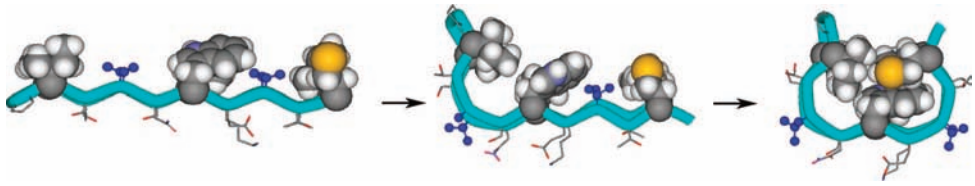
ACCOUNTS
of chemical research

Protein Design with L- and D-α-Amino Acid Structures as the Alphabet

SUSHEEL DURANI
Department of Chemistry, IIT Bombay, Powai, Mumbai 400076, India

RECEIVED ON DECEMBER 1, 2007

CONSPECTUS



Summarizing the implications of homochiral structures in interpeptide interactions, not only in the topology but also possibly in the physics of protein folding, this Account provides an overview of the concept of shape-specific protein design using D- and L-(α)amino acid structures as the alphabet. The molecular shapes accessible in de novo protein design are stereochemically defined. Indeed, the defining consideration for shape specificity in proteins to be α-helix/β-sheet composites is the L configuration of the α-amino acid structures. The stereospecificity in shapes implies that protein shapes may be diversifiable stereochemically, that is, designable de novo, using D and L structures as the alphabet. Indeed, augmented with D enantiomers, Nature's alphabet will expand greatly in the diversity of polypeptide stereoisomers, for example, from 1^{30} to 2^{30}—that is, from one to ca. one billion—for a modestly sized 30-residue polypeptide. Furthermore, with each isomer having conformers stereospecific to its structure, molecular folds of specific shapes may be approachable sequentially when D and L structures are used as the alphabet. Illustrating the promise, 14−20-residue bracelet-, boat-, canoe-, and cup-shaped molecular folds were designed stereochemically or implemented as specific sequence plans in the D- and L-α-amino acid alphabet. In practical terms, canonical poly-L peptide folds were modified to the desired shapes via stereochemical mutations invoking enantiomer symmetries in the Ramachandran ϕ,ψ space as the logic. For example, in designing the boat-shaped fold, the canonical β-hairpin was reengineered in its flat planar structure via multiple coordinated L-to-D mutations in its position specific cross-strand neighbor residues, upturning its ends enclosing six side chains in a molecular cleft. While affirming the generality of the approach, the 20-residue molecular canoe and the 14-residue molecular cup are also presented as examples of the scope of functional design. The canoe, possessing alkali cation-specific catgrips in its main chain, and the cup, featuring an organic cation-specific aromatic triad in its side chains, do indeed display desired specificities in their ligand binding. Stereochemistry is, therefore, the crucial specifier of protein shapes and valuable as the tool for shape-specific protein design. Proteins in general, whether poly-L or mixed-D,L, require sequence effects of amino acid side chain structures for their stability, if not also for specifying them conformationally. The principles underlying these phenomena remain a puzzle, but studies invoking a stereochemical mutation approach to the problem have suggested that the poly-L structure may be crucial to the principles of sequential encoding of protein structures in amino acid side chains as the alphabet.

Introduction

The molecules crucial to life are the heteropolymers specific in their building block alphabets. The algorithm of sequence level molecular programming, manifesting its scope and versatility in protein evolution, evokes interest for *de novo* design. Whether designing proteins in Nature's alphabet[1,2] or foldamers in artificial alphabets,[3] the principles of protein folding assume critical importance. The principles remain a puzzle,[4−6] and modifying the Nature's alphabet in specific structural detail may illuminate the puzzle as well as extend *de novo* protein design in its scope. The specific structural detail in the Nature's alphabet of interest in this

Published on the Web 07/22/2008 www.pubs.acs.org/acr
10.1021/ar700265t CCC: $40.75

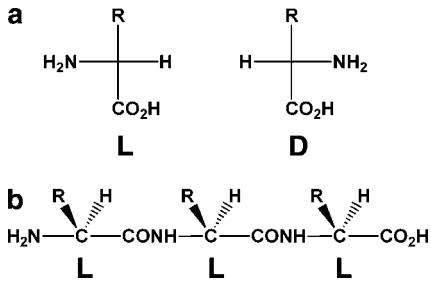


FIGURE 1. (a) The α-amino acids with R ≠ H are asymmetric with the possibility of L or D configurational structure. (b) The chains defining proteins are in identical L configuration in every substituted (R ≠ H) α-carbon.

Account is that of stereochemistry, specifically, the L configuration in proteinogenic α-amino acid structures (Figure 1). The question of how the natural parity between the D- and L-α-amino acid enantiomers became violated in the course of life's origin is important and has had a history of research.[7] The question of what purpose was served in, for instance, the translational logic of protein synthesis, the topological logic of protein structure, and the physical logic of protein folding evokes interest. Proteins are α-helix/β-sheet composites, and clearly the homochiral, poly-L peptide structure is the reason.[8] Could stereochemistry be relevant as the tool to probe protein-folding principles and to extend the scope in *de novo* protein design rationally? These are the issues we address in this Account. Introducing protein design in the D- and L-α-amino acid alphabet, we discuss parenthetically the notion that the sequential encoding of protein conformation may have a stereochemical basis.

