Molecular dynamics study of the interaction of Aβ(13-23) with β-sheet inhibitors

Belquis Mothana, Samir Roy, and Arvi Rauk*

Department of Chemistry, University of Calgary, Calgary, AB, Canada   T2N 1N4
E-mail: rauk@ucalgary.ca

This paper is dedicated to a good colleague and friend, Ted Sorensen, on the occasion of his 75th birthday

Abstract

The region encompassing residues 13-23 of the amyloid beta peptide (Aβ(13-23)) of Alzheimer's disease is the self-recognition site that initiates toxic oligomerization and fibrillization, and also is the site of interaction of Aβ with many other proteins. Peptidic compounds intended to act as β-sheet inhibitors targeted to Aβ(13-23) have been shown to inhibit fibrillization of Aβ and also to reduce its neurotoxicity. We describe herein a study by molecular dynamics (MD) of the complexes between Aβ(13-23) and three (pseudo)peptidic β-sheet inhibitors, as well as its homodimer. The monomers of all systems exist predominantly as extended β-strands, with Aβ(13-23) having the greatest flexibility to adopt other conformations. The dimers of all systems exist almost exclusively as stable antiparallel β-sheets anchored at the C-terminus of Aβ(13-23) by salt bridges to the C-terminal residues, Glu22 and Asp23. We also employ an MD technique called "atomic force microscopy" (AFM) to examine the dynamics of dissociation of the complexes in water. Each ligand attached to Aβ(13-23) begins dissociation by peeling back from its C-terminus, breaking interstrand H-bonds, and losing the β-sheet character. The salt bridges are the last to release, and presumably are the first to form in the reverse process of aggregation. The free energy profiles of the dissociation as a function of the separation of the centers-of-mass of all systems show plateau regions in which separation takes place with relatively little or no rise in free energy. For each system the dissociation profile does not have a maximum and reaches a flat plateau. By implication, the reverse process of assembly does not have a barrier. This and the plateau regions in the dissociation profile are examples of entropy-enthalpy compensation that arise naturally during the AFM-MD simulation.

Keywords: Amyloid beta, beta sheet inhibitors, molecular dynamics, atomic force microscopy, Gromacs, peptide-peptide interactions, beta sheets, hydrogen bond strength, Alzheimer's disease

Introduction

Alzheimer's disease (AD) is the most common cause of dementia among people age 65 or older. Although the etiology of AD is not completely understood, the amyloid beta peptide (Aβ), the major component of senile plaques found in AD brains, is implicated in the pathogenesis.[1,2,3] The Aβ peptide is a byproduct of proteolysis of the amyloid-β precursor protein (APP)[4] and is mostly comprised of 40 or 42 amino acid residues. Aβ aggregates to form fibrils with an ordered β-sheet pattern[5 6,7,8] in AD brains. The actual form of Aβ that causes the damage is still uncertain, but is most likely a small dimeric[9] or higher oligomeric species[10] with internal β-sheet structure, and most likely binding Cu^{2+}. Significant efforts of research are directed to the development of compounds capable of inhibiting and reversing the aggregation process. It has been well demonstrated that compounds intended to block β-sheet formation[11,12] inhibit the formation of both amyloid fibrils and smaller oligomers, and block Aβ toxicity toward neurons in *in vitro*[13,14,15] and in animal models.[16,17]

Aβ binds to itself at residues 12-23.[18,19] This region is also the binding site for cholesterol,[20] apolipoproteinE (apoE),[21] the α7 nicotinic acetylcholine receptor (α7nAChR),[22,23] and amyloid beta-peptide binding alcohol dehydrogenase (ABAD).[24] A model of the Aβ peptide consisting of amino acids 13-23 (Aβ(13-23)), was shown to initiate oligomerization. This region has been the basis for the synthesis of numerous small peptides as β–sheet blockers.[11,12,13,14,15] We also have used Aβ(13-23) as a model Aβ receptor, R = *N*-AcH13H14Q15K16L17V18F19F20A21E22D23NHMe, for this purpose. Several small compounds have been designed to bind specifically to Aβ(13–23) using computer modeling.[25] A design constraint is that the compounds be pseudopeptidic in order to resist hydrolysis and to make them easy to synthesize. Two of the compounds, designated SG1 and SG2, have been synthesized and assays were carried out to test their effectiveness at preventing fibril formation.[26] SG1 is acetyl-dab1-Orn2-meL3-F4-meF5-L6-P7-bA8, where dab = ε-diaminobutyric acid, Orn = ornithine, meL = *N*-methylleucine, meF = *N*-methylphenylalanine, and bA = beta alanine. SG2 is acetyl-dab1-Orn2-meL3-V4-meF5-F6-A7-E8-NH₂. Both compounds are predicted to bind strongly to Aβ, and not to self-oligomerize. The initial assay results show that both SG1 and SG2 are capable of stopping high-order fibrils from forming.[26]

SG1 and SG2 are intended to prevent the aggregation and fibrillization of Aβ(1-40) and Aβ(1-42) by binding more strongly to Aβ than Aβ to itself and their ability to do this has thus far been assessed by an empirical scoring algorithm, BHB (Binding/Hydrogen bonding/Buriedness).[27] BHB incorporates the binding energy of the adduct as assessed by the MMFF94x forcefield in MOE,[25] as well as additional terms that account for hydrophobic interactions and salt-bridge H-bonding. Scores assessed by BHB were normalized to yield a score of 100 for a complex between R and the pentapeptide KLVFF (=Aβ(16-20) = *N*-AcK16L17V18F19F20NH₂). For reference, BHB scores of the R-R homodimer, and the R-SG1 and R-SG2 heterodimers, are 327, 381, and 331, respectively. Thus, to the extent that R is

representative of full length Aβ, both designed pseudopeptides are predicted to be competitive inhibitors of Aβ aggregation.

In part to improve the predictive capability of BHB, we have adapted a method based on Molecular Dynamics (MD) to determine the binding affinities of the candidate drug/Aβ complexes as well as the Aβ/Aβ interactions, using R (=Aβ(13-23)) as a model of the full length Aβ. The method is a "potential of mean force" (PMF) method conceptually analogous to atomic force microscopy (AFM), and we shall refer to it as AFM-MD. Implementation consists of two steps. First a homo- or heterodimer structure is equilibrated in a box of water. Then the two species are pulled apart in such a way that the free energy change can be monitored as a function of separation. At a sufficiently large separation the two species are no longer in interaction and the cumulated free energy change is the free energy of dissociation, ΔG_{dis} in water. Determination of ΔG_{dis} for the R-R dimer, and R complexed to SG1, SG2, and KLVFF in this manner should provide a more realistic value for the binding affinity term of the BHB scoring algorithm. Feedback from the calculations, combined with experimental assaying, would be used to improve the design of the drug-candidates and further our understanding of their binding to the Aβ peptide. An additional benefit of the approach is that it provides a detailed picture of the relative strengths of the various amide H-hydrogen bonds and salt bridges in an explicit aqueous environment. The methodology employed, the structures of the peptidic complexes, and the relative strengths of the interactions are the subject of this report.

Methods

Structural models. Initial structures for the molecular dynamics simulations were obtained from docking results using the MMFF94x forcefield in MOE software.[25] The Aβ receptor, R, i.e. Aβ(13-23), is acetylated at the N-terminus, residue 13 (His13) and N-methylated at residue 23 (Asp23). The two histidine residues, His13 and His14, were not charged. MD simulations are performed using the Gromacs 3.1.4[28] suite of software.

