

2011

# Immunoglobulin Structure Exhibits Control over CDR Motion

Michael T. Zimmermann

*Iowa State University*

Aris Skliros

*Iowa State University*

Andrzej Kloczkowski

*Iowa State University*

Robert L. Jernigan

*Iowa State University*, [jernigan@iastate.edu](mailto:jernigan@iastate.edu)

Follow this and additional works at: [http://lib.dr.iastate.edu/bbmb\\_ag\\_pubs](http://lib.dr.iastate.edu/bbmb_ag_pubs)

 Part of the [Biochemistry Commons](#), [Bioinformatics Commons](#), [Biophysics Commons](#), [Molecular Biology Commons](#), and the [Structural Biology Commons](#)

The complete bibliographic information for this item can be found at [http://lib.dr.iastate.edu/bbmb\\_ag\\_pubs/146](http://lib.dr.iastate.edu/bbmb_ag_pubs/146). For information on how to cite this item, please visit <http://lib.dr.iastate.edu/howtocite.html>.

---

This Article is brought to you for free and open access by the Biochemistry, Biophysics and Molecular Biology at Iowa State University Digital Repository. It has been accepted for inclusion in Biochemistry, Biophysics and Molecular Biology Publications by an authorized administrator of Iowa State University Digital Repository. For more information, please contact [digirep@iastate.edu](mailto:digirep@iastate.edu).

---

# Immunoglobulin Structure Exhibits Control over CDR Motion

## Abstract

Motions of the IgG structure are evaluated using normal mode analysis of an elastic network model to detect hinges, the dominance of low frequency modes, and the most important internal motions. One question we seek to answer is whether or not IgG hinge motions facilitate antigen binding. We also evaluate the protein crystal and packing effects on the experimental temperature factors and disorder predictions. We find that the effects of the protein environment on the crystallographic temperature factors may be misleading for evaluating specific functional motions of IgG. The extent of motion of the antigen binding domains is computed to show their large spatial sampling. We conclude that the IgG structure is specifically designed to facilitate large excursions of the antigen binding domains. Normal modes are shown as capable of computationally evaluating the hinge motions and the spatial sampling by the structure. The antigen binding loops and the major hinge appear to behave similarly to the rest of the structure when we consider the dominance of the low frequency modes and the extent of internal motion. The full IgG structure has a lower spectral dimension than individual Fab domains, pointing to more efficient information transfer through the antibody than through each domain. This supports the claim that the IgG structure is specifically constructed to facilitate antigen binding by coupling motion of the antigen binding loops with the large scale hinge motions.

## Disciplines

Biochemistry | Bioinformatics | Biophysics | Molecular Biology | Structural Biology

## Comments

This article is published as Zimmermann MT, Skliros A, Kloczkowski A, Jernigan RL. Immunoglobulin structure exhibits control over CDR motion. *Immunome Res* 2011; 7:5 Posted with permission.

## Creative Commons License



This work is licensed under a [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/).



RESEARCH

Open Access

# Immunoglobulin Structure Exhibits Control over CDR Motion

Michael T. Zimmermann<sup>1,2,3</sup>, Aris Skliros<sup>1</sup>, Andrzej Kloczkowski<sup>1,2,4,5</sup>, and Robert L. Jernigan<sup>1,2,3†</sup>

## Abstract

Motions of the IgG structure are evaluated using normal mode analysis of an elastic network model to detect hinges, the dominance of low frequency modes, and the most important internal motions. One question we seek to answer is whether or not IgG hinge motions facilitate antigen binding. We also evaluate the protein crystal and packing effects on the experimental temperature factors and disorder predictions. We find that the effects of the protein environment on the crystallographic temperature factors may be misleading for evaluating specific functional motions of IgG. The extent of motion of the antigen binding domains is computed to show their large spatial sampling. We conclude that the IgG structure is specifically designed to facilitate large excursions of the antigen binding domains. Normal modes are shown as capable of computationally evaluating the hinge motions and the spatial sampling by the structure. The antigen binding loops and the major hinge appear to behave similarly to the rest of the structure when we consider the dominance of the low frequency modes and the extent of internal motion. The full IgG structure has a lower spectral dimension than individual  $F_{ab}$  domains, pointing to more efficient information transfer through the antibody than through each domain. This supports the claim that the IgG structure is specifically constructed to facilitate antigen binding by coupling motion of the antigen binding loops with the large scale hinge motions.

## Background

Immunoglobulin Gamma (IgG) is one of the principal players in the adaptive immune system and is commonly referred to as an antibody. It is produced in a huge array of diverse antigen binding forms by B-cells through a combinatorial process called V(D)J recombination. This produces molecules comprised of two heavy and two light chains with highly variable complementary determining regions CDRs. The heavy chains of a given antibody are identical in sequence to one another and are comprised of four immunoglobulin folds (a two layer sandwich of 7 anti-parallel beta strands), one of which is variable in sequence. Light chains are also identical in sequence and consist of two immunoglobulin folds, one of which is variable. The two light chain and the first two heavy chain (including the variable) immunoglobulin domains come together to form an antigen binding domain (one for each pair). The remaining four (two per chain) heavy chain immunoglobulin folds interact with each other to form a third domain that is often referred to as the constant region. IgG has been studied by proteolysis which cuts the structure into the three described fragments. For this reason the two types of domains which comprise the IgG are referred to as the antigen binding fragments, or  $F_{ab}$ , and the constant fragment, or  $F_c$ . The region connecting the  $F_{ab}$

domains to the  $F_c$  is known to be a highly flexible hinge. This hinge region has 2-4 disulfide bonds bridging the heavy chains. Previously it has been excised from IgG and used in protein design as a molecular linker. This sequence has been extensively characterized and even synthesized [1,2]. While the primary characteristic of this region is its hinge flexibility, which was an impediment to resolving the structure early on, it has also been shown to have somewhat unique hydrophobic binding that allows it to be selectively bound to a stationary membrane so that either the  $F_{ab}$  or  $F_c$  fragments can be cut off with different proteases and recovered [3].