The Chemical Scripts of Protein Structure

In designing a protein *de novo*, the polypeptide structure needs to be considered in its conformational space and in its alphabet to control conformation. The protein conformational canvas is in polypeptide main chain, while its design alphabet is in amino acid side chains (Figure 1). The peptide groups are planar due to amide resonance[9] and define the protein conformational space in the rotational freedom of their bonds joining tetrahedral α-carbon (Figure 2).[10] Side chains constrain conformational freedom in the bonds, defining the zones of access in ϕ,ψ space; stereochemistry is important in causing the zones to be D or L specific (Figure 3). Protein structures correspond to the L specific zones of ϕ,ψ space; *de novo* protein design is implemented typically in the L specific zones of ϕ,ψ space. Interestingly, the D specific zones were accessed, and the HIV protease in mirror-image relation of the natural poly-L variant was approached artificially.[11] Distinct from the proteins poly-L or poly-D in their stereochemical structure, here we are focused on the proteins that are complex in their diastereomeric structures due to D and L amino acids being the alphabet.

With unsubstituted α-carbons, polyglycines are random coils that could populate the ϕ,ψ space statistically in a centrosymmetric distribution (Figure 3). Side chains reduce the ϕ,ψ space stereospecifically.[10] Poly-L in their stereochemical structure, proteins populate in ϕ,ψ space in the L specific $^{L}\alpha_R$, $^{L}\beta$, and $^{L}\alpha_L$ basins, except in conformationally flexible glycine and constrained L-proline. Overall, ϕ values tend to be negative, while in ψ, there is the choice of sign corresponding to $^{L}\alpha$-helix[12] (negative ϕ and negative ψ) and $^{L}\beta$-strand structures[13] (negative ϕ and positive ψ) (Figure 4), the building blocks of protein structure. Being structurally invariant in the building blocks and limited in the topological options to join them, the protein structure remains greatly restricted in morphological possibilities of its tertiary structure.[14] Stereochemistry, quite clearly, is the reason.

Possible Stereochemical Logic of the Protein Chemical Scripts

The definition of protein conformational choices is clear in the principles, but selection of the choice specific for each protein remains a puzzle.[4–6] The principle are in the energetics of protein structures as the native structures are the sequentially specified minima of free energy.[15] Side chains specify the minima, but the apparatus responsive to their effects is in the system of linked peptides. That the apparatus may be in homochiral structure of the interactions in the system was implied in a puzzling observation of Flory[16,17] as clarified in the studies of Ramakrishnan et al.,[18,19] briefly as follows. Having noted that poly-L peptide random coils will be about three times the size of poly-L ester random coils, in contravention of almost the rule based dependence on excluded volumes, Flory implicated Coulomb interaction between peptide dipoles as the specific reason.[16] His contention that α_L/α_R conformation is disfavored by interpeptide electrostatics was affirmed in the subsequent findings that polypeptide random coils and denaturant-unfolded proteins are in largely $^{L}\beta$ conformation, specifically in the closely related ^{L}PPII conformation (Figure 3).[20–22] Having thus implicated interpeptide electrostatics in critical role, Flory found that mutated to alternating-L,D structure, poly-L peptide random coil collapses 10-fold, surprisingly with chain length and to the size smaller than that of a theoretical abstraction lacking both excluded volume and fixed bond angle constraints.[16]

Ramakrishnan et al.[18,19] addressed the puzzle as well as the controversy whether the heptalanine model peptide XAO

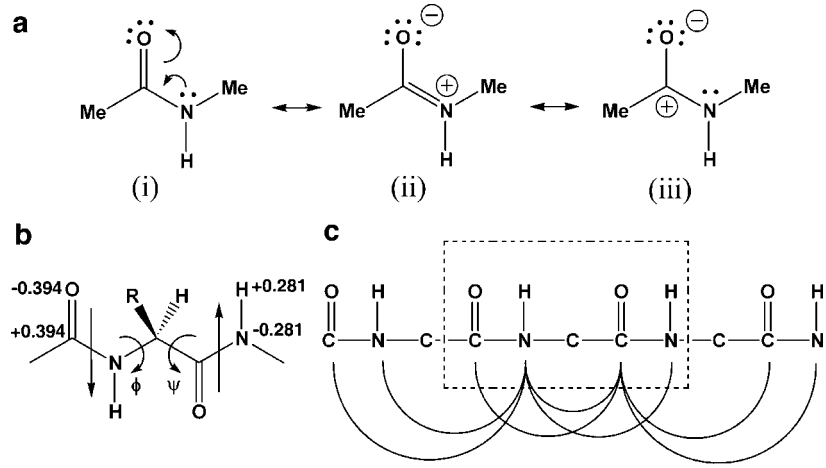


FIGURE 2. (a) Nitrogen loan pair and carbonyl π-bond resonance imposes planarity and defines peptide amide bond polarity. The traditional resonance model was modified to include the canonical structure iii.[9] (b) The bond rotations defining ϕ,ψ space in polypeptide structure. The partial charge assignments of Flory[16] are shown that define Coulombic interactions between peptides. (c) The interacting groups in a dipeptide and its neighbors defining the local Coulomb interactions, as distinct from those nonlocal including hydrogen bonds between peptides.

