J|A|C|S
ARTICLES
Published on Web 01/11/2007

Dual Binding Modes of Congo Red to Amyloid Protofibril Surface Observed in Molecular Dynamics Simulations

Chun Wu, Zhixiang Wang, Hongxing Lei, Wei Zhang, and Yong Duan*

Contribution from the UC Davis Genome Center and Department of Applied Science, University of California, Davis, California 95616

Received August 29, 2006; E-mail: duan@ucdavis.edu

Abstract: Congo red has been used to identify amyloid fibrils in tissues for more than 80 years and is also a weak inhibitor to both amyloid-β fibril formation and toxicity. However, the specificity of the binding and its inhibition mechanism remain unclear. Using all-atom molecular dynamics simulations with the explicit solvent model, we have identified and characterized two specific binding modes of Congo red molecules to a protofibril formed by an amyloidogenic fragment (GNNQQNY) of the yeast prion protein Sup35. The observation of dual-mode was consistent with the experimentally observed dual-mode binding to Aβ fibrils by a series of compounds similar to Congo red. In the primary mode, Congo red bound to a regular groove formed by the first three residues (GNN) of the β-strands along the β-sheet extension direction. Comparative simulations demonstrated that Thioflavin T also bound to the grooves on KLVFFAE protofibril surface. Because of the ubiquitous long grooves on the amyloid fibril surface, we propose that this binding interaction could be a general recognition mode of amyloid fibrils by Congo red, Thioflavin T, and other long flat molecules. In the secondary mode, Congo red bound parallel to the β-strands on the edge or in the middle of a β-sheet. The primary binding mode of Congo red and GNNQQNY protofibril was more stable than the secondary mode by -5.7 kcal/mol as estimated by the MM-GBSA method. Detailed analysis suggests that the hydrophobic interactions play important roles for burial of the hydrophobic part of the Congo red molecules. Two potential inhibition mechanisms of disrupting β-sheet stacking were inferred from the primary mode, which could be exploited for the development of non-peptidic amyloid-specific inhibitors.

Introduction

The discovery of a class of amyloid fibril-related diseases[1−4] has stimulated intensive research interest. These diseases affect a large human population worldwide and lack effective treatment. The identification of amyloid fibrils and, more likely, protofibrillar species as the primary sources of toxicity[5−9] has raised the hope that effective therapeutic agents could be developed in the near future. Yet, the drug-design community is facing tremendous challenges. First, although the early ordered profibrillar oligomers might be potential drug targets, the lack of molecular level knowledge on the interactions between small molecules and the early ordered oligomers hinders the progress. This is critical because many of the peptides involved in the formation of amyloid fibrils lack stable fold and are generally considered as "natively disordered". Second, unlike protein targets with specific binding sites, these fibrillar and protofibrillar species do not have well-formed cavities. Rather, they typically have large surface available for binding. Third, the interactions between ligand and different aggregate species are hard to differentiate experimentally.[10] This study is motivated by both the need to unveil the molecular mechanisms of binding between (weak) inhibitors and the protofibrillar oligomers and the need to understand the surface property of the peptide oligomers.

Congo red (CR) has been reported to reduce the neurotoxicity of amyloid-β (Aβ) of Alzheimer's diseases by binding to preformed fibrils.[11] Because of the inhibition activity, CR and its analogues have been screened as potential therapeutics of amyloid fibril formation.[12−15] CR is a long and flat molecule with both apolar and polar parts (Figure 1). The apolar part is comprised of a biphenyl group at the center and is extended by a diazo and two flanked naphthalene groups. The polar amino group and the negatively charged sulfate group are linked to a naphthalene ring. CR binds to amyloid fibrils formed by both

(1) Thirumalai, D.; Klimov, D. K.; Dima, R. I. *Curr. Opin. Struct. Biol.* **2003**, *13*, 146−159.
(2) Rochet, J. C.; Lansbury, P. T. *Curr. Opin. Struct. Biol.* **2000**, *10*, 60−68.
(3) Dobson, C. M. *Trends Biochem. Sci.* **1999**, *24*, 329−332.
(4) Kelly, J. W. *Curr. Opin. Struct. Biol.* **1998**, *8*, 101−106.
(5) Hashimoto, M.; Rockenstein, E.; Crews, L.; Masliah, E. *Neuromol. Med.* **2003**, *4*, 21−35.
(6) Lorenzo, A.; Razzaboni, B.; Weir, G. C.; Yankner, B. A. *Nature* **1994**, *368*, 756−760.
(7) Hardy, J.; Selkoe, D. J. *Science* **2002**, *297*, 353−356.
(8) Bucciantini, M.; Giannoni, E.; Chiti, F.; Baroni, F.; Formigli, L.; Zurdo, J. S.; Taddei, N.; Ramponi, G.; Dobson, C. M.; Stefani, M. *Nature* **2002**, *416*, 507−511.
(9) Bucciantini, M.; Calloni, G.; Chiti, F.; Formigli, L.; Nosi, D.; Dobson, C. M.; Stefani, M. *J. Biol. Chem.* **2004**, *279*, 31374−31382.
(10) Cairo, C. W.; Strzelec, A.; Murphy, R. M.; Kiessling, L. L. *Biochemistry* **2002**, *41*, 8620−8629.
(11) Lorenzo, A.; Yankner, B. A. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 12243−12247.
(12) Rudyk, H.; Vasiljevic, S.; Hennion, R. M.; Birkett, C. R.; Hope, J.; Gilbert, I. H. *J. Gen. Virol.* **2000**, *81*, 1155−1164.
(13) Lee, V. M. Y. *Neurobiol. Aging* **2002**, *23*, 1039−1042.
(14) Caughey, B.; Ernst, D.; Race, R. E. *J. Virol.* **1993**, *67*, 6270−6272.
(15) Caughey, B.; Race, R. E. *J. Neurochem.* **1992**, *59*, 768−771.

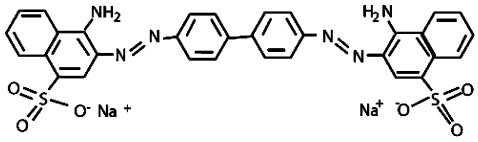

Figure 1. Structure of Congo red.

hydrophobic peptides (such as Aβ peptide) and hydrophilic peptides (such as GNNQQNY) and induces green-yellow birefringence under polarized light.[16−19] Two binding modes have been proposed.[19−24] One mode is β-sheet specific in which the CR molecules are aligned along the fibril axes. In this mode, the CR molecules are thought to be stabilized either by the electrostatic interactions between the negatively charged sulfate groups of CR and the positively charged amino acid residues of proteins[19] or by the hydrophobic effect as inserting CR into the grooves on the β-sheet surface.[22−24] The other mode is β-strand specific in which a CR molecule is parallel to the β-strand and is thought to intercalate between two β-strands in the antiparallel β-sheets.[20,21] Dual binding modes have been inferred from low-resolution spectroscopic experiments.[25−28] However, the specificity and the stabilities of these binding modes and their roles in amyloid fibril detection and inhibitions remain elusive.

