5. Dimer-octamer equilibrium in arylsulfatase A

Human arylsulfatase A (ASA) is an enzyme that functions as hydrolytic catalyst in the lysosome of the human cell. Lysosomes have acidic pH values. ASA occurs as homooctamer in an acidic environment. At higher pH values it dissociates to dimers. This pH-regulated association behavior indicates that it is perhaps due to the deprotonation of one or more titratable groups, that the octamers are less stable at higher pH values.

In this chapter I will describe how the pH-dependent association free energies between the dimer subunits of ASA can be investigated on the basis of electrostatic interactions. To introduce the topic I will first outline what is known about the role of electrostatic forces in protein-protein association in the following section.

5.1. Electrostatic interactions and protein-protein association

Many biological processes require that proteins bind to each other in a specific manner. It is important to know which forces drive the protein-protein interaction and by which means the association of proteins can be regulated by external factors. Electrostatic forces have an influence on the strength and the specificity of binding as well as on the rates of association.

The association of proteins has always been compared to the folding process of proteins. This comparison is obvious, because both processes are driven by the interaction between single amino acids as well as between secondary and tertiary structure entities. The physical properties of the interacting parts determine the fold of the protein and the binding between proteins.

Protein-protein association is a very specific process, that occurs via well defined interfaces that have evolved to accomplish the recognition and binding process. The composition of amino acids at protein-protein interfaces differs from the one within protein cores.[114,115,116,117] In protein cores hydrophobic forces dominate and are most relevant for the folding process and the stability of proteins. Accordingly, charged and polar groups are rare in protein cores.

In protein-protein interfaces, however, charged and polar groups are buried and the interfaces are more similar to protein surfaces than to the interior of proteins. This indicates that the forces in protein folding and in protein-protein association processes come from different sources and that both processes occur under different energetic circumstances.

Another fundamental difference between both processes is the amount of entropy loss that accompanies them. The configurational entropy penalty associated with protein folding is much larger than the one for protein-protein association, where only the translational and rotational entropy vanishes. Moreover the entropy loss of sidechains at the protein-protein interfaces has to be considered, but it is small compared to the number of residues that loose configurational entropy during protein folding. Therefore, the forces that control the association of proteins do not have to be as large as the forces that are responsible for protein folding. The hydrophobic forces in protein-protein association are generally smaller.

It is known, that salt bridges and hydrogen bonds, oppose protein folding.[118,119,120,121,15] The reason for this phenomenon is the desolvation energy: The formation of an ion pair within the protein interior requires the removal of the charged or polar moieties from water, including

the destruction of a solvent shell. The desolvation costs associated with this process cannot be compensated by the pairwise Coulomb interaction.[122]

Initially it was assumed that for the same reasons electrostatic interactions also oppose protein-protein association processes. But growing evidence suggests, that polar groups can contribute in some cases favorably to the stabilization of protein-protein complexes.[123]

Theoretical investigations showed that electrostatic interactions can have a stabilizing effect on protein-protein binding.[124] There it was argued that the polar environment in protein-protein interfaces allows a more favorable electrostatic interaction than it is possible in hydrophobic protein cores. This is comparable to the situation in hyperthermophilic proteins, where networks of electrostatic interactions account for the enhanced thermal stability[125] of the protein.

Mutagenesis experiments showed that in fact replacing polar and charged groups by alanine leads to the destabilization of protein protein complexes.[126] However, such mutagenesis experiments do not show, if polar or charged interactions stabilize or destabilize protein-protein complexes. If one residue of an ion pair is removed and the complex is destabilized, it demonstrates that the corresponding residue is stabilizing the complex when the other polar group is present. But the actual comparison has to be made with the reference state consisting of two isolated charged or polar monomers.

Also in the discussion of protein folding it has been shown that the removal of only one member of an isolated hydrogen-bonded pair or ion pair destabilizes the protein, even if the hydrogen bond or ion is itself destabilizing.[127] Moreover, there are theoretical studies supporting the initial hypothesis on charged and polar groups in protein-protein association by showing that electrostatic interactions can have a destabilizing effect in protein-protein binding.[128,129]

So it is still unclear to which extent electrostatic interactions can contribute to the stabilization of protein-protein complexes. On the other hand, it is known that electrostatic fields around proteins can enhance the rates of protein-protein association[130,131,132] and that electrostatic interactions have an influence on the specificity of protein-protein binding.

I will address this topic for a special protein-protein pair in the next section, where I will describe an investigation on the electrostatic interactions between dimers of human arylsulfatase A.

5.2. Introduction to arylsulfatase A

Human arylsulfatase A (ASA, EC 3.1.6.8) has its biological function as a hydrolytic enzyme in the lysosome of the human cell. Its major physiological substrate is cerebroside 3-sulfate, a sphingolipid sulfate ester, that is included in the myelin sheet (Fig. 5.1). The name arylsulfatase comes from the capability to hydrolyze also artificial arylsulfates with even higher rates than natural substrates.

ASA belongs to the sulfatase family. The members of this family hydrolyze sulfate ester bonds of a wide variety of compounds, including mucopolysaccharides, cerebrosides and steroids. Among all sulfatases ASA has been studied most extensively. With the other eight known human sulfatases it shares a high sequence similarity of 47-59 %.[133] Seven human disorders are known to be caused by a deficiency in one of the sulfatases.[134] The deficiency of arylsulfatase A leads to metachromatic leucodystrophy (MLD), where sulfate esters of galacto-cerebrosides accumulate in several organs[135,136,137]

The key residue for the activity of sulfatases is a C_α-formylglycine (FGly) that results from a posttranslational modification of a cysteine residue.[138] The newly synthesized enzyme is inactive if the posttranslational modification of this amino acid is missing.