Here we compute a hinge map of IgG using Elastic Network Models (ENMs), show the extensive spatial freedom of the unrestrained  $F_{ab}$  domains that presumably facilitates binding, analyze the internal changes of the structure and how they affect the hypervariable CDRs, and apply a recently derived normal mode based kinematic simulation to generate motions of the structure. Motions of the structure are analyzed and a description of the high mobilities of the CDRs is provided.

## Materials and methods

### Elastic Network Model

To study the kinematics of IgG we use NMA (normal mode analysis) within the context of ENMs. Biological structures are often represented by  $C^\alpha$  atoms connected with harmonic springs. This represents the protein structure as an elastic network. The Gaussian Network Model (GNM), the earliest and one of the simplest of the elastic network models, is used to compute the relative

<sup>1</sup>L. H. Baker Center for Bioinformatics and Biological Statistics

<sup>2</sup>Department of Biochemistry, Biophysics and Molecular Biology

<sup>3</sup>Bioinformatics and Computational Biology

Iowa State University, Ames, IA 50011, USA

<sup>4</sup>Battelle Center for Mathematical Medicine, The Research Institute at Nationwide Children's Hospital

<sup>5</sup>Department of Pediatrics

The Ohio State University College of Medicine, Columbus, OH 43205, USA

†jernigan@iastate.edu



magnitudes of motion. It was originally proposed by Bahar, Haliloglu and Erman for coarse-grained models in 1997 [4,5], who applied the assumption postulated by Tirion [6] for atoms that both bonded and non-bonded contacts in proteins can be represented by a single universal spring. The Anisotropic Network Model (ANM) proposed in [7], can be used to compute the directions of motions of all points in the structure. We employ the ANM model throughout the present analyses. To generate an ANM model we first construct a Laplacian (or Kirchhoff) matrix using Equation 1 where  $r_c$  is a cutoff radius (typically 10-13Å),  $d_{ij}$  is the distance between atoms  $i$  and  $j$ , and  $\gamma$  is the spring constant. The potential energy of such a system is given by Equation 2. We then compute a matrix of second derivatives of the potential energy (see [7] for details), the eigenvectors ( $Q_i$ ) of which are called normal mode shapes, and the eigenvalues ( $\omega_i$ ) are the corresponding squared frequency. For a given normal mode we can then compute fluctuations of the structure with Equation 3.

$$\Gamma = \begin{cases} -\gamma & d_{ij} \leq r_c \\ 0 & d_{ij} > r_c \end{cases} \quad (1) \quad V = \frac{\gamma}{2} \Delta R^T \Gamma \Delta R \quad (2)$$

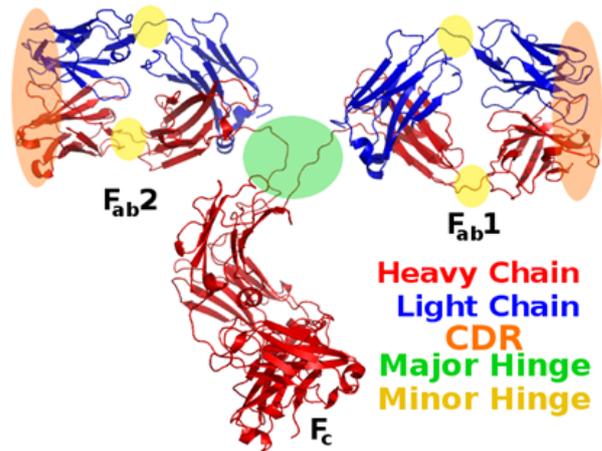
$$-\sum_{k=1, k \neq i}^N \Gamma_{ik} \quad i = j \quad \Delta R_i = Q_i \cos(\omega_i t) \quad (3)$$

Extensive applications of NMA to biological and chemical systems have been discussed in Cui and Bahar [8], Jernigan and Kloczkowski [9], and Sen et al. [10]. Successes with these methods make it clear that functionally important motions of biomolecules are usually governed by packing density. These and many other studies have enabled computations of the important motions on time scales beyond the usual reach of atomic molecular dynamics (MD). ENMs can be generated for small and medium sized proteins in seconds or minutes; a huge gain in comparison to the extremely long computational times required for corresponding MD studies. Work by Bakan and Bahar suggests that ANM may even sample conformation space more thoroughly than classic MD [33]. It has been demonstrated that extremely large molecular assemblages can be even further coarse-grained without loss of the major important motions [11]. More detailed analyses are available by use of elastic models that employ mixed-resolution models, where most of the structure is coarse-grained but with the regions of special interest remaining in atomic detail.

The dominance of the low frequency normal modes is universal, and usually there are only a few of these characteristic motions that are truly important. Here we represent the mean-square fluctuations by using the lowest frequency non-zero modes.

#### Kinematics of Proteins

Our method for solving the kinematics of coarse-grained protein structures is based on the Lagrangian equation for the potential and kinetic energy of the system, as described by Chirikjian and coworkers [12-16]. First, a rigid body translation and rotation of the structure is performed to place the origin of the coordinate system at the center of mass and so that the moment of inertia tensor is



**Figure 1:** The structure of IgG (PDB structure 1IGT). Heavy and light chains are distinguished from one another by colors. The  $F_{ab}$  domains have orange ellipses indicating the locations of the hypervariable loops (the CDRs) and green and yellow circles identifying the major and minor hinges.  $F_{ab}$  domains consist of one light chain and half of a heavy chain and are connected to the  $F_c$  and each other by the major hinge.

diagonal. The potential energy of the system of  $N$  points can then be written as in Equation 1. Note that  $\Gamma$  is the  $3N$  dimension square stiffness matrix of the system.

The displacement vector of the system,  $\Delta R(t)$ , could be calculated as

$$\Delta R(t) = \sum_{i=1}^{3N} \begin{bmatrix} \frac{1}{\sqrt{\omega_i}} \sin(t\sqrt{\omega_i}) Q_i Q_i^T \Delta \dot{R}(0) + \\ \cos(t\sqrt{\omega_i}) Q_i Q_i^T \Delta R(0) \end{bmatrix} \quad (4)$$

This facilitates performing time-dependent kinematic simulations with the ENMs using any desired combination of normal modes fixed by the index  $i$  and by choosing appropriate phase angles to describe the displacements between the phases of the different normal modes.

