Published in final edited form as:

J Mol Biol. 2018 October 19; 430(21): 3997–4012. doi:10.1016/j.jmb.2018.07.009.

Perspectives on structural molecular biology visualization: from past to present

Arthur J. Olson

Department of Integrative Structural and Computational Biology The Scripps Research Institute

Abstract

Visualization has been a key technology in the progress of structural molecular biology for as long as the field has existed. This perspective describes the nature of the visualization process in structural studies, how it has evolved over the years, and its relationship to the changes in technology that have supported and driven it. It focuses on how technical advances have changed the way we look at and interact with molecular structure, and how structural biology has fostered and challenged that technology.

Introduction

Visualization has been an enabling technology in structural molecular biology for as long as the field has existed. The process of visualization, itself, implies human involvement since its ultimate impact is insight. In general, visualization can aid in three functions that are essential components of the scientific enterprise and fundamental to our understanding of the “unseen” world of structural molecular biology, synthesis, analysis and communication:

Synthesis involves the process of creating a model to embody or integrate experimental observation or to develop and encompass a structural hypothesis.

Analysis entails the examination and exploration of either data or models to explain observed phenomena, derive new hypotheses, or suggest new experiments.

Communication involves the imparting or sharing of information and knowledge about the systems or methods under study.

These activities pre-date the advent of computer technology, as do the use of visualization tools in their service. Moreover, the evolution of molecular graphics and the visualization of biomolecular structure demonstrate a significant symbiosis between science and the technology that enables it. There are many good reviews of the history of molecular graphics and illustration^{2; 3; 4} We focus here on the how technical advances have changed the way we look at and interact with molecular structure. This perspective is not meant to be exhaustive, but rather one person’s perspective on a continually evolving field.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Looking back on the development of what originally was termed “molecular graphics” we see a history of invention and innovation driven by a need to comprehend the complex nature of the molecular world. Graphics and visualization tools have been critical in enabling and catalyzing the remarkable progress of structural molecular biology from the past half century to today. Conversely over this time, the field of structural biology has helped to drive development and commercialization of computer graphics and visualization technologies.

The Past

Origins—Physical models were the earliest interactive three dimensional molecular visualization tools. Examples such as the CPK (Corey Pauling Koltun)⁵ and Dreiding⁶ models built upon earlier work going back to the mid-19th century, when chemists such as Hoffmann and Kekule² built wooden ball-and-stick or wire models demonstrating element combining ratios or bond connectivity of simple chemical compounds. As X-ray crystallography began to produce three-dimensional molecular structures during the first half of last century, the need to represent them in accurate and comprehensible form increased the use of these physical models. In the 1950s they began to be used to model biological molecules. Pauling, who helped develop the CPK model and adopted the use of physical models throughout his career, used the simple process of folding a paper with a drawing of a polypeptide chain, coupled with his knowledge of the hydrogen bonding potential of amino acids to deduce the alpha helix as a folding motif for proteins⁷. Likewise, Watson and Crick, playing with brass models of nucleotide base structure, were able to explore geometric combinations to explain Franklin’s fiber diffraction data and synthesize a model of the structure of DNA that revealed its genetic function. In a real sense both examples demonstrate how physical visualizations can promote exploratory interaction and serve as “analogue computers” leading to new insights^{8,9}.

In the 1960’s computer technology became a critical catalyst driving progress in crystallography and structural molecular biology. Initially the computational power was dedicated to deducing electron density maps from the x-ray diffraction data. Typically these maps were displayed using computer line printers to produce pages of two-dimensional arrays of density values ranging from 0 to 9. Each page was hand-contoured by isovalue and transferred onto balsa wood sheets which were cut out and stacked to make a physical representation of the electron density^{10; 11} (see Figure 1). As computer-driven X-Y pen plotters became available the electron density contours could be drawn out automatically on transparent plastic sheets that could be stacked with appropriate spacers to give a direct 3D spatial view of the electron density map. Also at that time pen-plotters were being used to draw small molecule structures. The ORTEP (Oak Ridge Thermal Ellipsoid Plot) program by Carrol Johnson^{12; 13} was the most widely used and cited computer program of its day. It enabled the accurate plotting of stereoscopic molecular structures and introduced the concept of thermal ellipsoids, which visualized the anisotropic temperature factors of each atom in the small molecule structure. However, the creation of 3D molecular visualizations of protein structures from electron density maps relied upon the building of physical models, which was tedious and time-consuming. An ingenious apparatus devised by Fred Richards at Yale, utilized stacked transparent sheets with plotted electron density contours and brass

Kendrew model parts to be screwed together to build polypeptide structure¹⁴. Model support rods were placed in front of a large half-silvered mirror and with the appropriate lighting the Kendrew parts, as seen in the mirror could be built into the electron density as seen through the mirror (Figure 2). Once the physical model was built, the 3D coordinates of each atom in the model had to be measured and recorded (X-Y by projection of sight line, and z with a plumb-bob).

In the mid-1960's, using a one-off oscilloscope-like line drawing system at MIT, Cyrus Levinthal and Robert Langridge were able to display and rotate models of lysozyme and myoglobin created from their known 3D coordinates¹⁵. This event marked the beginning of interactive computer-generated molecular graphics. It took several years before the technology matured to a commercial stage, but from those early applications, interactive molecular graphics was seen as a driving application for the technology.

As new protein crystal structures were solved, the opportunity to communicate structural molecular biology opened up. Probably the most influential book on protein structure from that decade was "The Structure and Action of Proteins" by Dickerson and Geis¹⁶, published in 1969. The artistic input that illustrator Irving Geis used to represent molecular structure and protein folds in that book and in a Scientific American article¹⁷ brought protein structures to life and made them accessible to a broad audience. Geis's drawings also served as an inspiration for the now ubiquitous ribbon diagrams of protein topology made popular by Jane Richardson in the following decade¹⁸. The evolution of such molecular illustration work is covered by Goodsell and Jenkins [This issue]. It should be noted that much of what was originally developed for illustration and communication purposes, has been incorporated in scientific molecular visualization tools. Initially there was a strong distinction between "working graphics," where rapid feedback and high interactivity are paramount, and "publication graphics" where clarity of presentation is key. As technology has improved over the decades, the distinction between these modes has blurred somewhat.

In the early 1970s interactive monochrome CRT vector displays became commercially available. Several small companies such as Vector General, Evans&Sutherland, and Adage started up to produce and market these devices, initially for military and aviation applications. These monochrome displays used move and draw operations on an electron beam to create lines and dots on a phosphor screen at interactive frame rates. Remarkably the beam could address the equivalent of 8000×8000 pixels - a resolution that is only now being approached in color 8K displays. Depth cueing (dimming lines farther away) and interactive rotation enhanced three-dimensional perception. Meanwhile, with the advent of more powerful and available computers, the field of biomolecular crystallography expanded, and by the mid 1970s several groups began to develop software to move the protein model-building task from the physical Richards Box onto interactive computer displays. Programs, at the time termed "Electronic Richards Boxes," such as Bilder¹⁹, Frodo²⁰, MMSX²¹ and Grip²² were developed by crystallographers and computer scientists to interactively build polypeptide structures into the 3D computed iso-contoured electron density maps on the CRT display (Figure 3). These codes were able to speed up the process of protein model building several fold and, in addition, automatically produce more reliably captured atomic coordinates. The commercial availability of the display technology along with the purpose-

built electron density fitting programs brought more scientists into the field of protein crystallography and grew the demand for interactive computer display devices. Thus, technical advances in display hardware and molecular software not only greatly facilitated rapid growth and progress in structural molecular biology, but also provided an expanding market for those early and very expensive graphics devices. In fact molecular graphics was a “killer app” that helped sustain and build the nascent computer graphics industry²³. By the early 1980s there were close to 100 labs around the world that had purchased interactive display systems for biomolecular graphics²⁴.