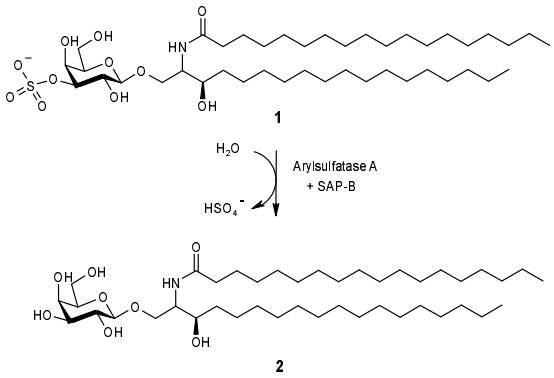


Figure 5.1.: Hydrolysis of the natural substrate of ASA, cerebroside 3-sulfate, (**1**) to galactosyl ceramide (**2**) A sphingolipid activator protein (SAP-B) is needed as a cofactor to enlarge the solubility of the sphingolipid.

In an acidic environment (pH $\approx$ 5), where the enzyme shows its optimal catalytic activity, human ASA occurs as an octamer, whereas at neutral pH it exists as dimer. A similar dissociation behavior is reported for ox liver arylsulfatase[139,140] and rabbit liver arylsulfatase.[141] It has been shown that the association between dimers also occurs between ASA monomers that were isolated from different species.[142] This finding suggests that the interface structure may be conserved.

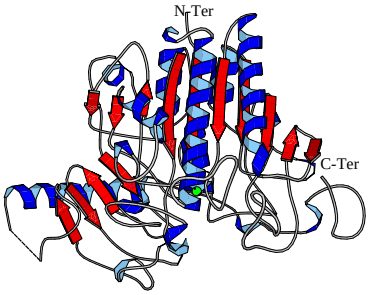
In 1998 the crystal structure of ASA was published.[143] The structure gave rise to a discussion about the catalytic mechanism of ASA as well as about the regulation of the dimer-dimer association that yields the octamer. Both topics will be addressed in this work. In this chapter the computation of protein-protein association energies of on the basis of electrostatic energies will be presented. In chapter 6 a strategy to simulate the catalytic process in ASA is outlined.

5.3. The structure of ASA

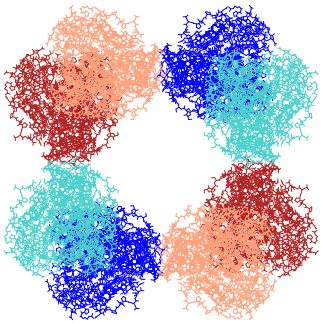
The ASA monomer (Fig. 5.2 (a)) consists of 489 residues. The globular protein is hat-shaped with a 70 Å $\times$ 45 Å base and a height of 50 Å. The secondary structure of ASA is of the mixed α/β-type. The enzyme contains an Mg^{2+}-ion in the active site. The formylglycine residue, which is most important for the catalytic activity, is located at the bottom of a cavity. The ASA monomer has a significant structural analogy to alkaline phosphatase (AP),[144,145] which is also a hydrolytic enzyme containing two zinc ions in the active site. The active sites of both enzymes are located similarly.

The x-ray structure[143] was solved from crystals that were grown at acidic pH (5.0-5.4). They consist of ASA homooctamers (Fig. 5.2 (b)). The octamers are composed of four dimers, $(\alpha_2)_4$. The monomers (α) are related by the symmetry elements of point group 422 (D_4) in the octamer. A two-fold crystallographic axis relates two ASA monomers to the homodimer.

In the octamer exist two different symmetrical monomer-monomer contact surfaces: The larger one has an area of 1600 $Å^2$ and and forms the dimers (Fig. 5.3). Within this large contact surface there are 2 $\times$ 3 direct hydrogen bonds between the monomers and 2 $\times$ 25 hydrogen bonds mediated by water molecules. Despite this small number of direct interaction the resulting interaction is very strong, such that previously the dimer was even described as a monomer.[146] The large part of stabilization of the dimer results from hydrophobic interactions.



(a)



(b)

Figure 5.2.: The ASA monomer (a) and the ASA octamer (b). In (a) the Magnesium ion in the active site is depicted as a green sphere. The dimers in (b) are colored blue and red and the monomers in the dimers are light or dark. The dimer-dimer interfaces are always between blue and red monomers.

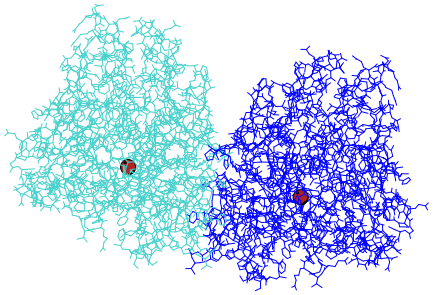


Figure 5.3.: The ASA dimer. The contact surface between the dimers is 1600 $Å^2$. There are 2 $\times$ 3 direct hydrogen bonds between the two monomers: His328N_ε $\cdots$ OPro42 (3.4 Å), Ser43O_γ $\cdots$ OSer432, (2.6 Å) and Tyr439O_η $\cdots$ O_γThr408 (2.9 Å) In addition 2 $\times$ 25 water molecules in the contact region mediate hydrogen bonds. The red spheres in each monomer are the magnesium ions in the active sites.

The contact surface that allows for the dimer-dimer cohesion covers 900 Å^2. It involves the hydrophobic interaction between the aliphatic amino acid residues of two nearly antiparallel α-helices (see Fig. 5.4) Moreover there are 2 × 3 direct hydrogen bonds between the dimers and additional hydrogen bonds mediated by 2 × 9 trapped water molecules.

The dimer octamer equilibrium is regulated via the pH value. It has been proposed earlier that this regulation is exerted by a protonation-deprotonation equilibrium of a titratable carboxylate or imidazole side chain.[147,146] The only acidic side chain at the dimer-dimer interface is Glu424. The two Glu424 of both dimers are only 5 Å apart. In the electron density map they are found two-fold disordered. One of the carboxylate positions is suitable for a hydrogen bond to the backbone oxygen of Phe398 of the other dimer. This hydrogen bond requires a protonated carboxylate side chain and would stabilize the octamer. In the other conformation the carboxylate side chain is involved in an intramolecular hydrogen bond to the N_ε atom of Gln460. This conformation would favor the dimer and the side chain of Glu424 needs to be unprotonated to exert the intramolecular hydrogen bond. Both conformations of Glu424 are depicted in Fig. 5.5.

