queens of the other genotypes in 15 of the 18 colonies containing *Pgm-3^a/-3^a* queens, a pattern that departs significantly from the null expectation of equal frequencies of higher and lower median weights for *Pgm-3^a/-3^a* queens (sign test; $P < 0.005$). Thus, weight differences among the genotypes are important at the colony level as well as at the population level.

33. At the colony level, the proportion of *Pgm-3^a/-3^a* queens shedding their wings invariably was greater than the proportion of nestmate queens of the other genotypes that shed their wings in the 15 colonies containing two or more *Pgm-3^a/-3^a* queens (sign test; $P < 0.001$).
34. Colony of origin and *Pgm-3* genotype interacted significantly to influence the number of eggs laid ($F = 2.68$; df = 30; $P < 0.001$). However, *Pgm-3^a/-3^a* queens invariably laid as many or more eggs than the median number laid by nestmate queens of the alternate genotypes (sign test comparing median numbers of eggs laid by *Pgm-3^a/-3^a* queens and other nestmate queens; $P < 0.005$), showing that the effect of genotype on fecundity is important in individual colonies as well as the entire population. All statistical analyses on egg counts were done on values transformed as $(x + 0.5)^{1/2}$ [R. R. Sokal and F. J. Rohlf, *Biometry: The Principles and Practice of Statistics in Biological Research* (Freeman, San Francisco, ed. 2, 1981)].
35. Nests 1 to 4 were collected from a monogynous population in northern Georgia in the spring of 1991 and reared for 10 days in the laboratory. These four nests were selected on the basis of having the genotype *Pgm-3^a/-3^a* segregating with *Pgm-3^a/-3^b* among winged daughter queens. The single reproductive queens in monogynous nests mate with a single haploid male (*22*), so that alternate homozygotes at single loci are never present in the same nest (this explains why the genotype *Pgm-3^b/-3^b* is not represented in these four nests). Queens from the mating flight were collected in the summer of 1991 from the same monogynous population as nests 1 to 4 (the polygynous form does not occur in this area). Mating flights of *S. invicta* are synchronized over large areas, and sexuals are capable of flying considerable distances during these flights [G. P. Markin, J. H. Dillier, S. O. Hill, M. S. Blum, H. R. Hermann, *J. Ga. Entomol. Soc.* **6**, 145 (1971)], with the result that these queens are likely to have originated from many different nests (this explains why all three genotypes are represented in this sample).
36. Funded in part by the Swiss National Science Foundation (grant 823A-0283650), the Janggen-Pöhn Stiftung, the National Geographic Society, and the Georgia Agricultural Experiment Stations of the University of Georgia. We thank J. J. Boomsma, D. Haig, N. E. Pierce, H. K. Reeve, and E. O. Wilson for comments on the manuscript.

5 October 1992; accepted 4 February 1993

Kinetics of Folding of the All–β Sheet Protein Interleukin-1β

Paul Varley,* Angela M. Gronenborn,† Henriette Christensen, Paul T. Wingfield, Roger H. Pain,‡ G. Marius Clore†

The folding of the all–β sheet protein, interleukin-1β, was studied with nuclear magnetic resonance (NMR) spectroscopy, circular dichroism, and fluorescence. Ninety percent of the β structure present in the native protein, as monitored by far-ultraviolet circular dichroism, was attained within 25 milliseconds, correlating with the first kinetic phase determined by tryptophan and 1-anilinonaphthalene-8-sulfonate fluorescence. In contrast, formation of stable native secondary structure, as measured by quenched-flow deuterium-hydrogen exchange experiments, began after only 1 second. Results from the NMR experiments indicated the formation of at least two intermediates with half-lives of 0.7 to 1.5 and 15 to 25 seconds. The final stabilization of the secondary structure, however, occurs on a time scale much greater than 25 seconds. These results differ from previous results on mixed α helix–β sheet proteins in which both the α helices and β sheets were stabilized very rapidly (less than 10 to 20 milliseconds).

