MODELLING CORRELATED MOTIONS IN PROTEINS AND IMPLICATIONS FOR PROTEIN REFINEMENT

F. R. Salemme+ and C. E. Kenknight*
+Genex Corporation
16020 Industrial Drive
Gaithersburg, Maryland 20877
*Department of Biochemistry
University of Arizona
Tucson, Arizona 85721

ABSTRACT

Refined protein structures frequently yield atomic temperature factors suggesting RMS atomic displacements of several tenths of Ångstroms or larger. Such large amplitude displacements result predominantly from rotations of groups about about single covalent bonds. Computer modelling studies of extended β-sheet structures found in proteins shows that correlated backbone torsional displacements can produce collective motions that propogate over extended regions of the protein. The present work models the time-average x-ray scattering associated with some representative collective motions in β-sheets and discusses how some of the unusual effects observed relate to protein x-ray refinement strategies.

I. INTRODUCTION

Numerous studies indicate that proteins in solution are dynamical molecules that undergo significant displacements in atomic positions at ambient temperatures.[1] Many properties that manifest protein dynamical behavior persist even when proteins are crystallized. For example, protein neutron diffraction studies show show that ostensably solvent inaccessible protons forming secondary structural hydrogen bonds can nevertheless exchange with protons from surrounding solvent.[2,3] Although the relatively large amplitude displacements leading to hydrogen exchange may only occur infrequently, they are indicative of the flexibility proteins possess even when constrained by crystal lattice interactions.

X-ray crystallographic

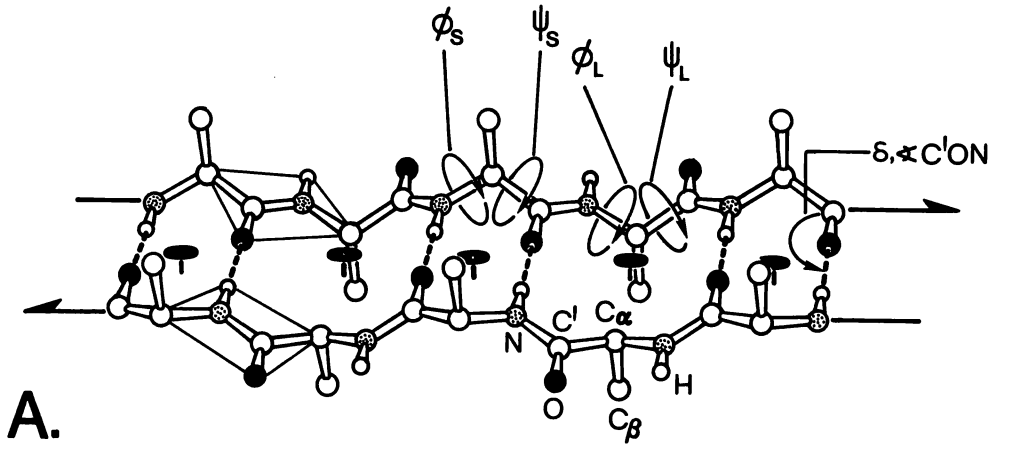
studies give direct evidence for the relative flexibility of proteins as these typically yield refined temperature factors suggesting dynamical displacements of several tenths of Angstroms.[4,5] Theoretical studies of protein dynamics indicate that such large amplitude displacements, while in part reflecting static disordering effects,[4] result mainly from torsional displacements about single bonds that characterize most of the proteins covalent structure.[6,7] Indeed, temperature factors computed by time averaging displacements along incremental steps of a molecular dynamics trajectory are generally in good agreement with crystallographically observed temperature factors.[8] More recently, Brooks and Karplus[9] have shown that many aspects of molecular motion apparent in molecular dynamics runs on pancreatic trypsin inhibitor (PTI) also manifest themselves in the normal mode behavior of the molecule. A feature of interest in the current context is the appearance of a number of low frequency vibrational modes corresponding to collective motions of the structure as a whole.

