FULL LENGTH	Iranian Biomedical Journal 23 (4): 253-261 July 2019

Physicochemical Position-Dependent Properties in the Protein Secondary Structures

Ehsan Saghapour[1] and Mohammadreza Sehhati[1,2*]

[1]Department of Bioelectronic and Biomedical Engineering, School of Advanced Technologies in Medicine, Isfahan University of Medical Sciences, Isfahan, Iran; [2]Medical Image and Signal Processing Research Center, Isfahan University of Medical Sciences, Isfahan, Iran

Received 11 August 2018; revised 22 September 2018; accepted 3 October 2018

ABSTRACT

Background: Establishing theories for designing arbitrary protein structures is complicated and depends on understanding the principles for protein folding, which is affected by applied features. Computer algorithms can reach high precision and stability in computationally designed enzymes and binders by applying informative features obtained from natural structures. **Methods:** In this study, a position-specific analysis of secondary structures (α-helix, β-strand, and tight turn) was performed to reveal novel features for protein structure prediction and protein design. **Results:** Our results showed that the secondary structures in the N-termini region tend to be more compact than C-termini. Decoying periodicity in length and distribution of amino acids in α-helices is deciphered using the curve-fitting methods. Compared with α-helix, β-strands do not show distinct periodicities in length. Also, significant differences in position-dependent distribution of physicochemical properties are shown in secondary structures. **Conclusion:** Position-specific propensities in our study underline valuable parameters that could be used by researchers in the field of structural biology, particularly protein design through site-directed mutagenesis. ***DOI: 10.29252/ibj.23.4.253***

Keywords: Algorithms, Amino acids, Physicochemical, Protein structure

Corresponding Authors: Mohammadreza Sehhati
Department of Bioelectronic and Biomedical Engineering, School of Advanced Technologies in Medicine, Isfahan University of Medical Sciences, Isfahan, Post Code: 81746-73461, Iran; Tel.: (+98-31) 37923854; E-mail: mr.sehhati@gmail.com

INTRODUCTION

Understanding the relationship between position-specific properties of amino acids sequence and the secondary structure formation is vital for protein structure prediction and *de novo* protein design. The first ideas of protein structure prediction and *de novo* protein design come from very early studies on the correlation of amino acid distribution in protein structures[1-4]. It has been shown that the occurrence of amino acids in local structures, e.g. secondary structure, is position-dependent[5-9]. Recent studies have explored more details of amino acid distribution in secondary structures[5,10-15] and their functional roles[16,17]. In addition, the physicochemical bases that dictate the preference or avoidance of the amino acids for the secondary structure formation have been reported in a number of investigations[18,19]. These properties would be useful for designing algorithms to encode the molecular structures of natural proteins, which would improve the stability and precision of the resulting proteins[20]. However, the lack of comprehensive studies on position-specific evolutionary conservation and physicochemical properties of amino acids in secondary structures have motivated us to investigate these matters in the current research work.

In this study, to extract rules governing position-specific preference or avoidance of amino acids in secondary structures, an extensive analysis was performed based on position-specific distribution and conservation of amino acids in secondary structures, as

[Downloaded from ibj1.pasteur.ac.ir on 2026-10-20]
[DOR: 20.1001.1.1028852.2019.23.4.7.3]
[DOI: 10.29252/ibj.23.4.253]

well as based on the position-specific physicochemical properties of amino acids in secondary structures. This analysis was conducted on a database of secondary structure segments, including helical segments, β-strands, and tight turns (δ-turns, γ-turns, β-turns, α-turns, and π-turns). Our result introduces novel rules that govern formation and stabilization of secondary structures.

MATERIALS AND METHODS

Utilized dataset

The PDB database was culled at 25% sequence identity by PISCES webserver[21]. The structures of the selected proteins were determined via X-ray crystallography with resolution higher than 2 Å and R-factor value lower than 0.3. The sequences were excluded for proteins smaller than 40 amino acids. Additionally, we discarded the PDB files containing protein chains with chain break(s) and/or high frequency of nonstandard residues. This attempt resulted in a database containing 5362 non-redundant protein chains, corresponding to a total number of 1,197,533 amino acid residues. We used the standard method of definition of secondary structure of proteins to derive the secondary structure information from the remaining PDB files[22]. Using this database, the secondary structure information for α-helix (H), β-strand (E), and tight turn (T) were selected for further investigation. We separated tight turns based on their classification into subclasses δ-turn, γ-turn, β-turn, α-turn, and π-turn[23]. Consequently, three main subsets were formed with 34422, 63279, and 53192 sequences corresponding to α-helix, β-strand, and tight turn, respectively.

Definition of specific positions in secondary structures

Amino acid positions in secondary structures were annotated as N_i and C_i, where N and C are the N-terminus and C-terminus of the secondary structure, respectively, and *i* is the position number of the amino acid with respect to distance from reference terminus, i.e. N and C. For helices and strands, N-cap and C-cap are referred to the first residue that precedes and succeeds the helix or strand, respectively; both residues do not participate in the conformation[13,24,25]. The notation used for different secondary structures is illustrated in Figure 1.

Amino acid propensities in the secondary structure elements

In order to investigate the relationship between a specific position in a secondary structure and the amino acid residue located at this position, we defined position-specific propensity (PSP_{ij}^{SS}) as follows:

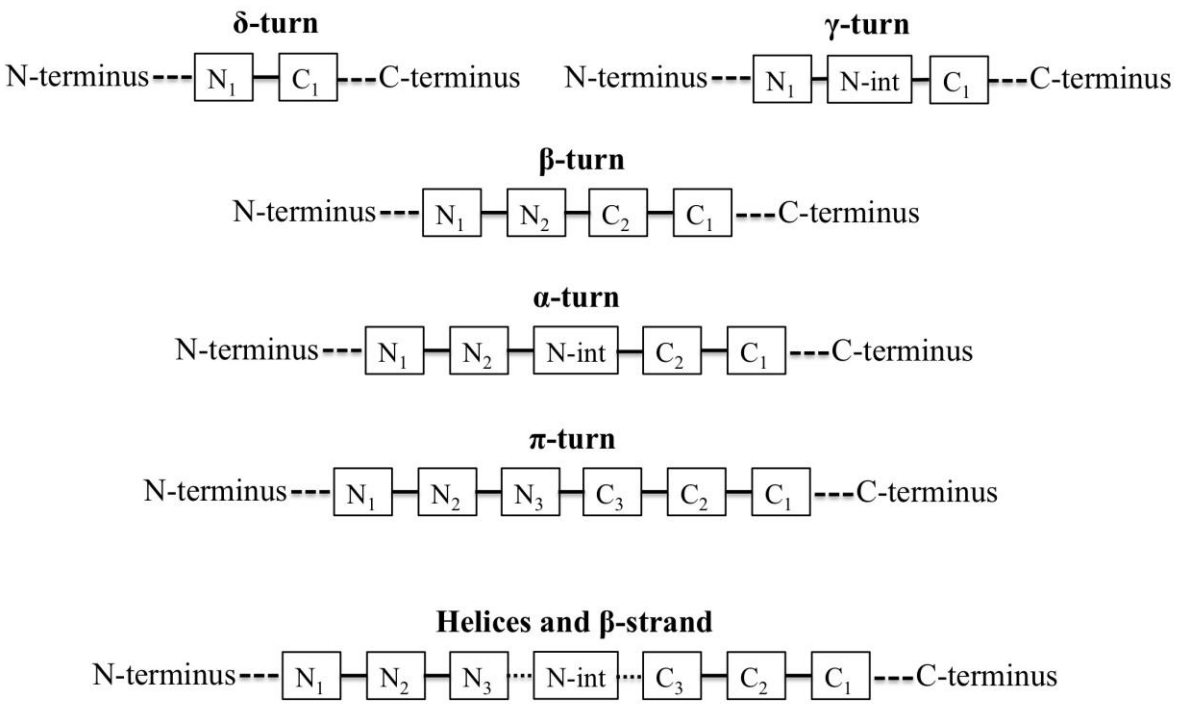

Fig. 1. A schematic representation of amino acid positioning in different secondary structures considered in this study. int, intermediate

[Downloaded from ibj1.pasteur.ac.ir on 2026-10-20]
[DOR: 20.1001.1.1028852.2019.23.4.7.3]
[DOI: 10.29252/ibj.23.4.253]

$$PSP_{ij}^{SS} = \frac{f_{ij}^{SS}}{f_i} = \frac{n_{ij}^{SS} / \sum_i n_{ij}^{SS}}{N_i / \sum_i N_i},$$

where n_{ij}^{SS} and f_{ij}^{SS} are the frequency and fraction of a given amino acid residue (type *i*) in position *j* of secondary structure, respectively. Also, N_i and f_i are the frequency and fraction of a given amino acid (type *i*) over the entire database.

Position-specific physicochemical properties of amino acids

Up to now, the majority of statistical studies of position-specific secondary structure properties have been focused on distribution of amino acid residues in the secondary structure[10,12,26,27]. In this study, we analyzed diverse evolutionary and physicochemical properties of amino acid residues in the secondary structure elements including conservation, compactness, planarity of side chains, crystal contact, B-factor, and surface accessibility. Most of these features were extracted from PDBFIND2 (ftp://ftp.cmbi.ru.nl/pub/molbio/data/pdbfinder2/), using in-house written programs.

Approximation of appropriate function for the obtained data

The Curve Fitting Toolbox of MATLAB V7.14 (R2012a) was applied to fit our data to smooth equations.

RESULTS AND DISCUSSION

In this study, the provided database includes the sample size larger than that found in the literature such as Bhattacharjee and Biswas' work[10], with only 2586 non-redundant protein chains. Also, as illustrated in Figure 1, a more comprehensive analysis was performed on all the secondary structures, including helical segments, β-strands, and tight turns (δ-turns, γ-turns, β-turns, α-turns, and π-turns) in contrast to the few limited structures that were considered in the similar studies[10-15].

α-helices

Our database includes a large number of α-helices (n = 34422). The size of this database supports the reliable analysis of length distribution of α-helices and position-specific distribution of amino acids and physicochemical propensities in α-helices.

Decaying periodicity in distribution of α-helices length

The number of different lengths of α-helices that observed in our database is illustrated in Figure 2. Only α-helices shorter than 20 amino acids occur more than 500 times. Besides, α-helices with 10 amino acids length are the most frequent. The number of residues in each helix is fitted to a smooth equation and supports previous reports[13,28]. We found that a vertically shifted Gaussian with two terms gave a much better fit than that reported earlier ($R^2 = 0.9953$)[13,28].

The weighted residuals, as shown in the inset of Figure 2, confirm preference for certain helical lengths reported by other works[13,28]. In agreement with those works, the preferences are periodic, showing ~3.6 residue periodicity. However, the weighted residuals in Figure 2 highlighted a neat two-sided decaying periodicity in the region of α-helices shorter than 20 amino acids length, which cover over 70% of our large database. The weighted residuals defined in terms of the observed frequency $f_i(O)$ and the Gaussian fit $f_i(P)$ are as follows: $WR_i = (fi(O) - f_i(P)) / f_i(P)$.

Periodicity in position-specific propensities of amino acids in α-helices

Figures 3 and 4 show the average propensities within helices for each amino acid, grouped based on the physicochemical properties of the amino acid. We examined position-specific propensities for the first 15 positions at both N-cap and C-cap in α-helices. As

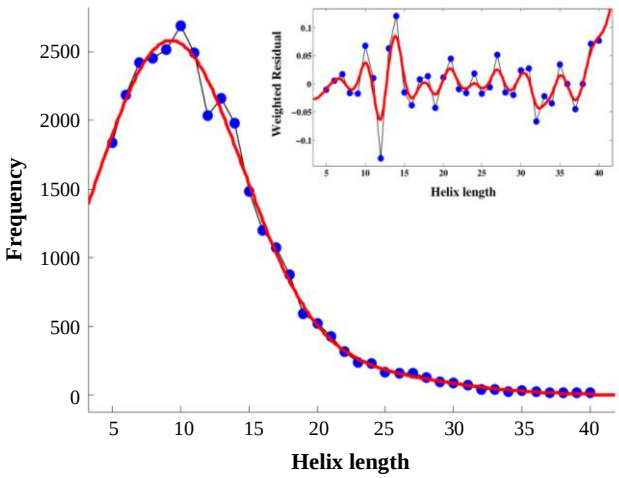

Fig. 2 The distribution of helix length in the non-redundant protein database. Periodicity with a period corresponds to a ~3.6 residue repeat is apparent. The inset plot of weighted residuals demonstrates a dramatic preference for certain helical lengths. Data are fitted to the equation:

$$f(x) = 2532 \times exp(-((x-9.195)/7.623)^2) + 174 \times exp(-((x-21.01)/10.51)^2).$$