It has been observed that helix formation can occur within the low millisecond time scale when unfolded proteins are transferred into refolding conditions (*1*). This was perhaps not surprising because the hydrogen bonding interactions involved in helix formation are local, extending over only four residues for one turn of helix. Whereas both α helices and β sheets form secondary structure elements within a protein, hydrogen bonding in β sheet formation involves interactions between distant parts of the polypeptide chain. Hence, it was less expected when quenched-flow deuterium-hydrogen (D-H) exchange protection experiments on mixed α helix–β sheet proteins showed that the kinetics of formation of stable β sheet are similar to those of helices (*2*). Theoretical studies have suggested that only when there is significant stabilization of the β structure will its rate of formation be compatible with the rates of folding of the proteins referred to above (*3*). Whether this rapid stabilization is adequately accounted for by hydrophobic collapse and compaction (*4*) or whether other more specific interactions are required is an important question in protein folding.

To further elucidate the mechanisms of β structure formation during the folding process, we studied the folding of the all–β protein interleukin-1β (IL-1β) (*5*) with techniques sensitive to the conformation, stability, and environment of the secondary structure elements. Interleukin-1β was chosen as a model system for three reasons: (i) complete ^{1}H, ^{15}N, and ^{13}C resonance assignments have been obtained (*6*); (ii) the high-resolution three-dimensional structure has been determined by NMR spectroscopy and x-ray crystallography (*7–9*); and (iii) the equilibria and conditions of its reversible unfolding have been established (*10*). Whereas equilibrium unfolding is two-state, the protein refolds kinetically through a significantly populated, compact intermediate or molten globule that appears to exhibit 80 to 90% of the secondary structure content present in the native protein, as monitored by far-ultraviolet circular dichroism (far-UV CD) ellipticity, but little or no stable tertiary structure (*11*).

The kinetics of β structure formation were monitored by stopped-flow far-UV CD (*12*). At an absorbance of 225 nm, the unfolded and native proteins exhibited mean ellipticities of −3000 and −1200 deg cm^2 dmol^{-1}, respectively, so that the ellipticity becomes less negative as the protein refolds from the unfolded to the native conformation (*10*). On refolding at 4°C, the earliest observed value of ellipticity (after a mixing artifact) was about −1400 deg cm^2 dmol^{-1}, showing that ~90% of the ellipticity was regained within 25 ms in a process with a half-life ($t_{1/2}$) that was less than the dead time (10 ms) of the stopped flow apparatus. Subsequently, there was a slow phase of low amplitude (~10%) with a $t_{1/2} \sim 6$ s.

The rapid formation of β structure as monitored by CD was paralleled by the fluorescence of the single tryptophan residue (Fig. 1A) and bound 1-anilinonaphthalene-8-sulfonate (ANS) (Fig. 1B) (*13*). In the native structure, the fluorescence spectrum of Trp^{120} has a maximum emission wavelength of 344.5 nm (*10*), characteristic of its partially buried state (*7*). Analysis of the tryptophan and ANS fluorescence data suggested a folding pathway with formation of three intermediates

P. Varley, Laboratory of Chemical Physics, Building 2, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health (NIH), Bethesda, MD 20892, and Department of Biochemistry and Genetics, University of Newcastle-upon-Tyne, Newcastle NE2 4HH, United Kingdom.
A. M. Gronenborn and G. M. Clore, Laboratory of Chemical Physics, Building 5, Room 132, National Institute of Diabetes and Digestive and Kidney Diseases, NIH, Bethesda, MD 20892.
H. Christensen and R. H. Pain, Department of Biochemistry and Genetics, University of Newcastle-upon-Tyne, Newcastle NE2 4HH, United Kingdom.
P. T. Wingfield, Protein Expression Laboratory, Building 6B, Office of the Director, NIH, Bethesda, MD 20892.

*Present address: National Institute for Biological Standards and Control, Blanche Lane, South Mimms, Potters Bar, Hertfordshire EN6 3QG, United Kingdom.
†To whom correspondence should be addressed.
‡Present address: Unité de Biochimie Cellulaire, Institut Pasteur, 28 Rue du Dr. Roux, 75724 Paris Cedex 15, France.

