RSC Advances

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/advances

Biomimetic Design: A Programmed Tetradecapeptide Folds and Homodimerizes as a Stereochemically Articulated Receptor Protein

Punam Ghosh, Deepa Pednekar, and Susheel Durani[a*]

[a]Department of Chemistry Indian Institute of Technology Bombay, Mumbai-400076.

(*Corresponding author E-mail: sdurani@iitb.ac.in Phone number: 91-22-25767164)

Abstract

Protein-quaternary structure may in mimicry of biogeny of the structure, be evolved to desired specificity—as fold, assembly, and sequence—in a hierarchy of independent design steps. In demonstration of the algorithm, octapeptides were accomplished to desired structural specificity as folds, assemblies, and sequences over a hierarchy of design steps; the approach was shown to extend the structure by allowing stereochemistry as a design variable, while it was proven to simplify design by allowing the programming to involve a hierarchy of independent design steps. Illustrating the above approach with stereochemistry as the variable, we now report functional design: a fourteen-residue sequence is accomplished to the desired binding specificity as a receptor protein, first in fold stereochemically, then in the assembly spatially, and finally in the sequence chemically.

Key words: protein stereochemistry, evolutionary design, β-hairpin fold, π-π interaction, self-assembly

Introduction

Biology rests on versatility of protein structure as the catalyst, regulator, and building block, etc., critical for the living cell. The basis may lie in the structure adapting in a hierarchy of independent selections, of conformation as folds, of spatial structure as assemblies, and of chemical structure as sequences. Having the higher level adapt over the lower one in an independent step, the hierarchy may have been crucial to biogeny by allowing the useable structures to recur, it would challenge design in the combinatorials possible. Design too may be in folds, assemblies, and sequences implemented in a hierarchy of selections. That a level of the natural hierarchy indeed may be programmed over the lower one in an independent step was considered for having the folds and assemblies of interest programmed as the sequences of desired specificity as chemical structures[1]. Given successes of the algorithm[2,3], we considered its extension for having not only sequences but also folds and assemblies adapted to the desired specificity computationally[4]. To test the algorithm, chirality in α-carbon was considered, in addition to conformation in main chain and chemical structure in side chains, as a variable for having the scope extended beyond what is possible with α-helix and β-sheet, the biological building blocks frozen as poly-L sequences. L- and D-α-Amino acids were introduced as a design alphabet,[5] which was proven to offer, fundamentally, the possibility of shape-specific design—bracelet-,[6] boat-,[7] canoe-,[8] and cup-shaped[9] folds were achieved as illustrations of the diversity of shapes possible. Stereochemistry may extend but will also complicate design in the combinatorials possible. To tackle the problem design was approached as a hierarchy of steps, programming first folds—in isolation of spatial and chemical variables—stereochemically, then assemblies—in isolation of conformational and chemical variables—spatially, and finally sequences—in isolation of conformational and spatial variables—chemically. With this algorithm, a shigatoxin-like C5 homopentamer could be accomplished as a protein remarkably small in its stereochemically articulated eight-residue folds contrasted with sixty-eight-residue folds of the natural protein poly-L in structure.[4] To further test the algorithm and stereochemistry as the variable, we now report functional design: a fourteen-residue sequence is accomplished to the desired binding specificity as a receptor protein, first in the fold stereochemically, then in the assembly spatially, and finally in the sequence chemically.

Bis-nitrophenylcarbonate (bNPC) and bis-nitropheylphosphate (bNPP) are harnessed as the model ligands. Bilateral symmetry of the structures is exploited for approaching the receptor structure biomimetically as an assembly C2 in symmetry, like the HIV protease active site[10]. A homooligomer of specific stoichiometry and point-group symmetry—cyclic, dihedral, or cubic—may be designed on basis of packing specificity of the monomer as the building-block fold. Packing specificity of the fold will depend on its shape, which in protein monomers is stereospecific for their topology as the poly-L folds[5]. With α-carbon chirality diversified, an assembly of interest may be adapted to the desired topological specificity in its building-block fold stereochemically. A desired assembly was thus approached by computational selection of stereochemistry and conformation in main chain, shape in fold of main chain, and chemical structure in side chains as the hierarchy of design variables[4]. Stereochemically diverse

homopolypeptides modeled computationally as random folds were thus first selected, with SYMMDOCK,[11] as the desired assemblies, and then inverse programmed, with IDeAS[21], as the desired sequences. Implemented over the random folds derived from the 256 stereoisomers possible for blocked octaleucine (Ac-Leu_8-NHMe), the method gave the sequences that folded and assembled to the desired specificity as homopentamers spontaneously[4]. With shape in fold directing its self assembly, the pentamer of interest was easy to achieve as the monomer fold of interest. Similar design of monomer for both self assembly and ligand binding is a tougher challenge for computation. For tackling the challenge we approach the desired receptor structure biomimetically. A β-hairpin ordered as a homodimer biomimetically is adapted for the desired binding specificity first stereochemically and then chemically.

Assemblies of desired structure can be specified with the protomer fold interacting over either its side chains or main chain. The C5 structure was specified with the interactions mainly of side chains[4]. The homodimer is in this study specified with the interactions of main chain. A type II' β- hairpin, centered on its conformational nucleator in $^{D}Pro_7$-Gly_8 structure, is targeted for self assembly over the interactions of main chain.[12,13] The dimer mutated in the sequence positions 4 and 11 from L to D structure, transforms, on well documented reasoning[5], as a boat-like homodimer. As illustrated in Figure 1, the dimer has a cleft for possible design as the desired receptor structure chemically. Only a limited number of interactions hold the monomers in the desired intermolecular assembly; thus there is interest to determine if the dimer will assemble and bind the desired ligand.

