PROTEINS: Structure, Function, and Bioinformatics 56:768–781 (2004)

The Occurrence of C—H…O Hydrogen Bonds in α-Helices and Helix Termini in Globular Proteins

K. Manikandan[1] and S. Ramakumar[1,2]*
[1]*Department of Physics, Indian Institute of Science, Bangalore, India*
[2]*Bioinformatics Centre, Indian Institute of Science, Bangalore, India*

ABSTRACT **A comprehensive database analysis of C—H…O hydrogen bonds in 3124 α-helices and their corresponding helix termini has been carried out from a nonredundant data set of high-resolution globular protein structures resolved at better than 2.0 Å in order to investigate their role in the helix, the important protein secondary structural element. The possible occurrence of 5 → 1 C—H…O hydrogen bond between the *i*th residue CH group and (*i* − 4)th residue C═O with C…O ≤ 3.8 Å is studied, considering as potential donors the main-chain Cα and the side-chain carbon atoms Cβ, Cγ, Cδ and Cε. Similar analysis has been carried out for 4 → 1 C—H…O hydrogen bonds, since the C—H…O hydrogen bonds found in helices are predominantly of type 5 → 1 or 4 → 1. A total of 17,367 (9310 of type 5 → 1 and 8057 of type 4 → 1) C—H…O hydrogen bonds are found to satisfy the selected criteria. The average stereochemical parameters for the data set suggest that the observed C—H…O hydrogen bonds are attractive interactions. Our analysis reveals that the Cγ and Cβ hydrogen atom(s) are frequently involved in such hydrogen bonds. A marked preference is noticed for aliphatic β-branched residue Ile to participate in 5 → 1 C—H…O hydrogen bonds involving methylene Cγ 1 atom as donor in α-helices. This may be an enthalpic compensation for the greater loss of side-chain conformational entropy for β-branched amino acids due to the constraint on side-chain torsion angle, namely, χ_1, when they occur in helices. The preference of amino acids for 4 → 1 C—H…O hydrogen bonds is found to be more for Asp, Cys, and for aromatic residues Trp, Phe, and His. Interestingly, overall propensity for C—H…O hydrogen bonds shows that a majority of the helix favoring residues such as Met, Glu, Arg, Lys, Leu, and Gln, which also have large side-chains, prefer to be involved in such types of weak attractive interactions in helices. The amino acid side-chains that participate in C—H…O interactions are found to shield the acceptor carbonyl oxygen atom from the solvent. In addition, C—H…O hydrogen bonds are present along with helix stabilizing salt bridges. A novel helix terminating interaction motif, X-Gly with Gly at C_{cap} position having 5 → 1 Cα—H…O, and a chain reversal structural motif having 1 → 5 Cα-H…O have been identified and discussed. Our analysis highlights that a multitude of local C—H…O hydrogen bonds formed by a variety of amino acid side-chains and Cα hydrogen atoms occur in helices and more so at the helix termini. It may be surmised that the main-chain Cα and the side-chain CH that participate in C—H…O hydrogen bonds collectively augment the cohesive energy and thereby contribute together with the classical N—H…O hydrogen bonds and other interactions to the overall stability of helix and therefore of proteins. Proteins 2004;56:768–781.**

Key words: C—H…O hydrogen bond; α-helix; hydrogen bonding potentials; solvent shielding; side-chain conformational entropy; helix propensity; helix capping; chain reversal motif

INTRODUCTION

The unconventional C—H…O hydrogen bond, though not strong in terms of energy, is ubiquitous, constitutes an important class among weak interactions, and is known to stabilize protein structures, protein–ligand complexes, protein– protein interfaces, DNA base pairs, protein– DNA complexes, and collagen helices.[1–11] The existence of C—H…O hydrogen bonds in crystal structures of organic molecules is known and has been studied extensively.[12–15] It has been demonstrated in our laboratory that the compaction of a de novo design helical hairpin supersecondary structure is mediated by side-chain to main-chain C—H…O interactions at the helix interface.[16] C—H…O hydrogen bonds in β-sheets,[17] proline-containing helices,[18] and capping of the C-terminus of α-helix by proline[19] have been well established. Cα—H…O hydrogen bonds have been suggested as being important determinants for stability and specificity in transmembrane helix–helix interactions[20] and their functional role in Photosystem I[21] has been discussed. Recently, 1 → 6 Cα—H…O hydrogen bond–mediated polypeptide chain reversal motif at the

The Supplementary Materials referred to in this article can be found at http://www.interscience.wiley.com/jpages/0887-3585/suppmat/index.html

Grant sponsor: CSIR (India) (fellowship to K. Manikandan).

*Correspondence to: S. Ramakumar, Department of Physics, Indian Institute of Science, Balgalore 560 012, India. E-mail: ramak@physics.iisc.ernet.in

Received 10 December 2003; Accepted 21 February 2004

Published online 11 June 2004 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.20152

C-terminus of helices in proteins has been reported.[22,23] However, no comprehensive analysis has been carried out so far about the occurrence and the role played by local C—H…O interactions in helices and helix termini of globular proteins. A number of C—H…O hydrogen bonds have been observed in helices of the ultrahigh resolution (0.89 Å) structure of a thermostable xylanase[24] (from *Thermoascus aurantiacus*) solved in our laboratory, which further encouraged us to carry out the systematic study on such interactions involving CH groups as donors in helices of globular proteins. We focus on the occurrence of local C—H…O hydrogen bonds in helices and helix termini, which are short range interactions intrinsic to the helix. A better understanding of the factors that are intrinsic to the helix should be useful in assessing their importance to protein folding and stability in relation to extrinsic factors that may be due to solvation or long-range interactions. Hence, it is of interest to investigate the possible occurrence of attractive intrahelical C—H…O interactions in helices, apart from the conventional N—H…O hydrogen bonds.

We report here a detailed analysis of intrahelical C—H…O hydrogen bonds in a data set of 3124 α-helices from 529 high-resolution (≤ 2.0 Å) and nonhomologous globular protein structures with crystallographic *R*-factor ≤ 20% from the Protein Data Bank (PDB).[25] We have carried out the analysis by exploring the possibility of involvement of all potential carbon atoms (main-chain Cα and side-chain Cβ, Cγ, Cδ, Cϵ, Cζ, and Cη) that have donatable hydrogen atoms.

Many instances of Cα, Cβ, Cγ, Cδ, and Cϵ hydrogen atom(s) systematically participating in C—H…O interactions in helices have been found and discussed. A novel helix-terminating interaction motif, with Gly at C_{cap} position and a chain reversal structural motif both additionally stabilized by Cα—H…O interactions, have been identified. Our work suggests that there is a compelling evidence for C—H…O interactions between main-chain as well as side-chain CH group of residue *i* and the carbonyl oxygen of the residues *i* − 4/*i* − 3 in helix. Finally, we discuss some energetic aspects of these interactions, C—H…O hydrogen bonds occurring concurrently with other local interactions present in the helix, a comparison of the propensity of amino acids to be involved in C—H…O interactions in helices with helical propensity, and the possible role of C—H…O interactions in helix stabilization.

MATERIALS AND METHODS

A nonhomologous (≤25% sequence identity) and high-resolution (≤2.0 Å) data set of 529 representative proteins with residual *R*-factor ≤ 20% were obtained from the PDB_SELECT[26] April 2002 list. The α-helices in the polypeptide chains were identified using the program PROMOTIF.[27] The "hgen" and "angles" programs from CCP4[28] suite were used to fix the hydrogen(s) to their parent atom and to calculate the backbone as well as side-chain torsion angles, respectively. The nomenclature

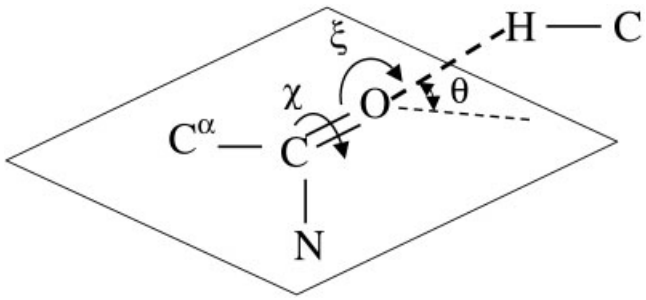

Fig. 1. Stereochemical parameters used for the identification of C—H…O hydrogen bond. The limits are C…O ≤ 3.8 Å; H…O ≤ 2.7 Å, C—H…O angle ≥ 90°, and H…O═C angle (ξ) ≥ 90°. The hydrogen-atom elevation angle (θ) is obtained from the relation,[1] sin(θ) = sin(χ) * sin(ξ), where the χ defines the torsion angle of Cα − C = O…H about the C═O bond.

for helical residues (N″-N′-N_{cap}-N1-N2 -…-C2-C1-C_{cap}-C′-C″, etc.) is as described by Aurora and Rose.[29] The selected hydrogen-bond criteria are similar to those used in an earlier analysis by Derewenda et al.[1] Briefly, they are as follows: C…O ≤ 3.8 Å, H…O ≤ 2.7 Å, the C—H…O angle ≥ 90°, and the H…O═C angle ≥ 90° (Fig. 1). The frequency distribution histograms of hydrogen-bonding parameters were "cone-corrected." In order to apply the cone correction, the frequency distributions of hydrogen-bonding parameters were weighted (multiplied) by the geometric factor, $1/r^2$ for H…O distance, 1/sinα for C—H…O angle α and 1/cosθ for hydrogen-atom elevation angle θ, where *r*, α, and θ correspond to the middle value of each distance and angle bin.[3,30] The accessible surface area of amino acids was calculated using the program NACCESS.[31]

The following formulae[32] were used to calculate the propensity and the associated standard deviation for the individual residues having 5 → 1 C—H…O hydrogen bonds in α-helices:

Propensity of *X*-type residue = $(n_X/n_T)/(N_X/N_T)$.

