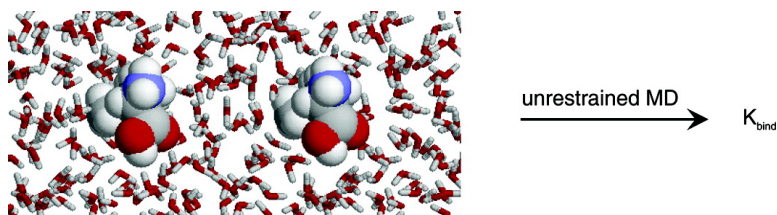## Communication

Association Lifetimes of Hydrophobic Amino Acid Pairs Measured Directly from Molecular Dynamics Simulations

Hongbo Yang, and Adrian H. Elcock

More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 3 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

Association Lifetimes of Hydrophobic Amino Acid Pairs Measured Directly from Molecular Dynamics Simulations

Hongbo Yang and Adrian H. Elcock*

Department of Biochemistry, University of Iowa, Iowa City, Iowa 52242-1109

Received July 1, 2003; E-mail: adrian-elcock@uiowa.edu

Interactions between the side chains of hydrophobic amino acids are crucial to the stabilities of proteins and their complexes. To fully understand the kinetics of processes such as protein folding, it is important to know the lifetimes of typical hydrophobic interactions, since they can have two opposing consequences. On one hand, for side chains that will be situated adjacent to each other in the native structure, a long contact lifetime might be advantageous.[1a,b] On the other, for side chains that do not form contacts in the folded state, the formation of long-lived interactions might be disadvantageous;[1c,d] this may be why stretches of contiguous hydrophobic residues are rare in nonmembrane proteins.[1e] As one way of measuring the lifetimes of hydrophobic contacts, we report here the results of six 70-ns molecular dynamics (MD) simulations of different amino acid pairs in aqueous solution. Our results provide estimates for association constants and demonstrate that current computational resources are now sufficient for full sampling of association events to be achieved for small molecules in aqueous solution. The latter point indicates that *direct* testing and parametrization of molecular mechanics force fields to reproduce aqueous phase binding data is now possible.

There have been many efforts to simulate the association equilibria of pairs of hydrophobic molecules,[2] most of which have involved computing the "potential of mean force" (PMF) for association. Although PMF calculations can yield important thermodynamic data, they have two limitations for the present context. First, being a purely thermodynamic property, the PMF does not provide direct information on the kinetics and lifetimes of association. Second, to compute a PMF, one must first assume a "reaction coordinate" for association; the associating molecules are restrained to points along this coordinate during the simulations. In an effort to avoid these drawbacks, the simulations reported here have been performed without the imposition of restraints, thereby allowing the kinetics and thermodynamics of association to be observed simultaneously.

Six MD simulations, each of 70-ns duration[3] were conducted of pairs of leucine, methionine and valine molecules in the following combinations: leu:leu, leu:met, met:met, leu:val, met:val, and val:val. The time dependence of the distance between the two separate molecules is plotted in Figure 1 for a typical simulation (note that this plot covers only half of the total 70 ns). The key feature is the rapid variation of the intermolecular distance: periods of association (low distances) are interspersed by periods in which the molecules are more than 15 Å apart. The large separation achieved between association events, and the fact that consecutive events often involve molecule copies that originated in different periodic cells, indicate that the events are independent. Intriguingly, association of the two molecules does not appear to be driven solely by side-chain–side-chain interactions. In the leu:leu simulation for example, on the occasions that the molecules are within 4 Å, the closest contact is a side-chain–side-chain contact only 35% of the time: backbone–backbone contacts account for ~30% of the closest contacts.
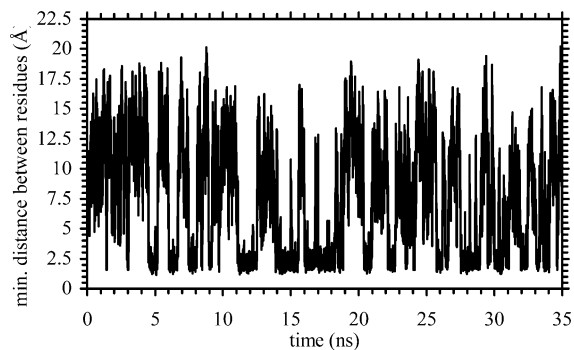

Figure 1. Plot of the intermolecular distance between atoms of the leu:leu pair as a function of time.