Eisenberg and co-workers[29] have recently solved the high-resolution crystal structures of amyloid-like fibrils by X-ray crystallography. One of the amyloid-like fibrils they studied was formed by a seven-residue peptide fragment (GNNQQNY) from yeast prion protein Sup35. The fibril shows all of the common characteristics of amyloid fibrils, including binding to the flat dyes such as CR and Thioflavin T (TFT) and exhibiting the characteristic green-yellow birefringence of CR. In this study, we investigate the binding of CR to a GNNQQNY dodecamer organized as a pair of six-strand β-sheets. The specificity and the stability of the possible binding modes were illuminated in atomic detail. The potential roles that the different binding modes may play in amyloid detection and inhibitions are discussed. To examine the generality of the conclusions, we simulated binding of TFT to a 16-peptide protofibrillar oligomer of a seven-residue peptide fragment (KLVFFAE) from Aβ peptide.

We have successfully exploited the strategy of using higher solute concentration and elevated temperature (320 K) to reduce the computational cost associated with simulating peptide association[30−32] and ligand binding.[33] Here, we applied the same strategy in this study. The concentration of the protofibril and CR (~mM) was about 1000-fold higher than the typical experimental concentrations (~μM). The increased concentration significantly reduced the computational cost in two ways. First, higher concentration allows the use of a smaller water box in simulations, which results in direct reduction in the computation cost. Second, the increased solute concentration reduces the time needed for the slow diffusion step. The combined reduction in the computational cost due to the elevation of concentration was expected to be close to 6 orders of magnitude. In this study, we focus on the binding modes and evaluating their relative stability rather than characterizing the kinetic process of the binding.

Methods

System Preparation. Eight simulations were performed on the GNNQQNY oligomer with CR starting from two different configurations. Each simulation system consisted of a 12-peptide protofibrillar oligomer, two CR molecules, four sodium ions, and 7268 water molecules. The 12-peptide protofibrillar oligomer was constructed from the crystallographic fibril structure (PDB ID 1yjp)[29] of GNNQQNY peptide fragment from yeast prion protein Sup35. The 12 peptides were arranged in a double-layered β-sheet structure with each layer consisting of 6 parallel in-register β-strands. The two β-sheets were related to one another by a 2_1 screw axis along the β-sheet extension (hydrogen bond) direction according to the X-ray structure. The dimensions of the protofibrillar oligomer were $22 \times 25 \times 18$ Å^3 along, respectively, β-strand, β-sheet extension (main-chain hydrogen-bond direction), and β-sheet stacking directions (perpendicular to the β-sheet surface). Four simulations on each of the two initial CR configurations were performed. In the first configuration, the two CR molecules were initially placed ~15 Å away from the 12-peptide oligomer along the β-strand extension direction and were on the opposite sides of the protofibril. In the second configuration, the two CR molecules were placed along the β-sheet stacking direction. Four positive sodium ions (Na^+) were added to neutralize the four negative charges carried by the two CR molecules. The solute molecules were immersed into a rectangle box of 7268 water molecules with dimensions of either $58 \times 101 \times 51$ Å^3 or $59 \times 60 \times 84$ Å^3. The periodic water box was constructed in such a way that the solute was at least ~12 Å away from the box surface and the minimum distance between the solute and the image was ~24 Å. The effective concentrations of CR and the protofibril oligomer were ~14 and ~7 mM, respectively. The advantage of using two CR molecules in each simulation system allows sampling of two conformations in a single trajectory and thus enhances sampling in comparison to systems with a single CR molecule.

The Duan et al. all-atom point-charge force field[34] (AMBER ff03) was chosen to represent the peptide. The basic form of CR (pK_a = 4.0) was used in this study because it is the dominant form at neutral pH condition, and the structure is shown in Figure 1. After geometry optimization at HF/6-31G* level, the partial charges were derived by fitting to the gas-phase electrostatic potential calculated at HF/6-31G* level of quantum mechanical theory using the RESP (restrained electrostatic potential) method.[35] Other force parameters of CR molecule

(16) Benhold, H. *Muenchen. Med. Wochenschr.* **1922**, 1537−1538.
(17) Divry, P. *J. Neurol. Psychiatry* **1927**, *27*, 643−657.
(18) Klunk, W. E.; Jacob, R. F.; Mason, R. P. *Anal. Biochem.* **1999**, *266*, 66−76.
(19) Klunk, W. E.; Pettegrew, J. W.; Abraham, D. J. *J. Histochem. Cytochem.* **1989**, *37*, 1273−1281.
(20) Turnell, W. G.; Finch, J. T. *J. Mol. Biol.* **1992**, *227*, 1205−1223.
(21) Carter, D. B.; Chou, K. C. *Neurobiol. Aging* **1998**, *19*, 37−40.
(22) Romhanyi, G. *Virchows Arch. A: Pathol. Anat.* **1971**, *354*, 209−222.
(23) Cooper, J. H. *Lab. Invest.* **1974**, *31*, 232−238.
(24) Krebs, M. R. H.; Bromley, E. H. C.; Donald, A. M. *J. Struct. Biol.* **2005**, *149*, 30−37.
(25) Miura, T.; Yamamiya, C.; Sasaki, M.; Suzuki, K.; Takeuchi, H. *J. Raman Spectrosc.* **2002**, *33*, 530−535.
(26) Elhaddaoui, A.; Delacourte, A.; Turrell, S. *J. Mol. Struct.* **1993**, *294*, 115−118.
(27) Sajid, J.; Elhaddaoui, A.; Turrell, S. *J. Mol. Struct.* **1997**, *408*, 181−184.
(28) Pigorsch, E.; Elhaddaoui, A.; Turrell, S. *Spectrochim. Acta, Part A* **1994**, *50*, 2145−2152.
(29) Nelson, R.; Sawaya, M. R.; Balbirnie, M.; Madsen, A. O.; Riekel, C.; Grothe, R.; Eisenberg, D. *Nature* **2005**, *435*, 773−778.
(30) Wu, C.; Lei, H.; Duan, Y. *Biophys. J.* **2004**, *87*, 3000−3009.
(31) Wu, C.; Lei, H.; Duan, Y. *Biophys. J.* **2005**, *88*, 2897−2906.
(32) Wu, C.; Lei, H.; Duan, Y. *J. Am. Chem. Soc.* **2005**, *127*, 13530−13537.
(33) Wu, C.; Lei, H.; Wang, Z. X.; Zhang, W.; Duan, Y. *Biophys. J.* **2006**, *91*, 3664−3672.
(34) Duan, Y.; Chowdhury, S.; Xiong, G.; Wu, C.; Zhang, W.; Lee, T.; Cieplak, P.; Caldwell, J.; Luo, R.; Wang, J.; Kollman, P. A. *J. Comput. Chem.* **2003**, *24*, 1999−2012.
(35) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269−10280.

were taken from AMBER GAFF[36] parameter set. The solvent was explicitly represented by the TIP3P[37] water model.