#### Computing Changes in Internal Distances

We also consider the displacements of the positions of points in the structure with ANM. The mean-square change in internal distance (**MSID**) is computed as

$$\langle (\Delta R_i - \Delta R_j)^2 \rangle = \langle \Delta R_i^2 \rangle + \langle \Delta R_j^2 \rangle - 2 \langle \Delta R_i \cdot \Delta R_j \rangle \quad (5)$$

These values are obtained directly from the inverse of the Hessian matrix,  $\Gamma^{-1}$ , from which the normal modes are derived:

$$\langle (\Delta R_i - \Delta R_j)^2 \rangle = (3k_B T / \gamma) * [\Gamma_{ii}^{-1} + \Gamma_{jj}^{-1} - 2\Gamma_{ij}^{-1}] \quad (6)$$

where  $k_B$  is the Boltzmann constant,  $T$  absolute temperature, and  $\gamma$  the ANM spring constant. We can also consider the normalized change in internal distances. This metric can be used to compare the magnitude of internal distance changes.

$$\langle (\Delta R_i - \Delta R_j)^2 \rangle' = \frac{\langle (\Delta R_i - \Delta R_j)^2 \rangle}{\sqrt{\Delta R_i^2} \sqrt{\Delta R_j^2}} \quad (7)$$

### Fractal and Spectral Dimension

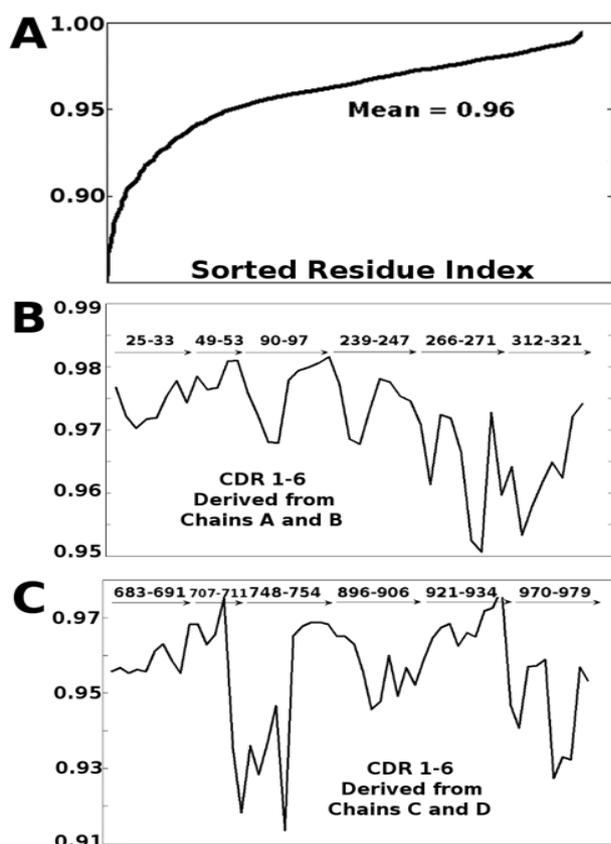
As early as 1980 the fractal dimension of myoglobin was studied [17], and it was found to be about 1.65. Experimental analysis of the spectral dimension of lysozyme was recently performed [18]. This study revealed not only that proteins may exhibit a mix of phonon (exhibiting discrete vibrational modes) and fractal character but also that the spectral dimension is relatively low and shows only moderate sensitivity to temperature. This finding provides an explanation for the efficient information transfer through protein structures. More recently, Granek and Klafter showed mathematically that certain fractal structures (and not uniform lattices) will experience the type of autocorrelation decay that is observed in protein experiments [19]. The compactness of protein structures is represented by a fractal dimension between 2.3 and 2.7 (see Enright and Leitner [20]). Investigation of the spectral dimension of elastic networks and explaining its relation to real structures has also been carried out [21]. Spectral and fractal dimensions were related to one another in recent papers by Reuveni and colleagues [22,23]. An equation was proposed that relates the two dimension metrics that fits well with the 5794 surveyed protein structures [24]. Here, we utilize the methods described in Ref. [22] for calculating the spectral and fractal dimensions of IgG and relate these findings to the CDR motions.

The fractal dimension describes how the mass captured within concentric spheres scales with the radius of these spheres. It is calculated here by finding the ten points closest to the proteins center of mass. Concentric spheres with incremental radii of 1Å are constructed and the total mass captured within each is recorded. Linear regression is performed ten times, once for each of the points closest to the protein's center. The average slope of the log-log plot of sphere radius versus mass captured is taken as the fractal dimension. The spectral dimension describes how the frequencies of vibration for the structure scale with the density of modes. That is, one performs a linear regression against the log-log plot of frequency versus the cumulative number of modes at each frequency. The spectral dimension is then the slope of this regression.

### Results and Discussion

We seek quantification of the motions of IgG in its dominant normal modes, which correspond to the flexing about the major hinge, particularly to see how this affects the spatial freedom of the CDR, both overall and internally.

Figure 2 shows the impact of the six slowest normal modes on the motion of the IgG and points out the CDRs. We see that these six normal modes account for nearly all of the motion of the Immunoglobulin, above 85% of the total motion for all residues and greater than 90% for the majority. This means that residues of the immunoglobulin move in a highly coordinated motion and that the loops do not act as in polymers, to randomly sample their dihedral angles. A large body of evidence shows that the ENM generates low modes that correspond with known biochemical functions of proteins. This gives us confidence



**Figure 2:** The first six modes of motion capture most of the total motion. Mean correlations of the motion derived from the first six normal modes with the total motion for (A) the entire structure, (B) the six CDR loops from chains A and B, (C) the six CDR loops from chains C and D. 11GT has 1316 residues in total. The mean correlation over all residues is 0.96 showing that the slowest modes strongly dominate the intrinsic motions of the structure. Behaviors for each of the individual CDRs are shown.

