Creating novel protein scripts beyond natural alphabets

Anil Kumar · Vibin Ramakrishnan

Received: 18 November 2010 / Accepted: 3 February 2011 / Published online: 1 March 2011
© Springer Science+Business Media B.V. 2011

Abstract Natural proteins are concatenated amino acids with definite handedness or chirality, with their spatial orientation being preferentially left handed or *L*-chiral. This paper discusses the biophysics of stereo-chemical perturbation to proteins using *D*-(α) amino acid and its utility as an additional design alphabet while scripting novel protein structures.

Keywords Protein design · Polypeptide stereochemistry · Protein structure energetics · Heterochiral proteins

Introduction

The molecular systems chosen by Nature for the expression of biologically encoded information have been in focus of intense scientific scrutiny for deciphering their operational logic. The chemical information encoded in a sequential arrangement of amino acids in protein structures is transduced into executable form during the process of folding (Anfinsen 1973; Anfinsen and Scheraga 1975). The deciphering of the rules and of the underlying physics has been an enterprise of tremendous scientific value (Chou and Fasman 1974; Cooper et al. 2010; Rohl et al. 2004). It is not only motivated by the need for accurate prediction algorithms for 3-dimensional shapes of protein from gene sequences, but

also for possible excursions with Nature's design algorithm beyond the constraints of *L*-(α) amino acid as preferred natural alphabets.

The interactions within protein backbone (main-chain) are integral to the sequential encoding of conformations in proteins. It has been clear from the time of Pauling's pioneering studies that the main-chain structure in polypeptides is having a defining role in the nature of protein secondary structure (Pauling and Corey 1950; Pauling et al. 1951). The constraints of planarity and polarity of peptide bonds, and hydrogen bonds between them are responsible for α -helix and β -sheet motifs. Similarly, local sterics, by limiting accessible conformational space, was shown by Ramachandran and Sasisekharan to also favor α -helix and β -sheet regions as the minima on residue level ϕ , ψ maps (Ramachandran et al. 1963; Ramachandran and Sasisekharan 1968). Thus the general principles responsible for secondary structure are well laid out, but the energetic trade off involved in sequence dependent selection of conformation either lacks a consensus about its real physical nature or has not been de-convoluted to all its contributing elements. According to one hypothesis, backbone (main-chain) conformational states are sensitive to the differential changes in side-chain entropy upon folding (Creamer and Rose 1992; Leach et al. 1966). Another hypothesis attributes the bias in backbone conformation to the localized interactions between main-chain and side-chain elements (Presta and Rose 1988; Richardson and Richardson 1988). Yet another line of enquiry imputes the electrostatics of main-chain response to solvent and side-chain structure, as the basis for fold selection (Avbelj and Baldwin 2002; Avbelj and Moult 1995). Clearly the main-chain roles are critical, but are made inscrutable by the immutability of backbone structure in proteins, because chirality of amino acids in protein structure is always

A. Kumar
Department of Chemistry, University of Toronto, 80 St. George Street, Toronto, ON M5S 3H6, Canada
e-mail: chemanil@gmail.com

V. Ramakrishnan (✉)
Rajiv Gandhi Centre for Biotechnology, Thiruvananthapuram, Kerala 695014, India
e-mail: vibin@rgcb.res.in

predominantly *L*-configurational. The origin of this apparent natural stereochemical filtration in protein biogenesis remains unclear; however, its consequences on protein structure and folding can be investigated. The folding of protein to its native state (operative form) is a cumulative effect of mutually cooperative inter-atomic interactions. Unfolding or misfolding is assumably a result of the absence of this cooperativity or its disruption. Stereochemistry may be a relevant perturbant of polypeptide structure to achieve such an objective. This review focuses on the possible impact of stereo-chemical mutation on the folding determinants of a polypeptide molecule and explore possible ways to design novel architectures by tune up of constituent energies using *D*-(α) amino acids.

Chirality defining chain stiffness of a polypeptide

Chain stiffness is a term used to describe the degree of angular correlation between neighbouring bonds. In general, the effects of bond correlations can be described by $\langle r^2 \rangle = C_N N b^2$; where C_N is called characteristic ratio (Flory 1969). Characteristic ratio of a random flight chain is 1. Chains having restricted angular diversity will have higher values of C_N . The value of C_N in polypeptide chains is in the range of 9.0–9.5, and for poly Gly it is as low as 2.0. The relatively high value of C_N in polypeptide chains shows that it is more towards ‘rod’ like than ‘random flight’ like. The reason can be directly inferred from the angular diversity of polypeptide chain. The overall conformation in a polypeptide chain is determined by two dihedral angles associated with main-chain, ϕ and ψ obtained by rotations along N-C α and C α -C bond in a polypeptide chain. The dihedral angles distribution of polypeptide chain can be described with the help of Ramachandran map. Due to excluded volume effect the accessible conformational space available for a normal polypeptide chain in either *L*- or *D*-configuration of α -amino acid is a maximum of 21% of the total conformational space available (Fig. 1).

