*Periodica Mathematica Hungarica Vol. 39 (1–3), (1999), pp. 241–252*

# RIGIDITY IN GLASSES AND PROTEINS

M. F. THORPE and LESLIE A. KUHN (East Lansing)

**Abstract**

We review recent progress in applying the theory of rigidity to glassy networks and to proteins. These three dimensional systems require a generalization of Laman's theorem, which we have used to develop a technique called the *Pebble Game* which allows the rigid regions (containing both isostatic and overconstrained parts) and the flexible joints between them, to be found. We show that a flexibility index, which measures the local density of floppy modes, is useful in characterizing the network. A sampling of recent results is given for network glasses, where we show how the glass structure can self-organize to produce an intermediate phase that is stress-free and contains a percolating isostatic cluster. In proteins, we show how maps of the rigid regions and flexible joints, as well as maps of the flexibility index, can help to elucidate the connection between structure and function.

## 1. Introduction

Atomic models of glasses, proteins, and other materials can be represented by a network or graph of length constraints. These length constraints represent the local geometry and chemistry by fixing covalent and hydrogen bond lengths and bond angles, and by inhibiting dihedral angle rotations around double bonds. Our goal is to determine the rigid regions and the flexible joints that separate them in these networks [1–7], which can lead to a better understanding of the relation between the structure and various observable physical properties.

Although no rigorous theorems exist for this specific combination of constraints in 3D, there is a rigorous theory for a slightly more generic set of bodies with hinge constraints [3–5]. Moreover, Tay and Whiteley [2,7] have conjectured (the *Molecular Framework Conjecture*) that this body and hinge theory does extend to the less generic molecular models in which all bonds (hinges) of an atom pass through a single central point of the atom. This conjecture, in turn, is precisely equivalent to the Laman type counting assumptions built into our 3D pebble game.

*Mathematics subject classification numbers,* 2000. Primary 74E35, 92C40; Secondary 68R10.

*Key words and phrases.* Rigidity, networks, glasses, proteins.

*0031-5303/99/$5.00*

*Akadémiai Kiadó, Budapest*
*Kluwer Academic Publishers, Dordrecht*

In studying hundreds of complex structures we have never found an exception to the applicability of these Laman type counts in 3-space for these molecular structures, but a rigorous proof would be highly desirable [1–7].

We have used a bookkeeping device, the *Pebble Game*, for counting constraints in the network that associates three pebbles with each atom and moves the pebbles around locally so as to balance degrees of freedom against constraints. This procedure is an integer algorithm and is extremely fast: the rigid region decomposition for a million-site (atom) network can be computed in a few seconds on a personal computer. This algorithm scales linearly with the size of the system in most situations, because the pebbles are usually redistributed locally.

Once the constraints are identified, the search for rigid regions and the intervening flexible joints can begin. For small molecules, the rigid regions and flexible joints can be found as follows. An unconstrained point (atom) has 3 degrees of freedom in a 3-dimensional space. For the $r$ atoms in an $r$-fold ring, there are $3r$ degrees of freedom, which are reduced by the $r$ covalent bond-length constraints in the ring and by another $r$ covalent bond-angle constraints. This leaves a residual $r$ degrees of freedom, of which 6 are the macroscopic rigid body motions involving the whole molecule. Therefore, there are $r-6$ internal bond-rotational degrees of freedom, or floppy modes associated with a ring. These are zero-frequency continuous deformations of the molecule consistent with the constraints and therefore do not cost any energy. When $r=6$, the number of floppy modes is zero, and the region is said to be isostatically rigid. This ring has two realizations, the chair and the boat, which are conformations with no continuous deformation path between them. A transition between two realizations in such a rigid cluster will be referred to as a *flip*. For $r>6$, the ring is underconstrained, or flexible. Thus a seven-fold ring has a single floppy mode that can be visualized as a rolling motion around the perimeter, and we will refer to this type of motion as a *roll*. A five-fold ring has one constraint more than is needed for rigidity, and so is overconstrained with one redundant constraint, creating stress within the ring. The flexibility of each bond in the ring can be quantified by a flexibility index, $f=(r-6)/r$, giving the number of floppy modes per bond within the ring. For an overconstrained ring, the flexibility index will be negative, giving the number of redundant constraints per bond. Thus the flexibility index represents the *density* of floppy modes (or redundant bonds) within a region.

