Protein Science (1997), *6*: 2606-2616. Cambridge University Press. Printed in the USA.

ARTICLE

Simulating the minimum core for hydrophobic collapse in globular proteins

JERRY TSAI, [1] MARK GERSTEIN, [1], [2] and MICHAEL LEVITT [1]

[1] Department of Structural Biology, Stanford University, Stanford, California 94305-5126

(Received May 13, 1997; Accepted August 14, 1997)

Reprint requests to: Jerry Tsai, Department of Structural Biology, Fairchild Building D109, Stanford University, Stanford, California 94305-5126;
e-mail:jotter@potential.stanford.edu.

[2]Present address: Molecular Biophysics & Biochemistry Department, Yale University, P.O. Box 208114, New Haven, Connecticut 06520.

Abstract (The NLM-formatted bibliographic entry is also available.)

To investigate the nature of hydrophobic collapse considered to be the driving force in protein folding, we have simulated aqueous solutions of two model hydrophobic solutes, methane and isobutylene. Using a novel methodology for determining contacts, we can precisely follow hydrophobic aggregation as it proceeds through three stages: dispersed, transition, and collapsed. Theoretical modeling of the cluster formation observed by simulation indicates that this aggregation is cooperative and that the simulations favor the formation of a single cluster midway through the transition stage. This defines a minimum solute hydrophobic core volume. We compare this with protein hydrophobic core volumes determined from solved crystal structures. Our analysis shows that the solute core volume roughly estimates the minimum core size required for independent hydrophobic stabilization of a protein and defines a limiting concentration of nonpolar residues that can cause hydrophobic collapse. These results suggest that the physical forces driving aggregation of hydrophobic molecules in water is indeed responsible for protein folding.

Keywords: aqueous solutions; Delauney Tessellation; hydrophobic collapse; isobutylene; methane; molecular dynamics simulation; protein hydrophobic core; urea; Voronoi polyhedra

Article Contents

(You can also go directly to the beginning of the text.)

Introduction

In 1959, Kauzmann proposed that the hydrophobic effect was one of the principal forces stabilizing a protein's structure. In general, three-dimensional structures of proteins have been found to possess a core of hydrophobic residues, confirming Kauzmann's hypothesis. During protein folding, the aggregation of these nonpolar amino acids into a stable core has been termed hydrophobic collapse. This process is not unique to proteins; nature makes use of hydrophobic collapse in the formation of many other biological structures, such as lipid membranes.

Because of such clear biological importance, a great deal of work has been done to understand the basis of hydrophobicity, especially as it relates to proteins (Dill, 1990). Many of these studies modeled simple, binary solutions of a single nonpolar solute molecule in water. Experiments on the free energy transfer of a nonpolar solute from a pure to an aqueous solution indicate that, at room temperature, the partition of the solution into aqueous and nonpolar phases is driven primarily by entropy (Privalov & Gill, 1988). Prior simulation studies of nonpolar solutes (usually Lennard-Jones spheres or methane) in solution agree with this description of hydrophobicity; however, most of them were unable to reproduce stable hydrophobic aggregation under normal conditions (Geiger et al., 1979; Panagali et al., 1979; Rapaport & Scheraga, 1982; Watanabe & Andersen, 1986; Laaksonen & Stilbs, 1991). Instead, the nonpolar solutes were usually separated by one water molecule (the solvent-separated pair). Not finding aggregation, some investigators (Rapaport & Scheraga, 1982; Laaksonen & Stilbs, 1991) suggested that stable aggregation

at room temperature required larger systems because all of the previous simulations involved fewer than four solute molecules. Another possible weakness of the earlier work was their limited sampling: all of the aforementioned dynamics runs simulated no more than a few hundred picoseconds. Addressing both of these shortcomings, Wallqvist (1991a, 1991b) performed molecular dynamics on 18 Lennard-Jones spheres in water for more than 1 ns and reported that the solute molecules formed a cylindrical aggregate. Instead of longer simulations, some studies have tried to increase sampling by slightly elevating the simulation temperature and have shown that transient aggregates form in molecular dynamics of methane (Skipper, 1993) and ethane (Mancera & Buckingham, 1995). Moreover, in long (>5 ns) simulations of two methane molecules in water, more recent work found that the potential of mean force between the methane molecules favored association over the solvent-separate pair (Smith & Haymet, 1993). Corroborating these results, a Monte Carlo study confirmed that aggregation is increasingly preferred in systems of water containing 2, 3, and 14 methane molecules (Rank & Baker, 1997).

As part of a previous study on urea solvation, we also discovered aggregation of a hydrophobic solute (isobutylene), but only at higher concentrations and longer simulation times (Tsai et al., 1996). That result motivated further studies in hopes of understanding hydrophobic collapse in protein folding. In the present study, we run a comprehensive set of the isobutylene simulations covering a wide range of concentrations from 0.26 M to 6.64 M (Table 1).

Table 1. *Isobutylene simulation parameters*

No. isobutylenes	No. waters	Box volume (Å^3)	Box side (Å)	Density (g/mL)	Molarity	Sim. length (ns)
1	211	6,428	18.59	0.996	0.26	0.5
2	210	6,519	18.68	0.991	0.51	0.5
3	208	6,580	18.74	0.987	0.76	0.5
4	204	6,582	18.74	0.983	1.01	0.5
5	196	6,464	18.63	0.978	1.28	0.5
6	198	6,645	18.80	0.975	1.50	0.5
7	195	6,676	18.83	0.971	1.74	0.5
8	195	6,797	18.94	0.967	1.95	0.5
9	189	6,739	18.89	0.962	2.22	0.5
10	192	6,950	19.08	0.960	2.39	0.5
11	195	7,160	19.27	0.957	2.55	0.5
12	197	7,341	19.44	0.954	2.71	0.5
13	193	7,343	19.44	0.950	2.94	0.5
14	191	7,404	19.49	0.947	3.14	0.5
16	194	7,736	19.78	0.942	3.43	0.5
18	186	7,739	19.78	0.935	3.86	0.5
20	183	7,891	19.91	0.929	4.21	0.5
22	186	8,223	20.18	0.925	4.44	0.5
24	178	8,226	20.19	0.918	4.84	0.5
25	179	8,377	20.31	0.916	4.96	0.5
26	177	8,439	20.36	0.913	5.12	0.5
27	179	8,619	20.50	0.912	5.20	0.5
28	168	8,412	20.34	0.907	5.53	0.5
29	168	8,533	20.43	0.905	5.64	0.5
30	173	8,803	20.65	0.904	5.66	0.5
31	173	8,924	20.74	0.902	5.77	0.5
32	175	9,105	20.88	0.901	5.84	0.5
34	169	9,168	20.93	0.896	6.16	0.5
36	160	9,141	20.91	0.889	6.54	0.5
38	164	9,503	21.18	0.888	6.64	0.5

We also study the simpler and more classic molecule, methane, by simulating concentrations from 0.26 M to 11.17 M (Table 2). To ensure adequate sampling (Tsai et al., 1996), all simulations were run for at least 0.5 ns.

