**COMPUTER GENERATED TANGIBLE INTERFACES FOR MOLECULAR BIOLOGY**

**INTRODUCTION**

The use of physical models in the molecular sciences has a long and successful history. Early ball-and-stick models of the first chemical structures helped scientists visualize and understand the spatial relationships between atoms. Chemists have used the visual, tactile, and kinematic characteristics of physical models to help them understand and design new molecular structures. The widely-used Dreiding models helped in conceptualizing the notions of torsional freedom and bond strain. The CPK or space-filling models developed by Pauling and co-workers served as analog computers to determine molecular shape, stereochemistry, and steric hindrance of large organic compounds. The famous DNA double helix built by Watson and Crick catalyzed thinking about the nature of chemical complementarity and molecular information storage. Our more recent explorations of the complex chemistry and shape of life’s molecular machinery have necessitated the intensive use of computation and computer graphics, far outstripping hand-made physical models. Therefore, visualization and simulation of the molecular biological world has unfortunately been largely confined behind the computer screen in intangible models. The recent development of computer controlled automated fabrication (e.g., stereolithography, laminated object manufacturing, etc.) has enabled new possibilities for the inexpensive production of detailed and accurate physical models. Up to this point, however, automated fabrication has been used primarily in engineering and manufacturing for rapid prototyping, with relatively little research into use for scientific visualization, interaction, and collaboration.

The goal of this proposal is to couple advances in the computational creation and integration of physical objects with the rapidly growing knowledge base from structural molecular biology. We will create new modes of interacting with biomolecular data. These modes encompass novel educational and research uses. The use of “physical visualizations” will enhance the experience and perception of the molecular nanoworld for all, but especially for the sight-impaired. Augmented Reality (AR) applications, where the physical models are used as haptic, real-world input devices in conjunction with computation and computer projected (visual, audio, haptic) output, will enhance exploration and understanding of biomolecular structure and interaction, by providing dynamic information overlay, private and public data display, context-sensitive visual appearance, and physically-based interactions. “Smart” models with embedded processors and communications will provide a new mode of querying and interacting with data in a natural and portable fashion for mobile and collaborative research and educational activities.

**GOALS AND OBJECTIVES**

Understanding the molecular basis of life is one of the most important scientific and computational challenges on our current horizon. The benefits to society could be enormous — ranging from improved health and longevity, to the development of new human potential. The challenges lie not only in the enormous complexity of biological systems, but in our abilities to interact with, comprehend, and communicate the beauty and intricacies of this submicroscopic world. As the Human Genome Initiative and post-genomic projects bring this world closer to us all, we need to develop new ways to explore and learn from it. From scientific research to general public understanding, new computational technologies can be used to make the invisible accessible and familiar.

We now know that the "life force" resides in the chemistry of biological molecules. As with all chemical phenomena, this is a science of shape. Molecular recognition and interaction is a reflection of chemical shape complementarity. This was the key finding of Watson and Crick in discovering the information-carrying capacity of DNA. The past 40 years have emphasized this lesson again and again in the structures of the more than 10,000 proteins that have been solved to atomic resolution by x-ray crystallography and NMR spectroscopy. Shape is a geometric, three-dimensional property. It is also a tangible characteristic, which can be felt and manipulated as well as seen. Computer generated physical models present intuitive access to the complex shapes and relationships of biological molecules and their interactions. Coupling these models with computational input and output will provide a natural interface between the user and the wealth of data and simulation coming from the structural biology community. Such access will be of value to the scientist as well as the student and the general public.

The specific research objectives of this proposal:

1. **Geometric representations appropriate for creating physical models of biomolecular structures.** We will integrate CAD/CAM and Augmented Reality interfaces with computational molecular biology and bioinformatics tools, providing means to create the geometric models necessary to fabricate and augment the physical models.

2. **Automated design and production of complex molecular models.** We will develop CAD/CAM algorithms for process-planning and fabrication providing means to integrate physical models with augmented reality and embedded computing.

3. **Fabrication of models that embody physico-chemical characteristics.** We will develop novel uses, combinations, and variations of automated fabrication technologies for purposes of tangible scientific visualization and interaction.

4. **Multi-modal Augmented Reality interfaces using tangible molecular models.** We will develop real-time object-registered computer graphics, haptics, and sound for interactive manipulation, query, and simulation in structural molecular biology.

5. **Application of tangible interfaces in molecular biology.** We will use the human-computer interface of physical models for interacting with structural and bioinformatic data from genomics and proteomics (e.g., amino acid sequence conservation and variation) and computational studies (e.g., electrostatics, mechanics, dynamics, free energy).

6. **Evaluation of physical models.** We will evaluate the effectiveness of different tangible representations of biomolecular structures in research, education, and the public understanding of science.

The anticipated impact of the proposed research:

1. The creation of physical models as tangible visualizations of important structures and processes in molecular biology for the purposes of research and education.

2. New accessibility for the visually impaired community, scientists, and the general public to multidimensional information on biological macromolecules and their interactions.

3. A new and far more intuitive mode of interacting with biological information by means of physical objects that are true molecular analogues for computer query, calculation, simulation, and visualization, bringing the human-computer interface to a higher level of use.

4. Innovations and improvements in the use of computer-automated fabrication technology for broader scientific visualization and collaboration purposes, driven by the complexity inherent in biology.

5. Influence on the external technological development of automated fabrication and embedded computing toward the needs of biological research and education.

