Normal modes of symmetric protein assemblies

Application to the tobacco mosaic virus protein disk

Thomas Simonson* and David Perahia‡

*Howard Hughes Medical Institute and Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06511, USA; and Lure Université de Paris-Sud, Orsay, France; and ‡Laboratoire d'Enzymologie Physico-Chimique et Moléculaire, Université de Paris-Sud, Bât 430, 91405 Orsay, France

ABSTRACT We use group theoretical methods to study the molecular dynamics of symmetric protein multimers in the harmonic or quasiharmonic approximation. The method explicitly includes the long-range correlations between protein subunits. It can thus address collective dynamic effects, such as cooperativity between subunits. The n lowest-frequency normal modes of each individual subunit are combined into symmetry coordinates for the entire multimer. The Hessian of the potential energy is thereby reduced to a series of blocks of order n or $2n$. In the quasiharmonic approximation, the covariance matrix of the atomic oscillations is reduced to the same block structure by an analogous set of symmetry coordinates. The method is applied to one layer of the tobacco mosaic virus protein disk in vacuo, to gain insight into the role of conformational fluctuations and electrostatics in tobacco mosaic virus assembly. The system has 78,000 classical, positional, degrees of freedom, yet the calculation is reduced by symmetry to a problem of order 4,600. Normal modes in the 0–100 cm^{-1} range were calculated. The calculated correlations extend mainly from each subunit to its nearest neighbors. The network of core helices has weak correlations with the rest of the structure. Similarly, the inner loops 90–108 are uncorrelated with the rest of the structure. Thus, the model predicts that the dielectric response in the RNA-binding region is mainly due to the loops alone.

INTRODUCTION

Computer simulation techniques are powerful tools for the study of protein dynamics (1). Normal mode calculations are one such technique (2–5). The harmonic approximation provides a simple analytical description of the dynamics, from which thermodynamic properties can be directly calculated and quantum effects estimated. Although protein vibrations contain significant anharmonicity (6, 7), a harmonic analysis can give qualitative information about flexibility and low-frequency collective motions (7–12). Normal modes have also been used to perform Monte-Carlo simulations of one protein (13) and to aid in B-factor refinement in protein crystallography (14). Furthermore, the method can be refined by adopting the less-restrictive quasiharmonic approximation (15, 16). This approximation takes into account low-order anharmonicity and can incorporate solvent effects to an extent, while retaining the analytical power of the harmonic method. Mean-field treatments of the solvent have also been combined with normal mode calculations (17).

Some of the most interesting biological systems are large macromolecular assemblies, such as viruses and many allosteric proteins, whose low-frequency, collective, vibrations undoubtedly have functional importance. Due to their size, these systems are precisely the most difficult to study either by standard molecular dynamics or normal mode methods. Constraints on bond lengths and angles have been used in normal mode calculations to reduce the size of the problem (2, 4, 8), but it would be desirable to go beyond this approximation. Molecular dynamics simulations can be performed on a local region of interest, either omitting the surroundings or replacing them by an average reaction potential (18, 19). This approach either neglects long-range correlations or takes them into account in a simplified, average, way. Another possibility is to constrain a molecular dynamics simulation to respect the average symmetry of the multimer. At each time-step, the forces on one monomer are calculated, along with the corresponding displacement. A symmetric displacement is then imposed on the other monomers. The system is thus restricted to the space of completely symmetric fluctuations. This is strictly analogous to the periodic boundary method for liquids. However, the completely symmetric motions represent only a fraction of the degrees of freedom (1/17 in the case of the TMV protein disk for example). The others are certainly important in many situations. For example, conformational transformations that lower the symmetry begin as symmetry-breaking fluctuations. The dislocation of the TMV protein disk is one such transformation.

In this paper, we present a method to study the dynamics of large symmetric protein assemblies. We use well-known group theoretical ideas, along with a simple reduction of the number of degrees of freedom of the system. The method explicitly includes the long-range correlations between protein subunits. It is thus designed to address directly collective dynamic effects,

0006-3495/91/02/410/18 $2.00

such as cooperativity between subunits. The importance of long-range correlations and cooperative motions in many proteins is well-known. Striking examples are the quaternary rearrangements in allosteric proteins such as hemoglobin, or breathing motions in proteins such as hexokinase (20, 21).

The method is illustrated below by a study of the harmonic fluctuations and the dielectric properties of the disk of protein of tobacco mosaic virus (TMV). It was our interest in this system that initially led us to develop the method. TMV assembly (22–26) is initiated in physiological conditions by the binding of a specific RNA sequence to an aggregate of coat protein, known to contain about 34 protein molecules (27). According to the mechanism postulated by Butler and Klug (27), this aggregate is the 34-subunit, two-layered, 17-fold symmetric, disk of coat protein (Fig. 1); a specific RNA hairpin is inserted into the central hole of the disk, binding between its two layers; the disk then dislocates, to form a two-turn helical fragment. Growth of the viral helix follows. In the absence of RNA, dislocation of the disk can be triggered by proton binding. At pH < 6.5, a proton binds on the inner face of the disk, whereon the disk dislocates and equilibrium is shifted to a long, virus-like, helical aggregate (28, 29). The transition from the disk to these helical forms is accompanied by ordering of the flexible loop 90–108 on the inner face of the disk, which binds the RNA in the virus, and binds the critical proton in the viruslike helix. Thus, the conformational flexibility, and the electrostatic properties, of the protein disk play a role in the equilibrium between the different aggregates of TMV protein. The initial microscopic steps of viral assembly and their energetics are not known in detail. Several hypotheses have been proposed (e.g., [26, 27, 30, 31]). One can speculate for example, that when RNA binds to the protein disk, it plays an active role in lowering the energy barrier for dislocation, thereby accelerating assembly. This would imply that the RNA significantly perturbs the fluctuations of the disk, between the moment when it is bound and the moment of dislocation. A prime candidate to mediate this perturbation is the long-range Coulomb interaction.

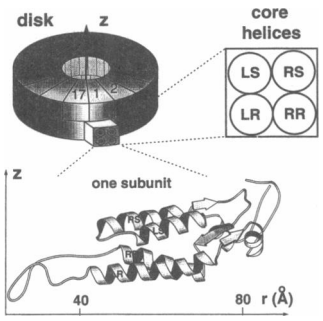

FIGURE 1 Schematic view of the two-layered disk and of a single protein subunit. The disk is schematized (*upper left*) with one subunit artificially protruding. The enlarged outer face of this subunit (*upper right*) shows the arrangement of the four-helix bundle. A ribbon trace of the subunit is shown below, produced with the program MolScript (56).

Molecular dynamics simulations of the TMV protein disk can provide information about the intrinsic flexibility and fluctuations of the disk, as well as about its electrostatic and dielectric properties. These properties determine the disk's response to a set of perturbing charges. The initial binding of RNA is a collective process, with eighteen bases binding to twelve of the protein subunits (32). Dislocation of the disk is also a collective process. Because of this, it is of interest to investigate the fluctuations of the disk as a whole. This can only be done within the framework of a harmonic or quasiharmonic approximation, which exploits fully the 17-fold symmetry of the structure. This approach is not meant to describe conformational changes, but rather the intrinsic fluctuations and dielectric properties of the disk itself. Due to computer limitations, only one layer of the two-layered disk could be simulated by this method. We expect that much of the dynamics of the individual layers of the disk will still be qualitatively reproduced by this partial model. There is very little contact between the two layers in the RNA binding region, most of the interactions being located near the disk's outer surface. The lateral contacts between subunits are much more numerous. These contacts should not be too strongly affected by our approximation. Other simulation approaches are currently being used to study portions of the two layers simultaneously, and to test this approximation.

The structure of the disk and virus are discussed briefly in the next section. The theory of harmonic and quasiharmonic molecular dynamics of symmetric protein multimers is then presented. In the last section results on the dynamic and dielectric properties of the TMV disk are described.

STRUCTURE OF TMV AND THE TMV PROTEIN DISK

TMV protein exists in a number of states of aggregation, depending on the solution conditions (24). In physiological conditions, the disk is found, possibly coexisting with short lockwasher helices. Below pH 6.5, long protein helices form, very similar to the intact virus (28, 29). Although titration experiments indirectly suggest that the lockwasher form may be the nucleating aggregate

(31, 33), the disk has been directly seen by electron microscopy in the conditions of assembly (57), and it has been seen passing through the lockwasher form during formation of helices (34).

The disk contains 34 protein molecules arranged in two layers, with 17-fold symmetry (35, 36). It is 40 Å thick, with an 80 Å radius and a central hole of 30 Å radius. The structure is schematized in Fig. 1. The central core of each molecule is an antiparallel bundle of four α-helices. The left radial (LR) helix contains residues 114–134. The right radial (RR) helix contains residues 74–87. The left slewed (LS) and right slewed (RS) helices contain residues 20–31 and 38–48. These helices run approximately radially, in the plane of the disk. The helices of neighboring subunits pack closely. Thus a close-packed, periodic, anti-parallel, network of 136 helices is formed. This network is sensitive to electrostatic perturbations. Changing the ionic strength in crystals of disks causes relative motions of the core helices; binding low concentrations of hexanucleotides has the same effect (37, 38). Higher concentrations cause the crystals to crack, presumably because the disk-helix transition is activated.

A second key region of the disk is also sensitive to electrostatic perturbations. In the disk's central hole, the loop 90–108 is directly involved in binding the viral RNA. It contains five charged residues, and is partially disordered in the disk (39, 35, 36) due to electrostatic repulsion. When the pH is lowered below 6.5, a proton binds in this region (40, 31), the loop becomes ordered and long protein helices form. The proton binding thus functions as an electrostatic switch for helix formation. The titrating site appears to be the pair of glutamates 95 and 106 from lateral neighbor subunits, binding a proton between them to form a Caspar pair. Alternatively, binding of RNA can also provide the energy needed to order the loop: the loop is ordered in the virus, and closed involved in RNA binding.

The virus structure (30, 31, 41) is very similar to the disk: within a subunit, the root mean square shift of the Cα positions between the virus and the disk is only 0.7 Å. RNA binds at a radius of ~40 Å, completely surrounded by protein. Three bases bind to each protein subunit, lying flat against the LR helix above the RNA. The three phosphate groups interact electrostatistically with the helices RR, LR and RS below. Phosphates 1 and 2 hydrogen bond to Arg90 and 92 respectively. Phosphate 2 binds a calcium ion together with Asp116.

It seems likely that conformational fluctuations of the protein disk, and its sensitivity to electrostatic perturbations, play a part in TMV assembly. The abundant, closely-packed, secondary structure of the disk suggests that it is somewhat rigid, and therefore a harmonic or quasiharmonic model of the dynamics is not unreasonable. This approach will permit us to treat the disk as a whole, and analyze cooperativity between subunits rigorously. We will focus on the flexibility of the disk, the interparticle correlations, and the interactions of the disk with perturbing charges. If, for example, the RNA charges induce a mechanical deformation of the disk, then modes of the disk that have a component along this deformation should exhibit a strong coupling to the RNA charges. Such modes can be likened to the soft modes of structural phase transitions (42). Simulations of B-DNA oligomers exhibit such a coupling between perturbing charges and conformational fluctuations: a drastic softening of one low-frequency mode is seen as the ionic strength increases to the critical value for the B–Z transition (17).

