Perspectives in Biochemistry

Kinetics versus Thermodynamics in Protein Folding

David Baker[‡] and David A. Agard[*]

Howard Hughes Medical Institute and the Department of Biochemistry and Biophysics, University of California at San Francisco, San Francisco, California 94143-0448

Received March 4, 1994; Revised Manuscript Received April 11, 1994[⊛]

ABSTRACT: Until quite recently it has been generally believed that the observed tertiary structure of a protein is controlled by thermodynamic and not kinetic proceses. In this essay we review several recent results which call into question the universality of the thermodynamic hypothesis and discuss their implications for the understanding of protein folding.

The thermodynamic hypothesis of protein folding has had a long and venerable history. Christian Anfinsen in his 1972 Nobel prize acceptance lecture described the "thermodynamic hypothesis" of protein folding as follows (Anfinson, 1973):

"This hypothesis states that the three-dimensional structure of a native protein in its normal physiological milieu (solvent, pH, ionic strength, presence of other components such as metal ions or prosthetic groups, temperature, and other) is the one in which the Gibbs free energy of the whole system is lowest; that is, that the native conformation is determined by the totality of interatomic interactions and hence by the amino acid sequence, in a given environment."

Here, thermodynamic stability is equated with the idea that the native conformation is determined by the amino acid sequence. Twenty-eight years later, in a 1990 review article, Kim and Baldwin stated "the evidence is good that the final, three dimensional structure of a protein is under thermodynamic, not kinetic control" (Kim & Baldwin, 1990). They cited the experimental fact that for many small proteins, folding and unfolding reactions reach an apparent equilibrium. A *Biochemistry* perspectives article in 1990 reviewed experimental work from 1931 (the reversibility of hemoglobin folding) to 1990 and concluded that "the thermodynamic hypothesis has now been widely established" (Dill, 1990).

Since the writing of the latter two reviews, several experimental results have cast doubt on the universality of the thermodynamic hypothesis. In this essay we review these results and discuss their implications for the understanding of protein folding.

Thus, it now becomes necessary to separate the concept that the native conformation is determined solely from the amino acid sequence from the notion of thermodynamic stability.

THE THERMODYNAMIC HYPOTHESIS

The thermodynamic hypothesis holds that the native conformations of proteins are at global free energy minima relative to all other states having identical bonded chemistry. The experimental evidence cited in support of the thermodynamic hypothesis is that the folding/unfolding reactions of many small proteins are reversible. It is also pointed out that the same native state is reached *in vivo* as *in vitro*, despite the fact that in vivo folding starts from the completed chain. In fact, this apparent distinction between in vitro and in vivo folding may not be valid. Due to the rapid association and slow release of hsp70-type chaperones with the nascent chain (Beckmann et al., 1990), folding in vivo may effectively begin only after synthesis is complete. In any event, these experimental data only argue that the native state is the lowest energy state within a neighborhood of conformational space that includes all kinetically accessible states. However, as conformations outside of this neighborhood cannot be accessed experimentally under normal conditions, the thermodynamic hypothesis is in a sense not falsifiable by experiment. Although in principle the thermodynamic hypothesis could be tested by comparing the lowest energy states found in an exhaustive computer survey of conformational space, such a project is doubly unachievable: current potential functions are highly inaccurate, and the size of the space is far too large to be sampled in our lifetimes.

There are good reasons to think that the native states of proteins may not be at global energy minima. Cyrus Levinthal pointed out many years ago that proteins fold in only a tiny fraction of the time required for an exhaustive search (Levinthal, 1968). Hence, only a tiny fraction of the total possible conformations available to a polypeptide chain can be sampled during folding; this subset of conformations may be viewed as a kinetic pathway. While such pathways must necessarily lead to conformations which are low in energy relative to other accessible states, there is no particular reason, given the vast size of the space, that these low-energy conformations will be global energy minima. There may be large regions of conformational space that are kinetically inaccessible in which a more stable state might exist.

