Testing membrane protein native state discrimination by FACTSMEM

Evaluierung von FACTSMEM bezüglich des Unterscheidungsvermögens für den nativen Zustand von Membranproteinen

Bachelor Thesis

Heinrich-Heine-Universität

handed in by
Matthias Henzgen

25th of February, 2014
Düsseldorf

Supervisor
Jun.-Prof. Dr. Birgit Strodel
Examiner
Jun.-Prof. Dr. Birgit Strodel
Co-Examiner
Prof. Dr. Christel Marian

Declaration

I, Matthias Henzgen, born on the 22nd of April, 1991, declare, that this thesis is the outcome of my own work. I did not use any other source than those mentioned and every citation is identifiable as such.

Düsseldorf, 25th of February, 2014
Matthias Henzgen

Acknowledgement

First of all I want to thank Junior Professor Birgit Strodel for her kind and elaborate aid and guidance during my Bachelor Thesis. The same goes for Professor Christel Marian, whom I would like to thank for introducing me to theoretical and computational chemistry and for acting as second reviewer of this Bachelor Thesis.
To the Multiscale Modelling Group - in particular to my office mates Martin, Falk and Philipp: Thank you for the interesting discussions about science and computational chemistry and for helping me through the tricky parts of my work.
Last but not least to my family and girlfriend: I am very grateful for the motivation and encouragement you gave me during my studies.

Abstract

The recently developed implicit membrane model FACTSMEM is tested for its ability of native state discrimination of transmembrane (TM) proteins. Five TM proteins are included in the test, and 450 decoys generated by the structure prediction program ROSETTA are considered. The performance of FACTSMEM is evaluated via relative potential energies and Z-scores. The results are compared to the ones obtained with the membrane models IMM1, GBSAIM, GBSW and HDGB according to the benchmark of Yuzlenko and Lazaridis published in their paper "Membrane Protein Native State Discrimination by Implicit Membrane Models".

However, it was not possible to reproduce the results of Yuzlenko and Lazaridis when using the same membrane models and simulation specifications as in their work. Therefore, it is not possible to make any conclusions with regard to the ability of FACTSMEM (and the other implicit membrane models) to discriminate the native state of membrane proteins. The reason for this are structural problems inherent to the decoys. Therefore, to assess the performance of FACTSMEM, this work should be repeated using another set of decoys.

Contents

1 Introduction

1.1 Transmembrane Proteins

Proteins form one of the most important type of biological macromolecules. They are of utmost importance in catalysing chemical reactions in organisms, in passing information and regulating the transport of all kind of compounds throughout an organism. Transmembrane proteins in particular play the key role in the latter processes.[1]
Therefore, revealing their complex structures and understanding folding and aggregation processes are central issues when trying to resolve basic processes in organisms.

Knowledge about them is for example crucial in designing drugs. Many pathogenic germs like virus depend on host cells to survive and grow. To inhibit them from entering cells is therefore a promising way to cure illnesses and this fact makes transmembrane proteins the main target in drug design.[2]

1.2 The problem

The question now is how TM proteins can be studied. Because of the complexity of protein structures and fast cell processes, it is challenging to study them via experiments. As the name "transmembrane protein" already suggests, TM proteins are always surrounded by membranes. This anisotropic environment increases complexity even more. Therefore structure elucidation is connected to various experimental difficulties and especially aggregation and folding processes could only be resolved for a few TM proteins (e.g. bacteriorhodopsin).[3]

Theoretical models and approaches like computer simulations are a possible alternative to work on this topic. Folding and aggregation processes can, for example, be studied via molecular dynamics (MD) simulations.
Here, the problem is the size of the simulation system and the time scale one has to reach (at least several micro seconds). Studying TM proteins one has not only to simulate the protein itself, but also the cell membrane as it is of high importance in determining the structure. Taking this into account and performing MD simulations in explicit membrane solvent, the system size grows rapidly and the simulation effort rises drastically.
To save CPU time and the cost of such MD simulations describing a cell membrane via implicit solvent models, which only consider the physico-chemical appearance of lipid membranes, is a promising possibility.[4]

1.3 The aim of this thesis

FACTSMEM, recently developed by Martin Carballo Pacheco et al.,[5,6] is one example for such an implicit membrane model. Using implicit solvent or membrane approaches to predict protein structure and folding, one has to make sure that these models are able to reproduce experimental observations in a proper way and that the assumptions do not influence the results in an inappropriate manner.

Therefore, this thesis aims at testing the discriminative power of FACTSMEM in distinguishing between native state and misfolded decoys of membrane proteins. Furthermore, the performance shall be compared to the results Yuzlenko and Lazaridis (YaL) obtained in their work on testing the membrane models IMM1,[7] GBSAIM,[8] GBSW[9] and HDGB[10] published in their paper "Membrane Protein Native State Discrimination by Implicit Membrane Models"[11] to benchmark FACTSMEM.

2 Background

2.1 Membranes and Proteins

Membranes are phospholipid bilayers (figure 1) which are very important for the building of cells as they form the outer barrier. They protect the interior of cells and inhibit the free mass transport in and out of the latter.

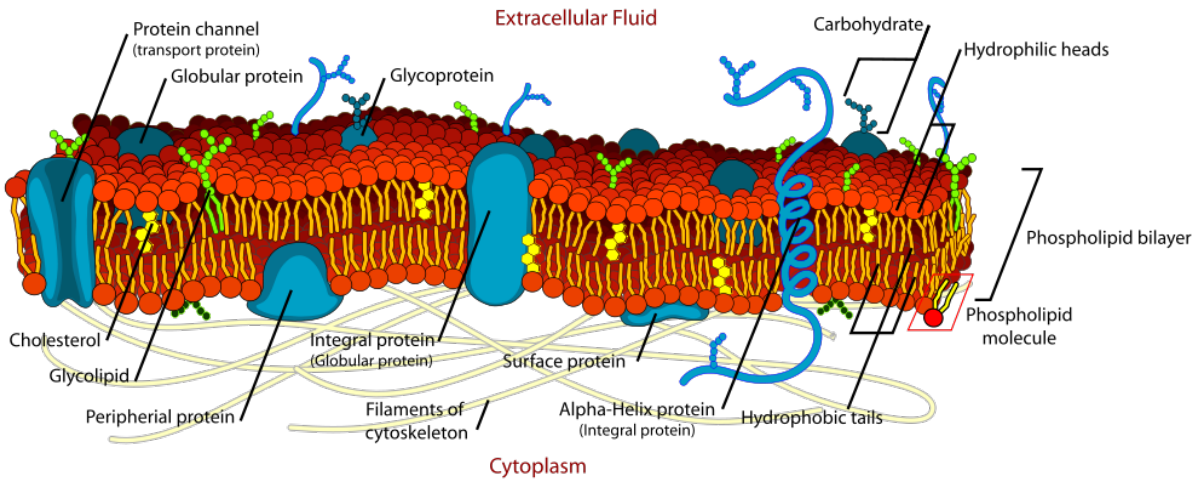

Figure 1: A cell membrane and its components is illustrated. The figure is taken from Wikipedia.[12]

Membranes are built of many different types of lipids, for example phospholipids (figure 2) which are amphiphilic molecules made by esterification of phosphoric acid, glycerine and fatty acids. Amphiphilic means hereby that they consist of a hydrophilic head group (the phosphoric acid) and hydrophobic tails (the fatty acids).[1]

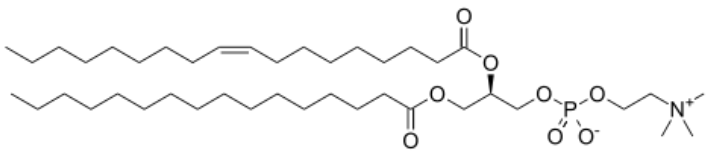

Figure 2: 1-Palmitoyl-2-oleoyl-phosphatidylcholine is shown as a sample phospholipid. The figure is taken from Wikipedia.[13]

Proteins are macromolecules consisting of an amino acid sequence, which in organisms is encoded in the genes. They are from extraordinary high importance for any biological process concerning life. First of all catalyzing cell reactions - without that most chemical reactions would not be possible at life friendly conditions - is to be named, but also signal transduction in intercell communication, mass transport throughout an organism and in

and out of cells are important tasks of proteins.
Transmembrane proteins form the group of proteins being responsible for substance passing in and out of cells. They form ion channels to control the cell potential or ATP-dependent proton pumps, to name just two of a vast number of examples. As the name "transmembrane proteins" already suggests, they span entire membranes, either via α-helices (figure 3) or β-barrels (figure 4) to form pores in the cell membrane.[1]

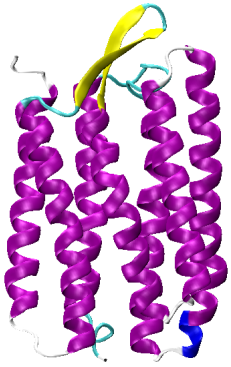

Figure 3: A chain of bacteriorhodopsin illustrates exemplarily an alphahelical TM protein. The figure was produced with VMD.[14]

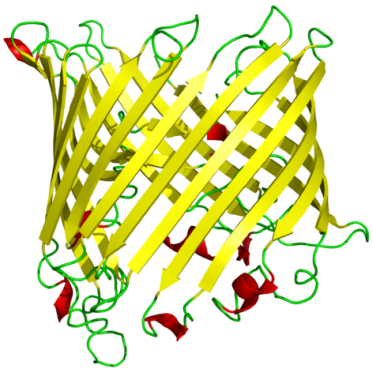

Figure 4: A beta-barrel TM protein is illustrated. The figure is taken from Wikipedia.[15]

In summary, the various essential functions of cell membranes are enabled via the unique interactions of membranes and TM proteins.

2.2 Molecular Dynamics

One possibility to study folding and aggregation processes of proteins are molecular dynamics (MD) simulations. They are based on Newton's equation of motion (1):[16,17]

$$F_i = m_i \dot{v}_i \tag{1}$$

with F_i representing the force acting on particle i and m_i and v_i describing the mass and velocity of this particle.
If one knows all the coordinates and velocities of the atoms composing a molecule as well as all the forces affecting them, one can calculate the future atomic positions and thereby their motions. The main problem thus is to identify these forces and to quantify their size and direction. Considering a molecule, there are numerous forces acting upon the atoms which derive, for example, from bonded interactions or electrostatic interactions between the atoms. So-called force fields are the method of choice to calculate the energy and forces of biomolecules.

There are many different types of force fields available. They differ in their theoretical approach and underlying assumptions as well as in their accuracy.[18] To categorize them one can divide them into knowledge-based force fields, which derive the energy from statistics of experimental data, and physics-based force fields. The latter aim to model the physics of molecules via particle-based mathematical functions, which are summed up to obtain the overall energy and force. One divides them into all-atom force fields, which deal with every single atom of the molecule of interest, and coarse-grained force fields, which combine several atoms into beads. For example, a methyl group would be described via three hydrogen and one carbon atom in the first case, whereas in the latter one would account just for the methyl group as a whole. Which kind of force field to choose depends pretty much on the simulation problem. One has to know its strengths and weaknesses of the various force fields and apply the theoretical level according to the accuracy one wants to reach.[18]

But how to identify all the forces? To this end, the potential energy of the system is systematically calculated. First, the potential energy is divided into bonded and nonbonded interactions between atoms. In general this is represented via the following equation (2):

$$V = V^{bond} + V^{elec} + V^{vdW} \tag{2}$$

where V^{bond} describes the bonded interactions, while the nonbonded interactions derive from electrostatic (V^{elec}) and van der Waals (V^{vdw}) contributions.

In this thesis the CHARMM22 force field is used.[19] It is an all atom force field, which is based on the following equation (3):[19–21]

$$
\begin{aligned}
V^{charmm} = & \sum_{i \in bonds} k_{b,i}(b_i - b_{0,i})^2 + \sum_{i \in angles} k_{\theta,i}(\theta_i - \theta_{0,i})^2 \\
& + \sum_{i \in dihedrals} k_{\phi,i}(1 + cos(n\phi_i - \delta)) + \sum_{i \in impropers} k_{\omega,i}(\omega_i - \omega_{0,i})^2 \\
& + \sum_{Urey-Bradley} k_{\mathrm{UB}}(S - S_0)^2 + \sum_{residues} V_{\mathrm{CMAP}}(\phi, \psi) \\
& + \sum_{i \in atoms} \sum_{j \in atoms} \frac{q_i q_j}{4\pi\epsilon_0 \epsilon r_{ij}} + \sum_{i \in atoms} \sum_{j \in atoms} \epsilon_{ij} \left[\left(\frac{r_{ij}^{vdW}}{r_{ij}} \right)^{12} - 2 \left(\frac{r_{ij}^{vdW}}{r_{ij}} \right)^{6} \right]
\end{aligned}
\tag{3}
$$

The first six terms account for bonded interactions, including bond vibrations (term 1), angle bending (term 2), angle torsion (term 3) and out-of-plane bending (term 4). The parameters $k_{x,i}$ are specific constants (e.g. in term 1 the force constant of a bond) with $x = b, \theta, \phi, \omega$ for bonds, angles, dihedral angles and improper angles and the subscript 0 marks the values for equilibrium. The Urey-Bradley potential[20] (term 5) as well as the CMAP potential[21] (term 6) are correction terms especially important for simulations of proteins. The Urey-Bradley term, depending on the distance S between atoms A and C (where A is bound to B and B to C), improves the description of out-of-plane motion. The CMAP correction is used to erase systematic errors in the representation of the protein's backbone depending on the dihedral angles ϕ and ψ.
The last two terms (7 and 8) describe the nonbonded interactions, the former the electrostatic contribution derived from Coulomb's equation and the latter accounts for the van der Waals interactions calculated via a Lennard-Jones potential. r_{ij} measures the distance between atoms i and j with partial charges q_i and q_j, while r_{ij}^{vdW} is the sum of the corresponding van der Waals radii.

Bonded interactions are straight forward to describe if one knows the connectivity of a molecule. In the case of bonds, as shown in equation (3), bond oscillations are typically described via a harmonic potential and the resulting force is calculated.
The challenging tasks are the nonbonded interactions. The problem is, that the number of atoms, which have to be taken into account, is not clear, because one does not know, how far these interactions reach. For example, to describe the van der Waals interactions 100% correctly, one would need to include all atoms of a system, because these interactions approach zero but never completely disappear. However, at some point the influence of one atom on another is not measurable any more and can thus be neglected.
But where has this boarder to be set? First, for answering this question one has to know, which accuracy one needs. The longer the cut-offs of nonbonded interactions are set, the higher the accuracy, but as well the higher the computational effort. Depending on the simulation problem and the setup of the system, the influence of different cut-offs has to

be examined to finally choose the best settings. The work of Steinbach et al.[22] gives an insight in the methodology how to introduce cut-offs in MD simulations and discusses the influence of the cut-off size.
Another problem is, that one does not want a force field to describe one special molecule only, but various molecule types, such as proteins. Therefore, force field parametrization plays an important role to define the parameters occurring in equation (3).

Having the force field, one can now calculate the forces on a molecule by differentiating the potential with respect to all coordinates of the system. But how to calculate the motion of it? One cannot just calculate the velocities according to Newton's equation (1) to determine the position and conformation of the molecule in one, two or three seconds, because during the motion the forces change. Therefore, one has to calculate the forces at a given time, calculate the resulting new positions of the atoms for a small time step Δt, and then recalculate the forces, and so on. To resolve the fastest motions (vibrations) Δt should be about 1 to 2 fs. In this thesis the original Verlet integrator of the CHARMM program, which is based on the Verlet algorithm, is used to iteratively solve equation (3). The positions x of the atoms at the current time step t and at $t - \Delta t$ as well as the forces of the current time step are used for calculating the new positions at $t + \Delta t$. That can be derived by formulating the positions via a forward and backward Taylor expansion (equations (4) and (5)):[16,17]

$$x(t + \Delta t) = x(t) + \Delta t\dot{x}(t) + \frac{1}{2}\Delta t^2\ddot{x}(t) + \frac{1}{6}\Delta t^3\dddot{x}(t) + O(\Delta t^4) \tag{4}$$

$$x(t - \Delta t) = x(t) - \Delta t\dot{x}(t) + \frac{1}{2}\Delta t^2\ddot{x}(t) - \frac{1}{6}\Delta t^3\dddot{x}(t) + O(\Delta t^4) \tag{5}$$

Ending the Taylor expansion after the third term one obtains an error of order $O(\Delta t^4)$. By summing equations (4) and (5) one obtains the new positions:[16,17]

$$x(t + \Delta t) = 2x(t) - x(t - \Delta t) + \Delta t^2\ddot{x}(t) + O(\Delta t^4) \tag{6}$$

As the velocities are not used for calculating the new positions, they have to be calculated from the resulting positions via:[16,17]

$$v(t) = \frac{x(t + \Delta t) - x(t - \Delta t)}{2\Delta t} + O(\Delta t^2) \tag{7}$$

Thus, for knowing the current velocities one needs to know the positions of the next time step.

2.3 Implicit Solvent - FACTSMEM

The previous subsection about "molecular dynamics" dealt with the problem of describing the motion of a molecule. The underlying assumption was, that it moves in vacuum. But of course, to describe real biological and chemical problems, one typically has to deal with a system in solution or an even more complex environment.
One possibility would be, to set up a system with, for example, a protein and put it in a box full of water molecules. Then the protein motion can be described as explained above, but also every single solvent atom has to be taken into account. In this case one explicitly describes the motion of the solvent and its effects on the protein, therefore, it is called an explicit solvent model.
The problem is, that by doing so, one needs a tremendous amount of CPU time only for calculating the motion of water, in which one actually is not interested in. Finding alternatives to describe solvents can thereby remarkably reduce the costs and increase the speed of simulations.

Implicit solvent models are a possible alternative.[4] Here the solvent atoms are not considered explicitly, but the physico-chemical appearance of a solvent is described to model the nonbonded polar and nonpolar interactions between the solvent and the solute, as one is merely interested in the effect of the solvent on the motion of the solute. The potential energy of the system can then be represented as:

$$V = V^{bond} + V^{elec} + V^{vdW} + \Delta G^{solv} \tag{8}$$

The first three terms are the same like simulating the solute in vacuum (equation (2)), while the ΔG^{solv} term accounts for the solvation free energy to include the effects of the solvent. This part can be divided into:

$$\Delta G^{solv} = \Delta G^{elec} + \Delta G^{nonpolar} \tag{9}$$

The implicit solvent model FACTS[23] represents one possibility to calculate ΔG^{solv} for water. Within this model the polar contribution to the solvation free energy is calculated via the generalized Born formalism[24] and the nonpolar part via the solvent accessible surface area ($SASA$) model.[23]
The main physico-chemical quantities describing a solvent within the generalized Born[25–28] and $SASA$[23] formalisms are the dielectric constant, the effective Born radii and the surface tension. The latter are a measure for the burial of an atom inside the protein and are needed for the calculation of ΔG^{elec}. Considering water as the solvent one has the advantage of an isotropic environment. The dielectric constant, the effective Born radii

and the surface tension do not change when changing the position of the solute. But if more complex solvents like membranes have to be described, one has to deal with an anisotropic medium where the named parameters depend on the position of the solute. Therefore, an extension to the FACTS model is needed.
One possible solution is given by FACTSMEM.[5,6] This implicit membrane model, recently developed by Martin Carballo Pacheco et al., extends FACTS to anisotropic environments. It is based on the same theory as FACTS using the generalized Born approach for the polar contribution and the $SASA$ model to describe the nonpolar part of the solvation free energy. It considers the needed changes in the dielectric constant, surface tension and the effective Born radii.

Figure 5 shows the geometry of a membrane as modelled by FACTSMEM. The surface of the membrane is orientated perpendicular to the z-axis and spreads over the whole x, y-plane. The membrane is surrounded by water. Figures 6 and 7 show how the dielectric constant and the surface tension, which change with the z-coordinate, are treated to represent the membrane. It has to be remarked that for describing the water phase outside the membrane FACTSMEM is identical to FACTS.

