Quantitative Analysis of Hydrophobically Induced Folding in a Minimal Model System

Robb R. Gardner, Laurie A. Christianson, and Samuel H. Gellman*

Department of Chemistry, University of Wisconsin Madison, Wisconsin 53706

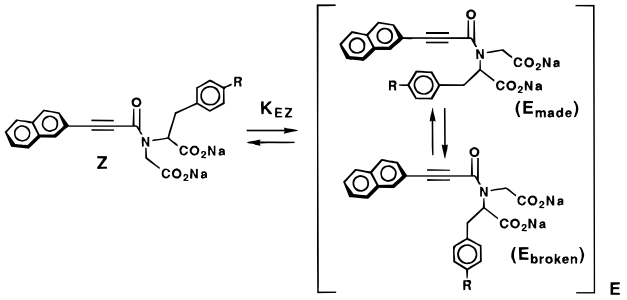
Received February 19, 1997

The extent to which hydrophobic interactions contribute to the stability of biopolymer folding patterns and biopolymer−ligand complexes is a topic of ongoing debate.[1] We are approaching this problem by examining minimum increments of hydrophobically induced folding in small molecules.[1l,q] In principle, such systems allow one to know which surfaces are buried in the folded state and, since the network of conformation-directing forces is relatively simple, to identify the forces responsible for folding. The folding model system approach is complementary to two more commonly employed methods for evaluating minimum hydrophobic increments: site-directed protein mutagenesis[1f] and aqueous/nonpolar partitioning of small solutes.[1a] Interpretation of mutagenesis data can be hampered by the extensive cooperativity among the noncovalent forces that underlie protein tertiary structure. Interpretation of partitioning data can be ambiguous because of variation in the nonaqueous component and because of the lack of structural insight on solute−solvent interactions.

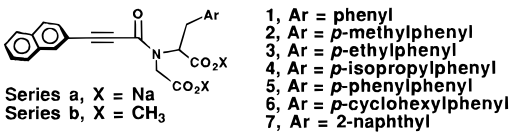
NMR spectroscopy is powerful for assessing solution conformation, but it is often difficult to use NMR to quantify small molecule folding equilibria, because most forms of molecular flexibility lead to conformationally averaged signals. This limitation prevented quantitative analysis of model systems we have previously examined.[1l,q] This difficulty can be overcome by the "slow rotation strategy": two interacting sites are separated by a linker containing a bond that rotates slowly enough to allow direct NMR observation of two rotamer populations but rapidly enough for equilibration.[2] Ideally, the interacting sites would make contact in only one rotamer. Here we use the slow rotation strategy to evaluate minimum increments of hydrophobic interaction (compounds **1**−**7**).

Our design is illustrated in Scheme 1.[3] The tertiary amide C−N bond provides the slow rotation. Carboxylates, required for aqueous solubility, are placed on both of the amide nitrogen's substituents, in order to reduce solvation differences between

Scheme 1



E and *Z* rotamers. Any residual rotamer solvation difference is presumably constant among **1a**−**7a**.[4] Molecular mechanics calculations (AMBER*[5]/MacroModel[6] 5.0; GB/SA[7] aqueous



solvation; calculations conducted on carboxylic acid forms) indicated that the distal edge of the phenyl side chain of **1** can just reach the naphthyl surface in low energy *E* conformers and that hydrocarbon surface area[8] burial in these conformers increases with para substitution on the phenyl group (**2**−**6**). Monte Carlo[9] and Monte Carlo/stochastic dynamics[10] (MCSD) analysis of **5** and **7** (separate MCSD runs for *E* and *Z* rotamers), using the GB/SA model for aqueous solvation, indicated that the most highly populated conformation in each case is an *E* rotamer with the naphthyl and aryl (biphenyl or naphthyl) groups clustered.[11] These structures are quite similar to the solid state conformations of diesters **5b** and **7b** (Figure 1). In contrast to the naphthyl-aryl clustering observed for the *E* rotamers, no direct contact between the naphthyl and aryl groups was observed for the *Z* rotamers of **5** or **7** in the MCSD runs; similar observations were made for **1**, **2**, and **4**. Monte Carlo simulations indicate that direct contact between the naphthyl and aryl groups of **1**−**7** would require substantial twisting about the amide C−N bond (20−30°). These computational results suggest that there is little or no contact between aryl and naphthyl groups of the *Z* rotamers in the real molecules **1a**−**7a**.[12]

