Proteins and their Ligands:

Their Importance and how to Crystallize them

Astrid Hoeppner[#], Lutz Schmitt* and Sander H. J. Smits*[¶]

[#]*X-ray Facility and Crystal Farm, Heinrich Heine University, Universitaetsstr. 1, D-40225 Duesseldorf, Germany*

**Institute of Biochemistry, Heinrich Heine University, Universitaetsstr. 1, D-40225 Duesseldorf, Germany*

[¶]corresponding author, email: Sander.Smits@hhu.de

1. Introduction

The importance of structural biology has been highlighted in the past few years not only as part of drug discovery programs in the pharmaceutical industry but also by structural genomics programs. Although the function of a protein can be studied by several biochemical and or biophysical techniques a molecular understanding of a protein can only be obtained by combining functional data with the three-dimensional structure. In principle three techniques exist to determine a protein structure, namely X-ray crystallography, nuclear magnetic resonance (NMR) and electron microscopy (EM). X-ray crystallography contributes over 90 % of all structures in the protein data bank (PDB) and emphasis the importance of this technique. Crystallization of a protein is a tedious route and although a lot of knowledge about crystallization has been gained in the last decades, one still cannot predict the outcome. The sometimes unexpected bottlenecks in protein purification and crystallization have recently been summarized and possible strategies to obtain a protein crystal were postulated [1]. This book chapter will tackle the next step: How to crystallize protein-ligand complexes or intermediate steps of the reaction cycle?

A single crystal structure of a protein however, is not enough to completely understand the molecular function. Conformational changes induced by for example ligand binding cannot be anticipated *a priori*. The determination of particular structures of one protein, for example with bound ligand(s) is required to visualize the different states within a reaction cycle. Ideally, one would trap an open conformation without any ligand, an open ligand-bound and a closed form with the bound molecule as well as the closed ligand-free protein to visualize the conformational changes occurring during catalysis in detail.

Within this chapter, the structural conformational changes induced by ligand binding with respect to the methods chosen for the crystallization are described. Here three distinct protein families are exemplarily described: first, where one substrate or ligand is bound, second, a protein with two or more bound substrates and finally, the structures of proteins, in which the product of the reaction cycle is present in the active site.

Within this chapter the conformational changes induced by ligand(s) of three different protein families, with respect to their crystallization possibilities are described. Specific methods or expressions written in ***bold italics*** are explained in the glossary box at the end of the chapter.

1.1 General approaches to obtain crystals with bound ligands and how to prepare the ground

Often the knowledge of the structure of a protein or enzyme without bound ***ligand(s)*** is not sufficiently significant since there is no or only little information provided about the catalytic mechanism. To gain further insights, it is important or at least helpful to obtain a ***binary*** or ***ternary structure*** of the protein of interest.

In theory there are different approaches to reach this goal even though it can be a difficult task in reality. All of them have in common that the naturally catalysed reaction must not occur. Apart from reporting all possible attempts we would like to give a general overview about several ***co-crystallization***/***soaking*** strategies first, followed by selected examples described within this bookchapter.

Possible **co-crystallization** or **soaking** trials:

(In order to keep it simple and coherent the expression „***ligand***“ in the following paragraph is used in terms of „***substrate***“, „***cofactor***“ or „binding partner“.)

- first ligand without second ligand
- second ligand without first ligand
- first ligand with product of the second ligand

- product of the first ligand with second ligand
- substrate analogue/inhibitor or non-hydrolysable cofactor
- application of substances that mimic transition state products (e. g. AlF_3 which imitates a phosphate group)
- usage of catalytically inactive mutants with bound ligand(s)
- creation of an environment (i. g. buffer condition) which shifts the equilibrium constant so that the reaction cannot occur

The most important point concerning preparing ***co-crystallization*** trials is the knowledge of the corresponding kinetical parameters. Proteins bind their natural ligand(s) with high ***affinity***, which means in the nM- up to low mM range. To successfully crystallize a protein with the ligand(s) bound, the affinity needs to be determined. There are numerous biophysical techniques to achieve this, for example ***Intrinsic Tryptophan Fluorescence***, ***Isothermal Calorimetry***, ***Surface Plasmon Resonance*** and many others. In principle the affinity is determined by the size of a ligand as well as the property of the binding site of the protein. As first approximation, one can state that affinity increases with a decrease in ligand size.

The application of a too low concentration of the ligand can lead to an inhomogeneous protein solution, which means that not all of the protein molecules are loaded with ligand (and this can prevent crystallization). It is also possible, that a low level of ***occupancy*** causes an undefined electron density so that the ligand cannot be placed or which even makes a structure solution impossible. As a rule of thumb the concentration of the ligand(s) should be applied to the crystallization trial about 5-fold of the corresponding K_M value.
Beyond that all requirements for the protein solution itself remain valid as described in [1] in more detail.

2. Binding protein with one ligand – how to crystallize and what can be deduced from the structure

A typical class of a protein binding one ***ligand*** are ***substrate***-binding proteins (SBPs), and substrate-binding domains (SBDs) [2]. They form a class of proteins (or protein domains) that are often associated with membrane protein complexes for transport or signal transduction. SBPs were originally found to be associated with prokaryotic ATP binding cassette (ABC)-transporters, but have more recently been shown to be part of other membrane protein complexes as well such as prokaryotic tripartite ATP-independent periplasmic (TRAP)-transporters, prokaryotic two-component regulatory systems, eukaryotic guanylate cyclase-atrial natriuretic peptide receptors, G-protein coupled receptors (GPCRs) and ligand-gated ion channels [2].

Structural studies of a substantial number of SBPs revealed a common fold with a bilobal organization connected via a linker region [2]. In the ligand-free, open conformation, the two lobes or domains are separated from each other, thereby forming a deep, solvent exposed cleft, which harbors the substrate-binding site. Upon ligand binding, both domains of the SBP move towards each other through a hinge-bending motion or rotation, which results in the so-called liganded-closed conformation. As a consequence of this movement, residues originating from both domains generate the ligand-binding site and trap the ligand deeply within the SBP [3]. In the absence of a ligand, unliganded-open and unliganded-closed states of the SBP are in equilibrium, and the ligand solely shifts this equilibrium towards the liganded-closed state. This sequence of events has been coined the “Venus-fly trap mechanism” [4-6]; it is supported by a number of crystal structures in the absence and presence of a ligand [7, 8] and other biophysical techniques [3].

For the maltose binding protein (MBP) from *Escherichia coli* [9], it has been shown that both domains are dynamically fluctuating around an average orientation in the absence of the ligand [10]. NMR spectroscopy of MBP in solution revealed that the ligand-free form of MBP consists of a predominantly open species (95 %) and a minor species (5 %) that corresponds to a partially closed state; both forms co-exist in rapid equilibrium [11]. The open form of MBP observed by NMR is similar to the crystal structure of the unliganded-open conformation [12]. However, the partially closed species detected by NMR [11] does not correspond to the ligand-bound, fully closed form found in crystallographic studies. Instead, it represents an intermediate, partially closed conformation [13], suggesting that the substrate is required to reach the final, liganded-closed conformation.

Upon ***substrate*** binding, the closed conformation is stabilized, and the ***ligand*** is trapped within a cleft in between the two domains [14-16]. In principle one can divide the conformational changes in four (I-IV) states (highlighted in Figure 1). State I is the „open-unliganded“ where the protein adopts an open conformation and no substrate is bound to the protein. State II is the „closed-liganded“ conformation where the substrate is bound and induced a conformational change of both domains towards each other. This is likely the state within the cell before delivery of the substrate to its cognate transporter. Two other states are known to be present in solution although less frequent and the equilibrium is shifted towards the open-unliganded conformation. These forms are state III, the „closed-unliganded“ state and state IV, the „semi-closed-unliganded“ state. These are unfavorable conformations of the SBP, which occur due to the flexibility of the linker region in between both domains.

To fully understand the function as well as the structural changes happening upon ***ligand***/***substrate*** binding it would require structural information of at least states I and II, preferably also of states III and IV.

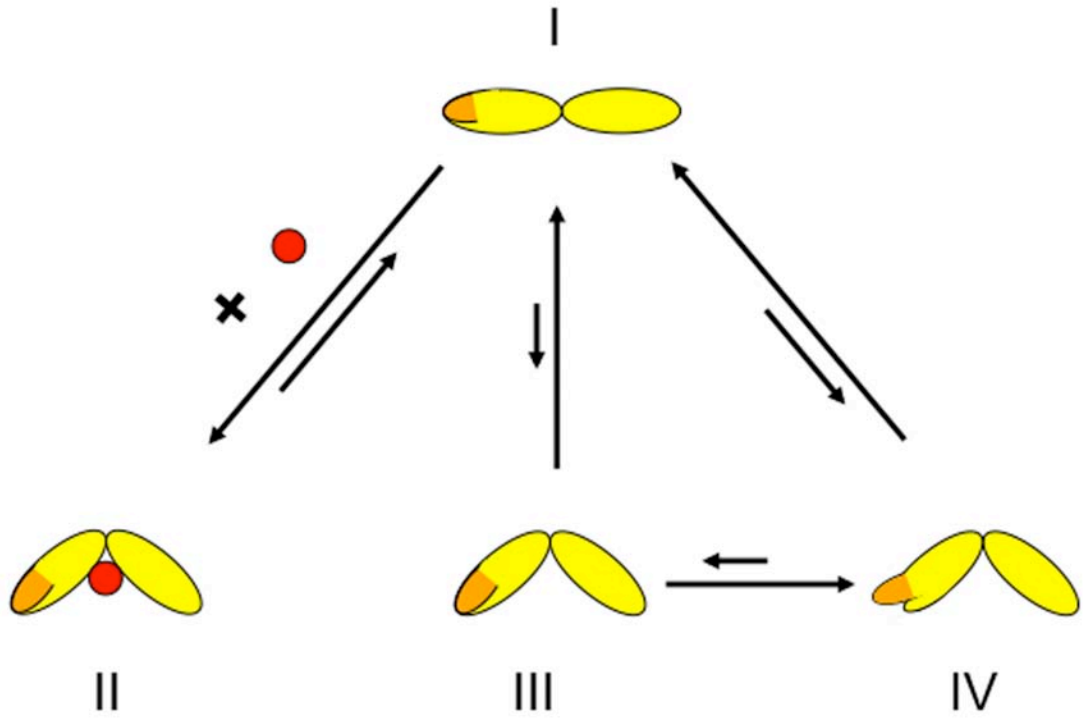


Figure 1. Substrate binding proteins exist in four major conformations: I) unliganded-open II) liganded-closed III) unliganded-closed and IV) unliganded-semi-open. All states are in equilibrium with each other. In solution states I and II occur most frequently. To fully understand the opening and closing mechanism of the protein however snapshots of every state are needed to gain full knowledge.

2.1 Crystallization of the open-unliganded conformation (state I)

The crystallization of an open conformation of a rather flexible protein is not straight forward and most of the success came from „trial-and-error" approaches. After purification of the protein, a reasonable concentration of the protein is taken to set up crystallization trials. Most commonly the ***vapor diffusion*** method with the ***hanging or sitting drop*** is used. SBPs mainly exist in the open-unliganded conformation in the absence of the ***substrate*** whereas only a small fraction is in a closed-unliganded conformation [5, 11, 17]. Thus, basically a standard crystallization approach is used to obtain crystals suitable for structure determination. This is reflected by the large number of structures solved in the unliganded-open conformation (see [2] for a recent summary of the available SBP structures). The open conformation basically gives an overall picture of the protein structures and in the case of SBPs the bilobal fold of the protein can be observed. In this conformation the binding site of the substrate is laid open and a detailed picture on how the substrate is bound cannot be deduced.

Most of the times the open conformation crystallizes differently from the ligand bound state. This is reflected in the different crystallization conditions as well as in changes of the crystal parameters (unit cell and/or spacegroup). One example is given below for the glycine betaine binding protein ProX.

2.2 Crystallization of the substrate bound closed conformation (state II)

The vast majority of ***substrate*** binding proteins have been crystallized in the closed-***ligand*** bound conformation (for a detailed list see [2]). This is mainly due to the fact that the substrate bound protein adopts a stable conformation and possesses a drastically reduced intrinsic flexibility. In principle there are four methods to include the substrate into the crystallization trials: 1) ***co-crystallization*** 2) ligand ***soaking*** 3) ***micro or macro seeding*** 4) endogenously bound ligands.

The first method is ***co-crystallization***. Here, normally the ***substrate*** is added prior to crystallization to the protein solution. As listed in Table 1 this is the method used the most in SBP crystallization trials. Knowledge about the ***affinity*** of the ***ligand*** is important, since the bilobal SPBs exist in equilibrium between the open and closed state in solution and the addition of substrate directs this equilibrium towards the latter. Exemplary, 11 SBPs are listed in Table 1 where the affinity of the corresponding ligand(s) as well as the concentration used in the crystallization trials is highlighted. In principle the concentrations used are 10-1000 times above the K_d.

Table 1. Solved structures of selected SBPs. Listed are the proteins, the host organism, the substrate, whether the structure was solved in the unliganded-open and/or liganded-closed state, the highest resolution, the biochemically determined affinity, the used substrate concentration during crystallization and the method used: 1) co-crystallization 2) soaking 3) seeding 4) endogenously bound substrates.

Protein	Organism	Ligand(s)	Open-un-liganded	Closed-liganded	Reso-lution (Å)	Max. affinity	Used Conc.	Method	Ref.
BtuF	*E. coli*	vitamin B12	Y	Y	2	15 nM	5 mM	1	[18]
Lbp	*S. pyogenes*	zinc	-	Y	2.45	~10 µM	-	4	[19]
GGBP	*S. typhimurium*	D-glucose	Y	Y	1.9	0.5 µM	3 mM	1	[20]
MBP	*E. coli*	oligosaccharide	Y	Y	1.67	0.16 µM	2 mM	1	[21]
RBP	*E. coli*	D-ribose	Y	Y	1.6	0.13 µM	1 mM	1	
OppA	*L. lactis*	oligo peptide	Y	Y	1.3	0.1 µM	0.5-5 mM	1 and 4	[22]
ProX	*A. fulgidus*	glycine betaine, proline betaine	Y	Y	1.8	50 nM	1 mM	1	[23]
PotD	*T. pallidum*	spermidine	-	Y	1.8	10 nM	n.n	2	[24]
SiaP	*H. influenzae*	sialic acid	Y	Y	1.7	58 nM	5 mM	1	[25]
UehA	*S. pomeroyi*	ectoine	-	Y	2.9	1.1 µM	10 mM	1	[26]
ChoX	*S. meliloti*	choline	Y	Y	1.8	2.7 µM	2 mM	1 and 3	[14, 15]

1) Co-crystallization to obtain the ligand bound structure

The method of ***co-crystallization*** ensures the presence of only the substrate bound conformation of the SBP in solution. One major advantage of co-crystallization is the possibility to add different ligands into the crystallization trial. A prominent SBP member where several crystal structures were solved is the maltose binding protein (MBP). This protein binds a maltose molecule and delivers it to its cognate ABC transporter, which imports maltose into the cell for nutrient purposes. Here, the substrates ranges from maltose, maltotriose, beta-cyclodextrin and many other sugar derivatives. All these structures were solved by using the addition of the substrates to the protein. Another example is the ectoine binding EhuB protein of *S. meliloti* [27]. Here, the structure was solved with both ligands, ectoine and hydroxyectoine, which yielded two high-resolution structures. The different binding modes of the substrates could be detected and the difference in affinity explained. The latter example was only crystallized in the closed-liganded state and no crystals could be obtained when the crystallization solution was depleted of substrate. This highlights the flexibility of the SBPs and the presence of multiple conformations of the SBPs in solution and in presence of the ligand. In many cases the ligand-closed conformation was crystallized under conditions, which differ greatly from the unliganded-open conformation also indicating the flexibility in the protein.

2) Ligand soaking to obtain the ligand bound state

The second method, which can be used to obtain a ***ligand*** bound protein structure, is ***ligand soaking*****.** Soaking crystals with ligands is often the method of choice to obtain crystals of protein-ligand complexes owing to the ease of the method. However, there are several factors to consider. The crystals may be fragile and soaking in a stabilization buffer or cross-linking may be required. The soaking time and inhibitor concentration need to be optimized, as many protein crystals are sensitive to the solvents used to dissolve the ligands. Although for other proteins ligand soaking is successfully applied, for SBPs this method is not very commonly used as reflected by the low number of structures solved using this method. This is likely due to the fact that upon ***substrate*** binding the two domains undergo a relative large conformational change. Since crystal contacts are fragile and are disrupted easily, large confomational changes induced by soaking can damage crystal contacts resulting either in a massive drop in the resolution of the diffraction or the crystals crack/dissolve completely.