Evolution—The 1980s saw an explosion of applications and innovation in structural biology and molecular graphics. The enhanced ability to build 3D biomolecular models on the computer drove new applications for the synthesis, analysis and communication of these previously unimagined protein structures. One innovative and very useful structure-based analytical representation developed during this time was the solvent-based molecular surface, originally conceived by Lee and Richards in 1973²⁵. This surface describes the interface between a protein’s atomic structure with its surrounding water solvent, and gives a quantitative measure of effective surface area for these structures and their interactions. Michael Connolly then at UC Berkeley/UCSF, developed the first program to compute and display the “solvent excluded” molecular surface²⁶. With the line-drawing graphics systems available at the time, the surface was visualized by placing dots on it. This representation had the benefit of allowing the underlying bond structure to be seen in the context of the surface, and became a standard molecular visualization technique (Figure 4). Also during this time, molecular dynamics trajectories of a protein’s motion added the time dimension to the static view of protein structure. While these simulations took hours or days to compute, the resulting protein trajectories could be re-played and viewed interactively.

With the fields of structural biology and computer graphics co-evolving over this period, the representation and communication of biomolecular structures and their function developed rapidly. Color line-drawing displays became available in the early 1980s and were quickly adopted for molecular graphics. Color added an extra visual dimension and enabled clear visual discrimination of atom-types, chain identities and multiple molecules in complexes. During this time color raster displays became more widely available. These devices utilized a frame buffer to build up a picture using a raster scan-line like the television, and an early use was the display of shaded surfaces of large molecular structures²⁷. The resolution of the available color displays was as high as 512×512 pixels (Figure 5). However for molecular graphics the color raster displays were used primarily to produce still images for presentation since interactive rotation and display of a protein with hundreds to thousands of atoms represented as spheres or surfaces was beyond computational capabilities at the time. Raster-based graphics enabled more “realistic” molecular renderings by simulating the effects of light on solid surfaces. Shaded and highlighted surfaces provided additional visual cues about shape, and could approach the appearance of plastic CPK physical models²⁸.

Initially molecular graphics programs were developed for specific functions, such as electron density fitting or display of dynamics trajectories. As an increasing number of protein structures became available, more general molecular graphics programs were developed that enabled exploration, analysis and communication. Three of the earliest

general purpose programs were MIDAS²⁹ HYDRA³⁰ and GRAMPS/GRANNY^{31;32}. These programs went beyond electron density fitting to provide general graphics functionality for examining molecular structure and properties. GRAMPS was developed as an interpreted graphics language that had interactive features built into it that enabled the composition of animations. It was used to produce a number of molecular movies throughout the 1980s that detailed the structures of viruses, antibodies, and enzymes. These 16mm films were shot from the computer display onto 16mm movie film with an animation camera triggered by a solenoid to advance each frame. In one instance, GRAMPS was used to produce animated segments for an OmniMAX film shown at Disney's EPCOT Center³³. The molecular segments depicted traveling along the major groove of DNA (Figure 6) and the assembly of viral subunits as viewed from inside. The segment was produced for large format IMAX film (equivalent of 3 frames of 70mm movie film) projected onto the OminMax spherical dome. This meant that the line-drawn images had to be computationally distorted³⁴. They were then written out frame-by-frame onto 6250bpi magnetic tapes and sent to a film recorder. The two-minute segment required over 60 nine-inch diameter tapes (roughly 6GB).

Over the decade of the 1980s the remarkable evolution of computer hardware - the advent of microprocessors, very large scale integration (VLSI), personal computers, and scientific workstations - increased the accessibility of molecular graphics. Initially structural researchers themselves wrote the molecular modeling software. By the middle of the decade, the demand was such that several commercial companies such as Biosym, Molecular Design and Tripos had been established to market molecular graphics and modeling software. By the end of the decade structural scientists in academic and industrial research settings had a wide variety of use-tested hardware and software platforms with which to do molecular modeling and visualization.

Also during these early decades computer graphics became a separate field of study within the discipline of computer science. The Association for Computer Machinery (ACM) initiated a special interest group devoted to computer graphics and interaction (SIGGRAPH) which had its first Annual Conference in 1974 with 600 people in attendance³⁵. In the 1990's SIGGRAPH Conferences had up to 45,000 attendees. During this decade the field of visualization, itself, became a distinct discipline, focusing more on the development of methods to apply in scientific visualization, information visualization and visual analytics. Over this time novel graphics and interactive technologies were developed with a number of them driven by biological problems. New interactively capable raster-based displays supplanted the line drawing displays, and graphics algorithms focused on photorealism, geometric modeling, volume rendering, and texture mapping. Along with the more powerful graphics systems, new scientific visualization environments such as the Application Visualization System (AVS)³⁶ were developed and marketed that used a graphical programming paradigm with data flow networks of computational modules to enable rapid implementation of custom scientific visualization applications.

The 1990s saw the rapid growth in the number of protein structures deposited in the protein data bank (PDB). By the end of the decade it contained over 10,000 structures. The development of programs such as Molscript³⁷ enabled the creation of illustrative shaded drawings of protein structures, including ribbon diagrams. Rasmol³⁸ created an easy and

accessible user interface to interactively view biomolecules on raster devices. Importantly, this decade saw the wide-use of the Internet, which made the ever-growing number protein structures more accessible and fostered the creation of web-based browser molecular graphics, such as Chime and Jmol³⁹ New modes of visualization and interaction were introduced and tested on molecular applications, including force-feedback haptic displays⁴⁰ data sonification, and virtual and augmented reality⁴¹. As with much of technology, the time between invention and broad utilization can take many years or even decades. One problem has been the traditional communication barriers between the computer science and scientific communities. Too often researchers in computer science develop tools and algorithms no real application in mind, while scientists view the details of their application as paramount and treat computer scientists as simply programmers. There are few, if any, conferences that are equally attended by both communities and publications and proceedings are rarely shared. In addition, not all tools and techniques fit the needs of a particular scientific discipline. Thus acceptance and usage of a given graphics technology or algorithm for a particular application domain, such as molecular visualization, is never guaranteed.

The Present

In the intervening years, over 130,000 protein structures have been deposited in the Protein Data Bank (<http://www.rcsb.org/stats/summary>) and we now have high definition color graphics displays in our pockets. Utilizing a growing arsenal of biophysical techniques such as electron microscopy and fluorescence light microscopy, we are generating, at accelerating rates, bio-structural data from experimental approaches that span from atoms to cells. Computational power and Graphics Processing Units (GPUs) now enable real time interactive display of billions of atoms, millions of proteins. GPUs are now considered “a sixth computer technology revolution”⁴². Currently available high-end GPUs have upwards of 5,000 processors and 24 GB of shared memory. While GPUs were originally developed for rendering images for 2D display devices with a mapping of pixels to processors, their utility is applicable to any process that can be broken down into many independent processes that can run in parallel. Today, applications that range from molecular dynamics to autonomous vehicles utilize the power of these processors. How have these advances in science and technology affected the still evolving world of molecular visualization and interaction? How has the commodification of 3D interactive computer graphics, and its broad adoption in society changed the way we look at molecules?

While in the early days of computer-based molecular graphics crystallographers and other molecular scientists developed their own software to visualize, explore and analyze the structures they were solving, today computer graphics and interactive techniques are being developed and utilized by a large number of different communities that range from scientific visualization to entertainment film and game production. Academic and commercial development has spread to fields ranging from computer graphics algorithm development, scientific and information visualization and from hardware design to the user interface, digital aesthetics, art, education and communication. The implication of this is that while billions of dollars are now spent on the technology and development of digital visualization and interaction, molecular graphics is no longer a principal driver. It has now become the task of the structural biology community to explore and adapt the technologies from these

other fields to serve the needs and expand the capabilities of molecular graphics and visualization in synthesis, analysis and communication. On the other hand, since biology is so complex, its modeling, visualization and analysis continue to challenge the technology. Here we survey some of the current approaches and the utility of new developments in the service of our science.