To validate the force field parameters of CR, a simulation (10 ns) with the TIP3P water model was conducted on CR alone at 320 K. The bond lengths and bond angles were well maintained in the MD simulations in comparison with QM structure (data not shown). The torsion angles of the rotatable bonds between rings were well sampled (data not shown). The parameter file in AMBER format is available upon request.

MD Simulation. The AMBER 8 simulation package[38,39] was used in both molecular dynamics simulations and data processing. The ligand−protofibril−water system was subjected to periodic boundary conditions via both minimum image and discrete Fourier transform as part of the Particle-Mesh Ewald method.[40] After the initial energy minimization, a total of eight simulations (four runs for each of the two CR configurations) were performed with different initial random velocities. The initial velocities were generated according to the Maxwell−Boltzmann's distribution at 500 K. The simulations started after a 10.0 ps run at 500 K to randomize the orientations and positions of the two CR molecules. A short 1.0 ns molecular dynamics at 320 K in the NPT ensemble (constant pressure and temperature) was performed to adjust system size and density, and to equilibrate the solvent. The simulations were continued at 320 K for 19 ns in the NVT ensemble (constant volume and temperature). The elevated temperature (10 K higher than 37 °C) was used in the simulations to enhance the hydrophobic effect. The Particle-Mesh Ewald method[40] was used to treat the long-range electrostatic interactions. SHAKE[41] was applied to constrain all bonds connecting hydrogen atoms, and a time step of 2.0 fs was used. To reduce the computation, nonbonded forces were calculated using a two-stage RESPA approach[42] where the forces within a 10 Å radius were updated every step and those beyond 10 Å were updated every two steps. Temperature was controlled at 320 K by using Berendsen's algorithm[43] with a coupling constant of 2.0 ps. The center of mass translation and rotation were removed every 500 steps. Studies have shown this removes the "block of ice" problem.[44,45] The trajectories were saved at 2.0 ps intervals for further analysis.

Binding Free Energy Calculation. The binding free energies of CR to the protofibril in two binding modes were evaluated on the representative structures of the respective clusters using the MM-GBSA (Molecular Mechanics-Generalized Born/Surface Area) method[46] as implemented in AMBER. In this method, the total binding free energy in water is approximated by $\Delta G_{TOT} = \Delta E_{GAS} + \Delta G_{GB} + \Delta G_{SUR}$. The ΔE_{GAS} is the energy difference of the solutes in the two (bound and unbound) states. The ΔG_{GB} is the polar part of the solvation free energy represented by the Generalized Born approach. The ΔG_{SUR} is the surface area term, approximating the apolar part of the solvation free energy (i.e., creation of a cavity inside the solvent). In this formula, the conformational entropy of the solute is not explicitly considered, although the solvent entropy is implicitly considered in the ΔG_{GB} and ΔG_{SUR}. Although the MM-GBSA calculations may overestimate the absolute binding free energy due to the missing terms (e.g., conformational entropy change of the solute upon binding), they usually give a reasonable estimate on the relative binding free energy when the conformational entropy changes of two binding modes are comparable.

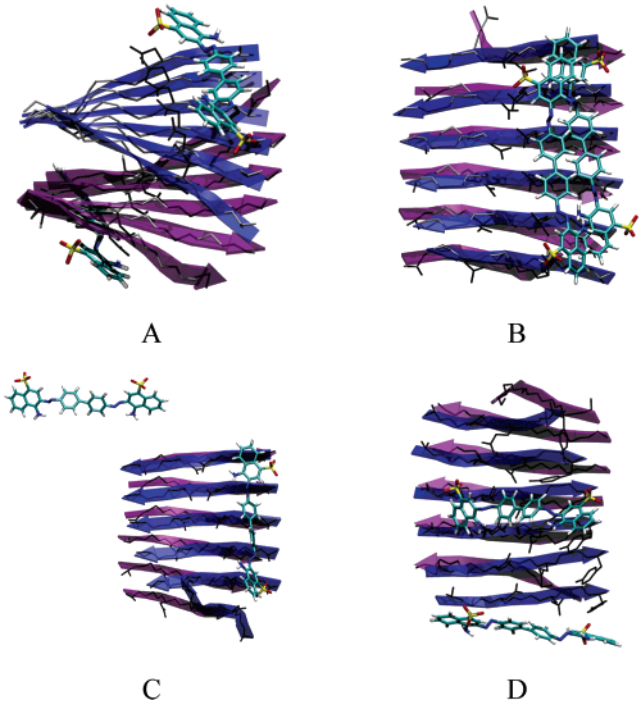

Figure 2. Four representative structures at the end of the eight simulations. (A) CR bound to the N-termini of the β-sheets (three trajectories). (B) CR dimer bound to the N-termini (three trajectories). (C) One CR bound to the N-termini of a β-sheet and the other was in unbound form (one trajectory). (D) CR bound parallel to β-strands in one trajectory. For clarity, water molecules are not shown. The backbones of each β-strand are shown in arrows with one β-sheet in blue and the other in purple. The side-chains (only heavy atoms) and CR molecules are in licorice. H, O, C, and S elements of CR are in white, red, cyan, and yellow, respectively.

Simulations of TFT and the KLVFFAE Protofibril. Two simulations were performed to study the binding of TFT to the KLVFFAE protofibril. Each system consisted of a 16-peptide protofibrillar oligomer, four TFT molecules, four chloride ions, and 9102 water molecules. NMR experiment[47] has shown that the KLVFFAE fragment from amyloid-β peptide forms antiparallel in-register β-sheets. The 16-peptide protofibrillar oligomer was arranged in a double-layered β-sheet structure with each layer consisting of 8 antiparallel in-register β-strands.[47] The four TFT molecules were initially placed ~15 Å away from four faces of the 16-peptide oligomer. Four chloride ions (Cl^-) were added to neutralize the system (TFT molecules are positively charged). These solute molecules were immersed into a rectangle box of 9102 water molecules with dimensions of 94 × 56 × 72 Å^3. The effective concentrations of TFT and the protofibril oligomer were ~18 and ~4 mM, respectively. The simulations were performed, and the parameters were obtained by following procedures identical to those described earlier.