C. Project Description page 2
Accomplishing these aims will require a multi-disciplinary team. Molecular biologists must be brought together with experts in CAD/CAM and Human–Computer Interfaces. Integrating the research and development in these disparate fields necessitates a single coordinated effort such as this, rather than several individual projects. This project comprises a unique collaboration between three world-recognized groups of researchers, each contributing expertise and experience in addressing the following research challenges.

### 1. Challenges and Approaches to Molecular Modeling and Software Development for Tangible Visualizations in Molecular Biology.

Work in the Molecular Graphics Laboratory (MGL) at The Scripps Research Institute (TSRI) has focused on software development for molecular computation and visualization for almost 20 years. The group has pioneered new computational approaches to molecular modeling, and since 1991 has been generating physical molecular models by means of automated fabrication techniques. This project will combine that experience with the new challenges of designing tangible visualizations that convey a perceptually expanded sense of the shapes and interactions of biological molecules.

Within this proposal the driving problem of structural molecular biology is one with significant scientific challenges and societal impact. A principal activity of the molecular modeling component will be to identify important structural data and to design and prototype tangible models representing that data, to be used in environments that range from forefront research in molecular biology to training and education. These activities will be driven by scientific studies from within TSRI and by interactions with scientists and educators in the broader community. Although the training and educational uses of tangible interfaces will be based upon our current and evolving knowledge, it is difficult to predict, at this early point, which uses and activities creative scientists will desire or demand in pursuit of their research. An important role of the TSRI group will be to explore novel biomolecular representations and uses within this context.

**Defining challenging and productive prototype problems.** We will develop and explore the utility of tangible models for a number of key systems that underlie the molecular structures and mechanisms of life. Protein recognition and interaction underpin most of the basis for biological function. A deeper understanding of molecular recognition, signaling, and self assembly will have profound impacts on such areas as cancer treatment, immune system control, and the battle against viral diseases. Tangible models that embody shape and the distribution of chemical and physical properties will aid in teaching the principles that are currently understood, and will help in developing and testing new scientific hypotheses.

Molecular self-assembly is the process that confers spatial organization in biological systems — from protein folding to viral replication to cell motility and division. Work in the laboratory of Jack Johnson, at TSRI, on the assembly of virus capsids, provides an atomistic view of various assembly states [Natarajan and Johnson, 1998]. However, algorithms for computationally assembling a virus from its components are still lacking. The ability to produce physical models of complex assemblies and assembly components will provide an enhanced perceptual context within which to explore ideas of self-organization [Bailey 1998]. Virus structure presents a tantalizing arena for the use of replicated parts, since there are typically only a small number of protein types that comprise the assembled viral capsid built from hundreds of copies. The design and fabrication of a set of viral components that "snap together" to form an accurate model of capsid architecture and flexibility will foster new approaches to understanding the principals of viral assembly, both for teaching and for research collaboration.

Models of the intact capsids could also be designed and fabricated (no assembly required), to show spatial and kinematic relationships within the assembled viral architecture. Such models may require built-in articulation (joints, hinges, tethers, etc.), to capture the essential dynamics of the subunits and the assembly. The analysis of rigidity and flexibility of biomolecules is an area of very active computational and experimental research. Incorporating "essential modes" [Kitao and Go, 1999] at the appropriate
scale for the representation being built will be a modeling challenge. The results would produce new ways of exploring kinematic and dynamic possibilities through physical manipulation, analogous to using Dreiding models to explore bond strain, or CPK models to explore steric hindrance in small molecule structures, but at a more general level of scale. One of Johnson’s viral structures, DNA bacteriophage HK97[Wikoff 1999], displays a "chain-mail" interlocking topological relationship between protein subunits in its spherical shell [Johnson 1998]. The mechanical and physical properties of such nano-structures have yet to be fully explored but could be modeled by direct manipulation in a physical visualization.

The understanding of biomolecular shape and interaction is also important in the area of drug and vaccine design, and will have significant impact on protein engineering and the emerging field of bio-nanotechnology. A tactile interface to the process of molecular design can serve to facilitate and expand the discovery process. Analysis and prediction of protein docking and interface formation can be guided and perceived through real-world "icons", representing physically accurate characteristics of the structures and interactions. Computer augmented data retrieval and computational steering via tangible models will facilitate interactive exploration of docking possibilities. Work in the Olson laboratory has focused on the prediction of protein interactions, both in the ligand binding drug design application (AutoDock [Goodsell and Olson 1990; Morris 1998]), and in the characterization of protein–protein docking (SurfDock [Duncan 1996; Strynadka 1996]). Preliminary tangible models of drug targets such as HIV protease [Olson 1997] and docking partners, such as the beta lactamase/BL inhibitory protein [Strynadka 1996] have already been produced. They have proven useful in understanding the nature of protein interactions, and as easily conveyed aids in collaborative discussions with other scientists. Enhancing the functionality of such models with improved representations and augmented reality approaches will expand our ability to explore complex systems. An example of such an exploration is in the biochemical initiation of blood coagulation — a process that can lead to thrombosis, the most common cause of death in western society. The Olson laboratory collaborates with biochemists and experimental molecular biologists in developing a molecular understanding of the protein interactions responsible for this process [Ruf 1999]. Although individual protein structures may be obtained experimentally by x-ray crystallography, more complex assemblies of proteins must be modeled from the component x-ray or homology built structures. Articulated models can embody constraints between rigid protein domains of the flexible, multi-domain protein factors, such as Tissue Factor, FactorVIIa and Factor X, which form the early initiation complex in the coagulation cascade. Such physical constraints can guide thinking about possible inter-molecular interactions and their implications in the overall assembly. As the complexity of these biological assemblies increases, the need for improved interface technology becomes more apparent.

