

Seeing the PDB

Received for publication, April 3, 2021, and in revised form, April 26, 2021. Published, Papers in Press, May 4, 2021.
<https://doi.org/10.1016/j.jbc.2021.100742>

Jane S. Richardson^{1,*}, David C. Richardson¹, and David S. Goodsell^{2,3,*}

From the ¹Department of Biochemistry, Duke University, Durham, North Carolina, USA; ²Department of Integrative and Computational Biology, The Scripps Research Institute, La Jolla, California, USA; ³Research Collaboratory for Structural Bioinformatics Protein Data Bank, Rutgers, the State University of New Jersey, Piscataway, New Jersey, USA

Edited by Karin Musier-Forsyth

Ever since the first structures of proteins were determined in the 1960s, structural biologists have required methods to visualize biomolecular structures, both as an essential tool for their research and also to promote 3D comprehension of structural results by a wide audience of researchers, students, and the general public. In this review to celebrate the 50th anniversary of the Protein Data Bank, we present our own experiences in developing and applying methods of visualization and analysis to the ever-expanding archive of protein and nucleic acid structures in the worldwide Protein Data Bank. Across that timespan, Jane and David Richardson have concentrated on the organization inside and between the macromolecules, with ribbons to show the overall backbone “fold” and contact dots to show how the all-atom details fit together locally. David Goodsell has explored surface-based representations to present and explore biological subjects that range from molecules to cells. This review concludes with some ideas about the current challenges being addressed by the field of biomolecular visualization.

Across most of the history of the PDB, we have shared a common passion for understanding and communicating the beauty and functional complexity of macromolecular 3D structures. Our approaches to this endless and satisfying challenge differ in a fundamental way, but because of that are highly complementary. Goodsell is fascinated with creating intuitive representations of what a single protein or a set of interacting molecules would look like, at a range of size scales, if it were possible to see their surfaces directly. The Richardsons are fascinated with how to show the internal organization—the connected backbone “fold” in 3D and the specific atom–atom contacts that determine the anatomy and individuality of the molecule, or the amazing specificity of interaction with another molecule. To celebrate the 50th anniversary of the Protein Data Bank (PDB) archive, we will give personal tours of our own experiences in seeing the PDB.

The Richardson perspective

Hand-drawn ribbons

In the beginning, even before the PDB, visualizations of proteins in scientific reports were mostly either smoothed

blobs of density to show helices and connectivity or full-detail Watson–Kendrew brass stick models (1, 2), photographed from the physically built models, and limited to black and white. The professional artist and scientific illustrator who first turned that information into appealing, convincingly understandable drawings was Irving Geis. Figure 1A shows Geis’ painting of the myoglobin model at 2-Å resolution done for *Scientific American* (3), which conveys the depth cueing and 3D relationships, atom types, H-bonds, and even hydrogens much more effectively than a photograph of the actual brass model. Figure 1B shows his representation of the deoxy-hemoglobin tetramer (4), with the numbered C α trace inside a partly transparent, smoothed tube of the low-resolution electron density map. He also drew oxyhemoglobin separately, but since animation was not possible on the printed page, he could only suggest the large conformational change by outlining in red the smaller central opening in the oxy form.

Those myoglobin and hemoglobin structures were what motivated Dave Richardson to take on solving the structure of Staphylococcal nuclease as his PhD project in inorganic (!) chemistry with Al Cotton at MIT, where Jane then joined the group as a technician. Stories about our punch-card days in the 1960s can be found in (5, 6).

Later, when Chris Anfinsen persuaded Jane to take on the task of drawing and classifying all the extant protein structures for *Advances in Protein Chemistry* (7), she got a number of useful tips from Geis, such as showing backbone and side chain in different colors, and that sometimes inconsistent representations can work better: e.g., Geis’ full-bond red for backbone carbonyls and all of the heme *versus* half-bond colors elsewhere (see Fig. 1A) and Jane’s thick β ribbons *versus* thin α ribbons *versus* round-rope loops (see ribbon figures below). In the other direction, Geis later made a painting for his friend Fred Richards that included a moonlit ribonuclease S in Jane’s ribbon style (8). His ribonuclease by itself can be seen in the digital Geis archive at the RCSB (<http://pdb101.rcsb.org/sci-art/geis-archive/gallery/rcsb-0002-ribonuclease-s>), and the entire painting can be seen in (8), by Richards’ grandson Ben Lillie who owns the painting.

When Dave and Jane Richardson built a brass model of their Staphylococcal nuclease structure (1sns) (9), they were struggling to show the backbone “fold” clearly. Chris Anfinsen, for whom the nuclease was a disulfide-free model system to study protein folding, suggested we tie 1/4" tygon tubing along the

* For correspondence: David S. Goodsell, goodsell@scripps.edu; Jane S. Richardson, jsr@kinemage.biochem.duke.edu.



Jane S. Richardson and David C. Richardson are Professors of Biochemistry at Duke University and are scientific and life partners since before the PDB.



David S. Goodsell is a Professor of Computational Biology, Department of Integrative Structural and Computational Biology, Scripps Research and is a Research Professor at the RCSB PDB, Rutgers University.

All three authors have throughout their impactful careers created many new and intuitively satisfying visual representations of macromolecular structure. Their paradigm-shifting images of the three-dimensional organization within and between individual macromolecules, molecular machines, and cellular contexts have helped empower biological and biomedical research and have enhanced public interest and understanding.

backbone, fill it with fluorescent dye, and view it under UV light. (Fig. 2, A and B). The result looks much like the earliest computer graphics. Jane made a primitive “worm” drawing from that UV-lit image (Fig. 2C), later done in a form more transitional toward ribbons for a few small structures such as insulin (Fig. 2D), and Dave learned to draw a worm of the hemoglobin tetramer on the chalkboard for his class. Those were precursors to the ribbon drawings, further inspired by analogy with MC Escher’s “ribbon” or “rind” drawings of human heads that alternate a spiral strip of surface with an equal opening, cleverly placed to include just enough of eye, nose, mouth, hair, and shoulders so that you perceive the face as intact and also see the back side of the head (see explanation in (10), or just google “MC Escher ribbon head”). Jane worked until she could see, and then hopefully make viewers see, her β ribbons as continuously H-bonded β sheets, not as a handful of separate arrows.

Not being even an amateur artist, Jane spent an entire year developing effective conventions for the ribbons, practicing how to draw them, and drawing the 75 distinct protein domain structures then known, plus another year writing the long review article they illustrate (7). Fortunately, she was then a nearly invisible Associate whose time was not closely scrutinized except by Dave. She had built or helped build the initial models for four proteins (1sns staphylococcal nuclease, 1gch γ -chymotrypsin, 1sod/2sod Cu,Zn superoxide dismutase, and 1ebx erabutoxin B), and so she was able to visualize in her head the peptide orientations and H-bonding from stereo images of the α -trace. The pioneering computer graphics system in Richard Feldman’s laboratory at National Institutes of Health in Bethesda allowed interactive choice of viewing direction for

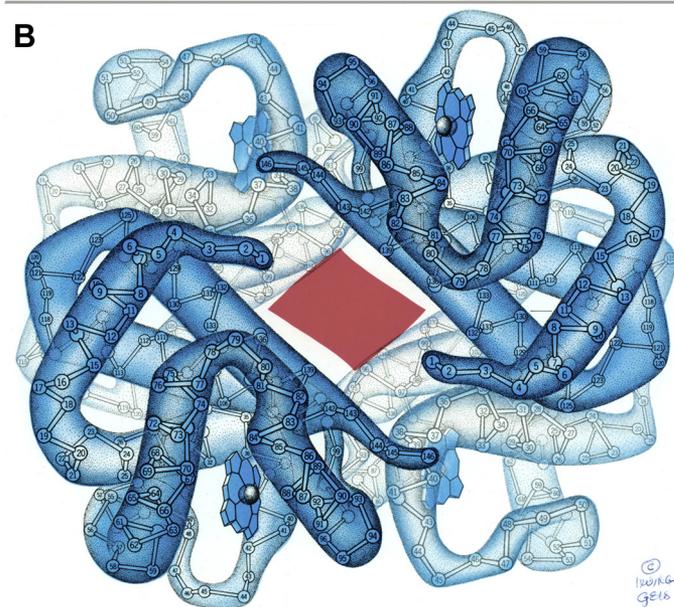
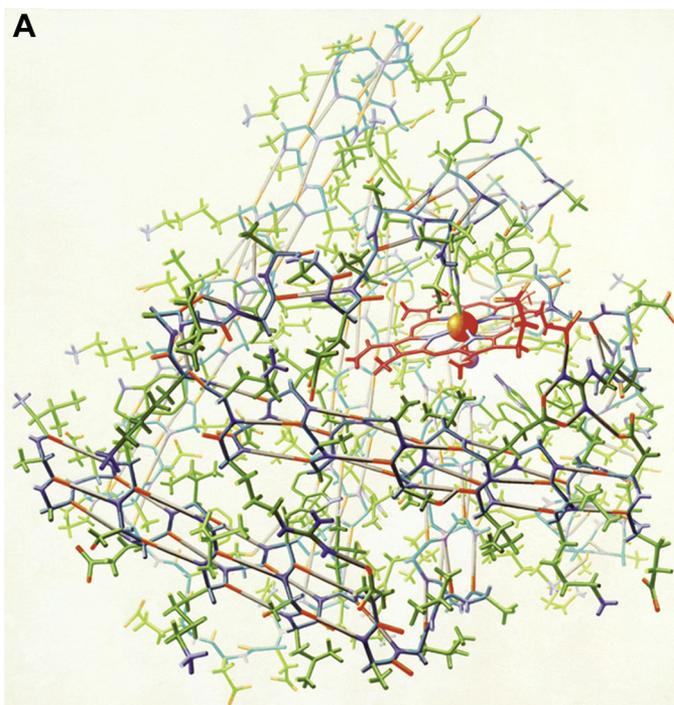


Figure 1. Pre-PDB protein visualizations by Irving Geis. A, painting of the all-atom stick model for myoglobin (including hydrogens), with strong depth cueing and emphasis on the heme ligand and on the helical H-bonds. B, lower-resolution representation of the four chains of deoxy hemoglobin, with the numbered α -trace inside the smoothed electron density map in blue. The red patch shows the narrower central opening in the oxy form. Used with permission from the Geis archive at Howard Hughes Medical Institute (www.hhmi.org); all rights reserved.

a α -trace, mono printout at a consistent scale, plus printout of small α stereo pairs. Jane drove up there often from Duke, displaying coordinates of different structures either from the PDB or from Feldman’s microfiche atlas (11) and bringing home the printouts.

The ribbon-drawing conventions for α helix, β sheet, and nonrepetitive loops are shown in Figure 3A, with the α -trace

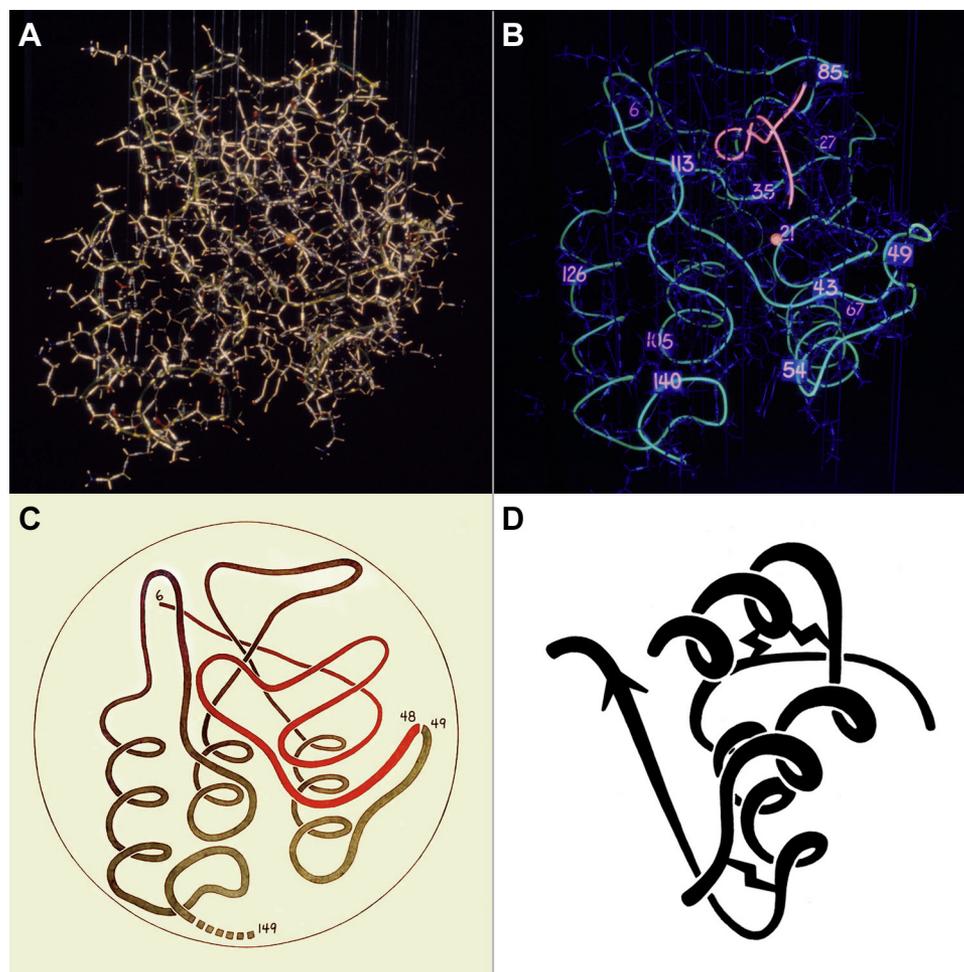


Figure 2. Preribbon Richardson representations. *A*, photo of the brass model of Staphylococcal nuclease in room lighting. *B*, photo in UV lighting of fluorescent-dye-filled tygon tubing tied along the backbone of that model. *C*, hand-drawn “worm” of the nuclease backbone, as cleaved into two parts that can still fold up. *D*, more sophisticated worm drawing of the two-chain monomer of insulin.

behind for the loop case; [Figure 3B](#) shows the pencil sketch for Staphylococcal nuclease on tracing paper over the mono Ca printout; each such sketch involved much trial-and-error erasing and redrawing until it looked right. The final pen-and-ink drawing was done on heavy tracing paper over the pencil drawing, then photographed by Dave to make the high-contrast, touched-up negative used for production of the review article ([Fig. 3D](#) shows a page of similar Greek-key fold structures) and distributed by us as a coloring book. Looking right also involves allowing for perceptual illusions caused by binocular vision, such as seeing both sides of an edge-on ribbon at once and seeing a bit under each side of the front strand when a strand behind crosses at a low angle ([Fig. 3C](#)); both of those cases occur in the drawing of our Cu,Zn superoxide dismutase (“SOD”; (12)), shown in [Figure 3D](#).

