführung eines Phenylalaninrestes würde durch aromatische Wechselwirkungen diese Stabilisierung noch vergrößern. Dazu wären aber geringe konformative Veränderungen nötig, damit eine gute Wechselwirkungsgeometrie eingestellt werden kann. Beide Mutationsvarianten könnten also eine allgemeine Affinitätsverbesserung darstellen.

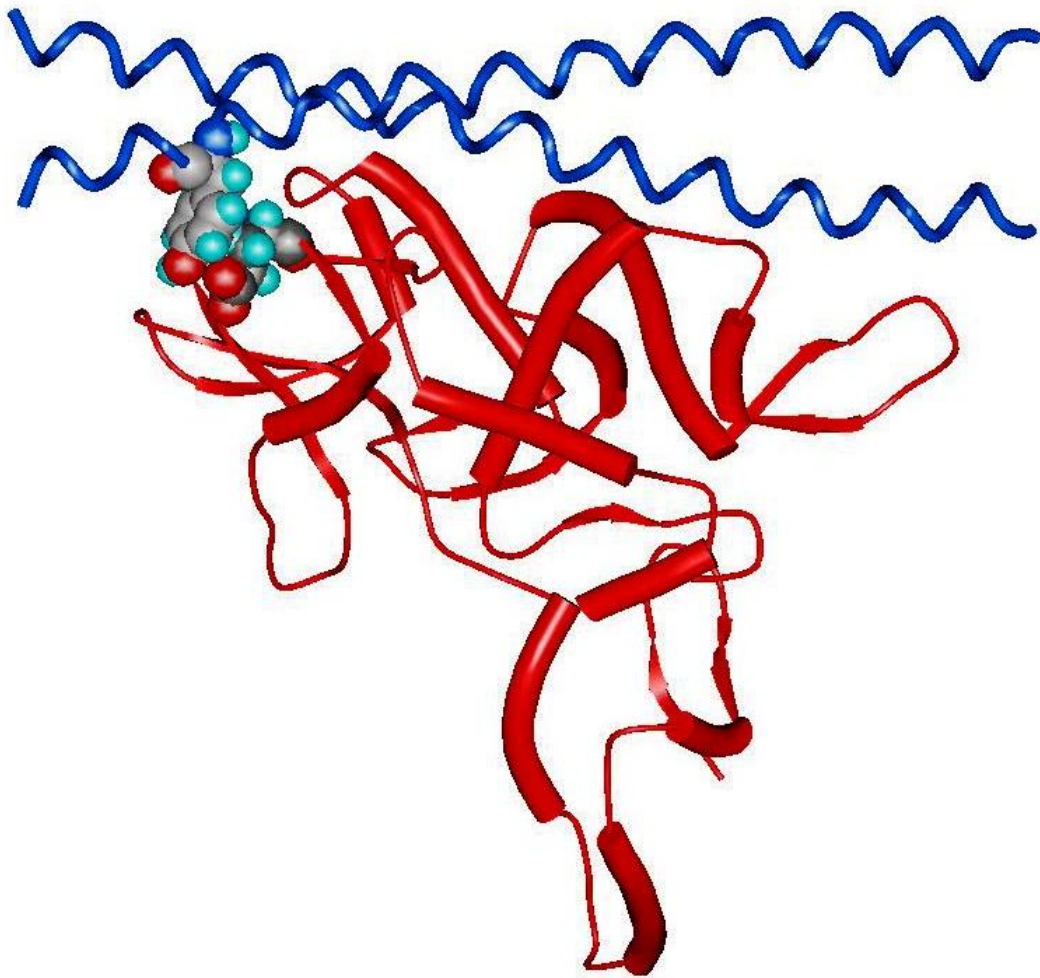


Abbildung 4.20 A: Mutation 3
Dargestellt ist der Proteinkomplex „LEGK“ aus Domäne B (rote Sekundärstruktur) und Keratin (blaue Helices). Beim Wildtyp befindet sich E2422 (Domäne B) in Nachbarschaft zu einem Y am Haarkeratin.

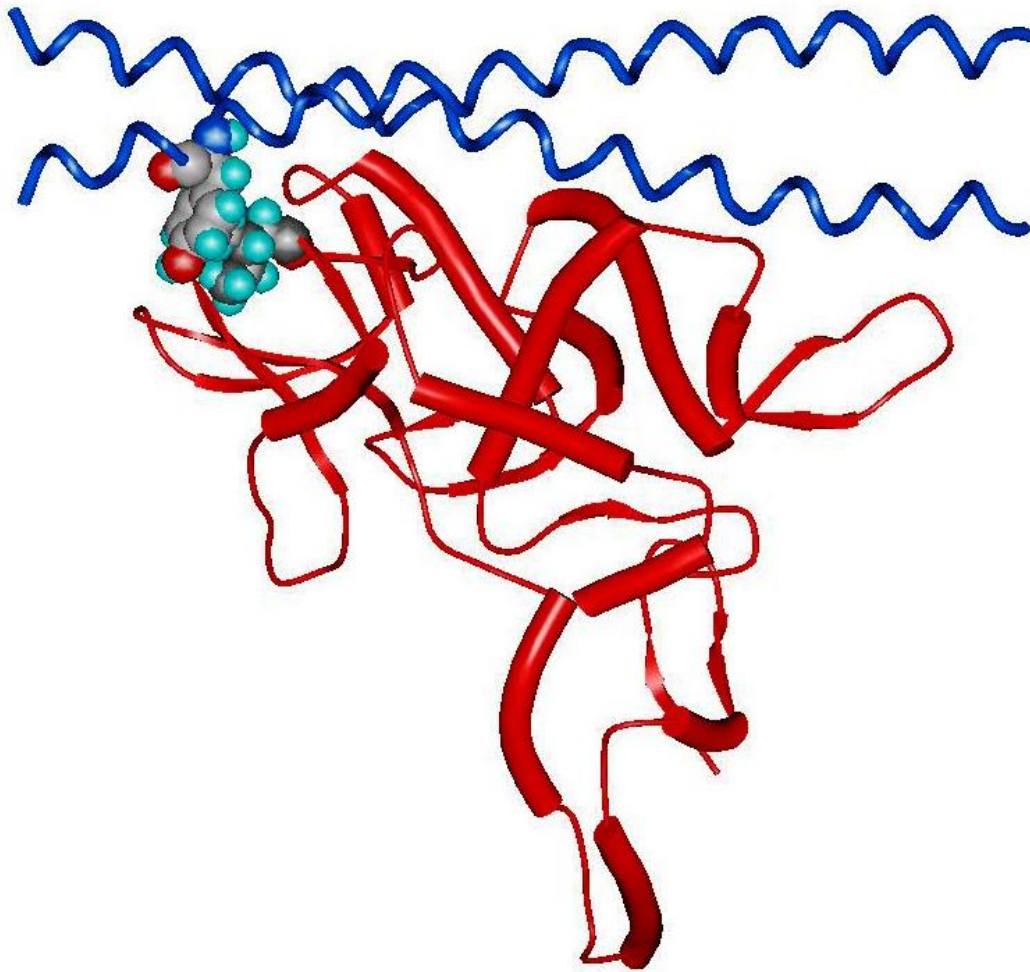


Abbildung 4.20 B: Mutation 3
Dargestellt ist der Proteinkomplex „LEGK“ aus Domäne B (rote Sekundärstruktur) und Keratin (blaue Helices). In der Mutante ist das E2422 durch ein I ersetzt.

Mutation 4 (drei Varianten): G2362M oder I oder Q

Der Aminosäurerest G2362 (Domäne B) befindet sich im Abstand von 5 Å zu einem konservierten Phenylalaninrest (Keratin) sowie in 6 Å Entfernung zu einem konservierten Valinrestes (Keratin) (siehe Abbildung 4.21 A). In der Domäne C befindet sich an dieser Stelle ein Glutamatrest (verwandte Proteine tragen dort auch ein Lysin, Arginin oder Glutamin). Durch die Einführung einer Seitenkette könnte somit Kontakt zum Keratin hergestellt werden. Lipophile

Reste (M oder I, wobei M besonders gute van der Waals Kontakte herstellen kann) könnten sich mit den zwei oben beschriebenen apolaren Keratinresten in Deckung bringen (siehe Abbildung 4.21 B). Dies könnte dann wieder ein Grund für eine erhöhte Stabilisierung in wässriger Umgebung sein. Für die Einführung von Glutamin spricht das geringe Risiko einer Fehlfaltung. Alle Mutationen würden jedoch ganz allgemein die Wechselwirkungsenergie erhöhen.