The dimer is targeted for chemical-level design as the desired receptor. Every sequence position, except $^{D}Pro_7$-Gly_8, is tested for a side chain for the fold, assembly, and complex to defined a minimum-energy sequence. Side chain polarity and aromaticity are applied as a heuristic to maximize hydrophobic gain and π-π interaction in the designed sequences. Side chain coordinates are drawn from a database of natural rotamers.[14,15] The structures are applied directly in the sequence positions of L stereochemical structure and, after a suitable symmetry transform, as the enantiomers in the positions of D stereochemical structure. Specific positions in the dimer loaded with the active site ligand are tested for the sterically compatible rotamers, which are evaluated as minimum energy pairs over the interacting positions in the bound complex using Dead-End-Elimination calculation.[16] Minimum energy sequences are computed over the selected rotamers with Monte-Carlo algorithm.[17] Orientation-dependent hydrogen bond energies,[18] solvent-accessibility based hydrophobic gain,[19] and steric and dispersion-interaction calculations involving Lennard-Jones potential[20] are the energy functions applied. The search methods and energy functions are implemented with IDeAS.[21] **P1**, **P2**, and **P3** are the sequences chosen for verifying design. The interactions likely to define the dimer and its bound complex as the minimum-energy sequences are summarized in Table 1. In **P1** Ile_2 substitutes Tyr_2 of **P2** and **P3** for hydrophobic effect or π-π interaction in ligand binding. In addition, $Trp_{6/9}$ and $His_{9/14}$ in all the sequences and Tyr_2 in **P2** and **P3** are for holding the ligand in π-π interaction.

P1: Ac-S_1-I_2-N_3-$^{D}E_4$-S_5-L_6-$^{D}P_7$-G_8-W_9-T_{10}-$^{D}K_{11}$-D_{12}-V_{13}-H_{14}-NH_2
P2: Ac-D_1-Y_2-N_3-$^{D}E_4$-S_5-W_6-$^{D}P_7$-G_8-I_9-T_{10}-$^{D}K_{11}$-S_{12}-L_{13}-H_{14}-NH_2
P3: Ac-N_1-Y_2-S_3-$^{D}E_4$-D_5-W_6-$^{D}P_7$-G_8-H_9-T_{10}-$^{D}K_{11}$-E_{12}-S_{13}-K_{14}-NH_2

Results and Discussion

MS and NMR

P1, P2, and P3 are made by manual solid-phase synthesis [26] and are judged as > 87 % pure with HPLC (Figure S5 †ESI). Appearance of requisite peaks in ESI-MS spectrum validates the structures (Figure S6 †ESI). ^{1}H NMR spectra in 90% D_2O-H_2O mixture reflect requisite resonances for the aromatic-H, peptide-NH, and other specific protons expected (Figure S7 †ESI). The spectra are assessed for effect of diluting the solutions from 5.0 to 0.5 mM concentration. Appreciable changes occur in selective peptide-NHs, which argue for possible involvement of the NHs in assembly on basis of inter-molecular hydrogen bonding. Selective aromatic-Hs in P2 and P3, but not in P1, are also affected by dilution, arguing for possible involvement of the aromatic structures with folding and/or self-assembly of the sequences (Figure S8 †ESI). Considering other evidences as shall be discussed, the assembly implied in mM concentration regime with NMR may reflect a higher order aggregation of the sequence as a homodimer structure. This is no surprise considering the notoriety of β-hairpin fold in its susceptibility to aggregation given that several peptides of main chain are unsatisfied in hydrogen bonds.

Circular Dichroism and Fluorescence

The sequences are assessed for assembly in 20-200 μM concentration range with CD and Fluorescence spectroscopy. As noted in Figure 2 Panel A, all sequences are similar in CD at 20 μM concentration but are remarkably structure specific in the high concentration CDs. A coupled minimum and maximum of ellipticity appears with concentration in P2 and P3, and possibly is a coupled exciton of aromatic side chains in the sequences.[27,28] The exciton is noted to be clearly absent in P1, which implies that Tyr_2 is involved in producing the excitons in P2 and P3; presumably, the interacting side chains are brought close by concentration dependent folding and/or self-assembly of the sequences. The observed CD change in P1 (Figure S9 †ESI) may reflect the ordering of main chain in the sequence following its concentration promoted self-assembly. The high-concentration CD in this case has deceptive appearance of α-helix CD but has the minima (200 and 215 nm) at the wavelengths lower than the requirement at ~208 and ~222 nm. Moreover, with its multiple D residues, including a Pro, P1 cannot adapt the α -helix fold due to unfavorable steric effect of the D residues. The lowest concentration CDs being similar suggest that the different sequences may be unordered as monomers that fold with self-assembly cooperatively. Conformed to the possibility, molar ellipticities has been plotted against concentration at 225 nm for all the sequences (Figure S10 †ESI). The protomer folds may be comparable in intermolecular strength, conformed to largely similar structures of the sequences. Fluorescence anisotropy is sigmoidal with concentration and has the midpoints coincide for all the sequences as is noted in Figure 2

Panel B. Concentration dependent change of anisotropy conforms to increase of mass, which directly evidences assembly of the structures. Thus on combined evidence of CD and fluorescence, the designed sequences are unordered monomers that fold with self-assembly on increase of concentration.

Dynamic Light Scattering

The sequences were assessed with DLS at 40 and 100 μM concentrations. At 40 μM concentration, a single peak of hydrodynamic radius ~ 4 nm appears (Figure 3). The size is in correspondence of the expectation from either the unordered monomer or the ordered dimer; the equilibrium cannot be diagnosed on this basis. The observation, however, does rule out any higher order aggregates at the given concentration. At 100 μM concentration, particles of hydrodynamic radius ~15 and ~100 nm appear; higher-order aggregates of the dimer may be involved due to the peptides of main chain that are unsatisfied in hydrogen bonds. On evidence of NMR and DLS, the sequences appear to aggregate in > 100 μM concentration. On combined evidence of DLS, CD, and Fluorescence, the sequences could be unordered monomers at < 20 μM concentration. In the intermediate concentration, CD, fluorescence, and DLS evidences argue that the unfolded monomer equilibrates with the folded dimer and at higher concentrations they fold to form oligomers.