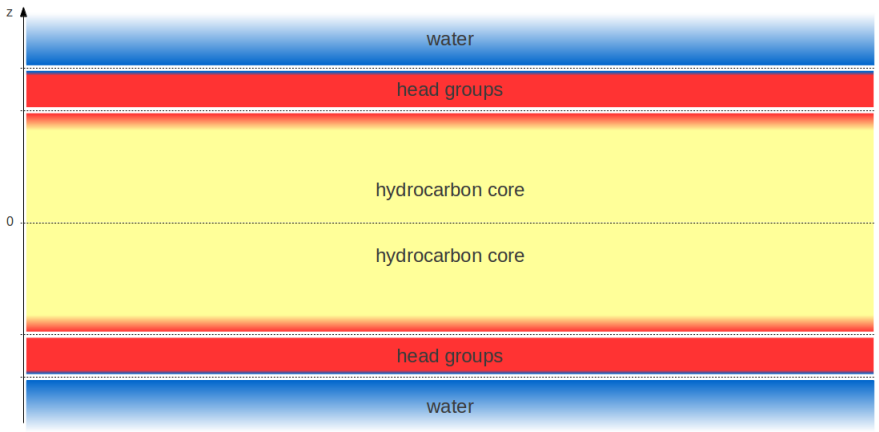

Figure 5: The appearance of a membrane as modelled by FACTSMEM and its placement in the Cartesian room is shown. Note that the plot is not true to scale. The different colours represent the characteristical parts of a membrane. This figure was adapted from Carballo Pacheco et al.[6]

In FACTSMEM the nonpolar contribution of the solvation free energy is calculated as[5,6]

$$\Delta G^{nonpolar} = \gamma \sum_{i=1}^{N} S(z_i) SASA_i \tag{10}$$

with γ representing the surface tension in water and $S(z)$ scaling it to the accurate value along the membrane normal z to model the surface tension of the membrane. $SASA$ is

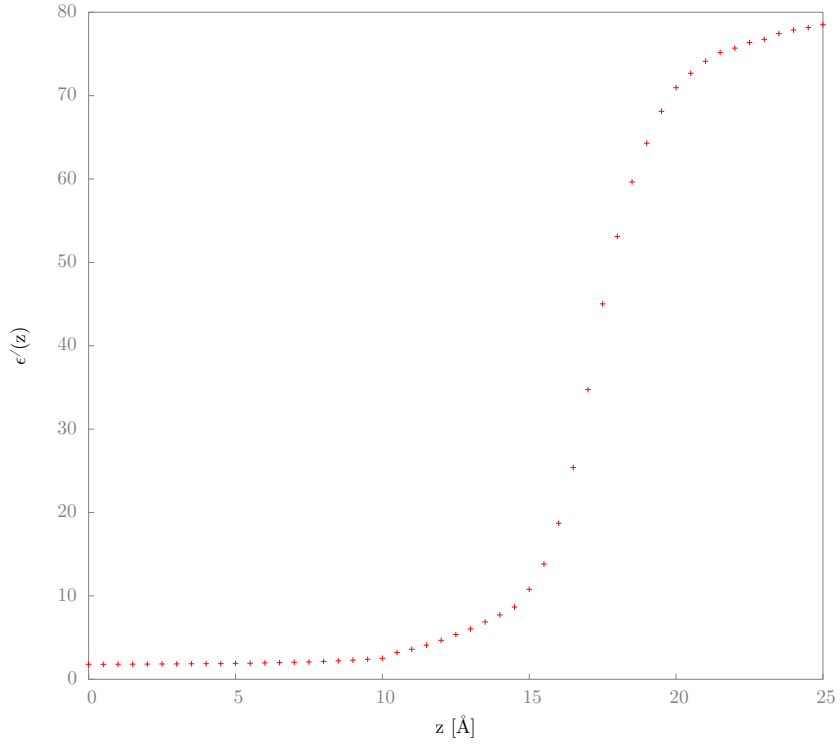

Figure 6: The effective dielectric constant is plotted as a function of the membrane normal z. The figure is reproduced from Carballo Pacheco.[5]

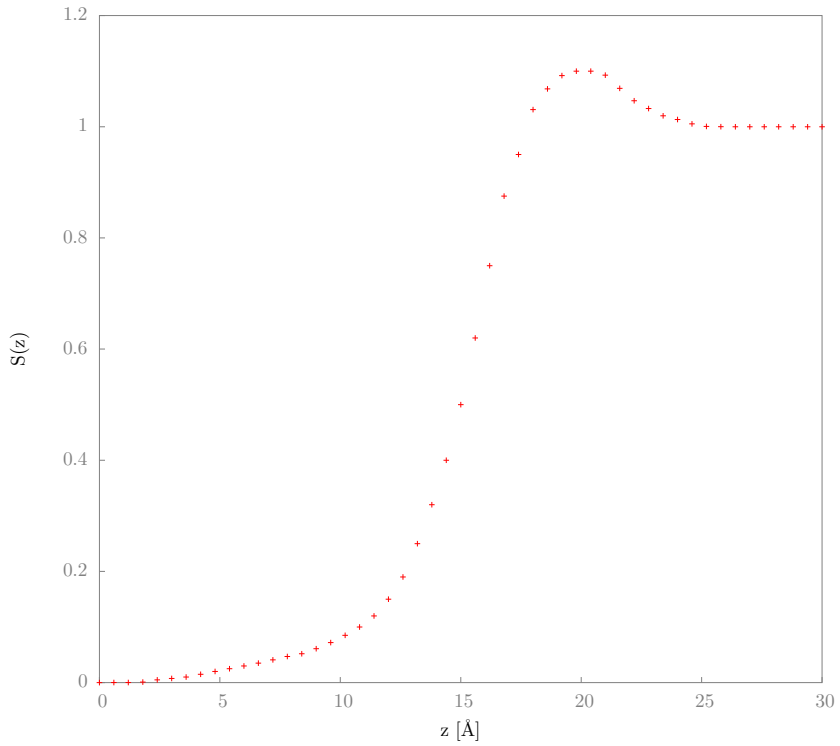

Figure 7: The nonpolar profile S(z) is plotted as a function of the membrane normal z. The figure is reproduced from Carballo Pacheco.[5]

calculated as[5,6,23]

$$SASA_i = c_0 + \frac{c_1}{1 + e^{-c_2(D_i - c_3)}} \tag{11}$$

where D_i is proportional to the burial of atom i and c_0, c_1, c_2 and c_3 are parameters being fitted to the exact values of $SASA$.[29]

The polar contribution is calculated according to the generalized Born formalism[24] by the following equation[5,6]

$$\Delta G^{elec} = -\sum_{j=1}^{N}\sum_{i=1}^{N}\tau(z)\frac{q_i q_j}{\sqrt{r_{ij}^2 + R_i(z)R_j(z)exp(-r_{ij}^2/\kappa R_i(z)R_j(z))}} \tag{12}$$

where $R_i(z)$ and $R_j(z)$ are the effective Born radii calculated via equation (14), r_{ij} measures the distance between atoms i and j, κ is a constant which is in general equal to 4 and $\tau(z)$ is given as[5,6]

$$\tau(z) = \frac{1}{\epsilon_p} - \frac{1}{\epsilon_{ij}(\epsilon'(z_i), \epsilon'(z_j))} \tag{13}$$

with $\epsilon_{ij} = \frac{\epsilon'(z_i)+\epsilon'(z_j)}{2}$.[5,6] ϵ_p is the dielectric constant of the solute and ϵ' gives the effective dielectric constant of the solvent in dependency on the z-coordinate.

The effective Born radii, which depend on the dielectric constant, are derived by calculating them with FACTS for water ($R_{i,w}$) and then estimating them for other dielectric constants $\epsilon \neq 78.5$ ($R_{i,m}$) as[5,6]

$$R_{i,m} = R_{i,w}\frac{\left(1 + \frac{\beta_m}{2}\right)}{\left(1 + \frac{\beta_w}{2}\right)} \tag{14}$$

Here, $\beta_w = \frac{\epsilon_p}{\epsilon_w}$ with $\epsilon_w = 78.5$ and $\beta_m(z) = \frac{\epsilon_p}{\epsilon'(z)}$ account for the dielectric constant ratios of protein-water and protein-membrane.

If one formulates equations (10) and (12) independent from z, one actually gets back to the original formulae used in the FACTS model for water.[23]

3 Methods

For all simulations in this work the CHARMM version $c35b4$ is employed with the CHARMM22 all atom force field with CMAP corrections and the implicit membrane model FACTSMEM.
As the results shall be compared to the results of Yuzlenko and Lazaridis,[11] their simulation procedure was strictly followed.

3.1 The test set

To assess the performance of FACTSMEM in discriminating the native state of membrane proteins, a test set of 450 decoys in total for five TM proteins is used. The five membrane proteins are bacteriorhodopsin (BRD7), rhodopsin (RHOD), V-ATPase (VATP), lactose permease (LTPA) and fumarate reductase (FMR).

BRD7 is a light sensitive proton pump[30] and RHOD activates visual photo transduction.[31] These opsins play an important role in optical processes in humans and animals. Their complete amino acid sequence is considered here.
V-ATPase is a special type of ATPase, which can for example be found in the vacuoles of plants and in Golgi-derived vesicles.[32] This TM protein regulates the pH-value and is e.g. responsible for the transport of small molecules across the membrane.[32]
LTPA is a symporter protein which transfers lactose using free energy stored in a proton gradient into a cell.[33]
FMR is important in catalyzing the reduction of fumarate to succinate and is therefore a key enzyme for organisms living in anaerobic conditions.[34] This TM protein consists of three subunits. Here, the membrane spanning unit C is considered only.

For each of the proteins except RHOD 100 decoys were used for the test ensemble, while for RHOD only 50 decoys were included. For the number of decoys used the work of Yuzlenko and Lazaridis[11] was followed on order to warrant comparability. The decoys were obtained from Yarov-Yarovoy, who generated them via the protein structure prediction program ROSETTA.[35]

3.2 Preparing the protein structures

At first, the native structures obtained from the Protein Data Bank (PDB)[36] were modified such, that their sequence fit completely the sequence of the decoys. As the crystal structures are often dimers or trimers, here only one of the proteins, i.e. a monomer, is considered. Moreover, as the decoys usually contain less residues than the corresponding native structures, the PDB structures had to be truncated.

For BRD7 residues 5 to 231 of *chainA* were chosen (the numbers refer to specifications in the PDB file), comprising 7 transmembrane α-helices. Retinals and crystal waters were deleted. The same was done for RHOD, keeping residues 33 to 310. For VATP *chainA* was considered, deleting everything but residues 12 to 156. FMR was truncated to residues 21 to 237 using *chainC*, while for LTPA residues 1 to 190 were taken into account. Table 1 provides a summary of the chosen protein sequences.

Table 1: Specifications of the native protein structures. The entries for "Chain" and "Residues" refer to the nomenclature of the PDB files.

Protein	Code	PDB-ID	Chain	Residues	TM helices	No. of residues
Bacteriorhodopsin	BRD7	1PY6	A	5-231	7	227
Rhodopsin	RHOD	1U19	A	33-310	7	227
Fumarate reductase	FMR	1QLA	C	21-237	5	217
Lactose permease	LTPA	1PV6	A	1-190	6	190
V-ATPase	VATP	2BL2	A	12-156	4	145

Having the correct input structures for the native states, for every protein except RHOD 100 decoys were chosen, while for the latter only 50 decoys were considered. This choice was made in order to make the test ensemble the same size as YaL. As there were 5000 decoys for each protein obtained from Yarov-Yarovoy, one has to find a criterion for the specification of the test ensemble. As there are no remarks in the YaL paper, this thesis provides several approaches to deal with that problem.
The first was to select the decoys by their root mean square deviation (RMSD) to the initial native structure. This approach ensures that one stays as close as possible to the native state in the energy landscape, which makes it more difficult for FACTSMEM to discriminate the native state. Other approaches were to choose the test set via the lowest initial energy of the decoys in vacuum, via the lowest initial energy in FACTSMEM and finally considering the decoys with the lowest energy among those with a RMSD smaller than 25 Å to the initial native structure.
The validity of these approaches is discussed below, as well as the problems connected to them.

3.3 Molecular Dynamic specifications

In the setup procedure, which needs to be performed for all 25,000 decoys, the missing hydrogen atoms were included using the HBUILD command of CHARMM. The termini were considered to be charged via the NTER and CTER command. Finally, the proteins

were orientated parallel to the membrane normal (z-axis of the system).

After setting up the systems, the RMSD for all 50,000 decoys was calculated and the 450 test decoys selected using the RMSD criterion described above. These structures were then submitted to a 1,000 step adopted basis Newton Raphson (ABNR) minimization, employing the FACTSMEM model at a default width of 28.5 Å of the hydrocarbon core. No constraints were applied, in disagreement to the work of YaL, because the FACTSMEM model does not allow any constraints.
Afterwards, a 0.1 ns MD simulation was performed, using the original Verlet integrator, a time step of 1 fs and a temperature profile heating from 240 to 310 K. Any bonds containing hydrogen atoms were fixed using the SHAKE command of CHARMM with a tolerance of 10^{-10} kcal/mol. The nonbonded cut-off was 12 Å, using a switching function from 10 Å. From the second half of the resulting trajectories the structure with the lowest energy was selected for each trajectory and submitted to a final energy minimization using 1,000 ABNR steps.

To determine the influence of the membrane width, the same simulation procedure was repeated for widths of 23.1 Å and 30.4 Å to match the range of membrane thickness applied in the YaL paper. According to Sayadi et al.[37] different membrane widths were obtained by simply scaling the default FACTSMEM width (28.5 Å) to the desired value, i.e. by scaling $\epsilon'(z)$ and $S(z)$ shown in figures 6 and 7.

The energies obtained after the final minimization were then used for further evaluation.

3.4 Criteria for assessing the discriminative power

3.4.1 Relative potential energy

According to YaL the ability of discriminating the native structure is first of all measured by the relative energy of the decoys with respect to the native structure, which is obtained by subtracting the energy of the corresponding native structure from the decoy energy.
According to the thermodynamic thesis, the native state is the lowest energy structure of the possible protein conformation space.[38] Thus, the relative decoy energies should be positive in order for FACTSMEM being an effective energy function and discriminating the native state correctly.

If one plots the relative energies against the RMSDs of the decoy structures from the native state, one should therefore obtain "funnel-like" plots.[39] That means, the closer the decoy structure gets to the native state, the lower the relative energy should be. A sample for such a "funnel-like" plot can be seen in figure 8.

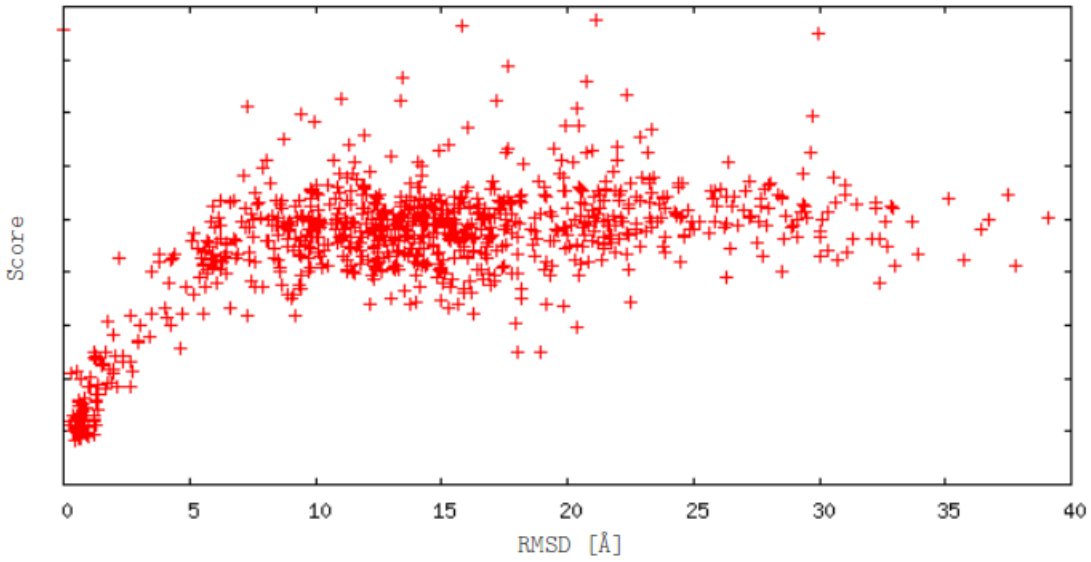

Figure 8: Shown is an example for a perfect funnel in folding simulations. The figure was taken from the ROSIE website.[40] An energy score is plotted against the RMSD in Å.

3.4.2 Z-score

A quantitative method to measure the discriminative power of FACTSMEM is calculating the Z-score according to equation (15):[41]

$$Z = \frac{\langle E_{decoys} \rangle - E_{native}}{\sigma_{decoys}} \tag{15}$$

with $\langle E_{decoys} \rangle$ representing the average energy of the decoy ensemble, E_{native} the energy of the corresponding native state and σ_{decoys} giving the standard deviation for the energy of that decoy ensemble. Higher Z-score means hereby higher discriminative power.

In summary, for assessing the discriminative power of FACTSMEM the relative potential energy of each decoy with respect to its corresponding native state is calculated as well as the Z-score for each protein according to equation (15). To see, if the results give "funnel-like" plots, the relative potential energy of the decoys is plotted against their RMSDs. The RMSD is computed for each decoy structure from its native correspondent after final minimization of these structures for backbone atoms.

4 Results and discussions

As already explained in the previous "Methods" section, a quality requirement for an effective energy function for proteins is its ability to discriminate between native structures and misfolded decoys. According to the thermodynamic assumption of the native state being the lowest energy structure,[38] the relative decoy energies should be positive.

4.1 The first decoy ensemble - initial RMSD

The results for the test ensemble chosen via the initial RMSD from the native state for a membrane width of 28.5 Å are shown in figure 9.

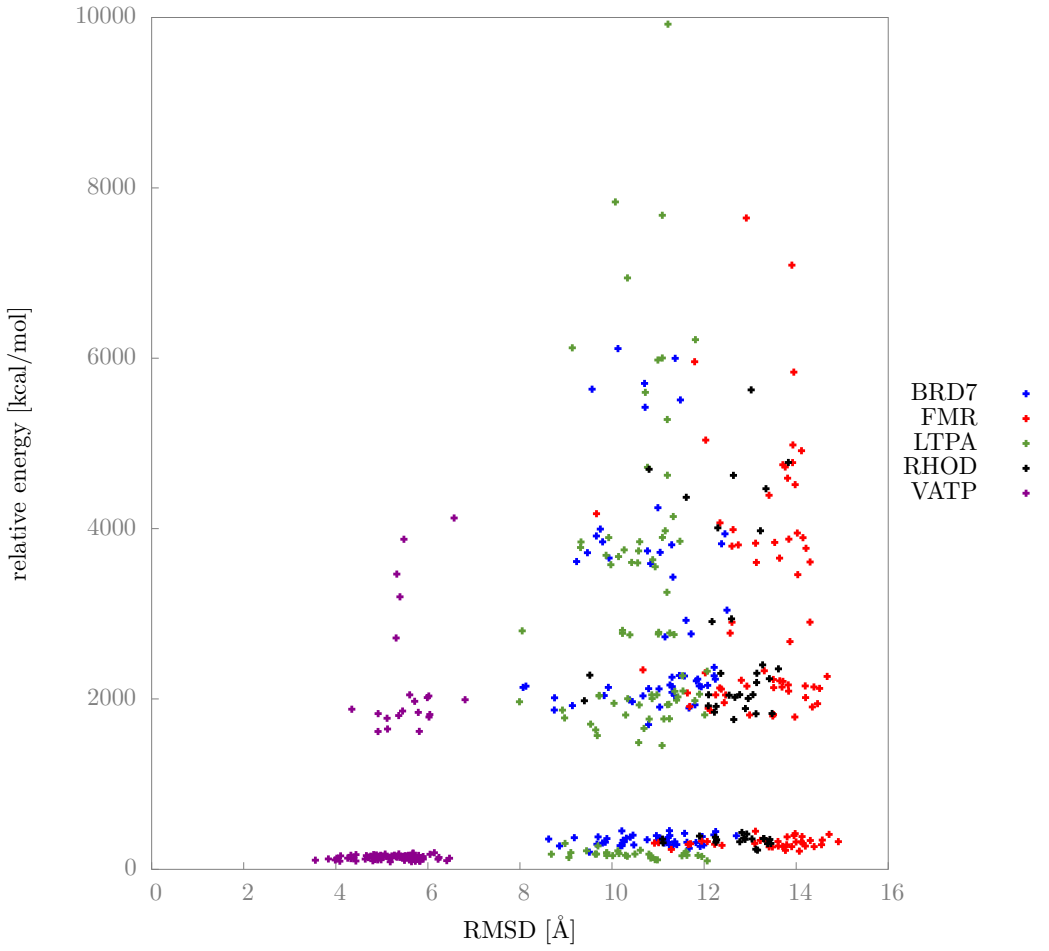

Figure 9: For the first ensemble the relative potential energies in kcal/mol are plotted against the initial RMSD measured in Å. BRD7 is shown in blue, FMR in red, LTPA in green, RHOD in black and VATP in purple.

As one can see, all the relative energies are greater than zero for all proteins. That implies a complete discrimination of the native state. But the Z-score averaged over all proteins and decoys is only 1.1. This is a surprising finding in comparison to the results of YaL. For none of the membrane models and proteins they find a 100% discrimination,

but observe larger Z-scores ranging from 1.5 to 2.9 depending on the membrane model. That can be explained by the very large standard deviation observed in this work. The energy range for the current ensemble is about 20 times higher (it varies from 0 to about 10000 kcal/mol) than in the YaL results (it varies between -50 and 400 kcal/mol). Therefore, it has to be analysed where these high energies come from.

To approach this problem, one decoy for BRD7 showing a very high relative potential energy is studied in detail. Looking at the energy contributions to the potential energy at the end of the simulation procedure, it can be seen, that the high energy structures have extraordinary large bond energies (see Appendix 2: energy development of the final minimization). Analysing the energy of every single bond of the decoy, one finds three bonds with energies over 600 kcal/mol, one of them is even about 880 kcal/mol. So these three bonds account for 70% of the total bond energy of that decoy.
One of the problematic bonds is visualized using the VMD program[14] in figure 23 in Appendix 3. This figure reveals why the bond energies are way too large: these bonds are far too long carbon-carbon bonds, almost double than the typical length.

How can that be explained? Checking the initial minimization procedure one gets the following impression: The by far largest energy term at the beginning is the van der Waals contribution (almost 100% of the total potential energy). That can only be explained if some atoms are too close together, alomst overlapping. In the specific example, two carbon atoms of residues 134 and 185 (both tryptophane of two neighbouring α-helices) in the BRD7 decoy are too close to each other (to compare these residue numbers with the native BRD7 nomenclature in the PDB file, one has to add 4 to the current residue number). Their distance measures only 0.33 Å (see figure 21 in Appendix 3).
In the first minimization steps, the minimizer tries to resolve these bad interactions and rapidly reduces the van der Waals energy. By doing so, the bond energy increases by about 25,000 kcal/mol. In the ongoing minimization, the minimizer indeed decreases both terms, but arrives at a structure, which can hardly be further optimized with a final bond energy of about 3,000 kcal/mol higher than at the beginning (compare Appendix 2: energy development of the initial minimization).
Remarkable is, that the potential energy after the initial minimization is more than 1,400 kcal/mol lower than at the beginning of the final minimization, i.e., the energy increases by about 65% during the MD simulation.
In the final minimization the potential energy can be reduced by about 2,000 kcal/mol. But looking at the development of the bond energy, one sees that this contribution cannot be considerably optimized (only by about 400 kcal/mol, i.e. 20% of the improvement, which is not much regarding the fact, that this is the largest contribution to the potential energy). In summary, after the simulation procedure all the strongly repulsive van der

Waals interactions are removed. There are no overlapping atoms any more. But there are still far too long carbon-carbon bonds (see Appendix 2: energy development of the final minimization). Thus it seems like optimizing the van der Waals interactions is done at the cost of the bond energies.
This assumption is supported by the fact, that the high energy bond discussed above occurs in residue 184 - a valine - between C_α and C_β. It is located in the neighbourhood of residues 185 and 134 (tryptophanes), which show the overlapping carbon atoms in the initial structure before the simulation procedure. It seems as if the minimizer has to rearrange the side chain to enable the neighbouring tryptophanes to move away from each other (compare figure 22 in Appendix 3).

