by Mark Gerstein and Samuel Flores

Molecular Motions

A database of proteins' twists and turns could enhance precision in pharmaceuticals.

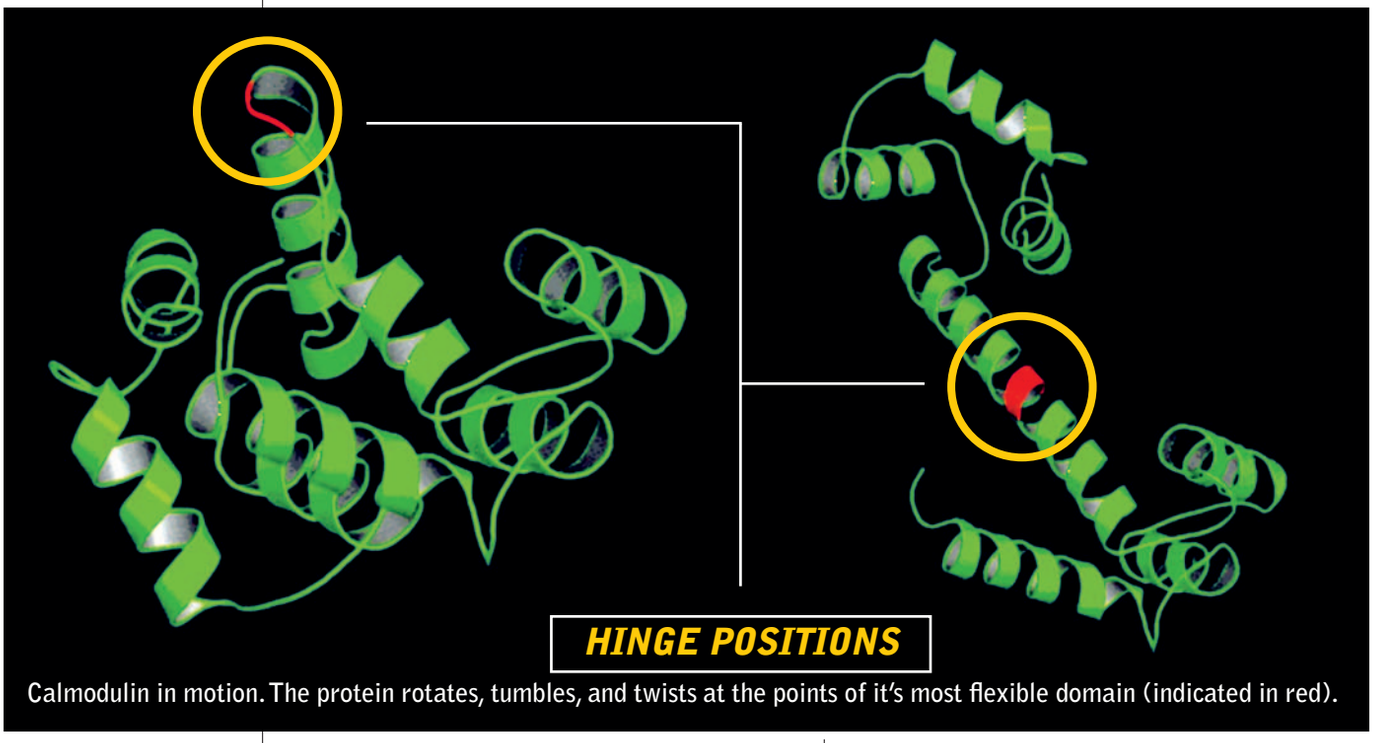
Mark Gerstein
is the Albert L. Williams Professor of biomedical informatics at Yale University

Samuel Flores
is a graduate student in physics at Yale University.

Our Database of Macromolecular Movements, www.molmovdb.org, compiles hypothesized protein motions. Proteins are large biological molecules that perform most of the jobs that keep life going—everything from building scaffolding in a cell to metabolizing food. In order to do these jobs, proteins usually have to bend, wiggle, move, and change their shapes. Sometimes, such as when digesting nutrients, proteins move to expose an active site, which is the part of a protein that starts a chemical reaction. At other times, such as when the muscle-protein myosin grasps actin in order to pull itself along, proteins move to conform to the shape of another structure that they need to grasp. There are many other examples of proteins moving to do their jobs. The point is: We have to know how proteins move in order to understand how they work. Ultimately this understanding will lead to more basic knowledge of biology and practical information that helps us design drugs and cure disease.

Unfortunately, proteins are incredibly tiny. They can't be seen in a microscope, for instance. So, it's hard to get information about how proteins are really moving. What we can do is use a technique called X-ray crystallography in which X-rays are fired at a crystallized protein, and the diffraction pattern can be used to solve the structure. This gives us only one snapshot of the protein. We can solve the structure again under different conditions, say, bound to another molecule. That gives us two snapshots of the protein. We can then interpolate between these two snapshots to get an idea of how the protein naturally moves. Think about a baseball player swinging a bat. Photographs of a player at the beginning and end of a swing can be used to surmise that at intermediate times the bat was at interpolated positions between the starting and ending points. That's essentially what we did to produce the motions stored in our database.

One of the primary motivations for studying protein flexibility is its application to computational drug design. If we know where the flexible domains are in a protein, we can allow flexure and rotation only there in a model, keeping the rest of the protein rigid, at great savings in computational cost. The resulting model should obtain a better estimate of the binding energy of a drug to a protein than would the currently available programs, which usually assume the protein is completely rigid. Our efforts are thus oriented toward using our database to advance knowledge of flexibility in proteins. We are excited about the prospects of the field and are always pleased to help workers in the field use the information our database can offer.



Calmodulin in motion. The protein rotates, tumbles, and twists at the points of it's most flexible domain (indicated in red).

ILLUSTRATION: ORANGE GRAPHICS