For interlocking rings and other complex, branched networks, it would be extremely difficult to determine manually which constraints are independent. This is because constraints on distant bonds may influence local rigidity if the bonds are coupled through a ring or interconnected rings, and the bookkeeping for this becomes very complex. However concepts from graph rigidity theory, using the *pebble game* can be applied to count dependent and independent constraints in macromolecules of the size and complexity of proteins. The techniques used for the study of proteins are very similar to those used in glasses — the major change being the necessity of also modeling hydrogen bonds in the network.

Rigidity in 3D is much more complex than in 2D, as it is much more nonlocal in character. Ridid regions no longer consist of contiguous units and there

is no three dimensional generalisation of Laman's theorem for the general case. However in this work we use the conjecture discussed earlier. We have considerable confidence in this as our implementation, using the a *Pebble Game* based on it, has internal consisitency checks that would fail if the conjecture were wrong, and this has never happened. We can therefore proceed, while waiting for a rigoroius matematical proof.

In this review we give examples of our work in glasses [10], where we show that by eliminating stressed regions (overconstrained regions containing redundant constraints), the glass undergoes *two* phase transitions. Initially there is a second order phase transition from a floppy phase to a rigid but unstressed phase (with an isostatic percolating backbone), and subsequently there is a first order transition to a rigid and stressed phase (with an overconstrained percolating backbone).

For proteins, we show how this method can be applied to find biologically important flexible regions [11], using HIV protease as an example. A *flexibility index*, calculated as the local density of floppy modes (deformations that are consistent with the constraints), gives a useful quantitative measure of the magnitude of the local flexibility.

## 2. Glasses

The study of the structure of network glasses has progressed steadily since the initial work of Zachariasen [12] in 1932 that introduced the idea of the Continuous Random Network (CRN). Zachariasen envisaged such networks maintaining local chemical order, but by incorporating small structural distortions, having a topology that is non-crystalline. This seminal idea has met some opposition over the years from proponents of various microcrystalline models, but today is widely accepted, mainly as a result of careful diffraction experiments from which the radial distribution function can be determined [10]. The CRN has been established as the basis for most modern discussions of covalent glasses, and this has occurred because of the interplay between diffraction experiments and model building. The early model building involved networks with $\sim$ 500 atoms constructed from a seed with free boundaries in a roughly spherical shape [13]. Subsequent efforts have refined this approach and made it less subjective by using a computer to make the decisions and incorporating periodic boundary conditions. The best of these approaches was introduced by Wooten, Winer and Weaire [14] and consists of restructuring a crystalline lattice with a designated large unit supercell, until the supercell becomes amorphous. The large supercell contains typically $\sim$ 5000 atoms. Both the hand-built models and the Wooten, Winer and Weaire models are relaxed during the building process using a potential. The final structure is rather insensitive to the exact form of the potential and a Kirkwood [15] or Keating [16] potential is typically used. Such models can be examined using the *pebble game* and some typical results are shown in figure 1. These models have been very successful when compared with experimental diffraction data on glasses [10].