Isothermal Titration Calorimetry and AutoDock

Ligand binding was evaluated with fluorescence quenching, Isothermal Titration Calorimetry (ITC), and AutoDock.[29] Titration with bNPC and bNPP promotes quenching of fluorescence of the sequences as is noted in (Figure S11 †ESI). The quenching is stronger at 120 μM concentration of the sequences than at 20 μM concentration. Presumably, the receptor orders with concentration due to self-assembly. The quenching data from 120 μM concentration experiment gave non-linear Stern Volmer plots (Figure 4 Panel A; Figures S12 Panel A †ESI). Static and dynamic modes of quenching may be involved. Fits to a modified Stern-Volmer equation given under the Materials and Methods Section are linear as is noted in Figure 4 Panel B, and Figure S12 Panel B †ESI. According to the derived interaction parameters given in Table 2, bNPC and bNPP are comparable in binding strength with the receptor structure.

A model for the ligated complex recovered as central member of the largest cluster prepared with MD was evaluated with AutoDock.[29] The calculated energies of ligand binding, given in Table 2, agree with the estimates from fluorescence quenching. It is possible that the modeled homodimer is the receptor binding the desired ligand. Interaction between **P3** and bNPP was assessed with ITC. The results are given in Figure 5. The interaction parameters derived by assuming 1:1 stoichiometry between the dimer structure and the ligand are given as inset in Figure 5. The binding strength compares well with the estimate from fluorescence quenching and AutoDock. The homodimer in **P3** may be the receptor structure binding the ligand. According to ITC, ligand binding is endothermic being positive in enthalpy. This may manifest desolvation penalty in the phosphate group of bNPP. Given the penalty, ligand binding is weak and may be driven

by entropy of hydrophobic gain. The comparable interaction strength of diverse sequences with geometrically distinct carbonate and phosphate ligands supports the possibility that being driven by hydrobhobic effect, ligand binding is low in structural specificity.

Molecular dynamics

The designs are verified with molecular dynamics in GROMACS package[22,23] with gromos-96 force field that is modified for treatment of D residues. Simulations initiated in a desired fold, assembly, and ligated complex are run for 50 ns each in a periodic box of SPC water.[24] Harvested at 10 ps intervals, the sequences are clustered to 0.15 nm RMSD cutoff over Cα of polypeptide structure using Duara *et al.* algorithm.[25] Thermodynamic stability of a fold, assembly, and complex is judged on basis of the number of clusters to which the structure drifts and the size of minimum-energy cluster taking it to approximate statistical weight of the desired structure in the quasi Boltzmann-like ensemble modeled. The results in Figures S1, S2, S3, and S4 †ESI justify the designs, but also imply that the structures may be rather weak in ordering as the desired folds, assemblies, and ligated complexes.

Conclusion

The basis for adaptive versatility of protein structure, in its hierarchic construction plan, may be the basis for the protein-design challenge a well, in the combinatorials possible. Protein design was greatly simplified by having chemical level of the natural hierarchy optimized in an independent step while having the conformational and spatial levels frozen as the fold and assembly of interest[1-3]. The possibility of having the levels optimized in steps prompted us to explore the biomimetic algorithm to have protein design extended with stereochemistry as an added variable. Exploring the algorithm and the added variable, two successful designs have been accomplished: a small C5 structure in mimicry of the much larger natural protein shiogatoxin[4], and now a small receptor structure of C2 symmetry in mimicry of the much larger natural protein HIV protease.

Experimental Section

Fmoc-protected amino acids, reagents for solid phase synthesis, Rink Amide AM resin, solvents, bis-nitrophenylcarbonate, and bis-nitrophenylphoshate were purchased from Sigma-Aldrich or Novabiochem-Merck.

Peptide Synthesis

Synthesis was performed manually on Rink Amide AM resin using standard Fmoc chemistry and HOBt/DIC as the coupling reagents. Each coupling, monitored with Kaiser and chloranil tests, typically required about 6 hrs. Deprotections were with 30% (v/v) piperidine-DMF. N-terminus was acetylated (-$NHCOCH_3$) with Ac_2O: DIPEA:DMF in 1:2:20 ratio. The cleavage of final polypeptide and deprotection of side chains were achieved with reagent K (82.5% TFA/5% dry-phenol/5% thioanisole/2.5% ethandithiol/5% water). The product precipitated with anhydrous diethyl ether was lyophilized from 1:4 H_2O:^{t}BuOH solution as a white powder. The final purification of the peptide was accomplished with HPLC over RP-18 (10 μm, 10 mm ×250 mm; Merck) eluting with ACN \ H2O \ 0.1%TFA \ 0-100% gradient.

ESI-MS

Mass spectra were recorded on QTOF-ESI Mass Spectrometer. Positive ions were detected in linear / reflectron mode.

Nuclear Magnetic Resonance

NMR experiments were performed on Bruker 800 MHz instrument equipped with cryoprobe, at 298 K in 9:1 H_2O and D_2O with 0.1 mM 2,2-dimethyl-2-silapentane-5-sulfonate sodium salt (DSS) as internal reference at 5 mM concentration. 1H NMR data has been processed with TOPSPIN software of Bruker.