Table 2. *Methane simulation parameters*

No. isobutylenes	No. waters	Box volume (Å^3)	Box side (Å)	Density (g/mL)	Molarity	Sim. length (ns)
Standard, reference simulations						
1	214	6,454	18.62	0.995	0.26	1.0
2	212	6,451	18.62	0.991	0.51	1.0
3	214	6,568	18.73	0.986	0.76	1.0
4	211	6,535	18.70	0.981	1.02	1.0
5	209	6,533	18.69	0.977	1.27	1.0
6	212	6,680	18.83	0.973	1.49	1.0
7	204	6,498	18.66	0.967	1.79	1.0
8	206	6,615	18.77	0.963	2.01	1.0
9	206	6,672	18.83	0.959	2.24	1.0
10	204	6,669	18.82	0.954	2.49	1.0
11	201	6,637	18.79	0.949	2.75	1.0
12	200	6,664	18.82	0.945	2.99	1.0
13	200	6,721	18.87	0.941	3.21	1.0
14	195	6,629	18.79	0.935	3.51	1.0
15	193	6,626	18.78	0.931	3.76	1.0
16	190	6,594	18.75	0.926	4.03	1.0
17	188	6,591	18.75	0.921	4.28	1.0
18	185	6,559	18.72	0.916	4.56	1.0
19	188	6,705	18.86	0.913	4.71	1.0
20	184	6,643	18.80	0.908	5.00	1.0
21	188	6,820	18.96	0.906	5.11	1.0
22	184	6,757	18.91	0.900	5.41	1.0
23	183	6,785	18.93	0.896	5.63	1.0
24	180	6,752	18.90	0.891	5.90	1.0
25	182	6,869	19.01	0.889	6.04	1.0
26	182	6,926	19.06	0.885	6.23	1.0
27	182	6,983	19.11	0.882	6.42	1.0
28	172	6,742	18.89	0.873	6.90	1.0
29	173	6,829	18.97	0.870	7.05	1.0
30	172	6,856	19.00	0.866	7.27	1.0
31	167	6,764	18.91	0.860	7.61	1.0
32	166	6,791	18.94	0.856	7.83	1.0
33	168	6,908	19.04	0.854	7.93	1.0
34	168	6,965	19.10	0.851	8.11	1.0
35	163	6,873	19.01	0.844	8.46	1.0
36	160	6,840	18.98	0.839	8.74	1.0
37	169	7,166	19.28	0.842	8.57	1.0
38	163	7,044	19.17	0.835	8.96	1.0
39	162	7,071	19.19	0.831	9.16	1.0
40	160	7,069	19.19	0.827	9.40	1.0
41	160	7,126	19.24	0.824	9.55	1.0
42	160	7,183	19.29	0.821	9.71	1.0
43	155	7,091	19.21	0.815	10.07	1.0
44	157	7,208	19.32	0.813	10.14	1.0
45	156	7,235	19.34	0.810	10.33	1.0
46	154	7,232	19.34	0.805	10.56	1.0
47	157	7,379	19.47	0.805	10.58	1.0
48	147	7,138	19.25	0.794	11.17	1.0
Half-box simulations						
3	106	3,340	14.95	0.973	1.49	1.0
6	100	3,332	14.94	0.945	2.99	1.0
12	90	3,376	15.00	0.891	5.90	1.0
24	73	3,554	15.26	0.793	11.21	1.0

Table 2. *Methane simulation parameters*

Our analysis uses a novel method to measure the degree of hydrophobic collapse that is based on an elegant mathematical construct developed by Voronoi (1908). Previously, these constructs have been used successfully both on proteins (Richards, 1974; Finney, 1978; Gerstein et al., 1995) and with simulations (Shih et al., 1994; Gerstein et al., 1995; Tsai et al., 1996). Using the Delauney Tessellation (Delauney, 1934) to uniquely define an atom's neighbors, this method completely divides the simulation box volume into polyhedra surrounding each atom. From these polyhedra, we find the volume and surface area of each atom and use the area of the face shared by two atoms to measure their degree of association (Fig. 1). This information allows us to unambiguously characterize the state of aggregation and the local environment of each individual molecule.

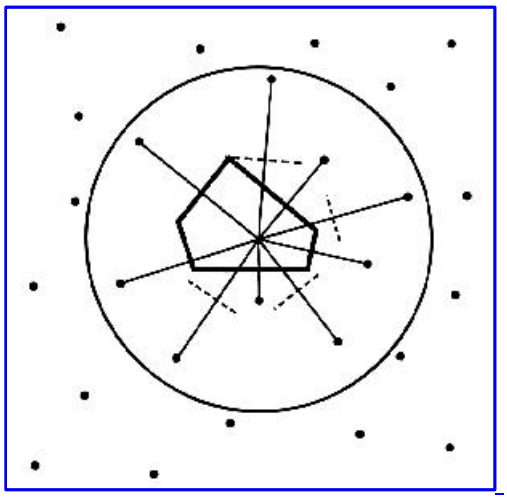

Fig. 1. Description of the Voronoi calculation. To provide a more intuitive description of what occurs at the molecular level, we calculate aggregation using Voronoi polyhedra (Voronoi, 1908). This geometrical construction associates with each atom a unique, limiting polyhedron so that all points within the polyhedron are closer to the enclosed atom than any other. The faces of Voronoi polyhedra are equidistant from two neighboring atoms as defined by the Delauney Tessellation and, therefore, they can be identified uniquely by these two atoms. In the figure, the points denote centers of atoms in a simulation of different-sized atoms, and we have constructed a polyhedra around a central atom. Within a certain cutoff distance (represented by the circle), atoms sharing a solid-line polyhedron face with the central atom are in contact. Atoms sharing broken-line faces are not.

Past simulations relied predominantly on simpler methods such as radial distribution functions or distance cutoffs to measure aggregation. Although generally adequate, we have found that this type of analysis is unable to describe how molecules pack with respect to each other (Tsai et al., 1996). Radial distribution functions measure the number of neighbors at different separations, averaged over the length of a simulation. These functions are useful because they allow direct comparison with scattering experiments. Because radial distribution functions are averages over the local environment, they fail to characterize the instantaneous manner in which molecules pack. In addition, distance alone fails to identify neighbors in solutions containing asymmetric molecules or molecules of mixed sizes. Molecules in such solutions arrange themselves in a nonuniform way that depends upon their shape, size, and chemistry. Situations can occur where two atoms within a neighbor distance cutoff are actually occluded from each other by other atoms, and, conversely, atoms farther away might be in contact (Fig. 2). As a result, distance cutoffs poorly estimate the correct number of contacts.

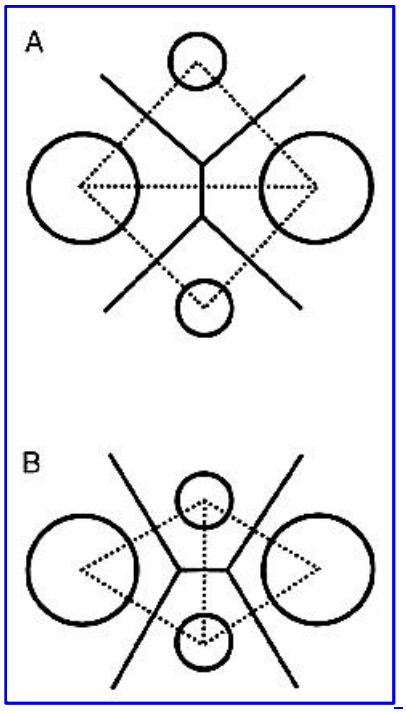

Fig. 2. Neighbors. A: Four atoms are shown by the spheres. Dark lines represent polyhedra faces shared between atoms and broken lines show the Delauney Tessellation. The two large atoms are in contact and occlude the small atoms from touching. B: Same as A, but the two small atoms are moved in closer so that they are now in contact (the Delauney Tessellation clearly shows this as do the polyhedra faces). Although the two larger atoms have not moved from the previous example, they are no longer in contact. Therefore, the two smaller atoms have affected the contact of the two larger atoms without changing the larger atoms' distance of closest approach.

Results and discussion

We compute the fraction of a solute's surface area contacting other solute molecules. For simplicity, we will refer to this quantity as the burial. It is a direct measure of how much the solute interacts with itself. Figure 3 shows the distribution of burial values for urea, methane, and isobutylene from simulations containing 32 solute molecules.