The idea of developing customized "molecular Legos" to physically construct models of molecular assemblies would appeal to the playful creativity of many scientific researchers. Ultimately, embedding intelligence into such components will create an extensible tool for computing the implications of these new structures. Tangible models made in collaboration with the Tainer laboratory at TSRI on the structures of DNA repair complexes[Slupphaug 1996; Parikh 1997], have shown utility in understanding the surprising plasticity of shape in nucleic acid structures. Exploration of the shapes of RNA structures has come to the research fore in the last few years [Williamson 1994]. With the increasing detail coming from structural work on ribozymes and assemblies as large as the ribosome, it becomes feasible to envision plug-together models of RNA folding motifs (canonical helices, loops, and stems) to represent and explore the range of known and possible RNA structures. Williamson’s lab at TSRI uses NMR and crystallography to discover such motifs [Agalarov 2000]. Similarly, recent work by Bourne at SDSC, and others on the analysis of the protein data bank holdings, points toward a relatively small number (under 100) of protein “folding units” or motifs that are smaller than protein domains but larger than secondary structural elements [Shindyalov 2000]. Physical models of such units may be useful in designing feasible shapes for novel protein molecules.
Modular software for model creation and utilization. Although generated from the same data, the requirements of such physical molecular models will differ from those of the virtual 3D models represented on 2D computer displays. Integrating computational methods across three different disciplines, computational biology, CAD/CAM, and augmented reality, represents a real-world challenge in the development of a software environment. To achieve this objective, we will apply a strategy that is proving effective in computational biology. This approach is based on building computational components around an interpreted language that serves as a glue-layer to flexibly inter-connect different components at a very high level. We have chosen Python [Lutz 1996] because we feel it is the language that presents the largest set of desirable characteristics, including: Open-Source, interpreted, high level, Object-Oriented, platform independent, extensible, efficient, and with introspection capabilities. Because of these qualities, Python has been rapidly gaining acceptance in academic and government research communities and beyond.

We will utilize our extensible Python Molecular Viewing environment [Sanner 1998; Sanner 1999] to integrate a variety of molecular modeling approaches into exportable geometric representations for downstream fabrication and interaction. We have developed a generic 3D visualization component (DejaVu), which provides a high level interface to the OpenGL library with objects such as a Viewer which is in itself a simple, but fully functional geometry viewing application, providing control over most of OpenGL’s capabilities. DejaVu is written exclusively in Python which implies that: the exact same code runs on a PC, a Sun, or an SGI computer; any aspect of the viewer can be dynamically modified by an application or a user; and the powerful scripting capabilities of Python can be used to operate on a DejaVu Viewer itself or on the geometries that it displays.

In addition to the components written in Python, we also have access to existing C, C++, and Java libraries, including: MSMS [Sanner 1996] for the calculation of molecular surfaces, Babel [www.eyesopen.com/babel.html] for molecular file format conversions, atomic hybridization assignment, Gasteiger partial charge calculations, GLE [www.linas.org/gle] for the extrusion of arbitrary 2D shapes along an arbitrary 3D path, RAPID [Gottschalk 1996] for the fast detection of intersections between polygonal models.

We will incorporate relevant computational methods for surface decimation [Garland 1998], smoothing, and filtering [Guskov 1999]. Constructive solid geometry (CSG) capabilities will be required in order to graft connectors and sensor anchors on the geometry representing molecular models. We will integrate functionality developed at University of Utah to address these needs. The architecture and the concepts underlying the development of their Alpha_1 software are very similar to our approach, greatly simplifying the integration. We will extend our environment with drivers for the VR and AR devices from the HIT laboratory. These developments will include: drivers to feed back results of computations into these devices, such as calculated forces into a Phantom; drivers to read data from 3D tracking devices, to update our computational molecular modeling and data retrieval.

In order to generate the range of physical models and interfaces described above, we will utilize and extend the advanced molecular modeling environment that has been developed within the Molecular Graphics Laboratory at TSRI. We will address representational issues in the specification of molecular shape, chemical topology, chemical and physical properties, flexibility, and relative scale.

Multi-representational molecular geometries. This environment will be capable of creating molecular surfaces, extruded volumes, backbone ribbons, and atomic ball–and–stick molecular representations. It will also produce and integrate arbitrary labeling and property information for annotation and query. We have found that a key to the utility of physical models is producing a physical context for the meanings of molecular shapes. Thus our most successful models to date have combined shapes within a functional context, such as molecular docking (SOD dimers, RTEM–Blip), or the combination of two structural features, such as solvent-accessible molecular surface with an internal hollow extruded polypeptide backbone (HIV protease). The challenge will be to devise new, insightful representations that convey the physical meaning of the structures and interactions.
**Representing flexibility.** Biological molecules move at all scales, from bond vibrations, to side–chain motion, to loop flexibility, to domain re-arrangement. Thus, the representation of potential motion and flexibility of a tangible model must be designed for the scale and use of the model. Monolithic models can be useful in representing the unchanging shape of a molecule at a given scale. Averaging, or smoothing surfaces for side–chain mobility, for instance, can represent the complementarity of molecular interactions at coarser than atomic resolution. Our Harmony code [Duncan 1993] uses multi–resolution surface representations for this purpose. Experimental data and computational analysis can be used to describe the hierarchy of motions available to molecular structures [Kitao 1999]. Such information can be used to specify and represent a model of flexibility at a given level of detail. Such specifications could abstract characteristic motions to hinges and tethers between rigid components. Our molecular modeling tools will be used to identify and specify the geometry of these components. This analysis will feed the fabrication process model (see below).