Equilibration studies. The individual MOE-generated structures, R, SG1, SG2 and KLVFF, as well as the homodimer, R-R, and the heterodimers, R-SG1, R-SG2, and R-KLVFF were equilibrated with Gromacs.[28] Each system was placed in the center of a cubic box whose dimensions were taken as 18 Å greater than the diameter of the solute (largest distance between atoms). The box was filled with SPC water molecules. Sodium or chloride ions were added to ensure electroneutrality in the system. The GROMOS96 53a5 force field was used for the MD simulations. The neighbor searching was performed using a twin-range approach and 1.4 nm for van der Waals-cutoff and 1.0 nm for electrostatic interactions. Particle mesh Ewald (PME) summation was applied for longer-range electrostatic interactions with a grid spacing of 0.12 nm.[29] The neighbor list update was performed every 5 steps. The time step was 2 fs, using LINCS to constrain all bond lengths.[30] We used NPT conditions (i.e., constant number of particles, pressure, and temperature) in the simulation, and periodic boundary conditions. A

constant pressure of 1 bar was applied with a coupling constant of 1 ps; system and water/ions were coupled separately to a temperature bath at 300 K with a coupling constant of 0.1 ps.

The following steps have been taken to perform the simulations: initially, an energy minimization of the system in solvent is carried out. Secondly, a short 100 ps MD run with position restraints on the system is performed. Finally a full MD without restraints is carried out. The trajectories were analyzed grouping structurally similar frames by cluster analysis using an RMSD cutoff of 3.0 Å.[31] The frame with the largest number of neighbors is denoted the "middle" structure and is the structure representing that cluster. The VMD suite of software[32] was extensively used for visualization, and secondary structure analysis was determined by the STRIDE method.[33] The "middle" equilibrated complexes (R-R, R-SG1, R-SG2, and R-KLVFF) from the cluster analysis are taken as the initial structures for carrying out the potential of mean force calculations to measure the binding affinity as described in the next section.

Figures displaying the cluster analysis of each system, and the STRIDE analysis of the dimeric systems are provided in Supporting Information and are numbered with "S", e.g., Figure S1, etc.

Atomic Force Microscopy (AFM-MD) calculations. Before carrying out the AFM calculations, a sequence of minimization and equilibration of the 'middle' equilibrated complexes had to be performed since a new water box was used. A cubic box of 9 nm dimensions was used for the AFM-MD calculations. The box was filled with approximately 25,000 SPC water molecules. The procedure adopted for the minimization sequence is similar to the protocol used by Mobley *et al.*[34] First a 5000 step L-BFGS (Broyden-Fletcher-Goldfarb-Shanno approach) minimization[35,36] is performed followed by a 500 step steepest descent minimization. This is followed by a short 10 ps constant volume equilibration and then a 100 ps constant pressure equilibration using the md integrator. A constant pressure of 1 bar was applied with a coupling constant of 0.5 ps; system and water were coupled separately to a temperature bath at 300 K with a coupling constant of 0.1 ps. Finally, a constant-volume production run of 50 ns is carried out to measure the binding affinity of R-drug and R-R complexes. The neighbor searching was performed using a twin-range approach and 0.9 nm for van der Waals-cutoff and 1.0 nm for electrostatic interactions. Particle mesh Ewald (PME) summation was applied for longer-range electrostatic interactions with a grid spacing of 0.1 nm.[29] The neighbor list update was performed every 10 steps. The time step was 2 fs, using LINCS to constrain all bond lengths.[30] The PMF is calculated using the pull code in Gromacs software with the AFM pulling option. In AFM-pulling, a spring (S) with known force constant (**k**) is attached to the center-of-mass of the ligand (L). The position of the center-of-mass of the Aβ receptor (R) is fixed and the ligand is pulled away slowly using the spring at constant rate. The force on the spring can be determined using Hooke's law, where $\Delta F_{LS} = \mathbf{k}\Delta r_{LS} = \Delta F_{RL}$. By integrating ΔF_{RL} as a function of distance, we can determine the free energy change, $\Delta G(r_{RL})$:

$$\Delta G(r_{\mathrm{RL}}) = \sum_{0 \to r_{\mathrm{RL}}} \Delta F_i \Delta r_i .$$

The cumulated free energy change represents the dissociation energy, ΔG_{dis}, in kJ/mol for the ligand-receptor complex. For our systems, the pulling dimension was set to all direction (x,y,z). The AFM rate of the spring used is 0.00006 nm/ps. This was chosen based on trial calculations which determined that a 3 nm distance between the ligand and the receptor is needed to completely separate them. In addition, equilibration results concluded that 50 ns is an acceptable time to have a stable conformation. Therefore, the rate constant was determined based on the 50ns simulation run to separate the ligand and the receptor by 3 nm. The force constant used for the calculation is 10,000 kJ/(mol nm^2). To start the simulation with zero initial force on the pulled group (ligand), the initial position of the spring relative to the center-of-mass of the receptor is set to the position of the center-of-mass of the ligand relative to the center-of-mass of the receptor. In order to get pulling along the direction of the centers-of-mass of the receptor and ligand, the unit vector from the initial position of the spring is used for the AFM direction of pulling. The PMF graph is plotted using the Grace software.[37]

Results and Discussion

Equilibration of structures
R, SG1, SG2, and KLVFF monomers. R: An 80 ns MD simulation of the receptor R = Aβ(13-23) shows that the monomeric receptor is quite flexible (see Figure S1). Cluster analysis performed with a RSMD cutoff of 3.0 Å found 20 distinct conformations over the 80 ns trajectory, with two dominant structures, accounting for 80% of the total population, in a ratio of 6:1 . The secondary structure profile (not shown) shows that the middle structure of the dominant cluster adopts an extended β-strand conformation, shown in Figure 1a.
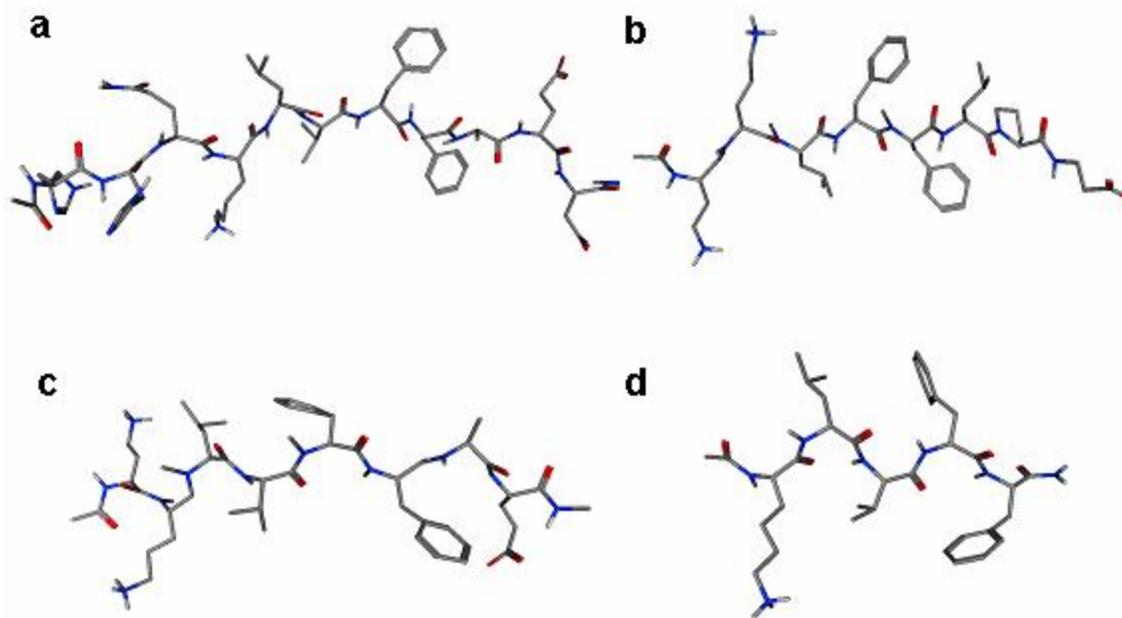
Figure 1. The central structures of the most populated cluster of the monomers after equilibration: **a** R = AcHHQKLVFFAEDNHMe; **b** SG1 = Ac-dab-Orn-meL-V-meF-F-P-bA; **c** SG2 = Ac-dab-Orn-meL-V-meF-F-A-E-NH$_2$; **d** AcKLVFFNH$_2$.

