Journal of Biomolecular Structure & Dynamics, ISSN 0739-1102
Volume 17, Issue Number 5, (2000)

HELANAL: A Program to Characterize Helix Geometry in Proteins

http://www.adeninepress.com

Manju Bansal*, Sandeep Kumar† and R. Velavan
Molecular Biophysics Unit,
Indian Institute of Science,
Bangalore – 560 012, India
†Present Address:
Laboratory of Experimental
and Computational Biology,
Bldg. 469, Rm. 151, NCI-FCRDC,
Frederick MD 21702, USA

Abstract

A detailed analysis of structural and position dependent characteristic features of helices will give a better understanding of the secondary structure formation in globular proteins. Here we describe an algorithm that quantifies the geometry of helices in proteins on the basis of their C^{α} atoms alone. The Fortran program HELANAL can extract the helices from the PDB files and then characterises the overall geometry of each helix as being linear, curved or kinked, in terms of its local structural features, *viz.* local helical twist and rise, virtual torsion angle, local helix origins and bending angles between successive local helix axes. Even helices with large radius of curvature are unambiguously identified as being linear or curved. The program can also be used to differentiate a kinked helix and other motifs, such as helix-loop-helix or a helix-turn-helix (with a single residue linker) with the help of local bending angles. In addition to these, the program can also be used to characterise the helix start and end as well as other types of secondary structures.

Introduction

α-helix is one of the most common secondary structural motifs in proteins (1,2). In case of a regular helix, the three-dimensional path traveled by its axis will be perfectly linear. However, α helices have been observed to be distorted due to a variety of reasons, such as, occurrence of proline (3,4), solvent induced distortions (5), peptide bond distortions (6,7). Due to these distortions, the three dimensional path of an α helix often becomes non-linear (8). An early systematic data base analysis of α helices in high resolution protein crystal structures available from Protein Data Bank (PDB) (9) revealed that most of the α helices in globular proteins are curved (4). Recent analyses of 64 long α helices and all 1131 α helices in the 205 non-homologous protein chains, present in the June 1995 release of PDB select (10) also confirmed that 73% of the α helices are curved (11,12). Knowledge about the details of α helical structures and their distortions is essential to understand completely the protein structure, function and folding. Precise determination of the three dimensional path followed by an α helix axis, i.e. helix geometry, is also required to study helix packing with the rest of the protein globule and formation of super secondary structures, such as, coiled coils and helix bundles, formed by packing of two or more helices.

In this paper we report the details of HELANAL, a Fortran 77 program developed in our laboratory to characterise the geometry of a helix. It computes local helical twist, rise, virtual torsion angle, local helix origins and local helix axes for each turn of the helix. Angles between successive local helix axes are used to estimate the bending in the helix. Local helix origins are reoriented in the X-Y plane and fitted with least squares line and circle. HELANAL classifies the overall geometry of a helix as being linear, curved or kinked based on the value of the maximum bending angle, relative goodness of the line and the circle fitted to the local helix origins and the square of linear correlation coefficient.

*Phone: 91-80-3092534;
Fax: 91-80-3600535;
E-mail: mb@mbu.iisc.ernet.in

Methods

HELANAL is coded in FORTRAN 77. It has been successfully compiled and executed using RM/FORTRAN compiler RMFORT ver 2.10 and Lahey FORTRAN F77L ver 4.10 compilers under DOS 6.0 environment on IBM-PC compatible machines, Fort77 compiler under LINUX 5.0 on Pentium 100MHz IBM-PC compatible-machine, XL FORTRAN compiler under AIX ver 4 on IBM RS-6000 workstations (590 series) and F77 compiler under IRIX Release 6.0 on Silicon Graphics O2 workstation.

Algorithm

Geometry of a helix is characterised by calculating the local helix axis and the local helix origin for four contiguous C^{α} atoms, following the method of Sugeta and Miyazawa (13). This window of four C^{α} atoms slides along the length of the helix, one C^{α} atom at a time. The angles between successive local helix axes can identify local bends or kinks as well as smooth curvature in helix. A matrix, whose elements M(i,j) are the bending angles between local helix axes i and j, is obtained to get a qualitative idea about the overall geometry of the helix. Unit twist and unit height of the helix have been calculated, for each set of four C^{α} atoms to analyse uniformity of the helix. The local helix origins trace out the path described by a helix in the three dimensional space. These local helix origins are projected into the X-Y plane and fitted with a circle and a line by the least squares method. HELANAL can work only with helices containing 9 or more residues in a helix, when calculating these parameters, since this ensures availability of sufficient data points to deduce its overall geometry.