**Integration of molecular simulations with physical model manipulation.** Unlike the macroscopic objects commonly manipulated in industrial CAD/CAM assembly models, biological molecules interact with each other at significant distances relative to their sizes. We will integrate data from tracked models to create an intuitive interface to computational simulations of molecular properties and forces, such as the electrostatic attractions between docking molecules. Changing features will be displayed by visual, haptic, or auditory output through the augmented reality interfaces. Inexpensive analogue approaches such as placing magnets, velcro, or other materials to simulate physico–chemical properties will be explored as well.

## 2. Challenges and Approaches in the Design and Manufacture of Physical Models

Newly emerging automated layered fabrication machines have been embraced as having the potential to revolutionize the fabrication process. In layered fabrication (also called "rapid prototyping"), objects are built up as a fused stack of thin layers of material, finally resulting in a free form solid (or partly hollow) object. These layers are shaped from a polyhedral model of the object that is derived from the CAD representation.

In collaboration with researchers at the University of North Carolina and TSRI, the Geometric Design and Computation Research Group at the University of Utah produced the first automated fabrication of a molecular model in 1991. These grapefruit–sized models of the molecular surfaces of two Superoxide Dismutase monomers were among the first natural objects to be made by layered fabrication, and demonstrated how the antioxidant proteins packed together in a dimer.

Over the past ten years we have made over a dozen physical molecular models to investigate various representational issues. The layered fabrication has been done by the Telemanufacturing Facility at the San Diego Supercomputer Center, by Allied–Signal Corporation, or by Design Concepts in 3D, a manufacturing service bureau. These experiments have added new representational capabilities to the molecular modeling repertoire, including...
multiple tangible models of the same molecule each at a different level of representational detail, ribbon–style protein backbones, solid partially transparent models with embedded hollow tubes that have been colored to represent the molecular skeleton, and a model that embeds internal atomic bonds by over–curing the polymer to locally darken it.

Potential layered fabrication processes for tangible molecular models include:

3D Systems: SLA – the stereolithography process that cures resin with a laser. Because of the chemical properties of the resins and the solvents required for cleaning, the process tends to be used only by commercial facilities.

Helisys: Laminated Object Manufacturing (LOM) – uses a relatively high powered laser to cut the outline of the part being produced on a thin sheet of paper. The parts have the look and feel of carved wood. It is not good for thin wall or hollow parts.

Stratasys: Fused Deposition Model (FDM) – extrudes either a molten ABS plastic or wax filament to build up a layered solid model by fusing each new layer to layer preceding it. The parts are opaque, and it is relatively good at prototyping thin walled parts.

Z–Corp: A new 3D printing process suitable for office environments, that uses a sugar water mixture applied with an ink–jet type print head to bind powdered starch. The material is opaque and the surface finish is relatively rough, but the process is fast and the materials are inexpensive.

Objet Geometries: Quadsra (not yet on the market). This system uses an inkjet type printer to deposit thin layers of a UV curable resin. Each resin layer is cured by UV lamps as it is placed. The models produced have similar physical properties to stereolithography parts, but it is a much faster process.

Research in manufacturing automation has for the most part been focused on mechanical shapes that consist largely of planar surfaces with some holes, or on convex sculptured shapes of low curvature. However, designing and producing parts or molds for complex objects, such as physical molecular models, requires significant expertise, and its automation will present significant research challenges in fabrication process planning and templating.

The Geometric Design and Computation Research Group at the University of Utah has been in the forefront on research in computer graphics, geometric modeling, and manufacturing. Their role in this project will be crucial in translating computational molecular models into realizable physical objects. The Advanced Manufacturing Laboratory at Utah is one of the few University sites to have incorporated both a layered fabrication system, currently a Stratasys FDM, and conventional CNC mills and lathes, together with sinking and wire Electro–Discharge Machining centers and plastic injection molding equipment. We will design appropriate sequences of these operations for each of the molecular tangible interfaces to be produced.

A significant research challenge is to develop geometric computation analysis techniques, process planning capabilities, and manufacturing process models, for broad classes of molecular models. Exploring novel types of tangible models for scientific visualization can lead to new insights in characterizing types of fabrication processes and geometric analyses needed. Fabrication methodologies will depend both on the purpose and the number of physical models needed. For each molecular model, one may want: multiple visualizations (surface, skeleton, labels, textures, sensor querying, etc.) at multiple resolutions of geometric detail and complexity at multiple scales (of structure and articulation). Moreover, unlike most industrial applications, many of our pieces will be used directly instead of for mold–making, thus making durability, coloring, and first–copy cost more important.
Over the course of this project, we expect that there will be a stream of new commercially available automated fabrication technologies, some of which will sit at deskside. With the development of the proposed computational geometric analysis tools, the scientist could design, fabricate and interact with many classes of tangible models almost immediately. However, for more complex models, such as those requiring multiple components and fabrication steps, new process plans must be developed. Research herein must develop techniques that simplify and automate the process of moving from molecular computer model to multiprocess fabrication model. Further, the same must be done for the various automated fabrication and traditional fabrication processes that are used. Finally, the cost of creating small lot tangible molecule models can be lowered using the results of the research we propose, so they can be an integral part of molecular research and education.