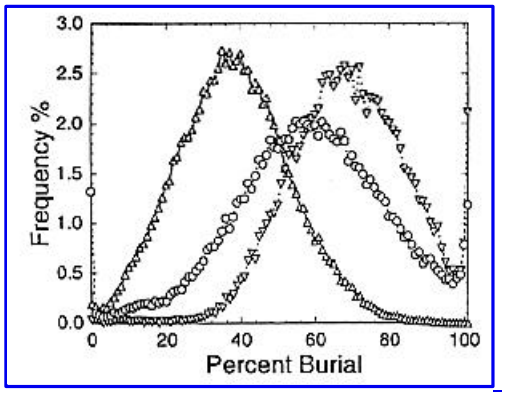

Fig. 3. Burial of solutes. Distributions of percent burial (as defined in the caption to Fig. 1) are shown for simulations containing 32 solute molecules of urea (triangles on a solid line, 6.71 M concentration), isobutylene (inverted triangles on a dashed line, 5.84 M concentration), and methane (circles on a dashed line, 7.83 M concentration).

From the peaks in the distribution, the solute molecules can be divided into three distinct populations: solvated, buried, and exposed. Solvated solute molecules are 0% buried, being completely surrounded by water. Buried solute molecules are 100% buried and completely surrounded by a separate solute phase. Exposed solute molecules have intermediate levels of burial, remaining partly exposed at the interface between water and solute. The methane and isobutylene simulations both contain a population of buried molecules and a population of exposed molecules with distributions centered above 50% burial. These two features in the distribution indicate that these solutes have aggregated into one or more compact structures, which is corroborated by the buried population and by the average number of clusters shown in Figure 4B.

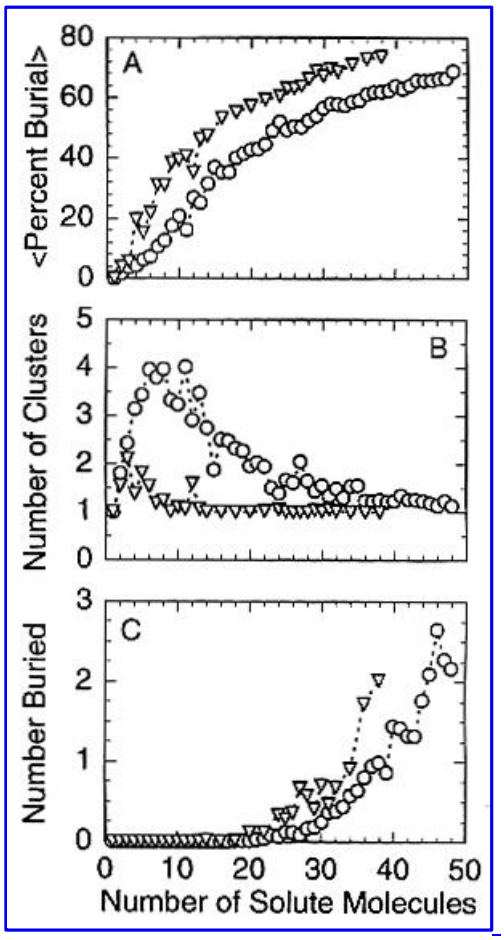

Fig. 4. Mean burial, clusters, and number buried. A: We present the mean percent burial from each methane (open circles on a dashed line) and isobutylene simulation (inverted triangles on a dashed line). The sigmoidal shape of the curves indicates that a transition occurs from free solute to a stable aggregate, and we divide the curve into dispersed, transition, and collapsed stages. The initial portion of both curves is the dispersed stage. The subsequent steep rise marks the transition stage. Here, the isobutylene simulations have a steeper slope than the methane simulations, indicating that the isobutylene molecules aggregate at lower concentrations than methane. The collapsed stage begins as the curves become less steep. B: We show the average number of clusters for each simulation. Again, methane simulations are shown by the open circles and the isobutylene by inverted triangles. These curves clearly indicate the boundaries between the three stages shown above and help describe what is occurring in the simulation. C: Average number of buried solutes are shown. Although isobutylene aggregates sooner than methane, it does not bury molecules at much lower concentrations. This is most likely due to isobutylene's larger size. Symbols represent solutes as explained previously.

Unique to the methane distribution, the solvated population persists at this concentration of 7.83 M. The small size of methane still allows many molecules to be solvated by water. The isobutylene simulation, on the other hand, has no solvated molecules. It only contains the exposed and the buried populations, indicating that the 32 isobutylene molecules form a single stable cluster.

In contrast to the aggregation seen with the hydrophobic solutes, the distribution from the urea simulation (Tsai et al., 1996) shows no buried molecules and only a very small fraction of solvated ones. We find no buried urea at this concentration. The predominant population is exposed and indicates that most of the urea molecules touch each other. Figure 3 also shows that the exposed population is centered below 50% burial, so all of the urea molecules have more than half of their surface area exposed to water molecules. Cluster analysis shows that all the ureas at this concentration spend most of their time in a single aggregate, yet they all touch the water solvent. To satisfy these seemingly opposing views, the urea must form a very open network of contacts, which allows each urea to touch water, but inhibits either complete burial or solvation. Thus, the distribution from the urea simulation describes a uniformly dispersed solute interacting well with the water solvent. This well-mixed solution of urea in water provides a useful contrast to the separation of solute and water seen in the solutions containing hydrophobic molecules.

Each of the three solute populations exhibits certain noteworthy properties. For the solvated population, the simulations require on average 21 waters to envelop a methane molecule and 32 to envelop an isobutylene. Only 8% of a shell water's surface area contacts a solvated methane, whereas 11% contacts an isobutylene. For both sets of solute simulations, we find the existence of the solvent-separated pair as reported by earlier work (Geiger et al., 1979; Panagali et al., 1979; Rapaport & Scheraga, 1982; Watanabe & Andersen, 1986; Laaksonen & Stilbs, 1991). However, it is not a stable configuration in any of the simulations. At higher concentrations, the solvated species sometimes shares waters with the solvation shell of a large cluster, but this situation also does not persist for long periods during the simulation. At the other extreme, between 12 and 13 solute molecules are necessary to surround a buried methane or isobutylene, and each contacts the buried solute with only 8% of its surface area. Of course, isobutylene is a larger molecule, and 8% of its surface area is greater than 8% of a methane's. Because it exists at the solute/solvent interface, the exposed population's contact properties depend on the solute concentration. As the concentration of both isobutylene and methane increases in the simulations, the peak of the exposed population moves steadily toward higher burial values as these solute molecules contact less water. This increase in burial results from the collapse into a single aggregate: as a solute cluster becomes larger, the interface becomes flatter and each solute molecule contacts less solvent.

From our distributions of burial values, we can compute the average percentage burial of a solute molecule as a function of solute concentration (Fig. 4A). For both solutes, the curves have a sigmoidal shape indicative of a transition. Each can be divided into three stages. At low solute concentration, the water solvent is able to accommodate the solute molecules easily, as shown by a near 1:1 correspondence in the average number of clusters and the number of solute molecules in the box (Fig. 4B). We identify this range of concentrations as the dispersed stage, where the solute molecules mix well with the water. This increase in the average number of clusters peaks at three isobutylene and six methane molecules, respectively; we define this as the end of the dispersed stage.