The first shaded versions were done with stick-on plastic cut to fit with an exacto knife, in halftone dots for black and white and in color for slides. Stick figures were often added for ligands or for critical sidechains. Dave’s mother, a professional artist (see (13) for an example of her work), encouraged Jane to try other media: pastels, which are wonderful for enlarged

versions but were done only for TIM ([Fig. 4A](#); 1tim; (14)) and for SOD ([Fig. 3D](#)) because it is so hard to keep from smudging them; scratchboard ([Fig. 4B](#)); stained glass; and even a small sculpture of glued-together pieces cut from a 1 ¼-in diameter wooden closet pole with our miter-bevel saw ([Fig. 4C](#)), which made the cover of *Biophysical Journal* (15). Other people also make ribbon sculptures, such as Byron Rubin (<http://5reed.edu/reed-magazine/articles/2019/byron-rubin-protein-sculptor.html>), Julian Voss-Andrea (<http://julianvossandrea.com>), and Bathsheba Grossman (<https://bathsheba.com/crystal/>). A recent Duke Library exhibit (<https://exhibits.library.duke.edu/exhibits/show/invisible/case04>) about our work showed that ribbon drawings go surprisingly well with gothic arches ([Fig. 4D](#)).

Many of our hand-drawn ribbons and our computer graphics images are available with open license on Wikimedia Commons at <https://commons.wikimedia.org/wiki/User:Dcrjsr>. The TIM barrel drawing ([Fig. 4A](#)) was Wikipedia Picture of the Day on November 19, 2009. More information and links are available on Wikipedia pages for Ribbon diagram, Kinemage, and Jane S. Richardson.

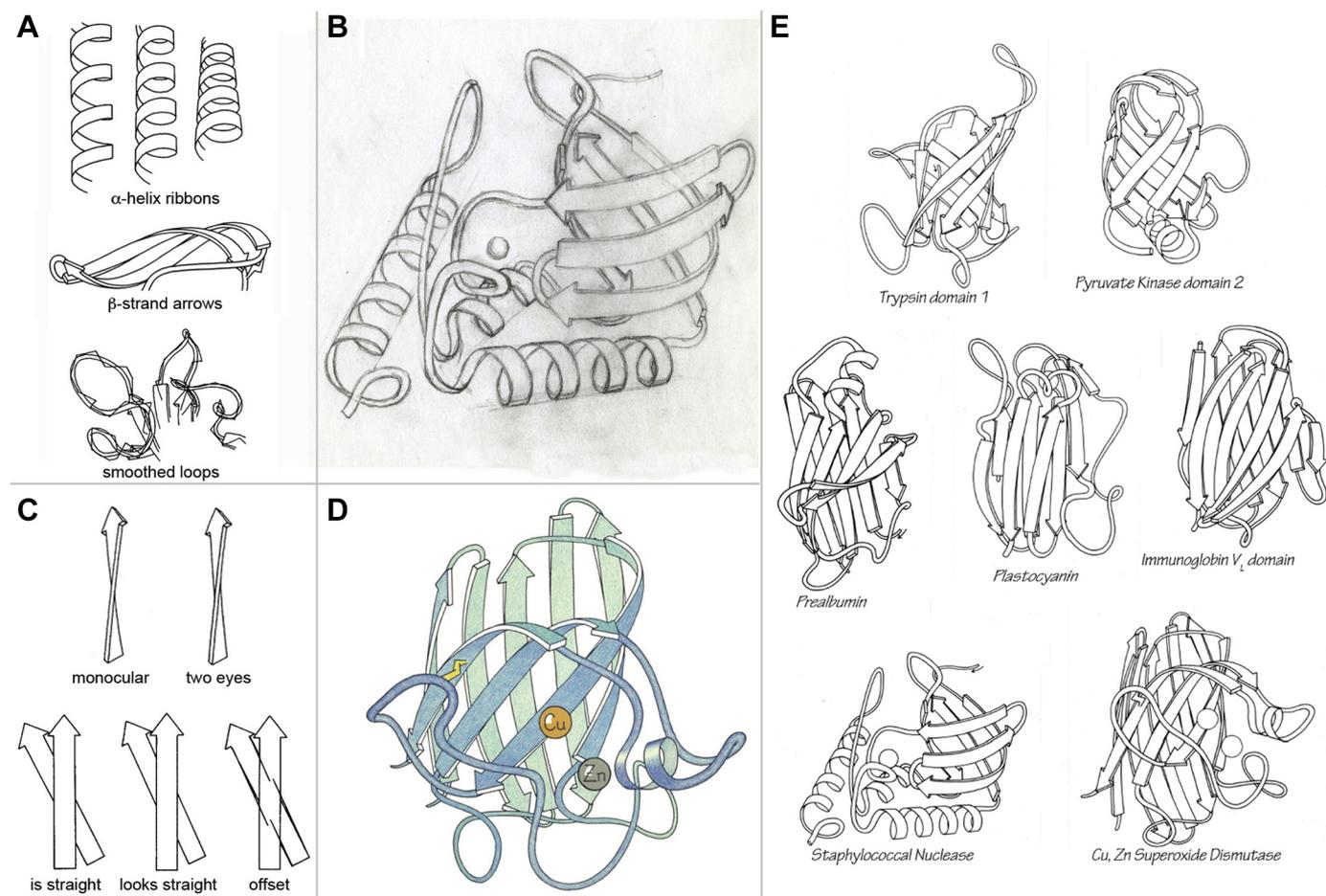


Figure 3. Jane Richardson's ribbon drawing process as of 1980. A, the different conventions for drawing α helix, β sheet, and loops. B, the pencil sketch for Staphylococcal nuclease, done on tracing paper over the scaled Ca plot. C, two optical illusions compensated for to draw convincing ribbons. D, the pastel-colored ribbon for Cu,Zn superoxide dismutase (SOD) that corrects for both illusions in (C). E, page 270 from reference (7), including the Greek key folds of Staph. nuclease, of SOD, and of an immunoglobulin V_L domain.

Computer graphics: Ribbons and all-atom contact dots

Our first exposure to interactive use of computer graphics for crystallography itself was the graphics laboratory in Fred Brooks' computer science department at UNC Chapel Hill. The user console of GRIP-75 (Fig. 5A) had both stereo and smooth rotation, and a wonderful array of five knob, slider, and "toothpick" controls, the first computer system capable of fitting an atomic model into an electron density map (16). The calculations were done on an IBM360 that filled the room behind. We were guinea pigs in its development, and our 2-Å resolution structure of Cu,Zn superoxide dismutase (2sod; "SOD") was the first protein crystallographic model built on computer graphics before building a physical model (17, 18). Figure 5B is a snapshot from the GRIP-75 screen, showing the single SS bond in SOD. A parade of users followed until GRIP-75 was gradually superseded by Frodo (19), which could run on hardware accessible to individual laboratories. We stayed on for 20 years, providing driving problems for the UNC graphics laboratory and collaborative graphics projects. That was enormous fun and taught us a great deal about making our own graphics systems.

When we got our own Evans&Sutherland PS300 calligraphic display, people were starting to try out a variety of

ribbon-like representations, but most were too literal and choppy. The first version we really liked was in Mike Carson's Ribbons program (20). He gave us his algorithm for smooth B-splines with offsets for helix and turn radius, and Dave coded it as multistrand vector ribbons such as the intertwined dimer of Trp repressor in Figure 6A. Then in 1990 we realized that the first Apple home desktop computer could smoothly rotate 500 vectors—enough to show interesting molecular visualizations in 3D or simple animations of conformational change, accessible to anyone. On a hiking trip in the Sierras we worked out the framework for "kinemage" format, including that it is human editable and that any display point clicked on will tell you its identity. Dave then wrote the Mage program to display kinemages, with many interactive viewing and even editing options (21); about the same time RasMol was also developed for home computers (22). Kinemages became a major part of the then-revolutionary electronic supplement to the new *Protein Science* journal, distributed each month on 4K diskettes. The first examples were very simple: an active site or overall C α s, such as animating the hinge motions of T4 lysozyme domains to illustrate a paper in that journal issue (Fig. 6B). As the capabilities of personal computers rapidly



Figure 4. A gallery of ribbons drawings in other media. *A*, the large, pastel-colored drawing of the $(\alpha/\beta)_8$ "TIM" barrel of triose phosphate isomerase (1tim). *B*, lighted drawing of ribonuclease S, done on scratchboard. *C*, wooden model of expanded Ca-trace for basic pancreatic trypsin inhibitor (BPTI); pieces cut and glued from closet pole. *D*, an enlarged piece from the TIM drawing, on the wall of a Duke Library exhibit about the Richardson's work over 50 years at Duke.

evolved, larger models and even ribbons could be shown interactively (Fig. 6C).

These days, ribbon graphics are limited in size and complexity mainly by what can be usefully perceived by the viewer, rather than by what can be calculated and displayed interactively, and other new representations have taken over for very large molecules, complexes, and even whole cells (see the Goodsell perspective). A gallery of some favorite mid-size ribbons of proteins and complexes is shown in Figure 7, A–F, from a small trimer to protein–DNA complexes and

membrane proteins. The first two of these were created in Mage, and the rest in KiNG (23, 24), a Java kinemage viewer for online as well as local use and with many additional editing and remodeling features. An interesting sideline is that the kinemage format is not limited to molecules but is expressed as geometric primitives (line, point, triangle, etc.). Kinemages have been made for social networks, food webs, geodesic domes, a map of the world, etc. One example plotted measurements of a vulture flight path in 3D, which unexpectedly showed that vultures do not actually circle but fly in

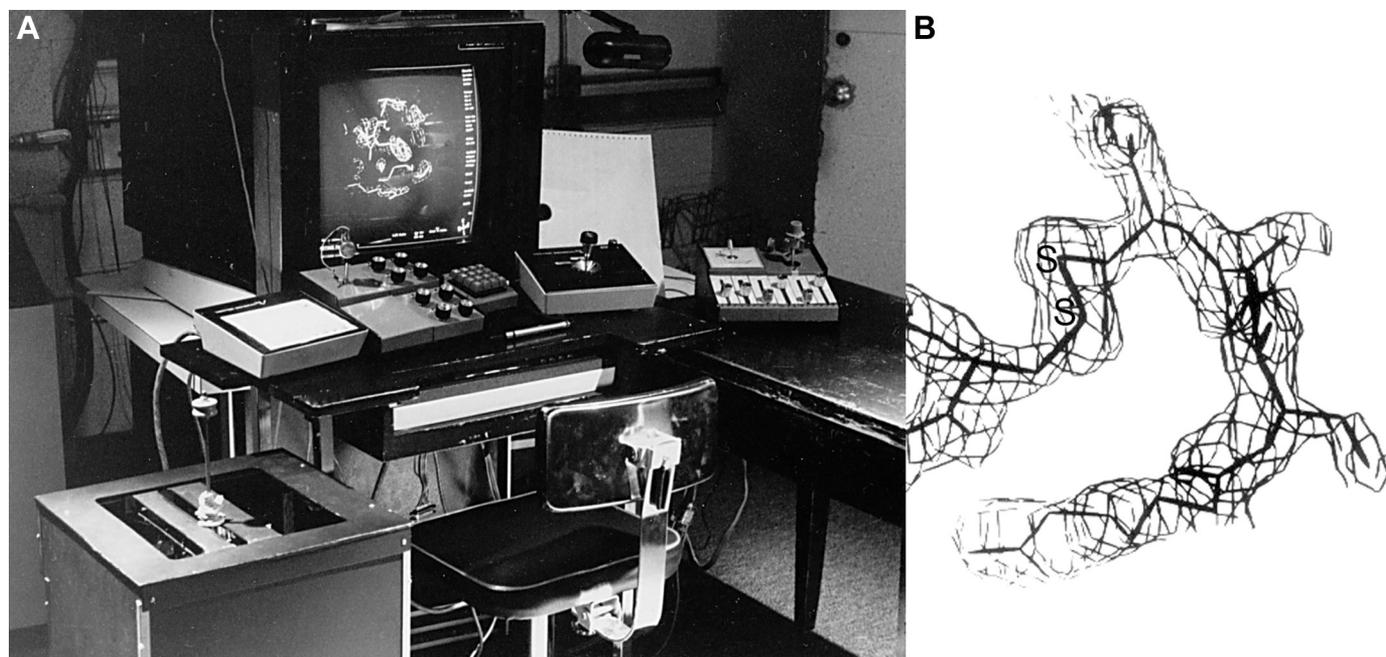


Figure 5. The GRIP-75 interactive protein model-building system at the UNC computer graphics laboratory. *A*, the user console during a session of fitting a model into electron density, as first done for Cu,Zn SOD. *B*, a closeup of GRIP-75 model and density for the single SS bond in SOD.

squared-off patterns parallel and perpendicular to wind direction, just as a small plane would.

