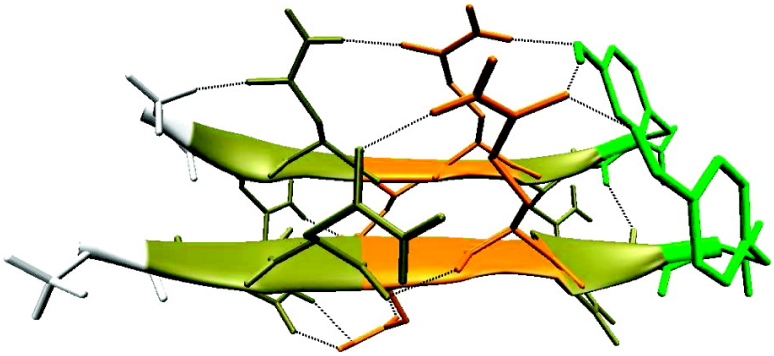
Article

Thermodynamics and Kinetics of Aggregation for the GNNQQNY Peptide

Birgit Strodel, Chris S. Whittleston, and David J. Wales

J. Am. Chem. Soc., **2007**, 129 (51), 16005-16014 • DOI: 10.1021/ja075346p



J|A|C|S
ARTICLES
Published on Web 12/04/2007

Thermodynamics and Kinetics of Aggregation for the GNNQQNY Peptide

Birgit Strodel, Chris S. Whittleston, and David J. Wales*

University Chemical Laboratories, Lensfield Road, Cambridge CB2 1EW, U.K.

Received July 18, 2007; E-mail: dw34@cam.ac.uk

Abstract: The energy landscape of the monomer and dimer are explored for the amyloidogenic heptapeptide GNNQQNY from the N-terminal prion-determining domain of the yeast protein Sup35. The peptide is modeled by a united-atom potential and an implicit solvent representation. Replica exchange molecular dynamics is used to explore the conformational space, and discrete path sampling is employed to investigate the pathways that interconvert the most populated minima on the free energy surfaces. For the monomer, we find a rapid fluctuation between four different conformations, where a geometry intermediate between compact and extended structures is the most thermodynamically favorable. The GNNQQNY dimer forms three stable sheet structures, namely in-register parallel, off-register parallel, and antiparallel. The antiparallel dimer is stabilized by strong electrostatic interactions resulting from interpeptide hydrogen bonds, which restrict its conformational flexibility. The in-register parallel dimer, which is close to the amyloid β-sheet structure, has fewer interpeptide hydrogen bonds, making hydrophobic interactions more important and increasing the conformational entropy compared to the antiparallel sheet. The estimated two-state rate constants indicate that the formation of dimers from monomers is fast and that the dimers are kinetically stable against dissociation at room temperature. Interconversions between the different dimers are feasible processes and are more likely than dissociation.

1. Introduction

Many human diseases arise as a consequence of protein misfolding, and the largest group of these is associated with protein aggregation.[1−3] Amyloid aggregates are known to be fibrillar in nature with a cross β-sheet structure, where the β-strands are perpendicular to the long fiber axis, with the backbone hydrogen-bonds stabilizing the sheets propagating along the direction of the fiber.[4−8] Although considerable progress has been made in recent years toward the elucidation of the structure and properties of amyloid fibrils, less is known about the structure and dynamics of the oligomers that are involved. The precise origin of pathogenicity in amyloid diseases remains elusive, although current evidence suggests that the soluble oligomeric precursors, rather than the fibrils themselves, are the cytotoxic species.[9−12]

The aim of the present work is to understand the very first step of oligomerization for the peptide GNNQQNY, namely dimerization, and to characterize the monomeric conformations that precede the formation of dimers. GNNQQNY is a polar heptapeptide from the N-terminal prion-determining region of the yeast prion protein Sup35, which exhibits similar amyloidogenic properties to the full-length Sup35.[13] The atomic resolution crystal structure of GNNQQNY has recently been determined by Eisenberg et al. using X-ray microcrystallography.[14] There is strong evidence that the microcrystals formed by amyloidogenic peptides or proteins are closely related to their amyloid fibrils.[15] The crystal structure of GNNQQNY suggests three levels of organization. First, the monomers align to give in-register parallel β-sheets, and then pairs of sheets bind to form a dry interface. The latter process is expected to be much slower due to the steric demands caused by interdigitation of the side-chains to form a "steric zipper".[14,15] Pairs of sheets then associate to form the fibril. The importance of the steric zipper interactions in stabilizing the GNNQQNY fibrils is supported by recent molecular dynamics (MD) simulations of several

(1) Selkoe, D. J. *Nature* **2003**, *426*, 900.
(2) Chiti, F.; Dobson, C. M. *Annu. Rev. Biochem.* **2006**, *75*, 333.
(3) Uversky, V. N.; Fink, A. L. *Protein misfolding, aggregation and conformational disease;* Springer: New York, 2006.
(4) Sipe, J. D.; Cohen, A. S. *J. Struct. Biol.* **2000**, *130*, 88−98.
(5) Eanes, E. D.; Glenner, G. G. *J. Histochem. Cytochem.* **1968**, *16*, 673−677.
(6) Geddes, A. J.; Parker, K. D.; Atkins, A. D.; Beighton, E. *J. Mol. Biol.* **1968**, *32*, 343−358.
(7) Sunde, M.; Serpell, L. C.; Bartlam, M.; Fraser, P. E.; Pepys, M. B.; Blake, C. C. F. *J. Mol. Biol.* **1997**, *273*, 729−739.
(8) Makin, O. S.; Serpell, L. C. *Fib. Diffr. Rev.* **2004**, *12*, 29−35.
(9) Walsh, D. M.; Klyubin, I.; Fadeeva, J. V.; Cullen, W. K.; Anwyl, R.; Wolfe, M. S.; Rowan, M. J.; Selkoe, D. J. *Nature* **2002**, *416*, 535−539.
(10) Bucciantini, M.; Giannoni, E.; Chiti, F.; Baroni, F.; Formigli, L.; Zurdo, J.; Taddei, N.; Ramponi, G.; Dobson, C. M.; Stefani, M. *Nature* **2002**, *416*, 507.
(11) Lee, E. B.; Leng, L. Z.; Zhang, B.; Kwong, L.; Trojanowski, J. Q.; Abel, T.; Lee, V. M.-Y. *J. Biol. Chem.* **2006**, *282*, 4292−4299.
(12) Lesné, S.; Teng Koh, M.; Kotilinek, L.; Kayed, R.; Glabe, C. G.; Yang, A.; Gallagher, M.; Ashe, K. H. *Nature* **2006**, *440*, 352−357.
(13) Balbirnie, M.; Grothe, R.; Eisenberg, D. S. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 2375−2380.
(14) Nelson, R.; Sawaya, M. R.; Balbirnie, M.; Madsen, A. O.; Riekel, C.; Grothe, R.; Eisenberg, D. *Nature* **2005**, *435*, 773−778.
(15) Sawaya, M. R.; Sambashivan, S.; Nelson, R.; Ivanova, M. I.; Sievers, S. A.; Apostol, M. I.; Thompson, M. J.; Balbirnie, M.; Wiltzius, J. J.; McFarlane, H. T.; Madsen, A. Ø.; Riekel, C.; Eisenberg, D. *Nature* **2007**, *447*, 453−457.

GNNQQNY assemblies of different sizes.[16] It was shown that, when equilibrated, the GNNQQNY protofibril twists while maintaining the steric zipper motif between β-sheets.