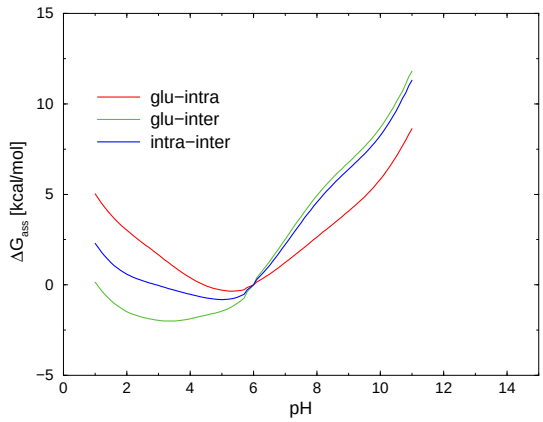


Figure 5.4.: The contact region between the two dimers. Two antiparallel α-helices are the major constituents of the adhering surfaces. The surface covers an area of 900 Å^2. There are 3 × 2 direct hydrogen bonds between the two dimers: Lys457N_ζ ⋯ OGlu382 (3.0 Å), Lys457$_\zeta$ ⋯ OAsp467, (3.0 Å) and Glu424O_ε ⋯ OPhe398 (2.9 Å) In addition there are 2 × 9 water molecules in the contact region that mediate hydrogen bonds.

The dimer-octamer equilibrium may be explained by a switch function of Glu424: At around pH 5 Glu424 is protonated and stabilizes the octamer due to the intermolecular hydrogen bond to the backbone oxygen of Phe398. At pH 7 Glu424 is unprotonated and the hydrogen bond to Phe398 is no longer possible. Moreover the octamer will be destabilized by the electrostatic repulsion between the two charged glutamic acids that are only 5 Å apart. In the unprotonated form, the preferred conformation of Glu424 is the one that allows for the intramolecular hydrogen bond to Gln460. With this argument the electrostatic interactions and the protonation-deprotonation equilibria are the most important factors for the regulation of the association process. To investigate the association behavior within the light of this hypothesis it is necessary to account for the protonation probability of Glu424 in the dimer and in the octamer and the free energy of association at various pH values.

5.4. Electrostatic energies of dimer-dimer association

For the electrostatic computations of the dimer-dimer association only a part of the whole molecular system consisting of eight monomers was considered in this study. The association of dimers occurs at the 900 Å^2 interface of the subunits. So the computations of the dimer-dimer association were actually done with monomers associating via the proper interface. This association leads to a "secondary dimer", which has to be distinguished from the "primary dimer", that is in equilibrium with octamers and predominating at higher pK$_a$ values as explained in the previous section. The secondary dimers do not exist isolated in solution but serve as a model system in the present computations.

This procedure shrinks the system to a size with limited computational demand. The calculation of the association energy requires to solve the LPBE of the protein numerically several times. In this procedure all protein charges are mapped onto a grid. A suitable grid for the octamer would have too many points too yield results in reasonable time. The limitation of the system size makes it nevertheless possible to investigate the important forces in the interface that are responsible for the association.

For the sake of simplicity I will from here on use the (chemically incorrect) term "conformation"when referring to the monomers and the artificial dimer. This is done for reasons of consistence with the notations of the method that is applied to sample the protonation and the conformational states of the protein. In the theoretical outline of this method in section 2.4 the term conformation is already used. The difference between isolated monomers and the associated artificial dimer is now categorized as a conformational difference to stick to the notions used in section 2.4.

The electrostatic energy of association is calculated by two methods in this study. First, the populations of the two conformations are investigated by an MC importance sampling of conformations, yielding the population distributions between the monomers and the artificial dimer. With the inclusion of correct reference energies this leads in principle to the absolute free energies of association. As an alternative approach the application of the proton linkage model is presented here. With this method the free association energy in dependence on the pH value is determined. With appropriate experimental informations also in this approach absolute values of association energies can be calculated for different pH values.

5.4.1. MC sampling of conformations

The protonation probability of a titratable group depends on the pH value as can be seen from Eq. 2.18. Also the conformation of side chains may depend on the pH value. The conformation itself has an influence on the protonation state of a titratable group. This becomes obvious by inspecting Eq. 2.25, where the energy difference between a protein conformation and a reference conformation contributes to the total energy of a protein state.

The consideration of different conformations is applied in this study to the case of different association states of a protein. The protonation pattern of two protein monomers can be established for the isolated monomers and for the monomers associated to the corresponding dimer. With the Monte Carlo importance sampling of states one gets the population probability of each association state. The free energy difference between the isolated monomers and the dimer is then given by

$$\Delta G_{ass} = -RT \ln \frac{\langle monomor \rangle}{\langle dimer \rangle} \tag{5.1}$$

Where ΔG_{ass} is the association energy and $\langle monomer \rangle$ and $\langle dimer \rangle$ are the population probabilities of the monomer and the dimer, respectively.

5.4.2. The proton linkage model

The association constant of a protein binding process can be measured via the proton release of the system accompanying the binding. This has been done frequently in experimental investigations.[148,149,150,151,152] The proton release can also reveal the free energy changes going along with processes like redox reactions[153,154] and protein unfolding.[155,28,156]

The proton release going along with a process can be used to determine free energies by applying the so-called *proton-linkage-model*. This model relates the pH-dependence of an equilibrium constant to proton release. So the proton-linkage model can be used to investigate the pH-dependence of processes like redox-reactions, conformational transitions and also the binding of proteins.