Recent work in the author's laboratory has focussed on the nature of cooperative motions that might occur in extended β-sheet secondary structures in proteins. This work illustrated that double-strand antiparallel sheets, such as occur in PTI and many other proteins, are potentially very flexible structures.[10] In fact, cooperative deformation of the β-sheet appear as determinative components of the lowest frequency molecular vibrational modes derived from the PTI normal mode analysis. Several questions then emerge concerning the nature and consequences of the time-average x-ray scattering from such structures in protein crystals. These specifically include how time-averaging the complex trajectories associated with cooperative motions may affect apparent structural geometry and temperature factor behavior. This is relevant to the interpretation of protein structural results because most current protein refinement approaches both constrain structural geometry and are limited to refining isotropic temperature facfactor behavior, a situation arising owing to the statistically unfavorable ratio of observed reflection data to

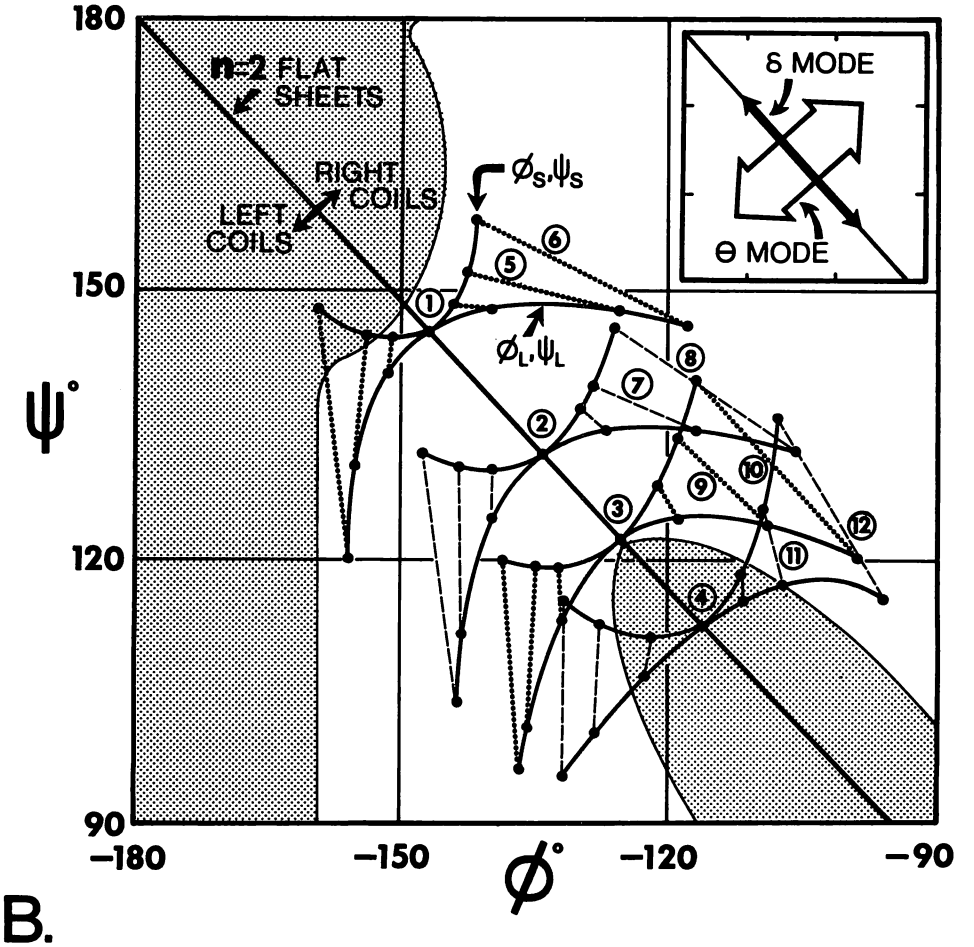
refined variables. Here we describe a computer modelling study that examines the effects of time-averaging some cooperative motions motions in a double-strand antiparallel β-sheet.