3) Seeding - a method to obtain the ligand bound state with unusual substrates

In some cases the ***ligand*** used for crystallization cannot be crystallized in a closed conformation. This occurs for example when the ligand is not stable during the time of crystallization. One such example is acetylcholine. During

crystallization of the choline binding protein ChoX from *S. meliloti*, it became evident that besides the natural ligand choline also actylcholine is bound by this SBP [8]. To understand the binding properties of ChoX, a structure determination of ChoX in complex with acetylcholine was undertaken. For this purpose the protein was subjected to ***co-crystallization*** experiments. Acetylcholine presents a chemical compound, which is easily susceptible to hydrolysis especially at non-neutral pH values. Although the crystallization of ChoX was done at low pH values, a co-crystallization with intact acetylcholine was achieved. However, subsequent structural determination showed that the ***substrate*** was hydrolyzed to choline in the setup during the time of crystal growth. To overcome this limitation, a ***micro seeding*** strategy was devised. The application of micro seeding helped to crystallize ChoX complexed with acetylcholine within 24 hours. Structural analysis revealed that acetylcholine was not hydrolyzed in the drop during this short period of time required for crystal growth. Thereby, it was possible to solve the structure of ChoX in complex with acetylcholine. The quality of the crystals was good, resulting in diffraction up to 1.8 Å [28]. However, one drawback encountered, when crystals of ChoX were obtained by seeding, was that they all showed a high twinning fraction (up to 50 %). This effect is possibly due to the rapid growth process where crystals reach their final size within a day allowing the formation of merohedral twins, a phenomenon one havs to take into account when using the streak seeding method.

4) Endogenously bound ligands

During purification of some proteins with high ***affinity*** for their ***substrate*** often the ligand is co-purified. Here, OppA from *L. lactis* is an excellent example. OppA belongs to peptide binding subgroup of the family of SBPs and is involved in nutrient uptake in prokaryotes and binds peptides of lengths from 4 to at least 35 residues and with no obvious specificity for a certain peptide sequence. These peptides bind so tightly that they remain associated with the protein throughout purification. The crystallization of the closed-liganded state therefor is relatively easy since the protein will stay only in the closed-liganded conformation. This results in a liganded bound structure. To obtain more different states of the protein one has to remove the ligand first, and afterwards add the wanted substrate. In the case of OppA the peptide was removed prior to crystallization and incubated either with a different ligand or no ligand to obtain a ligand free structure. In the case of OppA, the endogenous peptides can be removed from the protein only by partly unfolding using guanidium chloride, which generates ligand-free OppA. This removal of endogenous peptides was required to allow the binding of defined peptides which was used for crystallization. By this tour de force Bertnsson et al. were able to solve several structures with different ligands bound as well as a ligand free structure, explaining the substrate binding specificitiy of this protein in great detail [22].

2.3 Crystallization of the closed-unliganded state (state III)

The intermediate states of SBPs have been crystallized as well, although only a couple of structures have been reported. This energetically unfavorable state has been crystallized not on purpose in most cases. The choline binding protein ChoX from *S. melioti* has been crystallized in the absence of a ***ligand*** via micro seeding to gain structural insights into the open, ligand-free form of this binding protein. These attempts were not successful. Instead, the obtained crystals revealed a closed but ligand-free form of the ChoX protein. Nevertheless many structures are known of ***substrate*** binding proteins in either their unliganded-open or liganded-closed states [15].

2.4 Crystallization of a semi-open or semi-closed state (state IV)

During our efforts to solve the crystal structure of the choline-binding protein ChoX from *S. meliloti* we used the technique of ***micro seeding*** [15] to obtain ChoX crystals in the ligand-free form. To our surprise, a ligand-free structure, which was different from those that were expected for the ligand-free closed and/or open forms of SBPs described so far, was obtained. Here, ChoX was present in a ligand-free form whose overall fold was identical to the closed-unliganded structure. This structure however, represented a more open state of the ***substrate*** binding protein, which had not been observed before. From the crystal parameters such as the dimensions of the unit cell is was already obvious that the conformation of the protein had changed, since one axis of the unit cell appeared to be significantly larger (35 Å) when compared to the unliganded-closed crystal form of ChoX.

The structure revealed that the domain closure upon ***substrate*** binding does not occur in one step. Rather, a small subdomain in one of the two lobes is laid open and closes only after the substrate is bound. This observation was in line with data observed for the maltose importer system MalFEGK. Here it was observed that the ATPase activity of the ABC transporter was not stimulated by the maltose substrate binding protein when it was added in the unliganded-closed conformation. This is likely due to the fact that the subdomain is not fully closed and rotated outward, which does not activate the transporter. Thus, this biochemical phenomenon could only be explained by the semi-open/semi-closed structure of ChoX [14].

2.5 State I-IV: What do they tell about conformational changes

Substrate binding proteins are flexible proteins, which consist of two domains, which constantly fluctuate between several states of which the open and fully closed state are the most populated ones. Both domains together build up a deep cleft, which harbors the substrate-binding site. As described above the structural work on these proteins has been successful and in the next part a general outcome will be given of what these different states actually tell us about function and mode of action of this protein family.

The unliganded substrate binding proteins are thought to fluctuate between the open and closed state. The angle of opening varies between 26° up to 70° as observed in several open-unliganded structures, suggesting that the extent of opening is likely influenced by crystal packing. This has been observed very nicely for the ribose binding protein of which three different crystal structures have been described. Here the opening of the two domains varies between 43° and 63°. This suggests that the opening can be described as a pure hinge motion. The variation of the degree of opening has been elucidated by NMR in solution for the maltose binding protein MalF. Here 95 % of the protein adopts an open conformation fluctuating around one state with different degrees of opening.

Open and closed - an overall structure view

As an example for the closing movement observed when comparing the open-unliganded and closed-liganded structure the glycine betaine (GB) binding protein ProX from *A. fulgidus* is highlighted in more detail. ProX has been crystallized in different conformations: a liganded-closed conformation in complex either with GB or PB (proline betaine) as well as in an unliganded-open conformation [23]. From the crystallographic parameters it was already anticipated that the crystal obtained differ in the conformation of the protein. ProX crystals were grown using the ***vapor diffusion*** method. The authors attained four different crystal forms depending on the presence or absence of the ligand (hint 1). Liganded ProX crystallized in ***hanging drops*** using a reservoir solution containing 0.2 M $ZnAc_2$, 0.1 M sodium cacodylate, pH 6.0-6.5, 10-12 % (w/v) PEG 4000 and they belonged to the space group $P2_1$ (crystal form I). In a different setup, liganded ProX crystallized in ***sitting drops*** equilibrated against a reservoir containing 30 % (w/v) PEG 1500 and belong to the space group $P4_32_12$ (crystal form II). Unliganded ProX crystallized in hanging drops against a reservoir solution containing 0.3 M $MgCl_2$, 0.1 M Tris, pH 7.0-9.0, 35 % (w/v) PEG 4000. The first crystals appeared after 2-3 months, and belong to space group C2 (crystal form III). Again using a different setup, unliganded ProX crystallized in hanging drops equilibrated against a reservoir containing 0.1 M $ZnAc_2$, 0.1 M MES, pH 6.5, 25-30 % (v/v) ethylene glycol. These crystals grew within 4 weeks, reached a final size of 200 × 150 × 20 μm^3, and belong to space group $P2_12_12_1$ (crystal form IV) [23]. Thus, the different crystallization conditions as well as space group already suggested that several different conformations had been crystallized. Initial phases were obtained by two-wavelength anomalous dispersion of ProX-PB crystals of form IV. All other structures were determined by molecular replacement.

In Figure 2 the opening and closing of the glycine betaine binding protein ProX from *A. fulgidus* is highlighted. Here domain II was taken as an anchor point.

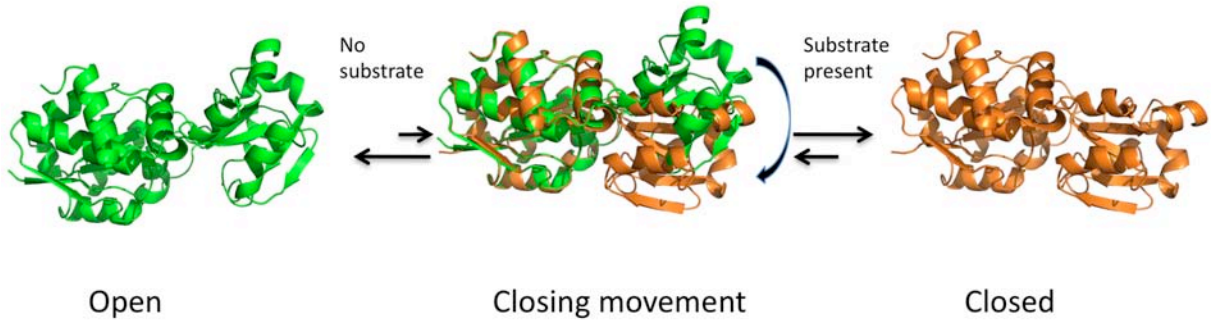


Figure 2. Equilibrium between the open and closed states of substrate binding proteins (ProX from *A. fulgidus)*. The unliganded structure (highlighted in green) of an SBP is fluctuating between the open and closed state (highlighted in orange). In the absence of substrate this equilibrium is pointing towards the open conformation. In the presence of the substrate this equilibrium is changed towards the closed conformation. Here the two domains are close together and side chains of both domains bury the substrate in a deep cleft in between them. (PDB entries: 1SW2, 1SW5). All Figures containing structures were prepared with pymol ("www.pymol.org").

Figure 2 highlights the open conformation (green), which is in equilibrium with the closed state although only a small percentage will be present in the closed unliganded state. Upon the addition of glycine betaine a stable closed conformation is reached and the equilibrium is shifted towards this state. Besides the crystal structure of the ***substrate*** bound state with glycine betaine, proline betaine and betaine as a substrate also the open conformation was crystallized. This allowed a detailed analysis of the closing and opening motion mediated by the hinge region between both domains. The comparison of the ligand-free and liganded conformation of other binding proteins showed an approximate rigid body motion of the two domains highlighting a total rotation of domain II by ~ 58° with respect to domain I (Figure 2). The total rotation has two components: 1) the hinge angle between the two domains of ~ 40° with its axis going through the above-mentioned hinges in the polypeptide and 2) a rotation perpendicular to the hinge axis of ~ 42°. Although the domains behave more or less as rigid bodies, there are a few changes of the ***binary complex*** in two regions of ProX. If one succeeds in crystallizing several conformation of a protein one can search for and visualize small distinct changes in the overall structure. This has been also observed in ProX, the α-helical conformation (in the open form) of residues 144–148 (domain I) change either to an isolated-β-bridge or to a turn conformation (in the closed form). This conformational change may be caused by the proximity to Arg149, which plays an important role in ligand binding as

discussed below. Furthermore, residues 222–225 (domain II), which are in turn and 3_{10}-helix conformation (in the open form), become rearranged to a short α-helix in the closed form. These structural changes highlight an important point in the function of such a protein (more detail below).

Open and closed - an active site view

A closer look at the binding site or the amino acids involved in ***substrate*** binding shows that small but distinct conformational changes of the amino acids involved in ligand binding occur upon substrate binding. Again as an example the glycine-betaine binding protein ProX from *A. fulgidus* is used.

The binding site is located in the cleft between domains I and II and can be subdivided into two parts, one binding the quaternary ammonium head group and the other binding the carboxylic tail of these compounds. The quaternary ammonium head group is captured in a box formed by Asp109 and the four tyrosine residues Tyr63, Tyr111, Tyr190, and Tyr214 being oriented almost perpendicular to each other. The tyrosine side chains provide a negative surface potential that is complementary to the cationic quaternary ammonium head group of GB. The carboxylic tail of GB is pointing outward of this partially negatively charged environment forming interactions with Lys13 (domain I), Arg149 (domain II), and Thr66 (domain I), respectively. Furthermore the structure was solved at a resolution sufficient to locate water molecules. An important water molecule was observed, which was held in place by residues Tyr111 and Glu145, and stabilizes domain closure. Here it is important to mention that this water molecule was not observed in the open unliganded structure and its importance would therefore be easily overlooked when no comparison between the two states were possible.

The superposition of the open-unliganded form and the closed-liganded form of ProX allowed an unambigouos identification of residues of domain II that are involved in ligand binding. They show virtually the same orientation in the open and closed forms (see Figure 3). Residues Tyr63, Tyr214, Lys13, and Thr66 superimpose very well. Only the main chain carbonyl of Asp109 from domain I is slightly out of place compared to the closed form because of the enormous main chain rearrangement between Asp110 and Tyr111 upon domain closure. The residues contributed by domain II behave quite differently. Tyr111 and Tyr190 are not only moved as parts of domain B but they undergo a major conformational change to adopt the conformation of the closed-liganded binding site. The side chain conformation of Arg149 shows only small changes between the open and closed conformations although it undergoes a large movement as part of domain II.

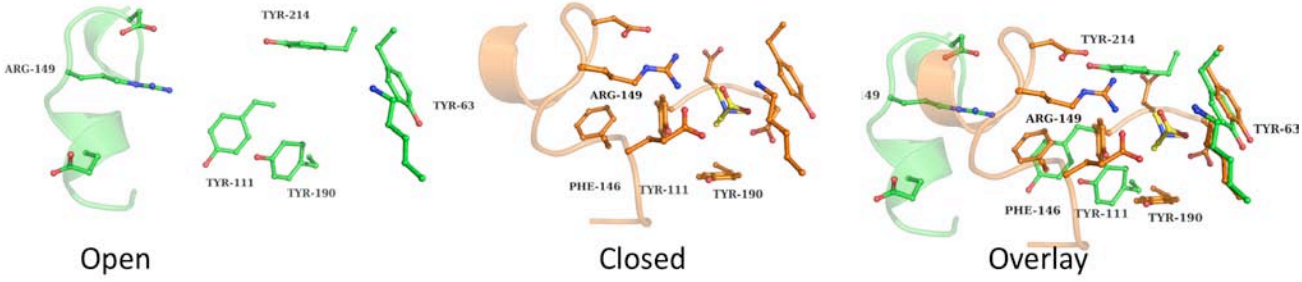


Figure 3. The binding site of ProX is highlighted in the open (depicted in ball and stick in green-left picture) and the substrate bound closed conformation (depicted in ball and stick in orange middle picture). As observed some of the ligand binding amino acid change their conformation. The right picture shows an overlay of both structures to visualize these conformation changes (PDB entries: 1SW2, 1SW5 and 3MAM).

Recently, another structure of ProX was solved in the liganded but open conformation [29]. This conformation represents a state of which only very few crystal structure are known. In other words, the protein has a ligand bound and is on its way to close up the binding site. This structure provided an even more detailed picture on the function of ProX and finally highlighted the crucial role of Arg149. In addition to the direct interaction with GB and residues that are part of the substrate-binding pocket (Tyr111, Thr66), Arg149 is a major determinant in domain-domain interactions in the closed structure. As such, Arg149 interacts with Val70 (domain I) and Asp151 (domain II), thereby acting as a linking element between the two domains enforcing stable domain closure. These interactions complement those mediated by Pro172 of domain II, where Pro172 interacts via its Cα-atom and a water molecule with Glu155 of domain II. Together, this provides a further explanation for the crucial role of Arg149 for the stability of the liganded-closed state, which has been observed in mutagenis studies. Here, the binding ***affinity*** of GB was dramatically lowered when Arg149 was mutated to alanine, a phenomenon that could not be explained since the aromatic cage which dominates the binding affinity was still present to bind glycine betaine. This suggested that Arg149 is the final amino acid to interact with the ***substrate*** and, thereby, terminate the molecular motions that result in the high affinity closed state of ProX. Besides this crucial role of switching from a low affinity to a high affinity state via the interaction of Arg149 the open liganded structure also shed light on the movement that the amino acids undergo during closure of the protein. In the open-liganded structure the presence of glycine betaine is communicated to Arg149 through interactions of the side chains of Tyr190, Tyr111, and Phe146 via a

side-chain network [29]. Interestingly when comparing the open and closed structures of other SBPs, the maltose binding protein (MBP) [9] and the ribose binding protein (RBP) from *E.coli* and the N-Acetyl-5-neuraminic acid binding protein (SIAP) from *H. influenza* [25] a similar network can be identified in these proteins, something which had not been identified before due to the lack of an open-liganded structure.