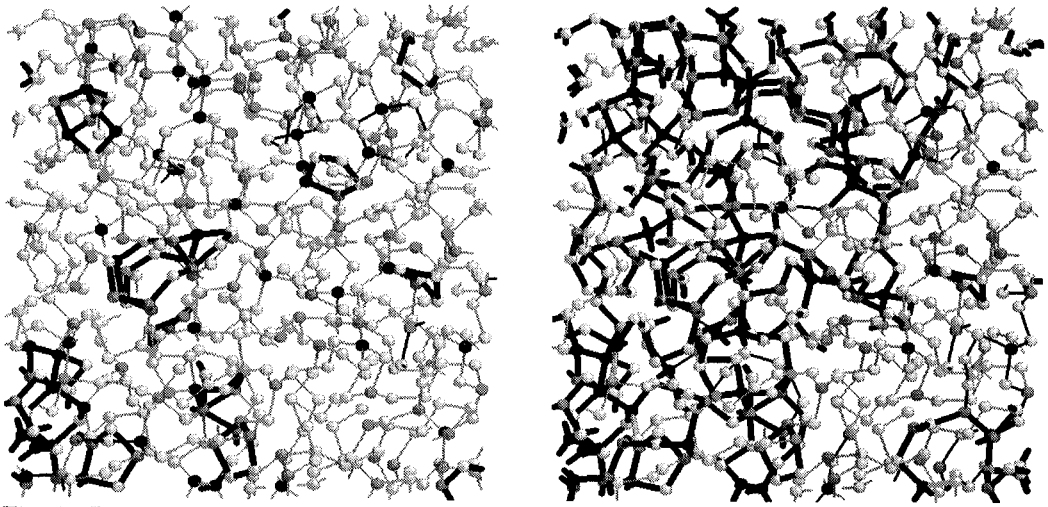

Fig. 1. Typical section of a random network with low mean coordination (2.37), where rigidity has not percolated (left) and with a high mean coordination (2.40), where rigidity has percolated (right). The wide black bonds are in overconstrained regions, while the thin black bonds are isostatic. The hinge joints are shown as grey lines.

Despite this success in understanding the glass structure, some concerns remain. Perhaps the most serious of these is that the network cannot be truly random. Even though bulk glasses form at high temperatures where entropic effects are dominant, it is clearly not correct to ignore energy considerations that can favor particular local structural arrangements over others. A simple example of this is local chemical separation, where, for example, bonding between like atoms is favored over bonding between unlike atoms. This can lead to chemical thresholds. A more subtle effect of interest here is how the structure itself can incorporate non-random features in order to minimize the free energy at the temperature of formation. This is even more important in amorphous solids, which are usually formed at lower temperatures, and so energy considerations are relatively more important and one can expect more non-random local structural arrangements. Such subtle structural correlations, which we refer to as *self-organization* [17], will almost certainly not show up in diffraction experiments, but may have other manifestations.

A very simple approximation is due to Maxwell and we will refer to this as Maxwell counting [10]. This involves a single global count, where the total number of degrees of freedom is set equal to the number of constraints. This is clearly not correct, but provides a very useful first approximation in networks that are rather homogeneous. It is convenient to express the result in terms of the mean coordination, which is defined as the mean number of nearest neighbors per site. For networks where the sites have individual coordiantions of 2, 3 and 4, Maxwell counting predicts that the mean coordination at the transition from floppy to rigid is 2.4, and that the number of floppy modes is linear in the mean coordination, going to zero at a value of 2.4 as shown in figure 2.

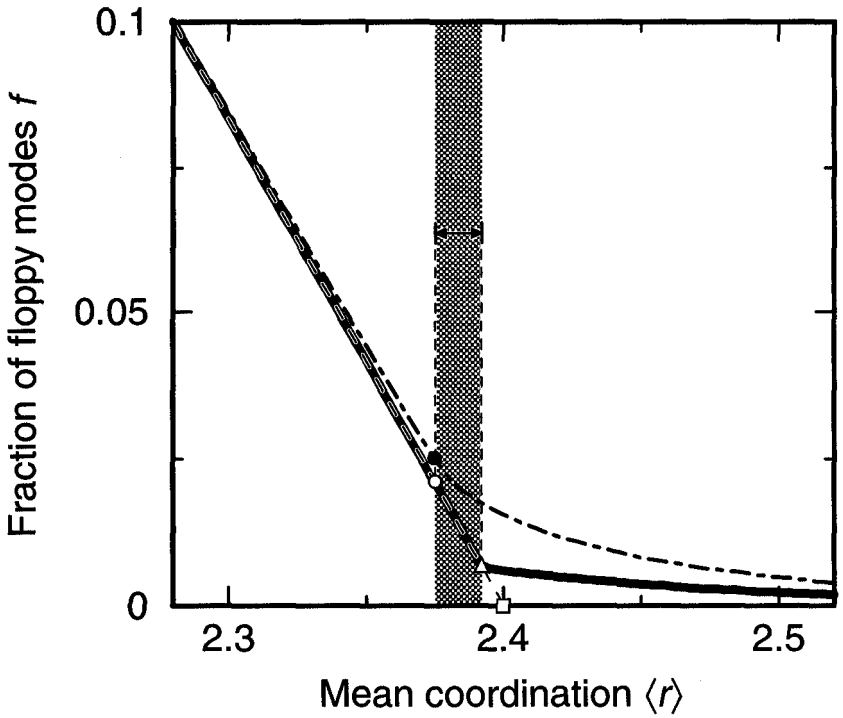


Fig. 2. The number of floppy modes plotted against the mean coordination for Maxwell counting (shown as a dashed line) with the associated mean field transition shown by the open square at $\langle r \rangle = 2.4$, and for a randomly diluted diamond lattice using a dot-dash line, where the second order transition is indicated by the solid circle at $\langle r \rangle = 2.375$. The self-organized model follows the Maxwell curve, and is shown by a solid line, and gives a second order transition at $\langle r \rangle = 2.375$ (open circle) from a floppy to an unstressed rigid state and a first order transition at $\langle r \rangle = 2.392$ (open triangle) to a stressed rigid state. The range of $\langle r \rangle$ over which the intermediate phase exists is indicated by the grey panel.