II. MODELLING CORRELATED MOTIONS IN ANTIPARALLEL β-SHEETS.

Figure 1A shows a short section of double-strand antiparallel β-sheet. The classical flat structure[11] has polypeptide chains that are 2-fold helices interconnected by straight interchain hydrogen bonds. Owing to the antiparallel N to C sense of the chains, residues situated on opposite sides of each of the "large" and "small" hydrogen bonded rings are related by a diad symmetry axis normal to the average sheet plane. In known protein structures, such sheets are invariably coiled in a right-handed sense.[10] This chiral coiling results in part from the structures composition of L-amino acids, which causes extended chains with local left-handed twist to be energetically favored relative to the 2-fold helical chains of a flat sheet.[9] Detailed studies of the conformational flexibility of this structure show that the ring diad symmetry elements and interchain hydrogen bonds can be maintained for a wide variety of alternative conformational states. These alternative states result from coupled alterations in both the polypeptide torsional angles and interchain hydrogen bond geometries (Fig. 1B) and, in fact, represent a continuum of cooperatively interconvertable coiled sheet conformations. As shown in Figure 2, the cooperative behavior can be factored into a cooperative compression of the structure along it's long axis (Qδ) and a cooperative helical coiling (Qθ). Using methods described elsewhere[12], we have computed the changes in potential energy associated with the interconversion of these structures to obtain the potential energy surface shown in Figure 3. As evident from inspection, the potential energy surface is reasonably continuous and smooth, with a low energy ravine corresponding to highly right-coiled states such as are actually observed in known protein structures.[10] Given the potential surface associated with the cooperative coiling of the structure allows the evaluation of a one-dimensional potential function associated with any desired path


ϕ_S
ψ_S
ϕ_L
ψ_L
δ,∝C'ON
N
C'
Cα
O
Cβ
H

A.


180
150
120
90
ψ°
−180
−150
−120
−90
φ°
n=2 FLAT SHEETS
RIGHT COILS
LEFT COILS
ϕ_S,ψ_S
ϕ_L,ψ_L
δ MODE
Θ MODE
1
2
3
4
5
6
7
8
9
10
11
12

B.

Fig. 1

of conformational interconversion between specific coiled conformations. For example, Figure 3 shows two displacement trajectories from the minimum potential energy structure. The first involves an excursion along the bottom of the low potential ravine that structurally corresponds to a spring-like axial compression of the β-sheet coil as a whole. The second involves a displacement of greater or less coiling, so corresponding to cooperative coiling and uncoiling of the structure as a whole. The one-dimensional potential functions associated with these motions can be used to evaluate the probabilities and amplitudes of thermally induced oscillations along a given path on the potential surface. This, in turn, allows an estimation of the time-average scattering behavior associated with the corresponding collective motions of the structure.

III. RESULTS

A. Displacement Trajectories

Figures 4 a-b show displacement trajectories associated with cooperative axial compression and coiling motions along the two potential surface paths shown in Figure 3. The displacements shown are based on least-squares superpositions of 2 X 12 residue double-strand sheet structures that minimize the positional deviation at the structures centers-of-mass (at the center of the sheet). The figures show only one-half of the structure for clarity. Consequently, displacement amplitudes shown increase from one end of the structure (actually the center of the 24 residue sheet) to the other. These illustrate the following points. a) Owing to the asymmetric nature of the energy surface governing excursions from the minimum potential structure (shown with full atomic backbone structure), oscillations approximating a given kinetic energy produce correspondingly different and asymmetric displacement amplitudes from the starting struc-

Figure 1. Part A shows a short section of flat double-strand antiparallel β-sheet. The structure can be viewed as an interconnected set of "large" and "small" hydrogen bonded rings formed of residues on adjacent chains related by two-fold symmetry axis in the sheet plane. Part B shows a ϕ,ψ plot that illustrates how the torsional angles of the "large" and "small" rings change in a correlated fashion on coiling the sheet, starting from representative flat structures (on the n=2 line) with different hydrogen bond geometries and polypeptide repeat periods.