Therefore the angular distribution along a polypeptide chain is very much restricted, resulting in high amount of stiffness or rod like behaviour. Achiral α -amino acid glycine can sample around 57% of the total Ramachandran map, and the result is directly evident in its characteristic ratio value (2.0). The situation is even worst in structures like collagen with a sequence of (Gly-Pro-Pro) $_n$, where the conformational space accessible for Pro residue is severely restricted around -60° , $+140^\circ$ due to pyrrolidine ring as side-chain of Pro. Consequently collagen triple helix is very much rod like. Therefore it is clear that C_N depends indirectly on the R group of amino acid, and it would be of interest to know, how the stereo-isomeric R group influences the chain stiffness

Fig. 1 **a** *L*- and *D*- enantiomers of α -amino acids where R refers to the side chains. **b** Ramachandran maps of ϕ and ψ space for enantiomeric *L*- (gray) and *D*- (magenta) amino acids. The labeled light shaded areas are freely accessible, while the outer contours enclose the partially accessible regions. All other areas are energetically forbidden

and overall fold of a protein chain. A pioneering analysis along this line is a work done by P. J. Flory (1969), where he calculated the limiting values of the characteristic ratio for random copolymers of *L*-Ala, Gly, and *L*- and *D*-Ala. The ratio measures the average physical dimension of a polymer as compared to the value calculated statistically assuming free rotation about all its backbone (main-chain) bonds (Flory 1969). The characteristic ratio is about 2 for typical polymers, as a consequence of steric effects that constrain the bond rotations compared with the hypothetical free-rotation abstraction. The characteristic ratio was found to be a “puzzling” value of ~ 9 for poly-glutamate and poly-lysine in water, for which Brant and Flory (1965a, b) proposed “dipolar repulsion” between amide as an explanation, leading to preference for locally “extended” rather than locally “folded” conformers in the polypeptide backbone.

The “chain stiffness” has also been measured in “denatured” proteins (Kohn et al. 2004; Tanford 1968). The values are large but not nearly as large as measurements of Brant and Flory (1965a, b) for “unfolded” polypeptides. It has been proposed that the difference could be an effect of Glycyl residues in proteins, which were shown by Miller et al. (1967) to indeed diminish the characteristic ratio in a glycyl-glutamyl copolymer. Flory made a calculation of the characteristic ratios in random *L*-Ala, *D*-Ala copolymers or different mole ratios and found a reduction in the characteristic ratio to a minimum value of 2 for a 50% mole ratio of the enantiomers, while the ‘poly *L*’ isomer of both *L*- and *D*-chiral alanine were with characteristic ratio of ~ 9 . While a twist was recently added, that the values need not necessarily imply random coil statistics (Fitzkee and Rose 2004), nearly four and a half scale reduction in the

characteristic ratio was also confirmed in a copolymer *D*-Glutamyl and *L*-Glutamyl (Miller et al. 1967). The significant reduction of “stiffness” in random *L*, *D* and alternating *L*, *D* polypeptides could be an apparent consequence of the “harmonious” nature of short and intermediate range electrostatics (Kumar et al. 2009; Ramakrishnan et al. 2006; Ranbhor et al. 2006), both favouring locally folded states of the polymer. This observation offers a new possibility to the designer; creating more stable secondary structures that can resist conformational perturbation.

Chirality defining secondary structure of a polypeptide

Fundamentally, proteins are linear polymers of planar N-methylacetamide (NMA) units carrying amino acid side-chain R groups as substituent on tetrahedral α -carbons. When R group is other than H, it is a stereocenter that can exist in *L*- or *D*-configurational forms. Polymer chains with all stereocenters of exclusively *L*-(*S*) or *D*-(*R*) configurational type are called as ‘*isotactic*’ or ‘*homotactic*’ or ‘*homochiral*’ polymers (Billmeyer 1984). If the stereocentres are of combined *L*- and *D*-stereochemical nature the polymer structure may be of either stereoregular or stereoirregular in nature. If the stereocentres in a stereoregular polymer are of alternately enantiomeric nature, the polymer is called as ‘*syndiotactic*’, and if the stereocentres are randomly distributed, it is called ‘*heterotactic*’ or ‘*atactic*’ (Fig. 2a–c). Made of asymmetric amino acids of

always *L*-chiral stereochemical structure, proteins are stereoregular ‘*isotactic*’ polymers of ‘*homochiral*’ structure.

The homochiral nature of backbone structure is fundamental to folding because local rotameric preferences of amide planes are stereochemically controlled. The bias in ϕ , ψ is dependent on residue configuration while at motif level, it is dependent on nature of chain tacticity (Ramachandran and Sasisekharan 1968). The pivotal role that stereochemistry plays in biasing both residue and motif level conformations in polypeptides was anticipated by, respectively, Ramachandran and Pauling based on modelling considerations. The seminal conclusions reached by these pioneers are strongly affirmed in the database of solved protein structures (Pauling and Corey 1950; Ramachandran et al. 1963). A fundamentally important mode of inter-segmental association in Pauling’s description of α -helix and β -sheet motifs, is that of hydrogen bonding. α -helix and β -sheets take the major share of protein secondary structure found in nature (Pauling and Corey 1950). Other secondary structure patterns are much less in abundance. An unusual one, and of stereospecific nature, occurs in alternating *L*, *D* polypeptides, of alternating *L*- and *D*-chiral stereochemical structure, with gramicidin- β -helix as a major stereochemically feasible secondary structure (Szabo and Urry 1979; Urry 1971; Wallace 1998, 2000).