We focus on the mechanical properties and critical mechanical thresholds, as this is where it is easiest to make theoretical progress at this time. How can such an idea be developed theoretically? A proper procedure might be to use a very large supercell containing $\sim$ 5000 atoms and use a first principles approach, like that of Car and Parrinello [18] to form the glass at the appropriate temperature. This could possibly lead to self-organization of the kind discussed above; however this is unlikely, in the same way that the superconducting state would be hard to find from a brute force solution of the Schrödinger equation for a solid. We therefore need to consider other ways of generating self-organized networks. One promising approach is that of Barkema and Mousseau [19], who explore the energy landscape of a glass by moving over saddle points to search for successively lower minima. We look at even more simplified approaches that show what kinds of effects self-organization, and the resulting non-randomness, can lead to.

The first of these approaches asks: what would happen to the properties of a

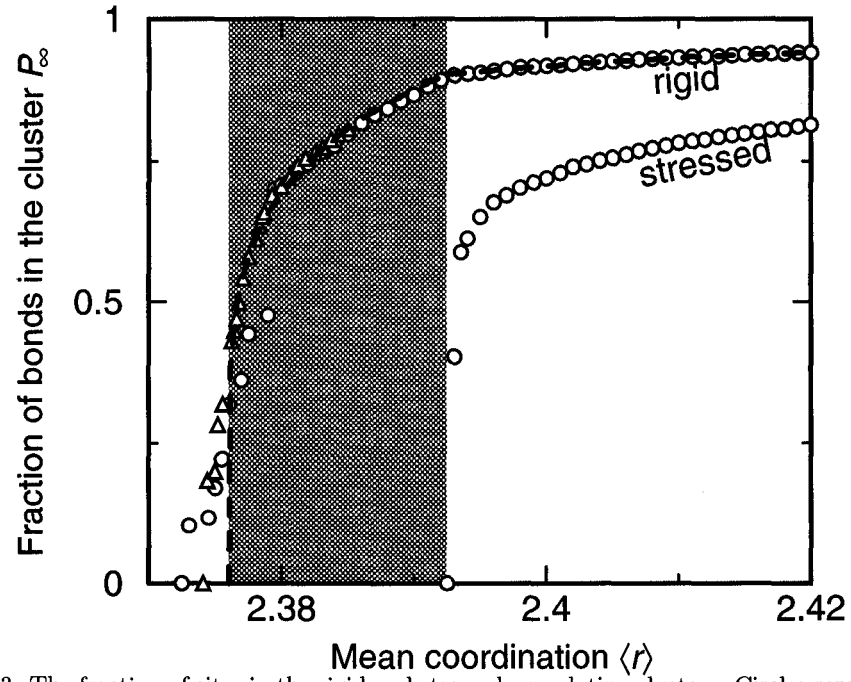


Fig. 3. The fraction of sites in the rigid and stressed percolating clusters. Circles represent the average over 4 networks with 64,000 sites, and triangles represent averages over 5 networks with 125,000 sites. The thicker dashed lines indicate a power law dependence before and after the stress transition. Note the break in slope of the dashed line at the first order transition at $\langle r \rangle = 2.392$. The intermediate phase, which is rigid but unstressed, exists for $2.375 < \langle r \rangle < 2.392$.

CRN if rings formed by bonds were eliminated as much as possible? Network models can be built where the smallest ring is 10-membered, and this can dramatically alter some properties of the network. In particular, the mechanical transition from a rigid to a floppy network, which occurs as the mean coordination is reduced below 2.4, seems to become first order rather than second order in character [10].