$$U \rightarrow X_1 \rightarrow X_2 \rightarrow X_3 \rightarrow N \quad (1)$$

where U and N are the fully unfolded and native states, respectively, and X_i is the intermediate along the folding pathway. The measured half-lives and amplitudes for the first three phases were, for tryptophan and ANS fluorescence, respectively, 3.5 ms (35% of the total fluorescence change) and 4 ms (55% of the maximum fluorescence intensity attained during folding), 46 ms (3%) and 32 ms (21%), and 2.7 s (25%) and 1.1 s (24%). The ANS molecule does not bind to either the unfolded state in 2 M guanidinium hydrochloride (GuHCl) or to native IL-1β, so that in the fourth, slow phase ($t_{1/2} \sim 20$ min) the fluorescence trace returned to its initial value of zero (*14*). Tryptophan fluorescence continued to increase, corresponding to the slow transition from the kinetic molten globule intermediate, X_3, to the native state (*11*) with a $t_{1/2} \sim 20$ min and an amplitude of 37% (*14*).

The kinetic data show that a series of activation energy barriers exist on the folding pathway, resulting in several intermediates being significantly populated. The first major phase in the recovery of far-UV CD ellipticity either corresponds to or is faster than the first step in the above pathway. At the end of this first step, 90% of the β structure of the native protein has been formed together with a major collapse of the polypeptide chain, as indicated by the change in environment of Trp^{120} and by the formation of sufficiently extensive hydrophobic clusters to bind ANS. The second phase of change in far-UV CD ellipticity corresponds to step three in scheme (*1*), by which time Trp^{120} is substantially buried and the compact molten globule intermediate X_3 (*11*) has reached its maximum capacity for ANS binding. Beyond this stage in folding, no further change in far-UV CD ellipticity at 225 nm is detected. The native, high-density packing of nonpolar residues within the hydrophobic core as monitored by the desorption of ANS, on the other hand, takes place on a time scale of many minutes ($t_{1/2} \sim 20$ min) (*15*).

To correlate the CD and fluorescence kinetics with structure, we followed the protection of backbone amide groups (ND) against exchange with solvent (H_2O) by means of quenched-flow D-H exchange experiments (*1, 2*) in combination with two-dimensional (2D) ^1H-^{15}N correlation spectroscopy. In this experiment, fully unfolded protein in D_2O at pH 5 is allowed to fold for a length of time T, after which it is subjected to a 16-ms exchange pulse in H_2O at pH 9.3, followed immediately by a quench pulse at pH 5 (*16*). Backbone amide groups that participate in stable native hydrogen bonds at the time the exchange pulse is applied are protected against D-H exchange, thereby permitting the time course of formation of the various hydrogen bonds in the native structure that are associated with very slowly exchanging amide groups ($t_{1/2} > 10^3$ min for H-D exchange in the native protein at pH 5) (*17*) to be followed.

A series of ^1H-^{15}N NMR correlation spectra of samples in which the labeling pulse was taken at T = 0, 1, and 25 s after initiation of folding from the unfolded state at 4°C is shown in Fig. 2. At T = 0 s, obtained after initiation of folding by diluting unfolded protein in D_2O directly into pulse buffer, no protection occurred, and all 47 slowly exchanging amide protons of the native protein (out of a total of 144 NH protons) were observed at their expected intensities (Fig. 2A). Hence, no stable backbone hydrogen bonds were formed during the dead time (3 to 5 ms) of the quenched-flow apparatus. Similarly, no protection was observed after 50 ms of refolding. Therefore, the rapid formation of β structure observed by far-UV CD and the hydrophobic collapse observed by fluorescence were not correlated with the formation of stable β sheet structure. At 1 s, however, there were 21 NH protons with resonance intensities that decreased by 50 to 70% (Fig. 2B). At 25 s, these NH protons could no longer be observed, and

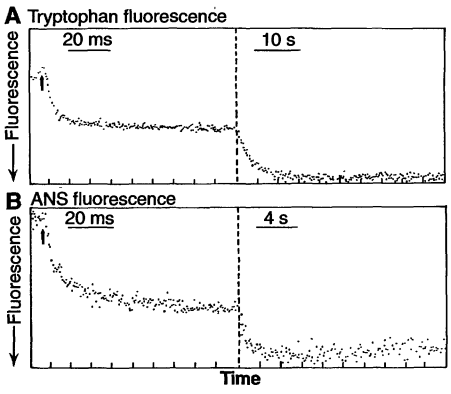

Fig. 1. Refolding of IL-1β at 4°C monitored by (**A**) tryptophan and (**B**) ANS fluorescence. The fluorescence is recorded in arbitrary units. The zero time of the experiment is marked with a vertical arrow. In both cases the fluorescence intensity extrapolates to a value characteristic of the unfolded protein at t = 0 s. The fluorescence intensity change shown in (A) represents 63% of the total increase in fluorescence observed during refolding. The remaining 37% of the increase in Trp fluorescence is associated with a slow phase with $t_{1/2} \sim 20$ min.