In summary, the "Venus fly trap" model describes the opening and closing of SBPs. Here the equilibrium between these two conformations is shifted towards the closed state upon ***substrate*** binding. Many crystal structures of SBPs have been solved in the unliganded-open, liganded-closed, and, more rarely, in the liganded-open or unliganded-closed state [3, 14, 15, 23]. The crystal structure of one of these states will give information on the overall structure of the protein as well as the ***ligand*** binding site. Several SBPs have been crystallized in two or more states and quite clearly the increasing amount of states will shed a more detailed look on how domain closure is occurring. Thus, although crystallization is trial and error and sometimes tedious, it is worth to search for crystals in the liganded-closed conformation as well as crystals of another state since every stage will visualize the cleverness of nature to use conformational changes for the formation of ligand binding sites.

3. Protein with multiple ligands – how to crystallize the different ligand bound intermediate states

Besides proteins that bind one substrate, a large number of enzymes are binding two or more substrates and convert these into a product. Here, the crystallization of the ***apo-enzyme*** often reveals the binding site of these ligands. However, the exact influence of the binding of these ligands can only be deduced from several structures, where different ligands are bound or one structure with all ligands bound. The different states are called apo-enzyme, when the enzyme is depleted of all ligands, the ***binary complex*** when the first substrate is bound, the ***ternary complex*** when the second ligand is bound as well. A quaternary complex would describe the protein with three ligands bound.

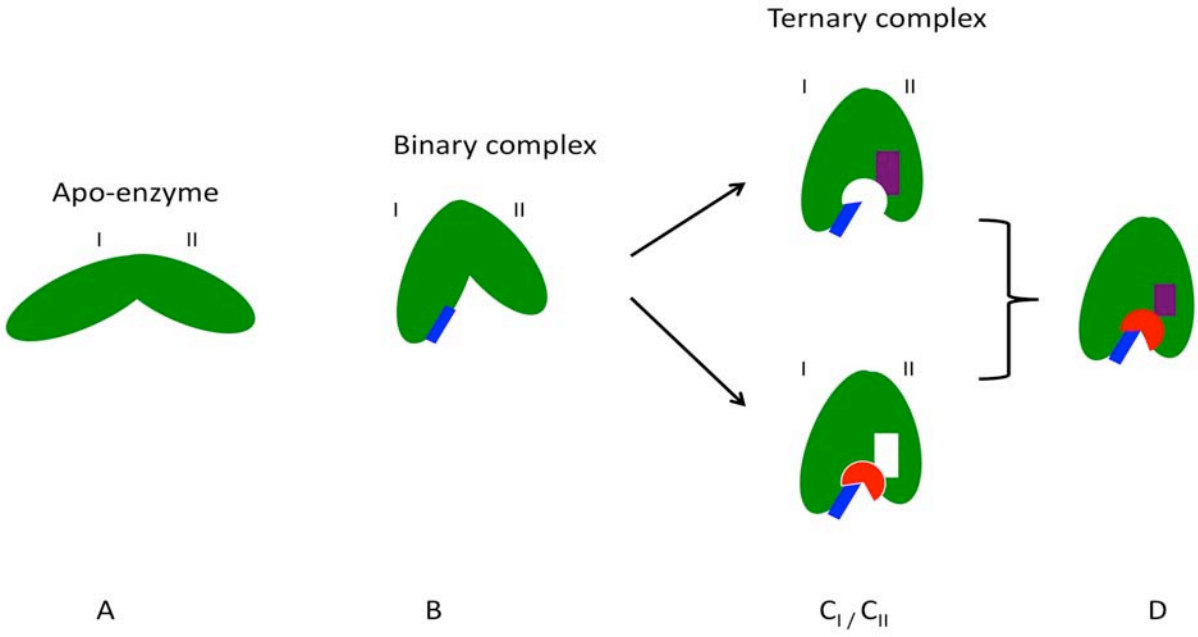


Figure 4. Overview of the conformations a protein can adopt with multiple ligands. A) The apo-enzyme B) binary complex where the first ligand is bound. This ligand with the highest affinity induces a stable conformation of the enzyme which allows the binding of the second ligand (ternary complex C_I or C_{II}). D) Enzyme complex where all ligands are bound.

Most of these proteins are enzymes. In reactions mediated by enzymes, the molecules at the beginning of the process, called ***substrates***, are converted into different molecules, called products. Almost all chemical reactions in a biological cell need enzymes in order to occur at rates sufficient for life. Since enzymes are selective for their substrates and speed up only a few reactions from among many possibilities, the set of enzymes synthesized in a cell determines which metabolic pathways are utilized. Obtaining a snapshot of the substrate bound enzyme is difficult, because the enzymatic reaction will proceed immediately after substrate binding. One "trick" mostly used to solve this problem is to inhibit the reaction by either the reaction condition, meaning by varying pH of the buffer to a value where the reaction is not occurring. Another approach often appied in crystallography is to use a mutant, which cannot catalyze the reaction anymore; however it is still capable of binding the substrate. This has been proven to be successful in many cases. For example the catalytic cycle of nucleotide binding domains has been unraveled by such a mutation. In the latter case the ATP hydrolysis, in the wild type the measure for activity, has been abolished by mutation of a crucial amino acid, which still allowed binding of ATP but prevented hydrolysis. Thereby the dimeric state of the protein was stabilized and the active form of the NBD could be crystallized in the presence of ATP [30-32].

Below the structural studies of the octopine dehydrogenases (OcDH) from *P. maximus* will be described in more detail. This enzyme catalyses the reductive condensation of L-arginine with pyruvate forming octopine under the simultaneous oxidation of NADH. This oxidation of NADH is the terminal step in the anaerobiosis, meaning the generation of ATP when organisms are suffering from low oxygen levels. A prominent member of these terminal pyruvate oxidoreductases is the lactate dehydrogenase, which catalyzes the transfer of a hydride ion from NADH to pyruvate, with produces NAD^+ and lactate. Thereby the redox state in vertebrates is maintained during functional anaerobiosis. OcDH fulfills the same function in the invertebrate *P. maximus*.

This enzyme has been chosen due to the fact that three ***substrates*** need to be bound simultaneously for the reaction, in contrast to the lactate dehydrogenase, which has only two subtrates, NADH and pyruvate. Furthermore this enzyme was crystallized as wildtype protein and in all substrate bound states (***binary*** and ***ternary complex*** C_I and C_{II}) and the corresponding structures were elucidated. The state where all substrates were present did not yield a structure due to the immediate conversion to the product. However, the other structure allowed a detailed view on how the latter state might look like.

In 2007 Mueller and co-workers achieved cloning and heterologously expression of this enzyme using *E. coli* as expression system [33]. After the purification the enzyme was characterized and the authors proposed a sequential binding mode of the ***substrates***. Here, NADH was bound first followed by either L-arginine or pyruvate. The order of the last two was not revealed by the enzymatic analysis. Furthermore, a catalytic triad was proposed consisting of three highly conserved amino acid, building up a protein rely-system for the reduction of NADH. This triad has been observed in the sequence and structure of the lactate dehydrogenase as well. Sequence analysis of different proteins from this family revealed that the protein contained two distinct domains where domain I contained the characteristic Rossmann-fold, a domain responsible for the binding of NADH. Domain II was assigned as octopine dehydrogenase domain, which is specific for this protein family and was suggested to contain the binding site for both L-arginine and pyruvate. Both domains are connected via a linker region of 5-8 amino acids suggesting that these domains might undergo large conformational changes.

3.1 The crystallization of apo-enzyme and the binary complex

Parallel to the biochemical characterization, the crystallization of the enzyme was started. Due to the two-domain structure OcDH can adopt multiple conformations in solution, which prevents crystal formation. However, purified OcDH-His_5 yielded small crystals that appeared to be multiple on optical examinations (Figure 5 A). They diffracted to a resolution of 2.6 Å. However the diffraction showed multiple lattices in one diffraction image and could not be used for structure determination (Figure 5 A) [34]. All attempts to improve these crystals using for example ***seeding***, temperature ramping or various crystallization conditions failed. Finally, the primary ***ligand***, NADH, was added prior to crystallization. This produced crystals under conditions similar to those in the absence of NADH. Here, the incubation temperature appeared to be critical and needs to be kept at 285 K. The crystals obtained were single and diffracted to 2.1 Å resolution, which allowed processing of the data and subsequent structure determination (Figure 5 B). The structure of OcDH was solved as binary complex with NADH [34, 35].

Cofactors like NADH are often observed to be co-purified. This was assumed to be the case for OcDH as well, however, no activity was ever observed without NADH, but in the presence of the other two substrates. This implies that OcDH is not homogenous and multiple conformations exist as observed in the multiple crystal lattices of the diffraction image. This is in line with the only other available three-dimensional structure of an enzyme of the OcDH superfamily, the apo-form of N-(1-D-carboxylethyl)-L-norvaline dehydrogenase (CENDH) from *Arthrobacter sp.* strain 1C [36]. CENDH catalyzes the NADH-dependent reductive condensation of hydrophobic L-amino acids such as L-methionine, L-isoleucine, L-valine, L-phenylalanine or L-leucine with α-keto acids such as pyruvate, glyoxylate, α-ketobutyrate or oxaloacetate with (D, L) specificity [37]. The structure of the ***binary complex*** of CENDH with NAD^+ was determined to a resolution of 2.6 Å. Although NAD^+ was added in the crystallization trials the cofactor could not be observed unambiguously in the electron density. This was likely due to the concentration of NAD^+, which was below the K_d. As a result not all proteins had the ***substrate*** bound, which led to a not very well defnied electron density. Only the nicotinamide ribose moiety was of moderate quality and the density of the nicotinamide ring was very weak. This has been attributed to low NAD^+ ***occupancy*** in this crystal, hence the ***co-factor*** has been omitted from the high resolution refinement [36].
This highlights the importance to verify the ***affinity*** of substrate prior to crystallization. Since NAD^+ is the product of the reaction and to ensure the release of the product, the affinity of NAD^+ must be lower than the affinity of NADH. In a recent study on the OcDH the affinities have been determined to be 18 µM for NADH and 200 µM for NAD^+ [38]. As described above the additions of substrate in crystallization trials need to be at least a 10-fold above the K_d. For OcDH 0.8 mM NADH was used for the crystallization of the binary complex, which represents a 40-fold excess.

The structure of the OcDH-NADH ***binary complex*** revealed why the initial crystallization attempt of the ***apo-enzyme*** failed. NADH is bound by the Rossmann-fold located in domain I as well as by an arginine residue in domain II. Thereby the OcDH captured in a state which enables the binding of the other ***substrates***, pyruvate and L-arginine (see below) [34, 35].

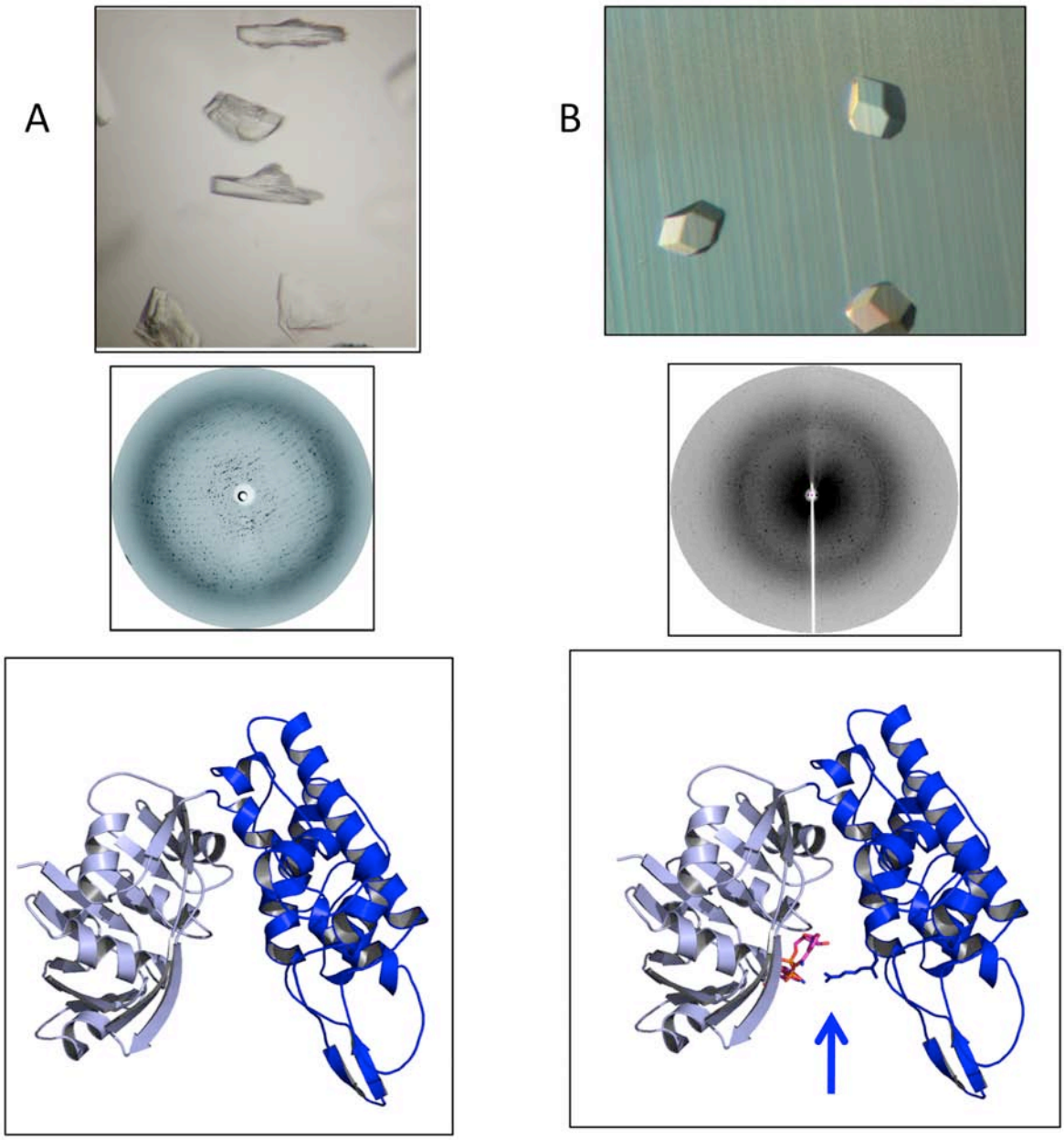


Figure 5. Crystallization of OcDH in the absence and presence of NADH. A) Absence of NADH. The crystals obtained are multiple (upper panel) and the diffraction pattern yielded showed several lattices (middle panel). The structure of OcDH shows two distinct domains connected by a flexible linker, which can rotate freely in the absence of NADH (lower panel). B) Crystals obtained in the presence of NADH (upper panel). The diffraction showed a single lattice diffracting up to 2.1 Å (middle panel). The structure revealed the binding site of NADH and an interaction of an arginine residue from domain II with NADH, which locks OcDH in one stable conformation (lower panel) (PDB entries: 3C7A and 3C7D).

In summary, the apo-state of multiple ***ligands*** binding enzymes is difficult to crystallize when the enzyme undergoes large conformation changes. In the case of the OcDH only the binary complex in the presence of NADH could be crystallize. Here crystals were of sufficient quality to determine the structure.

3.2 The crystallization of the ternary complexes C_I and C_{II}

OcDH catalyzes the condensation of L-arginine with pyruvate to form octopine under the oxidation of NADH. Biochemical analysis as well as the crystal structure revealed that NADH is the first ***substrate*** to bind to OcDH. The structure of this ***binary complex*** exhibited a stable conformation of the protein in solution with an Arg-sensor, which binds NADH, and thereby stabilizes the protein in one conformation (see above).

So the next step was to determine the structure of the OcDH in the presence of the second and third ***substrate***, L-arginine and pyruvate, respectively. Initially, the protein and the substrate were mixed and an extensive search for suitable crystallization conditions was started. However, no crystals were obtained for OcDH in the presence of L-arginine and/or pyruvate. Instead only needles were grown which were multiple and very fragile similar to the crystals obtained for the ***apo-enzyme***. This is in line with the biochemical data, which highlights the order of substate binding which show that NADH has to be bound prior to binding of L-arginine as well as pyruvate [38, 39]. Here the authors used two other techniques, NMR and ***ITC*** repectively, to show that L-arginine only binds after saturation of the ***apo-enzyme*** with NADH. Pyruvate was shown to be bound only after L-arginine binding to the enzyme. This suggests that OcDH undergoes a conformational change when NADH is bound and thereby the binding site of L-arginine is formed. Furthermore the binding site for pyruvate is only created when L-arginine is bound.

