

# Functional Motions of Influenza Virus Hemagglutinin: A Structure-Based Analytical Approach

Basak Isin,<sup>\*†</sup> Pemra Doruker,<sup>\*‡</sup> and Ivet Bahar<sup>†‡</sup>

<sup>\*</sup>Polymer Research Center and Chemical Engineering Department, Bogazici University, Bebek 80815, Istanbul, Turkey, and <sup>†</sup>Center for Computational Biology and Bioinformatics, and Department of Molecular Genetics and Biochemistry, School of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania 15213 USA

**ABSTRACT** Influenza virus hemagglutinin (HA), a homotrimeric integral membrane glycoprotein essential for viral infection, is engaged in two biological functions: recognition of target cells' receptor proteins and fusion of viral and endosomal membranes, both requiring substantial conformational flexibility from the part of the glycoprotein. The different modes of collective motions underlying the functional mobility/adaptability of the protein are determined in the present study using an extension of the Gaussian network model (GNM) to treat concerted anisotropic motions. We determine the molecular mechanisms that may underlie HA function, along with the structural regions or residues whose mutations are expected to impede function. Good agreement between theoretically predicted fluctuations of individual residues and corresponding x-ray crystallographic temperature factors is found, which lends support to the GNM elucidation of the conformational dynamics of HA by focusing upon a subset of dominant modes. The lowest frequency mode indicates a global torsion of the HA trimer about its longitudinal axis, accompanied by a substantial mobility at the viral membrane connection. This mode is proposed to constitute the dominant molecular mechanism for the translocation and aggregation of HAs, and for the opening and dilation of the fusion pore. The second and third collective modes indicate a global bending, allowing for a large lateral surface exposure, which is likely to facilitate the close association of the viral and endosomal membranes before pore opening. The analysis of kinetically hot residues, in contrast, reveals a localization of energy centered around the HA2 residue Asp112, which apparently triggers the solvent exposure of the fusion peptide.

## INTRODUCTION

Influenza virus hemagglutinin (HA) is a glycoprotein integral to the influenza virus envelope. HA is involved in two major functions: recognition of target cells, by binding to their sialic acid-containing receptors, and fusion of the viral and the endosomal membranes succeeding endocytosis (White et al., 1997). Understanding the structure and dynamics of HA is essential for designing novel antiviral agents that can potentially inhibit its binding or fusogenic activities.

The x-ray structure of HA has been determined (Wilson et al., 1981) and refined (Weis et al., 1990b) by Wiley and coworkers. HA is a cylindrically shaped trimer ~135 Å long, varying between 35 and 70 Å along the radial direction. It is composed of three identical monomers (Fig. 1 *a*) assembled into a central  $\alpha$ -helical coiled coil that forms the stem-like domain, and three globular heads containing the sialic acid-binding sites (Fig. 1 *b*). The monomers originate upon cleavage of the individual chains of the fusion-inactive precursor (HA0) into two polypeptides each, HA1 and HA2. The polypeptides HA1 and HA2 are linked by two intramonomer disulfide bridges that are presumably formed during the folding of HA0 in the endoplasmic reticulum

(ER). The free ends at the cleavage site, forming the C-terminus of HA1 and the N-terminus of HA2, are snapped apart by 20 Å (Fig. 1 *a*). Each monomer is anchored in the viral membrane by a helical transmembrane peptide of 27 amino acids near the C-terminus of each HA2 chain. This region is not seen in the x-ray structure.

The globular heads of HA are formed by the HA1 residues 116–261 folded into a jelly-roll motif of eight antiparallel  $\beta$ -strands. The distal tips of the globular heads contain the highly conserved shallow pockets that are the receptor binding sites (Weis et al., 1988), shown in green in Fig. 1 *a*. The virus exploits the sialic acid residues in carbohydrate side chains of cellular receptor proteins. The receptor binding pocket of HA is surrounded by antigenically variable antibody binding sites. Antibodies bound to these sites should, in principle, block the binding to receptor proteins. Yet, their binding may be ineffective due to the antibody-selected variation of the nonconserved residues on the edges of their target site (Bizebard et al., 1995; Wiley et al., 1981; Weis et al., 1988).

Viral infectivity requires a second activity of HA, membrane fusion, in addition to binding to host cell receptors. Viruses bound to the plasma membrane are endocytosed in coated pits and vesicles, and delivered to endosomes. Proton pumps in the membranes of endocytic vesicles induce an accumulation in protons, and a consequent lowering of the pH inside the vesicles. Acidic pH (between pH 5 and 6) triggers a conformational change in HA, such that the HA molecule becomes fusogenic (White and Wilson, 1987). The structure displayed in Fig. 1 *c* has been shown to be a stable structure at low pH by both x-ray crystallography

Submitted February 15, 2001, accepted for publication October 29, 2001.

Address reprint requests to Ivet Bahar, University of Pittsburgh School of Medicine, Center for Computational Biology & Bioinformatics, Suite 601, Kaufmann Bldg., 3471 Fifth Ave., Pittsburgh, PA 15213. Tel.: 412-648-6671; Fax: 412-648-6676; E-mail: bahar@pitt.edu.

<sup>‡</sup>Both authors contributed equally to this work.

© 2002 by the Biophysical Society

0006-3495/02/02/569/13 \$2.00

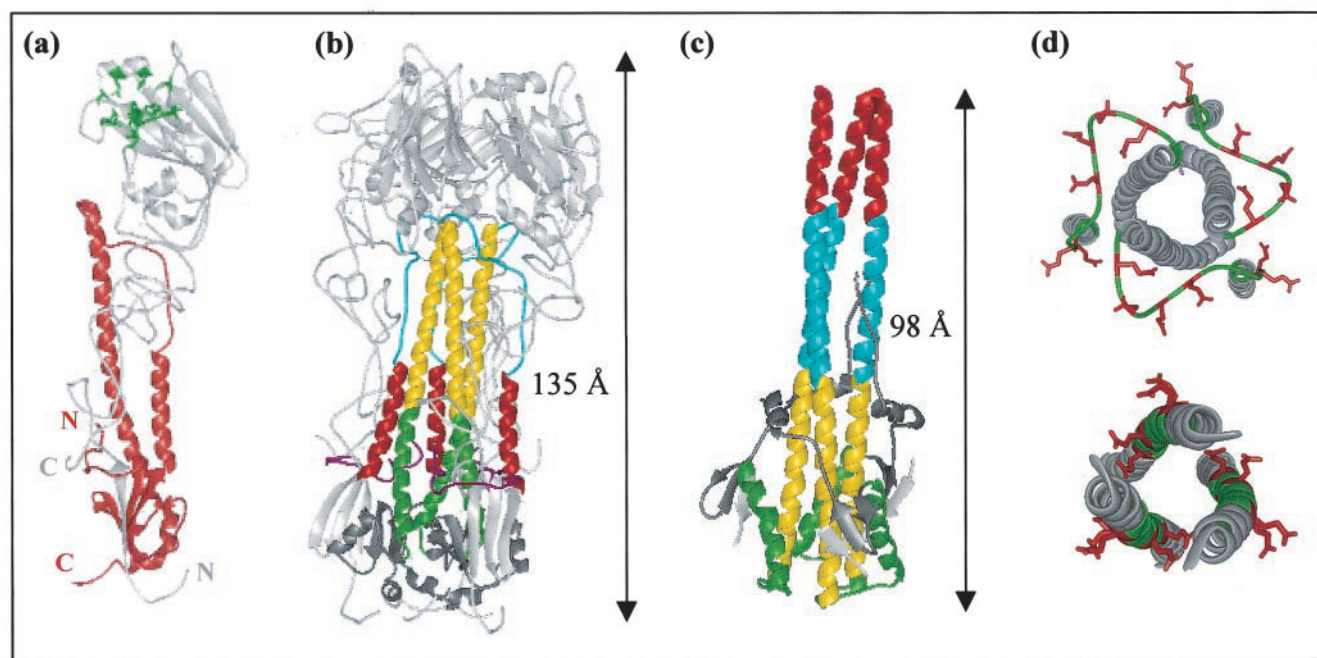


FIGURE 1 Ribbon diagrams of (a) the polypeptide chains HA1 (gray) and HA2 (red) of a single monomer of the influenza virus HA in the bromelain-digested neutral pH form (BHA) (Weis et al., 1990b; PDB code: 2HMG); (b) the trimeric structure of HA in the same neutral pH form; and (c) the trimeric fragment (TBHA<sub>2</sub>) composed of residues 38–175 of the HA2 chain disulfide linked to residues 1–27 of the HA1 chain at low pH (Bullough et al., 1994; PDB code: 1HTM). In (a), the N- and C-termini of both chains are indicated, and the receptor binding residues are displayed in green. The fusion peptides of the individual monomers are colored magenta in panel b. In panels b and c, the HA2 segments colored red (helix A; residues 38–55), blue (coil in b, helix B in c; residues 56–75), green (helix D; residues 106–129), and dark gray (residues 130–175 including helices E–H), undergo a substantial conformational change at low pH. The yellow portion (helix C, residues 76–105) preserves its original (left-handed  $\alpha$ -helical coiled coil) conformation. Panel d displays the top views of the conformation of the HA2 residues 38–105 in BHA (top) and TBHA<sub>2</sub> (bottom). Glu residues (explicitly shown in red) located in the BHA loop (green) (upper) play a dominant role in stabilizing the TBHA<sub>2</sub> B-helix (lower).

(Bullough et al., 1994), cryoelectron microscopy (Shang-guan et al., 1998) and theoretical arguments based on heptad repeat patterns (Carr and Kim, 1993; Carr et al., 1997). This low pH form (TBHA<sub>2</sub>) necessitates the reconfiguration of an entire loop (blue in Fig. 1, *b–c*) into a helix (helix B), and the overall translation and reorientation of the helix A (red) to extend the original triple-stranded helical coil (Fig. 1 *d*). This substantial rearrangement also requires prior opening of the globular heads to avoid steric clashes. Such a large-scale conformational change has been pointed out to be a rather slow process, which requires even a longer time than the fusion process, itself (Shang-guan et al., 1998). The identification of TBHA<sub>2</sub> with the fusion-competent form of HA is therefore open to discussion. In a broader context, a major issue still unresolved and addressed here is to understand the molecular mechanism of fusion, or to assess the cooperative conformational rearrangements of the trimer that are most likely conducive to fusion.