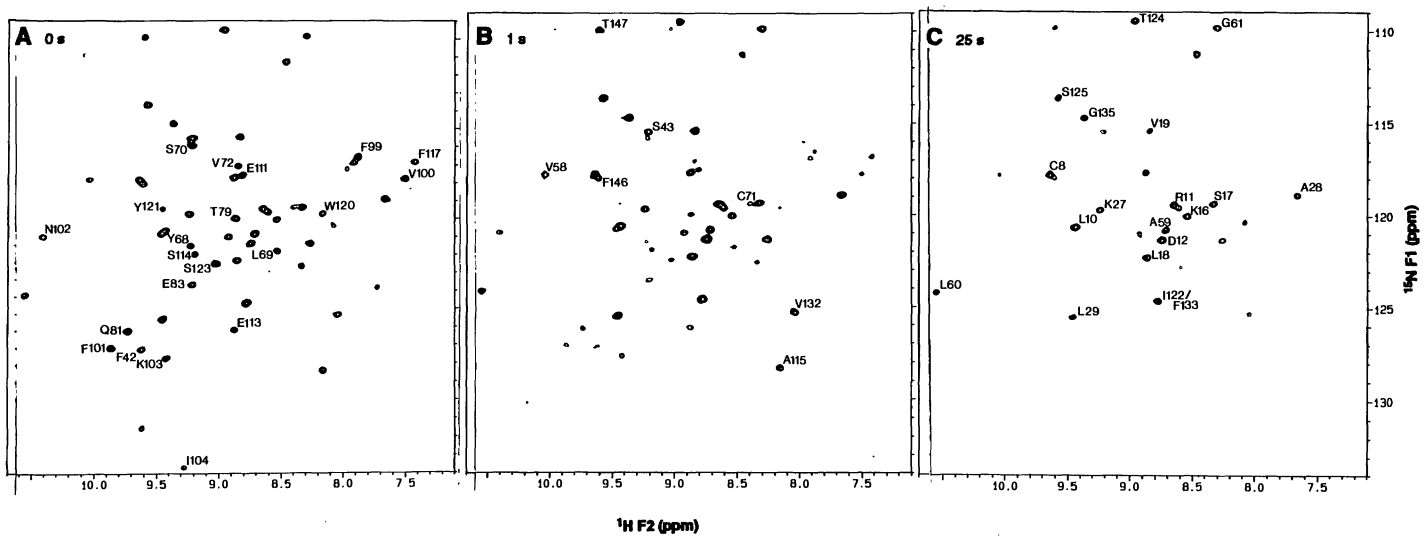

Fig. 2. ^1H-^{15}N correlation spectra of IL-1β showing the extent of amide group protection from D-H exchange at (**A**) 0 s, (**B**) 1 s, and (**C**) 25 s after initiation of protein folding at 4°C. The residues with a $t_{1/2}$ of 0.7 to 1.5, 15 to 25, and much greater than 25 s are labeled in the respective spectra.

the resonance intensities of another seven NH protons decreased by 50 to 70% (Fig. 2C). Also seen at 25 s were 19 NH protons with resonance intensities that remained essentially unchanged and that were therefore not protected at all against D-H exchange at this time, although at longer times the control experiments showed that they became protected as with the native protein. Thus, the NMR data provide evidence for at least two distinct intermediates formed with half-lives of 0.7 to 1.5 and 15 to 25 s that correlate with the formation of intermediate X_3 and another intermediate, X_4, between X_3 and the final native state N in the kinetic scheme (1) based on the fluorescence results (that is, $X_3 \rightarrow X_4 \rightarrow N$).