Since crystallization was not successful the next step was to use ***co-crystallization*** with the OcDH protein and L-arginine and/or pyruvate to obtain structural information of the ***ternary complex*** (C_I and C_{II}). This yielded crystals of OcDH only in the presence of NADH and no additional density was observed for neither L-arginine nor pyruvate. So, ***soaking*** the ***ligand*** into preformed OcDH-NADH crystals was the last method chosen. Crystallization trials were carried out using the ***hanging-drop vapor diffusion method*** and crystals of OcDH were grown in the presence of 0.8 mM NADH. L-arginine-bound crystals were obtained by soaking NADH bound OcDH crystals in 100 mM MES pH 7.0, 1.15 M Na-citrate, 0.8 mM NADH containing 10 mM L-arginine for at least 24 hours. Pyruvate-bound crystals were obtained also by soaking the crystals in 100 mM MES pH 7.0, 1.15 M Na-citrate, 0.8 mM NADH and 10 mM pyruvate for at least 8 hours. Both concentrations were chosen relatively high but they resemble the *in vivo* concentration as well as were backed up by the ***affinity*** observed for both ***substrates*** in biochemical and biophysical studies, being 5.5 mM L-arginine and 3.5 mM pyruvate, respectively. During soaking a cracking of the crystals was observed after the first minutes. However, the crystals recovered completely from this cracking within the following hours and showed no fissures or other damages after that soaking procedure. Desprite this, the diffraction analysis revealed a loss in diffraction. Initally the crystals diffracted to 2.1 Å. After soaking in L-arginine or pyruvate the diffraction potential was reduced to 3.0 Å and 2.6 Å, respectively. The phenomenon of crystal cracking and decline of the diffraction already was a good indication that the substrates diffused into the crystal. A dataset was collected from crystals where either one of the ligands was soaked in and besides the decrease in diffraction potential also the unit cell parameters changed (see Table 2).

Table 2: Crystallographic parameters of the unit cell of the binary OcDH-NADH complex and after soaking of the ternary complex C_I: OcDH-NADH/L-arginine and C_{II}: OcDH- NADH/pyruvate

Crystal Complex	Unit cell parameters (a,b,c in Å)
OcDH-NADH	99.8, 99.8, 126.5
OcDH-NADH/L-arginine	95.9, 95.9, 117.9
OcDH-NADH/pyruvate	95.0, 95.0, 120.2

The change in unit cell parameters suggested that a conformational change occurred during the ***soaking*** with the ***ligand***. This was further observed after the structure was resolved and electron density was clearly defined for L-arginine in the one and for pyruvate in the other dataset. The structure of OcDH-NADH/L-arginine showed a rotational movement of domain II towards the NADH binding domain I, and a stronger interaction of the Arginine residue with NADH. A domain closure was also observed in the pyruvate bound structure. So stable binding of NADH to the Rossman fold of domain I, the first step in the reaction sequence of OcDH, occurs without participation of domain II. A comparison of the OcDH-NADH (colored light-purple in Figure 6) and the OcDH-NADH/L-arginine complexes revealed a 42° rotation of domain II towards the NADH binding domain (domain I) in the latter complex. This domain closure is triggered by the interaction of Arg324 (domain II) with the pyrophosphate moiety of NADH bound to the Rossman fold in domain I.
A comparison of the two ***ternary complexes*** suggests that both, pyruvate and L-arginine, are capable to trigger domain closure to a similar extent. However, in the OcDH-NADH/pyruvate complex, pyruvate partially blocks the entrance for L-arginine, while in the OcDH-NADH/L-arginine complex, the accessibility of the pyruvate binding site is not restricted by L-arginine [34, 35]. From these structures it could be deduced that L-arginine binds to the OcDH-NADH complex in a consecutive step and induces a rotational movement of domain II towards domain I. This semi-closed active center, which is further stabilized using the pyrophosphate moiety of the bound NADH and by interactions of L-arginine with residues from both domains is then poised to accept pyruvate and consequently the product octopine can be formed. With regard to the structures it was proposed that instead of a random binding process, an ordered sequence of substrate binding in the line of NADH, L-arginine and pyruvate will occur.

This ordered sequence of ***substrate*** binding was then biochemically proven by ***ITC*** studys where the binding affinities of the substrates were measured. Here, the binding of L-arginine was only observed when NADH was bound primarily and the binding of pyruvate only when the complex was preloaded with L-arginine [38, 39]. Furthermore this ordered binding mechanism explains why no lactate is found in side *P. maximus* which is normally formed when NADH and pyruvate is bound by lactate dehydrogenses. Here, it is worth mentioning that the conformational changes induced by ***ligands***

soaking into the crystal were also observed in NMR studies that were perfomed in solution. So the apparent conformational changes in the crystal resemble the changes the protein undergoes in solution.

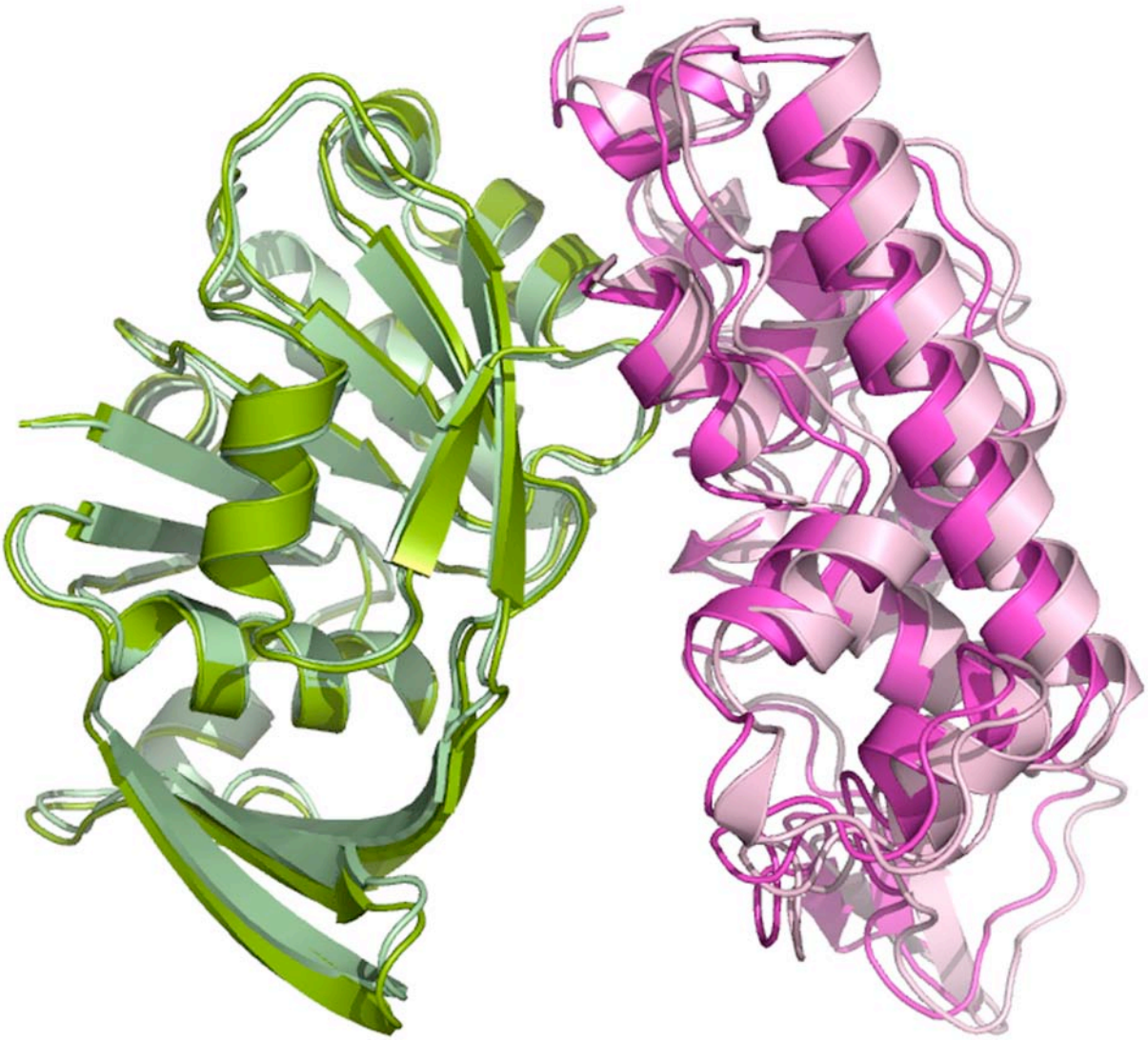


Figure 6: Overlay of the OcDH-NADH binary complex with the OcDH-NADH/L-arginine ternary complex C_I. As seen in the superposition the binding of L-arginine induces a conformational change. Domain II is rotated towards domain I which is thereby creating the pyruvate binding site. In the overlay the pyruvate structure is not shown due to clarity (PDB entries: 3C7A and 3C7D).

The crystal structures of the different states of OcDH, delivered snapshots elucidating for the first time the precise and very distinct binding order [35]. Unfortunally the crystals with the endproduct octopine did not diffract X-ray with a resolution and quality high enough for structure determination. The same hold true for a complex with all three substrates present at once. This is likely due to the fact that the immediate condensation occured and the product was formed.

To show how proteins can be crystallized with their enzymatic endproducts we chose another enzyme family as example and will describe the different procedures during the next paragraphs.

4. Enzymatic products in protein structure – how to crystallize this rather unfavored states

The state found to be important within an enzyme reaction cycle is supposedly the product bound state. After the reaction occurs the product is still sitting within the protein and will be released. Often these product have a low(er) ***affinity*** to the protein than the ***substrates*** and are therefor less often found to be successfully crystallized.

Examples of prosperous structure determination however are the shikimate dehydrogenase (SDH or AroE) of *Thermus thermophilus* (*Tth*SDH), *Aquifex aeolicus* (*Aae*SDH) and the recently deposited structures of the SDH of *Helicobacter pylori* (*Hpy*SDH) as well as the bifunctional dehydroquinase-shikimate dehydrogenase (*Ath*DHQ-SDH) from *Arabidopis thaliana* which were crystallized with its reaction product shikimic acid ([40-43]. Similar to that the closely related quinate dehydrogenase of *Corynebacterium glutamicum* (*Cgl*QDH) was structurally characterized in four different states: as ***apo-enzyme*** and at atomic resolution with bound ***cofactor*** NAD^+ as well as in complex with quinic acid (QA) and the reduced cofactor or shikimic acid (SA) and NADH [44].

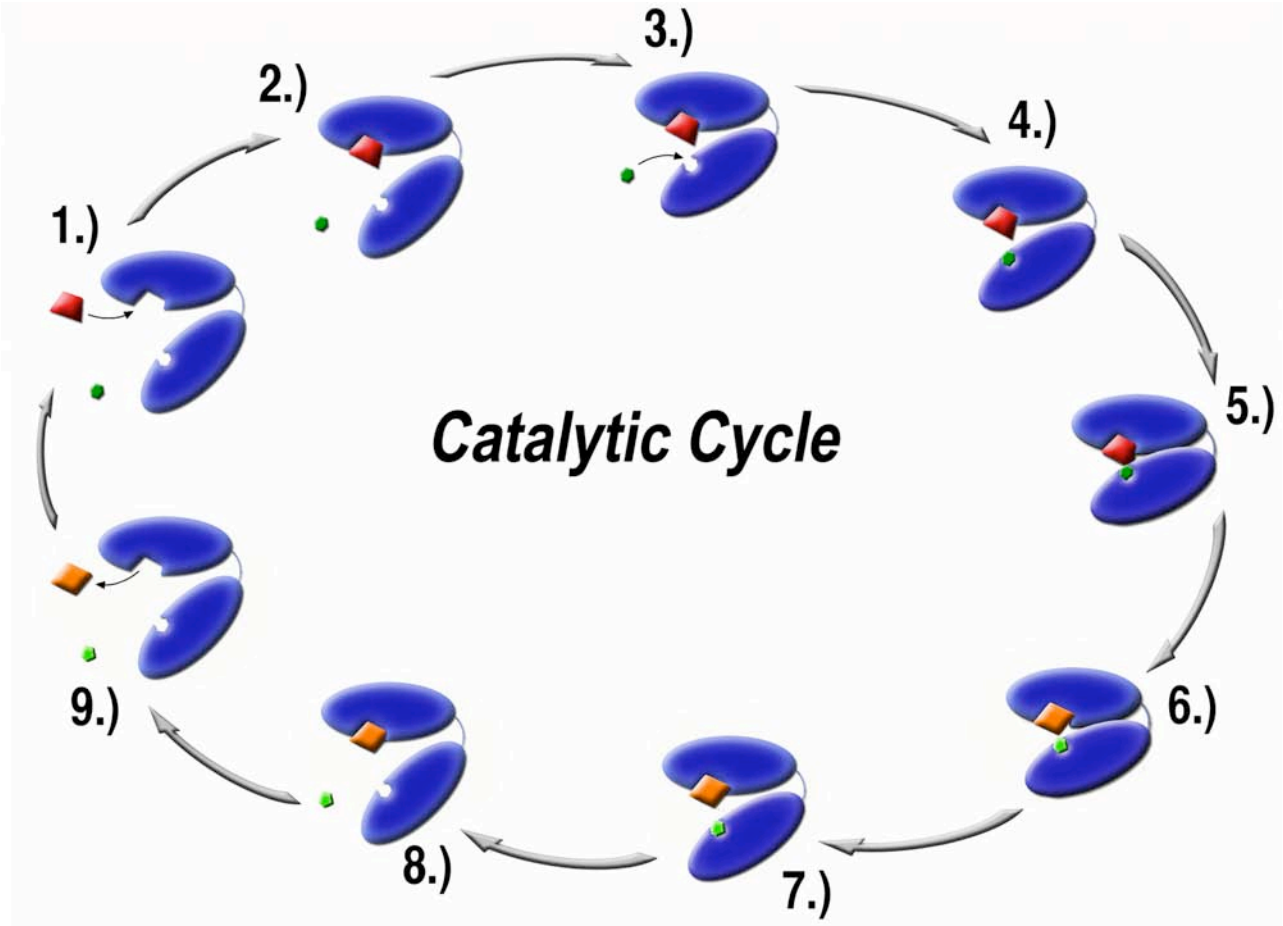


Figure 7. Schematic diagram of the conformational changes within a protein (blue ellipses) during the catalyzed reaction. 1.) Before a substrate (red trapezium) is bound the proteins exhibits an open conformation. 2.) – 4.) Binding of the substrate induces a slight domain closure before the cofactor (green hexagon) is bound. 5.) + 6.) In order to facilitate the conversion from substrate to the product (orange rhombus) both protein domains need to be in close contact. 7.) -9.) A stepwise domain opening allows the changed cofactor (light green pentagon) to leave the protein domain, followed by the product. The protein itself is not modified at all during the whole reaction and that mostly all steps are reversible.

Shikimate-/quinate dehydrogenases belong to the superfamily of NAD(P)-dependent oxidoreductases whereas SDHs catalyse the reversible reduction of 3-dehydroshikimate to shikimate under oxidation of NAD(P)H and QDHs the oxidation of quinate to 5-dehydroquinate with reduction of NAD(P), respectively. The overall fold consists of a N-terminal or ***substrate*** binding domain and a C-terminal or cofactor-binding domain and is highly conserved within that subfamily (schematically shown in Figure 7). Compared to other proteins, like the above-mentioned SBPs, the structural changes occurring during catalysis are less prominent and comprise a movement of the two domains against each other in a range of several Ångstrom.

4.1 Shikimate dehydrogenase from *Aquifex aeolicus*: Crystals of the native (apo-) *Aae*SDH were obtained with non-His-tagged protein, whereas the ***ternary complex*** crystals were obtained with His-tagged SDH. To get these complexes the protein solution was mixed with ***substrate*** and cofactor (i. e. with both natural products) to final concentrations of 5.0 mM shikimic acid and 5.0 mM $NADP^+$ before crystallization. The hanging-drop vapor diffusion method was used for crystallization trials. The drops were prepared by mixing 3 µl of the protein-ligand solution with 1 µl of well solution [41].