Biochemical, and molecular biological studies have established amino acids 1–20 of HA2 as the fusion peptide of the influenza virus HA (Wilson et al., 1981; Daniels et al., 1985; Weis et al., 1990a; White, 1992) (shown in magenta in Fig. 1 *b*). The fusion peptide is

relatively nonpolar. It is rich in glycines, which impart high flexibility. In the neutral pH form, the N-termini of three fusion peptides are buried in the interface between the subunits of the trimer  $\sim 30$  Å from where the protein inserts into the lipid bilayer of the virus envelope. At low pH, this region undergoes a conformational change to allow for the exposure of the fusion peptides, which is essential for fostering a hydrophobic association between the HA molecule and the target membrane. The fusion activity of HA is altered, or abolished, by site-specific mutations within the fusion peptide (Daniels et al., 1985; Gething et al., 1986; Qiao et al., 1998) or by inhibiting the exposure of fusion peptide with designed disulfide linkages (Godley et al., 1992).

It is clear that the processes of binding to receptor sites and membrane fusion require both substantial conformational flexibility from the part of the HA molecule. Knowledge of HA structure has been informative for visualizing the domains involved in specific functions. However, the biological activity of HA is closely coupled to the dynamics of the molecule, i.e., its ability to undergo collective conformational changes adjusting to different environments and

functions. A rigorous analysis of the dynamics of HA is made here.

We use a structure-based analytical approach that has proven to disclose the mechanisms of functional motions for a number of other biomolecular complexes (Bahar and Jernigan; 1998, 1999; Bahar et al., 1999; Jernigan et al., 1999; Bahar, 1999; Keskin et al., 2000). The structure is modeled as an elastic network, the nodes of which are the individual residues. The overall topology of the network, or the distribution of native contacts, is accounted for by a Kirchhoff connectivity matrix (Bahar et al., 1997), which has found wide applications in graph theory (Harary, 1971). The model bears a unique analytical solution. It permits us, by a mode decomposition technique, to identify the collective modes leading to global conformational changes and the critical sites controlling these modes. Quaternary structures of the order of  $10^3$  residues, which would necessitate molecular dynamics simulations of the order of weeks if examined at atomic resolution, can be explored within hours using this method. We note that HA is a trimer of  $\sim 1500$  residues.

Both the neutral form of the HA trimer (Fig. 1 *b*) and the fragment crystallized at low pH (Fig. 1 *c*) the coordinates of which are deposited in the Protein Data Bank (PDB) (Berman et al., 2000) are considered here. The former, obtained by the treatment of the X:31 virus with the protease bromelain to produce the water-soluble ectodomain, is shortly referred to as BHA (Weis et al., 1990b); and the latter, prepared from BHA by digestion with trypsin and thermolysin, is called TBHA<sub>2</sub> (Bullough et al., 1994). After a verification of the validity of the model by comparing the calculated fluctuation spectrum with that experimentally observed, we will filter out the uninteresting modes (Ama-dei et al., 1993) and focus on the shapes of the extracted dominant motions. The results will be combined with existing experimental data and other theoretical results to propose a molecular mechanism of action underlying membrane fusion and pore opening. Target sites whose mutation would be expected to disrupt the function or stability of HA will be identified.

## THEORY AND METHOD

In the elastic network model of proteins (Tirion, 1996; Bahar et al., 1997), the interactions between residues in close proximity are represented by harmonic potentials with a uniform spring constant  $\gamma$ . The physical basis and the analytical method for examining the collective dynamics of such a network (Haliloglu et al., 1997; Bahar et al., 1998) bear close similarities with the elasticity theory of random polymer networks (Flory, 1976). The junctions of the network are conveniently identified here with the C $^\alpha$  atoms. Any deformation from native-state coordinates is resisted by the linear springs that associate the close-neighboring (bonded and nonbonded) residues. Residues are subject to Gaussianly distributed fluctuations, hence the name Gaussian network model (GNM).

## General formulation

In the GNM, the fluctuations  $\Delta\mathbf{R}_{ij}$  in the separation  $\mathbf{R}_{ij}$  between the *i*th and *j*th residues obey the Gaussian distribution,

$$W(\Delta\mathbf{R}_{ij}) = [\gamma/(2\pi k_B T)]^{3/2} \exp\{-(\gamma/2)\Delta\mathbf{R}_{ij}^2/k_B T\}, \quad (1)$$

where  $k_B$  is the Boltzmann constant, and  $T$  is the absolute temperature. The configurational partition function for a protein of  $N$  residues is expressed with analogy to the theory of random networks (Flory, 1976) as

$$Z_N = \int \exp\left(-\frac{\gamma}{2k_B T} \{\Delta\mathbf{R}^T\} \Gamma \{\Delta\mathbf{R}\}\right) d\{\Delta\mathbf{R}\}. \quad (2)$$

Here,  $\{\Delta\mathbf{R}\}$  represents the  $3N$ -dimensional ( $3ND$ ) column vector  $\{\Delta\mathbf{R}_1, \Delta\mathbf{R}_2, \Delta\mathbf{R}_3, \dots, \Delta\mathbf{R}_N\}$  of the fluctuations in the position vectors of the individual residues, the superscript  $T$  denotes the transpose, and  $\Gamma$  is the  $N \times N$  Kirchhoff matrix. The elements of  $\Gamma$  are (Bahar et al., 1997) defined as

$$\Gamma_{ij} = \begin{cases} -1 & \text{if } i \neq j \text{ and } \mathbf{R}_{ij} \leq r_c \\ 0 & \text{if } i \neq j \text{ and } \mathbf{R}_{ij} > r_c \\ -\sum_{i \neq j} \Gamma_{ij} & \text{if } i = j \end{cases} \quad (3)$$

based on a cutoff distance  $r_c$  for inter-residue interactions. A lower bound for  $r_c$  is the first interaction shell radius of 7.0 Å extracted from statistical analyses (Miyazawa and Jernigan, 1985, 1996; Bahar and Jernigan, 1997) of PDB structures. The force constant  $\gamma$  is the only adjustable parameter of the theory for a fixed cutoff distance. Its magnitude, usually of the order of 1 kcal/(mol Å<sup>2</sup>), is found by rescaling the theoretically predicted mean-square (ms) residue fluctuations so that their average value matches that indicated by the experimental B-factors.  $\gamma$  affects only the absolute size of fluctuations. The relative fluctuation amplitudes of the individual residues, or the distribution of fluctuations are not affected by the value of the parameter  $\gamma$ .

## Equilibrium dynamics and modal contributions to collective motions

The equilibrium correlations between the fluctuations of residues *i* and *j* are obtained from

$$\begin{aligned} \langle \Delta\mathbf{R}_i \cdot \Delta\mathbf{R}_j \rangle &= (1/Z_N) \int (\Delta\mathbf{R}_i \cdot \Delta\mathbf{R}_j) e^{-V/k_B T} d\{\Delta\mathbf{R}\} \\ &= (3k_B T/\gamma) [\Gamma^{-1}]_{ij}, \end{aligned} \quad (4)$$

where  $V$  is the total potential energy

$$V = (\gamma/2) \{\Delta\mathbf{R}^T\} \Gamma \{\Delta\mathbf{R}\}, \quad (5)$$

and  $[\Gamma^{-1}]_{ij}$  is the *ij*th element of the inverse of  $\Gamma$ . The ms fluctuations  $\langle (\Delta\mathbf{R}_i)^2 \rangle$  of individual residues are evaluated from the last equality in Eq. 4, using  $i = j$ .

The eigenvalue decomposition of  $\Gamma$  reads  $\Gamma = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T$  where  $\mathbf{U}$  is an orthogonal matrix whose columns  $u_i$ ,  $1 \leq i \leq N$ , are the eigenvectors of  $\Gamma$ , and  $\mathbf{\Lambda}$  is the diagonal matrix of the eigenvalues ( $\lambda_i$ ), usually organized in ascending order, i.e.,  $\lambda_1 \leq \lambda_i \leq \lambda_{N-1}$ , and  $\lambda_N = 0$ . The *i*th eigenvector reflects the shape of the *i*th mode as a function of residue index, and the *i*th eigenvalue represents its frequency (Haliloglu et al., 1997; Bahar and Jernigan, 1998; Bahar, 1999). The correlations between the fluctuations of



residues  $i$  and  $j$  result from the superposition of  $(N - 1)$  modes in the GNM. The contribution of the  $k$ th mode,  $\langle \Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_j \rangle_k$ , is given by

$$\langle \Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_j \rangle_k = (3k_B T / \gamma) [\lambda_k^{-1} \mathbf{u}_k \mathbf{u}_k^T]_{ij} \quad (6)$$

and the cumulative effect of a subset  $k_1 \leq k \leq k_2$  of modes is found by summing up the above equation over the investigated subset.