A simple schematic depicting possible conformational free energy surfaces for a hypothetical protein under either thermodynamic or kinetic control is shown in Figure 1. In reality the free energy surface for a protein would be an extremely high dimensional space; for simplicity a one-dimensional cross section through such a surface is shown. Figure 1a depicts a situation where there is a single, global energy minimum that is accessible from any point on the energy surface. Here, the outcome of the folding reaction is independent of the starting configuration, and the reaction can be seen to be under thermodynamic control. By contrast,

[‡] Current address: Department of Biochemistry, University of Washington at Seattle, Seattle, WA 98195.

[⊛] Abstract published in *Advance ACS Abstracts,* June 1, 1994.

0006-2960/94/0433-7505$04.50/0

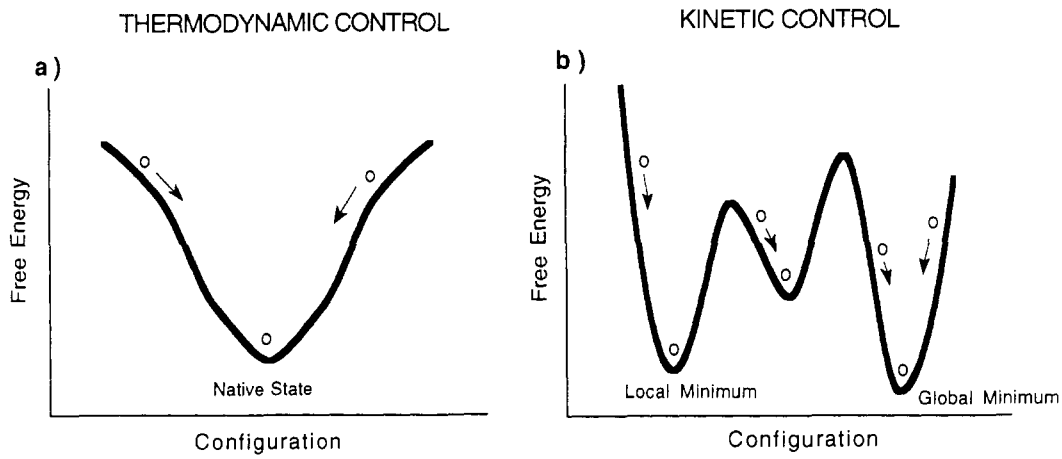

FIGURE 1: Schematic diagram of one-dimensional cross sections through the free energy surfaces of protein folding reactions contrasting two extremes: thermodynamic vs kinetic control. A simple folding surface with a single free energy minimum is shown in panel a. Such a molecule would fold under thermodynamic control, seeking out the most stable state. This is to be contrasted with the considerably more convoluted energy surface in panel b. Because of the high barriers, starting at different locations could lead to different final conformations.

Figure 1b depicts a considerably more convoluted energy surface with multiple energy minima. If the energy barriers are sufficiently high, the outcome of the reaction will depend strongly on the starting point. Molecules starting on the left may very well end up trapped in a local energy minimum, whereas those on the right will rapidly access the global energy minimum. In fact, the actual native state could correspond to either a local or a global minimum. The dependence on initial conditions indicates a reaction under kinetic control.

While it can be difficult to demonstrate that a process is under thermodynamic control, demonstrating kinetic control may be easy. If the final state of a system depends on the initial conditions, the process is kinetically determined (the observation that the same final state is reached from a finite number of different starting conditions, however, does not unequivocally make the case for thermodynamic control). In the following we consider several cases where kinetic control plays a crucial role in determining structure.

INSTANCES OF KINETIC CONTROL

Pro Region Dependent Folding. An increasing number of proteases have been found to be synthesized as proenzymes; that is, the catalytically active protease region is cleaved out of a larger precursor polypeptide. Essentially all known extracellular bacterial proteases are made as pre-proproteins (the pre region being a signal sequence), and a growing number of intracellular and extracellular eukaryotic proteases have also been shown to be synthesized as proenzymes. Pro regions can be amino-terminal extensions, carboxy-terminal extensions, or a combination of the two, amino-terminal extensions being the most common. In addition, there is tremendous variability in the size of the pro regions, ranging from ~40 amino acids to ~60 kDa [for a review, see Baker et al. (1993)].