All-atom contact analysis

After working on early protein design, which could then achieve correct secondary structures and topologies but not well-ordered native-like structure (15), we realized there was one of the inherent problems we could work to overcome: that if we changed from the standard of implicit “united-atom” treatment to including all explicit H atoms, and then paid attention to their contacts, we could provide a realistic, quantifiable, and visualizable analysis of detailed atom–atom internal packing. The atom–atom contact dots we use are

more or less the inverse of Connolly dots that outline the accessible surface of molecules (25). Our Reduce program places all explicit H atoms and optimizes their rotations and interaction networks for H-bonds, van der Waals (vdW), and clashing overlaps (26). This all-atom contact analysis visualizes how the atomic surfaces interact, rather than traditional energetics that is pairwise from atom centers and is only numerical, not visual.

Figure 8A shows the undulating, interdigitated dots for vdW contacts around the H atoms on a proline sidechain at high resolution (1 ϵ m; (27)). Green dots are for perfect atom–atom contact, blue for very close but not quite touching, and yellow for just a bit tight. Figure 8B shows a transition-state inhibitor

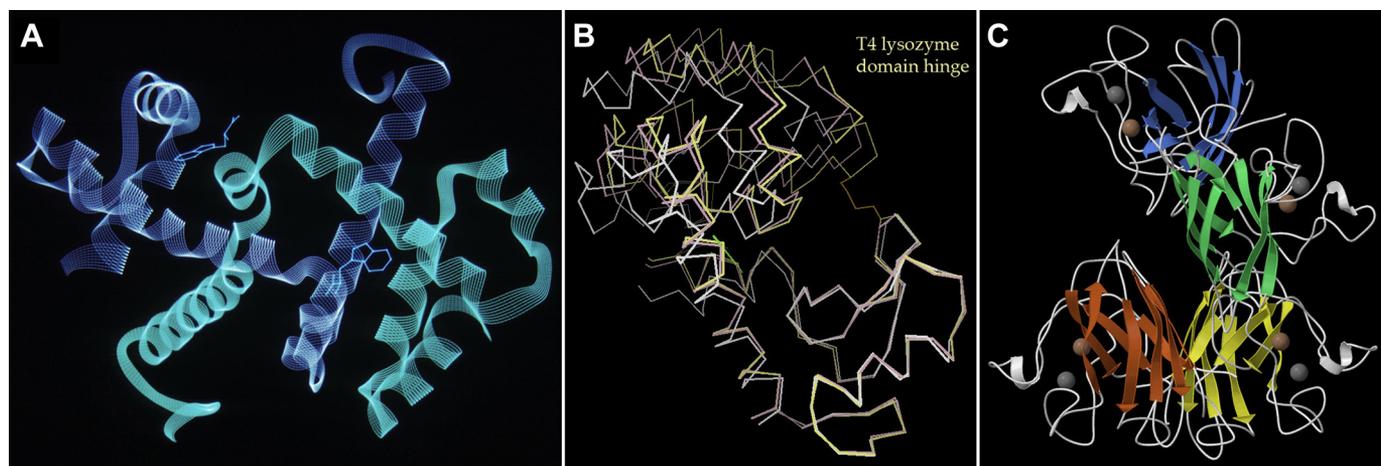


Figure 6. Dave Richardson’s early computer graphics. *A*, multistrand ribbons for the intertwined Trp repressor dimer, on the Evans & Sutherland PS300 calligraphic display. *B*, an animation of conformational change in T4 lysozyme, in his Mage program, for the *Protein Science* journal and the first Apple computers in 1992. *C*, a later Mage display of ribbons for the two SOD dimers in the crystal asymmetric unit (metals as zinc-colored and copper-colored balls).

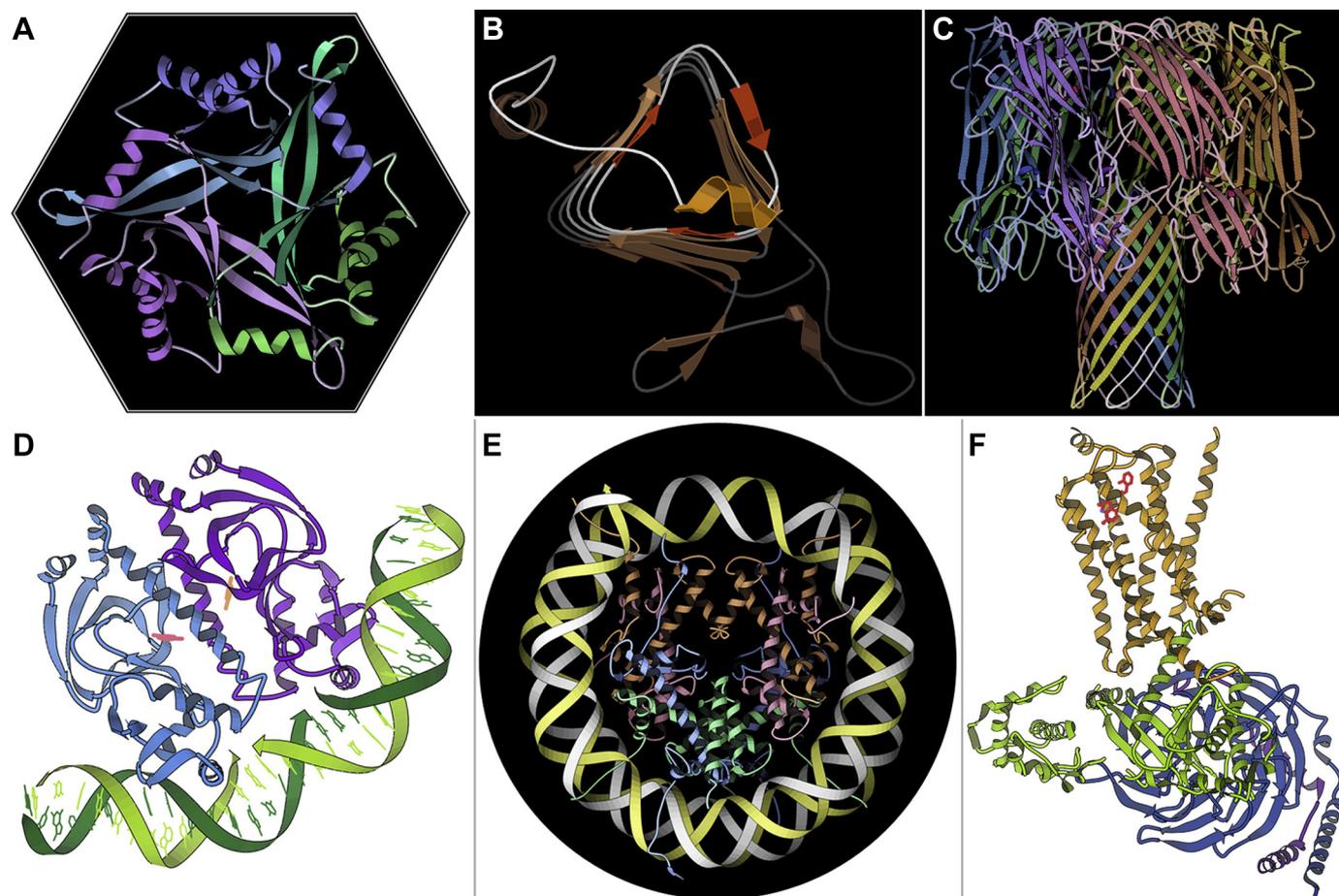


Figure 7. A gallery of computer-drawn ribbons, the first two from Mage and the rest from KiNG. *A*, interdigitated trimer of human CUTA, with each chain in a different color; 1xk8. *B*, three-sided, left-handed β helix of an archaeal carbonic anhydrase, emphasizing the helical loop that protects this end from β - β aggregation; 1qre. *C*, α -Hemolysin 7-mer, with its transmembrane β -barrel pore at bottom; 7ahL. *D*, CAP protein dimer binding and bending DNA; 1cgp. *E*, nucleosome, with central core of histone proteins (green, blue, peach, pink) encircled by two turns of DNA (yellow, white); 1aoi. *F*, seven-helix transmembrane β 2-adrenergic receptor (gold) with its in-cell signaling G protein chains (green, blue, purple); 3sn6.

held very tightly in ribonuclease (1ruv; (28)) by 13 good H-bonds (pillows of pale green dots). Figure 8C shows the base-stacking vdW contacts and base-pair H-bonds in an RNA double helix of a mutant G riboswitch at 1.35-Å resolution (4fen; (29)). The all-atom contact system is confirmed by the fact that the well-ordered parts of high-resolution experimental structures have almost no clashes, good density of H-bonds, and interdigitated hydrogens to give excellent vdW contacts. That system also provided a new and powerful way to find misfit local conformations in the experimental models, primarily by “clashes,” defined as unfavorable overlaps ≥ 0.4 Å. That discovery motivated our establishment of the MolProbity validation web service, discussed below.

RNA in seven dimensions

After the first ribosomes came out in 2000, our new student Laura Murray led us into revisiting our early fascination with RNA at MIT in the 1960s, when our friend Sung-Huo Kim solved the first structure of a complexly folded RNA (Fig. 9A), dramatically revising expectations about RNA structure beyond flat secondary-structure layouts such as the tRNA “cloverleaf.” The ribosome added more than an order of magnitude more data, enabling a realistic analysis of the

hard-to-see and inherently high-dimensional RNA backbone conformations. There are six variable dihedral angles in a PO_4 -to- PO_4 RNA nucleotide, but we discovered that the seven-dihedral sugar-to-sugar suite parsing of the backbone shows better clustering because it includes more backbone interaction than the nucleotide as well as being directly influenced by the interaction of successive bases. After trying a system of cross-picking between two 3D windows, Dave in Mage and Vincent Chen in KiNG developed an integrated high-dimensional technique for displaying and analyzing that data. It works by defining “views” for specific triples of parameters, choosing and coloring the points in an apparent cluster (Fig. 9B), and then checking cluster validity in other 3D combinations and also in parallel coordinates (Fig. 9C). That work culminated in a collaboration of five different laboratories to agree on a consensus library of RNA backbone conformers with two-character names or !! (bang-bang) for outliers (30). The 7D graphics are for doing the research, whereas the results allow anyone to meaningfully name and visualize an RNA motif such as a stack switch, an S-motif, or a GNRA tetraloop (Fig. 9D). When combined with all-atom contacts, such an image is even more informative and can guide correction of RNA validation outliers.

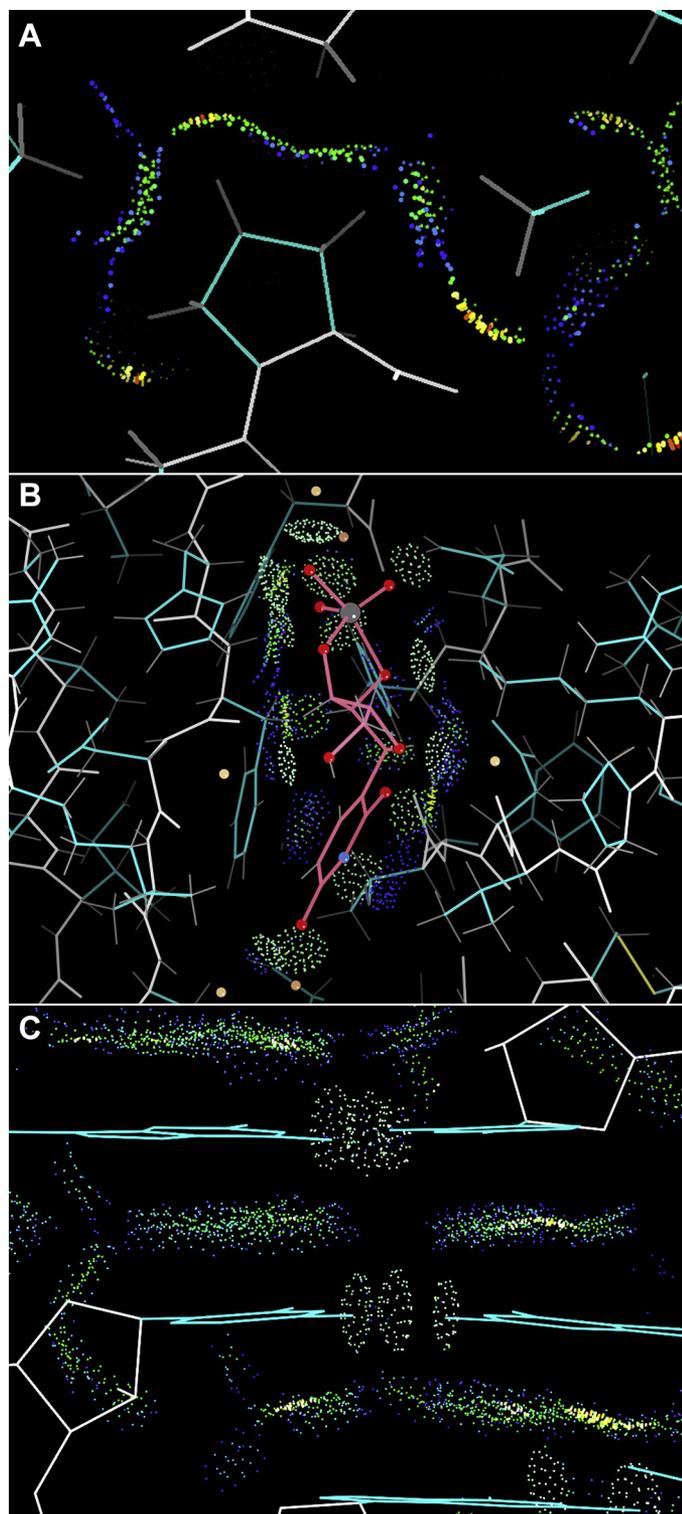


Figure 8. Using all-atom contact dots to understand an interaction. *A*, well-matched, interdigitated, hydrophobic van der Waals surfaces of the explicit hydrogens on a proline ring and its neighbors in elastase at 1.5 Å; 1ezm. *B*, uridine vanadate transition-state inhibitor (pink) in ribonuclease A at 1.3 Å is tightly held by 13 good H-bonds; 1ruv. *C*, large areas of van der Waals contact for stacked base pairs in a G-riboswitch at 1.35 Å; 4fen.