The locations of the three classes of amide protons within the secondary and tertiary structures of the native protein are summarized in Figs. 3 and 4, respectively. IL-1β displays threefold pseudosymmetry comprising three topological units, A, B, and C, each consisting of four antiparallel β strands (*7, 8, 17*). The earliest event ($t_{1/2} \sim$ 0.7 to 1.5 s) observed by NMR is associated with the formation of most of topological unit B, the interface between topological units B and C, and part of topological unit C, involving 11, 4, and 5 hydrogen bonds, respectively, that are associated with slowly exchanging amide protons in the native protein. In addition, a single hydrogen bond between β strands 1 and 4 in topological unit A is observed. In the second event ($t_{1/2} \sim$ 15 to 25 s), two additional hydrogen bonds are added to topological unit B, two to the interface between topological units A and C, and one each to topological unit C and the B-C and A-C interfaces. The last event ($t_{1/2} >>$ 25 s) principally involves the formation of topological unit A (11 hydrogen bonds), the A-B (2 hydrogen bonds) and A-C (1 hydrogen bond) interfaces, and the completion of topological unit C (5 hydrogen bonds). Of the 19 amide protons that remain unprotected against D-H exchange at 25 s, four are involved in bridging backbone hydrogen bonds associated with tightly bound internal water molecules (*7, 18*): in particular, Leu^{10} NH (strand 1) to Val^{40} O (strand 4) and Leu^{18} NH (strand 2) to Leu^{10} O (strand 1), which stabilize the internal strands of topological unit A; Leu^{60} NH (strand 5) to Leu^{69} O (strand 6) and Phe^{99} O (strand 8), which stabilize the internal strands of topological unit B; and Ser^{125} NH (strand 10) to Pro^{131} O (strand 11), which stabilizes the turn between β strands 10 and 11 in topological unit C.

The difficulties in assessing the content of β structure—and in particular in equating it to native-like β sheet structure—from the CD spectra of proteins that also contain helix, together with the "overshoot" anomalies seen with some proteins during refolding (*19, 20*), suggest caution in detailed structural interpretation in such cases. With IL-1β, however, an all–β sheet protein that exhibits no overshoot phenomena and that has no disulfide bonds, the rapid acquisition of 90% of the total far-UV CD ellipticity can only be ascribed to the formation of unstable β-like secondary structure, the precise form of which must remain the subject of further investigation. It is clear from the NMR data that any backbone hydrogen bonds that are formed in the early β structure seen by far-UV CD are not stable either because of local breathing or sliding of one strand relative to another, resulting in the rapid formation and breaking of hydrogen bonds. The lifetime of these hydrogen bonds for D-H exchange must be less than 16 ms, the duration of the high pH exchange pulse. This is more than five orders of magnitude shorter than the D-H exchange lifetimes observed in the native protein at the same pH (*21*).

In contrast, the stabilization of native hydrogen bonds within the β sheets of IL-1β occurs much more slowly, on the order of seconds, in a process that involves at least two intermediates. The first is formed with a $t_{1/2}$ of 0.7 to 1.5 s, which correlates with the third phase in the ANS binding kinetics, and is associated with a substantial increase in ANS fluorescence. Thus, although a significant amount of native secondary structure is formed at this stage, comprising most of topological unit B and part of topological unit C, large clusters of hydrophobic residues are available for ANS binding. The second intermediate is formed with an apparent $t_{1/2}$ of 15 to 25 s, which is somewhat slower than the second phase in the far-UV CD kinetics ($t_{1/2} \sim$ 6 s). During this time frame, desorption of ANS begins to occur, implying that hydrophobic residues are being buried within the protein core. The complete formation of stable secondary structure, which is principally associated with the formation of topological unit A, has not yet taken place at 25 s and probably occurs on a time scale similar to the late phase of packing of Trp^{120} into its native environment and the final desorption of ANS ($t_{1/2} \sim$ 20 min).