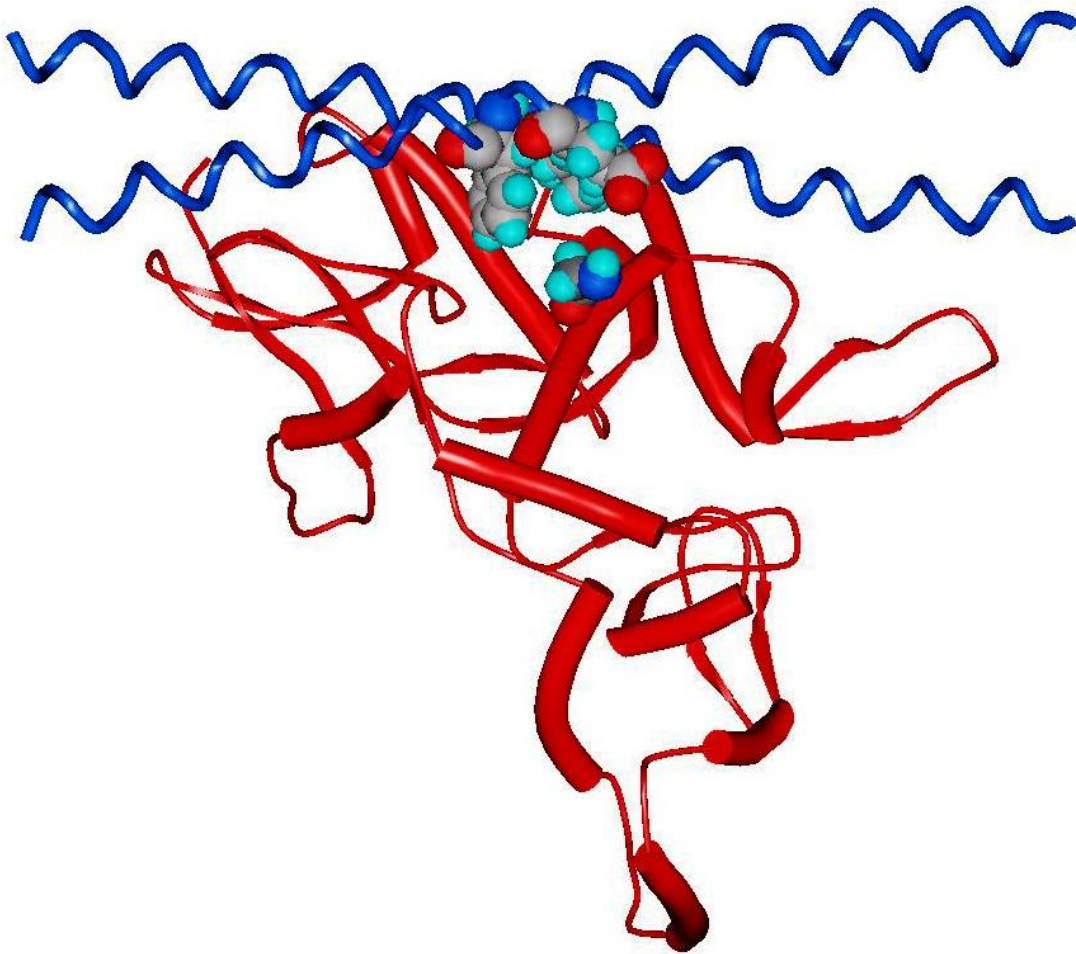


Abbildung 4.21 A: Mutation 4
Dargestellt ist der Proteinkomplex „LEGK“ aus Domäne B (rote Sekundärstruktur) und Keratin (blaue Helices). Beim Wildtyp befindet sich G2362 (Domäne B) in Nachbarschaft zu einem F, V und einem E am Haarkeratin.

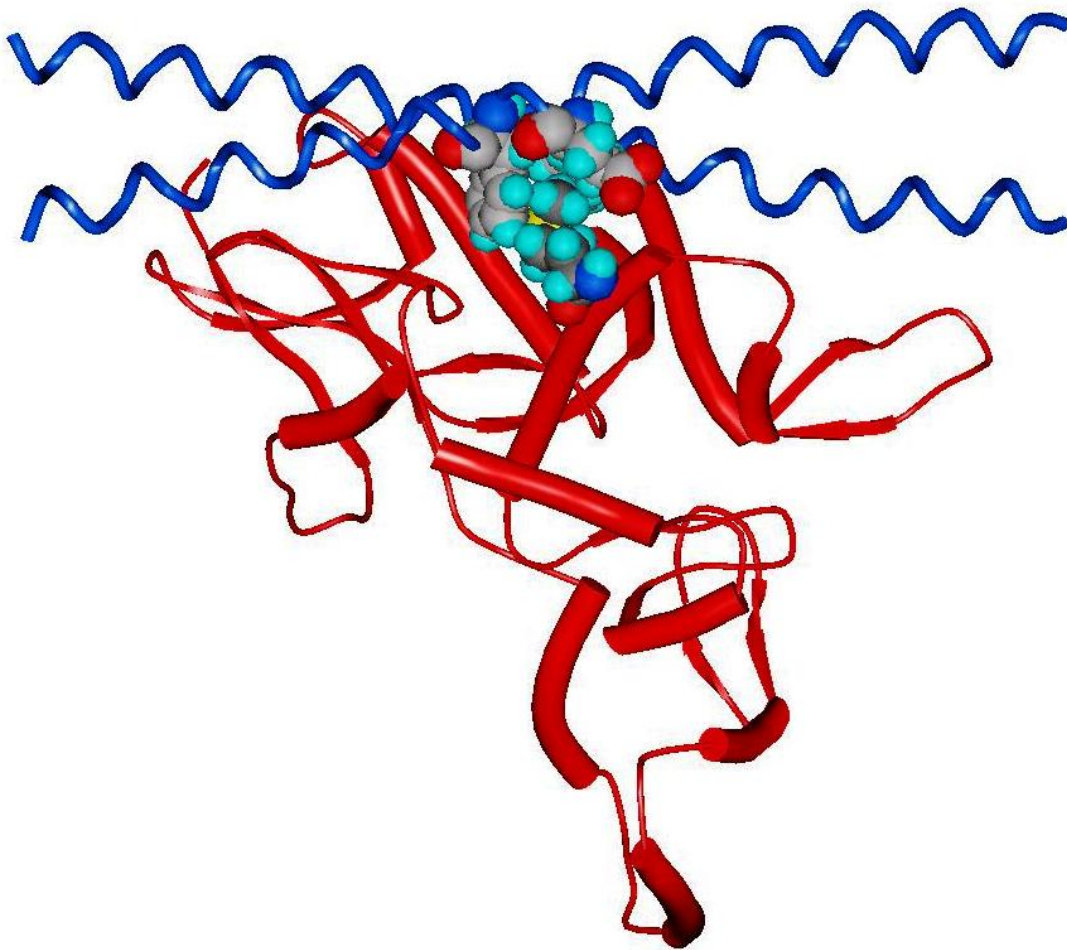


Abbildung 4.21 B: Mutation 4
Dargestellt ist der Proteinkomplex „LEGK“ aus Domäne B (rote Sekundärstruktur) und Keratin (blaue Helices). Durch die Mutation G2362M wird zusätzlicher van der Waals Kontakt vermittelt.

Mutation 5: H2363F

Der Aminosäurerest H2363 befindet sich in Nachbarschaft zum S177 (K18, bzw. ein L bei Haarkeratinen, siehe Mutation 2 - siehe Abbildung 4.22 A). Diese Position ist aber im Gegensatz zu E2293 nicht konserviert. Man findet an dieser Stelle Alanin-, Lysin-, Methionin-, Valin- und Serinreste. In der Wechselwirkungsgeometrie erscheint aber nur die Mutation zu Phenylalanin sinnvoll, da einerseits der Austausch durch andere Reste sterisch nicht möglich ist, andererseits verringert sich durch kleinere Reste auch die Wechselwirkungsenergie (siehe Abbildung 4.22 B).