The association between two proteins is described by

$$A + B \rightleftharpoons AB; \qquad K_A = \frac{[AB]}{[A][B]} \tag{5.2}$$

where K_A represents the association constant. The proteins A and B and the protein-protein complex can exist in the protonation states L_A, L_B and L_{AB}, respectively. The total concentration of each species is then given as the sum over all protonation states and Eq. 5.2 becomes

$$K_A = \frac{\sum_{i=0}^{L_{AB}} [AB]_i}{\sum_{i=0}^{L_A} [A_i] \sum_{i=0}^{L_A} [B_i]} \tag{5.3}$$

where $[A_i]$, $[B_i]$ and $[AB]_i$ denote the concentration of A, B and AB in the respective protonation state i. As the the concentrations $[A_i]$, $[B_i]$ and $[AB]_i$ vary with the pH, also K_A is pH-dependent. It was shown that the equilibrium constant K_A depends on the total proton release via[157]

$$\lg K_A(pH_2) = \lg K_A(pH_1) + \int_{pH_1}^{pH_2} \Delta n(pH)\, dpH \tag{5.4}$$

Eq. 5.4 indicates, that if the association constant $K_A(pH_1)$ at pH_1 is known, the association constant $K_A(pH_2)$ at pH_2 can be obtained by integrating over the number of protons $\Delta n(pH)$ that are released upon association. The free energy of the process is then given by

$$\begin{aligned} \Delta G_{12} &= G(pH_2) - G(pH_1), \\ \text{where } G &= -RT \ln K(pH) \end{aligned} \tag{5.5}$$

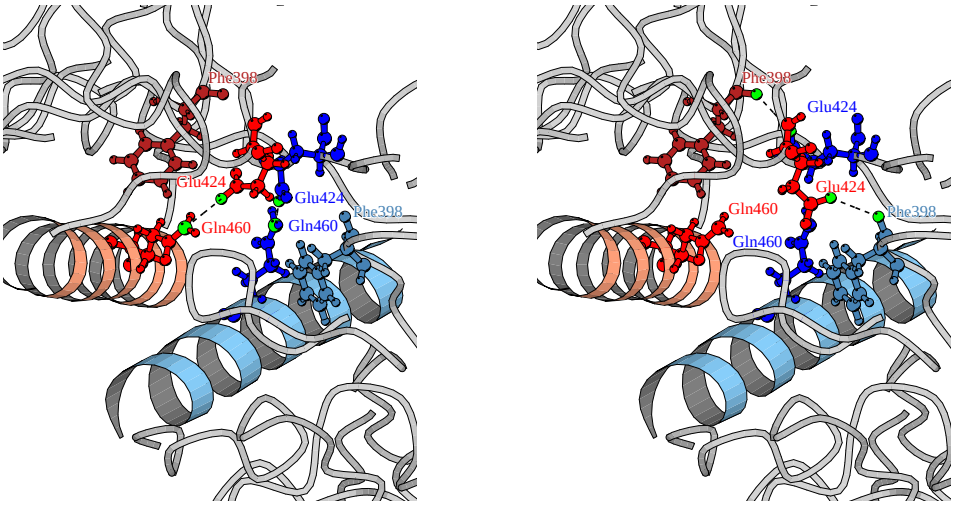
With the knowledge of an equilibrium constant at one pH-value, one is able to calculate equilibrium constants at other pH values, provided that the protonation states at other pH values are accessible. With this method the pH dependence of the free energy of association can be calculated.

5.4.3. The model system

Arylsulfatase A dimers in solution associate to octamers, when the pH value is low enough (pH $\approx$ 5). In this work the association of dimers was modeled by an association of two monomers via the dimer-dimer interface as mentioned above. That means, I considered the association of monomers to dimers only. The association was restricted to the association of monomers.

To investigate the role of Glu424 upon the association of the protein monomers, both possible conformations of this residue, as they are shown in Figure 5.5, wre considered.

1. Glu424 in the conformation suitable for the intramolecular hydrogen bond to Gln460.
2. Glu424 in the conformation suitable for the intermolecular hydrogen bond to Phe398.



(a) (b)

Figure 5.5.: Glu424 hydrogen bonds at the dimer-dimer interface. In figure (a) and figure (b) the residues involved in intra- and intermolecular hydrogen bonds at the dimer dimer interface are shown. The residues of the two subunits are colored red and blue respectively. In figure (a) the Glu424 residues of both subunits are in the conformation suitable for the intramolecular hydrogen bond with the side chain of Gln460 of the same subunit. The oxygen atom of the Glu424 side chain and the nitrogen of the Gln460 side chain, which are the the non-hydrogen atoms of the hydrogen bond are colored green. In (b) the conformation of Glu424 changed by a rotation around the C_β-C_γ bond and is now suitable for a hydrogen bond to Phe398 of the other subunit. The intramolecular hydrogen bond in figure (a) needs no protonated carboxyl group in the Glu424 side chain, whereas for the intermolecular hydrogen bond in figure (b) the protonation of the carboxylate side chain of Glu424 is prerequisite.

In the crystal structure 2×9 water molecules were found that mediate hydrogen bonds within the interface between dimers. These water molecules were not considered explicitly in this study. Although their function might contribute to binding it is not necessary to include them explicitly in electrostatic calculations. The binding sites of these water molecules were treated with a dielectric constant of $\varepsilon_s = 80$ in the monomers as well as in the secondary dimer. With this

procedure the reference states are properly defined, because one does not have to consider the problem whether the water molecules are bound also to the monomer or just in the secondary dimer, when the subunits have already associated. With the chosen procedure the number of explicitly treated atoms is the same in the monomer and in the secondary dimer. All solvent molecules were treated via the dielectric constant.

The first results of the present calculations were the titration curves of Glu424, that is expected, as mentioned above, to show a preference for the protonated state in the secondary dimer and for the unprotonated state in the monomer. Additionally, the protonation probability should vary significantly with pH to demonstrate the switch function postulated for Glu424.

Besides the protonation behavior of Glu424 at various pH values in the investigated conformations, it was important to see if the association energy between monomer and secondary dimer in the model system changes with varying the pH. This should hopefully agree with the pH regulation of the association process.

5.4.4. Methods

Structures and Charges

The crystal structure of ASA was solved at a resolution of 2.1 Å and is deposited as the entry 1auk in the Protein Data Bank[158] with the atomic coordinates for one monomer. In the structure the first 18 amino acids are missing. These residues form a signal peptide, which is not present in the active protein. So they were not considered in this study. Moreover the residues Gly444 to Gly447 and Asp504 to Ala507 could not be identified from x-ray structure analysis. They were modeled into the structure with the program CHARMM.[91] The hydrogens were added to the structure using the HBUILD facility of CHARMM. The structure was then energy minimized with the CHARMM force field[92] by a protocol that fixed all atoms whose positions are known from the crystal structure. So only atoms that were added by modeling found new positions according to the known structure and the force field.

