A HYBRID APPROACH TO THEORETICAL ANALYSIS OF BOVINE PANCREATIC TRYPSIN INHIBITOR

THOMAS THACHER AND HERSCHEL RABITZ
Department of Chemistry, Princeton University, Princeton, New Jersey 08544

ABSTRACT A static analysis of bovine pancreatic trypsin inhibitor (BPTI) is presented based on a new discrete/continuum approach to modeling the dynamics of biomolecules. This hybrid method utilizes knowledge of the intramolecular potential and molecular configuration to generate a field of elastic modulus tensors. These tensors, which relate the local stress and strain for each atom in the biomolecule, can be used to judge the local rigidity as well as indicate regions of high stress. Comparing the tensor fields for an unrelaxed and a relaxed configuration, the microscopic structure of BPTI is found to be anisotropic and to have regions of stress even when it is relaxed in the potential field. However, when these fields are averaged over the whole protein or over individual residues the structure becomes more isotropic and the stressed regions vanish. Using these averaged tensors, we calculated bulk properties such as Young's modulus and the Lamé constants and they agreed with previously reported values.

I. INTRODUCTION

Proteins are dynamical systems, and although the static configuration of a protein provides useful insight into the functions of proteins, if considered alone it is incapable of explaining all observed properties and functions. For that task a more integrated approach must be taken in which both the statics and dynamics are considered together. Whereas this point of view has long been recognized for processes such as protein folding and diffusion, it has only recently been emphasized for other processes such as conformational transformation, active and passive transport, and enzymatic catalysis (1–6).

This is due in part to the inherent complexity of studying the dynamics of large biomolecules. A typical biomolecule may have on the order of 10,000 atoms, so 30,000 equations of motion must be integrated to treat the dynamics properly. Because the biomolecule will not in general be symmetric, the number of degrees of freedom cannot be reduced through symmetry considerations. As a result methods that are used to study dynamics of small molecules, such as normal mode analysis (7) and molecular dynamics (8), are computationally expensive if not intractable for many biomolecules (9–11) unless constraints are imposed on them to reduce the number of degrees of freedom.

In addition to many degrees of freedom, the frequencies associated with the dynamical processes occurring in a biomolecule span an extremely wide spectrum ranging from 3,000 to 3.0×10^{-8} cm^{-1}. Although this complicates the complete description of the dynamics, it suggests an approximate piecewise approach with the dynamics being separated into regimes that are uncoupled from each other because of large differences in frequencies. If such a separation is valid, then each regime may be treated by a technique suitable for that particular regime (9). Whereas this separation may be used effectively on the very high and very low frequency regimes, its application to the midrange regimes is not as straight forward because the separation of neighboring regimes may be ambiguous. Further, it is not clear that the complete dynamical picture of the biomolecule may be constructed from the segmented descriptions.

The objective of this paper was to illustrate some of the features of a new hybrid approach to modeling macromolecules. The purpose of this approach is not to supplant the already established molecular dynamics techniques but rather to augment them so that they can be applied to larger macromolecular systems. Furthermore, the approach has the flexibility to allow one to match the complexity of the model with the sensitivity of the calculated output to the details of the model. By taking advantage of this, one can simplify the model in a systematic manner to yield a more efficient simulation. This hybrid method begins with an atomistic picture of the molecule in which the configuration and the intramolecular potential are given for the n atoms. Using this description, we calculate a field of position-dependent elastic modulus tensors (12, 13). By indicating the degree of local rigidity, the field can be used as a probe to lump the molecule into regions with uniform elastic properties. Regions of the molecule undergoing large amplitude anharmonic motion which cannot be properly described by the continuum model (i.e., the active site) are left unlumped and treated using the original atomistic picture. Thus, the molecule is represented by an amalgamation of both the discrete and continuum regions all linked together. These latter continuum regions may or may not have differing elasticities. The amount of lumping depends on the kind of information

being sought as well as the dynamic behavior of the molecule. For example the calculation of bulk elastic properties such as the Lamé constants may allow lumping over the entire molecule because these properties probe the average behavior of the molecule. On the other hand, to characterize some higher frequency behavior that is more sensitive to local structure, a lumping scheme that is less global in nature must be used. This lumping scheme can be adapted to be responsive to the local changes in the elastic properties and thus made suitable to model the behavior.

The elastic tensors also have specific symmetry properties (13) associated with the invariance of the unstressed continuum with respect to rigid body rotation and translation. Tensors that exhibit deviations in these properties indicate that the biomolecule is stressed at those points. These properties have an important application to the common occurrence of a discrepancy in the true static structure of the biomolecule and the equilibrium configuration associated with the intramolecular potential. To overcome this difference the field of tensors can be symmetrized so that they obey the correct symmetry properties. The resulting continuum equations of motion correspond to a self-consistent equilibrium configuration and intramolecular potential. This method of obtaining self-consistent equations of motion is better than previous approaches because it requires virtually no additional computational effort. However, it should be pointed out that the symmetrized continuum equations of motion are not the same as the self-consistent atomistic equations of motion and there is no direct way of obtaining the latter from the former.

After presenting a brief discription of the formalism, the field of elastic modulus tensors is calculated for two different configurations of BPTI, one of which is relaxed in the potential force field. To illustrate the flexibility of the hybrid approach, we analyzed the tensor field for the discretized continuum, the continuum, and a partially lumped intermediate model. In this latter model, which is only one of many possible lumping schemes, the residues are averaged over in a uniform manner. In addition to examining the traces and deviations in the Voight symmetry properties for all the models, bulk properties such as Young's modulus are calculated for the continuum and the lumped intermediate models. It should be emphasized that even though the field of tensors may be used in the above manner to describe both the local and global structure of the biomolecule, the primary purpose of the field is to serve as a means to lump the biomolecule in a consistent manner so that dynamic simulation becomes practical. While the latter calculations are relegated to future studies, an estimation of the dynamical response of the protein limit is determined for the continuum limit.

