LETTERS TO NATURE

FIG. 3 a, Electron density representation of the active site residues Ser 60 and Lys 97, with a contour level of 1.2σ. A stick model of the structure has been put into the density to show the relevant atoms and bonds. Water molecules are shown as red stars. b, A surface representation of the UmuD’ protein. The cleft leading to the active-site residues is clear in this representation. The surface contributed by the Ser 60 residue is coloured red and that by the Lys 97 residue is coloured blue. The residues that contribute to the molecular dimer interface have their surface contributions coloured violet (Tyr 52, Val 54, Ile 87, Gly 92, Gly 93, Phe 94, Phe 128), c, Worm representation of the UmuD’ backbone (yellow) with the sidechains of Ser 60 and Lys 97 in red and blue, respectively. The backbone carbonyl oxygen of residue Val 96 is in red. Two segments of the TEM1 β-lactamase structure (in magenta). The first fragment includes residues 69–74 with the Ser 70 and the Lys 73 side chains in red and blue, respectively. The second fragment includes residues 129–133 with the Asn 132 side chain in yellow. An r.m.s. deviation of 0.12 Å is found when superimposing the Ser 60/Ser 70 Oy, the Lys 97/Lys 73 N, and the Val 96 O with Asn 132 Oy atoms. In the reaction proposed in ref. 17, Ser 70 Oy attacks the carbonyl carbon of the β-lactam ring after activation by proton transfer to Lys 73 N, 2.7 Å away. The Lys 73 N is positioned by hydrogen bonds to the Asn 132 Oy and the Ser 130 Oy, holding the amino group in the proper conformation for catalysis. The TEM1 β-lactamase structure coordinates were kindly provided by N. Stynadka. Figures were generated in O2 (a) and GRASP(b, c).

Received 23 October 1995; accepted 4 March 1996.


ACKNOWLEDGEMENTS. We thank J. Newman, L. Shapiro and C. Ogata for help with data collection, analysis and subsequent structure solution. This work was supported in part by a grant from the NIH, Roscoe B. Jackson 444 at the National Synchrotron Light Source, a DOE facility, is supported by the Howard Hughes Medical Institute. Coordinates have been deposited in the Protein Data Bank, Brookhaven (accession number 1umu).

CORRESPONDENCE and requests for materials should be addressed to W.A.H. (e-mail: hendwl@ouhsc. hihm.columbia.edu)

Context-dependent secondary structure formation of a designed protein sequence

Daniel L. Minor Jr* & Peter S. Kim†

Howard Hughes Medical Institute, Whitehead Institute for Biomedical Research, Departments of *Chemistry and †Biology, Massachusetts Institute of Technology, Nine Cambridge Center, Cambridge, Massachusetts 02142, USA

PROTEIN secondary structures have been viewed as fundamental building blocks for protein folding, structure and design. Previous studies indicate that the propensities of individual amino acids to form particular secondary structures are the result of a combination of local conformational preferences and non-local factors. To examine the extent to which non-local factors influence the formation of secondary structural elements, we have designed an 11-amino-acid sequence (dubbed the ‘chameleon’ sequence) that folds as an α-helix when in one position but as a
The ‘chameleon’ sequence (Chm) was designed to replace α-helix residues 23–33 and β-sheet residues 42–52 (Fig. la) of GB1. Our aim was to preserve the hydrophobic nature of the residues that constitute the interface between each of these secondary structure elements and the core of GB1 (Fig. 1b). As the characteristic hydrophobic/hydrophilic patterning of buried residues in α-helices and β-sheets is very different, with periodicities of 3.6 and 2 residues respectively, creation of a sequence consistent with both patterns posed a number of design problems.

Fig. 1a, GB1 secondary structure. The primary amino-acid sequences of GB1, Chm-α and Chm-β are positioned below the corresponding secondary structures. b, Positions 23–33 and 42–52 of GB1 and Chm sequences. Changes from wild-type sequence are indicated in bold. Residues involved in the interface between each secondary structure and the remainder of the protein are shaded. c, Ribbon diagram showing the Chm sequence in Chm-α (left) and Chm-β (right); the Chm sequence is shown in yellow. The diagram was drawn using the coordinates of wild-type GB1 (ref. 21).