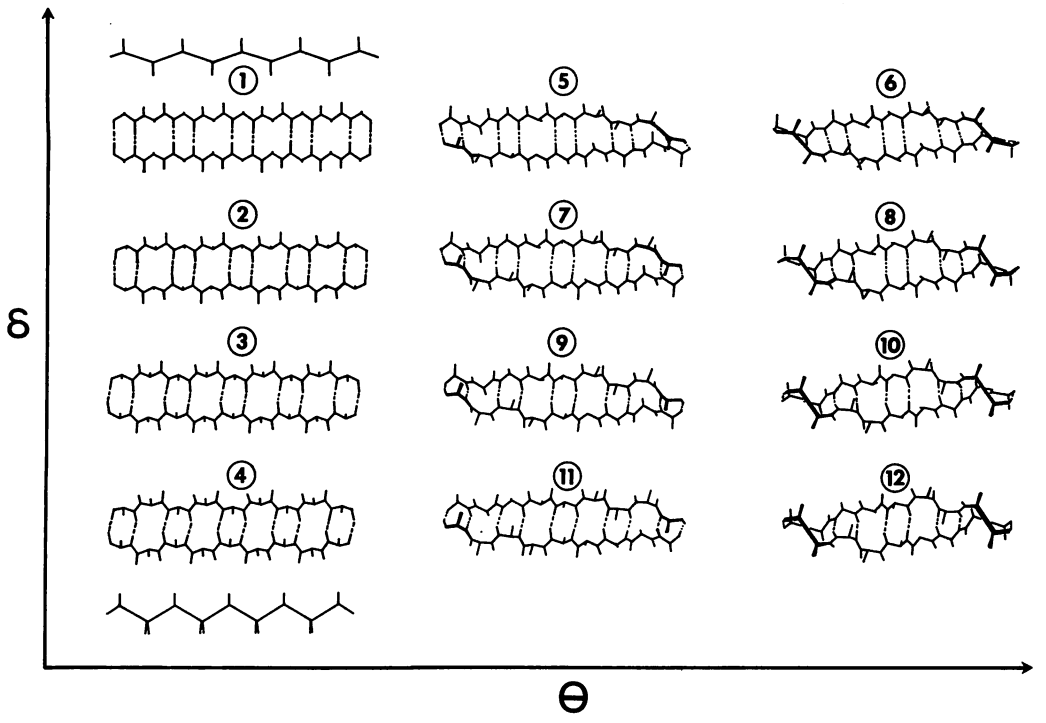


Figure 2. Variations of flat and coiled conformations in double-strand antiparallel sheets. The vertical (δ) axis shows structures that differ differ in extent of axial compression. The horizontal (θ) axis shows variation in extent of coiling. Numbers correspond to conformations shown in Figure 1B. Note that the repeating unit of the coiled structures is a dipeptide, characterized by a pair of (ψ,ϕ) values on the conformation plot.

ture. b) The trajectories followed by atoms may be curved. c) There is no uniform correlation between the direction of displacement vectors and features of local geometry, or throughout the structure as a whole. Note, for example, that while large amplitude motions at the sheet ends reflect the cooperative motions, small displacements near the structures superimposed centers of mass (see above) show no obvious correlation with displacement vectors elsewhere in the structure.

B. Modelling X-ray Scattering From Correlated Motions

Evaluation of an approximate partition function associated with the motions shown in Figure 4a-b allows a statistical estimation of the relative frequencies of occurance of displacements with different