Virtual reality in the DiVE: NMR RDC curves

The DiVE (Duke Immersive Virtual Environment) 3-m³ virtual reality lets us explore whether immersive virtual reality could

be set up to enable molecular research as well as to enhance intuition through a compelling visual and interactive experience. Working with the Visualization Technology group at Duke, we jointly programmed KinImmerse display of the kinemage format in the DiVE, with atom selection, identification, and measurements (Fig. 10A; (31)). Our test case was bringing to NMR structural ensembles the direct visual comparison of model to data that we rely on in crystallography by looking at the model within the electron density map. Basic NMR data measures very local atomic relationships: atom-pair distance by nuclear Overhauser effect (NOE) values and atom-pair orientation to the magnetic field by RDC values, and so is well suited to an aligned and zoomed-in perspective within the ensemble. We provided that perspective by a simple but novel tool called cocentering (on a specific atom type, which can be toured through the sequence) to then compare in each residue how well each model fits the relevant local data. NOE data are easily represented as a vector between the two specific atoms, but an RDC (residual dipolar coupling) measurement means the atom–atom orientation lies somewhere along a pair of curves. Those curves can be drawn around one of the atoms in the ensemble of models (such as the backbone N for an NH RDC), but that was never done until cocentering made it understandable. Figure 10B shows the elegant RDC curves around the cocentered N of Glu 36 in 2jng. In this case some of the models must be wrong, since the H atoms should all be on the same branch of the RDC curve. Which group of models is correct (the white ones) can be decided by other validation criteria such as H-bonding, clashes, and Φ, ψ values. This system was later ported back to single-screen viewing, but might not have been thought of there.

Validation markup and corrections, including SARS-CoV-2 molecules

Our MolProbity web service adds other validation measures to the all-atom contacts (bond length, bond angle, and C β deviations, Ramachandran Φ, ψ dihedrals, sidechain rotamers, and RNA sugar puckers and backbone conformers). In addition to sequence-based charts and summary statistics, it outputs visual markup for the local outliers on an interactive image of the 3D model. Each type of outlier has a distinct color and shape, and they stand out clearly on a simple C α -backbone model (Fig. 11A). The user can then zoom in and overlay the electron density map to see the full details, especially for a cluster of multiple outliers close in 3D, such as the helix start in Figure 11B.

MolProbity gradually became considered state of the art in model validation. However, model validation, if considered just as flagging problems, would be a very depressing and unpopular activity—what makes it positive and even exciting is the prospect of helping to fix many of those problems. The cluster of clashes and rotamer outliers in 1LpL (32) at upper left in Figure 11A, and expanded in Figure 11B along with the map and difference-density peaks, was diagnosed as an incorrect switch of backbone *versus* sidechain at the 136 C α atom. Local rebuilding with the correct switch choice at 136 and alternate conformations for Lys 137 removed those outliers and enabled building an additional turn of α helix into the previously

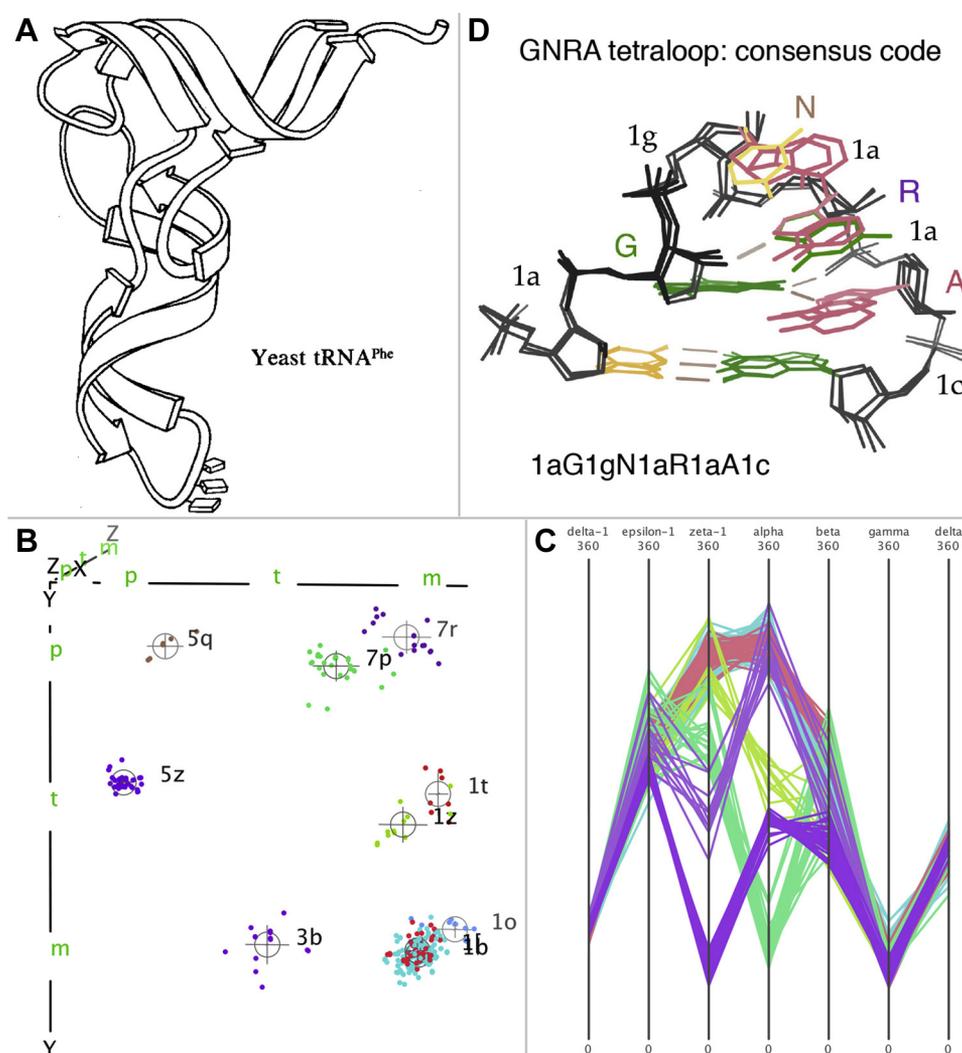


Figure 9. Visualizations of RNA structure and backbone conformations in 3D and 7D. *A*, hand-drawn ribbon of yeast tRNA Phe, with its two pairs of stacked stem-loops and the anticodon triplet bases at the *bottom*. *B*, ten colored and labeled clusters in the ζ , α , γ choice of 3D view for seven-dihedral RNA backbone "suite" conformers (the **1b** and **1l** clusters and their labels overlap completely in this view, because they differ only in the β dimension); displayed in KING. *C*, the same ten clusters shown in parallel coordinates, accessible by clicking the "/" key. *D*, three superimposed examples of GNRA tetraloops labeled with their two-character conformer names alternated with their base identity, thus neatly describable as 1aG1gN1aR1aA1c.

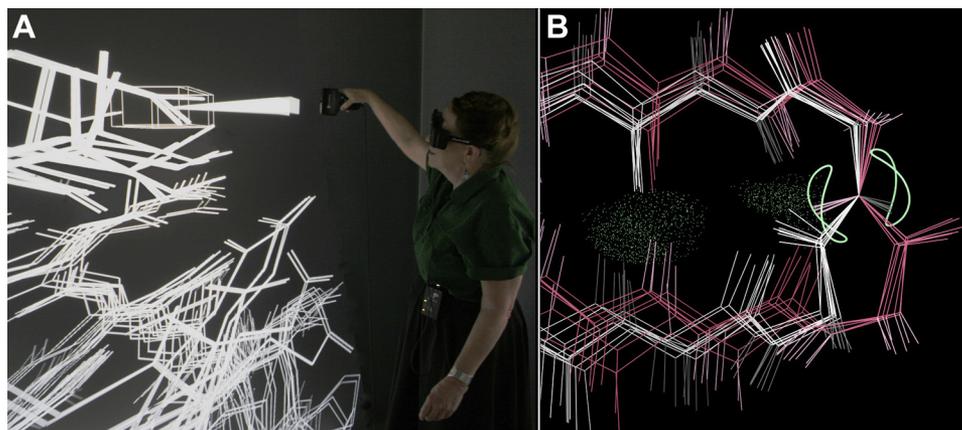


Figure 10. Doing macromolecular research in immersive VR (virtual reality). *A*, Jane selecting one bond of a multimodel NMR ensemble, using an early version of our software in the Duke DIVE. *B*, in KinImmerse (31), the NMR ensemble of 2jng cocentered on a backbone N atom to analyze fit to the elegant two-branch curve (pale green) that represents an experimental RDC (residual dipolar coupling) measurement. In this case, the NH atoms of the pink models lie on the wrong branch of the curve and the white ones are correct (see text).

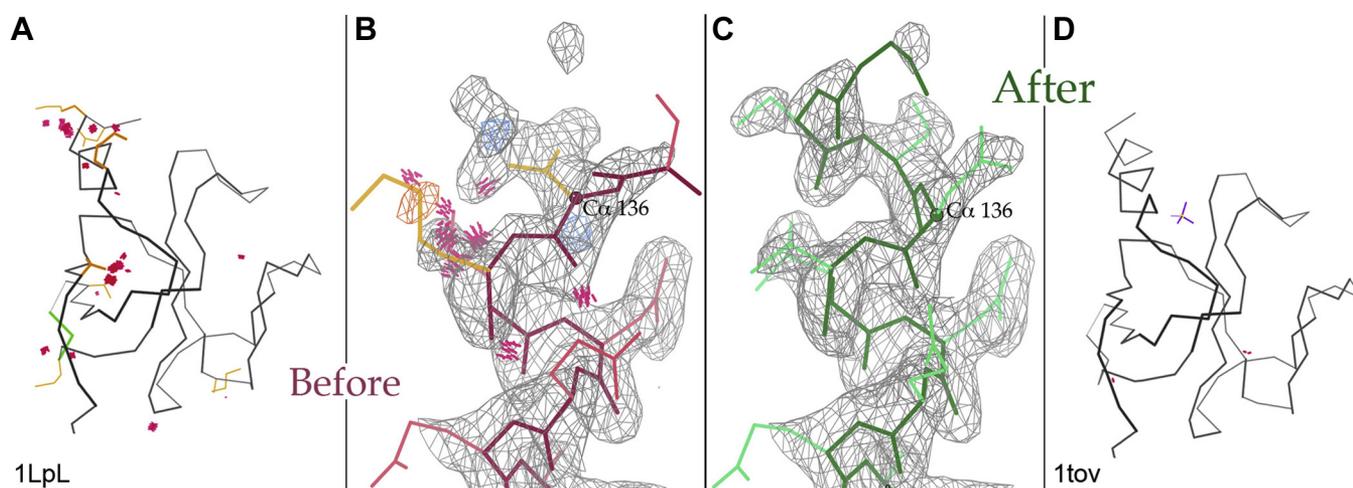


Figure 11. Using MolProbity validation outliers to guide structure improvement. *A*, Kinemage display of all-atom clashes (clusters of hot pink spikes) and sidechain rotamer outliers (gold) in context of the Ca backbone in 3D, for 1LpL. *B*, zooming in on the largest group of outliers, along with full model, electron-density map (gray), and difference density peaks (blue and orange). At Ca 136, the first residue modeled, density extends well beyond the Asp sidechain and is too short for the backbone. *C*, after correction of that sidechain-mainchain switch, the outliers are gone and another turn of helix can be modeled. *D*, the final result (deposited as 1tov) has only two small clashes, a new sulfate, and a 4% lower Rfree.

unoccupied density (Fig. 11C). Our proof-of-concept correction of this 1.77-Å structure lowered the R_{work} by only 1% but the R_{free} cross-validation measure by 4% (33), indicating that the model changes are true improvements: only two tiny clashes and a new sulfate replacing a tight cluster of waters (Fig. 11D).

Over the past 20 years, this system of validations and visualizations has enabled both individual structural biologists and automated programs to significantly improve the quality of worldwide PDB (wwPDB) depositions at the most common resolutions from sub-1 to about 2.5 Å. Recently, however, both X-ray crystallography and most especially cryo-EM (34) have succeeded in solving large numbers of huge, dynamic “molecular machines,” mostly at resolutions poorer than 2.5 Å.

These new structures make revolutionary contributions to our knowledge of biology, but that lower resolution introduces new systematic errors insensitive to traditional validation metrics. We and others are developing new types of validation that can detect local errors even when the traditional outliers have been artificially refined away, often pushing them to a legal but incorrect conformation. Our most effective new metric so far is CaBLAM, which tests whether the modeled peptide orientations are consistent with the surrounding five-residue stretch of Ca backbone (35).