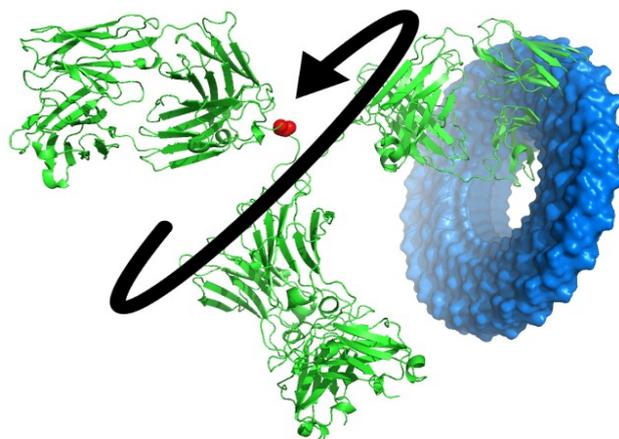
to conclude that correlations of motion within a low frequency mode are pertinent to the function of the IgG.

We have also computed correlations between mean-square fluctuations calculated using only the six lowest frequency modes and using all modes, for all residues, and the CDR (see Figure 2). Similarly, we compare correlations computed by using all normal modes in equation 4 with those obtained by using only the slowest modes (see details in the Methods section).

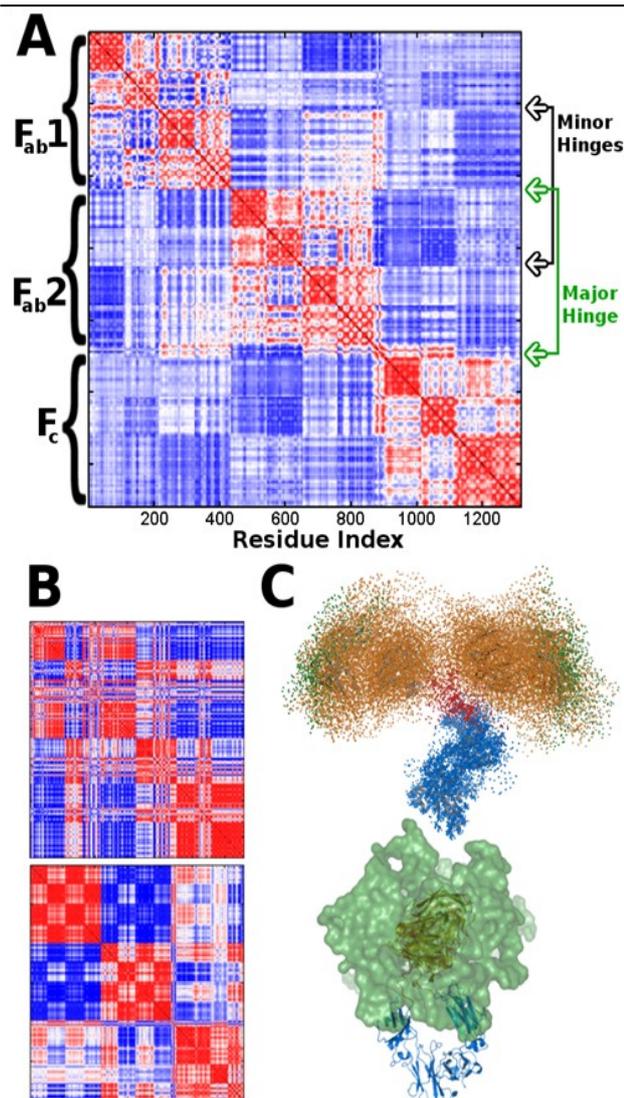
Figure 2 shows that the mean impact of the first 6 normal modes on the total motion is about 96% and that the lowest frequency modes do provide an excellent representation of the overall motions of the system. From a visual inspection we see the low frequency modes are associated with domain motions, a behavior that is usually seen in multidomain structures. For this reason normal modes have been used to identify hinges within structures [25]. We perform similar computations to confirm the presence of the hinges within IgG (see Figure 3). In order to determine the extent of CDR sampling we generate conformers using the normal modes. The magnitude of deformation in each mode is set by choosing the largest deviation that does not substantially deform the sequential virtual bond lengths. The lowest frequency modes corresponding to the global motions are collective in nature and exhibit comparatively low virtual bond stretching. We find that

overall the three domains are anti-correlated with one another. Figure 3 part C displays a representation of CDR sampling after following normal modes. It is apparent that the structure of IgG is designed to span the maximal space for the CDRs, presumably to aid binding. A similar approach for understanding the spatial freedom that the structure can sample would be to alter dihedral angles; Figure 4 illustrates this mapping approach, which is only preliminary. Other considerations would be required for more realistic dihedral sampling (steric hindrances, backbone constraints, energetic and inertial effects, etc.). Normal mode calculations are often performed to elucidate which residues or atoms in a molecular structure are the most mobile. Mobile active site residues may play roles in binding or substrate selection, whereas rigid regions are more likely to play key stabilization roles in the structure as a whole, as in a scaffold. An important exception to this occurs for the catalytic residues within an active site cleft that are relatively rigid.

Another quantity that is informative about internal conformational changes is the mean-square internal distance (MSID) changes,  $\langle (\Delta R_i - \Delta R_j)^2 \rangle$ , given by equation 5. MSID changes can be calculated directly from the Hessian matrix that is used to generate normal modes in ANM with equation 6. This quantity describes the changes



**Figure 4:** The sampling of hinge forms on the position of the CDR. In this figure the red amino acid's psi-angle at the major hinge is varied in 15° increments and the resulting position of the CDR on the right side is accumulated. The collection of all of these CDR coordinates is shown as the blue volume similar to Figure 3(C). This visualization could be useful for IgG hinge analysis, but would require inclusion of the limitations imposed by torsion angle availability (Ramachandran space) and steric clashes.