Software Development and Integration

One major challenge facing the molecular graphics and visualization community is how to best take advantage of all of the existing visualization, interaction, and animation advances that have been developed for other purposes. These advances include highly capable modeling techniques, such as geometric meshing, Boolean operations on complex geometries, rendering techniques that include realistic or special lighting effects, texture and bump mapping on complex surfaces, and interaction and animation approaches that include collision detection, kinematics and other physically based phenomena. In general there are two different ways to face this challenge. One path is to utilize existing high-end software; the other is to re-implement the desired technology or feature into existing molecular modeling programs. The re-implementation approach involves replicating prior work, which can be difficult and costly. The algorithms themselves may be tricky to reconstruct, and the implementation can be labor intensive in terms of coding time and integration into the existing molecular modeling package. On the other hand, utilization of the existing software packages can incur significant expense, if the codes are commercial, and in any case, the existing codes do not contain any domain knowledge of molecular structure, representation, or analysis. A number of high-end commercial modeling, rendering and animation packages have been developed for film and other entertainment industries. Among the most widely used are Maya⁴³, Cinema4D⁴⁴, 3DStudio Max⁴⁵, FormZ⁴⁶, and Houdini⁴⁷. Some other highly capable graphics software packages exist, such as Blender⁴⁸ and Sketchup⁴⁹, which are either open source or free for non-commercial use. Similarly, so called “game engines” such as Unity⁵⁰ and Torque⁵¹ have been developed to facilitate development of gaming software for real-time interactions utilizing fast rendering, physical simulations and adaptable user interfaces. The issue with using these high-end packages for molecular visualization is that they do not encompass any specific domain knowledge. Once a molecular structure is imported into this type of high-end software, it leaves behind its molecular and chemical data and its uses within the software. Thus, this creates an issue of continual import and export of structures between the graphics environment and the molecular modeling environment. To overcome this obstacle several approaches have been developed. One approach is to build molecular “plug-in” toolkits to extend a specific high-end graphics program with molecular modeling capability. Toolkits such as Molecular Maya⁵² and BioBlender^{53;54}, fall into this category. These toolkits make it easy to utilize the rendering and animation capabilities of the host graphics program while maintaining the ability to use molecular data, select molecular entities, such as atoms, amino acid residues and polypeptide chains, and specify conventional molecular representations, such as CPK, ball and stick, and molecular surfaces. Such toolkits have been successfully used in generating “Hollywood quality” animations of molecular structures and mechanisms (see, for instance, the work of Janet Iwassa^{55;56} (Figure 7). One major drawback of these toolkits is that they require a new implementation of the molecular visualization capability for each

specific high-end graphics environment. Typically these implementations do not have the full capabilities of mature and widely used molecular visualization environments such as Chimera⁵⁷, PyMOL⁵⁸, PMV⁵⁹, or VMD⁶⁰. A second approach to building upon the capabilities of existing high-end graphics program is to “embed” an existing molecular visualization package into the host graphics program. This becomes possible when the high-end host program contains within it a widely used interpreter programming language, such as Python, which happens to be the case for many of these codes, such as Maya, Cinema4D, 3DStudio Max and Blender. We have taken advantage of the fact that PMV (the Python Molecular Viewer) was developed in an object-oriented, component based Python environment, and that the molecular visualization components could be easily separated from the graphical user interface (GUI) of PMV. Thus with a simple open-source connector plug-in, an embedded Python Molecular Viewer (ePMV)⁶¹ can be inserted into any number of high-end graphics codes without having to re-implement the large existing code-base for the molecular visualization components.

Molecular representation and rendering—Structural molecular biology is now over 60 years old, and while the breadth and range of structural information has advanced remarkably over this time, the fundamental representations of molecular Structures has evolved into a relatively standard visual vocabulary. CPK and ball and stick atomic representations (with standard coloring) tend to be used for the details of atomic resolution structures; ribbon diagrams⁶² are used to highlight the backbone-level secondary and tertiary structural aspects of proteins and polynucleotide structures. A variety of geometric surface representations, including solvent excluded molecular surfaces⁶³ and coarse molecular and “blobby” isosurfaces⁶⁴ are used to convey the overall shapes of biological molecules and their assemblies. Molecular properties, such as hydrophobicity, polarity, electrostatics and sequence conservation, are represented by color encoding onto the geometric representations. Volumetric properties such as electron density and electrostatic fields are typically represented either as spatial isosurfaces at selected property values, or using direct volume rendering techniques, where the voxel values of the property of interest are translated into the optical characteristics of opacity and color⁶⁵ or as electrostatic field lines surrounding the molecular surface^{66;67}.

While this representational vocabulary has remained fairly static, advances in structural science have extended its range of applicability, and graphics technology has improved on its rendering and interactivity in numerous ways. Experimental NMR constraints and molecular modeling techniques, such as molecular dynamics, docking, and homology modeling can produce large ensembles of biomolecular structures and complexes. Static superposition and representation of these ensembles presents a problem in visual complexity, which has been addressed in a number of approaches, typically by using “clouds” of structures⁶⁸, variable diameter “sausage” backbone tubes, or volumetric rendering based on spatial occupancy⁶⁹. Alternatively, using interactive graphics with sequential display of structural information or trajectories opens up the time domain for comparison and analysis of dynamical information. “Flip-book” comparisons of several individual superposed structures can give a clearer picture of differences.

As the size and complexity of molecular structures has grown, the need for graphic clarity has become even more important. Here, advances in computer graphics rendering technology have had significant impact. The push toward realism in computer-generated imagery for film and games has resulted in a number of advanced rendering techniques. Depth cueing, or aerial perspective, has long been used in molecular graphics to give static and dynamic images a sense of dimensionality. The rendering of shadows can be helpful in interpreting depth and the relationship of component structures to one another, but can sometimes produce cluttered or confusing images. More recently, especially with surface representations, the development and use of ambient occlusion (AO) rendering has advanced the perceptual interpretation of shape. True ambient occlusion calculates how exposed each point on a surface is to ambient lighting or, put another way, how much of the surrounding environment can be seen from each point on a surface. For a typical, complex molecular surface this can be a time consuming calculation, and cannot be updated in real-time. However an approximation, called screen space ambient occlusion (SSAO), using pixel-based depth information is independent of scene complexity and can work in real time for dynamic representations (Figure 8). The program QuteMol⁷⁰ was the first molecular modeling package to incorporate SSAO. Subsequently many popular packages have implemented this type of rendering. This and other lighting effects that work on the screen-space image rather than the geometry can take advantage of the computer's built-in graphics processing units (GPUs) to speed up rendering and interactive display.

The speed of transferring, rendering and interacting with complex molecular systems becomes a significant issue when the number of elements (atoms, chains, molecules, assemblies) becomes very large. A ribosome contains over 100,000 atoms, viral capsids can contain over a million atoms. Current structures in the Protein Data Bank now contain assemblies as large as the capsid core particle of HIV, with more than 1300 protein chains and over 2 million independent atomic coordinates. The new MMTF binary file format from the PDB enables the efficient storage and transfer of such large structures⁷¹. Very large numbers of atomic spheres represented as triangulated polygonal surfaces is a challenge for molecular modeling software. However, screen space rendering methods, termed texture-based or imposter rendering, developed in the late 1990's⁷² for GPUs utilize camera-facing two-dimensional glyphs of shaded spheres or cylinders and pixel depth information for each atom or bond - greatly reducing the polygon count and enabling interactive display of very large molecular complexes⁷³. Other molecular representations for large-scale structures, using volume rendering and iso-surfaces created by meta-balls, can also be rendered in real time using GPUs⁷⁴. Advanced GPU-based imposter rendering techniques using WebGL, such as NGL⁷⁵ and LiteMol⁷⁶, enable interactive viewing of these massive molecular structures in a web browser (Figure 9).