Circular Dichroism (CD)

CD measurements were performed on JASCO J-815 CD spectropolarimeter calibrated with d_{10}-camphoursulphonic acid. Data were collected at 298 K in 0.2 cm path-length quartz cell with 2 nm bandwidth in 195 - 240 nm range. Scanning at 100 nm /min with 1.0 s time constant in 1 nm steps, five scans were averaged after baseline correction for solvent. Peptides were prepared in the range of 20-200 μM by optical measurements in aqueous solution of water-methanol mixture at pH 7.0. The observations in millidegrees were converted to mean residue ellipticity using a reported relation.[30]

Fluorescence

Fluorescence measurements were taken on a Varian Cary Eclipse fluorimeter. Data were collected at 298 K in 1cm cell, with excitation wavelength of 280 nm. The excitation and emission slits were 5 nm and scanning speed was 120 nm/min. The peptide

solutions were prepared in 10-100 μM concentration range by optical measurements of the peptide and of the standard NATA (SIGMA chemicals). Binding experiments were carried out at 20 and 120 μM concentration of the peptide and 100 μM concentrations of bNPC and bNPP ligands. The binding of the ligands with the peptide dimers are accompanied by quenching of Trp fluorescence. The fluorescence quenching data were analyzed with Stern-Volmer equation:

$F_0/F = 1 + K_{sv}[Q]$…… (1);

in which F_0 and F are the fluorescence intensities observed in absence and presence of quencher, [Q] is quencher concentration and K_{SV} is the Stern-Volmer quenching constant. In event of nonlinear Stern Volmer plots, both static and dynamic quenching could be involved according to the relation:

$F_0/F = (1 + K_{SV}[Q])\,(1 + K_a[Q])$……(2);

in which case the modified 'Modified Stern-Volmer' equation is:

$F_0/F_0\text{-}F = (1/f_a) + (1/f_aK_a[Q])$……(3);

where f_a is the fraction of fluorophore accessible to the quencher and K_a is the association constant of the quencher to the fluorophore. F_0/F_0-F against the reciprocal of ligand concentration [Q], is a linear equation which gives association constant K_a and binding free energy calculated as:

$\Delta G = -RT \ln K_a$……(4)

Fluorescence Anisotropy

Steady state fluorescence anisotropy measurements were recorded with Varian Cary Eclipse fluorimeter equipped with an emission-excitation polarizer. All the experiments are performed at 25^0C.

Dynamic Light Scattering

The size-distributions of the peptide samples were assessed by DLS using a Nano-ZS (Nanoseries) Malvern Instrument with a 35 mW diode laser source. The measurements were at 25°C and involved 90° scattering angle. The samples were prepared in water using 0.22 m filter with four hours of sonication. Processing of the fluctuating signal with digital autocorrelator gave particle's diffusion coefficient, particle size with the Stokes-Einstein equation. The resulting correlation function was processed to size-distribution for the particles using Malvern Software.

Isothermal Titration Calorimetry (ITC)

The experiments were performed on a VP-ITC micro calorimeter (Microcal, Inc.) at 298 K. Sample cell contained peptide in 250 μM concentration, as determined by optical measurement, and reference cell contained water. The 12.5 mM bis-nitrophenyl-

phosphate (ligand) solution loaded in 250mL syringe was titrated into peptide solution in 10 μL aliquots in 25 steps at 4 min intervals. The change in enthalpy (DH) due to dilution was determined by titrating ligand into solvent as well as solvent into peptide solution. These backgrounds were subtracted from DH obtained for the corresponding ligand-peptide binding experiments, prior to curve fitting. The background-subtracted data was fitted to a model describing single binding site using MicroCal software. The binding enthalpy (DH), entropy (DS), and association (K_a) constant were thus calculated.

Computational Methods

Inverse Algorithm for Sequence Design

The inverse-design package IDeAS implements combination of deterministic pruning algorithm Dead End Elimination and stochastic search method Monte Carlo. Energy functions for sequence solution involve accessibility-based solvation energy, H-bond energy, coulomb energy, and entropy in form of probability of side chain usage. The software uses natural database of rotamers and their suitable symmetry transforms as the rotamers of D structure. IDeAS has been validated in other examples of protein design.[31,32]

Molecular Dynamics

MD was performed using gromos-96-43a1 force field in GROMACS package. Simulations were under NVT with periodic boundary condition in a 4.8 nm cubic box for monomer and a 5.5 nm box for dimer. SPC water model was used. Bond lengths were constrained with SHAKE.[33] Time step of 2 femto second was used. Non-bonded interactions were cut-off at 1.4 nm using shift function. The potential energy of the system, i.e., peptide and water, was minimized using steepest descent algorithm with a tolerance limit of 100 KJ mol^{-1} nm^{-1} in sufficient number of steps to achieve convergence. Position restrained molecular dynamics was performed for 200 ps at 298 K to equilibrate water molecules. Initial velocities were drawn from Maxwellian distribution. The peptide and solvent were separately coupled to Berendsen temperature bath using a time constant of 0.1 ps. Conformational clustering was performed to RMSD cut-off ≤ 0.15 nm over Ca atoms.

Molecular Docking

Flexible docking algorithm was implemented with AutoDock 4.0. Central member of the top cluster populating the MD trajectory over aromatic residues comprising the biding site was used as the receptor structure for calculation of binding energy of the receptor and its ligand. Using genetic algorithm for ligand docking and an RMSD tolerance of 2 Å, structurally distinct conformational clusters of the ligand were ranked in increasing energy.

Acknowledgments

We acknowledge DST (SR/S1/OC-74/2008) Government of India and Department of Chemistry, IIT Bombay for the financial support and supercomputing facility "Corona".

