Articles

Synthesis and Biophysical Characterization of Engineered Topographic Immunogenic Determinants with $\alpha\alpha$ Topology†

Pravin T. P. Kaumaya,*,‡ Kurt D. Berndt,§ Douglas B. Heidorn,∥ Jill Trewhella,∥ Ferenc J. Kezdy,⊥ and Erwin Goldberg‡

Department of Biochemistry and Molecular Biology and Cell Biology, Northwestern University, Evanston, Illinois 60208, Department of Obstetrics and Gynecology, College of Medicine, The Ohio State University, Columbus, Ohio 43210, Department of Biochemistry and Molecular Biology, The University of Chicago, Chicago, Illinois 60637, Life Sciences Division, Los Alamos National Laboratories, Los Alamos, New Mexico 87545, and The Upjohn Company, Kalamazoo, Michigan 49001

Received February 27, 1989; Revised Manuscript Received August 9, 1989

ABSTRACT: Model peptides with predetermined secondary, tertiary, and quaternary conformation have been successfully designed, synthesized, and characterized in an attempt to mimic the three-dimensional structure of an antigenic determinant. This work is a continuing effort to map the antigenic structure of the protein antigen lactate dehydrogenase C_4 (LDH-C_4) to develop a contraceptive vaccine. A putative topographic determinant with $\alpha\alpha$ topology which associates into four-helix bundles was designed on the basis of the framework model of protein folding. An idealized amphiphilic 18-residue sequence (α_1) and a 40-residue $\alpha\alpha$ fold (α_3) have been shown to form stable 4-helix structures in solution with a free energy of association on the order of −20.8 kcal/mol (tetramerization of α_1) and −7.8 kcal/mol (dimerization of α_3). Both α_1 and α_3 form stable monolayers at the air–water interface. The CD spectra of Langmuir–Blodgett monolayers are characteristically α-helical. Both CD and FTIR spectroscopic studies reveal a high degree of secondary structure. The SAXS data strongly suggest that the helices are arranged in a four-helix bundle since the radius of gyration of 17.2 Å and the vector distribution function are indicative of a prolate ellipsoid of axial dimensions and molecular weight appropriate for the four-helix bundle. The major contribution to the formation and stabilization of α_1 and α_3 is believed to be hydrophobic interaction between the amphiphilic α-helices. The displayed heptad repeat, helix dipole, ion pairs, and the loop sequence may have also contributed to the overall stability and antiparallel packing of the helices. A detailed structural analysis of a relevant topographic immunogenic determinant will elucidate the nature of antigen–antibody interactions as well as provide insight into protein folding intermediates.

In another paper, a strategy to produce an antigenic peptide which assumes the conformation of the corresponding segment in the native protein was advanced, and the detailed immune responses to these conformational peptides at both the B cell and T cell levels are described (Kaumaya et al., 1989). Antibodies elicited by immunization with short peptides encompassing the flexible, accessible, and evolutionary variable segments of lactate dehydrogenase C_4 (LDH-C_4) bind poorly with greatly reduced affinity to the native protein (Hogrefe et al., 1989). This is presumably due to the fact that short linear peptides are unlikely to assume a conformation which accurately mimics the corresponding site in the native protein. The use of synthetic peptides to raise antipeptide antibodies of predetermined specificity (Lerner, 1982, 1984) for studying the mechanism of antigen–antibody interaction as well as T cell recognition of globular protein argues that rules which allow them to do so be defined structurally. As in the case of globular proteins, the conformational aspects of peptide structure must be deemed critical for their biological and immunological reactivity. The structural stability of small peptides and the physical interactions accounting for their stability are of current interest in order to understand the biological actions of peptides as well as the mechanisms of protein folding and protein structure in solution (Baldwin, 1986; Jaenicke, 1987). Recent work by Shoemaker et al. (1987) and Dyson et al. (1985, 1988a,b) has demonstrated that some short peptides are not without their own conformations and may possess recognizable secondary structures in aqueous solution. However, this demonstration cannot be generalized for all peptides.

This work extends our efforts to probe the antigenic structure of LDH-C_4 by engineering a topographic determinant to develop a contraceptive vaccine with useful antigenic and immunogenic characteristics. The design as illustrated in Figure 1 utilizes protein architecture as derived by X-ray analysis, protein stability as determined by factors including hydrophobic interactions (Kauzmann, 1959), electrostatic ion pairs (Sundaralingham et al., 1987), helix dipole model (Shoemaker et al., 1987), amphiphilicity (Kaiser & Kezdy, 1984), and salt bridges (Marqusee & Baldwin, 1987), and a set of empirical structure predictive algorithms (Yada et al., 1988). It seeks to use the framework model of protein folding (Ptitsyn & Rashin, 1975; Kim & Baldwin, 1982) and a set of rules that can be used to design a peptide able to fold in a predictable manner, thus mimicking the secondary structure (α-helix, β-sheet, turns, and loops) of a protein antigen in an incremental and hierarchical manner. In brief, the basic idea of our experimental approach is to approximate a nativelike structure by the hierarchic condensation (Fetrow et al., 1988) of these primitive "nascent" structures which associate in a

† This work was supported in part by National Institutes of Health Research Grants AI25790 (to P.T.P.K.) and HD23771 (to E.G.)

* To whom correspondence should be addressed at The Ohio State University.
‡ Northwestern University.
§ The University of Chicago.
∥ Los Alamos National Laboratories.
⊥ The Upjohn Co.

0006-2960/90/0429-13$02.50/0

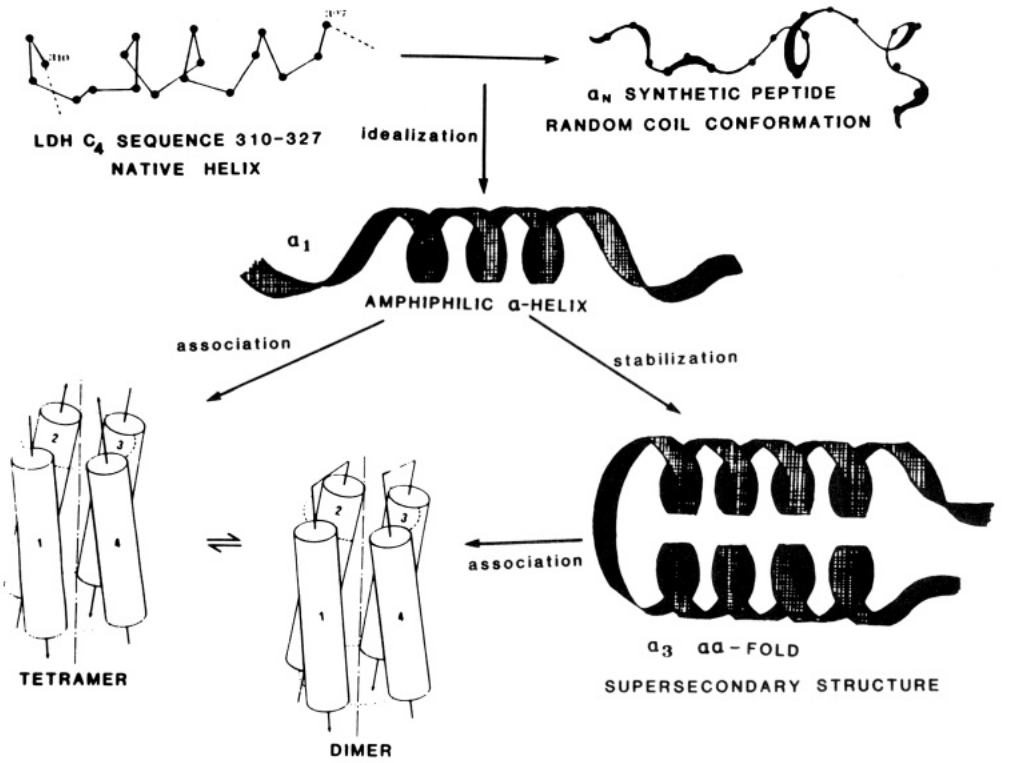

FIGURE 1: Schematic illustration of the design of a topographic determinant using the basic principles of the framework model of protein folding. The synthetic peptide mimic α_N is depicted in random coil conformation, wheres sequence 310–327 is shown as an ortep representation (three-dimensional structure from atomic ordinates of LDH-C_4) showing its native α-helical configuration. The idealization to an amphiphilic sequence is the first stage in the design followed by the choice of a loop sequence and the antiparallel α-helix to yield an $\alpha\alpha$ supersecondary structural motif (α_3), conformationally more stable than its single α-helical peptide (α_1,). Further stabilization is postulated to occur via association, in the case of the single helix by tetramerization and in the case of $\alpha\alpha$ fold by dimerization, to yield structures comparable to four α-helical bundles.

stepwise fashion into folding units or supersecondary structures ($\alpha\alpha$, $\beta\alpha\beta$, $\beta\beta\beta$, $\alpha\beta$, or $\beta\alpha\beta\alpha\beta$) (Levitt & Chothia, 1976). This approach will, we hope, allow us to demonstrate whether or not the important determinants recognized by the immune system for induction of antibodies are assembled from amino acid residues from distant part of the protein chain brought together in close proximity by folding of the polypeptide chain (Benjamini et al., 1984).

