Quaternion Contact Ribbons: a New Tool for Visualizing Intra- and Intermolecular Interactions in Proteins

Kurt Albrecht†, John Hart†, Alex Shaw‡, and A. Keith Dunker‡

†Department of Electrical Engineering and Computer Science, ‡Department of Biochemistry and Biophysics, Washington State University, Pullman, WA 99164-4660

ABSTRACT: Protein side chain interactions between residues separated by at least one loop or turn or break in the amino acid sequence are called 'nonlocal contacts' in this manuscript, and contiguous sets of such interactions located between segments of secondary structure are called 'contact zones.' A new interactive program, the quaternion contact ribbon tool, has been developed to help protein chemists identify, straighten if twisted, and display contact zones between two neighboring segments of helix.

INTRODUCTION

Nonlocal contacts, which involve side chain interactions between residues separated by at least one turn, loop or break in the backbone, are located within the heirarchy of protein structure between elements of secondary structure and side chain conformations. The forces and energies acting across such contacts, including the hydrophobic effect, van der Waals interactions, electrostatic interactions, salt bridges, hydrogen bonds, and polarization effects, literally hold proteins together. Thus, the nonlocal contacts, and contiguous sets of such interactions that we are calling 'contact zones,' are literally at the nexus of protein structure and stability. Yet, there is a notable lack of identification and display of such non-local contacts, perhaps because appropriate tools for this purpose are lacking. Development of a computer tool for this purpose therefore would seem to be very important.

For the development of new computer tools, it is useful to apply the prototypes to well-understood systems. The contact zones between two almost parallel-helices have been selected for our intial studies because these are probabably the best-studied and best-understood of all the contact zones in proteins[1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14].

Previous Work: As early as 1953 Crick proposed that fibrous helical proteins would exhibit knobs-into-holes packing and would wind around each other into coiled-coils[15]. Later, we generalized the principles of knobs-into-holes packing[2] for idealized helical bundles intended to be possible mod-

els for structure and function of membrane proteins[3, 4, 6]. This exercise gave us considerable experience with helix/helix contact zones and provided us with the rationale for the approach to be presented here. Also, there are a number of interesting proteins, and peptides derived from larger proteins, that contain prominant helix/helix contact zones. These include several membrane proteins, the various proteins based on the "4-helix bundle motif"[16], the leucine zipper peptide[13, 17, 18] and its relatives[14, 19], and numerous other proteins. Thus, even for a tool restricted to the special case of helix/helix interactions, there are many interesting possible applications for the tool, once developed. Finally, there has been a recent burst of activity in the study of coiled-coils joined by knobs-into-holes packing[20] using a combined approach of chemical synthesis and physical characterization of the resulting structures[14, 19, 21, 22, 23, 24]. A well-designed tool for displaying helix contact zones should complement these studies.

Overview: In our previous modeling studies we developed a planar projection for the originally twisted contact zones in coiled-coils[2]. This was possible because the models were designed to fit this projection. Here we describe the quaternion contact ribbon tool, which has been developed to yield our planar projections of coiled-coil packing zones[2], but now the projection is applied to real structures rather than to idealized models. The purpose of this paper is to present a description of this new tool. Detailed studies of various helix/helix interactions using this this tool will be presented elsewhere.

MATERIALS

An SGI Indigo, using the VIEW software[25], was used to display the out-puts of our programs. All programs were written in the VIEW system language[25]. Color versions of the figures presented here are available for viewing at

`http://www.eecs.wsu.edu/~kalbrech/Q_Contact_Ribbons.`

The coordinates of GCN4[18], which is a typical leucine zipper[13] and which is used here to illustrate the method, were obtained from the Protein Data Bank[26, 27].