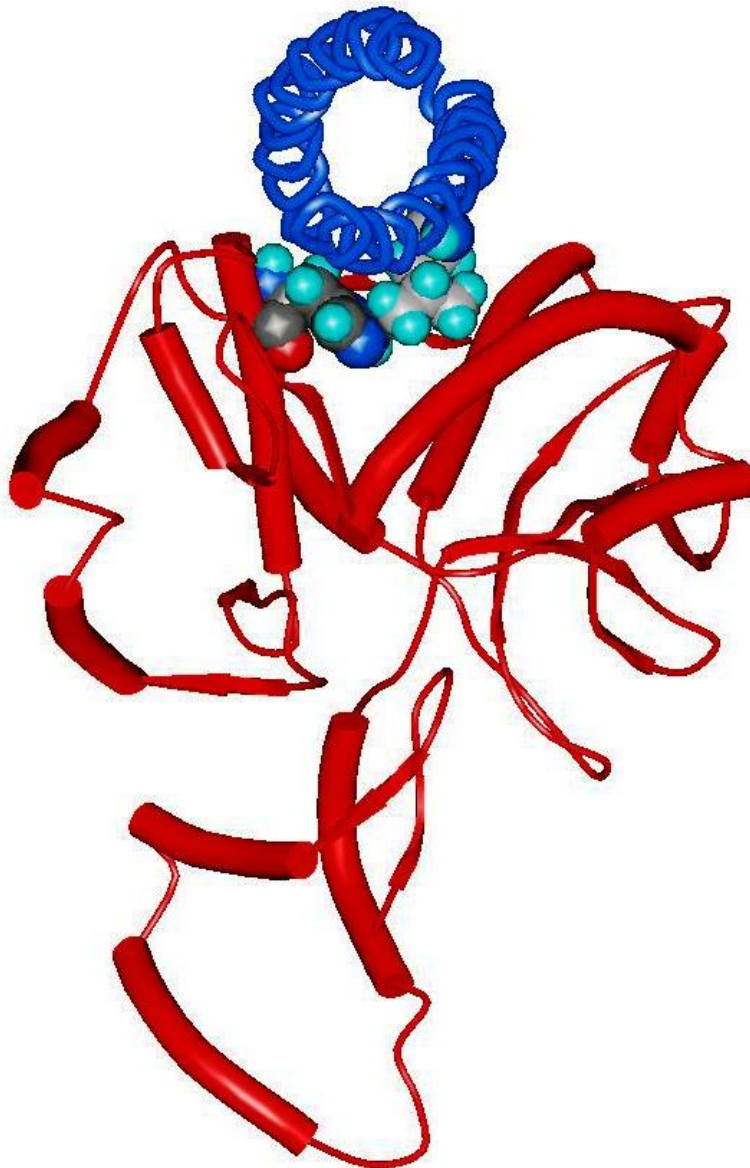


Abbildung 4.22 A: Mutation 5
Dargestellt ist der Proteinkomplex „LEGK“ aus Domäne B (rote Sekundärstruktur) und Keratin (blaue Helices). Beim Wildtyp befindet sich H2363 (Domäne B) in Nachbarschaft zu einem L am Haarkeratin.

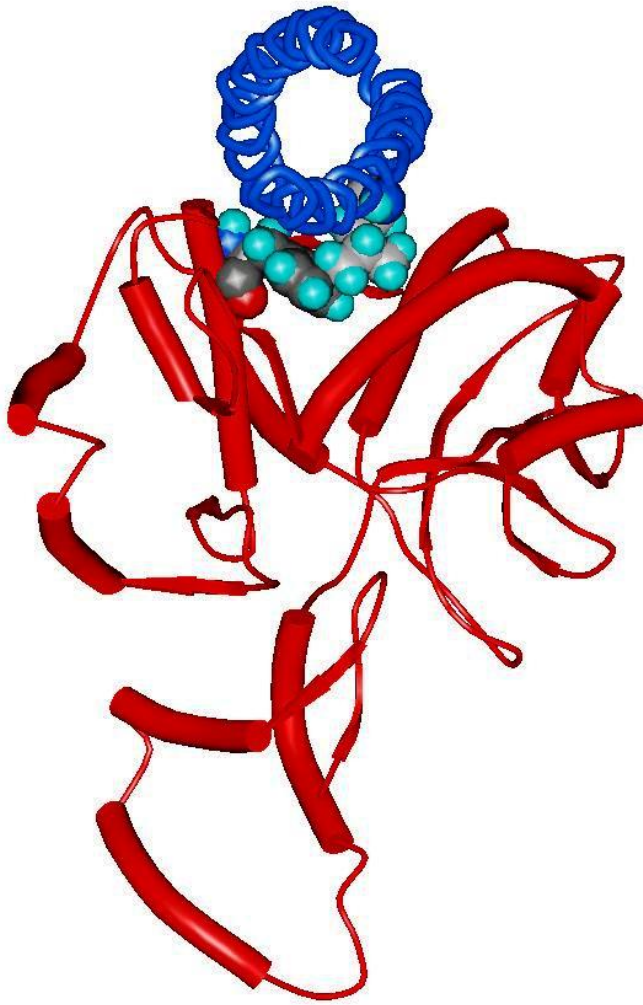


Abbildung 4.22 B: Mutation 5
Dargestellt ist der Proteinkomplex „LEGK“ aus Domäne B (rote Sekundärstruktur) und Keratin (blaue Helices). In der Mutante ist das H2363 durch ein F ersetzt.

Mutation 6 (zwei Varianten): E2421I oder E2421M

Der exponierte, nicht konservierte Aminosäurerest E2421 hat Kontakt zu T192 auf K18 (L bei Haar- und Hautkeratinen), zu T195 auf K18 (T und A bei Hautkeratinen, C bei den Haarkeratinen K33, K 35, K 40 und G bei dem Haarkeratin K39) sowie zu Y204 und K207 auf K8 (beide sind hoch konserviert) - (siehe Abbildung 4.23 A). Durch diese Mutationen könnten apolare Bereiche zur Deckung gebracht werden, welche den Komplex in wässriger Umgebung stabilisieren würden. Die Möglichkeit zur Ausbildung einer Salzbrücke zwischen E2421 und K207 wäre dann allerdings nicht mehr gegeben (O-N Abstand 5 Å).

Da sich diese Salzbrücke in wässriger Umgebung bilden würde, besitzt sie wahrscheinlich geringere Stabilität als in verborgener Proteinumgebung. Außerdem kann K207 eine (selten vorkommende) intrahelikale Salzbrücke zu E211 bilden (O-N Abstand 3,3 Å). Zudem kann E2421 auch eine intramolekulare Salzbrücke zu R2284 ausbilden (O-N Abstand 5,4 Å). Beides würde sich zusätzlich destabilisierend auf die mögliche Salzbrücke zwischen E2421 und K207 auswirken. Ein Isoleucinrest (Methioninrest) könnte dagegen durch Kontakt mit der lipophilen Lysinseitenkette und durch die anderen Reste in einen „apolaren Käfig“ integriert werden, was sich in einem verstärkten "Hydrophoben Effekt" bemerkbar machen müsste. (siehe Abbildung 4.23 B).