In 2020, structural biologists mobilized to produce and release structures of the SARS-CoV-2 molecules with unprecedented speed, and we used all of our tools to identify the occasional errors at important sites in those structures that were provably correctable, in collaboration with Tristan Croll and his ISOLDE program described below. The problems included stretches of sequence misalignment, ions with the wrong charge, a mismodeled SS bond that distorted the interface with an antibody, and many individual peptide reorientations or *cis-trans* isomerizations that mattered. Visual communication was essential to the success and speed of this process, by convincingly illustrating the recommended changes to the depositors, encouraging them to make those changes in their own models and update their PDB entry. Recently the PDB made it possible for the depositor to update coordinates with a new version number not a new PDB code. They prioritized the new SARS-CoV-2 versioning, and the whole process took just a few weeks from the initial release to an improved model available for download (36). Figure 12 shows an example of documentation for a simple change, where two successive CaBLAM outliers (magenta lines) suggest a large rotation of the central peptide (red ball on the carbonyl O), allowing the loop to settle into a near-ideal β hairpin and enabling two strong H-bonds to another chain in this polymerase complex that produces the new viral genomes (7bv2; (37)).

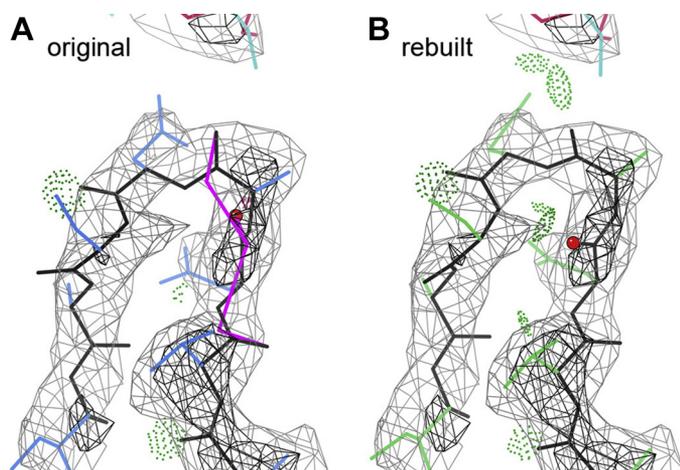


Figure 12. CaBLAM outliers to flag problems at 2.5 to 4 Å, applied to SARS-CoV-2 structures. *A*, original model of a loop in Nsp8 of the 7bv2 RNA-dependent RNA polymerase structure, with two adjacent CaBLAM outliers and only three H-bonds. *B*, with a large reorientation of the central peptide of those outliers, this loop forms a near-ideal β -hairpin with seven H-bonds including two to an adjacent chain in the complex (7bv2 v2.0).

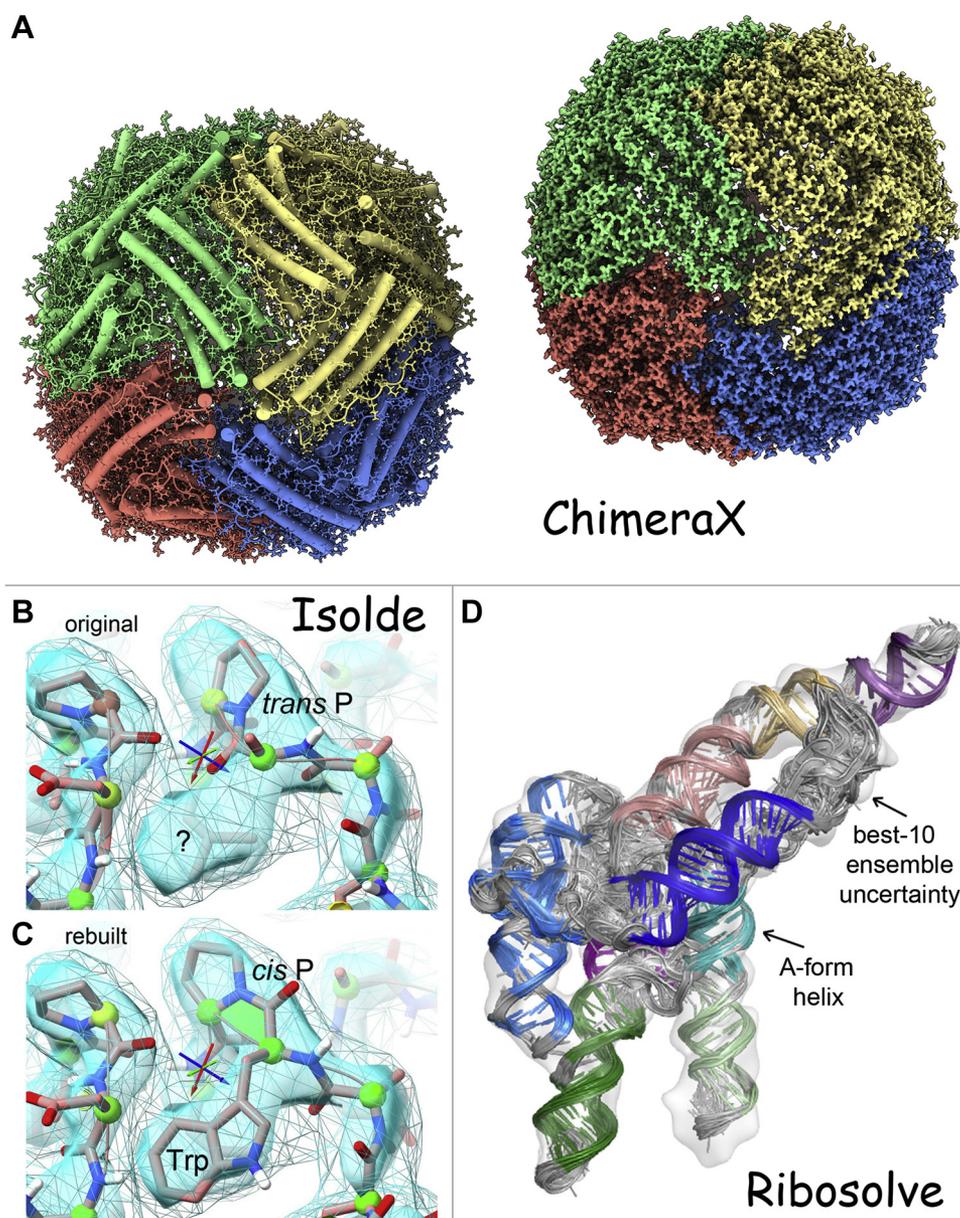


Figure 13. Favorite current visualization types. *A*, from ChimeraX, fast-ambient-occlusion rendering of surface shape and shadows for helix-cylinders or density map surface of apoferritin in 7a4m at 1.23-Å resolution. *B* and *C*, from ISOLDE user-guided molecular dynamics on the 7bv2 SARS-CoV-2 polymerase, correction of a *trans* Pro in *B* that should have been *cis*, allowing the adjacent Trp sidechain to occupy its very clear density in *C*. *D*, a visualization of model uncertainty from the new Ribosome pipeline that distinguishes fairly clear A-form helix from uncertainty (gray) in the best ten diverse Rosetta models out of thousands for the hc16 ligase product cryo-EM structure (6wLn), at 10-Å resolution and with an estimated ensemble accuracy of 6.3 Å.

Favorite current visualizations

- Fast-ambient-occlusion in ChimeraX (38): Their recently implemented fast-ambient-occlusion computer rendering provides beautifully and comprehensibly shaded surfaces that can be perused at leisure or smoothly rotated in real time for atomic models as big as a ribosome, or for the complex detail of the density-map surface of 7a4m apoferritin at 1.23-Å resolution (39) seen in Figure 13A.
- Tristan Croll's visual model rebuilding in ISOLDE (40), which thoroughly achieves what we were trying to do decades ago (41): Its user-guided molecular dynamics in context of the density map and with continuously updated validation flags (Fig. 13, *B* and *C*) provides a uniquely

powerful way to improve protein model quality at resolutions poorer than about 2.5 Å, including problems ranging from individual *cis-versus-trans* peptides to regions of sequence misalignment (Croll 2021).

- Rhiju Das' visualization of RNA model uncertainty (42): The Ribosome pipeline output includes a visual ensemble of the top ten sufficiently distinct models out of thousands (by a pretty realistic Rosetta score), represented as shown in Figure 13D. Ensembles are common, but their breadth is usually either unrealistically tight or unrealistically broad. The especially challenging aspect going forward will be learning to separate actual motion from model ambiguity and then visualizing those separate components.

- Our all-atom contacts in KiNG: These dot-contact images (see above) are still the best way to show, correct, and comprehend the details of atom–atom contact within or between molecules.
- David Goodsell’s representations: For individual proteins and complexes such as the RCSB’s “Molecule of the month” series (which we helped make available open-license on Wikimedia Commons and from there in many Wikipedia articles), they convey surface shape especially well. For larger molecular environments, his drawings changed the perceptions of both scientists and laymen to appreciate how crowded a cell really is.

The Goodsell perspective

When I was applying for graduate school in 1982, I had the great fortune of having abundant material to spark my interest in structural biology. At the time, the PDB archive contained just over 100 structures. A trip to the library yielded inspiring books like Dickerson and Geis “Structure and Action of Proteins” (4). During one of my interviews for graduate school, one of the professors handed me her well-thumbed photocopy of Jane’s “Protein Anatomy,” and after reading that cover to cover, I knew that I wanted to pursue a career in structural biology.

It was a very exciting time to be in structural biology, and in visualization of structural biology in particular. Affordable hardware for computer graphics was just then becoming available, at a cost that made it accessible to individual laboratories or departments (43). For example, during my graduate work with Richard E. Dickerson at UCLA, I had ready access to an Evans and Sutherland MultiPictureSystem, various pen plotters, and down the hall, a (quite slow) AED raster display. Amazingly, apart from occasional ORTEP diagrams on the pen plotter, this technology was virtually unused. This was no surprise, since effective software was still being developed and was largely the domain of visualization experts. So I started looking for available programs, and also writing my own programs, for displaying and presenting the DNA structures that the laboratory was determining (44). Also, artwork with traditional media has always been a big part of my life, so when the topics were too complex for my budding Fortran skills, I drew diagrams by hand. Two examples are included at upper right in Figure 14.

After meeting Arthur J. Olson at a meeting of the Molecular Graphics Society, I moved to the Scripps Research Institute to focus on computational biology and molecular graphics for my postdoctoral work. Art’s laboratory, to this day, is highly exploratory, using the newest developments in graphics hardware to explore biomolecules in fresh, new ways. When I joined, he had built a strong body of work in molecular animation and in representing biomolecules with surface-based approaches, and in the years since then, he has explored new media such as 3D printing, virtual reality, and augmented reality. During my postdoctoral work, I added some methods for volume rendering and also started work on nonphotorealistic cartoony methods of rendering (45). Perhaps most

importantly, I also became an active member of a creative community of structural biologists, computer scientists, and scientific artists all focused on revealing the wonders of the biomolecular world. This has led to my life-long interest in depicting the size and shape of molecules, and from that, how they function in living systems.

Surface metaphors to explore shape and size

What does a molecule look like? This question has fascinated structural biologists ever since it became clear that molecules have a defined structure, and an active community of artists and scientists have charted an ever-changing path through answering it. The central challenge, of course, is that this question has no direct answer. Given the physics of light and the scale of molecules, we can never “see” an individual molecule, at least as anything more detailed than a point of light. Instead, we are forced to create metaphors for what we might see if the physics were more amenable and use these metaphors to synthesize images of molecules. It is in these metaphors that the fun begins, since the goal is to create an image that somehow captures the properties and personality of the molecule, so that the image will provide a way to learn more about the molecule, its structure, its interactions with other molecules, and ultimately, its function.

As Jane and Dave discuss above, a host of effective metaphors have been used for exploring the internal structure, folding, and detailed atomic interactions of biomolecules, most notably, Jane’s ribbon diagrams and many flavors of covalent bond representations. Most of my work has focused on the overall shape and size of molecules and how they fit together into higher-order assemblies and ultimately into functional viruses and cells. Space-filling metaphors and their many variants are my favorite way to explore these aspects of biomolecules. The power of this metaphor was amply demonstrated by its creator, Linus Pauling, who used it, among other things, to develop the detailed model of the alpha helix and its structural constraints (46). Today, all molecular graphics programs provide options for creating representations of the steric surface of a molecule. A few have shown lasting utility, and a few are seeing increased use as structural biology moves to larger and larger assemblies.

Linus Pauling’s original concept was to center a sphere on each atom, with a size that represents the typical contact distance between noncovalently interacting atoms. When combined with a standard coloring scheme, this simple idea provides a powerful and flexible method for representing the steric bulk of a molecule and understanding how they exclude or fit together with other molecules. The simplicity of the approach also makes them quite easy to incorporate into molecular graphics software.