Figure 1 outlines the procedure used to classify the overall geometry of a helix as linear, curved or kinked. The helix geometry is assigned taking into consideration the values of several parameters, as outlined below.

Maximum local bending angle in the helix and the root mean square deviations for the best line (rmsL) and circle (rmsC) fitted to the local helix origins, as well as r^2, the square of linear correlation coefficient for the line fit. If the value of any local bending angle is > 20°, then the $i \rightarrow i - 4$ hydrogen bond gets broken, which implies that a maximum bending angle of 25° can be used as a criteria to define the helix start and end. Distortion at a single residue in an α helix is sufficient for the helix to be kinked. Hence, if the maximum bending angle is ≥ 20° at any C^{α} atom within the helix, then the helix is classified as **kinked**. However, if the kink angle exceeds 60° within the helix, then we treat the fragment as a motif consisting of two separate helices constituting a helix-loop-helix or a helix-turn-helix (a single residue linker). If the maximum local bending angle is < 20°, then the values of other three parameters (rmsC, rmsL and r^2) determine the helix geometry to be linear or smoothly curved as described in Figure 1. If both rmsC and rmsL have values greater than 1Å, the helix may have large random distortions. In this case, no geometry type is assigned to the helix. If either one or both of the parameters rmsC and rmsL are less than 1Å, then the ratio rmsL/rmsC is computed. A value of greater than one for the ratio implies that the circle fit to the local helix origins is better than the line fit. In this case, the helix is classified as **curved**. If the ratio is ≤ 0.7, it implies that the line fits to the reoriented local helix origins better than the circle and if in addition, the value of $r^2 \geq 0.8$, then the helix is assigned as linear, else no geometrical type is assigned to the helix. Value of 0.7 to 1.0 for the ratio represents a zone where both line and the circle fit almost equally well to the local helix origins. In this zone, if $r^2 \geq 0.8$, that is, the local helix origins significantly correlate with a line, then helix geometry is classified as **linear**. But, if $r^2 \leq 0.5$, i.e., correlation between local helix origins and the line is weak, then the helix geometry is classified as **curved**. If value of r^2 lies in the range between 0.5 – 0.8, then the geometry of the helix is ambivalent and in this case also, we do not assign any geometry type to the helix.

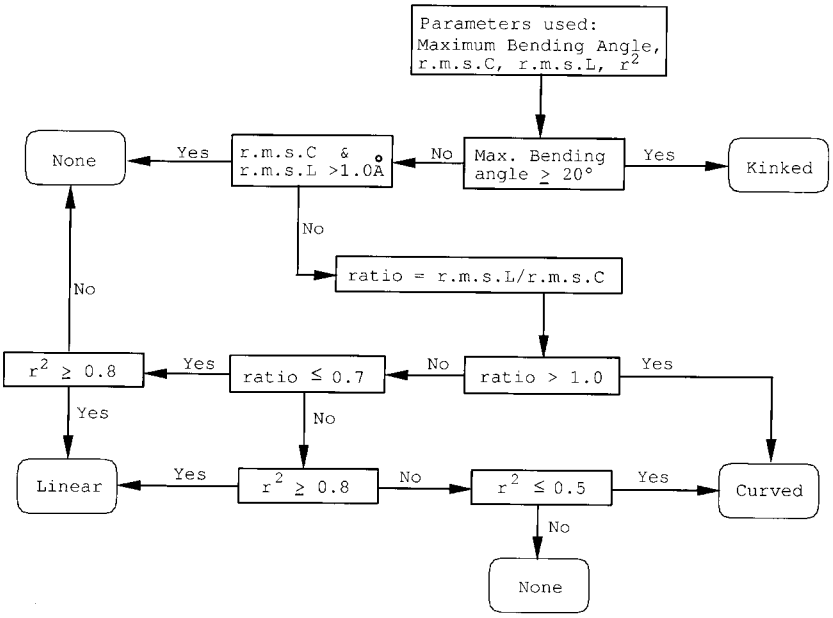

Figure 1: Flow chart describing the algorithm for geometrical classification of helices. Four parameters, namely, maximum bending angle, root mean square deviation values from line (rmsL) and circle (rmsC), fitted by least square method to local helix origins reoriented in the X-Y plane and square of linear correlation coefficient (r^2), are computed by HELANAL for each helix and used to classify its geometry as linear, curved and kinked or none of these.