(1) For leading references, see: (a) Makhatadze, G. I.; Privalov, P. L. *Adv. Protein Chem.* **1995**, *47*, 307. (b) Spolar, R. S.; Livingstone, J. R.; Record, M. T. *Biochemistry* **1992**, *31*, 3947. (c) Sharp, K. A.; Nicholls, A.; Friedman, R.; Honig, B. *Biochemistry* **1991**, *30*, 9686. (d) Muller, N. *Trends Biochem. Sci.* **1992**, *17*, 459. (e) Reynolds, J. A.; Gilbert, D. B.; Tanford, C. *Proc. Natl. Acad. Sci U.S.A.* **1974**, *71*, 2925. (f) Pace, C. N. *J. Mol. Biol.* **1992**, *226*, 29. (g) Spolar, R. S.; Record, M. T. *Science* **1994**, *263*, 777. (h) Dill, K. A. *Biochemistry* **1990**, *29*, 7133. (i) Dill, K. A. *Science* **1990**, *250*, 297. (j) Herzfeld, J. *Science* **1991**, *253*, 88. (k) Blokzijl, W.; Engberts, J. B. F. N. *Angew. Chem., Int. Ed. Engl.***1993**, *32*, 1545. (l) Newcomb, L. F.; Gellman, S. H. *J. Am. Chem. Soc.* **1994**, *116*, 4993. (m) Friedman, R. A.; Honig, B. A. *Biophys. J.* **1995**, *69*, 1528. (n) Gellman, S. H.; Haque, T. S.; Newcomb, L. F. *Biophys. J.* **1996**, *71*, 3523. (o) Honig, B. A.; Friedman, R. A. *Biophys. J.* **1996**, *71*, 3525. (p) Wiley, R. A.; Rich, D. H. *Med. Res. Rev.* **1993**, *13*, 327. (q) Newcomb, L. F.; Haque, T. S.; Gellman, S. H. *J. Am. Chem. Soc.* **1995**, *117*, 6509.

(2) Previous uses of the slow rotation strategy to probe noncovalent attractions: (a) Carter, R. E.; Stilbs, P. *J. Am. Chem. Soc.* **1976**, *98*, 7515. (b) Oki, M. *Acc. Chem. Res.* **1990**, *23*, 351. (c) Paliwal, S.; Geib, S.; Wilcox, C. S. *J. Am. Chem. Soc.* **1994**, *116*, 4497. (d) Kemp, D. S.; Allen, T. J.; Oslick, S. L. *J. Am. Chem. Soc.* **1995**, *117*, 9941. (e) Boyd, D. R.; Evans, T. A.; Jennings, W. B.; Malone, J. F.; O'Sullivan, W.; Smith, A. *J. Chem. Soc., Chem. Commun.* **1996**, 2269.

(3) All new compounds were synthesized and characterized by standard methods, which will be described in a subsequent report.

(4) Recent discussions of the interplay among hydrophobic surfaces, hydrophilic surfaces, and aqueous solvation shells: (a) Meng, E. C.; Kollman, P. A. *J. Phys. Chem.* **1996**, *100*, 11460. (b) Streefland, L.; Blandamer, M. J.; Engberts, J. B. F. N. *J. Am. Chem. Soc.* **1996**, *118*, 9539.

(5) AMBER force field: Weiner, S.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230. AMBER* modification: McDonald, D. Q.; Still, W. C. *Tetrahedron Lett.* **1992**, *33*, 7747.

(6) Mohamdi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. *J. Comput. Chem.* **1990**, *11*, 440.

(7) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127.

(8) Hubbard, S. J.; Thornton, J. M. NACCESS, computer program; Department of Biochemistry and Molecular Biology, University College London, 1993.

(9) Chang, C.; Guida, W. C.; Still, W. C. *J. Am. Chem. Soc.* **1989**, *111*, 4379, and references therein.

(10) Still, W. C.; Guarnieri, F. *J. Comput. Chem.* **1994**, *15*, 1302.

(11) MCSD simulations were run for 1000 ps in 1 fs steps at 300 K. Torsion angles with low energy barriers (except for those involving methyl groups) were allowed to vary at MC steps. Backbone torsion angles were monitored continuously and tabulated at 200 ps intervals through the runs to verify that the distribution of these angles did not drift. In three cases (conjugate acid forms of *E*-**1a,** *Z*-**4a**, and *Z*-**7a**) runs were repeated from different starting conformations, and the results of these runs were checked for convergence.