In order for custom tangible models encoding complex scientific information to be produced on demand, we will perform research in several areas of representations and geometric computing, including to:

1. develop algorithms to automatically transform from scientific model to fabrication model to tangible model,
2. develop algorithms to automatically generate templates for fabrication of tangible models requiring multiple passes and processes,
3. characterize and develop automatic analysis techniques to determine which classes of tangible models: can be automatically generated and embedded “deskside; or can be automatically transformed to the fabrication representation, specified and planned, but require more sophisticated equipment or processes than locally and immediately available to the user, and so could be done with “web” brokering and short turn–around; or require fabrication technology outside the scope of this research effort and hence will require significant manual specification and expert process design.

Many of our early experiments have used the stereolithography process and have resulted in solid models which have required hand finishing to obtain transparency and a smooth surface. Within the course of this research, we will explore the potential of other fabrication processes to create tangible models with different characteristics for embedding information. Potential tangible models include:

1. opaque smooth objects, such as those that result from the LOM process. Such objects resemble wood, and would work well with a computer projector system to show force fields and molecular interactions in color directly on the objects,
2. molecules with thin clear resin walls with an opaque skeleton,
3. models enhanced with additional opaque substructures,
4. models with embedded sensors, indicator lights, or motion actuators,
5. models that explore representing and fabricating hinge and flex regions within the molecules
6. creating assembled compound molecules, with jointed, flexible behaviors
7. models that can be put together or taken apart to show internal detail or to show how molecular components assemble or rearrange.
8. Replicated models for reusable components, symmetrical assemblies, or broader dissemination
9. Molecular "legos" for exploring modular structures.

The idea of embedding sensors and actuators gives rise to intriguing applications. For example, when illuminated by a laser pointer, we imagine embedded sensors notifying a computational database, which, in turn, gives an audio self-describing output. It could report atom or residue identification, or give statistics on sequence conservation, or specific functional or struc-
tural characteristics. It could be used as a driver to the immersive image rendering component for labeling, coloring, or further characterizing the molecule, as the human manipulates the tangible model.

The integration of embedded fixtures and articulated features into the tangible molecular model requires that the fabrication model be significantly richer than the standard rapid prototyping input model. Desirable automatic operators for tangible molecular model fabrication include:

1. Offset — if the tangible model is to be used to get a feel for docking and force field manipulations, it is important that the potential mating surfaces be reasonably accurate. Thus, offsetting such surfaces inward to account for material wall thickness or surface finishing material thickness is required.

2. Bond and virtual bond radius specification — the molecular backbone can be represented by cylinders or extrusion paths of arbitrary shape and diameter. However, if they are too thick, they may self intersect or generate cusps. If they are too thin, they may not be able to accommodate embedded lights, wires, or other features.

3. Clearance distances of bonds to molecular surface. It is desirable to control the relationships of the various physical representations so that surfaces are not unintentionally pierced, or structures are not too thin and fragile. The system should automatically check that the distances between hollow tubes and the surface (and other sections of the skeleton) are larger than some preset requirement. Otherwise, it may not be possible to fabricate the molecule.

4. Generate desired positions for attaching sensors or placing external fiducial markers given constraints of the molecular geometry and intended use.

5. Generate armatures and other fixtures for holding sensors or lights in place.

6. Generate mold design by analyzing geometry, nesting sequences, identifying the parting planes for the different mold sections, and verifying how they assemble.

7. Fastener design and assembly to automatically design fasteners for models that are to be taken apart and reassembled.

8. Flex analysis to simulate the constraints based upon molecular connectivity and conformational degrees of freedom.

9. Flexible connector design to construct physical constraints on the relative positions of component parts.

Some of the proposed models will be considerably more complex to fabricate than previous models, because of the internal details and added functionality. If multiple copies are needed, molds can be used to reduce cost. Since the geometry of a complex molecular model does not allow a simple two part rigid mold, a silicone rubber mold that permits the part to be removed must be used. Such a mold is made as a transfer mold from an automated fabrication produced positive. Geometric algorithms must be available to compute suitable parting lines and perhaps to add vent features to the mold at each local, orientation dependent, high spot, in addition to a gate feature to allow the resin to be poured into the mold. A hollow clear surface can be made by spin casting resin instead of pouring. This eliminates the need for vent features. The skeleton can be made using FDM, and then embedded into the hollow molecular boundary surface model. Automated fabrication can be used to make models that can be taken apart, either to show internal detail or to show how molecules may transform. This requires geometric computation algorithms to decompose the original single molecular model into the assembly model. Further, it requires augmenting the fabrication
model with assembling fasteners, designed automatically by algorithms based on design rules. For multiple copies, this will require a sequence of molding operations to produce parts that nest together. Finally, moving towards automatic fabrication requires that, once given the scientific model and user specifications for the tangible model, the fabrication module must augment the fabrication geometry as appropriate, decompose model realization into a proper sequence of stages (perhaps only one), and ensure that it is as straightforward a plan as possible.

Developing and implementing the rules and computational methodology for the geometric and process analyses for freeform, complex tangible molecular models for scientific visualization is significant research challenge that will advance the state of the art in computer aided design and manufacture.