## Mechanisms of collective motions

The mechanisms of motions are determined by the extension of the GNM to the 3ND space of collective modes, using the anisotropic network model (ANM) (Doruker et al., 2000; Atilgan et al. 2001). Whereas the inter-residue distances are controlled by harmonic potentials in the GNM, ANM adopts the further assumption that the three ( $x$ ,  $y$ , and  $z$ ) components of the inter-residue separation vectors obey Gaussian dynamics. ANM thus involves the inversion of a  $3N \times 3N$  Hessian matrix  $\mathcal{H}$  that replaces the  $N \times N$  Kirchhoff matrix  $\Gamma$ .  $\mathcal{H}$  is composed of  $N \times N$  super elements  $\mathbf{H}_{ij}$  ( $1 \leq i, j \leq N$ ) each of size  $3 \times 3$ , given by

$$\mathbf{H}_{ij} = \begin{bmatrix} \partial^2 V / \partial X_i \partial X_j & \partial^2 V / \partial X_i \partial Y_j & \partial^2 V / \partial X_i \partial Z_j \\ \partial^2 V / \partial Y_i \partial X_j & \partial^2 V / \partial Y_i \partial Y_j & \partial^2 V / \partial Y_i \partial Z_j \\ \partial^2 V / \partial Z_i \partial X_j & \partial^2 V / \partial Z_i \partial Y_j & \partial^2 V / \partial Z_i \partial Z_j \end{bmatrix}, \quad (7)$$

for  $i \neq j$ . See Atilgan et al. (2001) for explicit expressions. The diagonal super elements  $\mathbf{H}_{ii}$  are found from the negative sum of the off-diagonal terms in the same row. Inversion of  $\mathcal{H}$  yields  $3N - 6$  normal modes, the frequencies and shapes of which are defined by the nonzero eigenvalues and corresponding eigenvectors of  $\mathcal{H}$ . The  $k$ th eigenvector, for example, is a 3ND array of  $N$  blocks, each (of size  $3 \times 1$ ) representing the fluctuation vector of an individual residue in the  $k$ th mode.

## Limitations of the GNM/ANM

It is worth emphasizing that the major advantage of the GNM (and ANM) is the fact that it lends itself to an analytical solution. The solution is deterministic for a given architecture; it is a unique function of the particular topology of contacts. It does not require an energy minimization before calculations, not even inclusion of specific energy parameters, in contrast to conventional normal mode analyses. In contrast, a coarse-grained model, usually single site per residue, is adopted, and structural information at atomic level is lost.

Both GNM and ANM yield the equilibrium fluctuation dynamics of globular proteins around the native state. The fluctuations predicted by these models do not include the anharmonic changes such as local conformational isomeric jumps or large-scale conformational changes of proteins that have more than one equilibrium state. Molecular dynamics (MD) simulations, in contrast, can capture the anharmonic motions. However, individual MD trajectories often exhibit different time evolutions and multiple MD runs need to be carried out to overcome MD sampling problems. An apparent harmonic behavior usually emerges as a result of averaging over MD multiple trajectories. The neglect of anharmonic contributions has indeed negligible influence on the frequency spectrum of fluctuations around a stable conformation (Hinsen et al., 1999; Hinsen and Kneller, 1999). In a recent study, we showed that the millisecond fluctuations and the cross-correlations between domain movements obtained from GNM/ANM are consistent with MD simulations results (Doruker et al., 2000).

The folded structure is assumed here to be a fully elastic network, or a mechanical device, whose topology is defined by the native-state inter-residue contacts. The theory does not distinguish between the type and stiffness of specific interactions. All deformations in inter-residue distances are assumed to be subject to the same force constant. In principle, one can assign different spring constants to different pairs and use them readily in

the Kirchhoff matrix, without any mathematical complication. The adoption of different force constants is equivalent to changing the curvature of the energy curve near the minimum (equilibrium state), rather than changing the absolute height of the minimum. We tested the effect of differentiating between different types of interactions, by repeating the calculations with different force constants for bonded and nonbonded pairs. No detectable improvement in the agreement with experiments was observed, which would justify the introduction of one or more new energy parameters (Bahar et al., 1997). Likewise, the global modes have been verified to remain unaffected by increasing the stiffness of disulfide bridges of HA by up to a factor of five. The robustness and reproducibility of the global modes in spite of the changes in the details of the models are indeed fundamental features in support of mode analyses performed with low resolution models (Kitao and Go, 1999; Hinsen et al., 1999; Freire, 2000; Doruker et al., 2000).

## RESULTS AND DISCUSSION

### Equilibrium fluctuations: comparison with experiments

Figure 2 compares the Debye–Waller factors (or B-factors) of HA1 (part *a*) and HA2 (part *b*) residues measured for BHA by x-ray crystallography (Weis et al., 1990b) (*bold solid curve* in both cases) with those predicted by the GNM and ANM theories (*thin continuous* and *dotted curves*, respectively). The three monomers exhibit practically the same fluctuation behavior in experiments; and their almost indistinguishable behavior is confirmed by the theoretical calculations. Thus, we report the fluctuations averaged over the three monomers, although the calculations have been done for the trimeric structure of 1509 residues. The experimental Debye–Waller factors are related to the ms fluctuations of individual residues ( $1 \leq i \leq N$ ) as  $B_i = (8\pi^2/3) \langle \Delta R_i^2 \rangle$ , which, upon substitution from Eq. 4, can be expressed in terms of the diagonal elements of  $\Gamma^{-1}$  as

$$B_i = (8\pi^2 k_B T / \gamma) [\Gamma^{-1}]_{ii}. \quad (8)$$

The GNM curves in Fig. 2 are calculated using the above equation and uniformly rescaling all the results to match the experimentally observed fluctuation range. In the ANM,  $\gamma^{-1}[\Gamma^{-1}]_{ii}$  is replaced by the trace of the  $3 \times 3$  diagonal block elements of  $\mathcal{H}^{-1}$ .

The agreement observed in Fig. 2 between theory and experiments is quite good in that no residue specificity and nonlinear effects are included in the theory, on the one hand, and, the experiment is subject to uncertainties, on the other. Note that the resolution of the examined crystal structure was relatively low (3 Å), and the ms fluctuations would usually be subject to even larger experimental uncertainties than the mean positions. Correlation coefficients of 0.78 (0.73) and 0.85 (0.70) are obtained for the respective correlations of the ANM and GNM results with experimental data for HA1 (HA2). The difference in the fluctuation behavior of the protein in the crystallized form may be associated with other effects such as intermolecular constraints in the tight crystal packing or static disorder.

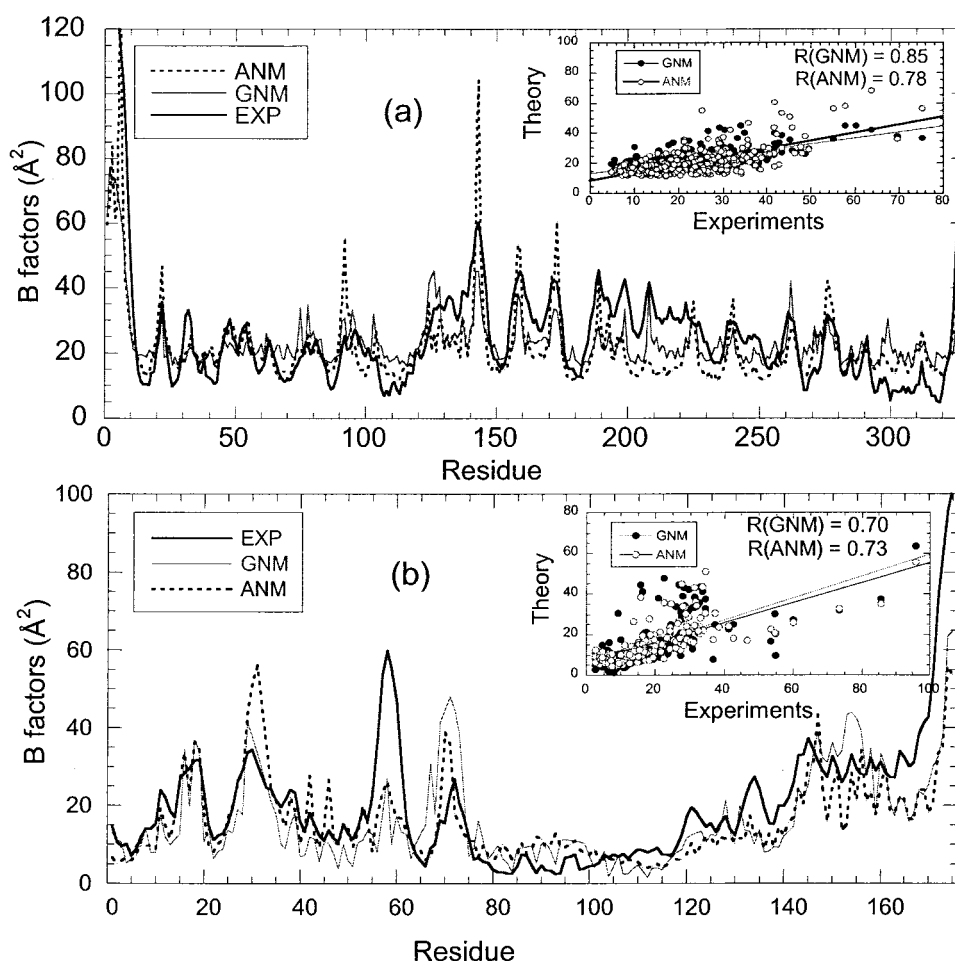


FIGURE 2 Comparison of theoretically calculated B-factors (curves labeled GNM and ANM) with those measured by x-ray crystallography (labeled EXP), for (a) HA1, and (b) HA2 residues in the neutral pH trimeric structure (Weis et al., 1990b). The insets display the correlation between theoretical and experimental data.

### Significance of low frequency collective motions

The low frequency modes, also called *global* modes, provide insights about the mechanisms of the cooperative molecular motions relevant to function. Our previous studies (Bahar and Jernigan, 1998, 1999; Bahar et al., 1999; Demirel et al., 1998; Bahar, 1999; Jernigan et al., 1999) indicate that the minima in the global mode shapes (i.e., in the distributions of residue fluctuations as driven by the lowest frequency modes) coincide with the hinge sites of the molecule, whereas the maxima usually correspond to substrate recognition sites. The hinge sites are thus fixed in space in the global modes, whereas substrate recognition sites sample a large space. The mechanical stability of the hinges is a requirement for them to serve as a swivel, a pivot, or a shaft, about which the collective motions are actuated, and the high flexibility of the recognition sites facilitates the recognition and optimal binding of substrates. The two lowest frequency modes of HA are interpreted

below with regard to these features, and with special attention to their relevance to the binding and fusogenic activities of the protein.