It was tried to resolve this problem by first elongating the minimization procedure and secondly by using different minimizers - without success. Of course there is the possibility, that there are some unphysical structures among the 5000 decoys for each protein because of misfolding during the generation of the decoys with ROSETTA, which cannot be successfully treatened by geometry optimization. According to this, these high energy structures should not be taken into account to really validate the effectiveness of FACTSMEM in discriminating the native state. Therefore, a new test set of decoys is needed.

4.2 The second decoy ensemble - initial energy in vacuum

In a second approach, the initial energy of the decoys was calculated for all 25,000 decoys in vacuum to leave out effects of the membrane. The 100 decoys with the lowest energy were considered for the new test ensemble for each protein (again only 50 for RHOD).
Then the first minimization step was carried out as described in "Methods" with FACTSMEM at a width of 28.5 Å. The energy distribution was checked right after this initial minimization. These results are shown in figure 10.

Again, there is the problem of a too large energy range. There are still structures with very high energies, and the largest contribution is again the bond energy. For comparison the maximum energy range when doing the minimization in vacuum for choosing the ensemble was found for RHOD. Here the 50 input structures differed only by about 300 kcal/mol. When shifting low energy structures from vacuum to the membrane, the relative energy changes a lot - the energy range is ten times the range in vacuum - and again unreasonable bond energies arise.
The reason for that is not obvious. It might be a problem to choose the decoys based on the initial energy in vacuum. Of course, the already folded decoys behave differently in vacuum and membranes, as in vacuum there are no external effects stabilizing the decoy structures via solvation.

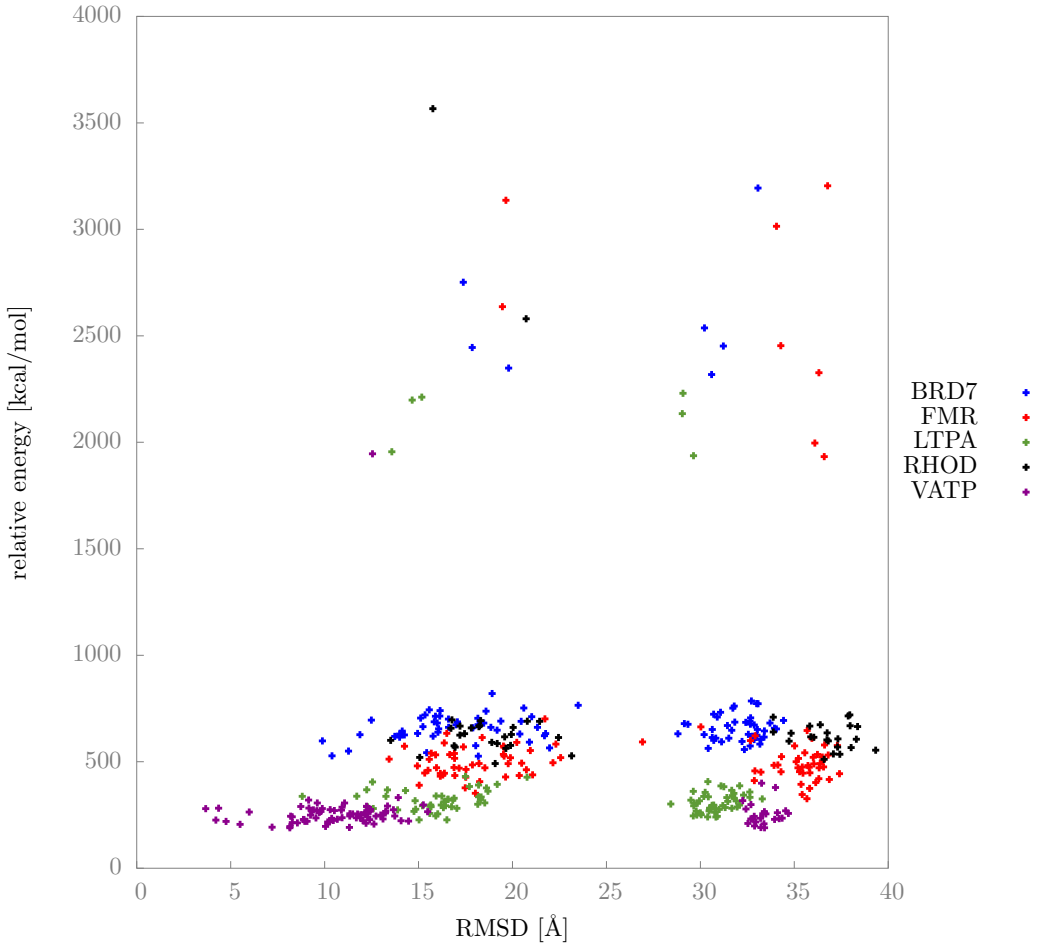

Figure 10: For the second ensemble the relative potential energies in kcal/mol are plotted against the initial RMSD measured in Å after the initial minimization. BRD7 is shown in blue, FMR in red, LTPA in green, RHOD in black and VATP in purple.

Given this energy change during the simulation procedure in the membrane it is of no value to continue with this decoy ensemble, because that would not improve the results considerably.

4.3 The third decoy ensemble - initial energy in FACTSMEM

In a third approach the decoys were chosen via their initial energy in FACTSMEM at the default width of 28.5 Å. The 450 decoys with the lowest energy are chosen. Again the complete simulation procedure as described in "Methods" was followed. The final results are shown in figure 11.

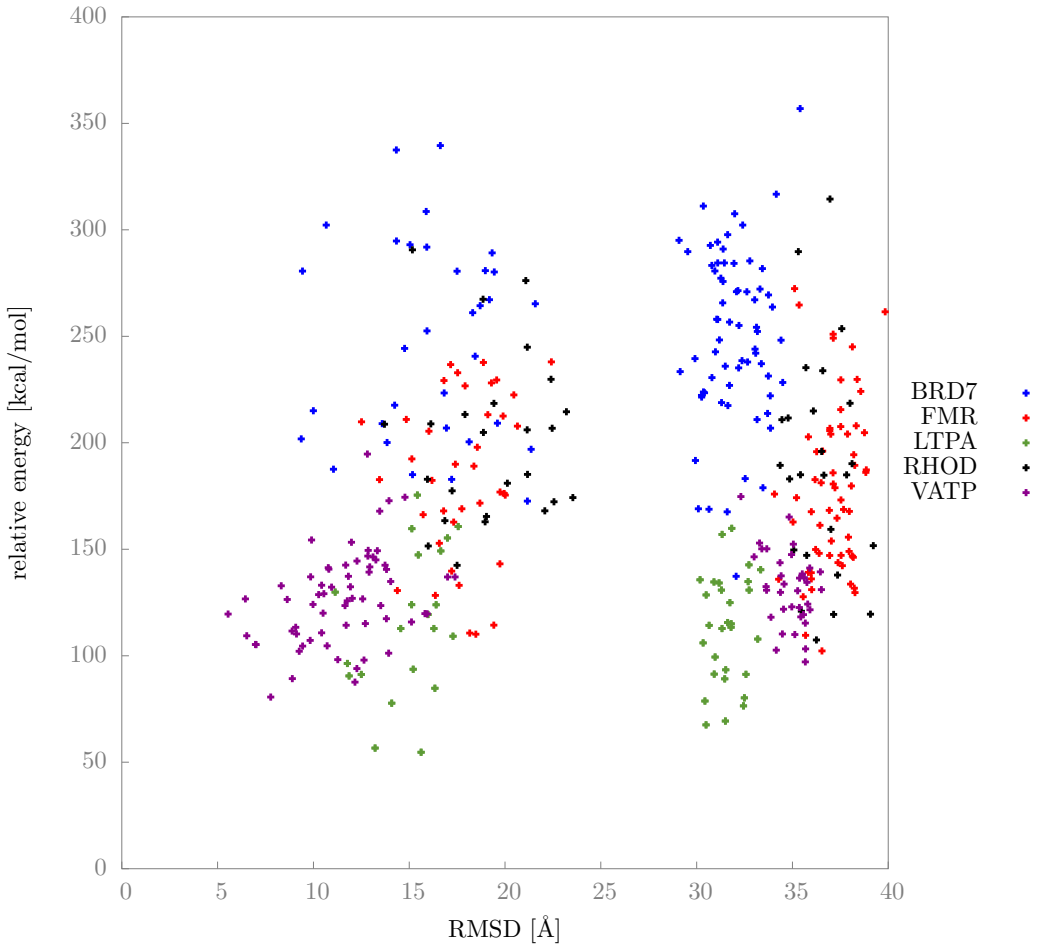

Figure 11: For the third ensemble using FACTSMEM with a width of 28.5 Å the relative potential energies in kcal/mol are plotted against the initial RMSD measured in Å. BRD7 is shown in blue, FMR in red, LTPA in green, RHOD in black and VATP in purple.

Here, the energy range has a reasonable distribution between 50 and around 400 kcal/mol. According to this, again a 100% discrimination rate is achieved. The large Z-score of 5.0 averaged over all proteins and decoys also shows a high discriminative power. This value

is 1.7 times higher than that achieved by the best model in the YaL paper (GBSW). But as seen above, the decoy ensemble influences the results a lot.

Following the work of YaL, in a next step it was tested, whether the membrane width influences the results. Therefore, this ensemble was submitted to different membrane widths using FACTSMEM with widths of 23.1 Å and 30.4 Å. The results are shown in figures 12 and 13.

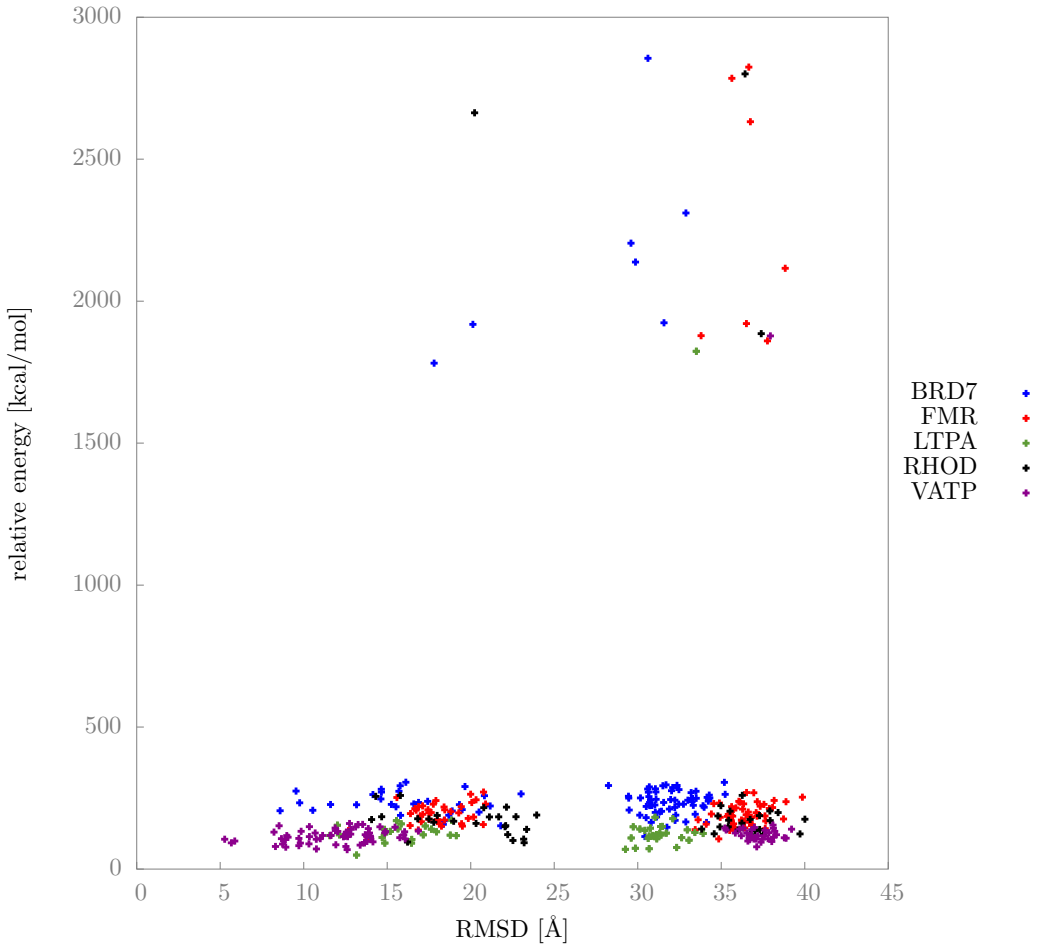

Figure 12: For the third ensemble using FACTSMEM with a width of 23.1 Å the relative potential energies in kcal/mol are plotted against the initial RMSD measured in Å. BRD7 is shown in blue, FMR in red, LTPA in green, RHOD in black and VATP in purple.

These plots reveal that the energy distribution drastically increases, ranging from 0 to 4,500 kcal/mol in the worst case. Thus, changing the membrane width leads to the same problems as before. Of course, one expects some change in the solvation free energies. For instance, reducing the membrane width some nonpolar amino acids might have to rearrange because being otherwise exposed to a more polar solvent region. But this is not the explanation here as the energy change again is far too large. Like already seen

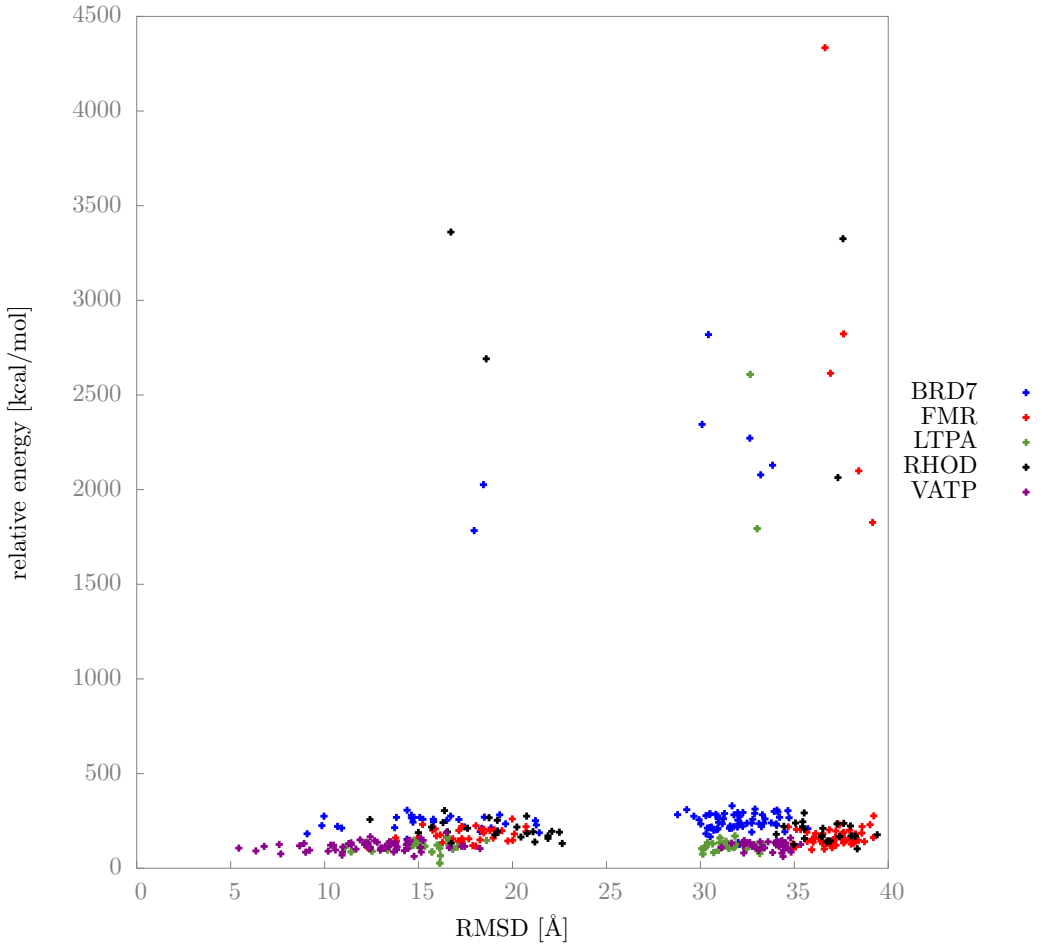

Figure 13: For the third ensemble using FACTSMEM with a width of 30.4 Å the relative potential energies in kcal/mol are plotted against the initial RMSD measured in Å. BRD7 is shown in blue, FMR in red, LTPA in green, RHOD in black and VATP in purple.

before, also by only changing the membrane width the bond energies of some decoys increase a lot. Having reasonable values at a width of 28.5 Å, for the widths of 23.1 Å and 30.4 Å some of the decoys experience a few too long carbon-carbon bonds. The hypothesis made above that the different environments of vacuum and a membrane could explain the rather different relative energies is therefore not the explanation.

FACTSMEM shows again a 100% discrimination for the changed membrane widths. However, given the large energy fluctuations with respect to the chosen set of decoys and membrane model, one has to doubt the validity of the obtained results. Regarding the RMSD range of the obtained results, the set of considered decoys is still different to the one in the YaL paper. Thus, another criterion for choosing the test ensemble was tried.

4.4 The fourth decoy ensemble - initial RMSD and initial energy in FACTSMEM

In a next step two criteria for choosing the decoys were used: First only decoys with an RMSD lower than 25 Å (this is the upper RMSD limit in the YaL results) are considered. From these the ones with the lowest initial energy in FACTSMEM at a width of 28.5 Å are chosen. This approach was only tested for BRD7 and VATP in order to evaluate whether acceptable results can be obtained.

Figure 14 shows the results for the new test ensemble. The results are similar to the ones obtained with the last test set at 28.5 Å in FACTSMEM for BRD7 and VATP, when only the initial energy was used as criterion. The energy range is the same, but because the RMSD range is smaller, the density of points in the populated RMSD-energy area is higher. Again, a 100% discrimination is found. Comparing to the YaL plots, the BRD7 results from FACTSMEM now are very similar to the results obtained with GBSW (figure 15). In order to test the reliability of these results, they are double-checked before extending the test to other proteins.

4.5 Comparing FACTSMEM and IMM1

To double-check the latest results, the last two decoy ensembles - chosen via the initial energy in FACTSMEM and via the RMSD-initial energy criterion respectively - are exposed to the IMM1 model[7] at a width of 28.5 Å for BRD7. This implies a change in the force field: IMM1 can only be used together with CHARMM19. All other simulation parameters are the same as before.

Figures 16 and 17 show the results after the simulation procedure, which indicate the

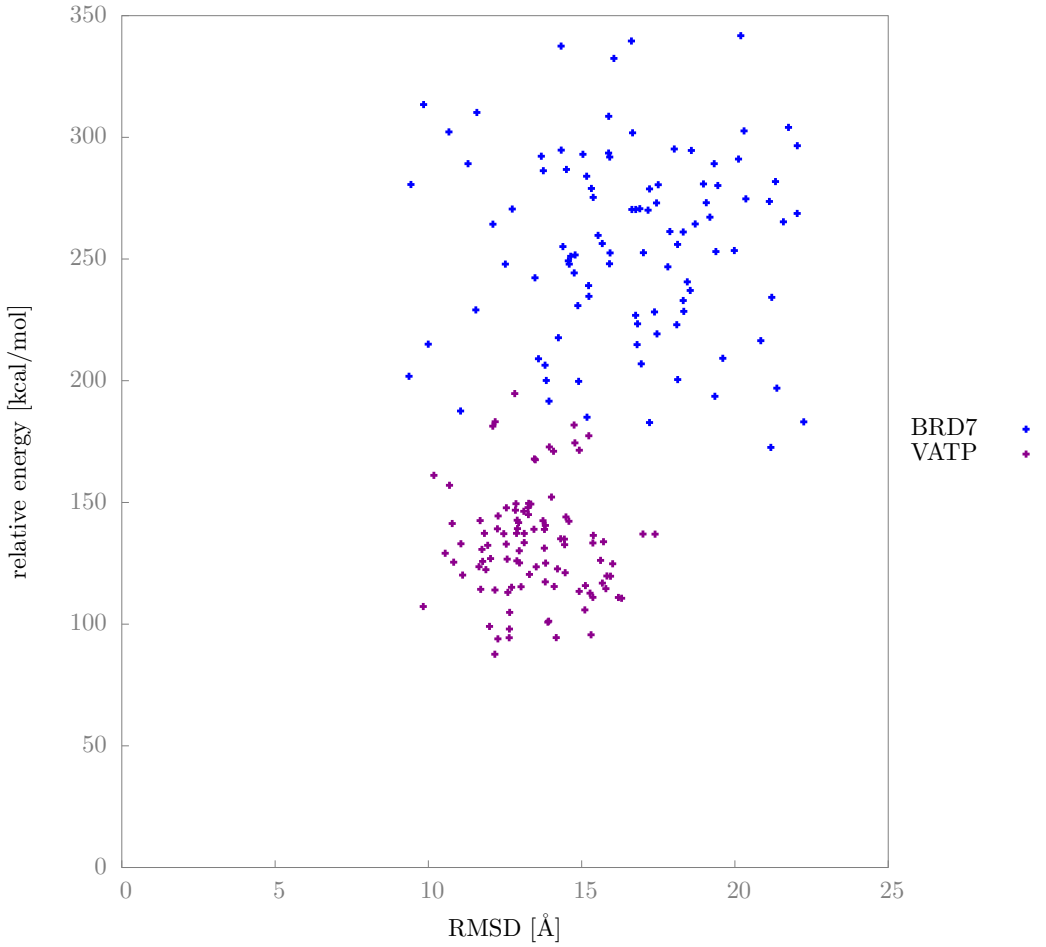

Figure 14: For the fourth ensemble the relative potential energies in kcal/mol are plotted against the initial RMSD measured in Å. BRD7 is shown in blue and VATP in purple.