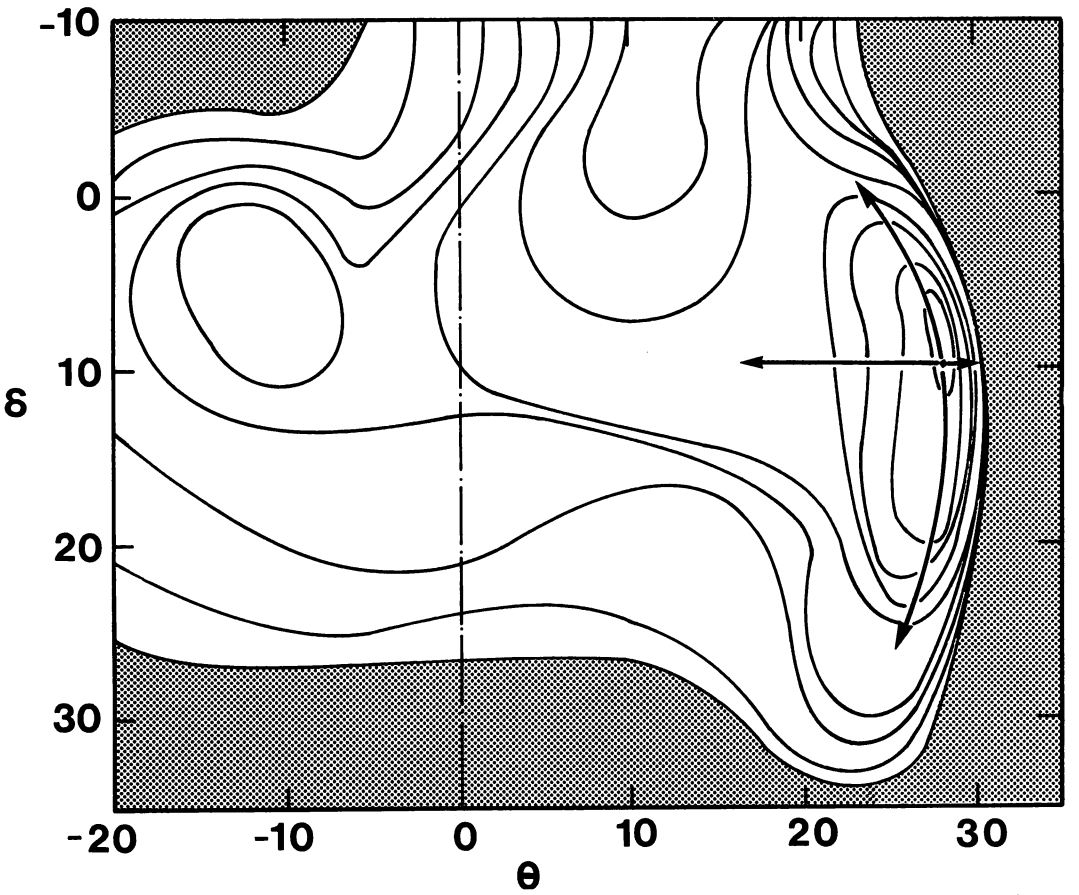


Figure 3. The potential energy surface associated with cooperative coiling in double-strand antiparallel β-sheet, given in terms of compression (δ) and coiling (θ), corresponding to Figure 3. The value θ=0 corresponds to flat sheets, and positive θ values to right-handed coils. The classical structure (Figure 1A) lies at 0,0. The shaded boundary is contoured at an energy of 2 kcal/mole residue, with succeeding contours at 0.1 kcal decrements. Excursions from the minimum potential structure occuring in the low potential ravine at the right of the plot are shown as double-ended arrows. The curved path corresponds to an axial compression of the structure approximating the interconversion of sheets 5, 8, 10, and 11 in Figure 2. The straight path approximates a coiling interconversion between structures 9 and 10 in Figure 2.

kinetic energies and amplitudes.[13] However, owing to the asymmetric and anharmonic nature of the potentials, an atom undergoing an oscillation of given amplitude will generally exhibit a correspondingly variable velocity. As a result, approximation of the time-average x-ray scattering incorporates both a statistical description of the distribution of states resulting in different amplitude oscillations, together with an estimate of dynamic atomic occupancies at given positions along the displacement trajec-