### First mode: global twisting around the longitudinal axis

Figure 3 *a* displays the shape of the first (slowest) global mode of motion, determined by the GNM analysis of BHA. The three-fold repeated pattern in the distribution of residue fluctuations reflects the identical behavior of the three monomers in this mode. In parts Fig. 3, *b* and *c*, the portions of the curve for the respective HA1 and HA2 chains are displayed as averages over the three monomers. The HA1 residues identified (Wiley et al., 1981; Weis et al., 1988) as recognition sites (Tyr98, Gly134–Ala138, Trp153, Thr155, His183, Ser186, Glu190, Leu194, Tyr195, Leu226, Ser228) are indicated by the open circles in Fig. 3 *b*. These residues

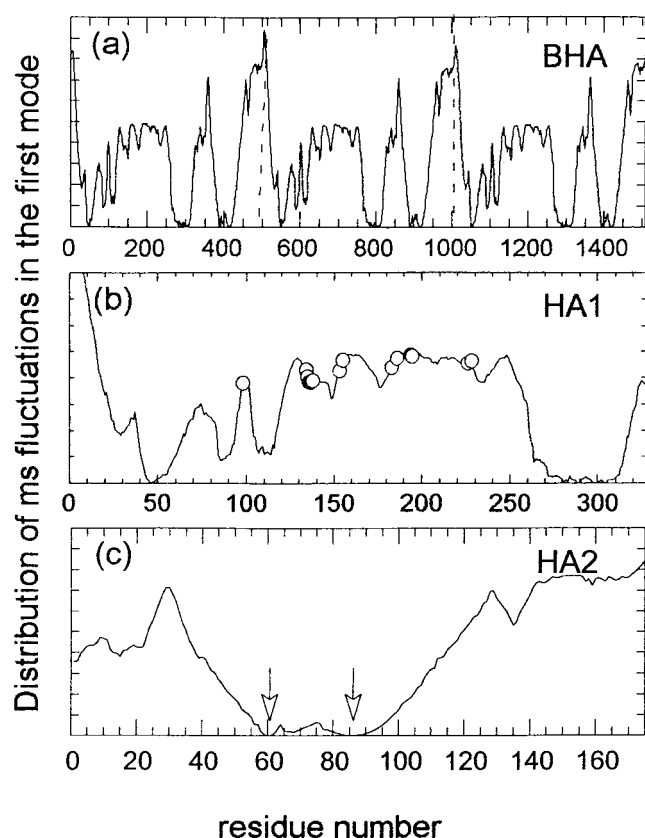


FIGURE 3 Residue mean-square fluctuations driven by the first (global) mode of the HA trimer in BHA. Minima usually indicate hinge sites controlling the collective fluctuations, and maxima are recognition sites. Results are displayed for (a) the trimer, (b) the HA1 chain, and (c) the HA2 chain. The curve in panel *a* shows that the three monomers exhibit the same mechanism of motion in the global mode. The HA1 residues recognizing (and binding to) the host cell receptors are shown by the open circles in panel *b*. There is a broad minimum in panel *c*, indicative of the central hinge site on the stem-like domain of the HA2 monomer. All minima coincide with the region colored green in Fig. 4 *a*.

are located in the conserved pocket near the distal end of the globular domains (green in Fig. 1 *a*). All of these residues occupy maxima positions in the global mode curve. Their enhanced mobility in the global modes conforms to the general behavior observed for recognition sites in other substrate-binding proteins (Keskin et al., 2000). On the other hand, two lowermost minima (residues 43–55 and 272–311) are in close agreement with the residues 38–64 and 266–302 pointed out by White and Wilson (1987) to form the hinge region of HA. See also the Fig. 3.2 *A* in the review of White et al. (1997).

As to the global mode shape of HA2 subunits (Fig. 3 *c*), a broad minimum is observed for BHA, with the lowermost points centered around Glu61 and Glu85, as indicated by the arrows. This region includes a large part of the loop (residues 56–75) that is reconfigured as helix B at low pH, and the N-terminus of helix C. We note that the mutation E81G

breaks a salt link that otherwise stabilizes the overall trimeric form (Daniels et al., 1985). The residues at the minima in the global mode shape participate in the hinge region shown in green in Fig. 4 *a*. The same region also includes the HA1 hinge residues 43–55 and 272–311 identified above. The most severely constrained residues in mode 1 are thus clustered together in a well-defined hinge region toward the upper part of the stem domain, although they are sequentially noncontiguous. The most flexible region of HA2, in contrast, is its C-terminal segment. This is the connection to the transmembrane domain of the glycoprotein. The high mobility of this region is functional for the translocation across viral membrane and opening of fusion pore.

Figure 4 displays the ribbon diagrams of HA, color-coded according to the amplitudes of motions in *a*, mode 1; *b*, modes 2 and 3; *c*, mode 4; and *d*, modes 9 and 10. The color code in the order of increasing mobility is green, cyan, blue, magenta, red, orange, and yellow.

Although Fig. 4 provides information on the size of motions driven by the global modes, the associated mechanisms of motion are as yet unidentified. The mechanism of mode 1 is identified using the ANM to be a global twisting of the three monomers around the cylindrical axis of symmetry. The results are shown in Fig. 5. The figure shows the (a) top, (b) bottom, and (c) side views of the trimer, both in the wild-type (*left*) and a distorted (*right*) conformation visited by the action of mode 1, when the latter is fully developed. One of the monomers is colored black to trace its conformational changes.

Before proceeding to the interpretation of Fig. 5, two points need to be clarified. First, the size of fluctuations depends (linearly and inversely) on the force constant  $\gamma$ ; and the present diagrams are obtained by adopting  $\frac{1}{2}$  of the  $\gamma$  value deduced (Fig. 2) from crystallographic B factors. Thus, the amplitudes of motions are magnified here by a factor of two compared to that predicted for crystals, which allows for an easier visualization, and also accounts for the usual higher flexibility of molecules in solution. Second, the displayed distorted (*right*) conformation represents one of the accessible fluctuating conformations visited by the action of mode 1; there exists a second symmetrically related conformation (not shown) driven by the same mode, and the protein fluctuates back and forth between these two alternative conformations.

Figure 5 *a* displays the motion of the globular heads. We observe a slight enlargement of the globular heads along the radial direction (i.e., away from the central axis of cylindrical symmetry), accompanied by an angular rotation of  $\sim \pm 30^\circ$  around the central axis. Remarkably, an even larger relaxation along the radial direction is driven at the base of the stem domain (i.e., near the transmembrane region), as the comparison of the left and right panels in Fig. 5 *b* reveals. More importantly, the base portions of the monomers are rotated by  $\sim 60^\circ$  about the cylindrical axis.



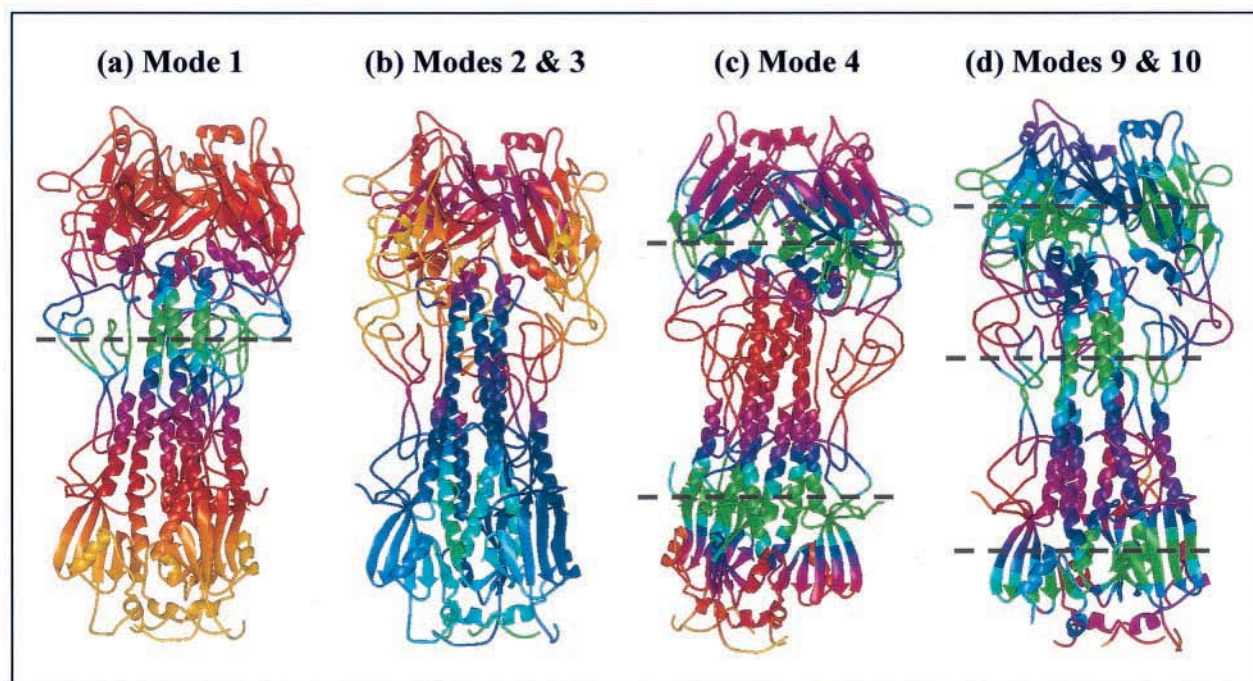


FIGURE 4 Ribbon diagrams of HA, color-coded according to the amplitudes of motions driven by low frequency collective modes. Panel *a* refers to the slowest mode (mode 1); (*b*) describes the joint effect of modes 2 and 3; (*c*) mode 4; and (*d*) joint effect of modes 9 and 10, elucidated by GNM. The color codes are green, cyan, blue, magenta, pink, and yellow, in the order of increasing mobility. Dashed lines represent the locus of hinge-bending sites. These are stationary regions in the specific modes illustrated. Panel *a* indicates a central hinge bending in the stem region of HA. Panel *b* shows that the HA1 globular domains enjoy large size fluctuations while the fibrous stem region remains almost intact. Panels *c* and *d* show the increase in the number of hinge sites, or the decrease in the size of the substructures that undergo coupled, coherent movements, as higher frequency modes are observed.