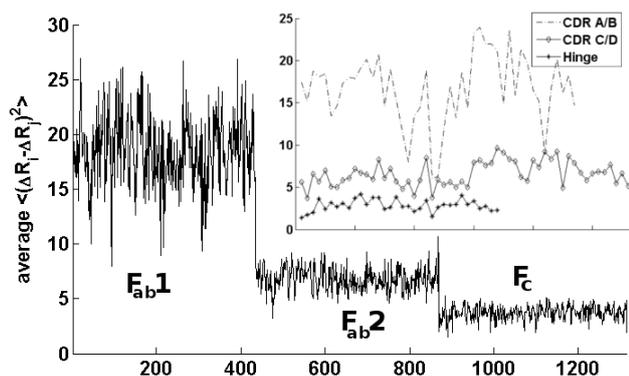


within a structure; how the normal modes stretch, compress, or otherwise alter the pair-wise distances between points in the structure. If this change in internal distance is zero for a given (i,j) pair, then the two points move together rigidly (the distance between them remaining unchanged). We have analyzed structures and seen that (data not shown) the areas of a protein with the smallest internal mean square distance changes are the cores of domains with these values increasing further away from stable cores. We have employed ANM models built with uniform springs with cutoffs ranging from 10-15 Å and with springs having inverse square dependences on distance. All of these yield similar results. Figure 5 shows this quantity averaged across all pairs of points within 7Å of one another. The CDR and major hinge are shown separately. We see that the CDR and hinge regions do not have significantly lower or higher average RMSID. We find that the F<sub>ab</sub> domains experience more internal motion than the F<sub>c</sub>, but that the two F<sub>ab</sub>s are not symmetric in their motions. This is likely due to the asymmetry in the initial structure. But, other feasible structures might be expected to actually behave in a symmetric way. F<sub>ab2</sub> is closer to the F<sub>c</sub> than F<sub>ab1</sub> and has more connections (higher stiffness) with it. Notably, we find that the internal distance changes at the hinge, as usual, are relatively small.

For many proteins, it is common to compare motions from the computations with the crystallographic temperature factors, the B-factors. The B-factors describe the uncertainty assigned to a given atom, usually by assuming it originates from relevant internal fluctuations. Rigid body contributions are often removed by the crystallographer,

**Figure 3:** Hinges and the motion behaviors identified by computations. (A) Correlation matrices (dot products between normalized pairwise displacement vectors). All values fall in the range [1,-1]. The average over the first 9 modes is displayed, with red corresponding to the motion of C<sup>α</sup> pairs positively correlated, blue for negative correlations, and white uncorrelated. From the block structure of the diagram and the changes in sign we can identify the hinge regions within the structure. (B) Similar to (A), but for individual modes; 1 (top) and 9 (bottom). (C) We use the 12 slowest modes to compute conformers of IgG because the majority of the motions in these modes are localized at the three prominent hinge regions. The F<sub>c</sub> is aligned in all conformers. Structure coloring shows the F<sub>c</sub> in blue, the major hinge in red, CDRs in green, and the remainder of the F<sub>ab</sub> in orange (in two perpendicular views).

but the successes of the TLS [26,27] and vGNM [28] methods provide strong evidence that B-factors often contain significant rigid body contributions. In the case of the 1IGT crystal we find that the B-factors may not be representative of the solution dynamics since the CDR of each  $F_{ab}$  is strongly bound to the  $F_c$  of another IgG (see Figure 6). While the experimental B-factors do highlight the major hinge as the most flexible part of the structure it is important to note that this is not because it is allowed to flex in the protein crystal, since the molecules are highly restrained by intermolecular interactions.



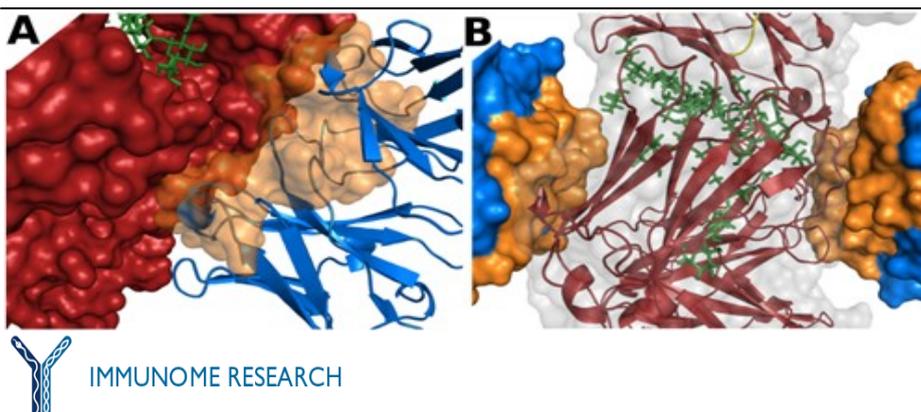
**Figure 5:** Mean-square internal distance changes within the entire immunoglobulin. (Inset) The CDRs from Fab1 are comprised of the variable loops in chains A and B, and the Fab2 CDRs are in chain C and D of 1IGT. We find that  $F_{ab}$  domains experience more internal motion than the  $F_c$ , but that these are not symmetric. The inset shows the same quantity but specifically for the two CDR regions and the major hinge.

Intrinsic disorder in proteins is a topic of growing popularity. Two disorder predictors were applied to the IgG structure, DisEMBL [29] and POODLE [30]. Interestingly, both methods predict the most mobile part of the structure, the CDR, to be the least likely to be disordered. Both methods have components in their scoring scheme that are knowledge-based; learned from scanning the PDB. We believe the CDR is predicted to be so stable because of the abundance of IgG structures and because the CDR is rarely unbound. Because the CDR is almost never free to move, it is always ordered in the known structures. DisEMBL predicts disorder while the POODLE prediction predicts 3 quantities; the secondary structure as coil, the residues un-resolved in a crystal, and the probability of residues having a high B-factor. Both methods employ a probability cutoff of 0.5; any residue above the threshold is deemed to be disordered. Interestingly, the CDR is predicted to be the least disordered part of the structure. It is possible, due to the ambiguity that remains in defining protein disorder and the complexity of crystal B-factors, that some knowledge based disorder