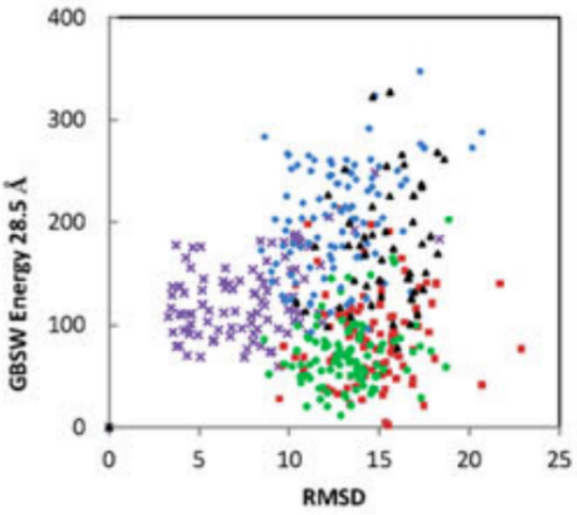

Figure 15: Shown is the plot of the relative potential energies in kcal/mol against the initial RMSD measured in Å for the GBSW model at a width of 28.5 Å. BRD7 is shown in blue, FMR in red, LTPA in green, RHOD in black and VATP in purple. The figure was taken from Yuzlenko and Lazaridis.[11]

occurrence of the same problems as before.
Looking at the results for the ensemble chosen via the initial energy in FACTSMEM (figure 16) the energy ranges from -500 to roughly 8000 kcal/mol. So changing to IMM1 the energy range is suddenly 20 times the energy of the FACTSMEM results (compare section 4.3).
Regarding the results for the fourth decoy ensemble (figure 17) changing to IMM1 the energy range is 30 times the range of the corresponding FACTSMEM ensemble ranging from -500 to about 5500 kcal/mol now (for FACTSMEM it was between 150 and 350 kcal/mol, figure 14). What is very confusing is the fact, that the input structures all had a RMSD lower than 25 Å. But after the IMM1 simulation, RMSDs up to about 47 Å occur. Looking at such a decoy with VMD[14] (figure 18) one finds it leaving the membrane. Before the simulation, the decoys span the entire membrane. Then this problematic sample starts moving out up to a point, where about 50% of the decoy are outside the membrane. That explains the huge RMSDs, but still this behaviour is surprising.
Furthermore, it is remarkable that in these IMM1 simulations relative decoy energies lower than zero are found. That was not observed for any FACTSMEM simulation.

In summary, changing the membrane model for the same decoy set changes the relative energy distribution extraordinarily, what cannot be explained via the change of the model. Even though these models estimate the effect of a membrane in different ways,

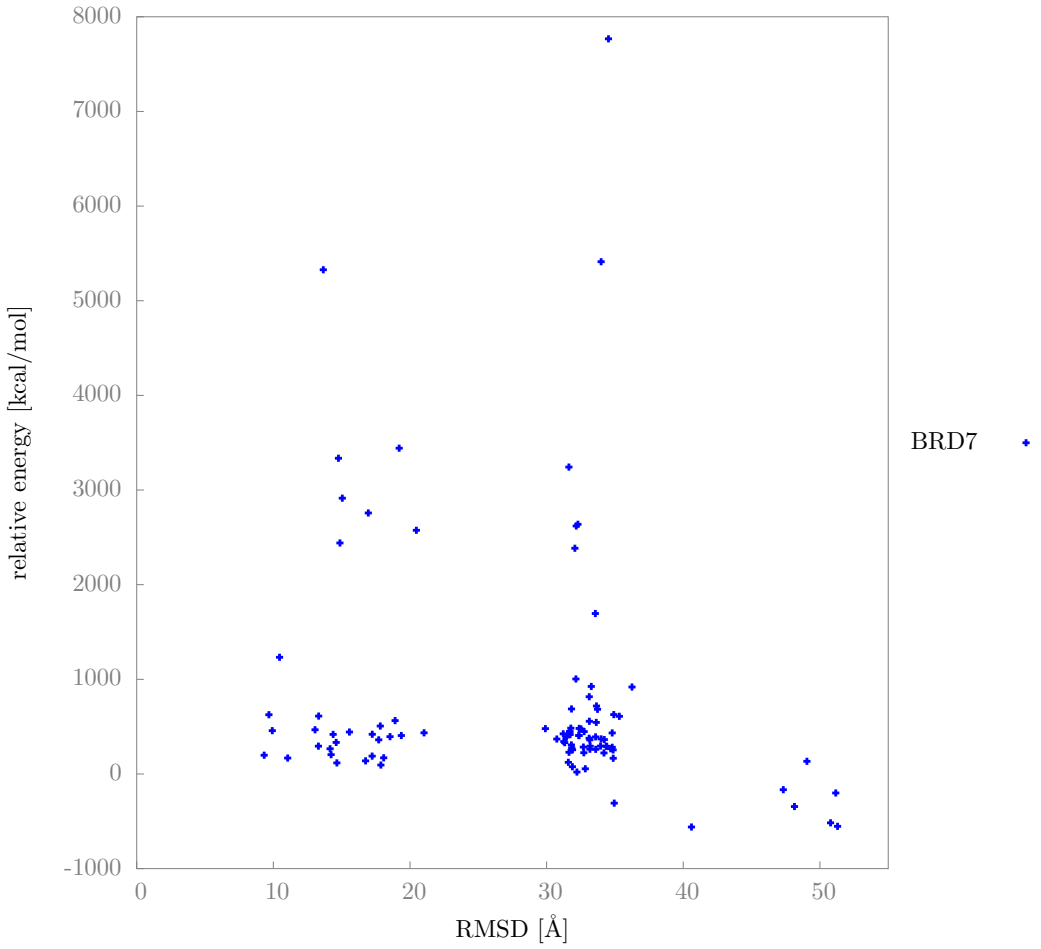

Figure 16: For the third ensemble using IMM1 at a width of 28.5 Å the relative potential energies in kcal/mol are plotted against the initial RMSD measured in Å.

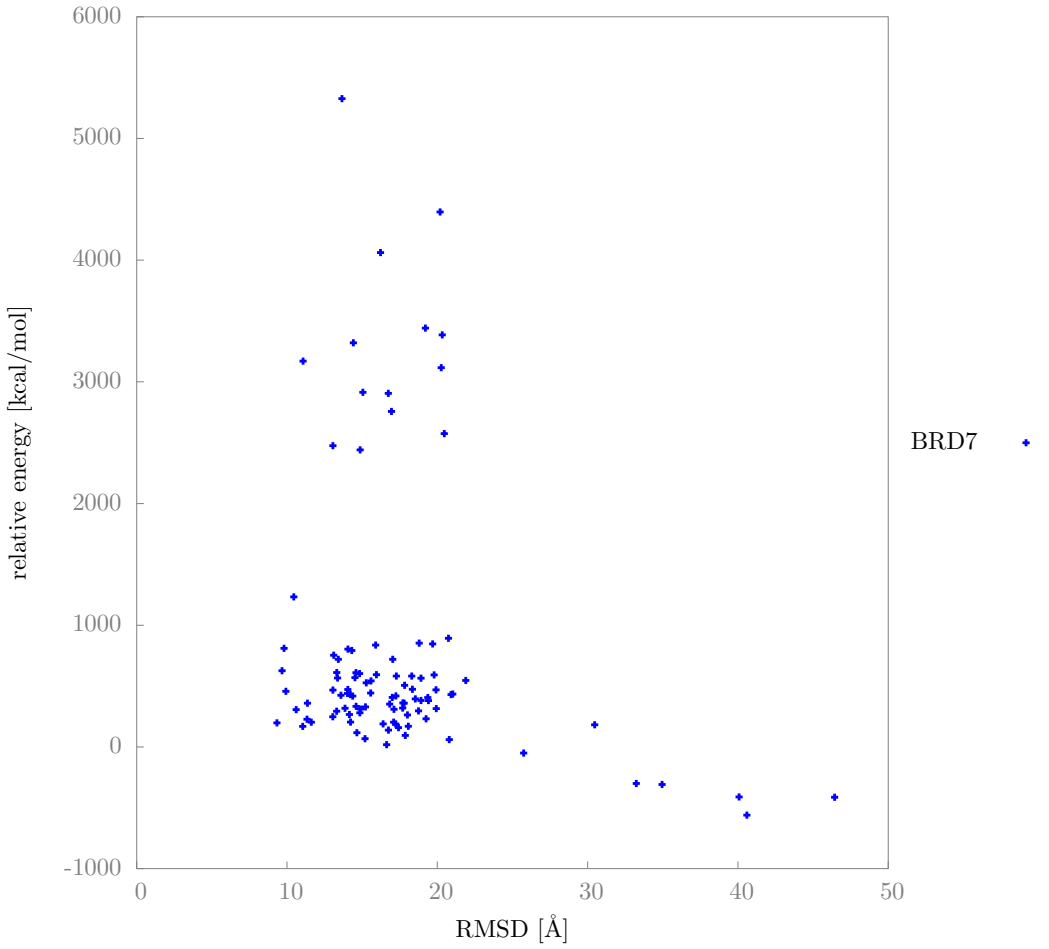

Figure 17: For the fourth ensemble using IMM1 at a width of 28.5 Å the relative potential energies in kcal/mol are plotted against the initial RMSD measured in Å.

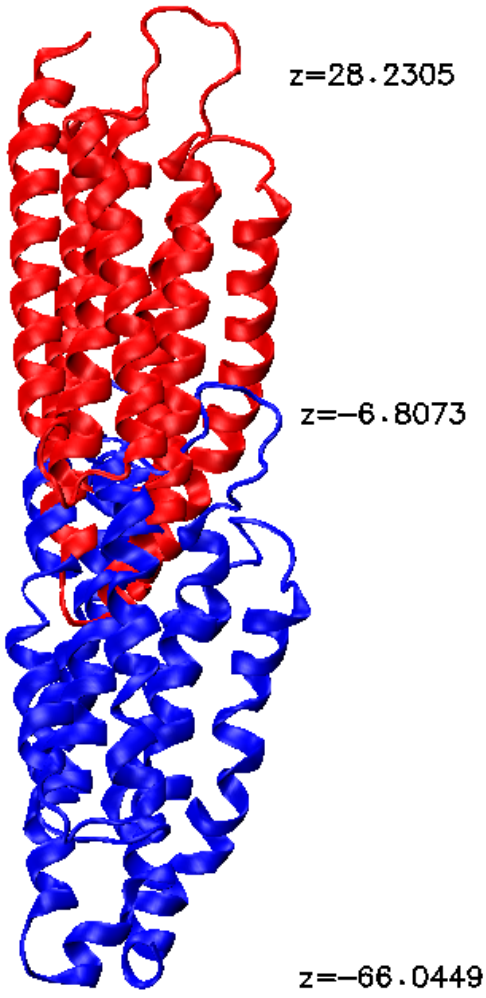

Figure 18: This figure shows one of the decoys from the fourth ensemble, which moved out of the membrane during the simulation with IMM1 at a width of 28.5 Å leading to a large rise of the RMSD. The labels refer to the membrane normal (values measured in Å). The red structure shows the decoy before the simulation procedure and the blue one the decoy after the simulation. The figure was produced with VMD.[14]

the difference in the assumptions and methods cannot explain an energy change of nearly 8000 kcal/mol for some decoys.
The problems arise for the same reasons already discussed for FACTSMEM above. For some decoys there is a tremendous rise in relative potential energy because of some high energy bonds. Again, they result from optimizing bad van der Waals interactions of the input structure.

The latest results tell us two things: First even the results achieved with the fourth decoy ensemble (section 4.4) cannot be verified. It cannot be ensured, that the good results obtained for FACTSMEM are really because of the model's discriminative power or due to unphysical decoy structures. Therefore, it does not make sense to extend the tests for this ensemble to the missing proteins FMR, LTPA and RHOD.
Second the IMM1 results show that the observed problems are not a problem of FACTSMEM itself but of the applied decoys. It is not explainable to that point, why the same decoy, giving reasonable results for one membrane model, ends up in an unphysical high energy structure for another model. It can be assumed that because of the enormously high van der Waals energies in the initial structures - what is true for all 25,000 decoys - the minimization is very unstable. In one case the minimizer "finds the right direction" to optimize the structure, but changing the conditions such as the solvation, the minimization can fail.

In summary it was not possible to reproduce the results of the YaL paper, neither with FACTSMEM nor with IMM1. That is surprising, as the simulation procedure was identical.

4.6 Evaluating the performance of FACTSMEM for native state simulations

The structural problems with the current decoys make it impossible to measure the performance of FACTSMEM with respect to native state discrimination. As these problems also occurred with the IMM1 model, they are not supposed to be connected to FACTSMEM, but to originate from the minimization problems because of the tremendous high van der Waals energies of the input structures. Therefore, in a last step the performance of FACTSMEM in simulating native state proteins was studied to ensure the correctness of the energy terms arising from it.

Table 2 shows the absolute potential energy of the native state for each protein for FACTSMEM at different widths after the final minimization. The bond potential is explicitly named to check whether in the native simulations high energy bonds occur. The individual nonbonded contributions to the potential energy are shown as well.

From these results one can see that FACTSMEM gives stable results. Shifting the native states to different membrane widhts does not cause an unexpected energy rise. As one can see the problems of unreasonable high potential energies due to unrelaxed bonds do not occur, they stay nearly at the same value for all membrane widths. The differences in potential energy for a given protein in the different membranes are always in a range of 1 to 2%. Depending on the protein bigger or smaller membranes are favoured. The biggest changes arise from ΔG^{elec}, here the change between the different membrane widths is about 20% on average. Compared to the fluctuations in bonded energies the relative change in the polar contribution to the solvation free energy is on average the double. As changing the membrane width mainly should influence the protein-solvent interactions this is a reasonable result.

As the proteins are just thermally fluctuating in the membrane, the solvation free energy should be rather constant during a MD simulation. Figure 19 shows the development of the potential energy and the nonbonded contributions during the MD simulation of the native state for BRD7 in FACTSMEM at a width of 28.5 Å. Changes of the nonbonded contributions to the potential energy - which are the terms being influenced by FACTSMEM - stay in a range of 200 kcal/mol (about 5% of the total energy), corresponding to small structural fluctuations as expected.

The initial rise in the potential energy comes from increased bond and angle potentials because of suddenly heating the system in the MD run without any equilibration beforehand. Thus, the velocity of the protein's atoms increases leading to geometry changes. Afterwards, the protein equilibrates again, yet at a higher temperature than before. Plots for the other proteins can be found in Appendix 4.

In summary, the native state simulations are carried out without problems and only little energy differences result from the change of the membrane width. It was thus possible to show that the energy terms arising from FACTSMEM give reasonable values for the native states. Together with the tests shown in Carballo Pacheco's work[5] no problems with the model itself were revealed. Regarding the fact that the same problems were observed with IMM1 this also indicates that there must be some structural problems in the decoy set obtained from Yarov-Yarovoy.

Table 2: Energies (in kcal/mol) of the native states after the simulation procedure in FACTSMEM. The meanings of abbreviations are: pot = potential energy, bond = bond energy, vdW = van der Waals energy in vacuum, elec = electrostatic interactions in vacuum.

	BRD7			FMR			LTPA			RHOD			VATP		
Membrane width [Å]	23.1	28.5	30.4	23.1	28.5	30.4	23.1	28.5	30.4	23.1	28.5	30.4	23.1	28.5	30.4
pot	-3720	-3708	-3690	-4156	-4129	-4105	-2636	-2614	-2616	-3624	-3652	-3644	-1266	-1260	-1243
bond	216	215	215	210	212	211	196	198	198	273	277	276	118	118	117
vdW	-964	-974	-977	-864	-826	-846	-725	-696	-700	-1093	-1083	-1097	-607	-613	-602
elec	-2697	-2799	-2712	-2786	-2818	-2717	-1998	-2066	-2176	-2797	-2955	-2797	-529	-579	-573
ΔG^{elec}	-1570	-1419	-1473	-1996	-1960	-2037	-1179	-1101	-983	-1725	-1597	-1717	-977	-913	-903
$\Delta G^{nonpolar}$	109	87	80	127	120	110	92	72	65	146	125	124	76	66	61
elec + ΔG^{elec}	-4267	-4218	-4185	-4782	-4778	-4754	-3177	-3167	-3159	-4522	-4522	-4514	-1506	-1492	-1476
vdW + $\Delta G^{nonpolar}$	-855	-887	-897	-737	-706	-736	-633	-624	-635	-947	-958	-973	-531	-547	-541
elec + vdW + ΔG^{elec} + $\Delta G^{nonpolar}$	-5122	-5105	-5082	-5519	-5484	-5490	-3810	-3791	-3794	-5469	-5510	-5487	-2037	-2039	-2017
pot - elec - vdW - ΔG^{elec} - $\Delta G^{nonpolar}$	1402	1397	1382	1363	1355	1385	1174	1177	1178	1845	1858	1843	771	779	774

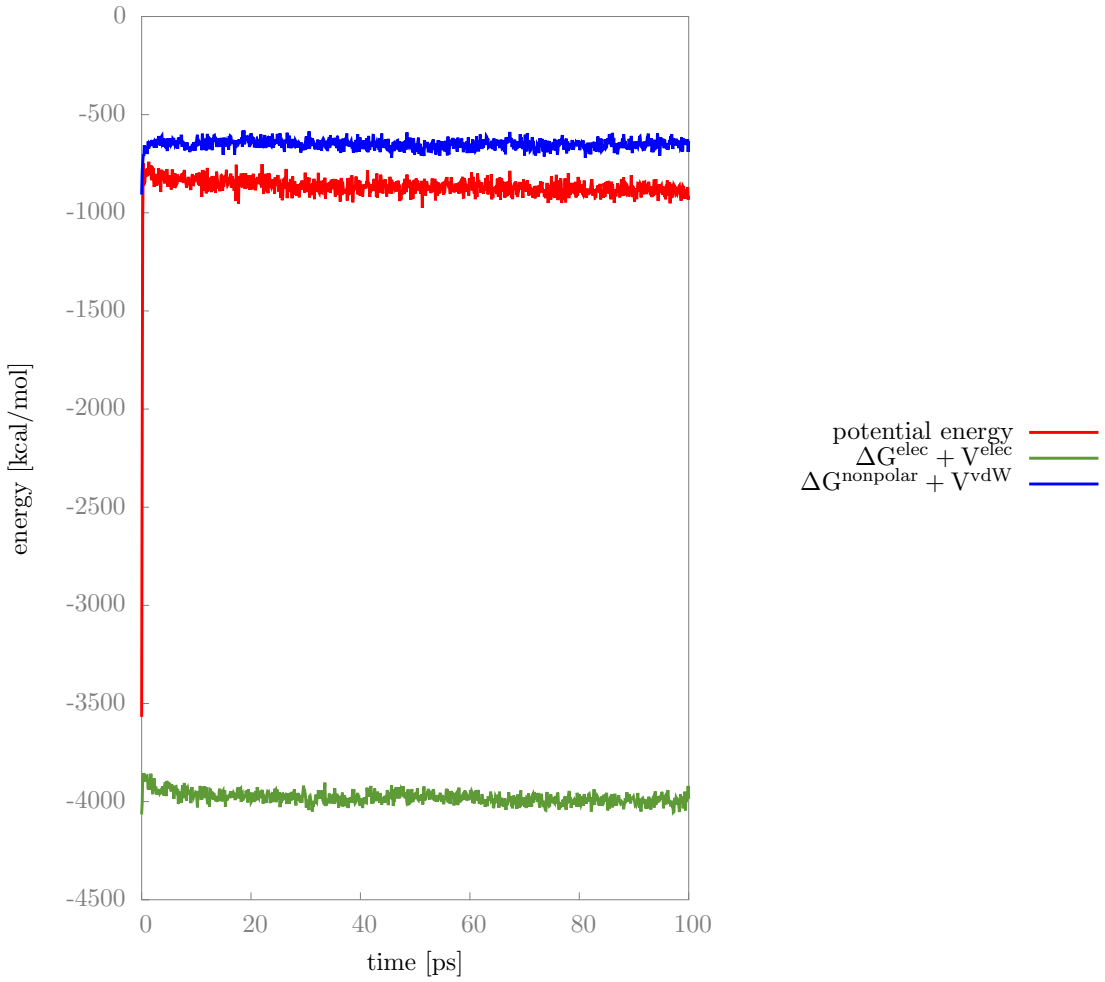

Figure 19: The energy development in kcal/mol for the native state of BRD7 using FACTSMEM at a width of 28.5 Å is plotted against the time in ps. The potential energy is shown in red, the nonpolar contributions to the nonbonded interactions are shown in blue and the polar contributions to the nonbonded interactions are shown in green.

5 Conclusion and outlook

Testing the discriminative power of the implicit membrane model FACTSMEM is an important benchmark in order to see if this model can be used effectively in folding simulations of membrane proteins, and to ensure that it will predict the correct stable structures close to the native state.