Ramachandran plots (not shown) of this structure, which we adopt as R for the purposes of the AFM studies, show that all residues fall in the β-sheet region with the exception of the C-terminal Asp23 residue (we will use the Aβ numbering for the corresponding residues of R). The second most populated cluster evolves around 35 ns and persists till around 50 ns. This structure folds on itself at His14 and at Leu17, allowing the side chain of Lys16 and Phe20 to interact. The rest of the structure from Val18-Asp23 adopts an extended β-strand conformation. The cluster analysis profile (Figure S1) shows that the receptor model, Aβ(13-23) equilibrates within a few ns, with many transitions between conformers over 80 ns. The evident flexibility of Aβ(13-23) is in sharp contrast to that found by the same RMSD criterion for Aβ(1-42), which equilibrated to a single conformation that persisted for the last 1000 ns of a 1300 ns trajectory.[38]

SG1. An 80 ns MD simulation of SG1 shows that its structure is quite stable and rigid due to the presence of the two amide N-methyl groups. The cluster analysis performed with a RSMD cutoff of 3.0 Å found only 2 clusters over the 80 ns trajectory (Figure S2), with one accounting for 99.9% of the population. The middle structure (Figure 1b) of the dominant cluster of SG1 adopts an extended β-strand conformation. Ramachandran plots (not shown) of the middle structure of SG1 also show that all residues fall in the β-sheet region.

SG2. A 60 ns MD simulation of SG2 shows that its structure is also quite stable and rigid. The cluster analysis performed with a RSMD cutoff of 3.0 Å found only 3 clusters over the 60 ns trajectory (Figure S3), with the dominant one accounting for 99.7% of the population. The middle structure (Figure 1c) of SG2 adopts an extended β-strand conformation and the Ramachandran plots (not shown) also show that all residues fall in the β-sheet region.

KLVFF. A 30 ns MD simulation of KLVFF shows that its structure is stable and forms an extended β-strand conformation. Only one cluster was found with a RSMD cutoff of 3.0 Å (Figure S4).   All residues of the middle structure (Figure 1d) fall in the β-sheet region in the Ramachandran plots (not shown).

R-R, R-SG1, R-SG2 and R-KLVFF complexes

R-R. The initial structure for the MD simulation of the homodimer of R (R-R) was taken from the MOE-docked structure, an antiparallel β-sheet. A 30 ns simulation of R-R surrounded by explicit water molecules was carried out. The cluster analysis performed with a RSMD cutoff of 3.0 Å found only 4 clusters of similar structure over the 30 ns trajectory (Figure S5). Secondary structure analysis (Figure S6) shows that R-R adopts an extended antiparallel β-sheet conformation throughout the time of simulation. The N-terminal residues, His13 and His14, are not involved in the β-sheet but adopt a β-strand conformation. Ramachandran plots (not shown) of the middle structure of the most populated cluster of R-R (Figure 2a) show that all residues fall in the β-sheet region.
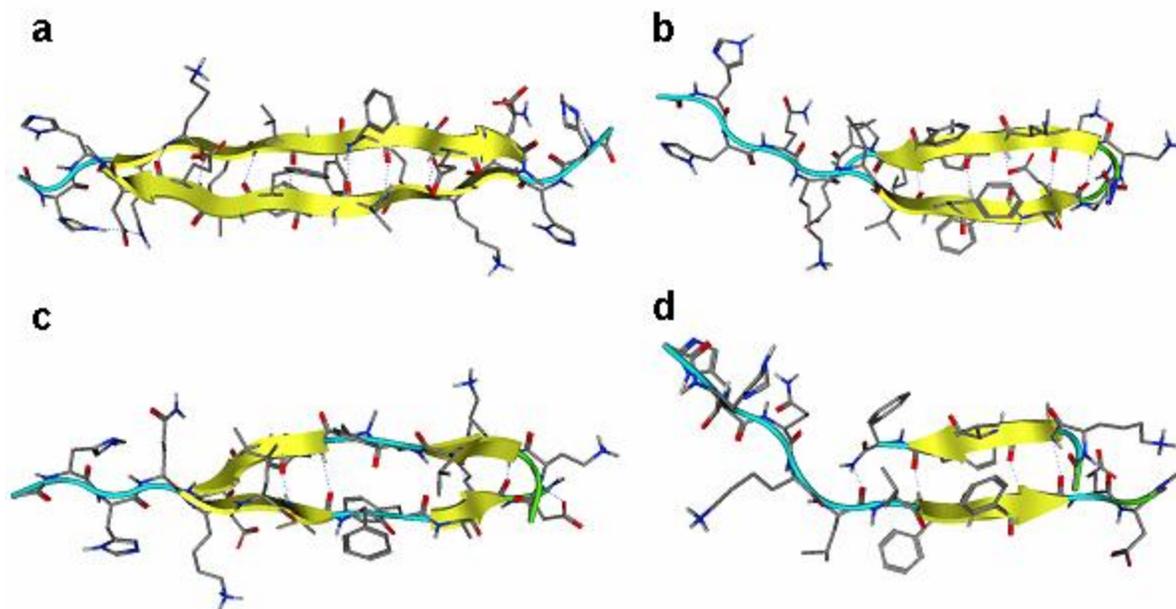
Figure 2. The central structures of the most populated cluster of the dimeric complexes after equilibration: **a** R-R; **b** R-SG1; **c** R-SG2; **d** R-KLVFF. In each case, R is shown as the lower strand.

An analysis of the number of interstrand hydrogen bonds in the β-sheet of R-R indicates that the average number of hydrogen bonds is 9.7 out of the 14 (10 inter-strand hydrogen bonds between the backbones of R-R and the other 4 are from charged side chain interactions) that formed during the 30 ns simulation. There are salt bridges formed between charged side chains of both Lys16-Glu22 pairs. There is also hydrogen bonding between the side chains of Gln15-Asp23 pairs. The anti-parallel bound dimer of R-R is not as flexible as the R monomer.

R-SG1. The initial structure for the MD simulation of the heterodimer, R-SG1, was taken from the MOE-docked structure, an antiparallel β-sheet with salt bridges at both ends. A 70 ns MD simulation of R-SG1 shows that its structure is also quite stable and rigid. The cluster analysis performed with a RSMD cutoff of 3.0 Å found 9 clusters over the 70 ns trajectory (Figure S7). The middle structure of R-SG1 adopts an extended β-sheet conformation (Figure 2b). The timeline for the secondary structure of the complex R-SG1 trajectory (Figure S8) also shows that the β-sheet persists during the entire simulation time. Moreover, Ramachandran plots (not shown) of the middle structure of R-SG1 also show that all residues fall in the β-sheet region except for Asp23 and dab1, which fall in the α-helix region. The average number of interstrand hydrogen bonds formed between R and SG1 during the 70 ns simulation time is 7.3. Salt bridges also form between charged side chains of Lys16-bA8, Glu22-Orn2, and Asp23-dab1.