The two most extensively studied cases of pro region assisted folding involve evolutionarily unrelated prokaryotic serine proteases, α-lytic protease and subtilisin. While both have amino-terminal pro regions, α-lytic protease has a 166 amino acid pro region (Silen et al., 1988), whereas subtilisin has a 70-residue pro region (Wells et al., 1983; Stahl & Ferrari, 1984). Proper folding of the mature protease domains *in vivo* in both cases requires the corresponding pro region, supplied either in *cis* as in the natural protease precursor (Ikemura et al., 1987; Silen et al., 1989) or in *trans* as a separate polypeptide chain (Silen & Agard, 1989). These *in vivo* results have been reproduced *in vitro*; both proteases refold to their native states after denaturation only in the presence of their pro regions (the pro region requirement for subtilisin refolding is not as absolute as for α-lytic protease refolding) (Baker et al., 1992a; Zhu et al., 1989). The pro regions do not function simply by stabilizing the folded form of the enzyme since the native states are very stable even after removal of the pro region.

Recent work has shed some light on the mechanism by which pro regions facilitate folding. Protein folding in general involves a kinetic competition between on-pathway reactions leading to the folded state and nonproductive pathways leading to aggregation. Thus pro regions could function either by increasing the rate of the forward folding reaction or by decreasing the rate of aggregation. The molecular chaperones, which include such proteins as hsp70 and groEL/groES, are thought to function by suppressing protein aggregation [for recent discussions, see Agard (1993) and Martin and Hartl (1993)].

Several lines of evidence suggest that, in contrast to the chaperones, pro regions function by directly increasing the rate of the forward folding reaction. First, pro regions generally interact strongly with the product of the folding reaction, the native state (K_i's range from 10^{-11} to 10^{-7} M) (Baker et al., 1992a; Ohta et al., 1991; Winther & Sorenson, 1991). A second, stronger piece of evidence is that pro regions are required for folding under conditions in which off-pathway reactions are almost completely suppressed. Denatured α-lytic protease cannot refold to the native state in the absence of the pro region. Instead, upon removal of denaturant, the protein folds to an intermediate state that has substantial secondary structure but little organized tertiary structure. This intermediate is stable for months in buffer with no detectable conversion to the native state. However, upon addition of the pro region, the intermediate is rapidly converted to the native state (Baker et al., 1992b). Since folding competence is maintained for an extended period of time, off-pathway reactions are clearly minimal. Instead, the intermediate appears to be kinetically trapped, and the pro region seems to function by directly reducing the free energy of the rate-limiting barrier which blocks access to the native state. Very similar observations have recently been reported for subtilisin, suggesting a similar role for the subtilisin pro region in folding (Eder et al., 1993; Strausberg et al., 1993).

The folding reactions of both subtilisin and α-lytic protease are clearly under kinetic and not thermodynamic control. The intermediate and native states of α–lytic protease are stable

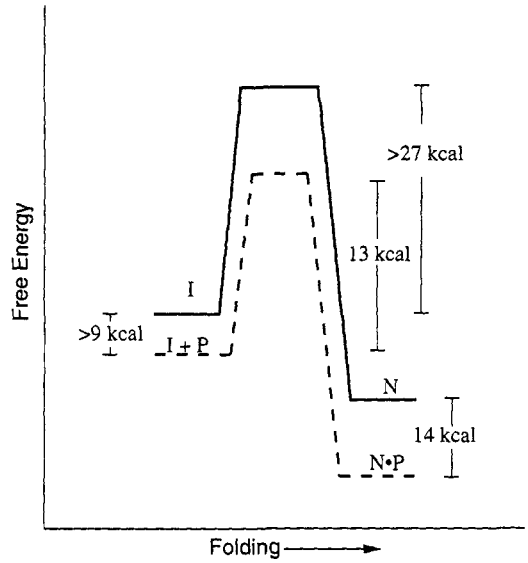

FIGURE 2: Free energy of folding for the trapped α-lytic protease folding intermediate (I). The barrier between I and the native state (N) is extremely high (>27 kcal/M) and cannot be crossed in a finite time, thus rendering N inaccessible. Pro region binding decreases the barrier height by at least 23 kcal/M (with respect to I), thereby speeding folding at least 10^7 -fold.

for an extended period of time under identical conditions with no detectable interconversion; for α-lytic protease they are separated by an energy barrier of at least 27kcal/M (see Figure 2). The native states of both proteases are striking illustrations of the fact that thermodynamic stability need not guarantee kinetic accessibility: in both cases the exceptionally stable native states are not accessible in the absence of the pro regions.