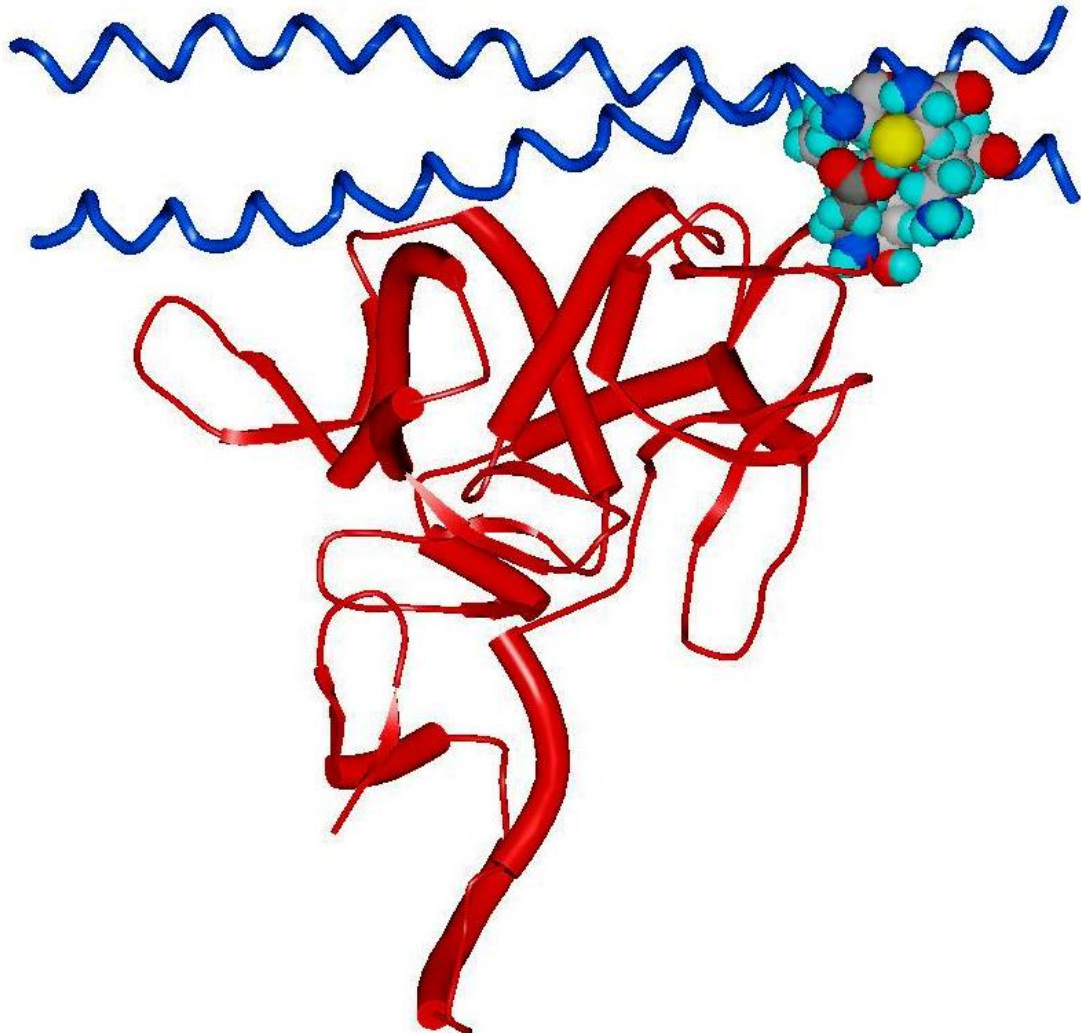


Abbildung 4.23 A: Mutation 6
Dargestellt ist der Proteinkomplex „LEGK“ aus Domäne B (rote Sekundärstruktur) und Keratin (blaue Helices). Beim Wildtyp befindet sich E2421 (Domäne B) in Nachbarschaft zu einem L, C, Y und K am Haarkeratin.

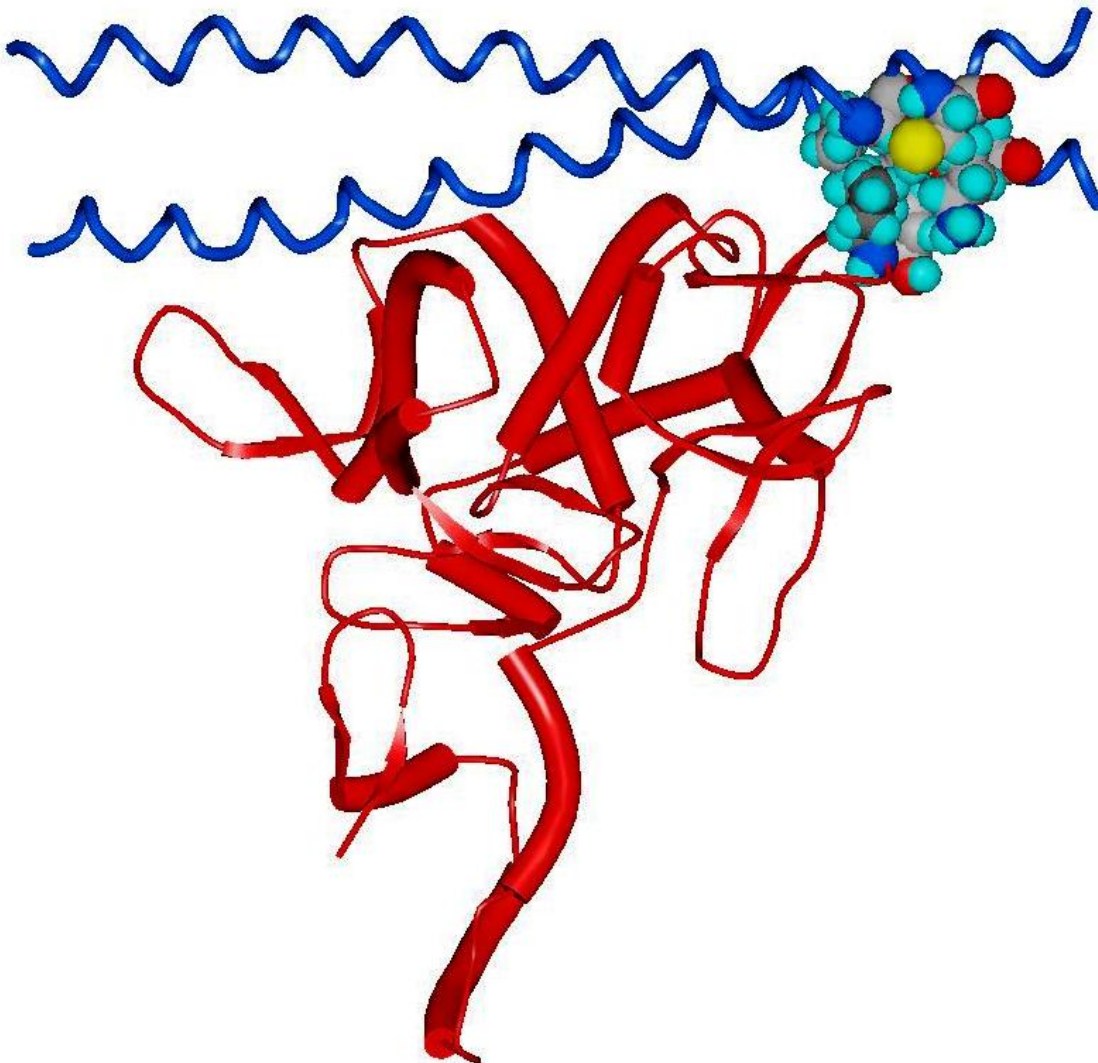


Abbildung 4.23 B: Mutation 6
Dargestellt ist der Proteinkomplex „LEGK“ aus Domäne B (rote Sekundärstruktur) und Keratin (blaue Helices). In der Mutante ist das E2421 durch ein I ersetzt.

5 Diskussion

5.1 Allgemein

Intermediärfilamente (IF) sind ein wesentlicher Bestandteil des Zytoskeletts in Eukariotischen und Prokariotischen Zellen (nicht bei Plantae und Arthropoden). Eine wichtige, in Zellen häufig anzutreffende Unterklasse von IF-Proteinen, bilden die Zytokeratine vom kombinierten IF-Typ I/II. Der exakte molekulare Aufbau sowie die funktionale Wirkungsweise bzw. Wechselwirkung mit anderen Zellstrukturen von IF, sind bis heute nur in Ansätzen bekannt/verstanden (siehe Abschnitt 1). Zelluläre Proteine, welche ein Potential (aufgrund grober experimenteller Ergebnisse) besitzen, mit IF-Proteinen in nichtkovalent-bindende-Protein-Protein-Wechselwirkungen (NBPPW) zu treten, werden als IF-assoziierte Proteine (IFAP) bezeichnet. Desmoplakin, ein IFAP, trägt innerhalb seines C-terminalen Endes drei globuläre, homologe Domänen (A, B und C), welche unabhängig voneinander die NBPPW zu IF-Proteinen vermitteln können. Völlig unklar ist jedoch der Proteininterface-Bereich (also die Aminosäurereste, die an der Bindung beteiligt sind) sowohl der IF-Proteine als auch der drei globulären, homologen Domänen. Für IF-Proteine konnte gezeigt werden, dass dieses Interface innerhalb der „rod“ liegt (siehe Abschnitt 1.4.1). Die IF-„rod“ von Vertebraten besteht aus etwa 620 Aminosäureresten. Da sich in enger Nachbarschaft zu einer IF-„rod“ weitere „rod“-Abschnitte befinden (nichtkovalente, nicht exakt verstandene Polymerisierung der IF-Grundbausteine), erscheint es daher möglich, dass diese benachbarten „rod“-Abschnitte ebenfalls an der Bindung zu den drei globulären, homologen Domänen beteiligt sind. Die Domäne B kann sowohl an Vimentin (IF-Typ III), verschiedene Zytokeratine und Haarkeratine (beides der kombinierte IF-Typ I/II) binden. Es ist nicht bekannt, ob die Domäne B an ein bestimmtes, oder an verschiedene Haarkeratine binden kann. Die Kristallstruktur (Homodimer) der Domäne B besteht aus insgesamt 596 Aminosäureresten. Es war vor meiner Arbeit nicht genau bekannt, ob dieses Homodimer ein funktionales Dimer (im Sinne einer Bindung an IF-Proteine) oder nur ein Artefakt des Kristallisationsprozesses darstellt.