R-SG2: Similarly, a 70 ns MD simulation of complex R-SG2 shows that its structure is stable and rigid. The cluster analysis found 6 clusters over the 70 ns trajectory (Figure S9). The middle structure of R-SG1 adopts an extended antiparallel β-sheet conformation (Figure 2c). Moreover, Ramachandran plots (not shown) of the middle structure of R-SG2 also show that all residues fall in the β-sheet region except for dab1, which falls in the α-helix region. The average number of hydrogen bonds formed between the receptor and SG2 during 70 ns simulation time is 7.8. Salt bridges also form between charged groups Lys16-Glu3, Glu22-Orn2, and Asp23-dab1.

R-KLVFF. A 30 ns MD simulation of complex R-KLVFF shows that its structure is quite flexible since KLVFF is a shorter segment and only binds to a small segment of R. The rest of the R group then is more mobile. The cluster analysis performed with a RSMD cutoff of 3.0 Å found 11 clusters over the 30 ns trajectory (Figure S11). The middle structure of R-KLVFF (Figure 2d) adopts an extended β-sheet conformation for the interactions of the hydrophobic residues, residues F19, F20, and A21 from R and F19, V18, and L17 from KLVFF. The rest of the residues form turn and coil conformations. The timeline for the secondary structure analysis of the complex R-KLVFF trajectory (Figure S12) also shows that the β-sheet persists during the simulation time. Moreover, Ramachandran plots (not shown) of the middle structure of R-KLVFF show that all residues fall in the β-sheet region with the exception of Lys16 of both R and KLVFF, and His14 falling in the α-helix region, and Glu22 in the left-handed helix region. The average number of hydrogen bonds formed between the receptor and KLVFF during simulation time is 4.9. One Salt bridge forms between charged groups Glu22-Lys16.

AFM-MD Results and Discussion: R-R, R-SG1, R-SG2 and R-KLVFF
All the calculations of the binding affinities using AFM-MD method were run for 50 ns with a pull rate of 0.00006 nm/ps. At this rate, all internal parameters except the pull parameter remain equilibrated over the 50 ns simulation time. As indicated above, all monomeric species and complexes with R equilibrated in much less than 50 ns. Equilibration must be maintained during the AFM-MD simulation to ensure the energy change is an acceptable approximation to the desired free energy change. Some contributions to the change in entropy, namely translational and rotational entropies of the separated species, will be underestimated. As a consequence, the calculated free energy change will be an upper bound to the actual value for each system, but the relative values should be well reproduced.

Looking at the results of the AFM-MD graphs (Figures 5 and 6), we can see that the systems produce reasonably smooth curves, which is an indication that the ensemble is properly sampled during the course of the simulation. Moreover, once the complex has separated, a flat curve is produced with zero free energy change, which also confirms that the system is equilibrated and that there is only a negligible viscosity component to the free energy change. Cluster analyses on the part of the curves after separation shows that the separated R structure resembles either the dominant equilibrated structure, or the second-most populated one. The separated SG1, SG2

and KLVFF species have structures corresponding to the equilibrated middle structures discussed above.

The dissociation of the complexes occurs as the center-of-mass of R and the center-of-mass of the ligand (R, SG1, SG2 or KLVFF) move apart. The centers-of-mass of the receptor and ligand were chosen as the "pull points" in order to provide as unbiased way as possible to investigate the mechanism of dissociation of the complexes.
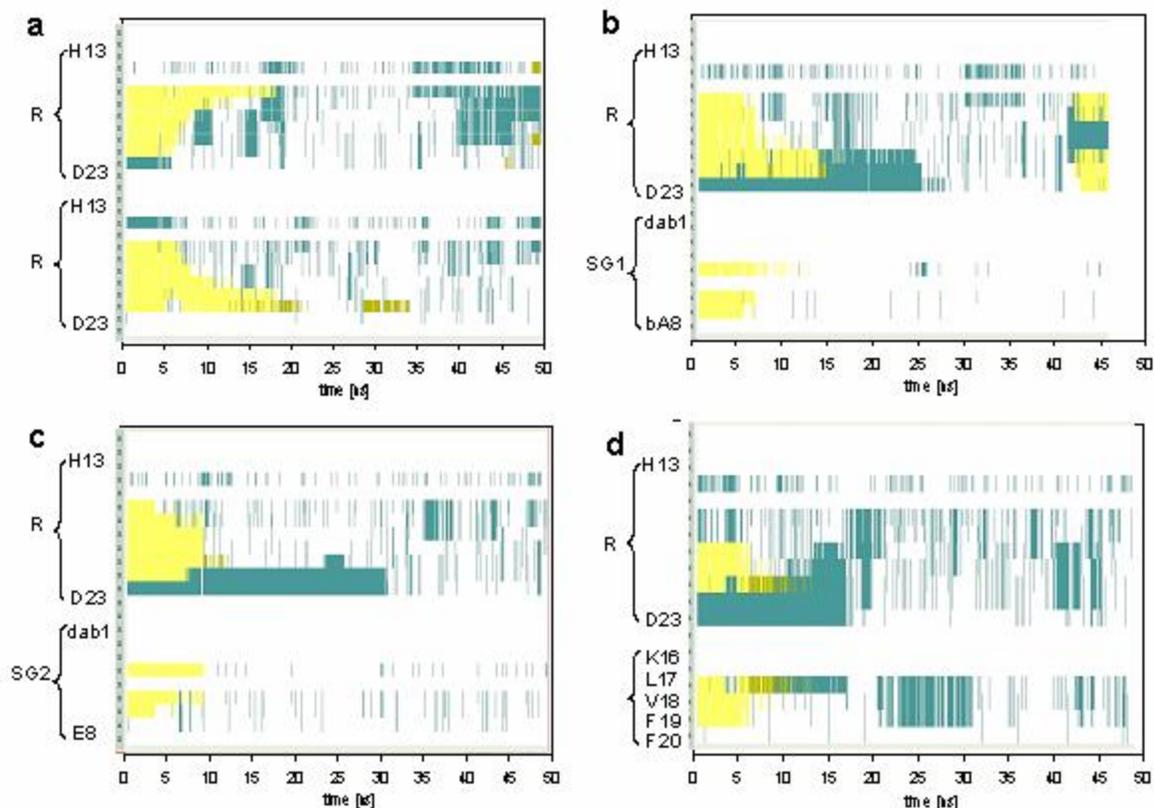
Figure 3. STRIDE analysis of the dimer structures during the 50 ns of the AFM trajectory during which the separation of the centers of mass was increased by 30 Å at a constant rate of 0.6 Å/ns. Yellow bars indicate residues involved in β-sheet secondary structure, green indicates turns, and white denotes random coil. **a** R-R; **b** R-SG1; **c** R-SG2; **d** R-KLVFF.