From the test series in this thesis, this question cannot be answered reliably. It was not possible to identify a subset of decoy structures leading to stable and reasonable relative energy ranges. With the given decoys it was not feasible to reproduce the results of the YaL paper. The same problem was observed when the IMM1 model was applied. It occurred, because the decoys react very differently for the different solvent models. For one model one obtains expectable results, which can drastically change when the same decoy is minimized in a different solvent model, turning into an unphysical structure with some far too long carbon-carbon bonds. This behaviour results from improving the van der Waals interactions which in all decoys are initially very large due to atom clashes.
For all of the 25,000 decoys there are extraordinarily high van der Waals energies at the beginning, which make up about 99% of the whole potential energy. While decreasing that value the bond energies increase considerably. In most cases finally both van der Waals and bond energies get to reasonable values with the bond energies being lower than before minimization. But in some cases, including decoys for which the minimization in other solvents was successful, the geometry optimization fails. Currently, this observation cannot be explained. It can only be assumed that the tremendous high energies for the initial van der Waals interactions cause an unstable minimization procedure.

It has to be asked, how Yuzlenko and Lazaridis achieved their stable small energy range using the same decoys. It can be assumed, that they e.g. changed the decoy ensemble for every model and membrane width. By doing so, one can reproduce their results. But if this is true, one still has to question the results, because then the results for different membrane models and widths are no longer comparable.

In fact, FACTSMEM gives reasonable values for the energy terms arising from it, like a stable solvation free energy. But it cannot be ensured that the results for the perfect discrimination, which was obtained here, are correct or arise from bad decoy structures being hardly minimizable. Therefore, these tests have to be repeated with new decoys in a future work.
Secondly, it would still be interesting to determine, why the minimizers are not able to relax the bond energies properly for some decoys and why the ability for a correct geometry optimization changes that much when only slightly changing the external conditions, such as a change in membrane width.

6 Appendix

6.1 Appendix 1 - CHARMM input scripts

The following scripts show the input files used for the CHARMM simulations. They are in the order setup procedure, initial minimization, MD run, search of the lowest energy structure in the second half of the trajectory, final minimization and then RMSD calculation.

Setup procedure:

```
* Setup of the decoy structure
*

!Set definitions
set top "~/mhenzgen/charmmlazaridis/ch45memfacts/toppar/ -
top_all22_prot_cmap.inp"
set par "~/mhenzgen/charmmlazaridis/ch45memfacts/toppar/gbsw/ -
par_all22_prot_gbsw.inp"

!Read topology and parameter files
read rtf card name @top
read param card name @par

!Read sequence from PDB file
open unit 12 card read name @inputpdb
read sequ pdb unit 12

!Generate PSF and the IC table
generate setu dcoy first NTER last CTER

rewind unit 12

bomlev -1
read coor pdb unit 12
bomlev 0

close unit 12

!Print atoms with undefined coordinates
define testa select segid ausg .and. ( .not. hydrogen ) -
.and. ( .not. init ) show end

!Bulid missing atoms
ic para
ic fill preserve
ic build
```

```
!print still missing coordinates
define testb select (.not. hydrogen) -
.and. (.not. init) show end

!Add H atoms
coor init sele type H* end
hbuild sele all end

!Print atoms with still undefined coordinates
define testall select .not. init show end

!Translate and rotate peptide until it is in the right position
coor orient
coor rota ydir 1.0 phi 90. sele all end

!calculate center of mass
coor rgyr sele all end

!write psf and coordinate file in crd
write psf card name @outputpsf
* PSF
*

write coor card name @outputcrd
* Coords
*

stop
```

Initial minimization:

```
* Initialminimization of the Inputstructure
*

!Set definitions
set top "~/mhenzgen/charmmlazaridis/ch45memfacts/toppar/ -
top_all22_prot_cmap.inp"
set par "~/mhenzgen/charmmlazaridis/ch45memfacts/toppar/gbsw/ -
par_all22_prot_gbsw.inp"

!Read topology and parameter files
read rtf card name @top
read param card name @par

! Read PSF and coordinates from file
read psf card name @inputpsf
read coor card name @inputcrd

! Set up FACTSMEM
!open unit 25 name eps.dat read form
!open unit 26 name np_inp.dat read form
```

```
set diele 1.0
nbond nbxmod 5 atom cdiel eps @diele shift vatom vdistance vswitch -
      cutnb 14.0 ctofnb 12.0 ctonnb 10.0 e14fac 1.0 wmin 1.5
scalar wmain = radius
fctmembrane tcps 22 tkps 4.0 teps @diele gamm 0.015 tavw
!           uneps 25 uns 26

!minimization
minimize abnr nstep 1000 nprint 20

write coor card name @outputcrd
* Coords
*

stop
```

MD run:

```
* Setup of the MD run
*

!Set definitions
set top "~/mhenzgen/charmmlazaridis/ch45memfacts/toppar/ -
top_all22_prot_cmap.inp"
set par "~/mhenzgen/charmmlazaridis/ch45memfacts/toppar/gbsw/ -
par_all22_prot_gbsw.inp"
set temp 310.0
set isee 53634

!Read topology and parameter files
read rtf card name @top
read param card name @par

! Read PSF and coordinates from file
read psf card name @inputpsf
read coor card name @inputcrd

! Set up FACTSMEM
!open unit 25 name eps.dat read form
!open unit 26 name np_inp.dat read form
set diele 1.0
nbond nbxmod 5 atom cdiel eps @diele shift vatom vdistance vswitch -
      cutnb 14.0 ctofnb 12.0 ctonnb 10.0 e14fac 1.0 wmin 1.5
scalar wmain = radius
fctmembrane tcps 22 tkps 4.0 teps @diele gamm 0.015 tavw
!           uneps 25 uns 26

shake bonh parameters tol 1.0e-10

! Verlet dynamics
! Dynamics specifications obtained from Themis Lazaridis
```

```
open unit 41 write card name @outputres
open unit 31 write file name @outputdcd
dynamics strt verlet -
      timestep 0.001 nstep 100000 nprint 1000 iprfrq 200 -
      ieqfrq 100 ichecw 1 -
      firstt 240 finalt 310 twindh 2.0 -
      iunwri 41 iuncrd 31 -
      iasors 0 iasvel 1 nsavc 1000 inbfrq -1
ioform extended

write coor card name @outputcrd
* Coords
*

stop
```

Searching lowest energy structure:

```
* Searching lowest energy structure from second half of trajectory
*

!Reading an existing trajectory

!Set definitions
set top "~/mhenzgen/charmmlazaridis/ch45memfacts/toppar/ -
top_all22_prot_cmap.inp"
set par "~/mhenzgen/charmmlazaridis/ch45memfacts/toppar/gbsw/ -
par_all22_prot_gbsw.inp"

!Read topology and parameter files
read rtf card name @top
read param card name @par

! Read PSF and trajectory from file
read psf card name @inputpsf
open unit 22 file name @inputdcd

TRAJECTORY FIRSTU 22 NUNIT 1 IREAD BEGIN 51000 STOP 100000 SKIP 1000

TRAJ READ
UPDATE INBFrq 1 IHBFrq 1

! Set up FACTSMEM
!open unit 25 name eps.dat read form
!open unit 26 name np_inp.dat read form
set diele 1.0
nbond nbxmod 5 atom cdiel eps @diele shift vatom vdistance vswitch -
      cutnb 14.0 ctofnb 12.0 ctonnb 10.0 e14fac 1.0 wmin 1.5
scalar wmain = radius
fctmembrane tcps 22 tkps 4.0 teps @diele gamm 0.015 tavw
!           uneps 25 uns 26
```

```
GETE
SET A 51
SET B ?ENER
COOR COPY COMP
SET OUT @A
INCR A BY 1

LABEL LOOP
   TRAJ READ
   UPDATE INBFrq 1 IHBFrq 1
   GETE
   IF B LT ?ENER THEN GOTO NEXT
        SET OUT @A
        COOR COPY COMP
        SET B ?ENER
   LABEL NEXT
   INCR A BY 1

IF A LE ?NFILE GOTO LOOP

OPEN WRITE CARD UNIT 12 NAME @outputcrd
WRITE COOR COMP CARD UNIT 12
* structure with the lowest energy
* frame number @OUT with energy @B
*

STOP
```

Final minimization:

```
* Finalminimization of the lowest energy structures
*

!Set definitions

set top "~/mhenzgen/charmmlazaridis/ch45memfacts/toppar/ -
top_all22_prot_cmap.inp"
set par "~/mhenzgen/charmmlazaridis/ch45memfacts/toppar/ -
gbsw/par_all22_prot_gbsw.inp"

!Read topology and parameter files
read rtf card name @top
read param card name @par

! Read PSF and coordinates from file
read psf card name @inputpsf
read coor card name @inputcrd

! Set up FACTSMEM
!open unit 25 name eps.dat read form
```

```
!open unit 26 name np_inp.dat read form
set diele 1.0
nbond nbxmod 5 atom cdiel eps @diele shift vatom vdistance vswitch -
      cutnb 14.0 ctofnb 12.0 ctonnb 10.0 e14fac 1.0 wmin 1.5
scalar wmain = radius
fctmembrane tcps 22 tkps 4.0 teps @diele gamm 0.015 tavw
!           uneps 25 uns 26

!minimization
minimize abnr nstep 1000 nprint 20

GETE
set ENERGY = ?Ener

write title name @outputenergy
*@ENERGY @decoy
*

write coor card name @outputcrd
* Coords of the lowest energy structure
* @ENERGY
*

stop
```

RMSD calculation:

```
* Compute final RMSD between native and decoys
*

!Set definitions
set top "~/mhenzgen/charmmlazaridis/ch45memfacts/toppar/ -
top_all22_prot_cmap.inp"
set par "~/mhenzgen/charmmlazaridis/ch45memfacts/toppar/ -
gbsw/par_all22_prot_gbsw.inp"

!Read topology and parameter files
read rtf card name @top
read param card name @par

!Read PSF and native coordinates
read psf card name ../ausgangsstruktur.psf
read coor card name ../ausgangsstruktur.crd

coor copy comp

!Read another set of coordinates
read coor card name @decoycrd

!Compute backbone RMSD values
DEFIne backbone SELEct TYPE N  .OR. TYPE CA .OR. TYPE C .OR. -
```

```
                      TYPE HN .OR. TYPE HA .OR. TYPE O .OR. -
                      TYPE NT .OR. TYPE HT1 .OR. TYPE HT2 .OR. -
                      TYPE CAY .OR. TYPE HY1 .OR. TYPE HY2 .OR. -
                      TYPE HY3 .OR. TYPE CY .OR. TYPE OY END
coor rms sele backbone end
set rmsd = ?rms

write title name rmsd.dat
*@rmsd @name
*

stop
```

6.2 Appendix 2 - Minimization files

With the information below one can follow the development of the initial and final minimization of a sample BRD7 decoy, which shows high energy bonds. Given is the energy output of the CHARMM minimization run with energies in kcal/mol.

Initial minimization:

```
MINI MIN: Cycle      ENERgy      Delta-E         GRMS    Step-size
MINI INTERN:          BONDs       ANGLes       UREY-b    DIHEdrals    IMPRopers
MINI CROSS:           CMAPs
MINI EXTERN:        VDWaals         ELEC       HBONds          ASP         USER
MINI FCTPOL:         FCTPOL
MINI FCTNPL:         FCTNPL
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>        0  0.71173E+12  0.00000E+00  0.35408E+12  0.00000E+00
MINI INTERN>      410.11957    461.91223    117.09170   1052.42566      7.23852
MINI CROSS>      -232.84505
MINI EXTERN>    0.71173E+12 -0.13215E+04  0.00000E+00  0.00000E+00  0.00000E+00
MINI FCTPOL>    -1732.19538
MINI FCTNPL>      103.73926
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>       20  0.10677E+10  0.71066E+12  0.19674E+09  0.25361E-01
MINI INTERN>      431.74665    467.63112    118.73430   1055.10995      7.28612
MINI CROSS>      -232.79012
MINI EXTERN>   1067716395.8      -1328.2          0.0          0.0          0.0
MINI FCTPOL>    -1732.32931
MINI FCTNPL>      103.73168
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>       40    1920591.3 1065794695.4      19216.0          0.2
MINI INTERN>     5483.37919   1361.87942    337.26213   1196.22484     47.90569
MINI CROSS>      -229.83718
MINI EXTERN>   1915416.5318   -1400.3188       0.0000       0.0000       0.0000
MINI FCTPOL>    -1725.34693
MINI FCTNPL>      103.59457
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>       60   61510.3679 1859080.9069     154.9097       1.8207
MINI INTERN>    25679.25472   3789.04545   2296.85238   1647.41246    233.37831
MINI CROSS>      -223.48889
MINI EXTERN>    31189.97622  -1476.63789      0.00000      0.00000      0.00000
MINI FCTPOL>    -1728.36282
MINI FCTNPL>      102.93795
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>       80  23416.89869  38093.46919     36.55596      3.51038
MINI INTERN>     9870.44506   3939.50922   1316.66702   1617.97899    234.83335
MINI CROSS>      -228.82420
MINI EXTERN>     9872.32220  -1566.72883      0.00000      0.00000      0.00000
MINI FCTPOL>    -1741.75998
```

```
MINI FCTNPL>      102.45588
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      100  13567.99575   9848.90294     13.05299      1.54014
MINI INTERN>     5642.25774   3136.59926    801.91864   1480.66380    126.53361
MINI CROSS>      -229.96014
MINI EXTERN>     5865.87197  -1603.23595      0.00000      0.00000      0.00000
MINI FCTPOL>    -1754.83856
MINI FCTNPL>      102.18538
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      120   6828.39893   6739.59682      5.34944      0.77056
MINI INTERN>     3869.65938   1946.25666    441.52320   1342.76747     65.20445
MINI CROSS>      -217.69380
MINI EXTERN>     2698.76487  -1645.18665      0.00000      0.00000      0.00000
MINI FCTPOL>    -1774.95117
MINI FCTNPL>      102.05452
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      140   4872.87499   1955.52395      3.79482      0.15614
MINI INTERN>     3639.55797   1389.22065    364.28016   1226.16663     47.77113
MINI CROSS>      -224.75507
MINI EXTERN>     1844.89295  -1730.04751      0.00000      0.00000      0.00000
MINI FCTPOL>    -1786.59528
MINI FCTNPL>      102.38335
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      160   3987.62808    885.24691      1.80317      0.24412
MINI INTERN>     3491.24888   1173.06196    356.69046   1179.01714     47.64121
MINI CROSS>      -237.59130
MINI EXTERN>     1457.37279  -1791.03307      0.00000      0.00000      0.00000
MINI FCTPOL>    -1791.33494
MINI FCTNPL>      102.55494
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      180   3367.05909    620.56899      1.49135      0.08244
MINI INTERN>     3429.64119    977.61467    342.63868   1156.18808     49.55201
MINI CROSS>      -252.03824
MINI EXTERN>     1156.18592  -1786.17074      0.00000      0.00000      0.00000
MINI FCTPOL>    -1809.59711
MINI FCTNPL>      103.04464
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      200   3031.11665    335.94244      1.41754      0.05156
MINI INTERN>     3408.74776    889.66272    330.89890   1124.46443     46.58433
MINI CROSS>      -262.57815
MINI EXTERN>      993.00672  -1783.96879      0.00000      0.00000      0.00000
MINI FCTPOL>    -1819.21765
MINI FCTNPL>      103.51638
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      220   2825.93369    205.18296      1.99452      0.22392
MINI INTERN>     3378.27673    841.13744    327.50405   1093.20098     44.62183
MINI CROSS>      -266.34512
MINI EXTERN>      939.68346  -1814.31716      0.00000      0.00000      0.00000
MINI FCTPOL>    -1821.80452
MINI FCTNPL>      103.97599
```

```
 ----------      ---------    ---------    ---------    ---------    ---------
MINI>      240   2716.09174    109.84194      1.18213      0.07376
MINI INTERN>     3404.14739    816.19766    323.05349   1099.23201     41.37623
MINI CROSS>      -274.28202
MINI EXTERN>      852.58711  -1828.53949      0.00000      0.00000      0.00000
MINI FCTPOL>    -1822.22684
MINI FCTNPL>      104.54619
 ----------      ---------    ---------    ---------    ---------    ---------
MINI>      260   2634.26961     81.82213      0.69086      0.02768
MINI INTERN>     3383.69762    801.46753    324.84739   1076.95901     40.93012
MINI CROSS>      -277.87387
MINI EXTERN>      837.88922  -1837.08038      0.00000      0.00000      0.00000
MINI FCTPOL>    -1821.26059
MINI FCTNPL>      104.69357
 ----------      ---------    ---------    ---------    ---------    ---------
MINI>      280   2571.59511     62.67450      0.64356      0.02597
MINI INTERN>     3361.30892    789.59346    325.97419   1051.11845     42.11387
MINI CROSS>      -280.99755
MINI EXTERN>      846.74747  -1847.99144      0.00000      0.00000      0.00000
MINI FCTPOL>    -1821.11450
MINI FCTNPL>      104.84224
 ----------      ---------    ---------    ---------    ---------    ---------
MINI>      300   2529.36244     42.23267      0.92577      0.03609
MINI INTERN>     3355.41146    798.55151    322.64745   1053.46593     39.33336
MINI CROSS>      -282.96418
MINI EXTERN>      831.79175  -1879.59015      0.00000      0.00000      0.00000
MINI FCTPOL>    -1814.25557
MINI FCTNPL>      104.97088
 ----------      ---------    ---------    ---------    ---------    ---------
MINI>      320   2499.06262     30.29983      0.51244      0.02031
MINI INTERN>     3366.88165    782.95025    324.14618   1044.18793     39.63263
MINI CROSS>      -282.41120
MINI EXTERN>      816.58768  -1887.34434      0.00000      0.00000      0.00000
MINI FCTPOL>    -1810.59716
MINI FCTNPL>      105.02899
 ----------      ---------    ---------    ---------    ---------    ---------
MINI>      340   2456.20764     42.85497      0.84868      0.03176
MINI INTERN>     3393.20159    761.07174    325.33270   1038.14407     39.39102
MINI CROSS>      -282.65525
MINI EXTERN>      790.88632  -1914.61330      0.00000      0.00000      0.00000
MINI FCTPOL>    -1799.87374
MINI FCTNPL>      105.32250
 ----------      ---------    ---------    ---------    ---------    ---------
MINI>      360   2406.90781     49.29984      1.35647      0.05517
MINI INTERN>     3379.15785    764.21466    322.33592   1037.69421     38.31755
MINI CROSS>      -284.75195
MINI EXTERN>      790.47754  -1944.01153      0.00000      0.00000      0.00000
MINI FCTPOL>    -1802.14362
MINI FCTNPL>      105.61717
 ----------      ---------    ---------    ---------    ---------    ---------
```

```
MINI>      380   2379.13877     27.76903      0.71984      0.03105
MINI INTERN>     3379.28720    766.14896    323.18313   1037.23902     37.35539
MINI CROSS>      -285.44198
MINI EXTERN>      778.84148  -1956.93430      0.00000      0.00000      0.00000
MINI FCTPOL>    -1806.40545
MINI FCTNPL>      105.86532
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      400   2365.18900     13.94977      0.47772      0.01748
MINI INTERN>     3383.92364    761.31100    323.82734   1034.04882     37.26867
MINI CROSS>      -285.96326
MINI EXTERN>      770.91941  -1959.82375      0.00000      0.00000      0.00000
MINI FCTPOL>    -1806.22921
MINI FCTNPL>      105.90634
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      420   2318.14016     47.04884      0.67709      0.03036
MINI INTERN>     3351.08965    765.12862    323.99813   1035.13042     37.63858
MINI CROSS>      -285.27844
MINI EXTERN>      788.74135  -1992.15405      0.00000      0.00000      0.00000
MINI FCTPOL>    -1812.57331
MINI FCTNPL>      106.41921
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      440   2300.68981     17.45035      0.42821      0.01709
MINI INTERN>     3365.40995    763.63186    325.03344   1033.65524     37.28153
MINI CROSS>      -285.33560
MINI EXTERN>      770.75770  -2003.52676      0.00000      0.00000      0.00000
MINI FCTPOL>    -1812.68479
MINI FCTNPL>      106.46723
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      460   2282.01447     18.67534      0.64397      0.02672
MINI INTERN>     3383.99616    758.07126    325.91950   1028.90709     36.68347
MINI CROSS>      -284.69157
MINI EXTERN>      752.65898  -2014.82933      0.00000      0.00000      0.00000
MINI FCTPOL>    -1811.16101
MINI FCTNPL>      106.45992
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      480   2271.27392     10.74055      0.30285      0.01504
MINI INTERN>     3361.85644    757.26913    324.74224   1028.31490     36.92808
MINI CROSS>      -284.72794
MINI EXTERN>      772.56944  -2022.34740      0.00000      0.00000      0.00000
MINI FCTPOL>    -1809.81322
MINI FCTNPL>      106.48224
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      500   2257.48679     13.78713      0.31122      0.01568
MINI INTERN>     3369.60372    757.19548    324.21511   1029.92494     36.70571
MINI CROSS>      -284.41187
MINI EXTERN>      762.09735  -2034.76931      0.00000      0.00000      0.00000
MINI FCTPOL>    -1809.65861
MINI FCTNPL>      106.58426
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      520   2238.37201     19.11478      0.42100      0.01470
```

```
MINI INTERN>     3370.21211    748.81214    324.12704   1029.50622     36.04584
MINI CROSS>      -285.04674
MINI EXTERN>      759.99039  -2037.28794      0.00000      0.00000      0.00000
MINI FCTPOL>    -1814.70944
MINI FCTNPL>      106.72240
 ----------      ---------    ---------    ---------    ---------    ---------
MINI>      540   2228.80642      9.56559      0.25669      0.01379
MINI INTERN>     3364.20276    752.12837    323.83020   1027.39717     36.21897
MINI CROSS>      -285.02916
MINI EXTERN>      763.70444  -2043.79785      0.00000      0.00000      0.00000
MINI FCTPOL>    -1816.62048
MINI FCTNPL>      106.77201
 ----------      ---------    ---------    ---------    ---------    ---------
MINI>      560   2220.72703      8.07939      0.32334      0.01294
MINI INTERN>     3373.94481    745.83093    324.40226   1024.90794     35.97205
MINI CROSS>      -284.82573
MINI EXTERN>      751.69402  -2039.47916      0.00000      0.00000      0.00000
MINI FCTPOL>    -1818.50541
MINI FCTNPL>      106.78532
 ----------      ---------    ---------    ---------    ---------    ---------
MINI>      580   2210.64250     10.08453      0.98297      0.03330
MINI INTERN>     3381.14624    749.30932    324.62737   1022.11026     36.00417
MINI CROSS>      -284.47552
MINI EXTERN>      743.37596  -2049.51038      0.00000      0.00000      0.00000
MINI FCTPOL>    -1818.81302
MINI FCTNPL>      106.86811
 ----------      ---------    ---------    ---------    ---------    ---------
MINI>      600   2195.23303     15.40947      0.40879      0.02083
MINI INTERN>     3360.53624    749.78095    323.19227   1017.75776     35.36639
MINI CROSS>      -285.21228
MINI EXTERN>      760.78513  -2052.82203      0.00000      0.00000      0.00000
MINI FCTPOL>    -1821.17376
MINI FCTNPL>      107.02237
 ----------      ---------    ---------    ---------    ---------    ---------
MINI>      620   2186.46475      8.76829      0.28267      0.01172
MINI INTERN>     3368.29571    745.53873    323.91699   1016.74026     35.21980
MINI CROSS>      -285.97167
MINI EXTERN>      751.38479  -2052.60995      0.00000      0.00000      0.00000
MINI FCTPOL>    -1823.14158
MINI FCTNPL>      107.09166
 ----------      ---------    ---------    ---------    ---------    ---------
MINI>      640   2178.22301      8.24174      0.29475      0.01100
MINI INTERN>     3360.72937    741.85734    323.27748   1017.01329     35.13726
MINI CROSS>      -287.05307
MINI EXTERN>      756.92974  -2050.74749      0.00000      0.00000      0.00000
MINI FCTPOL>    -1826.05493
MINI FCTNPL>      107.13402
 ----------      ---------    ---------    ---------    ---------    ---------
MINI>      660   2171.44908      6.77393      0.27371      0.01032
MINI INTERN>     3355.58049    740.53030    322.80733   1015.49501     34.94826
```

```
MINI CROSS>      -287.50810
MINI EXTERN>      762.05632  -2050.67286      0.00000      0.00000      0.00000
MINI FCTPOL>    -1828.94248
MINI FCTNPL>      107.15480
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      680   2166.62457      4.82451      0.24968      0.01075
MINI INTERN>     3366.06467    739.88056    323.73543   1011.15286     34.99401
MINI CROSS>      -287.24743
MINI EXTERN>      750.13178  -2045.30174      0.00000      0.00000      0.00000
MINI FCTPOL>    -1833.96960
MINI FCTNPL>      107.18403
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      700   2162.50502      4.11955      0.11412      0.00605
MINI INTERN>     3362.18932    739.10078    323.37606   1003.48688     34.94649
MINI CROSS>      -286.91210
MINI EXTERN>      751.99645  -2029.94906      0.00000      0.00000      0.00000
MINI FCTPOL>    -1842.95309
MINI FCTNPL>      107.22330
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      720   2157.06470      5.44033      0.44389      0.01577
MINI INTERN>     3366.85869    738.65792    323.79028   1000.99786     35.17595
MINI CROSS>      -286.73020
MINI EXTERN>      747.32248  -2025.62774      0.00000      0.00000      0.00000
MINI FCTPOL>    -1850.67203
MINI FCTNPL>      107.29148
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      740   2149.11323      7.95146      0.43797      0.01480
MINI INTERN>     3365.93693    737.33617    323.37126   1000.24866     34.78205
MINI CROSS>      -286.96814
MINI EXTERN>      745.12314  -2026.20334      0.00000      0.00000      0.00000
MINI FCTPOL>    -1851.91673
MINI FCTNPL>      107.40322
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      760   2145.01348      4.09975      0.17430      0.00833
MINI INTERN>     3361.97056    737.01541    323.44981    999.76001     34.55115
MINI CROSS>      -286.86227
MINI EXTERN>      748.84689  -2029.09057      0.00000      0.00000      0.00000
MINI FCTPOL>    -1852.06782
MINI FCTNPL>      107.44031
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      780   2140.51877      4.49471      0.25328      0.01302
MINI INTERN>     3363.97959    735.53105    323.53048   1000.60900     34.36929
MINI CROSS>      -286.68916
MINI EXTERN>      744.66186  -2033.53503      0.00000      0.00000      0.00000
MINI FCTPOL>    -1849.42074
MINI FCTNPL>      107.48243
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      800   2139.02658      1.49219      0.12115      0.00440
MINI INTERN>     3363.92213    735.30545    323.49184   1000.08750     34.42053
MINI CROSS>      -286.73180
```

```
MINI EXTERN>      745.91571  -2038.23361      0.00000      0.00000      0.00000
MINI FCTPOL>    -1846.61387
MINI FCTNPL>      107.46270
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      820   2135.17087      3.85571      0.32670      0.01146
MINI INTERN>     3354.36643    736.22212    322.69778    999.54867     34.86432
MINI CROSS>      -286.49128
MINI EXTERN>      755.79913  -2048.59779      0.00000      0.00000      0.00000
MINI FCTPOL>    -1840.66030
MINI FCTNPL>      107.42179
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      840   2128.79303      6.37784      0.26281      0.01075
MINI INTERN>     3359.20262    737.37900    322.92746   1000.27543     34.43210
MINI CROSS>      -285.62156
MINI EXTERN>      748.47351  -2058.55012      0.00000      0.00000      0.00000
MINI FCTPOL>    -1837.11477
MINI FCTNPL>      107.38937
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      860   2125.54012      3.25291      0.10652      0.00605
MINI INTERN>     3360.97191    737.11571    323.17316    999.66220     34.47545
MINI CROSS>      -285.53101
MINI EXTERN>      747.25792  -2062.26351      0.00000      0.00000      0.00000
MINI FCTPOL>    -1836.72116
MINI FCTNPL>      107.39944
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      880   2122.95777      2.58235      0.27507      0.00946
MINI INTERN>     3365.86475    735.58692    323.61364    999.47330     34.27601
MINI CROSS>      -285.59296
MINI EXTERN>      743.97393  -2065.29639      0.00000      0.00000      0.00000
MINI FCTPOL>    -1836.38496
MINI FCTNPL>      107.44354
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      900   2118.90545      4.05233      0.37576      0.01479
MINI INTERN>     3369.96269    734.63711    323.97191    998.99385     34.17668
MINI CROSS>      -285.57544
MINI EXTERN>      737.82817  -2066.56933      0.00000      0.00000      0.00000
MINI FCTPOL>    -1835.96799
MINI FCTNPL>      107.44779
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      920   2114.68916      4.21628      0.18113      0.00832
MINI INTERN>     3365.92495    734.10640    323.52919    998.43139     34.26985
MINI CROSS>      -286.12432
MINI EXTERN>      742.47283  -2071.32841      0.00000      0.00000      0.00000
MINI FCTPOL>    -1833.97983
MINI FCTNPL>      107.38712
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      940   2112.18279      2.50638      0.14234      0.00469
MINI INTERN>     3362.91450    734.50624    323.49495    997.96605     34.11804
MINI CROSS>      -286.26466
MINI EXTERN>      744.38918  -2072.62428      0.00000      0.00000      0.00000
```

```
MINI FCTPOL>    -1833.73779
MINI FCTNPL>      107.42054
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      960   2109.74495      2.43783      0.20715      0.00733
MINI INTERN>     3365.29942    735.23992    323.88368    996.69043     33.97704
MINI CROSS>      -286.36114
MINI EXTERN>      742.21171  -2074.27268      0.00000      0.00000      0.00000
MINI FCTPOL>    -1834.41919
MINI FCTNPL>      107.49577
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      980   2108.10673      1.63823      0.11418      0.00412
MINI INTERN>     3362.20465    734.55646    323.51961    997.23954     33.92698
MINI CROSS>      -286.37095
MINI EXTERN>      743.94730  -2073.49554      0.00000      0.00000      0.00000
MINI FCTPOL>    -1834.89399
MINI FCTNPL>      107.47265
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>     1000   2105.47386      2.63287      0.51113      0.01791
MINI INTERN>     3367.14727    734.43633    323.94456    997.52412     33.90728
MINI CROSS>      -286.58711
MINI EXTERN>      739.39521  -2077.63922      0.00000      0.00000      0.00000
MINI FCTPOL>    -1834.12343
MINI FCTNPL>      107.46884
 ----------       ---------    ---------    ---------    ---------    ---------

 ABNER> Minimization exiting with number of steps limit ( 1000) exceeded.

ABNR MIN: Cycle      ENERgy      Delta-E         GRMS    Step-size
ABNR INTERN:          BONDs       ANGLes       UREY-b    DIHEdrals    IMPRopers
ABNR CROSS:           CMAPs
ABNR EXTERN:        VDWaals         ELEC       HBONds          ASP         USER
ABNR FCTPOL:          FCTPOL
ABNR FCTNPL:          FCTNPL
 ----------       ---------    ---------    ---------    ---------    ---------
ABNR>     1000   2105.47386      2.63287      0.51113      0.01612
ABNR INTERN>     3367.14727    734.43633    323.94456    997.52412     33.90728
ABNR CROSS>      -286.58711
ABNR EXTERN>      739.39521  -2077.63922      0.00000      0.00000      0.00000
ABNR FCTPOL>    -1834.12343
ABNR FCTNPL>      107.46884
 ----------       ---------    ---------    ---------    ---------    ---------
```