All criteria involve the use of threshold values, which have been deduced from an in depth analysis of 1131 α helices from the Brookhaven Protein Data Bank (11,12). The analysis characterised 256 helices as linear, 821 as curved, 45 as kinked and only 9 helices (all of length between 9 and 14 residues) could not be assigned any geometry. The criteria for identifying kinked helices (*viz*. the maximum bending angle ≥ 20°) was successful in identifying 25 of the 28 proline containing kinked helices, in addition to 20 other kinked helices in the database. It was also found to agree with additional manual checks, such as large values for average bending angle and poor fit to a line (as indicated by rmsL and r^2 values). More than 99% of linear and curved helices have values of less than 0.5 for rmsL and rmsC respectively, even though the filter applied allows values up to 1.0 for these parameters. It was also found that even for helices with the large radius of curvature (radius ≥ 150Å) a distinction could be made between curved and linear helices, with 5 out of 24 helices in this category being unambiguously identified as being linear, unlike some of the earlier methods, which classified all helices with large radius of curvature, including coiled coils as linear (4).

Implementation

The HELANAL program has been designed to run interactively and the input files to this program depend on the options to be used.

HELIX record in each PDB file defines the extent of all α helices (identified by the crystallographers) in that protein. However, these records depend on the definition used to characterise an α helix. Various criteria, such as occurrence of hydrogen bonds - DSSP package (14), backbone torsion angles φ and ψ (15,16,17) and α carbon position (18,19,20) have been used by different workers to define α helices in protein crystal structures. This particularly affects the identification of helix termini, because an amino acid residue at the helix termini may or may not be included in the helix depending upon the criteria used. Since it is desirable to use a consistent definition for helix termini, HELANAL provides an option for using a user-defined input for helix identification and characterisation.

HELANAL can directly use the HELIX records, as defined in the PDB files, for information about the helices whose geometry is to be assigned. If the user wishes

to use the HELIX record in the PDB files, the name of the PDB files can either be keyed in consecutively or read in through another input file (e.g. ***.fil in (5x,a11) format) containing a list of PDB files. If the user does not wish to use the HELIX records from the PDB files, but prefers a user defined helix start and end, then a separate input file containing helix start and end information, equivalent to the PDB HELIX records with the first field "HELIX" replaced by the PDB filename can be used. This file can also be in a different format, in which case the format defining PDB file name, starting residue name, chain label, starting residue number, ending residue name, chain label and ending residue number should be typed in the first line of the input file. This allows analysis of secondary structures, other than α helices, to be carried out using HELANAL. A sample input file 'example.inp' is shown in Figure 2 and includes representatives of left handed ω helix, as well as β strands, in addition to right-handed α helices. The program can also take care of a non-PDB coordinate file. In this case, the format defining the fields for atom name (e.g. N, CA, CB, etc.), residue name (e.g. ALA, LEU, etc.), chain name (e.g., A, B or blank), residue number and Cartesian coordinates of the atoms has to be typed in by the user when HELANAL is executed.

```
(a11,4x,2(a3,1x,a1,1x,i4,2x))
pdb1rib.ent     PRO A  102   ILE A  129
pdb1bge.ent     ALA B  144   LEU B  169
pdb1lis.ent     THR     44   ASP     74
pdb1ars.ent     MET    397   VAL    405
pdbcoil.ent     ALA      1   ALA     28
pdb2zta.ent     ARG A    1   VAL A   30
pdb2zta.ent     ARG B    1   VAL B   30
pdbomgl.ent HA  ALA      1   ILE     10
pdbstrn.ent S1  ALA      1   ILE     10
pdb1lkk.ent     GLU A  176   ASN A  185
```

Figure 2: A sample input file called 'example.inp'. Each line in this file defines the name of the coordinate file and information about the first and last residue in a helix.

HELANAL program requires only C^{α} co-ordinates to define and characterise the helices and hence works with an input co-ordinate file containing only the C^{α} atoms. However, files containing all the atoms can also be used as input.