R-R

The dissociations of the complexes, R-R, R-SG1, R-SG2 and R-KLVFF follow a similar course. The R-R complex has pseudo-C_2 symmetry. The β-sheet structure starts unzipping by breaking interstrand hydrogen bond interactions from the C-terminus of one strand and its point of attachment near the N-terminus of the other. Figure 3a shows the secondary structure of R-R as it changed over the course of the 50 ns simulation. The β-sheet character was lost when the separation of the centers-of-mass reached around 2.14 nm (corresponding to 23 ns time) from its

initial value of 0.72 nm. The initial average of 9.7 interstrand hydrogen bonds (Figure 4a) is reduced to 1 H-bond at around 2.25 nm separation (at about 25 ns) at a cost of about 105 kJ/mol in free energy (Figure 5). At this point, there is a shoulder region of the free energy curve that extends to 2.50 nm (about 29 ns). At the end of the shoulder region, the unraveling strand flips end to end and as a result forms an average of 3 H-bonds with R receptor, which can be seen in Figure 4a at around 30 ns.
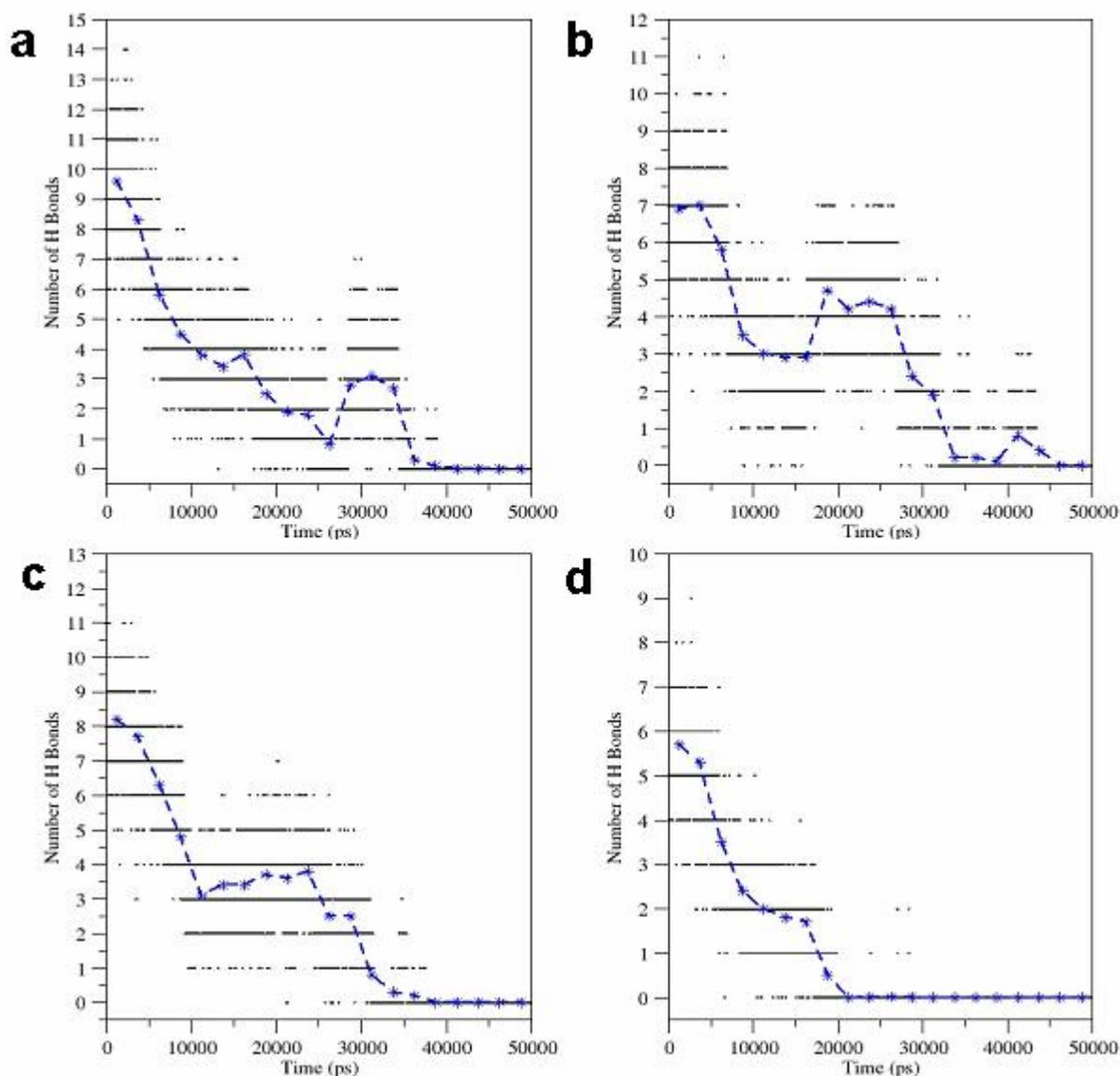
Figure 4. Count of interstrand H-bonds at 10 ps intervals for the dimer structures during the 50 ns of the AFM trajectory during which the separation of the centers of mass was increased by 3 nm at a constant rate of 0.06 nm/ns. Connected points are averages over successive 2.5 ns windows. **a** R-R; **b** R-SG1; **c** R-SG2; **d** R-KLVFF.

The last interstrand hydrogen bonds and salt bridges to break are those at the N-terminus of one strand (residues H13, H14, Q15 and K16) and C-terminus of the other (residues D23, E22 and A21). These residues maintain contact till the center-of-mass separation is large enough to separate them completely, at around 2.9 nm (corresponding to 37 ns). The last bonds required a free energy of 25 kJ/mol free energy to completely separate. The complex is completely separated at about 2.9 nm (37 ns).
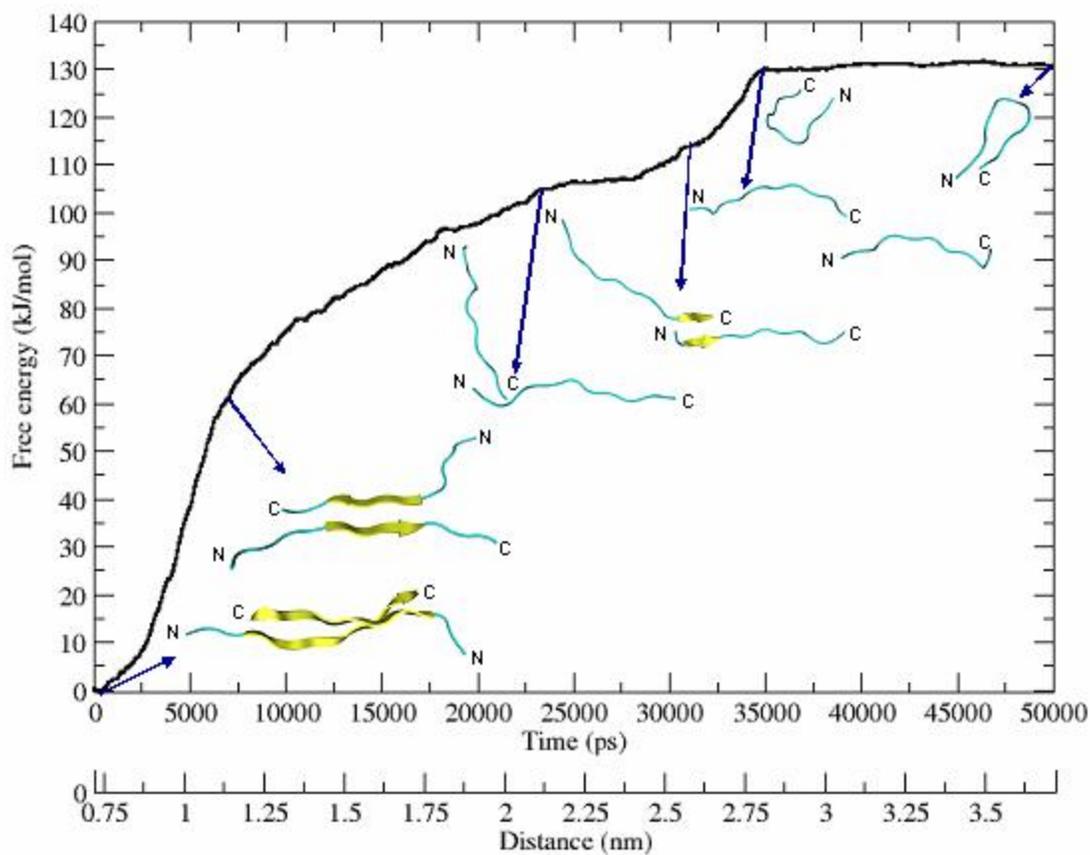
Figure 5. Free energy profile for the separation of R-R over the 50 ns of the AFM-MD trajectory during which the separation of the centers of mass was increased from its initial value of 0.72 nm by 3 nm at a constant rate of 0.06 nm /ns. The insets show snapshots of the structure at the designated points along the trajectory.