K_M values were determined to be 42.4 µM for both ligands, which means that there was a 100-fold excess in the crystallization drop. The bound products SA and $NADP^+$ in the protein could be explained by the low activity of the enzyme and the equilibrium constant favoring the formation of SA and $NADP^+$, both of which are caused by the low pH. The equilibrium constant ([SA][$NADP^+$]/[DHSA][NADPH]) was determined by Yaniv and Gilvarg (1955) to be 27.7 at pH 7 and 5.7 at pH 7.8 [45]. As of any dehydrogenase reaction, the equilibrium position of the *Aae*SDH-catalysed reaction depends on the hydrogen ion concentration of the environment. The pH of the well solution (0.2 M ammonium acetate, 30 % w/v PEG 4000, 0.1 M sodium acetate) was 4.6 and therefor the drop became more acidic during crystallization. They estimated the equilibrium constant at pH 5 to be around 3000 in favor of the formation of SA and $NADP^+$. The geometry of $NADP^+$ is not distinguishable from that of the NADPH at this resolution (2.2 Å) but the geometry of SA containing a tetrahedral (sp^3) C3 atom is distinct from that of DHSA, in which the geometry of C3 is planar (sp^2) [41].

There were eight (apo) and four (***ternary complex***) crystallographically independent *Aae*SDH molecules in the asymmetric unit of apo-*Aae*SDH or *Aae*SDH-$NADP^+$-SA, respectively. According to the structure of the ***apo-protein*** and the ternary complex a fully open (molecule F in apo-*Aae*SDH) and a closed conformation with bound ligands (molecule D in *Aae*SDH-$NADP^+$-SA; Figure 8) were observed as well as several intermediate states. From the fully open to the closed form there is a movement of three loops in the catalytic domain towards the NADP-binding domain by around 5 Å sealing the active site of the enzyme. SA and $NADP^+$ are brought in close contact in that cavity: the C3-O3 bond of SA is parallel to the C4-C5 bond of the nicotine amide ring of $NADP^+$, and the distance between the two bonds is 3.5 Å. This represents a typical distance for a hydride transfer.
The open conformation therefor represents the protein structure in state 9.) (or 1.), respectively) in Figure 7, the closed conformation correlates to state 6.) in that scheme.

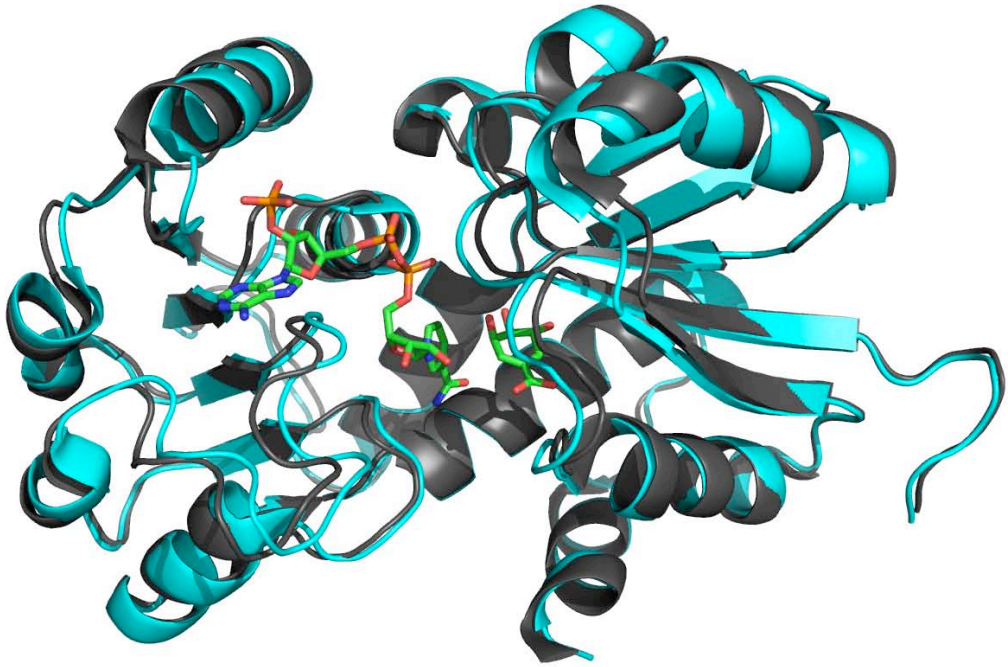


Figure 8. Shikimate dehydrogenase from *Aquifex aeolicus*. The cartoon depicted in cyan represents the open (apo) conformation of the enzyme (PDB entry: 2HK8), the structure coloured in black illustrates the closed conformation (PDB entry: 2hk9) with the bound ligands shikimic acid and $NADP^+$, shown as sticks.

4.2 Shikimate dehydrogenase from *Thermus thermophilus*: In case of *Tth*SDH, crystals of the native protein were grown in ***microbatch*** plates. ***Co-crystallization*** trials were only successful with added $NADP^+$ but failed with shikimate. To obtain complexes with bound shikimate crystals of the ***apo-protein*** or the SDH-$NADP^+$ complex were ***soaked*** for several seconds in cryosolution supplemented with shikimate. The final concentration of all added ***ligands*** was 5 mM. Although the kinetical parameters were not determined prior to crystallization, all K_M values of closely related SDHs are in a µM range so that there was at least a 20-fold excess of ***substrate*** and ***cofactor*** [40].
Evaluation of the complex structures revealed an open and a closed conformation of the two domains but neither the binding of shikimate nor $NADP^+$ seem to induce that conformational change. Shikimate could bind to the closed as well as to the open form, whereas $NADP^+$ was found only in closed conformation. As described for *Aae*SDH, the crystallization condition was in an acid range of about pH 4.6, which explains that the reaction did not occur. An alignment of the three structures (apo-SDH, SDH-SA, SDH-SA-$NADP^+$) of *T. thermophilus* illustrates the domain closure while/after SA and/or $NADP^+$ binding (Figure 9). Surprisingly there seems to be no further movement of the substrate binding domain against the NADP(H) binding domain when the cofactor is bound. Thus, the apo-structure represents

state 9.) (or 1.) in Figure 7 and both the ***binary*** and the ***ternary complex*** may match a state between 6.) and 7.) of that scheme.

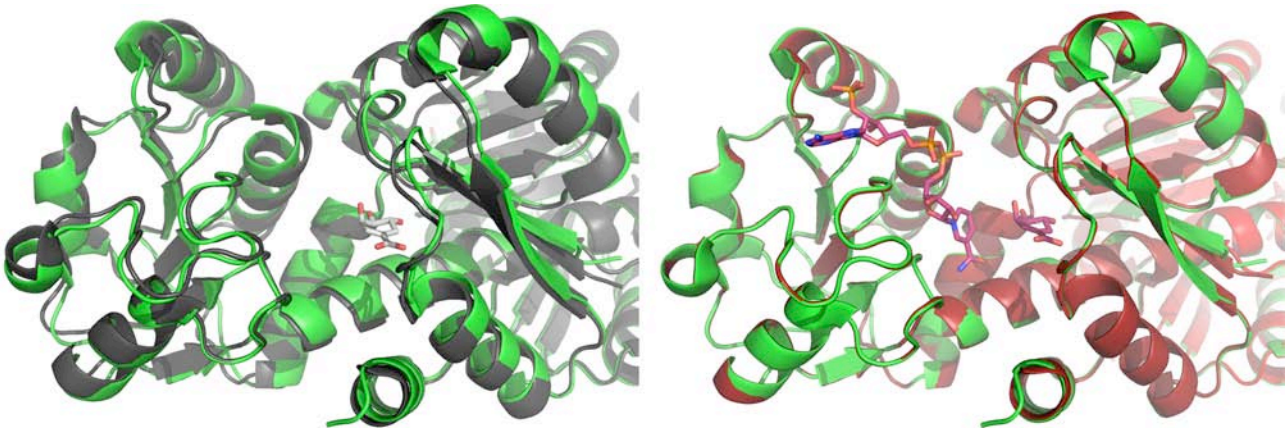


Figure 9. Shikimate dehydrogenase from *T. thermophilus*. The cartoons depicted in green (left and right side) represent the open (apo) conformation of the enzyme (PDB entry: 1WXD), the structure coloured in black illustrates the closed form with bound shikimic acid (PDB entry: 2D5C), whereas the red one corresponds to the ternary complex (PDB entry: 2EV9) with shikimic acid and $NADP^+$, shown as sticks.

4.3 Bifunctional dehydroquinase-shikimate dehydrogenase (*Ath*DHQ-SDH) from *Arabidopis thaliana*: Remarkable are the ***co-crystallization*** trials of Singh and Christendat with the bifunctional enzyme dehydroquinase-shikimate dehydrogenase from *Arabidopsis thaliana* (*Ath*DHQ-SDH). First crystals were obtained with the product shikimate at the SDH site and tartrate as a ***substrate*** analogue at the DHQ site. Later they could crystallize *Ath*DHQ-SDH with its natural products shikimate and $NADP^+$.

For the shikimate-tartrate complex crystals they used the ***vapor diffusion hanging-drop*** technique. Protein solution with a final concentration of 1 mM of shikimate was mixed with the reservoir solution containing 0.4 M potassium sodium tartrate tetrahydrate [42]. To obtain ***ligand*** bound crystals of the three different protein conditions were tested: protein only, protein with 1 mM shikimate or protein with 1 mM $NADP^+$. The protein-shikimate approach was the only one that yielded crystals (under the same conditions as mentioned above). To gain crystals of the ***ternary complex*** a further treatment was necessary: The above-mentioned crystals were soaked with a $NADP^+$ solution (final concentration 10 mM) for about 8 hours at pH 5.8. The K_M values were determined to be 0.6 mM for shikimic acid and 0.13 mM for $NADP^+$, respectively [42].
Not only that the closed conformation of the enzyme after binding of both products could be demonstrated (Figure 10) but also the activity of that ternary complex was proven as the oxidation of shikimate was evidenced by the generation of dehydroshikimate, – the product of the DHQ moiety – found in the DHQ site [43].

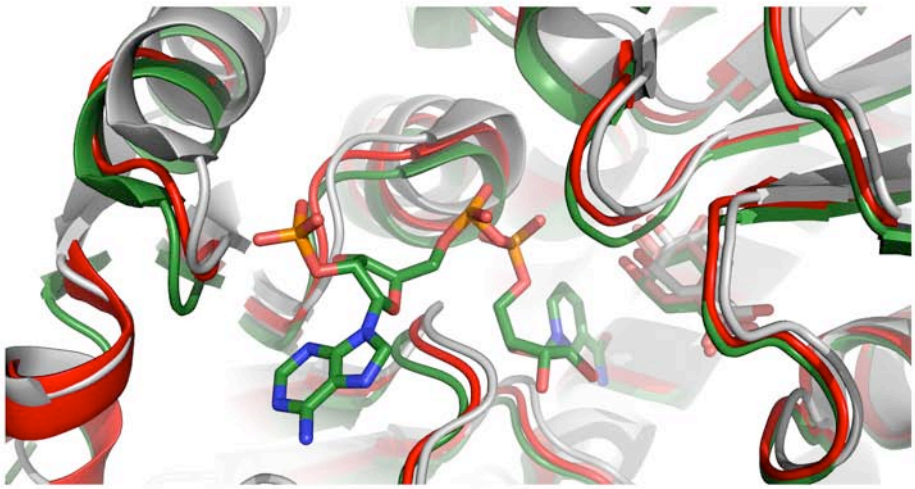


Figure 10. Bifunctional dehydroquinase-shikimate dehydrogenase (*Ath*DHQ-SDH) from *Arabidopis thaliana*. The cartoon coloured in grey reveals the binary complex (PDB entry: 2GPT) with the bound product shikimate (grey lines), the structure depicted in red shows the protein with bound substrate dehydroshikimate (red lines; PDB entry: 2O7Q), whereas the cartoon in green represents the ternary complex (PDB entry: 2O7S) with bound dehydroshikimate and the cofactor NADP(H).

The structures of the *Ath*DHQ-SDH ***binary complexes*** with bound product shikimate or ***substrate*** dehydroshikimate illustrate therefor the states 8.) or 2.), while the ***ternary complex*** corresponds to the transition state 5.) in Figure 7.

4.4 Shikimate dehydrogenase from *Helicobacter pylori*: Recently three different catalytic states of the *Hpy*SDH were deposited in the PDB. Unfortunately the results are not published so that detailed information about the crystallization trials are lacking. Apparently they obtained all crystals by means of the ***hanging-drop*** vapor diffusion method.
However, the structure is ideally suited to visualize structural changes during ***cofactor*** binding Figure 11.

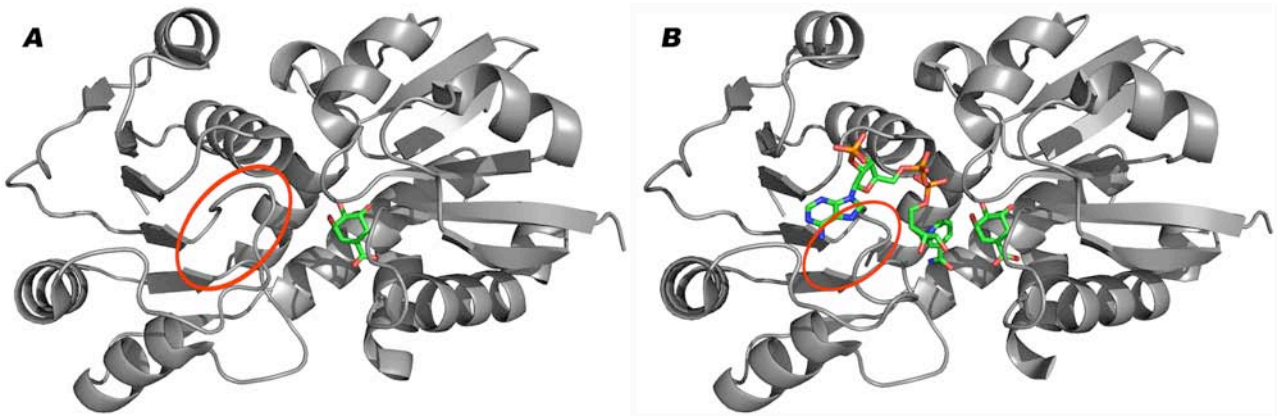
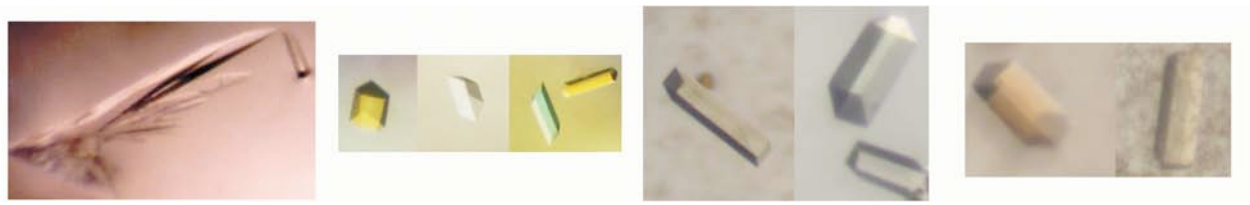


Figure 11. Binary (left; PDB entry: 3PHI) and ternary structure (right; PDB entry: 3PHH) of the shikimate dehydrogenase from *Helicobacter pylori*. The substrate dehydroshikimate and the cofactor NADP(H) are presented as sticks. The red circle indicates the loop region in the N-terminal domain which acts as a lid during cofactor binding.

In the ***binary*** structure of the *Hpy*SDH with bound shikimate there is a large loop in the C-terminal domain that obstructs the entrance to the cofactor-binding cleft and virtually acts as a lid. For cofactor binding this loop has to move away from the cleft in order to create space for NADP(H). Comparing these two structures with the overall conformation of the ***apo-protein***, these two conformational stages represent stages 2.)/3.) or 6./7.), in the catalytic cycle shown in Figure 7.

4.5 Quinate dehydrogenase from *Corynebacterium glutamicum*: Last but not least the bacterial quinate dehydrogenase of *C. glutamicum* could be structurally solved in four different catalytic states: apo-enzyme, with bound ***cofactor*** NAD^+ and in complex with quinate (QA) and the reduced cofactor or shikimate (SA) and NADH, i. e. with the natural ***substrate*** and the natural cofactor as product of the reaction.
For growing the crystals of the apo-form the protein solution was mixed with the reservoir solution and a NADH solution (2 µg/ml) in a drop ratio 1:1:1. The reduced cofactor could not be detected in the electron density due to the very low concentration [46].
For the ***co-crystallization*** trials (with the cofactor NAD^+, the substrate quinate (QA) and the reduced cofactor or shikimate (SA) and NADH) the kinetical parameters were determined first in order to get an idea of the concentrations necessary for successful ***ligand*** binding. The K_M values for NAD^+, QA and SA are 0.28 mM, 2.37 mM and 53,88 mM, respectively (Hoeppner et al.; publication in progess).