The most ubiquitous definition of secondary structure is in the topological patterns of inter-peptide hydrogen bonds, i.e., the sequential separation between the acceptor (*n*) and donor (*n* $\pm$ *i*) residues (*n* to *n* $\pm$ *i* or succinctly *n* $\pm$ *i*) of a

Fig. 2 Schematic representation of polypeptide chains with isotactic (a), syndiotactic (b) and heterotactic (c) structures. d The hydrogen bond registry between inter peptide bonds of polypeptide structures

hydrogen bond (Fig. 2d). The hydrogen bond topologies are stereospecific (Eisenberg 2003). The topologies are predominantly $n + i$ type in polypeptides of isotactic stereochemistry, but may be both $n + i$ and $n - i$ type in syndiotactic or heterotactic polypeptides.

The short ranged $n + 2$ type hydrogen bonds characterize the β turns, which may be of L - or D -stereochemical type. The $n + 3$ type hydrogen bonds characterize the β turns, which may be either homochiral or heterochiral depending on the stereochemistry of its central two residues (Rose et al. 1985). Trains of consecutive $n + 3$ (3_{10} helix), $n + 4$ (α helix) or $n + 5$ (helix) type hydrogen bonds are possible when all residues in a chain segment are of same stereochemical type and in α_R (right handed helix) or α_L (left handed helix) regions of Ramachandran diagram corresponds to A_L and A_D regions of ϕ, ψ space as noted in Fig. 3 (Donohue 1953).

The $n + 5$ type isotactic helix is the postulated (π -helix) that has actually only recently been observed experimentally (Weaver 2000). The syndiotactic helix corresponds to B_L/B_D regions of ϕ, ψ space (Fig. 3), the region of extended conformation for poly L -chains. Unlike poly L -chiral helices, the ϕ_L, ψ_L and ϕ_D, ψ_D values differ slightly for a syndiotactic helix, and the difference determines the screw sense of the helix. The helix is with a registry of alternating $n + 7$ and $n - 5$ type hydrogen bonds (Fig. 2d) and occurs in Gramicidin A, a microbial polypeptide of syndiotactic stereochemistry (Szabo and Urry 1979; Urry 1971; Wallace 1998, 2000). Extended conformation of isotactic stereochemistry, in mutually hydrogen-bonded parallel or antiparallel arrangement, characterize the β sheet motifs of proteins, in which the hydrogen bond topologies are on average $>n + 6$. The extended conformation of syndiotactic stereochemistry occurs in A_L/A_D regions of ϕ, ψ space, which is otherwise the region of isotactic helix.

Fig. 3 ϕ, ψ distribution plots due to the R group effects in L - (gray) and D - (magenta) regions of Ramchandran plot. The yellow (A_L/A_D), blue (B_L/B_D) and pink (C_L/C_D) regions of the ϕ, ψ map are used for the classification of residue conformations in this study. The sterically allowed Ramachandran regions (α_L, α_R and β) are outlined and colored as in Fig. 1

Chirality defining energetic trade-off of secondary structure

To understand the energetic influences in defining stereochemically engineered fold, we did systematic grid search to discover low energy conformations for the model system as isotactic and syndiotactic octa-alanine polypeptides (Ac-Ala₈-NHMe) in folded state. Break up of their total energy has also been computed to determine the roots of elective influences in conformational selection. Grid Search Method is a simple method to explore conformational space by systematic changes in conformation (Leach 2001). In this method, initially all rotatable bonds that can effect conformational change are identified. Each bond is then systematically rotated through 360° with fixed increments. Each junction point in the grid corresponds to one single structure corresponding to that point in the conformational space. Bond lengths and bond angles may be kept fixed in order to minimize the computational load. The search stops when all possible combinations of torsion angles have been generated. Potential energy of conformational variants corresponding to each junction points was calculated with the help of GROMOS forcefield (van Gunsteren et al. 1996). Unlike isotactic folds, the structures with syndiotactic stereochemistry are with torsional angles of alternate signs due to successive sampling in L - and D -chiral regions of Ramachandran map.

Exhaustive search of the conformational space in both isotactic and syndiotactic polypeptide structures was undertaken with Grid search method. The octa-alanine model Ac-Ala₈-NHMe of both isotactic and syndiotactic stereochemical structure was modelled in specific ϕ, ψ combinations, and the potential energy was calculated for each conformer. The energy calculations were done with GROMOS forcefield (van Gunsteren et al. 1996), while the polypeptide modelling was performed with public domain Ribosome program (Srinivasan). Systematic variations in ϕ, ψ was undertaken at 10° intervals through 360° . A total of 1,369 grid points were analysed in the $360^\circ \times 360^\circ \phi, \psi$ space. The energy landscape thus generated was compared with that of the alanine-dipeptide model Ac-Ala-NHMe.

For isotactic structures, the low energy basins are principally located in three regions, the right handed α helix region, extended β region and polyproline II region; and this was expected. (Table 1 and Fig. 1). Unlike isotactic structures, syndiotactic structures can never attain folded conformation by propagating with one single dihedral angle pairs. The folding chain will run over itself resulting in very high values of repulsive LJ energy. Hence, syndiotactic folds in nature like gramicidin helices are found to be staggered in ϕ_L, ψ_L and ϕ_D, ψ_D values to avoid these steric clashes.