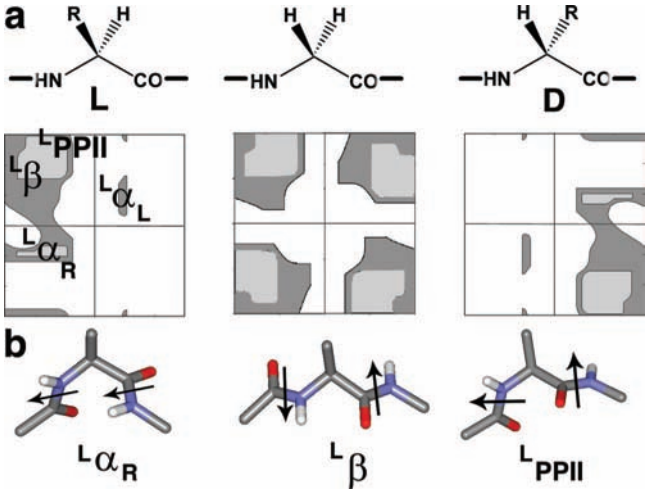


FIGURE 3. (a) Steric filtration of ϕ and ψ space in glycine dipeptide (R = H) (middle panel) to that in α-substituted dipeptides (R ≠ H) is stereospecific for the configuration in α-carbon. (b) The dipole orientations in a dipeptide unit manifest, besides mutual electrostatics, the hydrogen bonding interaction characterizing the proteins in folded ($^{L}\alpha_R$ and $^{L}\beta$ options) and unfolded conformation (^{L}PPII option).

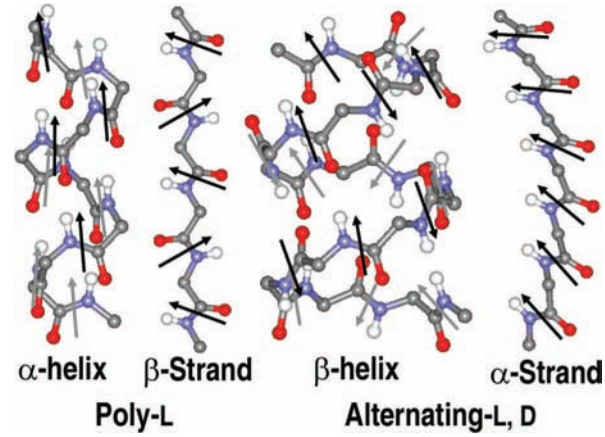


FIGURE 4. (a) The α-helix ($^{L}\phi = -57°$; $^{L}\psi = -47°$) and β-strand ($^{L}\phi = -120°$; $^{L}\psi = 120°$) motifs of poly-L structure and the β-helix ($^{L}\phi \approx -120°$; $^{L}\psi \approx 120°$; $^{D}\phi \approx 120°$; $^{D}\psi \approx -120°$) and α-strand ($^{L}\phi = -57°$; $^{L}\psi = -47°$; $^{D}\phi = 57°$; $^{D}\psi = 47°$) motifs of alternating-L,D structure are contrasted in the peptide dipole arrangements, mutually parallel in α conformation and antiparallel in β conformation. Mutual hydrogen bonding and Coulomb interactions between peptides is in harmony or conflict dependent upon polypeptide stereochemistry.

was a PPII helix or an ensemble of β-hairpin-like folds.[23,24] Studies involving molecular dynamics resulted in homochiral structure being implicated as possibly the apparatus of solvent and sequence control of conformation due to the electrostatic frustrations that if eliminated would cause alternating-L,D polypeptides to not only collapse but also lose sensitivity to solvent.[18] Poly-L structure in the chain of linked peptides, defining the protein backbone, was implicated in a critical role for electrostatic and steric reasons, briefly as follows.

The packing arrangements and the interactions among peptides within folded proteins, via short-range hydrogen bonds and long-range Coulombic interactions, are stereochemically defined. The interactions are in mutual conflict when the structure is poly-L but are in harmony when the structure is alternately L,D (Figure 4).[19] Thus, stereochemistry determines whether protein folding, while possibly driven by hydrogen bonding, will be favored or disfavored by interpeptide electrostatics, that is, will or will not be electrostatically frustrated. The interactions among peptides entail frustrations in the choice of ϕ as well due to homochiral structure in the peptide chain. The stereochemically less favored ϕ values, critical for protein globularity, are required in the linkers between

α-helix and β-sheet motifs[8] and necessitate either conformationally mobile glycine or L amino acids, which can penalize a fold for entropic or steric reasons. In their molecular dynamics experiments of polypeptide structure, Ramakrishnan et al. observed that the sampling in ϕ was stereochemically frustrated when the structure was poly-L but not when it was alternately L,D.[19] Given their spread in ϕ,[25] proteins are liable to be conformationally frustrated for entropic and steric reasons, besides electrostatic reason.

In summary, the options of conformation defining proteins involve interpeptide interactions that are liable to be electrostatically, sterically, or entropically frustrated, and specifically due to homochiral structure in the peptide chain. The effects may be involved in mediating the effects of solvent and sequences in protein conformation. Dielectric arbitration in electrostatically conflicted interpeptide interaction is a possible specific mechanism according to the electrostatic screening model of Avbelj.[26,27] The frustrated electrostatics of α-helix vs β-sheet selection, implied in the model, is an effect of poly-L structure according to Ramakrishnan et al.[18,19]

Stereospecificity of Protein and Polypeptide Conformation

Stereochemistry defines proteins sterically in the options of ϕ and ψ. The ϕ values are enantiospecific, while the ψ values present a choice of selectable options (Figure 3). ^{L}α-Helix and ^{L}β-sheet, the minima of free energy in folded proteins, are defined sterically in the constraint of poly-L structure in the hydrogen bonding among peptides.[12,13] In unfolded or denatured protein, ^{L}PPII is the minimum of free energy due to the steric, electrostatic, and solvation effects of the peptide chain of homochiral structure.[20,21] ^{L}α-Helix, ^{L}β-strand, and ^{L}PPII, the minima of free energy characterizing folded and unfolded proteins, are conformers of the peptide chain defined by its homochiral structure (see Figures 3 and 4).