To obtain the ***binary*** and both of the ***ternary complexes*** the protein solution was mixed with either NAD^+ or QA plus NADH or SA plus NADH to a final concentrations of 1 mM for NAD^+ or NADH and 35 mM for QA or SA. These mixtures were incubated on ice for about 1 hour prior to crystallization. All substrates and the cofactor were bound during co-crystallization experiments by means of the ***sitting drop method*** with drop size of 2-4 µl in 1:1 ratio of protein and reservoir solution.



Apo *Cgl*QDH	*Cgl*QDH-NAD^+	*Cgl*QDH-QA-NADH	*Cgl*QDH-SA-NADH
100 mM sodium acetate pH 4.6, 200 mM NaCl, 20 % (v/v) 2-methyl-2,4-pentanediol (MPD)	1.6 M sodium citrate tribasic dihydrate pH 6.9, plus up to 62 mM $CoCl_2$	24 % (w/v) PEG 6000, 360 to 400 mM $CaCl_2$, 100 mM Tris-HCl pH 8.0 to 9.5	

Figure 12. Comparison of the crystal shapes of the four different catalytical states of the *Cgl*QDH and the corresponding crystallization conditions.

Crystals of the binary and ternary complexes were different in shape compared to the crystals of the ***apo-enzyme*** and grew under diffenrent conditions (Figure 12), which was a hint to (structural) changes within the protein molecules.

The crystals of all three *Cgl*QDH complexes diffracted to atomic resolution and allowed us to assign the position of all ***ligand*** atoms unambiguously within the electron density (Figure 13).

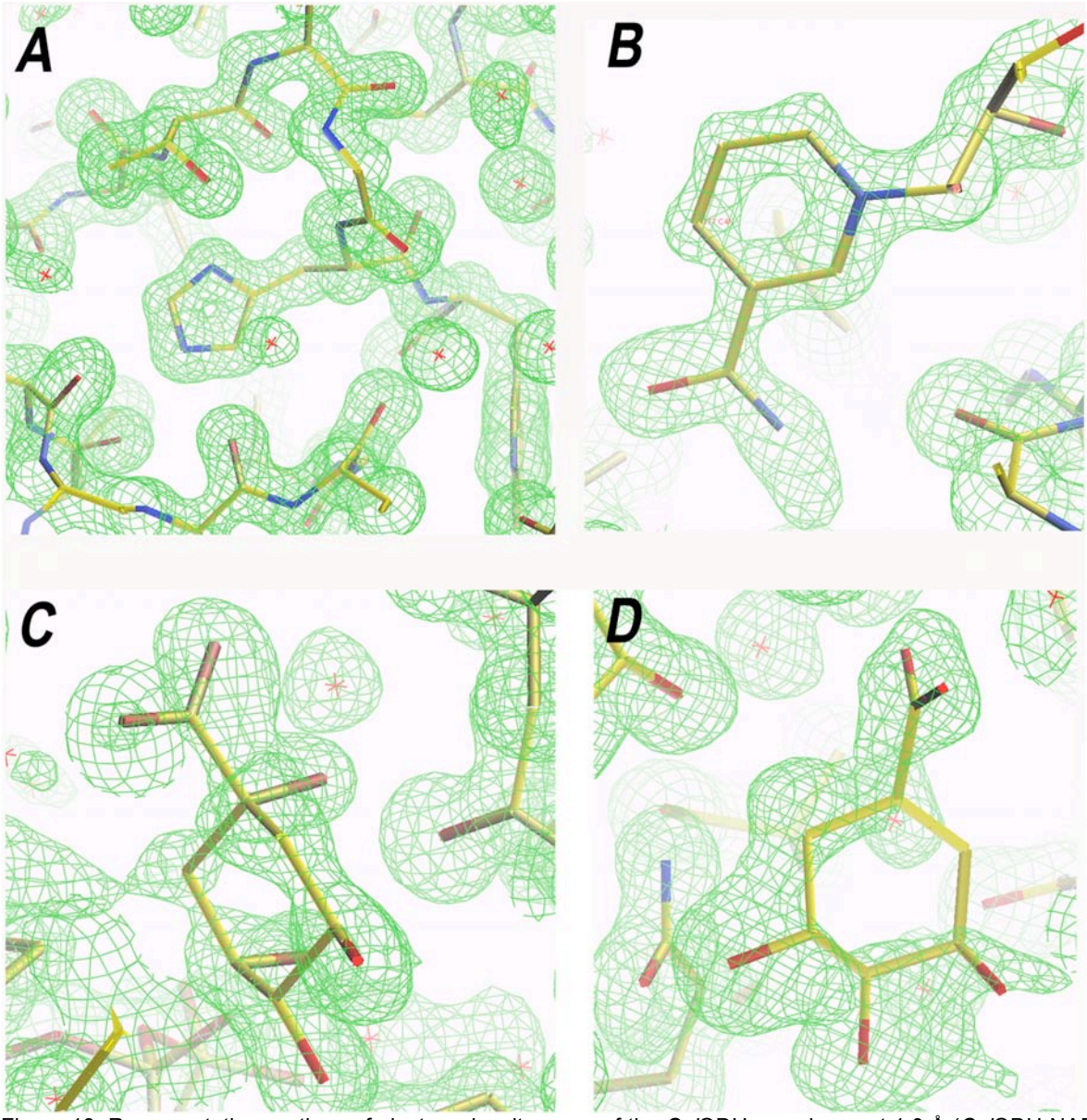


Figure 13. Representative sections of electron density maps of the *Cgl*QDH complexes at 1.0 Å (*Cgl*QDH-NAD^+) or 1.16 Å (*Cgl*QDH-QA-NADH and *Cgl*QDH-SA-NADH) resolution. A) electron density defining protein side chains, B) density around the nicotinamide ring of the cofactor NAD(H), C) bound substrate quinate, D) bound substrate shikimate (electron density maps in A)-C) contoured at 1 σ and in D) 0.7 σ).

By comparing the overall structures of all these states an open, a semi-open and a closed conformation of the enzyme (Figure 14) was observed. Surprisingly, the apo-structure of the *Cgl*QDH exhibits the closed form although one would intuitively expect the open conformation. But it is possible that these findings were a crystallization artifact since the reservoir solution was quite acidic (pH 4.6) compared to the *Cgl*QDH pH optimum, which is 9.0-9.5 for quinate and 10.0-10.5 for shikimate (Hoeppner et al.; publication in progess).

Within the ***cofactor*** binding domain of *Cgl*QDH the glycine rich loop, which is highly conserved within SDH proteins and represents a classical Rossmann fold, is flapped down towards the cofactor binding cleft in the apo-structure, but moved away when the cofactor is bound. With regard to the overall arragement of the apo-state compared to the NAD^+-bound

state there is a clearly visible opening of the two domains. After forming the ***ternary complex*** the two domains are brought closer to each other, if only more slightly compared to the apo-conformation and thus adopt a semi-closed conformation (Hoeppner et al.; publication in progess).

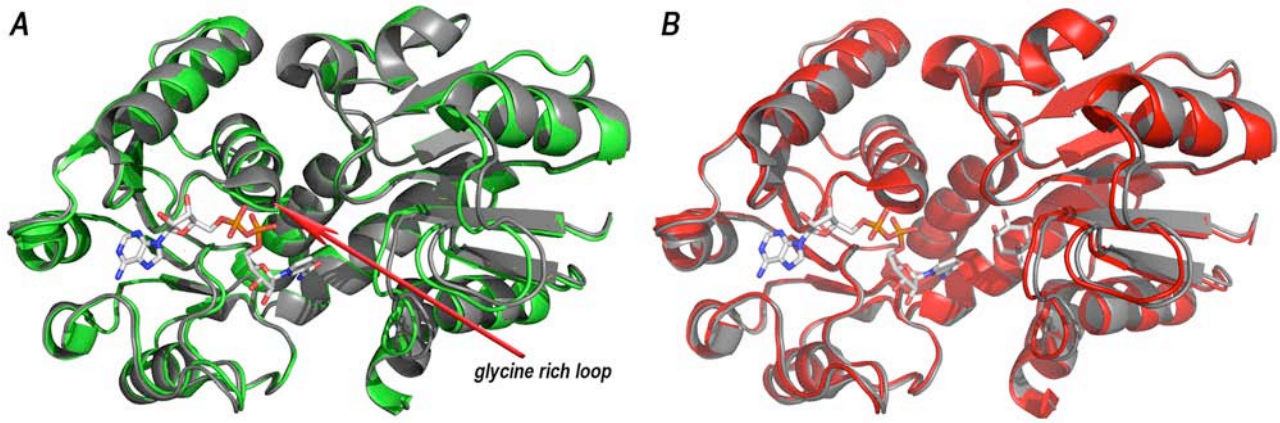


Figure 14. Structural alignments of the binary structure of *Cgl*QDH with bound NAD^+ (black; PDB entry: 3JYO) and A) the apo-protein (green; PDB entry: 2NLO) or B) the ternary structure with bound quinic acid and NADH (red; PDB entry: 3JYP). The substrate and the cofactor NAD(H) are presented as sticks. The red arrow indicates the conformational changes within the glycine rich loop.

4.6 Insights into the structural changes during catalysis and elucidation of substrate and cofactor specificity, using the example of *Cgl*QDH

Structure overview of C. glutamicum QDH

All *Cgl*QDH structures presented here are determined from crystals that were nearly isomorphous and belong to the same space group C2. The unit-cell parameters are very similar with one monomer per asymmetric unit.

The 282 residues in the QDH molecule form two structural domains (Figure 15): the N-terminal or catalytic domain (residues 1 to 113 and 256 to 283), which binds the ***substrate*** molecule, and the C-terminal or nucleotide binding domain (residues 114 to 255). The catalytic domain forms an open α/β sandwich, which is characteristic for enzymes of the S/QDH family but different from all other known proteins. The domain consists of a six-stranded, mainly parallel β sheet (strand order β2, β1, β3, β5, β6 and β4, where β5 is antiparallel. This β sheet is flanked by helices α1 and α11 at

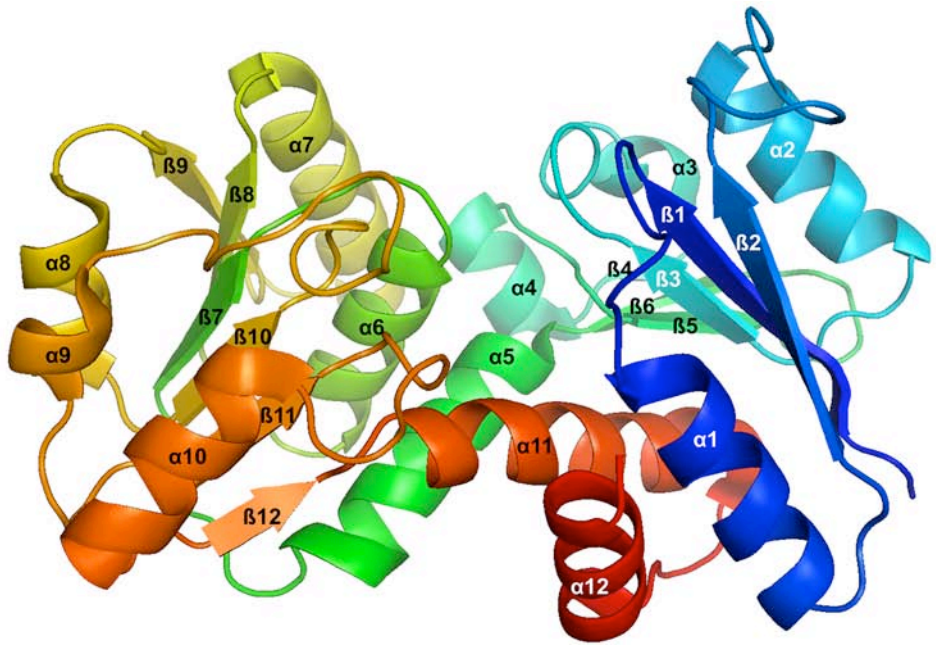


Figure 15. Schematical overview of the *Cgl*QDH fold

one side and α2, one 3_{10}-helix and α4 at the other. The C-terminal domain contains a six-stranded parallel β sheet (strand order β9, β8, β7, β10, β11 and β12) sandwiched by three helices (α7, α6, α5) on one face and by helices α8, α10 and a 3_{10}-helix on the other. The nucleotide binding domain exhibits a glycine rich loop with the sequence motive GXGGXG. The overall fold of this functional domain is very similar to that observed for other SDH proteins [47, 48] and represents the classical Rossmann fold. Both domains are linked together by helices α5 and α11. The arrangement of these two domains creates a deep active site groove in which ***cofactor*** and substrate are located.

Description and analysis of QDH active site

Cofactor Binding Site: The electron densities for NAD(H) were of high quality and allowed us to assign the position of these ***ligand*** unambiguously at 1.0 Å. *Cgl*QDH crystallizes in the presence of NAD^+ in the same space group with similar unit cell dimensions, but under different crystallization conditions compared to the ***apo-enzyme***. With regard to the overall structure we found that the catalytic domain moves away from the nucleotide-binding domain after cofactor binding making the interdomain cavity larger. Concerning the steric configuration of the residues there are only little but fundamental variances, especially in the glycine rich loop. In comparison to the QDH apo-enzyme (PDB entry 2NLO) the residues of the loop (Gly136-Val138) move out of the cavity after cofactor binding and therefor clear space for the NAD(H) molecule (Figure 14). Cofactor binding occurs in an extended groove between the N-terminal and C-terminal domain, whereas most of the molecular interactions result from the C-terminal domain. The adenine part of the adenosine moiety form hydrogen bonds only to some water molecules, while the ribose is bound by the side chains of Asp158 and Arg163. The phosphate moiety contacts the glycine rich loop and forms hydrogen bonds to Arg163 and the backbone nitrogen atom of Val138. The following ribose moiety again interacts only with water molecules, whereas the nicotinamide moiety is cramped by the backbone nitrogen of Ala255 and backbone oxygens of Val228 and Gly251, respectively. Gly251 and Ala255 are the only residues of the N-terminal domain involved in cofactor binding (Figure 16). The nucleotide-binding motive GXGGXG comprises the residues Gly134-Ala135-Gly136-Gly137-Val138-Gly139 (Hoeppner et al.; publication in progess).

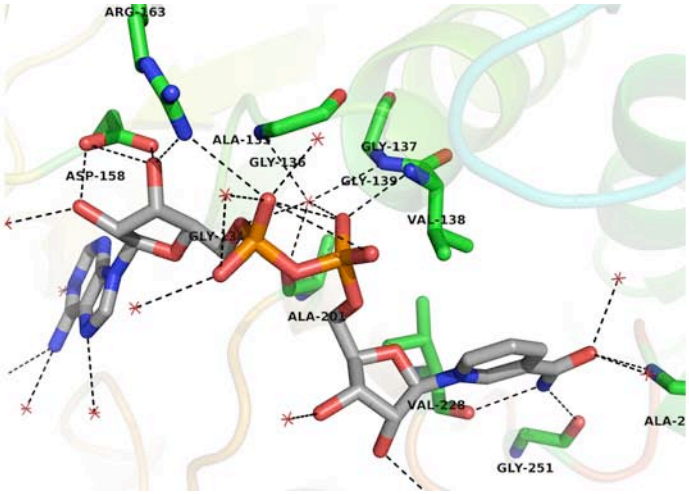


Figure 16. Interactions between the cofactor NAD(H) and *Cgl*QDH. Residues involved in hydrogen bonds (dotted lines) and bound ligand are shown as sticks, water molecules are depicted as red stars.

The strict specificity for NAD(H) is determined by the negatively charged aspartate residue 158, the neutral Leu159 and the bulk side chain of Arg163, which would result in steric hindrance with the additional phosphate group in the NADP(H) molecule.

Substrate Binding Site: We examined the ***substrate*** binding site of *Cgl*QDH by analysis of the two different ***ternary complexes*** QDH-QA-NADH and QDH-SA-NADH. The substrate binding site is located in the N-terminal domain, close to the nicotinamide ring of the ***cofactor***, and is characterized by a number of highly conserved residues.