THEORY: HARMONIC AND QUASIHARMONIC DYNAMICS OF A SYMMETRIC PROTEIN MULTIMER

The group theory presented here is well-known. It is recalled for completeness, and phrased in terms specific to protein assemblies.

Consider a symmetric assembly of asymmetric monomers. Let G be the symmetry group of the multimer. Let n_G be the order of G, which is also equal to the number of monomers. Let N be the number of degrees of freedom of one monomer. By symmetry, the normal mode calculation of the multimer can be reduced from a problem of order $n_G N$ to a series of problems of order at most $2N$, using elementary group theory (43, 44). The first step is to construct a convenient basis of symmetry coordinates for the multimer. The second is to write the Hessian matrix of the potential energy in this basis, thereby reducing it to a series of diagonal blocks of order at most $2N$.

Let ϵ be the configuration space of the multimer. Let T be the mechanical representation of G in ϵ (44), and $\{T_p; p = 1, p_G\}$ the irreducible representations of G. For $g \in G$, let $\tau^p(g)$ be the matrix representing g in T_p. Let s_p be the dimension of T_p, and m_p the number of times T_p appears in T. Let $\{\epsilon_{pj}; j = 1, m_p\}$ be the irreducible subspaces of ϵ that transform according to T_p, and let ϵ_p be their sum:

$$\epsilon = \sum_{p=1}^{p_G} \epsilon_p = \sum_{p=1}^{p_G} \sum_{j=1}^{m_p} \epsilon_{pj}. \quad (1)$$

An arbitrary displacement of the multimer can be projected onto the symmetry subspace ϵ_p by the s_p Vosko projectors (44),

$$P_{ps} = \sum_{g \in G} \tau^p_{ss}(g)^* T(g); \quad s = 1, s_p. \quad (2)$$

The asterisk denotes complex conjugation. Let $\{\mathbf{u}_k; k = 1, N\}$ be an orthonormal basis of the configuration space of monomer 1. To each $\mathbf{u}_k$ there corresponds a displacement $\mathbf{U}_k = (\mathbf{u}_k, \mathbf{0}, \mathbf{0}, \ldots, \mathbf{0})$ of the multimer, where monomer 1 moves along $\mathbf{u}_k$ and the others are immobile. We apply the Vosko projectors to the $\mathbf{U}_k$, projecting them onto the symmetry subspaces ϵ_p, to obtain the vectors:

$$\mathbf{q}_k^{ps} = P_{ps}\mathbf{U}_k. \tag{3}$$

The $\mathbf{q}_k^{ps}$ are symmetry coordinates, and form an orthonormal basis of ϵ. Denoting g_i the symmetry element that transforms monomer 1 into monomer i, the symmetry coordinate $\mathbf{q}_k^{ps}$ corresponds to a displacement of monomer i of the form:

$$\mathbf{r}_i = \tau_{ss}^p(g_i)^* g_i \mathbf{u}_k. \tag{4}$$

Expression of the Hessian in symmetry coordinates

Let $\mathbf{V}$ be the mass-weighted Hessian matrix of the potential energy function of the multimer, expressed in the basis of cartesian displacements. Let $\boldsymbol{\nu}$ be the Hessian matrix in the basis of symmetry coordinates $\mathbf{q}_k^{ps}$. The general matrix element $\nu_{k,k}^{ps,\pi\sigma}$ is given by:

$$\nu_{k,\kappa}^{ps,\pi\sigma} = {}^t\mathbf{q}_k^{ps}\mathbf{V}\mathbf{q}_\kappa^{\pi\sigma *}. \tag{5}$$

The raised t denotes transposition.

Let $\mathbf{V}^{(ij)}$ be the block in $\mathbf{V}$ corresponding to the interaction of monomers i and j. Suppose $T(g)$ transforms monomers I and J into monomers i and j respectively. By symmetry

$$\mathbf{V}^{(IJ)} = {}^t g \mathbf{V}^{(ij)} g. \tag{6}$$

Inserting 4 into 5, using 6, and using the orthogonality properties of the matrices τ^p (44), we obtain the desired expression:

$$\nu_{k,\kappa}^{ps,\pi\sigma} = \frac{n_G}{s_p}\,{}^t\mathbf{u}_k \left\{ \sum_{j=1}^{n_G} \tau_{ss}^p(g_j)\mathbf{V}^{(1j)} g_j \right\} \mathbf{u}_\kappa \delta_{s\sigma}\delta_{p\pi}. \tag{7}$$

In the basis of the symmetry coordinates, the Hessian matrix is thus reduced to a series of blocks of order N. If the irreducible representation p is not one-dimensional, each line in the matrix τ^p contributes one block. The s_p blocks corresponding to τ^p are all identical, so that there is an s_p-fold degeneracy of the spectrum for each irreducible representation T_p.

If the irreducible representation T_p is complex, then the complex conjugate subspaces ϵ_p and ϵ_p^* must be combined to give a physically meaningful symmetry subspace of double the dimension. The basis vectors $\mathbf{q}_k^{ps}$ and $\mathbf{q}_k^{ps*}$ are replaced by their real and imaginary parts, and the complex blocks $\boldsymbol{\nu}^{ps,ps}$ and $\boldsymbol{\nu}^{ps,ps*}$ combine to form a real block of order $2N$.

As expected from the multimer's symmetry, the Hessian matrix is completely determined in Eq. 7 by the blocks $\mathbf{V}^{(1j)}$ that couple monomer 1 to its neighbors. If there is no coupling between subunits $\boldsymbol{\nu}^{ps,ps}$ does not depend on p, and the spectrum is n_G-fold degenerate.

The methodology presented above corresponds to the harmonic approximation. It can also be applied to the quasiharmonic approximation (15). In the quasiharmonic approximation the potential energy is assumed to be of the second order with respect to the atomic displacements, but with coefficients that depend on the temperature. This temperature dependence effectively takes into account the anharmonicity of the potential function. The covariance matrix $\boldsymbol{\sigma}$ of the atomic displacements is estimated from a molecular dynamics simulation. Then an effective, quasiharmonic, Hessian matrix is derived from the equation:

$$\mathbf{V} = kT\boldsymbol{\sigma}^{-1}, \tag{8}$$

where k is Boltzmann's constant and T is the temperature. In symmetry coordinates, the Hessian $\boldsymbol{\nu}$ is made up of diagonal blocks $\boldsymbol{\nu}^{pp}$ of order N. Separating out the six external degrees of freedom, the remaining blocks are positive definite at the energy minimum, therefore invertible. The inverse $\boldsymbol{\nu}^{-1}$ has the same block-diagonal structure as $\boldsymbol{\nu}$. Therefore the transformation 7 that reduces $\mathbf{V}$ to block-diagonal form reduces $\mathbf{V}^{-1}$, and therefore $\boldsymbol{\sigma}$, to the same block-diagonal form. Thus the method applies without modification to the quasiharmonic approximation.

Elimination of the high-frequency modes of vibration

The magnitude of the atomic displacements, as well as the static dielectric properties of the multimer are mainly determined by its low-frequency modes of vibration. We therefore seek to eliminate the high-frequency modes at an early stage of the calculation. The natural way to do this is to elminate the high-frequency modes of the individual monomers. At this point, the basis $\{\mathbf{u}_k; k = 1, N\}$ is arbitrary. Let us now take the $\mathbf{u}_k$ to be the N eigenvectors of the block $\mathbf{V}^{(11)}$ of the Hessian $\mathbf{V}$. This means that $\mathbf{u}_k$ is a normal mode of monomer 1, with its neighboring monomers artificially held rigid. Let us discard the high-frequency modes from this set, retaining the n lowest-frequency modes $\{\mathbf{u}_k; k = 1, n)\}$ Constructing the corresponding symmetry coordinates, we reduce $\mathbf{V}$ to a series of diagonal blocks of order n or $2n$. For large systems such as proteins, only the lowest fraction of the spectrum need be retained, so that $n \ll N$. Then the calculation is essentially reduced to the

initial determination of the $\mathbf{u}_k$, a single problem of order N. The model can be improved gradually by including more and more modes in the basis set.

Other choices can be made for the reduced basis set $\{\mathbf{u}_k\}$, consistent with physical intuition. The use of dihedral angles as degrees of freedom is an example (2, 4, 8). Our choice of low-frequency normal modes has the advantage that angle flexibility is included. Bond angles are known to be important in determining the thermodynamic properties of macromolecules (e.g., 15).

Implementation of the method

The sequence of operations to obtain the normal modes of the multimer is the following: (*a*) Calculate the Hessian matrix of monomer 1 (or the covariance matrix in the quasiharmonic approach) in the presence of its rigid neighbors, $\mathbf{V}^{(11)}$; (*b*) diagonalize this matrix, retaining only the n lowest-frequency modes, $\{\mathbf{u}_k, k = 1, n\}$; (*c*) calculate the other blocks of the Hessian (or the covariance matrix in the quasiharmonic approach), coupling monomer 1 to its neighbors, $\mathbf{V}^{(1,j)}$; (*d*) contract and symmetrize these blocks by applying Eq. 7, to obtain small blocks of order $\leq 2n$, $(\nu^{ps,ps}_{k,\kappa})$; (*e*) diagonalize these small blocks to obtain the normal modes of the multimer, expressed in symmetry coordinates; and (*f*) convert these modes back to cartesian coordinates.

Each block $\mathbf{V}^{(1,j)}$ contributes separately to $\boldsymbol{\nu}^{(ps,ps)}$ (Eq. 7), so that each one can be calculated and processed separately, without storing more than one in computer memory at any time. Because $\boldsymbol{\nu}^{(ps,ps)}$ is typically much smaller than $\mathbf{V}^{(1,j)}$, the memory requirements are essentially the same as for a normal mode calculation of a single monomer. C.p.u. requirements are very small (<5 min for a TMV protein monomer on a Fujitsu VP200 supercomputer using highly vectorized subroutines).

Application to the disk of TMV protein

In the case of the disk of TMV protein, the symmetry group is C_{17}. The complex, one-dimensional, irreducible representations are listed in Table 1. The real, physically irreducible, representations are:

$$T_1, T_2 + T_{17}, T_3 + T_{16}, T_4 + T_{15}, \cdots, T_9 + T_{10}.$$

The real symmetry subspaces will be denoted $\epsilon_1, \epsilon_2, \cdots, \epsilon_9$. ϵ_1 constains $N - 2$ degrees of freedom. ϵ_2 contains $2N$–4 degrees of freedom. The seven other subspaces contain $2N$ degrees of freedom each.