De novo peptide and protein design to yield molecules of predetermined specificity and well-defined three-dimensional structures has met with some success. The tools for the design have relied on empirical protein secondary structure prediction algorithms and molecular modeling. For instance, oligopeptides having the potential to form amphiphilic α-helices and β-sheets as well as β-turns have been synthesized (Kaiser & Kezdy, 1983; Taylor & Kaiser, 1986; DeGrado & Lear, 1985). Supersecondary structural motifs such as α-helical coiled-coils (Lau et al., 1984), $\beta\alpha\beta$ folding unit (Mutter et al., 1986), four-helix bundles (Ho & DeGrado, 1987), antiparallel β-barrel (Daniels et al., 1988), and template-assisted synthetic peptides (Mutter et al., 1988a,b) have been designed and synthesized to investigate the properties by which these secondary structures pack together. These model studies have added enormously to our understanding of the problem of stability and folding of peptides and proteins. Of particular and relevant interest to our own strategy is the experimental approach adopted by DeGrado and co-workers (Ho & DeGrado, 1987; DeGrado et al., 1989). In these studies, artificial synthetic peptide models with minimal sequence homology to the parent were designed in order to probe the structural characteristics necessary for the formation of a four-helical bundle motif. They have successfully designed a highly stable amphiphilic 16-residue α-helical peptide that self-associates into tetrameric aggregates. A loop sequence was designed to connect two identical α-helical sequences into an antiparallel orientation. Experimental evidence was provided by size-exclusion chromatography and by analyzing the concentration dependency of their CD spectra, indicating that the 35-residue peptide formed dimers in solution with a structure comparable to four-helix bundles (Weber & Salemme, 1980). A different experimental approach to construct this bundle motif was advanced by Mutter et al. (1988a). Rather than linking amphiphilic helices via loops, they chose to assemble the four helices by attachment to a template. A parallel arrangement of the template–helices was deduced from spectroscopic data. Hodges and co-workers have designed a series of two-stranded coiled-coils (Lau et al., 1984; Hodges et al., 1989), a common motif found in another class of α-helical proteins as exemplified by fibrous proteins such as tropomyosin. There are many similarities between the arrangements of α-helices in four-helix bundles and coiled-coils. The role of amphiphilic helices in the stabilization of these structures in which hydrophobic interactions provide the major driving force for folding is clear. Hodges and more recently O'Shea et al. (1989) have shown that these disulfide-bonded coiled-coils associate to form very stable dimers of α-helices which are in register and in parallel.

A major goal of our studies is to determine whether an immunogenic peptide can be engineered to adopt a stable conformation in aqueous solution. Thus, one aspect of this investigation involves studying the forces which stabilize the secondary and supersecondary structural motifs in an attempt to mimic the three-dimensional structure of a protein antigenic site. To that end, a variety of biophysical techniques were used to determine the conformation of the model peptides as single techniques are usually biased to interpretation. In this study, peptide monolayers, circular dichroism (CD), and Fourier-

transform infrared (FTIR) were used to characterize and correlate the amphiphilic secondary and supersecondary structures with $\alpha\alpha$ topology. In addition, we extend the original characterization of a four-helical bundle topology (Ho & DeGrado, 1987) by small-angle X-ray scattering (SAXS) measurements.

Experimental Procedures

Unless otherwise stated, all chemicals and solvents were of the highest grade available. Dimethylformamide (DMF) was stored over molecular sieve (4A) for 2 weeks with occasional N_2 flushing and dinitrofluorobenzene (DNFB) tested before use. Most *N*-(9-fluorenylmethoxycarbonyl)-amino acids FMoc-amino acids were purchased from Cambridge Research Biochemicals Ltd. (Harston, England) and were all of the L configuration. All amino acid derivatives used for peptides were checked for purity by thin-layer chromatography (TLC) on precoated silica gel [F254, 0.25-mm-thick plates (E. Merck)]. The following solvents were used for TLC: chloroform/methanol/acetic acid (85:10:5 v/v, system A); benzene/dioxane/acetic acid (95:24:4 v/v, system B). Spots were visualized under ultraviolet light (254 nm). Purity was further checked by high-pressure liquid chromatography (HPLC) with a μBondapak C_{18} column, 30 cm × 3.9 mm (Waters Associates, Milford, MA), using a solvent system consisting of water [0.1% trifluoroacetic (TFA)] and acetonitrile (0.1% TFA) and a gradient of 10% A → 100% B, 20 min, 1.5 mL/min effluent being monitored at 260 nm.

HPLC was carried out by using a Waters Associates ternary gradient system consisting of 2M 6000 A pumps, one M46 pump, a Model 720 system controller, a Model 730 data module, and an automatic wisp injector. Ultraviolet detection and temperature control were provided by a Perkin-Elmer LC-65-T spectrophotometer.

Analytical HPLC was carried out on Waters μBondapak C_{18} (3.9 mm × 30 cm) and Vydac C_4 (4.6 mm × 25 cm, 5 μM, 300 Å) columns. Crude peptides were chromatographed in the presence of 0.05% TFA, and elution was accomplished by using a 20–30-min gradient of 10–50% aqueous acetonitrile containing 0.05% TFA at a flow rate of 1 mL/min.

Semipreparative HPLC was carried out on μBondapak C_{18} (19 mm × 15 cm) and Vydac C_4 (10 mm × 25 cm) columns. Elution was accomplished by using a 10–15-min gradient of 10–50% aqueous acetonitrile containing 0.05% TFA at a flow rate of 5 mL/min. Typical loadings per run were between 10 and 20 mg of peptide.

Hydrolysis of peptides was carried out by using 6 N HCl containing 1% phenol in evacuated sealed tubes at 110 °C for 24 h. The hydrolyzed peptides were derivatized by using phenyl isothiocyanate (PITC).

Amino acid analysis of derivatized hydrolysate was performed on a Pico-Tag system using the HPLC system described above.

High-voltage paper electrophoresis was performed at pH 6.5 (pyridine/acetic acid/water 100:3:900 v/v) and pH 1.8 (formic acid/acetic acid/water 1:4:45 v/v). Electrophoresis was carried out at 60 V/cm for 45 min on Whatman 1.

Peptide sequencing was carried out with an Applied Biosystems Model 477A/120 instrument using the on-line 120A phenylthiohydantoin analyzer. A sample of the peptides was subjected to 18/40 cycles of automated quantitative sequence analysis to further assess purity.

Peptide Synthesis. All synthetic peptides were assembled semimanually by stepwise FMoc-*tert*-butyl solid-phase synthesis on a Vega 1000 coupler with 4-methylbenzhydrylamine resin as the solid support (0.54 mmol of Cl/g) (Barany & Merrifield, 1979; Stewart & Young, 1984; Atherton & Sheppard, 1983). Side chain protection was as follows: *tert*-butyl esters for aspartic acid and glutamic acid; *tert*-butyl ethers for serine and threonine; *tert*-butyloxycarbonyl (Boc) for lysine. The resin support was neutralized with 10% diisopropylethylamine (DIEA) in dichloromethane (DCM). The acid-labile linkage agent 4-(hydroxymethyl)phenoxyacetic acid, activated as its pentafluorophenyl (pfp) ester (5-fold), was attached to the resin by double coupling (30 min each). During the last 15 min, 3 equiv of catalyst [1-hydroxybenzotriazole (HOBt)] was added. Ninhydrin Kaiser test indicated that no amino groups were left unreacted. The C-terminal amino acid (FMoc-Glu-O-*t*-Bu) was esterified to the linker-activated support by double coupling of the preformed Fmoc pentafluorophenyl ester for 1 h each time in the presence of 4-dimethylaminopyridine (DMAP, 0.1 equiv) as catalyst. Unreacted hydroxymethyl groups on the resin were left uncapped.

A typical coupling cycle was as follows: the N-protected peptide resin was treated with 20% piperidine in DMF for 3 min followed by a second treatment for 7 min to remove the FMoc protecting group. Batchwise washing of the peptide resin was with DMF, DCM, and 2-propanol. Couplings were performed by using preformed pentafluorophenyl esters (5 equiv) of the next amino acid in DMF. The preformed FMoc-pfp esters were prepared manually at the bench 1 h prior to coupling by dissolving the FMoc-amino acid in ethyl acetate together with pentafluorophenol. The reaction mixture was cooled to 0 °C prior to addition of *N*,*N*-dicyclohexylcarbodiimide (DCC). After 30 min, the activated amino acid solution was allowed to come to room temperature before the dicyclohexylurea was filtered off. The filtrate was evaporated to dryness and taken up in DMF before addition of the peptide resin and allowed to react for 1 h. After further washing, the coupling cycle was repeated with a fresh batch of activated FMoc pentafluorophenyl ester. All amino acids were coupled in similar fashion. Usually, HOBt (1 equiv) was added to the reaction mix during the latter stages of coupling. After the chain was elongated, the N-terminal FMoc group was removed with the usual piperidine treatment and the resin washed, transferred, and dried in a desiccator. The side chain protecting groups and the peptide resin anchoring bond were cleaved by treatment with 60% trifluoroacetic acid/dichloromethane (TFA/DCM) containing 10% anisole (4.5-h treatment). The resin was filtered and washed well with neat TFA, and the combined filtrate was rotary-evaporated to an oily residue. The peptide was precipitated with ice-cold ether and washed with ether by decantation. The crude peptide was taken up in 1 M acetic acid and extracted with diethyl ether. The ether extract was back-extracted with water, and the combined aqueous acidic phase was lyophilized to a dry powder.

Peptide Monolayer Measurements and Analysis at the Air–Water Interface. The surface pressure (π) of peptide monolayers at the air–water interface was monitored by using a Lauda film balance, as a function of the area of the monolayer (A). Peptide solutions (0.7–2.0 mM) in 5% acetic acid were spread onto 1 mM potassium acetate/0.16 M KCl, pH 5.0 and 7.0, and compressed and expanded between 600 and 100 cm^2 at the rate of 2.2 cm^2/s.