References

[1] D.N. Bolon and S.L.Mayo, *PNAS*, 2001, **98**, 14274.
[2] J.B. Siegel, A. Zanghellini, H.M. Lovick, G. Kiss, A.R. Lambert, J. L. St.Clair, J. L. Gallaher, D. Hilvert, M. H. Gleb, B. L. Stoddard, K. N. Houk, F. E. Michael and D. Baker *Science,* 2010**, 329**, 309.
[3] H.K. Privett, G. Kiss, T.M. Lee, R. Blomburg, R.A. Chica, L.M. Thomas, D. Hilvert, K.N. Hour and S.L. Mayo, *PNAS*, 2012, **109**, 3790.
[4] P. Ghosh, A. U. Mushtaq and S. Durani, *RSC Adv.*, 2012, **2**, 3242.
[5] S. Durani, *Acc. Chem. Res.,* 2008, **41**, 1301-1308.
[6] S. Rana, B. Kundu and S. Durani, *Chem. Commun*., 2004, **21**, 2462.
[7] S. Rana, B. Kundu and S. Durani, *Chem. Commun*., 2005, **2**, 207.
[8] S. Rana, B. Kundu and S. Durani, *Bioorg. Med. Chem*., 2007, **15**, 3874.
[9] S. Rana, B. Kundu and S. Durani, *Biopolymers*, 2007, **87**, 231.
[10] J. Erickson, D. J. Neidhart, J. VanDrie, D. Kempf, X. C. Wang and D. W. Norbeck, *Science*, 1990, **249**, 527.
[11] D. S. Duhovny, Y. Inbar, R. Nussinov and H. J. Wolfson, *Nucleic Acids Research*, 2005, **33**, 363.
[12] S. M. Noble, V. E. Carnahan, L. B. Moore, T. Luntz, H. Wang, O. R. Ittoop, J. B. Stimmel, P. R. Davis-Searles, R. E. Watkins, G. B. Wisely, E. LeCluyse, A. Tripathy, D. P. McDonnell and M. R. Redinbo, *Biochemistry*, 2006, **45**, 8579.
[13] P. B. Stranges, M. Machius, M. J. Miley, A. Tripathy and B. Kuhlman, *Proc. Natl. Acad. Sci. U. S. A*., 2011, **108**, 20562.
[14] R. L. Dunbrack, *Curr. Opin. Struct. Biol*., 2002, **12**, 431.
[15] S. C. Lovell, J. M. Word, J. S. Richardson and D. C. Richardson, *Proteins*, 2000, **40**, 389.
[16] R. F. Goldstein, *Biophys. J.*, 1994, **66**, 1335.
[17] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. J. Teller, *Chem. Phys.*, 1953, **21**, 1087.
[18] D. B. Gordon, S. A. Marshall and S. L. Mayo, *Curr. Opin. Struct. Biol.*, 1999, **9**, 509.
[19] B. I. Dahiyat and S. L. Mayo, *Prot. Sci.*, 1996, **5**, 895.
[20] D. N. Bolon, J. S. Marcus, S. A. Ross and S. L. Mayo, *J. Mol. Biol.*, 2003, **329**, 611.

[21] R. Ranbhor, A.Tendulkar, A. Kumar, V. Ramakrishnan, K. Patel, K. R. Srivastava and S. Durani, *Unpublished*

[22] E. Lindahl, B. Hess and D. van der Spoel, *J. Mol. Mod.*, 2001, **7**, 306.

[23] W. F. van Gunsteren, S. R. Billeter, A. A. Eising, P. H. Hunenberger, P. Kruger, A. E. Mark, W. R. P. Scott and I. G. Tironi, *Biomolecular Simulation: The GROMOS96 manual and user guide*, Hochschulverlag AG an der ETH Zurich, Zurich, Switzerland, 1996.

[24] H. J. C. Berendsen, , J. P. M. Postma, W. F. van Gunsteren, J. Hermans and B. Pullman, *Interaction models for water in relation to protein hydration. Intermolecular Forces, Reidel Publishing Company: The Netherlands, 331, 1981.*

[25] X. Daura, W. F. van Gunsteren and A. E. Mark, *Proteins: Struct. Funct. Genet*., 1999, **34**, 269.

[26] W. C. Chan, P. D. White, *Fmoc Solid Phase Peptide Synthesis: A Practical Approach*, IRL Press Oxford, UK, 1989.

[27] M. L. Waters, *Curr. Opin. Chem. Biol*., 2002, **6**, 736.

[28] C. A. Hunter, K. R. Lawson, J. Perkins and C. J. Urch, *J. Chem. Soc.*, 2001, **2**, 651.

[29] G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell and A. J. Olson, *J. Comput. Chem.*, 2009, **30**, 2785.

[30] G. D. Fasman, *Circular Dichroism and the Conformational Analysis of Biomolecules,* Plenum Press: New York, 1996.

[31] K. Patel, B. Goyal, A. Kumar, N. Kishore and S. Durani, *J. Phys. Chem.*, *B* 2010, **114**, 16887.

[32] K. Patel, K. R. Srivastava and S. Durani, *Bioorg. Medicinal Chem.*, 2010, **18**, 8270.

[33] J. P. Ryckaert, G. Ciccoti and H. J. C. Berendsen, *J. Comp. Phys.*, 1977, **23**, 327.

Figure Legends

Figure 1: Stereochemically bent β-hairpin as a C_2-symmetric homodimer (Panel A) complexed with bis-nitrophenyl-phosphate (bNPP) as the binding ligand (Panel B).

Figure 2: Coupled exciton in CD spectrum for **P3** (Panel A) and the associated changes of ellipticity at 225 nm in **P1, P2,** and **P3** (Inset Panel A) and concomitant increase of fluorescence anisotropy (Panel B) are sigmoidal functions of concentration.

Figure 3: Results of DLS showing hydrodynamic radius of **P3** at two concentrations of the peptide.