The complex symmetry coordinate $\mathbf{q}^p_k$ (Eqs. 3 and 4 with $s = 1$) has the form:

$$\mathbf{q}^p_k = \begin{pmatrix} \mathbf{u}_k \\ \omega^{(p-1)*}\mathbf{R}\mathbf{u}_k \\ \omega^{2(p-1)*}\mathbf{R}^2\mathbf{u}_k \\ \downarrow \\ \omega^{16(p-1)*}\mathbf{R}^{16}\mathbf{u}_k \end{pmatrix}. \quad (9)$$

R represents the rotation of $2\pi/17$ around the axis of the disk, and $\omega = \exp(2i\pi/17)$. The real symmetry coordinates have the form:

$$\mathbf{Q}^p_k = \frac{1}{\sqrt{2}}(\mathbf{q}^p_k + \mathbf{q}^{p*}_k).$$

$$\mathbf{R}^p_k = \frac{1}{i\sqrt{2}}(\mathbf{q}^p_k + \mathbf{q}^{p*}_k). \quad (10)$$

Consider a normal mode in subspace ϵ_2 for example, of frequency ν. It produces a displacement of subunit n of the form

$$\cos(2\pi\nu t)R^{n-1}\sum_k (A_k \cos(n-1)\alpha + B_k \sin(n-1)\alpha)\mathbf{u}_k,$$

where $\alpha = 2\pi/17$ and A_k and B_k are constants. A second, orthogonal, mode is degenerate with this one. Rearranging the cosine and sine terms of these modes gives a sum of two waves, with a time dependence of the form $\cos[2\pi\nu t - (n-1)\alpha]$ and $\cos[2\pi\nu t + (n-1)\alpha]$ respectively. Each normal mode, in which all atoms of the disk vibrate in phase, can be viewed as a stationary wave, formed by superimposing two waves propagating around the disk in opposite directions, with wave-numbers $\pm 2\pi/17$. In general, each irreducible representation T_p corresponds to motions that propagate around the disk with a wave number $\pm 2(p-1)\pi/17$. Thus ϵ_1 contains completely symmetric displacements, where the subunits are in phase. ϵ_2 contains displacements having a phase shift of $2\pi/17$ between neighboring subunits. ϵ_3 contains displacements having a phase shift of $4\pi/17$ between neighboring subunits, and so on.

TABLE 1 **Irreducible representations of C_{17}**

g	I	R	R^2	R^3	→	R^{16}
T_1	1	1	1	1	→	1
T_2	1	ω	ω^2	ω^3	→	ω^{16}
T_3	1	ω^2	ω^4	ω^6	→	ω^{32}
T_4	1	ω^3	ω^6	ω^9	→	ω^{48}
↓	↓	↓	↓	↓	↘	↓
T_{17}	1	ω^{16}	ω^{32}	ω^{48}	→	$\omega^{16\times 16}$

R represents the rotation of $2\pi/17$ around the axis of the disk, and $\omega = exp(2i\pi/17)$.

Simulation protocol

The CHARMM/PARAM19 force field (45) was used. This force field includes all the heavy atoms of the protein and all the polar hydrogens explicitly; CH, CH_2 and CH_3 groups are treated as extended atoms. Hydro-

gen bonds are treated as an electrostatic interaction between point charges. Charge-dipole induction forces are neglected. The charges of solvent accessible side-chains of charged residues were scaled by a factor of 0.3, to mimic solvent screening in a simple way. The corresponding charge-charge interactions are thus scaled by a factor of ≈ 0.1. Nonbonded interactions were gradually switched to zero over interparticle separations of 8–9 Å. This short cutoff distance was chosen to ensure that each monomer of the disk only interacted with its nearest neighbors. Thus the only nonzero blocks of the form $\mathbf{V}^{(1,j)}$ in the Hessian were the two blocks $\mathbf{V}^{(1,2)}$ and $\mathbf{V}^{(1,17)}$ (which are related by symmetry). It would be desirable to compare this rather arbitrary scheme to other truncation schemes, for example switching over the 5–9 Å region combined with a $1/r$ dielectric screening function, following Longarich and Brooks (46).

To reduce the computational burden, only the upper layer of the disk was included in the model. This results in a total of 1,529 atoms and 4,587 degrees of freedom per subunit. The starting model was the X-ray structure of the disk (36). The segment Ile94-Glu106 is disordered and missing from the x-ray model. It was built in on a graphics system in a helix-like conformation, and annealed by molecular dynamics and minimization. Then the entire structure was minimized, with the constraint imposed by the 17-fold symmetry.

Normal mode calculations were performed; complete quasiharmonic calculations are underway. We extended the molecular modelling program CHARMM (45) to perform normal mode calculations for the cyclic point groups C_n. Extending the code to treat other point groups would be fairly straightforward. Calculations were performed on a Fujitsu VP200 supercomputer. The bisection method, followed by inverse iteration, was used to diagonalize the Hessian matrix. The set of $\mathbf{u}_k$ vectors we used included the 780 lowest-frequency normal modes of monomer 1. The corresponding cutoff frequency is 148 cm^{-1}. This resulted ultimately in a useful frequency spectrum of 0–100 cm^{-1} for the whole disk (see Results), with 543 modes per monomer. This should be sufficient to describe the amplitude of atomic displacements reasonably well. The dielectric properties of several proteins in the harmonic approximation are also accurately described by this frequency range (53). Quantum effects will be negligible in this range (3). Because of computer memory limitations, further optimization of the program is necessary before the cufoff frequency can be increased.

A short, 50 ps, molecular dynamics simulation of three subunits in vacuo was also performed, to estimate the stability of the loop conformation and for general purposes of comparison. The same cutoff scheme was used, and equilibration lasted 40 ps.

RESULTS AND DISCUSSION

Minimized structure

Our starting model was the 2.8 Å resolution x-ray structure of the lower layer of the protein disk. The two layers are very similar, superimposable with an r.m.s. shift of only 0.7 Å for the α carbons, or 0.8 Å for all heavy atoms (excluding the loop 90–108). The best superposition corresponds roughly to an 11° rotation of the molecule around an axis in the plane of the disk, perpendicular to its radius.

Conjugate gradient energy minimization was first done with both layers of the disk present. This led to an r.m.s. shift of 1 Å away from the x-ray coordinates, including all heavy atoms except the loop 90–108. Then the lower layer of the disk was minimized alone. The final r.m.s. shift was 1.1 Å away from the x-ray structure. Backbone atoms shift by only 0.85 Å, side chains by 1.2 Å, hydrophobic side chains by only 0.8 Å. The r.m.s. shifts upon minimization correlate nicely with the experimental mobility of each class of atoms, as measured by the crystallographic B-factor. Fig. 2 *a* shows the backbone of the starting and final structures. The changes are smoothly distributed throughout the structure, except for the dangling C-terminus, which has a fairly large shift. The four core helices have shifted slightly and more or less as a body. The helix packing must be a strong structural restraint, since the helix interfaces are strongly conserved in different strains of TMV (40). The hydrogen bonds in the core helices shorten very slightly on minimization, and one H-bond is broken, Thr118-Arg122. The arginine side chain (NH1-HH12) displaces its backbone amide group as the H-bond donor. An inter-subunit salt bridge between this residue and Asp88 on the *R* side is maintained despite this shift.

The main structural differences are at the inner ends of the *RR* and *LR* helices, in the region of the central hole. The segment 94–105 is not included in the x-ray model. In the virus, it forms a series of reverse turns (31). 360 MHz H^1-NMR measurements (39) show that in solution, this segment is disordered on a nanosecond time-scale, but it may exchange between conformations that are ordered on a shorter time-scale. In our model, the loop started out in a mostly-helical conformation, and was annealed by 10 ps of molecular dynamics at room temperature, with the rest of the structure fixed; then energy minimization was begun. The final structure is an extended loop. Because there are five charged residues in the segment 90–108, the secondary structure of the loop is probably determined by electrostatic interactions between side chains and with the surrounding water and counterions. The resolution of the x-ray structure is too low to indicate ordered water molecules,

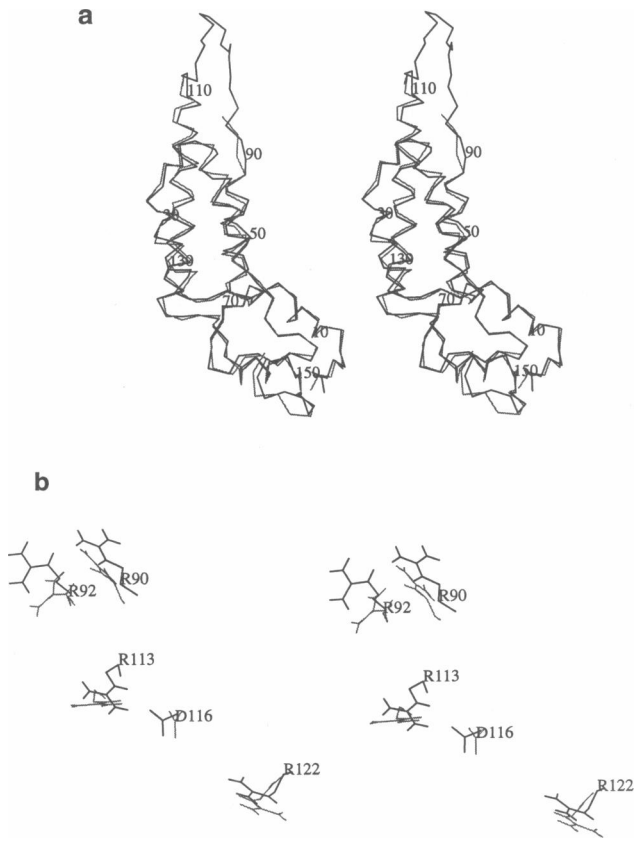

FIGURE 2 (*a*) Stereo view of the Cα trace of one subunit in the x-ray (*dashed*) and minimized (*solid*) structures. The view is from above, perpendicular to the plane of the disk (compare to Fig. 1.) (*b*) Stereo view of one subunit, showing the region at the inner ends of helices LS and RS, including several charged side chains. x-ray structure: *dashed;* minimized structure: *solid.*

either at the disk's surface or at the subunit interfaces, and therefore our model did not include any. This obviously limits our ability to predict the structure of the loop, and the configuration obtained is simply one plausible possibility among several.

Each subunit contains eight aspartates, seven glutamates, two lysines, and 11 arginines, all ionized at physiological ionic strength and pH > 7. This makes 28 charged residues out of 158. Half of these are in the inner loop region. Fig. 2 *b* shows the effect of minimization on the side chains at the inner end of the core helices, at the base of the disordered loop. Some of the charged side chains have moved by riding on the shifting backbone (e.g. Arg90 and Arg113). Others have reoriented slightly (e.g. Asp88, Arg92 and Asp115). The hydrophobic side chains have very small shifts. The pair Asp115-Asp116 may be the carboxyl pair that titrates anomalously at pH 7; Glu95-106 has also been suggested (31). Asp88-Arg122 and Arg113-Asp115 are highly conserved intersubunit salt bridges, whereas Arg90 and 92 are highly conserved due to their binding of the RNA phosphates. The small shifts upon minimization may reflect these evolutionary constraints.