Um die Zielsetzung (siehe Abschnitt 2.2) dieser Arbeit zu erfüllen, war es notwendig, dass genaue Interface beider Proteinkomplex-Partner bzw. eine

möglichst exakte Protein-Protein Wechselwirkungsgeometrie (IF-„rod“ und Kristallstruktur der Domäne B) zu bestimmen. Betrachtet man die IF-„rod“ und die Kristallstruktur der Domäne B zusammen, so liegt das mögliche Interface in mindestens 1216 Aminosäureresten verborgen. Um diese Geometrie zu bestimmen, habe ich eine theoretische Betrachtungsweise verwendet, die auf wissenschaftlichen, biologisch-chemisch-physikalischen Grundlagen basiert. Zur Unterstützung wurden verschiedene Methoden der theoretischen Chemie verwendet. Zuvor musste ich jedoch noch ein Strukturmodell der IF „rod“ durch „Homologie Modelling“ erstellen, da hierfür keine vollständigen Strukturkoordinaten bekannt sind. Ich konnte in der Literatur keine vergleichbare Arbeit finden, die wie meine Forschungsarbeit versucht hat, Protein-Protein-Wechselwirkungen vorauszusagen (ohne Hinweise auf die Interfacelokalisation durch Angabe von einzelnen Aminosäuren), welche dann experimentell nachvollzogen werden konnten. Es gibt jedoch allgemein betrachtet, keinen Unterschied (ausgenommen von der Systemgröße) von Protein-Protein-Wechselwirkungen im Vergleich zu Protein-Ligand-Wechselwirkungen, da sich diese immer auf die räumlich verteilten, zeitlich-dynamisch veränderlichen, molekular-verteilten, physikalischen Wechselwirkungsfelder (bedingt durch die jeweilige chemische Struktur) zurückführen lassen. Deshalb habe ich mich unter anderen an den theoretischen Arbeiten zu Protein-Ligand-Wechselwirkungen meines Mentors Prof. Höltje orientiert.

Bedingt durch die in dieser Forschungsarbeit betrachteten enormen Systemgröße aus beiden Proteinpartnern im wässrigen, ionenhaltigen Milieu, konnten (auf Grund vom Rechenzeitaufwand/Nutzen-Verhältnis) keine exakten theoretischen Methoden (quantenchemische Methoden) verwendet werden. Die Verwendung von weniger exakten Methoden war jedoch möglich. Weniger exakt bedeutet dabei, dass diese Methoden verschiedene Näherungsverfahren verwenden, die das Rechenzeitaufwand/Nutzen-Verhältnis in einen vertretbaren Zeitrahmen versetzen und somit eine Verwendung von theoretischen Methoden

überhaupt erst ermöglichen. Diese Näherungsverfahren bedingen jedoch, dass die berechneten Ergebnisse die Realität oft nicht exakt genug widerspiegeln. Es ist deshalb außerordentlich wichtig, dass diese so erhaltenen Ergebnisse einer kritischen Analyse bzw. Bewertung durch den Anwender unterzogen werden. Auf die Verwendung von Kraftfeldmethoden möchte ich an dieser Stelle im Besonderen noch einmal eingehen, da sie in dieser Arbeit einen hohen Stellenwert besaßen.

5.2 Kraftfeldmethoden und Mutationsvorschläge

Geometrieoptimierungen, die auf Kraftfeldmethoden (im Vakuum) basieren, konnten in dieser Forschungsarbeit schnell und zuverlässig gute und brauchbare Ergebnisse (in Form von energetisch günstigen Konformationen) liefern. Das globale Minimum kann so jedoch gerade bei großen und komplexen Systemen, wie diesem, nicht näherungsweise berechnet werden. Dieses globale Minimum ist zudem von weiteren Faktoren, wie z. B. beteiligte Wassermoleküle und Ionen, abhängig. Für eine repräsentative Verteilung von Seitenkettenkonformationen an der Proteinoberfläche ist dieses globale Minimum aber nicht von größerem Interesse. Denn bedingt durch die thermische Energie (im in Frage kommenden Temperaturbereich für Lebewesen - weit oberhalb von Temperaturen, die bei Röntgendiffraktometrie verwendet werden) kommt es zu einer ständigen, dynamischen Veränderung dieser Verteilung. Es ist deshalb vollkommen ausreichend gewesen, dass ich geometrieoptimierte Momentaufnahmen, welche aus Moleküldynamik-Simulationen erhalten wurden, betrachtet habe. Aus diesen konnte ich einen Überblick von den räumlich verteilten, zeitlich-dynamisch veränderlichen, molekular-verteilten, physikalischen Wechselwirkungsfeldern (bedingt durch die jeweilige chemische Struktur) gewinnen. Die Anwendung von Moleküldynamik-Simulationen (in wässriger, ionenhaltiger Umgebung; inklusive „simulated annealing“), welche ebenfalls auf Kraftfeldmethoden basieren, habe ich vor allem aus dem Grund verwendet, um in einer grob arrangierten Wechselwirkungsgeometrie in die Nähe des globalen Minimums (bezüglich der Seitenkettenkonformation im Wechselwirkungsinterface) zu kommen. Dies ist

zum Ableiten von Mutationen (im Sinne der Zielsetzung, siehe Abschnitt 2.2) aus einer Wechselwirkungsgeometrie wichtig, da die Seitenkettenbeweglichkeit im Proteininterface stark eingeschränkt ist. Das globale Minimum entspricht dabei der Realität. Wenn also die Konformation einer besonders langen Seitenkette, wie z.B. in Lysin falsch ist, so kann dies negative Folgen bei der Auswahl von Mutationsvorschlägen haben. Ich habe dennoch versucht, diesen Sachverhalt so gut wie möglich mit einzubeziehen. Die Voraussagen für das Zwischenziel konnten experimentell nachvollzogen werden (siehe Abschnitt 4.2.7). Diese Mutationen waren dazu bestimmt, die Bindungsaffinität der Domäne B am Keratin herabzusetzen, um somit meine Hypothese über das Interface auf der Domäne B zu verifizieren. Die vorgeschlagenen Mutationen im vermuteten Interface, führten zu einer Verminderung der Lipophilie, zu veränderten Ladungen oder zur Einführung einer sterischen Hinderung. Solche Veränderungen in der Bindungsregion verursachten offensichtlich eine Abnahme der Bindungsaktivität. Eine relativ exakte Wechselwirkungsgeometrie zum Ableiten dieser Mutationsvorschläge wurde hierfür nicht benötigt. Im Gegensatz zu den Mutationsvorschlägen des Hauptzieles. Diese wurden aus der für mich wahrscheinlichsten Wechselwirkungsgeometrie „LEKG“ abgeleitet. Die „LEKG“ wurde im Wesentlichen durch die visuelle Analyse der Strukturmodelle im Stereomodus bestimmt. Die Komplex-Strukturkoordinaten selbst wurden durch gezielte Dockingversuche mit Hex erhalten, da so die optimalen intermolekularen Abstände gut eingestellt werden konnten. Diese Geometrie konnte durch eine Moleküldynamik-Simulation mit dem Programm GROMACS genauer studiert und weiter verbessert werden. Die Verwendung von Protein-Docking-Programmen zur systematischen Suche der optimalen Wechselwirkungsgeometrie lieferte in diesem Zusammenhang keine brauchbaren Ergebnisse. Beispielsweise konnte das Programm Hex erst durch die vollständige Einschränkung des Suchraumes die Kristallstruktur des Dimers der Domäne B wieder finden (siehe Abschnitt 4.2.1), obwohl die Seitenketten in der für das Dimer richtigen Konformation vorlagen. Viele andere (falsche), vom Programm vorgeschlagene Komplexgeometrien hatten zu dem stärkere Wechselwirkungsenergien (das gilt sowohl für das Hex-„Shape“ als auch für

das Elektrostatische Potential). Ich gehe deshalb davon aus, dass man die auf Grundlage von Kraftfeldrechnungen, im Vakuum vorgenommene Bewertung von Proteinkomplexen kritisch betrachten muss. Außerdem ist die richtige Konformation der Seitenketten in der gesuchten (richtigen) Wechselwirkungsgeometrie im Normalfall unbekannt. Aus diesen Gründen wurde von einer weiteren Verwendung des Programms Hex abgesehen. An Stelle dessen wurde das Programm HADDOCK benutzt. Es scheint geeignet zu sein, (in Verbindung mit experimentellen Daten) realitätsnahe Wechselwirkungsgeometrien zu erzeugen [43]. Dieses Programm erlaubt zumindest in bereits gefundenen Wechselwirkungsgeometrien („rigid body docking“) eine Veränderung der Seitenkettenkonformation über lokale Minima hinaus („simulated annealing“; siehe Abschnitt 3.8.2). Die endgültige Bewertung der Komplexe durch ein Kraftfeld findet zudem in wässriger Umgebung statt. Ein Nachteil liegt jedoch in sehr langen Rechenzeiten. Eine systematische Suche war somit ausgeschlossen.