II. THEORY

Because a detailed theoretical formulation of the approach is given elsewhere (14), only a brief summary of it will be presented here. To obtain the elastic modulus tensors in terms of the intramolecular potential, two assumptions are made. First, the biomolecule is assumed to be sufficiently close to equilibrium to allow the use of a harmonic intramolecular potential of the form

$$V = V' + \frac{1}{2}\sum_{ia}\sum_{jb} V_{ia,jb}\, u_{ia}\, u_{jb}, \tag{2.1}$$

where V' is the equilibrium potential energy and u_{ia} is the ath component of the displacement of the ith atom. In this expression, the term $V_{ia,jb}$ is the second derivative of the potential, with respect to the ath displacement component of i and the bth displacement component of the atom j, evaluated at the equilibrium configuration. This term may be interpreted as the force on the ith atom in the a direction when the jth atom is displaced a unit distance in the b direction. In writing equation 2.1 it has been assumed that the term which is linear in the displacement vanishes because at equilibrium there are no net forces on the atoms.

Second, the displacement of each atom in the biomolecule is assumed to be small with respect to the distance from its neighboring atoms so that the bth component of the displacement of the jth atom may be expanded in a Taylor series about the displacement at the ith atom,

$$u_{jb} = u_b(r) + \left[\sum_c \frac{\partial}{\partial X_c} u_b(r)\right]\Delta X_{jc} + \frac{1}{2}\sum_c\sum_d \left[\frac{\partial^2}{\partial X_c X_d} u_b(r)\right]\Delta X_{jc}\Delta X_{jd}. \tag{2.2}$$

In writing Eq. 2.2, we have replaced the discrete variable i by the continuous variable r. The terms, ΔX_{jc} and ΔX_{jd}, are the cth and dth components, respectively, of the distance between r (corresponding to the ith atom) and the jth atom. These restrictions may be relaxed if necessary, but at the present they provide a good working model. By using a classical atomistic description of the biomolecule, the equations of motion for the displacement of the ith atom in the ath direction can be written in terms of the intramolecular potential as

$$\rho(r)\,\ddot{u}_a(r) = -\frac{1}{v}\sum_j\sum_{bc}[V_{a,jb}\,\Delta X_{jc}]\frac{\partial}{\partial X_c}u_b(r) + \frac{1}{2v}\sum_{bc}\sum_d [abcd]\frac{\partial^2}{\partial X_c \partial X_d}u_b(r), \tag{2.3}$$

where the bracketed quantity is defined as

$$[abcd] = -\sum_j V_{a,jb}\,\Delta X_{jc}\Delta X_{jd}. \tag{2.4}$$

Also, in this expression v is an appropriate characteristic volume such that $\rho(r)$ is the local mass density at point r given by $m(r)/v$.

Eq. 2.3 is valid in the long wavelength limit where the wavelength of the displacement of the atoms is much larger than the distance between the atoms. In this limit the

biomolecule may be considered to be an elastic continuum and an equation of motion for the displacement of the continuum at point r can be written in terms of a local elastic modulus tensor as

$$\rho(r)\ddot{u}_a(r) = \frac{1}{2}\sum_b\sum_{cd}\frac{\partial}{\partial X_b}\left[C_{abcd}(r)\frac{\partial}{\partial X_d}u_c(r) + C_{abcd}(r)\frac{\partial}{\partial X_c}u_d(r)\right], \quad (2.5)$$

where C_{abcd} is an element of the elastic modulus tensor. If the elastic continuum is unstressed so that it is invariant to rigid body rotation and translation, the tensors obey the so-called Voight symmetry properties: $C_{abcd} = C_{bacd} = C_{abdc} = C_{cdab}$ (13). Using these properties, the above expression may be simplified to

$$\rho(r)\ddot{u}_a(r) = \sum_b\sum_{cd}\frac{\partial}{\partial X_d}C_{adcb}(r)\frac{\partial}{\partial X_c}u_b(r). \quad (2.6)$$

In addition, the symmetry properties allow an equivalent form of Eq. 2.5 to be written as

$$\rho(r)\ddot{u}_a(r) = \sum_b\sum_{cd}\frac{\partial}{\partial X_c}C_{acbd}(r)\frac{\partial}{\partial X_d}u_b(r), \quad (2.7)$$

so that the ath component of the volume force at r is given as

$$\rho(r)\ddot{u}_a(r) = \sum_{bc}\sum_d\left[\frac{\partial}{\partial X_d}C_{adcb}(r)\right]\frac{\partial}{\partial X_c}u_b(r) + \frac{1}{2}\sum_{bc}\sum_d[C_{acbd}(r) + C_{adcb}(r)]\frac{\partial^2}{\partial X_c\partial X_d}u_b(r). \quad (2.8)$$

By equating expressions 2.3 and 2.8, the local elastic modulus tensor for each atom can be written in terms of the intramolecular potential as

$$C'_{abcd} = [acbd] + [bcad] - [abcd], \quad (2.9)$$

where the volume-modified elastic modulus tensor C'_{abcd} is defined by

$$C'_{abcd} = vC_{abcd}. \quad (2.10)$$

Note that the summation implicit in the coefficients in Eq. 2.10 are over all atoms except i. In practice, the summation extends to some cutoff distance (~8.0 Å) where contributions to the sum from interactions beyond this distance are negligibly small. Finally, using Eq. 2.9 and the Voight symmetry properties associated with elastic modulus tensors, it can be shown that

$$[abcd] = [cdab]. \quad (2.11)$$

This symmetry property, known as the Huang condition (12), follows from the invariance of the potential to rigid body rotation.

Each of the elastic modulus tensor elements represents the constant of proportionality between elements in the stress and strain tensors. Thus the field of elastic modulus tensors contains a great deal of information about the direction and magnitude of local stresses throughout the molecule. By using the former two symmetry properties, the original 9 × 9 tensor can be rigorously reduced to a 6 × 6 tensor, C'_{ab}. According to the normal convention the elements of the reduced tensor are ordered as follows: $\{1,1\} \rightarrow [1]$, $\{2,2\} \rightarrow [2]$, $\{3,3\} \rightarrow [3]$, $\{2,3\} = \{3,2\} \rightarrow [4]$, $\{1,3\} = \{3,1\} \rightarrow [5]$, $\{1,2\} = \{2,1\} \rightarrow [6]$. Furthermore, if the equilibrium symmetry property of the reduced tensor (i.e., $C'_{ab} = C'_{ba}$ or equivalently $C_{abcd} = C_{cdab}$) is also employed, then the resulting field of tensors corresponds to the field that would be found if the configuration matched the equilibrium configuration for the intramolecular potential.