We also introduce a self-organized model of a random network in which configurations that are stressed are avoided if possible. This leads to two phase transitions and an intermediate phase that is rigid but stress-free. Preliminary results show that the phase transition at the lower mean coordination is second order and at the upper mean coordination is probably first order. There has been some recent evidence of a first order transition being seen in Raman scattering of chalcogenide glasses as the composition is varied [20], and also some evidence for an intermediate phase using differential scanning calorimetry [21]. These results are illustrated in figure 2, where the intermediate phase is shown. In the region to the left, with low mean coordination, the network is floppy, and to the right, at high mean coordination, the network is overconstrained, contains redundant bonds, and is therefore stressed. In the narrow intermediate phase, the network is isostatically rigid and

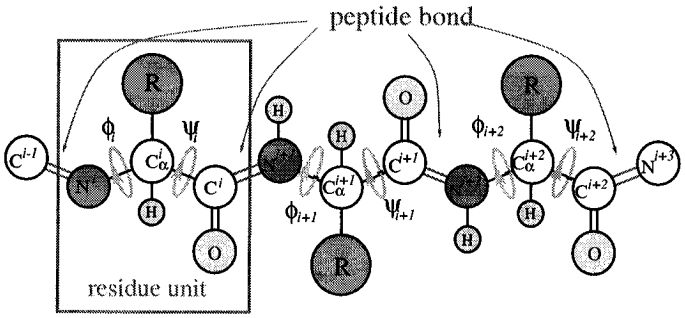


Fig. 4. Showing a piece of the peptide mainchain containing; comprising a backbone and sidegroups which are then crosslinked to form the folded protein. Double lines represent the peptide bond about which we do not allow dihedral angle rotation. Residues are denoted by R.

therefore unstressed. This is further illustrated in figure 3, where the isostatic rigid cluster and the stressed rigid cluster phases are shown.

## 3. Proteins

A new insight into modeling protein flexibility is that a protein structure can be reduced to its essentials, and viewed as a mechanical system of points (atoms) whose motion is limited by distance and angle constraints representing the interatomic bonds, with no explicit interatomic potentials required. When applied to the covalent and hydrogen-bond network of a single static protein structure, this approach can predict the major biologically important rigid and flexible features of proteins like HIV protease, which we use as an illustrative example here. We have also applied this approach to many other proteins (including dihydrofolate reductase, adenylate kinase, and a lysine-arginine-ornithine binding protein), and obtained a similar quality of predictions of experimental observables.

Predicting flexibility in proteins has proven elusive. While normal modes analysis [22] can extract the low frequency diffusive motion from molecular dynamics (MD) simulations, in practice it is not possible to run these simulations with realistic potentials for long enough (milliseconds) to sample the large-scale motions observed by experimental techniques. This motion is essential for the biological function of many proteins, such as the binding and processing of proteins required for HIV replication [23]. Hence, the development of techniques for studying protein flexibility remains an important outstanding problem.

Proteins are held together by several kinds of forces, of which the most important are the covalent forces that determine many bond lengths and angles, including the dihedral angles associated with peptide bonds. In addition, hydrogen bonds are responsible for forming the secondary structure of proteins, principally alpha-helices and beta-sheets, and stabilizing the higher-order structures in which

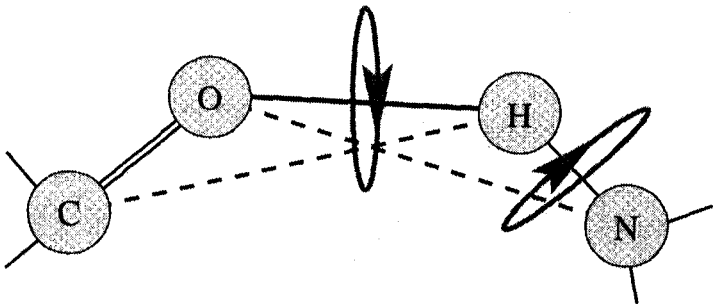


Fig. 5. A diagram showing a hydrogen bond inloving a donor and an acceptor atomtaken here as an nitrogen and oxygen respectively. It is modeled as three generic distance constraints, consisteing of a nearest neighbor central-force constraint shown as a thick solid line, and two next nearest neighbor bond bending force constraints shown as dashed lines. Each constrained hydrogen bond is also associated with three a priori rotatablke angles indicated by the arrows.

helices and sheets fold together into the complex units responsible for the diverse biological activity of proteins.