A good way to visualize association events is to plot the time dependence of the total solvent accessible surface area (SASA) of the molecule pairs (Figure 2). For the majority of time, the two molecules in each simulation are not closely associated; variations in the total SASA in this state result from internal conformational fluctuations. Frequently, however, the total SASA drops abruptly, indicating an association event. A first important result from Figure 2 is that the longest lifetime of any single association event is 3–4 ns. This suggests (perhaps not surprisingly) that the formation of individual pairs of non-native hydrophobic contacts is not an impediment to protein folding—although clearly it leaves open the possibility that formation of *clusters* of non-native contacts might be. A second important result however is that in all simulations, *many* association events are observed, suggesting that sampling is sufficient to compute equilibrium thermodynamic data. This idea receives support when frequency histograms of the SASA values are constructed; the results for the leu:leu, met:met, and val:val simulations are illustrated in Figure 3. For all six molecule pairs studied, the distributions obtained from the simulations (circles) can be accurately described as a sum of two Gaussian functions (lines)—one each for the bound and unbound states. Interestingly, the standard deviation (σ) of bound state SASA values is in each case larger than that of the corresponding unbound values (Table 1): this presumably reflects the fact that the structures of associated molecular pairs are more diverse than those of two separate molecules. It is especially striking that the disparity between bound and unbound σ values is larger for simulations involving methionine; this implies a consistently greater degree of structural heterogeneity in the latter's complexes.

The separation of the SASA data into two states allows the equilibrium constant (K_{eq}) for association to be calculated (Table 1); errors can be estimated with bootstrap sampling methods.[4] Although the computed values are subject to some degree of uncertainty, it appears that association of the leu:leu and met:met pairs is stronger (~0.2 kcal/mol) than that of the other four pairs.
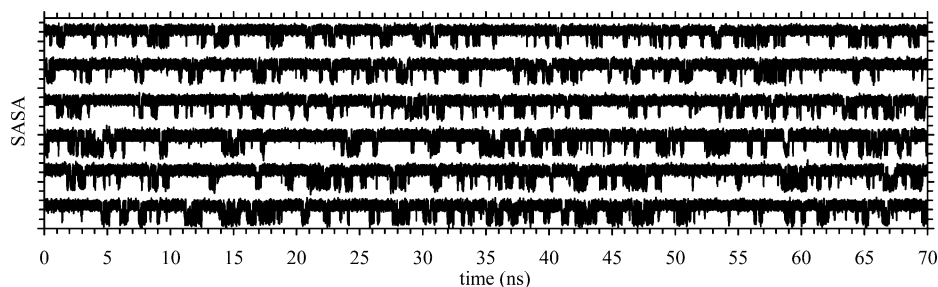
Figure 2. Solvent accessible surface area of the molecule pairs in the six simulations plotted as a function of time. Results shown are for (from top): val:val, val:met, val:leu, met:met, leu:met, and leu:leu.
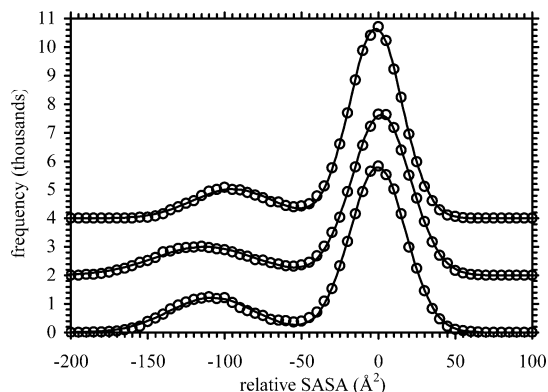

Figure 3. Frequency histograms of data shown in Figure 2. Open circles are from the simulations; solid lines are fits to the data (see text). Results shown are for (from top): val:val, met:met, and leu:leu.

Table 1. Computed Equilibrium Constants for Association

res. pair	K_{eq} (M^{-1})	$\langle\Delta$SASA$\rangle$ (Å^2)	σ_{bound}	$\sigma_{unbound}$	$\sigma_{bound}/\sigma_{unbound}$
val:val	4.5 ± 1.2	93	26.1	17.1	1.53
val:met	4.7 ± 1.4	100	30.0	17.7	1.70
val:leu	4.6 ± 1.1	100	27.1	17.8	1.52
met:met	6.3 ± 2.2	115	33.2	19.0	1.75
leu:met	4.3 ± 1.5	109	32.3	18.9	1.71
leu:leu	6.6 ± 1.9	108	28.1	18.3	1.54