One of the first computational studies of GNNQQNY examined the behavior of the trimer using constant temperature MD at 330 K.[17] The in-register parallel sheet structures observed were stabilized by side-chain hydrogen-bonding and π-stacking of the aromatic tyrosine residues. Another MD simulation suggested that β-sheet formation is preceded by a condensation stage including disordered aggregates.[18] Reaction path annealing has been employed to simulate the coil $\rightarrow$ β-sheet and α-helix $\rightarrow$ β-sheet transition, both in the absence and presence of a cross-β amyloid nucleus.[19] The results indicate that the presence of a preformed amyloid-like aggregate significantly reduces the energy barrier for the transition to the β-sheet, suggesting that aggregation is autocatalysed by a nucleus, as expected for prion-like peptides. In the same study, it was also found that the cross-β amyloid structure is stabilized by the favorable orientation of backbone dipoles, i.e. head-to-tail between strands, but antialigned within strands and between sheets. A further MD study of a range of oligomer sizes investigated the stability of such aggregates and concluded that there is a marked increase in oligomer stability from the dimer to the trimer, suggesting that the critical nucleus may be as small as three or four peptides,[20] which is in agreement with experiment.[14] This conclusion is further supported by a recent MD simulation of the GNNQQNY peptide investigating the process of oligomer formation for the dimer, trimer and tetramer. In the latter study, various stable parallel and antiparallel β-sheets have been found for the dimer, trimer, and tetramer, but no short double layer has been reported.[21]

In the present work we restrict ourselves to the GNNQQNY monomer and dimer to investigate the very first step of β-sheet formation along with the structures that may serve as a nucleus for β-sheet elongation. Analysis of steric zipper formation will be considered in future work, because it would probably require a system containing at least five, and perhaps eight GNNQQNY peptides.[14,20,21] The importance of the monomeric structure for amyloid formation has been demonstrated in a recent study by Pellarin and Caflisch, who used a coarse-grained model with varying stability of the β-prone state to investigate the kinetics and pathways of fibril formation.[22] They found that different β-propensities give rise to different aggregation mechanisms, namely a disordered aggregation pathway for the β-unstable state; fibril formation via oligomeric on-pathway intermediates for a more stable β-state; and fibril formation without intermediates in the case of a very β-stable state.[22] Although experiments do not indicate a β-strand for the monomer,[13] it was suggested that GNNQQNY belongs to the class of β-stable peptides, and thus aggregation should proceed via fibrillogenesis, rather than disordered aggregation followed by reorganization of the oligomer.[22]

One aim of the present work is to test the above hypothesis for GNNQQNY using an atomistic representation and an implicit solvent potential. We employ replica exchange molecular dynamics (REMD)[23] to survey the free-energy landscape of the monomer and dimer and discrete path sampling (DPS)[24−27] to analyze the kinetics of selected processes. For the monomer, we investigate the thermodynamic and kinetic stability of the β-strand. For the GNNQQNY dimer we find three different β-sheets from REMD. These structures are analyzed in detail in terms of thermodynamic stability conferred by favorable side chain interactions. We then employ the DPS method to investigate the dimerization pathways leading from the monomer to either of the three stable dimers identified from REMD. We also check how stable these dimeric structures are with respect to direct interconversion.

2. Methods

The GNNQQNY peptide (NH_3^+-Gly-Asn-Asn-Gln-Gln-Asn-Tyr-COO^-) was represented by a united-atom force field, CHARMM19,[28] and an implicit solvation potential, EEF1.[29]

2.1. Replica Exchange Molecular Dynamics. In a preliminary exploration of the conformational space, we used REMD,[23] as implemented in the MMTSB tool set.[30] Within the REMD protocol, multiple simulations of the same system (replicas) are run simultaneously at different temperatures. Every τ time steps an attempt is made to swap temperatures between two different replicas i and j. The exchanges are accepted with probability $\min\{1, \exp(-\Delta)\}$, where $\Delta = (1/k_B T_i - 1/k_B T_j)(E_j - E_i)$, and k_B is the Boltzmann constant, T_i and E_i denote the temperature and potential energy of replica i, respectively, and j is typically replica $i + 1$ at the same time step. The temperatures for the replicas were exponentially spaced between a minimum value T_{min} and a maximum value T_{max}. This procedure allows for an improved sampling of the conformational space at low temperatures, because crossings of potential energy barriers are facilitated at higher temperatures, and the resulting conformational changes migrate into the lower T replicas.[31] The dynamics were propagated using the Langevin method, with a time step of $\delta t = 1$ fs and a friction coefficient of $\zeta = 20$ ps^{-1}.

For the GNNQQNY dimer we confine the two peptides by repulsive walls to model constant concentration conditions using a spherical container as developed by Klimov et al.[32] The confining potential experienced by an atom located at $\mathbf{r}$ is then

$$V_c(r) = \frac{4\pi\epsilon R}{5r}\left[\left(\frac{\sigma}{r-R}\right)^{10} - \left(\frac{\sigma}{r+R}\right)^{10}\right], \quad (1)$$

where $r = |\mathbf{r}|$ and $\mathbf{r} = \mathbf{0}$ is the center of the sphere with radius R. The parameters σ and ϵ were chosen to be 1 Å and 1 kcal mol^{-1}, respectively.[33] The potential (1) guarantees short-range repulsive interactions between peptide atoms and the inner walls of the container. The container radius was chosen to be $R = 43$ Å, which corresponds to a molecular concentration of 10 mM. This concentration is only

(16) Esposito, L.; Pedone, C.; Vitagliano, L. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 11533−11538.
(17) Gsponer, J.; Haberthür, U.; Caflisch, A. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 5154−5159.
(18) Cecchini, M.; Rao, F.; Seeber, M.; Caflisch, A. *J. Chem. Phys.* **2004**, *121*, 10748.
(19) Lipfert, J.; Franklin, J.; Wu, F.; Doniach, S. *J. Mol. Biol.* **2005**, *349*, 648−658.
(20) Zheng, J.; Ma, B.; Tsai, C.-J.; Nussinov, R. *Biophys. J.* **2006**, *91*, 824−833.
(21) Zhang, Z.; Chen, H.; Bai, H.; Lai, L. *Biophys. J.* **2007**, *93*, 1484−1492.
(22) Pellarin, R.; Caflisch, A. *J. Mol. Biol.* **2006**, *360*, 882−892.
(23) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **1999**, *314*, 141−151.
(24) Wales, D. J. *Mol. Phys.* **2002**, *100*, 3285−3306.
(25) Wales, D. J. *Energy Landscapes*; Cambridge University Press: Cambridge, 2003.
(26) Wales, D. J. *Mol. Phys.* **2004**, *102*, 891−908.
(27) Wales, D. J. *Int. Rev. Phys. Chem.* **2006**, *25*, 237−282.
(28) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comp. Chem.* **1983**, *4*, 187−217.
(29) Lazaridis, T.; Karplus, M. *Proteins: Struct., Funct., Genet.* **1999**, *35*, 133−152.
(30) Feig, M.; Karanicolas, J.; Brooks, C. L. *J. Mol. Graphics Modell.* **2004**, *22*, 377−395.
(31) Swendsen, R. H.; Wang, J.-S. *Phys. Rev. Lett.* **1986**, *57*, 2607−2609.
(32) Klimov, D. K.; Newfield, D.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 8019−8024.
(33) Baumketner, A.; Shea, J.-E. *Biophys. J.* **2005**, *89*, 1493−1503.

about 1 order of magnitude higher than the typical concentrations of experiments in which the formation of amyloid fibrils are investigated. We have verified that the conformational space explored during the dynamics simulation is not affected by the confining sphere.

Several order parameters, such as the potential energy and radius of gyration, were collected during the REMD simulations and analyzed using the weighted histogram analysis method (WHAM).[34] The WHAM methodology combines the data from all replicas to extract ensemble averages and calculates the heat capacity and free energy at a given temperature as a function of suitable order parameters.