The secondary dimer that results from the association of monomers via the dimer-dimer interface was generated with the program PDBSET from the CCP4 program suite by applying the corresponding two fold-symmetry operation. After the creation of the artificial dimer all hydrogen atoms were energy minimized again such that the hydrogen atoms adopted their position in correspondence to the dimeric structure.

Although the structure was solved at an acidic pH value (pH 5.0-5.4), where the protein exists as octamer, the conformation of Glu424 in the PDB entry is the one corresponding to the dimer, where Glu424 forms an intramolecular hydrogen bond to Gln460. The intermolecular hydrogen bond was modeled with the program CHARMM by rotating the side chain around the torsion angles C_β-C_γ and C_γ-C_δ, such that the distance between the carboxylate oxygen of Glu424 and the backbone oxygen of Phe398 became 3.1 Å. This is done in both monomers. The distance of 3.1 Å is the shortest possible distance between the partners of the intermolecular hydrogen bond without putting some additional restraint on the system. The side chain of Glu424 was then also energy minimized while fixing the atoms of all residues but Glu424.

The described modeling procedure yielded four different structures. The abbreviations in capital letters are used in figures below, see also Fig. 5.5 (a) and (b)

- two monomers: (a) with Glu424 in a conformation corresponding to the intramolecular hydrogen bond (MONO:GLU-INTRA). (b) with a Glu424 conformation corresponding to the intermolecular hydrogen bond (MONO:GLU-INTER).

- two secondary dimers, which consist of monomers with the same Glu424 conformations (DIMER:GLU-INTRA), (DIMER:GLU-INTER).

Reference energies of different conformations

As outlined in section 2.4, a complete description of the state of a protein must comprise its conformation and its protonation state. For the sampling of protonation states and conformations Eq. 2.25 has to be evaluated. For that purpose, ΔG^{l}_{conf}, the energy difference of a conformation l to an arbitrarily chosen reference conformation r, has to be calculated for the same protonation state. As pointed out in the last section, four different conformations are considered. To calculate association energies it is necessary to sample the distribution of the population between two monomers on one side and the corresponding secondary dimer on the other side. The monomers that have all their titratable groups in the uncharged protonation state are then always chosen to be the reference conformation r.

For all four conformations, the energies necessary to determine ΔG^{l}_{conf} in Eq. 2.26 are calculated. For a detailed description of all terms see section 2.4. They were all calculated with all titratable groups in their uncharged protonation state. Only all histidines were chosen to be charged to overcome the ambiguity problem of the neutral His, where the proton can be bound to N_δ or N_ε. The energy difference to the actual reference state, where all residues have to be neutral, is accounted for afterwards when the states are sampled with the Monte-Carlo approach. The solvation energy ΔG^{l}_{S} was calculated with the program SOLVATE from the MEAD software package.[159,160] The homogeneous dielectricum was chosen to have a dielectric constant of $\varepsilon_S = 4.0$ for the surrounding medium and a dielectric constant $\varepsilon_P = 4.0$ for the protein as well. The ionic strength for the homogeneous dielectricum was 0.0 M. In the case of the heterogeneous dielectricum a dielectric constant $\varepsilon_S = 80.0$ was chosen for the solvent and a dielectric constant $\varepsilon_P = 4.0$ was chosen for the protein interior. The ionic strength was 0.1 M. The solvation energy was calculated in two focusing steps: For the dimer the side length of the grid was 242 Å and 140 Å with a grid spacing of 2.0 Å and 0.5 Å, respectively. The grid for the monomer had a side length of 202 Å and 101 Å with a grid spacing of 2.0 Å and 0.5 Å respectively. For the molecular system in homogeneous dielectric medium the electrostatic energies were calculated by simply calculating the Coulomb energy (Eq. 2.1) for all atom pairs with the charges from the CHARMM22 force field.

Electrostatic potentials and Monte Carlo titration

The electrostatic potentials, which were required for the titration of all structures, were calculated with the program MULTIFLEX from the MEAD software package. The dielectric constant of the solvent was set to $\varepsilon_S = 80.0$. The dielectric constant of the protein interior was set to $\varepsilon_P = 4.0$. The electrostatic potentials in the protein were calculated with three focusing steps: For the dimer the grid size was reduced from 242 Å via 107 Å to 18.75 Å with a grid spacing of 2.0 Å, 1.0 Å and 0.25 Å respectively. For the monomer grid sizes of 150 Å, 75 Å and 18.75 Å were applied, the grid spacing was 2.0 Å, 1.0 Å and 0.25 Å respectively. For the model compounds a grid size of 15 Å with a grid spacing of 0.25 Å was applied. The ionic strength was set to 0.1 M.

MC sampling of conformations

In each Monte Carlo sampling setup, two corresponding conformations were investigated to establish the distribution of the population between them. Three pairs resulted from this strategy:

1. Two monomers with Glu424 in the conformation suitable for the intramolecular hydrogen bond and the secondary dimer with Glu424 in the same conformation. 2. Two monomers with Glu424 in the conformation that corresponds to the intermolecular hydrogen bond and the secondary dimer with the same conformation. 3. Two monomers with Glu424 in the conformation that corresponds to the intramolecular hydrogen bond and the secondary dimer with the conformation of Glu424 corresponding to the intermolecular hydrogen bond. This last pair accounts for the hypothesis that in the monomer the conformation suitable for the intramolecular hydrogen bond should be favorable and in the dimer the conformation corresponding to the intermolecular hydrogen bond is predominating.

The titration calculations were done with the program KARLSBERG.[161] Each pair of conformations was introduced into KARLSBERG simply as two conformations. As the program requires that all treated conformations have the same number of titratable groups and to make sure that the reference energies refer to the proper reference state, which is *two* monomers in solution and one secondary dimer, the monomers were also represented as dimers of infinitely distant monomers.

The titration was made at eight different pH values for each pair ranging from pH 3 to pH 11. Within one MC titration scan (see section 2.5) two conformational moves were made. The sampling of conformations turned out to have a very low efficiency. Therefore, the conformation that was poorly populated was given a bias energy to raise the population and to increase the sampling efficiency and to reduce the statistical error. From the populations of both conformations the free energy difference between them can be calculated with Eq. 5.1. This energy still includes the used bias energy but it can be rebiased simply by subtracting the used bias energy since it is additive.