Leider kann ich keine experimentell unterstützten Aussagen zu den Mutationsvorschlägen des Hauptzieles treffen, da die entsprechenden Experimente noch nicht durchgeführt worden sind. Deshalb kann ich zu diesem Zeitpunkt nicht beurteilen, ob die Kombination aus visueller Betrachtung im Stereomodus und Moleküldynamik-Simulation geeignet ist, realitätsnahe Wechselwirkungsgeometrien zu erzeugen.

Im Folgenden möchte ich auf eine der ersten Herausforderungen in meiner Arbeit eingehen, da ich eine von mir getroffene, theoretische Voraussage später durch ein experimentelles Ergebnis des Kooperationspartners überprüfen konnte.

5.3 Vervollständigung der Kristallstruktur der Domäne B

Die eingefügten „loop“-Bereiche in die Kristallstruktur 1LM7 (Homodimer) befinden sich im N-terminus eines jeden Monomers der Domäne B. Dieser N-Terminus ist aber für die Bindung zum Keratin nicht notwendig. Dieser mit

Unsicherheiten behaftete Teil der Arbeit hat somit keinen größeren Einfluss auf die erzielten Ergebnisse gehabt. Interessant war jedoch, die eingefügten „loops“ mit den analogen Bereichen in der neuen, nicht publizierten Kristallstruktur des Kooperationspartners BASF zu vergleichen (hier sind alle Aminosäurereste molekular aufgelöst). Der kleinere eingefügte „loop“ (des einen Monomers) findet sich dabei, im Gegensatz zum größeren „loop“ (des anderen Monomers), mit großer Ähnlichkeit in der Kristallstruktur wieder. Die Vermutung liegt deshalb nahe, dass mit zunehmender Länge des eingefügten „loop“-Bereiches auch die Unsicherheit ansteigt bzw. die Qualität der eingefügten Strukturkoordinaten abnimmt. Es wären demzufolge bessere Ergebnisse erzielt worden, wenn man einen kompletten, passenden Strukturbereich, in welchem der kleinere „loop“ involviert ist, von dem einen Monomer auf das andere Monomer übertragen hätte.

Im Folgenden möchte ich auf die Konstruktion der Strukturmodelle durch „Homologie Modelling“ näher eingehen.

5.4 Konstruktion der Strukturmodelle von Keratinen der Haut und des Haares

Die Konstruktion der entsprechenden Strukturmodelle konzentrierte sich nur auf die Bereiche, welche die Realität mit hoher Wahrscheinlichkeit präzise widerspiegeln können. Diese Bereiche umfassten die vier „coiled coil“-Segmente der IF „rod“. Der Versuch, die „linker“-Bereiche zu modellieren, stellte sich als reine Spekulation heraus und wurde deshalb nicht weiter verfolgt. Gleiches gilt natürlich auch für die „head“- bzw. „tail“-Domänen. Es war aus diesem Grund zu Beginn meiner Forschungsarbeit von Wichtigkeit in Erfahrung zu bringen, dass die Bindung der Domäne B zumindest auf den „rod“-Bereich beschränkt ist.

Der Versuch, die weitere nichtkovalente Polymerisierung dieser IF-Grundstruktur zur nächsten übergeordneten Struktur auf der Basis der vorhandenen experimentellen Daten zumindest teilweise theoretisch nachzuempfinden, erscheint möglich. Angesichts der weiteren dramatischen Vergrößerung des ohnehin schon großen Systems, stellt dieses Vorhaben

besondere Anforderungen an die technischen und zeitlichen Voraussetzungen. Ich konnte diese Herausforderung in meiner Forschungsarbeit nicht bewältigen, obgleich ein daraus resultierendes Ergebnis im Hinblick des Erreichens der Zielsetzung (siehe Abschnitt 2.2) von großem Interesse gewesen wäre. Man hätte daraus Einschränkungen bezüglich der für die Domäne B freizugänglichen IF „rod“-Bereiche treffen können. Außerdem hätte man abschätzen bzw. definieren können, inwiefern benachbarte IF „rod“-Bereiche an der Bindung zur Domäne B beteiligt sind. Dies hätte den Auswahlbereich für Mutationsvorschläge (im Sinne der Zielsetzung; siehe Abschnitt 2.2) vergrößern können.

Abschließend möchte ich auf die experimentellen Ergebnisse eingehen.

5.5 Experimentelle Ergebnisse des Kooperationspartners BASF

Die für diese Arbeit wichtigen Mutationsstudien, die aus den theoretischen Voraussagen resultierten, wurden vom Kooperationspartner BASF durchgeführt. Genauere Informationen zu diesen Experimenten, als die in dieser Arbeit von mir angegebenen (siehe Abschnitt 4.2.7), liegen mir nicht vor. Diese Ergebnisse können deswegen nicht von mir diskutiert werden.

6 Zusammenfassung

6.1 Zusammenfassung

Die vorliegende Arbeit untersuchte die bindenden Wechselwirkungen auf molekularer Ebene zwischen KBD-B, einem globulären Protein aus der Plectinfamilie und stabförmigen IF-Proteinen aus der Familie humaner Intermediärfilamente. Die dazu benötigten atomaren Strukturmodelle wurden im Fall der KBD-B durch die Vervollständigung eines nichtaufgelösten „loop“-Bereichs der Kristallstruktur 1LM7 erhalten. Die Strukturmodelle der IF-Proteine wurden durch „homology modelling“ generiert. Die Ausgangslage zur Erforschung der Protein-Protein-Interaktionen war in diesem Fall besonders schwierig, da es keine experimentellen Daten gab, welche das Proteininterface beider Wechselwirkungspartner näher einschränkte. Aus der theoretischen bzw. „computational chemistry“ unterstützten Betrachtungsweise resultierte zunächst eine Hypothese von der Bindungsregion auf der KBD-B. Diese Region erschien aufgrund der guten „shape“-Komplementarität besonders geeignet zu sein, mit den stabförmigen „coiled coil“-Bereichen der IF-Proteine stabile Proteinwechselwirkungen ausbilden zu können. Auf der Basis dieser Vorstellung wurden Vorschläge für Mutationen abgeleitet, welche zu einer Störung in der Proteininteraktion und somit zu einer Verringerung der Bindungsaffinität führen sollten. Diese Hypothese konnte anschließend durch Mutations/Bindungsstudien des Kooperationspartners BASF experimentell verifiziert werden. Im weiteren Verlauf wurde nach einer potentiellen Bindungsstelle der gefundenen KBD-B-Bindungsregion an IF-Proteinen gesucht. Hierzu wurde die gleichmäßige, stabförmige Oberfläche der IF-Proteine nach komplementären molekularen Wechselwirkungsfeldern abgesucht. Die Komplementarität von apolaren Bereichen, welche einen starken „Hydrophoben Effekt“ bedingen, war dabei das entscheidende Suchkriterium. Elektrostatische und sterische Eigenschaften wurden ebenso berücksichtigt. Eine in dieser Hinsicht besonders attraktive Wechselwirkungsgeometrie konnte letztendlich definiert werden. Die daraus abgeleiteten Mutationsvorschläge für verbesserte Bindungseigenschaften möchte der Kooperationspartner BASF durch ein Patent schützen lassen.