Final minimization:

```
MINI MIN: Cycle      ENERgy      Delta-E         GRMS    Step-size
MINI INTERN:          BONDs       ANGLes       UREY-b    DIHEdrals    IMPRopers
MINI CROSS:           CMAPs
MINI EXTERN:        VDWaals         ELEC       HBONds          ASP         USER
MINI FCTPOL:          FCTPOL
MINI FCTNPL:          FCTNPL
 ----------       ---------    ---------    ---------    ---------    ---------
```

```
MINI>        0   3508.67489      0.00000     14.24081      0.00000
MINI INTERN>     3750.35311   1539.43413    462.99564   1121.55206     91.93114
MINI CROSS>      -289.60584
MINI EXTERN>      703.14985  -2363.89732      0.00000      0.00000      0.00000
MINI FCTPOL>    -1620.52305
MINI FCTNPL>      113.28516
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>       20   1832.31065   1676.36424      1.33572      0.07721
MINI INTERN>     3347.10419    763.71810    327.88196    972.26048     30.90271
MINI CROSS>      -298.26253
MINI EXTERN>      621.79487  -2413.34008      0.00000      0.00000      0.00000
MINI FCTPOL>    -1632.96294
MINI FCTNPL>      113.21388
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>       40   1711.72098    120.58967      2.16520      0.07243
MINI INTERN>     3365.32098    720.70405    327.56686    930.15712     28.13431
MINI CROSS>      -299.11810
MINI EXTERN>      588.21904  -2426.07778      0.00000      0.00000      0.00000
MINI FCTPOL>    -1636.35655
MINI FCTNPL>      113.17104
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>       60   1662.50753     49.21344      0.43662      0.02446
MINI INTERN>     3351.57951    711.46796    325.06167    917.52412     27.36159
MINI CROSS>      -298.72801
MINI EXTERN>      586.68976  -2434.29127      0.00000      0.00000      0.00000
MINI FCTPOL>    -1637.33909
MINI FCTNPL>      113.18129
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>       80   1632.75127     29.75627      0.64685      0.02295
MINI INTERN>     3333.19408    708.22061    323.90920    907.63232     27.24067
MINI CROSS>      -297.83445
MINI EXTERN>      599.40158  -2442.61635      0.00000      0.00000      0.00000
MINI FCTPOL>    -1639.62893
MINI FCTNPL>      113.23253
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      100   1622.01245     10.73882      0.18848      0.00775
MINI INTERN>     3348.20224    703.39246    324.33452    904.91860     26.86973
MINI CROSS>      -298.48492
MINI EXTERN>      584.04814  -2443.53319      0.00000      0.00000      0.00000
MINI FCTPOL>    -1640.99267
MINI FCTNPL>      113.25754
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      120   1606.82743     15.18502      0.64549      0.02019
MINI INTERN>     3366.21202    701.04899    325.21553    902.17467     26.72258
MINI CROSS>      -299.92603
MINI EXTERN>      561.89252  -2446.01243      0.00000      0.00000      0.00000
MINI FCTPOL>    -1643.79172
MINI FCTNPL>      113.29131
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      140   1601.64475      5.18268      0.17509      0.01137
```

```
MINI INTERN>     3349.82908    700.20104    323.87646    900.22021     26.82200
MINI CROSS>      -300.00909
MINI EXTERN>      580.01881  -2448.50445      0.00000      0.00000      0.00000
MINI FCTPOL>    -1644.09775
MINI FCTNPL>      113.28844
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      160   1594.96443      6.68032      0.25626      0.01066
MINI INTERN>     3344.96431    700.64568    323.62673    898.49930     26.99000
MINI CROSS>      -300.04497
MINI EXTERN>      587.79987  -2455.77529      0.00000      0.00000      0.00000
MINI FCTPOL>    -1645.00986
MINI FCTNPL>      113.26868
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      180   1583.44239     11.52204      0.40799      0.01667
MINI INTERN>     3333.52592    700.96212    322.76834    896.85722     26.86722
MINI CROSS>      -300.20119
MINI EXTERN>      599.64682  -2464.45996      0.00000      0.00000      0.00000
MINI FCTPOL>    -1645.74024
MINI FCTNPL>      113.21615
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      200   1579.66115      3.78124      0.19151      0.00938
MINI INTERN>     3347.07291    698.46989    324.06629    896.15565     26.84154
MINI CROSS>      -300.53948
MINI EXTERN>      584.70098  -2465.43814      0.00000      0.00000      0.00000
MINI FCTPOL>    -1644.88689
MINI FCTNPL>      113.21840
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      220   1577.02717      2.63398      0.19802      0.00880
MINI INTERN>     3345.37307    698.07196    323.84948    896.67010     26.83722
MINI CROSS>      -300.75766
MINI EXTERN>      584.23589  -2465.78602      0.00000      0.00000      0.00000
MINI FCTPOL>    -1644.66654
MINI FCTNPL>      113.19967
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      240   1572.60930      4.41787      0.41299      0.01376
MINI INTERN>     3348.43700    699.04081    323.40745    896.18563     26.84155
MINI CROSS>      -301.05053
MINI EXTERN>      581.07992  -2468.69465      0.00000      0.00000      0.00000
MINI FCTPOL>    -1645.78454
MINI FCTNPL>      113.14667
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      260   1571.47439      1.13491      0.10425      0.00465
MINI INTERN>     3345.35736    698.61835    323.68856    895.37476     26.83425
MINI CROSS>      -300.71619
MINI EXTERN>      584.59468  -2470.45189      0.00000      0.00000      0.00000
MINI FCTPOL>    -1644.96835
MINI FCTNPL>      113.14285
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      280   1570.29661      1.17778      0.42468      0.01211
MINI INTERN>     3349.26046    698.98069    324.12901    895.47234     26.85641
```

```
MINI CROSS>      -300.89217
MINI EXTERN>      579.56243  -2470.22826      0.00000      0.00000      0.00000
MINI FCTPOL>    -1645.97191
MINI FCTNPL>      113.12759
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      300   1568.33007      1.96654      0.13026      0.00409
MINI INTERN>     3350.68102    698.77999    323.96951    894.85349     26.81674
MINI CROSS>      -300.71592
MINI EXTERN>      579.81019  -2473.66755      0.00000      0.00000      0.00000
MINI FCTPOL>    -1645.29656
MINI FCTNPL>      113.09917
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      320   1567.41104      0.91903      0.16714      0.00639
MINI INTERN>     3351.27585    698.57474    323.82070    894.57668     26.84944
MINI CROSS>      -300.51367
MINI EXTERN>      579.39812  -2474.86753      0.00000      0.00000      0.00000
MINI FCTPOL>    -1644.77790
MINI FCTNPL>      113.07461
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      340   1561.58920      5.82184      0.27882      0.01666
MINI INTERN>     3349.17447    699.97553    323.50587    894.86340     26.97252
MINI CROSS>      -300.12612
MINI EXTERN>      583.24190  -2484.43159      0.00000      0.00000      0.00000
MINI FCTPOL>    -1644.45728
MINI FCTNPL>      112.87049
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      360   1560.01181      1.57739      0.09512      0.00338
MINI INTERN>     3345.52598    699.69286    323.48699    894.19743     26.97362
MINI CROSS>      -300.22314
MINI EXTERN>      586.78018  -2484.52080      0.00000      0.00000      0.00000
MINI FCTPOL>    -1644.78703
MINI FCTNPL>      112.88572
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      380   1559.36237      0.64943      0.07126      0.00317
MINI INTERN>     3348.88560    699.49796    323.78859    893.84767     26.95778
MINI CROSS>      -300.28525
MINI EXTERN>      583.25935  -2484.28561      0.00000      0.00000      0.00000
MINI FCTPOL>    -1645.18745
MINI FCTNPL>      112.88373
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      400   1558.32967      1.03270      0.06658      0.00297
MINI INTERN>     3346.43309    700.13328    323.61655    893.40112     26.93874
MINI CROSS>      -300.33892
MINI EXTERN>      584.61113  -2483.53318      0.00000      0.00000      0.00000
MINI FCTPOL>    -1645.78797
MINI FCTNPL>      112.85582
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      420   1557.97445      0.35521      0.14348      0.00465
MINI INTERN>     3348.35269    699.81062    323.68307    893.38805     26.91362
MINI CROSS>      -300.27702
```

```
MINI EXTERN>      582.81813  -2484.27510      0.00000      0.00000      0.00000
MINI FCTPOL>    -1645.28463
MINI FCTNPL>      112.84501
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      440   1557.09279      0.88166      0.11585      0.00436
MINI INTERN>     3346.26241    699.95129    323.63714    893.43347     26.93402
MINI CROSS>      -299.94055
MINI EXTERN>      585.30193  -2486.87727      0.00000      0.00000      0.00000
MINI FCTPOL>    -1644.37792
MINI FCTNPL>      112.76827
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      460   1554.86493      2.22786      0.31727      0.01136
MINI INTERN>     3355.13699    699.80917    323.86271    892.93419     26.99065
MINI CROSS>      -299.92985
MINI EXTERN>      575.86128  -2487.45328      0.00000      0.00000      0.00000
MINI FCTPOL>    -1645.03682
MINI FCTNPL>      112.68990
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      480   1553.99922      0.86571      0.04936      0.00230
MINI INTERN>     3346.77577    700.48043    323.49834    893.05882     26.94922
MINI CROSS>      -299.92460
MINI EXTERN>      583.65651  -2488.27388      0.00000      0.00000      0.00000
MINI FCTPOL>    -1644.89098
MINI FCTNPL>      112.66961
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      500   1552.84045      1.15878      0.35842      0.00999
MINI INTERN>     3359.26425    700.05265    324.21666    893.51417     27.04996
MINI CROSS>      -300.14578
MINI EXTERN>      570.72423  -2489.94442      0.00000      0.00000      0.00000
MINI FCTPOL>    -1644.48408
MINI FCTNPL>      112.59281
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      520   1552.40216      0.43829      0.04706      0.00202
MINI INTERN>     3348.39598    700.68360    323.70010    893.16117     26.98423
MINI CROSS>      -299.99919
MINI EXTERN>      581.61347  -2490.26080      0.00000      0.00000      0.00000
MINI FCTPOL>    -1644.46666
MINI FCTNPL>      112.59026
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      540   1552.09560      0.30656      0.05498      0.00190
MINI INTERN>     3349.57168    700.41757    323.68348    893.17011     27.00510
MINI CROSS>      -299.90781
MINI EXTERN>      580.35473  -2490.34013      0.00000      0.00000      0.00000
MINI FCTPOL>    -1644.43513
MINI FCTNPL>      112.57599
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      560   1551.03293      1.06267      0.24441      0.00825
MINI INTERN>     3353.43946    700.46029    323.84018    893.42410     27.08416
MINI CROSS>      -299.61148
MINI EXTERN>      577.65385  -2494.96355      0.00000      0.00000      0.00000
```

```
MINI FCTPOL>    -1642.74954
MINI FCTNPL>      112.45547
 ----------      ---------    ---------    ---------    ---------    ---------
MINI>      580   1550.37967      0.65326      0.11352      0.00464
MINI INTERN>     3346.05175    700.99288    323.48468    893.28003     27.09795
MINI CROSS>      -299.68811
MINI EXTERN>      584.33166  -2494.46251      0.00000      0.00000      0.00000
MINI FCTPOL>    -1643.16648
MINI FCTNPL>      112.45782
 ----------      ---------    ---------    ---------    ---------    ---------
MINI>      600   1549.97211      0.40756      0.06323      0.00261
MINI INTERN>     3346.07257    700.97847    323.52388    893.27498     27.08958
MINI CROSS>      -299.78709
MINI EXTERN>      583.68725  -2494.01528      0.00000      0.00000      0.00000
MINI FCTPOL>    -1643.30393
MINI FCTNPL>      112.45168
 ----------      ---------    ---------    ---------    ---------    ---------
MINI>      620   1549.61983      0.35228      0.15005      0.00409
MINI INTERN>     3345.73114    701.14458    323.50432    893.47531     27.10250
MINI CROSS>      -299.89998
MINI EXTERN>      583.63148  -2494.09561      0.00000      0.00000      0.00000
MINI FCTPOL>    -1643.40362
MINI FCTNPL>      112.42972
 ----------      ---------    ---------    ---------    ---------    ---------
MINI>      640   1549.09641      0.52342      0.08327      0.00383
MINI INTERN>     3348.87228    701.01095    323.69081    893.55877     27.09629
MINI CROSS>      -299.68371
MINI EXTERN>      580.55248  -2495.44877      0.00000      0.00000      0.00000
MINI FCTPOL>    -1642.93665
MINI FCTNPL>      112.38397
 ----------      ---------    ---------    ---------    ---------    ---------
MINI>      660   1548.42142      0.67500      0.10489      0.00360
MINI INTERN>     3350.60669    700.99802    323.78871    893.53689     27.12326
MINI CROSS>      -299.62947
MINI EXTERN>      578.93531  -2496.89582      0.00000      0.00000      0.00000
MINI FCTPOL>    -1642.38114
MINI FCTNPL>      112.33896
 ----------      ---------    ---------    ---------    ---------    ---------
MINI>      680   1547.73968      0.68173      0.08768      0.00337
MINI INTERN>     3348.75734    700.98869    323.67698    893.66100     27.13761
MINI CROSS>      -299.63517
MINI EXTERN>      580.61200  -2497.52230      0.00000      0.00000      0.00000
MINI FCTPOL>    -1642.25028
MINI FCTNPL>      112.31380
 ----------      ---------    ---------    ---------    ---------    ---------
MINI>      700   1547.26311      0.47657      0.08478      0.00316
MINI INTERN>     3349.39913    700.60732    323.71232    893.82050     27.14907
MINI CROSS>      -299.52590
MINI EXTERN>      579.65187  -2497.48445      0.00000      0.00000      0.00000
MINI FCTPOL>    -1642.35497
```

```
MINI FCTNPL>      112.28824
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      720   1545.79484      1.46828      0.10165      0.00495
MINI INTERN>     3347.53089    701.41634    323.56206    894.73300     27.15315
MINI CROSS>      -299.11785
MINI EXTERN>      581.35281  -2501.07597      0.00000      0.00000      0.00000
MINI FCTPOL>    -1641.88962
MINI FCTNPL>      112.13003
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      740   1544.33198      1.46285      0.11818      0.00464
MINI INTERN>     3351.01304    701.48345    323.86552    894.34162     27.15941
MINI CROSS>      -298.81282
MINI EXTERN>      578.27252  -2503.87369      0.00000      0.00000      0.00000
MINI FCTPOL>    -1641.18078
MINI FCTNPL>      112.06371
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      760   1543.86722      0.46476      0.07422      0.00261
MINI INTERN>     3349.97720    700.99562    323.72257    894.28852     27.16698
MINI CROSS>      -298.78549
MINI EXTERN>      579.11490  -2503.72167      0.00000      0.00000      0.00000
MINI FCTPOL>    -1640.94131
MINI FCTNPL>      112.04990
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      780   1543.68602      0.18120      0.06192      0.00245
MINI INTERN>     3346.07059    701.09609    323.47454    894.18024     27.19926
MINI CROSS>      -298.83561
MINI EXTERN>      582.90038  -2503.60419      0.00000      0.00000      0.00000
MINI FCTPOL>    -1640.83932
MINI FCTNPL>      112.04404
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      800   1543.41502      0.27100      0.04052      0.00138
MINI INTERN>     3347.47649    700.72924    323.57494    893.94775     27.15787
MINI CROSS>      -298.89784
MINI EXTERN>      581.25060  -2503.37156      0.00000      0.00000      0.00000
MINI FCTPOL>    -1640.47703
MINI FCTNPL>      112.02454
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      820   1543.16418      0.25084      0.06373      0.00216
MINI INTERN>     3349.60648    700.82119    323.74642    893.91996     27.16093
MINI CROSS>      -298.75347
MINI EXTERN>      579.03513  -2503.47191      0.00000      0.00000      0.00000
MINI FCTPOL>    -1640.91166
MINI FCTNPL>      112.01111
 ----------       ---------    ---------    ---------    ---------    ---------
MINI>      840   1543.00094      0.16324      0.03625      0.00202
MINI INTERN>     3347.98045    700.93121    323.62185    893.93685     27.14942
MINI CROSS>      -298.79927
MINI EXTERN>      580.23287  -2503.39719      0.00000      0.00000      0.00000
MINI FCTPOL>    -1640.66104
MINI FCTNPL>      112.00578
```

```
 ----------      ---------    ---------    ---------    ---------    ---------
MINI>      860   1542.86478      0.13616      0.05739      0.00190
MINI INTERN>     3346.01241    700.77962    323.54399    893.86524     27.13676
MINI CROSS>      -298.81027
MINI EXTERN>      582.57726  -2503.89758      0.00000      0.00000      0.00000
MINI FCTPOL>    -1640.33467
MINI FCTNPL>      111.99202
 ----------      ---------    ---------    ---------    ---------    ---------
MINI>      880   1542.48683      0.37795      0.07943      0.00297
MINI INTERN>     3345.43502    700.79209    323.44674    894.15587     27.13306
MINI CROSS>      -298.73585
MINI EXTERN>      582.73897  -2504.58739      0.00000      0.00000      0.00000
MINI FCTPOL>    -1639.82456
MINI FCTNPL>      111.93288
 ----------      ---------    ---------    ---------    ---------    ---------
MINI>      900   1541.48276      1.00407      0.12896      0.00464
MINI INTERN>     3344.32111    700.85144    323.53141    894.85369     27.12085
MINI CROSS>      -298.76933
MINI EXTERN>      582.04091  -2505.13512      0.00000      0.00000      0.00000
MINI FCTPOL>    -1639.16361
MINI FCTNPL>      111.83141
 ----------      ---------    ---------    ---------    ---------    ---------
MINI>      920   1540.93052      0.55224      0.10773      0.00435
MINI INTERN>     3347.26009    700.55195    323.70324    894.99611     27.14947
MINI CROSS>      -298.72239
MINI EXTERN>      579.13177  -2505.86574      0.00000      0.00000      0.00000
MINI FCTPOL>    -1639.09155
MINI FCTNPL>      111.81756
 ----------      ---------    ---------    ---------    ---------    ---------
MINI>      940   1540.67688      0.25363      0.03572      0.00147
MINI INTERN>     3346.81718    700.63766    323.64575    894.96069     27.11734
MINI CROSS>      -298.69570
MINI EXTERN>      579.55630  -2506.07008      0.00000      0.00000      0.00000
MINI FCTPOL>    -1639.10840
MINI FCTNPL>      111.81614
 ----------      ---------    ---------    ---------    ---------    ---------
MINI>      960   1540.47209      0.20480      0.05428      0.00230
MINI INTERN>     3347.32229    700.65084    323.62086    894.91412     27.12932
MINI CROSS>      -298.71262
MINI EXTERN>      579.19510  -2506.38658      0.00000      0.00000      0.00000
MINI FCTPOL>    -1639.07474
MINI FCTNPL>      111.81350
 ----------      ---------    ---------    ---------    ---------    ---------
MINI>      980   1540.24611      0.22598      0.05246      0.00216
MINI INTERN>     3348.00969    700.80472    323.75255    895.21641     27.12899
MINI CROSS>      -298.66127
MINI EXTERN>      579.15024  -2508.31851      0.00000      0.00000      0.00000
MINI FCTPOL>    -1638.63246
MINI FCTNPL>      111.79576
 ----------      ---------    ---------    ---------    ---------    ---------
```

```
MINI>     1000   1539.75440      0.49171      0.04606      0.00202
MINI INTERN>     3346.73386    700.97916    323.64930    895.71239     27.13918
MINI CROSS>      -298.70025
MINI EXTERN>      580.81246  -2510.50371      0.00000      0.00000      0.00000
MINI FCTPOL>    -1637.83537
MINI FCTNPL>      111.76737
 ----------       ---------    ---------    ---------    ---------    ---------

 ABNER> Minimization exiting with number of steps limit ( 1000) exceeded.

ABNR MIN: Cycle      ENERgy      Delta-E         GRMS    Step-size
ABNR INTERN:          BONDs       ANGLes       UREY-b    DIHEdrals    IMPRopers
ABNR CROSS:           CMAPs
ABNR EXTERN:        VDWaals         ELEC       HBONds          ASP         USER
ABNR FCTPOL:         FCTPOL
ABNR FCTNPL:         FCTNPL
 ----------       ---------    ---------    ---------    ---------    ---------
ABNR>     1000   1539.75440      0.49171      0.04606      0.00303
ABNR INTERN>     3346.73386    700.97916    323.64930    895.71239     27.13918
ABNR CROSS>      -298.70025
ABNR EXTERN>      580.81246  -2510.50371      0.00000      0.00000      0.00000
ABNR FCTPOL>    -1637.83537
ABNR FCTNPL>      111.76737
 ----------       ---------    ---------    ---------    ---------    ---------
```