Standard deviation

$$\sigma_X = 1/(N_X/N_T) * \{[(n_X/n_T)\,(1 - (n_X/n_T))]/n_T\}^{1/2}.$$

X belongs to one of the 20 amino acids. n_X and n_T are the total number of 5 → 1 C—H…O hydrogen bonds by *X*-type residue and all residues, respectively. N_X and N_T are the total number of *X*-type residues and all residues, respectively, from position *N*4 of all helices in the data set. The same formulae were used for calculating the propensity and the standard deviation of individual amino acids having 4 → 1 C—H…O hydrogen bonds by considering all residues from position *N*3 in helices. The propensity of amino acids to occur in helices has been derived from the helices in the data set using the formula defined by Chou and Fasman,[32] and it has been compared with the propensity of amino acids to participate in C—H…O interactions. Methyl group C—H…O hydrogen bonds are discussed separately. The figures of molecules were drawn using MolScript.[33] In the figures, only those hydrogen atoms of carbon atoms, which are participating in C—H…O hydrogen bonds and the corresponding helical segments are shown.

TABLE I. The Average Stereochemical Parameters and Corresponding RMSDs for intrahelical C—H…O Hydrogen Bonds

H Bonds	H…O (Å)	C…O (Å)	C—H…O (°)	H…O═C (°)	H Elevation (°)
4 → 1	2.56	3.40	143	124	31
C—H…O's	(0.12)	(0.13)	(10)	(16)	(13)
5 → 1	2.50	3.30	140	138	34
C—H…O's	(0.15)	(0.15)	(15)	(12)	(13)
4 → 1 & 5 → 1	2.53	3.35	141	131	33
C—H…O's	(0.14)	(0.15)	(13)	(15)	(13)
5 → 1	2.05	2.98	156	148	26
N—H…O's	(0.19)	(0.14)	(12)	(10)	(12)
4 → 1 & 5 → 1	2.52	3.35	142	132	34
Methyl C—H…O's	(0.15)	(0.15)	(13)	(15)	(13)
Aromatic ring[a]	2.46	3.29	143	126	38
C—H…O's	(0.16)	(0.16)	(16)	(15)	(18)
Aliphatic[b]	2.54	3.36	141	130	31
C—H…O's	(0.14)	(0.15)	(12)	(16)	(11)

[a]Aromatic ring CH groups of His, Phe, Trp, and Tyr

[b]Aliphatic side-chain CH groups: *Arg*, *Asn*, *Asp*, *Gln*, *Glu*, *His*, *Ile*, *Leu*, *Lys*, *Met*, *Phe*, *Trp*, *Tyr*, *and Val* *C*β; *Arg*, *Gln*, *Glu*, *and Lys* *C*γ; *Lys* *C*δ

RESULTS

The Geometry of C—H…O Hydrogen Bonds in α-Helices

A total of 9310 and 8057 C—H…O hydrogen bonds of type 5 → 1 and 4 → 1, respectively, are found to satisfy the chosen criteria (Fig. 1) in the data set containing 3124 α-helices from 529 independent globular protein structures. The mean values and the calculated standard deviations of H…O and C…O distances, C—H…O, H…O═C and the hydrogen-atom elevation angles for 4 → 1 and 5 → 1 C—H…O hydrogen bonds are given in Table I.

It is known that the hydrogen-bond geometry of conventional intrahelical N—H…O hydrogen bonds deviates from ideal values because of the constrained helix geometry.[34] In particular, the hydrogen-atom elevation angle (or out-of-plane angle, as in Baker and Hubbard[34]), the angle θ between the peptide plane and the O…H vector (Fig. 1), assumes approximately 30° for the backbone amide hydrogen atom instead of assuming the ideal value of 0°. The average elevation angles for 4 → 1 and 5 → 1 C—H…O hydrogen bonds are found to be 31° and 34°, respectively, and compare favorably with those observed for the intrahelical 5 → 1 N—H…O hydrogen bonds (26°) in α-helices (Table I). In the data set, about 40% of 4 → 1, 35% of 5 → 1, and 92% of 4 → 1, 83% of 5 → 1 C—H…O hydrogen bonds are present, with the elevation angle being less than 30° and 45°, respectively. The frequency distributions of "cone-corrected" hydrogen-bonding parameters are depicted in Figure 2. The frequency distribution of the C—H…O angles is expected to be isotropic if the C—H…O contacts are merely van der Waals contacts. However, the observed distribution of C—H…O angles in the data set [Fig. 2(b)] is not isotropic, as may be expected for hydrogen bonds, which are inherently directional in nature.[30] Thus, it may be reasonably assumed that the observed C—H…O contacts are hydrogen bonds and collectively contribute cohesive energy for the stabilization of helices and their termini.

CH Groups Involved in 5 → 1 and 4 → 1 C—H…O Hydrogen Bonds

The C—H…O hydrogen bonds found in helices are predominantly of type 5 → 1 or 4 → 1 (Table II). Better than 50% (9310 out of 17,367) of the total number of C—H…O hydrogen bonds identified in helices are of 5 → 1 type as the helix geometry brings the residues i and $i - 4$ closer together. As may be seen (Table II), the Cγ hydrogen atom is involved in more 5 → 1 side-chain to backbone C—H…O interactions (63%; constituting a 16-member hydrogen-bonded ring) than other types of carbon atoms. Only fewer interactions (4%) are possible for the relatively more acidic Cα hydrogen atom(s), because the helical (φ, ψ) values limit such interactions, unlike in β-sheets, where the extended conformation facilitates the Cα—H…O interactions between the adjacent β-strands, as reported earlier.[17] The β-carbon atom is positioned by the tetrahedral angle at Cα and is a little away from the helix backbone, resulting in the identification of Cβ—H…O hydrogen bonds (23%), more so than Cδ—H…O hydrogen bonds (10%) (Table II). In particular, the side-chain δ-carbon atom(s) of the aromatic residues (Phe, Tyr, His, and Trp) are involved in a number of 5 → 1 Cδ—H…O hydrogen bonds. The constraint on side-chain torsion angle by the helix backbone favors 5 → 1 Cγ—H…O interactions, as discussed later in this article.

4 → 1 C—H…O hydrogen bonds are observed between the residues i and $i - 3$ in α-helices. This defines a C—H…O hydrogen-bonded, 12-membered side-chain to main-chain ring in helices. The 8057 interactions that fall into this type comprise 45% of the total (Tables I and II).

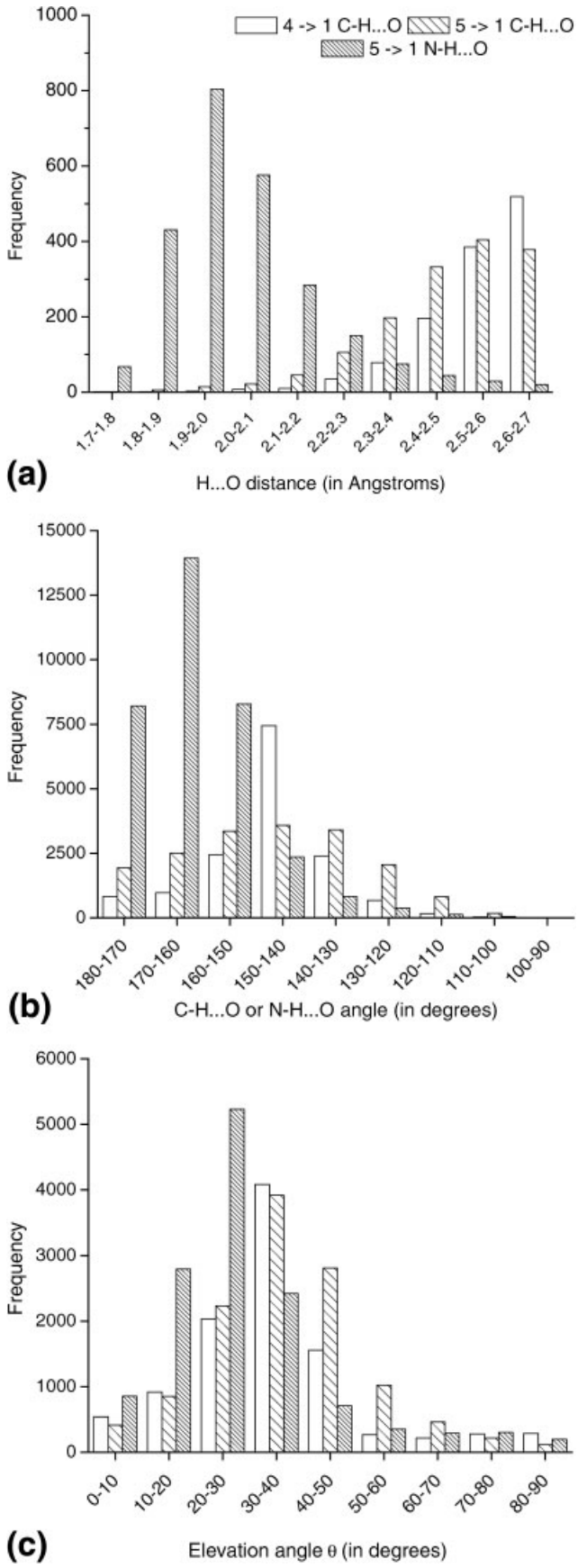

Fig. 2. Histograms show "cone-corrected" frequency distribution of the hydrogen-bond geometrical parameters for 4 → 1 and 5 → 1 C—H…O interactions: (**a**) H…O distance; (**b**) C—H…O angle; and (**c**) hydrogen-atom elevation angle. For comparison, helical 5 → 1 N—H…O geometry calculated with our data set is also shown.

From Table II, it can be seen clearly that more β-carbon atoms are primarily (88%) involved in such 4 → 1 C—H…O hydrogen bonds in the helix than other type carbon donors. This may be attributed to the fact that helix geometry spatially positions the Cβ hydrogen atom(s) of residue i closer to residue $i - 3$ C═O group than that of residue $i - 4$. The percentages of Cγ and Cδ hydrogen atom(s) that take part in 4 → 1 C—H…O interaction are relatively less, being 6% and 4%, respectively.

Propensities of Amino Acids

Based on the possibility of the carbon atom under consideration having donatable hydrogen atom(s) and participating in local C—H…O hydrogen bonds in helices, the propensities and their standard deviations of individual amino acids have been calculated for all residues (except Gly) using the formulae defined in the Materials and Methods section. The propensities of amino acids are found to be significant in comparison with their corresponding standard deviations. It is evident (Fig. 3) that Pro has the highest propensity for both 4 → 1 and 5 → 1 C—H…O hydrogen bonds; their occurrence in helices has been well documented by others[18,19] and therefore is not discussed here.