METHODS

Given a coiled-coil pair of helical segments (figure 1) and their axes (figure 2), development of the quaterninion contact ribbon consists of the following steps:

I Construction of chords joining the two helix axes (figure 3);

II Determination of the set of local coordinate frames to specify the contact ribbon (figure 4);

III Construction of the contact ribbon from the set of local coordinate frames (figure 5);

IV Transformation of the twisted contact ribbon into a flat surface using quaternions (figure 6);

V Display of the resulting quaternion contact ribbon with its associated, interacting side chains, with views down the three principal axes (figure 7, 8, 9).

Each of these steps will be considered in turn.

Construction of the connecting chords: Kahn's algorithm[28] for finding a helix axes yields a series of discrete points, one for each α-carbon in a peptide chain. We have developed a simpler version of Kahn's method (described in detail elsewhere[29]), which also represents helical axes as sets of points (figure 2).

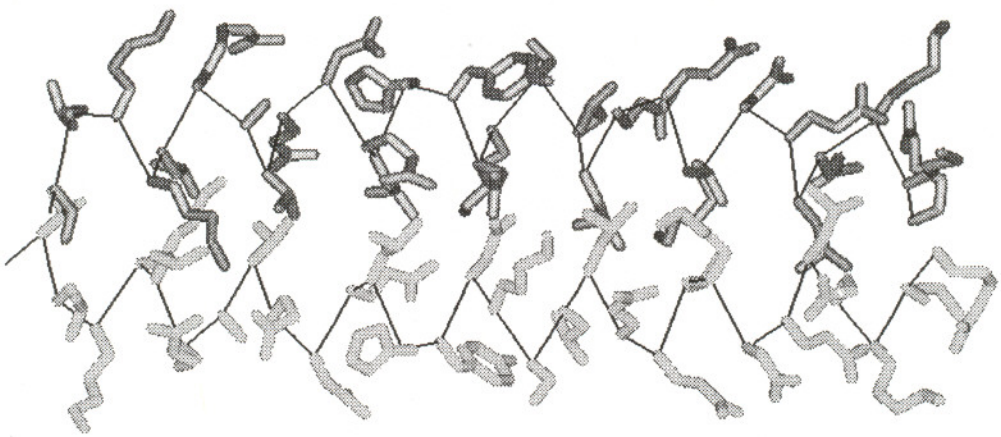

Figure 1: The contacts of a coiled-coil

A pair of helix axis is considered to be in local contact if nearby points on each axis are not too distant from each other (we set our threshold at 15 Å) and if the line through them is nearly perpendicular to the helix axes (we set this threshold to be within 18^o of perpendicular).

Once a contact pair has been found, the extent of the contact is determined by iterating down the helices simultaneously until a pair of helix points are either too far apart or the connecting chord is too divergent from right angles to the helix axes. A new chord is added to the set of connecting chords for each pair of helix points that meet the contact criteria. Such a collection of chords defines a helix/helix contact zone (figure 3).