3. CHALLENGES AND APPROACHES IN COUPLING PHYSICAL OBJECTS INTO HUMAN–COMPUTER INTERFACES

The "transparent computer interface" is a goal driving a growing activity in such fields as immersive displays, object tracking, haptics and numerous other technologies for virtual reality. An important objective of this development is the creation of a sense of user presence during a primarily computational interaction. Much of the activity has been focused on far and mid–field tasks, such as motion simulation, navigation, and "walkthroughs", where the user is immersed in the simulated environment.

Near–field activities, including such traditional human tasks as toolmaking, model building and close inspection have been advanced through the use of 3D computer graphics for over 40 years. Much of the recent work in near–field object presence [Barfield 1993] and interaction has focused on improved rendering, stereoscopy (e.g. “holographic workstation”, virtual workbench”), force feedback, and 6D object manipulation techniques [Poupyrev 1997]. Augmented Reality has also been brought to bear on near–field interactions in such applications as diagnostic medicine and surgical planning. More recently, the use of real–world proxies, or physical icons ("phicons") has begun to be explored in AR applications to increase the illusion of real interaction [Ishii 1997; Billinghurst 1999a,b; Underkoffler 1999]. Brooks has identified the area of haptics as being critical to the sense of presence for near–field activities [Brooks, 1999].

The Human Interface Technology (HIT) Lab at the University of Washington has contributed advances in several core virtual environment technologies that are relevant to the objectives of this project. Researchers at UW will further advance and integrate these technologies around the problem of a tangible object interface for structural molecular biology. A unique emphasis in the HIT Lab has been on collaborative AR environments, in which multiple participants experience and interact with shared virtual objects [Billinghurst 1996, 1997, 1998]. Participants in these environments may be co–located in the same physical space, or some may be connected as remote participants, demonstrated at the 1999 SIGGRAPH conference as part of the Emerging Technologies area [Billinghurst, 1999b]. In this Shared Space demonstration, head–mounted displays with attached cameras were used in video see–through mode, with computer graphics overlaid on a video view of the real world. Precise registration between real–world objects and 3D graphical objects, animations and live video streams was achieved using computer vision techniques for parsing the position, orientation, and identity of physical fiducial markers. In this system the physical markers provided a compelling interface for virtual object manipulation, and required essentially no training to be used successfully.

Three primary research issues will be addressed by this aspect of the project:

1. Tightly coupled real–space registration of AR computer graphics with the tangible models.
2. Multi–modality augmented reality interaction support (visual, auditory, and haptic).
3. Embedded intelligent tangible interfaces.
Spatial Tracking and Registration of physical models. Tightly-coupled registration of computer-generated graphics with the physical molecular models will require fast, accurate and robust spatial tracking of each model in both position and orientation (6 degrees of freedom). The HIT Lab has a variety of spatial tracking technologies in house, as well as new methods in development. We will explore the efficacy of several candidate approaches for the tangible molecular models applications developed under this effort.

Our initial approach will be a simple extension of the computer vision–based object registration methods used in our Shared Space applications [Billinghurst 1996, 1997a, 1998]. Small fiducial markers which support 6D tracking and object identification will be attached to the models (either by short rods, or in a redundant fashion on the surface) such that unconstrained spatial manipulation is afforded. Computer graphic enhancements to the model surfaces and the space surrounding them will be superimposed appropriately from the tracking data. Tight registration between the physical objects and virtual overlays is assured by the video compositing method, and the proper placement and optical parameters of the HMD–mounted video camera assures seamless interaction from the user’s perspective. For this application users will easily adapt to the 200 msec processing delay as they manipulate the models. Since this application is "near–field" object manipulation, our current HMD–based technology will need to be adapted for stereoscopic display.

We will also explore methods using traditional spatial tracking technologies (e.g., electro–magnetic trackers, ultrasonic triangulation, and direct mechanical coupling), but our long–term objective will be to incorporate untethered, unobtrusive, and portable object tracking methods that can be easily applied to newly fabricated models. Since the models are generated from known geometry and the identity of the objects is known at run time (or can be easily derived from visual or ultrasonic tags), computer–vision based object tracking methods (based on object geometry, rather than marker placement) are feasible. We will extend the computer vision methods used for hand–tracking in our Hi–SPACE shared desktop technology to parse and track the geometric features of each object.

Multi–modality integration. In addition to our primary focus on visual augmentation of the physical models, auditory and haptic augmentation will be explored. These modalities will be especially salient for the visually–impaired user community. We will explore the utility of each interface modality (singly and in combination) for this
domain, including information query, multiple object interaction, force display, and interactive visualization of dynamic spatially distributed attributes.

Sonification [Kramer, 1994a,b] of the relevant molecular surface and field parameters will require exploration of the efficacy of various mappings between molecular data and sound parameters. In addition, real-time spatialization of individual sound sources (perceptually co-located with the physical models) will be explored.

Molecular interactions have long been a fruitful area for haptic research [Ouh–Young 1988]. In our research, haptic interaction will be explored using the Phantom [Massie 1994] force feedback devices in use at the HIT Lab for other research applications [Anderson 1998; Berkeley 1999]. Tangible models automatically provide accurately registered hard-surface interactions, leaving only the softer, long-range forces to be generated computationally. As a first step, we will compute an electrostatic field around a molecule (in the form of a 3D grid of vector values), and use an "electron probe" at the tip of the PhanTOM to explore the field around the molecular model. The model may be stationary on a table or held in one hand and spatially tracked, while the other hand probes the field in the surrounding space. A second research activity will entail attaching the models themselves to the force–feedback devices and exploring the changing electrostatic (or other) fields and interactions between the models as they change relative positions.