METHODS. GB1 mutants were derived from a synthetic GB1 gene by site-directed mutagenesis, verified by dideoxynucleotide sequencing, expressed in E. coli and purified by IgG-affinity chromatography and reverse-phase HPLC (ref. 22). In the GB1 homologue GB2 (ref. 23), position 57 is incorporated as part of the fourth β-sheet; GB1, with residue K57 added to the 56-residue construct GB1-Thr1 (ref. 22), had a Tm ~ 2 °C higher than GB1-Thr1 and was used as the parent construct for all Chm proteins. Molecular identities were verified by matrix-assisted laser desorption/ionization (MALDI) mass spectrometry (Finnigan MAT Lasermat or PerSeptive Biosystems Voyager) and found to be within 2 daltons of the expected mass for the 57-residue protein. The Chameleon sequence positions 23, 26, 30 and 33 were judged to be important for the helix–protein interface, whereas positions 43, 45, 51 and 52 were judged to be important for the β-sheet–protein interface (see text). For all class I sites, the buried residue was used, with the exception of the 29/48 pair. The buried residue of the 29/48 pair was not used as T49F was a necessary change in the turn between strands III and IV of Chm-β to preserve the buried Phe from the 30/49 class I pair, and we did not wish to cause two sequential amino-acid changes in this turn. We tested a number of substitutions for pair 26/45 in GB1*: The Tm (ΔH°) values from thermal unfolding were: A26Y, < 0 °C (not determined); A26I, 59.6 °C (42.5 kJ mol⁻¹); A26V, 66.5 °C (50.9 kJ mol⁻¹); Y45A, 53.4 °C (40.4 kJ mol⁻¹); Y45V, 61.1 °C (47.5 kJ mol⁻¹); Y45I, 61.8 °C (44.5 kJ mol⁻¹); thus Val was chosen for the 26–45 pair. Constructs containing a 10-residue Chm sequence at positions 23–32 and 42–51 of GB1* were seen to fold, so the 33/52 pair was examined. Substitution of Y33F lowered the Tm of the α-helix Chm by 0.6 °C, whereas F52Y lowered the Tm of the β-sheet Chm by 8.7 °C. Phe was chosen for the 33/52 pair in Chm-α and Chm-β.
In comparing the structural environments of positions 23–33 and 42–52, we encountered three main types of environments based on accessible surface area: class I, sites where a residue was buried in one secondary structure but exposed in the other (pairs 23/42; 24/43; 29/48; 30/49 and 32/51); the buried residue is underlined; class II, sites where a residue occupied a buried position in both structures but was very different in size or polarity in each structure (pairs 26/45 and 33/52); and class III, sites that had no conflicts in terms of size, polarity or burial (pairs 25/44; 27/46; 28/47 and 31/50). For class I sites, the buried residue from each pair was used for the Chm sequence, with the exception of pair 29/48 (Fig. 1 legend). For class II sites, a series of hydrophobic
residues were substituted at the positions of the residue pairs in the wild-type GB1 background in order to find a common residue for both environments that did not disrupt protein stability too drastically (Fig. 1 legend). After determining the residues that would enable a single sequence to fulfill the tertiary requirements of both positions 23–33 and 42–52 in GB1 using this procedure, the sequence AWTVEKAFKTFL was introduced at positions 23–33 or 42–52 of GB1 by site-directed mutagenesis, thereby creating the proteins Chm-α and Chm-β, respectively (Fig. 1c).

