IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES
ISSN PRINT 2319 1775 Online 2320 7876
 Volume 11, Iss 10, 2022

Computational Analysis of Sixteen Residue Polypeptide Chain for Its Structure and Interaction Specificity

Kirti Patel
Department of Chemistry, N. B Mehta Science College, Bordi, Tal. Dahanu, Dist. Palghar, Maharashtra, India – 401701
Affiliated to University of Mumbai.

Abstract:

Computational study of 16-residue polypeptide chain is reported for its stability in structure and specificity in chemical interaction using Gromacs software. The results are observed to be in agreement with the experimental results on the same polypeptide chain.

Index Terms:**Aromatic π–π interactions, Clustering,Hydrogen bonds, IDeAS software, Molecular Dynamics, Radius of gyration, Salt bridge interactions, Simulated Annealing.**

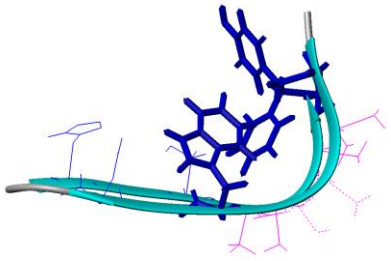
I. INTRODUCTION

Nature uses proteins as workhorses to perform various functions of the cell. Proteins are polypeptides with side chains as the variable alphabet that codes for the function associated with it. (Branden & Tooze, 2012) Understanding the codes means knowing the energetic link between structure and function. The energetics lies in the hydrogen bonds and electrostatics of polar and ionizable side chains and solvent-entropy-mediated interactions of non-polar side chains.(Avbelj &Moult, 1995; Avbelj & Baldwin, 2002; Chan & Dill, 1991; Dill, 1990; Dill, 2005; Kauzmann, 1959) The aromatic side chains were historically considered as belonging in this "polar-apolar" division, but were later recognized as chemically distinctive in their interactions (Dougherty, 1996; Ma & Dougherty, 1997; Meyer, 2003). The understanding of folding in proteins and catalysis in enzymes is thus a continuing quest to find for the basis in energetics. Much has been learnt, yet much remains unknown as the enzyme design remains a subject but in infancy. The studies in this paper are aimed at the computational studies of the stability of the structure using molecular dynamics simulations of the previously reported 16 residue polypeptide chain exhibiting mild hydrolase activity (Patel, 2010).

The polypeptide fold as poly-leucine was designed as β-hairpin using Simulated Annealing Molecular Dynamics SAMD method as described by Durani and group (Ramakrishnan, 2004; Ramakrishnan, 2005; Ramakrishnan, 2006). The selected fold was then given side chain variability using in-house software IDeAS (Ranbhor, 2018). The IDeAS software was used separately to design sequence in catalytic pocket for hydrolase activity and binding pocket for the substrate using aromatic π–π interactions. The rest of the sequence was adopted from whole set of amino acid side chains. The peptide was synthesized using solid phase synthesis and the experimental studies are reported (Patel, 2010).

Polypeptide sequence:

$AcY_1E_2{}^DL_3N_4W_5S_6D_7{}^DP_8G_9H_{10}T_{11}S_{12}A_{13}{}^DV_{14}K_{15}F_{16}NH_2$



ISSN PRINT 2319 1775 Online 2320 7876
Research paper **Volume 11, Iss 10, 2022**

Fig. 1: Cartoon representation of peptide fold depicting the design of aromatic cluster, catalytic active site and the D amino acids at 3, 8 and 14 positions.