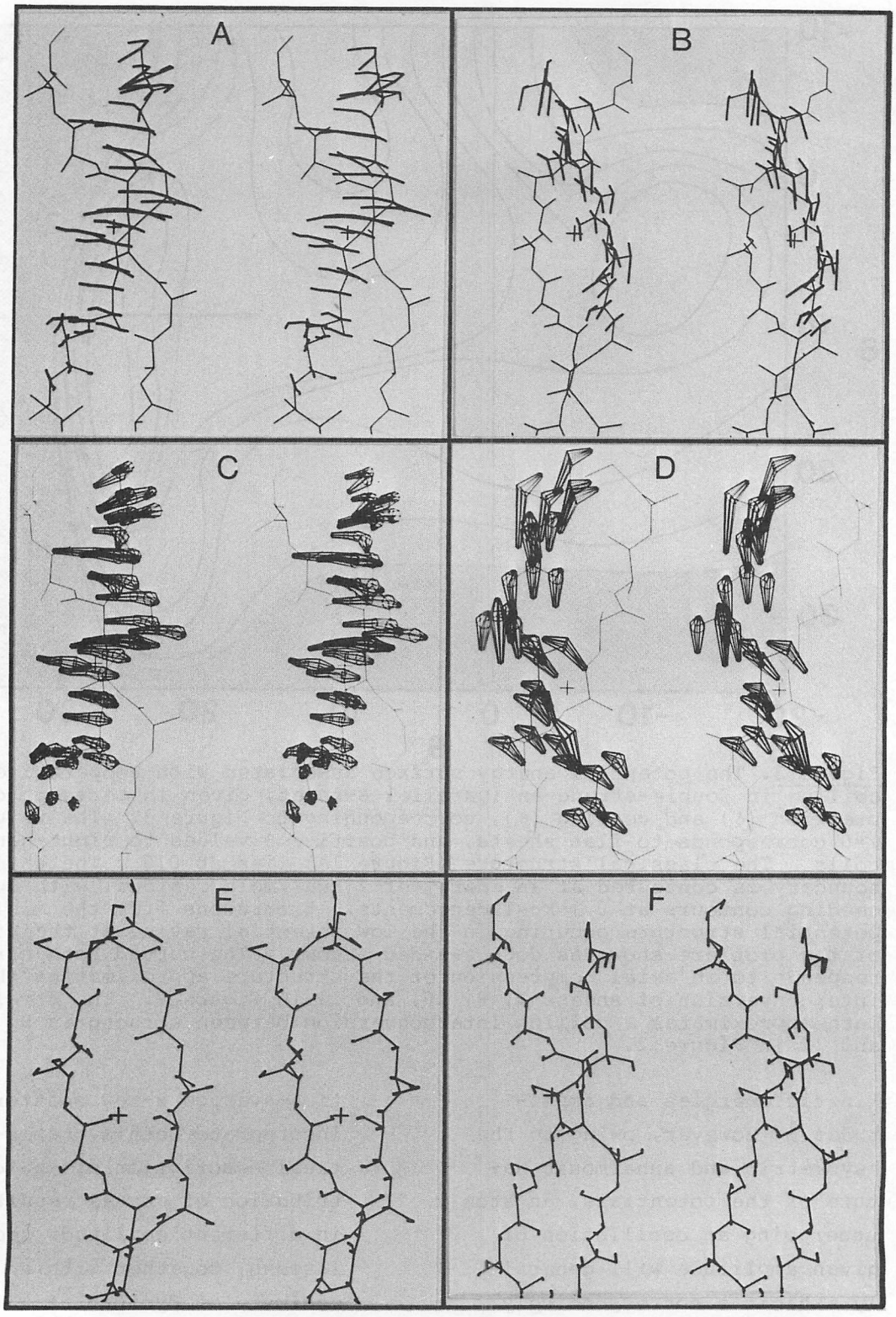


Fig.4

tory. Figures 4 b-c show the result of these averaging effects for the motions described above. Most notable are the unusual shapes of the time-average scattering distributions.

C. Geometrical Effects

Owing to the unfavorable ratio of observable reflection data to refinable structural parameters, most protein refinement protocols constrain structural geometry to be regular and model thermal motion isotropically.[14] However, it might be anticipated that the peculiar scattering behavior described above could produce time-average geometries that differed from regularity.[15] In order to test this possibility, the centers of mass of the scattering distributions shown in Figure 4c-d were connected to produce a structure approximating a result obtained from refinement incorporating isotropic temperature factors. Although some systematic bond-shortening effects were observed,[15] the resulting geometry was nevertheless quite regular, with RMS deviations on bond lengths of 0.02 to 0.03Å and 1 to 3° on bond angles. These values are generally within the constraint limits observed for highly refined protein structures.[4,5,14] However, when the scattering center-of-mass averaged structure was compared to the input minimum energy geometry (i.e. the starting geometry for the correlated displacements), RMS positional differences of 0.2 to 0.3 Å were observed (Figures 4 e-f).

IV. SUMMARY

The preceeding study was initially carried out to ascertain whether correlated motions in extended β-sheet structures gave rise to characteristic geometrical or temperature factor effects that might serve to identify regions undergoing cooperative motions in refined protein structures. The results presented suggest that this is not possible since spherically approximating the scattering from even the highly

Figure 4. Stereoscopic views showing the effects of correlated motions in β- sheets along trajectories shown in Figure 3. Parts a and b show the atomic trajectories followed by structures undergoing coiling (θ) and compression (δ) motions respectively. Parts c and d illustrate associated time-average scattering, and e and f show the positional differences between the minimum potential geometry and the centers-of-mass of the scattering distributions in c and d.