The usage of a bias was not sufficient to yield a sampling efficiency that lead to an error level which is acceptably small. In addition, the parallel tempering approach[32] had to be applied to raise the sampling efficiency. The sampling was done at the following temperatures: 300 K, 400 K, 530 K, 710 K, 950 K, 1266 K and 1688 K. Within one MC titration scan 20 tempering moves were made. For the sampling, 1000 full MC scans, including all titratable groups were made. Groups that remained in a pure protonation state (fully protonated or unprotonated) and did not change their protonation state during the first 1000 MC scans were kept at their protonation state and were excluded from the list of titratable groups. With the resulting reduced set of titratable groups another 10000 MC scans were made.

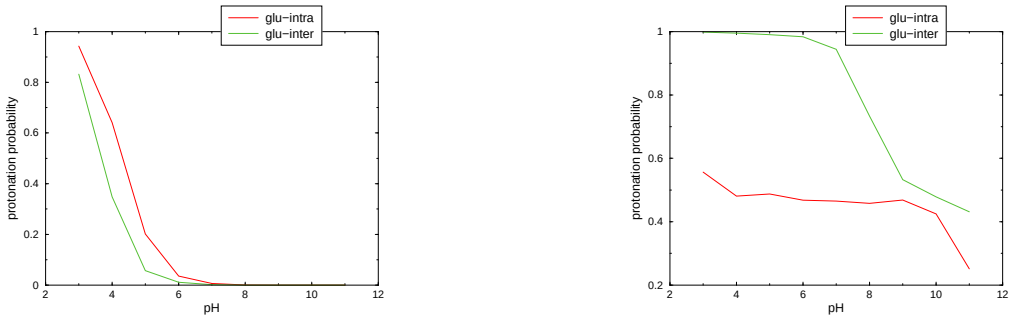
Titration for the proton linkage investigation

To investigate the pH dependent free energy of association by the proton linkage model, the monomer and the secondary dimer are titrated separately by the program KARLSBERG. The same number of full and reduced scans as in the conformational sampling approach were made. The number of protons released upon association is simply the difference of the total number of protons bound to the secondary dimer and to the monomer. The number of protons bound to the monomer has to be doubled to be comparable to the number of protons bound to the dimer. The titration was made for the same six conformations that were investigated in the conformational sampling approach. Each titration comprised 80 pH values ranging from 3.0 to 11.0 with intervals of 0.1.

5.5. Results and Discussion

5.5.1. The protonation state of Glu424

The suggested switch function of Glu424 for the dimer-dimer association should result in a corresponding titration behavior of this residue in the four conformations that were investigated in this study. The titration curves for the four considered conformations are shown in Fig. 5.6.



(a) (b)

Figure 5.6.: The protonation states of Glu424 in the monomer (a) and the dimer (b). Monomer and dimer occur in two different forms each. Glu-intra is the conformation of Glu424 corresponding to the intramolecular hydrogen bond. Glu-inter is the conformation of Glu424 corresponding to the intermolecular hydrogen bond.

The titration properties of the two monomer structures are very similar. The protonation probability decreases from a value larger than 0.8 at pH 3.0 to a value of 0.0 at pH 7.0. The significance of the conformation of Glu424 is very small for the titration behavior of the monomer. The conformation that includes an intramolecular hydrogen bond behaves the same as the conformation where Glu424 points out of the structure.

The titration curves of the dimers show a different behavior. The conformations of Glu424 suitable for the intermolecular hydrogen bond show a significantly higher protonation probability than the conformation that corresponds to the intramolecular hydrogen bond. For the intermolecular hydrogen bond with the backbone oxygen of Phe398 a protonated Glu424 is required. Hence, the calculated titration curves support very well the formation of an intermolecular hydrogen bond with protonated Glu424.

When Glu424 is in the conformation, where an intramolecular hydrogen bond is formed, it should be deprotonated in the secondary dimer at higher pH values. The behavior should be comparable to the corresponding monomer. But as in the associated conformation both glutamic acids are only 5 Å apart, this would lead to an energetically unfavorable situation, because of two neighbored negative charges. So the deprotonation would have to be accompanied by a dissociation of the secondary dimer. As long as the monomers are associated to the secondary dimer, as is the case in the present calculation, a complete deprotonation of Glu424 is not probable.

The importance of Glu424 for the association process is also underlined by the fact that no other titratable group changes its protonation state significantly upon association. If one compares the protonation probability of any titratable group in the monomer with the one in the dimer at

any pH value, the change will be always smaller than 0.1. Only for Glu424 the protonation probability varies considerably (dependent on the pH and the conformation) by up to 0.9.

5.5.2. The association free energy

MC sampling of conformations

To calculate the association energy on the basis of the population of the two different conformations evaluated by Monte Carlo sampling, the reference energies of the two conformations have to be considered. Between each pair of conformations the reference energy difference has to be calculated according to the description in section 2.4. The reference energies do not depend on the pH and the differences between each pair of conformations were found to be rather similar as can be seen from Table 5.1. The values are between -23.0 kcal/mol and -24.3 kcal/mol. This similarity occurs even though the individual contributions to the reference energies are large numbers as can be seen from Table 5.2. For the proportionality factor γ of the surface-dependent non-polar contribution ΔG_{NE} in Eq. 2.28 I chose a value of 20 cal mol^{-1} Å^{-2}. As the buried surface of the interaction between the monomers in this study covers 900 Å^2, the change in solvent accessible surface upon association is 1800 Å^2 leading to a contribution of 36 kcal/mol. This value is positive for the monomers, *e. g.* it favors the association as the surface decreases upon association. The chosen value of 20 cal mol^{-1} Å^{-2} for the proportionality factor is in correspondence with data from literature,[24,25] but these empiric data may not be valid for all systems. Therefore the chosen value is only a first approximation and a correction will be suggested below.

conformational pair	$\Delta\Delta G_{ref}$
MONO:GLU-INTRA → DIMER:GLU-INTRA	-24.3
MONO:GLU-INTER → DIMER:GLU-INTER	-23.0
MONO:GLU-INTRA → DIMER:GLU-INTER	-23.2