Normal mode spectrum

Due to the number of degrees of freedom of the disk, it was not possible to minimize the structure completely, so that the calculated spectrum contained seven negative eigenvalues per subunit. These eigenvalues were eliminated from the analysis. The corresponding eigenvectors are local in nature, and represent small structural rearrangements required to reach a true minimum. We assume that the overall motions will not be sensitive to these rearrangements.

The reduced basis set $\{\mathbf{u}_k\}$ is made up of the normal modes of a single subunit, with its neighbors fixed, with frequencies below 148 cm^{-1}. This gives a total of 780 basis vectors, and a spectrum for the disk that extends up to very nearly 148 cm^{-1}. Because we neglect the degrees of freedom of each individual subunit other than those in the set $\{\mathbf{u}_k\}$, the spectrum obtained for the disk is approximate. The accuracy decreases as the frequency approaches the cutoff 148 cm^{-1}. Calculations with a cutoff of 100 cm^{-1} were done for comparison. The two calculations give similar spectra, but somewhat different displacements, even in the low-frequency range (see below). This approach, of breaking down the molecule into parts and retaining only some of the degrees of freedom of each part in the overall Hessian, is not restricted to symmetric molecules. It has been systematically tested on several non-symmetric systems, including BPTI, a scorpion neurotoxin, and lysozyme (Mouawad, L., and D. Perahia, manuscript in preparation). The success of the methods varies. However when a natural subdivision into compact parts exists, the results are accurate up to over ½ of the cutoff frequency of the reduced basis set. Fig. 3 shows the spectrum of the disk, giving the number of modes in 10 cm^{-1} windows of frequency. The shape of the histogram up to 100 cm^{-1} is similar to previous studies (8). There is a smooth increase, then decrease, in the number of modes per window up to 100 cm^{-1}. At this point, there is a sharp, narrow gap in the spectrum, followed by two windows with substantially more modes. This gap may be due to missing degrees of freedom in the reduced basis set, which are needed, in combination with the $\{\mathbf{u}_k\}$, to form modes in the 100–110 cm^{-1} range. This suggests that the main inaccuracies in our spectrum appear above 100 cm^{-1}. In what follows we have retained the modes from 0–100 cm^{-1} to estimate the atomic fluctuations and the dielectric properties of the disk. These quantities are

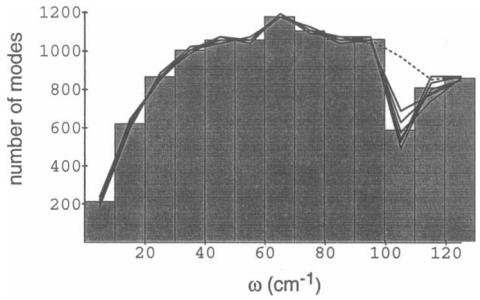

FIGURE 3 Frequency histogram giving the number of normal modes of the disk per 10 cm^{-1} frequency interval. The histograms of the individual symmetry subspaces are superimposed (*dashed line:* space 1; *solid lines:* spaces 2–9). They are normalized to the total number of modes per subunit, for comparison.

largely determined by the lowest-frequency modes. If there are errors in the upper part of the spectrum, it will affect them only slightly.

Our calculations give 543 normal modes per subunit for the disk in the 0–100 cm^{-1} range. The average frequency is 56.8 cm^{-1}. There are 543/890 = 0.61 modes per dihedral angle. Go (47) estimated that in an elastic model of pancreatic trypsin inhibitor, there are roughly as many elastic modes as dihedral degrees of freedom. An elastic mode here means a mode whose wavelength is greater than the size of an amino acid. The corresponding frequency range is 20–200 cm^{-1}. Levitt et al. found frequencies up to 200–250 cm^{-1} in four proteins, using only dihedral angles as degrees of freedom (8). It appears therefore that an accurate description of the low-frequency spectrum of a protein should include at least one mode per dihedral degree of freedom, and cover the range 0–200 or 250 cm^{-1}. Our more limited basis set and resulting spectrum should thus be regarded as a first approximation to the collective motions of the disk. Levitt et al. observed characteristic frequencies for the internal motions of different kinds of secondary structure. Internal motions of β sheets were produced by modes in the range 15–40 cm^{-1}, internal motions of turns by modes in the range 4–75 cm^{-1}, and internal motions of helices by modes in the range 55–135 cm^{-1}. Side chain torsions were produced by modes in the range above 40 cm^{-1}. Normal mode calculations on a single α-helix show however a low frequency region ranging from 20 to 40 cm^{-1}, corresponding to internal collective motions (7). The shift to higher frequencies is thus due to the surrounding protein. The extent of the shift presumably depends on the environment of any particular helix. Our present model should give a good description of the harmonic motion of sheets and turns, which are soft, but will limit side chain motions somewhat, and may produce overly rigid helices. These effects will occur especially at the subunit interfaces. Further optimization of the program is necessary to allow a greater cutoff frequency.

Quantum effects are virtually absent in the 0–100 cm^{-1} frequency range at room temperature. The classical and quantum vibrational entropies are given by (48):

$$S_{class} = k_B \sum_i (1 - ln u_i)$$

$$S_{quant} = k_B \sum_i \frac{u_i}{e^{u_i} - 1} - ln(1 - e^{-u_i}), \qquad (11)$$

where $u_i = \hbar\omega_i/k_BT$, k_B is Boltzmann's constant, and the sum is over all frequencies ω_i. Here the contribution of the 0–100 cm^{-1} frequency range, at room temperature T, is TS_{class} = 781 kcal/mol per subunit; the quantum correction is +1.2 kcal/mol per subunit (estimated by inserting the classical frequencies into the quantum expression.) Quantum entropic effects are small up to about 300 cm^{-1}. Quantum corrections to the rms displacements are also small up to 300 cm^{-1} (3).

The entropy obtained represents 1.44 kcal/mol for each degree of freedom included in the model. Levitt et al. found TS_{class} = 1.51 ± 0.03 kcal/mol per degree of freedom, in their analysis of four proteins (8), using only dihedral degrees of freedom. This agreement suggests that the degrees of freedom in our model are essentially linear combinations of dihedral degrees of freedom.

The spectra of the different symmetry subspaces are similar throughout the frequency range examined (Fig. 3). The only large differences appear >100 cm^{-1}. We calculated the contributions to the vibrational entropy from each subspace, over the 0–100 cm^{-1} range. The spread of values of TS_{class} is 0.03 kcal/mol per degree of freedom. This is just 2% of the total value, however it does represent 10% of the mean thermal energy at room temperature. If the blocks $\mathbf{V}^{i,i+1}$ in the Hessian matrix were zero, the spectrum would be 17-fold degenerate, with each mode appearing once in ϵ_1 and twice in each of the other subspaces. We have therefore grouped the modes in sets of seventeen, taking the first mode from ϵ_1 with the first two from each of the nine other subspaces, and so on through the whole spectrum. The splitting of the frequencies within each set of seventeen was then calculated, that is the difference between the maximum and minimum in each set. The average splitting is 0.56 cm^{-1}. The average relative splitting (splitting divided by the mean frequency of the set) is 2%. However at low frequencies the splitting is relatively important. In the range 0–20 cm^{-1}, the relative splittings are almost all greater than 5% (Fig. 4). This frequency range accounts for most of the atomic displacements (see below). The

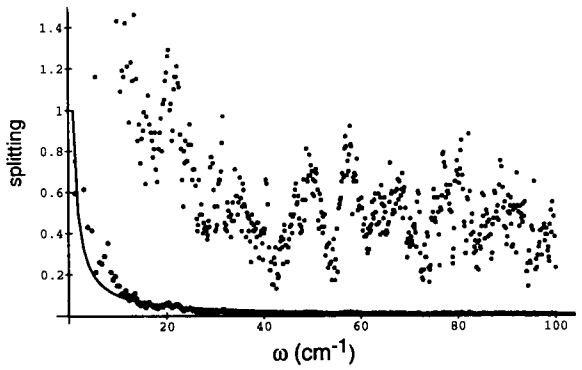

FIGURE 4 The frequency splitting within each successive group of 17 modes (see text). Absolute splitting (*black points*) and relative splitting (splitting divided by the mean frequency of the group; *gray points*) vs the mean frequency ω of the group. The solid line represents $1/\omega$.

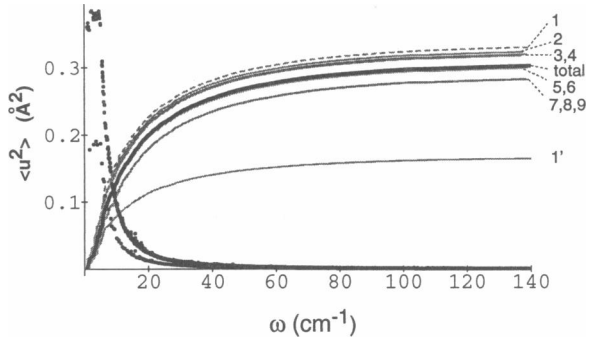

FIGURE 5 Mean square atomic displacements averaged over all atoms as a function of frequency; spectral density: *black dots;* integrated value: *solid curve.* The spectral density is multiplied by 50 for comparison. The contributions of the 9 symmetry subspaces, normalized to the total number of modes per subunit, are shown as grey lines. The labels on the right are the subspace numbers. The contribution of subspace 1 is dashed. Subspace 1 has ½ as many degrees of freedom as the others. The curve (1′) is equal to 0.5* (1), and thus gives the contribution of subspace 1 on the same scale as the other subspaces.

differences between symmetry subspaces suggest that the corresponding normal modes involve more than one subunit.

The sensitivity of the spectrum to the cutoff frequency of the basis set $\{\mathbf{u}_k\}$ was tested by performing calculations with a 100 cm^{-1} cutoff. With the smaller basis set, the frequencies of the disk are shifted upwards by 0.2 cm^{-1} on average. This shift is fairly uniform over the spectrum, affecting even the lowest modes.

Atomic fluctuations

The calculated r.m.s. displacement, averaged over all main-chain atoms, is 0.49 Å. The mean square displacement along each mode falls off rapidly with frequency, roughly as $1/\omega$, and the overall mean square displacement levels off above 60 cm^{-1} (Fig. 5). Modes <20 cm^{-1} account for almost all of the average displacement.

The r.m.s. displacements averaged over different classes of atoms are given in Table 2. As expected, the calculated mobility increases going from the backbone out to δ side chain atoms, whereas hydrophobic side chains are rather constrained. The loop 90–108, disordered in the x-ray structure, is predicted here to be no more mobile than the rest of the structure.