The following output files are created during each run of HELANAL:

RUN.ANS contains all the questions asked and the answers given, during a run of HELANAL. Output files corresponding to three different runs, are given in Appendix A, B and C and illustrate the usage of various options available in HELANAL. Appendix A shows the direct utilisation of PDB HELIX records, wherein the PDB file names are keyed in, consecutively, during an interactive run, whereas in Appendix B, the PDB HELIX records are used, but the PDB files names are acquired from a separate file (input.fil). The third mode is shown in Appendix C, where a separate input file (shown in Figure 2 as example.inp) contains HELIX start/end information. Examples of non-PDB formats, which can be used for a HELANAL run, are also shown in this case.

HELINFO.OUT contains a list of helix records and is created only when the original HELIX records in the PDB files are used. This file can be used as an input file for subsequent runs of HELANAL.

HELCA.OUT contains co-ordinates of the C^{α} atoms constituting the helices.

AXES.OUT contains direction cosines, l, m, n, of the local helix axes fitted to four consecutive C^{α} atoms, along with a matrix M(i,j) whose elements are the angles between all pairs of the local helix axes i and j.

ANGLE.OUT contains the angles between successive local helix axes with appro-

priate residue number and names in one letter code. It also contains values of mean local bending angle and maximum local bending angle for the helix.

ORIGIN.OUT contains the local helix origins for the helix along with the statistics obtained by fitting least squares plane, circle and line to the local helix origins. For the plane, its direction cosines, l, m, n, perpendicular distance from X-Y plane and root mean square deviation (r.m.s.d.) are listed. For the circle fitted to the local helix origins reoriented in the X-Y plane, its centre, radius and r.m.s.d. are listed. Finally, for the line fitted to the local helix origins reoriented in the X-Y plane, its slope, intercept, square of linear correlation coefficient (r^2) and r.m.s.d. are given.

NH.OUT contains the unit twist, unit height, number of units per turn, radius of the helix and virtual torsion angle, calculated for each set of four C^α atoms in the helix as well as the average values of these parameters for the whole helix.

***.PRM file contains a summary of the parameters obtained by HELANAL from a single PDB file or from a set of PDB files. It lists the sequence of the helix (in one letter code), average values of the helical parameters, mean and maximum bending angles, radius and r.m.s.d. for the least squares circle fitted to the local helix origins, r.m.s.d and square of linear correlation coefficient (r^2) for the line fitted to the local helix origins reoriented in the X-Y plane. The file is called 'example.prm', if the input file is 'example.inp', else it has the name ***.prm, where *** is the name of the input coordinate file.

***.TAB presents the overall results for all the helices in tabular form, along with the assignment of helix geometry (Linear 'L', Curved 'C', Kinked 'K' or none '*'). A file 'example.tab' that was created by running HELANAL on 'example.inp' file is shown in Figure 3.

Results and Discussion

Aim of the software package presented here is to determine the helix termini and to develop a self-consistent method for analysis of helix geometry. The algorithm presented here is able to define the helix start/end and to characterise the geometry of a helix in terms of the three-dimensional path described by the local helix origins for the helix traced by C^α atoms. Thus, the overall geometry of the helix emerges as a natural consequence of its local structural features. It requires four contiguous C^α atoms to compute a local helix axis and a local helix origin. Hence, it is ideally suited to characterise the geometries of α helices, which have 3.6 residues per turn. In order to identify the helix boundary, the helix can be arbitrarily extended by a few residues at either end, in which case large deviations in the local bending angles (local bending angle > 25° at more than one contiguous C^α atom as listed in angle.out) and local helical parameters twist and rise (in the file nh.out) from the ideal α helical values will give a clear identification of the helix termini. After the helix length reassignment or cross check of the helix termini, HELANAL can be iterated for the geometrical analysis of helices.