Surface pressure/area isotherms were analyzed by nonlinear least-squares regression of the experimental data (<1 dyn/cm) according to the equation (Adamson, 1976) of state:

$$\pi = nRT/(A - A_0) \quad (1)$$

where π is the observed surface pressure for a given area per

amount of peptide spread, R is the gas constant (8.31×10^7 erg mol^{-1} K^{-1}), n is the number of moles of applied peptide, and A_0 is the limiting surface area of the monomer. Linear extrapolation of the experimental data (>8 dyn/cm) to zero pressure will also yield an independent measure of A_0.

Langmuir–Blodgett monolayers of α_1 and α_3 were collected from the air–water interface of the Lauda film balance at constant pressure. The surface of the quartz glass slides (22-mm diameter × 1.25-mm width, Hellma) was made hydrophobic by treatment with Silicad (Clay Adams, Parsipanny, NJ). Monolayers were transferred to the slides by vertical insertion of the slide through the monolayer and back out (5 mm/min), with faces parallel to the moving barrier. The extent of transfer was quantitative as monitored by the isobaric change in area. CD spectra of the Langmuir–Blodgett monolayers were performed on stacks of five to seven freshly collected slides (Cornell, 1979).

CD Studies for Concentration Dependency Measurements and Analysis. The circular dichroism (CD) spectra of the peptides were measured at room temperature under a constant flow of nitrogen on a Jasco J-600 spectrapolarimeter attached to a Jasco DP-500N data processor. The instrument was calibrated with an aqueous solution of recrystallized ammonium *d*-10-camphorsulfonate (Johnson, 1988). Solutions of the peptides (1–400 μM in 0.01 M Tris/0.16 M KCl, pH 7.4) were measured in 0.1–1.0-cm path length, cylindrical quartz cuvettes (Hellma), and concentrations were determined by amino acid analysis of aliquots of stock solutions.

The experimental data were analyzed by nonlinear least-squares regression according to a cooperative monomer–oligomer equilibrium as expressed in eq 2 where θ_{exp} is the ex-

$$\ln [P] = \ln (\theta_{olig} - \theta_{mon}) + [1/(n-1)] \ln (K_D/n) + [1/(n-1)] \ln (r_{exp} - \theta_{mon}) - [n/(n-1)] \ln (\theta_{olig} - \theta_{exp}) \quad (2)$$

perimental mean residue ellipticity and θ_{mon} and θ_{olig} are the theoretical mean residue ellipticities of the monomer and oligomer, respectively. K_D is the apparent dissociation constant for the monomer–oligomer equilibrium, and P is the total concentration of peptide. The value of n can also be obtained from the slope of a ln [oligomer] vs ln [monomer] plot of the experimental data converted by eq 4a,b. The values for θ_{olig} and θ_{mon} were obtained from extrapolation of the data:

$$\ln [\text{oligomer}] = n \ln [\text{monomer}] - \ln K_D \quad (3)$$

$$[\text{oligomer}] = P(\theta_{exp} - \theta_{mon})/(\theta_{olig} - \theta_{mon})(1/n) \quad (4a)$$

$$[\text{monomer}] = P(\theta_{olig} - \theta_{exp})/(\theta_{olig} - \theta_{mon}) \quad (4b)$$

Fourier-Transform Infrared Spectroscopy (FTIR) Measurements and Analysis. α_1, α_N, and α_3 peptide fragments were lyophilized and exchanged in D_2O 3 times; the first exchange was for 48 h at 5 °C in order to assure complete exchange. The final D_2O-exchanged samples were brought up in 10 mM phosphate buffers at pH 4.8 and 7.0 (uncorrected meter readings). Protein concentrations of the stock solution were determined by weighing, and dilutions were done gravimetrically.

Infrared spectra were measured at 2 cm^{-1} resolution and signal-averaged over 1024 scans. The samples were pipetted into a Perkin-Elmer solution cell with CaF_2 windows and a 0.05-mm Teflon spacer. Spectra were collected on a Mattson Alpha Centauri Fourier-transfer spectrometer. To minimize interference in the amide I band (1620–1690 cm^{-1}) from the rotational fine structure bands resulting from trace amounts of water vapor, the optical bench was maintained under constant positive dry nitrogen pressure, and spectra were collected after a 20-min purge. Spectra of buffer blanks were collected for each buffer condition and subtracted from the sample spectra to give absorption spectra free of any contributions from the phosphate buffer.

The analysis of the FTIR data is described in detail elsewhere (Trewhella et al., 1989). The evaluation of the secondary structure content was done by fitting deconvolved amide I′ components and using the empirical assignments of Byler and Susi (1986). The methods described by Kauppinen et al. (1981) were used to deconvolve the spectra. Gaussian line shapes were used to fit the deconvolved components, and the intensity of each Gaussian was used to compute the relative amount of secondary structure component. The intensity of each band is directly proportional to the number of carbonyl groups contributing at that frequency.

Second-derivative spectra were also computed in order to assign frequencies in the amide I′ band. This procedure is independent of the deconvolution procedure, but the frequency assignments obtained from this method give excellent agreement with the frequencies obtained by fitting the deconvolved spectra.

Small-Angle X-ray Scattering (SAXS) Measurements and Analysis. α_1 and α_3 samples for SAXS experiments were prepared in a 10 mM phosphate buffer at pH 5.2 (α_1) and pH 7.4 (α_3). Concentrations of stock solutions were determined by amino acid analysis. Protein dilutions were done gravimetrically and ranged from 10 to 35 mg/mL for the α_1 sample and from 10 to 25 mg/mL for the α_3 sample.

Data were collected on a small-angle instrument at Los Alamos which is described in detail elsewhere (Heidorn & Trewhella, 1988). The net scattering from the protein molecules alone was calculated by subtracting a normalized buffer spectrum measured in the same sample cell. Typical data collection times were 6–16 h, depending on peptide concentration.

The X-ray data analysis procedures have been described in detail in a previous paper (Heidorn & Trewhella, 1988). Data were analyzed with the Guinier approximation (Guinier, 1939) as well as with an indirect Fourier-transform analysis of Moore (1980). The extrapolated scattering intensity at zero angle, I_0, is used to determine the molecular weight (mw) of the protein sample by

$$mw_s = I^s_0 mw_1 / I^1_0 \quad (5)$$

where mw_s, mw_1 and I^s_0, I^1_0 are the molecular weights and I_0 values of the sample and a standard protein, respectively. Lysozyme was used as a molecular weight standard because of its size, shape, and monodispersity even at high concentrations (Kringbaum & Kugler, 1970). Equation 5 assumes the standard protein and the protein of interest have the same electron density and hence partial specific volume, v_s. The partial specific volumes for most proteins lie in the range 0.69–0.75. The value for lysozyme has been determined experimentally as 0.72 (Kringbaum & Kugler, 1970). For the α_1 and α_3 peptides, v_s has been calculated. For each peptide, the computed value of v_s was 0.77. In the absence of experimental values of v_s for the peptides, we have refrained from adjusting the molecular weight determinations to account for the possible effects of a larger value for v_s. We note, however, that this is a source of uncertainty in the molecular weight determinations at the 5–10% level. The experimental I_0 may also be a few percent larger than that computed from the molecular weight due to the effects of counterions binding to the peptide.

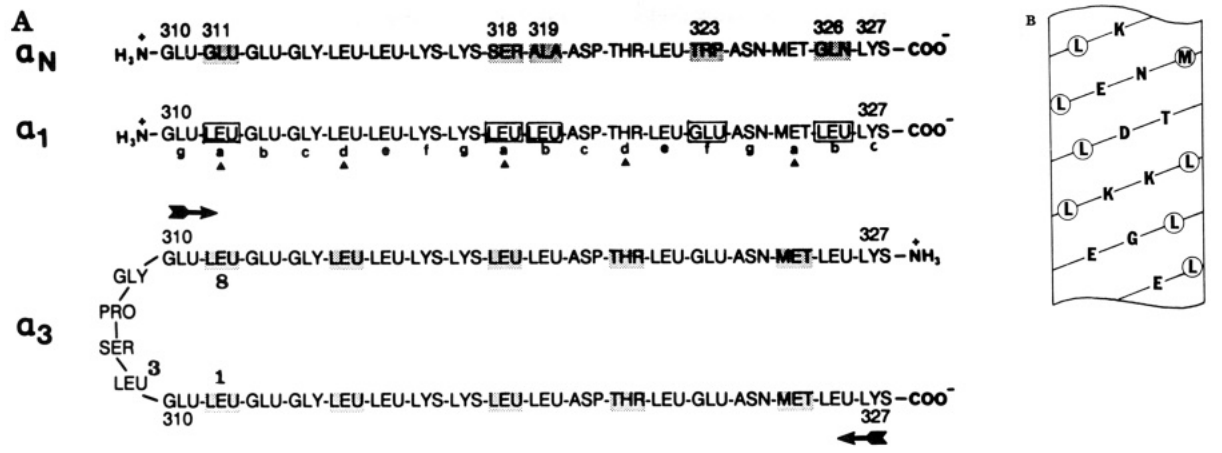

FIGURE 2: (A) Amino acid sequences of LDH-C_4 sequence 310–327 designated α_N and engineered synthetic peptides α_1 and α_3. The shaded amino acids in α_N were determined not to be involved in antibody recognition by using the algorithm of Novotny et al. (1986) using a 10-Å probe. These residues were mutated to yield an ideal amphiphilic sequence in α_1, residue substitutions are indicated as an open boxes. Residues in positions a and d are hydrophobic and display a heptad repeat. In α_3, the loop sequence was chosen such that hydrophobic residues cluster in positions 1, 3, and 8 as they point in the hydrophobic core. The hydrophobic residues comprising the heptad repeat are shaded. (B) Helical net diagram to depict the amphiphilic nature of the α_1 sequence.