It is of note that the free energy profile reaches a plateau at 35 ns while there is still an average of 3 interstrand H-bonds, and that it remains constant while these are lost in the next 2 ns. The total free energy required for complete dissociation of R-R is ΔG_{dis} = 131 kJ/mol. The last H-bonds to break are the salt bridges between C-terminal Glu22 and Asp23 of one strand, and Lys16 near the N-terminus of the other strand.

A maximum is not observed in the dissociation profile defined by the distance between the centers-of-mass. This suggests that there should be no barrier to the aggregation of the two

strands. Evidently the loss in entropy (the $-T\Delta S$ term) is almost exactly compensated by the gain in enthalpy from the formation of the initial salt bridges. The flat portions seen in the dissociation profiles of R-R and the other complexes (see Figure 6 and below) are additional instances of entropy-enthalpy compensation.[39] The phenomenon of entropy-enthalpy compensation in connection with the association/dissociation of biomolecules is well documented in the literature although there is uncertainty as to its origin.[40]
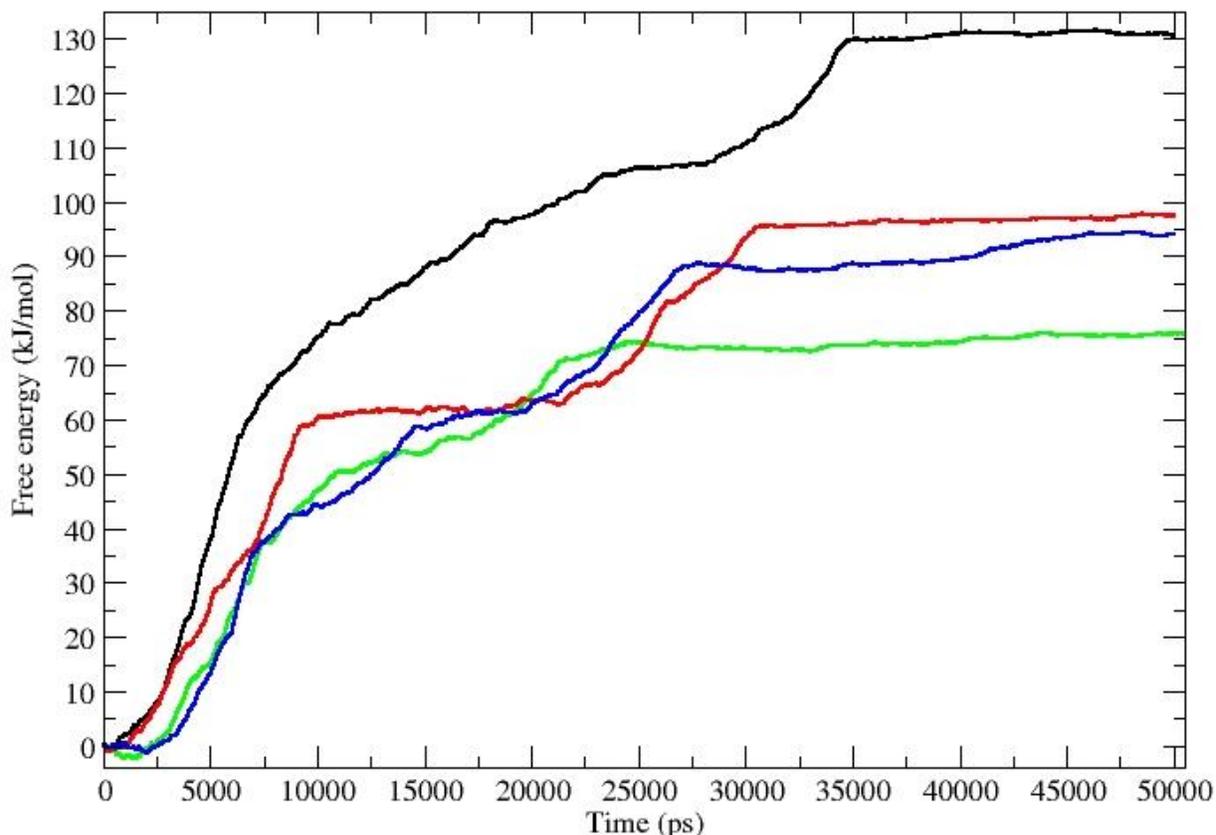
Figure 6. Free energy profiles for the separation of R-R (black line), R-SG1 (blue line), R-SG2 (red line), and R-KLVFF (green line) over the 50 ns of the AFM-MD trajectory during which the separation of the centers of mass was increased by 3 nm at a constant rate of 0.06 nm /ns. The initial separations of the centers of mass are (in nm): R-R, 0.72; R-SG1, 0.49; R-SG2, 0.52; R-KLVFF, 0.61.

R-SG1
In the dissociation of R-SG1, the C-terminus of SG1 starts separating from its point of attachment near the N-terminus of R. Secondary structure analysis over the span of the 50 ns simulation of R-SG1 (Figure 3b) shows the β-sheet character is completely lost at a separation of 1.5 nm (about 17 ns). The free energy curve of R-SG1 shows a small flat region around 10 ns and 45 kJ/mol by which time 4 of the initial 7 H-bonds had broken (Figure 4b). There is a second plateau at 17 ns and free energy of 60 kJ/mol, while interstand H-bonds rise to 4-5.

During the second flat region, the SG1 strand flips end for end while maintaining hydrogen bonds with the C-terminus of R (compare Figures 4b and 6). The last interstrand hydrogen bonds and salt bridges to come apart are residues at the C-terminus of R (E22 and D23) and N-terminus of SG1 (dab1 and Orn2). These residues stick together for quite longer time than R-R, or R-SG2, and R-KLVFF (see below), and come apart completely at around 3.23 nm (45 ns) with a free energy change of about 34 kJ/mol. The total free energy of the dissociation of the complex R-SG1 is $\Delta G_{dis} = 94$ kJ/mol (Figure 6).

Much more so than in the case of R-R, the free energy profile levels out long before all H-bonds are lost. At 2.23 nm separation (27 ns), while there are an average of 4 H-bonds connecting the strands, the energy profile becomes independent of separation. The H-bonds are mostly lost in the next 8 ns but electrostatic contact through transient H-bonds is maintained between dab1 and E22, D23 for another 10 ns. The rise of the free energy by only 4 kJ/mol over the plateau region must be attributed to the entropy-enthalpy compensation phenomenon mentioned above.

R-SG2

The complex R-SG2 also starts separating from the C-terminus of SG2 at its point of attachment near the N-terminus of R. The β-sheet structure is completely lost at 1.16 nm (after 10 ns) as can be seen in the secondary structure profile (Figure 3c). The β-sheet structure of R-SG2 is lost much faster than that of R-SG1. After only 1.16 nm (9 ns), the free energy curve (Figure 6) shows a much wider flat region compared to complex R-SG1. Similarly as in complex R-SG1, a free energy of about 60 kJ/mol was required to break 4 of the 8 initial interstrand hydrogen bonds at the C-terminus of SG2. In this flat region of the potential curve which extends to about 22 ns, interstrand hydrogen bonds (Figure 4c) and salt bridges between residues at the end of the C-terminus of R (E22 and D23) and N-terminus of SG2 (Dab15 Orn16 and Mleu17) maintain contact with each other. As with R-R and R-SG1, at the end of the flat region, the SG2 strand flips end over end while the same residues keep interacting. As the centers-of-mass move further apart, these interactions are broken and the free energy rises by 35 kJ/mol till it finally plateaus again at 30 ns with 2-3 H-bonds remaining. There is little cost in free energy to break these last H-bonds. The complex separates completely at around 2.42 nm (corresponding to 32 ns) with an additional free energy cost of about 2 kJ/mol (Figure 6). Complex R-SG2 separates much earlier (by 15 ns) than complex R-SG1, although the free energy of binding is similar. The total free energy required to dissociate the complex R-SG2 is $\Delta G_{dis} = 97$ kJ/mol.