III. RESULTS

Using Eqs. 2.9 and 2.4 and an intramolecular potential given by the following equation (15),

$$V = \sum_{bds} k_b(d - d_o)^2 + \sum_{bend\ \measuredangle} k_\theta(\theta - \theta_o)^2 + \sum_{tor\ \measuredangle} |k_\phi| - k_\phi\cos(n\phi) + \sum_{\substack{nb\\pairs}}\left\{\frac{A_{ii}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} + \frac{q_iq_j}{4\Pi\epsilon_o r_{ij}}\right\}, \quad (3.1)$$

the field of elastic modulus tensor was determined for two different configurations of BPTI. One configuration corresponds to the x-ray crystallographic structure (16). The other configuration was found by calculating the intramolecular forces on all the atoms in the crystallographic structure and then allowing them to move to minimize these forces. This minimized structure represents the local minima associated with the potential which is closest to the

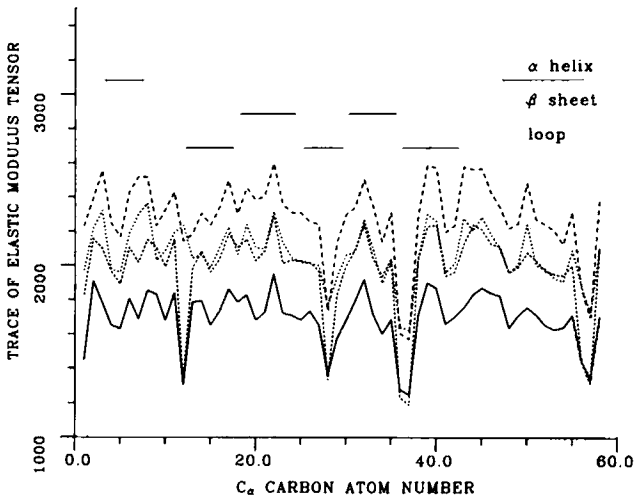

FIGURE 1 Traces of both the full and symmetrized elastic modulus tensors for α-carbons in the peptide backbone of the unrelaxed configuration of BPTI. In all the figures the region of α-helix, β-sheet, and loop are shown by labeled horizontal lines. (–) C'_{abcd} assuming valence interactions; (– – –) C'_{abcd} assuming valence and nonbonded interactions (· · · ·) C'_{ab} assuming valence interactions; (— – — –) C'_{ab} assuming valence and nonbonded interactions. Units are in kilocalories per mole.

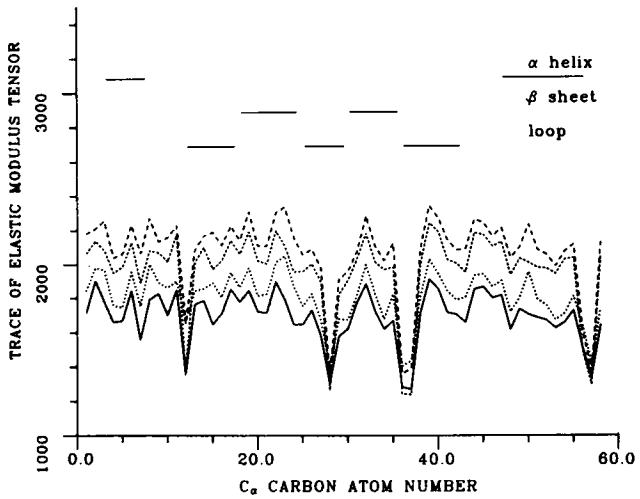

FIGURE 2 Traces of both the full and symmetrized elastic modulus tensors for α-carbons in the peptide backbone of the relaxed configuration of BPTI. (–) C'_{abcd} assuming valence interactions; (– – – –) C'_{abcd} assuming valence and nonbonded interactions; (· · · ·) C'_{ab} assuming valence interactions; (— — —) C'_{ab} assuming valence and nonbonded interactions. Units are in kilocalories per mole.

crystallographic structure. It should be emphasized that this configuration is not only potential dependent but it is also not unique to the potential because there are many local minima. In this respect, the configuration is somewhat arbitrary. However, the structure is representative of other local minima and can be used to illustrate the nuances of the approach. Hereafter we will refer to this structure as the relaxed configuration and the original crystal structure as the unrelaxed configuration. The minimization procedure (17) yielded a root-mean-square force of 6.054×10^{-1} kcal/Å when averaged throughout the molecule in contrast to a value of $2.396 \times 10^{+2}$ for the "unrelaxed" configuration.

Discretized Continuum Limit

Even though the trace of the elastic modulus tensor does not directly correspond to a physical property, it is a useful qualitative indicator of the tensor field which is invariant to the orientation of the biomolecule. The field of tensors for both the full and symmetrized tensor associated with the α-carbons in the peptide backbone are displayed in Figs. 1 and 2 for both configurations using two different forms of the potential. The valence potential calculation assumed only the first three terms in Eq. 3.1, whereas the full potential calculation considered the nonbonded interaction in addition to the valence terms. As is apparent from the figures and Table I, which lists the averages and standard deviations for the α-carbon traces, the valence contributions to the traces are much larger than the nonbonded contributions. If the valence forces are only considered, the field of traces for the full tensors show large changes in flexibility in the transition between looped regions and a β-sheet region whereas the smaller minima at 12 and 57 correspond to transitions out of an α-helix region. Beyond this there does not seem to be any correlation between the field of tensors and the substructure of the biomolecule. This lack of correlation is also apparent in the calculations assuming the complete interaction potential given by Eq. 3.1. These same conclusions can be drawn from plots of the elastic modulus tensor field for the amide nitrogens in the peptide backbone (Fig. 3). The large spikes in the plots which occur at the proline position indicate the increased rigidity because of the constraints of the ring structure. Thus at this level the tensors reflect the local bonding environment.

It had been anticipated that the different substructures within the protein could be characterized by different elastic properties which would be reflected in the tensor field. The reason that such a correlation is not found is due to the local nature of field. That is, the traces represent the rigidity on an atomistic level whereas the elastic properties describing the substructures are much more macroscopic in nature. Because of this a better comparison can be made by lumping the field into macroscopic regions and then contrasting the traces of these regions with the substructure. This approach is considered later in the paper.