The length distribution functions, $P(r)$, were calculated directly by Fourier transforming the scattering profiles. The maximum linear dimension, d_{max}, of the particle can be determined from the $P(r)$ function, and the R_g can be computed from the second moment of $P(r)$. In addition, the overall shape of the scattering particle can be inferred from the $P(r)$ curve.

Results

Design of an Amphiphilic α-Helix and an αα Fold. The rational and theoretical considerations underlying the strategy to mimic a topographic immunogenic determinant are discussed elsewhere (Kaumaya et al., 1989). To summarize briefly, sequence 310–327 of mouse LDH-C_4, designated α_N, was chosen (Figure 2A) to maximize its potential as a potent immunogen. In the crystal structure, α_N is part of an extended α-helix; however, the excised fragment is unlikely to adopt a similar conformation in solution according to the Zimm–Bragg theory (Zimm & Bragg, 1959). The delineation of the ends of the helices was therefore chosen to maximize the dipole across the helix with Glu^{310} at the N terminus and Lys^{327} at the C terminus, a property which has been shown to increase α-helical conformation (Shoemaker et al., 1987). The solvent-exposed hydrophilic residues incorporated stabilizing electrostatic ion pairs in positions i, $i + 3$ and i, $i + 4$. The α_N sequence was next idealized into an amphiphilic α-helix designated α_1 with rational sequence changes (Figure 2A,B) in such a way that the hydrophilic, solvent-exposed, and antibody contact residues accessible on the surface, as measured by a 1.4-Å and a 10-Å probe (Lee & Richards, 1971; Novotny et al., 1986; Hogrefe et al., 1987), were not disturbed. The assumption was made that this idealization of the sequence would not affect antibody specificity for the native protein since the spatial orientation of the hydrophilic antibody-accessible residues for epitope recognition would be preserved by formation of a stable helix structure.

This structure should be stabilized by association of the helices to yield a tetramer (Ho & DeGrado, 1987) of comparable structure to a four-helix bundle (Weber & Salemme, 1980). The helices should pack with their axes nearly antiparallel inclined by 20° if the sequence displays a seven-residue "heptad" repeat (Figure 2B). The design of a putative β-turn to connect the two helices was examined by CPK models and computer graphics (Evans & Sutherland, PS330) using the program FRODO. The 4-residue β-turn consisted of Leu-Ser-Pro-gly based on examination of 33 patterns in native proteins that code for αα corners (Efimov, 1984). The resulting αα corner, therefore, consists of a total of eight residues starting with the last hydrophobic residue of one helix (Leu^1) and ending with the first hydrophobic residue of the antiparallel helix (Leu^8). The short connection joining the two helices must have an extended conformation and be oriented approximately perpendicular to the axes of both helices. An artificial Leu was placed at residue 3 of the αα corner because its side chain in that position is completely buried in the hydrophobic core and must be hydrophobic (Figure 2). Ser^4 was chosen as the second residue as it can hydrogen bond with the free NH groups of the backbone. Additionally, if a β-turn was formed, the side chain of Ser^4 would interact favorably with the solvent. Pro and Gly were chosen to complete the turn because of the high propensity of their occurrence in that position in natural proteins. Another important feature is that the last hydrophobic residue of one helix (Leu^1) and the first hydrophobic (Leu^8) residue of the second helix necessary cluster form a pair in positions 1–8.

The choice of the antiparallel helix in α_3 was identical in sequence (α_1) with the exception that the helix was synthesized in an opposite orientation (N → C). This strategy should favor an antiparallel arrangement of the helices and a parallel allignment of the dipole. It would also allow the sequence to display a "7 residue heptad repeat", a sequence motif commonly found in coiled-coil systems (Lau et al., 1984). This arrangement allows tight packing of side chains protruding from the apolar face of the helices and also giving rise to an apolar surface stripe inclined around the axis of each α-helix at about 18° to each other (Cohen & Parry, 1986). In this arrangement, the stability of the αα fold would be dependent on intramolecular folding of the helices prior to intermolecular association to yield a quaternary structure. Finally, the amphiphilic nature of the two α-helices should act as a driving force for adoption of a folded conformation resulting in an αα fold. The structure should further be stabilized by dimeric association into an α-helical bundle (Ho & DeGrado, 1987).

Syntheses, Purification, and Characterization. Synthetic peptides (20–40 residues) assembled by solid-phase methods can result in very complicated mixtures in which the target peptide sequences would be very difficult to purify even by modern HPLC techniques. During the syntheses of over 30 different sequences of varying complexities and chemical

composition by a combination of different chemical procedures, we have found the following strategy to consistently yield crude peptides in excess of 80% purity. Thus, our method of choice at present consists of FMoc for temporary α-amino protection [deprotected by 20% piperidine in DMF (10 min)], *tert*-butyl ether. Double-coupling protocols are routinely carried out by using preformed FMoc-amino acid pentafluorophenyl esters. All couplings were performed exclusively in dimethylformamide to disrupt any sequence-dependent aggregation tendencies of the growing peptide chain.

The yield of protected peptide resin from the synthesis of α_1, α_3, and α_N was in the range of 75–85%. The yield of crude peptide from each cleavage ranged between 85 and 90% after multiple half-hour cleavage. One of the criteria used to determine the purity of synthetic crude peptide was analytical HPLC. The major peaks from each spectrum analyzed as approximately 70–80% of the desired product. Crude synthetic peptides were analyzed and purified by reverse-phase HPLC on either Waters C_{18} μBondapak or Vydac C_4 columns utilizing an acetonitrile/water gradient containing 0.05% trifluoroacetic acid. Fractions represenging the top portion of the principal peak were selected in order to obtain highly purified material. Gradients were optimized on analytical runs to resolve additional, close running impurities and transposed appropriately onto a semipreparative scale. The purified peptide when rerun on reverse-phase HPLC was greater than 95% single peak as detected by the UV absorbance at 214 nm.

Amino acid analysis of α_1, α_3, and α_N showed a close correlation with the theoretical composition, as indicated [observed (theory)]: α_1: Asx, (2); Glu, 3.04 (3); Gly, 1.17 (1); Thr, 0.90 (1); Leu, 6.9 (7); Met, 0.79 (1); Lys, 3.08 (3). α_3: Asx, 3.85 (4); Glu, 5.9 (6); Gly, 3.43 (3); Thr, 1.65 (2); Leu, 15.8 (15); Met, 1.32 (2); Lys, 5.8 (6); Pro, 1.2 (1); Ser, 0.88 (1). α_N: Asx, 2.05 (2); Glx, 3.95 (4); Gly, 1.1 (1); Lys, 3.1 (3); Leu, 3.2 (3); Ser, 0.86 (1); Thr, 0.79 (1); Trp, – (1); Met, 0.82 (1); Ala, 1.0 (1).

Confirmation that the correct sequence had been assembled by chemical synthesis was obtained by automated sequence analysius utilizing stepwise Edman degradation of 18 and 40 cycles on an Applied Biosystem analyzer. Further characterization was by fast atom bombardment mass spectroscopy.

Monolayer Studies of the Peptides at the Air–Water Interface. If a peptide forms stable monolayers at the air–water interface, then in the monolayer the molecule must assume an amphiphilic conformation (Adamson, 1976). Analysis of force–area (π/A) curves obtained by compression of these monolayers yields structural information on the surface-active species.

Figure 3 shows the force–area (π/A) curve obtained by the compression of the α_1 and α_3 monolayers. In the case of α_1, at low pressures (<1 dyn/cm), the π/A curve obeyed the equation describing a two-dimensional gaseous phase. A nonlinear least-squares fit of the data from this region yielded a limiting area of 935 Å^2/molecule. At pressures above 1 dyn/cm and below 18 dyn/cm, the π/A curve became a straight line, indicating that α_1 monolayer underwent a linear compression without serious structural reorganization. The collapse pressure of α_1 was found to be >20 dyn/cm. Extrapolation of the linear portion of the π/A curve to $\pi = 0$ yields a limiting area of 1039 Å^2/molecule (16 Å^2/AA), in good agreement with that obtained from analysis of the gaseous phase.

For α_3, a nonlinear least-squares fit of the data from $\pi <$ 1 dyn/cm yielded a limiting area of 2238 Å^2/molecule which has an area of 17 Å^2/AA, characteristic of an α-helix. At

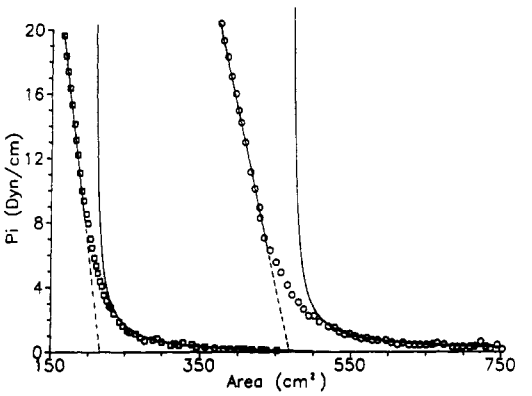

FIGURE 3: Force–area curves for α_1 (□) and α_2 (○). 12.5 nmol of each was spread onto 1 mM potassium acetate/0.16 M KCl, pH 5.0. Solid curve, generated from a nonlinear least-squares fit of the experimental data (<1.5 dyn/cm) to eq 1; solid line, linear extrapolation of experimental data (>8 dyn/cm) to zero pressure.

pressures above 1 dyn/cm and below 30 dyn/cm, the π/A curve became a straight line, indicating that α_3 did not undergo any serious reorientation but rather compression of the side chains. The collapse pressure of α_3 (data not shown) was about 35 dyn/cm. Extrapolation of the linear portion of the π/A curve to $\pi = 0$ yields a limiting area of 2250 Å^2/molecule (17 Å^2/AA), in good agreement to that obtained from analysis of the gaseous phase.