Integrative Modeling and Visualization

Today, experimental methods can characterize the structures and distributions of biomolecules and their complexes both in isolation and in the context of cellular environments. These data, combined with information from genomics, proteomics, and other systems-based methods are being integrated into mesoscale structural models of complex cellular environments down to molecular and atomic detail. These models can be used in

simulations to explore the dynamic and emergent behavior of crowded and heterogeneous cellular environments, and to test hypotheses regarding assembly and mechanism of multi-scale cellular processes. Building, visualizing and interacting with such large structural models is a significant challenge in the field of integrative structural and computational biology⁷⁷. Over the past several years there have been a number of efforts to build complete structural models of cellular environments at molecular detail. This type of work has typically focused on a particular portion of a cell for example *E.coli* cytoplasm⁷⁸, *M. genitalium* cytoplasm⁷⁹, bacterial division machinery⁸⁰, synaptic vesicles⁸¹, and an entire synaptic bouton⁸². Other efforts have endeavored to build software frameworks to enable the construction of cellular environments at molecular scale [eg. PackMol⁸³, MolTemplate⁸⁴, MegaMol⁸⁵]. CELLPack⁸⁶ is an advanced version of this class of software. It uses distance field grid to discretize and describe a cellular model in a volume using multiple modular algorithms to interoperate on the model. It combines these algorithms to integrate three different major localization modes - volumetric, surface, and procedural - into a unified model. CellPACK has been used to produce models of systems as large as a complete mycoplasma cell, with about 30,000 proteins, 200,000 lipids, and one million base pairs of DNA, comprising over 50 million atoms^{87,88} (Figure 10). Work on the GPU visualization of this type of mesoscale cellular model has demonstrated the ability to interactively navigate scenes with over 10 billion atoms at 60 frames per second⁸⁹ which is roughly the number of protein atoms in an *E. coli* bacterial cell. The idea of working with and interacting with terabytes of data has driven the field of big data analysis and visualization. A recent report in the visualization literature describes the interactive visualization of a dark matter cosmological simulation of a 32 terabyte data set with over a trillion particles by using advanced GPU technology⁹⁰. This effort involved several levels of data pre-fetching and caching from disk to RAM to GPU memory and simultaneously utilized two GPUs, one for visualizing the entire data set at coarse resolution using volume rendering, and the other for rendering detailed particle views. In terms of structural biology this size data represents about the number of protein atoms in a typical human cell. The ability to build an atomistic model of a complete human cell with all of its internal structures and interactions is, as yet, an unrealized and challenging goal for the structural biology community. However these efforts from the atom upwards coupled with those from the cell downwards, with projects like the Allen Cell Explorer⁹¹ make this goal appear feasible in the not too distant future.

Computer gaming has become a multi-billion dollar entertainment industry with a strong technical focus on visualization, interactivity and engagement. Over the last several years a number of “serious” games have been developed and deployed in the educational and citizen science communities. For example, Immune attack⁹² has been developed to teach fundamentals of immunology. It is a so called “first-person shooter game that requires players to activate specific proteins to cause targeted behaviors of various white blood cells. Fold-IT⁹³ is probably the best-known biostructure- based game. It is more of an interactive puzzle that challenges users to fold up a protein structure by manipulating the unstructured polypeptide chain. It has introduced thousands of non-scientists to the intricacies of protein folding. Game Engines, such as Unity3D and Torque3D, are component-based high-level software platforms designed to facilitate the development of a wide variety of sophisticated interactive computer games and experiences. Typically such development environments

include resources for graphics, sound, physics, and user interface device controls. The physics components tend to include procedures for collision detection and constraint specification, such as springs between objects. These game engines can serve as effective development environments for interactive visualization and manipulation of biomolecules in modeling tools and have been used to produce programs such as UDock⁹⁴ and UnityMol⁹⁵.

The Unity3D game engine was utilized in the creation of CellVIEW⁸⁹ an interactive environment for visualizing CellPACK models. For interactive performance on large mesoscale models, CellVIEW uses screen-space GPU rendering of sphere imposters at different levels of detail (LOD) depending on distance to the camera. The rendering is applied only to objects visible in the camera's field of view to keep the computational burden to a minimum. A scene with thousands of individual protein molecules can be difficult to interpret and navigate, so CellVIEW has a number of features to simplify and focus on specific characteristics and content. Visibility control of specified compartments or molecular ingredient types are provided. Clipping is also a useful tool to select only a portion of a larger scene or object type. CellVIEW has wide variety of clipping geometries available that can be applied to the entire scene or specific compartments or molecular ingredients to aid in selected viewing⁹⁶. Because of the multi-scale nature of these models the coloring of structural elements is also based upon level of detail⁹⁷. Thus when the camera is at a distance, the coloring of the scene is based upon the compartments or organelles in the view and, as the camera moves successively closer, the coloring can change to distinguish molecule types, then chains and finally atoms. The fact that CellPACK itself utilizes a volume-element (voxel) data structure has recently led to a new implementation that uses a novel set of GPU algorithms which unite modeling and visualization in the context of a game engine for rapid construction of mesoscale models of complex biological structures⁸⁷. These structures consist of multiple membrane-enclosed compartments including both soluble molecules and fibrous structures. Lipid membrane compartments are defined using volume voxelization of triangulated meshes, and are populated by patches of lipid bilayer pre-calculated from an atomistic molecular dynamics simulation. Soluble molecules are placed within compartments by distributing them randomly using a Halton sequence⁹⁸ on a space filling curve, which guarantees object centers will be no closer than a specified distance from one another. Fibrous structures, such as DNA, RNA or actin filaments, are created by procedure-based self-avoiding random walks. Procedural models, such as a bacterial nucleome⁸⁸, are typically treated as single ingredients and placed as such. Any resulting overlaps of molecules are resolved by a grid-localized forced-based system. The fact that the placement and relaxation of molecular ingredients is done in a highly parallel fashion on the GPU enables the interactive creation of mesoscale models of cellular environments and the production of large ensembles of such models, removing a significant barrier to structure based hypothesis generation and simulation.

CellVIEW was designed as an interactive viewer of CellPACK models, and not as a creation tool for generating new mesoscale models. In a complementary effort, using the Unity game engine with a computer painting metaphor, we have developed CellPAINT⁹⁹ to enable the production of two dimensional Goodsell style "mesoscape illustrations." With cross-section views, three depth layers, a pallet of 2D sprites of molecular images, and circle colliders for steric exclusion, a user can draw membrane compartments, spray or brush molecules into

them, and animate the scene with Brownian motion diffusion (Figure 11). Currently we are developing CellPAINT 3D in Unity, moving into a volumetric context and which could serve as a CellPACK building interface.

Perception and the User Interface—Over the last decade or so, with the continuing march of Moore's Law and the incorporation of cameras, sensors and displays into mobile computing technology, how we interact with our computational environment has changed dramatically. Moreover, with the advancing developments of virtual and augmented reality, 3D computer printing, and the Internet of things (IoT), the lines between the real and the virtual world have become blurred and our expectations of how we interact with information have expanded. While much of the fundamental technologies underlying these advances have been around for two or even three decades, they have only recently reached the level of wide-spread availability and acceptance. What, if anything, is the impact of these technologies on visualization and interaction with biomolecular structure?

Tangible and Haptic interfaces: The advent of computer autofabrication, now referred to as 3D printing, has provided an automated way to produce physical objects from computer data, and has opened up the possibility of having highly accurate material visualizations of complex molecular molecules¹⁰⁰ (see, for example, Figure 12). Physical models can convey spatial relationships and mechanisms in ways that images alone cannot. They engage perceptual and cognitive processes that go beyond the visual and bring a sense of reality and natural interaction into the process of exploration and understanding. Physical molecular models can also serve as "analogue computers" where spatial relationships between components in complex molecular interactions can be explored and manipulated. In addition, individual molecular components can be printed, and then used to build novel or flexible larger structures. For instance, this approach has been applied to create foldable polypeptide chain models^{101;102} and self-assembling virus models^{103;104}. To broaden the use and utility of physical models in structural biology, the NIH established a repository for 3D printable models¹⁰⁵. Of course physical molecular models are constrained in ways that computer graphics visualizations are not. Every 3D printing technology uses distinct materials, each with its own mechanical properties and limitations. Therefore the range of molecular representations or color choices greatly depends on the nature of the 3D print technology used. Moreover, a physical model constitutes a fixed representation of a molecular entity and lacks the unlimited representational flexibility and manipulability of a computer visualization. Thus physical molecular models, while valuable adjuncts for examination and exploration, play a relatively limited role in molecular visualization and tend to be used mostly for communication, either in a teaching or research collaboration setting.