[Downloaded from ibj1.pasteur.ac.ir on 2026-10-20]
[DOR: 20.1001.1.1028852.2019.23.4.7.3]
[DOI: 10.29252/ibj.23.4.253]

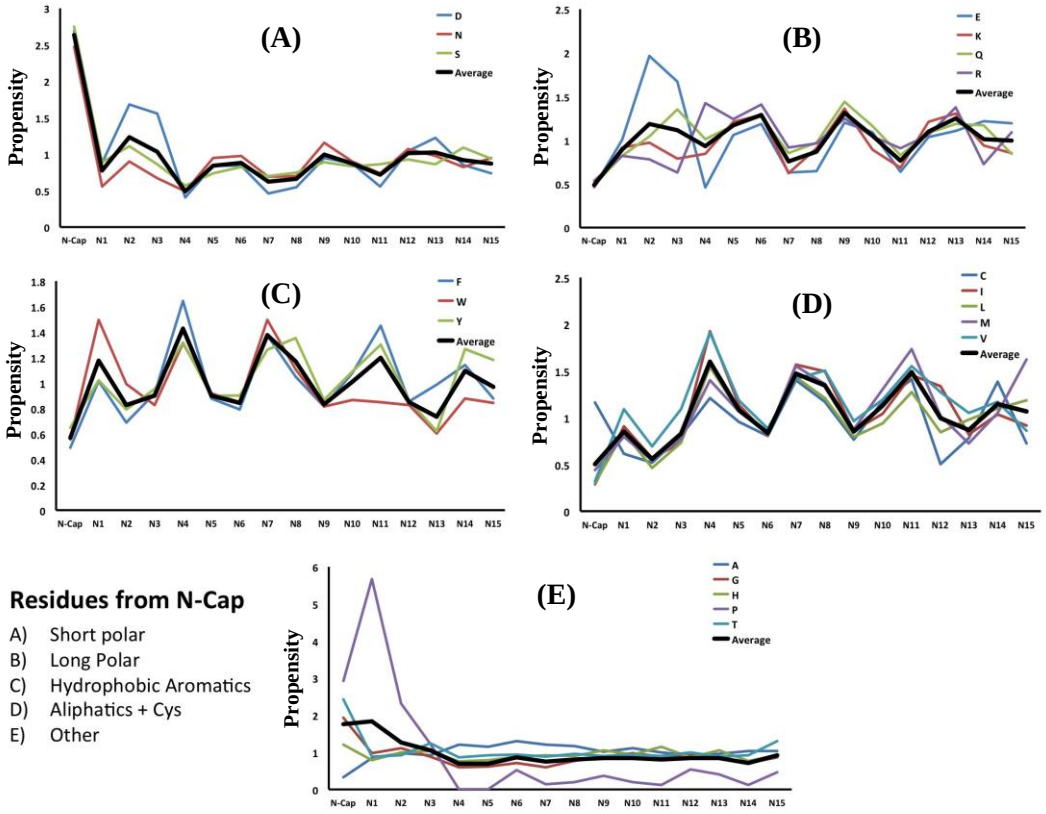

Fig. 3. Position-specific propensities for single amino acids and amino acids in different physicochemical groups in the first 15 residues located at the N-terminus of helices in different categories.

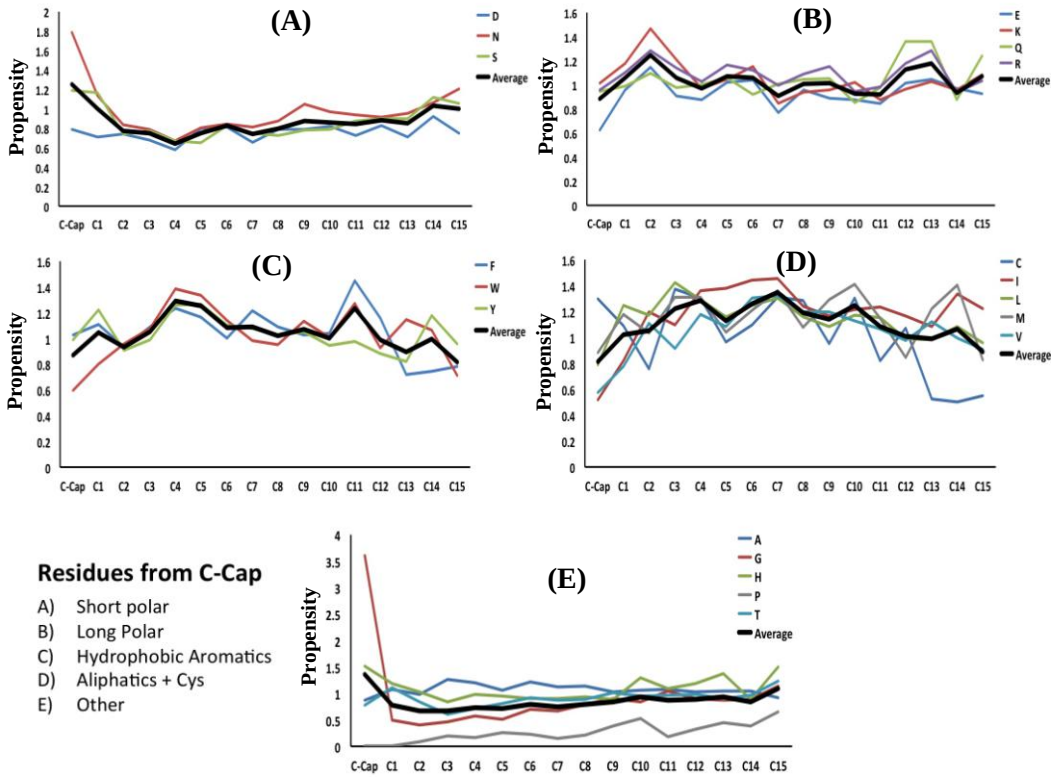

Fig. 4. Position-specific propensities for single amino acids and amino acids in different physicochemical groups in the first 15 residues located at the C-terminus of helices in different categories.