After quinate binding a slight closure of the N- and C-terminal domain of *Cgl*QDH so that the crevice becomes closer by about 0.5 Å was observed. The substrate quinate is anchored by numerous key interactions with these residues: the carbonyl group of quinate is bound by the hydroxyl groups of Ser17 and Thr19; the hydroxyl groups of the C1 and C3 atom of the substrate form hydrogen bonds to side chain of Thr69, whereas the nitrogen atom of Lys73 binds to the hydroxyl groups of C3 and C4, the latter furthermore interacts with the side chains of Asn94 and Asp110; the fourth hydroxyl group at C5 forms hydrogen bonds to the amide group of Asp110 and the oxygen atom of the Gln258 side chain, respectively. A total of eleven hydrogen bonds cause a forcipate anchorage of the substrate molecule (Figure 17 B).

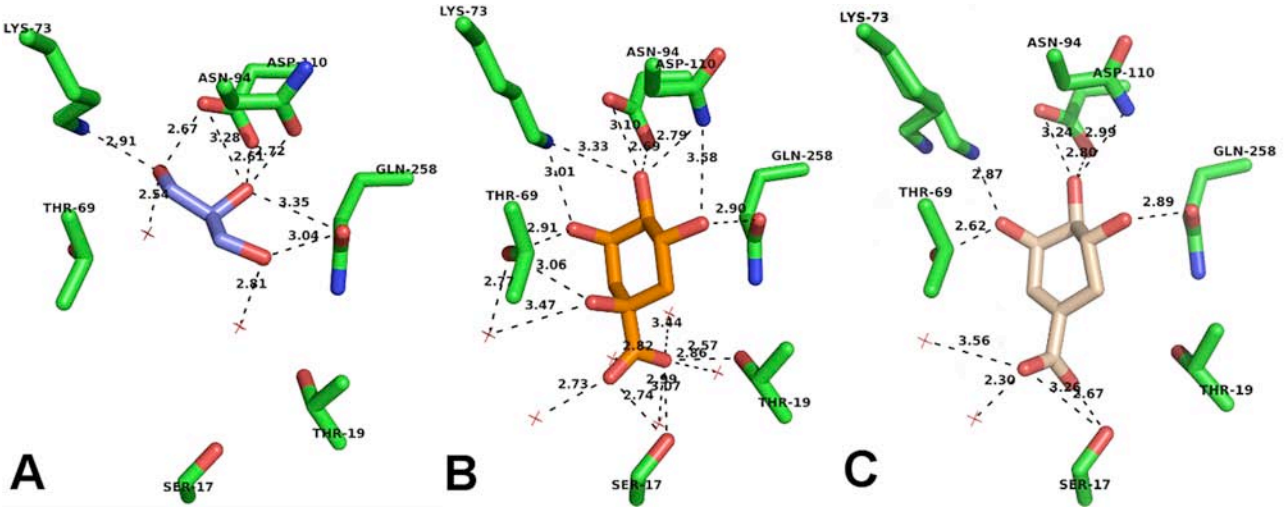


Figure 17. Active site residues of *Cgl*QDH. A) Apo-*Cgl*QDH (PDB entry: 2NLO) with bound glycerol (purple), B) ternary complex (PDB entry: 3JYP) with bound quinate (orange), C) ternary complex (PDB entry: 3JYQ) with shikimate (wheat). Residues involved in hydrogen bonds (dotted lines) and bound ligand are shown as sticks, water molecules are depicted as red stars.

In comparison to the ***apo-enzyme*** QDH (Figure 17 A) it is noteworthy that the side chain of Lys73 exhibits a sprawled conformation after quinate binding, which is required for interaction with the C3 and C4 hydroxyl groups of the substrate. For the hydride ion transfer from C3 of quinate to C4 of NAD^+ a particular distance between these atoms is very important. In the crystal structure the nicotinamide ring is located in a suitable orientation for the H^- transfer. After quinate binding and resulting closure of the domains the ***cofactor*** approaches to the substrate-binding site, whereby the distance of interest amounts to 4.27 Å.

In the case of shikimate binding a somewhat different situation was observed. In principle the above mentioned residues except Thr19 are involved in shikimate binding (Figure 17 C), but only eight polar interactions are achieved (compared to eleven when QA is bound), from which some are furthermore weaker pronounced: Thr19 is not involved in polar contacts to SA, Thr69 has contact only to the hydrogen group of C3, Asn94 is about 0.2 Å farer apart from the hydrogen atom of C4 and has no contact to the OH-group of C5. Remarkable is the appearance of an alternative side chain conformation of Lys73, as evidenced by the excellent electron density in this region. The first conformation of the Lys73 side chains in the crystal exhibits the sprawled conformation as found for the quinate binding; the second conformation reveals an angled rotamer as it occurs in apo-*Cgl*QDH. The latter conformation makes hydrophobic interactions with the shikimate molecule impossible (Hoeppner et al.; publication in progess). Furthermore the shikimate molecule exists in a half-chair conformation, whereas the quinate molecule adopts a chair conformation. Hence the distance of the C4 atom of the cofactor and the C3 atom of the substrate increases to 4.67 Å. All residues involved in cofactor and substrate binding identified here are consistent with these of further reported structures (i. e. *Tth*SDH, *Aae*SDH, *AthDHQ*-SDH).

Substrate and cofactor specificity and discrimination

All results of the structural analysis are also in excellent agreement with the findings of the kinetical assays. The higher ***affinity*** of *Cgl*QDH to the ***substrate*** quinate (as obvious by means of the clearly unequal K_M values) arises from the major quantity of hydrogen bonds between protein and the substrate quinate. Considering the known structures of shikimate dehydrogenases as SDH from *T. thermophilus* [40] SDH from *A. aeolicus* [41] or the SDH domain of *A. thaliana* [42, 43] they all possess at least eleven hydrogen bonds to the substrate molecule shikimate, comparable to the quinate binding in CglQDH. In contrast there are only eight polar interactions present between the enzyme and the shikimate molecule, because shikimate offers a somewhat different conformation (half-chair instead of chair) and exhibits no hydroxyl group at the C1 atom. Furthermore the formed hydrogen bonds between shikimate and the enzyme are accomplished weaker. The higher catalytic efficiency of *Cgl*QDH regarding to quinate (as obvious on the basis of significantly higher k_{cat}/K_M values) possibly results from the slightly lower distance between the C4 atom of the ***cofactor*** NAD(H) and the C3 atom of the quinate molecule (4.27 Å versus 4.67 Å) and the improved orientation of the substrate quinate.

A further occasion for the lower affinity and catalytic efficiency regarding shikimate results from the appearance of an alternative conformation of the Lys73 side chain (Figure 17 C), which leads to a loss of an important hydrogen bond. At last we compared the substrate binding residues of *Cgl*QDH with those of *Aae*QDH, *AthDHQ*-SDH and *Tth*SDH, which convert shikimate. We detected two differences possibly jointly responsible for quinate binding: in all the above-mentioned structures a tyrosine residue is involved in shikimate binding but not in *Cgl*QDH (Tyr230). Furthermore the second serine, which forms a hydrogen bond to the carbonyl group of shikimate, is replaced by a threonine residue in

*Cgl*QDH (Thr19). Since the carbonyl group of shikimate bound in *Cgl*QDH is twisted about 90° compared to the situation of the aforesaid enzymes, Thr19 cannot take part in polar interaction with the substrate shikimate. Concerning the usage of the cofactors NAD(H) and NADP(H) the classical dinucleotid fold were identified in the past. Characteristic for all nucleotide binding proteins is the glycine rich loop with the common sequence GXGXXG, in which the number of glycine residues changes [49]. Enzymes using NAD or FAD possess a well conserved negative charged amino acid at the C-terminus of the second β-strand of the nucleotide binding βαβ unit, mostly aspartate or glutamate. This residue interacts with the 2´-hydroxyl group of the ribose. In the majority of the NADP binding proteins this negatively charged residue is absent since the additional 2´-phosphate group is located at this position. Moreover, various NADP dependent proteins exhibit a charged amino acid (like arginine) in the position of the 2´-phosphate group which stabilizes the cofactor molecule. The *Cgl*QDH described here offers a negative charged residue (Asp158) forming a hydrogen bond to 2´-hydroxyl group of the ribose, followed by a neutral amino acid (Leu159) which is unable to interact with an additional phosphate group as present in NADP. Furthermore the bulk side chain of Arg163 constrict the cofactor binding site, which would result in steric hindrance with the additional phosphate group in a NADP(H) molecule. Due to all results of the kinetical and structural analysis we conclude that *Cgl*QDH is strictly NAD(H) specific and not able to bound NADP(H).

5. Conclusions

Crystal structures of proteins and enzymes are important to fully understand the mechanism and mode of action. Although the crystllization of a large number of proteins was successful and delivered valuable information the goal must be to fully understand the function. When crystallizing a protein a snapshot of the protein in a certain conformation is observed in the electron density. It is known that proteins are flexible and can obtain several states in solution.

Within that book chapter we explained the importance to acquire structural information of different catalytical states of proteins or enzymes, to fully understand how the protein behaves during catalysis or how the substrate bound state differs from the apo-enzyme.

The open and closed structures of the substrate binding protein ProX as apo-protein or with different substrates bound revealed enormous conformational changes during ligand binding and clearly visualzes how flexible a protein can be and elucidates the side chain movements within the substrate site upon ligand binding.

All described crystallization trials of the different transition states of the OcDH showed impressively that protein crystallization is a trial and error approach and that knowledge of the protein (especially the kinetical parameters beside others) is the essential thing to be successful. At best and as recompenses for ones effort one will achieve important insights that clearly explains the catalytic mechanism.

Last but not least the different structural information of the enzymes of the shikimate dehydrogenase family could bring to light how substrate and cofactor specificity and discrimination can be achieved throught detailed analysis of apo-, binary and ternary structure information about involved amino acids in substrate and cofactor binding.
So with these three examples the difficulties in crystallization on one hand and on the other hand the beauty of looking at proteins at work is shown.

PDB entries used:

Protein	PDB Code	Title
ChoX	2RF1	Crystal structure of ChoX in an unliganded closed conformation
	3HCQ	Structural analysis of the choline binding protein ChoX in a semi-closed and ligand-free conformation
	2REJ	ABC-transporter choline binding protein in unliganded semi-closed conformation
	2RIN	ABC-transporter choline binding protein in complex with acetylcholine
	2REG	ABC-transporter choline binding protein in complex with choline
ProX	1SW1	Crystal structure of ProX from *Archeoglobus fulgidus* in complex with proline betaine
	1SW4	Crystal structure of ProX from *Archeoglobus fulgidus* in complex with trimethyl ammonium
	1SW2	Crystal structure of ProX from *Archeoglobus fulgidus* in complex with glycine betaine
	3MAM	A molecular switch changes the low to the high affinity state in the substrate binding protein *Af*ProX
	1SW5	Crystal structure of ProX from *Archeoglobus fulgidus* in the ligand free form
CenDH	1BG6	Crystal structure of the N-(1-D-carboxylethyl)-L-norvaline dehydrogenase from *Arthrobacter sp*. strain 1C
OcDH	3C7C	A structural basis for substrate and stereo selectivity in octopine dehydrogenase (OcDH-NADH-L-arginine)
	3C7D	A structural basis for substrate and stereo selectivity in octopine dehydrogenase (OcDH-NADH-pyruvate)
	3C7A	A structural basis for substrate and stereo selectivity in octopine dehydrogenase (OcDH-NADH)

AroE	2HK8	Crystal structure of shikimate dehydrogenase from *Aquifex aeolicus*
	2HK9	Crystal structure of shikimate dehydrogenase from *Aquifex aeolicus* in complex with shikimate and $NADP^+$
	1WXD	Crystal structure of shikimate 5-dehydrogenase (AroE) from *Thermus thermophilus* HB8
	2D5C	Crystal structure of shikimate 5-dehydrogenase (AroE) from *Thermus thermophilus* HB8 in complex with shikimate
	2EV9	Crystal structure of shikimate 5-dehydrogenase (AroE) from *Thermus thermophilus* HB8 in complex with NADP(H) and shikimate
	3PHG	Crystal structure of the shikimate 5-dehydrogenase (AroE) from *Helicobacter pylori*
	3PHH	Crystal structure of the shikimate 5-dehydrogenase (AroE) from *Helicobacter pylori* in complex with shikimate and NADP(H)
	3PHI	Crystal structure of thesShikimate 5-dehydrogenase (AroE) from *Helicobacter pylori* in complex with dehydroshikimate
DHQ-SDH	2GPT	Crystal structure of *Arabidopsis* dehydroquinate dehydratase-shikimate dehydrogenase in complex with tartrate and shikimate
	2O7Q	Crystal structure of the *A. thaliana* DHQ-dehydroshikimate-SDH in complex with dehydroshikimate
	2O7S	Crystal structure of the *A. thaliana* DHQ-dehydroshikimate-SDH in complex with dehydroshikimate, NADP(H) and tartrate
QDH	3JYO	Quinate dehydrogenase from *Corynebacterium glutamicum* in complex with NAD
	3JYP	Quinate dehydrogenase from *Corynebacterium glutamicum* in complex with quinate and NADH
	3JYQ	Quinate dehydrogenase from *Corynebacterium glutamicum* in complex with shikimate and NADH
	2NLO	Crystal structure of the quinate dehydrogenase from *Corynebacterium glutamicum*

Glossary

Apo-protein/apo-enzyme
Enzymes that require a cofactor but do not have one bound are called *apo-enzymes* or *apo-proteins*. An apo-enzyme together with its cofactor(s) is called a *holoenzyme*.
Affinity
The dissociation constant is commonly used to describe the **affinity** between a ligand and a protein, i.e. how tightly a ligand binds to a particular protein. Ligand-protein affinities are influenced by non-covalent intermolecular interactions between the two molecules such as hydrogen bonding, electrostatic interactions, hydrophobic and Van der Waals forces. They can also be affected by high concentrations of other macromolecules, which causes macromolecular crowding. The smaller the dissociation constant K_d, the more tightly bound the ligand is, or the higher the affinity between ligand and protein.
Binary complex
A **binary complex** refers to a protein complex containing two different molecules which are bound together. In structural biology, the term binary complex can be used to describe a crystal containing a protein with one small molecule bound, for example the cofactor or the substrate; or a complex formed between two proteins.
Co-crystallization
Co-crystallization means that the protein solution is mixed with one or more ligand prior to the crystallization. Often the protein-ligand mixture is preincubated before setting up the crystallization drops.
Cofactor
A **cofactor** is a non-protein chemical compound that is bound to a protein and is required for the protein's biological activity. These proteins are commonly enzymes, and cofactors can be considered "helper molecules" that assist in biochemical transformations. Cofactors are either organic or inorganic. They can also be classified depending on how tightly they bind to an enzyme, with loosely bound cofactors termed **coenzymes** and tightly-bound cofactors termed prosthetic groups. Examples of widespread cofactors are ATP, coenzyme A, FAD, and NAD^+, vitamins or metal ions.
K_d
In chemistry, biochemistry, and pharmacology, a dissociation constant $\boldsymbol{K_d}$ is a specific type of equilibrium constant that measures the propensity of a larger object to separate (dissociate) reversibly into smaller components, as when a complex falls apart into its component molecules, or when a salt splits up into its component ions.