R-KLVFF

The complex R-KLVFF starts coming apart from the N-terminus of R and C-terminus of KLVFF. The β-sheet character is lost early at around 10 ns as can be seen in the secondary structure profile for complex R-KLVFF (Figure 3d). At a separation of 1.25 nm (10 ns), most of the interstrand hydrogen bonds that make up the β-sheet (Figure 4d) were lost at a free energy cost of about 45 kJ/mol (Figure 6). Unlike complexes, R-SG1 and R-SG2, no flat region is

observed in the free energy curve but small shoulders are observed. At around 1.62 nm (around 16 ns), the KLVFF strand flips over while maintaining the interaction between residues, K16 and L17, from KLVFF, and residues, E22 and D23, of R. These residues keep interacting till they come apart at 2.12 nm (around 24 ns) with an additional energy cost of about 31 kJ/mol. Complex R-KLVFF comes apart much faster than the other complexes since KLVFF is a shorter β-strand segment. Unlike the other systems, the free energy profile does not flatten out until all of the H-bonds are lost. The total free energy cost of the dissociation is $\Delta G_{\text{dis}} = 76$ kJ/mol.

Table 1. Characteristics of the complexes between Aβ(13-23) (= R) and three β-sheet inhibitors

Complex	E_{bind}[a] (kJ/mol)	Buriedness fraction	BHB Score	N[b]	ΔG_{dis}[c] (kJ/mol)
R-R	-1212	0.81	327	9.7	131
R-SG1	-1090	0.81	381	7.3	94
R-SG2	-1030	0.74	331	7.8	97
R-KLVFF	-552	0.68	100	4.9	76

[a] Binding energy given by the MMFF94x forcefield
[b] Number of interstrand hydrogen bonds averaged over the equilibration MD simulation step.
[c] Total free energy of dissociation from AFM-MD simulation – see Figures 5 and 6

Binding affinities comparison
The AFM-MD results are in good agreement with the BHB scores for binding of SG1, SG2, and KLVFF to R (Table 1). SG1 is predicted by BHB to bind most strongly, SG2 slightly less so, and both much more strongly than KLVFF. However, on the basis of the BHB score, both SG1 and SG2 were expected to bind slightly more strongly to R than R to itself, in disagreement with the AFM results which show that the self binding of R is the strongest. Complexes R-SG1 and R-SG2 have comparable binding affinities of 94 and 97 kJ/mol, respectively, and they have comparable BHB scores. The present AFM-MD results show that SG1 and SG2 are more competitive than KLVFF (76 kJ/mol) but not as competitive as R binding to itself (131 kJ/mol). As mentioned earlier in the report, the BHB score is based in the binding energy given by the MMFF94x forcefield but adds a term to account for hydrophobic interactions and gives extra weight to salt-bridge H-bonding. These extra terms are intended to improve the predictive ability to rank drug candidates for their interaction with larger biomolecules like proteins and nucleic acids. The predictive power of the BHB score as applied to Aβ is born out in the limited antifibrillization and aggregation tests that have been carried out with Aβ itself.[26] The present AFM-MD results for complexing to the model receptor, R, correlate best with the first term, the binding energy which is highest to R-R (Table 1). It is of considerable interest that the binding affinity results from AFM, ΔG_{dis}, are directly proportional to the number of interstrand hydrogen bonds, N, of the complexes averaged over the pre-AFM equilibration trajectories: $\Delta G_{\text{dis}} = 11.1\,N + 16.8$ (in kJ/mol), with $R^2 = 92.9\%$. These include not only H-bonds that form the antiparallel β-

sheet, but also interstrand salt bridges. From the slope of the line, each H bond contributes close to 11 kJ/mol to the free energy of binding in water. The intercept indicates a constant component of about 17 kJ/mol for each of the four complexes. It is likely that this constant component ensues largely from the interactions in the common hydrophobic core region, KLVFF of each complex. In the BHB algorithm, this is taken into account by the "buriedness", namely the fraction of each ligand's surface that is not accessible to solvent. This fraction lies in the narrow range 0.7 – 0.8 for the four ligands (Table 1).

The current AFM-MD determination of the average free energy of hydrogen bond formation in water, 11 kJ/mol, compares well with a wide range of determinations by a wide range of techniques. An early MD simulation of alanine dipeptide dimer yielded a value of 11.4 kJ/mol.[41] Calorimetric measurements on barnase yielded an estimate of 12 kJ/mol.[42] In globular proteins, an estimate of 9.2 kJ/mol has been determined,[43] similar to the value in human lysozyme, 8.9 ± 2.6 kJ/mol.[44] However, a survey of all classes of proteins yielded a much lower value, 0.95 kJ/mol.[45]

Conclusions

The MD simulations of Aβ(13-23) (= R) and three (pseudo)peptidic β-sheet inhibitor ligands, SG1, SG2, and KLVFF as well as dimeric complexes between them, have been carried out. The monomers are predominantly extended β-strands, with Aβ(13-23) having the greatest flexibility to adopt other conformations. All dimeric systems exist as antiparallel β-sheets that are stable in water. AFM-MD studies, in which the dimers are slowly pulled apart by their centers-of-mass yield high values for the free energy of dissociation, ΔG_{dis} (in kJ/mol): R-R, 131; R-SG1, 94, R-SG2, 97; R-KLVFF, 76. The average free energy contribution of an inter-amide hydrogen bond is found to be 11 kJ/mol, in good agreement with many but not all values reported in the literature. Each ligand attached to Aβ(13-23) begins dissociation by peeling back from its C-terminus, breaking interstrand H-bonds, and losing the β-sheet character. The anchoring salt bridges at the C-terminus of Aβ(13-23)and the N-terminus of the ligands are the last to release, and presumably are the first to form in the reverse process of aggregation. The free energy profiles of the dissociation as a function of the separation of the centers-of-mass of all systems show plateau regions in which separation takes place with relatively little or no rise in free energy. For each system the dissociation profile does not have a maximum and reaches a flat plateau. By implication, the reverse process of assembly does not have a barrier. This, and the plateau regions in the dissociation profile, are examples of entropy-enthalpy compensation that arise naturally during the AFM-MD simulations.

Acknowledgements

References

1. Glenner, G. G.; Wong, C. W. *Biochem. Biophys. Res. Commun.* **1984**, *120*, 885.

2. Glenner, G. G. ; Wong, C. W. *Biochem. Biophys. Res. Commun.* **1984**, *122*, 1131.

3. Hardy, J.; Selkoe, D. J. *Science* **2002**, *297*, 353.

4. Nunan, J.; Small, D. H. *FEBS Lett.* **2000**, *483*, 6.

5. Tjernberg, L. O.; Callaway, D. J. E.; Tjernberg, A.; Hahne, S.; Lilliehöök, C.; Terenius, L.; Thyberg, J.; Nordstedt, C. *J Biol. Chem.* **1999**, *274*,12619.

6. Petkova, A. T.; Yau, W.-M.; Tycko, R. *Biochemistry* **2006**, *45*, 498.

7.Sato, T.; Kienlen-Campard, P.; Ahmed, M.; Liu, W.; Li, H.; Elliott, J. I.; Aimoto, S.;

Constantinescu, S. N.; Octave, J.-N.; Smith, S. O. *Biochemistry* **2006**, *45*, 5503.