II. METHOD: MOLECULAR DYNAMICS

A Dell Linux server,Intel Core i5-3337U CPU@1.80GHz server, and aPARAM Padma supercomputer equipped with 248 (Power4@1GHz) processors and an aggregate memory of 512 GB werethe hardware platforms were used. Molecular dynamics was carriedout in the GROMACS package (version 2.1) using the gromos-96 43A1 force field(van Gunsteren, 1996) in abox of explicit solvent with periodic boundary conditions underNVT (constant number of particles, volume, and temperature).The nonbonded list cutoff was 1.4 nm with a shift at 0.8 nm,the integration step was 2 fs, the initial velocities were drawnfrom the Maxwellian distribution, and the temperature wascoupled to an external bath with a relaxation time constant of0.1 ps. Bond lengths were constrained with SHAKE (Ryckaert, 1997) to ageometric accuracy of 0.0001. First, the model peptideenergy minimized was placed in a periodic cubic box ofappropriate edge length and soaked in SPC water (Berendsen, 1981), to thedensity in correspondence of 1 atm at 298 K. The system wasenergy minimized first in solvent restraining the solute and thenin both solvent and the solute relieved of the restraint. Moleculardynamics was initialized, and the trajectory was sampled atintervals 5ps after allowing an initial3 ns for equilibration.Conformational clustering to achieve microstates of peptidestructure to 0.15 nm rmsd cutoff was performed with a reportedprocedure (Lazar, 2003). The difference Helmholtz free energy ($\Delta F_{A\text{-}B}$)between specific microstates was calculated from relativeprobabilities pA and pB of finding the system in microstates Aand B as $\Delta F_{A\text{-}B} = -RT \ln pB/pA$, with R as the gas constant,T as the temperature, and pA and pB the number of members inmicrostates A and B.

III. RESULTS

Conformational equilibria with molecular dynamics

The peptide modeled as β–hairpin was submitted to MD in a periodic box of SPC water model (Berendsen, 1981) under NVT, viz., constant number of particles, volume, and temperature. The simulation was run for 100ns specific duration after allowing peptide chain to equilibrate for 3ns. The conformers were harvested at 5ps interval. The ensembles were evaluated in the occupancy statistics over ϕ,ψ space and in hydrogen bonds characterizing secondary structure as shown in Figure 2. Thestructures harvested from the molecular dynamics trajectory, were clustered to 0.15 nm rmsd cutoff over Cα coordinatesas specific folds (Daura, 1999).The peptide populate in predominantly the β–basin of Ramachandran ϕ,ψ space, as noted in Figure 2. The central member of the most populated cluster forms β–hairpin structure as shown in the Figure 3. The populated folds are with an average of 5.69 hydrogen bonds over backbone and 6.73 over side chains, as noted in Table 2. The main chain hydrogen bonds are predominantly long ranged, being n–n≥6 type (~80%), as expected in a β-hairpin, with about a fifth of them n–n± 3,4 type (16–19%) as required for the β–turn as shown in the Figure 2.

Table 1: Statistics of conformers and microstates in molecular dynamics simulation of sixteen-residue peptide.

Length (ns)	No. of conformers	No. of microstates	% population in top 5 clusters				
			1	2	3	4	5
150	30001	13	92.70	4.03	2.32	0.49	0.25

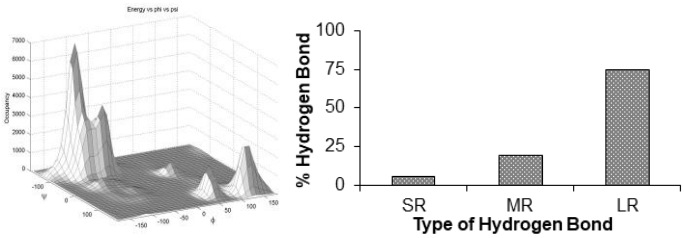


Fig. 2: ϕ, ψ distributions and percentage distribution of H-bonds of peptide over complete trajectory.

ISSN PRINT 2319 1775 Online 2320 7876
 Volume 11, Iss 10, 2022

Table 2: Hbond distribution over complete MD trajectory.

Type	Avg Hbonds per conformer	% population of Hbond type			
		n-n≤2	n-n±3,4	n-n±5	n-n≥6
Overall	**6.73**	**5.27**	**19.63**	**1.90**	**72.71**
Backbone	*5.69*	*3.54*	*16.65*	*0*	*79.81*

The quasi-equilibria characterizing MD ensembles were evaluated for the microstates in peptide structure, being the conformers clustered to 0.15 nm RMSD cutoff over backbone (Daura, 1991). The statistics of microstates are summarized in Table 1 and the central member of the most populous microstate populating the peptide is shown in Fig. 3. The growth in the number of clusters with time (ns) and % population of the most populated cluster are shown in the Figures 4 and 5 respectively. 13 number of microstates are observed for the peptide and the most populous microstate comprises 92.70% of the quasi-equilibria. The peptide is remarkably well ordered as shown in the Helmholtz free energy diagram in Figure 6.The highly ordered structure is also in concordance with experiment (Patel, 2010), which implies the convergence of favorable interactions over the distinctive chemical groups of main chain and side chains. Asp at position 7 is having strong cross-strand interaction with Ser at position 12. To assess the role of side chain interactions, radius-of-gyration distribution was evaluated over specific side-chain groups, the aromatic group of Tyr, Trp, and Phe, the catalytic group of Asp, His, and Ser, the Lys-Glu salt bridge and Ser-Thr cross-strand pair. The histograms over different distributions in Figure 7 and the mean RMSD over specific grouping in Table 3 illustrate the chemically specific groups in their tendency to contribute in peptide folding.