6.2 Summary

The present dissertation examined the binding interaction at molecular level between KBD-B, a globular protein from the plectin family and stick-shaped IF proteins from the family of human intermediate filaments. The required atomic structural models were received in case of the KBD-B by the completion of a missing loop region in the crystal structure 1LM7. The structural models of the IF proteins were generated by homology modelling. In this case the situation for starting the investigation of the protein protein interactions was especially difficult, because there was no experimental data which limited the protein interface of both interaction partners in more detail. The resulting first hypothesis for the interaction region on the KBD B was based on a „computational chemistry “ aided approach. This region seems to be able to form stable protein protein complexes with stick-shaped coiled coil regions of the IF proteins due to good shape complementarity. Based on this assumption suggestions were derived for the mutations which should lead to a disturbance in the protein protein interaction and therefore to a reduction of the binding affinity. Subsequently this hypothesis could be verified by mutations/affinity studies of the cooperation partner BASF experimentally. Next it was searched for a potential binding site of the detected KBD-B interface in IF proteins. For this purpose the homogenous, stick-shaped surface of the IF proteins was investigated for complementary molecular interaction fields. The complementarity of apolar molecular fields which presumably cause a strong hydrophobic effect was the determining searching criterion. Electrostatic and steric qualities were also taken into consideration. Finally an interaction geometry especially attractive in this regard could be defined. The cooperation partner BASF would like to protect by patent the derived mutation suggestions for improved affinity qualities.

7 Anhang

7.1 Abkürzungen

| | |
|-------|---|
| AIDS | Acquired Immune Deficiency Syndrome |
| DNA | Desoxyribonukleinsäure |
| DPCT | Desmoplakin C-terminal |
| GFAP | Glial Fibrillary Acidic Protein |
| HIV | Human Immunodeficiency Virus |
| IF | Intermediärfilamente |
| IFAP | IF-assoziierte Proteine |
| K | Keratin |
| KBD-B | Keratinbindedomäne B |
| LEKG | lipophil elektrostatisch komplementäre Geometrie |
| MD | Moleküldynamik |
| NBPPW | nichtkovalent-bindende-Protein-Protein-Wechselwirkungen |
| NF-H | schweres Neurofilament |
| NF-L | leichte Neurofilament |
| NF-M | mittelschweres Neurofilament |
| NMR | nuclear magnetic resonance |
| PBC | periodic boundary conditions |
| PDB | Protein Data Base |
| SDS | Sodiumdodecylsulfat |
| SPC | Single Point Charge |

7.2 Einheiten

- Å Ångström, Längeneinheit ($1 \text{ Å} = 10^{-10} \text{ m}$)
- K Kelvin, Temperatureinheit ($0 \text{ K} = -273,15^\circ \text{ Celsius}$)
- kJ Kilojoule, Energieeinheit ($1 \text{ kJ} = 0,239 \text{ kcal}$)
- ps Picosekunde, Zeiteinheit ($1 \text{ ps} = 10^{-12} \text{ s}$)
- mol Mol, Einheit der Stoffmenge ($1 \text{ mol} = 6 \cdot 10^{23} \text{ Teilchen}$)

7.3 Curriculum Vitae

Online nicht verfügbar!

matthiasvolkmar@gmx.de

7.4 Danksagung

An dieser Stelle möchte ich mich bei den Menschen persönlich bedanken, die besonders zum Gelingen dieser Arbeit beigetragen haben.

Ganz besonderer Dank gilt meinem Doktorvater Prof. Höltje, der mir die Möglichkeit gegeben hat, in seinem Arbeitskreis unter hervorragenden Arbeitsbedingungen zu promovieren.

Weiterhin möchte ich mich ganz herzlich bei Herrn Dr. Schleifer bedanken, der als Ansprechpartner im Rahmen der Kooperation mit der BASF jederzeit für konstruktive Diskussionen sowie als Zweitgutachter zur Verfügung stand.

Größten Dank möchte ich auch den Mitgliedern des Arbeitskreises aussprechen. Ihre Kooperationsbereitschaft bzw. ihr Interesse an wissenschaftlichen Diskussionen waren für neue Ideen und kreative Gedanken überaus hilfreich. Ihre Erfahrungen, Tipps und Tricks im Umgang mit verschiedenen Programmen haben mir zudem nicht nur den Einstieg in die Arbeit wesentlich erleichtert.

Ganz liebevoll möchte ich mich auch bei meiner Freundin Nadine Diekmann und bei meinen Eltern Gisela und Jürgen für ihre allgegenwärtige Unterstützung bedanken.

8 Literaturverzeichnis

1. Parry, D. A.; Steinert, P. M. Intermediate filaments: molecular architecture, assembly, dynamics and polymorphism. *Q Rev Biophys* **1999**, 32, 99-187.
2. Wang, Q.; Tolstonog, G. V.; Shoeman, R.; Traub, P. Sites of nucleic acid binding in type I-IV intermediate filament subunit proteins. *Biochemistry* **2001**, 40, 10342-9.
3. Herrmann, H.; Aebi, U. Intermediate filaments and their associates: multi-talented structural elements specifying cytoarchitecture and cytodynamics. *Curr Opin Cell Biol* **2000**, 12, 79-90.
4. Fuchs, E.; Karakesisoglou, I. Bridging cytoskeletal intersections. *Genes Dev* **2001**, 15, 1-14.
5. <http://www.interfil.org/aboutDB.php>. The Human Intermediate Filament Database, Centre for Molecular Medicine and the Bioinformatics Institute Singapore, 2007
6. Shoeman, R. L.; Huttermann, C.; Hartig, R.; Traub, P. Amino-terminal polypeptides of vimentin are responsible for the changes in nuclear architecture associated with human immunodeficiency virus type 1 protease activity in tissue culture cells. *Mol Biol Cell* **2001**, 12, 143-54.
7. Geisler, N.; Kaufmann, E.; Weber, K. Proteinchemical characterization of three structurally distinct domains along the protofilament unit of desmin 10 nm filaments. *Cell* **1982**, 30, 277-86.
8. Conway, J. F.; Parry, D. A. D. Intermediate Filament Structure .3. Analysis of Sequence Homologies. *International Journal of Biological Macromolecules* **1988**, 10, 79-98.
9. Strelkov, S. V.; Herrmann, H.; Geisler, N.; Wedig, T.; Zimbelmann, R.; Aebi, U.; Burkhard, P. Conserved segments 1A and 2B of the intermediate

filament dimer: their atomic structures and role in filament assembly. *Embo J* **2002**, 21, 1255-66.

10. Herrmann, H.; Aebi, U. Intermediate filaments: molecular structure, assembly mechanism, and integration into functionally distinct intracellular Scaffolds. *Annu Rev Biochem* **2004**, 73, 749-89.

11. Burkhard, P.; Kammerer, R. A.; Steinmetz, M. O.; Bourenkov, G. P.; Aebi, U. The coiled-coil trigger site of the rod domain of cortexillin I unveils a distinct network of interhelical and intrahelical salt bridges. *Structure* **2000**, 8, 223-30.

12. Parry, D. A.; Strelkov, S. V.; Burkhard, P.; Aebi, U.; Herrmann, H. Towards a molecular description of intermediate filament structure and assembly. *Exp Cell Res* **2007**, 313, 2204-16.

13. Steinert, P. M.; Marekov, L. N.; Fraser, R. D.; Parry, D. A. Keratin intermediate filament structure. Crosslinking studies yield quantitative information on molecular dimensions and mechanism of assembly. *J Mol Biol* **1993**, 230, 436-52.

14. Schweizer, J.; Bowden, P. E.; Coulombe, P. A.; Langbein, L.; Lane, E. B.; Magin, T. M.; Maltais, L.; Omary, M. B.; Parry, D. A.; Rogers, M. A.; Wright, M. W. New consensus nomenclature for mammalian keratins. *J Cell Biol* **2006**, 174, 169-74.

15. Hutton, E.; Paladini, R. D.; Yu, Q. C.; Yen, M.; Coulombe, P. A.; Fuchs, E. Functional differences between keratins of stratified and simple epithelia. *Journal of Cell Biology* **1998**, 143, 487-499.

16. Fuchs, E.; Cleveland, D. W. A structural scaffolding of intermediate filaments in health and disease. *Science* **1998**, 279, 514-9.

17. Swensson, O.; Langbein, L.; McMillan, J. R.; Stevens, H. P.; Leigh, I. M.; McLean, W. H. I.; Lane, E. B.; Eady, R. A. J. Specialized keratin expression pattern in human ridged skin as an adaptation to high physical stress. *British Journal of Dermatology* **1998**, 139, 767-775.