8.Luhrs, T.; Ritter, C.; Adrian, M.; Riek-Loher, D.; Bohrmann, B.; Dobeli, H.; Schubert, D.; Riek, R. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 17342.

9. Shankar, G. M.; Li, S.; Mehta, T. H.; Garcia-Munoz, A.; Shepardson, N. E.; Smith, I.; Brett, F. M.; Farrell, M. A.; Rowan, M. J.; Lemere, C. A.; Regan, C. M.; Walsh, D. M.; Sabatini, B. L.; Selkoe, D. J. *Nature Medicine* **2008**, *14*, 837.

10. Cohen, E.; Bieschke, J.; Perciavalle, R. M.; Kelly, J. W.; Dillin, A. *Science* **2006**, *313*, 604.

11. Bieler, S.; Soto, C. *Curr. Drug Targets* **2004**, *5*,: 553.

12. Findeis, M. A. *Biochim. Biophys. Acta.* **2000**, *1502*, 76.

13. Stott, K. et al. *US Patent Number* 7,060,671. June 13, **2006** .

14. Kokkoni, N.; Stott, K.; Amijee, H.; Mason, J. M.; Doig, A. J. *Biochemistry.* **2006**, *45,* 9906.

15. (a) Findeis, M. et al. US Pat. 6 831 066. Dec14, 2004. (b) Findeis, M. et al. US Pat. 6,689,752. Feb 10, 2004. (c) Adessi, C.; Frossard, M. J.; Boissard, C.; Fraga, S.; Bieler, S.; Ruckle, T.; Vilbois, F.; Robinson, S. M.; Mutter, M.; Banks, W. A.; Soto, C. *J. Biol. Chem.* **2003**, *278,* 13905.

16. Permanne, B.; Adessi, C.; Saborio, G. P.; Fraga, S.; Frossard, M. J.; Van Dorpe, J.; Dewachter, I.; Banks, W. A.; Van Leuven, F.; Soto, C. *FASEB J.* **2002**, *16,* 860.

17. Chacon, M. A.; Barria, M. I.; Soto, C.; Inestrosa, N. C. *Mol Psychiatry.* **2004**, *9,* 953.

18. Tjernberg, L. O.; Naeslund, J.; Lindqvist, F.; Johansson, J.; Karlstroem, A. R.; Thyberg, J.; Terenius, L.; Nordstedt, C.. *J. Biol. Chem.* **1996**, *271*, 8545.

19. Tjernberg, L. O.; Lilliehook, C.; Callaway, D. J.; Naslund, J.; Hahne, S.; Thyberg, J.; Terenius, L.; Nordstedt, C. *J. Biol. Chem*. **1997**, *272*, 12601.

20. Yao, Z.-X.; Papadopoulos, V. *FASEB J*. **2002**, *16*, 1677.

21. Sadowski, M. J.; Pankiewicz, J.; Scholtzova, H.; Mehta, P. D.; Prelli, F.; Quartermain, D.; Wisniewski, T. *PNAS* **2006**, *103*, 18787.

22. Wang, H. Y.; Lee, D. H. S.; Davis, C. B.; Shank, R. P. *J. Neurochem*. **2000**, *75*, 1155. As a cautionary note, the binding site of the α7AChR to Aβ(1-42) was determined by molecular modeling assuming an α-helical structure for Aβ.

23. A more recent paper claims there is no binding: Small, D. H.; Maksel, D.; Kerr, M. L.; Ng, J.; Hou, X.; Chu, C.; Mehrani, H.; Unabia, S.; Azari, M. F.; Loiacono, R.; Aguilar, M.-I.; Chebib, M. *J. Neurochem.* **2007**, *101*, 1527.

24. Lustbader, J. W.; Cirilli, M.; Lin, C.; Xu, H. W.; Takuma, K.; Wang, N.; Caspersen, X. Chen, S. Pollak, M. Chaney, F. Trinchese, S. Liu, F. Gunn-Moore, L.-F. Lue, C.; Walker, D. G.; Kuppusammy, P.; Zewier, Z. L.; Arancio, O.; Stern, D.; Yan, S. S.; Wu, H. *Science* **2004**, *304*, 448.

25. Garrett, M.; Morris, D. S.; Goodsell, R. S.; Halliday, R. H.; Hart, W. E.; Belew, R. K.; Olson, A. J. *J. Comp. Chem.* **1998**, 1640.

26. Roy, S.; Rauk, A., to be published. Initial results on SG1 and SG2 have been communicated: Roy, S.; Rauk, A. Design and Testing of Novel Potent Inhibitors of beta Amyloid Oligomerization. CSC2007 - 90th Chemistry Conference and Exhibition, Winnipeg, Manitoba, Canada , May 26-29, 2007.

27. Feher, M.; Deretey, E.; Roy, S. *J. Chem. Inf. Comput. Sci.* **2003**, *43,* 1316.

28. van der Spoel, D.; van Buuren, A. R.; Apol, E.; Meulenhoff, P. J.; Tieleman, D. P.; Sijbers, A. L. T. M.; Hess, B.; Feenstra, K. A.; Lindhal, E.; van Drunen, R.; Berensden, H. J. C. *Gromacs User Manual version 3.1.1*, Nijenborgh 4, 9747 AG Groningen, The Netherlands. Internet: www.gromacs.org , 2002.

29. Essmann, U.; Perera L.; Berkowitz M. L. *J. Chem. Phys.* **1995**, *103*, 8577.

30. Hess, B.; Bekker H.; Berendsen H. J. C.; Fraaije. J. G. E. M. *J. Comp. Chem.*, **1997**, *18*, 1463.

31. Daura, X.; van Gunsteren, W. F.; Mark, A. E. *Proteins: Structure, Function, and Genetics* **1999**, *34*, 269.

32. Humphrey, W.; Dalke, A.; Schulten, K.; "VMD - Visual Molecular Dynamics", *J. Molec. Graphics* **1996**, *14*, 33-38. <http://www.ks.uiuc.edu/Research/vmd/>

33. Frishman, D.; Argos, P. *PROTEINS: structure and function* 1995, *23*, 566

34. Mobley, D. L.; Chodera, J. D.; Dill, K. A. *J. Chem. Phys.* **2006**, *125*, 084902/1-084902/16.

35. Nocedal, J. *Math. Comput.* **1980**, *35*, 773

36. Liu, D. C.; Nocedal, J. *Math. Program*. **1989**, *45*, 503

37. The home page for GRACE: http://plasma-gate.weizmann.ac.il/Grace/

38. Raffa, D. F.; Rauk, A. *J. Phys. Chem. B.* **2007**, *111*, 3789-3799. This paper shows 850 ns of the trajectory. The simulation was continued for an additional 500 ns after publication.

39. Dunitz, J. D. *Chemistry and Biology* **1995**, *2*, 709.

40. Sharp, K *Protein Science* **2001**, *10*, 661.

41. Tobias, D. J.; Sneddon, S. F.; Brooks III, C. L. *J Mol. Biol.* **1992**, *227*, 1244.

42. Griko, Y. V**.;** Makhatadze, G. I**.;** Privalov, P. L.; Hartley, R. W**.** *Protein Science* **1994,** *3*, 669.

43. Myers, J. K.; Pace, C. N. Biophys. J. **1996**, *71*, 2033.

44. Takano, K.; Yamagata, Y.; Kubota, M.; Funahashi, J.; Fujii, S.; Yutani, K. *Biochemistry* **1999**, *38*, 6623.

45. Ragone, R. *Protein Sci.* **2001**, *10*, 2075.