Folded structures in syndiotactic stereochemistry are sampled in the region between -70 – -180 (ϕ) and

Table 1 (A) Potential Energy (kJ/mol) of the modelled folds in isotactic and syndiotactic stereochemistry and (B) comparison with standard secondary structure elements found in structures available in protein data bank

	Isotactic helix	Syndiotactic helix				
No	ϕ, ψ	P.E	L	D	P.E	Handedness
(A)						
1	-70, -40	-59.8161	-120, 140	120, -110	-70.2572	Right
2	-60, -40	-57.5839	-120, 150	110, -110	-68.4012	Right
3	-70, -30	-52.7579	-110, 150	110, -120	-67.3828	Right
4	-80, -30	-45.8933	-110, 120	110, -150	-67.0053	Left
5	-60, -50	-38.6062	-120, 110	120, -140	-66.937	Left

Secondary structure 2 fold ribbon α -helix 3_{10} helix (1) 3_{10} helix (2) π -helix Isotactic extend Syndiotactic extended $\beta - 6.3$ $\beta - 4.4$ Gramicidin

(B)

P.E	98.3	-59.7	-22.6	30.5	1,265.9	15.6	43.3	3.6	-42.2	-38.2
-----	------	-------	-------	------	---------	------	------	-----	-------	-------

The definition of different secondary structures used for calculation: 1) 2 fold ribbon ($\phi \sim -75, \psi \sim 70$); α -helix ($\phi \sim -57, \psi \sim -47$); 3_{10} helix (1) ($\phi \sim -60, \psi \sim -30$); 3_{10} helix (2) ($\phi \sim -74, \psi \sim -4$); π -helix ($\phi \sim -57, \psi \sim -70$); Isotactic extended ($\phi \sim -120, \psi \sim 120$); Syndiotactic extended ($\phi^{L/D} \sim -60/60, \psi^{L/D} \sim -50/50$); $\beta - 6.3$ ($\phi^{L/D} \sim -106/140, \psi^{L/D} \sim 122/-85$); $\beta - 4.4$ ($\phi^{L/D} \sim -80/125, \psi^{L/D} \sim 100/-85$); Gramicidin ($\phi^{L/D} \sim -120/120, \psi^{L/D} \sim 120/-120$)

+70+180 (β) for L - and +70+180 (ϕ) and -70-180 (ψ) for D -chiral amino acids, respectively. This will correspond to 144 grid points each in left and right quadrants of Ramachandran map. An all-against-all exhaustive enumeration was performed between these regions for residues of enantiomeric structure, giving a total of 20,736 (144×144) combinations of ϕ_L, ψ_L and ϕ_D, ψ_D values for syndiotactic octa-alanine model. The torsional angles of minimum energy conformations sampled in the staggered helical structures of syndiotactic stereochemistry and a comparative energetic analysis with standard secondary structure patterns found in natural proteins are shown in Table 1. The minimum energy conformations of syndiotactic helices are found to be comparatively more stable than α -helix. The helical folds in syndiotactic isomer of the octa-alanine model can be of four different stereochemical types, right handed -SRSR-, left handed -RSRS-, right handed -RSRS- and left handed -SRSR-. The four isomers comprise of two enantiomeric pairs of mutual diastereomeric nature (Fig. 4). Notably, the most stable one of the four matches is the right handed SRSR-stereochemical β -helix identical of gramicidin A structure.

The results of this investigation establish that the stereochemically governed tacticity of polypeptides not only defines the morphological nature of chain motifs, but also its potential energy on account of steric and electrostatic interaction. The isotactic motifs with the residues of β conformational type tend to propagate linearly while those with the residues of α conformational type tend to propagate helically. Practically every mode of chain propagation is therefore sterically accessible in polypeptides of isotactic stereochemistry, in accordance with Flory's isolated pair hypothesis (Flory 1969). In syndiotactic chains, most of β

Fig. 4 Low energy staggered syndiotactic helices of polyalanine (Ac-Alag-NHMe) simulated by alternatively sampling the β -region of L - and D - of Ramchandran map. The stereochemistry of alanine and the respective screw sense of the structures are indicated. The diagonally positioned structures are enantiomeric to each other

conformational region is sterically excluded due to steric clashes. The clashes are relievable by staggering the ψ_L and ψ_D values, resulting in helical mode of chain propagation. The steric accessibilities of chain propagation modes are thus stereochemically determined. Accordingly, the entropic costs in both chain propagation and folding could also be stereochemically determined. Prima facie, such costs could be much higher in isotactic helices, than in syndiotactic helices, because practically every ϕ, ψ combination and therefore practically every mode of chain propagation is sterically accessible.

Stereochemistry also defines the mutual orientations of amide-dipoles, and therefore the nature of electrostatic interactions.

In isotactic chains, adjacent dipoles are in electrostatically favorable antiparallel arrangement in extended conformation (Fig. 5) but in electrostatically unfavorable parallel arrangement in helical conformation. Exactly the opposite is true in syndiotactic chains (Ramakrishnan et al. 2006; Ranbhor et al. 2006). In this case the mutual arrangement of adjacent dipoles is electrostatically unfavorable in extended conformation but electrostatically favorable in helical conformation (Fig. 5). While designing novel protein molecules with both *L* and *D* amino acids, this observation is crucial. The effect may contribute to the relative energies of different chain motifs, and could be deterministic in folding kinetics.