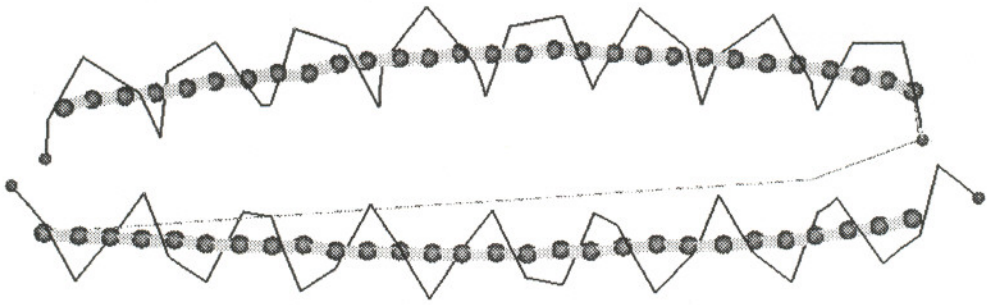

Figure 2: A pair of helical axes

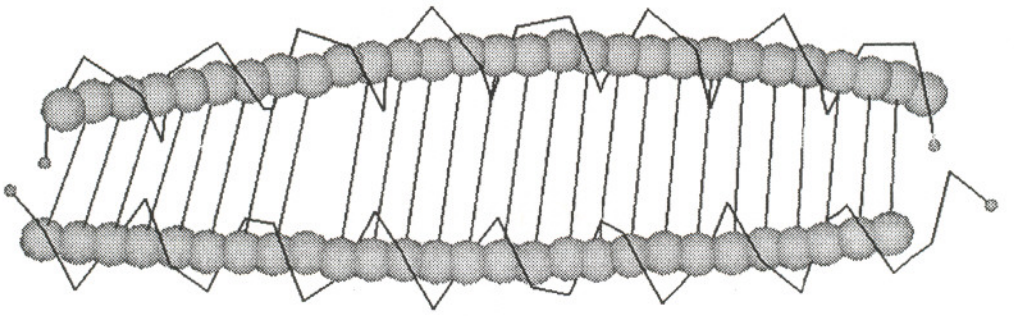

Figure 3: The contact zone displayed as a set of chords

Determination of the local coordinate frames for the contact ribbon: A local coordinate frame is a local cartesian coordinate system with an origin and three mutually orthogonal unit vectors. The three unit vectors are called the tangent, binormal, and normal vectors. For the ribbons drawn here the tangent vector lies along the center of the ribbon, the binormal vector points from the center toward an edge of the ribbon, and the normal vector points generally between the appropriate contact pair and is perpendicular to the ribbon. Figure 4 is an illustration of a set of local coordinate frames in a contact zone.

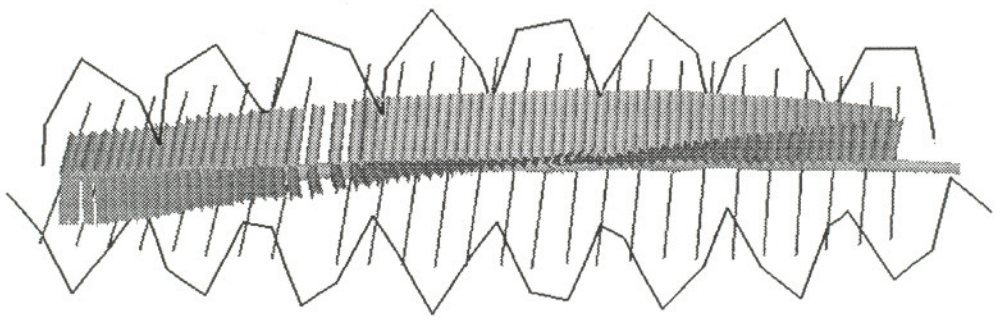

Figure 4: Local coordinate frames

Each chord in a contact zone yields a distinct local coordinate frame. Each midpoint of a contact chord is the origin of a local coordinate frame. One way to define a vector is as the difference of two points. The tangent vector is the difference between the midpoint of one chord and the midpoint of the next (or previous) chord. The binormal vector is the cross product of the tangent vector and the chord. The normal vector is the cross product of the tangent and binormal vectors.

Since a rough or jagged ribbon tends to distract from the scene, it is desirable to produce a smooth ribbon. To yield a smooth contact ribbon the set of local coordinate frames is interpolated. It is possible to interpolate the set of local coordinate frames by fitting interpolating splines through points appropriately offset from the chord midpoints. However, such a set of splines is difficult to manipulate and does not have a well defined relationship to the molecule under examination. It is preferable to interpolate the local coordinate frames themselves, which do have well defined relationships with individual atoms within the molecule and which are not difficult to manipulate.

Our scheme for interpolating a set of local coordinate frames is to first convert each frame into a point and an orientation. Since other methods of representing orientation have serious drawbacks and because the set of unit quaternions corresponds closely to the set of orientations, we represent each orientation with a unit quaternion[30, 31]. Secondly we interpolate the sets of points and quaternions separately, and finally convert the resulting sets of points and quaternions back into local coordinate frames. The result is a set of easily manipulated local coordinate frames that curve smoothly through the scene.