It is important to note that the expansion and counter rotations of the head and tail portions of the molecule leave almost unchanged the conformation of the central part of the coiled coil stem. See the region colored green in Fig. 4 *a*. The coordinates of residues in this region are almost unchanged. This region indeed serves as a hinge site for this global twisting of the trimer.

Notably, mode 1 can serve multiple purposes. First, there is a relaxation in the conformation of the globular heads. This mechanism can be useful for relieving the steric overlap that would resist the conformational rearrangement of the HA2 monomers at low pH. The rotational mobility of the lowermost part of the stem region, in contrast, could be a prerequisite for the efficient translocation or clustering of the glycoproteins on the viral membrane before fusion. The coupled torsional rotation of four clustered trimers would be operative in the fusion pore opening and dilation or in the fusing-membrane dimple formation through a biaxial deformation mechanism. These issues, of utmost importance for the biological function of the HA, will be further elaborated in the Discussion.

In the above computations, the glycosylated residues are also represented as single nodes located at their  $\alpha$ -carbon sites. Calculations repeated with the carbohydrates included as additional nodes in the elastic network yielded almost

indistinguishable mode shapes. The correlation coefficient between the ms fluctuations in the global modes of the ANM calculations with and without carbohydrates was above 99%. Because we could not detect any significant differences among the dominant modes of motion in the presence of carbohydrates, these results will not be elaborated further.

### Second dominant mode: global bending of the molecule

The second and third lowest frequency modes of HA are related by cylindrical symmetry. These can be described as the overall bending of the molecule. Basically, the globular heads of the HA1 subunits are bent, while the stemlike portion around the triple-stranded coiled coil remains practically unchanged. For a cylindrically symmetric structure, the radial displacement can be decomposed into two components ( $x$  and  $y$ , if the cylindrical axis is identified as the  $z$  axis), that complement each other. These two modes combine to yield a radial motion. The molecule can apparently bend in any given direction, should it be unoccupied by the neighboring integral membrane proteins.

Figure 4 *b* illustrates the flexibility of different regions of the molecule, as induced by the modes 2 and 3 of GNM.

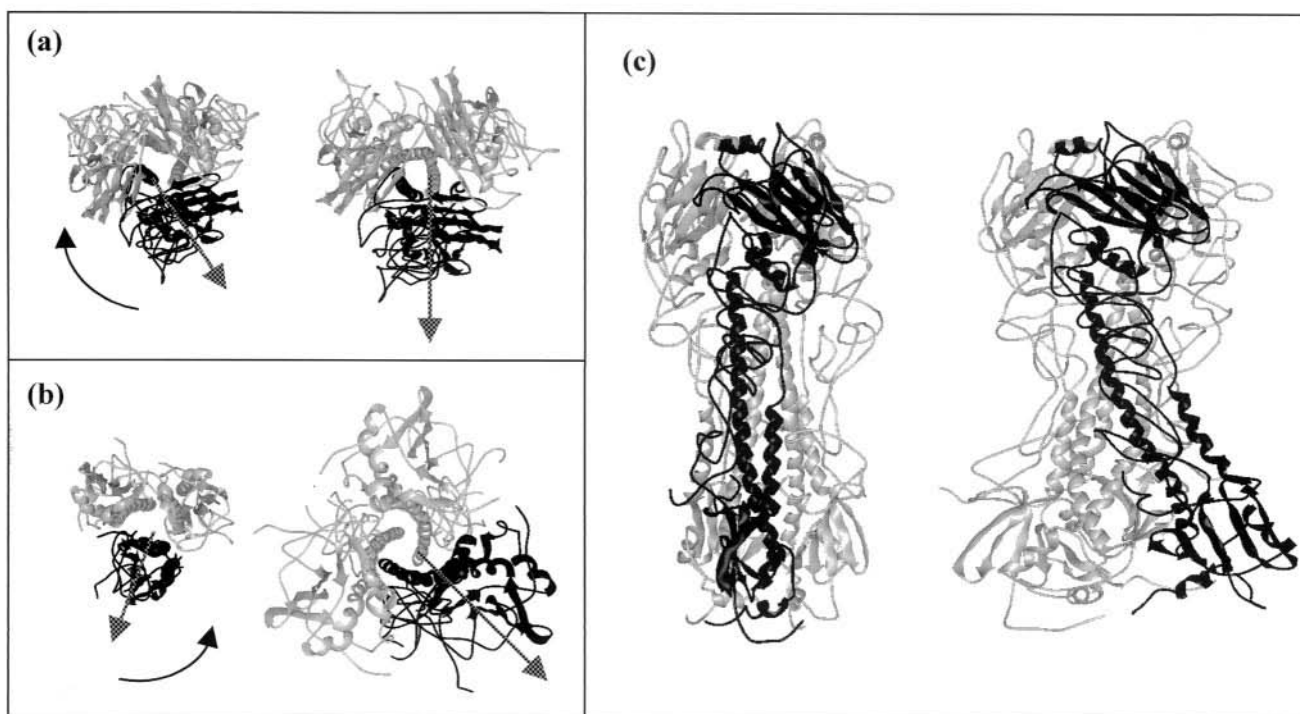


FIGURE 5 Mechanism of the first dominant mode of motion: global twist of the monomers about the cylindrical axis of symmetry. Panels *a*, *b*, and *c* display the top, bottom and side views, respectively, of the wild-type conformation (*left*) and its distorted form (*right*) visited by the action of mode 1. One monomer is shown in black in all diagrams to ease the comparison between the conformations. The top and bottom parts undergo opposite direction rotations. The apparent same direction rotations in (*a*) and (*b*) is a consequence of the opposite directions of observations (*top-down* and *bottom-up*) of the molecule in the two respective cases. This mode induces a partial relaxation in the globular heads, along with a significant angular and radial mobility at the lowermost regions, suggesting its possible implication in the displacement and clustering of HA molecules before membrane fusion and opening of the fusion pore.

The same color code as in Fig. 4 *a* holds. The HA1 molecules and, in particular, their globular heads enjoy a high mobility, whereas the HA2 stem is relatively stiff.

Figure 6 illustrates the molecular deformation induced by the action of mode 2, as predicted by the ANM. The conformational change induced by mode 3 is hardly distinguishable from that of mode 2, in view of cylindrical symmetry. Clearly, the molecule possesses a tendency to be bent along any radial direction by the joint action of these two global modes. As will be further discussed later, this bending flexibility is important for the association of the viral and endosomal membranes.

### Other collective motions

The number of hinge sites increases as higher frequency modes are considered, while the size of the structural elements undergoing en bloc movements decreases. Figure 4, *c* and *d*, illustrates the mobilities in the GNM mode 4, and the joint modes 9 and 10, respectively. We note that the region including the fusion peptide is active in the latter modes, suggesting a possible involvement in the solvent exposure of the fusion peptide before fusion.

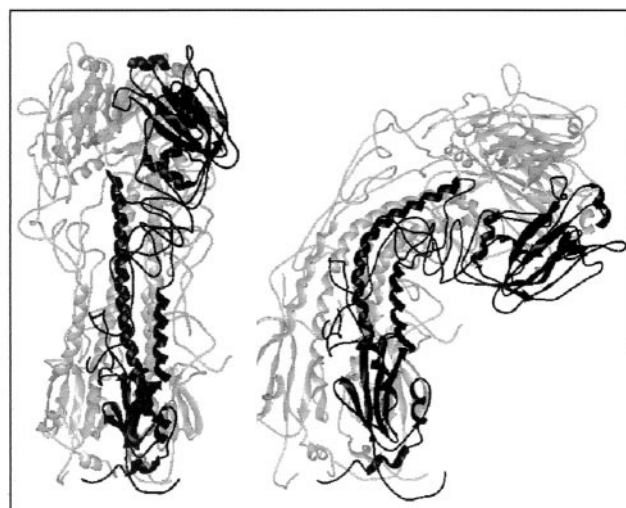


FIGURE 6 Mechanism of the second and third dominant modes of motion: bending of the globular heads while the stemlike region is almost rigidly affixed. Modes 2 and 3 are related by cylindrical symmetry. Their combination ensures the overall bending of the trimer along any radial direction.