Conformational Properties of Peptides in Aqueous Solution As Measured by CD. Spectra typical of α-helices were obtained for α_1 and α_3 with minima at 222 (n–π^*) and 208 nm (π–π^*) and a maximum at 193 nm. α_N showed a characteristic random coil minimum at around 200 nm (π–π^*). The concentration dependencies of each peptide were investigated over a wide range of concentrations (1–400 μM peptide) in 0.01 M Tris-HCl/0.16 M KCl, pH 7.4, or in 0.01 M potassium acetate/0.16 M KCl, pH 5.0. The spectra of α_1 and α_3 were strongly dependent on peptide concentration. At low peptide concentrations (<10 μM) for α_1 and (<6 μM) for α_3, the CD spectra showed low -helix content. At increasing peptide concentrations, the mean residue ellipticity became more negative, indicating progressive helix formation.

Figure 4A shows the concentration dependence of the mean residue ellipticity of α_1 at 222 nm. Analysis of the data (inset of Figure 4A) gives the degree of aggregation as 3.9 ± 0.2, consistent with tetramer formation. These data were analyzed in terms of a cooperative monomer–oligomer equilibrium according to eq 2. α_1 was optimally fit to a monomer–tetramer equilibrium with a value of −30 310 deg·cm^2/dmol for the ellipticity in the tetrameric form and −3013 deg·cm^2/dmol for the monomer and a dissociation constant of 7.42×10^{-16} M^3. The dissociation constant corresponds to a free energy of tetramerization of −20.9 kcal/mol.

Analysis of the data for α_3 (inset of Figure 4B) gives a degree of aggregation of 1.9 ± 0.1, corresponding to dimer formation. The data for α_3 were optimally fit to a monomer–dimer equilibrium with a value of −22 485 deg·cm^2/dmol for the ellipticity of the dimer and −5059 dec·cm^2/dmol for the monomer. A dissociation constant of 1.18×10^{-18} M^3 was calculated, which corresponds to a free energy of dimerization of −7.8 kcal/mol.

Secondary Structure Analysis by FTIR. Panels A and B of Figure 5 show the original spectra, the deconvolved spectra, and the second-derivative spectra of α_1 and α_3, respectively, in the amide I region. The spectra for α_1 and α_3 are similar to one another, and the deconvolution and second-derivative procedures give similar frequency assignments. The spectra

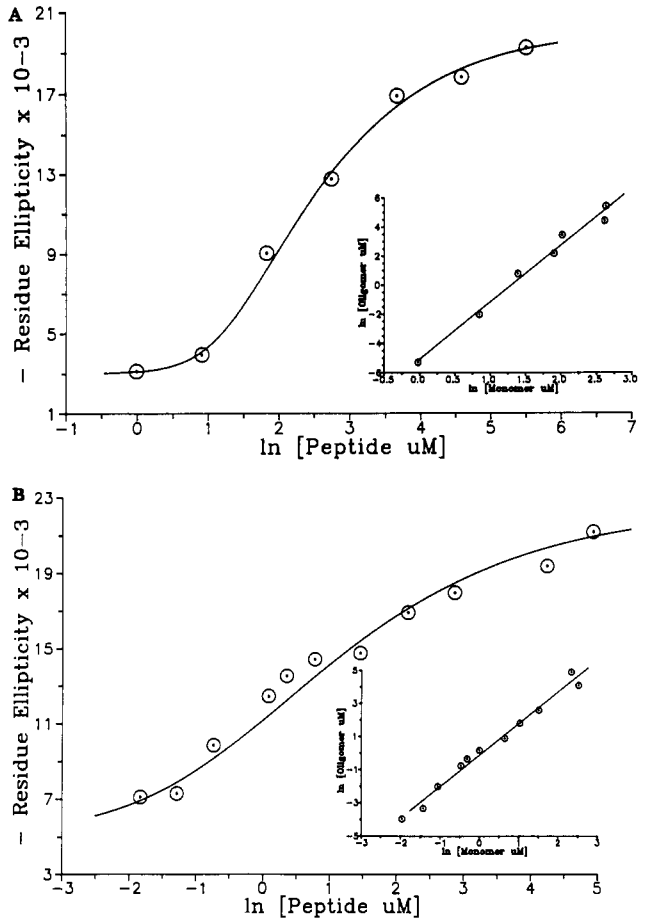

FIGURE 4: (A) Concentration dependence of the residue ellipticity (degrees centimeter squared per decimole) at 222 nm as a function of α_1 concentration. The theoretical curve was generated by analysis of the experimental data for a cooperative monomer/tetramer equilibrium as described by eq 2. $\theta_{monomer}$ = 3013 deg·cm^2 dmol^{-1}; $\theta_{tetramer}$ = 20 310 deg·cm^2 dmol^{-1}; $K_d = 7.42 \times 10^{-16}$ M^3. The inset shows the same experimental data converted to monomer and oligomer concentrations (eq 4a,b) analyzed by linear regression (eq 3). The slope of the regression line is 3.9 ± 0.2. (B) Concentration dependence of the residue ellipticity (degrees centimeter squared per decimole) at 222 nm as a function of α-3 concentration. The theoretical curve was generated by analysis of the experimental data for a cooperative monomer/dimer equilibrium as described by eq 2. $\theta_{monomer}$ = 5059 deg·cm^2 dmol^{-1}; θ_{dimer} = 22 485 deg·cm^2 dmol^{-1}; $K_d = 2.4 \times 10^{-16}$ M. The inset shows the same experimental data converted to monomer and oligomer concentrations (eq 4a,b) analyzed by linear regression (eq 3). The slope of the regression line is 1.9 ± 0.1.

for α_N in Figure 5C are quite different from either α_1 or α_3.

The existence of narrow bands within the amide I region of the deconvolved spectra indicates that the peptides have specific structural organization. The α_N spectrum shows a strong-component (59%) relative intensity at 1673 cm^{-1}, which is characteristic of extended structures, and much less intensity (19%) in the α-helix frequency assignment region. Other peaks are hard to distinguish. Deconvolution identifies peaks in the 1620–1660 cm^{-1} range which are distinct but not particularly well resolved. The deconvolved spectra for the α_1 and α_3 peptides are qualitatively similar to one another; both have a strong band at 1650 cm^{-1} which is assigned to the α-helix and a strong band at 1673 cm^{-1} assigned to extended chain structure. Both peptide spectra also exhibit a smaller peak at 1633 cm^{-1} which is assigned to the low-frequency component of the extended chain. The α_3 spectrum has a weak shoulder at 1660 cm^{-1} which is assigned to turns in the structure: this feature is not significant in the α_1 spectra. The FTIR data, which are summarized in Table I, confirm that both α_1 and α_3 peptides have a high α-helix content (50% for α_3 and 44% for α_1) as well as a significant amount of extended chain (45% for α_3 and 54% for α_1) and that the structural composition of each is independent of pD. In contrast, α_N exhibits low α-helix content (14%) and a high amount of extended chain (65%). The 1660 cm^{-1} turn assignment for α_3 is consistent with the fact that α_3 is expected to be a pair of adjacent α-helices arranged in antiparallel fashion connected by a single turn. This assignment is not seen in the α_1 spectrum, consistent with α_1 being a single helix.

Table I: Frequency Assignments and Relative Intensities for α_1, α_3, and α_N Amide I′ Bands[a]

band	ν (cm^{-1}) deconvolution	ν (cm^{-1}) second derivative	intensity (%)	secondary structure
α_3	1631.5 ± 0.8	1633	13.3 ± 2.4	extended chain
	1637.8 ± 0.8	1637	4.8 ± 1.9	extended chain
	1650.3 ± 0.2	1653	49.9 ± 2.5	α-helix
	1660.7 ± 1.1	1663	4.9 ± 1.9	turn
	1673.2 ± 0.2	1673	27.1 ± 1.7	extended chain
α_1	1634.5 ± 1.2	1635	16.4 ± 2.7	extended chain
	1642.5 ± 1.0	1643	3.3 ± 2.0	random coil
	1649.6 ± 0.6	1650	32.6 ± 3.0	α-helix
	1657.2 ± 1.2	1653	11.2 ± 2.1	α-helix
	1664.2 ± 1.7	1662	2.1 ± 2.0	turn
	1673.8 ± 0.2	1673	34.4 ± 2.3	extended chain
α_N	1620.7 ± 0.8		5.0 ± 1.6	extended chain
	1626.9 ± 0.8		3.3 ± 1.4	extended chain
	1632.3 ± 0.5		5.6 ± 1.4	extended chain
	1637.8 ± 0.7		4.0 ± 1.4	extended chain
	1643.0 ± 0.4	1642	14.9 ± 1.9	random coil
	1649.8 ± 0.3	1649	7.7 ± 1.3	α-helix
	1656.7 ± 0.3	1655	11.3 ± 1.6	α-helix
	1663.7 ± 0.5		7.5 ± 1.5	turn
	1673.7 ± 0.1	1673	40.7 ± 1.8	extended chain

[a] Data for pH 7.0 samples. Frequency assignments and intensities for pH 4.8 data not shown, but are identical with pD 7.0 data.