We interpolate the local coordinate frame origins using a modified Catmull-Rom spline[32, 33]. The Catmull-Rom spline is an easy to implement curve which interpolates its control points.

Since the orientation of a local coordinate frame can be represented as a quaternion we interpolated the set of quaternions using the same scheme mentioned above. It is important for quaternions describing orientations to have unit length. Because our spline method does not preserve length, we use the additional step of normalizing the resulting set of quaternions.

The Contact Ribbon: Offset points for each local coordinate frame are determined by finding the two points at a certain distance from the local coordinate frame origin lying along the extended binormal for that local coordinate frame. These points lay on the edge of the ribbon. A consecutive pair of local coordinate frames forms a quadrilateral with two ribbon points from each frame. The non-planar quadrilateral is filled in with a pair of triangles whose common edge is a diagonal of the quadrilateral. The ribbon is a collection of such pairs of triangles. Figure 5 is the contact ribbon for GCN4, a "leucine

zipper".

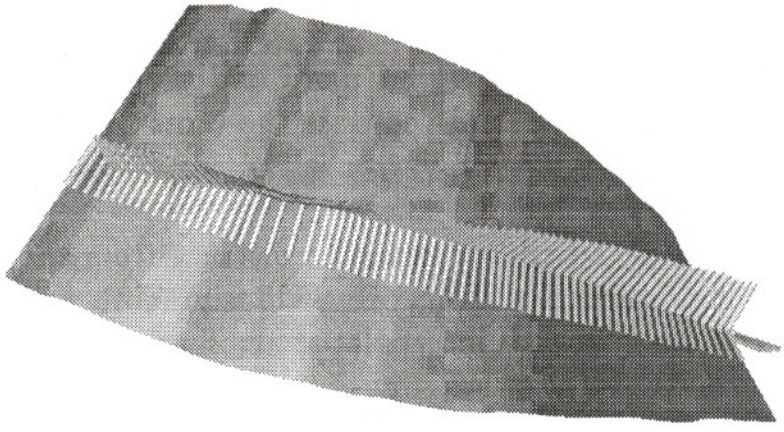

Figure 5: Formation of a contact ribbon

Transformation of the twisted contact ribbon into a flat surface using quaternions: Given a twisted contact ribbon described by an ordered set of local coordinate frames arranged along its spine, quaternions are used to rotate all the local frames into alignment thereby straightening the ribbon. The details of this transformation are given elsewhere[29].

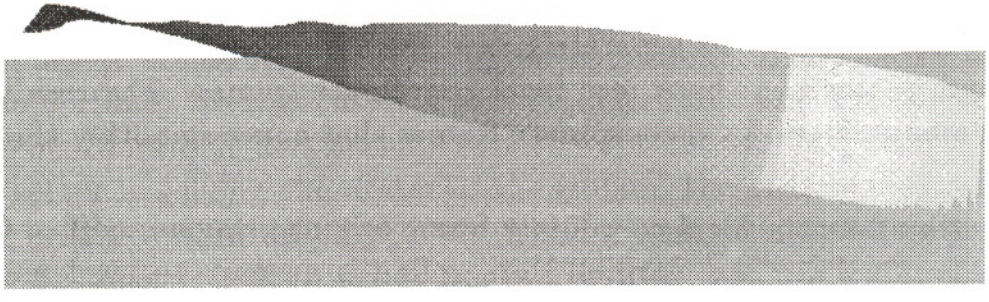

Figure 6: Flattening of a contact ribbon

Display of the quaterion contact ribbon with its associated, interacting side chains: In addition to their use in drawing manipulatable ribbons, the original set of local coordinate frames can contain information from the molecule.

Since each residue in a helix is associated with an helix point, and since each helix point in a contact is associated with a particular local coordinate frame, the position of each atom in a residue is mapped to a specific position with respect to its associated local coordinate frame.