The program unambiguously characterises geometries of a helices forming "coiled-coil" motif as being smoothly curved, in case of the fiber model for paramyosin (21) as well as for the leucine zipper solved by x-ray crystallography (22). The 'example.inp' file shown in Figure 2 contains representative examples of a linear helix from *E.Coli* ribonucleotide reductase protein R2 (pdb1rib.ent, chain A, Pro 102 – Ile 129), a curved helix from cytokine protein (pdb1bge.ent, chain B, Ala 144 – Leu 169), a kinked helix from fertilisation protein (pdb1lis.ent, Thr 44 – Asp 74), the coiled-coil model (pdbcoil.ent) for paramyosin (21) and the two α helices in the crystal structure of a leucine zipper motif (pdb2zta.ent) (22). Results obtained by running HELANAL on this file are summarised in the file called 'example.tab' shown in Figure 3. These results demonstrate that the program works well and give

confidence in the values of various parameters obtained by HELANAL for crystallographically observed α helices. Cα traces of the three a helices characterised as being linear, curved and kinked in Figure 3, are shown in Figure 4 as ribbon diagrams (23), along with the paths traced by their local helix origins. A representative example of a helix, which cannot be characterised as linear, curved or kinked is also included in the example file. As seen from Figure 3, the rmsC and rmsL values are both > 1.0Å for the helix geometry to be unassigned. The path traced by the local helix origins, as shown in Figure 5, is quiet irregular in this case.

PDB Code	Helix Start-End	n	h	Aver. vtor	Aver. BA	Max. BA	Radius	rmsC	rmsL	r2	Geometry
1rib A	102P-129I	3.61	1.52	50.5	4.9	12.2	140	.39	.16	1.00	L
	Std. Dev.	.13	.09	4.6	3.6	(106T)					
1bge B	144A-169L	3.68	1.51	48.9	6.6	15.0	73	.07	.93	.98	C
	Std. Dev.	.11	.11	5.0	4.2	(158L)					
1lis	44T-74D	3.59	1.55	52.0	12.3	54.4	25	1.36	10.03	.30	K
	Std. Dev.	.18	.21	10.9	14.5	(62W)					
1ars	397M-405V	3.67	1.49	48.6	5.3	5.7	4	1.03	1.78	.64	*
	Std. Dev.	.09	.07	3.6	.5	(402E)					
coil	1A-28A	3.60	1.50	50.5	1.7	2.1	170	.00	.62	1.00	C
	Std. Dev.	.01	.02	.8	.3	(6A)					
2zta A	1R-30V	3.61	1.51	50.4	4.4	7.0	141	.11	.51	.99	C
	Std. Dev.	.08	.07	3.0	1.8	(25R)					
2zta B	1R-30V	3.65	1.50	49.3	4.2	9.4	107	.08	.62	.94	C
	Std. Dev.	.09	.07	2.8	2.6	(14S)					
omgl	1A-10I	-4.00	1.32	-38.1	1.7	2.0	4022	.00	.00	1.00	L
	Std. Dev.	.00	.03	1.0	.5	(6Q)					
strn	11A-20I	-2.30	3.30	-152.8	.0	.0	417	.06	.00	1.00	L
	Std. Dev.	.00	.00	.0	.0	(17R)					
1lkk A	176E-185N	-2.38	3.23	-147.2	16.9	23.9	36	.15	1.17	.96	K
	Std. Dev.	.12	.16	16.8	9.6	(181Y)					

Figure 3: Example.tab, an output file of HELANAL which lists the various parameters and the overall helix geometry assignment, in a tabular form, for the helices listed in the example.inp file shown in Figure 2. Rows (i)-(iv) list examples of α helices with linear (pdb1rib.ent, chain A Pro 102 – Ile 129), curved (pdb1bge.ent, chain B Ala 144 – Leu 169), kinked (pdb1lis.ent, Thr 44 – Asp 74) and unassigned geometry (pdb1ars.ent, Met 397 – Val 405) as found in proteins. (v) coil corresponds to a helix with four heptad repeats of the coiled coil model of paramyosin (21). (vi) and (vii) are two α helices in a leucine zipper structure (pdb2zta.ent) solved by x-ray crystallography (22). (viii) and (ix) correspond to a left handed ω-helix model and β-strand whose coordinates are read in from the files pdbomgl.ent and pdbstrn.ent, respectively. (x) is an example of a kinked β-strand in a protein structure (pdb1lkk.ent, chain A Glu 176 – Asn 185).