α_1 and α_N each have small components at 1654–1660 cm^{-1} (11%) which are also in the frequency region assigned to the α-helix. This peak is a real feature and is not due to noise in the spectrum or to the presence of water vapor nor is it an artifact caused by overdeconvolution. The spectra were signal-averaged to reduce noise, and the nitrogen purge to remove residual water vapor and CO_2 was thorough. Also, the scans were repeated at a different pH, and the frequency assignments of the two absorption spectra agree, which would tend to preclude random errors such as noise. The problem of overdeconvolution was addressed by repeatedly deconvolving the spectra while simultaneously varying σ and L. The deconvolved spectra were plotted to verify that no artifacts appeared in the spectra as the parameters were changed. σ was checked and was kept smaller than the FWHM of the narrowest peak in the amide I spectrum to minimize side-band noise. Amide I′ contributions due to helix structure can be shifted in frequency up to ~10 cm^{-1} by variations in hydrogen bond strengths. Shifts to high frequency have generally been attributed to a weakening of the hydrogen bonds between backbone carbonyl and amide groups. Such an effect would be expected to strengthen the carboxy bond, hence, give rise to a higher frequency (Krimm & Dwivedi, 1982). α_1, α_N, and α_3 exhibit a strong absorption at 1673 cm^{-1} for pH 4.8 and 7.0. On the basis of the assignments of Byler and Susi (1986), this is due either to an extended chain or to β-sheet structure. None of the peptides are predicted to form a β-sheet; therefore, the absorption at 1673 cm^{-1} is assumed to be due to an extended chain. The percent composition of the extended chain is significant in the three peptides (54% in α_1, 45% in α_3, and 59% in α_N), and the contribution may arise from the N and C termini of the peptides. At the extremities, the peptide

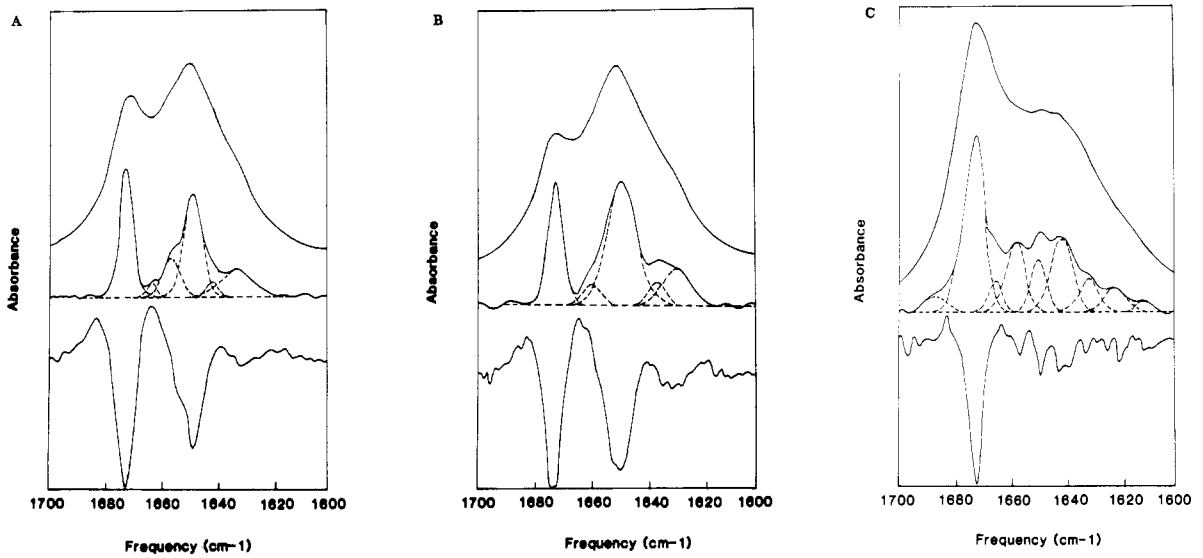

FIGURE 5: FTIR buffer-subtracted (top), deconvolved (middle), and second-derivative spectra (bottom) for (A) α_1, (B) α_3+, and (C) α_N peptides. The dashed curves in the deconvolved spectra are Gaussian fits to the data.

Table II: R_g Values for α_1 Determined by SAX Using Guinier and $P(r)$ Analyses[a]

concn (mg/mL)	R_r(Guinier) (Å)	$R_g[P(r)]$ (Å)
10	16.6 ± 0.5	16.3 ± 0.3
15	15.6 ± 0.4	16.1 ± 0.2
20	15.1 ± 0.3	15.8 ± 0.1
25	14.8 ± 0.2	15.3 ± 0.1
30	14.5 ± 0.2	15.2 ± 0.1
35	14.3 ± 0.2	14.5 ± 0.1
infinite dilution	16.7 ± 0.4	17.2 ± 0.1

[a] Interparticle interference at finite protein concentration contributes to an intensity loss of the scattering profiles at low angle so the R_g's are extrapolated to infinite dilution (Heidorn & Trewhella, 1988).

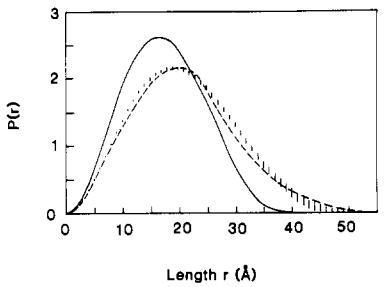

FIGURE 6: $P(r)$ curves for α_1, [vertical dash (|)], lysozyme [solid line (—)], and a prolate ellipsoid [short dash (- - -)] with semimajor and semiminor axes of 28 and 16 Å, respectively.

structure may not be well-defined.

Shape Analysis by SAXS. The small-angle data for the α_1 peptide indicate that the solution is monodisperse for all protein concentrations in the range 10–35 mg/mL at pH 7.4. An I_0/c calculation using lysozyme as a molecular weight standard gives a molecular weight for α_1 of 10570 ± 510, confirming that α_1 exists predominantly as a tetramer in solution since the calculate molecular weight of an α_1 tetramer is 8404. Although the molecular weights of lysozyme and the α_1 tetramer are nearly the same, their R_g's and $P(r)$ curves are quite different. The R_g for α_1 obtained from the $P(r)$ analysis is 17.2 ± 0.1 Å (16.7 ± 0.4 Å from the Guinier analysis, Table II), and the R_g for lysozyme is 14.0 ± 0.2 Å (compared with the published value of 14.3 ± 0.3; Kringbaum & Kugler, 1970). $P(r)$ is the frequency of vectors connecting small-volume elements within the entire volume of the particle weighed according to their X-ray scattering power. The rather symmetric $P(r)$ curve for lysozyme (Figure 6) therefore indicates it is a compact, globular protein with a spherical shape. The larger R_g of the α_1 tetramer indicates a more extended structure, and its $P(r)$ curve (Figure 6) can be modeled by using a prolate ellipsoid with a semimajor axis of 28 Å and a semiminor axis of 16 Å. This elongated structure is consistent with the proposed four-helix bundle for the α_1 tetramer.

The SAXS data for α_N were not analyzable because of the formation of large nonspecific aggregates. The aggregates dominated the scattering at low angle, hence making it impossible to derive parameters relevant to the monomer. The SAXS data for α_3 showed only a small degree of nonspecific aggregation that affected only the very lowest angle scattering data and hence could be effectively subtracted to leave a Guinier region. The I_0 derived from Guinier analysis gave a molecular weight for α_3 in phosphate buffer (50 mM), pH 7.4, of 8660 ± 200 (error derived from counting statistics). This agrees well with α_3 existing as a dimer (calculated molecular weight 9076) under these conditions. $P(r)$ curves and the R_g values indicate that the α_3 dimer has an overall shape very similar to the α_1 tetramer (data not shown).

Discussion

De novo peptide and protein design requires that rules by which the sequence of amino acids folds into a predetermined structure be known. However, since this so-called "second genetic code" has yet to be deciphered, we have adopted an incremental, knowledge-based approach to peptide engineering of immunogenic epitopes (Kaumaya et al., 1989). By using the framework model of protein folding (Ptitsyn & Rashin, 1975; Kim & Baldwin, 1982) as a starting hypothesis, as well as principles that determine the structure of proteins (Chothia, 1984; Richardson, 1981), and by considering the current knowledge about factors which enhance peptide and protein stability in aqueous environment, we have designed and synthesized various synthetic peptide models (Figure 1).

The framework model of protein folding proposes that fluctuating secondary structures (α-helices, β-sheets, β-turns, and loops) are first formed as nucleation sites induced by short-range interactions. Due to medium- and long-range interactions, these secondary structures associate to form supersecondary structural motifs ($\beta\alpha\beta$, $\alpha\alpha$, $\beta\beta\beta$, $\beta\alpha\beta\alpha\beta$, four-helix bundle) which then coalesce to form the tertiary structure. Thus, the formation of secondary structures plays a key role in the folding process (Mutter, 1985). The current dogma

was that short peptides do not show recognizable structures and are not predicted to have measurable helix formation in water according to the Zimm–Bragg theory (Zimm & Bragg, 1959). Accordingly, peptides of less than 20 amino acids are not helical in solution because the backbone hydrogen bonds do not provide the requisite stabilization. The fact is that some short peptides do have recognizable ordered secondary structures. Several reports have now appeared for linear peptides in water (Shoemaker et al., 1985; Brems et al., 1987; Gooley & MacKenzie, 1988; Dyson et al., 1985, 1988a,b). This seemingly expanding repertoire of defined secondary structures in short peptides which were hitherto considered structureless adds evidence for the feasibility of the framework model in which α-helices, β-sheets, or turns are postulated as "seeds" for initiation of folding (Baldwin, 1986).