The subsequent decrease in number of clusters (Fig. 4B) suggests that the solute molecules begin to favor larger clusters over smaller ones and explains the sharp rise in the slope of the percent burial starting at seven methane molecules and four isobutylene, respectively (Fig. 4A). This marks the beginning of the transition stage. Over the length of the simulations within this transition, we see that clusters involving

most of the solute molecules form transiently, but then break into smaller clusters or solvated solute molecules, only to reform again. This behavior might indicate that these simulations have not reached equilibrium and that, at longer simulation times, a single stable cluster would form. To test this, we reran the simulation with 16 methane molecules (4.03 M and within the transition stage) starting from a completely clustered configuration. We find this second simulation behaves just like our original simulation and the initial cluster does not persist. Figure 5A shows a distribution of cluster sizes for each of the two 16-methane simulations.

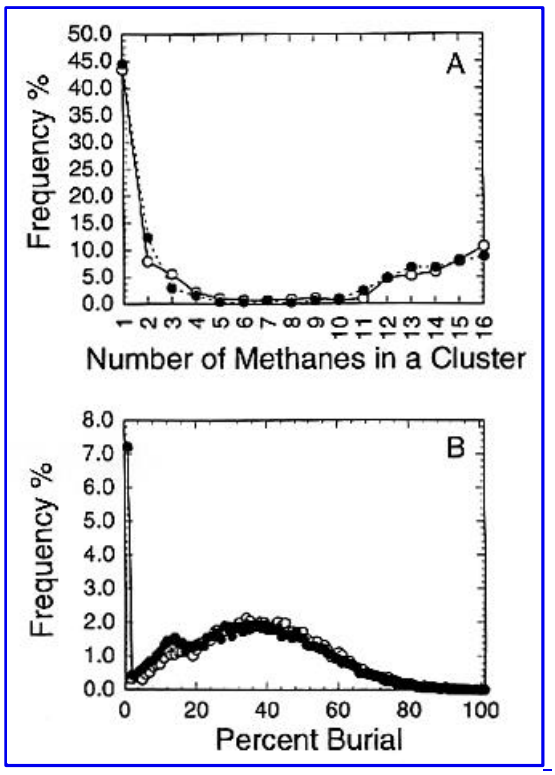

Fig. 5. Comparison of simulations with 16 methane molecules. In both parts, the original run begun from a distributed configuration is shown by the open circles connected by a solid line, and the run started from an aggregated configuration is shown by the filled circles connected by a broken line. A: Frequency of clusters sizes. B: Frequency of percent burial.

The curves almost duplicate each other. If the first simulation had not reached its equilibrium, then we would expect the distributions to differ in that the second simulation would favor formation of larger clusters. To make sure that these clusters form in a similar manner as in the initial simulation, we also compare percent burial for the two runs and find that these curves are also almost identical (Fig. 5B). The results prove that the simulations have reached equilibrium even in the transition stage.

At the highest concentrations, the slope of the burial curve levels off, indicating that the simulations have reached the collapsed stage (Fig. 4A). The average number of clusters decreases and approaches a minimum of 1.0 (Fig. 4B), which clearly shows that a single aggregate containing all the solute molecules dominates in these simulations. For the two solutes, somewhat different criteria mark the beginning of this last stage. Isobutylene, with more than twice the volume of a methane and a more complex structure, aggregates at lower concentrations than methane. Because isobutylene is larger than methane, it aggregates more readily, but we find no buried molecules until about 14 isobutylene molecules (Fig. 4C). This effect is probably due to the isobutylene's large size and planar, Y shape, which requires more molecules and more complex packing in order to bury itself. Simulations with more than nine isobutylene

molecules essentially contain only one cluster, and, by our definition, this concentration marks the start of collapsed stage. For the methane simulations, defining the beginning of the collapsed stage is a little more ambiguous. Methane aggregates more slowly with increasing concentration, and the simulations always contain some of the completely solvated population. As a result, the average number of clusters never reaches 1.0, even with 48 methane molecules in the simulation. This single methane is often dissociated from the cluster formed by all the other methane molecules, which yields two clusters in the simulation box. We therefore define the start of the collapsed stage when the average number of methane clusters drops below 2.0 (20 methane molecules in the box).

All of the previously discussed simulations were run with approximately the same box volume (Tables 1, 2). We expect that the aggregation seen in the above simulations is dependent on concentration and independent of volume. As a simple test, we chose four methane simulations representing different stages of hydrophobic aggregation. For each of these, we halved the volume of box, but kept the concentration the same. Figure 6 compares mean burial and number of clusters from these two set of runs.

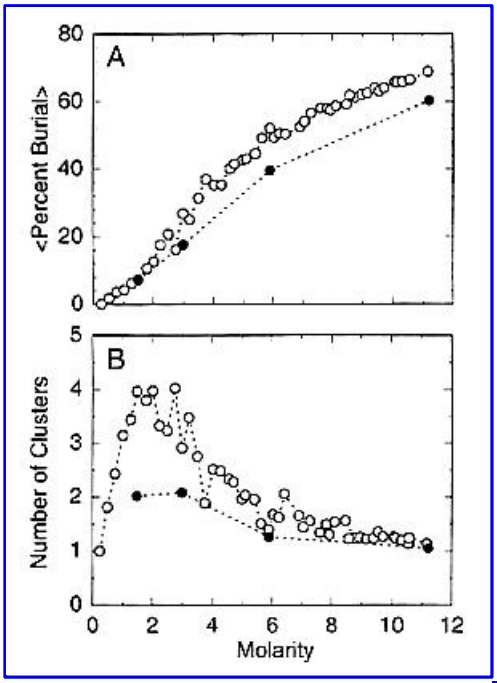

Fig. 6. Box volume comparison. Simulations using our standard reference box volume shown by open circles (connected by a dotted line) are compared to the four simulations at half the box volume but the same concentration shown by filled circles (connected by a broken line). Concentrations of all simulations shown in terms of their molarity. Note that the simulations at half the box volume contain half the number of solutes (methane) and water molecules (see Table 2). A: Mean burial. B: Number of clusters.

For both measures, the runs in smaller volumes basically mimic the shape of the curves using our standard box. This effect is due to the smaller number of solutes used in the simulation: with less solute molecules, the aggregates are of smaller size; the maximum level of burial or number of clusters found with a larger box (i.e., with more solute molecules) is unattainable for the smaller box at the same concentration. Apart from this, the half-box simulations produce the same characteristics as their counterparts run in boxes of twice the volume. This clearly shows that the hydrophobic collapse seen here depends on solute concentration and not on the size of the simulation box.

We present our interpretation of each of the three stages of collapse in Figure 7.