Figure 3(a) illustrates the propensity of individual residues to form 4 → 1 C—H…O hydrogen bonds in α-helices. Trp, Asp, Cys, Phe, and His (in descending order) have relatively higher propensity and notably, in these, except Cys, the donating carbon atom, Cβ, is covalently bonded to an sp^2 hybridized carbon atom that is comparatively more acidic than an sp^3 hybridized carbon atom.

The propensity values for amino acids involved in 5 → 1 C—H…O interactions are depicted in Figure 3(b), where it may be seen that aliphatic β-branched residue Ile (also Val, if methyl CH group hydrogen bonds are considered) has a higher propensity for 5 → 1 C—H…O interactions by donating the methylene Cγ 1 hydrogen atom(s). The residues, such as Met, Gln, Glu, Leu, Arg, and Lys, in decreasing order (with propensities from 1.54 to 1.13), have a strong preference to form 5 → 1 type C—H…O hydrogen bonds.

Preferred Side-Chain Conformation Observed for CH Donors

The side-chain rotameric states of the residues in the helix are highly restricted due to the constraint imposed by the helix backbone.[36] In particular, the g^- ($gauche^-$ refers to +60°) rotameric state for the torsion angle about Cα—Cβ bond is seen relatively less in helical residues because of steric hindrance and g^+ ($gauche^+$ refers to −60°) state is most preferred.[36,37] [The IUPAC-IUB convention,[38,39] according to which g^- refers to −60° and g^+ refers to +60°, is not used here.] The restriction imposed

TABLE II. Distribution of Intrahelical C—H…O Hydrogen Bonds Identified in 3124 α-Helices in the Data Set

Amino	Cα—H…O		Cβ—H…O		Cγ—H…O		Cδ—H…O		Cε—H…O		No. of	No. of
Acid	4→1	5→1	4→1	5→1	4→1	5→1	4→1	5→1	4→1	5→1	C—H…Os	Amino acids[a]
Ala	2	12	**847**[b]	**239**							1100	3500
Arg	7	14	466	172	65	618	19	50			1411	1692
Asn	5	14	302	235							556	1161
Asp	2	7	556	144							709	1158
Cys	1	1	174	61							237	399
Gln	2	19	507	114	67	674					1383	1443
Glu	4	16	651	154	62	866					1753	1971
Gly	6	125									131	1212
His	0	11	235	101			12	96	15	3	473	638
Ile	1	6	437	57	71 (**83**)	1123 (**70**)	**4**	**44**			1896	1762
Leu	11	12	920	273	56	1366	**24**	**113**			2775	3500
Lys	5	27	604	174	72	694	12	39	24	32	1683	2057
Met	4	5	279	46	38	465			**5**	**10**	852	835
Phe	3	3	428	116			40	300	9	6	905	1162
Pro	2	72	2	14	31	39	184	103			447	116
Ser	5	23	337	180							545	1325
Thr[c]	5	13	302	53	**47**	**255**					675	1255
Trp	0	2	175	49			24	56	10	10	326	422
Tyr	5	11	346	115			46	249	8	4	784	1049
Val	6	8	395	58	**200**	**1389**					2056	1799
	76	401	7963	2355	792	7559	365	1050	71	65	20,697	28,456
Total	477 (2.3, 2.8)[d]		10,318 (49.8, 57.4)		8351 (40.3, 58.7)		1415 (6.8, 19.8)		136 (0.65, 6.2)			

[a]Calculated by considering residues from N4 position of helices.
[b]Values in bold indicate C—H…O hydrogen bonds by methyl CH group.
[c]Though Thr is a β-branched residue, loss of side-chain conformational entropy can be compensated through hydrogen bond involving Oγ atom in helices as discussed by Eswar and Ramakrishnan.[35]
[d]The first value in the parentheses corresponds to the percentage occurrence of a given type of C—H…O with respect to all the C—H…O in helices. The next value stands for the percentage occurrence of a given type of C—H…O hydrogen bonds with respect to total number of the same type interaction found in all proteins in the data set, and it is more for Cβ—H…O and Cγ—H…O hydrogen bonds. There are a few occurrences of 3 → 1, 1 → 3, 6 → 1, and 1 → 6 C—H…O hydrogen bonds in helical region and also at N and C termini. The C—H…O interactions involving side-chain Cζ and Cη group hydrogens occur rarely. The details are not discussed.

for g^- and *t* states (*t* stands for trans or 180°) is especially severe for β-branched residues, since their Cγ atom bumps into the backbone in those side-chain conformational states.[37,40] The side-chain torsion angles χ_1, χ_2 for residues donating the Cγ hydrogen atom in helices are predominantly g^+ and *t* rotameric state, respectively [Fig. 4(a)], in which the Cγ atom and hence its hydrogen atom(s) are favorably positioned to participate in 5 → 1 Cγ—H…O interactions. Similarly, β-carbon hydrogen atoms are also well placed to form 4 → 1 type C—H…O hydrogen bonds in helices; as is evident from Figure 4(b), the side-chain rotamer is populated around the g^+ rotameric state. The preferred rotameric states of the side-chains in helices are compatible with the formation of local C—H…O hydrogen bonds.

C—H…O Hydrogen Bond in Helix-Stabilizing Interactions

Apart from the backbone N—H…O hydrogen bonds, numerous other factors contribute to helix stability, including solvent shielding of the backbone by side-chains, side-chain to side-chain noncovalent interactions, electrostatic interactions between charged side-chains, which are in opposition to the helix dipole, and helix-capping interactions. An attempt is made in this study to examine some of the aforementioned factors in view of the observed C—H…O hydrogen bonds in helices.

C—H…O hydrogen bond and solvent shielding

The stabilization of the α-helix by side-chain shielding of backbone polar atoms is known[41] and is illustrated here with two examples in the context of C—H…O hydrogen bonds. Two 5 → 1 Cγ—H…O hydrogen bonds, Glu 218 Cγ—Gly 214 O and Met 219 Cγ—Asp 215 O, in an amphipathic helix (residues 212–219 and sequence RVGDLCEM) of the enzyme λ integrase[42] (1ae9A), are considered. Glu 218 Ala mutation carried out in silico shows an increase in the solvent-accessible surface area of the Gly 214 carbonyl oxygen from 1.69 to 6.52 Å^2. For the Met 219 Ala mutation, also carried out in silico, the solvent-accessible surface area of the Asp 215 carbonyl oxygen increases from 1.22 to 4.21 Å^2 (similar trends were observed for other side-chains also). There is a loss of C—H…O hydrogen bond and the backbone carbonyl oxygen is solvent-exposed when the side-chain is mutated to Ala residue. It indicates that the side-chains, which are C—H…O hydrogen-bonded to backbone carbonyl oxygen atoms, make the carbonyl oxygen atom less accessible to the solvent and wrap the main-chain polar groups of the

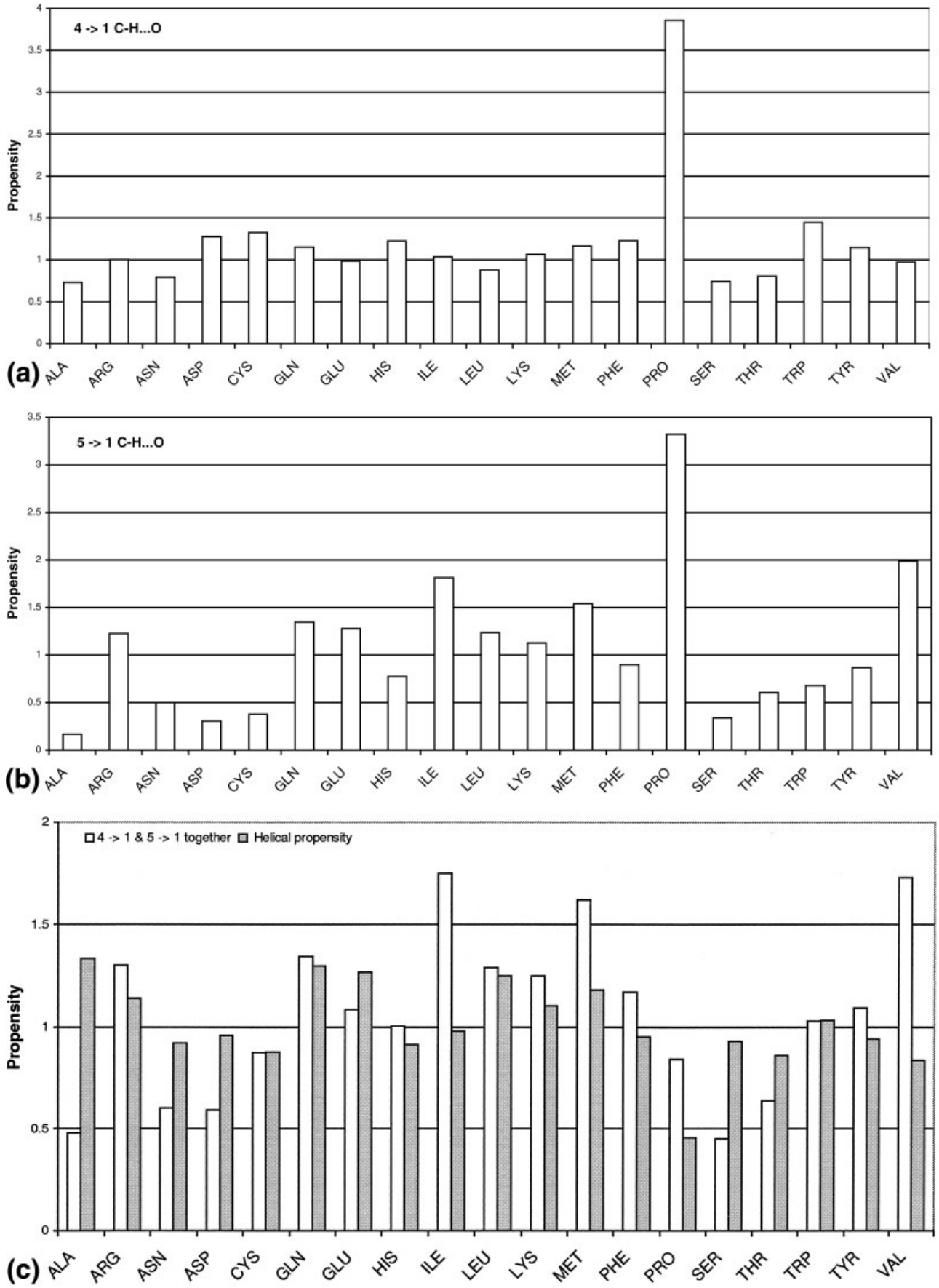

Fig. 3. Plot of propensities for amino acids participating in (**a**) 4 → 1 and (**b**) 5 → 1 C—H…O hydrogen bonds in α-helices. (**c**) The overall propensity to participate either in 4 → 1 or 5 → 1 C—H…O hydrogen bonds and Chou–Fasman type helical propensity of amino acids derived from the data set are shown.