Proteins by no means are exclusively homo-enantiospecific; motifs of mixed enantiospecific structures do occur and are important in their structural or functional roles. In considering these motifs, we need to distinguish enantiomerism in the stereochemical sense from that in the conformational sense. The enantiomerism of structure refers to the isomerism in the attachment stereochemistry of side chains (Figure 1). The enantiomerism of conformation, on other hand, refers to the isomers that have ϕ,ψ values inverted in sign (Figure 3). The ^{L}α-helix and ^{L}β-sheet motifs are the homo-enantiomeric motifs that are negative in all their ϕ values (Figure 4). The hetero-enantiospecific protein motifs involve either glycine or L residues in D enantiospecific conformation, that is, in positive ϕ.

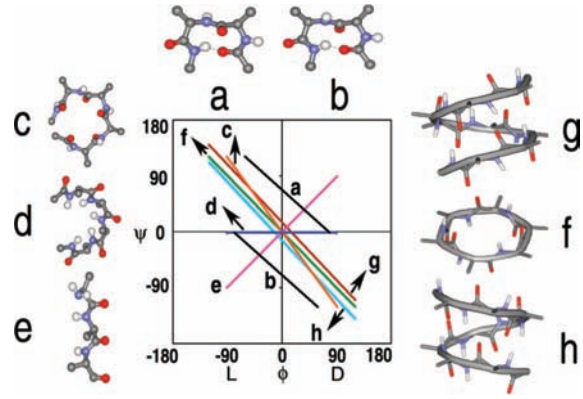


FIGURE 5. The ϕ,ψ values of mixed-L,D folds locally or globally periodic in stereochemical structure: (a, b) type II and II′ b-turns; (c) catgrip with the peptide C=O′s in cation binding mode; (d) anion nest with the peptide N−H′s in anion binding mode; (e) K^+ selectivity filter with peptide C=O′s in parallel arrays; (f) cyclic peptide building blocks of nanotubes; (g, h) left- and right-handed β-helix folds of gramicidin-A.

The best-recognized mixed enantiospecific protein motifs are the β-turns (Figure 5). β-Turns are important primarily for their structural role in the chain reversal for globularity.[28,29] The most common variants involve the residues in i and $i + 3$ sequence positions mutually hydrogen bonded and enclosing the $i + 1$ and $i + 2$ positional residues in specific ϕ,ψ values. The range of conformational and stereochemical possibilities for the overall tetrapeptide segments are LLLL (type I and type III β-turns), LDDL (type I′ β-turn), LLDL (type II β-turn), and LDLL (type II′ β-turn). The type I and I′ and the type II and II′ turns are in mutual enantiospecific relationship due to the middle position ϕ values being, respectively, L,L $(-\phi,-\phi)$ vs D,D $(+\phi,+\phi)$ and L,D $(-\phi,+\phi)$ vs D,L $(+\phi,-\phi)$. The longer tetrapeptide segments in conformational and stereochemical antisymmetry with the natural variants, namely, DDDD, DLLD, DDLD, and DLDD, are the targets potentially designable in L and D structures as the alphabet.

Unique in conformation, β-turns are position specific in their amino acid preferences,[28,29] which is the recipe to design a variant of interest.[30,31] The position crucial to hetero-enantiospecific β-turns is the D enantiospecific position. In proteins, the position has the option of glycine or an L residue. L-Proline is strong in its positional preference and can lock a specific variant due to its constrained ϕ.[28,30,31] D-Amino acids, ideal for D enantiospecific positions, are powerful aids in *de novo* design; L- and D-prolines will lock the desired enantiomer in a β-turn.[30,31] Artificial amino acids, modified pep-

tides, and even unrelated molecular structures capable of serving as β-turn mimetics or surrogates have been important subjects of research over decades.[32]

One interest in β-turns stems from their varied structural roles.[33] The hetero-enantiospecific variants, of specific interest in this perspective, can be specific in the structural roles. The examples include the type II′ turn as a helix nucleator[34] and the type I′ and type II′ turns as hairpin nucleators.[30,31] In fact, the type I′ and type II′ β-turns define the most common β-hairpins of the protein structure.[33] Hairpins are the motifs with extended β-stands side-by-side hydrogen bonded and joined in a central β-turn. The type I′ (LDDL type) and type II′ (LDLL) β-hairpins are morphologically distinct; the type I′ variant has a helical twist, while the type II′ variant is flatter. The β-hairpins stabilized with stereochemically constrained β-turns have been the focus of intense research being the simplest possible models of the β-sheet protein motif,[31] similar to the helix models stabilized with constrained residues, most notably with achiral α-aminoisobutyric acid (Aib).[35,36]

An important class of hetero-enantiospecific protein motifs features peptides as anion or cation recognition centers (Figure 5). Anion interaction centers termed anion nest,[37] cation interaction centers termed catgrip,[38] and potassium selectivity filter[39] are alternately enantiospecific, and the individual variants seem to be function specified in the ψ values as implied in Figure 5. Indeed, the chain topology (linear or curved, concave or convex) defining the interaction specificities of the motifs is specific in ψ values (Figure 5). In its ψ, the linear variant is in closer correspondence to α-helix, while the curved variant concave in the registry of NHs is in closer correspondence to β-sheet. Peptide carbonyls in mutual parallel alignment define the K^+ selectivity filter at the symmetry axis in a C_4 tetramer protein.[39] With curvature, the concave arrays in peptide C═O′s define catgrips, while the concave arrays in peptide N−H′s define anion nests (Figure 5).