The association lifetimes and equilibrium constants reported here should be considered estimates, since it is not yet clear that current molecular parameter sets are of quantitative accuracy. Instead, the main conclusion to be drawn here is that these numbers can be computed, with statistical confidence, directly from unforced MD simulations; this ability appears to be independent of the parameter set used (data not shown). This finding fits well with previous work by van Gunsteren and co-workers, indicating that the *intra*molecular conformational behavior of short β-peptides can be exhaustively explored using unforced MD simulations;[5a] the potential power of "brute force" simulations has also been demonstrated recently by McCammon and co-workers.[5b] It is worth noting that simulations such as those reported here can now be completed in ~3 weeks on current generation single-processor PCs. This means that it is now possible—by comparison with experimental measurements[6]—to

develop quantitatively accurate simulation models for both the binding and folding of amino acids in aqueous solution.

References

(1) (a) Northey, J. G.; Di Nardo, A. A.; Davidson, A. R. *Nat. Struct. Biol.* **2002**, *9*, 126. (b) Viguera, A. R.; Vega, C.; Serrano, L. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 5349. (c) Paci, E.; Vendruscolo, M.; Karplus, M. *Proteins* **2002**, *47*, 379. (d) Chowdury, S.; Zhang, W.; Wu, C.; Xiong, G.; Duan, Y. *Biopolymers* **2002**, *68*, 63. (e) Schwartz, R.; Istrail, S.; King, J. *Protein Sci.* **2001**, *10*, 1023.

(2) (a) Jorgensen, W. L.; Buckner, J. K.; Boudon, S.; Tirado-Rives, J. *J. Chem. Phys.* **1988**, *89*, 3742. (b) Rank, J. A.; Baker, D. *Protein Sci.* **1997**, *6*, 347. (c) Ludemann, S.; Abseher, R.; Schreiber, H.; Steinhauser, O. *J. Am. Chem. Soc.* **1997**, *119*, 4206. (d) Shimizu, S.; Chan, H. S. *J. Chem. Phys.* **2000**, *113*, 4683. (e) Czaplewski, C.; Rodziewicz-Motowidlo, S.; Liwo, A.; Ripoll, D. R.; Wawak, R. J.; Scheraga, H. A. *Protein Sci.* **2000**, *9*, 1235. (f) Raschke, T. M.; Tsai, J.; Levitt, M. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 5965. (g) Southall, N. T.; Dill, K. A. *Biophys. Chem.* **2002**, *101*, 295.

(3) Simulations were performed with GROMACS[7a] and maintained at 1atm pressure and 298 K, using the Parrinello−Rahman[7b] and Nose−Hoover[7c, d] restraint algorithms. The molecule pairs were initially separated by around 10 Å and then immersed in a cubic box of linear dimension 30 Å containing ~870 preequilibrated water molecules described by the SPC model.[7e] Periodic boundary conditions were applied. A short-range cutoff of 12 Å was applied to nonbonded interactions; long-range electrostatic interaction were evaluated using the particle mesh Ewald (PME) method.[7f] Simulations were conducted with a time step of 2 fs, with all bonds constrained using the LINCS algorithm.[7g] The GROMACS all-atom force field[7a] was used for all simulations; as a simple means of mimicking how the molecules might behave in a protein, their termini were assumed to be in neutral protonation states.

(4) Two separate lists were constructed: one of the lifetimes of association events and one of the lifetimes of dissociation events. By randomly sampling alternately from the two lists, a total of 1000 alternative "trajectories" was constructed, each of 70ns duration. Numbers reported in Table 1 are the standard deviations of K_{eq} values calculated from these 1000 samples.

(5) Daura, X.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. *J. Mol. Biol.* **1998**, *280*, 925. (b) Zhang, Y. K.; McCammon, J. A. *J. Chem. Phys.* **2003**, *118*, 1821.

(6) See, e.g., Blackburn, G. M.; Lilley, T. H.; Milburn, P. J. *J. Chem. Soc., Faraday Trans. 1* **1985**, *81*, 2191. (b) Habermann, S. M.; Murphy, K. P. *Protein Sci.* **1996**, *5*, 1229.

(7) (a) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. *Comput. Phys. Comm.* **1995**, *91*, 43. (b) Parrinello, M.; Rahman, R. *J. Appl. Phys.* **1981**, *52*, 7182. (c) Nose, S. *Mol. Phys.* **1984**, *52*, 255. (d) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695. (e) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. *Intermolecular Forces*; D. Reidel Publishing Company: Dordrecht, 1981. (f) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089. (g) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. *J. Comput. Chem.* **1997**, *18*, 1463.

JA037010V