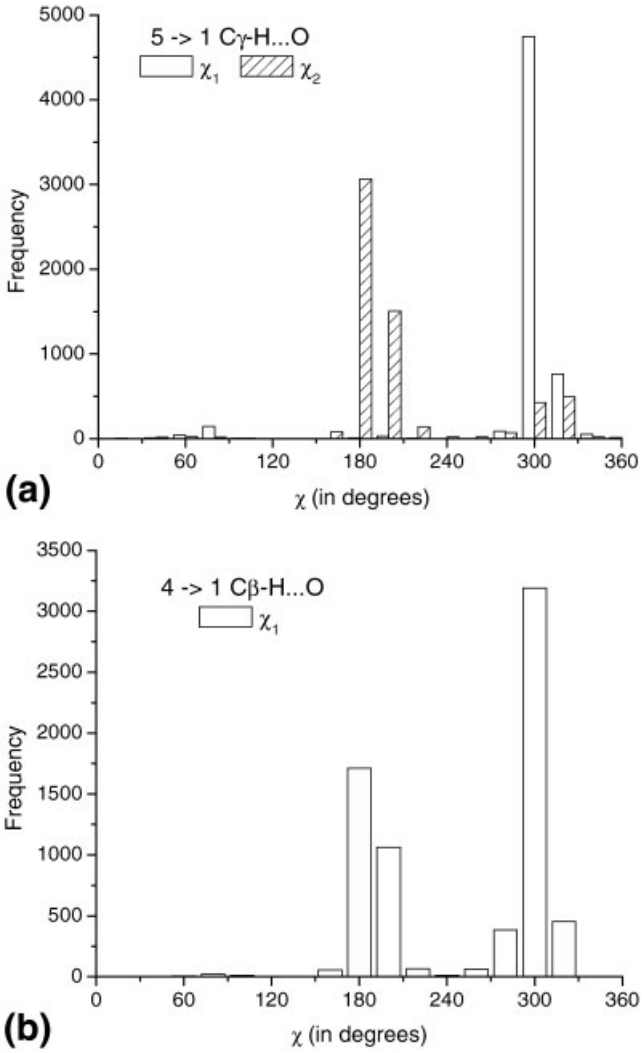

Fig. 4. Plot of frequency distribution of the side-chain torsion angle: (**a**) χ_1 and χ_2 for residues involved in 5 → 1 Cγ—H…O; (**b**) χ_1 for 4 → 1 Cβ—H…O hydrogen bonds.

solvent-exposed helix face against the destabilizing attack by water molecules.[43,44]

C—H…O hydrogen bond and van der Waals interactions

A study of side-chain–side-chain noncovalent van der Waals interactions between residues in helices has shown that the frequency of the occurrence of nonpolar amino acid pairs (LL, VL, and IL) at position $i, i-4$ is higher, and the next highest occurrence is of the charged residue pairs (KD and KE).[36] The calculation of frequency of amino acid pairs that have side-chain to backbone 5 → 1 Cγ—H…O hydrogen bonds between them in our data set showed very similar results, and the top 6 frequency pairs are listed in Table III. C—H…O hydrogen bonds may additionally enhance the van der Waals interactions in a concerted manner by bringing the side-chains of nonpolar residues closer together.

TABLE III. Frequency of Top 6 amino Acid Pairs in Helices (at position $i, i-4$) with 5 → 1 Cγ—H…O Hydrogen Bond

Amino Acid Pairs	Number of Pairs
Leu–Val	307
Leu–Ile	297
Glu–Lys	226
Leu–Leu	212
Val–Ile	188
Glu–Arg	163

C—H…O hydrogen bond may bring the paired residues closer together in helices, thereby improving the van der Waals interactions in the case of apolar pairs and electrostatic interactions in the case of charged pairs.

C—H…O hydrogen bond in salt bridges

Charged residues can stabilize the helix by the formation of salt bridge between them,[45] and such local interactions are shown to prevent the solvation of helix backbone polar groups (both C═O and N—H).[41] In as many as 132 examples, the C—H…O hydrogen bonds are also present along with ionic interactions, which are between charged residues, with their polarity being opposite to that of the helix dipole. While the atoms in the polar part of the side-chains are involved in salt bridge formation, the atoms in the aliphatic part of the side-chains, which are in the vicinity of polar atoms, participate in C—H…O interaction with the main-chain carbonyl oxygen atoms (Fig. 5). Hence, the C—H…O hydrogen bonds contribute additional cohesive energy and synergistically with salt bridges help to stabilize the helix.

C—H…O Hydrogen Bonds at the Helix Termini

Helices with minimum length of 12 residues have been considered to analyze the role of the C—H…O hydrogen bond at the helix termini. From a total of 3124 helices in the data set, 1762 such helices were pooled out. The choice of long helices (12 or more amino acids) helps to define the N_{cap} and C_{cap} regions clearly, without any overlap between them. The helix was dissected as N_{cap} to N5, and C5 to $C^{3'}$ as N- and C-terminal regions, respectively.[29]

The number of 4 → 1 C—H…O hydrogen bonds for N- and C-terminal regions of helices is found to be 1544 (25%) and 2473 (40%), respectively. The corresponding values for 5 → 1 C—H…O hydrogen bonds are 1189 (14%) and 4377 (51%), respectively. This indicates that C—H…O hydrogen bonds occur largely at the C-terminal region of helices, where the hydrogen bond accepting potential of C═O groups is less well satisfied compared to the C═O groups at the N-terminal region. Among 5 → 1 and 4 → 1 types, 5 → 1 C—H…O hydrogen bonds are found more at the C-termini. Overall, 65% of total C—H…O interactions are present in the helix terminal region. It is important to note (Table II) that nearly 75% of His C—H…O hydrogen bonds involving highly acidic Cϵ1 and Cδ2 groups were found to be at the C-terminal end of helices, as His has a higher propensity to occur at the helix C-terminal (Fig. 6).[29] In this context, the functional role and importance of His Cϵ1—H…O in the active sites of serine hydrolases may be

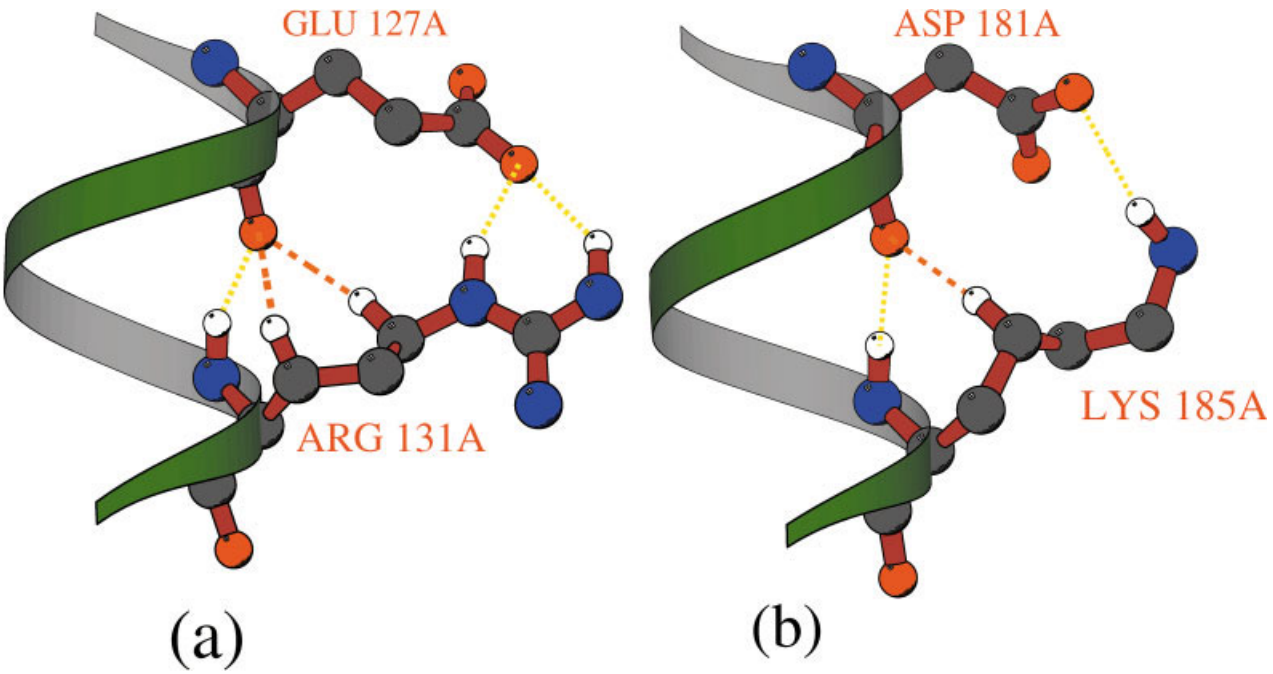

Fig. 5. Examples of C—H…O hydrogen bonds occurring along with helix-stabilizing salt bridges. (**a**) N-terminal salt bridge in the protein cystathionine gamma synthase (1cs1A).[46] (**b**) C-terminal salt bridge in the protein surface antigen PSAA (1pszA).[47] The conventional N—H…O hydrogen bonds and the C—H…O interactions are differentiated by different line types.

recalled.[50] The pronounced occurrence of C—H…O hydrogen bonds at helix termini may be due to the slight distortion in helical backbone torsion angle, and it appears that they have been recruited to provide additional stabilizing interactions in the capping regions.

Cα—H…O Hydrogen Bonds in Helices

Cα—H…O hydrogen bonds in the helix-terminating motif

Gly and Pro are known to occur at C-termini of helices as helix stop signals.[29,51] It has been observed earlier in an analysis of C—H…O hydrogen bonds in β-sheets[17] that Gly is not involved in Cα—H…O interactions despite Gly occurring recurrently in β-sheets. In contrast, in helices, the Cα—H…O hydrogen bond is present along with conventional N—H…O interactions with Gly as donor (Fig. 7).