[Downloaded from ibj1.pasteur.ac.ir on 2026-10-20]
[DOR: 20.1001.1.1028852.2019.23.4.7.3]
[DOI: 10.29252/ibj.23.4.253]

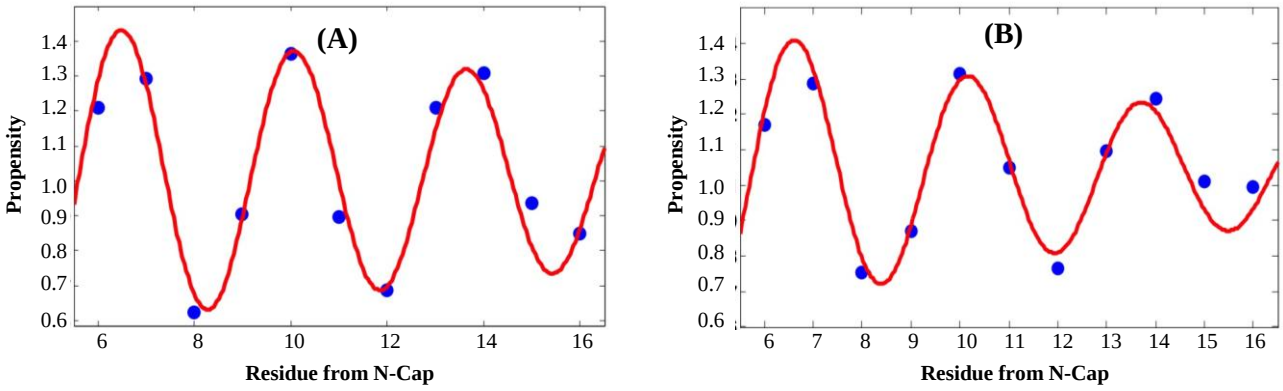

Fig. 5. Position-specific amino acid propensities (blue circles) fitted to decaying sinusoid equation (red curve). The data for the residue Lysine (A) and for long polar amino acid residues (B) in positions 6–16 prior to the N-terminus are fitted to a decaying sinusoid equation (R^2 = 0.94 and R^2 = 0.95, respectively).

depicted in these Figures, we grouped the amino acids into five categories, including short polar (Figs. 3A and 4A, D, N, S), long polar (Figs. 3B and 4B, E, K, Q, R), aromatic (Figs. 3C and 4C, F, W, Y), hydrophobic aliphatic, and Cys (Figs. 3D and 4D, C, I, L, V, M), and other residues that do not fall into any one of these categories (Figs. 3E and 4E, G, H, P, T, A). Our results, as demonstrated in Figures 3 and 4, showed position-specific independency and periodicity of the presence of amino acids in helix. Besides, in a few cases, the data for positions 6-16 were fitted to a decaying sinusoid with R^2 values over 0.9 (e.g. Fig. 5).

Position-specific physicochemical propensities in α-helices

In addition to sequence-based position-specific propensities, we have analyzed position-specific physicochemical properties, including relative side chain accessibility, B-factor, conservation, crystal contact, entropy, absolute inside/outside distribution, insertions and deletions, packing, and planarity of side chains for the first 15 positions at both N-cap and C-cap in α-helices (Figs. 6 and 7). The comparison of the curves demonstrates a periodic pattern of residue positioning regarding their physicochemical properties. This pattern is particularly observed for the 15 N-terminal residues within each helical conformation. Interestingly, the central positions of α-helices are highly conserved, and the N-terminus of α-helices is more compact compared to the C-terminus.

β-strands

Length distribution, position-dependent distribution of amino acids, and physicochemical propensities in β-strands were studied for 63279 β-strand in our database. The obtained results confirm diversity in rules in formation and stabilization of the secondary structures.

No periodicity in the distribution of β-strands length

A plot of occurrence of strands, as the function of the strand length in our database, is illustrated in Figure 8. In our database, strands with five amino acids length are the most frequent; the result is consistent with previous reports[10,29]. The occurrence level decreased sharply for *β*-strands longer than six residues.

Position-specific propensities of amino acids in β-strands

Except for both N1 and C1 positions in β-strands, the average propensities of amino acid residues in other positions show no significant fluctuation. Figure 9 shows the average propensities of amino acid residues in five selected positions, i.e. N1, N2, N-int, C1, and C2. From our results, three amino acids, including glutamine, lysine, and isoleucine demonstrated higher local and global propensity in β-strands. Also, these three amino acids avoided in N1 and C1 positions. Interestingly, some amino acid residues preferred in one or both N1 and C1 positions. For instance, leucine preferred in N1 position but not in other positions.