The symmeterized tensors shown in Figs. 1–3 were

TABLE I
AVERAGE VALUE AND CORRESPONDING STANDARD DEVIATION FOR TRACES OF BOTH THE FULL AND SYMMETRIZED ELASTIC MODULUS TENSOR

Atom type	Full tensor		Reduced tensor	
	Relaxed BPTI	Unrelaxed BPTI	Relaxed BPTI	Unrelaxed BPTI
	kcal/mol	*kcal/mol*	*kcal/mol*	*kcal/mol*
Valence interactions				
CH	1,734.8(65.1)	1,740.4(69.5)	2,040.1(65.4)	2,053.6(72.9)
NH	1,476.1(113.0)	1,469.5(123.2)	1,533.6(199.1)	1,504.7(207.6)
Valence and nonbonded interactions				
CH	1,844.0(71.5)	2,899.6(87.5)	2,140.8(69.6)	2,347.7(93.4)
NH	1,649.5(106.1)	1,691.5(133.8)	1,688.1(185.7)	1,829.6(207.7)

SD is shown in parentheses.

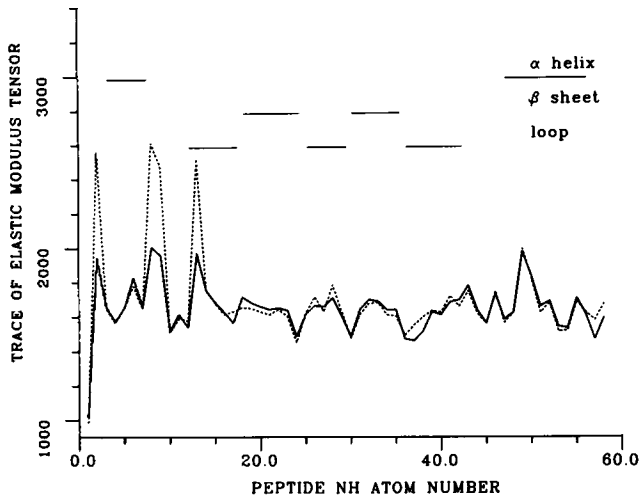

FIGURE 3 Traces of both the full and symmetrized elastic modulus tensors for the amide nitrogen atoms in the peptide backbone of the relaxed configuration of BPTI. (——) C'_{abcd} assuming valence and nonbonded interactions; (– – – –) C'_{ab} assuming the same interactions. Units are in kilocalories per mole.

determined by averaging the elements associated by the Voight symmetries. Whereas this averaging allowed the reduction of the 9 × 9 tensor to a 6 × 6, the overall behavior of the traces for the symmeterized tensors did not change markedly from those for the unsymmeterized tensors. The profile of the traces for the two configurations remained virtually unchanged in the relaxation process. Even though they show the same overall behavior, the traces of the full and symmetrized tensors will not in general be equal to each other. The trace of the full tensor is the sum of the elements in the upper left 3 × 3 submatrix of the symmetrized tensor. Consequently, the difference in the behavior of the two traces is due to the off-diagonal, elements in the submatrix, and the diagonal elements of the lower right 3 × 3 submatrix. It is satisfying that the overall behavior of the traces is preserved when the tensors are symmetrized. The result stems from the fact that the hybrid approach uses the trace as a scalar quantity to characterize the microscopic elasticity at that atom. Thus, if there were large changes in this quantity, symmetrization of the tensor field would be difficult to justify.

To summarize the information from the tensor fields, the averages of the traces and their associated standard deviations are listed in Table I for both the α-carbon and amide nitrogens. The average values for both the full and reduced tensors, calculated using only the valence interactions, did not change when the molecule was relaxed. However, when both valence and nonbonded interactions were considered the relaxation caused a reduction in the value of the trace. This suggests that the contribution to the local forces from the valence forces do not change as much as the nonbonded forces. A comparison of the amide nitrogen and the α-carbon values show that the former group has both smaller average trace values and larger

TABLE II
TRACE OF AND DEVIATIONS IN THE VOIGHT SYMMETRIES IN THE FULL 9 × 9 AND THE SYMMETRIC 6 × 6 AVERAGED ELASTIC MODULUS TENSOR FOR BPTI

Property	Unrelaxed BPTI		Relaxed BPTI	
	a	*b*	*a*	*b*
Symmetry				
$S_{abcd} - S_{bacd}$	0.011	0.011	0.011	0.011
$S_{abcd} - S_{abdc}$	0.161	0.190	0.087	0.091
$S_{abcd} - S_{cdab}$	0.121	0.143	0.066	0.070
Traces				
S_{abcd}	1.35	1.58	1.34	1.49
S_{ij}	1.51	1.72	1.51	1.64

Deviations × 10^3/454 kcal per mol. (*a*) Stretching interaction (3.1). (*b*) Bending, torsional, and nonbending interactions.

standard deviation about the average values. Thus, the amides are more flexible and have local environments that are more varied. This latter observation was found to be true even when the proline groups responsible for the sharp peaks in Fig. 3 were excluded from the averages.

Continuum Limit

Although valuable local information may be obtained by considering the traces for each tensor throughout the molecule, the calculation of the elastic properties from these tensors is not justified because these constants are macroscopic quantities. Thus to obtain bulk elastic constants for the molecule, the tensors are averaged over the whole field using

$$S_{abcd} = \frac{1}{V}\sum^{n} C'_{abcd}(i). \tag{3.2}$$

In Eq. 3.2, n is the number of atoms averaged over and V is the total volume of BPTI. Table II lists the deviations in the Voight symmetries for both the relaxed and unrelaxed configurations as a function of the two different potential

TABLE III
ELASTIC CONSTANTS FOR BPTI

Constants	Unrelaxed BPTI		Relaxed BPTI	
	a	*b*	*a*	*b*
Lamda	0.96	1.24	0.97	1.14
Mu	1.64	1.79	1.64	1.76
Bulk modulus	2.06	2.43	2.07	2.32
Young's modulus	3.88	4.30	3.88	4.20
χ_i	0.18	0.30	0.17	0.18
[Young's modulus]x	4.10	4.42	4.08	4.40
[Young's modulus]y	3.87	4.15	3.87	4.18
[Young's modulus]z	4.21	4.62	4.20	4.54
⟨Young's modulus⟩	4.07	4.40	4.05	4.38

Constants × 10^4 MPa. (*a*) Stretching interaction (3.1). (*b*) Bending, torsional, and nonbending interactions.

forms. Table III lists the corresponding elastic constants which will be discussed later in the paper.