Chameleon-α and Chm-β each display cooperative reversible thermal unfolding as measured by circular dichroism (CD) spectroscopy (Fig. 2a) and differential scanning calorimetry experiments (Fig. 2 legend). This type of thermal unfolding behaviour is a hallmark of compact single-domain globular proteins with uniquely packed hydrophobic cores. The NMR spectra of Chm-α and Chm-β have substantial chemical-shift dispersion (Fig. 2b). In addition, NMR-detected hydrogen-exchange experiments indicate that both proteins contain a set of backbone amides that have exchange rates equal to or slower than those expected for exchange occurring only from global unfolding (Fig. 2c, d). Together with the thermal unfolding data, these results strongly suggest that both Chm proteins have unique folded structures with well packed hydrophobic cores (see discussion in ref. 11).

Nuclear Overhauser (NOE) spectra indicate that, as designed, the Chm sequence is folded into an α-helix in Chm-α (Fig. 3a) and a β-strand/turn/β-strand in Chm-β (Fig. 3b). In both proteins, amide protons from residues in the Chm sequence are significantly protected from hydrogen exchange, indicating that the hydrogen bonds in each secondary structure formed by the Chm sequence are stable.

NOE patterns throughout each protein are consistent with those in wild-type GB1 (Fig. 4). Furthermore, Chm-α and Chm-β are capable of competing for Fc binding with wild-type GB1, although with somewhat reduced affinities (Fig. 4 legend). The amino-acid changes in Chm-α and Chm-β occur in regions of the protein that have been identified as part of the Fc binding interface, so it is likely that much of the affinity differences are the result of mutation of interface residues. Considered together with the NMR data, the fact that Fc binding can occur strongly suggests that both Chm proteins are folded into structures similar to wild-type GB1.

As some short isolated peptides can form significant amounts of secondary structure, an 11-residue peptide corresponding to the Chm sequence (Ac-AWTVEKAFKTFL-NH2, where Ac is acetyl) was synthesized to investigate whether this sequence could have any intrinsic conformational preferences. Both CD and NMR indicate that the Chm peptide is unfolded in isolation (Fig. 4 legend), suggesting that the Chm primary sequence itself has no strong preference for either the α-helix or β-sheet conformation. Thus, the secondary structure formed by the Chm sequence in both Chm-α and Chm-β proteins is specified by tertiary interactions.

The secondary structure of peptide sequences in some natural proteins, such as haemagglutinin and the serpins, undergo major conformational alterations following tertiary rearrangements induced by pH changes or proteolytic cleavage. Tertiary interactions play a dominant role in determining the propensity of individual amino acids to form β-sheets; they also affect the secondary structure conformation of identical short peptide sequences (up to 6 amino acids long) in proteins in the structural database. Our Chm sequence design demonstrates directly

---

that the information specifying α-helix or β-sheet secondary structures can be entirely non-local. Taken together, these results underscore the importance of context-dependent effects in protein folding.

Received 14 December 1995; accepted 23 February 1996.


ACKNOWLEDGEMENTS. We thank L. M. Mayr, C. J. McNight, Y. P. Peng, P. A. Petito, B. A. Schulman and T. N. M. Schumacher for advice and discussion, and A. G. Cochran, B. A. Schultz, T. N. M. Schumacher, R. Dukovski and R. T. Sauer for comments on the manuscript.

CORRESPONDENCE AND MATERIALS. Requests to be addressed to P.S.K.

RETRACTION

Long-term corretation of rat model of Parkinson's disease by gene therapy
Shoushuo Jiao, Vladimir Gurevich & Jon A. Wolff

J.A.W. writes — In my laboratory we have been attempting to extend the findings reported in this paper. During these efforts, it has come to my attention that the pertinent laboratory notebooks were replaced with edited text and data. An independent analysis of the remaining original data revealed that the published Fig. 2b and c contains errors that exaggerate both the reductions in the number of rotations after transplantation and the increments in the numbers of rotations following graft removal. Review of the protocol reported in the legend to Fig. 2 indicates that control transfections were done using TransfectACE (Promega) instead of Lipofectin (BRL), which may have affected the outcome of the experiments. Subsequent experiments have failed to replicate the original observations. Regrettably, therefore, I am unable to verify that the conclusions of this paper are correct.