Position-specific physicochemical propensities in β-strands

Position-specific physicochemical properties for five different positions in β-strands, i.e. N1, N2, N-int, C1, and C2, were analyzed (Table 1). We observed that the central positions of β-strands have less insertions and deletions and the relative side chain accessibility, similar to our obtained results for α-helices. In addition, entropy constantly decreased from N- to C-termini. Interestingly, packing and B-factor show higher values in the middle and lower values on both N1 and C1 positions. However, packing in C1 position is slightly lower than N1 position. As we mentioned

[Downloaded from ibj1.pasteur.ac.ir on 2026-10-20]
[DOR: 20.1001.1.1028852.2019.23.4.7.3]
[DOI: 10.29252/ibj.23.4.253]

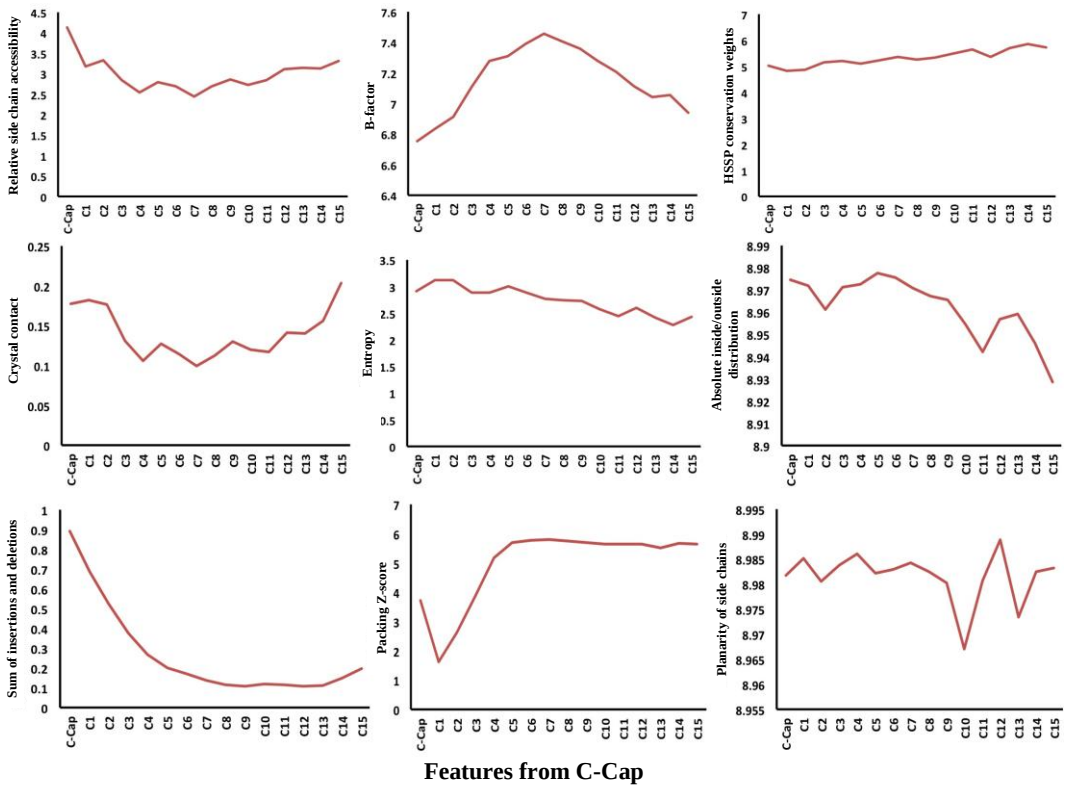

Fig. 6. Position-specific physicochemical propensities in the first 15 residues located at the C-terminus of helices.

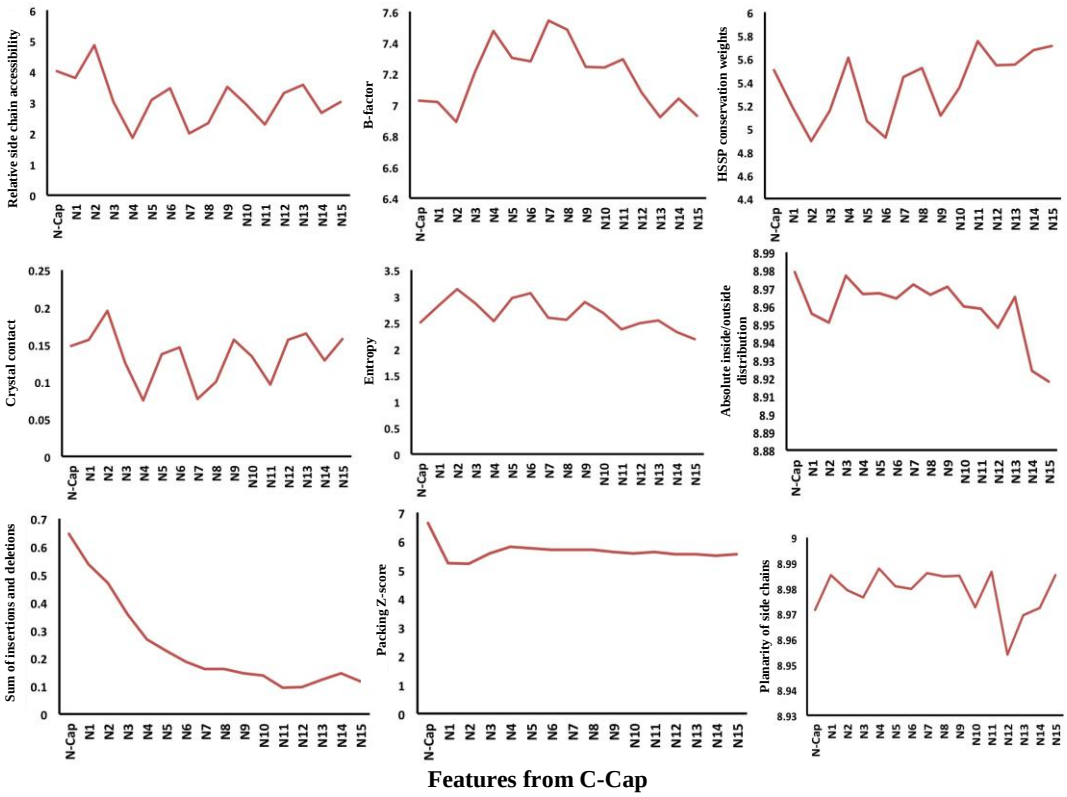

Fig. 7. Position-specific physicochemical propensities in the first 15 residues located at the N-terminus of helices.