The symmetry property that exhibits the smallest deviation is $S_{abcd} = S_{bacd}$, which is associated with the requirement

$$\sum_i \left\{ \frac{\partial}{\partial a_i}\frac{\partial}{\partial b_j} V(r)\Big|_{r'} - \frac{\partial}{\partial b_i}\frac{\partial}{\partial a_j} V(r)\Big|_{r'} \right\} \Delta X_{jc}\Delta X_{jd} = 0, \quad (3.3)$$

whereas $S_{abcd} = S_{cdab}$, which is larger, is associated with the Huang condition, Eq. 2.11,

$$\sum_j \left\{ \frac{\partial}{\partial a_i}\frac{\partial}{\partial b_j} V(r)\Big|_{r'} \Delta X_{jc}\Delta X_{jd} - \frac{\partial}{\partial c_i}\frac{\partial}{\partial d_j} V(r)\Big|_{r'} \Delta X_{ja}\Delta X_{jb} \right\} = 0. \quad (3.4)$$

The property that uniformly shows the largest deviation is $S_{abcd} = S_{abdc}$ and corresponds to both of the previous requirements. While the traces for the tensor sums, S_{abcd} and S_{ij}, of the relaxed and unrelaxed configurations are not appreciably different, the deviations in the latter two symmetry properties are significantly different. The former symmetry property, $S_{abcd} = S_{bacd}$, is initially very small for the unrelaxed configuration and seems to be insensitive to the relaxation process. This in fact should be expected because expression 3.3 is invariant to the change in the indices of the second derivative if the sum is over all atoms. Thus as the number of atoms summed over increases, the relative contribution from the term for the ith atom decreases and expression 3.3 approaches zero. Because the ratio of the third to the second symmetry property is constant for both the relaxed and unrelaxed configuration, either property can be used equally well to indicate the internal stress of the molecule. For both types of interactions, the deviations in the second and third symmetry properties decrease when the protein is relaxed. It is interesting to note that the larger decrease is seen when only the valence interactions are considered. This suggests that the degrees of freedom associated with bond stretching and bending converge to their equilibrium configuration more quickly than do the other degrees of freedom.

The Bulk and Young's moduli which are representative of the whole protein can be calculated from the averaged tensor by assuming that it is isotropic so that the Lamé constants, λ and μ, are given by,

$$\lambda = \tfrac{1}{3}\,[C_{12} + C_{23} + C_{13}] \quad (3.5)$$

and

$$\mu = \tfrac{1}{3}\,[C_{44} + C_{55} + C_{66}], \quad (3.6)$$

respectively (18). Judging from the symmetrized tensor in Table IV one can see that this is not an unreasonable

TABLE IV
FULL AND SYMMETRIZED ELASTIC MODULUS TENSOR FOR THE RELAXED CONFIGURATION OF BPTI ASSUMING VALENCE AND NONBONDED INTERACTIONS

C_{abcd}/454 kcal per mol

b	x	x	x	y	y	y	z	z	z
d	x	y	z	x	y	z	x	y	z
ac									
xx	405.7	−7.4	6.6	−7.4	145.4	6.9	6.6	6.8	154.5
xy	−7.0	95.4	4.8	145.0	−6.3	−0.9	7.1	−1.9	−5.2
xz	6.7	5.5	100.4	7.8	−0.3	−2.3	154.5	−5.3	5.5
yx	−7.4	145.4	6.9	95.8	−5.8	−0.8	4.6	0.1	−5.4
yy	145.0	−6.3	−0.9	−6.3	382.9	9.8	−0.9	9.9	137.2
yz	7.7	−0.4	−2.2	−1.3	9.5	91.2	5.3	137.1	9.3
zx	6.6	6.9	154.6	4.6	0.1	−5.5	100.3	−2.2	5.0
zy	7.1	−1.8	−5.2	−0.9	9.8	137.2	−2.0	91.3	9.6
zz	154.6	−5.4	5.6	−5.5	137.1	9.2	5.5	9.3	416.2

$C_{ab} \times 10$/454 kcal per mole

a-	xx	yy	zz	xy	xz	yz
xx	406.0					
yy	95.6	383.0				
zz	100.0	91.3	416.0			
xy	4.9	9.8	9.3	137.0		
xz	6.6	−1.5	5.9	−5.3	155.0	
yz	−7.3	−6.2	−2.2	−5.1	7.2	145.0

assumption. Clearly the off-diagonal 3 × 3 submatrix has elements which are of much lower magnitude than the diagonal submatrices. Further, the off-diagonal elements of the lower diagonal 3 × 3 matrix are much smaller than the diagonal elements. These constants are related to the Bulk and Young's moduli by

$$K = \lambda + \tfrac{2}{3}\mu \quad (3.7)$$

and

$$E = 9K\mu/(3K + \mu). \quad (3.8)$$

In addition, because the diagonal elements of the upper left 3 × 3 submatrix are equal to the quantity $(2\lambda + \mu)$, the following expression can be used to judge the isotropy of the protein:

$$\chi_i = \tfrac{1}{3}\,[(C_{11} + C_{22} + C_{33}) - 2\,(C_{12} + C_{23} + C_{13}) - (C_{44} + C_{55} + C_{66})]. \quad (3.9)$$

Using this criterion the protein can be considered to be reasonably isotropic because the value of χ_i is about an order of magnitude smaller than the constants for both configurations. Additionally, if the protein is isotropic the difference in the traces for the full and symmetrized tensors will be given by $3(2\lambda - \mu)$. As can be seen from Tables II and III there is very good agreement between this quantity and the difference in the traces. Young's moduli can also be calculated along each of the axes from the

expression (18)

$$E_j = \left\{ \sum_i [\underline{S}]_{ij}^{-1} \delta_{ij} \right\}, \tag{3.10}$$

where $i, j \epsilon \{x, y, z\}$. In general it is found that the spacially averaged Young's moduli in Table III are larger than the values calculated assuming that the protein is isotropic. However this difference is quite small, (<10%), so that the values are comparable with each other. The range in the values along each axis is ~3% of the average when only valence interactions were considered for both configurations. If all the interactions are considered the range for the unrelaxed configuration increases to 5% of the average value and then decreases to 4% when the configuration is relaxed.