Figure 4: Plots of fluorescence quenching data of **P3** with specific binding ligands according to Stern-Volmer equation and a modified form of the equation

Figure 5: Heat changes on addition of 10μL aliquots of bNPP (12.5 mM) into a solution of **P3** in 250 μM concentration (Upper Panel) and the integrated heat changes (Lower Panel). Fit for single-site binding model is shown as a solid line. The calculated thermodynamic parameters are shown as inset.

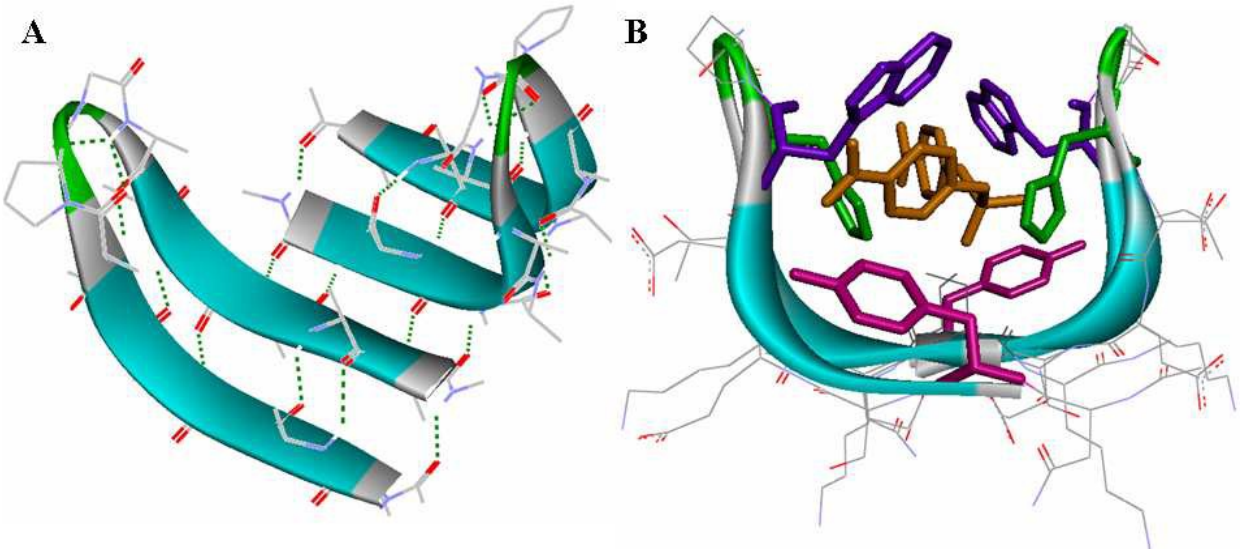


Figure 1

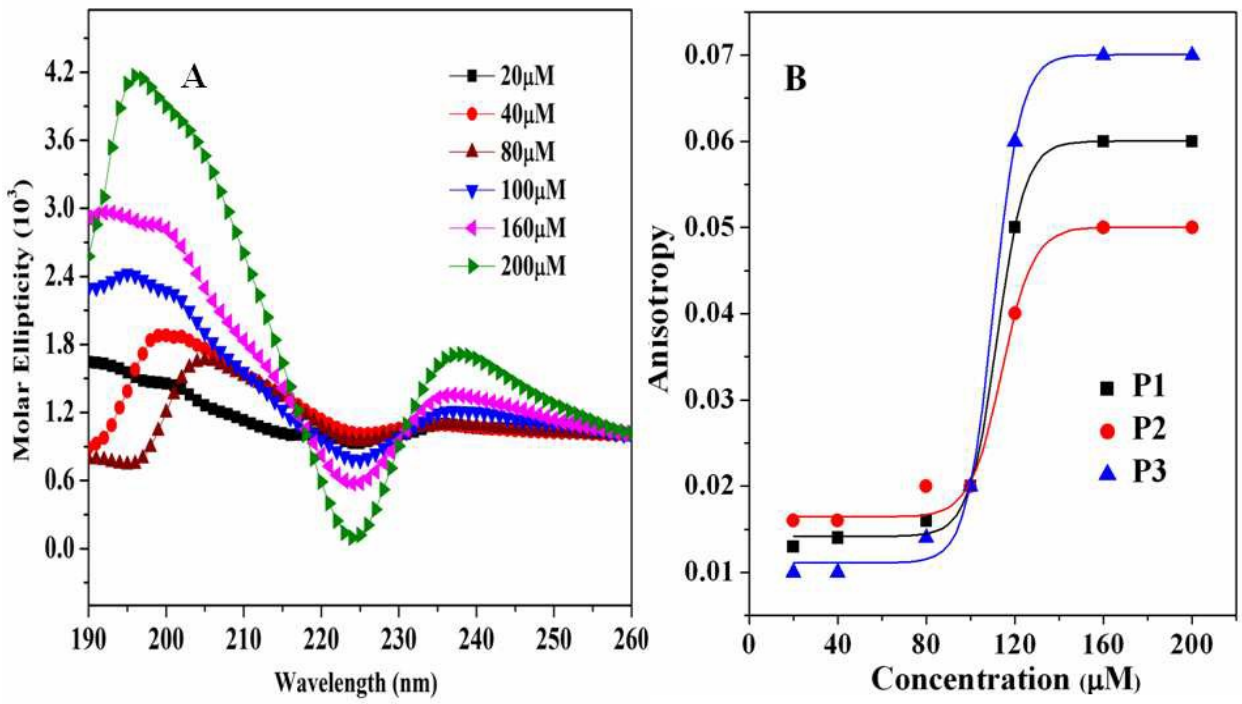


Figure 2

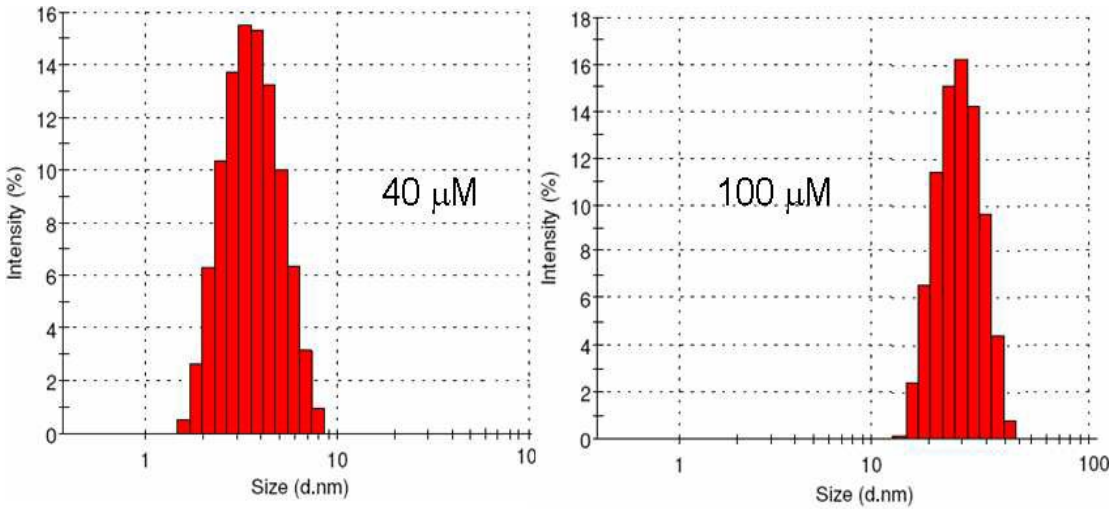


Figure 3

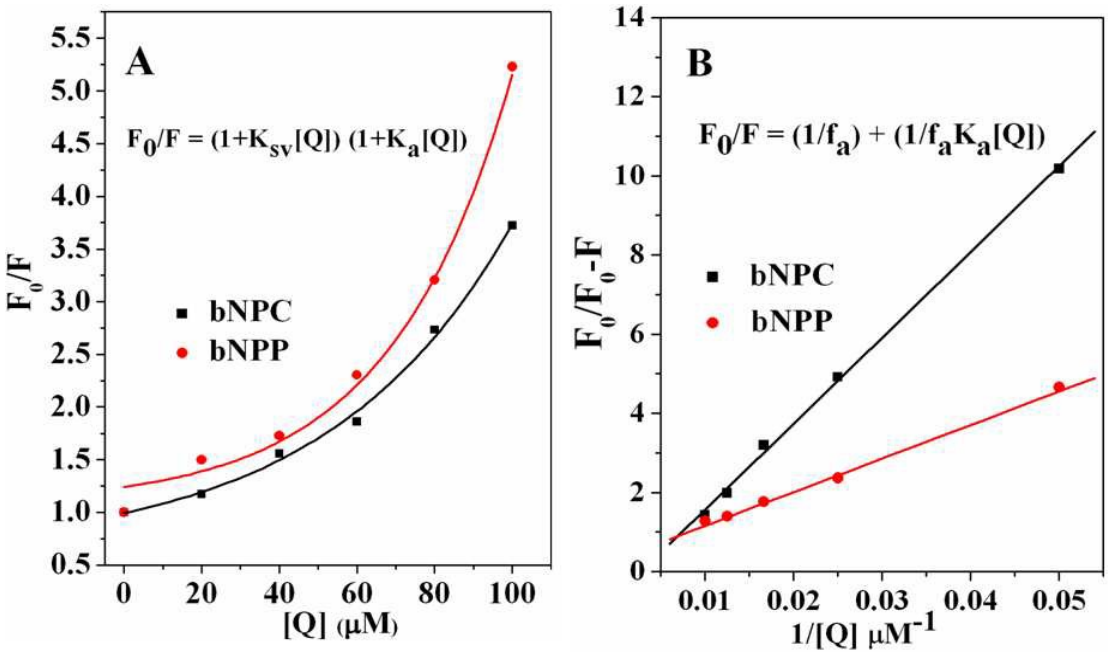


Figure 4

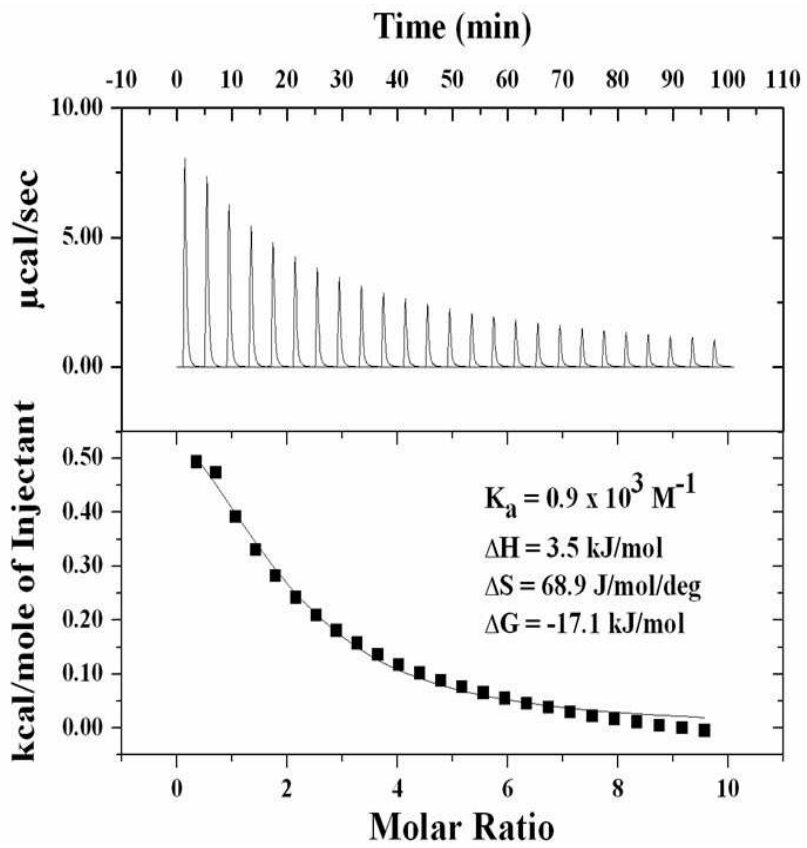


Figure 5

Table Legends

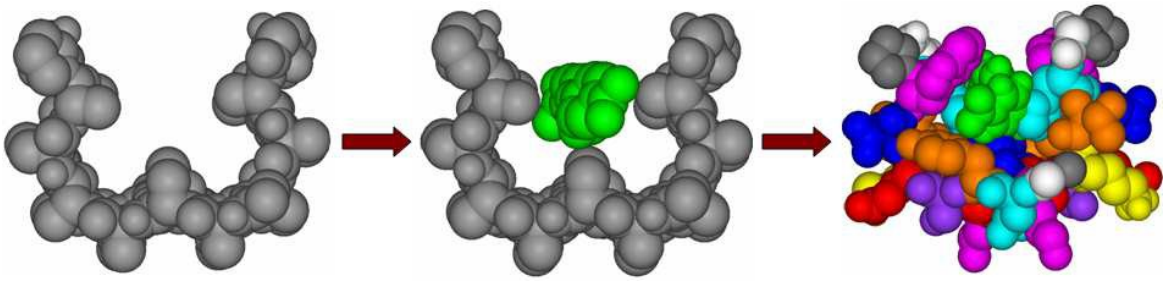
Table 1. Specific interactions in complexes of peptide homodimers with ligands.

Interactions	H-Bond	Salt Bridge	π –π/ CH-π/ Hydrophobic