6.3 Appendix 3 - Plots of high energy sites

The following plots show a BRD7 decoy having a high potential energy after the simulation procedure. Figure 20 shows the whole decoy, while figures 21 to 23 concentrate on a region of that decoy causing high energy terms.
The important residues - showing the overlapping carbon atoms and the high energy bond - are emphasized. Residue 184 is a valine, residues 134 and 185 are tryptophanes.

Figure 20: A VMD scene[14] of the BRD7 decoy before the simulation procedure. The important residues are coloured: residue 184 in tan, residue 185 in blue and residue 134 in silver.

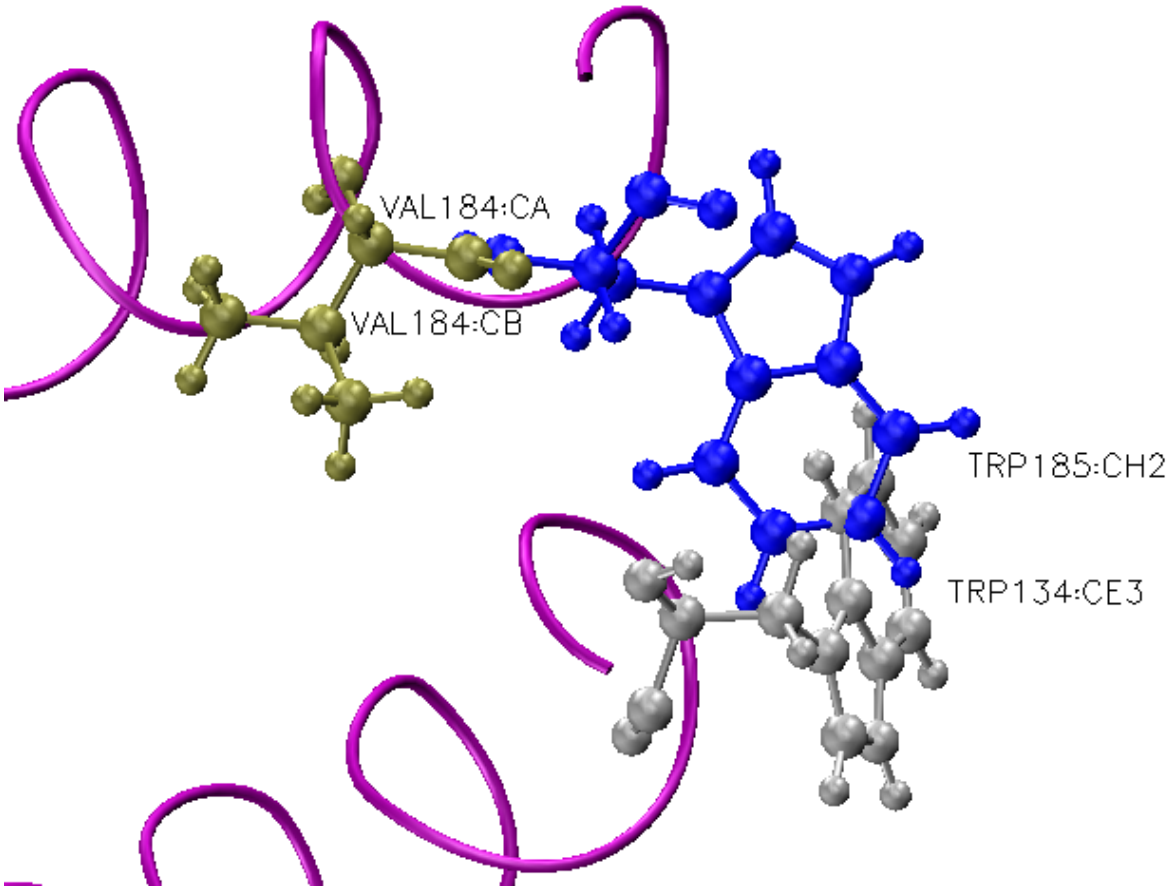

Figure 21: A cutout of the BRD7 decoy before the simulation procedure. Two carbon atoms of residues 134 and 185 overlap. The carbon bond in residue 184 still has the correct length. Residue 184 is coloured in tan, residue 185 in blue and residue 134 is shown in silver.

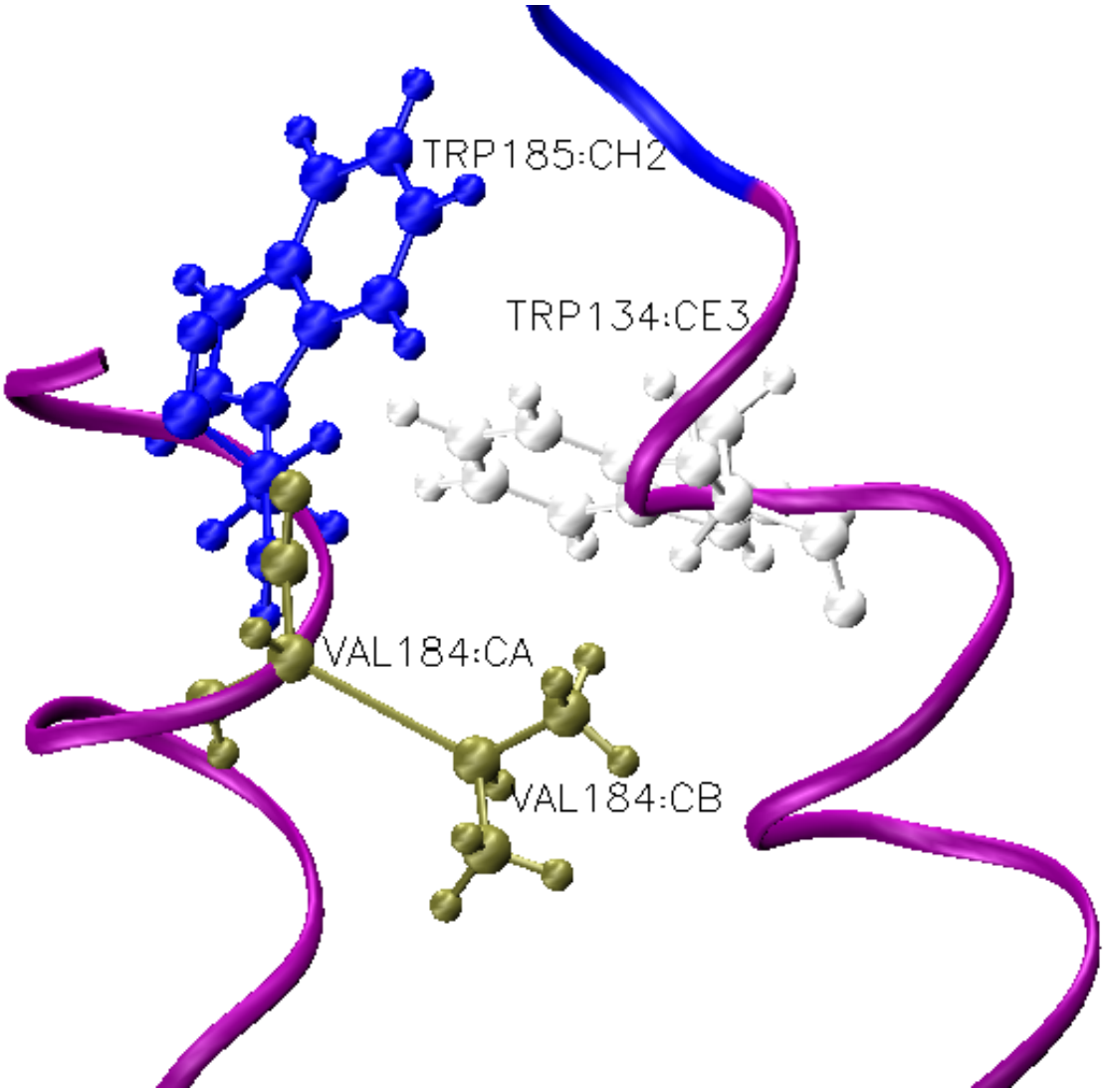

Figure 22: A cutout of the BRD7 decoy after the simulation procedure. Residue 184 shows the high energy bond - a too long carbon-carbon bond. Residues 134 and 185 do not overlap anymore. Residue 184 is coloured in tan, residue 185 in blue and residue 134 is shown in silver.

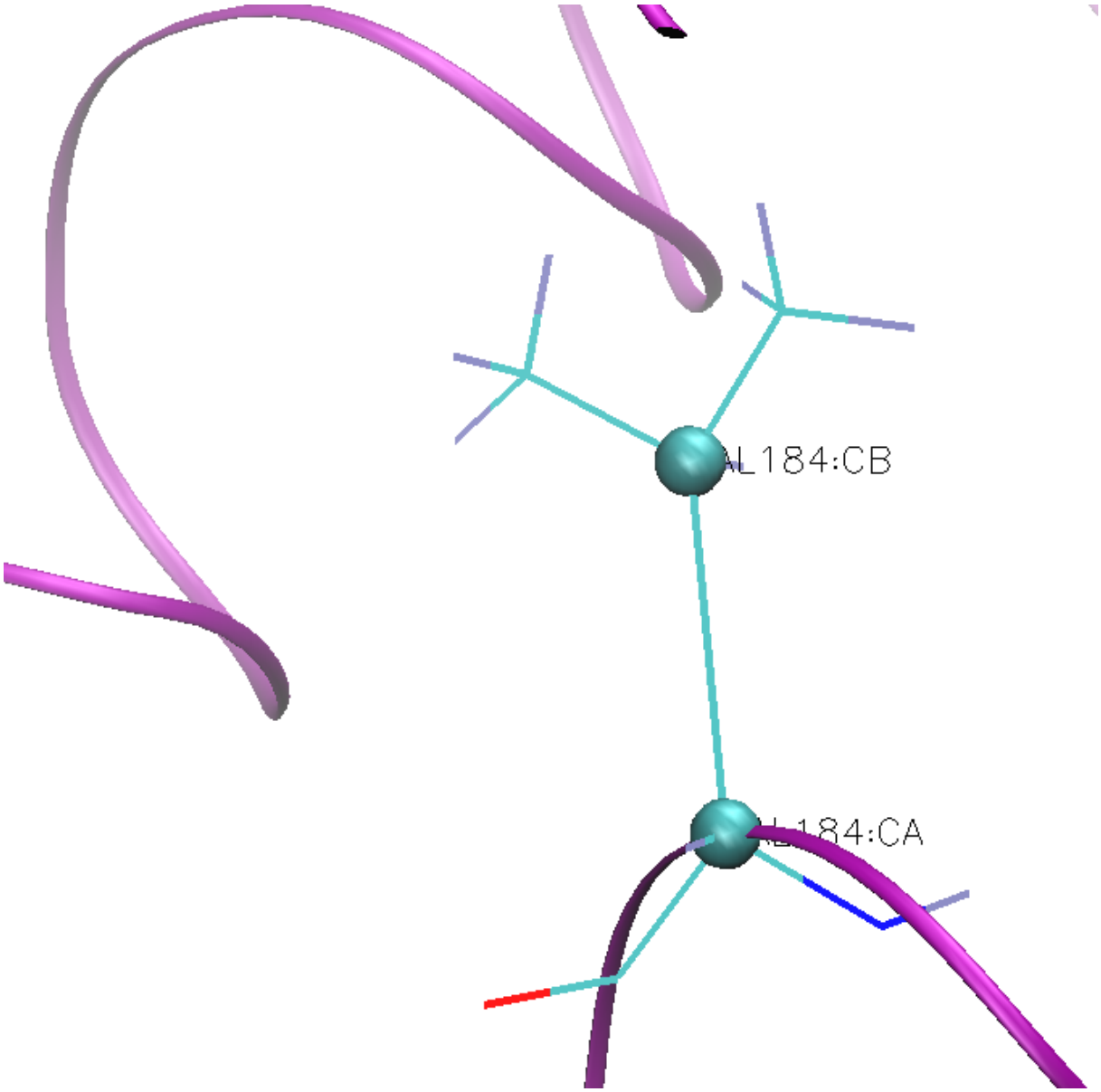

Figure 23: The BRD7 decoy after the simulation procedure showing the high energy bond in detail.

6.4 Appendix 4 - Energy development of the native state MD simulations

Figures 24, 25, 26 and 27 show the energy development during the MD simulations of the native states of FMR, LTPA, RHOD and VATP in FACTSMEM at a width of 28.5 Å.