For modeling protein flexibility using a bond network or graph, the covalent bond between two adjacent atoms, say $C_\alpha$–C in an amino acid, can be considered to fix the distance between these two points, such that all motions remain consistent with this constraint. An angular constraint, reflecting the bond angle specified by the molecular orbital, is represented by a constraint between second neighbor atoms, e.g., between the N and C neighbors tetrahedrally coordinated to a main-chain $C_\alpha$ atom. An additional constraint is introduced between third neighbor O and H atoms in the O–C–N–H group, to prohibit any rotation around the C–N peptide bond linking amino acids along the protein main chain. Such constraints restrict the possible motions of the main and side chains, and have also been used in molecular dynamics calculations [24,25] to reduce the number of dynamical degrees of freedom. A wide range of hydrogen-bond strengths are found in proteins, with a typical geometric criterion for their assignment [25] being a donor-acceptor distance less than 3.5 Å and a donor–H–acceptor angle greater than 120°. These hydrogen bonds are included in the network used to represent the protein for the graph theoretical analysis. The hydrogen-bond criteria can be made more stringent either by decreasing the distance or increasing the angular threshold, resulting in only the strongest hydrogen bonds being included.

As an illustration, the constraint-counting analysis is applied to identify the flexible and rigid regions in HIV protease (HIVP), a major target for development of inhibitory drugs for use in anti-AIDS therapy. The results (figure 6) show that the protein is dominated by a single rigid cluster, including the base and walls of the substrate and inhibitor binding site (cavity at center) and four flexible regions (in each half of the dimer) shown as bonds in various grays (each gray indicating a rigid micro-cluster within the flexible region). We will refer to the flexible regions

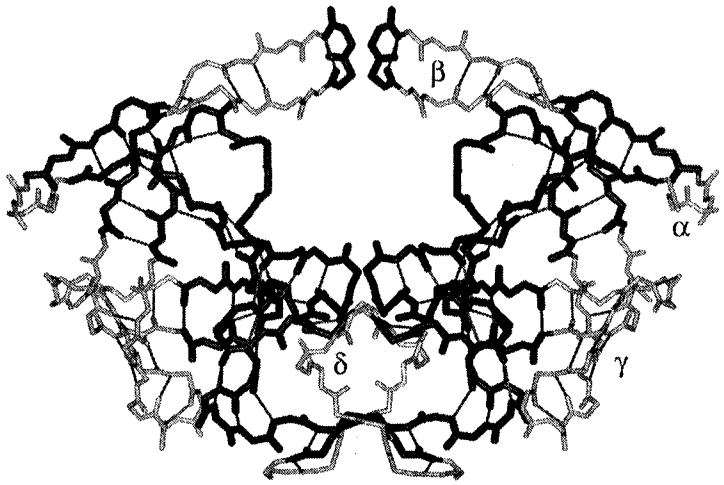


Fig. 6. Rigid and flexible region decomposition for HIVP. Each rigid cluster appears as a single color, including the dominant rigid cluster shown in black. Flexible regions appear as alternating grey bonds, and include the ends ($\beta$) and bases ($\alpha$) of the flaps that close over the substrate, inhibitor, and drug binding site (cavity at center). Additional flexible regions comprise the amino termini in the dimer interface at the bottom center ($\delta$) and in the side regions ($\gamma$). Light grey lines join the donors and acceptors of hydrogen bonds in this inhibitor-free form of HIVP (PDB entry 1hhp).

as $\alpha$, $\beta$, $\gamma$, and $\delta$ as shown in figures 6 and 7. The tips of the flaps (region $\beta$, residues 45–56, top-center) are known from crystallographic and nuclear magnetic resonance structures to be important for closing over and binding inhibitors [26], and appear as the most flexible regions in this analysis of a single, inhibitor-free structure (Protein Data Bank (PDB) entry 1hhp). Other flexible regions include the base of the flap (region $\alpha$, residues 39–42), the termini of the protein chains (region $\delta$, residues 1–8) at bottom center, and the side region (region $\gamma 1$, residues 13–20 and region $\gamma 2$, residues 60–74), consisting of discontinuous segments of the main chain (shown in figures 6 and 7).