Serpins. A second, dramatic example of kinetic control is provided by the serpin family of protease inhibitors. After synthesis *in vivo* or refolding from denaturant *in vitro*, plasminogen activator inhibitor 1 (PAI-1) folds first to a state which is an active protease inhibitor. Remarkably, this active form slowly converts to an inactive, latent form over a period of several hours (Franke et al., 1990). Structural insight into the conformational change has come from comparison of the structure of ovalbumin, a model for the initial active state (Stein et al., 1990), and the structure of latent plasminogen activator inhibitor 1 (PAI-1) (Mottonen et al., 1992). In ovalbumin, the residues which correspond to the reactive center of PAI-1 form an extended loop with a central α-helix which protrudes from the underlying β-sheet structure of the protein. The reactive site residues are thus poised to insert into the active sites of target proteases. A dramatic rearrangement of these residues is observed in latent PAI-1. The extended loop and helix are gone; instead about half of these residues now form the central strand of a six-stranded β-sheet. The essence of this remarkable conformational change is the insertion of an extended loop into the center of a β-sheet. The burial of the reactive site residues readily explains the lack of protease inhibitor activity of the latent form.

In other serpins, the conversion of the active inhibitor to the inactive state requires proteolysis of the reactive site by the cognate protease. The crystal structure of the cleaved form of the serpin α1-proteinase inhibitor suggests that a conformational change similar to that seen in PAI-1 occurs in other serpins upon proteolytic cleavage (Stein & Chothia, 1991). In addition, it now seems that either the latent or the cleaved forms act as potent cellular signaling molecules.

The latent form of PAI-1 can be converted back to the active inhibitory form by denaturation and renaturation. The inhibitory form is the lowest energy state *accessible* during folding, but since it slowly converts to the latent form, it is clearly not the lowest energy state. Presumably the extensive contacts made by the residues in the reactive loop with the adjacent strands of the β-sheet after insertion are responsible for the greater stability of the latent form. Why the latent form is less accessible than the inhibitory form is less obvious; perhaps the initial five-stranded sheet forms more rapidly and temporarily excludes the reactive loop. In any event, PAI-1 provides a dramatic illustration of a protein where a biologically relevant conformation is clearly not at a global energy minimum. One should also note that if the energy barrier blocking conversion to the latent form were only 1 or 2 kcal/mol higher, there would be no hint that the inhibitory form was metastable. Indeed, this may be the situation in the other serpins which are only inactivated by proteolysis.

Influenza Hemagglutinin. The folding of the influenza virus hemagglutinin (HA) may also be under kinetic control. HA is a trimeric viral envelope glycoprotein which undergoes a dramatic conformational change at low pH, triggering the fusion of endocytosed influenza virus with the endosomal membrane [for a recent review, see White (1993)]. The initially synthesized native conformation (HA-N) differs substantially from the conformation obtained after exposure to low pH (HA-L). The X-ray structure of HA-N shows three globular head domains which sit atop a fibrous stem built around a coiled-coil of three long α-helices, one from each monomer (Wilson et al., 1981). The low-pH-induced conformational change has been studied using a wide range of biochemical and biophysical techniques which taken together suggest a concerted and substantial dissociation of the globular head domains from each other (White, 1994). Recently, Carr and Kim (1993) have made a 27-residue peptide corresponding to a region of the protein predicted to have high coiled-coil propensity but that actually forms a loop in the HA-N structure. While disordered at neutral pH, this peptide forms a coiled-coil structure at low pH. The hypothesis is that this loop to coiled-coil transition is at the heart of the structural changes in HA induced by low pH. X-ray crystallographic data on a large proteolytic fragment of HA-L now coming out of the Wiley lab support this remarkable loop–helix conversion (F. Hughson and O. Wiley, personal communication).