The three dimensional structure adopted by a polypeptide is dependant on its environment, the temperature, the pressure, and the solvent or molecular surroundings. Under any given set of conditions the structure of a polypeptide will not be unique, but instead there will be an equilibrium distribution between different structures. If a particular conformation predominates under a given set of conditions the polypeptide is commonly referred to as being “folded” (Daura et al. 2001, 1999). Often, the folded conformation of a polypeptide is only marginally more stable than the lowest free energy “unfolded” conformation. Hence the macroscopic observables of small polypeptides are weighted with both the folded and unfolded states. Durani and coworkers examined the effects of dielectric and stereochemical perturbation by simulating an equilibrium ensemble of the octa-alanine model Ac-Ala₈-NHMe, by Molecular Dynamics techniques (Ramakrishnan et al. 2006; Ranbhor et al. 2006). There was appreciable ordering in thermodynamic ensemble of alternating *L,D* variant also

Fig. 5 Orientation of main-chain peptide dipoles in the extended (β -region) and folded (α -region) structures of isotactic and syndiotactic stereochemistry. Direction of peptide dipoles can be inferred from the orientation of oxygen (red) and polar hydrogen (white). Inter-peptide hydrogen bonding and coulomb interactions is in harmony or conflict dependent upon stereochemistry of polypeptide

with its major contributing microstates displaying β -helical conformation. The poly *L*-chiral model was found to be highly solvent sensitive, almost fully folded in water and fully unfolded in methanol, while the alternating *L, D*-chiral model is conformationally non-responsive to its dielectric stimulus, fully folded in both solvents. The observation was attributed to the conflicting nature of short- and intermediate range electrostatics, promoted by homochiral nature of polypeptide stereochemical structure, which enables the protein molecule to act as a chemosensory apparatus responsive to external stimuli (Kumar et al. 2009; Ramakrishnan et al. 2006; Ranbhor et al. 2006). Relative insensitivity to stimulus may not be ideal for a designed protein (peptide) molecule, because protein molecules are required to be sensitive to external stimulus. But for artificial biomaterials, this extra stability may be helpful in withholding structural integrity while subjected to fluctuations in temperature and pressure conditions to a certain extent.

De novo design of hetero-chiral proteins

Programmed heteropolymer folding is a potentially powerful approach for molecular design, and in protein folding there are useful tips for practical advancement of such an objective. Much has been learnt about protein folding (Baldwin 2007; Rose et al. 2006) and much accomplished in its implementation in de novo design (Allen and Mayo 2010; Baker 2010; Korendovych et al. 2010). The time is therefore ripe to explore excursions with Nature’s design algorithm beyond the realm of biological alphabets. Although seemingly limitless in functional versatility, natural proteins display molecular architectural plans that are highly conservative. The residue level choices limited to α and β , while the choice of structural scaffolds is restricted to α -helix and β -strand motifs. With such a limited range of building blocks, protein tertiary structure is restricted to only $\cong 10^3$ topologically possibilities (Chothia 1992; Zhang and DeLisi 1998). The feature of the biological alphabet with fundamental bearing on these numbers is its stereochemical composition. With nineteen residues of *L*-chiral structure, proteins are polymers of stereochemically frozen structure. It is of interest to examine monomer stereochemistry as an additional, more fundamental, sequence variable for de novo design.

There are adequate literature precedents that testify the value of sequence stereochemistry as a tool for polypeptide conformational design. The type II and II’ β -turns, among the most common β -turns of protein structure (Sibanda et al. 1989; Sibanda and Thornton 1985), are formally of heterochiral nature because their central two residues are in a conformation that is stereochemically favored in a locally

enantiomeric *LD* or *DL* chiral dipeptide sequence. While in proteins the position corresponding to the *D*-chiral residue is typically occupied by the achiral Gly, the type II and II' β -turns are among the few protein conformational motifs that can be made sequentially more resolute by the simple stereochemical logic of actually recruiting an *LD* or *DL* chiral dipeptide as the design aid. There has been wide spread use of this simple design strategy as a method for rational design of β -hairpin folds (Dhanasekaran et al. 1999; Mohanraja et al. 2003). Imperiali and coworkers used this device beneficially for rational design of a small and synthetic protein called BBA (Struthers et al. 1996) (Fig. 6). Milner-White and coworkers reported the occurrence in proteins of small polypeptides stretches of alternately enantiomeric conformation, describing so called “anion-nests” and “cation-grips”, due to the characteristic semicircular disposition of hydrogen bond donor or acceptor groups from either the main-chain alone or both main-chain and side-chains, that participate in recognition and binding of cationic or anionic hosts (Watson and Milner-White 2002b, a). Exploiting monomer chirality for conformational design, Bobde et al. (1993); (Fabiola et al. 1997) have reported two different heterochiral helix nucleating signatures while Mohanraja et al. (2003) reported an attempt to recruit these folding signatures for mechanism based design of a minimal helix-bundle protein. Even contrary to our presumption about exclusive natural selection of *L*-configurational α -amino acids as the building blocks of protein, there has been report that *D*-(α)-amino acids are also important content of many naturally occurring proteins in plants (Robinson 1976), microbes (Brueckner et al. 1993), peptidoglycans in bacterial cell walls (Martinez del Pozo et al. 1989; Billot-Klein et al. 1997) and peptide antibiotics (Jack and Jung 1998).

Gramicidin-A, an antimicrobial peptide of fungal origin is a 14-mer alternating *L,D* chiral peptide of syndiotactic stereochemistry (Fig. 6). It adopts a range of unusual conformational folds of β -helical and double β -helical nature of both parallel and antiparallel variety, which contribute in its function as a membrane active peptide with ion channeling activity, with the main-chain amides performing the dual role of both conformation specification and ion conductance (Urry 1971; Wallace 1998, 2000). In close correspondence with this example, Ghadiri and coworkers developed a series of alternating *L, D* chiral cyclic peptides as self assembling nanotubes cum channels due to their β -sheet type intermolecular hydrogen bonding (Ghadiri et al. 1994).