The procedure for straightening the ribbon yields the set of atomic coordinates of the interacting side chains, but moved to be consistent with the straightened ribbon. The resulting coordinates and the contact ribbon can then be displayed by any appropriate molecular graphics tool. The results of

this procedure applied to the leucine zipper peptide are shown in figures 7, 8, 9.

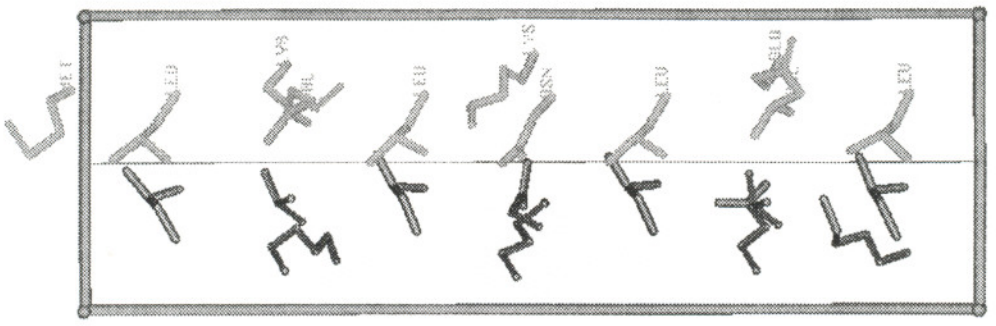

Figure 7: A flattened contact zone viewed edge on. The thin line at the center of the figure is the contact ribbon

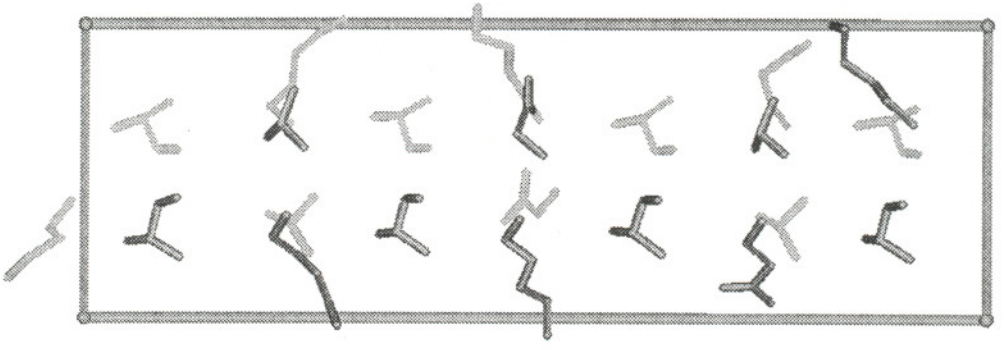

Figure 8: A flattened contact zone viewed face on, ie. the view is from one helix axis toward the other.

Discussion

Here we describe a method in which the geometry of difficult-to-visualize contact zones in a coiled-coil is captured by a twisted contact ribon, which is then flattened using quaternions. We have illustrated this approach using the well-studied GCN4 "leucine zipper" coiled-coil[12, 13, 14, 18, 19]

It has been noted previously that the luecine side chains of the GCN4 peptide have highly regular geometries owing to the overall symmetry of this particular helix contact[13, 17, 18]. The quaternion ribbon deformation obviously maintains the similarity of the leucine geometries as is evident in figure 3 - all the leucines appear to have essentially the same structure. We are in the process of carrying out a quatitative investigation of the amount of local distortion that is introduced by our new tool. Furthermore, in the studies