The present results obtained for IL-1β contrast with the very early stabilization (on the order of milliseconds) of native β structure observed in mixed α-β proteins

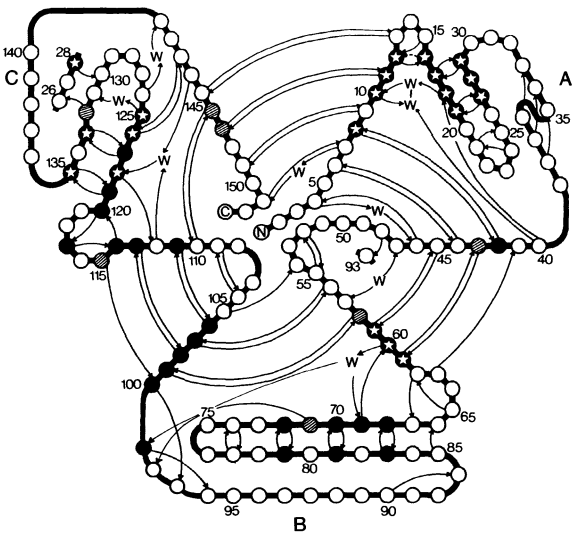

Fig. 3. Schematic diagram of the hydrogen bonding and topology of IL-1β illustrating the sequential formation of secondary structure as observed in the quenched-flow D-H amide protection experiments. The positions of amide groups that are protected during folding with a $t_{1/2}$ of 0.7 to 1.5, 15 to 25, and much greater than 25 s are indicated by the black circles, the hatched circles, and the black circles with a white star, respectively. The three pseudosymmetric topological units, A, B, and C, are formed by β strands 1 to 4, 5 to 8, and 9 to 12, respectively. Hydrogen bonds are marked by arrows from the donor backbone NH group to the acceptor group (backbone carbonyl or water), and waters involved in bridging hydrogen bonds are indicated by the letter W. The hydrogen bond from Lys^{16} NH to the side chain O^{δ} atom of Asn^{12} is indicated by a dashed line.

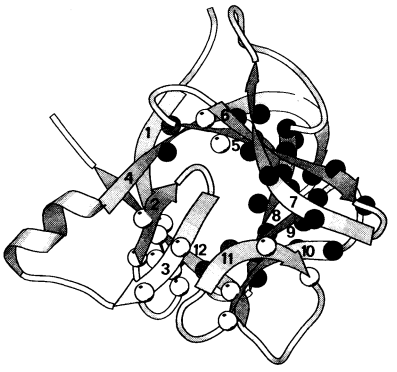

Fig. 4. Ribbon diagram of the polypeptide fold of IL-1β illustrating the spatial location of the amide groups protected during folding. The residues of amide groups protected with a $t_{1/2}$ of 0.7 to 1.5, 15 to 25, and much greater than 25 s are shown as black, gray, and white circles, respectively. Beta strands 1, 4, 5, 8, 9, and 12 form a six-stranded antiparallel β barrel that is closed off at one end by the remaining six strands. In the view shown, the open face of the barrel is located in the back of the structure. The ribbon drawing was made with the program Molscript (*22*).

(2). In all these cases, the regions of native β sheet that are detected first are those involved in tertiary interactions with a helix (or helices) that are also stably formed early in the folding process. In the native structure of the all–β sheet protein IL-1β, on the other hand, the three pseudosymmetric elements of β sheet do not form such interactions. Each β strand contributes two or three nonpolar residues to a hydrophobic core that depends for its stability on the hydrophobic and van der Waals interactions of a large number of tightly packed side chains. We therefore suggest that folding to the stable native structure for this type of protein involves the rapid formation of β structure around a nonpolar core, followed by the much slower stabilization of native secondary structure that accompanies the progressive final tight packing of the core groups or those groups external to the β sheets, or both.