Compared with these examples of small heterochiral motifs or rationally designed heterochiral peptides, a heterotactic protein of stereochemically defined architectural plan supported by an appropriate chemical sequence and structure, may be a far more ambitious research objective. For an isotactic protein there are only two stereoisomers

Fig. 6 Schematic structures of heterochiral folds: Structures of BBA, Gramicidin, bracelet, boat, canoe, pi-cup and Zn-finger hydrolase where *D*-(α)-amino acids are shown as dark spheres (magenta)

possible, poly-*L* and poly-*D*. With residue stereochemistry as a sequence variable, there are 2^{100} stereoisomers, and therefore $2^{100}/2$ enantiomeric pairs possible for a 100 residue heterotactic protein. Although an astronomical number, this is much smaller than 20^{100} , being the size of chemical sequence space corresponding to a 100 residue protein. Due to the lifting of residue level degeneracy of conformational choices, a huge multiplication of conformational possibilities can be expected for a “heterotactic” protein. The morphological range and the architectural plans for heterotactic proteins remain to be defined, but the conceptual foundations for diversification and downsizing structure seem clear. Ghadiri’s nanotube’s and gramicidin-A perhaps best exemplify the prospect (Ghadiri et al. 1994; Urry 1971; Wallace 2000). Likewise several reported hairpin peptides and mini-proteins directed or stabilized by one or more *D* chiral residues are examples of downsizing achieved by de novo design (Dhanasekaran et al. 1999; Durani 2008; Struthers et al. 1996). Possibly the first known heterochiral fold of complex sequence chirality was a hexapeptide reported by Fabiola et al. (1997). There have been more recent examples of β -hairpin peptides that were

stereochemically modified into bracelet (Rana et al. 2004) or boat (Rana et al. 2005) type architectural morphologies and cup-shaped molecules, designing the canoe as an alkali receptor (Rana et al. 2007a), pi-cup as an acetyl choline receptor (Rana et al. 2007b) and $\alpha\beta\beta$ fold as hydrolase catalyst (Patel et al. 2010) (Fig. 6). Developing computational (Joshi et al. 2006; Nanda and DeGrado 2006; Ramakrishnan et al. 2005) and experimental methodologies (Marahiel and Essen 2009; Golovine et al. 2004; Marahiel 2009) for de novo protein design with stereochemistry as the sequence level variable would hence be a profitable and much needed line of enquiry.

Conclusion

The disproportionate or rather complete excess of *L*-(α)-amino acids in naturally occurring proteins is argued to be preordained and eventually turned out to be one of the most powerful argument of creationists who believe that species on earth are not formed by evolution but by ‘intelligent design’. *How, why and when* the complete separation of stereo-isomers in living tissue remains an enigma and naturally this intellectual thunderbolt figured in the top 125 unanswered questions identified by science magazine in its 125th anniversary edition. In this context, it would be of interest to think of proteins beyond their homo-chiral structure, to stereo chemically diversified and customized hetero-chiral folds. Introduction of chirality as an additional sequence variable could greatly diversify the conformational template in polypeptides for de novo design of proteins. Energetically, it offers an opportunity to design artificial protein molecules optimised with all constituent energetic factors complimenting each other. This vast and potentially unexplored terrain in the protein universe is likely to be the face and future of synthetic biology in the years to come.

Acknowledgments Authors thank Prof. Susheel Durani for the core concept of this manuscript. VR thanks IYBA program of Dept of Biotechnology, Govt. of India for funding. The authors declare that they have no conflict of interest.