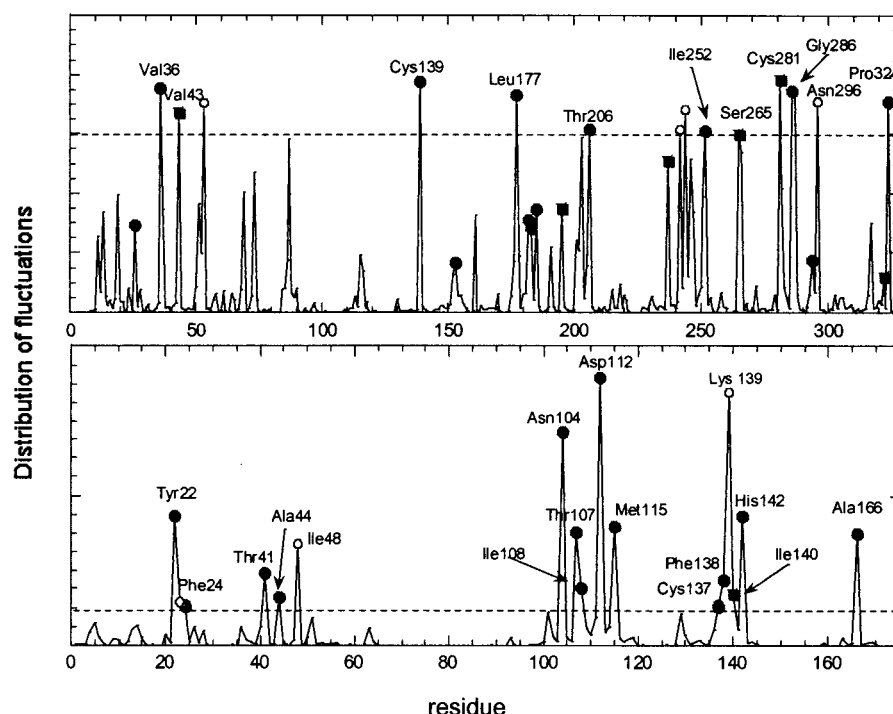


FIGURE 7 Distribution of fluctuations in the high frequency modes for (a) HA1 and (b) HA2 in the neutral pH trimeric form. The peaks refer to residues that play a critical role in the overall stability of the trimer. Residues exhibiting the strongest peaks (above the dashed lines) are labeled. All, except a few, indicated by the open circles, are found by independent multiple-sequence alignments to be highly conserved. The residues marked by the filled circles are fully or strongly conserved. Those indicated by filled squares are conserved in the HAs of type A, only.

Figure 4, *a*, *c*, and *d*, shows that the collective motions of the overall molecule in the slowest modes are reminiscent of the normal modes of a harmonic oscillator, which gradually exhibits smaller wavelength fluctuations and a larger number of nodes as the frequency of the operating mode increases. The hinge sites are marked in these ribbon diagrams by the dashed lines. The displacements of different parts of the molecule increase in size with increasing separation from the hinge plane, consistent with the longer moment arm of the more distant regions. Other collective modes (not shown), induce a shrinkage or expansion of the whole molecule along its longitudinal axis, or an overall S-shaped distortion around two central hinge planes perpendicular to the longitudinal axis.

### Kinetically hot residues

The predicted global modes are relatively insensitive to the details of the computational model and method. Their robustness has been confirmed in several theoretical studies (Kitao and Go, 1999). The high-frequency modes, in contrast, are sensitive to the detailed interactions at the atomic level. They usually contain white-noise contributions that need to be filtered out to extract physically meaningful information. Not surprisingly, these modes are referred to as “uninteresting modes” (Amadei et al., 1993). They usually drive isolated

fluctuations, as opposed to the correlated ones that underlie the intramolecular communication (Baysal et al., 1996).

The GNM results differ from those extracted from conventional simulations in that they are devoid of random noise effects. The high-frequency modes identified by the GNM are interesting: they indicate the most strongly constrained sites in the presence of the intricate coupling between all residues. The peaks emerging in these mode shapes have been referred to as “kinetically hot residues” in view of their high frequencies (Bahar et al., 1998; Demirel et al., 1998). These sites are usually implicated in the folding nuclei, or in the key tertiary contacts stabilizing the overall fold. As a consequence, they ought to be evolutionarily conserved among different members of a given family.

Our GNM examination of kinetically hot residues for HA yielded the distribution displayed in Fig. 7 for the ms fluctuations of residues. Panels *a* and *b* represent the weighted effect of the subset of 100 highest frequency modes, for the respective subunits HA1 and HA2 in BHA. The most pronounced peaks are labeled. First, we compared these peaks with the conserved residues identified by the multiple alignment of 78 sequences from the SWISSPROT and PDB (composed of 74 influenza HA molecules of type A and four of type B from different strains). The fully and strongly conserved residues were found using CLUSTALW algorithm (Thompson et al., 1994). Notably, out of the 14 residues

of HA1 distinguished in the top panel of Fig. 7 as the strongest peaks in the fast mode shapes (above the dotted line), 10 coincide with the fully or highly conserved residues detected by multiple alignments, the remaining four are shown by the open circles. In the case of HA2 monomers (lower panel in Fig. 7), all of the peaks, except Ile48 and Lys143, are highly conserved in the examined strains. Met115 is replaced by Val115 in two other strains (Japan/305/57 and PR/8/34) commonly used to study fusion.

Among the kinetically hot residues of the HA2 subunits, the segment Asn104–Met115 emerges as a critically important region. Interestingly, this segment coincides with the specific region (residues 108–115) that has been found to exhibit fusion activity by Shin and coworkers. (Kim et al., 1998). The sharpest peak occurs at Asp112 (Fig. 7*b*). Asp112 makes four hydrogen bonds that stabilize the fusion peptide at the N-terminus of HA2. Mutations at Asp112 give rise to fusion under less acidic conditions, due to the weakening of intramolecular interactions (Weis et al., 1990a, Daniels et al., 1985). Try22 also plays a key role in maintaining the stability of the core region near the fusion peptide, making a network of tertiary contacts with Ala44 and Met115 in the same chain. Finally, we notice that Pro324 in HA1 is also involved in close interactions with the fusion peptide, especially with the residues Asn12 and Glu15. Thus, a cluster of residues coordinating the fusion peptide is found here as engaged in high-frequency fluctuations, producing a local accumulation of energy in the neighborhood of the fusion peptide (Fig. 8). This stored energy is presumably relieved by the conformational rearrangements occurring at low pH. Not surprisingly, the initial rearrangement at the onset of fusion is proposed to be the exposure of the fusion peptide to the environment (White and Wilson, 1987). The stored high energy near the fusion peptide is certainly sufficient to drive such a local conformational rearrangement.

A recent study showed that Glu15 undergoes a conformational change at low pH (Han et al., 2001), which exposes this residue (and Asp19) to the membrane-solution interface at low pH, while the same residues insert into the lipid bilayer at neutral pH. Thus, in contrast to expectations, fusion activity at low pH is not necessarily related with a deeper insertion of acidic residues, when protonated, into the lipid membrane (Cohen and Melikyan, 2001). A more cooperative mechanism, involving a tight network of tertiary contacts (Fig. 8), triggers the HA-mediated fusion according to the present analysis; and it remains to be seen whether mutations at the hot spots displayed in Fig. 8 can indeed impair the fusion activity of HA.

### Efficiency of the analytical methods

ANM and GNM calculations require the inversion of a square matrix describing the topology of inter-residue contacts,  $\mathcal{H}$  in the ANM and  $\Gamma$  in the GNM. The respective

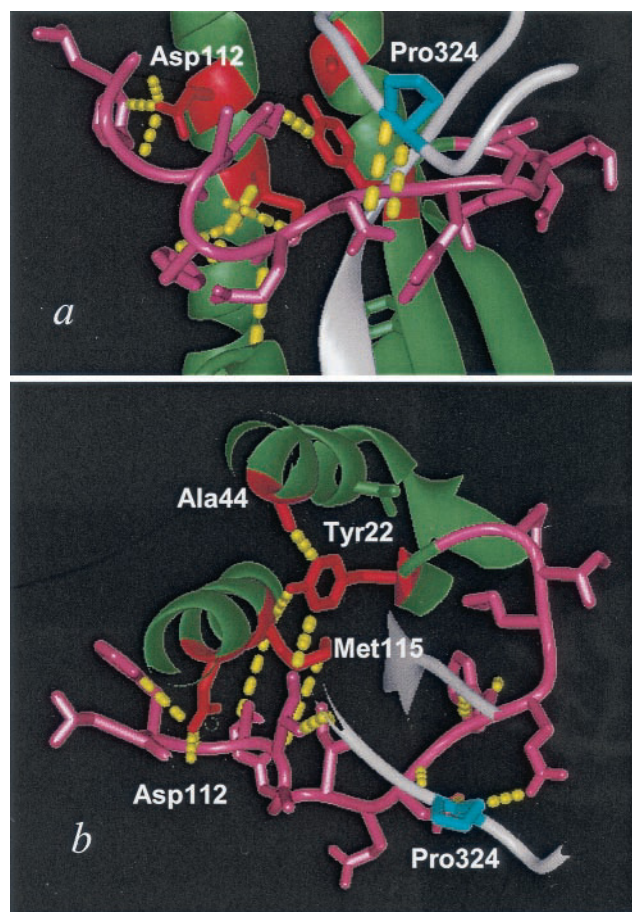


FIGURE 8 Kinetically hot residues near the fusion peptide. Ribbon diagrams of (*a*) side view of the close neighborhood of the fusion peptide (HA2 residues 1–20, pink) in one subunit of HA, and (*b*) cross-sectional top view of the same region. The HA2 and HA1 monomers are shown in green and gray, respectively. The HA2 residues (Asp112, Met115, Tyr22, and Ala44) distinguished in the fast modes (Fig. 7), which wrap across the fusion peptide, are colored red; their side chains are explicitly displayed. Pro324 of HA1 is colored blue. Dashed lines indicate interaction with the fusion peptide, including: Asp112 with Gly1, Phe3, and Gly4; Met115 with Ala5 and Phe9 and with Tyr22; Tyr22 with Ile6 and Ala44; and Pro324 with Glu15 and Asn12. Other interactions not shown are Asp112–Ala5, Asp112–Ile6, Met115–Ile6.

sizes of these matrices are  $N \times N$  and  $3N \times 3N$ , where  $N = 1509$ , in the case of trimeric HA. The inversion of these matrices is the most time-consuming part of the computations. The time cost on an R10,000 SGI O<sub>2</sub> workstation is approximately 20 min (real time) for the GNM, and about one order of magnitude longer for the ANM, the latter presenting the advantage of yielding the directions of fluctuations in addition to their size, and thereby elucidating the mechanism of collective motions. Both approaches are therefore extremely efficient compared to conventional MD simulations, the time cost of which would be of the order of weeks for a molecule of 1500 residues, using the same computational power.