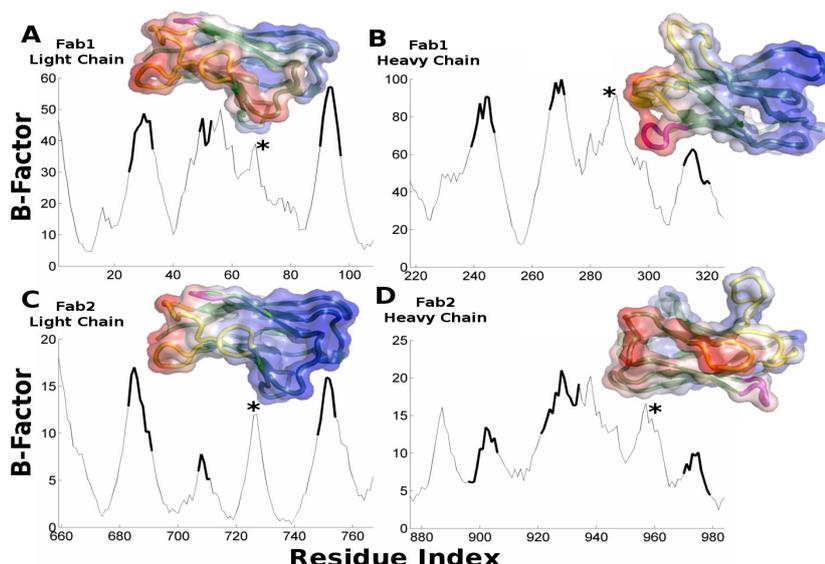
predictions may not be predicting exactly what one expects.

In Figure 7 we show the mean-square fluctuations of the IgG variable fragments computed with the ANM model. We find the parts of the structure that are most variable in sequence, the CDRs, are also the most mobile. It is interesting to note that ANM indicates 4 loops with high spatial mobility (and also the N-terminus), but there are only 3 CDR loops in the variable domain of each chain. The fourth loop is colored purple in the inset molecular images in Figure 7. Each  $F_{ab}$  domain then has two of these conserved mobile loops with one on either side of the CDR. It is interesting to note that the  $F_{ab}$ -like T-cell receptor (TCR) has the same spatial arrangement of loops, but the fourth loop found to be mobile here is also hypervariable in sequence [31]. While this loop does not usually bind antigen, it is involved in nonspecific antigen binding of TCRs.

To further investigate the motion patterns within the CDR and whether these may indicate that the IgG structure itself facilitates excursions of the hypervariable loops, we consider the difference in the mean square fluctuations, internal distance changes, as well as spectral dimensions and fractal dimensions of the full IgG structure and of the individual  $F_{ab}$  domains. GNM usually predicts the mean square fluctuations of each point more accurately (as judged with crystallographic B-factors) than ANM. Motions of the CDRs from the GNM are shown in figure 8. Curves from the full IgG structure and for the  $F_{ab}1$  domain have been scaled overall to match the range of the crystal B-factors. Including the the whole structure yields a correlation with experiment of 0.87 for the  $F_{ab}1$ , whereas utilizing only the structure of the  $F_{ab}1$  for the computations yields a correlation of 0.5. We have previously noted that the B-factors in this structure have the  $F_{ab}$  locked in a bound state. Interestingly, the computed CDR motion appears to be captured better by the full IgG model than by use of the Fab only. Internal distance changes, computed by using equation 5, indicate the extent of deformation between pairs of points in the structure. In figure 9, the effect of the full structure on the magnitude of change within and between hypervariable loops is summarized. The mean change within a (or between) loop(s) is plotted with error bars indicating one standard deviation. Before computing the extent of fluctuation, we rescale each mode to agree with equipartition wherein each internal degree of freedom would be assigned energy [32]. The full IgG structure amplifies internal distance changes within and between hypervariable loops, relative to a single  $F_{ab}$ . This is further indication



**Figure 6:** Intermolecular crystal packing in 1IGT. (A) One  $F_{ab}$  domain in 1IGT is shown in blue with its CDR as an orange surface. A symmetry related IgG is interacting with this CDR in the protein crystal. It is shown as a red surface and green sticks for the bound N-Acetyl-D-Glucosamine. (B) The  $F_c$  domain of 1IGT is shown as a red cartoon with gray transparent surface. Two intermolecular interacting  $F_{ab}$ s are shown in blue whose CDRs are highlighted in orange. These two views highlight crystal packing via CDRs contacting  $F_c$ .



**Figure 7:** The computed mean-square fluctuations in positions of the  $F_{ab}$  residues, with the CDR residues highlighted with a thicker line. Molecular structures are shown with a semi-transparent surface colored blue to red for low and high computed B-Factors, respectively. Each plot has 4 peaks. Three of these correspond to the CDR and are colored yellow while the fourth is marked with an asterisk and colored purple. The remainder of the structure is colored green. In A, B, and C the CDR faces to the left while in D it faces right. This has been done because in D the back side is less mobile and the yellow CDR loops are less distinguishable when looking through the blue surface. In the  $F_{ab}$ -like T-cell receptor the fourth loop (\*) is also variable in sequence.

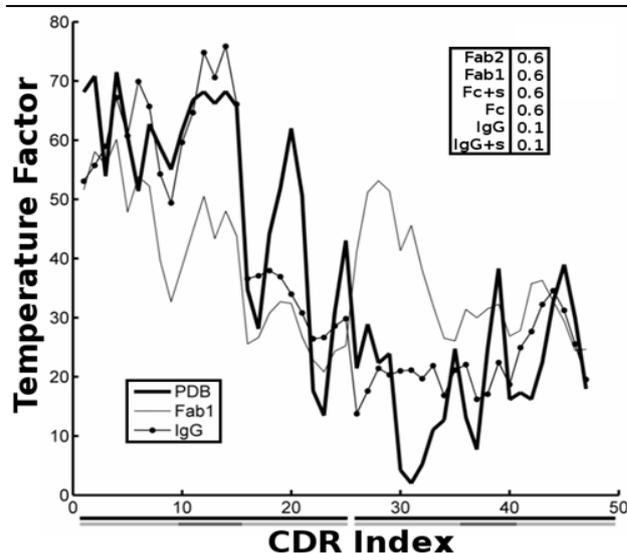
that the whole antibody structure may facilitate CDR configurational sampling so that a proper binding pose is found.

Information transfer within molecular structures has been the focus of numerous studies including the consideration of protein structures being fractal in nature (see Methods). Following previously established algorithms, we compute the fractal and spectral dimension of IgG and single  $F_{ab}$  domains (Figure 10). Briefly, the fractal dimension describes how the mass captured by concentric spheres scales with the radius of these spheres, and the spectral dimension describes how the frequencies of vibration for the structure scale with the density of modes. We find that the full length IgG behaves nearly like a 2D object in terms of its ability to transfer information from one part of the structure to another. Information transfer is thus significantly faster than one would expect from a uniform crystal lattice. This spectral dimension is in the range expected for proteins [24]. Interestingly, we find that the full structure has a lower spectral dimension than any single domain, again pointing to the possible utility of the whole structure for finding the right binding pose.