Results

Shown in Figure 2 are four representative structures taken among the final snapshots of the eight trajectories. Irrespective of their initial positions and orientations, CR molecules consistently bound to the protofibrils in two specific modes. This is quite different from the nonspecific binding of phenol red

(36) Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(37) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Comput. Phys.* **1983**, *79*, 926−935.
(38) Case, D. A.; et al. *AMBER 8*; University of California, San Francisco, 2004.
(39) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005**, *26*, 1668−1688.
(40) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T. A.; Lee, H.; Pedersen, L. G. *J. Comput. Phys.* **1995**, *103*, 8577−8593.
(41) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(42) Procacci, P.; Berne, B. J. *Mol. Phys.* **1994**, *83*, 255−272.
(43) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Comput. Phys.* **1984**, *81*, 3684−3690.
(44) Chiu, S. W.; Clark, M.; Subramaniam, S.; Jakobsson, E. *J. Comput. Chem.* **2000**, *21*, 121−131.
(45) Harvey, S. C.; Tan, R. K. Z.; Cheatham, T. E. *J. Comput. Chem.* **1998**, *19*, 726−740.
(46) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E., III. *Acc. Chem. Res.* **2000**, *33*, 889−897.
(47) Balbach, J. J.; Ishii, Y.; Antzutkin, O. N.; Leapman, R. D.; Rizzo, N. W.; Dyda, F.; Reed, J.; Tycko, R. *Biochemistry* **2000**, *39*, 13748−13759.

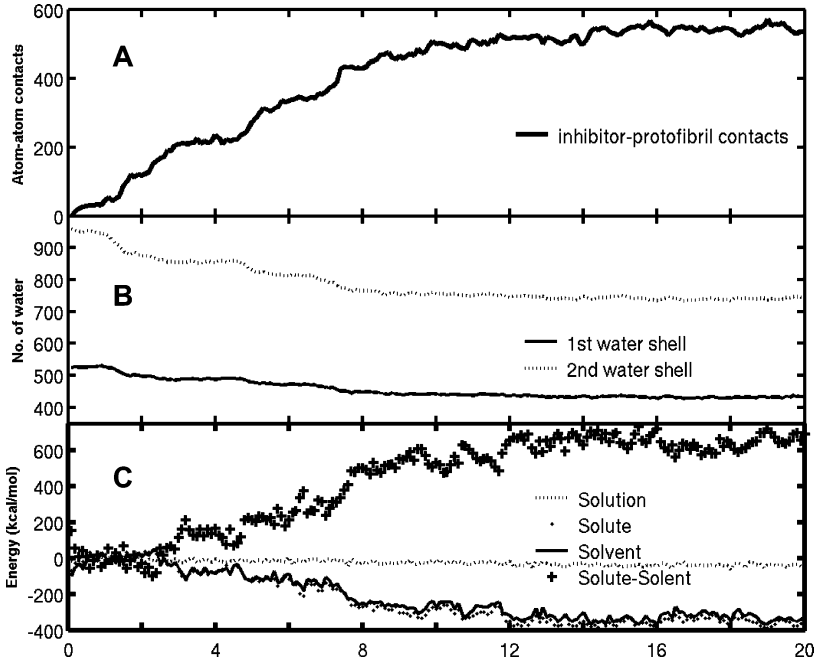

Figure 3. Time-development of properties of the protofibril−ligand−water system averaged over all simulations. (A) Atom contacts between the GNNQQNY protofibril oligomer and two CR molecules (within 5 Å cutoff). (B) Number of water molecules in the first (0.0−3.4 Å) and second (3.4−5.0 Å) solvation shells of the solute (the protofibril plus the CR molecules). (C) Potential energy changes of the solution, solute, solvent, and the interaction energy change of solute−solvent upon the binding of CR molecules to the protofibril.

molecule to the NFGAIL protofibril observed in our previous study.[33] The protofibrils maintained their overall structures and were stable in the simulations. However, the twist angle between neighboring β-strands increased from 0° in the crystal form to ~12° (e.g., Figure 2A) in the first 2.0 ns of the simulation (before binding to CR) and remained so throughout the rest of the simulations. This suggests that the perfect planarity of the β-sheets in the crystal structure (PDB ID 1yjp),[29] which is rare in protein structures, might be partially due to crystal packing.

Two binding modes were observed (Figure 2). In the first mode, CR molecules were parallel to the β-sheet extension (main-chain hydrogen bond) direction in the grooves formed by the repeating shallow concave surface, composed by the backbones of GNN and the side-chains of Gly1 and Asn3, along the main-chain hydrogen-bond direction (Figure 2A−C). Both the small side-chain of Gly1 and the parallel registration of β-strands led to the regular grooves on the surface. This mode was the primary binding mode, because 13 of the 15 bound CR molecules in the 8 simulations were in this mode. In the secondary binding mode, representing 2 of the 15 bound molecules, CR molecules were parallel to the β-strand direction, and the binding site was either next to the ending β-strand or on top of a β-strand in the middle of a β-sheet (Figure 2D). Because of its long and flat shape, binding of a CR molecule requires a groove of at least 6 continuous β-strands in a β-sheet. Such type of grooves, formed by peptide side-chains, is a hallmark of amyloid fibril surface. Thus, the primary binding mode is specific to amyloid fibrils. Furthermore, because the binding is facilitated by the grooves that are ubiquitous on amyloid fibrils, other long and flat molecules may also bind to the amyloid fibrils in the same binding mode. The secondary binding mode is β-strand-specific due to its parallel alignment to the β-strand direction, but it was less populated.

Among the total of eight trajectories, the two CR molecules bound respectively to the grooves at N-termini of the β-sheets of the protofibril in three trajectories (Figure 2A). In one trajectory, one of the two CR bound to the groove at the N-termini of a β-sheet, and the other was still free in solvent (Figure 2C). Alternatively, CR molecules were also found parallel to the β-strands and bound to either the edge of a β-sheet or the middle of a β-sheet (Figure 2D). In the remaining three trajectories, CR formed dimers and bound to the groove at the N-termini of a β-sheet (Figure 2B). This was consistent with the experimental observation[48] by Król and co-workers, which suggested that CR might bind to amyloid fibril in a supramolecular form.

As CR molecules approached the protofibril, the number of atomic contacts between CR molecules and the protofibril increased monotonically within the first 14 ns (Figure 3A). Meanwhile, water molecules were released from the complex surface, indicated by a reduction of the number of water molecules in the first two solvation shells (Figure 3B). In particular, the first solvation shell of the protofibril−ligand complex lost ~70 water molecules, and the second shell lost ~130 water molecules within the first 14 ns. After binding to the protofibril, CR molecules formed stable atom-contacts with the protofibril (~500), and the number of water molecules in the first two solvation shells became steady (~500 and ~870, respectively) (Figure 3A and B).

The energetic analysis (Figure 3C) indicates that the binding was driven by a small reduction in the total potential energy of the whole system. The total potential energy decreased by ~24.5 kcal/mol when averaged over the first and last 4.0 ns. While the solute (protofibril + CR) and the solvent energies were both favorable to binding, the solute−solvent interaction energy was against binding. The solvent energy decreased by −268.0 kcal/mol in favor of binding as a result of increased water−water

(48) Stopa, B.; Piekarska, B.; Konieczny, L.; Rybarska, J.; Spolnik, P.; Zemanek, G.; Roterman, I.; Krol, M. *Acta Biochim. Pol.* **2003**, *50*, 1213−1227.