2.2. Discrete Path Sampling. Having obtained a representation of the free energy landscape from REMD, the next step was to characterize pathways between populated conformations and to estimate the corresponding rate constants. Our approach here is the DPS methodology, which provides a systematic way to produce stationary point databases designed for analyzing kinetics.[24−27]

Details of the DPS approach can be found in a recent review.[27] In brief, a discrete path is here defined as a sequence of local minima together with the transition states that connect them on the potential energy surface. A local minimum is a stationary point where all the nonzero normal-mode frequencies are positive, whereas a transition state has a single negative Hessian eigenvalue,[35] which corresponds to the reaction coordinate. The connectivity of a transition state is then defined by the two minima reached by (approximate) steepest-descent paths commencing parallel and antiparallel to the reaction coordinate. Local minima were obtained using a modified version of the LBFGS procedure described by Liu and Nocedal.[36] The doubly nudged[37] elastic band[38−40] (DNEB) algorithm was used to produce initial transition state guesses to connect pairs of local minima. The resulting transition state candidates were then tightly converged using hybrid eigenvector-following.[41−43] The Dijkstra missing connection algorithm was used to select pairs of minima for DNEB calculations.[44]

The DPS procedure starts from an initial discrete path between states *A* and *B*, where *A* and *B* denote the reactant and product. An order parameter is required to classify local minima as *A*, *B*, or *I* in character, where the *I* (intervening) set contains all the minima outside the *A* and *B* regions. From the initial path, the stationary point database is grown by addition of minima and transition states from successive connection attempts for pairs of local minima already in the database. The principal difference between the present calculations and previous work lies in how these pairs of local minima were selected, as described below.

Assuming that the dynamics between adjacent local minima are Markovian, the $B \leftarrow A$ rate constant within the steady-state approximation for the intervening minima can be written as a sum over all discrete paths.[24,25,27] The resulting quantity, denoted k_{BA}^{SS}, is not the most accurate formula for the rate constant, but it provides a convenient way to select pairs of local minima for subsequent connection attempts. The discrete paths that make the largest contributions to k_{BA}^{SS} can be calculated[45] using Dijkstra's algorithm.[46] For brevity, we will refer to these discrete paths as the "fastest" paths, although it is important to note that the analysis includes the conditional occupation probability of the reactant minimum, and all the relevant branching probabilities. Hence, this approach does not suffer from problems that might be experienced using a framework based on multiple short MD trajectories.[47−51]

In previous work,[45,52] a maximum number of pairs of local minima were selected from a maximum number of discrete paths that made the largest contributions to k_{BA}^{SS}. In the present contribution, we applied a Dijkstra analysis to the stationary point database on-the-fly, as suggested by Evans and Wales.[45] Having determined the discrete path that makes the largest contribution to k_{BA}^{SS}, we then identified the largest potential or free energy barriers. We also ranked the local minima from this path that are not directly connected by a transition state according to the minimal distance between them.[53] Several schemes for growing the stationary point databases were then considered. The first SHORTCUT scheme selects minima from the fastest path according to their minimized Euclidean distance with respect to overall translation and rotation.[53] The pairs were also required to be separated by a minimum number of steps on the fastest path. The second scheme, referred to as SHORTCUT BARRIER, selects pairs of minima on either side of the largest potential or free energy barriers on the fastest path. The pairs were chosen to be an equal number of steps away from the barrier, up to a maximum value. This scheme provides an efficient way to find alternative paths that avoid the highest barriers, while the SHORTCUT scheme produces discrete paths with a smaller number of steps.

Growing a stationary point database according to either SHORTCUT approach will generally lead to an improvement in the overall rate constant if we only consider the fastest paths. However, it is likely to introduce kinetic traps into the database in the form of local minima that are found during the connection attempts, but lie behind high barriers. Such traps are artefacts of incomplete sampling, because low barrier paths to the rest of the database exist but have not been found. Furthermore, they lead to a reduction in the true two-state rate constants, which are affected by the slowest relaxation time in the database. To reflect the true global kinetics, we must also refine the database to remove such artificial traps. We have considered several ways to achieve this refinement. In the DIJPAIR scheme, we perform a Dijkstra analysis for all intervening minima as well as the reactant set. Candidate pairs are then chosen according to the highest barriers on these discrete paths, as for SHORTCUT BARRIER. However, the computational cost of the DIJPAIR scheme becomes prohibitive for larger databases. Instead, we therefore considered an approach based on barrier thresholds determined from a disconnectivity graph analysis.[54−57] The latter scheme, referred to as UNTRAP, was implemented by sorting first on the maximum potential or free energy barriers, and then on the minimum distance between unconnected minima in different superbasins. The final databases were produced using a combination of the above schemes.

For simplicity, we employed the harmonic approximation to calculate the local density of vibrational states for each stationary point. In principle, any consistent scheme for the densities of states and minimum-to-minimum rate constants could be used. However, the harmonic approximation seems to be appropriate given the uncertainties in the empirical potentials employed. For example, a recent study found that the harmonic transition state theory rate constants only differed by a factor of 2 from the anharmonic values for hydrogen abstraction

(34) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comp. Chem.* **1992**, *13*, 1011−1021.
(35) Murrell, J. N.; Laidler, K. J. *Trans. Faraday Soc.* **1968**, *64*, 371.
(36) Liu, D.; Nocedal, J. *Math. Program* **1989**, *45*, 503−528.
(37) Trygubenko, S. A.; Wales, D. J. *J. Chem. Phys.* **2004**, *120*, 2082.
(38) Henkelman, G.; Jónsson, H. *J. Chem. Phys.* **1999**, *111*, 7010−7022.
(39) Henkelman, G.; Uberuaga, B. P.; Jónsson, H. *J. Chem. Phys.* **2000**, *113*, 9901−9904.
(40) Henkelman, G.; Jónsson, H. *J. Chem. Phys.* **2000**, *113*, 9978−9985.
(41) Hildebrand, F. B. *Methods of Applied Mathematics*; Dover: New York, 1992.
(42) Munro, L. J.; Wales, D. J. *Phys. Rev. B* **1999**, *59*, 3969−3980.
(43) Kumeda, Y.; Munro, L. J.; Wales, D. J. *Chem. Phys. Lett.* **2001**, *341*, 185−194.
(44) Carr, J. M.; Trygubenko, S. A.; Wales, D. J. *J. Chem. Phys.* **2005**, *122*, 234903.
(45) Evans, D. A.; Wales, D. J. *J. Chem. Phys.* **2004**, *121*, 1080−1090.
(46) Dijkstra, E. W. *Numerische Math.* **1959**, *1*, 269−271.
(47) Snow, C. D.; Zagrovic, B.; Pande, V. S. *J. Am. Chem. Soc.* **2002**, *124*, 14548−14549.
(48) Snow, C. D.; Nguyen, H.; Pande, V.; Gruebele, M. *Nature* **2002**, *420*, 102−106.
(49) Snow, C. D.; Qiu, L.; Gai, F.; Hagen, S. J.; Pande, V. S. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 4077−4082.
(50) Fersht, A. R. *Proc. Nat. Acad. Sci. U.S.A.* **2002**, *99*, 14122−14125.
(51) Marianayagam, N. J.; Fawzi, N. L.; Head-Gordon, T. *Proc. Nat. Acad. Sci. U.S.A.* **2005**, *102*, 16684−16689.
(52) Carr, J. M.; Wales, D. J. *J. Chem. Phys.* **2005**, *123*, 234901.
(53) Kearsley, S. K. *Acta Cryst. A* **1989**, *45*, 208−210.
(54) Becker, O. M.; Karplus, M. *J. Chem. Phys.* **1997**, *106*, 1495.
(55) Wales, D. J.; Miller, M. A.; Walsh, T. R. *Nature* **1998**, *394*, 758.
(56) Krivov, S. V.; Karplus, M. *J. Chem. Phys.* **2002**, *117*, 10894−10903.
(57) Evans, D. A.; Wales, D. J. *J. Chem. Phys.* **2003**, *118*, 3891−3897.

from ethane.[58] The superposition approach was used to express the total partition function in terms of contributions from individual local minima, and standard transition state theory expressions were employed for the elementary rate constants between connected minima, as in previous work.[24,25,27]

Phenomenological two-state rate constants, k_{BA} and k_{AB}, were extracted from the stationary point databases using the graph transformation (GT) approach.[27,59,60] This method allows us to relax the steady-state approximation for intervening local minima and the assumption of local equilibrium within the *A* and *B* regions.[59] The GT approach can be used to calculate mean first passage times (MFPT's), and hence rate constants, that are equivalent to solving the master equation[61] for the time-dependent occupation probabilities of the local minima, or running kinetic Monte Carlo simulations.[27,60] We emphasize that these rate constants are formally exact within the Markov approximation for minimum-to-minimum transitions, given a complete database, exact densities of states for local minima, and correct rate constants for transitions between adjacent minima. Our additional approximations are therefore finite sampling, harmonic densities of states, and transition state theory. The fundamental Markov assumption is most likely to break down for minima separated by low barriers. The rate constants for these transitions may therefore be underestimated, but these are not likely to be rate-determining steps. Hence, the largest errors are probably related to the empirical potential employed.