Force feedback devices and other haptic modalities also engage our senses of touch and proprioception. Today we find haptic displays in the form of vibrations or small taps from smart phones and watches to prompt awareness, or in video game controllers to enhance certain actions or physical effects. Touch screens, accelerometer motion sensors and depth cameras can capture finger, hand or arm gestures and body movement as input for a wide variety of computer and visualization activities. Touch screens are now commonly used to

control molecular rotations, translations and scaling on laptops and mobile devices. Haptic output rendering of molecular processes, however, is more difficult. Over the years, elaborate force feedback devices have been explored for creating the feel of molecular interactions²². Such devices work both as 6 degree of freedom (DOF) input devices for positioning molecules and as 3 or 6 DOF output devices for rendering interaction forces. These devices are typically mechanical armatures with encoders and actuators in a nested succession of joints. By manipulating an attached wand moving a virtual molecule, it is translated and oriented while forces (and in some devices torques) computed from the molecule's interactions are fed back to the user. Typically both long-range electrostatic and short-range van der Waals forces are rendered^{40;106}. A different use of such force-feedback devices has been to remotely control and feel the forces at the tip of an atomic scanning probe microscope¹⁰⁷. Interesting exploratory work has been done using force feedback displays in educational settings^{108;109}. However, even with the cost of some haptic devices now below \$1000, force-feedback has not been widely adopted for molecular modeling and visualization. Their cost is still high compared to other input and output devices; the workspace of force-feedback arms is typically highly constrained; and the dynamic range of the feedback forces is limited, making it difficult to discriminate force magnitude. Hard surface rendering is especially difficult, requiring high frequency output, which can lead to uncontrolled mechanical jittering. Additionally, as a physical device that can create significant force, these arms can potentially injure a user. Recently, however, a new interest in haptics has arisen in gaming and virtual reality applications. Haptic gloves with hand and finger tracking and feedback are now becoming commodity devices¹¹⁰ but have yet to be applied in the context of molecular rendering. It still remains to be seen whether haptics and force feedback devices will be useful, useable, and adaptable to the needs of biomolecular visualization.

Virtual and Augmented Reality: Virtual Reality (VR) is typically defined as an artificial environment experienced through sensory stimuli generated by a computer. While the initial concepts and work on VR date back to the 1980s^{111;112} it has only recently become a consumer technology and new applications are appearing daily. Commercially available virtual reality headsets range in performance and price. At the low end, hand-held Google Cardboard and head-mounted smartphone display devices provide viewpoint tracking and a 360° immersive experience of pre-recorded content. At the other end are purpose-built head mounted displays (HMDs) along with motion and hand controller tracking that communicate with interactive virtual reality software running on high end GPU enabled computers. Some of the earliest explorations of VR were focused on biomolecular visualization. In one telling early experience in the 1990's at UNC Chapel Hill, a protein crystallographer was fitted with an HMD in a large space arrayed with tracking devices to enable walking around and into a large virtual version of a protein of interest. As Fred Brooks, the project PI, has described it, the first thing the scientist requested was a chair to sit down in. Clearly, walking around inside the protein was not how she wanted to interact with her molecule and, in fact found doing so disorienting.

An experience with VR in 1998 was somewhat different. Visiting Tom DeFanti's Electronic Visualization Lab (EVL) at U. Illinois at Chicago. I had imported our early mesoscale model

of a 75nm cube of blood plasma (Figure 13), into the CAVE¹¹³ a cubic space with stereoscopic projections on surrounding walls and ceiling. Exploring our mesoscape without the need of a bulky tethered HMD, made walking among the boulder sized proteins and sticking my head inside the poliovirus capsid an engaging and inspiring experience. However, as I couldn't afford the space for, or the cost of, a CAVE for our routine visualization environment, I didn't pursue its application to structural molecular biology at the time.

Even today, VR is groping around in the molecular world, to find a "killer app" for structural biologists. Would implementing a widely used molecular visualization environment like PyMOL or Chimera in a virtual reality environment entice a lot of users to drop their traditional molecular displays? On the plus side, the VR environment can provide a strong sense of presence with the molecular model. Employing hand tracking to position and pose a molecule in VR eases cognitive load, and can enable more accurate placement while facilitating "what-if" type exploration. Perceptually, the combination of spatial cues from head motion coupled with stereoscopy along with the person centric immersive perspective increase a sense of realism and true three dimensionality. On the negative side, the head sets themselves, even though they continue to improve, are still intrusive to wear, and are physically restrictive, with cables tethered to the computer. Also, for an immersive experience that doesn't leave one dizzy or motion sick, fast image update rates (over 90fps), high resolution, and low latency motion response, as well as correctly computed stereoscopy, are critical. Most importantly, the sense of removal from the real world when immersed in the virtual, can be disorienting, and feel disembodied and isolating after extended use.

Since game engines are designed to enable running software on multiple platforms, we decided to implement our prototype CellPAINT 3D program for VR headsets (Figure 14). We have found that the ability to place objects in an immersive volumetric space provides a unique user experience. CellPAINT VR is still in early development, and the painting metaphor is evolving into a 3D molecular mesoscale model-building environment. Many issues, such as the interactive building of molecular complexes and effective "way finding" and navigation in complex environments still need to be improved.

Augmented Reality (AR) blends the real and cyber world together using a number of different types of interfaces¹¹⁴. Mobile devices are now capable of running augmented reality applications, using their built in camera to recognize some aspects of the surrounding world, and then inserting computer generated text or 3D graphics into the scene. The computer augmentation is viewed on the hand-held display in the context of the real-world camera view. Depth information acquired by either camera motion or optics convincingly places the generated graphics into the scene. Pokemon Go, a popular game, introduced a wide audience to this technology. More capable AR systems have been developed with head-mounted displays that use computer projections onto a transparent visor in front of the eyes, with the real world still in view. The Microsoft HoloLens is an example of this class of AR display device. The key to augmented reality is the recognition and tracking of real-world objects or environments, and the ability to convincingly display computer generated material within that context. The strong advantage of AR over VR is that the user operates in the real world but can perceive and interact with computer generated visualizations.

Several years ago, our interest in the use of 3D printed molecular models, led us into the area of Augmented Reality. Through the use of an early open source augmented reality toolkit (ARToolkit¹¹⁵), with coded markers on objects for camera recognition, we developed Tangible Molecular Interfaces^{116;117}. By attaching these markers to molecular models, we could track their motions with a camera, and augment the molecular models with additional representations and computed properties, such as surrounding electrostatic fields. Since multiple molecular models could be tracked in the computer we could continuously monitor distances between them and interactively compute and display energetics of molecular interaction (Figure.15). To this point there has been no widespread adoption of Augmented Reality in the field of structural molecular biology. As costs continue to drop, and as form and function of the devices become less intrusive and easier to use, these technologies could be especially effective in scientific collaboration, where multiple people can view and interact with the same virtual molecular entities while still being engaged in the real physical environment.

Prospects/ Summary

Scientific and technical trends appear to be clear. Structural molecular biology continues to expand into more complex systems and integrative approaches, and the path from the atom to the living cell appears ever more approachable. Information visualization of non-structural data continues to evolve and combine with structural visualizations. Computer power and the utilization of that power with more effective algorithms and more capable physical devices continue unabated. Digital communication and the extent and pervasiveness of networked people and machines bring together new communities and collaborations.

Artificial Intelligence (AI) and machine learning are perhaps the most widely anticipated technologies for the near future, and are being used in ever-broader applications, including molecular biology. Since AI is useful for pattern discovery and image interpretation, we are beginning to see the coupling of AI algorithms with interactive recognition of various imaging technologies, such as electron tomographic annotation¹¹⁸. Pattern recognition can also help refine parameters for creating mesoscale molecular models of cellular features, such as membrane and fiber shapes and distributions. With the structural biology literature expanding at exponential rates, machine learning used in text and image processing could assist greatly in developing and curating consensus views of molecular componentry and distributions in cellular systems. However, since the function of visualization is to put human in the loop, will artificial intelligence make visualization obsolete? It is my belief that it will be a long time before human inspiration, insight and fresh concepts derived from visualization and interaction are no longer central to the progress of structural molecular biology.