The experimental x-ray r.m.s. displacement is 1 Å for all heavy atoms, double the calculated value. The experimental value includes translation libration of the entire disk, whereas the simulation only includes internal degrees of freedom. Analysis of crystallographic temperature factors (14, 50, 51), as well as diffuse x-ray scattering (52) indicate that translation-librations in protein crystals are approximately as large as the internal motions. Furthermore, a rigid-body model accounts qualitatively for the observed B-factors of several proteins (51). This shows that the contribution of translation libration to the B-factors varies along the polypeptide chain in about the same way as the intramolecular contribution. This is not surprising because the internal regions have both the smallest librations and the least flexibility. Without direct experimental information in the TMV case, we simply scale the x-ray displacements uniformly. The scaled displacements represent the internal motions of the protein. A residual R, based on residue-averaged B-factors, was defined by Sternberg et al. (50, 51). If we choose the scaling factor so as to minimize this residual, we obtain 0.5. Thus the normal mode calculation and the B-factors together predict that the rigid-body motions have the same amplitude as the internal motions. The correlation between the experimental and calculated B-factors, after scaling, is 0.45 (excluding the loop 90–108). For the residual R, we

TABLE 2 R.m.s. fluctuations of different classes of atoms

Atom class	X-ray	Calculated	ϵ_1	ϵ_9
Heavy atoms	1.00	0.49	0.53	0.46
Backbone	0.95	0.42	0.45	0.39
Side chains	1.05	0.55	0.59	0.52
Residues 90-108	>1.9	1.1	0.52	0.38
β atoms	1.06	0.48	0.52	0.46
γ atoms	0.90	0.57	0.61	0.55
σ atoms	0.90	0.57	0.62	0.55
Hydrophobic side chains	0.77	0.51	0.55	0.48

Results in Å. X-ray values (from the B-factors), normal mode result, normal mode result for the subspaces ϵ_1 and ϵ_9. The values for these subspaces are normalized to the total number of degrees of freedom of the disk.

obtain 0.85. For several proteins, studied by Kuriyan and Weis (51), fitting just the rigid-body model to the experimental B-factors gave R factors between 0.4 and 1.3, and correlation coefficients over 0.7.

The calculated and (scaled) experimental values of the r.m.s. displacements, averaged over each residue, are displayed in Fig. 6 as a function of residue number. The spatial variation of the displacements is similar to

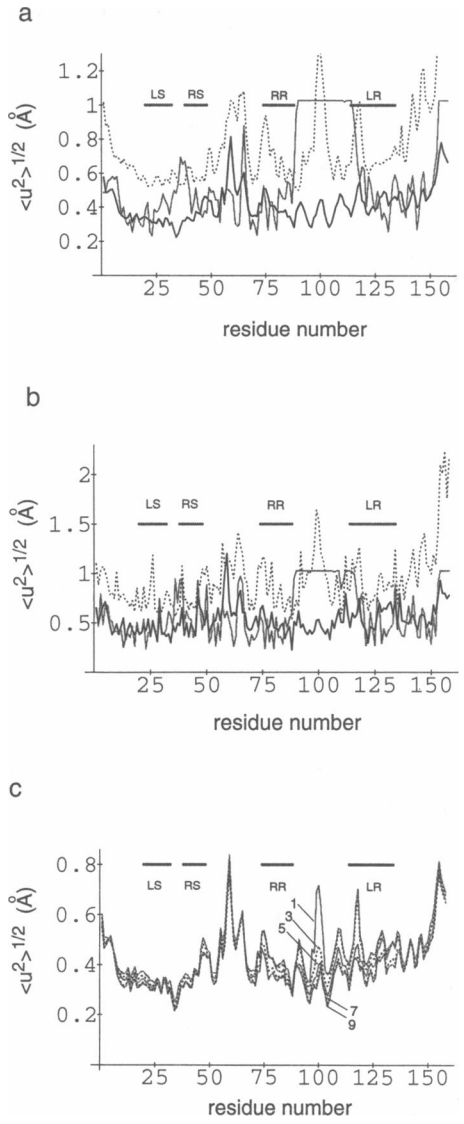

FIGURE 6 (a) R.m.s atomic displacements of backbone as a function of residue number for a given subunit in the TMV protein disk. Calculated: *black;* x-ray: *grey.* The core helices are highlighted. The MD result is also shown (*dotted*). (*b*) R.m.s. atomic displacements of side chains. (*c*) Contributions of the symmetry subspaces to the r.m.s. displacements of the Cα atoms. Subspaces 1, 3, 5, 7, and 9 are shown alternately as solid grey and dashed black curves, from bottom (ϵ_1) to top (ϵ_9).

that observed experimentally except in the loop region. The correlation coefficient between the two sets, excluding the loop 90–108, is 0.41 for the backbone atoms, and 0.48 for the side chains. The large mobility of the NH_2 and COOH-terminal regions is reproduced by the calculation, as well as the general shape of the curves. The mobility of the core helices is reproduced approximately. The largest discrepancies occur in (*a*) the inner loop 90–108, disordered in the x-ray structure; (*b*) residues Val58-Thr59, very mobile in the calculation; and (*c*) the segment 34–40, very mobile experimentally. In the real disk, Val58 and Thr59 are at the interface with the upper layer, near the outer surface of the disk. In our model, the upper layer of the disk is absent, and these residues are in a very different environment from the experimental one.

In the calculation, the loop 90–108 is not particularly mobile. The harmonic approximation only considers the shape of the energy surface immediately at the energy minimum. It will predict a low mobility if the curvature of the energy surface is large, even if the energy barrier for the loop to change conformations is less than its thermal energy. To investigate the stability of the minimized conformation, we did a short, 50 ps, molecular dynamics simulation of three subunits of the disk in vacuo. In this simulation, the loop occupies a different, but well-ordered conformation, with a 1.1 Å rms shift from the minimized loop; and it is only slightly more mobile than the rest of the structure (Fig. 6). This suggests that the disorder observed in the NMR and x-ray experiments may correspond to transitions on a sub-nanosecond time scale between conformations that are stable on a shorter time scale.

To measure the collective character of each mode, we count the number of α carbons having an r.m.s. displacement along that mode greater than the average value for the mode. By this criterion, all the modes in our calculation are collective modes, involving roughly ⅓ of one subunit. Gibrat and Go (12) found that modes up to 200 cm^{-1} in lysozyme are collective modes, which each involve over ⅓ of the molecule.

The fluctuations in the different symmetry subspaces are very similar on average, as seen in Fig. 5, which shows the contribution of each subspace to the average mean square displacement. The spectral density is shown, i.e. the contribution of each mode (*black dots*), as well as the integrated spectral density (*grey curves*). The spectral density appears as the superposition of two curves. The upper one corresponds to modes from spaces 2–9, which are two-fold degenerate. The lower one corresponds to modes from ϵ_1, which are nondegenerate. These displacements are normalized with respect to the number of degrees of freedom in each subspace, for purposes of comparison. The r.m.s. displacements of

different classes of atoms in the subspaces 1 and 9 are summarized in Table 2. On average, the amplitudes are greatest, and the motions softest, in the completely symmetric subspace 1, and hardest in subspace 9. Viewing the normal modes as waves propagating around the disk, the motions are more and more constrained as the wave number increases from 0 to $8\pi/17$, that is as we go from subspace 1 to subspace 9. This is what we would expect from an elastic continuum model of the disk.

Despite their average similarity, the subspaces make very different contributions in different parts of the molecule, as seen from Fig. 6 *c*. There are large differences in all four core helices and in the inner loop. For example subspaces 1–3 produce a large mobility in the RNA binding region, at the end of the LR helix and in the inner loop. The other subspaces contribute much smaller motions in this region. Here a simulation that restricts the system to the completely symmetric subspace 1 (see Introduction) would be seriously in error.

If we choose a smaller basis set $\{\mathbf{u}_k\}$, with a cutoff frequency of 100 cm^{-1} instead of 148 cm^{-1}, we saw that frequencies in the 0–50 cm^{-1} range are only slightly affected. The mean square displacement, on average, is lowered from 0.3 $Å^2$ to 0.29 $Å^2$. With this smaller basis set, the different subspaces are almost identical: the mean square displacements in all subspaces are between 0.3 $Å^2$ and 0.29 $Å^2$. With the larger basis set, the differences were much greater. Furthermore, they appeared in the 0–20 cm^{-1} range. Thus the effects of the frequency cutoff on the calculated displacements extend all the way down to the 0–20 cm^{-1} range. In the light of this, we cannot be sure that a cutoff of 148 cm^{-1} gives a good representation of even the slowest modes. This confirms the need for further calculations with a larger cutoff. Note that the cutoff of the basis set $\{\mathbf{u}_k\}$ should not be confused with the cutoff of the overall spectrum of the disk. The latter cutoff simply eliminates the high frequencies from the ensemble averages (r.m.s. atomic displacements, dielectric properties, et cetera). The former cutoff represents a set of constraints we impose on the individual subunits, to reduce the dimensionality of the system. It affects the entire calculated spectrum.

Correlations of the atomic motions

The mean correlations between the displacements of the α carbons are shown in Fig. 7. The correlation between two α carbons is represented by the shading of a point in a square matrix. Fig. 7 *a* shows the correlations within a subunit. The upper triangle of the matrix gives the correlations, the lower triangle the distances between the α carbons. Within each subunit, the correlation

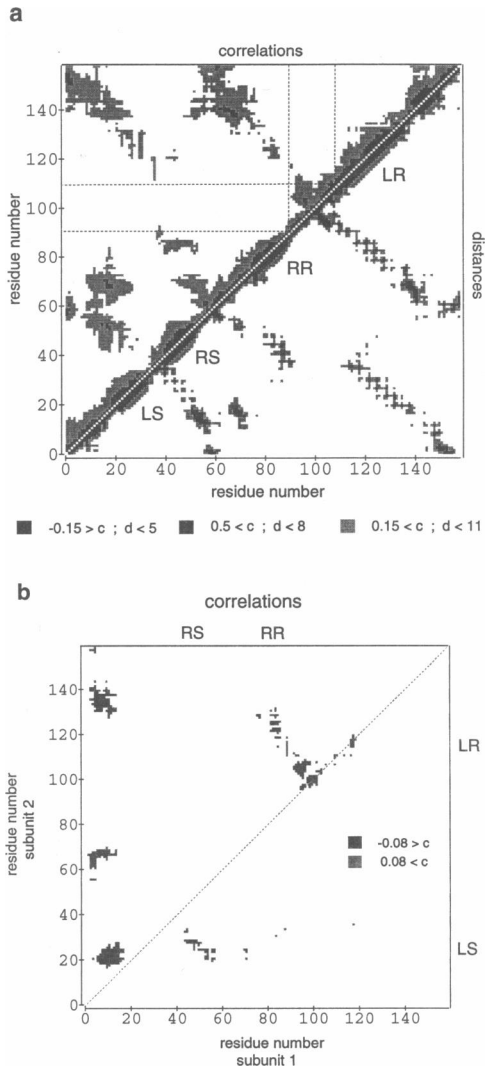

FIGURE 7 (*a*) Mean correlations (*upper triangle*) and distances (*lower triangle*) between α carbons within a single subunit. The diagonal is left blank for clarity. (*b*) Mean correlations between α carbons in subunits 1 and 2. A dashed line highlights the diagonal.

between two α carbons reflects closely the spatial separation between them. The core helices are visible (and labeled) along the main diagonal of the distance matrix. Within the helices the correlations are strong. The interfaces between the helices appear as four linear regions perpendicular to the main diagonal. As expected, the close packing of the helices leads to correlated motions. The flexible loop 90–108 extends between the LR and RR helices. It has no significant correlation with other parts of the polypeptide chain. The NH_2- and COOH-termini are linked by three hydrogen bonds, forming a very short two-stranded β-sheet, accounting

for the correlation observed in this region. The only significant negative correlations are between the segments 29–33 and 46–50, and between the segments 90–92 and 117–119; these range from −0.1 to −0.3. They are visible as tiny black rectangles in Fig. 7 *a*.