The calculated value 4.2×10^3 Mpa of Young's modulus for BPTI was found to be in excellent agreement with the $4–6 \times 10^3$ Mpa experimentally determined value for hen egg white lysozyme (19). Our value is better than the $1.2–2.3 \times 10^3$ Mpa range estimated from static correlation of the fluctuations in the dihedral angles of alpha-helicies (20).

Because of the agreement of the calculated modulus with the experimental one and the structure of the averaged tensor, (Table IV), it is appropriate to treat the protein on this level as being isotropic even though it is clearly anisotropic on a microscopic level (Figs. 1–3). To take the analysis a step further, a simple estimate of the molecules' dynamical response can be made from the first order assumption that the protein is a sphere of radius r. Taking the radius as 20 Å (21) then the frequencies of vibration are given by the roots of the dispersion relation

$$\tan\left(\frac{wr}{C_\ell}\right) = \left[\frac{wr}{C_\ell}\right]\left[1 - \left(\frac{wr}{2C_t}\right)^2\right]^{-1}. \tag{3.11}$$

C_ℓ and C_t in the above expression are longitudinal and transverse wave velocities related to the Lamé constants by

$$C_\ell = \sqrt{\mu/\rho} \tag{3.12}$$

and

$$C_t = \sqrt{2\lambda + \mu/\rho}, \tag{3.13}$$

respectively. Using the values determined for the relaxed configuration and assuming all interactions except stretching the lowest vibrational frequency is 11.2 cm^{-1}. Whereas this estimate is in reasonably good agreement with previously reported values (21–23), recent normal mode calculations find frequencies in the 3 cm^{-1} range. The discrepancy between these results and the present one is most likely due to the fact that BPTI is not spherical and the crude volume averaging used above can mask some types of motion.

Homogeneous Lumped Intermediate

The lumping or averaging of the tensor field throughout the protein represents the full continuum limit of the hybrid approach. As an illustration of another lumping scheme intermediate to this limit and the rather unphysical discrete limit where each tensor in the field is considered individually, we have chosen to calculate the elastic constants from the tensors averaged over each residue and its corresponding nearest neighbor residues. Because the profiles of the traces of both the full and symmetrized tensors are virtually the same for both configurations only the traces for the relaxed conformation are displayed in Fig. 4.

Comparing the results shown in Fig. 4 with Fig. 2, it is seen that there is a more marked correlation between the averaged traces and the substructures of the molecule. When the traces are averaged over the substructure regions it is found that the looped region (residues 25–29) is the most flexible. The β-sheet regions (residues 18–24 and 29–35) which flank this looped region are stiffer. Finally, the α-helical region (residues 47–55) is a little stiffer than the looped region (residues 12–17 and 36–42) but more flexible than the β-sheet regions.

Assuming that the protein is isotropic at each residue, Young's modulus may be calculated from the averaged tensor field using Eqs. 3.5–3.8. Although the individual moduli differ slightly for each configuration, the overall profile does not. This same trend is also exhibited in a comparison of the traces for the elastic modulus tensors (Fig. 4) and the elastic moduli (Fig. 5), which suggests that the traces are qualitatively proportional to Young's modulus if the protein is not too anisotropic.

The root-mean-square displacement of an atom in a

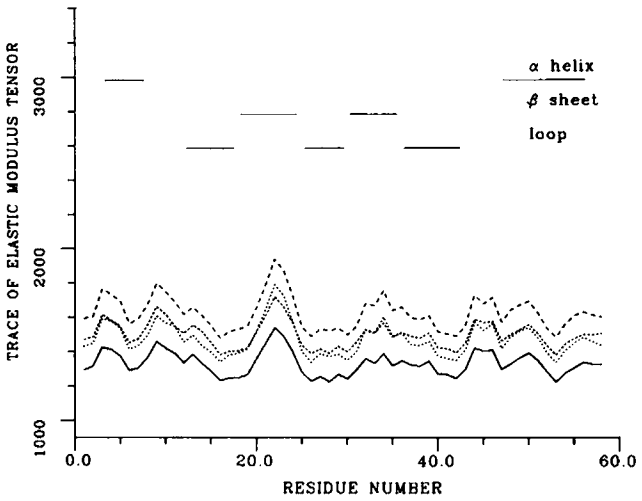

FIGURE 4 Traces of both the full and symmetrized elastic modulus tensors, averaged over the residues, for the relaxed configuration of BPTI. (——) C'_{abcd} assuming valence interactions; (– – – –) C'_{abcd} assuming valence and nonbonded interactions; (· · · ·) C'_{ab} assuming valence and nonbonded interactions. Units are in kilocalories per mole.

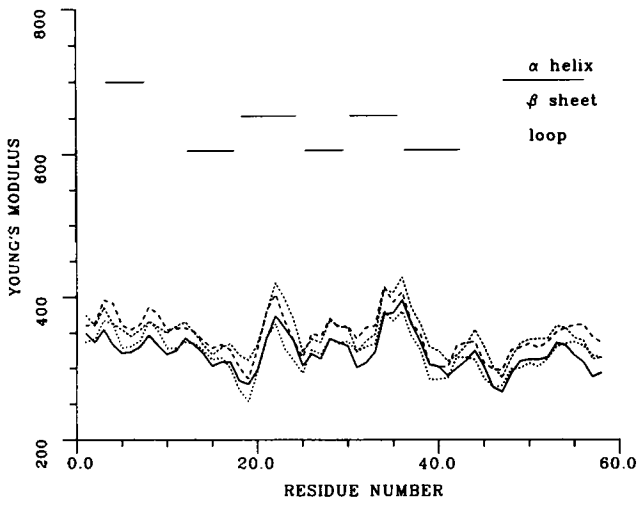

FIGURE 5 Young's modulus calculated from the symmetrized tensors, averaged over the residues. (——) The unrelaxed configuration assuming valence interactions; (– – – –) the unrelaxed configuration assuming valence and nonbonded interactions; (· · · ·) the relaxed configuration assuming valence interactions; (— — — —) the relaxed configuration assuming valence and nonbonded interactions.