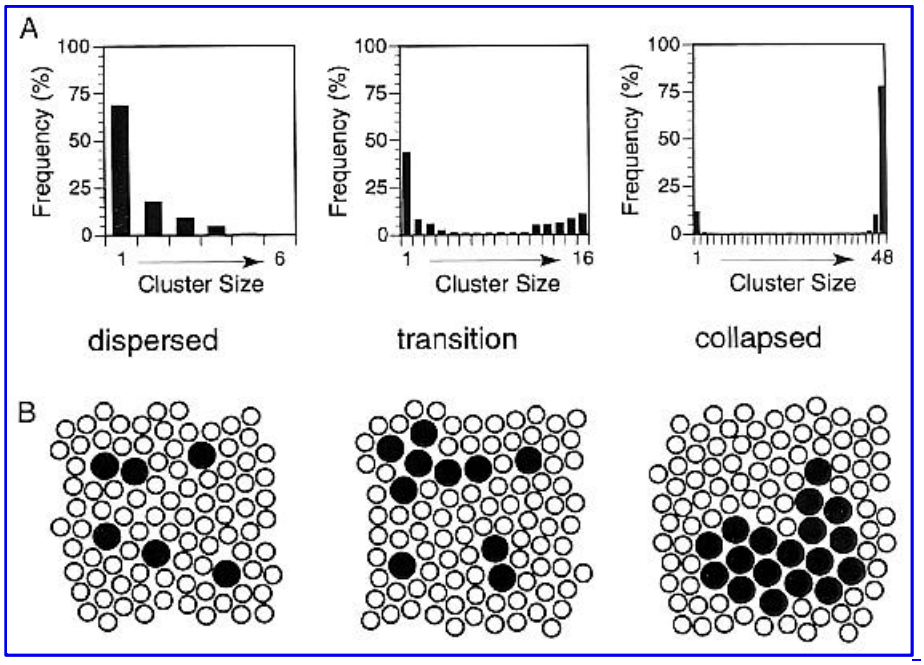

Fig. 7. Three stages of solute aggregation. A: Cluster distributions from representative methane simulations in each of the three stages. For the dispersed stage, we show the distribution from the 6-methane simulation, which clearly favors the solvated species. The distribution illustrative of the transition stage was calculated from the 16-methane simulation, which clearly shows the two overlapping populations located at the two extremes of the distribution. The existence of individual smaller clusters of solute molecules shows that the single aggregate is unstable. The distribution illustrative of the collapsed stage was calculated from the 48 methane simulation, which shows two populations well-separated from each other. None of the intermediate cluster sizes are populated. B: In two dimensions, we depict our interpretation of the three stages found in our simulations. These illustrations were made after extensive viewing of the simulations using space-filling representations of the moving solute molecules in stereographics. Free solute dominates the dispersed stage; some molecules do aggregate, but no large clusters form. We also find some existence of the solute separated pair as shown by the two solutes in the lower left-hand corner of the box. The transition stage cycles between the large, single cluster and a state with several clusters of intermediate sizes. The final collapsed stage has only one cluster, the stable hydrophobic aggregate. In the methane simulations, we would also find a single molecule solvated in solution, and for isobutylene molecules, there is only the single cluster.

The dispersed stage, which occurs at low concentrations, contains solute molecules that prefer to be individually solvated by water and mixed well in solution. The transition stage, which occurs at intermediate concentrations, contains an unstable transient aggregate. The collapsed stage, which occurs at high concentration, contains a single hydrophobic aggregate. In the methane simulations, we find it surprising that a solvated population appears at such high concentrations, because, with isobutylene, the solvated population quickly disappears at low concentrations. One interesting characteristic of all the methane and isobutylene simulations is that we never find water completely buried by solute even in the collapsed stage. This is probably due to the weaker binding free energy of methane to the cluster (Delta*G* = -1.33 kcal/mol for methane and -3.90 kcal/mol for isobutylene, see Table 4). Computing the burial of water by solute, we find that both solutes never cover more than 90% of a water molecule's surface. This result agrees with experimental findings on how difficult it is to bury water in a hydrophobic environment (Wolfenden & Radzicka, 1994).

Table 3 shows the distribution of cluster sizes for some of the different concentrations of isobutylene and

methane simulated here.

Table 3. *Concentration of different sized clusters in simulations of isobutylene and methane*[a]

Run	Box[b] conc.	V_{box} (Å^3)	Cluster size 1	2	3	4	5	6	7	8	9	10	11
Isobutylene													
1	257	6,428	256										
2	506	6,519	275	116									
3	753	6,580	361	110	57								
4	1,004	6,582	87	27	67	165							
5	1,275	6,464	186	48	30	79	114						
6	1,488	6,645	129	4		12	106	126					
7	1,729	6,676	42	2			2	42	205				
8	1,944	6,797	55					2	53	186			
9	2,205	6,739	2							2	242		
10	2,370	6,950	21								18	218	
11	2,530	7,160	13									13	216
Methane in groups of three[c]													
3	753	6,568	628										
6	1,482	6,680	960	37									
9	2,225	6,672	571	118	111								
12	2,971	6,664	450	7	61	170							
15	3,735	6,626	191	4	2	7	231						
18	4,528	6,559	329	7	7	10	35	206					
21	5,080	6,820	266	24	4	2	21	29	186				
24	5,864	6,752	112	4	2	2		2	4	229			
27	6,379	6,983	238	9	16	2	2	7	7	21	193		
30	7,220	6,856	149								14	223	
33	7,882	6,908	107	2			2	2				4	226

[a]Concentration is given in units of mmol/L (mM). To convert to the more natural units of molecules per box, these mM concentration must be divided by the factor (1,000,000/18) × (29.7/V_{box}), where the box of volume is V_{box}, and a mole of water occupies 18 mL (29.7 Å^3/molecule).

[b]Box concentration is the mM concentration of the solute in the box in the particular run.

[c]To simplify this table, three methane molecules are counted together so that a cluster of size 1 contains 1, 2, or 3 methane molecules, whereas a cluster cluster of size 11 contains 31, 32, or 33 methane molecules.

Table 3. *Concentration of different sized clusters in simulations of isobutylene and methane*[a]

The differences in cluster distributions between the different stages in the solute simulations suggest that the aggregation is cooperative with the isobutylene molecules, preferring to be either monomeric at low concentration or completely aggregated at high concentration. Similar behavior is seen for the methane solutions, but the onset of aggregation occurs at higher concentrations. To show this, we count consecutive groups of three methane molecules together to show the larger clusters. To quantify the aggregation and especially the apparent cooperativity, we attempt to model the system by simple equilibrium theory. Defining $[a_i]$ as the concentration in moles/liter of a cluster of size i and $k_{i,j}$ as the equilibrium constant for combining clusters of sizes i and j to form a cluster of size $i + j$, gives

$$[a_2] = k_{1,1}[a_1][a_1],$$

$$[a_3] = k_{1,2}[a_1][a_2],$$

$$[a_4] = k_{1,3}[a_1][a_3] + k_{2,2}[a_2][a_2],$$

$$[a_n] = \Sigma k_{i,n-i}[a_i][a_{n-i}] \quad \text{for } i = 1, n/2.$$

This derivation ignores the rare three-way collisions of clusters, but does allow for all the different combinations of assembling the cluster of size n from all possible pairs of smaller clusters. Because $[a_n]$ depends on the concentration of smaller clusters $[a_i]$, it is easy to express all values of $[a_n]$, in terms of $[a_1]$ and the matrix of equilibrium constants, $k_{i,j}$. The value of $[a_1]$ is then determined using the fact that the total number of solute molecules in the box remains constant at N. This gives

$$N \cdot 1{,}649{,}984/V_{box} = \Sigma i \cdot [a_i] \quad \text{for } i = 1, N,$$

where the scale factor converts concentration from molecules/box to millimoles/liter.

Using this theory, we first calculate the concentrations of the different-sized clusters expected theoretically for the concentration and number of solutes simulated here. We then determine values for the equilibrium constants, $k_{i,j}$, that give the best fit between the theoretical and the simulated cluster concentration. Rather than determine many different $k_{i,j}$ values, we use a very simple model in which

$$k_{i,j} = \exp(\Delta G_{ij}/RT),$$

where RT is the Boltzmann thermal energy at the temperature of the simulation (at 298 K, RT = 0.6 kcal/mol). DeltaG_{ij}, the free energy of association of two clusters of sizes i and j, is taken as

$$\Delta G_{ij} = \Delta G_0 + (i + j - 2)\Delta G_I \quad \text{for } (i + j) < n_o$$

$$= \Delta G_0 + (n_o - 2)\Delta G_I \quad \text{for } (i + j) \geq n_o.$$

These two equations model the energy for cooperative associations. DeltaG_0 is the basic energy for bringing two monomers together, whereas the DeltaG_I is the extra binding energy due to the cooperativity in clusters. n_o sets an upper size limit above which forming a larger cluster provides no further advantage. For example, the binding energy is DeltaG_0 for two monomers, DeltaG_0 + DeltaG_I for a monomer and a dimer, DeltaG_0 + 2DeltaG_I for two dimers, up to a maximum of DeltaG_0 + (n_o - 2)DeltaG_I. For noncooperativity (DeltaG_I = 0), all interactions have the same energy (DeltaG_{ij} = DeltaG_0).