3. Results

3.1. Thermodynamics and Kinetics of the Monomer. A REMD simulation with eight replicas exponentially distributed between 250 and 400 K was performed for the GNNQQNY monomer. The run consisted of 50 000 replica exchange cycles, each involving 2000 MD steps, resulting in a total simulation time of 100 ns. The acceptance ratio was around 0.5 throughout the considered temperature range. The simulation was started from a random structure for each replica, selected from a previous high-temperature MD run.

Before data were collected, 10 simulation cycles were applied for initialization with no exchanges, allowing each replica to equilibrate at the temperature it was assigned. At intervals of 1 ps, we saved the potential energy of the peptide, the radius of gyration, R_g, and the root-mean-square deviation, rmsd, based on the C_α positions of the β-strand shown in Figure 1a and the candidate geometry.

Monomers Prefer Coil Conformations. In Figure 2, the free energy surface (FES) for the monomer at 298 K is displayed as a function of R_g and rmsd. The minima can be divided into compact structures with $R_g < 4$ Å, extended structures with $R_g > 5$ Å, and coil structures for $4 \lesssim R_g \lesssim 5$ Å. The compact conformations are dominated by helix-like structures (region H) and geometries with a turn (region T). In agreement with experiment[13] and previous simulations,[19] we find that the coil structures are the highest populated, with the most prominent free energy minimum at $R_g = 4.5$ Å and rmsd = 3.8 Å (region C). The extended geometries in region E are strand-like with some structures having an rmsd value less than 2 Å from the β-strand.

The potential energy, *V*, and the conformational entropy, *S*, were compared within the harmonic approximation at $T =$ 298 K for some randomly selected structures belonging to each of the free energy minima (Supporting Information). The results show that the compact structures in regions T and H have a low potential energy but are stiff, leading to relatively high harmonic free energies. Conversely, in region E, the structures are floppy but have high potential energies. Region C has the best balance of low potential energy and soft normal modes, leading to the lowest harmonic free energies, which compares well with the REMD results.

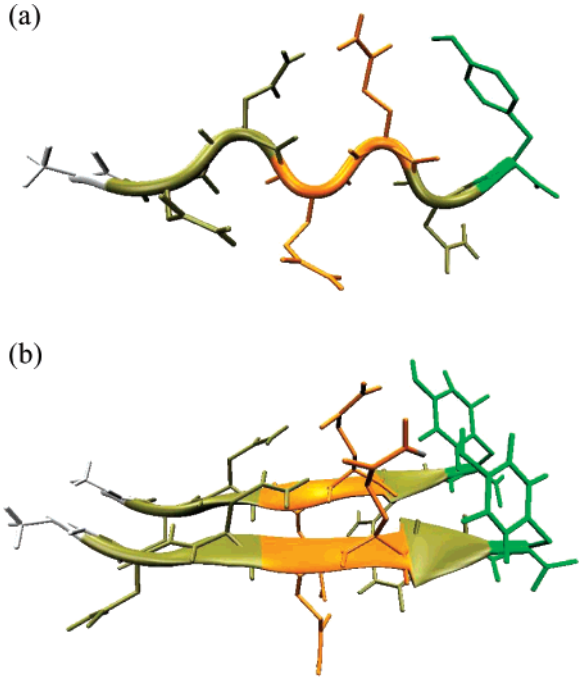


Figure 1. Structure of GNNQQNY in (a) a β-strand for the monomer and (b) a β-sheet for the dimer in-register parallel configuration.

Monomers Fluctuate Rapidly between Different Conformations. The barriers between the free energy minima in Figure 2 are low (around 2 kcal mol^{-1}), suggesting fast transitions between conformations. Using the DPS approach, two-state rate constants for transitions between regions T and C, H and C, and C and E were calculated. The final database of stationary points consisted of 1024 minima and 1542 transition states. All MFPT's at 298 K were found to lie in the nanosecond regime, confirming that the conformational changes for the GNNQQNY monomer are fast processes and would probably not be the rate-limiting step for the formation of amyloid fibrils if GNNQQNY follows the fibrillogenesis pathway. The entropic driving force toward an extended structure for the GNNQQNY monomer is a sign of its β-propensity, suggesting that GNNQQNY indeed belongs to the class of amyloidogenic peptides that form fibrils via fibrillogenesis rather than disordered aggregation, in agreement with previous work.[22]

3.2. Conformational Space of the Dimer. To further examine whether or not GNNQQNY forms protofibrils via fibrillogenesis, we investigated the formation of dimers as the first step of oligomerization. For the REMD simulation, 12 replicas between 270 and 500 K were considered. The total simulation time was 100 ns for 50 000 replica exchange cycles (excluding initialization and equilibration cycles) with 2000 MD time steps per cycle. To ensure that the simulation was not biased by a particular starting conformation, we initiated each replica by placing two peptides in random orientations. The

(58) Sturdy, Y. K.; Clary, D. C. *Phys. Chem. Chem. Phys.* **2007**, *9*, 2397−2405.
(59) Trygubenko, S. A.; Wales, D. J. *Mol. Phys.* **2006**, *104*, 1497−1507.
(60) Trygubenko, S. A.; Wales, D. J. *J. Chem. Phys.* **2006**, *124*, 234110.
(61) van Kampen, N. G. *Stochastic processes in physics and chemistry*; North-Holland: Amsterdam, 1981.

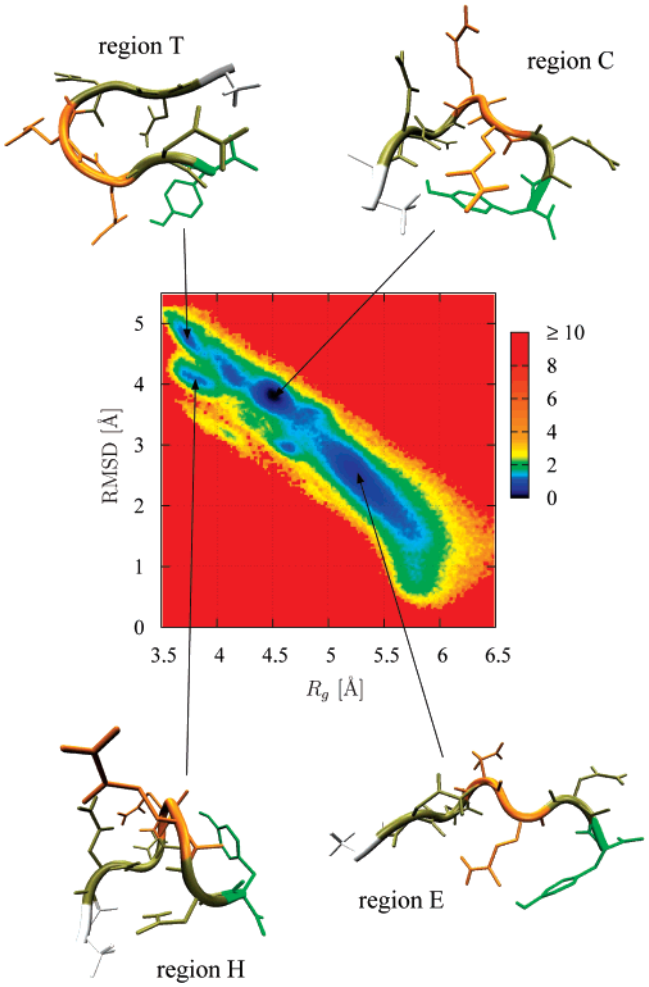


Figure 2. Free energy surface at 298 K for the GNNQQNY monomer in terms of the radius of gyration, R_g, and C_α-rmsd from the β-strand. The free energy scale (in kcal mol^{-1}) is given on the right. Typical structures for the free energy minima are shown.

acceptance ratio for the exchanges was seen to be uniform in the considered temperature range, with values between 0.35 and 0.5. The principal quantities calculated every 1 ps during the REMD simulation are the potential energy, the monomer−monomer interaction energy, E_{int}, consisting of electrostatic, $E_{\text{int}}^{\text{el}}$, and van der Waals, $E_{\text{int}}^{\text{vdW}}$, interactions, the radius of gyration, and the C_α-rmsd from the β-sheet. The β-sheet was constructed based on theoretical results for the amyloid structure of GNNQQNY,[16] and the equilibrated structure is shown in Figure 1b. The convergence of this simulation was tested by another REMD simulation under identical conditions, but each replica was started from the β-sheet in Figure 1b. This simulation yielded the same free energy surface and heat capacity as the one started from random structures.