Acknowledgements

I would like to acknowledge the past and present members of the Molecular Graphics Laboratory at The Scripps Research Institute, for their creativity, talent and hard work. I especially thank Michael Pique, Michel Sanner, David Goodsell, Ludovic Autin and Adam Gardner for their continuing efforts in the area of molecular visualization in making the unseen world of biology accessible, understandable, and beautiful.

I also acknowledge with great appreciation funding on visualization research from the National Institutes of Health (R33EB000798, P41-GM103426, U54GM103368); National Science Foundation (ITR 0121282, and DRL-1108896) and the U.S. Department of Education (IES R305A120047).

Literature Cited

1. Iwasa JH & Kirchhausen T (2012). Clathrin Mediated Endocytosis, 2018, http://www.youtube.com/watch?v=o_EUHu4OJus.
2. Perkins JA (2005). A history of molecular representation part one: 1800 to the 1960s. *The Journal of Biocommunication* 31.
3. Perkins JA (2005). A history of molecular representation part one: The 1960s - Present. *The Journal of Biocommunication* 31.
4. Olson AJ & Pique ME (1998). Visualizing The Future of Molecular Graphics. *SAR and QSAR in Environmental Research* 8, 233–247. [PubMed: 9522476]
5. Walter LK (1965). Precision space - filling atomic models. *Biopolymers* 3, 665–679. [PubMed: 4158989]
6. Fieser L (1963). Plastic Dreiding models. *Journal of Chemical Education* 40.
7. Pauling L & Corey RB (1950). Two hydrogen-bonded spiral configurations of the polypeptide chain. *Journal of the American Chemical Society* 72, 5349–5349.
8. Crick F (2008). *What mad pursuit* (Series APSF, Ed.), Basic Books, New York.
9. Watson JD (1968). *The Double Helix: A Personal Account of the Structure of DNA*, Weidenfeld and Nicolson, London.
10. Perutz MF, Rossmann MG, Cullis AF, Muirhead H & Will G (1960). Structure of hæmoglobin: a three-dimensional Fourier synthesis at 5.5-Å. resolution, obtained by X-ray analysis. *Nature* 185, 416–422. [PubMed: 18990801]
11. Kendrew JC, Bodo G, Dintzis HM, Parrish RG, Wyckoff H & Phillips DC (1958). A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature* 181, 662–666. [PubMed: 13517261]
12. Johnson C (1965). ORTEP. Report ORNL-3794 Oak Ridge National Laboratory.
13. Farrugia LJ (1997). ORTEP-3 for Windows - a version of ORTEP-III with a Graphical User Interface (GUI) *J. Appl. Cryst* 30, 565.
14. Richards FM (1968). The matching of physical models to three-dimensional electron- density maps: a simple optical device. *Journal of molecular biology* 37, 225–230. [PubMed: 5760491]
15. Levinthal C (1966). Molecular Model-building by Computer. *Scientific American* 214, 42–53. [PubMed: 5930597]
16. Dickerson R, Geis I (1969). *The Structure and Action of Proteins*, W.A. Benjamin Inc., Menlo Park, California.
17. Dickerson RE (1980). Cytochrome c and the Evolution of Energy Metabolism. *Scientific American* 242, 136–153.
18. Richardson JS (1985). Schematic drawings of protein structures In *Methods in Enzymology*, Vol. 115, pp. 359–380. Academic Press. [PubMed: 3853075]
19. Diamond R (1979). BILDER: An Interactive Graphics Program for Biopolymers, User's Guide MRC Laboratory of Molecular Biology, Cambridge, U.K.
20. Jones T (1978). A graphics model building and refinement system for macromolecules. *Journal of Applied Crystallography* 11, 268–272.
21. Molnar CE, Barry CD, Rosenberger FU (1976). MMSX Technical Memo Computer Systems Laboratory, Washington University St. Louis, St. Louis, MO, USA.
22. Brooks FP (1988). Grasping reality through illusion; interactive graphics serving science. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems ACM*, Washington, D.C., USA.
23. Olson AJ & Goodsell DS (1992). Visualizing Biological Molecules. *Scientific American* 267, 76–83. [PubMed: 1411463]

24. Olson AJ (1983). Computer Graphics in Biomolecular Science, NicoGraph '83, Japan Computer Graphics Society, Tokyo Japan.
25. Lee B & Richards FM (1971). The interpretation of protein structures: Estimation of static accessibility. *Journal of Molecular Biology* 55, 379–IN4. [PubMed: 5551392]
26. Connolly ML (1983). Solvent-accessible surfaces of proteins and nucleic acids. *Science* 221, 709–713. [PubMed: 6879170]
27. Thomas KP (1979). The shaded surface display of large molecules. In *Proceedings of the 6th annual conference on Computer graphics and interactive techniques ACM*, Chicago, Illinois, USA.
28. Max NL (1979). ATOMLLL: ATOMS with shading and highlights. *SIGGRAPH Comput. Graph.* 13, 165–173.
29. Ferrin TE, Huang CC, Jarvis LE & Langridge R (1988). The MIDAS display system. *Journal of Molecular Graphics* 6, 13–27.
30. Hubbard RE (1986). *Computer Graphics and Modeling*, Cold Spring Harbor Laboratory, New York.
31. O'Donnell TJ & Olson AJ (1981). GRAMPS - A graphics language interpreter for real-time, interactive, three-dimensional picture editing and animation. *SIGGRAPH Comput. Graph.* 15, 133–142.
32. Connolly ML & Olson AJ (1985). GRANNY, a companion to GRAMPS for the real-time manipulation of macromolecular models. *Computers & Chemistry* 9, 1–6.
33. Olson AJ (1983). Experience in Producing Computer Vector Animation in OMNIMAX Format, Nicograph '83, Japan Computer Graphics Society, Tokyo Japan.
34. Max LM (1983). Computer graphics distortion of IMAX and OMNIMAX projection, Nicograph '83, Tokyo, Japan.
35. Brown J and Cunningham S (2007). A history of ACM SIGGRAPH. *Commun. ACM* 50, 54–61.
36. Upson C, Faulhaber TA, Kamins D, Laidlaw D, Schlegel D, Vroom J, Gurwitz R & Dam A. v. (1989). The application visualization system: a computational environment for scientific visualization. *IEEE Computer Graphics and Applications* 9, 30–42.
37. Kraulis P (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *Journal of Applied Crystallography* 24, 946–950.
38. Sayle RA & Milner-White EJ (1995). RASMOL: biomolecular graphics for all. *Trends in Biochemical Sciences* 20, 374–376. [PubMed: 7482707]
39. Herraes A (2006). Biomolecules in the computer: Jmol to the rescue. *Biochemistry and Molecular Biology Education* 34, 255–261. [PubMed: 21638687]
40. Frederick P. Brooks, Jr, Ming O-Y, James JB. & Kilpatrick PJ. (1990). Project GROPEHaptic displays for scientific visualization. *SIGGRAPH Comput. Graph.* 24, 177–185.
41. Tek A, Laurent B, Piuze M, Lu Z, Chavent M, Baaden M, Delalande O, Martin C, Piccinali L, Katz B, Bourdot P, Autin L & Ferey N (2012). Advances in Human-Protein Interaction - Interactive and Immersive Molecular Simulations. In *Biochemistry, Genetics and Molecular Biology* “Protein-Protein Interactions - Computational and Experimental Tools” (Weibo C. & Hao H, eds.), pp. 27–65. InTech. [PubMed: 23216827]
42. Brooks JFP (2014). Impressions by a dinosaur - summary of Faraday discussion 169: molecular simulations and visualization. *Faraday Discussions* 169, 521–527. [PubMed: 25341079]
43. Autodesk. (2018). Maya -computer animation software, 2018, <http://www.autodesk.com/products/maya/overview>.
44. Maxon. (2018). Cinema 4D (Maxon, ed.), 2018, <http://www.maxon.net/en/products/cinema-4d/overview/>.
45. Autodesk. (2018). 3D Studio Max, <http://www.autodesk.com/products/3ds-max/overview>.
46. FormZ, Vol. 2018 AutoDesSys, <http://www.formz.com>.
47. SideFX. (2018). Houdini, <http://www.sidefx.com/>.
48. Blender. Blender, 2018, <http://www.blender.org/>.
49. Sketchup. Sketchup, <https://http://www.sketchup.com/products/sketchup-free>.
50. Unity. Unity 3D, <https://unity3d.com/unity>.
51. GarageGames. Torque 3D, 2018, <http://www.garagegames.com/products/torque-3d>.