We used a numerical method to find values of DeltaG_0, n_o, and DeltaG_I that optimizes the fit between the theoretical concentrations $[a_n]$ and data from the simulation (Table 4).

Table 4. *Theoretical association energies for different sized clusters derived to fit simulations of isobutylene and methane solutions*

Solute	Residual[a] (%)	n_o	ΔG_0	ΔG_1	ΔG_{11}	ΔG_{12}	ΔG_{13}	ΔG_{19}
			(kcal/mol)					
No cooperativity								
Isobutylene	41	—	−1.66	0	−1.66	−1.66	−1.66	−1.66
Methane	42	—	−0.55	0	−0.55	−0.55	−0.55	−0.55
Cooperativity								
Isobutylene	10	10	−0.34	−0.45	−0.34	−0.78	−1.23	−3.90
Methane	12	10	0.62	−0.24	0.62	0.38	0.13	−1.33

[a]Residual is the sum of squares of the difference of the simulated and theoretical concentrations expressed as a percentage of the sum of squares of the simulated concentrations. For methane, the fitting process used the results for all runs (up to 48 methanes in the box). For isobutylene, we only used the runs with up to 16 isobutylenes in the box so as not to give undue weight to the runs where aggregation is complete.

Table 4. *Theoretical association energies for different sized clusters derived to fit simulations of isobutylene and methane solutions*

It is clear that the fit to the simulation is much better when cooperativity is allowed. The residual error becomes similar to11% with cooperativity, compared to a value of similar to42% without cooperativity (if all theoretical concentrations were set to zero, the residual would be 100%). With a limiting cluster size n_0 of 10, the cooperativity is substantial, with much stronger binding of monomers to larger clusters compared to binding of two monomers (-3.90 versus -0.34 kcal/mol, respectively, for isobutylene, -1.33 versus 0.62 kcal/mol, respectively, for methane). As noticed qualitatively above, three methane molecules behave like one isobutylene, with limiting energies for large clusters of -1.33 kcal/mol for methane and -3.90 kcal/mol for isobutylene. These limiting energies are also reasonable in comparison to their heats of vaporization: 2.13 kcal/mol for methane and 5.74 kcal/mol for isobutylene (Weast, 1979). The cooperativity we observe and quantify here has clear implications for modeling of hydrophobic interactions. The interaction of a single hydrophobic solute with a large cluster is much stronger than that of a pair of such solutes. In fact, a pair of methane solutes do not like to associate, possessing an unfavorable binding energy of 0.62 kcal/mol. This helps explain the results of earlier simulations that failed to find hydrophobic aggregation (Geiger et al., 1979; Panagali et al., 1979; Rapaport & Scheraga, 1982; Watanabe & Andersen, 1986; Laaksonen & Stilbs, 1991).

These calculations show that hydrophobic aggregation is strongly cooperative and explain the energetics behind the three stages of aggregation. In the dispersed stage, the solute molecules do not favor aggregation. The energy of interaction cannot overcome the unfavorable entropy lost upon aggregation. This entropy due to mixing solute and solvent molecules should not be confused with the entropy driving hydrophobic collapse, which is included in the energy of solute interaction. In the transition stage, which occurs at higher concentration, the solute molecules now possess increased binding energy due to cooperativity that is comparable to the unfavorable entropy lost upon aggregation. These opposing energy terms are within kT of each other so that neither dominates. In effect, the solute molecules are not stable in either an aggregated or a dispersed state, but addition of each solute molecule increasingly favors aggregation. In the collapsed stage, the free energy of association has maximized. Now, the association energy is strong enough to form a stable aggregate in solution and can overcome the loss of mixing entropy. As the concentration of solute is increased, there are two different effects contributing to hydrophobic collapse. (1) The mixing entropy opposing aggregation is reduced. (2) The aggregation energy is increased as larger, more cooperative clusters are able to form.

From this description, both solutes begin to favor the aggregated state midway through the transition stage. This occurs when the average number of clusters is midway between the maximum value (2 for isobutylene and 4 for methane, respectively) and the asymptote of 1, i.e., 1.5 isobutylene and 2.5 methane clusters. These average numbers of clusters correspond to the simulations containing 6 isobutylene molecules (1.5 M) and 16 methane (4.0 M), respectively. At these limiting concentrations, the single aggregates formed by both of these solutes have surprisingly similar total volumes of 885 $Å^3$ (for 6 isobutylene molecules) and 1,075 $Å^3$ (for 16 methane molecules). This result suggests that a stable hydrophobic cluster may have a minimum volume that does not depend on the nature of the molecules in the cluster. Because these solute volumes approximate the minimum, stable, hydrophobic volume, we will define them as solute core volumes.

Because proteins show the same general segregation of nonpolar atoms from polar atoms and solvent as seen in these simulations, we were motivated to examine the volumes of the hydrophobic cores in folded, globular proteins and to compare them to the solute core volumes. Of course, proteins are much more complicated than simple hydrophobic solvents in water, including, as they do, many different nonpolar and polar side chains linked by the polypeptide backbone. Using the same methods applied to solution simulations, we calculated the core volumes in the X-ray structures of 31 small globular proteins (Fig. 8).

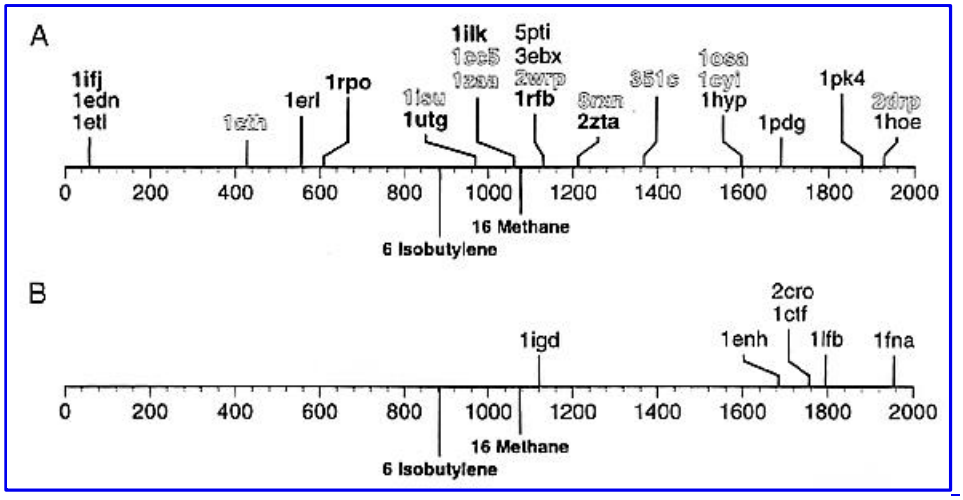

Fig. 8. Comparison with protein core volumes to minimal cluster volumes. To make the plot clearer, we placed all proteins possessing volumes within similar to25 $Å^3$ of each other on the same point. The volumes for the minimum stable hydrophobic aggregate from the 16-methane (1,075 $Å^3$) and the 6-isobutylene (885 $Å^3$) simulations are also shown in both parts. A: Proteins using extra types of stabilization are shown. Depending upon which type of interaction they possess, we classify the proteins as follows: (1) with disulfides (in italics,) (2) oligomers (in bold), and (3) with prosthetic groups (in outline). B: We show proteins without any sort of additional stabilization.