The heat capacity of the system, C_v, calculated from REMD is plotted together with R_g as a function of temperature in Figure 3. For temperatures above 450 K, the system occupies conformations with $R_g \approx 25$ Å. At such high temperatures, the system consists of two monomers rather than a dimer, and the distance between the two peptides is controlled mainly by the confining sphere. As the temperature decreases, the distance between the peptides starts to diminish, driven by favorable intermolecular interactions, reaching $R_g \approx 6-7$ Å for dimeric conformations at temperatures below 380 K. The resulting heat capacity curve

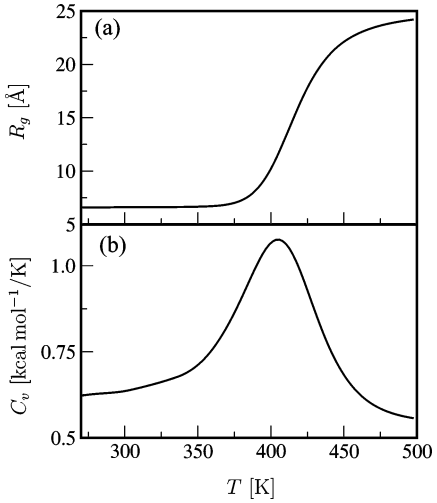


Figure 3. Temperature dependence of (a) the radius of gyration, R_g, and (b) the heat capacity, C_v, computed from replica exchange simulations for the GNNQQNY dimer.

in Figure 3b peaks at $T_d = 405$ K, which we will refer to as the dissociation temperature of the dimer. This high melting temperature for the dimer might be an artifact of the empirical potential and is observed for other potentials as well.[62] It would be interesting to compare these results with accurate calculations involving explicit treatments of electronic structure in future work.

GNNQQNY Dimers Adopt Three Stable β-Sheet Conformations. The FES at 298 K is plotted as a function of the radius of gyration and rmsd in Figure 4. It shows three pronounced free energy minima corresponding to different types of β-sheet structures, which involve both parallel and antiparallel alignments of the strands, in agreement with previous simulations.[17,21] One of these β-sheets is close to the amyloid structure in Figure 1b with rmsd values of 1.0 $\lesssim$ rmsd $\lesssim$ 1.5 Å. The difference between the amyloid structure and the in-register parallel conformation, IP, found in the simulation is the twist, which is not present in the IP sheet. Another parallel β-sheet, OP, occurs at 2.0 $\lesssim$ rmsd $\lesssim$ 2.5 Å, where the two strands are off-register by one residue. The structure at rmsd $\approx$ 8 Å is an antiparallel β-sheet, IA, where the two strands are somewhat twisted around each other.

Antiparallel β-Sheet is Stabilized by Hydrogen-Bonds. To further characterize the three dimer conformations, the interaction energy and hydrogen-bonds between monomers were examined in more detail. Typical structures from the minima of the FES were extracted and minimized, the electrostatic and van der Waals (vdW) interactions were calculated, and the existence of hydrogen-bonds was checked for all pairs of residues between the two peptides. For the identification of hydrogen-bonds, a distance cutoff of 3.2 Å was used with a maximum allowed deviation from linearity of 40.0°. The results are summarized in Table 1 and represented in more detail as residue−residue maps in the Supporting Information. The vdW energies mainly reflect the hydrophobic interactions, and the values in Table 1 reveal that the vdW energies are similar for the three dimers with attractive interactions in the range of $\sim$−40 kcal mol^{-1}. However, the stabilization originating from electrostatic forces varies significantly between the dimer

(62) Meinke, J. H.; Hansmann, U. H. E. *J. Chem. Phys.* **2007**, *126*, 014706.

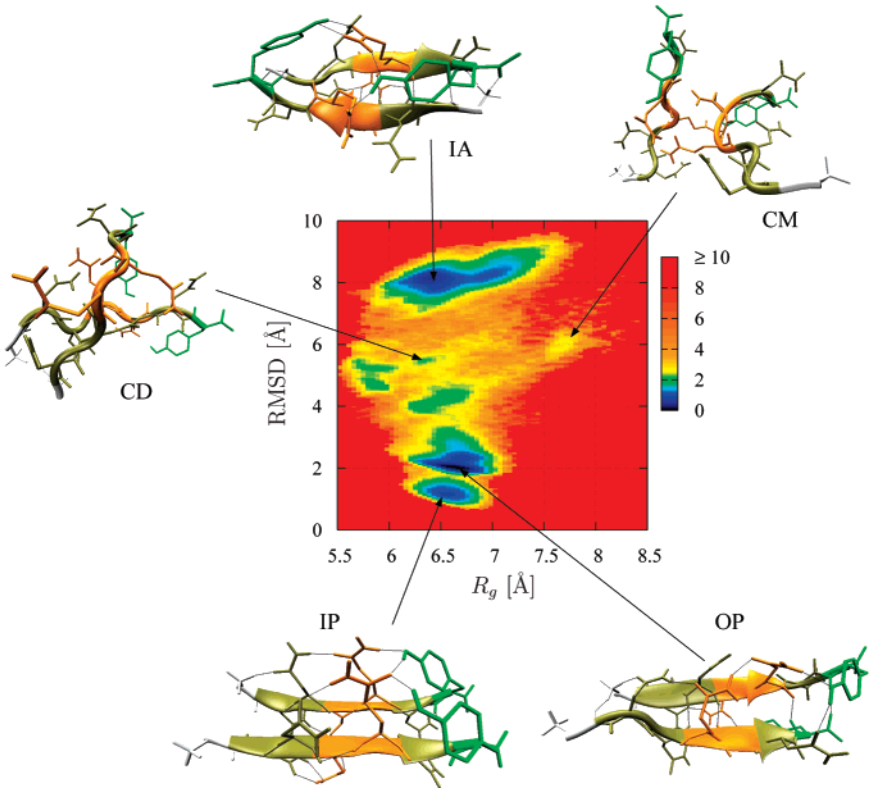


Figure 4. Free energy surface at 298 K for the GNNQQNY dimer in terms of the radius of gyration, R_g, and C_α-rmsd from the amyloid β-sheet. The free-energy scale (in kcal mol^{-1}) is given on the right. Typical structures for the free energy minima are shown. The hydrogen-bonds are depicted as black dotted lines for the three sheet structures IP, OP, and IA.

Table 1. Peptide−Peptide Interaction Energies and Interpeptide Hydrogen Bonds for GNNQQNY Dimers[a]

dimer	V	no. of H bonds	E_{int}^{tot}	E_{int}^{elec}	E_{int}^{vdW}
IA	−639.49	16	−130.06	−91.74	−38.33
OP	−633.57	13	−115.54	−74.81	−40.73
IP	−629.91	7	−84.69	−46.16	−38.58

[a] All energies are in kcal mol^{-1}.

conformations. For the IA dimer, the attraction between the two peptides is dominated by electrostatic interactions, whereas the IP geometry is almost stabilized to the same extent by electrostatic and vdW interactions. The electrostatic interpeptide attraction leads to a potential energy ordering of V(IA) < V(OP) < V(IP). A comparison between the patterns for the electrostatic residue−residue interaction and the residue−residue hydrogen bonds shows that the electrostatic attraction mainly originates from the formation of hydrogen-bonds. For each of the dimers, the hydrogen-bonds are similarly distributed between backbones and side chains. The twist in the IA dimer allows for the formation of 16 interpeptide hydrogen-bonds, that is, more than two per residue, but we expect that this twisted antiparallel geometry will be unable to participate in fibril formation because docking of a third strand appears structurally unfavorable.