REFERENCES AND NOTES

1. H. Roder, G. A. Elöve, S. W. Englander, *Nature* **335**, 700 (1988).
2. J. B. Udgaonkar and R. L. Baldwin, *ibid.*, p. 694; M. Bycroft, A. Matouschek, J. T. Kellis, A. R. Fersht, *ibid.* **346**, 488 (1990); M. S. Briggs and H. Roder, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 2017 (1992); J. Lu and F. W. Dahlquist, *Biochemistry* **31**, 4749 (1992); S. E. Radford, C. M. Dobson, P. A. Evans, *Nature* **358**, 302 (1992).
3. A. V. Finkelstein, *Proteins* **9**, 23 (1991).
4. K. A. Dill, *Biochemistry* **24**, 1501 (1985).
5. IL-1β is made up of 12 antiparallel β strands connected by turns, short loops, or long loops (*7*). The long loop (residues 30 to 39) connecting β strands 3 and 4 is unusual in that it contains a type II turn (residues 32 to 35) followed by 2.3 turns of a 3_{10} helix (residues 33 to 39) characterized by four CO(*i*)-NH(*i* + 3) hydrogen bonds; the NH protons of residues 36 to 39, however, are not slowly exchanging. Because this loop is surface-exposed and the short 3_{10} helix lies outside the core of the protein, IL-1β can still be considered to be an all–β sheet protein.
6. P. C. Driscoll, G. M. Clore, D. Marion, P. T. Wingfield, A. M. Gronenborn, *Biochemistry* **29**, 3542 (1990); G. M. Clore, A. Bax, P. C. Driscoll, P. T. Wingfield, A. M. Gronenborn, *ibid.*, p. 8172.
7. G. M. Clore, P. T. Wingfield, A. M. Gronenborn, *ibid.* **30**, 2315 (1991).
8. B. C. Finzel *et al.*, *J. Mol. Biol.* **209**, 779 (1989); J. P. Priestle, H. P. Schär, M. G. Grütter, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 9667 (1989); B. Veerpandian *et al.*, *Proteins* **12**, 10 (1992).
9. B. Shaanan *et al.*, *Science* **257**, 961 (1992).
10. S. Craig, U. Schmeissner, P. T. Wingfield, R. H. Pain, *Biochemistry* **26**, 3570 (1987).
11. O. B. Ptitsyn, R. H. Pain, G. V. Semisotnov, E. Zerovnik, O. I. Razgulyaev, *FEBS Lett.* **262**, 20 (1990).
12. Rapid kinetics of the change in far-UV CD were measured with a Biologic SFM-3 stopped-flow mixing system fitted with a cell of path length 1 mm installed on a Jobin Yvon Dichrographe Mk 6 (*14*). Folding at 4°C was initiated by mixing of 1 volume of unfolded IL-1β (1.5 mg/ml) in 2 M GuHCl with 5 volumes of 100 mM sodium acetate (pH 6.5). Typically, the results of 12 separate runs were accumulated. At the end of each experiment the cell was washed with acetate buffer. Controls were performed in which unfolded protein in GuHCl buffer was mixed with the same buffer.
13. Folding at 4°C was initiated by mixing of 1 volume of unfolded IL-1β (1 mg/ml) in 2 M GuHCl with 5 volumes of 10 mM MES, 100 mM NaCl, and, in the case of the ANS experiments, 0.1 mM ANS (pH 6.5) with a Hi-Tech stopped-flow apparatus. Tryptophan and ANS fluorescence were excited at 290 and 380 nm, respectively, and measured above 400 and 430 nm, respectively, with high-pass filters.
14. P. Varley *et al.*, unpublished results.
15. Previous experiments in which the protein was unfolded in GuHCl for different lengths of time before the initiation of refolding have indicated that this slow phase is not due to cis-trans proline isomerization giving rise to both fast and slow refolding populations of unfolded and partially folded molecules (*10*).
16. Quenched-flow mixing was performed with a Biologic QFM-5 module. IL-1β (~4 mg/ml and uniformly labeled to >95% with ^{15}N), prepared as described in (*7*), was unfolded in D_2O for at least 1 hour in 3 M guanidinium deuterochloride (GuDCl)–50 mM sodium d_3-acetate (pH 5.05) to allow complete denaturation and H-D exchange. The slightly lower pH than that used for the fluorescence and CD kinetics was necessitated by the greater tendency of IL-1β to aggregate at the higher concentrations required for the D-H exchange experiment. Refolding at 4°C was initiated by dilution with 7 volumes of folding buffer [50 mM sodium d_3-acetate (pH 5.05) in H_2O for folding times of ≤1 s and in D_2O for folding times >1 s]. After a selected time T of refolding, the sample was pulsed to rapidly exchange deuterons for protons on nonprotected amide groups by dilution with 5 volumes of pulse buffer [100 mM glycine-NaOH (pH 9.75)] to produce a final pH of 9.3. After 16 ms, the pulse was ended by a further 1:1 dilution with 0.5 M sodium d_3-acetate (pH 5.05) to give a final pH of 5.10, and the protein was allowed to fold to completion (as determined by NMR). For T = 25 s, mixing was carried out manually, and exchange of nonprotected amide groups, 25 s after the initiation of folding, was attained by dilution of the folding mixture with 10 volumes of 500 mM sodium acetate in H_2O (pH 5.05). The resulting samples were then concentrated and exchanged into 50 mM sodium d_3-acetate in D_2O (pH 5.05) with Amicon Centriprep 10 concentrators. Refolding was carried out at relatively low protein concentrations to avoid complications arising from aggregation. The extent of protection was then measured by recording a ^1H-^{15}N Overbodenhausen correlation spectrum [A. Bax *et al.*, *J. Magn. Reson.* **86**, 304 (1990)] on a Bruker AMX 600 spectrometer. The concentration of IL-1β in the NMR samples was 0.3 to 0.4 mM. To compare cross-peak intensities between samples taken at different refolding times T, we normalized the spectra (relative to the T = 0 s spectrum) to correct for the different concentrations of protein in the samples (as measured spectrophotometrically by the absorbance at 280 nm and by the intensity of the resonances in the aliphatic region of the ^1H-NMR spectrum) and the concentration of D_2O relative to H_2O at the time of the exchange pulse.
17. P. C. Driscoll, A. M. Gronenborn, P. T. Wingfield, G. M. Clore, *Biochemistry* **29**, 4668 (1990).
18. G. M. Clore, A. Bax, P. T. Wingfield, A. M. Gronenborn, *ibid.*, p. 5671.
19. K. Kuwajima, H. Yamaya, S. Miwa, S. Sugai, T. Nagamura, *FEBS Lett.* **221**, 115 (1987).
20. A. F. Chaffotte, Y. Guillou, M. E. Goldberg, *Biochemistry* **31**, 9694 (1992).
21. H-D exchange experiments on native IL-1β at pH 9.3, using ^1H-^{15}N Overbodenhausen correlation spectroscopy to monitor the NH → ND exchange as a function of time after taking up a lyophilized sample of IL-1β in D_2O, indicated that after 1 hour there was no change in the intensities of the 47 slowly exchanging NH protons (that is, those with a $t_{1/2}$ > 3000 min at pH 5).
22. P. J. Kraulis, *J. Appl. Crystallogr.* **24**, 946 (1991).