## CONCLUSION

### Critical sites for function and stability

The elastic network model allows for the determination of a number of critically important sites. These are classified in two categories. The first group is composed of the residues that are important for coordinating the cooperative motions of the overall molecule. Their mutation can impede function. These are identified from the minima of the global mode shapes (Fig. 3). The hinge sites, or the residues that act as swivels, shafts, levers, etc. lie in this group. The second group, kinetically hot spots, consists of residues that experience an extremely strong coupling to their close neighbors. They usually occupy regions of a high packing density, and their mutations can have a strong destabilizing effect. They can act as folding nuclei. They are usually engaged in networks of tertiary contacts that underlie the stability of the structure. The sharp peaks in the highest frequency motions (Fig. 7) correspond to such residues (see also Fig. 8). Both groups of residues are expected to be evolutionarily conserved, the former for function requirements, and the latter for folding and stability.

### Functional implications of global modes

Multistep mechanisms have been proposed for the HA-mediated membrane fusion (White and Wilson, 1987; Stegmann et al., 1990; White et al., 1997; Bentz, 2000a,b). The first step is usually the binding of HA1 globular heads to the sialic acid-containing receptors on the target membrane. Usually, several HA molecules are recruited to the binding site to form a fusogenic aggregate while a subset of these undergo a conformational change required to initiate bilayer destabilization (Mittal and Bentz, 2001). These reconfigured HAs cooperatively form a fusion intermediate that associates the two bilayers. They are usually envisioned in a tilted conformation at this intermediate step, and the bilayers are associated by the lateral contacts of the bridging HA molecules. The association step is usually succeeded by the opening of a fusion pore via the collective action of a number of closely clustered HA molecules. The mobility of the transmembrane domain is thought to play a crucial role in membrane association and pore opening.

The dominant modes identified here can be related to functional mechanisms as follows. The first mode is a global twisting of the HA trimer around its longitudinal axis, accompanied by a high rotational mobility and expansion of the lowermost part of the stem region. There could be several important functional implications of this motion. For example, the kinetic studies of Danieli et al. (1996) have indicated that HA-mediated fusion is a cooperative process that requires the concerted action of three to four HA molecules. Recent analysis of kinetic data (Melikyan et al., 1995) by Bentz (2000a) suggested that aggregates of at least

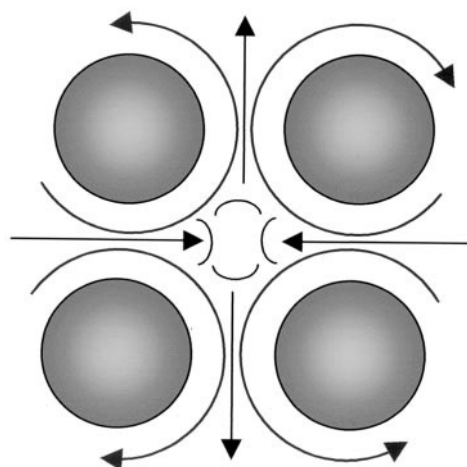


FIGURE 9 Schematic view of the cooperative global twisting of four clustered HA molecules about their longitudinal axes, to exert a biaxial deformation possibly facilitating a pore opening on the viral membrane.

eight HA molecules are necessary to form the fusion site, and only two or three need to undergo the essential conformational change for fusion. A subsequent stage in the fusion mechanism proposed by Bentz (2000b) is the formation of a lipidic stalk (e.g., Chenomordik et al., 1997) between the apposed membranes as a consequence of the essential conformational change. The global twisting motion of the HAs could play a crucial role in the formation of the stalk and especially in the opening of a fusion pore. Specifically, a coupled twisting of four HA molecules can greatly enhance pore opening and dilation by imposing a biaxial strain on the membrane, as schematically described in Fig. 9.

It is likely that the pore opening at the viral membrane is easier if the global torsional motion of the HAs is not restricted by the binding of the globular heads to the sialic acid-containing receptors on the target membrane. The recent studies showing that receptor binding delays the HA refolding into a fusion-competent form (Leikina et al., 2000), and that the fusogenic aggregate contains at least two unbound HAs (Mittal and Bentz, 2001) could be associated with the prevention of such a global motion that would otherwise initiate the fusion process.

The global twisting motion could also be functional in the mechanism proposed by White et al. (1997), in which the extended coiled-coil conformation is not a fusion-active intermediate in the opening of the fusion pore. Still, global twisting of intact HA molecules could facilitate the pore opening and dilation. It is worth noting that the global twisting presently observed is a molecular mechanism that becomes operative only in the trimeric form, and cannot be observed upon examination of single monomer's dynamics. Such a dominant mechanism specific to the particular architecture of the intact HA is likely to be involved in the specific biological function—fusion pore opening—of the



glycoprotein. Experimental studies can test whether the effective hindering of this mode, by mutating the residues acting as hinges, for example (minima in Fig. 3), would effectively reduce or impair the fusion activity of HA.

The second and third modes, in contrast, reveal the potential of the molecule to be bent at a position that has been suggested in previous experiments to be a hinge-bending site. This type of bending mechanism is essential for the close association of the apposed viral and target membranes, because it allows for a maximization of the contact surface. The tilting of the HA molecule during membrane association has indeed been proposed in the several studies. (Tatlian et al., 1995; White et al., 1997; Bentz, 2000b).

In summary, we find that a global torsional motion about the cylindrical axis of symmetry is the most favorable motion for the trimer from a mechanical point of view. This type of motion can be essential for the opening of a fusion pore via a biaxial strain of the viral membrane. A second mode, leading to the bending of the overall molecule, is found to complement this mode. Although this second mode has been suggested in several studies, it has been demonstrated for the first time to be one of the two dominant mechanisms of motion for the particular three-dimensional structure.

In addition to these global modes, a local reconfiguration of the fusion peptide on a significantly shorter time scale is favored because of the localization of high energy in its close neighborhood, as evidenced by the analysis of hot spots. This local reconfiguration can easily trigger the exposure of the fusion peptide to the environment.

The HA1 monomers play the important role of target-membrane recognition, and the GNM-predicted mobilities support this role. The picture that emerges is that first the target membrane is recognized by the several aggregating HA1 monomers. This brings into close proximity the viral and target membranes. The next step is the merging of the apposed membranes, which is facilitated by the fusion peptides exposed along the lateral surface of the cylindrical trimers bent by the action of mode 2. This mechanism would be followed by the slowest mode (mode 1, axial rotation) that drives the viral-membrane stretching and fusion-pore opening.

## Future work

It is now established from several studies that the analysis of collective dynamics provide information on the mechanisms of functional motions and the identity of sites actively engaged in controlling the function (Kitao and Go, 1999; Bahar et al., 1999; Doruker et al., 2000). This information can be exploited for devising control mechanisms to inhibit the protein function. A classical example is the docking of ligands to the hinge-bending site of HIV-1 protease, thus inhibiting the mechanism of action of this enzyme (Hodge et al., 1997). HA is likewise a good target for the development of novel antiviral agents. The influenza virus has two other integral membrane proteins,

M2 and neuraminidase. The only anti-influenza agents presently in clinical usage interfere with the structure and function (ion channel) of M2. There are no molecular therapies in use that block the binding of viruses (Air and Luo, 1997), but many efforts are underway. Insofar as blocking the activity of HA is concerned, polyvalent sialic-acid analogs have been tested as well as benzo- and hydroquinones that prevent the conformational change leading to fusion (Bodian et al., 1993). The former intervention has been limited by antigenic shifts and antigenic drifts, whereas the latter necessitates a clear assessment of the molecular mechanisms leading to fusion, and the identification of residues mediating these cooperative motions. The present results shed light into possible target sites whose mutation or complexation with small molecules could inhibit the HA function. A further extension would be to analyze the collective dynamics of other viral fusion proteins that also contain extended coiled-coil motifs to unravel the possible existence of common mechanisms of action of glycoproteins (Skehel and Wiley, 1998; Bentz, 2000b).

## REFERENCES

- Air, G. M., and M. Luo. 1997. Structure-guided therapeutic strategies. In *Structural Biology of Viruses*. W. Chiu, R. M. Burnett, and R. L. Garcea, editors. Oxford University Press, NY. 411–431.
- Amadei, A., A. B. M. Linssen, and H. J. C. Berendsen. 1993. Essential dynamics of proteins. *Proteins*. 17:412–425.
- Atilgan, A. R., S. R. Durell, R. L. Jernigan, M. C. Demirel, O. Keskin, and I. Bahar. 2001. Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophys. J.* 80:505–515.
- Bahar, I. 1999. Dynamics of proteins and biomolecular complexes: inferring functional motions from structure. *Rev. Chem. Eng.* 15:319–349.
- Bahar, I., A. R. Atilgan, and B. Erman. 1997. Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. *Fold. Des.* 2:173–181.
- Bahar, I., A. R. Atilgan, M. C. Demirel, and B. Erman. 1998. Vibrational dynamics of folded proteins: significance of slow and fast motions in relation to function and stability. *Phys. Rev. Lett.* 80:2733–2736.
- Bahar, I., B. Erman, R. L. Jernigan, A. R. Atilgan, and D. G. Covell. 1999. Collective motions in HIV-1 reverse transcriptase: examination of flexibility and enzyme function. *J. Mol. Biol.* 285:1023–1037.
- Bahar, I., and R. L. Jernigan. 1997. Inter-residue potentials in globular proteins and the dominance of highly specific hydrophilic interactions at close separation. *J. Mol. Biol.* 266:195–214.
- Bahar, I., and R. L. Jernigan. 1998. Vibrational dynamics of transfer RNAs. Comparison of the free and enzyme-bound forms. *J. Mol. Biol.* 281:871–884.
- Bahar, I., and R. L. Jernigan. 1999. Cooperative fluctuations and subunit communication in tryptophan synthase. *Biochemistry*. 38:3478–3490.
- Baysal, C., A. R. Atilgan, B. Erman, and I. Bahar. 1996. A molecular dynamics analysis of coupling between librational motions and isomeric jumps in chain molecules. *Macromolecules*. 29:2510–2514.
- Bentz, J. 2000a. Minimal aggregate size and minimal fusion unit for the first fusion pore of influenza hemagglutinin-mediated membrane fusion. *Biophys. J.* 78:227–245.
- Bentz, J. 2000b. Membrane fusion mediated by coiled coils: a hypothesis. *Biophys. J.* 78:886–900.
- Berman, H. M., J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shynalov, and P. E. Bourne. 2000. The protein data bank. *Nucleic Acids Res.* 28:235–242.
- Bizebard, T., B. Gigant, P. Rigolet, B. Rasmussen, O. Diat, P. Bösecke, S. A. Wharton, J. J. Skehel, and M. Knossow. 1995. Structure of