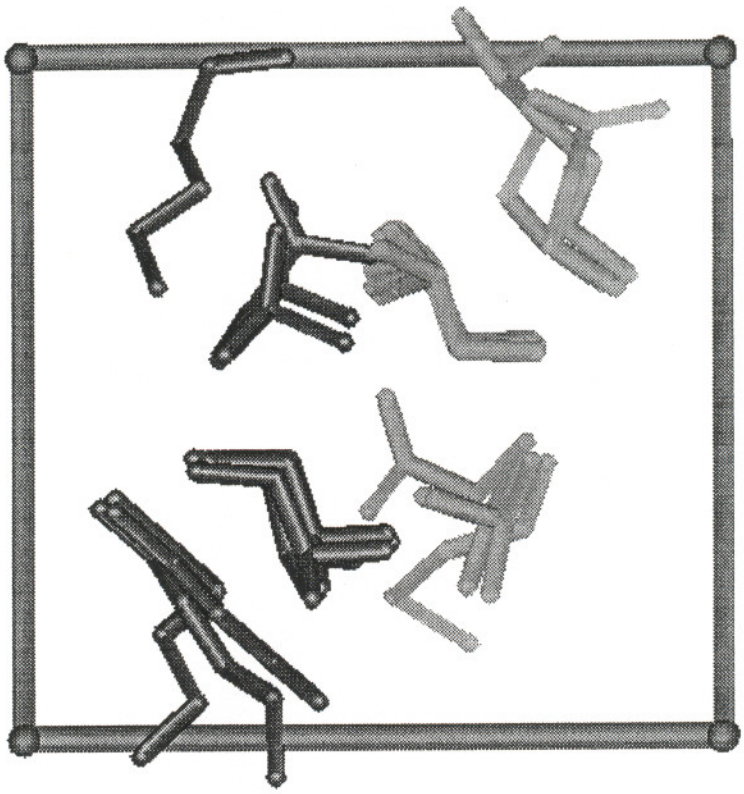

Figure 9: A flattened contact zone viewed end on

of the GCN4 contact, O'Shea et al[18] showed that there are two distinctive geometries for the orientation of a side-chain "knob" inserting into its corresponding hole. In one case, the bond from the α- to the β-carbon is almost perpendicular to the vector from residue (i) to (i + 1), which in turn is approximately parallel to the surface of the hole spanned by (i) and (i + 1), and in the second case the same bond is almost parallel to the same vector. Although not evident in the figures given here (and space limitations preclude additional figures to show this point, which will be given in detail elsewhere), these detailed geometries are well-preserved following the deformations induced by the quaternion contact ribbon tool. Thus, preliminary data suggest that the quaternion contact ribbon tool gives a reasonable preservation, not only of the local side chain contacts, but also of the geometric details of the interactions across the contacts.

Preliminary data suggests that, for a typical coiled-coil, the quaternion contact ribbon tool does not distort the local features beyond recognition, and so this tool should be useful for the investigation of the geometric details and organization of contact zones in proteins. But small distortions are not the same as no distortions, so any interesting feature observed in the simplified version of the contact should be checked out in the original, undistorted contact. For this reason we have constructed a feature that allows the user to define local

regions in the straightened contact zone and to click onto the corresponding region in the undistorted contact.

There have been previous attempts to identify and classify helix/helix contact zones[1, 2, 5, 7], but a rather limited data set was examined. It is evident from comparing figure 1 with figures 7, 7, 9 that our tool does greatly simplify the visualization of contact zones. It is noteworthy that the helix/helix contact zone shown in figures 7, 7, 9 for the "leucine zipper" gives a result identical to our original proposals for the planar projection of the coiled-coil interface[2].

A long term goal would be to modify our algorithms for other types of contact zones, thereby providing a basic tool that would enable the display and classification of all types of contact zones in proteins. If it were possible to represent all types of backbone structures by a set of points (one for each α-carbon, for example), then it would be straightforward to create the quaternion contact ribbons for all types of contact zones using the same logic we have presented here. There would have to be modifications to take into account that close pairs of β-strands could be due to hydrogen-bonding between the backbones or due to side chain interactions, but the orientations of the planar peptide groups would provide an easy way to distinguish these two possibilities. Recent papers show that Kahn like algorithms can be implemented to represent all types of backbone structure, yielding curves representing not only the axes of the helices but also the β-strands, turns, loops, etc[34, 35, 29]. Thus, we are confident that our basic approach for the development of contact ribbons can be extented for all types of contact zones within proteins.