In addition to making molecular biology more accessible to people with sensory disabilities, real-time integration of all three display modalities (visual, auditory and haptic), with each encoding a different set of molecular interaction parameters, will provide a unique tool for both education and research.

**Object–embedded interface.** We envision the models themselves as the focal point of the user’s interface with the data. Researchers and students will pick up models, query them, and explore their interactions with other models. The objective here will be to create and sustain the perception of direct interaction with (and among) the tangible molecular models. Specific embedded”interface approaches we will explore include:

1. model–embedded voice pickup and apparent reply localization
2. embedded ultrasonic, IR or RF emitters and receivers for tracking model position and orientation (relative to each other and to the user’s viewpoint)
3. surface touch sensing methods, such as ribbon microphone pickup of spark gap pulse epochs (emitted, for example, by a fingertip thimble source)
4. multi–sensory embedded output displays modulating model features such as light, sound, vibration, or surface temperature

**Evaluation of Tangible Model Interaction.** Our proposed research and technology development is based on the premise that the multi–modality interaction provided by this new molecular exploration paradigm will enhance research and education. Although the impact on research outcomes may be difficult to assess in the short term, the “goodness” of our tangible interface solutions and, in particular, their impact on learning will be systematically evaluated as part of this ITR project. These evaluation activities will be conducted during project years 3–5.

HIT Lab researchers have established a track record in assessing the effects of novel interface technologies on a number of goodness measures, including run–time performance [Pouppyrev, 1997], as well as content and concept learning. These latter educational evaluation efforts have included: summer programs with the Pacific Science Center in which elementary school students designed and constructed their own virtual worlds [Furness, 1997]; a multi–State program to bring VR technology to the schools via Virtual Reality Roving Vehicles (VRRVs); training programs to train school teachers in methods for encapsulating lessons into VR environments [Osberg, 1997]; and an ongoing NSF–sponsored project to
incorporate and assess environmental process learning into immersive exploration of a large multivariate database of the Puget Sound region.

At the level of graduate and professional training, the HIT Lab has been active in developing new methods for evaluating the efficacy of virtual environments, particularly in the area of simulation for biomedical training [Weghorst, 1998; Oppenheimer, 1999].

Building upon this base of evaluation experience and methods already in place at the HIT Lab [Winn 1997a and 1997b, Furness 1997, Weghorst 1998], we will evaluate the efficacy of the various tangible model interaction methods for teaching content and concepts appropriate to high school and college level students. Age-appropriate learning objectives, “goodness” measures, and evaluation methods will be developed via close collaboration between the HIT Lab’s evaluation group and the TSRI project team and their educational partners.

As our technology development efforts come to fruition we will determine the most appropriate evaluation questions and methods. Our evaluations will assess the effectiveness of the most promising interaction technologies, both relative to each other and in comparison with more traditional teaching methods. Students and teachers will be recruited as subjects from the HIT Lab’s existing broad pool in the Seattle area. As part of this evaluation effort we will assess the role of individual learning style in determining technique efficacy, and we will look for possible emergent learning strategies that arise when free-form multi-modality exploration is afforded.

We are particularly interested in assessing the educational impact of these tangible interface technologies on visually-impaired students. Subjects for this aspect of the evaluation activity will be recruited from three sources: (1) the HIT Lab’s subject pool for its ongoing research on low vision applications of the VRD (virtual retinal display) technology; (2) high school students participating in the University of Washington’s DO–IT program for disabled students (also sponsored by NSF and supported by HIT Lab personnel); and (3) interviews with blind scientist collaborator (see letters). Input from this potential user community will be an integral part of our development effort.

**TIME–LINE**

The goals of this proposal are ambitious, but do-able with sufficient coordination between the collaborators. The three groups share a unified vision of specific, desirable, and achievable milestones to gauge our progress. The overarching vision of creating advanced "information appliances" [Norman 1998] for molecular biology points toward the automated fabrication of manipulable models that embody physical analogue characteristics as well as embed and project digital intelligence.

The trend of our progress will go from the physically simple to the more complex, with an increasing degree of model autonomy, model to model interaction, and simplicity of use and accessibility. Specific fabrication and AR implementations will be based upon the requirements of the molecular modeling and educational activities. We anticipate producing 10–20 usable models each year (with many more prototypes tried), ramping up fabrication transfer technologies for inexpensive replication and larger lot sizes in the later years. As the tangible models and interfaces are produced they will be shared, as appropriate, with our collaborators for research and educational purposes (see letters).

**Year 1**

**Utah.** Establish access to variety of automatic fabrication processes. Design an exchange data-structure to integrate the essentials of Scripps’ molecular models into the Alpha_1 software environment. Start developing design synthesis algorithms for tangible model elements. Design multiprocess fabrication templates for models with embedded annotation structures. Produce initial models, get users’ feedback.

**Washington.** Extend the Shared Space techniques to track and augment representative tangible molecular models. Explore visual data presentation, haptic feedback, and sonification options for candidate molecular biology features and data.
**TSRI.** Demonstrate geometric vocabulary for representing various molecular biological structures, including atomic and low resolution from x-ray, NMR, electron microscopy, or other sources. Include protein and nucleic acid representations, such as viral subunits, enzymes, RNA and DNA structures. Demonstrate link to audio input and output for bioinformatic data retrieval in collaboration with blind researcher in the Forest laboratory (see letter). Generate 3D molecular field data for haptic feedback experiments.