anisotropic and asymmetric scattering distribution can produce structures of acceptable geometry. Although some improvement might be made by utilizing fully anisotropic thermal parameters in protein refinement, no clear correlation appears to exist between the directions of local displacement vectors (that would presumably emerge as major axes of refined anisotropic thermal ellipsoids) in the time-average structure that would readily identify a given type of cooperative motion. Indeed, the larger amplitude asymmetric and curvlinear scattering distributions would be poorly approximated by even a fully anisotropic thermal ellipsoid approximation. On the other hand, the observation that center-of-mass averaging such asymmetric scattering distribution can produce structures with orthodox geometry, that nevertheless differ from the "true" minimum potential energy geometry, provides some rationalization for the coordinate shifts typically observed when refined protein coordinates are energy minimized.[5] What ultimately seems to be required to accurately refine a dynamic protein structure is a means of approximating the time-average scattering with an orthogonal function set that may better reflect the actual collective motions in the protein. One possible, although computationally intensive, approach might expand the electron density as a superposition of displacements in the structures normal modes.[8] In this case it might be hoped that only a small fraction of the total number of modes would predominate in producing the large amplitude motions typically inferred from protein temperature factor behavior. This could, in turn, effect an improvement in the ratio of observed data to parameters in the refinement process.

ACKNOWLEDGEMENTS:

Work in this laboratory is supported by NIH Grants GM30393 and 33325. The graphics algorithm used in Figure 4cd was developed by M. Rouald.

REFERENCES

1. Richards, F. M., in Neutrons in Biology (B. Schoenborn, ed.) Plenum (1984) pp. 1-23.

2. Kossiakoff, A. A. Nature 296, 713721 (1982).

3. Wlodawer, A. & Sjolin, L. Proc. Natn, Acad. Sci. U.S.A. 79, 1412-1422 (1982).

4. Frauenfelder, H. Petsko, G. A., Tsernoglou, D., Nature 280, 558-562 (1979).

5. Artymiuk, P. J., Blake C. C. F., Grace, D. E. P., Oatley, S. J., Phillips, D. C. and Sternberg, M J. E., Nature 280, 563-568 (1979).

6. McCammon, J. A., Gelin, B. R. and Karplus, M. Nature 277, 585-580 (1977).

7. Levitt, M. Nature 294, 379-380 (1981).

8. Northrup, S. H., Pear, M. R., McCammon, J. A., Karplus, M., and Takano, T., Nature 287, 659-60 (1980).

9. Brooks, B., and Karplus, M., Proc. Nat. Acad. Sci. U.S.A. 80, 6571-6575 (1981).

10. Salemme, F. R., Prog. Biopys. Molec. Biol. 42, 95-133 (1983).

11. Pauling, L. & Corey, R. Proc. Nat. Acad. Sci. U.S.A. 37, 729-740 (1951).

12. KenKnight, C. E., and Salemme, F. R. (in preparation)

13. Peticolas, W. L. Meth. Enzymology 61, 425-456, (1979).

14. Hendrickson, W. A. & Konnert, J. H. in Computing in Crystallography (eds. Diamond, R., Ramaseshan, S. & Venkatsen, K.) 13.01-13.23 (Indian Academy of Sciences, 1980).

15. Busing, W. R., and Levy, H. A. Acta. Cryst. 17, 142-146 (1964).

DISCUSSION

Q. (F.L. Carter, Naval Research Laboratories) Karplus alluded earlier to some nonlinear effects. I have a question that either or both of you might address. Davydon, in 1976, suggested that solitons may be signal carriers (of ATP bond breaking) down the alpha helix. In these convoluted structures that have been discussed today does one see any signs of nonlinear effects such as solitons in the short α-helicies or β-sheets?

A: There are really two questions there. One is whether or not pro-

tein molecules are like -sheets or are aspects of protein molecules like -sheets? The second is would you expect to see such effects in the X-ray data. The fact is that the kinds of things that I have been describing probably represent statistically varying regions in the structures. Thus although they are interesting and significant in structural terms, it would not be appropriate, necessarily, to interpret in the terms you are asking about.

A: (Karplus, Harvard University) My feeling is that in most of the protein structures effects like this are lost. The model calculations that people have done have been highly idealized with very limited degrees of freedom so that there is little chance of seeing subtle effects.