Intra-molecular	$^{2}I_{NH}$-$^{13}V_{CO}$(1), $^{13}V_{NH}$-$^{2}I_{CO}$(1), $^{2}Y_{NH}$- $^{13}L_{CO}$(2), $^{13}L_{NH}$-$^{2}Y_{CO}$(2), $^{4}E_{NH}$-$^{11}K_{CO}$(1,2,3), $^{11}K_{NH}$-$^{4}E_{CO}$(1,2,3), $^{6}L_{NH}$-$^{9}W_{CO}$(1), $^{9}W_{NH}$-$^{6}L_{CO}$(1), $^{6}W_{NH}$-$^{9}I_{CO}$(2), $^{9}I_{NH}$-$^{6}W_{CO}$(2), $^{6}W_{NH}$-$^{9}H_{CO}$(3), $^{9}H_{NH}$- $^{6}W_{CO}$(3), $^{9}W_{NH}$-$^{7}P_{CO}$(1), $^{9}I_{NH}$-$^{7}P_{CO}$(2), $^{9}H_{NH}$-$^{7}P_{CO}$(3)	$^{D}E_{4}$-$^{D}K_{11}$ (1, 2, 3) S_1-H_{14} (1) D_1-H_{14} (2) H_9-S_{13} (3)	Y_2-W_6 (2, 3) W_6-H_9 (3) L_6-W_9 (1) W_6-I_9 (2) Y_2-L_{13} (2) I_2-V_{13}, I_2-L_6 (1) I_9-L_{13} (2)
Inter-molecular	$^{12}D_{NH}$-$^{14}H_{CO}$ (1) $^{14}H_{NH}$-$^{12}D_{CO}$ (1) $^{12}S_{NH}$-$^{14}H_{CO}$ (2) $^{14}H_{NH}$-$^{12}S_{CO}$ (2) $^{12}E_{NH}$-$^{14}K_{CO}$ (3) $^{14}K_{NH}$-$^{12}E_{CO}$ (3)	D_{12}-H_{14} (1) E_{12}-K_{14} (3)	W_9-W_9 (1) W_6-W_6 (2, 3) H_9-H_9 (3) V_{13}-V_{13} (1) L_6-L_6 (1) I_9-I_9, L_{13}-L_{13} (2)
Peptide-ligand	H_{NH}^{δ} (3)-CO L		W_9 (1)-L Y_2/W_6 (2)-L Y_2/W_6/H_9 (3)-L I_2/L_6/V_{13} (1)-L I_9/L_{13} (2)-L

Table 2. Binding parameters assessed with quenching of fluorescence and calculated with AutoDock.

Peptide	Ligand	K_a (x10^3 M^{-1})	ΔG^0 Fluorescence (kJ mol^{-1})	ΔG^0 AutoDock (kJ mol^{-1})
P1	bNPC	2.4	-19.29	-18.4±0.33
	bNPP	7.4	-22.08	-21.2±0.56
P2	bNPC	9.5	-22.68	-18.1±0.28
	bNPP	8.7	-22.48	-22.5±0.13
P3	bNPC	1.2	-17.51	-20.4±0.33
	bNPP	4.6	-20.89	-22.7±0.07

Graphical Abstract



Exploiting protein-structure evolution as the algorithm to select first in folds, then assemblies, and finally sequences, a fourteen residue polypeptide is accomplished to the desired binding specificity as a receptor protein of C2 symmetric structure much smaller than HIV protease as the natural example. The successful implementation of folding, self-assembly, and chemical optimization as a hierarchy of steps illustrates biomimetic approach as the design algorithm and stereochemistry as the design variable. A Stereochemically Articulated Receptor Protein is thus accomplished in biomimetic steps of a simplifying design algorithm