NL = 3; NC = 4; NK = 2; NA = 1; NH = 10
Helix: Number and name of first and last residues of the helix.
n: Average number of residues per turn of the helix.
h: Average unit height in the helix (Angstroms).
Aver. vtor: Average virtual torsion angle defined by four CA atoms (Deg.).
Aver. BA: Average Bending angle between successive local helix axes (Deg.).
Max. BA: Maximum Bending angle between succesive local helix axes (Deg.). Residue number and name given in parenthesis
Radius: Radius of least squares circle fitted to the local helix origins (Angstroms).
rmsC: Root Mean Square Deviation for least squares circle fitted to the local helix origins (Angstroms).
rmsL: Root Mean Square Deviation for least squares line fitted to the local helix origins (Angstroms).
r2: Square of linear correlation coefficient for least squares line fitted to the local helix origins.
Geometry: Overall geometry of the helix, Linear (L), Curved (C), Kinked (K) or unassigned (*).
Std. Dev.: Standard Deviations of the average parameters for the helix.
NL, NC, NK: Number of helices with overall linear, curved, and kinked geometry.
NA, NH: unassigned geometry and total number of helices, respectively.

HELANAL has been used to characterise the geometries of 1131 α helices (≥ 9 residues) in 205 non-homologous globular protein chains in June 1995 release of PDB select. The results obtained from these analyses have been used to carry out an extensive study of sequence structure correlation in α helices (12,24). The trace of local helix origins obtained by HELANAL for the α helices can be used to study helix-helix packing. Besides these applications, the program may be used to analyse subtle change in geometries of helices in protein structures, due to changes in physical conditions, e.g., temperature, pressure and humidity, the changes in helix geometry during unfolding simulations of proteins, etc. HELANAL can also be used to characterise the helices formed in the crystals of *de novo* designed peptides. Though HELANAL has been used by us extensively to characterise only α helices, it can also be used to analyse β-strands as well as other single stranded helical structures. Results for a left handed ω-helix model structure (pdbomgl.ent, n = -4.0, h = 1.32Å), a β-strand in an ideal twisted sheet structure (pdbstrn.ent, n = -2.3, h = 3.3Å), as well as in a tyrosine kinase/peptide complex (pdb1lkk.ent, chain

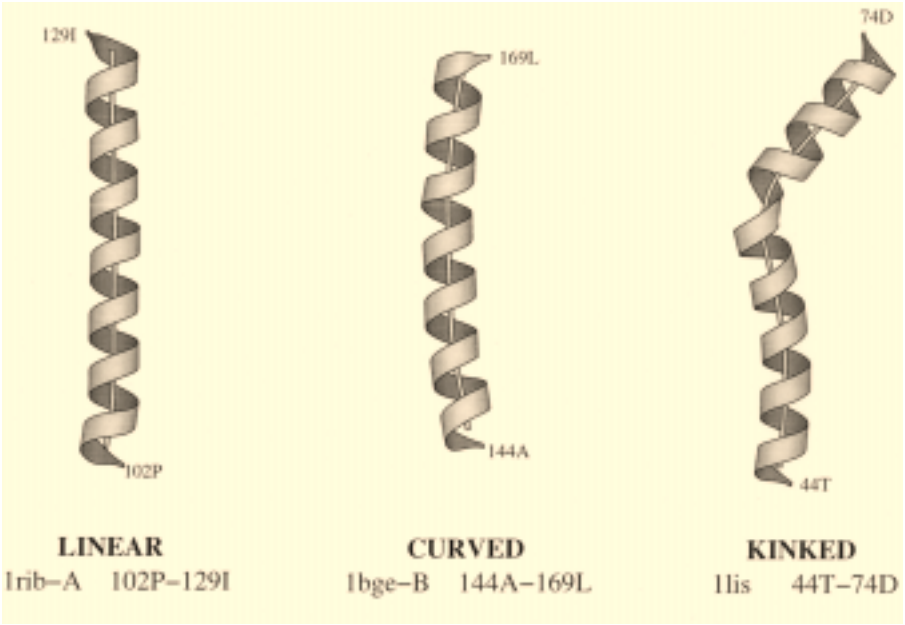

Figure 4: Representative examples of the three geometrical types of α helices (linear, curved and kinked), characterised using HELANAL. The MOLSCRIPT package (23) has been used to generate a ribbon diagram connecting C^{α} atoms in a helix with a line indicating the path traced by its local helix origins. Various parameters obtained by HELANAL for these helices are shown in Figure 3.

A, Glu 176 – Asn 185) are also shown in Figure 3. It is clear that HELANAL works well for all types of secondary structures. However, caution should be exercised in interpreting results for β-strands, since these structures tend to show local fluctuations in handedness, which is indicated by the large values for the bending angles. Similar large values are seen for bending angles at the ends of right handed α helices, when they terminate with a switch over to a left handed α helical structure, as in the well known Schellman motif. When the α helix is flanked by an extended structure, then the other helical parameters (such as rise and twist) give an indication of the helix termini.