The strongest correlations (>0.5) systematically occur at separations <9 Å. Most of these occur in the range 5–8 Å. In particular, the C_i–C_{i+4} correlations along the core helices are virtually all >0.5, and correspond to separations of 5–6 Å. A few strong correlations occur in the 8–9 Å range. All but four of these are $C_i - C_{i+3}$ or $C_i - C_{i+4}$ pairs. The other four are all in the outer part of the disk: 1–57, 2–154, 2–155, 64–142, with correlations of 0.5–0.6, and distances of 8–9 Å. Weaker correlations extend to considerable distances. 40 pairs of α-carbons have correlations >0.3 at distances of 12–14 Å. Some of these are pairs 6–8 residues apart in the sequence. The others connect different segments of the polypeptide chain, in the outer, high-radius, part of the disk. Levitt et al. (8) predicted significant correlations at separations up to 20 Å in several proteins.

Fig. 7 *b* shows the correlations between the α carbons in the neighboring subunits 1 and 2. Again, the pattern of correlations follows closely the pattern of distances (not shown). There are contacts as well as correlations between the "right" helices of subunit 1 and the "left" helices of subunit 2, i.e. between *RR*(1) and *LR*(2), *RS*(1) and *LS*(2), *RR*(1) and *LS*(2) (see Fig. 1). The *RR*(1)–*LS*(2) contacts are more limited, and the correlations are weaker, barely visible in Fig. 7 *b*. However at a lower contour level distinct features appear in this region. The loops 90–108 of the two subunits are in contact and have significant correlations. When comparing inter- and intrasubunit correlations, for a comparable inter-carbon separation, the inter-subunit correlations are weaker. This is probably an effect of our limited basis set, made up of normal modes of subunit 1, with its neighbors held rigid. This basis set neglects some important degrees of freedom involving the subunit interfaces.

The correlations between subunits 1 and 3, which are separated by subunit 2, are an order of magnitude weaker than the 1–2 correlations. There are some weak negative correlations (≈ −0.1) between the helices *RR*(1) and *LR*(3), which are separated by *LR*(2) and *RR*(2). The strongest correlations are in the loop region, with correlations of 0.17 between the two segments 99–100. The width of a subunit in this region is 9 Å.

Within each symmetry subspace, the intrasubunit correlations are similar to the total values (the sum over all subspaces). The intersubunit correlations however are completely different from one subspace to the next, and completely different from the total values (Fig. 8). In general, the intersubunit correlations are much larger than the total, and there are systematic cancellations between the subspaces. This is a natural consequence of the symmetry of the system. Fig. 8 shows the correlations in subspaces 4 and 7. These subspaces give rise to large dielectric susceptibilities in the regions 110 and 70–80 of the peptide chain respectively (see below). The correlations between subunits 1 and 2 are systematically large and positive in subspace 4, and large and negative in subspace 7. This is to be expected, since the phase shift between the subunits is $6\pi/17$ in subspace 4 and $12\pi/17$ in subspace 7: in subspace 4 the subunits are roughly in phase; in subspace 7 they are roughly in phase opposi-

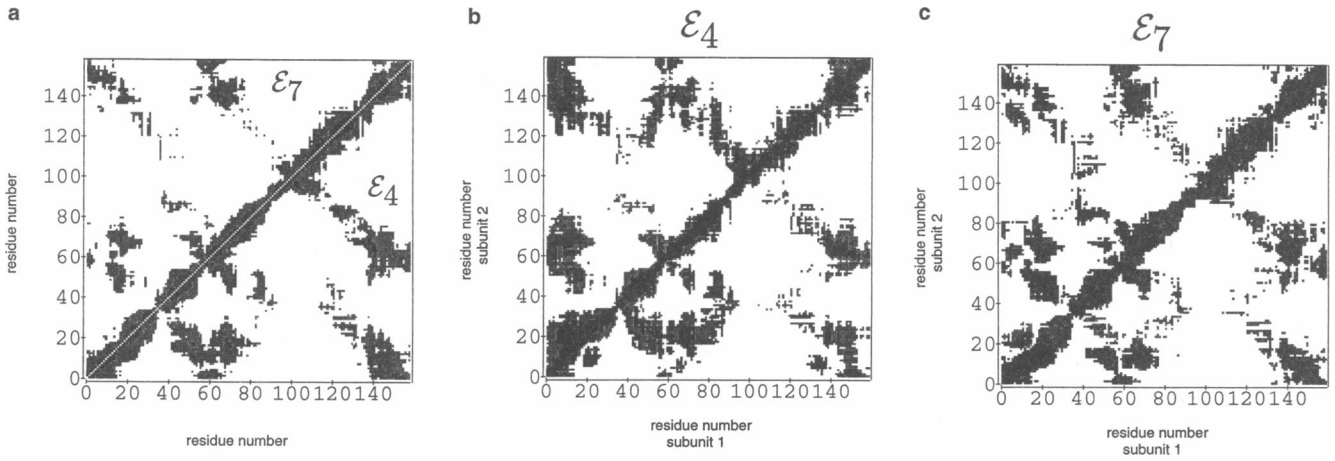

FIGURE 8 (*a*) Intrasubunit correlations in ϵ_4 (*lower triangle*) and ϵ_7 (*upper triangle*). Color coding as in Fig. 7a. (*b*) Correlations between subunits 1 and 2 in ϵ_4. Color coding as in Fig. 7*b*. (*c*) Correlations between subunits 1 and 2 in ϵ_7. Color coding as in *b*.

tion. On the other hand, the intrasubunit correlations are similar to the total values.

Levitt et al. (8) found mainly negative correlations between distant atoms in BPTI. Adjacent secondary structure units were seen to move mainly in opposite directions. In contrast we find few large negative correlations. We believe this is partly due to the low cutoff of both our reduced basis set and our overall spectrum. The low-frequency modes tend to have long wavelengths, so that atoms move approximately in phase over rather long distances. In particular, these modes tend to preserve the rather rigid secondary structure packing. Because our spectrum extends only to 100 cm^{-1}, it must omit many shorter wavelength motions, with core helices moving in phase opposition for example.

Interactions with perturbing charges

We noted earlier that the core helices and the loop 90–108 undergo conformational changes under the influence of external charges. These changes appear to play a role in viral assembly as well as the formation of protein helices. In particular, binding of a proton by a pair of carboxyl groups in the loop region can provoke ordering of the loop and helix formation. Binding of the viral RNA provokes loop ordering and formation of the viral helix.

The response of the disk to a set of external charges is determined by its dielectric properties. An approach that predicts these properties from a microscopic simulation was developed by Simonson et al. (53–55). When a small perturbing charge is introduced, the structure relaxes to lower the electrostatic energy, until the energy gained matches the elastic energy expended to deform the structure. The perturbation free energy can be split into a 'static' term and a relaxation term. The static term is equal to the interaction energy of the perturbing charges with the mean unperturbed structure. The relaxation free energy ΔA is related to the dielectric susceptibility tensor $\boldsymbol{\alpha}$:

$$\Delta A = -\frac{1}{2}\,{}^t\mathbf{f}\cdot\boldsymbol{\alpha}\mathbf{f}, \tag{12}$$

where $\mathbf{f}$ is the list of the perturbing fields $\mathbf{f}_i$ on the protein atoms i. The susceptibility tensor can be separated approximately into an electronic term, due to polarization of the electronic cloud, and a dipolar term, due to deformation of the 3-dimensional structure. Coupling between the two components gives a corrective term. All three terms can be calculated analytically in the harmonic or quasiharmonic approximation (53, 54). The electronic part is estimated using an atomic point polarizability model. The dipolar part is proportional to the mean dipole–dipole correlation matrix of the unperturbed protein.

A scalar susceptibility α is defined by

$$\Delta A = -\frac{1}{2}\alpha\mathbf{f}^2. \tag{13}$$

For a single perturbing point charge, α is a function of the perturbing charge's position, and can be thought of as a generalized polarizability. It has the form of a sum over the normal modes. The contribution of a mode or set of modes measures the coupling of these modes to the perturbation. If the kth mode contributes α_k to the scalar susceptibility, the perturbing charges will induce a deformation along the mode approximately proportional to α_k (see for example Eqs. 4 and 30 of reference 54). Because charge-dipole interactions are short-ranged, the susceptibility is largely a local property, expected to vary strongly from one region of the protein to another. Mobile, polar regions can deform easily and adjust their polar groups to screen a perturbing charge. They are thus expected to have large susceptibilities.

To probe the regions of the structure that have high or low susceptibilities, we consider a perturbing, point, test charge, which is placed successively on each α carbon. The electronic and dipolar parts of the scalar susceptibility are then calculated. The electronic contribution to the susceptibility is uniform throughout the structure, close to 0.8–0.9 $Å^3$. This is similar to other proteins (53), and appears to be due to the isotropic and nearly uniform atomic polarizabilities used. The dipolar susceptibility (Figs. 9, 10) varies markedly throughout the disk, with a mean value of 3.6 $Å^3$ and a standard deviation of

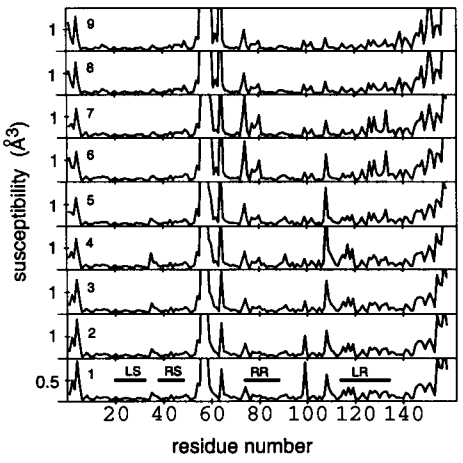

FIGURE 9 Dipolar susceptibility at each α carbon vs residue number; each panel shows the contribution of an individual symmetry subspace ϵ_i.

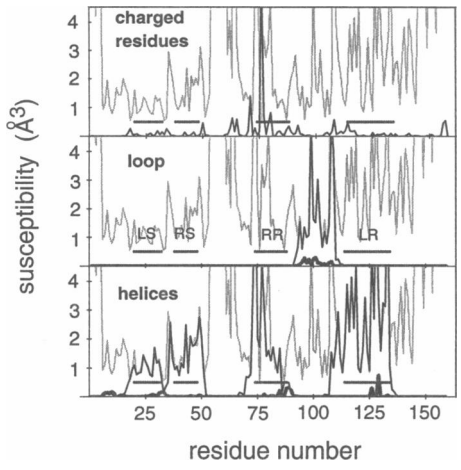

FIGURE 10 Susceptibilities produced in subunit 1 by different subsets of partial charges. (*Lower panel*) the charges in the core helices (of all subunits). The contribution of the core helices of all subunits except 1 is shown as a thick grey curve. (*Middle panel*) the charges in the loop 90–108, including all subunits. The contribution of subunits other than 1 is shown as a thick grey curve. (*Top panel*) the charges on ionized residues. The core helices are highlighted by horizontal lines. The total dipolar susceptibility is overlaid in light grey in each panel.