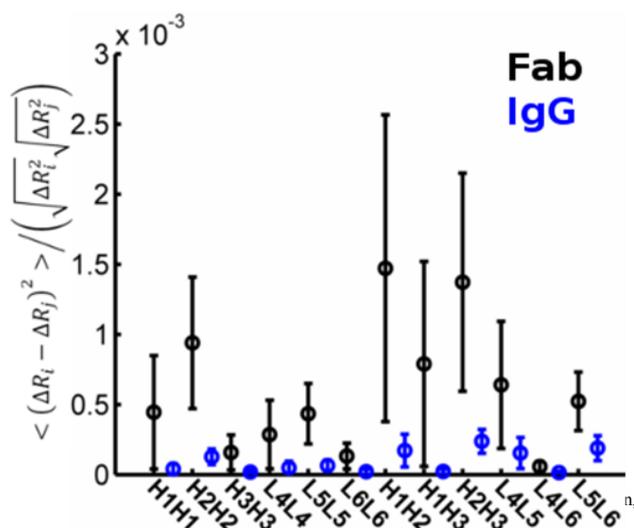
### Conclusion

Normal mode analysis using ANM are shown here to detect the hinge motions within the dominant low frequency motions, as well as the internal motions of the IgG structure. We have also evaluated the protein crystal and compared against the experimental temperature factors and disorder predictions. We find that the protein environment may be misleading in the crystal regarding the actual functional motions. Crystallographic temperature factors also reflect the crystal intermolecular interactions, which are extensive in this structure. Modeling approaches such as those applied here can provide a more comprehensive view of the biomolecule and its functional motions. The spectral dimension relates the density of vibrational modes to their frequency and can be used as a judge of the efficiency of energy transfer through a structure. Since this quantity is lower for IgG than for any individual domain, the hypothesis that hinge motions (the dominant

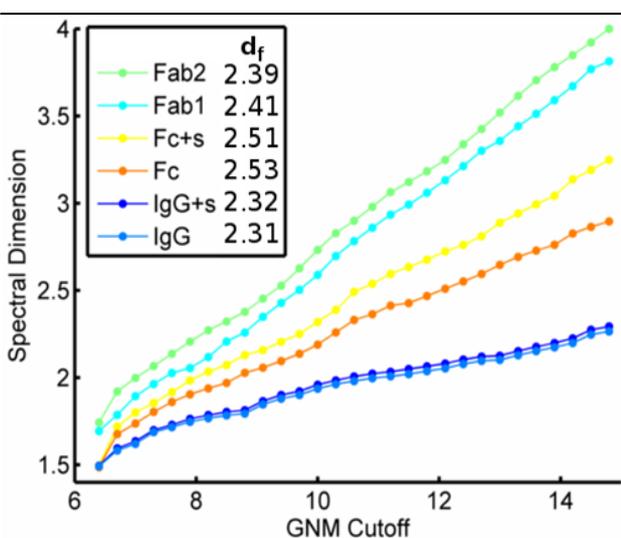
computed motions) facilitate CDR motion is strongly supported. We conclude that the IgG structure is specifically designed to facilitate large excursions of the  $F_{ab}$  domains, as shown with the present methods for computational evaluation of the extent of hinge motions and the spatial sampling by components of the structure. Normal modes derived from the simplest potential function afford a good approximation to the total hinge motion and predict the most sequence-variable regions also to be the most spatially mobile – facilitating the binding of the  $F_{abs}$ . Our results may impact immunology by suggesting ways to include flexibility in the docking to predict the bound structures of IgGs.



**Figure 8:** Correlation of experimental and computed temperature factors of the CDRs is stronger when the full IgG structure is considered than when parts of the structure are considered. Experimental temperature factors from the CDR are plotted in the heavy line. The first 25 residues plotted are from the heavy chain and the next 22 from the light. Heavy (left) and light (right) chains are delineated by black bars underneath the residue index and individual CDR hypervariable loops are shown by gray bars. CDR motion computed from the full IgG has a higher correlation (0.87) with experimental data than motion computed for the CDR using only the  $F_{ab}$  domain (0.50) when the 50 lowest frequency normal modes are used. The correlation between the two theoretical curves is 0.60. (inset) The correlations between temperature factors from the experimental B factors and computed from GNM models are shown. The magnitude of the correlation is not significantly affected by the choice of cutoff value. “+s” indicates the inclusion of coarse grained points from sugar molecules that attached to the  $F_c$ . It is evident that the motions available to IgG in the crystal environment are not likely to be identical to those in solution.



**Figure 9:** Inclusion of the full IgG structure diminishes the magnitude of internal distance changes. Normalized internal distance changes within the CDR calculated from ENM modes are scaled according to equipartition and constructed from the full IgG structure (blue) and from a single Fab domain (black). The IgG structure itself appears to facilitate significantly larger excursions of the CDR loops away from their native positions, but does so without any significant internal rearrangements. The non-normalized changes are significantly greater for IgG.



**Figure 10:** Spectral and Fractal dimension of IgG and its domains. The full length IgG has a lower spectral dimension no matter the cutoff employed. For the typical GNM of 7.3 the spectral dimension of IgG is 1.7. This is in good agreement with experimental measures on other proteins. The Fab domain alone has a larger spectral dimension of about 1.9, but all structures considered have a similar fractal dimension. “+s” indicates the inclusion of coarse-grained points from sugar molecules that are affixed to Fc.

**Acknowledgements:**

We gratefully acknowledge the financial support provided by the National Institutes of Health through grants R01GM081680, R01GM072014, and R01GM073095. We thank Saras Saraswathi for assistance in preparing figure 2.