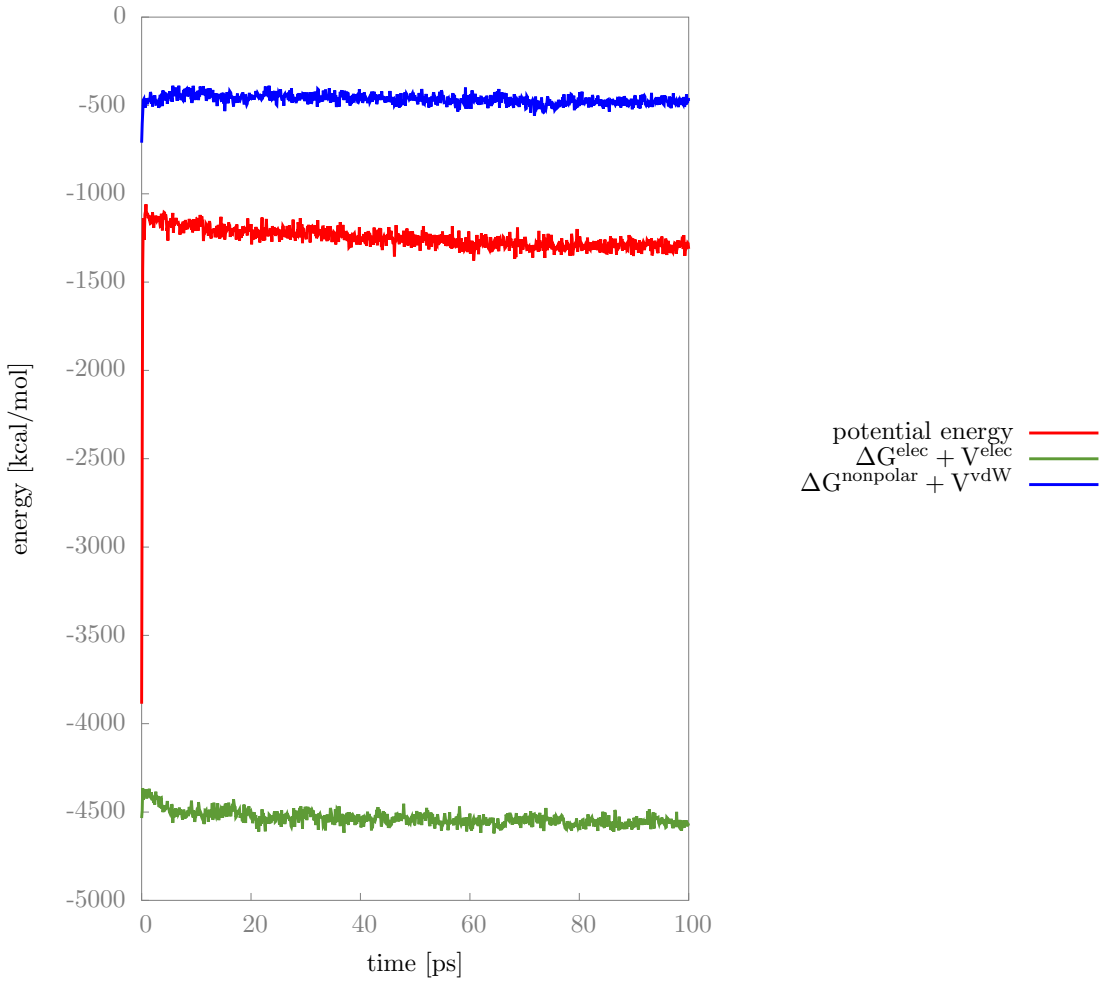

Figure 24: The energy development in kcal/mol of the native state of FMR is plotted against the time in fs. The potential energy is shown in red, the nonpolar contributions to the nonbonded interactions are shown in blue and the polar contributions to the nonbonded interactions are shown in green.

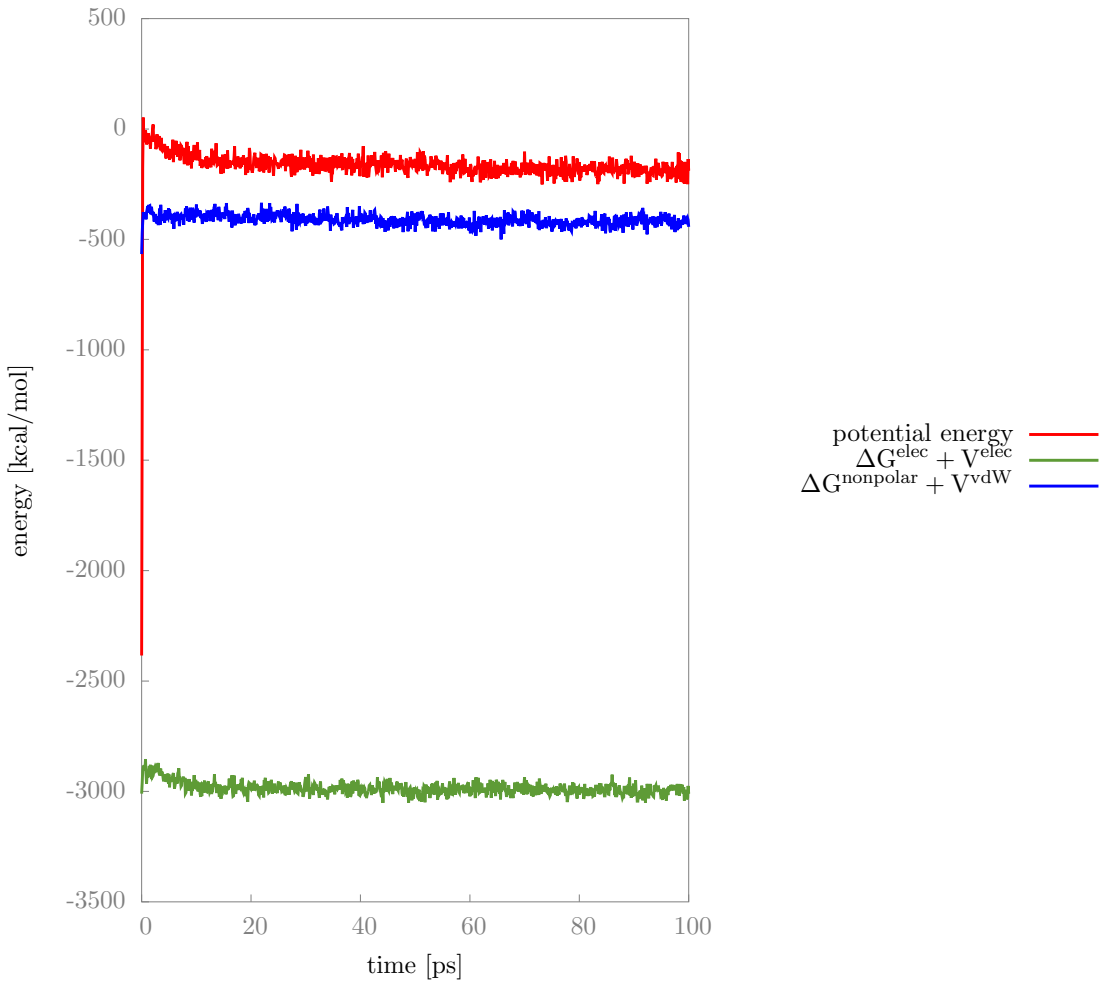

Figure 25: The energy development in kcal/mol of the native state of LTPA is plotted against the time in fs. The potential energy is shown in red, the nonpolar contributions to the nonbonded interactions are shown in blue and the polar contributions to the nonbonded interactions are shown in green.

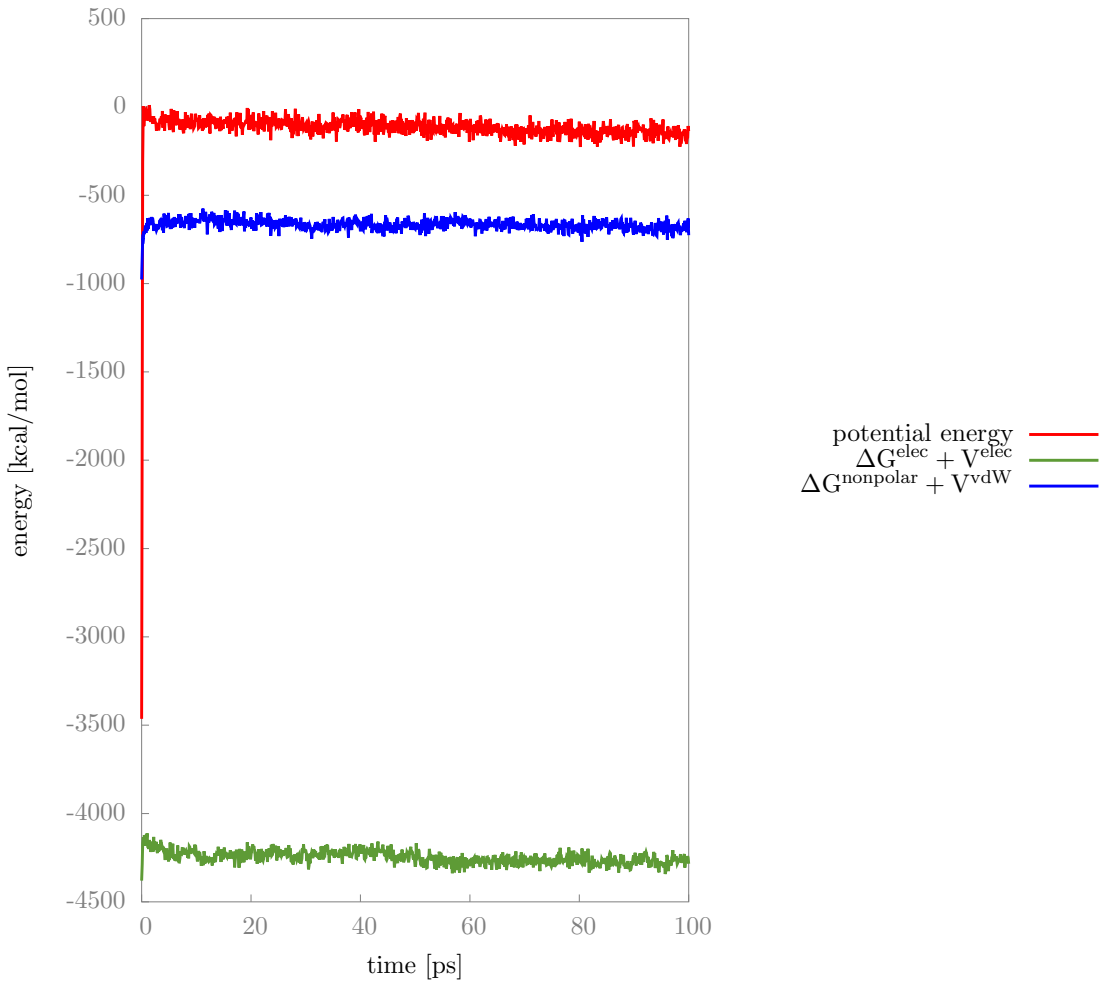

Figure 26: The energy development in kcal/mol of the native state of RHOD is plotted against the time in fs. The potential energy is shown in red, the nonpolar contributions to the nonbonded interactions are shown in blue and the polar contributions to the nonbonded interactions are shown in green.

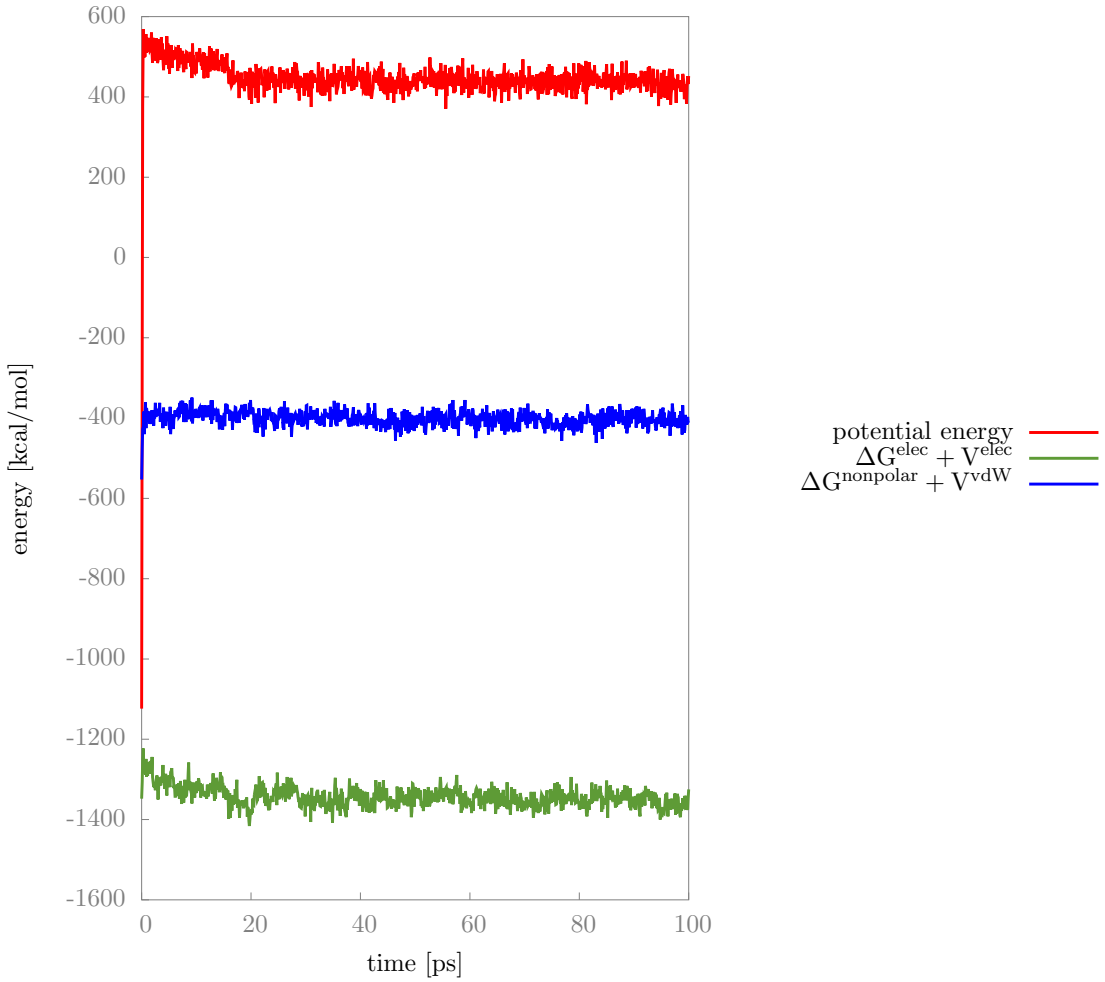

Figure 27: The energy development in kcal/mol of the native state of VATP is plotted against the time in fs. The potential energy is shown in red, the nonpolar contributions to the nonbonded interactions are shown in blue and the polar contributions to the nonbonded interactions are shown in green.

List of Tables

List of Figures

List of Abbreviations

RMSD	root mean square deviation
TM	transmembrane protein
YaL	Yuzlenko and Lazaridis' paper[11]
BRD7	bacteriorhodopsin
RHOD	rhodopsin
FMR	fumarate reductase
LTPA	lactose permease
VATP	V-ATPase
MD	molecular dynamics
SASA	solvent accessible surface area
PDB	Protein Data Bank[36]
ABNR	adopted basis Newton Raphson

References

[1] M. Luckey. *Membrane Structural Biology with Biochemical and Biophysical Foundations*. Cambridge University Press, New York, USA, 2008.

[2] A. L. Hopkins and C. R. Groom. The druggable genome. *Nature Reviews Drug Discovery*, pages 727–730, 2002.

[3] A. M. Seddon, P. Curnow, and P. J. Booth. Membrane proteins, lipids and detergents: not just a soap opera. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1666:105–117, 2004.

[4] M. Feig and C. L. Brooks III. Recent advances in the development and application of implicit solvent models in biomolecule simulations. *Current Opinion in Structural Biology*, 14:217–224, 2004.

[5] M. Carballo Pacheco. Extension of the FACTS implicit solvation model to membranes. Master's thesis, RWTH Aachen, Germany, 2013.

[6] M. Carballo Pacheco, I. Vancea, and B. Strodel. Extension of the FACTS implicit solvation model to membranes. submitted.

[7] T. Lazaridis. Effective energy function for proteins in lipid membranes. *Proteins: Structure, Function, and Bioinformatics*, 52:176–192, 2003.

[8] V. Z. Spassov, L. Yan, and S. Szalma. Introducing an implicit membrane in generalized born/solvent accessibility continuum solvent models. *The Journal of Physical Chemistry B*, 106:8726–8738, 2002.

[9] W. Im, M. S. Lee, and C. L. Brooks. Generalized born model with a simple smoothing function. *Journal of Computational Chemistry*, 24:1691–1702, 2003.

[10] S. Tanizaki and M. Feig. A generalized born formalism for heterogeneous dielectric environments: Application to the implicit modeling of biological membranes. *The Journal of Chemical Physics*, 122:–, 2005.

[11] O. Yuzlenko and T. Lazaridis. Membrane protein native state discrimination by implicit membrane models. *Journal of Computational Chemistry*, 34:731–738, 2013.

[12] Wikipedia. Cell membrane, 2014. http://en.wikipedia.org/wiki/File:Cell_membrane_detailed_diagram_en.svg.

[13] Wikipedia. Phosphatidylcholine, 2014. http://en.wikipedia.org/wiki/File:1-palmitoyl-2-oleoylphosphatidylcholine.svg.

[14] W. Humphrey, A. Dalke, and K. Schulten. Vmd: Visual molecular dynamics. *Journal of Molecular Graphics*, 14:33–38, 1996.

[15] Wikipedia. Beta barrel, 2014. http://en.wikipedia.org/wiki/File:Sucrose_porin_1a0s.png.

[16] D. Frenkel and B. Smit. *Understanding Molecular Simulation: From Algorithms to Applications*. Academic Press, San Diego, USA, 2002.

[17] D. C. Rapaport. *The Art of Molecular Dynamics Simulation*. Cambridge University Press, New York, USA, 2004.

[18] W. F. van Gunsteren and H. J. C. Berendsen. Computer simulation of molecular dynamics: Methodology, applications, and perspectives in chemistry. *Angewandte Chemie International Edition in English*, 29:992–1023, 1990.

[19] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. Charmm: A program for macromolecular energy, minimization, and dynamics calculations. *Journal of Computational Chemistry*, 4:187–217, 1983.

[20] B. R. Brooks et al. Charmm: The biomolecular simulation program. *Journal of Computational Chemistry*, 30:1545–1614, 2009.

[21] A. D. MacKerell, M. Feig, and C. L. Brooks. Improved treatment of the protein backbone in empirical force fields. *Journal of the American Chemical Society*, 126:698–699, 2004.

[22] P. J. Steinbach and B. R. Brooks. New spherical-cutoff methods for long-range forces in macromolecular simulation. *Journal of Computational Chemistry*, 15:667–683, 1994.

[23] U. Haberthr and A. Caflisch. Facts: Fast analytical continuum treatment of solvation. *Journal of Computational Chemistry*, 29:701–715, 2008.

[24] D. Bashford and D. A. Case. Generalized born models of macromolecular solvation effects. *Annual Review of Physical Chemistry*, 51:129–152, 2000.

[25] V. Z. Spassov, L. Yan, and S. Szalma. Introducing an implicit membrane in generalized born/solvent accessibility continuum solvent models. *The Journal of Physical Chemistry B*, 106:8726–8738, 2002.

[26] W. Im, M. Feig, and C. L. Brooks III. An implicit membrane generalized born theory for the study of structure, stability, and interactions of membrane proteins. *Biophysical Journal*, 85:2900–2918, 2003.

[27] S. Tanizaki and M. Feig. A generalized born formalism for heterogeneous dielectric environments: Application to the implicit modeling of biological membranes. *The Journal of Chemical Physics*, 122:124706, 2005.

[28] S. Tanizaki and M. Feig. Molecular dynamics simulations of large integral membrane proteins with an implicit membrane model. *The Journal of Physical Chemistry B*, 110:548–556, 2006.

[29] B. Lee and F. M. Richards. The interpretation of protein structures: Estimation of static accessibility. *Journal of Molecular Biology*, 55:379–IN4, 1971.

[30] H. G. Khorana, G. E. Gerber, W. C. Herlihy, C. P. Gray, R. J. Anderegg, K. Nihei, and K. Biemann. Amino acid sequence of bacteriorhodopsin. *Proceedings of the National Academy of Sciences*, 76:5046–5050, 1979.

[31] R. R. Franke, T. P. Sakmar, R. M. Graham, and H. G. Khorana. Structure and function in rhodopsin. studies of the interaction between the rhodopsin cytoplasmic domain and transducin. *Journal of Biological Chemistry*, 267:14767–74, 1992.

[32] T. H. Stevens and M. Forgac. Structure, function and regulation of the vacuolar (h+)-atpase. *Annual Review of Cell and Developmental Biology*, 13:779–808, 1997.

[33] J. Abramson, I. Smirnova, V. Kasho, G. Verner, R. H. Kaback, and S. Iwata. Structure and mechanism of the lactose permease of escherichia coli. *Science*, 301:610–615, 2003.

[34] J. J. Van Hellemond and A. G. M. Tielens. Expression and functional properties of fumarate reductase. *Biochemical Journal*, 304:321–331, 1994.

[35] V. Yarov-Yarovoy, J. Schonbrun, and D. Baker. Multipass membrane protein structure prediction using rosetta. *Proteins*, 62:1010–1025, 2006.

[36] F. C. Bernstein, T. F. Koetzle, G. J.B. Williams, E. F. Meyer Jr., M. D. Brice, J. R. Rodgers, O. Kennard, T. Shimanouchi, and M. Tasumi. The protein data bank: A computer-based archival file for macromolecular structures. *Journal of Molecular Biology*, 112:535–542, 1977.

[37] M. Sayadi, S. Tanizaki, and M. Feig. Effect of membrane thickness on conformational sampling of phospholamban from computer simulations. *Biophysical Journal*, 98:805–814, 2010.

[38] C. B. Anfinsen. Principles that govern the folding of protein chains. *Science*, 181:223–230, 1973.

[39] Ken A. Dill. Polymer principles and protein folding. *PRS*, 8:1166–1180, 1999.

[40] Welcome to ROSIE Rosetta Online Server that Includes Everyone. Rosie docking server documentation, 2014. http://rosie.rosettacommons.org/documentation/docking.

[41] L. Zhang and J. Skolnick. What should the z-score of native protein structures be? *Protein Science*, 7:1201–1207, 1998.