Table 5.1.: The energy differences between the conformations were calculated via the thermodynamic cycle in Fig 2.3. All energies are in kcal/mol

conformation	ΔG_S	ΔG_{FF}	ΔG_{NE}	$\sum$
MONO:GLU-INTRA	-2866.04	-41926.53	36.00	-44263.87
MONO:GLU-INTER	-2378.06	-41922.01	36.00	-44264.07
DIMER:GLU-INTRA	-2866.04	-41422.12	0.00	-44288.17
DIMER:GLU-INTER	-2867.14	-41922.01	0.00	-44287.07

Table 5.2.: The energies of the different monomer and dimer conformations investigated for the association process of ASA. ΔG_S is the solvation energy resulting from the transfer of the corresponding system from a homogeneous dielectric environment to a heterogeneous dielectric environment as explained in section 2.4. ΔG_{FF} is the energy resulting from changing the conformation in the homogeneous dielectric and is calculated as described in section 2.4. ΔG_{NE} is the energy change resulting from the change of solvent accessible surface upon association. All energies that belong to a monomer are doubled, such that they are comparable with the corresponding dimer. The conformations of the monomers and the dimers are named as described in the text. All energies are in kcal/mol

The association energies of the monomers via the dimer dimer interface, that result from the MC sampling including the reference energies, were found to be pH dependent. The values of the free association energy in dependence on the pH can be seen in figure 5.7.

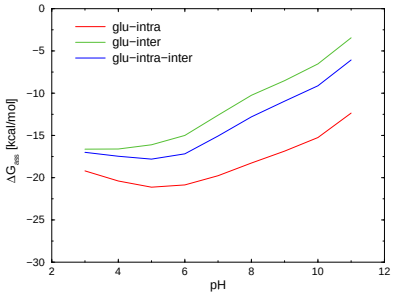


Figure 5.7.: The association free energies of the monomers via the dimer-dimer interface. All curves include the reference energies from Table 5.1. The association energies were calculated with Glu424 in two different conformations: The red curve shows the free energy of association of the system when Glu424 adopts the conformation suitable for the intramolecular hydrogen bond. The green curve shows the free association energy of the system, when Glu424 adopts the conformation suitable for the intermolecular hydrogen bond. The blue curve shows the association energy for the system, when Glu424 adopts the conformation suitable for the intramolecular hydrogen bond in the monomeric state and the conformation that is required for the intermolecular hydrogen bond in the associated state.

All association energies are negative. That signifies that the association to the artificial dimer is favored at each pH value. In correspondence with the experimental finding, that at higher pH values the octamer dissociates to dimers, the absolute value of the association energy decreases with an increasing pH value.

The association processes with Glu424 in the conformation suitable for the intramolecular hydrogen bond have a minimum value for the free energy around pH 6.0. A minimum at this pH value is not consistent with the hypothesis that around pH 6 the dimer should dissociate. The dissociation at pH 6 requires, that the association energy is negative at lower pH values and positive at higher pH values than 6. The shape of the present curve may indicate that the association will not occur with Glu424 in the conformation for the intramolecular hydrogen bond. In addition it can be seen from Fig. 5.6 that the protonation of both Glu424, which is favorable for the association process cannot be accomplished completely even at small pH values with Glu424 in the intramolecular hydrogen bond. So from the present calculations it is required that Glu424 adopts the conformation suitable for the intermolecular hydrogen bond to Phe398 to support the association process.

The free energy curves for the association process with Glu424 in a conformation that is in accordance with the intermolecular hydrogen bond confirm this: They show a behavior as one expects it for a dissociation at neutral pH. The association energy is more negative at small pH values than at higher pH values and shows no minimum, where the dissociation should occur.

The problem with these curves are the absolute values of the association energies. As the energies are negative at all pH values, a dissociation should not occur. That may be explained by the uncertainty of the values of the reference energies. From Table 5.2 it can be seen that

the contributions to the reference energies have large absolute numbers. They cancel to values of around -23 kcal/mol to -24 kcal/mol for the individual pairs. Especially the proportionality factor γ that is responsible for the non-electrostatic contribution to the solvation energy cannot be considered to be valid for all proteins as discussed above. To overcome the problem with the reference energies, the theoretical model is enlarged and the fact that at around pH 6.0 the octamer dissociates into dimers is taken into account. This is done by adjusting all association energy curves such that at pH 6.0 the association energy is 0.0 kcal/mol. These adjusted curves can be seen in Fig 5.8.

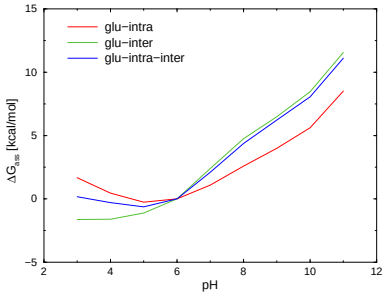


Figure 5.8.: The association energies shifted to reproduce the dissociation at pH 6.0

The three curves, GLU:INTRA, GLU:INTER, GLU:INTRA-INTER had to be shifted by +20.8 kcal/mol, +15.0 kcal/mol and +17.2 kcal/mol, respectively. That means, the reference energies would have to be less negative. The corrected reference energies are listed in Table 5.3. The correction of the reference energies is justified because of the uncertainty of the surface factor γ. The initially chosen value of 20.0 cal mol^{-1} Å^2 was parameterized for completely hydrophobic surfaces consisting of ethyl and propyl groups[24] The dimer-dimer interface of arylsulfatase consists of 21 residues. Only 12 of them are hydrophobic. This means that a reduction of the surface factor to 50-60 % of the initial value would be necessary. The three pairs of monomers and secondary dimers need the surface factors given in Table 5.3 to have a free association energy of zero at pH 6.0. These factors range from 8.4 cal mol^{-1} Å^2 to 11.6 cal mol^{-1} Å^2 Compared with the value for completely hydrophobic surfaces of 20 cal mol^{-1} Å^2 this corresponds well with a reduction of around 50 %. The nonpolar contribution of the association energy at the dimer-dimer interface of ASA amounts to approximately -18 kcal/mol with these adopted surface factors. As at pH 5 the association energy is around -2 kcal/mol it can be deduced that the electrostatic contribution opposes binding.