Although the range of hydrophobic core volumes is broad (56-1,955 $Å^3$ for proteins with 13-148 residues), it is striking that additional interactions stabilize all the proteins with core volumes smaller than those of solute cores. Figure 8A shows volumes for proteins with additional stabilization, which can be classified into one of three types: disulfide bonds, oligomerization, and prosthetic groups. Falling well below the solute core volumes, the two smallest proteins (1etl and 1edn) are less than 25 amino acids in length, have small hydrophobic cores (below 500 $Å^3$), and use two disulfides bonds for stability. The classically studied trypsin inhibitor (5pti) maintains its fold with three disulfides and possess a volume close to our solute cores. The inovirus coat protein (1ifj) has a very small hydrophobic core (96 $Å^3$) that

alone is unstable, but it oligomerizes with several thousand monomers to span the length of this filamentous phage's genome. The other oligomeric proteins only dimerize. They fall into three classes: homodimers (uteroglobin, 1utg), intertwined dimers (interferon gamma, 1rfb), and helical coiled-coils (GCN4's leucine zipper, 2zta). Only 1rfb and 2zta, which are dimers in the crystal, show volumes well above the solute core volumes; all the others solved as monomers fall around or below the cutoff. Proteins with prosthetic groups contain a heme (cytochrome *c*3, 1cth), iron clusters (rubredoxin, 8rxn), or zinc (ZIF268 immediate early gene, 1zaa).

In Figure 8B, we show core volumes of proteins without extra stabilizing interactions. All of these possess core volumes above our minimum solute core volumes. Of these cores, the smallest is the third domain of protein G, an immunoglobulin binding protein (1igd), and is quite stable, functioning even at pH 10 (Äkerström & Bjorck, 1986). The DNA binding domains of proteins 1enh, 2cro, and 1lfb do form dimers, but this dimerization is necessary for biological recognition, not for stability. These results show that our solute core volume roughly approximates the minimum hydrophobic core size that is needed to stabilize protein structure independently. Proteins with hydrophobic cores below this limit need additional forms of stabilization, and those with larger cores contain enough hydrophobic stability not to need any extra help.

As shown by the previously discussed comparison of methane simulations run with a smaller box volume and constant concentration, the aggregation seen in these simulations must depend on the concentration of the solute molecules. For both isobutylene and methane, the limiting concentration occurs when the total volume occupied by the solute is about a seventh of the total box volume, with the water solvent occupying six times the volume of the hydrophobic solute. At and above this concentration, the solute molecules have sufficient interaction energy to cause hydrophobic collapse. The fact that protein molecules have a minimum hydrophobic core volume close to what we find here suggests that similar concentrations should occur during protein folding. Measuring the effective concentration of hydrophobic side chains during folding is difficult, but recent small-angle X-ray scattering measurements of the radius of gyration of apomyoglobin (Eliezer et al., 1995) yielded values of 19 Å and 34 Å for the folded and unfolded states, respectively. The results indicate that the volume of the unfolded state is 5.7 $[(34/19)^3]$ times larger than that of the native state. Because the native folded state includes very little water, the concentration of residues in the unfolded state is surprisingly close to that found in our simulations. The expanded volume can be larger as long as the concentration of hydrophobic residues is the same or greater than the effective concentration.

This estimate of residue concentration in an unfolded protein is necessarily crude, neglecting as it does the fact that the residues are linked by the polypeptide chain, whereas our simulations consist of free, unlinked molecules. The polypeptide chain both brings residues together and hinders their closest approach. These two effects would tend to increase and decrease the effective concentration, respectively. In addition, the covalent connections along the polypeptide chain already keep the hydrophobic atoms close together. Upon collapse, a protein loses less entropy of mixing than free solute. On the other hand, the chain stereochemistry limits the freedom of core hydrophobic residues, which increases the penalty of forming an aggregate in comparison to the rather amorphous solute core. Again, these two effects would seem to balance out. Even so, we expect our results to overpredict the limiting concentration for hydrophobic collapse, which, as explained above, our results do. This last difference between proteins and solute cores raises an important point. Being dynamic and unstructured, the solute core better models molten globules. We chose to compare the solute cores with native protein cores instead because there is no comparable data available for molten globules. Although our assumptions are crude, that our results come close to experiment show that these approximations are reasonable for the simple comparison considered here.

Thus, the same sort of hydrophobic aggregation seen in our simulations is similar to the collapse of a protein from the unfolded to the folded state. Furthermore, theoretical modeling of this aggregation supports the view that the hydrophobic collapse in protein folding is a cooperative process (Creighton, 1995). Good agreement is also obtained for the energetics of the process. Using our theoretical model, the free energy of hydrophobic aggregation is between -6.4 and -10.6 kcal/mol (the sum of DeltaG_{ij} for 6 isobutylene molecules or 16 methane, respectively), which compares well with the experimental folding free energies of between 5 and 15 kcal/mol (Pace, 1975).

In conclusion, by simulating simple hydrophobic solutes at different concentrations, we are able to follow hydrophobic collapse and use a simple theoretical model to show that this aggregation is cooperative. Our analysis also allows us to determine the minimum volume above which the solutes favor aggregation into a single cluster. Comparison to protein hydrophobic core volumes shows that the volume of these solute clusters estimates the lower limit to the independent, hydrophobic stabilization of proteins. These results suggest that the hydrophobic collapse found in these simulations closely follows the collapse measured in protein folding studies. As such, this study provides additional support for Kauzmann's (1959) original proposal that hydrophobicity drives protein folding.

Materials and methods

Molecular dynamics simulations

We used the program ENCAD (energy calculation and dynamics) for all simulations. The program and potentials have been described previously (Levitt, 1983; Levitt et al., 1995). For consistency, energy parameters for atom types in the solute molecules were taken as the values used for the same atom types in protein simulations. Solute molecules were distributed evenly in a standard box of 216 waters, and all water molecules closer than 1.67 Å to a nonhydrogen atom of the solute were removed. The box was scaled to the appropriate volume, V_{box}, using the solution's experimental density, rho(M), at molarity M, according to the formula

$$V_{box} = \frac{m_{box}}{\rho(M)} \times \frac{V_w}{m_w},$$

where m_{box} is the mass of the atoms in the box in atomic mass units (a.m.u.), V_w is the volume of water in a pure solution (29.89 Å^3), and m_w is the mass of one water molecule (18 a.m.u.). We used the density of pure methane at the phase transition from liquid to gas to estimate the volume of methane (Wolf et al., 1984-1985), whereas we used volume increments given in Harpaz et al. (1994) to calculate the volume of isobutylene. These simulations parameters are outlined in Tables 1 and 2, respectively. The simulations used our normal protocol (Levitt et al., 1995) of a 2-fs time step, a periodic box, and smooth force-shifting truncation of the nonbonded interactions. The system temperature was equilibrated to 298 K, and coordinates were saved every 0.5 ps. All simulations of methane were run for 1 ns (Table 2), and, running on a DEC Alpha 3000 400 workstation, 1 ns of simulation required about 40 CPU hours. Isobutylene simulations were run for 0.5 ns (Table 1). Data were collected after the initial temperature equilibration (5 ps).

For the second simulation containing 16 methanes, we obtained the initial configuration from the original simulation, choosing a time step where the solute molecules formed a single aggregate. Apart from this, all conditions were the same as in the original simulation.

Half-box simulations were constructed using the simulations with 6, 12, 24, and 48 methane molecules as a reference. First, half of the solute molecules were used. From the box of 216 waters, an exclusion radius was used to obtain half the number of waters as in the reference simulations. This kept the concentration the same. The box size was then scaled accordingly. Simulations followed the same procedure as in the reference runs.