Figure 4. Distribution of CR molecules observed in the simulations. The GNNQQNY protofibril is shown as a cartoon, and CR molecules are represented by lines. (A) The random positions of CR molecules in the first 5 ns simulations. (B) Stereoview of the bound complexes in the simulations.

hydrogen bonds due to releasing water molecules to the bulk solvent. The solute energy (protofibril + CR) also decreased substantially (−292.9 kcal/mol) because of favorable interactions between CR molecules and the protofibril. In contrast, the interaction energy between the solute and water increased by 536.4 kcal/mol and was against the binding, mainly due to loss of hydrogen bonds with water in the desolvation of the polar groups. These three components added together to the total potential energy change of −24.5 kcal/mol in favor of binding. Obviously, in addition to these enthalpic factors, one also needs to consider the entropic terms of releasing water and binding of the solutes.

We superimposed the protofibril structures in all snapshots to reveal the interaction patterns between CR molecule and the protofibril. When CR molecules were far away from the protofibril (in the first 5.0 ns simulations), their positions and orientations were random as indicated by their distribution and random orientation (Figure 4A). In the bound state, CR molecules formed four tight and populated clusters (Figure 4B). The two stable binding modes revealed by these four clusters were the same as those observed in the final snapshots of the simulations (Figure 2).

The complexes were clustered into different structural groups using the root-mean-square deviation (rmsd) of the CR molecule with a cutoff of 5 Å after aligning the protofibril. The representative structures of the top 6 clusters from the combined 8 simulation runs are shown in Figure 5. The first 4 clusters belong to the primary binding mode with a combined total of 78% of all snapshots, while the other 2 belong to the secondary binding mode with a total of 11% snapshots. In the primary mode (Figure 5A1−A4), CR was aligned parallel to the β-sheet extension direction, and the binding site was in the groove at the N-termini of GNNQQNY β-strands. The main difference of CR conformation in these four clusters is the direction of SO3 and NH2 groups, which faced different directions in the clusters. The populations of clusters A1, A2, and A3 were similar to each other (24%, 23%, and 22%, respectively), while A4 was only ∼9%. Given the similar hydrophobic interactions among these four clusters, the reduced population in cluster A4 might be caused by the unfavorable electrostatic repulsion between the two positively charged NH2 groups of CR and the positively charged NH3 groups of Gly1. In the secondary mode, CR molecule was aligned along the β-strand direction, and the binding site was either next to the edge of the ending β-strand (Figure 5B1) or on top of a β-strand in the middle of the β-sheet (Figure 5B2). Based on the population difference between the primary and the secondary binding mode, 78% and 11%,

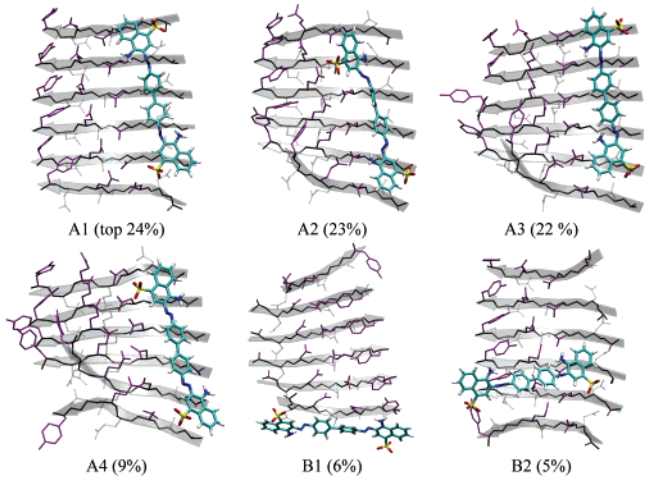

Figure 5. Representative complex structures from the most populated clusters. For clarity, only the β-sheet that bound to CR is shown. (A1−A4) CR is perpendicular to the β-strands, and the binding site is at N-termini of GNNQQNY β-strands. The upper and lower SO3 groups face different directions: (right, left), (left, right), (left, left), and (right, right) in A1−A4, respectively. (B1−B2) CR is parallel to the β-strands, and the binding site is either at the edge of the ending β-strand (B1) or on top of a β-strand (B2). The population is annotated in parentheses. The surface side-chains of Gly1, Asn3, Gln5, and Tyr7 are colored in purple, and the buried side-chains of Asn2, Gln4, and Asn6 are in silver. The backbone of each β-strand is in an arrow. The side-chains and CR molecules are in licorice. H, O, C, and S elements of CR are in white, red, cyan, and yellow color, respectively.

respectively, the primary binding mode was more stable than the secondary mode by −1.3 kcal/mol at 320 K.

A detailed analysis of molecular interactions between a CR molecule and the protofibril in the two binding modes is presented in Table 1. In the primary binding mode (Figure 5A1−A4), the apolar part of the CR (constituting 6 six-membered carbon rings plus the diazo groups) fitted in the groove at the N-termini of β-sheet very well while exposing the polar groups (NH2 and SO3) to solvent extensively. Because the groove is formed by the side-chains of Gly1 and Asn3 and the backbone, the apolar part of CR formed van der Waals contacts mainly with the first three residues and had ∼18 and ∼204 direct atom contacts with Gly1 and Asn3 side-chains, respectively. In total, the apolar part of CR lost ∼330 atom contacts with water molecules and gained ∼258 direct atom contacts with the peptide side-chains. In comparison, the polar SO3 and NH2 groups remained exposed to solvent in the bound state and maintained 13.5 (∼75%) of their 18 original hydrogen bonds with water molecules. The loss of ∼4.5 hydrogen bonds with water molecules was partially compensated by forming the hydrogen bonds to peptides (∼2) and to the other CR molecule (∼1). The contrasting behavior of the apolar and polar groups of the CR, the substantial loss of the contacts between the hydrophobic groups and the water, and the relatively small changes in the polar groups suggest that the main driving force of the primary binding mode was the hydrophobic interactions, although specific interactions such as salt bridge could also play important roles (Table 1).