A variety of specialized surfaces were proposed in the 1970s and 1980s to explore in more detail the way macromolecules interact (Fig. 15) (25, 47). These surfaces are generated by rolling a probe atom everywhere around the molecule and saving up all the points that it touches. Most often, the probe represents a

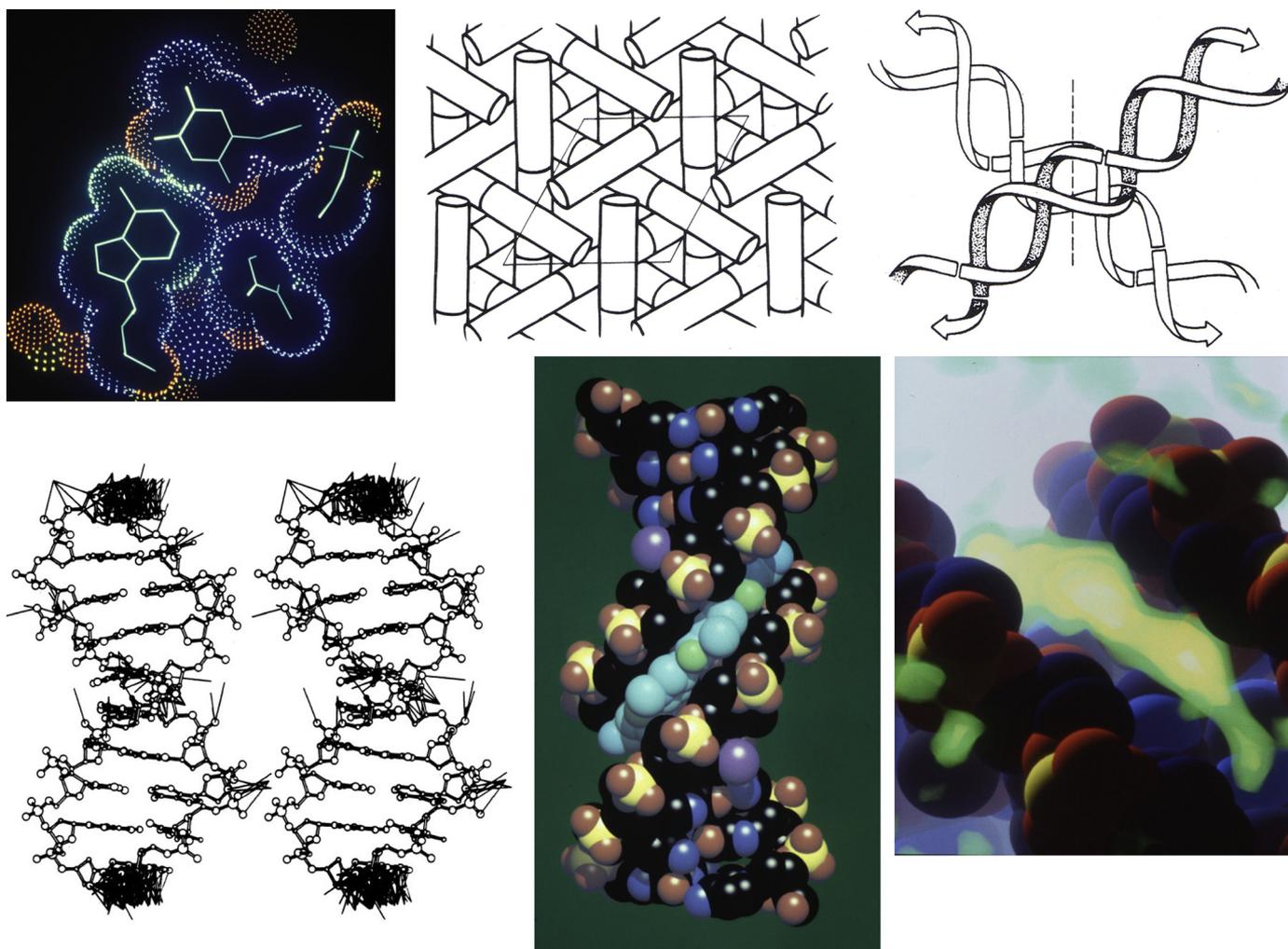


Figure 14. Early experiments in molecular graphics. *Top*, dot surface and covalent skeleton displayed interactively with the Evans and Sutherland MultiPicture System with custom software. Two hand-drawn illustrations of packing of DNA decanucleotides in crystal lattices. *Bottom*, stereo pair of a bond diagram created using ORTEP and output on a pen plotter, with extra lines depicting the contacts between neighboring molecules in a crystal lattice. Raster image of a drug complex with DNA using shiny Phong shading, and volume-rendered image of electron density of a drug in the minor groove of DNA, both created with custom computer graphics software.

water molecule, creating a “solvent-accessible” surface. These surfaces have many uses, such as defining the shape of ligand-binding sites and protein–protein interfaces and providing a quantitative measure of water interaction area for use in hydrophobicity calculations. Additional variants are possible for specialized tasks, for example, saving the surface drawn out by the center of the probe sphere or using spheres of multiple sizes to evaluate a “fractal dimension” of the surface (48).

Structures are getting larger and larger with the current advances in cryoelectron microscopy, and methods are needed for interactive display. A host of simplified surfaces have been conceived and implemented over the years and are increasingly being brought to bear on the problem. For example, Gaussian surfaces approximate the electron density of a molecule by summing Gaussians centered on each atom, then contouring the density. This approach can create representations reminiscent of solvent-accessible surfaces (Fig. 15) and can also be used to create simplified surfaces by tweaking the parameters. The web-based Mol* viewer (molstar.org) at the RCSB PDB site currently uses a smooth Gaussian surface as

the default for very large structures (Fig. 16). This has two advantages. First, since only the outer surface is seen, it greatly reduces the complexity of the image making the overall structure of the assembly interpretable. Second, the representation is compact enough to allow interactive performance.

“Seeing” molecules

Once we have chosen a metaphor for the shape and form of the molecule, we then need to turn our attention to metaphors for “seeing.” Given that these images are synthesized digitally, we have unfettered freedom to make choices based on the goals of the image. The field of computer graphics has long been focused on creating photorealistic depictions, so in the first decades of molecular graphics development, a distinctive “shiny plastic” look was widely used, using techniques of Phong shading (49) to generate images reminiscent of the physical space-filling models used by Pauling and later in many laboratories and classrooms. This style of rendering has proven to be quite effective and has been progressively refined with cast shadows and softer ambient occlusion shading, subtler approaches to the shiny specular

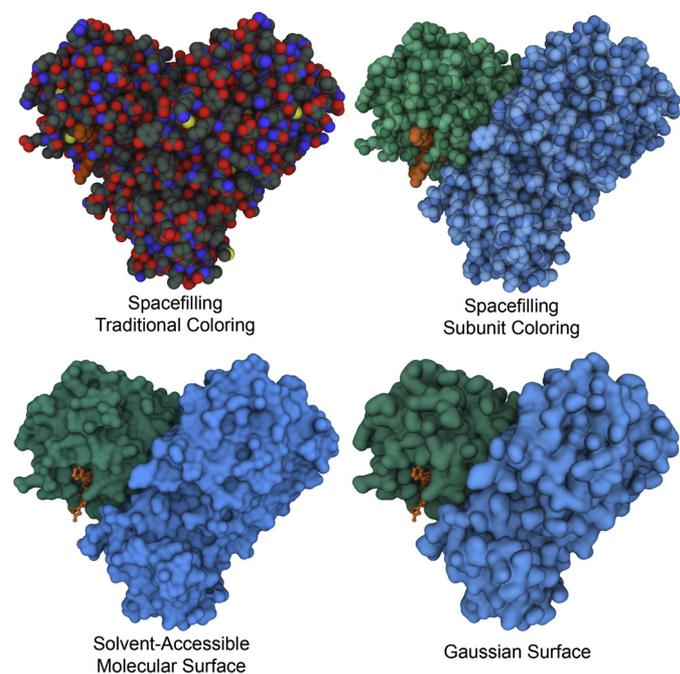


Figure 15. Several common surface representations used to display biological molecules. The coronavirus main protease is shown, with an inhibitor in orange blocking the active site (6Lu7; (66)). All images created with Mol* (67) at the RCSB PDB website ([rcsb.org](https://www.rcsb.org)).

highlights, depth cueing, and many other rendering tricks of the trade. Together, these generate images that highlight the three-dimensionality of the molecule. The images in Figures 15 and 16 are exemplary of the types of renderings that are provided by default in many contemporary molecular graphics packages using these techniques.

For many years, we and others have also explored non-photorealistic approaches to seeing molecules (50), most notably, as part of outreach work with the “Molecule of the Month” feature at PDB-101, the outreach portal of the RCSB Protein Data Bank (51). As perfectly demonstrated in Jane’s hand-drawn diagrams, cartoons are an effective way to simplify a scene, paring it down to the essential components. Our “Illustrate” software generates cartoony images using an imaging processing technique to generate outlines and flat, unshaded atoms to reduce visual clutter (52). The goal of these cartoon illustrations is to give an impression of the subunits in an assembly, rather than focusing on the details of each atom. As with the Gaussian surfaces in Figure 16, they scale nicely in larger scenes using highly simplified cartoons of the molecular shape. For example, I routinely combine digital images of individual molecular structures with paintings of cellular environments that use simpler cartoony shapes for each molecule (Fig. 17). Illustrative images also lend themselves easily to other graphical tricks for telling molecular stories, such as including schematic elements (as in the unobserved membrane-spanning portions of the spike protein) and overlaps and transparency to show processes occurring inside molecular assemblies (as in the three ribosome structures shown here).

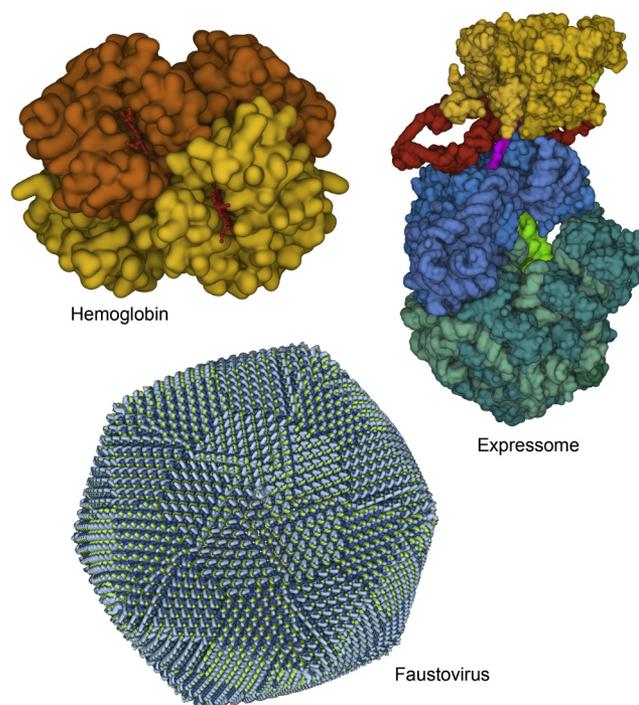


Figure 16. Gaussian surfaces provide a way to tune a representation based on the complexity of the image. For hemoglobin (2hhb; (68)), a detailed surface closely follows the contours of atoms. For the expressome (6x9q; (69)), a smoother surface is used to help focus attention on the different functional units in the assembly (RNA polymerase in yellow, ribosome in blue and cyan, tRNA in green, transcription factors in red, and a nascent mRNA in magenta). For faustovirus (5j7v; (70)), one of the largest complexes in the PDB archive, a highly smoothed Gaussian surface is provided by default in Mol* to highlight subunits while allowing interactive manipulation. All images created with Mol* interactively at the RCSB PDB website.

Seeing cellular environments

As the PDB archive continues to grow, we are rapidly approaching the time when we will have structural snapshots of every molecule in a cell. Combined with the ever-finer resolution of cryoelectron tomography, it is becoming feasible to create integrative models of entire cells at molecular detail. This goal poses several grand challenges, including the gathering and curation of data from multiple sources, the complexity of the higher-order assemblies and ultrastructure of cells, and the limited (but constantly improving) performance of available hardware, all of which are being addressed by multiple laboratories around the world (53). We are taking two approaches, designed for two different but complementary goals, and leveraging the powerful advances in graphics software from the gaming community (Fig. 18). The first is CellPACK, a suite of programs designed to create models of cellular environments based on a recipe of all the component molecules and where they reside in the cell (54). The second is CellPAINT, an interactive program much like a digital painting program that allows nonexpert users to build illustrations of cellular environments (55). This type of guided image creation is designed to give users some of the freedom that is possible in hand-drawn images, providing tools that enforce known relative sizes and behaviors of molecules while allowing

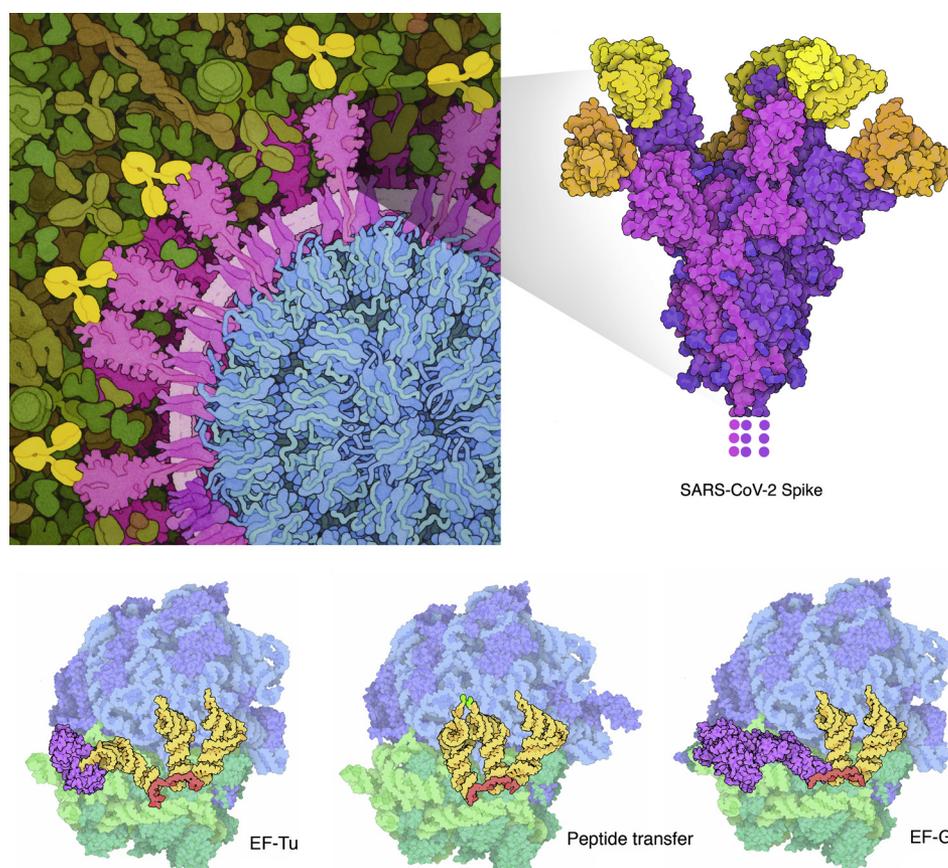


Figure 17. Nonphotorealistic illustration. *Top*, computer graphics illustration of coronavirus spike protein with bound antibody Fab domains (*right*, 7cwu; (71)) and a hand-drawn artistic conception of this spike-antibody interaction in the context of the whole virus in blood plasma (*left*). *Bottom*, computer graphics illustrations of ribosomes showing three key stages in elongation of protein synthesis (4v5g, 4v5d, 4v5f; (72–74)). Illustrations are part of outreach materials at PDB-101 (pdb101.rcsb.org).

easy, interactive construction of diverse subcellular scenes comprising these molecules.