(12) Small aryl side chain-dependent variations in chemical shift were observed for the proton at the naphthyl 1-position of the *Z* rotamers of **1a**−**7a** in D_2O: 8.27 ppm for **1a**−**4a** vs 8.24 ppm for **5a** and **6a** vs 8.25 ppm for **7a**. Furthermore, this proton appeared at 8.32 ppm in the analogous dicarboxylate bearing only a methyl group as the side chain (an alanine derivative). The upfield shifts observed for **1a**−**7a** relative to this alanine-derived dicarboxylate could arise in part from anisotropic effects of the aromatic side chains in noncollapsed conformations, since MCSD analysis of the *Z* rotamers indicates a small population of noncollapsed conformers in which the 1-naphthyl proton is within 5 Å of the aryl group (<10% of *Z* rotamers found).

S0002-7863(97)00531-3 CCC: $14.00

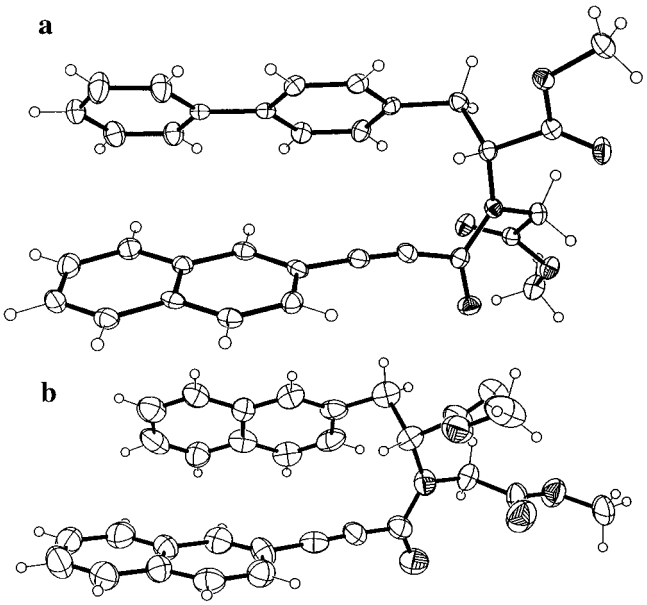


Figure 1. (a) Solid state conformation of dimethyl ester **5b** (*E* rotamer). (b) Solid state conformation of dimethyl ester **7b** (*E* rotamer).

Table 1

structure	K_{EZ} (D_2O)[a]	$\Delta\Delta G_{EZ}$ (D_2O; vs **1a**)	K_{EZ} ($CDCl_3$)[b]
1	1.4		1.5
2	1.8	−0.15 kcal mol	1.7
3	2.0	−0.21 kcal/mol	1.7
4	2.5	−0.34 kcal/mol	1.6
5	3.9	−0.61 kcal/mol	1.6
6	3.0	−0.45 kcal/mol	1.6
7	2.9		1.6

[a] *E*:*Z* ratio for dicarboxylates (series a) measured by NMR in D_2O at 298 K. Values for **1a**−**3a** and **7a** at 1 mM; values for **4a**−**6a** at 0.5 mM; control experiments indicated minimal aggregation under these conditions (see Supporting Information). For all compounds except **5a** and **7a**, K_{EZ} was measured independently from two or more resonances. Estimated uncertainty ±0.2. [b] *E*:*Z* ratio for dimethyl esters (series b) measured by NMR in $CDCl_3$. Estimated uncertainty = ±0.1.

The *E*:*Z* ratios in Table 1 were measured for **1a**−**7a** in dilute D_2O solution and for diesters **1b**−**7b** in $CDCl_3$, by integration of well-resolved ^1H NMR resonances. (Control experiments indicated that there is little or no aggregation of dicarboxylates **1a**−**7a** in D_2O at the concentrations at which the *E*:*Z* rotamer ratio measurements were made.[13]) Most of the dicarboxylates and diesters were also subjected to NOESY[14] analysis. Side-chain naphthyl NOEs were observed for only one rotamer, which was assigned as *E*. These assignments were supported by consistent trends among the chemical shifts of protons common to all compounds (the protons α and β to the nitrogen), for dicarboxylates **1a**−**7a** in D_2O and for diesters **1b**−**7b** in $CDCl_3$. Further support for the rotamer assignments came from dissolution of the crystalline diesters **5b** and **7b** in $CDCl_3$ at low temperature: only resonances previously assigned to the *E* rotamers were observed initially, as expected from the crystal structure (Figure 1). Resonances previously assigned to the *Z* rotamers appeared upon warming.