We have identified 5 → 1 Cα—H…O hydrogen bonds occurring in the helix-terminating interaction motif (Supplementary Table I), X-Gly [57 examples; Fig. 7(a)] and X-non-Gly [43 examples; Fig. 7(b)], with Cα donor Gly/non-Gly at C_{cap} position and the acceptor carbonyl oxygen at C4 position in the helix. Of 57 observed examples (Supplementary Table I), in 20 cases Gly assumes positive ϕ, and in the remaining cases it takes an α_L conformation with the preceding residue in the extended region of the Ramachandran (ϕ, ψ) map.[54,55] In the case of helix termination by non-Gly residues[55] that assume α_L conformation, the backbone Cα—H…O hydrogen bond presumably compensates for ϕ being positive, since it is in an energetically less favored region of the Ramachandran (ϕ, ψ) map for non-Gly residues [Fig. 7(b)].

1 → 5 Cα—H…O hydrogen bond in a chain-reversal motif

We have found 1 → 5 Cα—H…O hydrogen bonds in part of the helix-stopping Schellman motif[56] (56 examples; Supplementary Table II) (Fig. 8). In all these examples, Gly precedes the acceptor and adopts either positive ϕ or α_L conformation. The (ϕ, ψ) values of the residues in the segment, which succeed the acceptor, cluster either in the extended or in the helical region of the Ramachandran (ϕ, ψ) map (Fig. 9), suggesting that the Cα—H…O hydrogen bond stabilizes the polypeptide chain reversal with the succeeding segment, being either an α-helix or a β-strand. The 1 → 5 Cα—H…O defines a 16-atom backbone hydrogen-bonded ring. It is a novel motif, since it is a tighter reverse turn with one residue less, as compared to a similar motif reported by Babu et al.,[22] where it is a 1 → 6 Cα—H…O interaction. Nonetheless, in both cases, the local Cα—H…O hydrogen bond helps to register secondary structural elements, α-helix and β-strand, and further stabilizes the Schellman motif.

C—H…O Hydrogen Bonds by Methyl Group

The methyl CH group is considered to be a weakly activated donor. However, the contribution to the cohesive energy of C—H…O hydrogen bonds, where the hydrogen is from a weakly polarized methyl group, cannot be overlooked.[13] It has been found that the methyl hydrogen atom(s) of thymine bases of DNA participate in C—H…O hydrogen bonds mediating protein–DNA interactions.[9] Recently, a theoretical computational study[58] on the complex of 1,2-ethanediol-dimethyl sulfoxide provides evi-

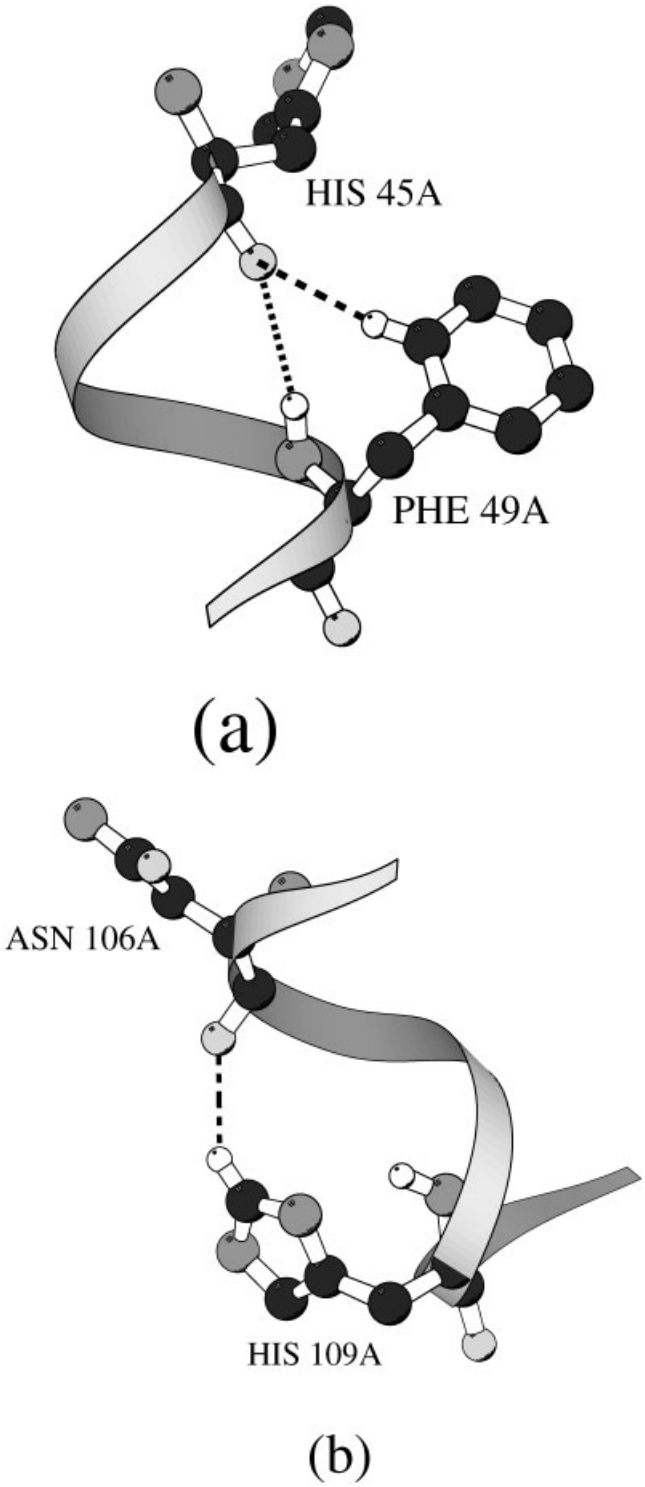

Fig. 6. Representative examples for C—H…O hydrogen bonds involving aromatic residues side-chain CH group, which is relatively more acidic than its aliphatic counterpart. (**a**) A 5 → 1 C—H…O hydrogen bond between Phe Cδ hydrogen atom and main-chain carbonyl [as in the protein human adenosine kinase (1bx4A)[48]]. (**b**) A 4 → 1 C—H…O hydrogen bond involving His Cε1 hydrogen atom at the distorted C-terminal of a helix of the protein Fmdv leader protease (1qmyA).[49]

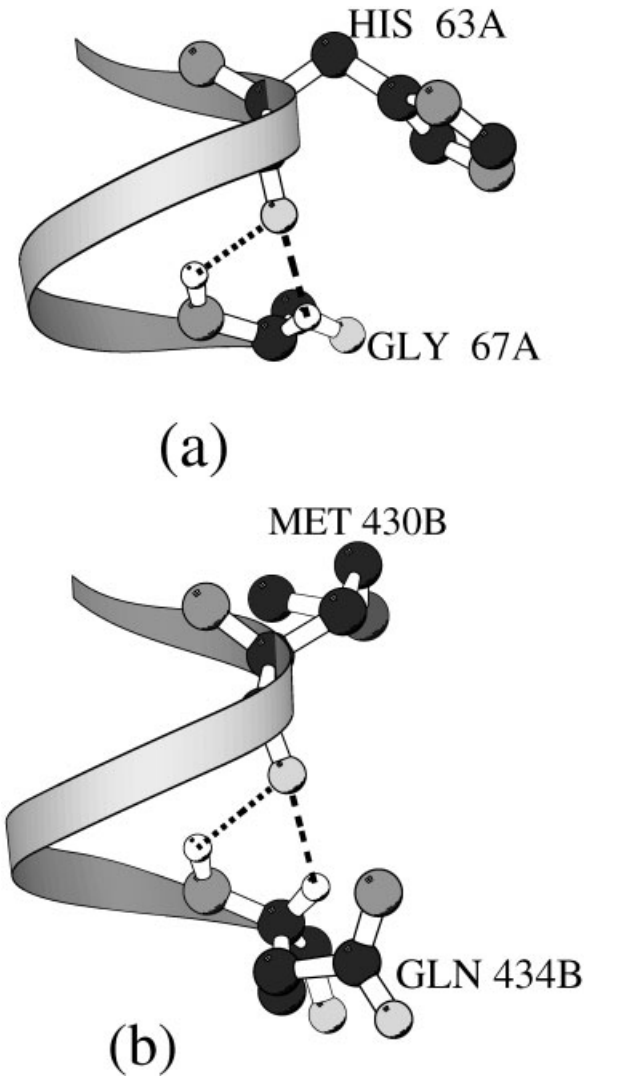

Fig. 7. Representative examples for helix termination by (**a**) Gly (as in Met apo-repressor, 1cmbA)[52] and (**b**) Gln (as in nitrogenase, 1qh8),[53] where 5 → 1 Cα—H…O hydrogen bond occurs concomitantly with conventional 5 → 1 N—H…O hydrogen bond. The donor residue adopts α_L conformation in both cases.

dence that methyl groups are capable donators of hydrogen bonds. It is also suggested that the weak C—H…O hydrogen bonds involving methyl groups exist cooperatively with the relatively strong N—H…O hydrogen bonds.[58]

We have identified a total of 3330 methyl C—H…O interactions with specified hydrogen-bond geometry (Tables I and II). The average C—H…O angle (142°) compares well with that (137°) found for small-molecule crystal structures analysis of CH_2–CH_3…O═C interactions (Table I).[30] Of these interactions, 85% and 35% have the elevation angle less than 45° and 30°, respectively.