18. Lodish, H. *Molecular Cell Biology* ISBN 0-7167-3136-3 **2000**.
19. Mizuno, Y.; Thompson, T. G.; Guyon, J. R.; Lidov, H. G.; Brosius, M.; Imamura, M.; Ozawa, E.; Watkins, S. C.; Kunkel, L. M. Desmuslin, an intermediate filament protein that interacts with alpha -dystrobrevin and desmin. *Proc Natl Acad Sci U S A* **2001**, 98, 6156-61.
20. Newey, S. E.; Howman, E. V.; Ponting, C. P.; Benson, M. A.; Nawrotzki, R.; Loh, N. Y.; Davies, K. E.; Blake, D. J. Syncoilin, a novel member of the intermediate filament superfamily that interacts with alpha-dystrobrevin in skeletal muscle. *J Biol Chem* **2001**, 276, 6645-55.
21. Goulielmos, G.; Gounari, F.; Remington, S.; Muller, S.; Haner, M.; Aebi, U.; Georgatos, S. D. Filensin and phakinin form a novel type of beaded intermediate filaments and coassemble de novo in cultured cells. *J Cell Biol* **1996**, 132, 643-55.
22. Choi, H. J.; Park-Snyder, S.; Pascoe, L. T.; Green, K. J.; Weis, W. I. Structures of two intermediate filament-binding fragments of desmoplakin reveal a unique repeat motif structure. *Nat Struct Biol* **2002**, 9, 612-20.
23. Fontao, L.; Favre, B.; Riou, S.; Geerts, D.; Jaunin, F.; Saurat, J. H.; Green, K. J.; Sonnenberg, A.; Borradori, L. Interaction of the bullous pemphigoid antigen 1 (BP230) and desmoplakin with intermediate filaments is mediated by distinct sequences within their COOH terminus. *Mol Biol Cell* **2003**, 14, 1978-92.
24. Meng, J. J.; Bornslaeger, E. A.; Green, K. J.; Steinert, P. M.; Ip, W. Two-hybrid analysis reveals fundamental differences in direct interactions between desmoplakin and cell type-specific intermediate filaments. *J Biol Chem* **1997**, 272, 21495-503.
25. Nikolic, B.; Mac Nulty, E.; Mir, B.; Wiche, G. Basic amino acid residue cluster within nuclear targeting sequence motif is essential for cytoplasmic plectin-vimentin network junctions. *J Cell Biol* **1996**, 134, 1455-67.

26. Langbein, L.; Rogers, M. A.; Winter, H.; Praetzel, S.; Beckhaus, U.; Rackwitz, H. R.; Schweizer, J. The catalog of human hair keratins. I. Expression of the nine type I members in the hair follicle. *J Biol Chem* **1999**, *274*, 19874-84.
27. Langbein, L.; Rogers, M. A.; Winter, H.; Praetzel, S.; Schweizer, J. The catalog of human hair keratins. II. Expression of the six type II members in the hair follicle and the combined catalog of human type I and II keratins. *J Biol Chem* **2001**, *276*, 35123-32.
28. Langbein, L.; Rogers, M. A.; Praetzel-Wunder, S.; Bockler, D.; Schirmacher, P.; Schweizer, J. Novel type I hair keratins K39 and K40 are the last to be expressed in differentiation of the hair: completion of the human hair keratin catalog. *J Invest Dermatol* **2007**, *127*, 1532-5.
29. Hutton, E.; Paladini, R. D.; Yu, Q. C.; Yen, M.; Coulombe, P. A.; Fuchs, E. Functional differences between keratins of stratified and simple epithelia. *J Cell Biol* **1998**, *143*, 487-99.
30. Wojcik, S. M.; Bundman, D. S.; Roop, D. R. Delayed wound healing in keratin 6a knockout mice. *Mol Cell Biol* **2000**, *20*, 5248-55.
31. Thompson, J. D.; Higgins, D. G.; Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **1994**, *22*, 4673-80.
32. Henikoff, S.; Henikoff, J. G. Amino acid substitution matrices from protein blocks. *Proc Natl Acad Sci U S A* **1992**, *89*, 10915-9.
33. *INSIGHT II*, Accelrys Inc.: San Diego, CA, USA, 2003.
34. Berger, B.; Wilson, D. B.; Wolf, E.; Tonchev, T.; Milla, M.; Kim, P. S. Predicting coiled coils by use of pairwise residue correlations. *Proc Natl Acad Sci U S A* **1995**, *92*, 8259-63.
35. SYBYL 7.1
Tripos Associates Inc.: St. Louis, USA, 2005.

36. Canutescu, A. A.; Shelenkov, A. A.; Dunbrack, R. L., Jr. A graph-theory algorithm for rapid protein side-chain prediction. *Protein Sci* **2003**, 12, 2001-14.
37. van Gunsteren, W. F. B., S. R.; Eising, A. A.; Hünenberger, P.H.; Krüger, P.; . Biomolecular Simulation: The GROMOS96 manual and user guide. *Hochschulverlag AG an der ETH Zürich* **1996**.
38. Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. GROMACS: fast, flexible, and free. *J Comput Chem* **2005**, 26, 1701-18.
39. David van der Spoel, E. L., Berk Hess, Aldert R. van Buuren, Emile Apol, Pieter J. Meulenhoff, D. Peter Tieleman, Alfons L.T.M. Sijbers, K. Anton Feenstra, Rudi van Drunen, Herman J.C. Berendsen. GROMACS USER MANUAL, Version 3.2. *Department of Biophysical Chemistry, University of Groningen. Nijenborgh 4*, 9747 AG Groningen, The Netherlands **2002**.
40. Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. Procheck - a Program to Check the Stereochemical Quality of Protein Structures. *Journal of Applied Crystallography* **1993**, 26, 283-291.
41. Morris, A. L.; MacArthur, M. W.; Hutchinson, E. G.; Thornton, J. M. Stereochemical quality of protein structure coordinates. *Proteins* **1992**, 12, 345-64.
42. Ritchie, D. W. Hex 4.1 User Manual, University of Aberdeen, Aberdeen AB24 3UE Scotland, UK
In 2003.
43. Dominguez, C.; Boelens, R.; Bonvin, A. M. HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *J Am Chem Soc* **2003**, 125, 1731-7.
44. Burkhard, P.; Steinmetz, M. O.; Schulthess, T.; Landwehr, R.; Aebi, U.; Kammerer, R. A. Crystallization and preliminary X-Ray diffraction analysis of the

190-A-long coiled-coil dimerization domain of the actin-bundling protein cortexillin I from dictyostelium discoideum. *J Struct Biol* **1998**, 122, 293-6.

45. North, A. C.; Steinert, P. M.; Parry, D. A. Coiled-coil stutter and link segments in keratin and other intermediate filament molecules: a computer modeling study. *Proteins* **1994**, 20, 174-84.

46. Romano, C.; Miller, J. K.; Hyrc, K.; Dikranian, S.; Mennerick, S.; Takeuchi, Y.; Goldberg, M. P.; O'Malley, K. L. Covalent and noncovalent interactions mediate metabotropic glutamate receptor mGlu5 dimerization. *Mol Pharmacol* **2001**, 59, 46-53.

47. Hebert, T. E.; Moffett, S.; Morello, J. P.; Loisel, T. P.; Bichet, D. G.; Barret, C.; Bouvier, M. A peptide derived from a beta2-adrenergic receptor transmembrane domain inhibits both receptor dimerization and activation. *J Biol Chem* **1996**, 271, 16384-92.

48. Furthmayr, H.; Marchesi, V. T. Subunit structure of human erythrocyte glycophorin A. *Biochemistry* **1976**, 15, 1137-44.

49. Bouvier, M. Oligomerization of G-protein-coupled transmitter receptors. *Nat Rev Neurosci* **2001**, 2, 274-86.

50. Cvejic, S.; Devi, L. A. Dimerization of the delta opioid receptor: implication for a role in receptor internalization. *J Biol Chem* **1997**, 272, 26959-64.

51. Barg, H.; Subkowski, T.; Lemaire, H.-G.; Bollschweiler, C.; Ptock, A. KERATIN-BINDENDE DESMOPLAKINPOLYPEPTIDSEQUENZEN; Weltorganisation für geistiges Eigentum, Internationales Büro; WO 2006/097432 A2; 2006.