Gramicidin-A is a 15 residue microbial peptide alternately L,D in structure (Figures 4 and 5).[40−42] The molecule is a cation channel in its β-helical fold, modeled in Figure 4, enclosing a pore lined with the peptide planes in β-sheet-type mutual hydrogen bond registry. Alternately enantiospecific, the β-helix is in close correspondence to β-sheet ϕ,ψ values. However, the chain with ϕ,ψ values alternating in sign, propagating a flat circular course, will be sterically overlapped in its termini (Figure 5); small displacements in ψ create helical aspect, right- or left-handed, as noted in Figure 5, defining β-helices of the opposite screw senses. Alternately L,D, the gramicidin β-helix is in harmony of hydrogen bonding and Coulombic interactions between its peptides, as noted in Figure 4. Possibly for this reason the peptide is in collapsed, that is, folded, conformation under diverse solvents even in the membrane interiors serving as ion channel.[40−42] The apparent solvent indifference of this alternately L,D fold is in accord with the notion that the solvent-mediated folding−unfolding of the natural protein structures may be an effect of poly-L structure.[19] Alternately-L,D peptides of cyclic structures are flat planar ring-like structures due to the alternately enantiospecific ϕ,ψ values (Figure 5). With peptide dipoles at right angles to the molecular plane and capable of hydrogen bonding intermolecularly, the peptides define self-organizing nanotube structures.[43]

Scripting Proteins Stereochemically

A total of 2^N polypeptide stereoisomers are possible over N α-carbons in L or D structure. This corresponds to 2^{30} ($\sim 10^9$) stereoisomers for even a modestly sized 30-residue polypeptide. The natural proteins correspond to but one of the astronomical possibilities, the poly-L diastereomer. The ϕ coordinate being stereospecific, the possibilities of protein structure are defined essentially in the options of ψ. Freeing up ϕ would diversify proteins, while enantiospecificity of the coordinate would place the protein shapes under the control of L and D structure. Hetero-enantiospecific protein motifs can be designed with D amino acids in D enantiospecific positions. β-Hairpins,[30,31] α-helices,[34] and even mixed-α,β miniproteins[44] have thus been accomplished invariably with mixed-L,D β-turns as the folding nuclei. Although hetero-enantiomeric in the folding nuclei, clearly the higher order folds in these designs remain primarily homo-enantiospecific (poly-L) in the overall shapes. The polypeptide folds in overall hetero-enantiospecific shapes occur in the microbial world. Gramicidin-A in an alternately enantiospecific fold.[40−42] A cyclic depsipeptide, alternately enantiomeric valinomycin, is a ring-shaped cation carrier.[45] Tolaasin, a mixed L,D polypeptide, is a golf-club-shaped molecule.[46]

Hetero-enantiomeric structures are a rarity in the biomolecular world. Ribosomal synthesis, incompatible with D amino acids, will produce only small, sequence-local elements within protein structures, with glycine or L amino acids serving as the D enantiospecific residues. Site-specific D amino acid incorporation has been made possible through the ribosomal route[47] but will not in its present capability support the unfettered use of D structure as the alphabet. The microbial heteroenantiomeric polypeptides are made by nonribosomal synthesis. Without the assistance from mRNA templates, the syntheses necessitate highly specialized, often complex, biosynthetic machinery. Given that such machinery exists, the L and D

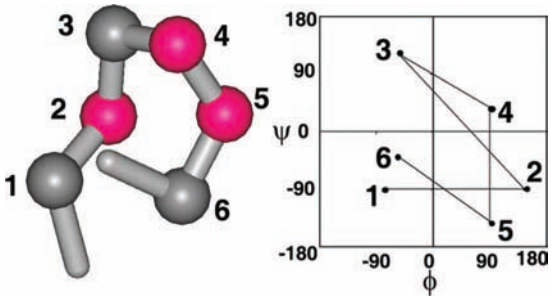


FIGURE 6. A designed peptide aperiodic in stereochemical and conformational structures, showing the fold topology (D residues in red color) and ϕ,ψ values in the crystal structure of Boc-^{L}Leu-^{D}Val-^{L}Pro-^{D}Asp-^{D}Val-^{L}Leu-OMe.[48]

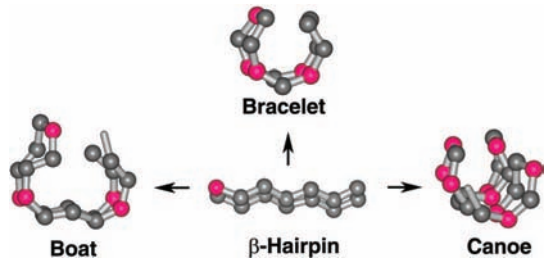


FIGURE 7. Schematics of the coordinated L to D mutations transforming canonical poly-L β-hairpin/β-sheet folds as bracelet-, boat-, and canoe-shaped molecules.

alphabet seems compelling in its evolutionary value; the complexities of nonribosomal biosynthesis, however, seem to restrict the scope to the hetero-enantiomeric structures either shorter and localized or longer but stereochemically periodic. Longer and stereochemically aperiodic sequences would seem biosynthetically impossible but are routine in chemical synthesis via the solid phase approach.

Fabiola et al.[48] reported arguably the first stereochemically and conformationally aperiodic polypeptide; an LDLDDL type hexapeptide was shown to be $^{L}\alpha^{D}\beta^{L}\beta^{D}\alpha^{D}\beta^{L}\alpha$ type fold in its crystal state conformation (Figure 6). Rana et al. reported the first stereochemically aperiodic proteins as bracelet-,[49] boat-,[50] canoe-,[51] and cup-shaped[52] molecules, designing the canoe as an alkali cation receptor and the cup as an acetylcholine receptor (Figures 7 and 8).