52. Molecular Maya, <https://clarafi.com/tools/mmaya/>.
53. Andrei RM, Callieri M, Zini MF, Loni T, Maraziti G, Pan MC & Zoppe M (2012). Intuitive representation of surface properties of biomolecules using BioBlender. *BMC Bioinformatics* 13, S16.
54. bioblender. Bioblender Making Visuals with Technology, 2018, <http://www.bioblender.org/>.
55. Iwasa JH (2010). Animating the model figure. *Trends in Cell Biology* 20, 699–704. [PubMed: 20832316]
56. Iwasa JH Visualization of molecular processes, Vol. 2018.
57. Pettersen EF, Goddard TD, Huang CC, G S, C, Greenblatt DM, Meng EC. & Ferring TE. (2004). UCSF Chimera -- A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* 25, 1605–1612. [PubMed: 15264254]
58. DeLano WL (2002). The PyMOL Molecular Graphics System on World Wide Web. CCP4 Newsletter On Protein Crystallography.
59. Morris GM, Huey R, Lindstrom L, Sanner MF, Belew RK, Goodsell DS, Olson AJ (2009) AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem* 30, 2009, 2785–2791. [PubMed: 19399780]
60. Humphrey W, Dalke A & Schulten K (1996). VMD: Visual molecular dynamics. *Journal of Molecular Graphics* 14, 33–38. [PubMed: 8744570]
61. Johnson Graham T, Autin L, Goodsell David S, Sanner Michel F. & Olson Arthur J. (2011). ePMV Embeds Molecular Modeling into Professional Animation Software Environments. *Structure* 19, 293–303. [PubMed: 21397181]
62. Carson M (1987). Ribbon models of macromolecules. *Journal of Molecular Graphics* 5, 103–106.
63. Sanner MF, Olson AJ, Spehner J-C (1996). Reduced surface: an efficient way to compute molecular surfaces. *Biopolymers* 38, 305–320. [PubMed: 8906967]
64. James FB (1982). A Generalization of Algebraic Surface Drawing. *ACM Trans. Graph.* 1, 235–256.
65. Goodsell DS, Mian IS & Olson AJ (1989). Rendering volumetric data in molecular systems. *Journal of Molecular Graphics* 7, 41–47.
66. Getzoff ED, Tainer JA, Weiner PK, Kollman PA, Richardson JS & Richardson DC (1983). Electrostatic recognition between superoxide and copper, zinc superoxide dismutase. *Nature* 306, 287. [PubMed: 6646211]
67. Zoppè M & Loni T (2015). The Representation of Electrostatics for Biological Molecules In *Computational Electrostatics for Biological Applications: Geometric and Numerical Approaches to the Description of Electrostatic Interaction Between Macromolecules* (Rocchia W & Spagnuolo M, eds.), pp. 215–225. Springer International Publishing, Cham.
68. Damm KL & Carlson HA (2006). Gaussian-Weighted RMSD Superposition of Proteins: A Structural Comparison for Flexible Proteins and Predicted Protein Structures. *Biophysical Journal* 90, 4558–4573. [PubMed: 16565070]
69. O'Donoghue SI, Goodsell DS, Frangakis AS, Jossinet F, Laskowski RA, Nilges M, Saibil HR, Schafferhans A, Wade RC, Westhof E & Olson AJ (2010). Visualization of macromolecular structures. *Nature Methods* 7, S42. [PubMed: 20195256]
70. Tarini M, Pignoni P & Montani C (2006). Ambient Occlusion and Edge Cueing for Enhancing Real Time Molecular Visualization. *IEEE Transactions on Visualization and Computer Graphics* 12.
71. Bradley AR, Rose AS, Pavelka A, Valasatava Y, Duarte JM, Prlic A & Rose PW (2017). MMTF—An efficient file format for the transmission, visualization, and analysis of macromolecular structures. *PLOS Computational Biology* 13, e1005575. [PubMed: 28574982]
72. Xavier D, Francois S, Gernot S & Julie D (1999). Multi - layered impostors for accelerated rendering. *Computer Graphics Forum* 18, 61–73.
73. Bajaj C, Djeu P, Siddavanahalli V (2004). Thane A, *Proceedings IEEE Visualization* 2004.
74. Chavent M, Levy B, Krone M, Bidmon K, Nomine J-P, Ertl T & Baaden M (2011). GPU-powered tools boost molecular visualization. *Briefings in Bioinformatics* 12, 689–701. [PubMed: 21310717]

75. Rose AS & Hildebrand PW (2015). NGL Viewer: a web application for molecular visualization. *Nucleic Acids Research* 43, W576–W579. [PubMed: 25925569]
76. Sehnal D, Deshpande M, Varekova RS, Mir S, Berka K, Midlik A, Pravda L, Velankar S & Koca J (2017). LiteMol suite: interactive web-based visualization of large-scale macromolecular structure data. *Nature Methods* 14, 1121. [PubMed: 29190272]
77. Im W, Liang J, Olson A, Zhou H-X, Vajda S & Vakser IA (2016). Challenges in structural approaches to cell modeling. *Journal of Molecular Biology* 428, 2943–2964. [PubMed: 27255863]
78. McGuffee SR & Elcock AH (2010). Diffusion, Crowding & Protein Stability in a Dynamic Molecular Model of the Bacterial Cytoplasm. *PloS Computational Biology* 6:e1000694 [PubMed: 20221255]
79. Feig M, Harada R, Mori T, Yu I, Takahashi K & Sugita Y (2015). Complete atomistic model of a bacterial cytoplasm for integrating physics, biochemistry, and systems biology. *Journal of Molecular Graphics and Modelling* 58, 1–9. [PubMed: 25765281]
80. Vendeville A, Larivière D & Fourmentin E (2011). An inventory of the bacterial macromolecular components and their spatial organization. *FEMS Microbiology Reviews* 35, 395–414. [PubMed: 20969605]
81. Takamori S, Holt M, Stenius K, Lemke EA, Grønborg M, Riedel D, Urlaub H, Schenck S, Brügger B, Ringler P, Müller SA, Rammner B, Gräter F, Flub JS, De Groot BL, Mieskes G, Moriyama Y, Klingauf J, Grubmüller H, Heuser J, Wieland F & Jahn R (2006). Molecular Anatomy of a Trafficking Organelle. *Cell* 127, 831–846. [PubMed: 17110340]
82. Wilhelm BG, Mandad S, Truckenbrodt S, Kröhnert K, Schafer C, Rammner B, Koo SJ, Claen GA, Krauss M, Flaucke V, Urlaub FI. & Rizzoli SO (2014). Composition of isolated synaptic boutons reveals the amounts of vesicle trafficking proteins. *Science* 344, 1023–1028. [PubMed: 24876496]
83. Martinez L, Andrade R, Birgin EG & Martinez JM (2009). PACKMOL: A package for building initial configurations for molecular dynamics simulations. *Journal of Computational Chemistry* 30, 2157–2164. [PubMed: 19229944]
84. Jewett AI, Zhuang Z & Shea J-E (2013). Moltemplate a Coarse-Grained Model Assembly Tool. *Biophysical Journal* 104, 169a.
85. Grottel G, Krone M, Muller C, Reina G & Ertl T (2014). MegaMol—A Prototyping Framework for Particle-Based Visualization. *IEEE Transactions on Visualization and Computer Graphics* 21, 201–214.
86. Johnson GT, Autin L, Al-Alusi M, Goodsell DS, Sanner MF & Olson AJ (2014). cellPACK: a virtual mesoscope to model and visualize structural systems biology. *Nature Methods* 12, 85. [PubMed: 25437435]
87. Klein T, Autin L, Kozlikova B, Goodsell DS, Olson A, Groller ME & Viola I (2018). Instant Construction and Visualization of Crowded Biological Environments. *IEEE Transactions on Visualization and Computer Graphics* 24, 862–872. [PubMed: 28866533]
88. Goodsell DS, Autin L & Olson AJ (2018). Lattice Models of Bacterial Nucleoids. *The Journal of Physical Chemistry B*.
89. Le Muzic M, Autin L, Parulek J & Viola I (2015). cellVIEW: a Tool for Illustrative and Multi-Scale Rendering of Large Biomolecular Datasets. *Eurographics Workshop on Visual Computing for Biomedicine 2015*, 61–70. [PubMed: 29291131]
90. Schatz K, Müller C, Krone M, Schneider J, Reina G & Ertl T (2016). 2016 IEEE 6th Symposium on Large Data Analysis and Visualization (LDAV).
91. Horwitz R & Johnson GT (2017). Whole cell maps chart a course for 21 st-century cell biology. *Science* 356, 806–807. [PubMed: 28546174]
92. Stegman M (2014). Immune Attack players perform better on a test of cellular immunology and self confidence than their classmates who play a control video game. *Faraday Discussions* 169, 403–423. [PubMed: 25340640]
93. Cooper S, Khatib F, Treuille A, Barbero J, Lee J, Beenen M, Leaver-Fay A, Baker D, Popovi Z & players F (2010). Predicting protein structures with a multiplayer online game. *Nature* 466, 756. [PubMed: 20686574]
94. Levieux G, Tiger G, Mader S, Zagury J-F, Natkin S & Montes M (2014). Udock, the interactive docking entertainment system. *Faraday Discussions* 169, 425–441. [PubMed: 25341068]