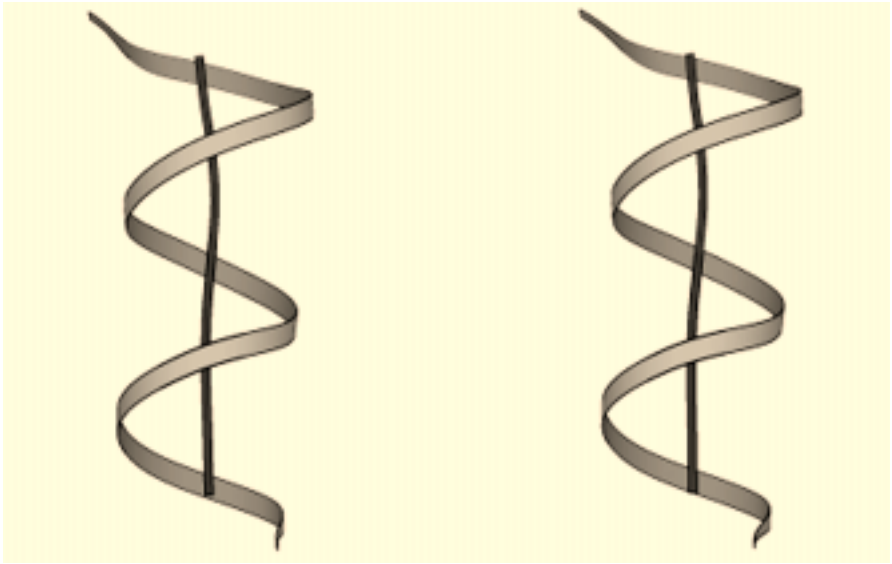

Figure 5: Stereo diagram of a helix (pdb1ars.ent, Met 397 - Val 405) showing the irregular nature of the local helix origins trace. This makes it difficult to characterise the geometry of the helix axis. Hence it is designated as 'unassigned' in Figure 3.

Acknowledgements

Sandeep Kumar acknowledges CSIR, India for fellowship. Velavan acknowledges DBT, India for Research associate fellowship.

References and Footnotes

1. L. Pauling, R.B. Corey and H.R.Branson, *Proc. Nat. Acad. Sci., U.S.A. 37*, 205-211 (1951).
2. J.C. Kendrew, G. Bodo, H.M. Dintzis, R.G. Parrish, H. Wycoff and D.C. Philips, *Nature 181*, 662-666 (1958).
3. W.F. Love, P.A. Klock, E.E. Lattman, E.A. Padhan, K.B.J. Ward and W.A. Hendrickson, Cold Spring Harbor Symp., *Quan. Biol. 36*, 349-357 (1971).
4. D. Barlow, and J.M. Thornton, *J. Mol. Biol. 201*, 601-619 (1988).

5. T.L. Blundell, D. Barlow, N. Barkakoti and J.M. Thornton, *Nature 306*, 281-283 (1983).
6. P. Chakarabarti, M. Bernard and D.C. Rees, *Biopolymers 25*, 1087-1093 (1986).
7. W.M. MacArthur and J.M. Thornton, *J. Mol. Biol. 264*, 1180-1195 (1996).
8. F.M. Richards and C.E. Kundrot, *Proteins 3*, 71-84 (1980).
9. F.C. Bernstein, T.F. Koetzle, G.J.B. Williams, E.F. Meyer, Jr., M.D. Brice, J.R. Rodgers, O. Kennard, T. Shimanouchi and M. Tasumi, *J. Mol. Biol. 112*, 535-542 (1977).
10. U. Hobohm and C. Sander, *Protein Sci. 3*, 522-524 (1994).
11. S. Kumar and M. Bansal, *Biophys. J. 71*, 1574-1586 (1996).
12. S. Kumar and M. Bansal, *Biophys. J. 75*, 1935-1944 (1998).
13. H. Sugeta and T. Miyazawa, *Biopolymers 5*, 673-679 (1967).
14. W. Kabsch and C. Sander, *Biopolymers 22*, 2577-2637 (1983).
15. G.N. Ramcahandran and V. Sasisekaran, *Adv. Prot. Chem. 23*, 283-437 (1968).
16. IUPAC – IUB commission on Biochemical Nomenclature. Abbreviations and symbols for the description of the conformation of polypeptide chain. Tentative rules, *J. Biol. Chem. 245*, 6489-6497 (1970).
17. R. Aurora and G.D. Rose, *Protein Sci. 7*, 21-38 (1998).
18. M. Levitt and J. Greer, *J. Mol. Biol. 114*, 181-239 (1977).
19. J.S. Richardson and D.C. Richardson, *Science 240*, 1648-1652 (1988).
20. T.J. Oldfield and R.E. Hubbard, *Proteins: Struct. Funct. Genet. 18*, 324-337 (1994).
21. D.A.D. Parry and E. Suzuki, *Biopolymers 7*, 189-197 (1969).
22. E.K. O'Shea, J.D. Klemm, P.S. Kim and T. Alber, *Science 254*, 539-544 (1991).
23. PJ. Kraulis, *J. Appl. Crystallogr. 24*, 946-950 (1991).
24. S. Kumar and M. Bansal, *Proteins: Struct. Funct. and Genet. 31*, 460-476 (1998).