The methyl groups of Val and Ala are found to make 48% and 33%, respectively, of total methyl C—H…O hydrogen bonds. It can be seen clearly (Table II) that the β-branched Val (for which the loss of side-chain conformational entropy is more when it is in the helix) overwhelmingly

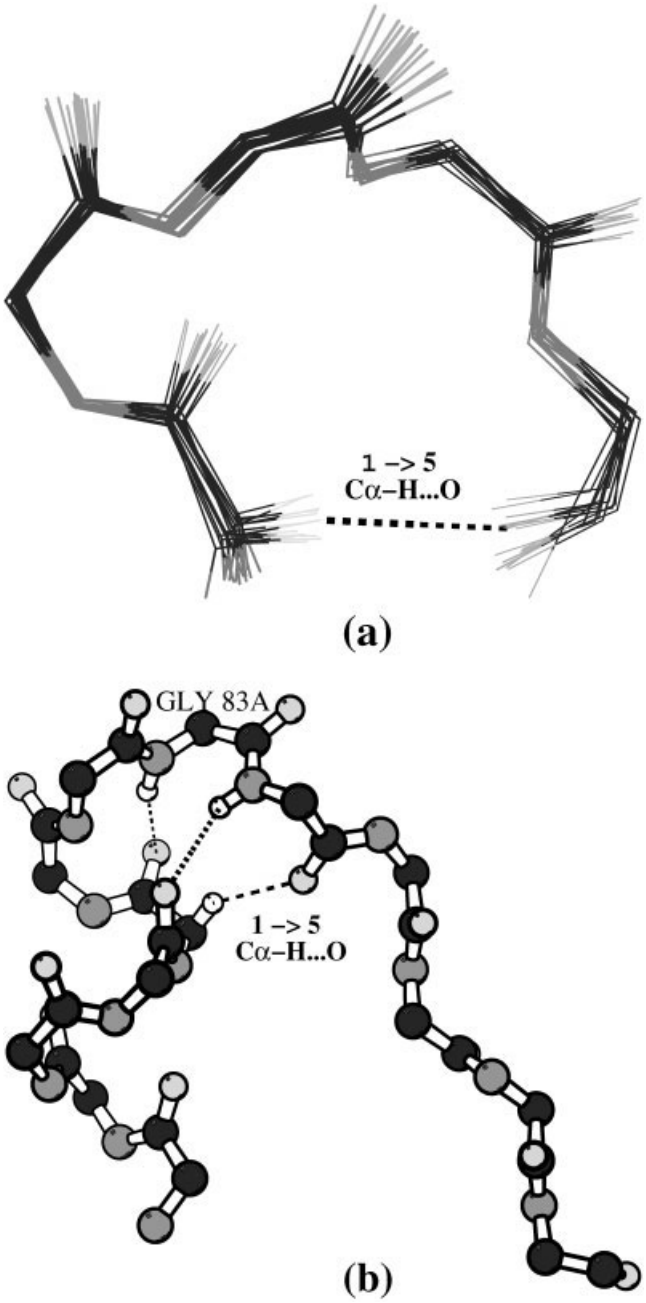

Fig. 8. (**a**) Structural superposition of the first 15 examples (Supplementary Table II) of 1 → 5 Cα—H…O hydrogen-bonded chain reversal motif. (**b**) 1 → 5 Cα—H…O hydrogen-bonded chain reversal motif seen in a helix (residues 76–81) of the protein RNA triophosphatase domain (1i9sA).[57] The 1 → 5 Cα—H…O interaction is between Glu 80 and Ile 84. The two backbone N—H…O hydrogen bonds, which represent the Schellman motif signature, are between residues Ile 84– Ile 79 (6 → 1), and Gly 83– Glu 80 (5 → 2), respectively. The succeeding residues (85–88), which are in a β-strand, are also depicted.

participates in 5 → 1 C—H…O interactions, and the 4 → 1 type is preferred for Ala. Even if it is argued that a few of the C—H…O contacts involving methyl groups are marginally destabilizing and have positive interaction energies,[30] on the whole, like other C—H…O hydrogen bonds, they also serve to satisfy the hydrogen-bonding acceptor potential of C═O groups, though they are energetically relatively weaker (−0.87 kcal/mol[3]) than those of other CH groups.

DISCUSSION

Energetics of C—H…O Hydrogen Bonds

The energy of a given hydrogen bond depends not only on the nature of the donor (D) and the acceptor (A) that form the hydrogen bond but also on the associated geometric parameters such as D…A and H…A distances, ∠D—H…A and the hydrogen-atom elevation angles.[14,30] However, it is pertinent to discuss the typical energies associated with different types of C—H…O hydrogen bonds analyzed in this report.

It has been theoretically calculated that with water as an acceptor, the strength of Cα—H…O interaction is −2.1 kcal/mol, nearly one-half of the energy of a conventional N—H…O hydrogen bond.[59,60] Pierce et al.[3] have shown that the interaction energy for the aromatic benzene and methane CH group donors with water is −1.47 and −0.87 kcal/mol, respectively. Similar estimated energy for the CH group, which is covalently attached to the electronegative nitrogen atom(s), as present in the amino acid His and Trp side-chains, is −2.4 kcal/mol.[61] With carbonyl oxygen as an acceptor, it is reported that the interaction energy is slightly weaker as compared to hydrogen bond to water.[59] At the least, one can, on average, expect interaction energy better than −1 kcal/mol for cohesive C—H…O hydrogen bonds. The ratio of the total number of local C—H…O hydrogen bonds (4 → 1 and 5 → 1 together) to the total number of intrahelical 5 → 1 N—H…O hydrogen bonds in the data set analyzed is found to be 0.63 (the ratio becomes 0.73 if methyl hydrogen bonds are included), indicating that for every 3 N—H…O hydrogen bonds, there are 2 C—H…O interactions present in the helix. It is found that an average-length helix has roughly 6 C—H…O hydrogen bonds that amount to stabilization energy of better than −6 kcal/mol per helix. Thus, together with classical N—H…O hydrogen bonds and other interactions, C—H…O hydrogen bonds appear to contribute cohesive energy to the stabilization of helices in proteins.

Possible Roles for C—H…O in Helix Stabilization

Satisfying the hydrogen-bonding potential of the C═O group

The carbonyl oxygen atom can accept more than one hydrogen atom. A study on lost hydrogen bonds[62] in proteins indicated that 70% of hydrogen bonds lost during the folding process involve main-chain carbonyl groups. McDonald and Thronton[63] have shown that in protein, on average, the C═O groups, for which the hydrogen bond–satisfying possibility with solvent is lost during folding, make 1.16 conventional hydrogen bonds. It is likely that the C—H…O hydrogen bonds, which occur in folded proteins, indeed at least partly compensate for such a loss.[1,63] Given this, hydrogen bonding with the amide hydrogen atom, as well as with CH donor(s), satisfies to a better extent the hydrogen-

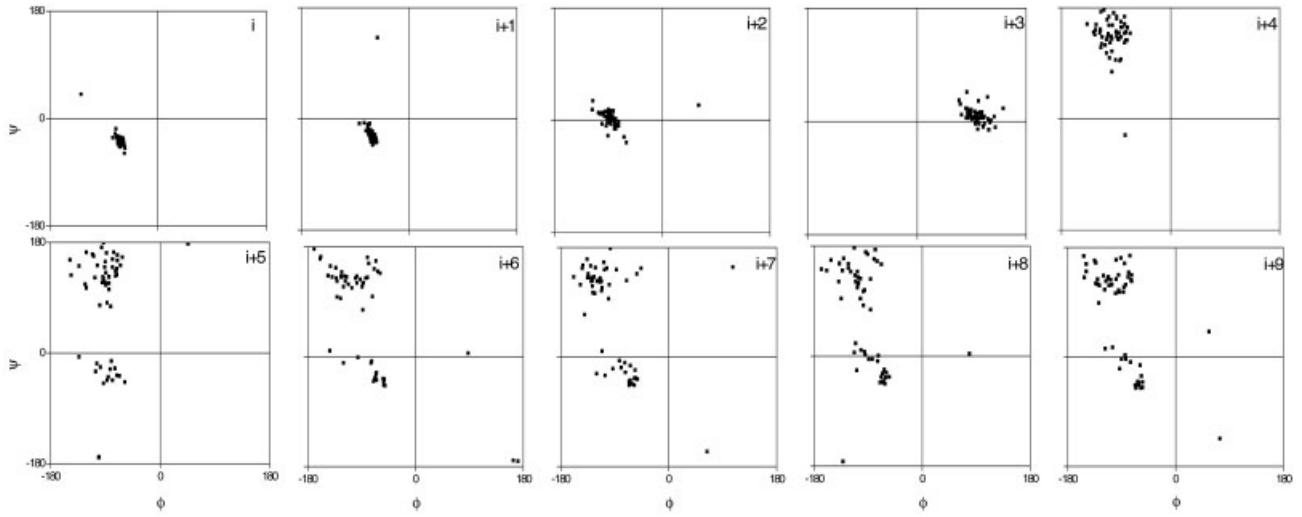

Fig. 9. Ramachandran plots of individual residues involved in the backbone 1 → 5 Cα—H…O hydrogen-bonded chain reversal motif [donor (*i*) to acceptor (*i* + 4) and the succeeding 5 residues (*i* + 5–*i* + 9) are also shown]. Gly (*i* + 3 residue), which precedes the acceptor, adopts α_L conformation with the acceptor (*i* + 4 residue), being in the extended region of the Ramachandran (ϕ, ψ) map. The backbone (ϕ, ψ) angles of the residues succeeding the acceptor cluster are either in the extended region or in the helical region.

bonding potential of the carbonyl oxygen atom, thereby contributing to helix stability.

Enthalpic compensation for the loss of side-chain conformational entropy

One of the known opposing forces in the protein-folding process is the loss of side-chain conformational entropy, and it has been estimated to be ~1 kcal/mol for the amino acid side-chains that fold into an α-helix.[64] Particularly in repetitive structures, helix and strand, the energetic "tug-of-war" is largely between side-chain conformational entropy and attractive local interactions such as hydrogen bonds.[65] It appears that the loss of side-chain conformational entropy for residues in helices is compensated at least to some extent through the enthalpic contribution due to the side-chain to backbone local C—H…O hydrogen bond. That the β-branched residues Ile and Val are involved in a larger number of C—H…O interactions [Fig. 3(c), Table II] than other residues is explained, as the side-chain conformational entropy loss is relatively more for these residues when they occur in the helix.[40,64]

Expanding hydrogen-bonding networks

The networks of hydrogen bonds are important for protein stability, functionality, and structural integrity. The participation of unconventional C—H…O hydrogen bonds, especially methyl-donated hydrogen bonds, in the hydrogen bonding networks has been shown recently in an organic complex molecule based on a theoretical study.[58] In this report, we have shown that the weak interactions formed by CH groups act together with conventional strong hydrogen bonds formed by backbone polar groups in helices and are thereby expected to expand the hydrogen-bonding network in helices.