95. Doutreligne S, Cragnolini T, Pasquali S, Derreumaux P & Baaden M (2014). 2014 IEEE 4th Symposium on Large Data Analysis and Visualization (LDAV).
96. Muzic ML, Mindek P, Sorger J, Autin L, Goodsell DS & Viola I (2016). Visibility Equalizer Cutaway Visualization of Mesoscopic Biological Models. *Computer Graphics Forum* 35, 161–170. [PubMed: 28344374]
97. Waldin N, Le Muzic M, Waldner M, Groller E, Goodsell D, Ludovic A & Viola I (2016). Chameleon: Dynamic Color Mapping for Multi-Scale Structural Biology Models. *Eurographics Workshop on Visual Computing for Biomedicine 2016*, 10.2312/vcbm.20161266.
98. Halton JH (1964). Algorithm 247: Radical-inverse quasi-random point sequence. *Commun. ACM* 7, 701–702.
99. Gardner A, Autin L, Barbaro B, Olson A & Goodsell DS (2018). CellPAINT: Interactive Illustration of Dynamic Mesoscale Cellular Environments. *Computer Graphics & Applications in Press*.
100. Gardner A & Olson A (2016). 3D Printing of Molecular Models. 2016 40.
101. Davenport J, Pique M, Getzoff E, Huntoon J, Gardner A & Olson A (2017). A Self-Assisting Protein Folding Model for Teaching Structural Molecular Biology. *Structure* 25, 671–678. [PubMed: 28380340]
102. Chakraborty P & Zuckermann RN (2013). Coarse-grained, foldable, physical model of the polypeptide chain. *Proceedings of the National Academy of Sciences* 110, 13368–13373.
103. Olson AJ, Hu YHE & Keinan E (2007). Chemical mimicry of viral capsid self-assembly. *Proceedings of the National Academy of Sciences* 104, 20731–20736.
104. Olson AJ (2015). Self-assembly gets physical. *Nature Nanotechnology* 10, 728–728.
105. Coakley MF, Hurt DE, Weber N, Mtingwa M, Fincher EC, Vsevelod A, Chen DT, Yun A, Gizaw M, Swan J, Yoo TS, Huyen Y (2014). The NIH 3D Print Eexchange: a public resource for bioscientific and biomedical 3D prints, 3D Print. *Addit. Manuf* 1, 137–140. [PubMed: 28367477]
106. Wollacott AM & Merz KM (2007). Haptic applications for molecular structure manipulation. *Journal of Molecular Graphics and Modelling* 25, 801–805. [PubMed: 16987676]
107. Guthold M, Falvo MR, Matthews WG, Paulson S, Washburn S, Erie DA, Superfine R, Brooks FP & Taylor RM (2000). Controlled manipulation of molecular samples with the nanoManipulator. *IEEE/ASME Transactions on Mechatronics* 5, 189–198.
108. Persson PB, Cooper MD, Tibell LAE, Ainsworth S, Ynnerman A & Jonsson BH (2007). 2007 IEEE Virtual Reality Conference.
109. Bivall P, Ainsworth S, Tibell LAE (2011). Do haptic representations help complex molecular learning? *Science Education* 95, 700–719.
110. VRgluv. VRgluv, Vol. 2018, <https://vrglu.com/>.
111. Fisher SS, McGreevy MJH Robinett W (1987). Proceedings of the 1986 workshop on Interactive 3D graphics, Chapel Hill, North Carolina, USA.
112. Alqahtani AS, Daghestani LF & Ibrahim FI (2017). Environments and System Types of Virtual Reality Technology in STEM: A Survey. *International Journal of Advanced Computer Science and Applications* 8, 77–89.
113. Cruz-Neira C, Sandin DJ, DeFanti TA (1993). Surround-screen projection-based virtual reality: the design and implementation of the CAVE In Proceedings of the 20th annual conference on Computer graphics and interactive
114. Milgram P, Takemura H, Utsumi A & Kishino F (1994). Telemanipulator and Telepresence Technologies, Boston, MA USA.
115. Brave S, Hiroshi I Andrew D (1998). Tangible interfaces for remote collaboration and communication. Proceedings of the 1998 ACM conference on Computer supported cooperative work; Seattle, Washington, USA: ACM;
116. Gillet A, Sanner M, Stoffler D, Goodsell D & Olson A (2004). IEEE Visualization 2004.
117. Gillet A, Sanner M, Stoffler D & Olson A (2005). Tangible interfaces for structural molecular biology. *Structure* 13, 483–491. [PubMed: 15766549]

118. Chen M, Dai W, Sun SY, Jonasch D, He CY, Schmid MF, Chiu W & Ludtke SJ (2017).
Convolutional neural networks for automated annotation of cellular cryo-electron tomograms.
Nature Methods 14, 983. [PubMed: 28846087]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Highlights

This perspective traces the evolution of structural molecular biology visualization from physical models of the earliest protein structures to immersive interaction of mesoscale models of cellular environments.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 1.
Balsa wood model of hemoglobin from Perutz Lab. ca.1960.

Figure 2.
Stacked electron density (left) and brass Kendrew protein model (right) in Richards Box. ca. 1975

Figure 3.
Screenshot of Bilder ca. 1977

Figure 4.
Dot Molecular Surface with protein backbone trace ca. 1983

Figure 5.
Image from AED 512 raster graphics display of protein-protein interaction between trypsin and trypsin inhibitor, ca. 1984

Figure 6.
Frame from Omnimax animation of trip down DNA major groove, for Disney's EPCOT Center, 1983

Figure 7.
Frame from Maya Animation on clathrin mediated endocytosis¹. Image courtesy of Janet Iwasa, U. Utah, structural data from Thomas Kirchhausen, Harvard Med.

Figure 8.
Comparison of shaded image (Left) and shaded image with SSAO (Right) of GroEL-ES chaperonin complex. Image from PMV.

Figure 9.
Interactive NGL ribbon image of HIV-1 capsid core particle.

Figure 10. CellPACK model of a complete mycoplasma cell interactively visualized in CellVIEW. Two clipping cubes reveal the interior (left) and the DNA alone (right)

Figure 11.
CellPAINTing of HIV entry into a T-Cell

Figure 12.
3D printed model of influenza virus.

Figure 13.
Mesoscale blood plasma model, that was imported into VR in the CAVE in 1998.

Figure 14. Virtual Reality implementation of CellPAINT 3D. Video screen shows part of immersive HMD display.

Figure 15.

Augmented reality display of 3D printed flexible polypeptide folding model. Tracking marker is seen near the top. Computer augmented amino acids can be manipulated and distances between them measured.