In-Register Parallel β-Sheet is Stabilized by Hydrogen-Bonds and Hydrophobic Interactions. The electrostatic interactions are characterized by some strong residue−residue attractions, $E_{int}^{el} \lesssim -10$ kcal mol^{-1}, but also some repulsions with $E_{int}^{el} \approx +2$ kcal mol^{-1}, whereas the majority of the vdW interactions are in the range of $-2.5 \lesssim E_{int}^{vdW} < 0$ kcal mol^{-1}. The only deviation from this trend is seen for the Tyr-Tyr interactions in IP with $E_{int}^{vdW} = -5.1$ kcal mol^{-1} and OP with $E_{int}^{vdW} = -3.9$ kcal mol^{-1}, as well as for the Tyr-Asn interaction in the OP sheet with $E_{int}^{vdW} = -3.9$ kcal mol^{-1}. It has been suggested that the π-stacking interactions between tyrosine rings are the major source of stability for the in-register parallel aggregates for GNNQQNY.[17,20] However, our comparison of the IP and OP sheets indicates that the Tyr-Tyr interactions are similar, despite the aromatic rings being on opposite faces of the OP sheet and separated by more than 8 Å. Hence, hydrophobic forces aside from π stacking can be as important for the stabilization of peptide aggregates, at least with the present potential. Furthermore, from *ab initio* calculations it is known that for a maximum $\pi-\pi$ interaction to occur, the optimal distance between the aromatic planes in a parallel-displaced configuration is 3.4 Å, giving rise to a binding energy of 2.4−2.7 kcal mol^{-1}.[63] Inspection of the configuration of the two Tyr rings in the IP sheet reveals that they are parallel-displaced with a separation of 4−5 Å, which does not allow for an optimum attractive $\pi-\pi$ interaction. Our results are supported by a recent MD study of the GNNQQNY dimer employing the GROMACS96[64] force field and an explicit water description, which finds that π stacking does not give directionality to the self-assembly process but stabilizes the parallel aggregates.[21] The stabilization was not quantified in the latter study but inferred from geometrical arguments.

Overall, for the present potential, hydrophobic interactions are as important as electrostatic interactions for the formation

(63) Sinnokrot, M. O.; Valeev, E. F.; Sherrill, C. D. *J. Am. Chem. Soc.* **2002**, *124*, 10887−10893.
(64) van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hünenberger, P. H.; Krüger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. *Biomolecular Simulation: The GROMOS96 Manual and User Guide*; Vdf Hochschulverlag AG an der ETH Zürich: Zürich, 1996.

Table 2. Estimated Rate Constants at 298 K for the Aggregation and Dissociation of the GNNQQNY Dimer as Well as for the Interconversion between Different Dimer Structures

$A \leftrightarrow B$	k_{BA}/s^{-1}	k_{AB}/s^{-1}
CM ↔ IP	1.6×10^{-1}	1.6×10^{-11}
CM ↔ OP	2.2×10^{3}	1.7×10^{-8}
CM ↔ IA	3.7×10^{-2}	1.1×10^{-11}
IP ↔ OP	5.8×10^{-7}	1.8×10^{-5}
IP ↔ CD	7.6×10^{-2}	2.0×10^{3}
OP ↔ CD	1.8×10^{-5}	4.4×10^{-6}
CD ↔ IA	9.5×10^{-5}	1.3×10^{-6}

of the IP sheet of GNNQQNY and they originate from aromatic, that is, $\pi-\pi$ stacking, and aliphatic interactions between the residues.

3.3. Kinetics of Dimerization. From the FES for the GNNQQNY dimer in Figure 4 it appears that the transition from the monomeric to the dimeric phase is a downhill process leading to one of the three dimer structures. Furthermore, it seems that the free energy barrier between the different dimers is relatively small (~3 kcal mol^{-1}), suggesting that direct conversion between them would be fast. To investigate this suggestion in more detail, and to quantify the kinetics of dimerization, several discrete path sampling runs were carried out to calculate two-state rate constants for processes involving the IP, OP, and IA conformations. As a starting structure for the aggregation pathway, a populated conformation (CM) with two monomers was chosen from the FES. To investigate the interconversion between the parallel and antiparallel dimers, an intermediate structure (CD) was considered where the peptide strands are approximately orthogonal to each other. The two structures CM and CD are shown in Figure 4 together with their location on the FES for the dimeric system.

The following pathways were considered: CM ↔ IP, CM ↔ OP, CM ↔ IA, IP ↔ OP, IP ↔ IA via CD, and OP ↔ IA via CD. The initial discrete paths for each case contained high barriers, which were easily removed by growing the DPS stationary point database as outlined in section 2.2. The DPS runs were deemed to be converged if the rate constants k_{BA}^{GT} and k_{AB}^{GT} were stable to within an order of magnitude. The final database of stationary points, containing 28 631 minima and 38 319 transition states, was used for the following analyses.

Association is Fast, Dissociation is Slow. In Table 2, the final rate constants at 298 K for all paths are summarized. The rate constants indicate that the three dimers IP, OP, and IA are stable against dissociation at 298 K. The estimated rate constants for the association of monomers leading to either of the three dimeric structures are about 10 orders of magnitude higher than their corresponding dissociation rates. The fastest pathways obtained by placing intervening minima in steady state (see section 2.2) are plotted in Figure 5 in terms of energy and integrated path length. The pathway starting from two monomers and leading to the OP dimer is almost completely downhill (Figure 5b) and does not involve any high potential or free-energy barriers. Hence, it is associated with a large rate constant. Both pathways in Figure 5a and c leading to the IP and IA dimers, respectively, are also mainly downhill but involve some potential energy barriers in the range of 7−10 kcal mol^{-1} , which give rise to a time scale of seconds rather than microseconds, as for the formation of the OP sheet. Analysis of these high barriers revealed that they originate from the breaking of several

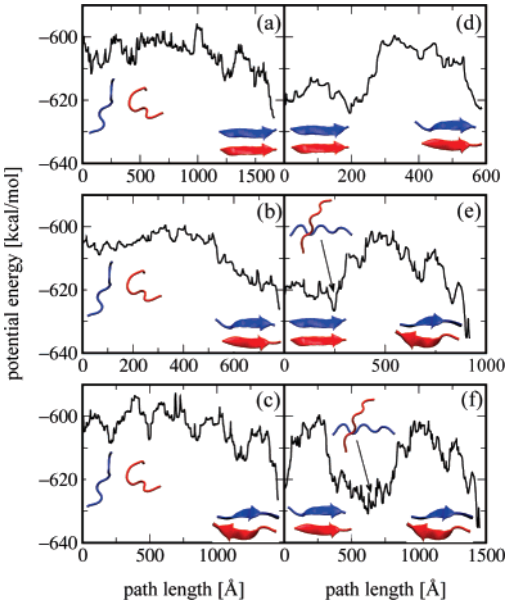


Figure 5. Fastest paths for the formation of a series of GNNQQNY dimer structures obtained from a discrete path sampling database. (a) CM → IP (b) CM → OP (c) CM → IA (d) IP → OP (e) IP → IA (f) OP → IA. Here, "fastest" paths are rated according to their contribution to the two-state rate constant calculated within a steady-state approximation for intervening minima, as explained in section 2.2.

hydrogen-bonds at once to allow for the rearrangement of the side chains and formation of new hydrogen-bonds.