**Reference List**

1. Wunsch E, Moroder L, Gohring-Romani S, Musiol HJ, Gohring W, Bovermann G: **Synthesis of the bis-cystinyl-fragment 225-232/225'-232' of the human IgG1 hinge region.** *Int J Pept Protein Res* 1988, **32**:368-383.
2. Niederhafner P, Gut V, Jezek J, Budesinsky M, Kasicka V, Wunsch E, Hlavacek J: **Synthetic study on cystinyl peptides using solution and solid phase methodology: human IgG1 hinge region.** *Amino Acids* 2010.
3. Yu D, Ghosh R: **Method for studying immunoglobulin G binding on hydrophobic surfaces.** *Langmuir* 2010, **26**:924-929.

4. Bahar I, Atilgan AR, Erman B: **Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential.** *Folding Des* 1997, **2**:173-181.
5. Haliloglu T, Bahar I, Erman B: **Gaussian dynamics of folded proteins.** *Phys Rev Lett* 1997, **79**:3090-3093.
6. Tirion MM: **Large amplitude elastic motions in proteins from a single-parameter, atomic analysis.** *Phys Rev Lett* 1996, **77**:1905-1908.
7. Atilgan AR, Durell S R, Jernigan RL, Demirel MC, Keskin O, Bahar I: **Anisotropy of fluctuation dynamics of proteins with an elastic network model.** *Biophys J* 2001, **80**:505-515.
8. Cui Q, Bahar I: **Normal Modes Analysis: Theory and applications to biological and chemical systems.** London, UK: Chapman & Hall/CRC; 2006.
9. Jernigan RL, Kloczkowski A: **Packing regularities in biological structures relate to their dynamics.** *Methods Mol Biol* 2007, **350**:251-276.
10. Sen TZ, Feng YP, Garcia JV, Kloczkowski A, Jernigan RL: **The extent of cooperativity of protein motions observed with elastic network models is similar for atomic and coarser-grained models.** *J Chem Thy Comp* 2006, **2**:696-704.
11. Doruker P, Jernigan RL, Bahar I: **Dynamics of large proteins through hierarchical levels of coarse-grained structures.** *J Comput Chem* 2002, **23**:119-127.
12. Kim MK, Jernigan RL, Chirikjian GS: **Efficient generation of feasible pathways for protein conformational transitions.** *Biophys J* 2002, **83**:1620-1630.
13. Kim MK, Li W, Shapiro BA, Chirikjian GS: **A comparison between elastic network interpolation and MD simulation of 16S ribosomal RNA.** *J Biomol Struct Dyn* 2003, **21**:395-405.
14. Kim MK, Jernigan RL, Chirikjian GS: **An elastic network model of HK97 capsid maturation.** *J Struct Biol* 2003, **143**:107-117.
15. Schuyler AD, Chirikjian GS: **Normal mode analysis of proteins: a comparison of rigid cluster modes with C-alpha coarse graining.** *J Mol Graphics Model* 2004, **22**:183-193.
16. Schuyler AD, Chirikjian GS: **Efficient determination of low-frequency normal modes of large protein structures by cluster-NMA.** *J Mol Graphics Model* 2005, **24**:46-58.
17. Stapleton HJ, Allen JP, Flynn CP, Stinson DG, Kurtz SR: **Fractal Form of Proteins.** *Physical Review Letters* 1980, **45**:1456-1459.
18. Lushnikov SG, Svanidze AV, Sashin IL: **Vibrational density of states of hen egg white lysozyme.** *Jep Letters* 2005, **82**:30-33.
19. Granek R, Klafter J: **Fractons in proteins: Can they lead to anomalously decaying time autocorrelations?** *Physical Review Letters* 2005, **95**.
20. Enright MB, Leitner DM: **Mass fractal dimension and the compactness of proteins.** *Physical Review e* 2005, **71**.
21. Burioni R, Cassi D, Fontana MP, Vulpiani A: **Vibrational thermodynamic instability of recursive networks.** *Europhysics Letters* 2002, **58**:806-810.
22. Reuveni S, Granek R, Klafter J: **Proteins: Coexistence of stability and flexibility.** *Physical Review Letters* 2008, **100**.
23. Reuveni S, Granek R, Klafter J: **Anomalies in the vibrational dynamics of proteins are a consequence of fractal-like structure.** *Proceedings of the National Academy of Sciences of the United States of America* 2010, **107**:13696-13700.
24. Leeuw M, Reuveni S, Klafter J, Granek R: **Coexistence of flexibility and stability of proteins: an equation of state.** *PLoS One* 2009, **4**:e7296.
25. Emekli U, Schneidman-Duhovny D, Wolfson HJ, Nussinov R, Haliloglu T: **HingeProt: automated prediction of hinges in protein structures.** *Proteins* 2008, **70**:1219-1227.
26. Verner Schomaker, K.N.Trueblood: **On the Rigid-Body Motion of Molecules in Crystals.** *Acta Crystallography B* 1968, **Acta Crystallography B**, **24**:63-76.
27. Kuriyan J, Weiss WI: **Rigid protein motion as a model for crystallographic temperature factors.** *Proc Natl Acad Sci U S A* 1991, **88**:2773-2777.
28. Song G, Jernigan RL: **vGNM: a better model for understanding the dynamics of proteins in crystals.** *J Mol Biol* 2007, **369**:880-893.
29. Linding R, Jensen LJ, Diella F, Bork P, Gibson TJ, Russell RB: **Protein disorder prediction: implications for structural proteomics.** *Structure* 2003, **11**:1453-1459.
30. Shimizu K, Hirose S, Noguchi T: **POODLE-S: web application for predicting protein disorder by using physicochemical features and reduced amino acid set of a position-specific scoring matrix.** *Bioinformatics* 2007, **23**:2337-2338.
31. Janeway CA, Travers P, Walport M, Shlomchik M.: **Immunobiology 5th edition.** Garland Science; 2001:Chapter 3-10.
32. Dobbins SE, Lesk VI, Sternberg MJE: **Insights into protein flexibility: The relationship between normal modes and conformational change upon protein-protein docking.** *PNAS* 2008, **105**:10390-10395.

33. **Bakan A, Bahar I: Computational generation inhibitor-bound conformers of P38 MAP kinase and comparison with experiments. *Proceedings of the Pacific Symposium on Biocomputing*, 2011, 181-192**