Thus, significant insights can be gained into flexibility from the analysis of a single protein conformation, in a few seconds of computational time and without employing interatomic potentials. This increase in speed means that flexibility calculations for large systems such as proteins, which were previously infeasible, can now be done in real time. This new algorithm can be used to instantaneously assess changes in protein flexibility due to natural or designed side-chain mutations, substrate or inhibitor binding, and interactions with other molecules, including crystal lattice neighbors and solvent molecules. Using this approach, information about protein flexibility can be extracted from a single snapshot of the protein structure, which can aid in drug design as well as our understanding of protein folding.

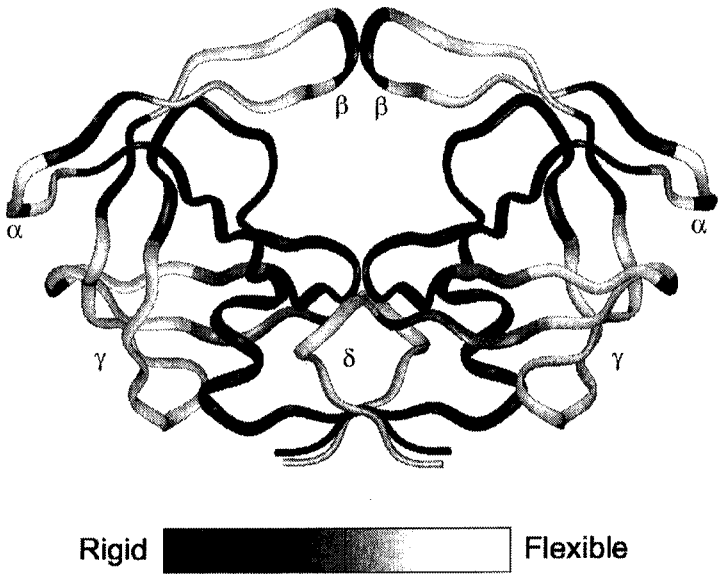


Fig. 7. The flexibility index is shown as a grey scale on the main-chain ribbon of the same HIVP structure shown in figure 6. The light regions are the most floppy (high flexibility) and the dark regions are the most rigid (high stability).

## 4. Conclusions

The work described here is a brief review of recent applications of rigidity theory to glasses and proteins. Further details can be found in the references cited, particularly the book containing references [10] and [11].

ACKNOWLEDGEMENTS. We should like to thank D. J. Jacobs, M. V. Chubynsky, B. Hespenheide, J. C. Phillips and A. J. Rader for their continuing contributions and discussions related to the work that has been reviewed here. We should also like to thank W. Whiteley for his many insightful comments. This work is partially supported by grants from the U.S. National Science Foundation.

## REFERENCES

[1] G. LAMAN, On graphs and rigidity of plane skeletal structures, *J. Engrg. Math.* **4** (1970), 331–340.

[2] T-S. TAY and W. WHITELEY, Recent advances in generic rigidity of structures, *Structural Topology* **9** (1985), 31–38.

[3] W. WHITELEY, Some matroids from discrete applied geometry, *Contemporary Mathematics* **197** (1996), 171–311.

[4] T-S TAY, Linking (n-2)-dimensional panels in n-space I: (k-1,k) graphs and (k-1,k) frames, *Graphs and Combinatorics* **7** (1991), 289–304.

[5] T-S TAY, Linking (n-2)-dimensional panels in n-space II: (n-2,2)-frameworks and body and hinge structures, *Graphs and Combinatorics* **5** (1989), 245–273.

[6] W. WHITELEY, The union of matroids and the rigidity of frameworks, *Siam J. Discrete Methods* **1** (1988), 237–255.

[7] W. WHITELEY, Rigidity of molecular structures: generic and geometric analysis, in *Rigidity Theory and Applications*, ed. by M. F. Thorpe and P.M. Duxbury, Kluwer Academic/Plenum Publishers, New York, 1999, pp. 21–46.

[8] D. J. JACOBS, Generic rigidity in three-dimensional bond-bending networks, *J. Phys. A: Math. Gen.* **31** (1998), 6653–6668.

[9] T-S TAY, On the generic rigidity of bar-frameworks, *Advances in Applied Math* **23** (1999), pp 14–28; B. SERVATIUS and H.SERVATIUS, Generic and Abstract Rigidity, in *Rigidity Theory and Applications*, ed. by M. F. Thorpe and P. M. Duxbury, Kluwer Academic/Plenum Publishers, New York, 1999, pp. 1–19; T-S Tay and W. Whiteley, Generating isostatic frameworks, *Structural Topology* **11** (1985), pp. 21–69.