[Downloaded from ibj1.pasteur.ac.ir on 2026-10-20]
[DOR: 20.1001.1.1028852.2019.23.4.7.3]
[DOI: 10.29252/ibj.23.4.253]

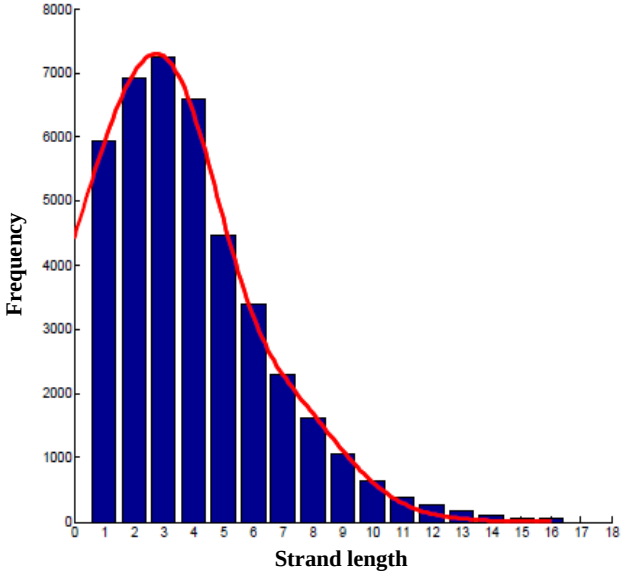

Fig. 8. The distribution of β-strand length in the non-redundant protein database.

that the distribution of amino acids in C1 and N1 is different in β-strands (Fig. 9); therefore, the difference in compactness between the termini is indeed encoded in the primary sequence. Contrary to recent works that have evaluated a limited list of properties (propensity, χ^2-values, hydrophobicity, and free energy) in β-strands[10,24], Table 1 provides a more complete picture of the secondary structures using physicochemical position-dependent properties in β-strands.

Tight turns

Physicochemical features and Position-specific propensities of amino acids were calculated for tight turns, including δ-, γ- β-, α-, and π-turns, also called as 2-, 3-, 4-, 5-, and 6-turns, respectively (Tables 2 and 3). The obtained results demonstrated significant differences in position-dependent distribution of amino acid residues and physicochemical properties in tight turns. Pattern of preference of physicochemical properties was completely different in tight turn subclasses. However, packing in C1 position was slightly lower compared to N1 position in β-, α-, and π-turns (Table 2). Significant differences observed in the distribution of amino acids confirm that the differences in compactness between the termini are indeed encoded in the primary sequence.

This work presents the most comprehensive analysis of position-dependent properties in protein secondary structures. An exhaustive study of the frequency of occurrence of individual amino acids and physico-chemical properties was carried out on a set of 34422, 63279, and 53192 sequences corresponding to α-helix, β-strand, and tight turns, respectively. The protein sample used in this study was very large, hence unbiased, giving high confidence to the obtained results, expressed in terms of global and local propensities. Some position-dependent physico-chemical features were also studied in α-helix, β-strand, and tight turns. The amount of information collected will need a further automatic analysis in order to obtain useful predictive rules. The physicochemical properties and the data concerning their individual and pair propensities generated in this work would be crucial to start the predictive modeling. With this approach, we aimed to find some general rules that can be applied to any amino acid sequence in order to predict the stability of secondary structures.

In summary, our results suggested more compactness in N-termini of α-helix, β-strand, and tight turns secondary structures compared to C-termini. We have observed decoying periodicity in position-specific propensities of amino acids in α-helices and the length of α-helices. Meanwhile, we have shown significant differences in propensities of amino acids in different positions, which could guide the the formation of secondary structures.

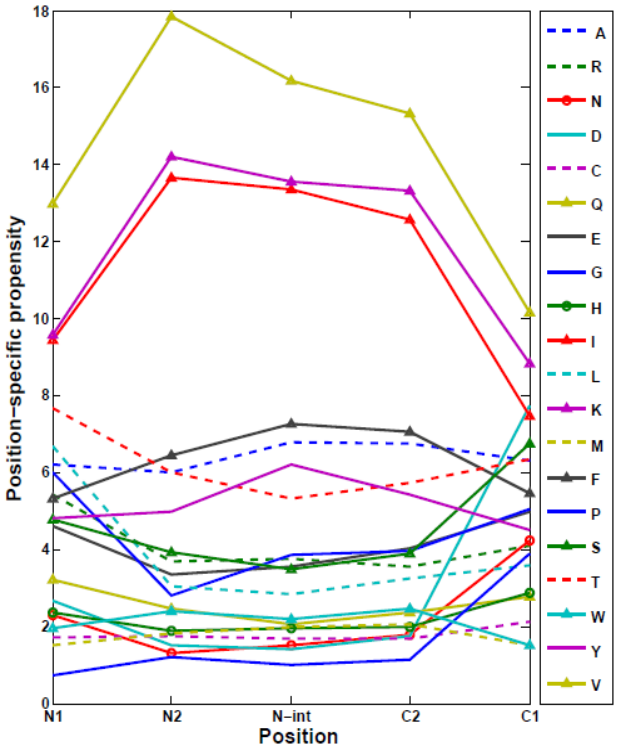

Fig. 9. Position-specific amino acid propensities in five selected positions in β-strands.

[Downloaded from ibj1.pasteur.ac.ir on 2026-10-20]
[DOR: 20.1001.1.1028852.2019.23.4.7.3]
[DOI: 10.29252/ibj.23.4.253]

Table 1. Physicochemical properties in β-strands

Physicochemical property	Position				
	N1	N2	N-int	C2	C1
Sum of insertion and deletions	0.487	0.3375	0.275	0.284	0.408
Entropy	2.764	2.552	2.498	2.457	2.268
HSSP conservation weight	5.281	5.654	5.768	5.849	5.941
Relative side chain accessibility	2.492	1.798	1.774	1.846	2.274
B-factor	7.541	7.733	7.786	7.743	7.599
Planarity of side chains	8.992	8.987	8.993	8.987	8.981
Packing Z-score	5.856	6.582	6.541	6.037	4.793
Absolute inside/outside distribution	8.982	8.989	8.984	8.986	8.981

*Each row colored separately based on related values.