Amino Acids Preferred in C—H…O Hydrogen Bonds and a Comparison With Helical Propensity

It is well known that the donatable strength of the CH group depends on the state of hybridization of the carbon atom, C (sp) > C (sp^2) > C (sp^3), and increases with the number of electron-withdrawing atoms covalently bonded to it.[14] The sp^2 hybridized carbon atoms of the aromatic residues (His, Phe, Trp, and Tyr) that are more acidic than sp^3 hybridized aliphatic CH_n groups make a considerable number of Cδ—H…O hydrogen bonds (Fig. 6). On average, all Ile, Met, and Val occurring in helices (from N3 position) are involved in at least one C—H…O hydrogen bond for each residue. In general, Trp, Cys, Asp, Phe, His, and Met, in decreasing order, are preferred for 4 → 1 C—H…O hydrogen bonds, and Val, Ile, Met, Gln, Glu, and Leu are favored for 5 → 1 type C—H…O hydrogen bonds.

Several factors may be invoked to rationalize the propensity of amino acids to occur in helices. However, it is noteworthy that the overall propensity of each amino acid to form either 4 → 1 or 5 → 1 C—H…O interactions shows that a majority of the helix-favoring residues with large side-chains—Met, Glu, Arg, Lys, Leu, and Gln—have significant propensity to form side-chain to main-chain C—H…O hydrogen bonds in the helix [Fig. 3(c)]. Several molecular dynamics (MD) studies[41,66,67] on Ala-rich peptides have pointed out the disruption of the enthalpic interaction between water and the helical peptide group by the amino acids side-chains as an important determinant of amino acid helical propensities. Ghosh et al.[68] have shown that the large side-chains play an important role in stabilizing α-helical structure of short peptides in aqueous solution through mediation of water access to backbone hydrogen bonds. In this study, we have shown that amino acids with large side-chains participate in more C—H…O hydrogen bonds in helices, as more donatable CH groups

are present in those side-chains. We have also shown that the C═O group in helices become more accessible to the solvent when large side-chains are replaced by Ala. The large side-chains may be contributing to the stability of the helix in two complementary ways: first through favorable enthalpic contribution in terms of C—H…O hydrogen bonds, which also better satisfy the hydrogen-bonding potential of carbonyl oxygen in the helix, and second by decreasing the solvent accessibility of the polar atoms of the peptide backbone, thereby preventing water from competing with helical hydrogen bonds. These arguments may provide at least qualitatively a rationale for the preference of amino acids with large side-chains to occur in helices.

Strain in Helices and C—H…O Hydrogen Bonds

In proteins, the peptide covalent geometry has an associated hidden strain, which has been shown to depend on the backbone conformation (ϕ, ψ) of residues.[69] There is a correlation between N-Cα-C angle (τ angle) and the conformational state of the residue.[70] An analysis from our laboratory on ultrahigh-resolution protein crystal structures has shown that the residues in helical conformational region show a larger τ angle than that found for the residues in an extended conformation, such as in the β-sheet.[24] A recent theoretical study[71] on peptides suggests that the formation of helices entails approximately 6.6 kcal/mol of strain within the backbone per hydrogen bond, and the hydrogen bond cooperativity is essential for the α-helix to become more stable than a corresponding β-strand. That same study also discusses the contribution of alkyl groups to the formation of helix via β-hydrogen atom(s) of Ile, Leu, and Val participating in C—H…O hydrogen bonds.[71] Here, we have shown that the amino acid side-chain γ-carbon hydrogen atom(s) are also equally found to be involved in C—H…O hydrogen bond in helices (Table II). Taken together, our analysis suggests that the C—H…O hydrogen bonds are additionally stabilizing the helix, which has backbone geometric strain as well as strain due to restricted rotameric states of side-chains.

C—H…O Hydrogen Bonds in Proteins Versus Helices

Hydrogen bonds, one of the most significant noncovalent interactions, play an important role in proteins, from the formation of secondary structures to the globular tertiary fold.[34] In the present analysis, we found that nearly 57% and 59% of the identified C—H…O hydrogen bonds (Table II) made by main-chain carbonyl oxygen atom with β- and γ-carbon atom(s), respectively, in proteins are present in helices, suggesting the preponderance of their occurrence in helices of globular proteins. These C—H…O hydrogen bonds are local intrinsic interactions favored by the helix geometry and the positioning of side-chains in the helix. In total, 35% of all C—H…O hydrogen bonds in proteins in the data set were found to be in helices and helix termini. Although only C—H…O interactions between the side-chain to main-chain carbonyl oxygen are studied in this report, there are a number of weak C—H…O interactions present between side-chain CH donors and side-chain polar atoms as acceptors (details not shown). In general, there is a prevalence of CH groups that have a range of carbon acidities that can participate in a multitude of noncovalent interactions such as C—H…O as compared to conventional potential polar atom donors in proteins. As seen in the helix, the C═O groups in proteins tend to better satisfy their hydrogen-bonding potential through alkyl and aromatic CH groups of amino acid main-chains as well as side-chains.

CONCLUSIONS

We have discussed in this report the role played by the various amino acid side-chains through local C—H…O hydrogen bonds and some of the other interactions in stabilizing helices and helix termini. Our analysis has brought to the fore the occurrence of attractive C—H…O interactions in helices and helix termini. The interactions studied are local in nature and are intrinsic to the helix. They occur in the middle region of the helix, at the N- and C-terminal regions, and are also present in the novel $1 \rightarrow 5$ Cα—H…O chain reversal structural motif. It emerges that the C—H…O hydrogen bonds are part of the repertoire of capping interactions present at the helix C-terminus.

The overall propensity for each amino acid to form either $4 \rightarrow 1$ or $5 \rightarrow 1$ C—H…O shows that the majority of the helix-favoring residues—Met, Glu, Arg, Lys, Leu, and Gln—that also have large side-chains with more donatable CH hydrogen atoms have significant propensity to form side-chain to main-chain C—H…O hydrogen bonds in the helix. The large side-chains are marked by their ability to shield from the solvent the polar atoms of the peptide backbone and at the same time participate in weak cohesive C—H…O interactions in the helix.

Collectively, C—H…O hydrogen bonds satisfy the hydrogen-bonding potential of C═O groups, to some extent compensate for the side-chain conformational entropy loss (especially for the β-branched residues Ile and Val for which the loss is greater when they occur in the helix), and occur along with helix-stabilizing salt bridges and helix-capping/terminating interactions. Furthermore, the relatively more strained helical conformation is expected to be additionally stabilized by C—H…O hydrogen bonds. In view of this, it may be suggested that C—H…O hydrogen bonds are likely to play a role in folded proteins and further, it may be speculated that they may even facilitate the very process of protein folding.

Even though our study has focused on the occurrence of local C—H…O interactions in the helical regions of globular proteins, it may be suggested that, in general, C—H…O hydrogen bonds expand the hydrogen-bond network and thereby contribute to the structural integrity and stability of folded proteins.

These observations may be useful to modulate the energetics of the helix by engineering hydrogen bonds involving side-chains, side-chain rotamer fixing to optimize C—H…O hydrogen bonds in homology modeling and refinement of experimentally determined protein structures, and de novo design of sequences that fold as helical

secondary structural elements, and can further enhance our understanding of the maximization of hydrogen bonds in the protein-folding process, as hydrogen bonds are the major noncovalent interactions influencing the stability and specificity in protein folding.

ACKNOWLEDGMENTS

Facilities at the Interactive Graphics-Based Molecular Modeling Facility and the Distributed Information Centre (both funded by DBT) were used in this work and are acknowledged. We are extremely grateful to the anonymous referees for a very thorough reading of the manuscript and valuable suggestions.