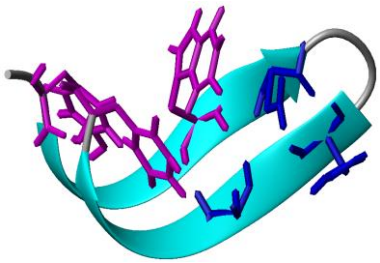


Fig. 3: Central member of most populous microstate of the peptide.

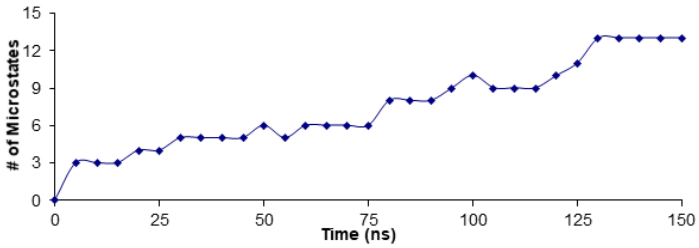


Fig. 4: Evolution of microstates as a function of time for the peptide.

ISSN PRINT 2319 1775 Online 2320 7876
 Volume 11, Iss 10, 2022

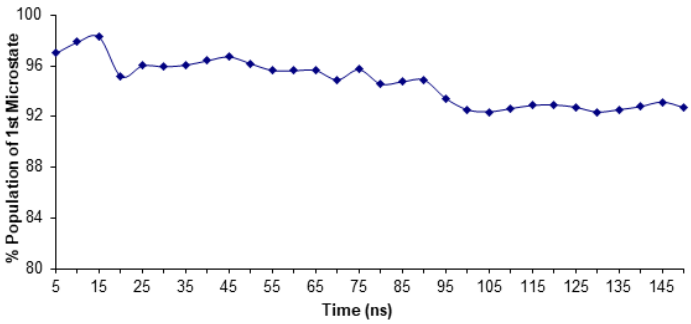


Fig. 5: Percentage population of first microstate as a function of time for the peptide.

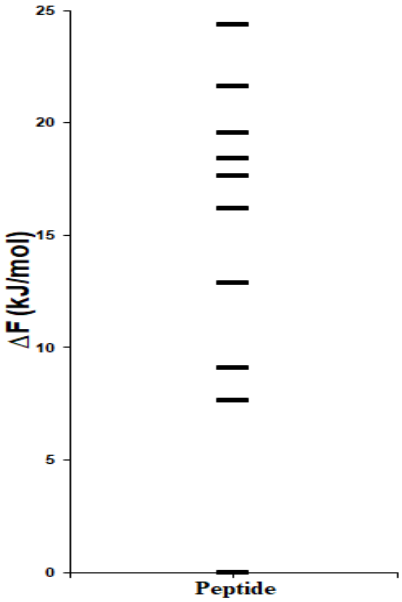


Fig. 6: Comparison of Helmholtz free energy of microstates with respect to first microstate of the peptide

Table 3: Statistics of radius of gyration values over fold (overall), backbone atoms, aromatic residue side chain atoms (Y-W-F), salt bridge residue side chain atoms (E-K), Serine-Threonine side chain atoms (S-T) and Active Site residue side chain atoms of the peptide.

Overall	Backbone	Y-W-F	E-K	S-T	Active Site
0.71 ± 0.03	0.67 ± 0.03	0.45 ± 0.03	0.37 ± 0.02	0.32 ± 0.02	0.38 ± 0.03

.