Is the transition from HA-N to HA-L under thermodynamic or kinetic control (see Figure 3)? In the former case, low pH would drive the conformational change by shifting the equilibrium in favor of HA-L (either by destabilizing HA-N or stabilizing HA-L). In the latter case, low pH would reduce the height of an energy barrier blocking access to HA-L from HA-N (in this scenario HA-L would be lower in energy than HA-N at both pH's).

Two lines of evidence suggest that HA-L is the lower energy state and hence that the initially adopted conformation HA-N is metastable. First, the conformational change induced by low pH is irreversible. Not only does HA-L not convert back to HA-N when the pH is raised, but it is more thermostable than HA-N even at the higher pH (Ruigrok et al., 1988). Second, a 52-residue peptide that contains the 27-mer described above folds by itself into a coiled-coil that is stable at neutral pH and even more stable at low pH (Carr & Kim, 1993). An attractive model is that folding at neutral pH leads to a structure under strain, and that lowering the pH allows the strain to be relieved with the adoption of coiled-coil structure by a portion of the chain distorted into a loop in the initial structure.

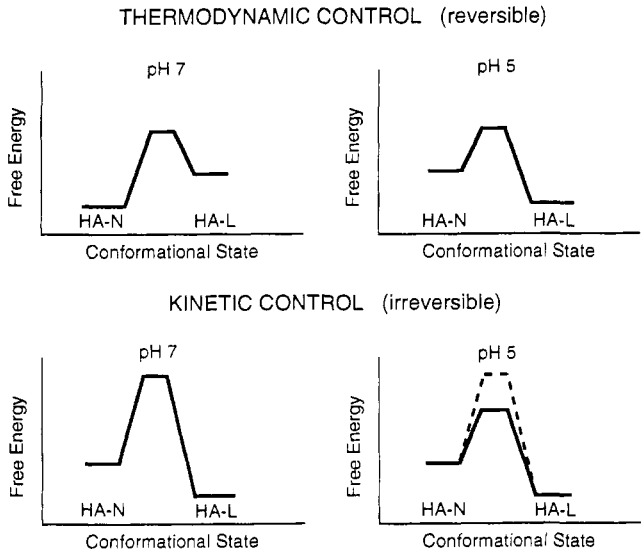

FIGURE 3: Possible free energy diagrams for the pH-induced transition in influenza hemagglutinin. The thermodynamic control situation is depicted in the top panel where the most stable state simply changes as a consequence of altering the pH. Such changes should be fully reversible. By contrast, the bottom panel depicts the system under kinetic control. Here, the initially formed structure at neutral pH (HA-N) is kinetically blocked from obtaining the most stable state (HA-L). Lowering the pH stabilizes the energy barrier, allowing the transition to HA-L. In this case, the processes would be irreversible.

Thus, as in the case of plasminogen activator inhibitor 1 described above, the initial product of folding of influenza hemagglutinin may well be a metastable state.

Luciferase. The folding of the heterodimeric enzyme luciferase also exhibits several striking hallmarks of kinetic control. The active enzyme consists of one α subunit and one β subunit. The β subunit, when allowed to fold in the absence of the α subunit, forms an exceptionally stable homodimer (β_2). Subsequent addition of the α subunit does not result in the formation of the active heterodimer ($\alpha\beta$) because of the very slow rate of disassembly of the β homodimer. However, when the α subunit is present during the folding of the β subunit, only the active heterodimer is observed (Sinclair et al., 1993; Clark et al., 1993; T. Baldwin, personal communication).

Why is the β_2 homodimer not also generated in the reaction $\alpha + \beta \rightarrow \alpha\beta$ given the efficiency of the reaction $\beta + \beta \rightarrow \beta_2$ in the absence of α? Were thermodynamic control to hold, the answer would be the greater stability of $\alpha\beta$ relative to β_2. However, since the formation of both β_2 and $\alpha\beta$ is essentially irreversible on the time scale of these experiments, the reaction $2\alpha + \beta_2 \rightarrow 2\ \alpha\beta$ does not come to equilibrium and hence the relative stability of the two species is simply not relevant. Instead, the formation of only the heterodimer is due to kinetic factors: the rate of the reaction $\alpha + \beta \rightarrow \alpha\beta$ is an order of magnitude faster than that of $\beta + \beta \rightarrow \beta_2$. Because neither $\alpha\beta$ nor β_2 dissociates significantly during the time scale of the experiments, the heterodimerization of luciferase is under kinetic and not thermodynamic control (Baldwin et al., 1993; T. O. Baldwin, personal communication).