In the secondary binding mode (Figure 5B1,B2), because CR was parallel to the β-strand, it could form atom contacts with all 7 residues (GNNQQNY) of a β-strand. However, because the side-chains of Asn2, Gln4, and Asn6 are buried deep inside the inner-core of the protofibril, CR formed much fewer contacts with them. For comparison, in the secondary mode, CR formed

Table 1. Atom−Atom Contacts and H-Bonds before/after Binding in Two Binding Modes

CR moieties	time (ns)	Gly	Asn	Asn	Gln	Gln	Asn	Tyr	H_2O
hydrophobic part[a]	0.0−0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	864.4
primary mode	15.0−20.0	17.6	2.3	203.9	0.0	33.6	0.0	0.4	534.0
secondary mode	15.0−20.0	6.0	16.0	87.0	0.0	77.0	0.0	33.0	695.5
2 SO3[b]	0.0−0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	15.7
primary mode	15.0−20.0	0.1	0.5	0.0	0.1	0.0	0.0	0.4	12.3
secondary mode	15.0−20.0	0.0	0.4	0.0	0.0	0.0	0.0	0.6	12.8
2 NH2[b]	0.0−0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.3
primary mode	15.0−20.0	0.1	0.0	0.0	0.0	0.0	0.0	0.2	1.2
secondary mode	15.0−20.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	1.6

[a] Atom−atom contacts between residue side-chains and the hydrophobic part (cutoff 4 Å). The hydrophobic part includes the 6 six-membered carbon rings plus the diazo groups. [b] Hydrogen bonds: donor−acceptor distance <3.5 Å and donor−hydrogen−acceptor angle >120°.

Table 2. Binding Free Energies Calculated by the MM-GBSA Method[a]

	ΔE_{GAS}	ΔG_{GB}	ΔG_{SUR}	ΔG_{TOT}
primary mode of CR[b]	−172.6	155.6	−5.8	−22.8 ± 1.6
secondary mode of CR[b]	−108.1	95.2	−4.2	−17.1 ± 1.4
difference	−64.5	60.4	−1.6	−5.7
primary mode of TFT[c]	−20.9	14.6	−1.7	−8.0 ± 2.3
secondary mode of TFT[c]	−26.2	23.8	−1.4	−3.8 ± 2.1
difference	−5.3	9.2	−0.3	−4.2

[a] All energy terms are in kcal/mol. ΔE_{GAS}: gas-phase interaction energy. ΔG_{GB}: polar part of the solvation free energy by the Generalized Born model. ΔG_{SUR}: apolar part of the solvation free energy represented by surface tension. $\Delta G_{TOT} = \Delta E_{GAS} + \Delta G_{GB} + \Delta G_{SUR}$ binding free energy in water. [b] CR-GNNQQNY complexes. [c] TFT-KLVFFAE complexes.

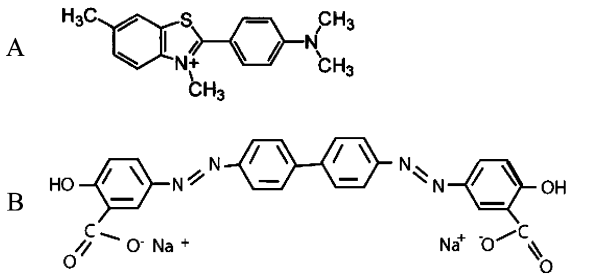

Figure 6. Structures of two amyloid dyes that are known to bind amyloid fibrils. (A) Thioflavin T (TFT). (B) Chrysamine G.

∼6, ∼87, ∼77, and ∼33 atom contacts with the side-chains of Gly1, Asn3, Gln5, and Tyr7, respectively, and only ∼16, ∼0, and ∼0 atom contacts with Asn2, Gln4, and Asn6, respectively (Table 1). The hydrophobic interactions in this binding mode were slightly weaker than those in the primary mode because the apolar part of CR had ∼40 fewer atom contacts with the side-chains and ∼161 more contacts with water molecules. Other interactions such as $\pi-\pi$ interaction between the naphthalene ring and Tyr ring could also have played roles in stabilizing the secondary binding mode as indicated by ∼33 atom contacts between Congo red and the Tyr side-chain (Table 1).

The binding free energies of the two modes were evaluated by the MM-GBSA method on the representative structures and are summarized in Table 2. The primary mode had stronger binding free energies (−22 kcal/mol) than that of the secondary mode (−17 kcal/mol), and the difference (−5.7 kcal/mol) is qualitatively consistent with the results obtained by population analysis (−1.3 kcal/mol, 320 K). More importantly, the observation of two binding modes was consistent with the experimental measurements on a series compounds that share a scaffold similar to that of CR.[49] Kang and Han found that a series of four compounds with similar scaffold, including Chrysamine G (Figure 6B), all exhibit dual-mode binding behavior to amyloid-β 1−40. Intriguingly, the binding free energy difference between the primary and the secondary modes obtained from the experiments ranged from −2.6 to −3.1 kcal/mol, which are within the range of our estimates (−1.3 to −5.7 kcal/mol). Apparently, considerable caution should be taken because of the substantial differences between the two types of amyloid fibrils. Nevertheless, it is possible that the two studies together suggest that CR may bind to amyloid fibrils in two distinct modes.

To examine the generality of the conclusions, particularly those related to the primary binding mode, we performed a set of simulations on the binding of TFT dye to the KLVFFAE protofibril with the same protocol as used in the CR-GNNQQNY simulations. The clustering analyses from all saved snapshots demonstrated that the binding can be partitioned into two modes. The primary mode, with population of ∼50% of all snapshots, was notably more favorable than the secondary mode (7%). These populations give an estimated −1.3 kcal/mol free energy difference for the TFT-KLVFFAE complexes. For the CR-GNNQQNY complexes, the population difference also results in an estimated −1.3 kcal/mol free energy difference between the two modes. On the other hand, from the MM-GBSA calculations, the binding free energies of the TFT-KLVFFAE complexes differ by −4.2 kcal/mol, which is close to that of the CR-GNNQQNY binding (−5.7 kcal/mol). In comparison to the CR-GNNQQNY system, three main differences can be noted. First, the peptide KLVFFAEA is hydrophobic and GNNQQNY is hydrophilic. Second, KLVFFAEA forms anti-parallel β-sheets and GNNQQNY forms parallel β-sheets. Third, CR is replaced by another linear dye TFT. Despite these differences, the two systems behave with remarkable similarity. Both systems exhibit dual binding modes with the primary mode notably more prevalent than the secondary mode. The binding free energy differences between the primary and the secondary modes (−1.3 to −5.7 kcal/mol vs −1.3 to −4.2 kcal/mol) are also remarkably consistent. In both cases, the primary binding modes were along the fibril direction (Figure 7A) and occurred in surface grooves formed by the sidechains (Figure 7B), and the secondary modes were along the β-strand direction. The high degree of similarity between the two systems underscores the potential generality of the observations.

Discussion

In this study, the protofibrils of both GNNQQNY and KLVFFAE (two-layered β-sheets) were modeled as a minimal

(49) Kang, J.; Han, K. *Bull. Korean Chem. Soc.* **2001**, *22*, 1065−1066.