In order to characterize the conformational properties of these model peptides, a variety of independent biophysical techniques were used in the present study. Examination of the X-ray crystal structure of LDH-C_4 shows that sequence 310–327, designated α_N, is in an extended α-helical configuration. Formation of α-helix by α_N was first investigated by circular dichroism in aqueous solution. The helicity of this fragment is barely measurable and typically could be described as random coil. A shift toward 208 nm was observed at increasing peptide concentration, indicating a small amount of helix formation that displayed no concentration dependency. The FTIR spectra of α_N show a much different absorbance spectrum than either α_1 or α_3+, and the amount of helix was estimated at approximately 14%. The SAXS data for α_N were not analyzable because of the nonspecific aggregation (30 mg/mL) which was not reversible by dilution. In conclusion, α_N shows very low α-helical content. These results are somewhat inconsistent with recent data obtained for some short peptides in aqueous solution in which the helix dipole model (Shoemaker et al., 1987) and electrostatic interactions (Shoemaker et al., 1985; Brems et al., 1987) have been invoked to explain their stability. One possible explanation for this discrepancy is that the free α-amino and α-carboxyl groups of α_N were not blocked with acetyl and amide groups; thus, the free charges may cause helix destabilization with the helix dipole. We are presently investigating this possibility and testing the pH dependency of helix formation to establish whether the electrostatic ion pairs (Glu^{312} and Lys^{316}, Glu^{323} and Lys^{327}, Lys^{317} and Asp^{320}) play any role in stabilizing the α-helix. A more plausible explanation for the lack of significant α-helical content in this fragment is the lack of medium- and long-range interactions, a property which is present in the native protein. The hydrophilic residues Glu^{311}, Ser^{318}, and Gln^{326} invade the hydrophobic face of the helix and thus could interfere with helix–helix packing. In a similar vein, the hydrophobic residue Trp^{323} would be solvent-exposed and thus cause destabilization of helical content.

On the other hand, helix formation by the idealized amphiphilic sequence α_1 is shown to be strongly dependent on peptide concentration, indicating that intermolecular interactions between hydrophobic surfaces ofthe amphiphilic surfaces is important for this high helix stability. Natural globular proteins are believed to fold by a similar mechanism involving hydrophobic interactions between adjoining segments of secondary structures. The fact that α_N does not form tetramers and α_1 does would seem to indicate that optimum complementarity of the interfaces is a prerequisite of stable self-association. Subunit association in natural proteins is highly specific and requires the subunits to be in their proper conformation before they coalesce to form the quaternary structure. The helical content of the α_1 monomer (−3013 deg·cm^2/dmol) is approximately one-sixth of the α_1 tetramer (−20 310 deg·cm^2/dmol). This corresponds to helical contents of approximately 13.3% and 86.1% for the monomers and the tetramers, respectively [compared to the ellipticity of poly(L-lysine)hydrochloride]. The free energies of tetramerization for α_1 were calculated to be −20.9 kcal/mol with a dissociation constant of 7.42×10^{-16} M^3. A direct comparison of the stabilities of our α_1 tetramer (18 residues corresponding to 5 turns of an α-helix, −20.9 kcal/mol) and the $\alpha_1\beta$ tetramer [16 residues corresponding to 4 turns of an α-helix, −21.9 kcal/mol, as reported by Ho and DeGrado (1987)] is not possible. The reported value of −21.9 kcal/mol was obtained by extrapolation to zero guanidine concentration, a chaotropic denaturant which had to be included in the buffer to allow dissociation of the peptide at a measurably experimental concentration.

The molecular areas of α_1 were further studied by peptide monolayers and analyzing monolayer surface pressure–area curves. Our results indicate that α_1 has an area of 17 Å^2/AA, characteristic of an α-helix, and does indeed form amphiphilic structures as shown by its behavior at the air–water interface. CD spectra of Blodgett slides collected from these monolayers are characteristically α-helical. The FTIR results are summarized in Table I and indicate that α_1 is approximately 40% α-helix at pH 7.4. This value may be reflective that the helix is localized to 40% of the molecule or that 40% of the molecules are completely helical. The end effects, free α-NH_3^+ and α-COO^- in α_1, may account for this result. Acetylation and amidation of the α-amino and -carboxyl groups, respectively, should cast light on this supposition. Our SAXS data indicate that the solution structure of α_1 is a tetramer whose shape is like a prolate ellipsoid with a radius of gyration of 17.2 ± 0.1 Å (compared to 14.0 ± 0.2 Å for lysozyme which has the same molecular weight but is more spherical in shape). The SAXS data are also consistent with α_1 being a tetramer.

In the next stage of the proposed framework model of experimental design, the stabilization of α_1 into a supersecondary $\alpha\alpha$ folding motif, α_3, was undertaken (Kaumaya et al., 1989). The success of the present design is dependent on a number of factors unlikely to be present in isolated short peptide fragments excised from segments of native protein. An inevitable question is the following: Would this experimental approach yield a compact structure which is stabilized by medium- and long-range interactions of the type found in globular proteins? A recognized factor in the folding of proteins is the burying of hydrophobic side chains. Packing of α-helices on α-helices ($\alpha\alpha$, $\alpha\alpha\alpha$), α-helices on β-sheets ($\beta\alpha$, $\alpha\beta$, $\beta\alpha\beta$), and β-sheets on β-sheets ($\beta\beta\beta$, $\beta\beta$) are classical examples of interdigitation of nonpolar side chains. For de novo protein design, it is therefore equally important to understand the nature of loops that connect pieces of secondary structures (Milner-White & Poet, 1986; Pabo & Suchanek, 1986). The connecting piece for joining the two helices into a β-hairpin turn (Leu-Ser-Pro-Gly) was based on the examination of 33 patterns in native proteins that code for $\alpha\alpha$ corners (Efimov, 1984). A more recent survey of 97 $\alpha\alpha$ loops (Thornton et al., 1988) has appeared. The choice of the hairpin turn was guided by a number of considerations designed to yield a stereochemically sensible conformation in which there were no disallowed close contacts, tight packing of the hydrophobic core, and exposure of polar charged groups to solvent. There is at present no direct evidence on the strengths and specificities of these interactions.

Ho and DeGrado (1987) designed a three-residue loop sequence consisting of an Arg-Arg-Pro to connect two α-helices

and provided evidence that these residues contributed to the stability and dimer formation of their $\alpha_2\beta$(PRR) sequence. This linking sequence in place of a single Pro prevented trimer formation. It would probably be more appropriate to consider the linking sequence as a five-residue loop because of the flanking glycines. Disulfide bridging (two Cys residue turn) has also been successfully used to link α-helices in parallel orientation in the coil-coiled systems (Hodges et al., 1989; O'Shea et al., 1989), and very stable dimers were formed. These results would seem to indicate that a variety of loop lengths are possible and they may be considered to play only a passive role in the folding process although they may be important in promoting the formation of a compact structure.

The choice of the antiparallel helix in our α_3 peptide was identical in sequence except it was synthesized in reverse orientation (N–C) such that it would display an antiparallel arrangement with the helix dipole running parallel. Disulfide-linked helices of the coil-coiled systems allow parallel arrangement of both helices and dipole. The linked helices reported by Ho and DeGrado are in a head-to-tail arrangement with the C terminus of one helix becoming the N terminus of the other helix and with the dipole running antiparallel. The successes of these various designs as assessed by their folding properties implicate that the folding is not tied down to a specific sequence of amino acid but that there is a fundamental commonality other than the parallel or antiparallel arrangements of the helices or dipoles.

Helix formation by α_3 is markedly dependent on peptide concentration. The helical content of α_3 monomer (−5059 deg·cm^2/dmol) is approximately one-fourth that of its dimer (−22 485 deg·cm^2/dmol) indicative that the dimer is considerably more stable than its monomer. These values correspond to helical contents of 23% and 99.8% for the monomer and dimer, respectively. The α_3 monomer is more stable than the α_1 monomer (−5059 vs −3013 deg·cm^2/dmol), and similarly the α_3 dimer is more stable than the α_1 tetramer (−22 485 vs −20 310 deg·cm^2/dmol). The free energy of dimerization of α_3 is −7.8 kcal/mol with a dissociation constant of 1.18×10^{-18} M^3. This value of association is consistent with intramolecular folding of α_3 prior to intermolecular stabilization by dimer formation, as the formation of the same tetrahelical aggregates releases less energy when starting from the dihelical α_3 monomer than when starting with the α_1 monomer. Ideally, the free energy of dimerization should be equal to half the free energy of tetramerization of α_1 (−10.4 kcal), but it is 75% in this case. Again, a direct comparison of the stability of the α_3 dimer (40 residues) with the Ho and DeGrado (1987) $\alpha_2\beta$(PRR) dimer (35 residues) is not possible. Other spectroscopic data are in agreement with the hypothetical folded structure. The amphiphilicity of α_3 was demonstrated by its ability to form a stable monolayer at the air–water interface consistent with dimer formation. The FTIR spectra show the expected bands for α-helix (1652 cm^2) and loop (1660 cm^{-1}) structure. The SAXS data for α_3 in phosphate buffer (50 mM), pH 7.4, indicate a molecular weight of 8000 and a radius of gyration of 17.3 Å, consistent with dimer formation.

Intermolecular association of α_1 (tetramerization) and α_3 (dimerization) appears to be the predominant force in stabilizing the postulated structures. The intermolecular association to yield a four-helix bundle tilted in such a manner that it has a left-handed twist is less well understood. A recent theoretical analysis (Chou et al., 1988) of the packing of pairs of α-helices has shown that the orientation angle $\Omega_0 = -154°$ is energetically the most favorable among several low-energy arrangements. Further helix-stabilizing forces such as charged group effect in the helix dipole model, electrostatic ion pairs, and salt bridges may also contribute to helix stability. As a deliberate design protocol, these forces were incorporated within α_1 and α_3. We have thus far been unable to determine to what extent any of these forces contribute to helix stability because these potential interactions are being shielded by the hydrophobic interaction resulting from intermolecular packing of the lipophilic surface of the amphiphilic helices. This factor is absent in α_N, and experiments are underway to dissect these components. These results support the claim that medium- and long-range interactions are necessary to stabilize ordered structures such as α-helices in proteins and demonstrate relevance to the framework model of protein folding. The implication of these observations for studying the mechanism of protein folding is also considerable. Because protein folding occurs in a spontaneous fashion, folding intermediates have been difficult to characterize. Thus, accessibility of these folding intermediates through design and synthesis should provide an invaluable tool for studying folding pathways when coupled with immunochemical methods.