K_M
In biochemistry, Michaelis-Menten kinetics is one of the simplest and best-known models of enzyme kinetics. The Michaelis constant $\boldsymbol{K_M}$ is the substrate concentration at which the reaction rate is half of $\boldsymbol{V_{max}}$ (which represents the maximum rate achieved by the system, at maximum (saturating) substrate concentrations).
Intrinsic tryptophan fluorescence
Binding of ligands to proteins frequently causes changes to their three-dimensional structure. Exampes of this include the binding of substrates, inhibitors, cofactors or allosteric modulators to enzymes or of hormons to receptors. If this structural change has an effect on the environment of an intrinsic or extrinsic fluorophore in the protein, this can result in measurable changes in the fluorescence spectrum. Provided that the fluorophore has a unique location in the protein, such changes of fluorescence at a particular wavelength can be used to determine the dissociation constant (K_d) of the protein for the ligand where K_d is a measure of the affinity of the protein for the ligand [50].
Isothermal Titration Calorimetry (ITC)
This technique is useful for protein concentrations in the range of mg/ml. A typical experiment involves measurement of heat change as a function of addition of small quantities of a reagent to the calorimeter cell containing other components of the system under investigation. For example, this reagent could be a protein ligand or substrate/ inhibitor of an enzyme. At the beginning of the experiment, there is a large excess of protein compared to ligand. This means that Δ *H* values associated with each aliquot can be individually measured. Initially, these values are large but, as aliquots are progressively added, eventually decrease to values similar to the Δ *H* of dilution of ligand into the solution in the calorimeter cell. The Δ *H* measured is the total enthalpy change which includes heat associated with processes such as formation of noncovalent bonds between interacting molecules and with other equilibria in the system such as conformational changes, ionization of polar groups (e.g. deprotonation) and changes due to interactions with solvent. ITC provides a useful method for studying binding processes such as those involving a protein and a ligand. It allows estimation of both the binding constant (K_b) and of the dissociation constant (K_d) [50].
Ligand
In biochemistry a **ligand** is a substance (usually a small molecule), that forms a complex with a biomolecule to serve a biological purpose. In the context of this chapter ligand is used as a more general expression for substrate, product or cofactor.
Ligand soaking
Ligand soaking means the addition of ligands into the mother liquid with preformed crystals. The idea is that the ligand diffuses into the crystals and binds at the active site. This technique was initially used for the incorporation of heavy atoms into protein crystals for phasing purposes.
Macro and Micro Seeding
During **Macro Seeding** the protein crystal is replaced into a freshly made mother liquid which allows the further enlargement of the crystals size. In **Micro Seeding** a suspension of microcrystals is prepared by either resuspending or crushing a protein crystal cluster or single crystals. These seeds are then used (often streaked through a new droplet of precipitant and fresh protein) to serve a crystallization starting point.
Microbatch
Microbatch is a method in which the molecule to be crystallized is mixed with the crystallizing agents at the start of the experiment. The concentration of the ingredients is such that supersaturation is achieved immediately upon mixing, thus the composition and the volume of a trial remain constant and crystals will only form if the precise conditions have been correctly chosen.
Occupancy
Occupancy means the degree of protein molecules in solution or in a crystal with bound ligand. If every second protein has attached a ligand the occupancy is 50 %.
Substrate
In biochemistry, a **substrate** is a molecule upon which an enzyme acts. Enzymes catalyze chemical reactions involving the substrate(s). In the case of a single substrate, the substrate binds with the enzyme active site, and an enzyme-substrate complex is formed. The substrate is transformed into one or more products, which are then released from the active site. The active site is now free to accept another substrate molecule. In the case of more than one substrate, these may bind in a particular order to the active site, before reacting together to produce products.

Surface Plasmon Resonance (SPR)
SPR is an optical technique which depends on changes in refractive index or mass changes near metal surfaces. When two surfaces, one a metal and the other a dielectric material are exposed to a beam of plane-polarized light of wavelength, λ, a longitudinal charge density wave (a surface plasmon) is propagated along the interface between them. This only happens when one of the surfaces is a metal and works best with silver, gold, copper and aluminium. This is because metals contain free oscillating electrons called plasmons. When light traveling through an optically dense medium such as glass arrives at an interface with a lower optical density (e.g. liquid), it is reflected back into the more optically dense medium, a phenomenon called total internal reflectance. Any process altering n_s (the refractive index of the dielectric medium) can be sensitively detected by SPR so the technique has found applications in the study of kinetics and thermodynamics of binding processes (e.g. protein-ligand, protein-protein) [50].
Ternary complex
A **ternary complex** refers to a protein complex containing three different molecules which are bound together. In structural biology ternary complex can be used to describe a crystal containing a protein with two small molecules bound, for example cofactor and substrate; or a complex formed between two proteins and a single substrate.
Vapor diffusion Hanging or Sitting drop
Two of the most commonly used methods for protein crystallization fall under the category of vapor diffusion. These are known as the **hanging drop** and **sitting drop** methods. Both entail a droplet containing purified protein, buffer, and precipitant being allowed to equilibrate with a larger reservoir containing similar buffers and precipitants in higher concentrations. Initially, the droplet of protein solution contains an insufficient concentration of precipitant for crystallization, but as water vaporizes from the drop and transfers to the reservoir, the precipitant concentration increases to a level optimal for crystallization. Since the system is in equilibrium, these optimum conditions are maintained until the crystallization is complete [51].

References:

1. Abts A, Schwarz CK, Tschapek B, Smits SH, Schmitt L. Rational and Irrational Approaches to Convince a Protein to Crystallize, Modern Aspects of Bulk Crystal and Thin Film Preparation, 2012 Nikolai Kolesnikov and Elena Borisenko (Ed.), ISBN: 978-953-307-610-2, InTech
2. Berntsson RP, Smits SH, Schmitt L, Slotboom DJ, Poolman B. A structural classification of substrate-binding proteins. FEBS Lett. 2010 Jun 18;584(12):2606-17.
3. Wilkinson J, Verschueren KHG. Crystal structures of periplasmic solute-binding proteins in ABC transport complexes illuminate their function. In: Holland IB, Cole SPC, Kuchler K, Higgins CF, editors. ABC proteins: from bacteria to man. London: Academic Press (Elsevier Science); 2003. p. 187-208.
4. Quiocho FA, Ledvina PS. Atomic structure and specificity of bacterial periplasmic receptors for active transport and chemotaxis: variation of common themes. Mol Microbiol. 1996 Apr;20(1):17-25.
5. Mao B, Pear MR, McCammon JA, Quiocho FA. Hinge-bending in L-arabinose-binding protein. The "Venus's-flytrap" model. J Biol Chem. 1982 Feb 10;257(3):1131-3.
6. Sack JS, Saper MA, Quiocho FA. Periplasmic binding protein structure and function. J Mol Biol. 1989;206:171-91.
7. Oh BH, Pandit J, Kang CH, Nikaido K, Gokcen S, Ames GF, et al. Three-dimensional structures of the periplasmic lysine/arginine/ornithine-binding protein with and without a ligand. J Biol Chem. 1993 May 25;268(15):11348-55.
8. Loh AP, Pawley N, Nicholson LK, Oswald RE. An increase in side chain entropy facilitates effector binding: NMR characterization of the side chain methyl group dynamics in Cdc42Hs. Biochemistry. 2001;40(15):4590-600.
9. Davidson AL, Dassa E, Orelle C, Chen J. Structure, function, and evolution of bacterial ATP-binding cassette systems. Microbiol Mol Biol Rev. 2008 Jun;72(2):317-64, table of contents.
10. Shilton BH. The dynamics of the MBP-MalFGK(2) interaction: a prototype for binding protein dependent ABC-transporter systems. Biochim Biophys Acta. 2008 Sep;1778(9):1772-80.
11. Tang C, Schwieters CD, Clore GM. Open-to-closed transition in apo maltose-binding protein observed by paramagnetic NMR. Nature. 2007 Oct 25;449(7165):1078-82.
12. Sharff AJ, Rodseth LE, Quiocho FA. Refined 1.8-A structure reveals the mode of binding of beta-cyclodextrin to the maltodextrin binding protein. Biochemistry. 1993 Oct 12;32(40):10553-9.
13. Spurlino JC, Lu GY, Quiocho FA. The 2.3-A resolution structure of the maltose- or maltodextrin-binding protein, a primary receptor of bacterial active transport and chemotaxis. J Biol Chem. 1991 Mar 15;266(8):5202-19.
14. Oswald C, Smits SH, Hoing M, Bremer E, Schmitt L. Structural analysis of the choline-binding protein ChoX in a semi-closed and ligand-free conformation. Biol Chem. 2009 Nov;390(11):1163-70.
15. Oswald C, Smits SH, Hoing M, Sohn-Bosser L, Dupont L, Le Rudulier D, et al. Crystal structures of the choline/acetylcholine substrate-binding protein ChoX from *Sinorhizobium meliloti* in the liganded and unliganded-closed states. J Biol Chem. 2008 Nov 21;283(47):32848-59.

16. Pittelkow M, Tschapek B, Smits SH, Schmitt L, Bremer E. The Crystal Structure of the Substrate-Binding Protein OpuBC from *Bacillus subtilis* in Complex with Choline. J Mol Biol. 2011 Aug 5;411(1):53-67.
17. Bermejo GA, Strub MP, Ho C, Tjandra N. Ligand-free open-closed transitions of periplasmic binding proteins: the case of glutamine-binding protein. Biochemistry. 2010 Mar 9;49(9):1893-902.
18. Borths EL, Locher KP, Lee AT, Rees DC. The structure of Escherichia coli BtuF and binding to its cognate ATP binding cassette transporter. Proc Natl Acad Sci U S A. 2002 Dec 24;99(26):16642-7.
19. Linke C, Caradoc-Davies TT, Young PG, Proft T, Baker EN. The laminin-binding protein Lbp from *Streptococcus pyogenes* is a zinc receptor. J Bacteriol. 2009 Sep;191(18):5814-23.
20. Zou JY, Flocco MM, Mowbray SL. The 1.7 A refined X-ray structure of the periplasmic glucose/galactose receptor from *Salmonella typhimurium*. J Mol Biol. 1993 Oct 20;233(4):739-52.
21. Quiocho FA, Spurlino JC, Rodseth LE. Extensive features of tight oligosaccharide binding revealed in high-resolution structures of the maltodextrin transport/chemosensory receptor. Structure. 1997 Aug 15;5(8):997-1015.
22. Berntsson RP, Doeven MK, Fusetti F, Duurkens RH, Sengupta D, Marrink SJ, et al. The structural basis for peptide selection by the transport receptor OppA. EMBO J. 2009 May 6;28(9):1332-40.
23. Schiefner A, Holtmann G, Diederichs K, Welte W, Bremer E. Structural basis for the binding of compatible solutes by ProX from the hyperthermophilic archaeon *Archaeoglobus fulgidus*. J Biol Chem. 2004 Nov 12;279(46):48270-81.
24. Machius M, Brautigam CA, Tomchick DR, Ward P, Otwinowski Z, Blevins JS, et al. Structural and biochemical basis for polyamine binding to the Tp0655 lipoprotein of Treponema pallidum: putative role for Tp0655 (TpPotD) as a polyamine receptor. J Mol Biol. 2007 Oct 26;373(3):681-94.
25. Muller A, Severi E, Mulligan C, Watts AG, Kelly DJ, Wilson KS, et al. Conservation of structure and mechanism in primary and secondary transporters exemplified by SiaP, a sialic acid binding virulence factor from *Haemophilus influenzae*. J Biol Chem. 2006 Aug 4;281(31):22212-22.
26. Lecher J, Pittelkow M, Zobel S, Bursy J, Bonig T, Smits SH, et al. The crystal structure of UehA in complex with ectoine-A comparison with other TRAP-T binding proteins. J Mol Biol. 2009 May 29;389(1):58-73.
27. Hanekop N, Hoing M, Sohn-Bosser L, Jebbar M, Schmitt L, Bremer E. Crystal structure of the ligand-binding protein EhuB from *Sinorhizobium meliloti* reveals substrate recognition of the compatible solutes ectoine and hydroxyectoine. J Mol Biol. 2007 Dec 14;374(5):1237-50.
28. Oswald C, Smits SH, Bremer E, Schmitt L. Microseeding - a powerful tool for crystallizing proteins complexed with hydrolyzable substrates. Int J Mol Sci. 2008 Jun;9(7):1131-41.
29. Tschapek B, Pittelkow M, Sohn-Bosser L, Holtmann G, Smits SH, Gohlke H, et al. Arg149 Is Involved in Switching the Low Affinity, Open State of the Binding Protein AfProX into Its High Affinity, Closed State. J Mol Biol. 2011 Aug 5;411(1):36-52.
30. Zaitseva J, Oswald C, Jumpertz T, Jenewein S, Wiedenmann A, Holland IB, et al. A structural analysis of asymmetry required for catalytic activity of an ABC-ATPase domain dimer. EMBO J. 2006 Jul 26;25(14):3432-43.
31. Karpowich NK, Huang HH, Smith PC, Hunt JF. Crystal structures of the BtuF periplasmic-binding protein for vitamin B12 suggest a functionally important reduction in protein mobility upon ligand binding. J Biol Chem. 2003 Mar 7;278(10):8429-34.
32. Chen J, Lu G, Lin J, Davidson AL, Quiocho FA. A tweezers-like motion of the ATP-binding cassette dimer in an ABC transport cycle. Mol Cell. 2003 Sep;12(3):651-61.
33. Muller A, Janssen F, Grieshaber MK. Putative reaction mechanism of heterologously expressed octopine dehydrogenase from the great scallop, *Pecten maximus* (L). Febs J. 2007 Dec 7;274(24):6329-39.
34. Smits SH, Mueller A, Grieshaber MK, Schmitt L. Coenzyme- and His-tag-induced crystallization of octopine dehydrogenase. Acta Crystallogr Sect F Struct Biol Cryst Commun. 2008 Sep 1;64(Pt 9):836-9.
35. Smits SH, Mueller A, Schmitt L, Grieshaber MK. A structural basis for substrate selectivity and stereoselectivity in octopine dehydrogenase from *Pecten maximus.* J Mol Biol. 2008 Aug 1;381(1):200-11.
36. Britton KL, Asano Y, Rice DW. Crystal structure and active site location of N-(1-D-carboxylethyl)-L-norvaline dehydrogenase. Nat Struct Biol. 1998 Jul;5(7):593-601.
37. Asano Y, Yamaguchi K, Kondo K. A new NAD+-dependent opine dehydrogenase from Arthrobacter sp. strain 1C. J Bacteriol. 1989 Aug;171(8):4466-71.
38. van Os N, Smits SH, Schmitt L, Grieshaber MK. Control of D-octopine formation in scallop adductor muscle as revealed through thermodynamic studies of octopine dehydrogenase. J Exp Biol. 2012 May 1;215(Pt 9):1515-22.
39. Smits SH, Meyer T, Mueller A, van Os N, Stoldt M, Willbold D, et al. Insights into the mechanism of ligand binding to octopine dehydrogenase from *Pecten maximus* by NMR and crystallography. PLoS One. 2010;5(8):e12312.
40. Bagautdinov B, Kunishima N. Crystal structures of shikimate dehydrogenase AroE from *Thermus thermophilus* HB8 and its cofactor and substrate complexes: insights into the enzymatic mechanism. J Mol Biol. 2007 Oct 19;373(2):424-38.
41. Gan J, Wu Y, Prabakaran P, Gu Y, Li Y, Andrykovitch M, et al. Structural and biochemical analyses of shikimate dehydrogenase AroE from *Aquifex aeolicus*: implications for the catalytic mechanism. Biochemistry. 2007 Aug 21;46(33):9513-22.
42. Singh SA, Christendat D. Structure of Arabidopsis dehydroquinate dehydratase-shikimate dehydrogenase and implications for metabolic channeling in the shikimate pathway. Biochemistry. 2006 Jun 27;45(25):7787-96.

43. Singh SA, Christendat D. The DHQ-dehydroshikimate-SDH-shikimate-NADP(H) Complex: Insights into Metabolite Transfer in the Shikimate Pathway. Cryst Growth Des. 2007;7(11):2153-60.
44. Schoepe J, Niefind K, Chatterjee S, Schomburg D. Cloning, expression, purification and preliminary crystallographic characterization of a shikimate dehydrogenase from *Corynebacterium glutamicum*. Acta Crystallogr Sect F Struct Biol Cryst Commun. 2006 Jul 1;62(Pt 7):635-7.
45. Yaniv H, Gilvarg C. Aromatic biosynthesis. XIV. 5-Dehydroshikimic reductase. J Biol Chem. 1955 Apr;213(2):787-95.
46. Schoepe J, Niefind K, Schomburg D. 1.6 angstroms structure of an NAD+-dependent quinate dehydrogenase from *Corynebacterium glutamicum*. Acta Crystallogr D Biol Crystallogr. 2008 Jul;D64(Pt 7):803-9.
47. Singh S, Korolev S, Koroleva O, Zarembinski T, Collart F, Joachimiak A, et al. Crystal structure of a novel shikimate dehydrogenase from *Haemophilus influenzae*. J Biol Chem. 2005 Apr 29;280(17):17101-8.
48. Michel G, Roszak AW, Sauve V, Maclean J, Matte A, Coggins JR, et al. Structures of shikimate dehydrogenase AroE and its Paralog YdiB. A common structural framework for different activities. J Biol Chem. 2003 May 23;278(21):19463-72.
49. Wierenga RK, De Maeyer MCH, Hol WGJ. Interaction of Pyrophosphate Moieties with a-Helixes in Dinucleotide Binding Proteins. Biochemistry. 1985;24:1346-57.
50. Sheenan D. Physical Biochemistry. Principles and Applications: Wiley; 2009.
51. McRee DE. Practical Protein Crystallography. San Diego: Academic Press; 1993.