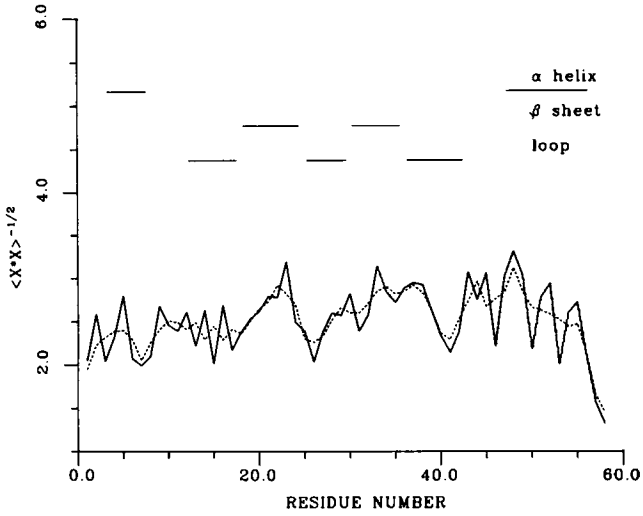

FIGURE 6 The normalized inverse of the root-mean-square displacement calculated from the temperature factors for BPTI. (——) The values averaged over one residue, (· · · ·) values averaged over three residues.

molecule is related to the temperature factor determined though x-ray diffraction by the following expression (3, 4).

$$\langle u_i^2 \rangle^{1/2} = \frac{1}{2\pi} (B_i/2)^{1/2}. \tag{3.14}$$

Here $\langle u_i^2 \rangle^{1/2}$ and B_i are the rms displacement and temperature factor, or the Debye-Waller factor for the ith atom. The implicit assumptions used in writing Eq. 3.14 is that the motion of each atom is harmonic in nature and that it is isotropic. In lieu of a rigorous calculation of the temperature factors, which requires the integration of the lumped continuum equations of motion, we compare the reciprocal of the rms displacement to the trace and Young's modulus. It is expected that because the latter quantities measure the local rigidity, they should be correlated to the former. The trace and Young's modulus, which are measures of the rigidity, will be correlated to the reciprocal of the rms displacement. To compare the two averaged quantities, the temperature factors (16) are averaged over the residues in the same manner as was done for the tensor field. These averaged values are shown in Fig. 6. The comparison of Fig. 6 to Fig. 1 and 2 demonstrates little correlation between the traces of the tensor fields and the inverse rms displacement. On the other hand there is general agreement for the averaged results, in Figs. 4 and 5. This latter agreement indicates that while the motion causing the rms displacement may not be harmonic (24, 25), a large component is.

In Fig. 7 the deviations in the three Voight symmetry properties are displayed for the relaxed configuration when all the interactions in the potential are considered. Although the magnitudes of the deviations increase when the unrelaxed configuration is considered the overall profile remains the same. These residual deviations indicate that while the biomolecule is relaxed at the continuum level (see Table II), it is not on the lumped residual level. Comparing the substructure of the protein with the deviations in the symmetry properties one sees a correlation between the peaks in the deviation and the transition regions separating the α-helical and looped regions. Additionally, it is found that the looped regions generally have larger deviations then the β- and α-helical regions. This latter result is due to the inherent stability of the α-helix and β-sheet which causes the regions containing those substructures to relax to an equilibrium configuration in

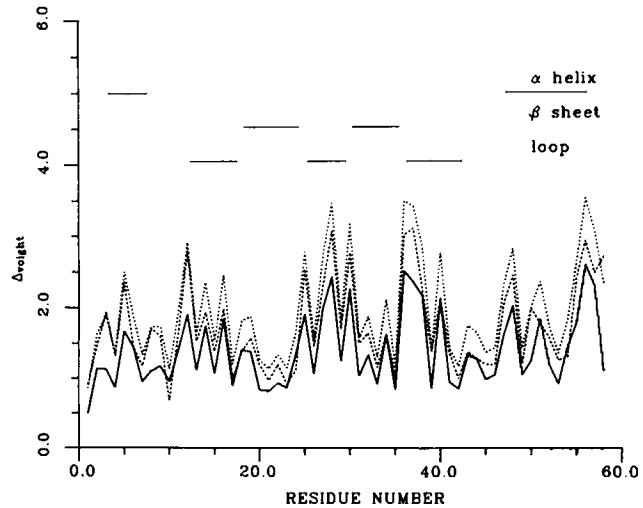

FIGURE 7 Deviations in the Voight symmetry properties of the full tensor averaged over the residues for the relaxed configuration assuming all the interactions. (——), (– – –), and (· · ·) denote $C_{abcd} - C_{bacd}$, $C_{abcd} - C_{abdc}$, and $C_{abcd} - C_{cdab}$, respectively.

fewer iterations. The looped and transition regions on the other hand are much more flexible and thus take more iterations.

As was mentioned in the Introduction the partial lumping scheme used here is only one of many possible ones. One could for example lump residues according to the underlying substructure. However, this scheme serves as a useful example of the lumping concept by showing that the calculated parameters refect the size of the lumped regions. That is, the parameters were more anisotropic than those for the continuum limit model but less than for the discrete continuum limit model.

IV. CONCLUSIONS

In this paper we have analyzed BPTI using a method which couples the atomistic and continuum descriptions of the protein. In this hybrid approach, a field of local elastic modulus tensors was calculated from a specified intramolecular potential and configuration. Whereas the primary purpose of calculating the field is to enable one to lump the molecule in a systematic and consistent fashion, the analysis of the field was also shown to be useful in understanding the elastic properties of the biomolecule.

The ability to obtain a potential and configuration which are self-consistent represents an important attribute of the hybrid approach. In most studies one begins with a crystal structure and by using some sort of minimization scheme finds a configuration that is relaxed in a chosen force field. Assuming that this relaxed configuration is representative of the true equilibrium structure of the molecule, one then calculates and compares theoretical with observed behavior. Through symmetrization the hybrid approach yields a balance between the chosen intermolecular potential and the experimentally determined structure. The potential and the structural information are put on the same footing. While the symmetrization resolves the inconsistency, it does so at the continuum level and not at the atomistic level. Thus, if a self-consistent set of equations is required at the atomistic level, the symmetrization method will not be useful.