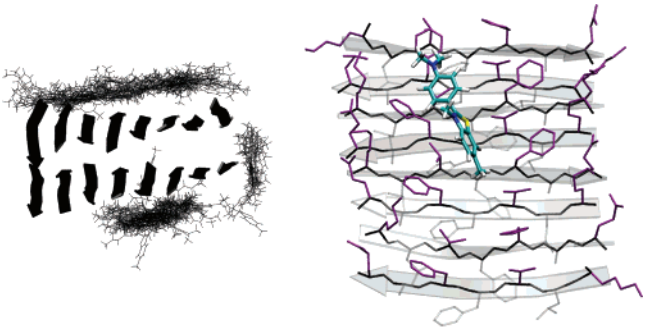

Figure 7. Binding of TFT to the KLVFFAE protofibril in the simulations. (A) TFT molecules (shown in lines) bound to the protofibril (in cartoons). (B) Representative complex structure of the primary binding mode. For clarity, only the β-sheet that bound to TFT is shown. The surface side-chains are colored in purple, and the buried side-chains are in silver. The backbone of each β-strand is in an arrow. The side-chains and thioflavin T molecules are in licorice. H, O, C, and S elements of TFT are in white, red, cyan, and yellow, respectively.

structural model of the amyloid fibril. CR and TFT are two of the most common amyloid dyes. The observed primary binding modes show how linear ligands can exploit the regular patterns (e.g., grooves) on the amyloid fibril surface. These observations may convey key information on the general recognition modes by linear amyloid dyes to amyloid fibrils formed by both hydrophobic and hydrophilic sequences.

Binding to the grooves on the amyloid fibril surface formed by side-chains is a result of the shape complementarity between the grooves and linear dyes. The alternation of side-chain directions in an extended β-strand leads to small regular concave and convex surfaces along the β-strand. Repetition of the pattern along the β-sheet direction allows formation of long grooves on the β-sheet surface. These long grooves, rarely seen in proteins, are common on and specific to the amyloid fibril surface. Therefore, the primary binding modes could be a general recognition mode of linear flat amyloid dyes such as CR, TFT, and their drivatives.[50−54] This is clearly demonstrated by the highly similar primary binding modes observed between CR and GNNQQNY and between TFT and KLVFFAE protofibrils, despite the notable differences between these two systems. Notably, the binding free energy differences between the primary and the secondary modes were quite similar in the two systems (−5.7 vs −4.2 kcal/mol from MM-PBSA calculations and −1.3 vs −1.3 kcal/mol from cluster populations). Experimental evidence supporting the primary binding modes comes from the confocal microscopy by Donald and co-workers,[24] who inferred that TFT may bind to the amyloid fibril formed by bovine insulin and bovine β-lactoglobulin along the axis of the fibril. Although the exact binding positions and binding free energies are likely sequence-specific for different amyloid fibrils, this mode is expected to be generally applicable to other amyloid fibrils.

The CR molecules were observed to aggregate and form dimers in the simulations before they bound to the protofibrils. At the effective concentration of 0.014 M, two stacking modes were observed: reversed-stacking where the sulfonic groups are on the opposite sides of the dimer and direct stacking (data not shown). These observations were in excellent agreement with the work of Roterman and co-workers,[55] who, based on their theoretical and experimental studies, found that CR may self-assemble into rod-like structures in water in both "direct" and "reversed" modes at low concentration (<0.05 M), ionic strength (>0.05 M), and temperature (<330 K).

Our analyses revealed that the hydrophobic effect was an important factor, in agreement with other studies[22−24] and complementary to that proposed by Abraham and co-workers,[19] who proposed that the electrostatic interaction was the main driving force. These observations were consistent with the experimental study on the interactions between Chrysamine G and Aβ fibril. When a $-CH_2$ group was added between the two central rings of Chrysamine G, the higher hydrophobicity led to a notable increase in the binding affinities by 5−10-fold.[49]

Carter and Chou[21] attempted to build a model of the CR bound to the core fragment (KLVFFAE) of Alzheimer Aβ fibrils. Their model was based on an X-ray complex structure of Finch and co-worker[20] where CR intercalates between two β-stands of the fragment (RGFFYTP) of pig insulin. In their modeling exercise, by sequence alignment and energy minimization, Carter and Chou constructed a model complex in which the CR intercalated into a peptide dimer. On the basis of this modeling result of a peptide dimer, they proposed that the CR could intercalate the amyloid fibrils. Our observed secondary binding mode in which the CR was parallel to the β-stands was partially consistent with the model proposed by Chou and co-worker. However, although the orientation of CR in Chou's model was the same as that in the secondary binding mode, the binding modes were notably different. While the binding site in Chou's model was between two β-strands, the binding sites observed in our simulations were either at the ends of the fibril or on top of the β-strand.

There are two possible scenarios that may lead CR molecules to intercalate between two β-strands. In the first scenario, a CR is directly inserted between the two neighboring β-strands after the formation of the fibril (or the β-sheets). This is unlikely because there could be a large (free) energy barrier to break the main-chain/side-chain hydrogen bonds between the neighboring strands and to move the two parts of a broken β-sheet outward to accommodate a CR molecule. In the second scenario, a CR molecule may bind to the end of a β-sheet and is later sandwiched by another β-strand. This could conceivably happen but depends on the competition between a CR molecule and a β-strand to bind to the end of a β-sheet. The second scenario is consistent with one of the secondary binding modes that a CR bound at one end of the protofibril. Therefore, our observation partially supports Chou's model, although the binding site of this binding mode on the surface of a β-sheet is also possible. However, judging from the results that the binding affinity of this mode was 1.3−5.7 kcal/mol weaker than the primary mode, we suggest that it might only play a secondary role in the binding of amyloids by CR.

Amyloid fibril formation is a multi-stage process involving different species at each stage. Thus, different inhibition

(50) Sabate, R.; Estelrich, J. *Biopolymers* **2003**, *72*, 455−463.
(51) Klunk, W. E.; Debnath, M. L.; Pettegrew, J. W. *Neurobiol. Aging* **1995**, *16*, 541−548.
(52) Styren, S. C.; Hamilton, R. L.; Styren, G. C.; Klunk, W. K. *J. Histochem. Cytochem.* **2000**, *48*, 1223−1232.
(53) Ashburn, T. T.; Han, H.; McGuinness, B. F.; Lansbury, P. T. *Chem. Biol.* **1996**, *3*, 351−358.
(54) Han, H.; Cho, C. G.; Lansbury, P. T. *J. Am. Chem. Soc.* **1996**, *118*, 4506−4507.
(55) Skowronek, M.; Stopa, B.; Konieczny, L.; Rybarska, J.; Piekarska, B.; Szneler, E.; Bakalarski, G.; Roterman, I. *Biopolymers* **1998**, *46*, 267−281.

strategies could be contemplated at each stage. For instance, in the early nucleation stage, small ligands can be used to block the conformational transitions of amyloid peptides by stabilizing their compact conformation and to block nucleation either by improving peptide solubility or by blocking key peptide−peptide interaction. In the growth stage, asymmetric ligands can be used to block β-sheet stacking by binding on top of β-sheet and to block β-sheet extension by binding to two ends of β-sheet. In later fibril assembly stage, one may try to block protofibril−protofibril association and to destabilize the fibril by inducing conformational change. Conversely, when small molecules like CR demonstrate amyloid inhibition property, the mechanism could be complicated because they may have the potential to interact with various species at various stages. In this study, we focus only on the interactions between ligands and protofibrillar-oligomer, which is one of the key species in amyloid fibril formation. Clearly, interactions with other species deserve further investigation.