Proton-Linkage Method

The critical parameter that is used in the proton linkage model is the proton release upon the association of the monomers. This is the difference between the number of protons bound to the monomer and the number of protons bound to the dimer. This difference in dependence on the pH value can be seen in Fig. 5.9 for the three systems described in section 5.4.3.

From Fig. 5.9 it can be seen that the proton release upon association is dependent from the pH value in all three systems. The integration of these curves between a pH value with a known

conformational pair	$\Delta\Delta G_{ref}$	γ
MONO:GLU-INTRA $\longrightarrow$ DIMER:GLU-INTRA	-3.5	8.4
MONO:GLU-INTER $\longrightarrow$ DIMER:GLU-INTER	-8.0	11.6
MONO:GLU-INTRA $\longrightarrow$ DIMER:GLU-INTER	-6.2	9.3

Table 5.3.: The corrected reference energies between the different conformations. The reference energies from Table 5.2 are corrected such that the association energy of the corresponding monomer-dimer pair is 0.0 kcal/mol at pH 6.0, which is roughly consistent with the experimental observations. All energies are in kcal/mol. The surface factor γ is in the unit of cal mol^{-1} Å^2

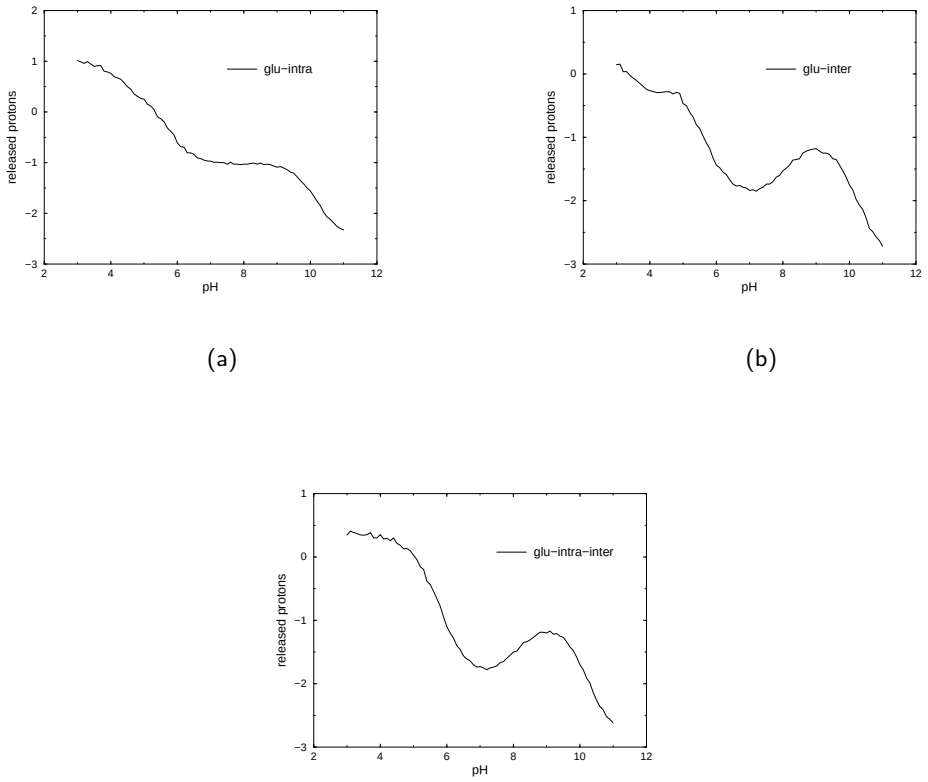


Figure 5.9.: The proton release upon association at different pH values. The systems in (a)GLU:INTRA (a)GLU:INTER and (c)GLU:INTRA-INTER correspond to the monomer-dimer pairs as described in section 5.4.3. A negative value for the proton release has to be considered as proton uptake upon association.

equilibrium constant (pH_1) and a chosen pH value (pH_2) leads to the equilibrium constant at pH_2 corresponding to Eq. 5.4. In the present example the association constant at pH 6.0 is chosen to be 1.0. That is the rough correlation to the experimental finding, that the octamer dissociates to dimers at a pH value between 5 and 7. The free energies of association are determined from the equilibrium constants via Eq. 5.5. The free energy curves for all systems are shown in Fig. 5.10.

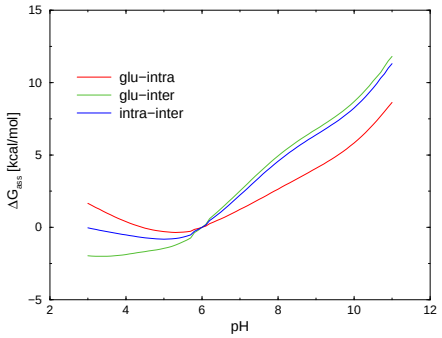


Figure 5.10.: The free energies of association evaluated by the proton linkage model. At around pH 6 the free energy of association is approximately 0.0 kcal/mol in all three systems. The association energy increases with the pH value

The results from the calculations with the proton-linkage model are similar to the results from the MC-sampling of conformations. The association free energy is zero kcal/mol at pH 6, which is implicated by the parameterization of the proton linkage method, where the equilibrium constant was set to 1 at pH 6.0. The curves for Glu424 in a conformation suitable for the intramolecular hydrogen bond show a minima around pH 6.0. This does not correspond to the experimental finding, that the secondary dimer should be stable at pH values smaller than pH 6.0 and unstable at higher pH values.

The present calculations suggest, that the association at the dimer-dimer interface occurs with Glu424 in a conformation suitable for the intermolecular hydrogen bond. This was also suggested from the crystal structure analysis. From these calculations it could not be shown, that Glu424 adopts the conformation corresponding to the intramolecular hydrogen bond in the monomer or in the dimer. That means that it could stay in the conformation corresponding to the intermolecular hydrogen bond even in the non-associated state. The switch function of Glu424 would then be restricted to the change in the protonation state, when changing the pH value.