Appendix A

QUESTIONS AND THEIR ANSWERS

Do you wish to use HELIX records in PDB files?
y
Do you wish to type in the PDB file names?
y
Key in a PDB file name: pdb1rib.ent
pdb1rib.ent DOES NOT CONTAIN HELIX RECORDS
Do you wish to use another PDB file?
y
Key in a PDB file name: pdb1bge.ent
pdb1bge.ent HELIX VAL A 49 LEU A 55
HELIX IGNORED: IT CONTAINS LESS THAN 9 RESIDUES.
pdb1bge.ent HELIX VAL B 49 LEU B 55
HELIX IGNORED: IT CONTAINS LESS THAN 9 RESIDUES.
RESULTS WRITTEN TO: pdb1bge.prm
AND IN TABULAR FORM TO: pdb1bge.tab
Do you wish to use another PDB file?
y
Key in a PDB file name: pdb1lis.ent..
Key in a PDB file name: pdb1lkk.ent
RESULTS WRITTEN TO: pdb1lkk.prm
AND IN TABULAR FORM TO: pdb1lkk.tab
Do you wish to use another PDB file?
n

Appendix B

QUESTIONS AND THEIR ANSWERS

Do you wish to use HELIX records in PDB files?
y
Do you wish to type in the PDB file names?
n
Key in name of the INPUT file containing PDB file names in (5x,a11) format e.g.
***.fil: input.fil

pdb1bge.ent HELIX VAL A 49 LEU A 55
HELIX IGNORED: IT CONTAINS LESS THAN 9 RESIDUES.
pdb1bge.ent HELIX VAL B 49 LEU B 55
HELIX IGNORED: IT CONTAINS LESS THAN 9 RESIDUES.
pdb1lis.ent HELIX PRO 116 ARG 123
HELIX IGNORED: IT CONTAINS LESS THAN 9 RESIDUES.
RESULTS WRITTEN TO: input.prm
AND IN TABULAR FORM TO: input.tab

Appendix C

QUESTIONS AND THEIR ANSWERS

Do you wish to use HELIX records in PDB files?
n
Key in name of the INPUT file containing helix start/end information
e.g. ***.inp: example.inp
Is the information in the same format as HELIX records in PDB files?
n
(a11,4x,2(a3,1x,a1,1x,i4,2x))
Is pdb1rib.ent a PDB file?
y
Is pdb1bge.ent a PDB file?
y
Is pdb1lis.ent a PDB file?
y
Is pdb1ars.ent a PDB file?
y
Is pdbcoil.ent a PDB file?
n
If not, key in the input data format: (Specify fields for atom name (4 characters), res. name, chain, res. number and XYZ coordinates. Use three letter code for residue names.)
(13x,a4,a3,1x,a1,i4,4x,3f8.3)
Is pdb2zta.ent a PDB file?
y
Is pdbomgl.ent a PDB file?
y
Is pdbstrn.ent a PDB file?
y
Is pdb1lkk.ent a PDB file?
y

RESULTS WRITTEN TO: example.prm
AND IN TABULAR FORM TO: example.tab

Date Received: December 8, 1999

Communicated by the Editor Ramaswamy H. Sarma