References

- Allen BD, Mayo SL (2010) An efficient algorithm for multistate protein design based on FASTER. *J Comput Chem* 31(5): 904–916
- Anfinsen CB (1973) Principles that govern the folding of protein chains. *Science* 181(96):223–230
- Anfinsen CB, Scheraga HA (1975) Experimental and theoretical aspects of protein folding. *Adv Protein Chem* 29:205–300
- Avbelj F, Baldwin RL (2002) Role of backbone solvation in determining thermodynamic beta propensities of the amino acids. *Proc Natl Acad Sci USA* 99(3):1309–1313
- Avbelj F, Moulton J (1995) Role of electrostatic screening in determining protein main chain conformational preferences. *Biochemistry* 34(3):755–764
- Baker D (2010) An exciting but challenging road ahead for computational enzyme design. *Protein Sci* 19(10):1817–1819
- Baldwin RL (2007) Energetics of protein folding. *J Mol Biol* 371(2):283–301
- Billmeyer FW (1984) Textbook of polymer science, 3rd edn. Wiley Interscience, New York
- Billot-Klein D, Legrand R, Schoot B, Van Heijenoort J, Gutmann L (1997) Peptidoglycan structure of *Lactobacillus casei*, a species highly resistant to glycopeptide antibiotics. *J Bacteriol* 179(19):6208–6212
- Bobde V, Beri S, Durani S (1993) Stable type II' reverse turn-310 helix conformation of Boc-D-Glu-Ala-Gly-Lys-Ala-Leu-OMe in apolar solvents. *Tetrahedron* 49(24):5397–5406
- Brant DA, Flory PJ (1965a) The configuration of random polypeptide chains. i. Experimental results. *J Am Chem Soc* 87(13):2788–2791. doi:10.1021/ja01091a002
- Brant DA, Flory PJ (1965b) The configuration of random polypeptide chains. II. Theory. *J Am Chem Soc* 87(13):2791–2800. doi:10.1021/ja01091a003
- Brueckner H, Becker D, Luepke M (1993) Chirality of amino acids of microorganisms used in food biotechnology. *Chirality* 5(5): 385–392
- Chothia C (1992) Proteins. One thousand families for the molecular biologist. *Nature* 357(6379):543–544
- Chou PY, Fasman GD (1974) Prediction of protein conformation. *Biochemistry* 13(2):222–245
- Cooper S, Khatib F, Treuille A, Barbero J, Lee J, Beenen M, Leaver-Fay A, Baker D, Popovic Z, Players F (2010) Predicting protein structures with a multiplayer online game. *Nature* 466(7307):756–760
- Creamer TP, Rose GD (1992) Side-chain entropy opposes alpha-helix formation but rationalizes experimentally determined helix-forming propensities. *Proc Natl Acad Sci USA* 89(13):5937–5941
- Daura X, van Gunsteren WF, Mark AE (1999) Folding–unfolding thermodynamics of a β -heptapeptide from equilibrium simulations. *Proteins: structure, function, and bioinformatics* 34:269–280. doi:10.1002/(sici)1097-0134(19990215)34:3<269:aid-prot1>3.0.co;2-3
- Daura X, Gademann K, Schafer H, Jaun B, Seebach D, van Gunsteren WF (2001) The beta-peptide hairpin in solution: conformational study of a beta-hexapeptide in methanol by NMR spectroscopy and MD simulation. *J Am Chem Soc* 123(10):2393–2404
- Dhanasekaran M, Fabiola F, Pattabhi V, Durani S (1999) A rationally designed turn-helix peptide. *J Am Chem Soc* 121(23): 5575–5576. doi:10.1021/ja984237v
- Donohue J (1953) hydrogen bonded helical configurations of the polypeptide chain. *Proc Natl Acad Sci USA* 39(6):470–478
- Durani S (2008) Protein design with L- and D-alpha-amino acid structures as the alphabet. *Acc Chem Res* 41(10):1301–1308
- Eisenberg D (2003) The discovery of the alpha-helix and beta-sheet, the principal structural features of proteins. *Proc Natl Acad Sci USA* 100(20):11207–11210
- Fabiola F, Pattabhi V, Rawale S, Raju EB, Durani S (1997) Configurationally guided peptide conformational motifs: crystal structure of a l[small alpha]d[small beta]l[small beta]d[small alpha] d[small beta]l[small alpha] type hexapeptide fold. *Chem Commun* 15:1379–1380
- Fitzkee NC, Rose GD (2004) Reassessing random-coil statistics in unfolded proteins. *Proc Natl Acad Sci USA* 101(34):12497–12502
- Flory PJ (1969) Statistical Mechanics of Chain Molecules. Interscience, New York
- Ghadiri MR, Granja JR, Buehler LK (1994) Artificial transmembrane ion channels from self-assembling peptide nanotubes. *Nature* 369(6478):301–304
- Golovine S, Hecht SM, Dedkova L (2004) Methods of protein synthesis incorporating D-amino acids using modified ribosomes