References

[1] F.H. Crick. The Packing of α-Helices: Simple Coiled-Coils. *Acta Cryst.*, 6:689–697, 1953.

[2] A.K. Dunker and D.J. Zaleske. Stereochemical Considerations for Constructing alpha-Helical Protein Bundles with Particular Applications to Membrane Proteins. *Biochem. J.*, 163:45–57, 1977.

[3] A.K. Dunker and T. Jones. Proposed Knobs-Into-Holes Packing for Several Membrane Proteins. *Membrane Biochemistry*, 2:1–16, 1978.

[4] A.K. Dunker and D.A. Marvin. A Model for Membrane Transport through alpha-Helical Protein Pores. *J. Theor. Biol.*, 72:9–16, 1978.

[5] A.V. Efimov. Packing of α-Helices in Globular Proteins. *Journal of Molecular Biology*, 134:23–40, 1979.

[6] N. Burres and A.K. Dunker. Membrane Transport Through alpha-Helical Bundles. IV. Preliminary Model Building Investigation of Helix-Helix Interactions. *J. Theor. Biol.*, 87:723–736, 1980.

[7] C. Chothia, M. Levitt, and D. Richardson. Helix to helix packing in proteins. *Journal of Molecular Biology*, 145:215–250, 1981.

[8] T.J. Richmond and F.M. Richards. Packing of alpha-Helices: Geometrical Constraints and Contact Areas. *J. Mol. Biol.*, 119:537–555, 1978.

[9] K. Chou, G. Nemethy, and H.A. Scheraga. Energetic Approach to the Packing of alpha-Helices. 2. General Treatment of Nonequivalent and Nonregular Helices. *J. Am. Chem. Soc.*, 106:3161–3170, 1984.

[10] M. Gerritsen, K. Chou, G. Nemethy, and H.A. Scheraga. Energetics of Multihelix Interactions in Protein Folding: Application to Myoglobin. *Biopolymers*, 24:1271–1291, 1985.

[11] C. Cohen and D.A.D. Parry. alpha-Helical Coiled Coils and Bundles: How to Design an alpha-Helical Protein. *Proteins: Structure, Function, and Genetics*, 7:1–15, 1990.

[12] A. Tropsha, J.P. Bowen, F.K. Brown, and J.S. Kizer. Do Interhelical Side chain-Backbone Hydrogen-bonds Participate in Formation of Leucine Zipper Coiled Coils? *Proc. Natl. Acad. Sci. U.S.A.*, 88:9488–9492, 1991.

[13] W.H. Landschulz, P.F. Johnson, and S.L. McKnight. The Leucine Zipper: A Hypothetical Structure Common to a New Class of DNA Binding Proteins. *Science*, 240:1759–1764, 1988.

[14] P.B. Harbury, T. Zhang, P.S. Kim, and T. Alber. A Switch Between Two-, Three-, and Four-stranded Coiled Coils in GCN4 Leucine Zipper Mutants. *Science*, 262:1401–1407, 1993.

[15] L. Pauling, R.B. Corey, and H.R. Branson. The Structure of Proteins: Two Hydrogen-bonded Helical Configurations of the Polypeptide Chain. *Proc. Natl. Acad. Sci. U.S.A.*, 37:205–211, 1951.

[16] N.L. Harris, S.R. Presnell, and F.E. Cohen. Four Helix Bundle Diversity in Globular Proteins. *J. Mol. Biol.*, 236:1356–1368, 1994.

[17] R.V. Talanian, C.J. McKnight, and P.S. Kim. Sequence-specific DNA Binding by a Short Peptide Dimer. *Science*, 249:769–771, 1990.

[18] E.K. O'Shea, J.D. Klemm, P.S. Kim, T. Alber. X-ray Structure of the GCN4 Leucine Zipper, a Two-Stranded, Parallel Coiled Coil. *Science*, 254:539–544, 1991.