**Year 2**

**Utah.** Design initial mold–based replication strategies. Create molds for multiple copies of some initial molecular models. Develop design algorithms for rigid models with richer embedded structures such as light cable, or optical sensors. Begin designing and making models with flexible components.

**Washington.** Undertake initial design study and development of model I/O and communications methods. Iterate on and refine the most promising techniques and concepts from Year 1. Integrate dynamic real–time molecular model attribute generation with augmented sensory display.

**TSRI.** Demonstrate expanded molecular shape vocabulary, including mating molecular surfaces for physical docking. Make molecule docking pairs, to model viral interfaces, protein complexes, protein–DNA and RNA interactions. Demonstrate analysis of molecular flexibility and incorporation into geometric description. Make prototype flexible models of antibodies, blood coagulation factors using domain models and custom or standard connectors. Demonstrate integration of physico–chemical properties onto model with analogue attachments such as magnets, velcro. Demonstrate interactive updating of changing properties, such as intermolecular forces, at tracking speeds.

**Year 3**

**Utah.** Develop design algorithms for modular molecular model components and fabricate prototypes. Experiment with design alternatives for attaching multiple materials, including elastomers, in modular structures. Test for durability. Establish design algorithms for flexible components to model realistic molecular kinematics. Develop mold design algorithms for snap–together models with transparent shells and embedded components.


**TSRI.** Enhance flexibility and shape vocabulary. Demonstrate enhanced modeling tools to describe molecular assemblies and flexibility to integrate with the CAD/CAM analysis. Demonstrate flexible models of expanding viral capsids, motions in allosteric protein oligomers, such as hemoglobin, plasticity of nucleic acid chains. Demonstrate computation and visualization of interactive molecular mechanics, dynamics and animated flow fields, in coordination with molecule model tracking. Demonstrate audio output reporting interactions between two molecular models. Produce prototype of multi–modal interaction in molecular docking scenario, with factors from the blood coagulation cascade, viral assembly and protein–nucleic acid interaction in collaboration with the Johnson and Williamson labs. Begin pilot placement into formal and informal science education settings through collaborations with Stewart at San Diego State University, Dahms at CSUPERB biotechnology education center, and Kirsch at the Fleet Science Center (see letters).

**Year 4**

**Utah.** Develop initial support for embedded active components in molecular models. Advance modular component molecular design algorithms. Incorporate feedback from use of tangible models in education collaboration and research.
Washington. Demonstrate prototypes of tangible object–embedded interface. Port desktop simulations to prototype interaction objects. Develop demonstration applications for education, research and public awareness. Undertake human factors analysis of goodness of modality–specific displays and inter-modality effects (e.g., visual capture, synesthesia).

TSRI. Demonstrate enhanced molecular modeling environment to aid in the design of "molecular Lego" components. Design and produce custom part sets for protein substructure components (with Phil Bourne, SDSC), and RNA substructure components (with Jamie Williamson, TSRI), to enable physical modeling of novel structures. Demonstrate enhanced computational steering, query, and rendering software to integrate with advanced AR approaches. Demonstrate keyboard–free, multi–modal (visual, audio, haptic) molecular modeling for drug–target interactions in systems such as HIV protease or integrate, and for protein assembly in systems such as viral capsids or pilin filament formation. Expand placement into formal and informal educational settings in collaboration with San Diego State University, CSUPERB, Fleet Science Center, and Byron Rubin, sculptor. Report informal evaluations to formal evaluation effort at Washington.

Year 5

Utah. Enhance software to support fabrication of more elaborate, more ambitious, tangible models of more molecules. Further develop snap–together high–level "Molecular LEGO" components.

Washington. Produce and demonstrate engineering prototypes of object–embedded interface. Evaluate human subjects for target applications (e.g., molecular biology education). Refine prototype technologies and applications for transfer to industrial, educational, and research concerns.

TSRI. Demonstrate GUI that consolidates and integrates modeling software with CAD/CAM analysis for deskside fabrication. Analyze successes and failures in representation, fabrication, and utilization. Design effective, replicable applications for research, collaboration, and education. The molecular systems to be determined based upon experiences to this point. Work with collaborators in formal and informal education (see Year 4) to optimize for their respective needs. Continue to develop custom research and collaboration applications.

Results from Prior NSF Support Most Closely Related to This Proposal

A. NSF: EIA8920219, $31,000,000, 2/1991 to 1/2002

B. NSF Science and Technology Center for Computer Graphics and Scientific Visualization.
Project Director: Richard F. Riesenfeld (Utah), co–PIs: Alan Barr (Caltech), Henry Fuchs (UNC–CH), Donald Greenberg (Cornell), Andries van Dam (Brown).

C. Significant research on topics related to this proposal have been done under STC auspices, including the first tangible molecular model made by an automated fabrication process (1991) in collaboration with UNC and TSRI, and exploring flexible transfer molds to replicate tangible models. Shape analysis and constraint research is important for determining whether proposed affordances can fit properly and for issues related to flexible molds. Results in feature–based modeling can be used as a basis for creating molecular fabrication features for joints. Process planning and fabrication results directly relate to creating tangible models.


E. F. Not Applicable