3.8 Å^3 (excluding residues 56–59, which have artificially large susceptibilities due to the absence of the upper layer of the disk). There are several regions of very high susceptibility. The most striking is at high radius on the outer face of the disk, including the NH_2- and COOH-termini, and the segment 56–59. The NH_2- and COOH-termini are very flexible, and there are a number of charged side chains at high radius, such as Arg61, Asp64,66, Lys68, Arg141 and Glu145. The segment 56–59 is also very mobile in the simulation, and exhibits enormous susceptibilities. In the real disk, this segment is restrained by the upper layer, and not very mobile. Large susceptibilities are also observed in the inner loop, and the LR and RR helices. This region contains many charged residues. The susceptibilities at the phosphate-binding arginines 90 and 92 are low, while those at the junction between the loop and the LR helix are very large.

The dielectric responses obtained in the different symmetry subspaces are quite different, as shown in Fig. 9. The nine mean values for the individual subspaces, normalized with respect to the number of degrees of freedom in each, vary from 3.5 Å^3 (ϵ_1, ϵ_2) to 5.2 Å^3 (ϵ_9). The lowest susceptibilities are thus obtained in the totally symmetric subspace. The susceptibilities around residue 60, as well as the COOH-terminus, increase smoothly going from subspace ϵ_1 to ϵ_9. Significant differences in the inner loop and the LR helix are observed. The susceptibility at the junction between the loop and the LR helix is largest in subspace ϵ_4, whereas the susceptibilities in the RR helix are largest in the subspaces ϵ_6 and ϵ_7. The differences between subspaces indicate the importance of collective motions involving several subunits. They are consistent with the different inter-subunit correlations in the different subspaces. The exact nature of the displacements responsible for these variations remains to be analyzed.

We can decompose the susceptibility into contributions of different groups of charges. Such a decomposition is merely approximate, as the susceptibility includes not only contributions from each group, but also cross-terms containing the correlations between the groups. Fig. 10 shows the susceptibilities produced within subunit 1 by three subsets of protein partial charges. The true normal modes are used, but only the subset of charges is included in the susceptibility calculation. The first panel shows the susceptibilities produced by the core helix charges, including all subunits. The contribution of the helices of subunits other than 1 is also shown, and is very small. The second panel corresponds to the charges in the loop 90–108. The third panel corresponds to the charges of all ionized side chains. The total susceptibility is overlaid on each panel for comparison. In each core helix, the total susceptibility is almost entirely accounted for by the charges of the core helices. The only exception is one susceptibility peak in the RR helix. This is consistent with the pattern of correlations of the displacements (Fig. 7): each helix has correlations mainly with itself and with its neighboring helices (including helices in neighboring subunits). Similarly, the charges in the loop 90–108 account very well for the susceptibilities in the loop. Neighboring subunits make only a small contribution. This is consistent with the small correlations between the loop and the rest of the structure (Fig. 7). Conversely, the loop and the core helices only contribute to the susceptibility locally. The susceptibilities produced by the core helices extend only slightly beyond the helices themselves; similarly for the loop.

The charged residues taken alone produce rather weak, localized, susceptibilities. One strong peak exists at Asp77, which is absent from the total susceptibility. The correlations of Asp77 with its surroundings produce destructive interference, which destroys the peak.

The susceptibility is a sum over the normal modes; the convergence of this sum is illustrated in Fig. 11. A test charge is placed in the vicinity of arginines 90 and 92 successively, and the susceptibility calculated with an increasing cutoff frequency. The test charges are placed along the C_α-C_ζ axis, 3.5 Å away from the amide groups. These positions are representative of counter-ions mak-

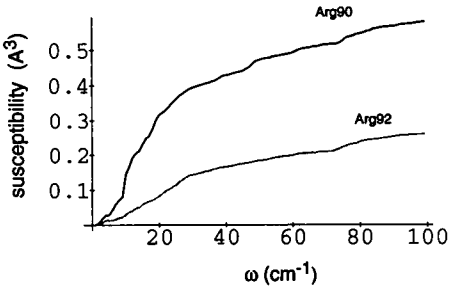

FIGURE 11 Susceptibilities in response to a test charge in the vicinity of Arg90 and 92 respectively (see text), including modes up to a cutoff frequency ω, as a function of ω. Arg90 test charge: *black*; Arg92 test charge: *grey*.

ing a salt-bridge with either arginine. We see that low-frequency modes dominate the susceptibilities at these two points. The susceptibilities increase rapidly at first, then more slowly, with convergence being reached at about 50 cm^{-1}. Thus many more modes are needed to describe the dielectric properties than to describe the average magnitude of motion. This was also observed in cytochrome c and deca-alanine (53): in these systems, convergence of the susceptibilities was only partially complete at 100 cm^{-1}.

Having placed a first test charge in subunit 1, consider a second test charge. If this charge is placed in subunit 3, the dielectric response will be nearly independent of the presence of the first test charge, because the correlations between subunits 1 and 3 are very small. The two susceptibilities will thus be additive. On the other hand, if it is placed in subunit 2, correlations between the subunits will add a corrective term.

CONCLUSION

We have presented a general harmonic and quasiharmonic method for studying the dynamics of large symmetric protein multimers. This method explicitly includes the long-range correlations between protein subunits. It is thus designed to address directly collective dynamic effects, such as cooperativity between subunits.

The group theoretical ideas we use are not all novel (44). For example symmetry coordinates were used to calculate the normal modes of gramicidin, which has C_2 symmetry (11). However, the form we give is completely general, yet phrased in terms of protein assemblies. It may help to make the use of symmetry more straightforward in this field. Because each block $\mathbf{V}^{ij}$ of the Hessian can be calculated, stored, and symmetrized, independently, the memory requirements for this kind of calculation are about the same as for a normal mode calculation of a single monomer, regardless of the size of the multimer. Cpu requirements are very modest. We propose a two-step procedure, which first selects the most important degrees of freedom of the individual monomers, that is their low-frequency modes, and eliminates the other degres of freedom; and second calculates the Hessian of the entire multimer in the reduced configuration space that results. In protein multimers, the monomers are usually small, and amenable to the individual normal mode calculation used in the first step here. This two-step procedure is not restricted to symmetric assemblies. In fact it can be generalized in an obvious way to a multi-step procedure, in which larger and larger subdivisions of the system are considered, and some of the degrees of freedom are eliminated at each step (Mouawad, L., and D. Perahia, manuscript in preparation). The basis of low-frequency modes of a monomer can be improved systematically by increasing the cutoff frequency; it incorporates automatically angle flexibility, which contributes significantly to the thermodynamic properties. Finally, we saw that for a quasiharmonic system, the covariance matrix of the atomic fluctuations has the same symmetry properties as the Hessian matrix, i.e. is reduced to the same block diagonal structure in a basis of symmetry coordinates. Thus the formulae given can be used directly in the quasiharmonic case.

The harmonic results presented for the TMV disk are preliminary, because the cutoff frequency on the basis set $\{\mathbf{u}_k\}$, 148 cm^{-1}, is not sufficient to include all the important degrees of freedom of the individual monomers. Our model includes 0.61 degrees of freedom per dihedral angle, and useful frequencies up to 100 cm^{-1} at most. Useful frequencies up to 200–250 cm^{-1} would be desirable. Other basis sets can be used, such as a set of normal modes of a monomer in vacuo. This would restrict motions at the subunit interfaces less. However the structure of a monomer minimized in vacuo would have to be introduced, which would differ significantly from a monomer in the disk. Our results also suffer from the general limitations of the harmonic approximation. Therefore, rather than extending the harmonic analysis to higher frequencies, or testing other basis sets, we are now focusing on a full quasiharmonic analysis.

Nevertheless, we believe that our spectrum is sufficiently accurate at this stage for some conclusions to be drawn, emphasizing their qualitative nature. First, the calculated displacements agree roughly with the x-ray displacements, with a correlation coefficient of 0.4–0.5. The internal loop 90–108 is not very mobile in the simulation, which suggests that conformation transitions on a subnanosecond time scale may account for the disorder observed by x-ray diffraction and NMR. The

overall translation-libration of the disk is estimated to have the same amplitude as the internal motions. The modes we obtain are all collective modes, which involve roughly one third of one subunit. We saw that a simulation that restricts the system to completely symmetric fluctuations would have made serious errors. Indeed, different subspaces make very different contributions to the displacements and susceptibilities of different parts of the structure.

In this harmonic model, the correlations of the small oscillations essentially extend from each subunit i to its nearest neighbors $i \pm 1$. Correlations between subunits i and $i \pm 2$ are an order of magnitude smaller. The correlations follow the pattern of distances closely, with most of the strong correlations occurring between nearby atoms. However significant correlations do extend to 8Å and even 12Å separations. This reflects fairly long-range mechanical coupling through the protein matrix. Furthermore, much longer-range correlations exist in the individual symmetry subspaces. The overall correlations are the result of systematic cancellations between these subspaces. The model does not allow for solvent-damping of the protein oscillations, or mixing between the normal modes due to anharmonicity in the potential energy function. These effects would reduce the long-range correlations. However on the time scale of a reactive event, damping and mixing can be small. If at a given time a sufficient phasing develops between the most important modes, then correlations of this magnitude will develop, which can presumably subsist over times that are shorter than characteristic damping and mixing times. Perturbing charges could amplify the correlations further. Perturbing charges in the region 70–80 of the chain, for example, tend to activate modes in spaces ϵ_6 and ϵ_7, since the susceptibility in this region is largely due to these subspaces. These modes exhibit negative correlations between neighboring subunits that are much larger than the total correlations.

The small oscillations of the inner loop 90–108 are correlated with the loops 90–108 of the neighboring subunits on either side, but they show very weak correlations with the rest of the structure. Similarly, the network of core helices has strong internal correlations, but weak correlations with the rest of the structure. It is not surprising therefore that in the loop region, the dielectric response is due to the direct effect of the loop alone. In particular, the model predicts that if perturbing charges, such as the RNA charges or a titrating proton, are introduced in this region, the loop will respond, but the rest of the structure will not be strongly affected. Obviously, the harmonic model assumes small oscillations around a single conformation. In reality other conformations exist for the loop, as shown by the NMR (39) and x-ray (36) data, and confirmed by our molecular dynamics simulation. The occupancies of the different loop conformations may have a much stronger correlation with other parts of the structure.

The low susceptibility at the two phosphate-binding arginines 90 and 92 indicates that soft elastic deformations will not be very effective at screening perturbing charges in this region. This suggests that a plastic deformation, such as a shift to a different conformation, is necessary to effectively screen a perturbing field.