New ways of seeing molecules for the challenges of today

The molecular graphics techniques in widespread use today—bond diagrams, ribbon diagrams, dot surfaces, and space-filling representations—are highly effective for research, dissemination, and education/outreach (43, 56–59). However, the field of structural biology continues to expand, and the classic graphical metaphors are providing the seeds for further exploration and development. To me, the most exciting aspect of molecular graphics is that there is always a new challenge. Here are a few selected examples of avenues that are calling for innovation in the near future and a few of the ways they are being addressed by the creative molecular graphics community.

The dynamic nature of biomolecules continues to be a central challenge. This is particularly problematic, and interesting, in the growing field of intrinsically disordered proteins, which undergo large structural transitions in their functional roles, or perhaps never adopt a defined structure at all. Dynamics is also an intrinsic part of nearly every biomolecular process, from the allosteric motions of hemoglobin to the many complex steps of ribosomal initiation, elongation, and

termination. Our visual tools for exploring and disseminating this dynamic information are woefully limited. A snapshot metaphor is often used for defined dynamic processes such as allosteric transitions and pathways of enzyme catalysis. In publication, these snapshots are often captured in a series of small multiple images. Kinemages (21), JSmols (60), and the like add personal agency to these snapshots, allowing users to flip between conformations and interactively explore various aspects of the structures. Animation also remains a primary tool for representing dynamics, borrowing metaphors from cinema to create molecular stories that capture salient dynamic events (61). These are necessarily scripted, since biomolecular processes typically occur over large numbers of time steps and key transitions are often lost in a sea of random motion. This requires much effort and artistry to focus attention and employs many approximations to create a comprehensible presentation of digestible length.

For decades, molecular graphics researchers have explored the newest technology to improve the user experience when interacting with molecules. As described above, molecular graphics was one of the driving applications for the first interactive computer graphics hardware and provided an ongoing test-bed for development of virtual reality applications. New metaphors were developed to streamline the user experience in these

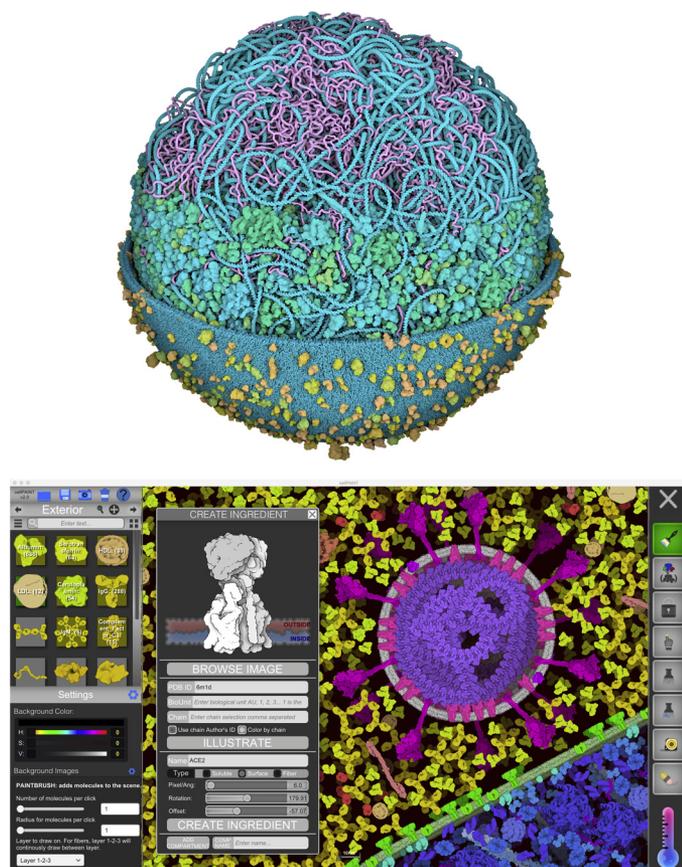


Figure 18. Modeling and visualization of molecules in cells. *Top*, model of a mycoplasma bacterium based on data from WholeCell models (75). In this computer graphics visualization, the whole model is shown at the base, then in progressive layers, the membrane is removed, and soluble molecules are removed to show only the DNA and RNA. Modeling and visualization by Ludovic Autin and Martina Maritan. *Bottom*, creating an illustration of SARS-CoV-2 using CellPAINT. Molecules are chosen from the palette on the left and interactively painted into the scene. In the pop-up “Create Ingredient” window, the user is reading coordinates for the cellular receptor ACE2 to add to the palette.

applications, as they require both comprehensible representations of the molecules (that are compact enough for interactive performance) and intuitive methods for navigating the space around them. All manner of grabbing, turning, walking, and teleporting metaphors have been tried, within spaces that represent a “Fantastic Voyage,” molecular museum, virtual laboratory, and many other imaginative scenarios (see (62) and the many projects cited there). Given the growing widespread availability of immersive headsets, and also interactive touch-sensitive tablets, we expect that there is a rich environment for continued experimentation. Currently, we also have the great advantage of being able to leverage the massive development effort that is being applied to the gaming industry, providing turnkey development environments such as Unity for creating our molecule-based applications.

We are also seeing molecular graphics researchers going back to their roots and turning once again to physical models as a way to provide tangible understanding of biomolecular structure and function. The growing availability of 3D printing technology has opened the door to creating tangible models based on atomic coordinate files (63, 64). This has posed new challenges for

molecular metaphors, since an orthogonal set of design constraints are required on top of the typical visual design constraints. For printed models, the representation must be strong enough to withstand printing and handling, often requiring the addition of support struts and other elements related to the physical structure. These elements must be carefully designed so as to not distract from understanding of the underlying biology.

Perhaps the most exciting short-term challenge facing the molecular graphics community is the current revolution in cryoelectron microscopy (34). Cryo-EM is revealing the structures of biomolecular assemblies of unprecedented size and complexity and is consequently breaking many of the familiar methods that we use for exploring molecular structure. New methods are needed to make these complex structures accessible, easily and intuitively (65). A multiresolution approach will undoubtedly be necessary, allowing users to flip between detailed exploration of protein folds and sidechain interactions and larger-scale understanding of the assembly of subunits and comparison with other assemblies. This will pose great challenges in the design of effective representations at all of the levels, methods for facile navigation between these levels, and implementation of these methods in infrastructure powerful enough to provide interactive performance. Solutions to these issues are already well under way in projects like ChimeraX (38) and Mol*, promising exciting new developments to come.

Acknowledgments



First, we want to express our enormous gratitude to the worldwide Protein Data Bank and all the contributing structural biologists, with a ribbon 50th-birthday bouquet created by Onisha Patel, WEHI, Melbourne. We thank all the talented people who have come through our laboratories and our many insightful collaborators.

Author contributions—All authors were involved in funding acquisition and execution of the work, and in conception, drafting, revision, and approval of this manuscript.

Funding and additional information—The Richardsons’ research and visualizations have been almost entirely supported by National Institutes of Health General Medical Sciences, from 30 years of R01 GM15000 (D. C. R.) to our current P01-GM063210 Project IV (J. S. R.) and R35-GM131883 (D. C. R.). Goodsell’s visualization work is supported by grant GM120604 from the National Institutes of Health and the RCSB Protein Data Bank (National Science Foundation DBI-1832184, National Institutes of Health GM133198, and US Department of Energy DE-SC0019749). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflict of interest—The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: 1sns, 1LpL, Four-character codes starting with a number are accession codes at the PDB archive; we use lower-case except for L, to avoid ambiguity in any font; 3D, three-dimensional; AED, Advanced Electronic Design; C α , Alpha carbon atoms in a protein chain; CaBLAM, C α -Based Low-resolution Annotation Method that uses peptide CO orientations to diagnose incorrect backbone conformations even if Ramachandran Φ, ψ values are restrained; Frodo, widely used, early model-to-map graphics on laboratory-accessible hardware; GRIP-75, the first model-to-map molecular graphics system, at UNC Chapel Hill; H, hydrogen atom (as in H-bond); Kinemage, a file using Dave Richardson's format for interactive molecular graphics; KiNG, Kinemage Next Generation, in Java, by Ian Davis and Vincent Chen; Mage, Dave's original program, in C, to display kinemage graphics; NOE, Nuclear Overhauser effect measurement of atom–atom distance, by NMR; ORTEP, Oak Ridge Thermal Ellipsoid Plot, a small-molecule line graphics system for drawing the ellipsoids of anisotropic temperature factors at each atom; PDB, Protein Data Bank, for experimental structures of macromolecules; PS300, or MPS, Evans & Sutherland calligraphic (vector-drawn) display workstation; RCSB, Research Collaboratory for Structural Biology; US branch of the wwPDB; RDC, Residual dipolar coupling measurement of atom–atom orientation, by NMR; SOD, Cu,Zn superoxide dismutase; SS bond, disulfide bond; Suite, the sugar-to-sugar, rather than nucleotide, parsing of RNA backbone (the Richardsons' best published pun); TIM, triose phosphate isomerase; vdW, van der Waals; wwPDB, worldwide PDB, including RCSB, PDBe, PDBj, BMRB, and EMDB.