The lack of any long-lived off- or on-pathway intermediate for the formation of dimer sheets, together with the higher rate constants for the processes of association and lower rate constants for the processes of dissociation, supports the idea that GNNQQNY forms amyloid fibrils via fibrillogenesis.[22] The association of GNNQQNY monomers to yield a dimeric β-sheet is also in agreement with the dock-lock mechanism, which has been suggested to explain the kinetics of fibrillogenesis for the Aβ peptide[65−67] and Sup35,[68] where monomers add to a pre-existing fibril. Here, the elongation of fibrils appears to consist of a fast, reversible "dock" phase of the monomer to the fibril, whereas in the slower second "lock" stage the deposited peptide becomes irreversibly associated with the fibril.[69] In the first, mainly diffusion-controlled step of the GNNQQNY dimerization, the two monomers form a loosely bound dimer driven by favorable interpeptide interactions. To maximize the interpeptide contact radius, the dimer is formed from extended monomers,[70] so that the early dimer is already sheet-like. In the second, rate-determining phase, the conformational space of the dimer is explored to find the optimal side chain arrangement. In accord with the dock−lock mechanism, dissociation is then slow at room temperature.

(65) Esler, W. P.; Stimson, E. R.; Jennings, J. M.; Vinters, H. V.; Ghilardi, J. R.; Lee, J. P.; Mantyh, P. W.; Maggio, J. E. *Biochemistry* **2000**, *39*, 6288−6295.
(66) Cannon, M. J.; Williams, A. D.; Wetzel, R.; Myszka, D. G. *Anal. Biochem.* **2004**, *328*, 67−75.
(67) Nguyen, P. H.; Li, M. L.; Stock, G.; Straub, J. E.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 111−116.
(68) Collins, S. R.; Douglass, A.; Vale, R. D.; Weissman, J. S. *PLoS Biol.* **2004**, *2*, 1582−1590.
(69) Massi, F.; Straub, J. E. *Proteins: Struct., Funct., Genet.* **2001**, *42*, 217−229.
(70) Dima, R. I.; Thirumalai, D. *Protein Sci.* **2002**, *11*, 1036−1049.

Reptation Move for Interconversion between Parallel Sheets. We also investigated whether direct interconversion between the low-energy dimer structures is likely. The rate constants in Table 2 suggest that these processes are slow, but more likely than dissociation.

The IP ↔ OP conversion is a so-called "reptation move" of one strand of the β-sheet with respect to each other. A reptation move was reported as a possible folding pathway for the second β-hairpin of the domain B1 of protein G.[71] Later, it was also discussed as a mechanism for oligomeric growth of the amyloidogenic peptide Aβ_{16-22}.[72−74] Although the reptation move requires only the movement of one of the peptides by one residue to switch between the in- and off-register alignments, and the C_α-rmsd between the IP and OP geometries is only 2.4 Å, the interconversion is relatively slow with estimated MFPT's of around 20 days for IP → OP and 15 h for OP → IP. The rate-limiting step for the reptation move is the disruption of hydrogen-bonds in the reactant dimer to allow the distance between the two strands to increase from its equilibrium value of 4.7 Å to about 8 Å. This change is necessary for the rearrangement of the sidechains and the formation of the hydrogen-bond pattern of the product dimer. The overall potential energy barrier for the IP ↔ OP reptation move is about 20 kcal mol^{-1}, which agrees well with the reported barriers of 12 kcal mol^{-1} for the β-hairpin in protein G[71] and 20 kcal mol^{-1} for Aβ_{16-22}.[72] A reptation move for the growth of amyloid fibrils was also detected experimentally for the PrP$_{109-122}$ peptide, and a time scale of about 8 h at 308 K was reported.[75] At this temperature, we find an MFPT of 5.5 days for the IP → OP conversion and 3.5 h for OP → IP. Given that we are comparing two different peptides, along with the uncertainties associated with the empirical force field and the approximations involved in calculating the rate constants, the time scales from experiment and theory compare quite well. Overall, the rate constants indicate that the reptation move is a likely mechanism to align GNNQQNY peptides in parallel β-sheet aggregates, and it seems that it is favored over a dissociation/reassociation mechanism.

Interconversion of Parallel and Antiparallel β-Sheets. Both interconversions IP ↔ IA and OP ↔ IA proceed via structure CD, where one of the peptide strands is rotated by about 90° with respect to the other. Inspection of the path in Figure 5e shows that the IP ↔ IA conversion is rate limited by the CD ↔ IA interconversion with an overall potential energy barrier of about 20 kcal mol^{-1}, whereas the first part of the pathway, for the IP ↔ CD conversion, involves only barriers below 5 kcal mol^{-1}. The OP ↔ CD conversion, on the other hand, implies an overall barrier of about 20 kcal mol^{-1}. The high barriers involving the structures IA and OP originate from their dense network of interpeptide hydrogen bonds. The fastest of the rotation pathways, IP → IA, has also been found for the GNNQQNY dimer in a recent MD study using the GROMACS96 forcefield and an explicit solvent description.[21] Our conclusion is that the interconversions between parallel and antiparallel conformations are feasible, with estimated MFPT's in the range of hours to days and they might provide a possible mechanism for the growth of amyloid β-sheets of GNNQQNY. Comparing with the rates for dissociation, it seems that a strand, which has docked in an antiparallel manner at the end of a parallel β-sheet, would rather rotate than dissociate/reassociate to align parallel to the β-sheet.

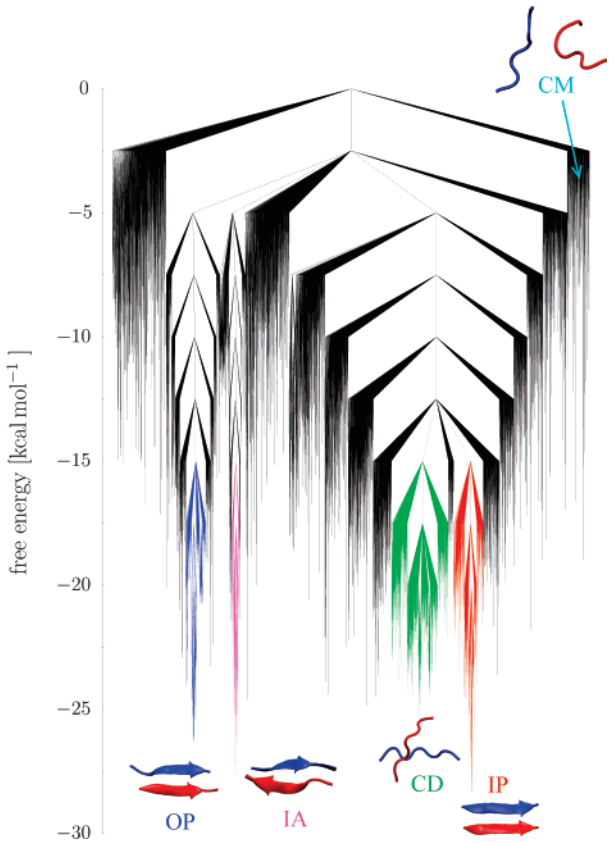


Figure 6. Free energy disconnectivity graph for the dimeric phase of GNNQQNY. The colored branches label the funnels corresponding to the IP, OP, IA, and CD structures clustered by C_α-rmsd < 1 Å from the lowest energy structure of the corresponding superbasin[54] with $F < -15$ kcal mol^{-1}.

3.4. Disconnectivity Graph. The combined stationary point database from all DPS runs is displayed in Figure 6 as a free energy disconnectivity graph.[56,57] The potential energy disconnectivity graph[54,55] is given in the Supporting Information for this paper. The minima in the database were classified by the C_α-rmsd from four reference structures corresponding to IP, OP, IA, and CD, where the reference structure for each cluster was taken as the one with the lowest potential energy in the set. Any minimum with C_α-rmsd < 1 Å from the given reference structure and a free energy $F < -15$ kcal mol^{-1} ($V < -620$ kcal mol^{-1} for the potential energy disconnectivity graph) is classified as belonging to that cluster. This grouping allows the identification of "funnels" in the disconnectivity graph, which correspond to the sets of minima IP, OP, IA, and CD. As expected, the CM structure is rather unfavorable: only 10 structures out of 28 631 in the database were classified as CM based on the requirement of rmsd < 1 Å from the reference structure shown in Figure 4.