Voronoi and Delauney calculations

We calculate Vornoi polyhedra and the Delauney Tessellation as implemented in Gerstein et al. (1995). Figure 2 shows that environment as well as distance determine whether two atoms are in contact. This point illustrates that the Voronoi method, which considers many-body packing, is better suited for measuring aggregation than distance cutoffs or radial distribution functions, which rely on pairwise distances. Furthermore, we can use the area of this polyhedra face to assess the degree of contact. For each atom, we sum the total area of its faces and the facial area covered by a particular molecule type. These sums are used to find the fraction of a molecule's polyhedra area covered by a given type of molecule (percent burial). To identify the first layer of water surrounding solute clusters, we considered the water molecules that share a polyhedron face with a solute molecule as part of the shell of that solute cluster. Clusters are defined as the set of all the molecules linked to at least one other molecule in the cluster. Links are defined by those pairs of molecules in contact by the Delauney Tessellation (Delauney, 1934). Clusters are built by starting at any solute molecule, which is taken as the start of a cluster. All solute molecules linked to it are marked as being in the same cluster. This is repeated iteratively, making sure that a molecule is not added to the growing cluster more than once. The procedure stops when no new molecules can be added and the entire procedure is then repeated for any molecule still not assigned to a cluster.

Protein set selection

To find a list of suitable proteins, we first began with the list of small proteins from the structural classifications of proteins, SCOP (Murzin et al., 1995). Only one protein from a family was used. This list was then augmented with other proteins containing less than 200 residues. We only used protein structures solved by crystallography, and those reporting all heavy atoms, which gave a total of 124 structures. Using Voronoi polyhedra, the hydrophobic core volumes were then summed for clusters of protein carbon atoms satisfying the following criteria: no exposed surface area, as found from a Connolly surface calculation (Connolly, 1983), and the selected atoms had to touch each other to form a cluster. No prosthetic groups were included. Clusters were found in a similar fashion as in the simulations.

Acknowledgments

This work is supported by the National Institutes of Health (grant number GM41455). We thank Robert Baldwin, Sebastian Doniach, and Doug Laurents for helpful discussions and input.

References (The NLM-formatted Reference List is also available.)

Äkerström B, Bjorck L. 1986. A physiochemical study of protein G molecule with unique immunoglobulin G-binding properties. *J Biol Chem 261*:10240-10247.

Connolly M. 1983. Solvent-accessible surfaces of proteins and nucleic acids. *Science 221*:709-713.

Creighton TE. 1995. An unfolding story: A unified and coherent picture for the mechanism of protein folding is emerging: The crucial factor in folding is the cooperativity of multiple interactions that is required for stability of the folded state. *Curr Biol 5*:353-356.

Delauney B. 1934. Sur la sphére vide. *Bull Acad Sci USSR (VII), Classe Sci Mat Nat*:783-800.

Dill KA. 1990. Dominant forces in protein folding. *Biochemistry 29*:7133-7155.

Eliezer D, Jennings PA, Wright PE, Doniach S, Hodgson KO, Tsuruta H. 1995. The radius of gyration of an apomyoglobin folding intermediate. *Science 270*:487-488.

Finney JL. 1978. Volume occupation, environment, and accessibility in proteins. Environment and molecular area of RNAse-S. *J Mol Biol 96*:721-732.

Geiger A, Rahman A, Stillinger FH. 1979. Molecular dynamics study of the hydration of Lennard-Jones solutes. *J Chem Phys 70*:263-276.

Gerstein M, Tsai J, Levitt J. 1995. The volume of atoms on the protein surface: Calculated from simulation, using Voronoi polyhedra. *J Mol Biol 249*:955-966.

Harpaz Y, Gerstein M, Chothia C. 1994. Volume changes on protein folding. *Structure 2*:641-649.

Kauzmann W. 1959. Some factors in the interpretation of protein denaturation. *Adv Protein Chem 14*:1-63.

Laaksonen A, Stilbs P. 1991. Molecular dynamics and NMR study of methane-water systems. *Mol Phys 74*:747-764.

Levitt M. 1983. Molecular dynamics of native protein. I. Computer simulation of trajectories. *J Mol Biol 168*:595-620.

Levitt M, Hirshberg M, Sharon R, Daggett. 1995. Potential energy function and parameters for simulation of the molecular dynamics of proteins and nucleic acids in solution. *Comp Phys Commun 91*:215-231.

Mancera RL, Buckingham AD. 1995. Further evidence for a temperature-dependent hydrophobic interaction: The aggregation of ethane in aqueous solutions. *Chem Phys Lett 234*:296-303.

Murzin AG, Brenner SE, Hubbard T, Chothia C. 1995. SCOP: A structural classification of proteins database for the investigation of sequences and structures. *J Mol Biol 247*:536-540.

Pace CN. 1975. The stability of globular proteins. *CRC Crit Rev Biochem 3*:1-43.

Panagali C, Rao M, Berne BJ. 1979. Hydrophobic hydration around a pair of apolar species in water. *J Chem Phys 71*:2975-2982.

Privalov PL, Gill SJ. 1988. Stability of protein structure and the hydrophobic interaction. *Adv Protein*

Chem 39:191-234.

Rank JA, Baker D. 1997. A desolvation barrier to hydrophobic cluster formation may contribute to the rate limiting step in protein folding. *Protein Sci 6*:347-354.

Rapaport DC, Scheraga HA. 1982. Hydration of inert solutes. A molecular dynamics study. *J Phys Chem 86*:873.

Richards FM. 1974. The interpretation of protein structures: Total volume, group volume distributions and packing density. *J Mol Biol 82*:1-14.

Shih JP, Sheu SY, Mou CY. 1994. A Voronoi polyhedra analysis of structures of liquid water. *J Phys Chem 100*:2202-2212.

Skipper NT. 1993. Computer simulation of methane-water solutions. Evidence for a temperature-dependent hydrophobic attraction. *Chem Phys Lett 207*:424-429.

Smith DE, Haymet ADJ. 1993. Free energy, entropy, and internal energy of hydrophobic interactions: Computers simulations. *J Phys Chem 98*:6445-6454.

Tsai J, Gerstein M, Levitt M. 1996. Keeping the shape but changing the charges. A simulation study of urea and its iso-steric analogues. *J Phys Chem 104*:9417-9430.

Voronoi GF. 1908. Nouveles applications des paramétres continus á la théorie de formes quadratiques. *J Reine Angew Math 134*:198-287.

Wallqvist A. 1991a. Molecular dynamics study of a hydrophobic aggregate in an aqueous solution of methane. *J Phys Chem 95*:8921-8927.

Wallqvist A. 1991b. Molecular dynamics study of hydrophobic aggregation in water/methane/methanol systems. *Chem Phys Lett 182*:237-241.

Watanabe K, Andersen HC. 1986. Molecular dynamics of the hydrophobic interaction in an aqueous solution of krypton. *Am Chem Soc 90*:795-800.

Weast RC. 1979. Heats of vaporization of organic compounds. In: Weast RC, eds. *CRC handbook of chemistry and physics*. Boca Raton, Florida: CRC Press, Inc. pp C727-C731.

Wolf AV, Brown MG, Prentiss PG. 1984-1985. Concentrative properties of aqueous solutions: Conversion tables. In: Weast RC, ed. *CRC handbook of chemistry and physics*. Boca Raton, Florida: CRC Press, Inc. pp D222-D272.

Wolfenden R, Radzicka A. 1994. On the probability of finding a water molecule in a nonpolar cavity. *Science 265*:936-937.

Return to *Protein Science* Articles Table of Contents

Return to *Protein Science* Home Page

---

prosci@cup.org