Table 2. Physicochemical properties in tight turns

Feature	δ-turn		γ-turn			β-turn				α-turn					π-turn					
	N1	C1	N1	N-int	C1	N1	N2	C2	C1	N1	N2	N-int	C2	C1	N1	N2	N3	C3	C2	C1
Sum of insertion and deletions	1.036	1.042	1.006	1.161	1.092	0.937	1.045	1.047	0.951	0.962	1.059	0.927	1.104	1.057	1.011	1.026	1.401	1.359	1.036	1.031
Entropy	3.082	2.824	2.761	3.151	2.815	2.786	2.932	2.914	2.729	2.649	2.812	2.513	2.845	2.696	2.891	2.536	2.646	2.453	3.031	3.062
HSSP conservation weight	4.804	5.159	5.222	4.806	5.122	5.225	5.049	4.93	5.233	5.285	5.163	5.442	5.234	5.265	5.139	5.554	5.189	5.828	5.108	4.898
Relative side chain accessinility	4.885	6.236	4.058	5.284	4.328	3.963	4.836	4.022	5.187	3.698	4.148	3.085	4.046	5.383	4.237	5.093	2.814	2.739	4.869	5.307
B-factor	6.711	6.708	6.847	6.515	6.786	6.933	6.743	6.821	6.914	7.011	6.911	7.082	6.943	6.844	7.116	7.065	7.321	7.279	6.897	6.939
Planarity of side chains	8.981	8.974	8.989	8.977	8.983	8.988	8.973	8.985	8.965	8.992	8.994	8.989	8.966	8.994	8.995	8.991	8.981	9.154	8.995	8.972
Packing Z-score	2.844	3.075	3.472	2.371	3.262	3.573	3.039	2.952	3.452	4.062	3.575	3.685	3.009	3.106	3.632	4.139	4.163	4.177	3.014	2.962
Absolute inside/outside distribution	8.957	8.965	8.967	8.941	8.967	8.969	8.932	8.971	8.967	8.969	8.937	8.975	8.961	8.966	8.901	8.967	9.431	8.963	8.921	8.991

Table 3. Propensities of amino acids in tight turns

Amino acid	δ-turn		γ-turn			β-turn				α-turn					π-turn					
	N1	C1	N1	N-int	C1	N1	N2	C2	C1	N1	N2	N-int	C2	C1	N1	N2	N3	C3	C2	C1
A	0.097	0.035	0.093	0.075	0.039	0.085	0.075	0.095	0.04	0.084	0.087	0.084	0.122	0.027	0.104	0.05	0.12	0.071	0.124	0.041
C	0.007	0.006	0.01	0.006	0.026	0.011	0.009	0.015	0.011	0.014	0.01	0.014	0.009	0.01	0.008	0.008	0.025	0.029	0.012	0.012
D	0.069	0.123	0.046	0.093	0.07	0.046	0.069	0.068	0.076	0.051	0.087	0.06	0.062	0.08	0.033	0.1	0.058	0.017	0.095	0.083
E	0.085	0.056	0.068	0.108	0.051	0.069	0.089	0.063	0.041	0.06	0.075	0.068	0.058	0.048	0.1	0.108	0.041	0.017	0.066	0.021
F	0.022	0.023	0.035	0.025	0.038	0.038	0.044	0.037	0.04	0.046	0.041	0.045	0.036	0.031	0.046	0.021	0.041	0.133	0.046	0.033
G	0.068	0.312	0.072	0.103	0.133	0.073	0.114	0.076	0.256	0.055	0.108	0.07	0.103	0.255	0.058	0.203	0.091	0.104	0.12	0.22
H	0.022	0.028	0.02	0.03	0.028	0.027	0.031	0.023	0.034	0.024	0.031	0.041	0.019	0.027	0.021	0.037	0.012	0.017	0.029	0.037
I	0.021	0.007	0.03	0.024	0.033	0.034	0.023	0.032	0.011	0.029	0.031	0.048	0.034	0.019	0.025	0.021	0.041	0.037	0.021	0.033
K	0.074	0.047	0.063	0.084	0.061	0.061	0.062	0.077	0.069	0.058	0.06	0.05	0.087	0.058	0.041	0.025	0.062	0.066	0.116	0.046
L	0.057	0.028	0.075	0.05	0.062	0.086	0.075	0.061	0.048	0.094	0.11	0.099	0.077	0.046	0.071	0.066	0.149	0.129	0.062	0.033
M	0.011	0.006	0.012	0.009	0.01	0.016	0.011	0.014	0.014	0.005	0.017	0.017	0.017	0.009	0.017	0.012	0.025	0.025	0.004	0.021
N	0.055	0.107	0.037	0.066	0.076	0.037	0.066	0.064	0.106	0.039	0.055	0.063	0.082	0.094	0.033	0.071	0.033	0.037	0.058	0.116
P	0.16	0.013	0.153	0.039	0.005	0.145	0.042	0.018	0.021	0.17	0.017	0.012	0.043	0.015	0.137	0.041	0.004	0.05	0.025	0.021
Q	0.038	0.032	0.029	0.049	0.034	0.027	0.04	0.032	0.051	0.033	0.039	0.029	0.039	0.033	0.033	0.029	0.017	0.025	0.037	0.033
R	0.049	0.037	0.054	0.054	0.05	0.044	0.042	0.057	0.052	0.039	0.05	0.038	0.046	0.063	0.062	0.025	0.029	0.041	0.054	0.037
S	0.07	0.06	0.066	0.066	0.068	0.054	0.057	0.084	0.04	0.063	0.031	0.065	0.05	0.051	0.054	0.091	0.062	0.037	0.05	0.066
T	0.032	0.039	0.049	0.046	0.118	0.046	0.057	0.093	0.029	0.048	0.038	0.07	0.031	0.055	0.062	0.033	0.071	0.033	0.037	0.037
V	0.029	0.009	0.043	0.036	0.045	0.04	0.04	0.038	0.019	0.039	0.036	0.039	0.041	0.029	0.05	0.008	0.066	0.033	0.012	0.037
W	0.009	0.01	0.01	0.011	0.013	0.017	0.018	0.014	0.012	0.01	0.021	0.027	0.017	0.01	0.004	0.017	0.008	0.021	0.004	0.025
Y	0.022	0.023	0.029	0.025	0.039	0.039	0.035	0.033	0.027	0.034	0.048	0.057	0.027	0.033	0.037	0.033	0.037	0.066	0.025	0.037

[Downloaded from ibj1.pasteur.ac.ir on 2026-10-20]

[DOR: 20.1001.1.1028852.2019.23.4.7.3]
[DOI: 10.29252/ibj.23.4.253]

ACKNOWLEDGEMENTS

This research was financially supported by the Medical Image and Signal Processing Research Center, Isfahan University of Medical Sciences, Isfahan, Iran (Grant no. 295236).