The folding of the isolated β subunit contains an additional kinetic twist. If the isolated β subunit is allowed to refold from denaturant at 35 °C, the β homodimer does not form and instead a monomeric molten globule-like state is observed. Remarkably, the high-temperature form does not dimerize when the temperature is subsequently lowered to 25 °C, nor is it capable of interacting productively with the α subunit. Thus folding at 35 °C leads to a state which is separated by a substantial energy barrier from the conformation of the β subunit competent for dimerization (T. O. Baldwin, personal communication).

IMPLICATIONS

The outcomes of all four of the folding reactions described above appear to be determined by kinetic rather than thermodynamic factors. In the first three cases at least two very different conformations of the polypeptide chain are observed, and the first low-energy state accessed during folding appears not to be the lowest energy state (the second state in each case is more thermostable).

How widespread is kinetic control—is it a peculiar quirk of these four examples, or is it a general feature of protein folding? On the one hand, the biology in each of the first three examples may have led to selection for kinetic control. The large energy barrier between the native and intermediate states of α-lytic protease is thought to enhance stability and to protect the native protease from autolysis. Thus, the high-energy barrier and the concomitant need for a disposable pro region to facilitate crossing the barrier may well be the result of selective pressures for maximal protease lifetime in a harsh environment. The self-inactivating property of the serpins is important to the proper regulation of the activity of key enzymes such as tissue plasminogen activator and urokinase. The irreversibility of the conformational change of the influenza hemagglutinin at low pH may be important in driving the fusion of the virus to the host cell endosomal membrane to completion.

On the other hand, the above examples make it clear that the free energy barriers in polypeptide chain conformational space can have appreciable magnitude. If such barriers were a general feature of folding free energy surfaces, then there would be little reason to expect that folding reactions must come to equilibrium.

The assumption that the native state is equivalent to the lowest energy state is implicit in almost all theoretical treatments of protein folding. Computational efforts almost always begin by simplifying a polypeptide chain to make the number of degrees of freedom tractable, and then using one of many different search strategies in an effort to find the global minimum of a specified potential function. Considerable ingenuity has gone into the development of a wide variety of methods for reducing the size of the conformational space, sampling the space more efficiently, and simplifying the potentials. The two main obstacles blocking such computational efforts have been considered to be the large size of the search space and the inaccuracy of current potential functions. However, even with infinite computer time and perfect potential functions, such efforts will only succeed in predicting the native state of a protein from its amino acid sequence if the native state is in fact the lowest energy state. If kinetic control is general, then the global energy minima may be simply irrelevant for biology. To the extent this is the case, computational efforts will be faced with the challenge of mimicking the essential features of protein folding kinetics. Needless to say, the relatively little that is known about kinetics from experiment makes this a perhaps even more formidable obstacle than the ones mentioned above.

Regardless of the question of global vs local minima, the above examples make it clear that conformational free energy surfaces can be considerably more complicated than depicted in Figure 1a. In the first three examples, the outcome of the folding reaction is quite dependent on the starting conditions and is more in keeping with the diagram shown in Figure 1b.

Kinetic control may extend far beyond these few examples. It should be pointed out that rather special conditions must hold for kinetic control to be observable. In order to experimentally detect two distinct states of a protein, either the height of the energy barrier separating them must be in the very narrow range corresponding to conversion times of approximately hours to days (a lower barrier would not lead to the accumulation of the first state, and a higher barrier would block access to the second state) or there must be an external factor (such as a pro region) that modulates the barrier height.

These results raise the intriguing possibility that natural selection may have had to ensure not only the stability of native states but also their accessibility. This suggests that amino acid sequences may have evolved to encode pathways to the native states of proteins.

REFERENCES

Agard, D. A. (1993) *Science 260*, 1903–1904.

Anfinson, C. (1973) *Science 181*, 223–227.

Baker, D., Silen, J. L., & Agard, D. A. (1992a) *Proteins 12*, 339–344.