- with 23S rRNA gene mutations. WO Patent 2003-US33108; 2004035757
- Jack RW, Jung G (1998) Natural peptides with antimicrobial activity. *Chimia* 52(1–2):48–55
- Joshi S, Rana S, Wangikar P, Durani S (2006) Computational design of proteins stereochemically optimized in size, stability, and folding speed. *Biopolymers* 83(2):122–134
- Kohn JE, Millett IS, Jacob J, Zagrovic B, Dillon TM, Cingel N, Dothager RS, Seifert S, Thiagarajan P, Sosnick TR, Hasan MZ, Pande VS, Ruczinski I, Doniach S, Plaxco KW (2004) Random-coil behavior and the dimensions of chemically unfolded proteins. *Proc Natl Acad Sci USA* 101(34):12491–12496
- Korendovych IV, Senes A, Kim YH, Lear JD, Fry HC, Therien MJ, Blasie JK, Walker FA, Degrado WF (2010) De novo design and molecular assembly of a transmembrane diporphyrin-binding protein complex. *J Am Chem Soc* 132(44):15516–15518
- Kumar A, Ramakrishnan V, Ranbhor R, Patel K, Durani S (2009) Homochiral stereochemistry: the missing link of structure to energetics in protein folding. *J Phys Chem B* 113(51):16435–16442
- Leach AR (2001) Molecular modelling: principles and applications, 2nd edn. Prentice Hall, England
- Leach SJ, Nemethy G, Scheraga HA (1966) Computation of the sterically allowed conformations of peptides. *Biopolymers* 4(4):369–407
- Marahiel MA (2009) Working outside the protein-synthesis rules: insights into non-ribosomal peptide synthesis. *J Pept Sci* 15(12):799–807
- Marahiel MA, Essen LO (2009) Chapter 13. Nonribosomal peptide synthetases mechanistic and structural aspects of essential domains. *Methods Enzymol* 458:337–351
- Martinez del Pozo A, Pospischil MA, Ueno H, Manning JM, Tanizawa K, Nishimura K, Soda K, Ringe D, Stoddard B, Petsko GA (1989) Effects of D-serine on bacterial D-amino acid transaminase: accumulation of an intermediate and inactivation of the enzyme. *Biochemistry* 28(22):8798–8803
- Miller WG, Brant DA, Flory PJ (1967) Random coil configurations of polypeptide copolymers. *J Mol Biol* 23(1):67–80
- Mohanraja K, Dhanasekaran M, Kundu B, Durani S (2003) Mechanism-based protein design: attempted “nucleation-condensation” approach to a possible minimal helix-bundle protein. *Biopolymers* 70(3):355–363
- Nanda V, DeGrado WF (2006) Computational design of heterochiral peptides against a helical target. *J Am Chem Soc* 128(3):809–816
- Patel K, Srivastava KR, Durani S (2010) Zinc-finger hydrolase: Computational selection of a linker and a sequence towards metal activation with a synthetic alphabeta protein. *Bioorg Med Chem* 2010:27
- Pauling L, Corey RB (1950) Two hydrogen-bonded spiral configurations of the polypeptide chain. *J Am Chem Soc* 72(11):5349. doi:10.1021/ja01167a545
- Pauling L, Corey RB, Branson HR (1951) The structure of proteins; two hydrogen-bonded helical configurations of the polypeptide chain. *Proc Natl Acad Sci USA* 37(4):205–211
- Presta LG, Rose GD (1988) Helix signals in proteins. *Science* 240(4859):1632–1641
- Ramachandran GN, Sasisekharan V (1968) Conformation of polypeptides and proteins. *Adv Protein Chem* 23:283–438
- Ramachandran GN, Ramakrishnan C, Sasisekharan V (1963) Stereochemistry of polypeptide chain configurations. *J Mol Biol* 7:95–99
- Ramakrishnan V, Ranbhor R, Durani S (2005) Simulated folding in polypeptides of diversified molecular tacticity: implications for protein folding and de novo design. *Biopolymers* 78(2):96–105
- Ramakrishnan V, Ranbhor R, Kumar A, Durani S (2006) The link between sequence and conformation in protein structures appears to be stereochemically established. *J Phys Chem B* 110(18):9314–9323
- Rana S, Kundu B, Durani S (2004) Stereospecific peptide folds. A rationally designed molecular bracelet. *Chem Commun (Camb)* (21):2462–2463
- Rana S, Kundu B, Durani S (2005) A small peptide stereochemically customized as a globular fold with a molecular cleft. *Chem Commun (Camb)* (2):207–209
- Rana S, Kundu B, Durani S (2007a) A double catgrip mixed L and D mini protein only 20 residues long. *Bioorg Med Chem* 15(11):3874–3882
- Rana S, Kundu B, Durani S (2007b) A mixed-alpha, beta miniprotein stereochemically reprogrammed to high-binding affinity for acetylcholine. *Biopolymers* 87(4):231–243
- Ranbhor R, Ramakrishnan V, Kumar A, Durani S (2006) The interplay of sequence and stereochemistry in defining conformation in proteins and polypeptides. *Biopolymers* 83(5):537–545
- Richardson JS, Richardson DC (1988) Amino acid preferences for specific locations at the ends of alpha helices. *Science* 240(4859):1648–1652
- Robinson T (1976) D-Amino acids in higher plants. *Life Sci* 19(8):1097–1102
- Rohl CA, Strauss CE, Misura KM, Baker D (2004) Protein structure prediction using Rosetta. *Methods Enzymol* 383:66–93
- Rose GD, Gierasch LM, Smith JA (1985) Turns in peptides and proteins. *Adv Protein Chem* 37:1–109
- Rose GD, Fleming PJ, Banavar JR, Maritan A (2006) A backbone-based theory of protein folding. *Proc Natl Acad Sci USA* 103(45):16623–16633
- Sibanda BL, Thornton JM (1985) Beta-hairpin families in globular proteins. *Nature* 316(6024):170–174
- Sibanda BL, Blundell TL, Thornton JM (1989) Conformation of beta-hairpins in protein structures. A systematic classification with applications to modelling by homology, electron density fitting and protein engineering. *J Mol Biol* 206(4):759–777
- Struthers MD, Cheng RP, Imperiali B (1996) Design of a monomeric 23-residue polypeptide with defined tertiary structure. *Science* 271(5247):342–345
- Szabo G, Urry DW (1979) N-acetyl gramicidin: single-channel properties and implications for channel structure. *Science* 203(4375):55–57
- Tanford C (1968) Protein denaturation. *Adv Protein Chem* 23:121–282
- Urry DW (1971) The gramicidin A transmembrane channel: a proposed pi(L, D) helix. *Proc Natl Acad Sci USA* 68(3):672–676
- van Gunsteren WF, Billeter SR, Hünenberger PH, Krüger P, Mark AE, Scott WRP, Tironi IG (1996) Biomolecular simulation: the GROMOS96 manual and user guide. Vdf Hochschulverlag AG an der ETH Zürich, Zürich
- Wallace BA (1998) Recent advances in the high resolution structures of bacterial channels: gramicidin A. *J Struct Biol* 121(2):123–141
- Wallace BA (2000) Common structural features in gramicidin and other ion channels. *Bioessays* 22(3):227–234
- Watson JD, Milner-White EJ (2002a) 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* 315(2):183–191
- Watson JD, Milner-White EJ (2002b) A novel main-chain anion-binding site in proteins: the nest. A particular combination of phi, psi values in successive residues gives rise to anion-binding

- sites that occur commonly and are found often at functionally important regions. *J Mol Biol* 315(2):171–182
- Weaver TM (2000) The pi-helix translates structure into function. *Protein Sci* 9(1):201–206
- Srinivasan R RIBOSOME V1.0. <http://roslab.jhu.edu/~raj/Manuals/ribosome.html>
- Zhang C, DeLisi C (1998) Estimating the number of protein folds. *J Mol Biol* 284(5):1301–1305