Besides being useful in the static analysis by providing a means of quantitatively judging how close a given configuration is to the equilibrium configuration, the traces can be used to reduce the number of equations of motion for the molecule so that an accurate dynamical analysis is practical. Regions characterized by tensors having approximately the same traces are considered as lumped regions with a single tensor which is an average of the tensors within the regions. The amount of lumping is in part controlled by the nature of the information sought about the molecule. If, for example, information on the very low-frequency breathing motion of the biomolecule is desired then the whole molecule may be lumped. However, if some other property is desired which is less global in nature and thus characterized by shorter wavelengths, then less lumping is allowed. In this presentation two lumping schemes were considered. By averaging over the whole tensor field it was found that the protein was fairly isotropic in nature and could be characterized by a Young's modulus which was very reasonable with respect to previously reported values. This is in contrast to a previous phenomenological approach in which the modulus is arbitrarily chosen. The other lumping scheme in which the field was averaged over the residues permitted a more detailed analysis of the protein. Generally even though the resulting tensors were less isotropic than in the previous scheme, the calculated Young's moduli correlated well with the rigidity of the protein as determined from the experimental temperature factors. Furthermore, the deviations in the symmetry properties also exhibited a dependence on the substructure of the protein.

Although the present work is a static analysis, the primary motivation of the development of this approach is to make dynamic analysis of proteins and other biomolecules tractable. The dynamics of an active site in a large enzyme provides a good illustration of a typical application. Here the active site undergoes large amplitude and high-frequency motion, which is nonharmonic in nature, whereas the bulk of the enzyme away from the active site experiences small amplitude harmonic motion. Using the hybrid approach, the field of elastic tensors would be calculated throughout the bulk. This field would then be symmetrized and lumped as dictated by the traces. Finally, by coupling the lumped equations for the bulk to the classical equations of motion for the active site, the dynamics of the whole enzyme can now be described in a very consistent manner.

We wish to thank Ronald M. Levy and Paul H. Weiner for helpful discussions and for providing the energy minimized configuration of BPTI.

This work was supported by the Office of Naval Research.

Received for publication 18 November 1987 and in final form 13 May 1988.

REFERENCES

1. Karplus, M., and J. A. McCammon. 1981. The internal dynamics of globular proteins. *CRC Crit. Rev. Biochem* 9:293–357.
2. Welsh, R. G., B. Somogy, and S. Damjanovich. 1982. The role of protein fluctuations in enzyme action: a review. *Prog. Biophys. Mol. Biol.* 39:109–146.
3. Debrunner, P. G., and H. Frauenfelder. 1982. Dynamics of proteins. *Annu. Rev. Phys. Chem.* 33:283–299.
4. Ringe, D., and G. A. Petsko. 1985. Mapping protein dynamics by x-ray diffraction. *Prog. Biophys. Mol. Biol.* 45:197–235.
5. Cooper, A. 1984. Protein fluctuations and the theromodynamic uncertainty principle. *Prog. Biophys. Mol. Biol.* 44:181–214.
6. Careri, G., P. Fasella, and E. Gratton. 1979. Enzyme dynamics: the statistical physical approach. *Annu. Rev. Biophys. Bioeng.* 8:69–98.
7. Wilson, E. B., J. C. Decius, and P. G. Cross. 1955. Molecular Vibrations. Dover Publishing Inc., New York.
8. Birge, R. R., and L. M. Hubbard. 1980. Molecular dynamics of

cis-trans isomerization in rhodopsin. *J. Am. Chem. Soc.* 102:2195–2205.

9. McCammon, J. A., and M. Karplus. 1980. Simulation of protein dynamics. *Annu. Rev. Phys. Chem.* 31:29–45.
10. Levy, R. M., A. R. Srinivasan, W. K. Olson, and J. A. McCammon. 1984. Quasi-harmonic method for studying very low frequency modes in proteins. *Biopolymers.* 23:1099–1112.
11. Van Zandt, L. L., K. C. Lu, and E. W. Profhofsky. 1977. A new procedure for refining force constants in normal coordinate calculations on large molecules. *Biopolymers.* 16:2481–2490.
12. Born, M., and K. Huang. 1954. Dynamical Theory of Crystal Lattices. Ch. 5. Oxford University Press, London.
13. Maradudin, A. A., E. W. Montroll, G. H. Weiss, and I. P. Ipotava. 1971. The Theory of Lattice Dynamics in the Harmonic Approach. Academic Press, New York.
14. Thacher, T., S. Ganesan, A. Askar, and H. Rabitz. 1986. A hybrid approach to modelling the dynamics of macromolecules. *J. Chem Phys.* 85:3655–3673.
15. Brooks, B. R., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. 1983. CHARMM: a program for macromolecular energy, minimization and dynamics calculations. *J. Comp. Chem.* 4:187–217.
16. Deisenhofer, J., and W. Steigemann. 1975. Crystallographic refinement of the structure of bowine pancreatic trypsin inhibitor at 1-5A resolution. *Acta Crystallogr. Sect. B.* 31:238–250.
17. Weiner, P. K., and P. A. Kollman. 1981. AMBER: assisted model building with energy refinement. A general program for modelling molecules and their interactions. *J. Comp. Chem.* 2:287–303.
18. Landau, L. D., and E. M. Lifshitz. 1970. Theory of Elasticity. Pergamon Press Ltd., New York.
19. Morozov, V. N., and T. Y. A. Morozova. 1981. Viscoelastic properties of protein crystals: triclinic crystals of hen egg white lysozyme in different conditions. *Biopolymers.* 20:451–467.
20. Suezaki, Y., and N. Go. 1976. Fluctuations and mechanical strength of A-helices of polyglycine and poly L-alanine. *Biopolymers.* 15:2137–2159.
21. Suezaki, Y., and N. Go. 1975. Breathing mode of conformational fluctuations in globular proteins. *Int. J. Pept. Protein Res.* 7:333–334.
22. Go, N., T. Noguti, and T. Nishikawa. 1983. Dynamics of a small globular protein in terms of low-frequency vibrational modes. *Proc. Natl. Acad. Sci. USA.* 80:3696–3700.
23. Pickover, C. A. 1984. Spectrographic representation of globular protein breathing motions. *Science (Wash. DC).* 223:181–223.
24. Fauenfelder, H., G. A. Petsko, and D. Tsernoglov. 1979. Temperature-dependent x-ray diffraction as probe of protein structural dynamics. *Nature (Lond.).* 280:558–563.
25. Brooks, B., and M. Karplus. 1983. Harmonic dynamics of proteins: normal mode and fluctuation in BPTI. *Proc. Natl. Acad. Sci. USA.* 80:6571–6575.