Towards a Knowledge Base of Structural Biology

James F. Brinkley
Department of Biological Structure SM-20
University of Washington
Seattle, Washington 98195

The long term goals of our group are to develop a large, distributed knowledge base of structural biology, and to develop computer modules which utilize this knowledge to solve problems in basic science, teaching and clinical medicine.

Structural biology can be defined as a basic medical science which studies the physical organization of the human body at levels ranging from gross anatomy to molecules. A knowledge base of structural biology would therefore contain the kind of information found in a textbook, namely spatial (or visual) knowledge about the morphologic shape and range of variation of structural objects; and symbolic (or textual) knowledge about the names of objects, their relationships, what they do, how they develop, and what can go wrong with them. Since structural biology is a fundamental framework upon which most of the basic medical sciences rest, and since basic medical science is in turn the foundation for clinical medicine, a knowledge base of structural biology should have important implications for clinical medicine, teaching and research.

Within our group the major activity has been the development of a 3-D spatial database of large portions of the human body obtained from serial sections. The database is primarily utilized for generating 3-D graphics reconstructions for teaching, but it will also be useful for many other applications, including the development of morphometric models comprising the spatial knowledge base of shape and range of variation.

More recently we have designed a framework for adding knowledge to the growing spatial database. Recognizing that representation of both spatial and symbolic knowledge is difficult, we have devised a distributed framework that we believe will allow incremental development while providing increasingly useful problem solving capabilities. Within this framework, separate modules run on different machines and communicate via remote procedure calls. The advantages of this approach are that different machines are appropriate for different tasks, and that modules may

be developed by different research groups at widely scattered locations. Each module may be accessed via a standalone user interface, or by a program interface making it accessible as a server to other modules. The modules that we have currently defined include symbolic and spatial database servers, symbolic and spatial knowledge base servers, low level imaging and graphics servers, and knowledge-based modules for graphics, image processing, tutoring, diagnosis, treatment, and research. The knowledge-based modules will access the structural knowledge base embodied in the knowledge servers, and these servers will in turn provide links to the database and other servers.

The reason we have been able to identify so many potential useful knowledge-based modules is that structure is fundamental to function and hence to medicine. Our research approach is to design the knowledge-base in response to an interrelated set of difficult real-world problems, in much the same way that expert systems were developed by looking at real-world problems. The problems that we are currently looking at are knowledge-based analysis of 3-D head and neck computed tomography images for radiation treatment planning, and knowledge-based information retrieval for neuroscience.

The hypothesis behind our approach to the image analysis problem is that knowledge of anatomy is necessary in order to interpret images. Thus, a solution to this problem will require development of the spatial knowledge base. On the other hand, the information retrieval problem will require development of the symbolic knowledge base of structural biology, as well models of expected users. Our approach to this problem is to initially develop hypermedia systems as front ends, and later to develop AI back ends that contain models of the user so that the hypermedia information is presented in the most appropriate manner.

In our poster we will present our current progress in each of these areas: 3-D reconstructions, knowledge-based image analysis, and knowledge-based information retrieval. We will also show how solutions to these problems will fit into the framework of a large, distributed, and reusable knowledge base of structural biology.