In conclusion, we have been able to design and synthesize a 40-residue synthetic peptide of defined supersecondary structure whose conformation has been confirmed by four independent biophysical techniques. These studies to assess quantitatively the stability of marginally stable peptides may open the way for a generalized scheme for analyzing these intermediate structures. In this respect, crystallization for X-ray studies and 2D NMR are also presently being investigated. The synthesis and characterization of structures like α_3 are not a serious problem because they are made up of one type of secondary structure. Stabilization of α_1 in $\alpha\beta$, $\beta\alpha\beta$, and other supersecondary structural motifs displaying parallel and antiparallel orientations, and consisting of identical and/or unidentical helices, have been synthesized, and their conformations are presently being investigated (P. T. P. Kaumaya, unpublished results). Studies on the physicochemical, conformational properties as described above and the interactions between different and secondary structural units would provide valuable information for de novo design of topographic immunogenic peptides which in turn could be correlated with their immunogenic capacities.

Added in Proof

A recent report (Gilson & Honig, 1989) indicates that the helix dipole is actually a destabilizing factor in the formation of helix bundles.

Acknowledgments

We appreciate the assistance of Martine Benoit in the preparation of the manuscript.

References

Adamson, A. W. (1976) *Physical Chemistry of Surfaces*, 3rd ed., Chapter 2, Wiley, New York.

Atherton, E., Caniezel, M., Fox, H., Harkiss, D., Over, H., & Sheppard, R. C. (1983) *J. Chem. Soc., Perkin Trans. 1*, 65–73.

Baldwin, R. L. (1986) *Trends Biochem. Sci. 11*, 6–9.

Barany, G., & Merrifield, R. B. (1979) *The Peptides* (Gross, E., & Meinehofer, J., Eds.) Vol. 2, Academic Press, New York.

Benjamin, D. C., Berzofsky, J. A., East, I. J., Gurd, F. R. N., Hannum, C., Leach, S. J., Margoliash, E., Michael, J. G., Miller, A., Prager, E. M., Reichlin, M., Sercarz, E. E., Smith-Gill, S. J., Todd, P. E., & Wilson, A. C. (1984) *Rev. Immunol. 2*, 67–101.

Brems, D. N., Plaisled, S. M., Kauffman, E. W., Lund, M., & Lehrman, S. R. (1987) *Biochemistry 26*, 7774–7778.
Byler, D. M., & Susi, H. (198) *Biopolymers 25*, 469.
Chothia, C. (1984) *Annu. Rev. Biochem. 53*, 537–572.
Chou, K. C., Maggiora, G. M., Nenethy, G., & Scheraga, H. A. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 4295–4299.
Cohen, C., & Parry, D. A. A. (1986) *Trends Biochem. Sci. 11*, 245.
Cornell, D. G. (1979) *J. Colloid Interface Sci. 70*(1), 167–179.
Daniels, S. B., Reddy, P. A., Albrecht, E., Richardson, J. C., & Richardson, D. C. (1988) in *Peptides: Chemistry and Biology, Proceedings of the 10th American Peptide Symposium* (Marshall, G. R., Ed.) p 383.
DeGrado, W. F., & Lear, J. D. (1985) *J. Am. Chem. Soc. 107*, 7684–7689.
DeGrado, W. F., Wasserman, Z. R., & Lear, J. D. (1989) *Science 243*, 622–628.
Dyson, H. J., Cross, K. J., Houghten, R. A., Wilson, I. A., Wright, P. E., & Lerner, R. A. (1985) *Nature 318*, 480–483.
Dyson, H. J., Rance, M., Houghten, R. A., Wright, P. E., & Lerner, R. A. (1988a) *J. Mol. Biol. 201*, 201–217.
Dyson, H. J., Rance, M., Houghten, R. A., Lerner, R. A., & Wright, P. E. (1988b) *J. Mol. Biol. 201*, 161–200.
Efimov, A. V. (1984) *FEBS Lett. 166*(1), 33.
Fetrow, J. S., Zenfus, M. H., & Rose, G. D. (1988) *Bio/Technology 6*, 167–171.
Gilson, M. K., & Honig, B. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 1524–1528.
Gooley, P. R., & MacKenzie, N. E. (1988) *Biochemistry 27*, 4032–4040.
Guiner, A. (1939) *Ann. Phys. (Paris) 12*, 161.
Heidorn, D. B., & Trewhella, J. (1988) *Biochemistry 27*, 909–915.
Ho, S. P., & DeGrado, W. F. (1987) *J. Am. Chem. Soc. 109*, 6751–6758.
Hodges, R. S., Semchuk, P. D., Taneja, A. K., Kay, C. M., Parker, J. M. R., & Mant, C. T. (1988) *Pept. Res. 1*, 19–29.
Hogrefe, H. H., Griffith, J. P., Rossmann, M. G., & Goldberg, E. (1987) *J. Biol. Chem. 262*, 13155–13162.
Hogrefe, H. H., Kaumaya, P. T. P., & Goldberg, E. (1989) *J. Biol. Chem. 264*, 10513–10519.
Jaenicke, R. (1987) *Prog. Biophys. Mol. Biol. 49*, 117–237.
Johnson, W. C., Jr. (1988) *Annu. Rev. Biophys. Chem. 17*, 145–166.
Kaiser, E. T., & Kezdy, F. J. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 1137–1148.
Kaiser, E. T., & Kezdy, F. J. (1984) *Science 223*, 249–255.
Kaumaya, P. T. P., O'Hern, P., & Goldberg, E. (1989) (to be submitted for publication).
Kauppinen, J. K., Mofatt, D. J., Mantsch, H. H., & Cameron, D. C. (1981) *Appl. Spectrosc. 35*, 271.
Kauzmann, W. (1959) *Adv. Protein Chem. 14*, 1–63.
Kim, P. S., & Baldwin, R. L. (1982) *Annu. Rev. Biochem. 51*, 459–489.
Krimm, S., & Dwivedi, A. M. (1982) *Science 216*, 407.
Kinngbaum, W. R., & Kugler, F. R. (1970) *Biochemistry 9*, 1216–1223.
Lau, S. Y., Taneja, A. K., & Hodges, R. S. (1984) *J. Biol. Chem. 259*, 13253–13261.
Lee, B. K., & Richards, F. M. (1971) *J. Mol. Biol. 55*, 379–400.
Lerner, R. A. (1982) *Nature 299*, 592–596.
Lerner, R. A. (1984) *Adv. Immunol. 36*, 1–144.
Levitt, M., & Chothia, C. (1976) *Nature 261*, 552–558.
Marqusee, S., & Baldwin, R. L. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8898–8902.
Milner-White, E. J., & Poet, R. (1986) *Biochem. J. 240*, 289–292.
Moore, P. B. (1980) *J. Appl. Crystallogr. 13*, 168.
Mutter, M. (1985) *Angew. Chem., Int. Ed. Engl. 24*, 639.
Mutter, M., Altmann, K. H., & Vorherr, T. (1986) *Z. Naturforsch. 41B*, 1315–1322.
Mutter, M., Altmann, E., Altmann, K. H., Hersperger, R., Koziej, P., Nebel, K., Tuchscherer, G., Vuilleumier, S., Gremlich, H. V., & Muller, K. (1988a) *Helv. Chim. Acta 71*, 835.
Mutter, M., Altmann, K. H., Tuchscherer, G., & Vuilleumier, S. (1988b) *Tetrahedron 44*, 771–785.
Novotny, J., Handschumacher, M., Haber, E., Bruccoleri, R. E., Carlson, W., Fanning, D., Smith, J. A., & Rose, G. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 226–230.
O'Shea, E. K., Rutkowski, R., & Kim, P. S. (1989) *Science 243*, 538–542.
Pabo, C. O., & Suchanek, E. G. (1986) *Biochemistry 25*, 5987–5991.
Ptitsyn, O. B., & Rashin, A. A. (1975) *Biophys. Chem. 3*, 1–20.
Richardson, J. S. (1981) *Adv. Protein Chem. 34*, 167–339.
Shoemaker, K. R., Kim, P. S., Brems, D. N., Marqusee, S., York, E. J., Chaiken, I., Steward, J. M., & Baldwin, R. L. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 2349–2353.
Shoemaker, K. R., Kim, P. S., York, E. J., Stewart, J. M., & Baldwin, R. L. (1987) *Nature 326*, 563–567.
Stewart, J. M., & Young, J. (1984) in *Solid Phase Peptide Synthesis*, 2nd ed., Pierce Chemical Co, Rockford, IL.
Sundaralingham, M., Sekharadu, Y. C., Yathindra, N., & Ravichandran, V. (1987) *Proteins: Struct., Funct., Genet. 2*, 64–71.
Taylor, J. W., & Kaiser, E. T. (1986) *Pharmacol. Rev. 38*, 291.
Thornton, J. M., Sibanda, B. L., Edwards, M. S., & Barlow, D. J. (1988) *Bioessays 8*, 63–69.
Trewhella, J., Liddle, W. K., Heidorn, D. B., & Strynadka, N. (1989) *Biochemistry 28*, 1294–1301.
Weber, P. C., & Salemme, F. R. (1980) *Nature 287*, 82–84.
Yada, R. Y., Jackman, R. L., & Nakai, S. (1988) *Int. J. Pept. Protein Res. 31*, 98–108.
Zimm, B. H., & Bragg, J. K. (1959) *J. Chem. Phys. 31*, 526–535.