- influenza virus haemagglutinin complexed with a neutralizing antibody. *Nature*. 376:92–94.
- Bodian, D. L., R. B. Yamasaki, R. L. Buswell, J. F. Stearns, J. M. White, and I. D. Kuntz. 1993. Inhibition of the fusion-inducing conformational change of influenza hemagglutinin by benzoquinones and hydroquinones. *Biochemistry*. 32:2967–2978.
- Bullough, P. A., F. M. Hughson, J. J. Skehel, and D. C. Wiley. 1994. Structure of influenza hemagglutinin at the pH of membrane fusion. *Nature*. 371:37–43.
- Carr, C. M., and P. S. Kim. 1993. A spring-loaded mechanism for the conformational change of influenza hemagglutinin. *Cell*. 73:823–832.
- Carr, C. M., C. Chaudhry, and P. S. Kim. 1997. Influenza hemagglutinin is spring-loaded by a metastable native. *Proc. Natl. Acad. Sci. U.S.A.* 94:14306–14313.
- Chernomordik, L. V., E. Leikina, V. Frolov, P. Bronk, and J. Zimmerberg. 1997. An early stage of membrane fusion mediated by the low pH conformation of influenza hemagglutinin depends upon membrane lipids. *J. Cell Biol.* 136:81–93.
- Cohen, F. S., and G. B. Melikyan. 2001. Implications of a fusion peptide structure. *Nature Struct. Biol.* 8:653–655.
- Danieli, T., S. L. Pelletier, Y. I. Henis, and J. M. White. 1996. Membrane fusion mediated by the influenza virus hemagglutinin requires the concerted action of at least three hemagglutinin trimers. *J. Cell Biol.* 133:559–569.
- Daniels, R. S., J. C. Downie, A. J. Hay, M. Knossow, J. J. Skehel, M. L. Wang, and D. C. Wiley. 1985. Fusion mutants of the influenza virus hemagglutinin glycoprotein. *J. Cell Biol.* 105:431–439.
- Demirel, M. C., A. R. Atilgan, R. L. Jernigan, B. Erman, and I. Bahar. 1998. Identification of kinetically hot residues in proteins. *Protein Sci.* 7:2522–2532.
- Doruker, P., A. R. Atilgan, and I. Bahar. 2000. Dynamics of proteins predicted by molecular dynamics simulations and analytical approaches: application to  $\alpha$ -amylase inhibitor. *Proteins*. 40:512–524.
- Freire, E. 2000. Can allosteric regulation be predicted from structure? *Proc. Natl. Acad. Sci. U.S.A.* 97:11680–11682.
- Flory, P. J. 1976. Statistical thermodynamics of random networks. *Proc. R. Soc. London A*. 351:351–380.
- Gething, M. J., W. D. Robert, D. York, and J. White. 1986. Studies on the mechanism of membrane fusion: site-specific mutagenesis of the hemagglutinin of influenza virus. *J. Cell Biol.* 102:11–23.
- Godley, L., J. Pfeifer, D. Steinhauer, B. Ely, G. Shaw, R. Kaufmann, E. Suchanek, C. Pbao, J. J. Skehel, D. C. Wiley, and S. Wharton. 1992. Introduction of intersubunit disulphide bonds in the membrane distal region of the influenza hemagglutinin abolishes membrane fusion activity. *Cell*. 68:635–645.
- Haliloglu, T., I. Bahar, and B. Erman. 1997. Gaussian dynamics of folded proteins. *Phys. Rev. Lett.* 79:3090–3093.
- Han, X., J. H. Bushweller, D. S. Cafiso, and L. K. Tamm. 2001. Membrane structure and fusion-triggering conformational change of the fusion domain from influenza hemagglutinin. *Nature Struct. Biol.* 8:715–720.
- Harary, F. 1971. Graph Theory. Addison-Wesley. New York.
- Hinsen, K., and G. R. Kneller. 1999. A simplified force field for describing the vibrational protein dynamics over the whole frequency range. *J. Chem. Phys.* 111:10766–10769.
- Hinsen, K., A. Thomas, and M. J. Field. 1999. Analysis of domain motions in large proteins. *Proteins*. 34:369–382.
- Hodge, C. N., T. P. Straatsma, J. A. McCammon, and A. Wlodawer. 1997. Rational design of HIV protease inhibitors. In *Structural Biology of Viruses*. W. Chiu, R. M. Burnett, and R. L. Garcea, editors. Oxford University Press, NY. 451–473.
- Jernigan, R. L., M. C. Demirel, and I. Bahar. 1999. Relating structure to function through the dominant modes of motion of DNA topoisomerase II. *Int. J. Quant. Chem.* 75:301–312.
- Keskin, O., R. L. Jernigan, and I. Bahar. 2000. Proteins with similar architecture exhibit similar large-scale dynamic behavior. Specificity resides in local differences. *Biophys. J.* 78:2093–2106.
- Kitao, A., and N. Go. 1999. Investigating protein dynamics in collective coordinate space. *Curr. Opin. Struct. Biol.* 9:164–169.
- Kim, C. H., J. C. Macosko, and Y. K. Shin. 1998. The mechanism for clustering of phospholipid vesicles carrying the HA2 ectodomain of influenza hemagglutinin. *Biochemistry*. 37:137–144.
- Leikina, E., I. Markovic, L. V. Chernomordik, and M. M. Kozlov. 2000. Delay of influenza hemagglutinin refolding into a fusion-competent conformation by receptor binding: a hypothesis. *Biophys. J.* 79:1415–1427.
- Melikyan, G. B., W. Niles, and F. S. Cohen. 1995. The fusion kinetics of influenza hemagglutinin expressing cells to planar bilayer membranes is affected by HA surface density and host cell surface. *J. Gen. Physiol.* 106:783–802.
- Mittal, A., and J. Bentz. 2001. Comprehensive kinetic analysis of influenza hemagglutinin-mediated membrane fusion: role of sialate binding. *Biophys. J.* 81:1521–1535.
- Miyazawa, S., and R. L. Jernigan. 1985. Energy functions that discriminate x-ray and near-native folds from well-constructed decoys. *Macromolecules*. 18:534–552.
- Miyazawa, S., and R. L. Jernigan. 1996. Residue-residue potentials with an unfavourable contact high packing density term, for simulation and threading. *J. Mol. Biol.* 256:623–644.
- Qiao, H., S. L. Pelletier, L. Hoffman, J. Hacker, R. T. Armstrong, and J. M. White. 1998. Specific single or double proline substitutions in the “spring-loaded” coiled-coil region of the influenza hemagglutinin impair or abolish membrane fusion activity. *J. Cell Biol.* 141:1335–1347.
- Shangguan, T., D. P. Siegel, J. D. Lear, P. H. Axelsen, D. Alford, and J. Bentz. 1998. Morphological changes and fusogenic activity of influenza virus hemagglutinin. *Biophys. J.* 74:54–62.
- Skehel, J. J., and D. C. Wiley. 1998. Coiled coils in both intracellular vesicle and viral membrane fusion. *Cell*. 95:871–874.
- Stegmann, T., J. M. White, and A. Helenius. 1990. Intermediates in influenza induced membrane fusion. *EMBO J.* 9:4231–4241.
- Tatullian, S. A., P. Hinterdorfer, G. Barber, and L. K. Tamm. 1995. Influenza hemagglutinin assumes a tilted conformation during membrane fusion as determined by attenuated total reflection FTIR. *J. Cell Biol.* 106:5514–5523.
- Tirion, M. M. 1996. Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. *Phys. Rev. Lett.* 77:1905–1908.
- Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighing, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22:4673–4680.
- Weis, W., J. H. Brown, S. C. Cusack, J. C. Paulson, J. J. Skehel, and D. C. Wiley. 1988. Structure of the influenza virus haemagglutinin complexed with its receptor, sialic acid. *Nature*. 333:426–431.
- Weis, W. I., S. C. Cusack, J. H. Brown, R. S. Daniels, J. J. Skehel, and J. D. Watson. 1990a. The structure of a membrane fusion mutant of the influenza virus hemagglutinin. *EMBO J.* 9:17–24.
- Weis, W. I., A. T. Brünger, J. J. Skehel, and D. C. Wiley. 1990b. Refinement of the influenza virus hemagglutinin by simulated annealing. *J. Mol. Biol.* 212:737–761.
- White, J. M., L. R. Hoffman, J. H. Arevalo, and I. A. Wilson. 1997. Attachment and entry of influenza virus into host cells. Pivotal roles of hemagglutinin. In *Structural Biology of Viruses*. W. Chiu, R. M. Burnett, and R. L. Garcea, editors. Oxford University Press, NY. 80–104.
- White, J. M., and I. A. Wilson. 1987. Anti-peptide antibodies detect steps in a protein conformational change: low-pH activation of the influenza virus hemagglutinin. *J. Cell Biol.* 105:2887–2895.
- White, J. M. 1992. Membrane fusion. *Science*. 258:917–924.
- Wiley, D. C., I. A. Wilson, and J. J. Skehel. 1981. Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. *Nature*. 289:373–378.
- Wilson, I. A., J. J. Skehel, and D. C. Wiley. 1981. Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature*. 289:366–373.