Considering the multiplicity of effects, evolutionary algorithms are the desired option in implementing shape-specific protein design. The computational methods of protein structure prediction and inverse design may be beneficial for the purpose. The accomplished hetero-enantiomeric proteins illustrate the inverse approach; the folds were designed in L and D alphabet, and then the side chains were selected in the elements of an evolutionary algorithm.[53] In-house software, computer-aided peptide modeler (CAPM), assisted in modeling the folds, while explicit solvent molecular dynamics assisted in

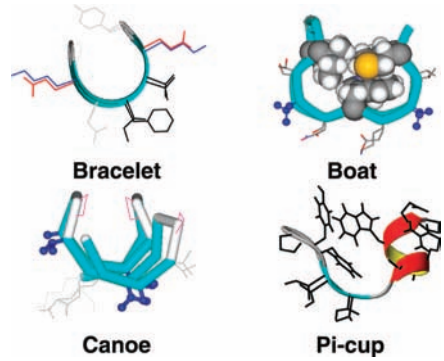


FIGURE 8. Schematic structures of the bracelet-, boat-, canoe-, and cup-shaped polypeptide molecules.

optimizing the side chains from *ad hoc* possibilities, as is typical in *de novo* protein design.

The reported examples of shape-specific protein design exploit a common approach; type II or II′ β-hairpin elements, nucleated appropriately in ^{L}Pro-Gly or ^{D}Pro-Gly β-turns, are mutated in the cross-strand neighbors from LL to DD structures so as to achieve a local changeover of the direction between the main chain and side chain elements, and thus to the required shapes as illustrated in Figure 7. Modified in their β-hairpin elements, several of the native-like poly-L peptide folds were reengineered in the shapes as desired (Figure 8). The bracelet, a molecule found to be well-ordered in both water and DMSO,[49] is a canonical 14-residue β-hairpin reengineered stereochemically to the bracelet type morphology. The canoe, with affinity for alkali and alkaline-earth metal ions having been designed as a double catgrip mimic,[51] is the stereochemical mutant of a 20-residue four-stranded all-β miniprotein. The boat, a molecule well-ordered in water and harboring an enzyme-like cleft,[50] is the reengineered variant of a canonical 20-residue β-hairpin fold. The cup, harboring an aromatic triad in a binding pocket specific for trimethylalkylammonium-type organic cations,[52] is a 14-residue canonical α,β-construct, with a stereochemically bracketed helix and a stereochemically nucleated β-hairpin joined in a flexible glycine linker, mutated stereochemically as desired. The peptide binds acetylcholine and has nanomolar affinity for the ligand.

Postscript

Sequential encoding of protein structure puzzles about the relevance of a homochiral alphabet. Stereospecificity of protein structure intrigues about the possibilities from a heterochiral alphabet. Our interest in heterochiral proteins amounted also to querying Nature in its logic for the use of the homochiral alphabet. Ranganathan Bharadwaj and Ranbir SinhaRoy, spending practically all vacations during their graduate pro-

gram at IIT Bombay in my laboratory, developed the modeling software CAPM. The bracelet, canoe, boat, and cup, the designs implemented by Soumendra Rana,[49−52] are examples of the capability of CAPM. Exploring stereochemistry as the tool to probe proteins in their folding principles, Vibin Ramakrishnan[18] produced the results that brought to our notice the Flory puzzle about the statistical coil properties of the polypeptide structure in relation to its stereochemical structure.[16] The concurrence of a plausible solution for the Flory puzzle with the electrostatic screening hypothesis of Avbelj[26] suggested the possibility that the peptides interacting in homochiral stereochemistry could be involved in the mediation of sequence control over protein conformation via screening effects in interpeptide electrostatics.[18,19] The implied criticality of homochiral structure to protein-folding principles is in line with the spirit of a recently proposed backbone-based hypothesis[6] and with the interpeptide interaction being implicated in a possible critical role.[4] The illumination of protein folding for a possible role of stereochemistry was but a serendipitous spin off in the protein design approach under exploration in my laboratory harnessing L- and D-(α)amino acid structures as the alphabet.

I am thankful to BRNS-DAE, CSIR, and DST for the continued funding of the research in my laboratory. Ranganathan Bharadwaj and Ranbir SinhaRoy deserve special thanks; this work would hardly have been possible without the assistance of CAPM. Anil Kumar helped me with figures and referencing of the manuscript.

BIOGRAPHICAL INFORMATION

Born and educated in Jammu and Kashmir, **Durani** obtained a Ph.D. (1977) at CDRI, Lucknow, and contributed in drug discovery research at the Institute before joining IIT, Bombay, in 1985. As Professor of Chemistry and Biotechnology, Durani headed the Biotechnology Centre during 1992 to 2002. His current research interests span protein structure, folding, and *de novo* design.

REFERENCES

1 Hill, R. B.; Raleigh, D. P.; Lombardi, A.; DeGrado, W. F. De Novo Design of Helical Bundles as Models for Understanding Protein Folding and Function. *Acc. Chem. Res.* **2000**, *33*, 745–754.

2 Kuhlman, B.; Dantas, G.; Ireton, G. C.; Varani, G.; Stoddard, B. L.; Baker, D. Design of a Novel Globular Protein Fold with Atomic-Level Accuracy. *Science* **2003**, *302*, 1364–1368.