[10] M. F. THORPE, D. J. JACOBS, N. V. CHUBYNSKY and A. J. RADER, Generic rigidity of network glasses, in *Rigidity Theory and Applications*, ed. by M. F. Thorpe and P. M. Duxbury, Kluwer Academic/Plenum Publishers, New York, 1999, pp. 239–277.

[11] D. J. JACOBS, L. A. KUHN and M. F. THORPE, Flexible and rigid regions in proteins, in *Rigidity Theory and Applications*, ed. by M. F. Thorpe and P. M. Duxbury, Kluwer Academic/Plenum Publishers, New York, 1999, pp. 357–384.

[12] W.H. ZACHARIASEN, The atomic arrangement in glass, *J. Am. Chem. Soc.* **54** (1932), 3841–3851.

[13] D.E. POLK, Structural model for amorphous silicon and germanium, *J. Non-Cryst. Solids* **5** (1971), 365–376.

[14] F. WOOTEN, K. WINER, D. WEAIRE, Computer generation of structural models of amorphous Si and Ge, *Phys. Rev. Lett.* **54** (1985), 1392–1395.

[15] J.G. KIRKWOOD, Skeletal modes of vibration of long chain molecules, *J. Chem. Phys.* **7** (1939), 506–509.

[16] P. N. KEATING, Effect of invariance requirements on the elastic strain energy of crystals with application to the diamond structure, *Phys. Rev.* **145** (1966), 637–645.

[17] M. F. THORPE, D. J. JACOBS, M. V. CHUBYNSKY and J. C. PHILLIPS, Self-organization in network glasses, accepted by *J. Non-Cryst. Solids.*

[18] R. CAR, M. PARRINELLO, Unified approach for molecular dynamics and density-functional theory, *Phys. Rev. Lett.* **55** (1985), 2471–2474.

[19] G.T. BARKEMA, N. MOUSSEAU, Event-based relaxation of continuous disordered systems, *Phys. Rev. Lett.* **77** (1996), 4358–4361.

[20] X. W. FENG, W. J. BRESSER, P. BOOLCHAND, Direct evidence for stiffness threshold in chalcogenide glasses, *Phys. Rev. Lett.* **78** (1997), 4422–4425; P. BOOLCHAND, X. W. FENG, D. SELVANATHAN, W.J. BRESSER, Rigidity transition in chalcogenide glasses, in *Rigidity Theory and Applications*, Kluwer Academic/Plenum Publishers, New York, 1999, pp. 279–295.

[21] D. SELVANATHAN, W.J. BRESSER, P. BOOLCHAND and B. GOODMAN, Thermally reversing window and stiffness transitions in chalcogenide glasses, *Solid State Comm.* **111** (1999), 619–624.

[22] A. AMADEI, A. B. M LINSSEN and H. J. C. BERENDSEN, Essential dynamics of proteins, *Proteins: Structure, Function, and Genetics* **17** (1993), 412–425.

[23] A. G. TOMASSELLI, W. J. HOWE, T. K. SAWYER, A. WLODAWER and R. L. HEINRIKSON, The complexities of AIDS: an assessment of the HIV protease as a therapeutic target, *Chimica Oggi* **9** (1991), 6–27.

[24] T. R. FORESTER and W. SMITH, Shake, rattle and roll: Efficient constraint algorithms for linked rigid bodies, *J. Comp. Chem.* **19** (1998), 102–111.

[25] S. M. HABERMANN and K. P. MURPHY, Energetics of hydrogen bonding in proteins: A model compound study, *Protein Science* **5** (1996), 1229–1239.

[26] L. K. NICHOLSON ET AL., Flexibility and function in HIV-1 protease, *Nature Structural Biology* **2** (1995), 274–280.

*(Received: December 2, 1999)*

M. F. THORPE
PHYSICS AND ASTRONOMY DEPARTMENT
MICHIGAN STATE UNIVERSITY
EAST LANSING, MI 48824
USA
E-MAIL: thorpe@pa.msu.edu

AND

LESLIE A. KUHN
BIOCHEMISTRY DEPARTMENT
MICHIGAN STATE UNIVERSITY
EAST LANSING, MI 48824
USA
E-MAIL: kuhn@agua.bch.msu.edu