REFERENCES

1. Derewenda ZS, Lee L, Derewenda U. The occurrence of C—H…O hydrogen bonds in proteins. J Mol Biol 1995;252:248–262.
2. Weiss MS, Brandl M, Suhnel J, Pal D, Hilgenfeld R. More hydrogen bonds for the (structural) biologist. Trends Biochem Sci 2001;26:521–523.
3. Pierce AC, Sandretto KL, Bemis GW. Kinase inhibitors and the case for C—H…O hydrogen bonds in protein–ligand binding. Proteins 2002;49:567–576.
4. Sarkhel S, Desiraju GR. N—H…O, O—H…O and C—H…O hydrogen bonds in protein–ligand complexes: strong and weak interactions in molecular recognition. Proteins 2004;54:247–259.
5. Klaholz BP, Moras D. C—H…O hydrogen bonds in the nuclear receptor RARγ—a potential tool for drug selectivity. Structure 2002;10:1197–1204.
6. Jiang L, Lai L. C—H…O hydrogen bonds at protein–protein interfaces. J Biol Chem 2002;277:37732–37740.
7. Ghosh A, Bansal M. C—H…O hydrogen bonds in minor groove of A-tracts in DNA double helices. J Mol Biol 1999;294:1149–1158.
8. Wahl MC, Sundarlingam M. C—H…O hydrogen bonding in biology. Trends Biochem Sci 1997;22:97–102.
9. Mandel-Gutfreund Y, Margalit H, Jernigan RL, Zhurkin VB. A role for C—H…O interactions in protein–DNA recognition. J Mol Biol 1998;277:1129–1140.
10. Denessiouk KA, Johnson MS. "Acceptor–donor–acceptor" motifs recognize the Watson–Crick, Hoogsteen and sugar "donor–acceptor–donor" edges of adenine and adenosine containing ligands. J Mol Biol 2003;333:1025–1043.
11. Bella J, Berman M. Crystallographic evidence for Cα—H…O═;C hydrogen bonds in a collagen triple helix. J Mol Biol 1996;264:734–742.
12. Taylor R, Kennard O. Crystallographic evidence for existence of C—H…O, C—H…N and C—H…Cl hydrogen bonds. J Am Chem Soc 1982;104:5063–5070.
13. Desiraju GR. The C—H…O hydrogen bond in crystals: what is it? Acc Chem Res 1991;24:290–296.
14. Desiraju GR, Steiner T. The weak hydrogen bonds. In: Structural chemistry and biology. Oxford, UK: Oxford University Press; 1999.
15. Gu Y, Kar T, Scheiner S. Fundamentals properties of the C—H…O interaction: is it a true hydrogen bond? J Am Chem Soc 1999;121:9411–9422.
16. Ramagopal UA, Ramakumar S, Sahal D, Chauhan VS. *De novo* design and characterization of an apolar helical hairpin peptide at atomic resolution: compaction mediated by weak interactions. Proc Natl Acad Sci USA 2001;98:870–874.
17. Fabiola GF, Krishnaswamy S, Nagarajan V, Pattabhi V. C—H…O hydrogen bonds in β-sheets. Acta Crystallogr D Biol Crystallogr 1997;53:316–320.
18. Chakrabarti P, Chakrabarti S. C—H…O hydrogen bond involving proline residues in α-helices. J Mol Biol 1998;284:867–873.
19. Bhattacharyya R, Chakrabarti P. Stereospecific interactions of proline residues in protein structures and complexes. J Mol Biol 2003;331:925–940.
20. Senes A, Ubarretxena-Belandia I, Engelman DM. The Cα—H…O hydrogen bond: a determinant of stability and specificity in transmembrane helix interactions. Proc Natl Acad Sci USA 2001;98:9056–9061.
21. Loll B, Raszewski G, Saenger W, Biesiadka J. Functional role of Cα—H…O hydrogen bonds between transmembrane α-helices in Photosystem I. J Mol Biol 2003;328:737–747.
22. Babu MM, Singh SK, Balaram P. A C—H…O hydrogen bond stabilized polypeptide chain reversal motif at the C terminus of helices in proteins. J Mol Biol 2002;322:871–880.
23. Singh SK, Babu MM, Balaram P. Registering α-helices and β-strands using backbone C—H…O interactions. Proteins 2003;51:167–171.
24. Natesh R, Manikandan K, Bhanumoorthy P, Viswamitra MA, Ramakumar S. Thermostable xylanase from *Thermoascus aurantiacus* at ultrahigh resolution (0.89 Å) at 100 K and atomic resolution (1.11 Å) at 293 K refined anisotropically to small-molecule accuracy. Acta Crystallogr D Biol Crystallogr 2003;59:105–117.
25. Westbrook J, Feng Z, Jain S, Bhat TN, Thanki N, Ravichandran V, Gilliland GL, Bluhm WF, Weissig H, Greer DS, Bourne PE, Berman HM. The Protein Data Bank: unifying the archive. Nucleic Acids Res 2002;30:245–248.
26. Hobohm U, Sander C. Enlarged representative set of protein structures. Protein Sci 1994;3:522–524.
27. Hutchinson EG, Thornton JM. PROMOTIF: a program to identify and analyse structural motifs in proteins. Protein Sci 1996;5:212–220.
28. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr 1994;50:760–763.
29. Aurora R, Rose GD. Helix capping. Protein Sci 1998;7:21–38.
30. Steiner T, Desiraju GR. Distinction between the weak hydrogen bond and the van der Waals interaction. Chem Commun 1998;891–892.
31. Hubbard SJ, Thornton JM. NACCESS. Computer program, Department of Biochemistry and Molecular Biology, University College of London.
32. Chou PY, Fasman GD. Conformational parameters for amino acids in helical, beta-sheet, and random coil regions calculated from proteins. Biochemistry 1974;13:211–222.
33. Esnouf RM. An extensively modified version of MolScript that includes greatly enhanced coloring capabilities. J Mol Graph Model 1997;15:132–134.
34. Baker EN, Hubbard RE. Hydrogen bonding in globular proteins. Prog Biophys Mol Biol 1984;44:97–179.
35. Eswar N, Ramakrishnan C. Deterministic features of side-chain main-chain hydrogen bonds in globular protein structures. Protein Eng 2000;13:227–238.
36. Walther D, Argos P. Intrahelical side-chain–side-chain contacts: the consequences of restricted rotameric states and implications for helix engineering and design. Protein Eng 1996;9:471–478.
37. Chakrabarti P, Pal D. The interrelationships of side-chain and main-chain conformations in proteins. Prog Biophys Mol Biol 2001;76:1–102.
38. IUPAC-IUB. Commission on biochemical nomenclature: abbreviations and symbols for the description of the conformation of polypeptide chains. J Mol Biol 1970;52:1–17.
39. Markley JL, Bax A, Arata Y, Hilbers CW, Kaptein R, Sykes BD, Wright PE, Wuthrich K. Recommendations for the presentation of NMR structures of proteins and nucleic acids. J Mol Biol 1998;280:933–952.
40. Aurora R, Creamer TP, Srinivasan R, Rose GD. Local interactions in protein folding: lessons from the α-helix. J Biol Chem 1997;272:1413–1416.
41. Garcia EA, Sanbonmatsu YK. α-Helical stabilization by side-chain shielding of backbone hydrogen bonds. Proc Natl Acad Sci USA 2002;99:2782–2787.
42. Kwon HJ, Tirumalai R, Landy A, Ellenberger T. Flexibility in DNA recombination: structure of the lambda integrase catalytic core. Science 1997;276:126–131.
43. Fernandenz A, Scheraga HA. Insufficiently dehydrated hydrogen bonds as determinants of protein interactions. Proc Natl Acad Sci USA 2003;100:113–118.
44. Fernandenz A, Kardos J, Goto Y. Protein folding: could hydrophobic collapse be coupled with hydrogen-bond formation? FEBS Lett 2003;536:187–192.
45. Kumar S, Nussinov R. Close range interactions in proteins. ChemBioChem 2002;3:604–617.
46. Clausen T, Huber R, Prade L, Wahl MC, Messerschmidt A.

Crystal structure of *Escherichia coli* cystathionine gamma-synthase at 1.5 Å resolution. EMBO J 1998;17:6827–6838.
47. Lawrence MC, Pilling PA, Epa VC, Berry AM, Ogunniyi AD, Paton JC. The crystal structure of pneumococcal surface antigen PsaA reveals a metal-binding site and a novel structure for a putative ABC-type binding protein. Structure 1998;6:1553–1561.
48. Mathews II, Erion MD, Ealick SE. structure of human adenosine kinase at 1.5 Å resolution. Biochemistry 1998;37:15607–15620.
49. Guarne A, Hampoelz B, Glaser W, Carpena X, Tormo J, Fita I, Skern T. Structural and biochemical features distinguish the foot-and-mouth disease virus leader proteinase from other papain-like enzymes. J Mol Biol 2000;302:1227–1240.
50. Derewenda ZS, Derewenda U, Kobos PM. (His)C^{ϵ}—H…O═C∠ hydrogen bond in the active sites of serine hydrolases. J Mol Biol 1994;241:83–93.
51. Gunasekaran K, Nagarajaram HA, Ramakrishnan C, Balaram P. Stereochemical punctuation marks in protein structures: glycine and proline containing helix stop signals. J Mol Biol 1998;275:917–932.
52. Rafferty JB, Somers WS, Saint-Girons I, Phillips SE. Three-dimensional crystal structures of *Escherichia coli* met repressor with and without corepressor. Nature 1989;341:705–710.
53. Mayer SM, Lawson DM, Gormal CA, Roe SM, Smith BE. New insights into structure–function relationships in nitrogenase: a 1.6 Å resolution X-ray crystallographic study of *Klebsiella pneumoniae* MoFe-protein. J Mol Biol 1999;292:871–891.
54. Ramachandran GN, Sasisekharan V. Conformations of polypeptides and proteins. Adv Protein Chem 1968;23:283–438.
55. Nagarajaram HA, Sowdhamini R, Ramakrishnan C, Balaram P. Termination of right handed helices in proteins by residues in left handed helical conformations. FEBS Lett 1993;321:79–83.
56. Rajashankar KR, Ramakumar S. π-Turns in proteins and peptides: classification, conformation, occurrence, hydration and sequence. Protein Sci 1996;5:932–946.
57. Changela A, Ho CK, Martins A, Shuman S, Mondragon A. Structure and mechanism of the RNA triphosphatase component of mammalian mRNA capping enzyme. EMBO J 2001;20:2575–2586.
58. Vergenz RA, Yazji I, Whittington C, Daw J, Tu Tran K. Computational evidence for methyl-donated hydrogen bonds and hydrogen-bond networking in 1, 2-ethanediol-dimethyl sulfoxide. J Am Chem Soc 2003;125:12318–12327.
59. Scheiner S, Kar T, Gu Y. Strength of the Cα—H…O hydrogen bond of amino acid residues. J Biol Chem 2001;276:9832–9837.
60. Vargas R, Garza J, Dixon DA, Hay BP. How strong is the Cα—H…O hydrogen bond? J Am Chem Soc 2000;122:4750–4755.
61. Scheiner S, Kar T, Pattanayak J. Comparison of various types of hydrogen bonds involving aromatic amino acids. J Am Chem Soc 2002;124:13257–13264.
62. Savage H, Elliot C, Freeman C, Finney J. Lost hydrogen bonds and buried surface area: rationalizing stability in globular proteins. J. Chem Soc Faraday Trans 1993;89:2609–2617.
63. McDonald IK, Thornton JM. Satisfying hydrogen bonding potential in proteins. J Mol Biol 1994;238:777–793.
64. Penel S, Doig AJ. Rotamer strain energy in protein helices—quantification of a major force opposing protein folding. J Mol Biol 2001;305:961–968.
65. Srinivasan R, Rose GD. A physical basis for protein secondary structure. Proc Natl Acad Sci USA 1999;96:14258–14263.
66. Avbelj F, Luo P, Baldwin RL. Energetics of the interaction between water and the helical peptide group and its role in determining helix propensities. Proc Natl Acad Sci USA 2000;97:10786–10791.
67. Vila JA, Ripoll DR, Scheraga HA. Physical reasons for the unusual α-helix stabilization afforded by charged or neutral polar residues in alanine-rich peptides. Proc Natl Acad Sci USA 2000;97:13075–13079.
68. Ghosh T, Garde S, Garcia AE. Role of backbone hydration and salt-bridge formation in stability of α-helix in solution. Biophys J 2003;85:3187–3193.
69. Karplus PA. Experimentally observed conformation-dependent geometry and hidden strain in proteins. Protein Sci 1996;5:1406–1420.
70. Esposito L, Vitagliano L, Sica F, Zagari GSA, Mazzarella L. The ultrahigh resolution crystal structure of ribonuclease A containing an isoaspartyl residue: hydration and sterochemical analysis. J Mol Biol 2000;297:713–732.
71. Wieczorek R, Dannenberg JJ. H-bonding cooperativity and energitics of α-helix formation of five 17-amino acid peptides. J Am Chem Soc 2003;125:8124–8129.