3 Cheng, R. P. Beyond De Novo Protein Design - De Novo Design of Non-Natural Folded Oligomers. *Curr. Opin. Struct. Biol.* **2004**, *14*, 512–520.

4 Baldwin, R. L. Energetics of Protein Folding. *J. Mol. Biol.* **2007**, *371*, 283–301.

5 Dill, K. A.; Ozkan, S. B.; Weikl, T. R.; Chodera, J. D.; Voelz, V. A. The Protein Folding Problem: When Will It Be Solved. *Curr. Opin. Struct. Biol.* **2007**, *17*, 342–346.

6 Rose, G. D.; Fleming, P. J.; Banavar, J. R.; Maritan, A. A Backbone-Based Theory of Protein Folding. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 16623–16633.

7 Bonner, W. A. Chirality and Life. *Origins Life Evol. Biosphere* **1995**, *25*, 175–190.

8 Chothia, C. Principles That Determine the Structure of Proteins. *Annu. Rev. Biochem.* **1984**, *53*, 537–572.

9 Wiberg, K. B. The Interaction of Carbonyl Groups with Substituents. *Acc. Chem. Res.* **1999**, *32*, 922–929.

10 Ramachandran, G. N.; Ramakrishnan, C.; Sasisekharan, V. Stereochemistry of Polypeptide Chain Configurations. *J. Mol. Biol.* **1963**, *7*, 95–99.

11 Milton, R. C. L.; Milton, S. C. F.; Kent, S. B. H. Total Chemical Synthesis of a D-Enzyme: The Enantiomers of HIV-1 Protease Show Demonstration of Reciprocal Chiral Substrate Specificity. *Science* **1992**, *256*, 1445–1448.

12 Pauling, L.; Corey, R. B. Configuration of Polypeptide Chains. *Nature* **1951**, *168*, 550–551.

13 Pauling, L.; Corey, R. B. Configurations of Polypeptide Chains with Favored Orientations around Single Bonds: Two New Pleated Sheets. *Proc. Natl. Acad. Sci. U.S.A.* **1951**, *37*, 729–740.

14 Zhang, C.; DeLisi, C. Estimating the Number of Protein Folds. *J. Mol. Biol.* **1998**, *284*, 1301–1305.

15 Anfinsen, C. B. Principles That Govern the Folding of Protein Chains. *Science* **1973**, *181*, 223–230.

16 Flory, P. J. *Statistical Mechanics of Chain Molecules*; Interscience Publishers: New York, 1969.

17 Miller, W. G.; Brant, D. A.; Flory, P. J. Random Coil Configurations of Polypeptide Copolymers. *J. Mol. Biol.* **1967**, *23*, 67–80.

18 Ramakrishnan, V.; Ranbhor, R.; Kumar, A.; Durani, S. The Link Between Sequence and Conformation in Protein Structures Appears To Be Stereochemically Established. *J. Phys. Chem. B* **2006**, *110*, 9314–9323.

19 Ranbhor, R.; Ramakrishnan, V.; Kumar, A.; Durani, S. The Interplay of Sequence and Stereochemistry in Defining Conformation in Proteins and Polypeptides. *Biopolymers* **2006**, *83*, 537–545.

20 Shi, Z.; Chen, K.; Liu, Z.; Kallenbach, N. R. Conformation of the Backbone in Unfolded Proteins. *Chem. Rev.* **2006**, *106*, 1877–1897.

21 (a) Baldwin, R. L. A New Perspective on Unfolded Proteins. *Adv. Protein Chem.* **2002**, *62*, 361–367. (b) Shi, Z.; Olson, C. A.; Rose, G. D.; Baldwin, R. L.; Kallenbach, N. R. Polyproline II Structure in a Sequence of Seven Alanine Residues. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 9190–9195.

22 Tiffany, M. L.; Krimm, S. Circular Dichroism of Poly-L-Proline in an Unordered Conformation. *Biopolymers* **1968**, *6*, 1767–1770.

23 Makowska, J.; Rodziewicz-Motowidlo, S.; Baginska, K.; Vila, J. A.; Liwo, A.; Chmurzynski, L.; Scheraga, H. A. Polyproline II Conformation Is One of Many Local Conformational States and Is Not an Overall Conformation of Unfolded Peptides and Proteins. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 1744–1749.

24 Zagrovic, B.; Lipfert, J.; Sorin, E. J.; Millett, I. S.; van Gunsteren, W. F.; Doniach, S.; Pande, V. S. Unusual Compactness of a Polyproline Type II Structure. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 11698–11703.

25 Pal, D.; Chakrabarti, P. On Residues in the Disallowed Region of the Ramachandran Map. *Biopolymers* **2002**, *63*, 195–206.

26 Avbelj, F.; Moult, J. Role of Electrostatic Screening in Determining Protein Main Chain Conformational Preferences. *Biochemistry* **1995**, *34*, 755–764.

27 Avbelj, F.; Baldwin, R. L. Origin of the Neighboring Residue Effect on Peptide Backbone Conformation. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 10967–10972.

28 Rose, G. D.; Gierasch, L. M.; Smith, J. A. Turns in Peptides and Proteins. *Adv. Protein Chem.* **1985**, *37*, 1–109.

29 Wilmot, C. M.; Thornton, J. M. Analysis and Prediction of the Different Types of Beta-Turn in Proteins. *J. Mol. Biol.* **1988**, *203*, 221–232.

30 Haque, T. S.; Gellman, S. H. Insights on β-Hairpin Stability in Aqueous Solution from Peptides with Enforced Type I′ and Type II′ β-Turns. *J. Am. Chem. Soc.* **1997**, *119*, 2303–2304.