References

- Perutz, M. F., Rossmann, M. G., Cullis, A. F., Muirhead, H., Will, G., and North, A. C. (1960) Structure of haemoglobin: A three-dimensional Fourier synthesis at 5.5-Å resolution, obtained by X-ray analysis. *Nature* **185**, 416–422
- Kendrew, J. C., Dickerson, R. E., Strandberg, B. E., Hart, R. G., Davies, D. R., Phillips, D. C., and Shore, V. C. (1960) Structure of myoglobin: A three-dimensional Fourier synthesis at 2 Å resolution. *Nature* **185**, 422–427
- Kendrew, J. C. (1961) The three-dimensional structure of a protein molecule. *Sci. Am.* **205**, 96–110
- Dickerson, R. E. (1969) *Structure and Action of Proteins*, Addison-Wesley, Boston, MA
- Richardson, J. S., and Richardson, D. C. (2013) Studying and polishing the PDB's macromolecules. *Biopolymers* **99**, 170–182
- Richardson, J. S., and Richardson, D. C. (2013) Doing molecular biophysics: Finding, naming, and picturing signal within complexity. *Annu. Rev. Biophys.* **42**, 1–28
- Richardson, J. S. (1981) The anatomy and taxonomy of protein structure. *Adv. Protein Chem.* **34**, 167–339
- Lillie, B. (2014) What Wikipedia taught me about my grandfather. *Atlantic Monthly*
- Arnone, A., Bier, C. J., Cotton, F. A., Day, V. W., Hazen, E. E., Richardson, D. C., Yonath, A., and Richardson, J. S. (1971) A high resolution structure of an inhibitor complex of the extracellular nuclease of *Staphylococcus aureus*. I. Experimental procedures and chain tracing. *J. Biol. Chem.* **246**, 2302–2316
- Richardson, J. S. (2018) Art and analogy help scientists “see” big biological molecules in 3D. *SciArt Mag.*
- Feldman, R. J. (1976) *AMSOM: Atlas of Macromolecular Structure on Microfiche*, Tracor Jitco, Inc, Rockville, MD
- Richardson, J. S., Thomas, K. A., and Richardson, D. C. (1975) Alpha-carbon coordinates for bovine Cu,Zn superoxide dismutase. *Biochem. Biophys. Res. Commun.* **63**, 986–992
- Chuse, A. R. (1935) *Costume Design*, Bridgman Publishing, Pelham, NY
- Banner, D. W., Bloomer, A. C., Petsko, G. A., Phillips, D. C., and Wilson, I. A. (1976) Atomic coordinates for triose phosphate isomerase from chicken muscle. *Biochem. Biophys. Res. Commun.* **72**, 146–155
- Richardson, J. S., Richardson, D. C., Tweedy, N. B., Gernert, K. M., Quinn, T. P., Hecht, M. H., Erickson, B. W., Yan, Y., McClain, R. D., and Donlan, M. E. (1992) Looking at proteins: Representations, folding, packing, and design. Biophysical Society National Lecture, 1992. *Biophys. J.* **63**, 1185–1209
- Britton, E. G., Lipscomb, J. S., and Pique, M. E. (1978) Making nested rotations convenient for the user. *SIGGRAPH Comput. Graph.* **12**, 222–227
- Richardson, D. C. (1977) The three-dimensional structure of Cu,Zn superoxide dismutase. In: Michelson, McCord, Fridovich, eds. *Superoxide and Superoxide Dismutase*, Academic Press, Cambridge, MA: 217–223
- Tainer, J. A., Getzoff, E. D., Beem, K. M., Richardson, J. S., and Richardson, D. C. (1982) Determination and analysis of the 2 Å structure of copper, zinc superoxide dismutase. *J. Mol. Biol.* **160**, 181–217
- Jones, T. A. (1978) A graphics model building and refinement system for macromolecules. *J. Appl. Crystallogr.* **11**, 268–272
- Carson, M., and Bugg, C. E. (1986) Algorithm for ribbon models of proteins. *J. Mol. Graph.* **4**, 121–122
- Richardson, D. C., and Richardson, J. S. (1992) The kinemage: A tool for scientific communication. *Protein Sci.* **1**, 3–9
- Sayle, R. A., and Milner-White, E. J. (1995) RASMOL: Biomolecular graphics for all. *Trends Biochem. Sci.* **20**, 374
- Davis, I. W., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2004) MOLPROBITY: Structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res.* **32**, W615–W619
- Chen, V. B., Davis, I. W., and Richardson, D. C. (2009) KiNG (Kinemage, Next Generation): A versatile interactive molecular and scientific visualization program. *Protein Sci.* **18**, 2403–2409
- Connolly, M. (1983) Solvent-accessible surfaces of proteins and nucleic acids. *Science* **221**, 709–713
- Word, J. M., Lovell, S. C., LaBean, T. H., Taylor, H. C., Zalis, M. E., Presley, B. K., Richardson, J. S., and Richardson, D. C. (1999) Visualizing and quantifying molecular goodness-of-fit: Small-probe contact dots with explicit hydrogen atoms. *J. Mol. Biol.* **285**, 1711–1733
- Thayer, M. M., Flaherty, K. M., and McKay, D. B. (1991) Three-dimensional structure of the elastase of *Pseudomonas aeruginosa* at 1.5-Å resolution. *J. Biol. Chem.* **266**, 2864–2871
- Ladner, J. E., Wladkowski, B. D., Svensson, L. A., Sjölin, L., and Gilliland, G. L. (1997) X-ray structure of a ribonuclease A-uridine vanadate complex at 1.3 Å resolution. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 290–301
- Stoddard, C. D., Widmann, J., Trausch, J. J., Marcano-Velázquez, J. G., Knight, R., and Batey, R. T. (2013) Nucleotides adjacent to the ligand-binding pocket are linked to activity tuning in the purine riboswitch. *J. Mol. Biol.* **425**, 1596–1611
- Richardson, J. S., Schneider, B., Murray, L. W., Kapral, G. J., Immormino, R. M., Headd, J. J., Richardson, D. C., Ham, D., Hershkovits, E., Williams, L. D., Keating, K. S., Pyle, A. M., Micallef, D., Westbrook, J., Berman, H. M., et al. (2008) RNA backbone: Consensus all-angle conformers and modular string nomenclature (an RNA Ontology Consortium contribution). *RNA* **14**, 465–481
- Block, J. N., Zielinski, D. J., Chen, V. B., Davis, I. W., Vinson, E. C., Brady, R., Richardson, J. S., and Richardson, D. C. (2009) KinImmerse: Macromolecular VR for NMR ensembles. *Source Code Biol. Med.* **4**, 3
- Li, S., Finley, J., Liu, Z.-J., Qiu, S.-H., Chen, H., Luan, C.-H., Carson, M., Tsao, J., Johnson, D., Lin, G., Zhao, J., Thomas, W., Nagy, L. A., Sha, B., DeLucas, L. J., et al. (2002) Crystal structure of the cytoskeleton-associated protein glycine-rich (CAP-Gly) domain. *J. Biol. Chem.* **277**, 48596–48601
- Arendall, W. B., Tempel, W., Richardson, J. S., Zhou, W., Wang, S., Davis, I. W., Liu, Z.-J., Rose, J. P., Carson, W. M., Luo, M., Richardson, D. C., and Wang, B.-C. (2005) A test of enhancing model accuracy in high-throughput crystallography. *J. Struct. Funct. Genomics* **6**, 1–11
- Chiu, W., and Downing, K. H. (2017) Editorial overview: Cryo electron microscopy: Exciting advances in CryoEM Herald a new era in structural biology. *Curr. Opin. Struct. Biol.* **46**, iv–viii
- Williams, C. J., Headd, J. J., Moriarty, N. W., Prisant, M. G., Videau, L. L., Deis, L. N., Verma, V., Keedy, D. A., Hintze, B. J., Chen, V. B., Jain, S.,

- Lewis, S. M., Arendall, W. B., Snoeyink, J., Adams, P. D., *et al.* (2018) MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315
36. Croll, T. I., Williams, C. J., Chen, V. B., Richardson, D. C., and Richardson, J. S. (2021) Improving SARS-CoV-2 structures: Peer review by early coordinate release. *Biophys. J.* **120**, 1085–1096
37. Yin, W., Mao, C., Luan, X., Shen, D.-D., Shen, Q., Su, H., Wang, X., Zhou, F., Zhao, W., Gao, M., Chang, S., Xie, Y.-C., Tian, G., Jiang, H.-W., Tao, S.-C., *et al.* (2020) Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. *Science* **368**, 1499–1504
38. Goddard, T. D., Huang, C. C., Meng, E. C., Pettersen, E. F., Couch, G. S., Morris, J. H., and Ferrin, T. E. (2018) UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25
39. Nakane, T., Kotecha, A., Sente, A., McMullan, G., Masiulis, S., Brown, P. M. G. E., Grigoras, I. T., Malinauskaitė, L., Malinauskas, T., Miehling, J., Uchański, T., Yu, L., Karia, D., Pechnikova, E. V., de Jong, E., *et al.* (2020) Single-particle cryo-EM at atomic resolution. *Nature* **587**, 152–156
40. Croll, T. I. (2018) Isolde: A physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr. D Struct. Biol.* **74**, 519–530
41. Surlis, M. C., Richardson, J. S., Richardson, D. C., and Brooks, F. P. (1994) Sculpting proteins interactively: Continual energy minimization embedded in a graphical modeling system. *Protein Sci.* **3**, 198–210
42. Kappel, K., Zhang, K., Su, Z., Watkins, A. M., Kladwang, W., Li, S., Pintilie, G., Topkar, V. V., Rangan, R., Zheludev, I. N., Yesselman, J. D., Chiu, W., and Das, R. (2020) Accelerated cryo-EM-guided determination of three-dimensional RNA-only structures. *Nat. Methods* **17**, 699–707
43. Olson, A. J. (2018) Perspectives on structural molecular biology visualization: From past to present. *J. Mol. Biol.* **430**, 3997–4012
44. Goodsell, D. S. (1988) RMS: Programs for generating raster molecular surfaces. *J. Mol. Graph.* **6**, 41–44
45. Goodsell, D. S., and Olson, A. J. (1992) Molecular illustration in black and white. *J. Mol. Graph.* **10**, 235–240
46. Corey, R. B., and Pauling, L. (1953) Molecular models of amino acids, peptides, and proteins. *Rev. Sci. Instrum.* **24**, 621–627
47. Lee, B., and Richards, F. M. (1971) The interpretation of protein structures: Estimation of static accessibility. *J. Mol. Biol.* **55**, 379–400
48. Lewis, M., and Rees, D. (1985) Fractal surfaces of proteins. *Science* **230**, 1163–1165
49. Phong, B. T. (1975) Illumination for computer generated pictures. *Commun. ACM* **18**, 311–317
50. Goodsell, D. S., and Jenkinson, J. (2018) Molecular illustration in research and education: Past, present, and future. *J. Mol. Biol.* **430**, 3969–3981
51. Goodsell, D. S., Dutta, S., Zardecki, C., Voigt, M., Berman, H. M., and Burley, S. K. (2015) The RCSB PDB “molecule of the month”: Inspiring a molecular view of biology. *PLoS Biol.* **13**, e1002140
52. Goodsell, D. S., Autin, L., and Olson, A. J. (2019) Illustrate: Software for biomolecular illustration. *Structure* **27**, 1716–1720.e1
53. Goodsell, D. S., Olson, A. J., and Forli, S. (2020) Art and science of the cellular mesoscale. *Trends Biochem. Sci.* **45**, 472–483
54. Johnson, G. T., Autin, L., Al-Alusi, M., Goodsell, D. S., Sanner, M. F., and Olson, A. J. (2015) cellPACK: A virtual mesoscope to model and visualize structural systems biology. *Nat. Methods* **12**, 85–91
55. Gardner, A., Autin, L., Barbaro, B., Olson, A. J., and Goodsell, D. S. (2018) CellPAINT: Interactive illustration of dynamic mesoscale cellular environments. *IEEE Comput. Graph. Appl.* **38**, 51–66
56. Johnson, G. T., and Hertig, S. (2014) A guide to the visual analysis and communication of biomolecular structural data. *Nat. Rev. Mol. Cell Biol.* **15**, 690–698
57. Kozlíková, B., Krone, M., Falk, M., Lindow, N., Baaden, M., Baum, D., Viola, I., Parulek, J., and Hege, H.-C. (2017) Visualization of biomolecular structures: State of the art revisited. *Comput. Graph. Forum* **36**, 178–204
58. Martínez, X., Chavent, M., and Baaden, M. (2020) Visualizing protein structures — tools and trends. *Biochem. Soc. Trans.* **48**, 499–506
59. O’Donoghue, S. I., Goodsell, D. S., Frangakis, A. S., Jossinet, F., Laskowski, R. A., Nilges, M., Saibil, H. R., Schafferhans, A., Wade, R. C., Westhof, E., and Olson, A. J. (2010) Visualization of macromolecular structures. *Nat. Methods* **7**, S42–S55
60. Hanson, R. M. (2010) Jmol — a paradigm shift in crystallographic visualization. *J. Appl. Crystallogr.* **43**, 1250–1260
61. McGill, G. (2008) Molecular movies...Coming to a lecture near you. *Cell* **133**, 1127–1132
62. Goddard, T. D., Brilliant, A. A., Skillman, T. L., Vergenz, S., Tyrwhitt-Drake, J., Meng, E. C., and Ferrin, T. E. (2018) Molecular visualization on the holodeck. *J. Mol. Biol.* **430**, 3982–3996
63. Coakley, M. F., Hurt, D. E., Weber, N., Mtingwa, M., Fincher, E. C., Alekseyev, V., Chen, D. T., Yun, A., Gizaw, M., Swan, J., Yoo, T. S., and Huyen, Y. (2014) The NIH 3D Print Exchange: A public resource for bioscientific and biomedical 3D prints. *3D Print. Addit. Manuf.* **1**, 137–140
64. Gardner, A., and Olson, A. (2016) 3D printing of molecular models. *J. Biocommun.* <https://doi.org/10.5210/jbc.v40i1.6626>
65. Goddard, T. D., and Ferrin, T. E. (2007) Visualization software for molecular assemblies. *Curr. Opin. Struct. Biol.* **17**, 587–595
66. Jin, Z., Du, X., Xu, Y., Deng, Y., Liu, M., Zhao, Y., Zhang, B., Li, X., Zhang, L., Peng, C., Duan, Y., Yu, J., Wang, L., Yang, K., Liu, F., *et al.* (2020) Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. *Nature* **582**, 289–293
67. Sehnal, D., Rose, A., Koca, J., Burley, S., and Velankar, S. (2018) Mol*: Towards a common library and tools for web molecular graphics. In: Byska, J., Krone, M., Sommer, B., eds. *Workshop on Molecular Graphics and Visual Analysis of Molecular Data*, The Eurographics Association, Aire-la-Ville, Switzerland: 29–33
68. Fermi, G., Perutz, M. F., Shaanan, B., and Fourme, R. (1984) The crystal structure of human deoxyhaemoglobin at 1.74 Å resolution. *J. Mol. Biol.* **175**, 159–174
69. Wang, C., Molodtsov, V., Firlar, E., Kaelber, J. T., Blaha, G., Su, M., and Ebright, R. H. (2020) Structural basis of transcription-translation coupling. *Science* **369**, 1359–1365
70. Klose, T., Reteno, D. G., Benamar, S., Hollerbach, A., Colson, P., La Scola, B., and Rossmann, M. G. (2016) Structure of faustovirus, a large dsDNA virus. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 6206–6211
71. Wang, N., Sun, Y., Feng, R., Wang, Y., Guo, Y., Zhang, L., Deng, Y.-Q., Wang, L., Cui, Z., Cao, L., Zhang, Y.-J., Li, W., Zhu, F.-C., Qin, C.-F., and Wang, X. (2021) Structure-based development of human antibody cocktails against SARS-CoV-2. *Cell Res.* **31**, 101–103
72. Schmeing, T. M., Voorhees, R. M., Kelley, A. C., Gao, Y.-G., Murphy, F. V., Weir, J. R., and Ramakrishnan, V. (2009) The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA. *Science* **326**, 688–694
73. Voorhees, R. M., Weixlbaumer, A., Loakes, D., Kelley, A. C., and Ramakrishnan, V. (2009) Insights into substrate stabilization from snapshots of the peptidyl transferase center of the intact 70S ribosome. *Nat. Struct. Mol. Biol.* **16**, 528–533
74. Gao, Y.-G., Selmer, M., Dunham, C. M., Weixlbaumer, A., Kelley, A. C., and Ramakrishnan, V. (2009) The structure of the ribosome with elongation factor G trapped in the posttranslocational state. *Science* **326**, 694–699
75. Karr, J. R., Sanghvi, J. C., Macklin, D. N., Gutschow, M. V., Jacobs, J. M., Bolival, B., Assad-Garcia, N., Glass, J. I., and Covert, M. W. (2012) A whole-cell computational model predicts phenotype from genotype. *Cell* **150**, 389–401