Antiparallel Sheet is Favored by Potential Energy and the In-Register Parallel Sheet is Favored by Entropy. The order of potential energy for the structure sets IP, OP, and IA agrees well with the interaction energies, as summarized in Table 1, with the IA conformation being most and the IP structure being

(71) Wei, G.; Derreumaux, P.; Mousseau, N. *J. Chem. Phys.* **2003**, *119*, 6403.
(72) Santini, S.; Wei, G.; Mousseau, N.; Derreumaux, P. *Structure* **2004**, *12*, 1245−1255.
(73) Santini, S.; Mousseau, N.; Derreumaux, P. *J. Am. Chem. Soc.* **2004**, *126*, 11509−11516.
(74) Derreumaux, P.; Mousseau, N. *J. Chem. Phys.* **2007**, *126*, 025101.
(75) Petty, S. A.; Decatur, S. M. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 14272−14277.

least stabilized by peptide−peptide interactions. The situation changes when taking conformational entropy into account. The lowest free energy structures for the IP and IA conformations have almost the same free energy, and the OP structures are only 1−2 kcal mol^{-1} above them. These free energies support the result from the REMD simulation, which predicts almost equal populations for the three dimers, as seen in Figure 4. Comparing the potential and free energies, we conclude that optimizing the interpeptide interaction comes at the cost of reduced conformational flexibility in the case of the IA sheet, whereas the weaker interpeptide interaction in the IP sheet increases its conformational entropy.

Free Energy Barriers for Dimer Interconversion. The free energy disconnectivity graph provides additional insight into the rate constants calculated from the DPS approach and listed in Table 2. The funnels for the IP and CD structures are separated by low-lying transition states, which explains the relatively high interconversion rates between them. The relative instability of the IP dimer was also observed in a molecular dynamics simulation at 330 K.[20] At this temperature, the IP dimer was not stable when modeled by the CHARMM22 potential and explicit TIP3P water. For the interconversion between the IP, OP, and IA structures, we find a free energy barrier of about 20 kcal mol^{-1}, leading to MFPT's in the range of hours to days. For dissociation of the dimeric structures, barriers of more than the 25 kcal mol^{-1} need to be overcome, making this an unlikely process.

Comparison between the FES in Figure 4 and the free energy disconnectivity graph in Figure 6 reveals that the predicted barriers from the FES for the interconversion between the dimeric sheets are lower than those calculated from the database of stationary points. However, the free energy surface projected in terms of R_g and the interaction energy, E_{int}, (not shown) produces barriers that are more consistent with the DPS databases; in fact, some of the FES barriers are then actually higher.

4. Conclusions

In this report, we have studied the energy landscapes of the monomeric and dimeric amyloidogenic heptapeptide GNNQQNY, which was modeled by a united-atom potential and an implicit solvent representation. We performed replica exchange molecular dynamics (REMD) to explore the conformational space of the GNNQQNY monomer. The free energy surface at room temperature is dominated by four minima, which belong to a helix-like state, a compact geometry with a turn, a coil structure, and an extended β-like conformation. There is no significant free energy barrier for the transitions between the different conformations, and the discrete path sampling (DPS) approach predicts a time scale of nanoseconds for these transitions at room temperature. In agreement with experiment[13] and previous simulations,[19] the thermodynamically favored structure is found to be a coil geometry. The compact structures with either a helix or a turn are disfavored by entropy resulting from the restriction of vibrational degrees of freedom, whereas the extended structures are disfavored by potential energy.

From the findings for the monomer, it seems likely that fibrillogenesis rather than disordered aggregation is the preferred pathway of amyloid formation of GNNQQNY, as suggested for β-prone peptides.[22] To further test this hypothesis, we investigated the process of dimerization. We started with a REMD simulation to explore the conformational space. The dimer was found to have three free energy minima, corresponding to in-register parallel, IP, off-register parallel, OP, and antiparallel, IA, sheets. All are populated at room temperature for the current potential. Examination of the potential energy reveals that the IA conformation is lowest and the IP structure highest in potential energy. From an analysis of the interpeptide interactions, it is clear that the difference in potential energy mainly originates from different electrostatic interactions as a result of interpeptide hydrogen-bonding. The hydrogen-bonds between peptides are the dominant source of stabilization for the IA dimer, resulting in a twisted structure that optimizes these interactions. For the IP sheet, on the other hand, hydrophobic interactions are as important as electrostatic interactions for the stability of the dimer. Accurate calculations involving explicit treatment of the electronic structure would be desirable in future work to check the reliability of the empirical forcefield.

We then employed the DPS method to investigate the association pathways leading to stable dimers, as well as the interconversions between them. Kinetic properties were extracted from the resulting database of stationary points using the graph-transformation approach.[59] Inspection of the association pathways reveals that the dimers form without long-lived intermediates, which supports the idea that GNNQQNY belongs to the class of amyloidogenic peptides that form fibrils via fibrillogenesis.[22] The dimerization process can also be understood as a two-stage dock−lock mechanism, which was introduced to describe the kinetics of amyloid fibril elongation by addition of monomers.[65−67,69]

The estimated rate constants indicate that the three dimers are formed from monomers within milliseconds to seconds and are kinetically stable against dissociation at room temperature. Interconversion between the dimers is possible on time scales ranging from hours to days at 298 K. The rearrangement between the IP and OP structures is predicted to occur via reptation[71] of one strand of the β-sheet with respect to the other, and the calculated time scale is comparable to experiment.[75] Both the reptation move and the rearrangement of antiparallel sheets into parallel sheets via rotation of one of the peptide strands seem likely mechanisms for the growth of in-register β-sheets for GNNQQNY. According to the estimated rate constants, these moves should be preferred over a mechanism consisting of dissociation and reassociation of a β-strand. However, for larger GNNQQNY oligomers, further investigation is needed to check whether the internal rearrangement pathways remain kinetically accessible.

We note that these time scales should be considered as estimates and that the relative values and the mechanistic features of the pathways are likely to be more reliable than the absolute values of the rate constants. Nonetheless, the calculated rate constants generally compare well with experimental values such as the observed 8 h time scale for the reptation move in the $PrP_{109-122}$ peptide[75] and the rate of about one monomer attachment every few seconds in fibril elongation for insulin.[76] The approximations involved in calculating time scales from stationary point databases are generally complementary to approaches based on explicit dynamics. For example, the

(76) Knowles, T. P.; Shu, W.; Devlin, G. L.; Meehan, S.; Auer, S.; Dobson, C. M.; Welland, M. E. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 10016−10021.

harmonic densities of states employed in the present work result in systematic errors in the minimum-to-minimum rate constants, whereas Langevin dynamics simulations are affected by the choice of a friction coefficient.[77] Because the rate-limiting steps for all considered dimer pathways are concentration independent, the rate constants for these processes should not be influenced by the 1 order of magnitude higher concentration compared to experiment. Furthermore, we have verified that the higher concentration does not influence the conformational space of the dimer sampled during the REMD simulation.

The DPS stationary point database of the GNNQQNY dimer was visualized as a free energy disconnectivity graph. Comparing the potential and free energy of the stable dimers shows that the IA sheet has the lowest potential energy, but its tight network of hydrogen bonds restricts the vibrational degrees of freedom. The IP sheet, on the other hand, is favored by entropy as a result of weaker interpeptide interactions. The barriers between the favorable structures for the dimeric sheets in the free energy disconnectivity graph are higher than the corresponding barriers in the free energy surface plotted as a function of the radius of gyration and the rmsd from the amyloid β-sheet. However, the barriers calculated from the DPS database are comparable or lower than those of the FES expressed in terms of the radius of gyration and the interaction energy between the two peptides.

Acknowledgment. We are grateful to Dr Michele Vendruscolo for helpful discussions. B.S. gratefully acknowledges the BBSRC for financial support.

Supporting Information Available: Table listing the radius of gyration, rmsd from the β-strand, potential energy, temperature times entropy, and free energy for structures drawn randomly from the minima of the FES for the GNNQQNY monomer; residue−residue interaction energies and hydrogen-bonds for the different β-sheets found for the GNNQQNY dimer; potential energy disconnectivity graph for the dimeric phase of GNNQQNY. This material is available free of charge via the Internet at http://pubs.acs.org.

JA075346P

(77) Feig, M. *J. Chem. Theory Comput.* **2007**, *3*, 1734−1748.