The x-ray coordinates of the disk were kindly provided by Sir A. Klug, A. C. Bloomer, and G. Bricogne. Gérard Bricogne initially pointed out to us the importance of electrostatics in tobacco mosaic virus assembly. He contributed considerable support, discussion, and criticism to this and ongoing work.

Received for publication 12 June 1991 and in final form 8 October 1991.

REFERENCES

1. Brooks, C., M. Karplus & M. Pettit. 1987. *Adv. Chem. Phys.* 71: 1–259.
2. Noguti, T., and N. Go. 1982. Collective variable description of small-amplitude conformational fluctuations in a globular protein. *Nature (Lond.).* 776–778.
3. Levy, R., D. Perahia, and M. Karplus. 1982. Molecular dynamics of an α-helical polypeptide: temperature dependence and deviation from harmonic behavior. *Proc. Natl. Acad. Sci. USA.* 79:1346–1350.
4. Levitt, M., C. Sander, and P. Stern. 1983. The normal modes of a protein: native bovine pancreatic trypsin inhibitor. *Int. J. Quant. Chem. Symp.* 10:181–199.
5. Brooks, B., and M. Karplus. 1983. Harmonic dynamics of proteins: normal modes and fluctuations in bovine pancreatic trypsin inhibitor. *Proc. Natl. Acad. Sci. USA.* 80:6571–6575.
6. Frauenfelder, H., G. Petsko, and D. Tsernoglou. 1979. Temperature-dependent x-ray diffraction as a probe of protein structural dynamics. *Nature (Lond.).* 280:558–563.
7. Perahia, D., R. Levy and M. Karplus. 1990. Motions of an α-helical polypeptide: comparison of molecular and harmonic dynamics. *Biopolymers.* 29:645–677.
8. Levitt, M., C. Sander, and P. Stern. 1985. Protein normal mode dynamics: trypsin inhibitor, crambin, ribonuclease and lysozyme. *J. Mol. Biol.* 181:423–447.
9. Brooks, B. and M. Karplus. 1985. Normal modes for specific motions of macro-molecules: application to the hinge-bending mode of lysozyme. *Proc. Natl. Acad. Sci. USA.* 82:4995–4999.
10. Nishikawa, T., and N. Go. 1987. Normal modes of vibration in bovine pancreatic trypsin inhibitor and its mechanical property. *Proteins.* 2:308–329.
11. Roux, B., and M. Karplus. 1988. The normal modes of the gramicidin-A dimer. *Biophys. J.* 53:297–309.
12. Gibrat, J. F., and N. Go. 1990. Normal mode analysis of human lysozyme: study of the relative motion of the two domains and characterization of the harmonic motion. *Proteins.* 8:258–297.

13. Noguti, T., and N. Go. 1985. Efficient Monte Carlo method for simulation of fluctuating conformations of native proteins. *Biopolymers.* 24:527–546.

14. Diamond, R. (1990) On the use of normal modes in thermal parameter refinement: theory and application to the bovine pancreatic trypsin inhibitor. *Acta Cryst.,* A46:425–435.

15. Karplus, M., and J. N. Kushick. 1981. Method for estimating the configurational entropy of macromolecules. *Macromolecules.* 14:325–332.

16. Levy, R., M. Karplus, J. Kushick, and D. Perahia. 1984. Evaluation of the configurational entropy for proteins: application to molecular dynamics simulations of an α-helix. *Macromolecules.* 17:1370–1374.

17. Garcia, A. E., and D. M. Soumpasis. 1989. Harmonic vibrations and thermodynamic stability of a DNA oligomer in monovalent salt solution. *Proc. Natl. Acad. Sci. USA.* 86:3160–3164.

18. Brooks, C., A. Brünger, and M. Karplus. 1985. Active site dynamics in proteins: a stochastic boundary molecular dynamics approach. *Biopolymers.* 24:843–65.

19. Head-Gordon, T., and C. Brooks. 1991. Virtual rigid body dynamics. *Biopolymers.* 31:77–100.

20. Bennett, W., and R. Huber. 1984. Structural and functional aspects of domain motions in proteins. *CRC Crit. Rev. Biochem.* 15:291–385.

21. Perutz, M. 1989. Mechanisms of Cooperativity and Allosteric Regulation in Proteins, Cambridge University Press, Cambridge. 101 pp.

22. Caspar, D., and A. Klug. 1962. Physical principles in the construction of regular viruses. *Cold Spring Harbor Symp. Quant. Biol.* 27:1–24.

23. Caspar, D. 1963. Assembly and stability of tobacco mosaic virus. *Adv. Prot. Chem.* 18:37–121.

24. Butler, P. J. and A. Durham. 1977. Tobacco mosaic virus protein aggregation and the virus assembly. *Adv. Prot. Chem.* 31:187–251.

25. Holmes, K. 1984. The structure determination of tobacco mosaic virus. *In* Biological Macromolecules and Assemblies. Vol. 1. Virus Structures. F. Jurnak and A. McPherson, editors. John Wiley and Sons, Inc. New York. 121–148.

26. Stubbs, G. 1984. Macromolecular interactions in tobacco mosaic virus. *In* Biological Macromolecules and Assemblies. Vol. 1. Virus Structures. F. Jurnak, and A. McPherson, editors. John Wiley and Sons, Inc., New York. 149–202.

27. Butler, P. J., and A. Klug. 1971. Assembly of the particle of tobacco mosaic virus from RNA and disks of protein. *Nature (Lond.).* 229:47–50.

28. Franklin, R. 1956. Location of the ribonucleic acid in the tobacco mosaic virus particle. *Nature (Lond.).* 177:928–930.

29. Mandelkow, E., G. Stubbs and S. Warren. 1981. Structures of the helical aggregates of tobacco mosaic virus protein. *J. Mol. Biol.* 152:375–386.

30. Namba, K., and G. Stubbs. 1986. Structure of tobacco mosaic virus at 3.6 Å resolution: implications for assembly. *Science (Wash. DC).* 231:1401–1406.

31. Namba, K., R. Pattanayek R. and G. Stubbs. 1989. Visualization of protein-nucleic acid interactions in a virus. Refined structure of intact tobacco mosaic virus at 2.9 Å resolution by x-ray fiber diffraction. *J. Mol. Biol.* 208:307–325.

32. Zimmern, D. 1977. The nucleotide sequence at the origin for assembly on tobacco mosaic virus RNA. *Cell.* 11:463–482.

33. Shalaby, R., and M. Lauffer. 1977. Hydrogen ion uptake on tobacco mosaic virus protein polymerization. *J. Mol. Biol.* 116: 709–725.

34. Klug, A., and A. Durham. 1972. The disk of tobacco mosaic virus protein and its relation to the helical and other modes of aggregation. *Cold Spring Harbor Symp. Quant. Biol.* 36:449–460.

35. Champness, J., A. Bloomer, G. Bricogne, P. J. Butler, and A. Klug. 1976. *Nature (Lond.).* 259:20–24.

36. Bloomer, A., J. Champness, G. Bricogne, R. Staden, and A. Klug. 1978. Protein disk of tobacco mosaic virus at 2.8 Å resolution showing the interactions within and between subunits. *Nature (Lond.).* 276:362–368.

37. Graham, J., and P. J. Butler. 1978. Location of tyrosine residues in the disk of tobacco mosaic virus protein and comparison of the subunit packing with that of the virus. *Eur. J. Biochem.* 83:523–528.

38. Graham, J., and P. J. Butler. 1979. Binding of oligonucleotides to the disk of tobacco mosaic virus protein. *Eur. J. Biochem.* 93:333–337.

39. Jardetzky, O., K. Asaka, D. Vogel, S. Morris and K. Holmes. 1978. Unusual segmental flexibility in a region of tobacco mosaic virus coat protein. *Nature (Lond.).* 273:564–566.

40. Altschuh, D., A. Lesk, A. Bloomer, and A. Klug. 1987. Correlation of coordinated amino acid substitutions with function in viruses related to tobacco mosaic virus. *J. Mol. Biol.* 193:693–707.

41. Stubbs, G., S. Warren, and K. Holmes. 1977. Structure of RNA and RNA binding site in tobacco mosaic virus from 4 Å map calculated from x-ray fiber diagrams. *Nature (Lond.).* 267:216–21.

42. Blinc, R., and B. Zeks, 1974. Soft Modes in Ferroelectrics and Antiferroelectrics. Elsevier Science Publishing Co., Inc., New York.

43. Wilson, E. B., J. C. Decius and P. C. Cross. 1955. Molecular Vibrations. McGraw-Hill, New York.

44. Lyubarskii, G. 1960. The Application of Group Theory in Physics. Pergamon Press, New York.

45. Brooks, B., R. Bruccoleri, B. Olafson, D. States, S. Swaminathan, and M. Karplus. 1983. CHARMM: a program for macromolecular energy, minimization, and molecular dynamics calculations. *J. Comp. Chem.* 4:187–217.

46. Loncharich, R., and B. Brooks. 1988. The effect of truncating long-range forces on protein dynamics. *Proteins.* 6:32–45.

47. Go, N. 1978. Shape of the conformational energy surface near the global minimum and low-frequency vibrations in the native conformation of globular proteins. *Biopolymers.* 17:1373–1379.

48. McQuarrie, D. 1976. Statistical Mechanics. Harper and Row, New York.

49. Westhof, E., D. Altschuh, D. Moras, A. C. Bloomer, A. Mondragon, A. Klug, and M. H. V. Van Regenmortel. 1984. Correlation between segmental mobility and the location of antigenic determinants in proteins. *Nature, (Lond.).* 311:123–126.

50. Sternberg, M., D. Grace, and D. C. Phillips. 1979. Dynamic information from protein crystallography. An analysis of temperature-factors from refinement of hen egg-white lysozyme structure. *J. Mol. Biol.* 130:231–253.

51. Kuriyan, J., and W. Weis. 1991. Rigid protein motion as a model for crystallographic temperature factors. *Proc. Natl. Acad. Sci. USA,* 88:2773–2777.

52. Doucet, J., and J. P. Benoit. 1987. Molecular dynamics studied by analysis of the x-ray diffuse scattering from lysozyme crystals. *Nature (Lond.)*. 325:643–646.

53. Simonson, T., D. Perahia, and G. Bricogne. 1991. Intramolecular dielectric screening in proteins. *J. Mol. Biol.* 218:859–886.

54. Simonson, T., D. Perahia and A. Brünger. 1991. Microscopic theory of the dielectric properties of proteins. *Biophys. J.* 59:670–690.

55. Simonson, T., D. Perahia, G. Bricogne, and A. Brünger. 1991. Dielectric properties of proteins: microscopic and macroscopic theories. *J. Chim. Phys. (Paris)*. In press.

56. Kraulis, P. 1991. *J. Appl. Cryst.* In press.

57. Butler, P. J., A. C. Bloomer, and J. T. Finch. 1992. Direct visualization of the "205" aggregate of coat protein of tobacco mosaic virus: the "disk" is the major structure at pH 7 and the proto-helix at lower pH. *J. Mol. Biol.* In press.