Baker, D., Sohl, J. L., & Agard, D. A. (1992b) *Nature 356*, 263–265.

Baker, D., Shiav, A. K., & Agard, D. A. (1993) *Curr. Opin. Cell Biol. 5*, 966–970.

Baldwin, T. O., Ziegler, M. M., Chaffotte, A. F., & Goldberg, M. E. (1993) *Journal of Biological Chemistry 268*, 10766–72.

Beckmann, R. P., Mizzen, L. E., & Welch, W. J. (1990) *Science 248*, 850–854.

Bryan, P., Alexander, P., Strausberg, S., Schwarz, F., Lan, W., Gilliland, G., & Gallagher, D. T. (1992) *Biochemistry 31*, 4937–4945.

Carr, C. M., & Kim, P. S. (1993) *Cell 73*, 823–830.

Clark, A. C., Sinclair, J. F., & Baldwin, T. O. (1993) *J. Biol. Chem. 268*, 10773–10779.

Dill, K. A. (1990) *Biochemistry 29*, 7133–55.

Eder, J., Rheinnecker, M., & Fersht, A. R., (1993) *Biochemistry 32*, 18–26.

Franke, A. E., Danley, D. E., Kaczmarek, F. S., Hawrylik, S. J., Gerard, R. D., Lee, S. E., & Geoghegan, K. F. (1990) *Biochim. Biophys. Acta 1037*, 16–23.

Ikemura, H., Takagi, H., & Inouye, M., (1987) *J. Biol. Chem. 262*, 7859–7864.

Kim, P. S., & Baldwin, R. L. (1990) *Annu. Rev. Biochem. 59*, 631–660.

Levinthal, C. (1968) *J. Chim. Phys. 65*, 44–45.

Martin, J., & Hartl, F. U. (1993) *Structure 1*, 161–164.

Mottonen, J., Strand, A., Symersky, J., Sweet, R. M., Danley, D. E., Geoghegan, K. F., Gerard, R. D., & Goldsmith, E. J. (1992) *Nature 355*, 270–273.

Ohta, Y., Hojo, H., Aimoto, S., Kobayashi, T., Zhu, X., Jordan, F., & Inouye, M., (1991) *Mol. Microbiol. 5*, 1507–1510.

Ruigrok, R. W. H., Aitken, A., Calder, L. J., Martin, S. R., Skehel, J. J., Wharton, S. A., Weis, W., & Wiley, D. C. (1988) *J. Gen. Virol. 69*, 2785–2795.

Silen, J. L., & Agard, D. A., (1989) *Nature 341*, 462–464.

Silen, J. L., McGrath, C. N., Smith, K. R., & Agard, D. A., (1988) *Gene 69*, 237–244.

Silen, J. L., Frank, D., Fujishige, A., Bone, R., & Agard, D. A. (1989) *J. Bacteriol. 171*, 1320–1325.

Sinclair, J. F., Waddle, J. J., Waddill, E. F., & Baldwin, T. O. (1993) *Biochemistry 32*, 5036–5044.

Stahl, M. L., & Ferrari, E. (1984) *J. Bacteriol. 158*, 411–418.

Stein, P., & Chothia, C. (1991) *J. Mol. Biol. 221*, 615–621.

Stein, P., Leslie, A. G., Finch, J. T., Turnell, W. G., McLaughlin, P. J., & Carrell, R. W. (1990) *Nature 347*, 99–102.

Strausberg, S., Alexander, P., Wang, L., Schwarz, F., & Bryan, P. (1993) *Biochemistry 32*, 8112–8119.

Wells, J. A., Ferrari, E., Henner, D. J., Estell, D. A., & Chen, E. Y. (1983) *Nucleic Acids Res. 11*, 7911–7925.

White, J. M. (1994) in *Receptor mediated virus entry into cells* (Wimmer, E., Ed.) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (in press).

Wilson, I. A., Skehel, J. J., & Wiley, D. C. (1981) *Nature 289*, 366–367.

Winther, J. R., & Sorensen, P. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 9330–9334.

Zhu, X. L., Ohta, Y., Jordan, F., & Inouye, M. (1989) *Nature 339*, 483–484.