23. Supported by the Intramural AIDS Targeted Antiviral Program of the Office of the Director of the National Institutes of Health (to G.M.C. and A.M.G.) and by a grant from the Science and Engineering Research Council (to R.H.P.).

7 December 1992; accepted 12 February 1993

Substrate Phage: Selection of Protease Substrates by Monovalent Phage Display

David J. Matthews and James A. Wells*

A method is described here for identifying good protease substrates among approximately 10^7 possible sequences. A library of fusion proteins was constructed containing an amino-terminal domain used to bind to an affinity support, followed by a randomized protease substrate sequence and the carboxyl-terminal domain of M13 gene III. Each fusion protein was displayed as a single copy on filamentous phagemid particles (substrate phage). Phage were then bound to an affinity support and treated with the protease of interest. Phage with good protease substrates were released, whereas phage with substrates that resisted proteolysis remained bound. After several rounds of binding, proteolysis, and phagemid propagation, sensitive and resistant substrate sequences were identified for two different proteases, a variant of subtilisin and factor X_a. The technique may also be useful for studying the sequence specificity of a variety of posttranslational modifications.

Proteolysis is a common form of posttranslational modification and is important in regulation and protein turnover (*1*). Knowledge of protease specificity aids in the identification of biologically relevant substrates, helps direct the design of specific inhibitors, and is useful in applying proteases for site-specific proteolysis (*2*). Substrate sequences for proteases often extend over seven or eight contiguous residues (*3*), and thus one protease can potentially interact with a vast number of possible substrates.

We present a method, called substrate phage, in which more than 10^7 potential substrates can be tested concurrently. Each substrate sequence is displayed as a single

Department of Protein Engineering, Genentech, South San Francisco, CA 94080.

*To whom correspondence should be addressed.