CONFLICT OF INTEREST. None declared

REFERENCES

1. Argos P, Palau J. Amino acid distribution in protein secondary structures. *International journal of peptide and protein research* 1982; **19**(4): 380-393.
2. Levitt M. Conformational preferences of amino acids in globular proteins. *Biochemistry* 1978; **17**(20): 4277-4285.
3. Chou PY, Fasman GD. Conformational parameters for amino acids in helical, β-sheet, and random coil regions calculated from proteins. *Biochemistry* 1974; **13**(2): 211-222.
4. Davies DR. a correlation between amino acid composition and protein structure. *Journal of molecular biology* 1964; **9**: 605-609.
5. Fooks HM, Martin ACR, Woolfson DN, Sessions RB, Hutchinson EG. Amino acid pairing preferences in parallel β-sheets in proteins. *Journal of molecular biology* 2006; **356**(1): 32-44.
6. Petukhov M, Muñoz V, Yumoto N, Yoshikawa S, Serrano L. Position dependence of non-polar amino acid intrinsic helical propensities. *Journal of molecular biology* 1998; **278**(1): 279-289.
7. Petukhov M, Uegaki K, Yumoto N, Yoshikawa S, Serrano L. Position dependence of amino acid intrinsic helical propensities II: Non-charged polar residues: Ser, Thr, Asn, and Gln. *Protein science* 1999; **8**(10): 2144-2150.
8. Richardson JS, Richardson DC. Amino acid preferences for specific locations at the ends of alpha helices. *Science* 1988; **240**(4859): 1648-1652.
9. Kumar S, Bansal M. Dissecting a-helices: position-specific analysis of a-helices in globular proteins. *Proteins* 1998; **31**(4): 460-476.
10. Bhattacharjee N, Biswas P. Position-specific propensities of amino acids in the β-strand. *BMC structural biology* 2010; **10**: 29.
11. Crasto CJ, Feng J. Sequence codes for extended conformation: A neighbor-dependent sequence analysis of loops in proteins. *Proteins* 2001; **42**(3): 399-413.
12. Fonseca NA, Camacho R, Magalhaes A. Amino acid pairing at the N-and C-termini of helical segments in proteins. *Proteins* 2008; **70**(1): 188-196.
13. Engel DE, DeGrado WF. Amino acid propensities are position-dependent throughout the length of α-helices. *Journal of molecular biology* 2004; **337**(5): 1195-1205.
14. Enkhbayar P, Hikichi K, Osaki M, Kretsinger RH, Matsushima N. 310-helices in proteins are parahelices. *Proteins* 2006; **64**(3): 691-699.
15. Iqbalsyah TM, Doig AJ. Pairwise coupling in an Arg-Phe-Met triplet stabilizes α-helical peptide via shared rotamer preferences. *Journal of the American chemical society* 2005; **127**(14): 5002-5003.
16. Sato T, Kawasaki T, Mine S, Matsumura H. Functional role of the C-terminal amphipathic helix 8 of olfactory receptors and other G protein-coupled receptors. *International journal of molecular sciences* 2016; **17**(11): 1930.
17. Liou YF, Huang HL, Ho SY. A hydrophobic spine stabilizes a surface-exposed α-helix according to analysis of the solvent-accessible surface area. *BMC bioinformatics* 2016; **17**(Suppl 19): 503.
18. Avbelj F, Moult J. Role of electrostatic screening in determining protein main chain conformational preferences. *Biochemistry* 1995; **34**(3): 755-764.
19. Street AG, Mayo SL. Intrinsic β-sheet propensities result from van der Waals interactions between side chains and the local backbone. *Proceedings of the national academy of sciences* 1999; **96**(16): 9074-9076.
20. Netzer R, Fleishman SJ. Programming protein shapes and specificities by modular design. *Science* 2016; **352**(6286): 657-658.
21. Wang G, Dunbrack RL. PISCES: a protein sequence culling server. *Bioinformatics* 2003; **19**(12): 1589-91.
22. Kabsch W, Sander C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 1983; **22**(12): 2577-2637.
23. Chou K-C. Prediction of tight turns and their types in proteins. *Analytical biochemistry* 2000; **286**(1): 1-16.
24. Dai Q, Li Y, Liu X, Yao Y, Cao Y, He P. Comparison study on statistical features of predicted secondary structures for protein structural class prediction: From content to position. *BMC bBioinformatics* 2013; **14**: 152-165.
25. Duan M, Huang M, Ma C, Li L, Zhou Y. Position-specific residue preference features around the ends of helices and strands and a novel strategy for the prediction of secondary structures. *Protein science* 2008; **17**(9): 1505-15012.
26. De Sousa MM, Munteanu CR, Pazos A, Fonseca NA, Camacho R, Magalhaes A. Amino acid pair-and triplet-wise groupings in the interior of α-helical segments in proteins. *Journal of theoretical biology* 2011; **271**(1): 136-144.
27. Fujiwara K, Toda H, Ikeguchi M. Dependence of α-helical and β-sheet amino acid propensities on the overall protein fold type. *BMC structural biology* 2012; **12**: 18.
28. Penel S, Morrison RG, Mortishire-Smith RJ, Doig AJ. Periodicity in α-helix lengths and C-capping preferences. *Journal of molecular biology* 1999; **293**(5): 1211-1219.
29. Sreerama N, Venyaminov SY, Woody RW. Estimation of the number of α-helical and β-strand segments in proteins using circular dichroism spectroscopy. *Protein science* 1999; **8**(2): 370-380.

[Downloaded from ibj1.pasteur.ac.ir on 2026-10-20]
[DOR: 20.1001.1.1028852.2019.23.4.7.3]
[DOI: 10.29252/ibj.23.4.253]