[19] P.B. Harbury, P.S. Kim and T. Alber. Crystal Structure of an Isoleucine-zipper Trimer. *Nature*, 371:80–83, 1 September 1994.

[20] C. Cohen and D.A.D. Parry. α-Helical Coiled Coils: More Facts and Better Predictions. *Science*, 263:488–489, 28 January 1994.

[21] I.L. Karle, J.L. Flippen-Anderson, K. Uma, M. Sukumar, and P. Balaram. Modular Design of Synthetic Protein Mimics. Crystal Structures, Assembly, and Hydration of Two 15- and 16-Residue Apolar, Leucyl-rich Helical Peptides. *J. Am. Chem. Soc.*, 112:9350–9356, 1990.

[22] P.C. Lyu, J.C. Sherman, A. Chen, and N.R. Kallenbach. alpha-Helix Stabilization by Natural and Unnatural Amino Acids with Alkyl Side Chains. *Proc. Natl. Acad. Sci. U.S.A.*, 88:5317–5320, 1991.

[23] N.E. Zhou, B.Y. Zhu, C.M. Kay, and R.S. Hodges. The Two-stranded alpha-Helical Coiled Coil is an Ideal Model for Studying Protein Stability and Subunit Interactions. *Biopolymers*, 32:419–426, 1992.

[24] B. Levejoy, S. Choe, D. Cascio, D.K. McRorie, W.F. DeGrado, and D. Eisenberg. Crystal Structure of a Synthetic Triple-stranded alpha-Helical Bundle. *Science*, 259:1288–1293, 1993.

[25] L.D. Bergman, J.S. Richardson, D.C. Richardson, and F.P. Brooks, Jr. VIEW - An Exploratory Molecular Visualization System with User-Definable Interaction Sequences. *Computer Graphics (Proceedings of SIGGRAPH 93)*, pages 117–126, 1993.

[26] 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. The Protein Data Bank: A Computer-based Archival File for Macromolecular Structures. *J. Mol. Biol.*, 112:535–542, 1977.

[27] E.E. Abola, F.C. Bernstein, S.H. Bryant, T.F. Koetzle, and J. Weng. *"Protein Data Bank" in Crystallographic Databases - Information Content, Software Systems, Scientific Applications*, pages 107–132. Data Commission of the International Union of Crystallography, Bonn/Cambridge/Chester, 1987.

[28] P.C. Kahn. Defining the axis of the helix. *Computers Chem.*, 13:185–189, 1989.

[29] K. L. Albrecht. The Use of Quaternions in the Modeling of Non-Local Contact in Proteins. Master's thesis, Washington State University, December 1995.

[30] K. Shoemake. Animating rotation with quaternion curves. *Computer Graphics (Proceedings of SIGGRAPH 85)*, 19:245–254, 1985.

[31] A. J. Hanson. Quaternion Frenet Frames: Making Optimal Tubes and Ribbons from Curves. Technical report, Computer Science Department, Indiana University, 1993.

[32] E. Catmull. *Subdivision Algorithm for the Display of Curved Surfaces.* PhD thesis, University of Utah, 1974.

[33] A. Watt and M. Watt. *Advanced Animation and Rendering Techniques Theory and Practice.* ACM press, 1992.

[34] H. Sklenar, C. Etchebest, and R. Lavery. Describing Protein Structures: A General Algorithm Yielding Complete Helicoidal Parameters and a Unique Overall Axis. *Proteins: Structure, Function, and Genetics*, 6:46–60, 1989.

[35] S. Swaminathan, G. Ravishanker, D.L. Beveridge, R. Lavery, C. Etchebest, and H. Sklenar. Conformational and Helicoidal Analysis of the Molecular Dynamics of Proteins: "Curves," Dials, and Windows for a 50 psec Dynamic Trajectory of BPTI. *Proteins: Structure, Function, and Genetics*, 8:179–193, 1990.