ISSN PRINT 2319 1775 Online 2320 7876
Research paper **Volume 11, Iss 10, 2022**

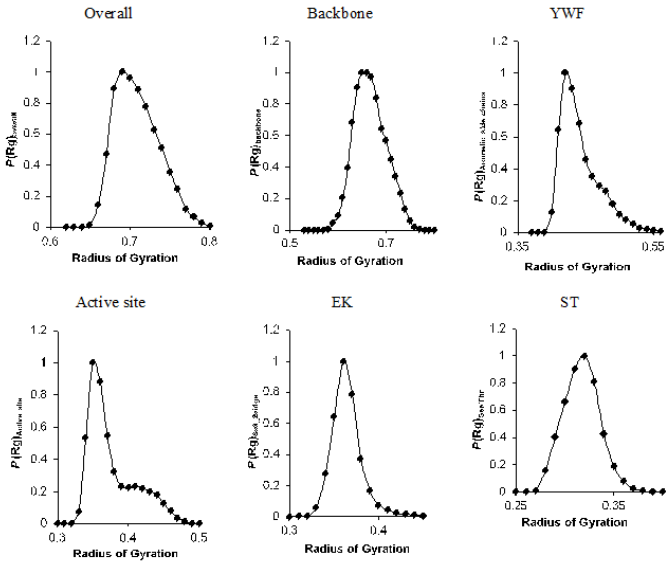


Fig. 7: Distribution of radius of gyration of the peptide fold. The radius of gyration calculated over all the atoms of peptide (overall), backbone atoms (backbone), active site residue side chain atoms (Active Site), aromatic residue side chain atoms (YWF), Glutamate and Lysine residue side chain atoms (EK) and Serine and Threonine residue side chain atoms (ST).

IV. DISCUSSION

A minimal acetylcholinesterase was targeted to have an active-site pocket of aromatic and acid-base-nucleophile side chains for possible catalyzed hydrolysis of acetylcholine. Enzymes as polypeptides are catalytic because of their folding under guidance of the sequential effects of side chains. The effects in protein folding and enzyme catalysis are not only chemical but also stereochemical in the bases. The chemical effects have direct concern with catalyses' while stereochemical effects are critical to the folds possible in approaching an enzyme. The stereochemical consideration central to protein folding is that of configuration in side chain attachment. The configuration remains L due to biological reason, causing protein folds to remain restricted as only but the possibilities of poly-L-peptide structures. Polypeptides, in principle, may be L or D in the residue-level stereochemistry, because of which it may be possible for proteins to be diversifiable stereochemically. Thus, while the natural proteins remain restricted in their folds, the fold possibilities may be widened with use the L and D structures as the alphabet. Though protein structure has been minimized drastically, the scope of functional design is also reduced drastically due to limited poly-L folds that are possible. Diversified stereochemically, the fold space may widen, expanding the scope in functional design. We tested the possibility by modifying poly-L hairpin to form a pocket for catalyzed hydrolysis of acetylcholine as a model substrate. Thus, the chemical elements of canonical protein structure were tested for catalyzing the hydrolysis in an artificially created active site pocket. Aromatic side chains were tested for substrate specificity while acid-base-nucleophile side chains were tested for chemical catalysis of substrate hydrolysis. Exploiting stereochemistry as the design tool, a sixteen-residue peptide was folded as an enzyme model. The model was shown to fold with effects of chemical and stereochemical structure.

The search for a minimal enzyme using the chemical groupings of protein structure elicits more than academic interest. The understanding of protein folding and enzyme catalysis will remain empirical and phenomenological unless verified and proven with theory. Statistical mechanics as the relevant theory remains difficult to apply due to the computational expense involved. Thus, minimal enzymes are potentially valuable aids theoretical investigations into enzyme catalysis. Leveraging its small size, we assessed our model enzyme in the equilibria concerning its folding. The similar characterization of enzyme catalysis is a formidable challenge, which may benefit from the use of minimal models.

The catalysis functions of enzymes include a role for the polypeptide main chain structure. The chain carries side chains always in L configuration. The poly-L structure constrains peptide dipoles enforcing their interactions as α–helix and β–sheet elements. The canonical β–hairpin is poly-L in structure, which defines the interaction

ISSN PRINT 2319 1775 Online 2320 7876
 Volume 11, Iss 10, 2022

fundamental to its structure. The β–hairpin has been a useful model for protein folding and misfolding and is actively being explored in its materials-design applications. The properties are remarkable for the hairpin of this size and must be considered stereochemical in their basis.

Aromatic side chains are important both in protein folding and enzyme catalysis. Interacting laterally and diagonally, aromatic side chains will lock together the stands of β–hairpin, making it an autonomous fold. We employed in our stereochemically modified hairpin aromatic residues for both its folding and substrate binding. Evidence was presented that the aromatic interactions contribute in the folding of our peptide, mainly in direct observation of the interactions with molecular dynamics.