31 Hughes, R. M.; Waters, M. L. Model Systems for β-Hairpins and β-Sheets. *Curr. Opin. Struct. Biol.* **2006**, *16*, 514–524.

32 Hummel, G.; Reineke, U.; Reimer, U. Translating Peptides into Small Molecules. *Mol. BioSyst.* **2006**, *2*, 500–508.

33 Sibanda, B. L.; Thornton, J. M. Beta-Hairpin Families in Globular Proteins. *Nature* **1985**, *316*, 170–174.

34 Dhanasekaran, M.; Fabiola, F.; Pattabhi, V.; Durani, S. A Rationally Designed Turn-Helix Peptide. *J. Am. Chem. Soc.* **1999**, *121*, 5575–5576.

35 Karle, I. L.; Balaram, P. Structural Characteristics of α-Helical Peptide Molecules Containing Aib Residues. *Biochemistry* **1990**, *29*, 6747–6756.

36 Marshall, G. R.; Hodgkin, E. E.; Langs, D. A.; Smith, G. D.; Zabrocki, J.; Leplawy, M. T. Factors Governing Helical Preference of Peptides Containing Multiple α,α-dialkyl amino acids. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 487–491.

37 Watson, J. D.; Milner-White, E. J. A Novel Main-Chain Anion-Binding Site in Proteins: The Nest. A Particular Combination of φ,ψ Values in Successive Residues

Gives Rise to Anion-Binding Sites That Occur Commonly and Are Found Often at Functionally Important Regions. *J. Mol. Biol.* **2002**, *315*, 171–182.

38 Watson, J. D.; Milner-White, E. J. The Conformations of Polypeptide Chains Where the Main-Chain Parts of Successive Residues are Enantiomeric. Their Occurrence in Cation and Anion-Binding Regions of Proteins. *J. Mol. Biol.* **2002**, *315*, 183–191.

39 Doyle, D. A.; Morais, C. J.; Pfuetzner, R. A.; Kuo, A.; Gulbis, J. M.; Cohen, S. L.; Chait, B. T.; MacKinnon, R. The Structure of the Potassium Channel: Molecular Basis of K+ Conduction and Selectivity. *Science* **1998**, *280*, 69–77.

40 Urry, D. W. The Gramicidin A Transmembrane Channel: A Proposed pi(L,D) Helix. *Proc. Natl. Acad. Sci. U.S.A.* **1971**, *68*, 672–676.

41 Veatch, W. R.; Fossel, E. T.; Blout, E. R. The Conformation of Gramicidin A. *Biochemistry* **1974**, *13*, 5249–5256.

42 Wallace, B. A. Gramicidin, a Simple Ion Channel. *Curr. Top. Membr. Transp.* **1988**, *33*, 35–50.

43 Ghadiri, M. R.; Granja, J. R.; Milligan, R. A.; McRee, D. E.; Khazanovich, N. Self-Assembling Organic Nanotubes Base on a Cyclic Peptide Architecture. *Nature* **1993**, *366*, 324–327.

44 Struthers, M.; Ottesen, J. J.; Imperiali, B. Design and NMR Analyses of Compact, Independently Folded BBA Motifs. *Folding Des.* **1998**, *3*, 95–103.

45 Pinkerton, M.; Steinrauf, L. K.; Dawkins, P. The Molecular Structure and Some Transport Properties of Valinomycin. *Biochem. Biophys. Res. Commun.* **1969**, *35*, 512–518.

46 Jourdan, F.; Lazzaroni, S.; Mendez, B. L.; Lo, C. P.; de, J. M.; Amodeo, P.; Iacobellis, N. S.; Evidente, A.; Motta, A. A Left-Handed Alpha-Helix Containing Both L- and D-Amino Acids: The Solution Structure of the Antimicrobial Lipodepsipeptide Tolaasin. *Proteins: Struct., Funct. Gen.* **2003**, *52*, 534–543.

47 Chin, J. W.; Cropp, T. A.; Anderson, J. C.; Mukherji, M.; Zhang, Z.; Schultz, P. G. An Expanded Eukaryotic Genetic Code. *Science* **2003**, *301*, 964–967.

48 Fabiola, F.; Pattabhi, V.; Rawale, S.; Raju, E. B.; Durani, S. Configurationally Guided Peptide Conformational Motifs: Crystal Structure of a $^{L}\alpha^{D}\beta^{L}\beta^{D}\alpha^{D}\beta^{L}\alpha$ Type Hexapeptide Fold. *Chem. Commun.* **1997**, 1379–1380.

49 Rana, S.; Kundu, B.; Durani, S. Stereospecific Peptide Folds. A Rationally Designed Molecular Bracelet. *Chem. Commun.* **2004**, 2462–2463.

50 Rana, S.; Kundu, B.; Durani, S. A Small Peptide Stereochemically Customized as a Globular Fold with a Molecular Cleft. *Chem. Commun.* **2005**, 207–209.

51 Rana, S.; Kundu, B.; Durani, S. A Double Catgrip Mixed L and D Mini Protein Only 20 Residues Long. *Bioorg. Med. Chem.* **2007**, *15*, 3874–3882.

52 Rana, S.; Kundu, B.; Durani, S. A Mixed-Alpha,Beta Miniprotein Stereochemically Reprogrammed to High-Binding Affinity for Acetylcholine. *Biopolymers* **2007**, *87*, 231–243.

53 Joshi, S.; Rana, S.; Wangikar, P.; Durani, S. Computational Design of Proteins Stereochemically Optimized in Size, Stability, and Folding Speed. *Biopolymers* **2006**, *83*, 122–134.