As CR molecules fill in the grooves on the amyloid protofibril surface, the surface may become flat or even convex, making it difficult to stack multiple layers of β-sheets. This could be crucial when surface complementarity plays an important role in the assembly of amyloid fibrils (e.g., as in the "zipper" model). The exposure of the polar SO3 and NH2 groups of CR to the solvent can improve the solubility of the protofibrils to hinder further growth of the protofibril. Thus, the presence of CR on the surface may block fibril growth along the β-sheet stacking direction both by disrupting the surface pattern and by increasing the solubility of small protofibrils. Caution should be taken, however, because this binding mode could also stabilize the early (single layer) β-sheets and favor formation of early β-sheet or oligomers by providing templates to align the β-strands. The net effect should be examined by further studies.

A number of peptidic inhibitors[56−62] have been developed, based on the core fragments of amyloid peptides. Many of these inhibitors are assumed to interact with the protofibrillar and fibrillar species through formation of the main-chain β-sheet hydrogen bonds. Presumably, their inhibition mechanism is to cap the β-sheet extension. In contrast, small molecule non-peptidic inhibitors have also been studied, including small aromatic molecules such as CR, rifampin, and others.[11,63−69] These non-peptidic molecules might exhibit different interaction patterns and inhibition mechanisms in comparison with the peptidic counterparts. Indeed, the primary interaction mode by CR observed in this study is along the β-sheet stacking direction rather than the β-sheet extension direction. Our previous study shows that the disruption of β-sheet stacking by mutating Phe to Ala could lead to formation of amorphous aggregates rather than amyloid fibril.[31] Disruption of β-sheet stacking by CR might have similar effects to inhibit amyloid fibril formation. Further study on the difference of the two types of inhibitors should provide more insight into the amyloid inhibition mechanism.

Conclusion

We have conducted a set of eight all-atom molecular dynamics simulations (20 ns of each) with explicit solvent representation on the binding of CR to the protofibrils of an amyloidogenic peptide GNNQQNY of the yeast prion protein Sup35. Independent from the initial positions and orientation of CR molecules, two β-sheet specific binding modes have been observed, consistent with the experiment measurements on a series of compounds that share a similar scaffold with CR.[49] The primary binding mode was along the β-sheet extension direction and occurred in a regular groove composed by the first three residues of the continuous β-strands in a long β-sheet. Because such long shallow grooves are common on the surface of amyloid fibril and are rare on protein surface, this binding mode could be the general recognition mode of amyloid fibrils by amyloid dyes. Simulations showed that TFT was mainly bound to the grooves on the antiparallel β-sheet surface of the KLVFFAE protofibril and the primary binding mode was highly similar to that of the CR-GNNQQNY complex, reaffirming the generality of the binding patterns. Analyses revealed that the hydrophobic interactions played an important role in this binding mode. CR molecules may reduce the ordered stacking between two β-sheets by changing the surface pattern via filling the surface grooves with their apolar parts and by improving the solubility of the protofibril via exposing the polar groups. The secondary mode was along the β-strands, in which a CR molecule was parallel to the β-strand and the bind site was either next to the ending β-strand or on top of a β-strand in the middle of a β-sheet. The secondary mode was less stable than the primary mode by 1.3−5.7 kcal/mol as estimated by population in clusters and by MM-GBSA calculation. The estimated binding free energy differences were also in qualitative agreement with the experimental data.[49]

Acknowledgment. Computer time was provided by the Pittsburgh Supercomputer Center (MCA06S027 to Y.D.), by the UC Davis Computer Science Department, and by the UC Davis Genome Center. This work was supported by research grants from the NIH (GM64458 and GM67168 to Y.D.). Use of graphics packages including VMD is gratefully acknowledged.

Supporting Information Available: Complete ref 38. This material is available free of charge via the Internet at http://pubs.acs.org.

JA0662772

(56) Kapurniotu, A.; Schmauder, A.; Tenidis, K. *J. Mol. Biol.* **2002**, *315*, 339−350.
(57) Tatarek-Nossol, M.; Yan, L. M.; Schmauder, A.; Tenidis, K.; Westermark, G.; Kapurniotu, A. *Chem. Biol.* **2005**, *12*, 797−809.
(58) Gilead, S.; Gazit, E. *Angew. Chem., Int. Ed.* **2004**, *43*, 4041−4044.
(59) Porat, Y.; Mazor, Y.; Efrat, S.; Gazit, E. *Biochemistry* **2004**, *43*, 14454−14462.
(60) Scrocchi, L. A.; Chen, Y.; Waschuk, S.; Wang, F.; Cheung, S.; Darabie, A. A.; McLaurin, J.; Fraser, P. E. *J. Mol. Biol.* **2002**, *318*, 697−706.
(61) Gordon, D. J.; Meredith, S. C. *Biochemistry* **2003**, *42*, 475−485.
(62) Gordon, D. J.; Tappe, R.; Meredith, S. C. *J. Pept. Res.* **2002**, *60*, 37−55.
(63) Aitken, J. F.; Loomes, K. M.; Konarkowska, B.; Cooper, G. J. S. *Biochem. J.* **2003**, *374*, 779−784.
(64) De Felice, F. G.; Vieira, M. N. N.; Saraiva, L. M.; Figueroa-Villar, J. D.; Garcia-Abreu, J.; Liu, R.; Chang, L.; Klein, W. L.; Ferreira, S. T. *FASEB J.* **2004**, *18*, 1366−1372.
(65) Lin, S. J.; Shiao, Y. J.; Chi, C. W.; Yang, L. M. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 1173−1176.
(66) Ono, K.; Yoshiike, Y.; Takashima, A.; Hasegawa, K.; Naiki, H.; Yamada, M. *Exp. Neurol.* **2004**, *189*, 380−392.
(67) Ono, K.; Hasegawa, K.; Naiki, H.; Yamada, M. *J. Neurosci. Res.* **2004**, *75*, 742−750.
(68) Ono, K.; Hirohata, M.; Yamada, M. *Biochem. Biophys. Res. Commun.* **2005**, *336*, 444−449.
(69) Tomiyama, T.; Kaneko, H.; Kataoka, K.; Asano, S.; Endo, N. *Biochem. J.* **1997**, *322*, 859−865.