The acid, base, and nucleophilic functional groups may be important in protein folding but they are central to enzymes as the chemical apparatus for catalysis. Ser, His and Asp side chains are the catalytic nucleophiles in the so-called proteases and were employed in the peptide.

V. CONCLUSION

The principles of sequence-based coding of protein folding were explored in a small peptide of sixteen residues. The stereochemistry of residue-level structure was explored as a novel design variable, and was shown to facilitate the design of an enzyme unusually small and active against a specific substrate. The small size of the model peptide facilitated analysis of its folding with molecular dynamics, allowing promising possibility of finding for the underlying energetics. The present study has provided a possible avenue for the design enzyme models minimal is structure and simple enough for possible theoretical studies.

ACKNOWLEDGEMENT

The author acknowledges the financial support given by UGC-WRO for this study under the project - F. - 47-200/12.

REFERENCES

Avbelj, F.; Baldwin, R. L. (2002) Role of backbone solvation in determining thermodynamic β propensities of the amino acids *Proc. Natl. Acad. Sci. U. S. A.,99*, 1309-1313.

Avbelj, F.; Moult, J. (1995) Role of electrostatic screening in determining protein main chain conformational preferences*Biochemistry*,34, 755-764.

Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., Hermans, J. (1981). Interaction Models for Water in Relation to Protein Hydration. In: Pullman, B. (eds) Intermolecular Forces. The Jerusalem Symposia on Quantum Chemistry and Biochemistry, vol 14. Springer, Dordrecht. Pp. *331-342*.

Branden, C.; Tooze, J. (2012)*Introduction to Protein Structure*, Garlan Publishing Inc., New York, USA.

Chan, H. S.; Dill, K. A. (1991) Polymer principles in protein structure and stability. *Annu. Rev. Biophys. Biomol. Chem.,20*, 447-490.

Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. (1999) Peptide folding: when simulation meets experiment.*Angew. Chem. Int. Ed., 38(1-2),* 236-240.

Dill, K. A. (1990) Dominant forces in protein folding.*Biochemistry,29*, 7133-7155.

Dill, K. A.; Truskett, T. M.; Vlachy, V.; Hribar-Lee, B. (2005) Modeling water, the hydrophobic effect and ion solvation.*Annu. Rev. Biophys. Biomol. Struct.,34*, 173-199.

Dougherty, D. A. (1996) Cation-π interactions in Chemistry and Biology: A new view of benzene, Phe, Tyr and Trp.*Science,271*, 163-168.

Kauzmann, W. (1959) Some factors in the interpretation of protein denaturation.*Adv. Prot. Chem.* **1959**, *14*, 1-63.

ISSN PRINT 2319 1775 Online 2320 7876
 Volume 11, Iss 10, 2022

Ma, J. C.; Dougherty, D. A. (1997) The cation-π interaction.*Chem. Rev.,97*, 1303-1324.

Meyer, E. A.; Castellano, R. K.; Diederich, F. (2003) Interactions with aromatic rings in chemical and biological recognition.*Angew. Chem. Int. Ed.,42*, 1210-1250.

Patel, K.; Goyal, B.; Kumar, A.; Nand Kishore; Durani, S. (2010) Cured of "Stickiness", poly-L β-hairpin is promoted with LL to DD mutation as a protein and a hydrolase mimic.*J. Phys. Chem. B,114*, 16887-16893.

Ramakrishnan, V.; Ranbhor, R.; Durani, S. (2004) Existence of "specific" folds in polyproline II ensembles of an "unfolded" alanine peptide detected by molecular dynamics.*J. Am. Chem. Soc., 126*, 16332-16333.

Ramakrishnan, V.; Ranbhor, R.; Durani, S. (2005) Simulated folding of 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*, 9314-9323.

Ranbhor, R.; Kumar, A.; Tendulkar, A.; Patel, K.; Ramakrishnan, V.; Durani, S. (2018) IDeAS: automated design tool for heterochiral protein folds.*Physical Biology,15*, 066005.

Ryckaert, J. P.; Ciccoti, G.; Berendsen, H. J. C. (1997) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comp Phys*, *23(3)*, 327-341.

van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hunenberger, P. H.; Kruger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. (1996) *Biomolecular Simulation: The GROMOS96 Manual and User Guide*, Hochschulverlag AG and the ETH Zurich: Zurich, Switzerland.