For compound **1a**, with just a phenylmethyl group as side chain, the preference for the *E* rotamer is very modest (Table 1), which was expected since computer modeling indicated that the phenyl group barely reaches the naphthyl group in the *E* rotamer. Introducing even small para substituents on the side chain phenyl, however, causes the *E*:*Z* ratio to increase. The variation in *E*:*Z* rotamer ratios among **1a**−**7a** in D_2O appears to represent a hydrophobic effect, because the diester analogues **1b**−**7b** all display very similar *E*:*Z* ratios in $CDCl_3$. Further, the *E*:*Z* ratio for **1a** in D_2O is similar to that for the analogous dimethyl ester, **1b**, in $CDCl_3$.

The *E*:*Z* ratio measured by NMR can be converted to ΔG_{EZ} in the usual way. ΔG_{EZ} is not exactly equal to the naphthyl-aryl interaction energy, because the covalent skeleton is not sufficiently rigid to enforce naphthyl-aryl contact in the *E* rotamers (as indicated by the two *E* substates in Scheme 1), and because the covalent skeleton may not allow optimal naphthyl-aryl juxtaposition. Nevertheless, ΔG_{EZ} should provide a useful comparative indication of the energetic consequences of bringing the aryl and naphthyl groups together. These consequences are expected to include (but are not necessarily limited to) a favorable contribution from hydrophobic surface desolvation and an unfavorable contribution from conformational entropy. This latter component would arise from restriction of motion about the several bonds with low torsional barriers, as a result of naphthyl-aryl contact. The *E*:*Z* ratio of 1.4 observed for **1a** in D_2O indicates a slight preference for the *E* rotamer that is apparently *not* a hydrophobic effect, since the *E*:*Z* ratio of diester **1b** in $CDCl_3$ is indistinguishable. We therefore subtract ΔG_{EZ} for **1a** from the ΔG_{EZ} values for **2a**−**6a** to generate the $\Delta\Delta G_{EZ}$ values; $\Delta\Delta G_{EZ}$ should be related to the energetic contribution to folding that results from interaction between the para substitutents on the phenyl rings (R in Scheme 1) and the naphthyl group.

Our system can reliably detect differences among R-naphthyl interactions that cause variations in K_{EZ} of ≥0.4. Conclusions drawn from Table 1 differ substantially from those drawn from thermodynamic parameters for transfer of simple hydrocarbons from the pure liquid phase to dilute aqueous solution at 25 °C. ΔG for transfer of toluene is 0.8 kcal/mol more favorable than for transfer of ethylbenzene,[15] while the *E*:*Z* rotamer ratios of **2a** and **3a** are indistinguishable (the drive for folding provided by the methyl-naphthyl and ethyl-naphthyl interactions differ by <0.2 kcal/mol in Gibbs free energy). ΔG for transfer of benzene to aqueous solution is 2.1 kcal/mol more favorable than for transfer of cyclohexane,[15] but the *p*-phenyl group of **5a** promotes folding *more strongly* than does the *p*-cyclohexyl group of **6a**. This result may reflect differences in the conformational preferences about the phenyl−phenyl bond of **5a** relative to the phenyl−cyclohexyl bond of **6a** or an intrinsic affinity of one aromatic group for another.[1a]

We have identified a versatile model system that allows quantitative evaluation of small increments of hydrophobic contact in aqueous solution. This type of minimal hydrophobic clustering has thermodynamic consequences that differ substantially from those of nonpolar-to-aqueous transfer of small hydrocarbons. Hydrocarbon phase transfer may be a good model for the complete surface burial experienced by an amino acid side chain in the core of a large protein,[16] but our model system is probably more relevant to the clustering of side chains at or near the surface of a native protein or the packing of a hydrophobic ligand against a biopolymer surface. The approach we have employed should allow examination of a wide range of noncovalent interactions that are thought to be important to biopolymer structure and function.

Acknowledgment. We thank Dr. Douglas Powell for crystallographic analysis and Dr. Ruth Saecker for valuable comments. This research was supported by the National Science Foundation (CHE-9622653). S.H.G. thanks the Alfred P. Sloan Foundation for a Research Fellowship.

Supporting Information Available: Variable concentration ^1H NMR data for selected compounds (7 pages). See any current masthead page for ordering and Internet access instructions.

JA970531W

(13) Data may be found in the Supporting Information.
(14) Macura, S.; Ernst, R. R. *Mol. Phys.* **1980**, *41*, 95.
(15) Privalov, P. L.; Gill, S. J. *Pure Appl. Chem.* **1989**, *61*, 1097, and references therein.
(16) (a) Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 8069. (b) Livingstone, J. R.; Spolar, R. S.; Record, M. T. *Biochemistry* **1991**, *30*, 4237.