Aesthetics and Understanding in Molecular Motion

Shareef M. Dabdoub dabdoub.2@osu.edu The Ohio State University Biophysics Program
Sheryl S. Justice, Ph.D. sheryl.justice@nationwidechildrens.org The Research Institute at NCH
William C. Ray, Ph.D. ray.29@osu.edu The Ohio State University Biophysics Program

Introduction

The key to understanding molecular function is often found through investigation of structure and changes in structure through motion. In the case of Parkinson s disease, the protein alpha–synuclein, normally present in an unfolded state, misfolds and accumulates as a filamentous mass in cells within the brain–stem neurons, causing premature cell death [14]. Furthermore, the dynamics of motion itself is often necessary for proper function and intramolecular signal transduction. Imaging static molecular structure is extremely important in such studies, but it is only through gathering and representation of the structural dynamics of motion that deeper insight will be gained [15].

Such data can be obtained through Nuclear Magnetic Resonance imaging or molecular dynamics (MD) simulations, both of which yield snapshots of molecular conformations ("poses") that represent changes in the structure of the molecule over time. Traditionally, visual representations of motion data display these poses either as a static series (or overlay) of conformations, much like a stroboscopic multiple–exposure picture of the motion, or as a computer animation [8]. Unfortunately, while each of these methods can provide useful insights, they both suffer from problems caused by an inherent mismatch with how viewers understand and percieve motion.

The rendering of motion data as a set of connected points sharing the same time–point is known as a timeline representation [8]. Each timeline represents the physical state of a molecule at a discrete moment along the axis of time; the advantage being sheer simplicity and portability. MD simulations produce data in this form as a matter of course, and a timeline rendering can be simply produced using any of the simulated time steps. The resulting images are easily publishable in digital or print format, and present the desired information: molecular structure as it changes over time. Additionally, structure at different time–points can be compared across multiple simulations or homologous molecules. Despite the seeming appropriateness, visualizations of this nature are unintuitive and clash with how we understand motion flow over time. Specifically, we comprehend motion as just that: continuous flow, not a discrete set of events. This disconnect becomes quickly apparent for anything other than simple motion with limited range and complexity (see Figure 1).

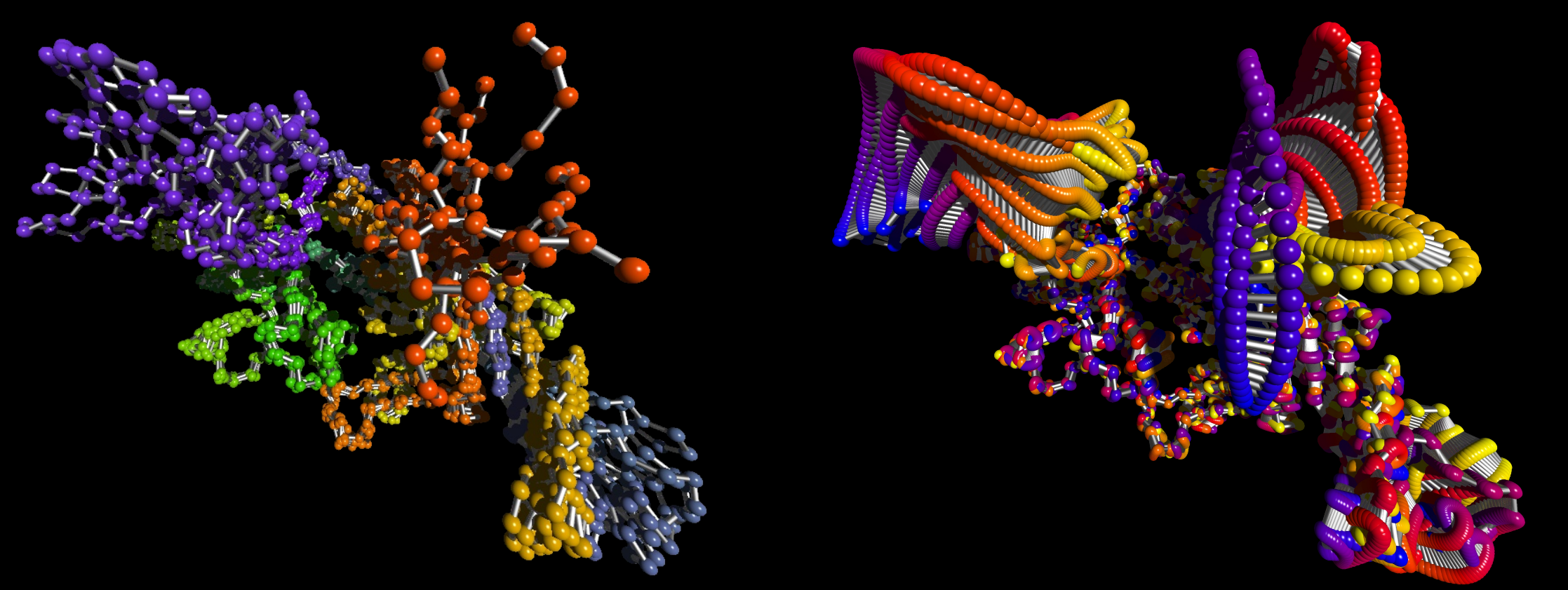

Figure 1: Illustrating the difference between timeline and pathline representations of molecular motion. (a) Timeline rendering of six alpha carbons from the capsid protein monomer of the Adeno Associated Virus, Type 2. Here the main impression left by this image is a jumble of atoms (b) The visually orthogonal pathline rendering of the same monomer. This method presents the motion data in a manner far more conducive to intuitive understanding.

In order to overcome the difficulties in understanding spatiotemporal motion from static snapshots, people turn to computer animations of the motion. This is, however, only a partial step in the right direction. An animation conveys both structure and actual motion, but simply amalgamating static timeline representations into an animation does not address the problem with perception. Specifically, the motion visualization presented by animations gives no indication as to where structures of interest have been or where they are going at each instant in the video. Such animations ultimately rest upon timelines as opposed to pathlines, and viewers have difficulty integrating temporally–separated information from beginning to end. Moreover, animations require active display and cannot be printed and disseminated with the ease of the previous technique. Thus there is a clear need for a technique by which molecular motion can be represented as a static image, while at the same time conveying the dynamic behavior of a molecule in a manner which humans intuitively perceive and comprehend.

Methods

Comparing the timeline representation (Figure 1a) of a system of motion to the flow–oriented visualizations (Figure 1b), it was clear that the pathlines of atom motion would provide the most useful mode of visualization. Thus, the heart of our Molecular Flow visualizations are formed by the pathlines of atoms salient to motion of a system. In the case of proteins, we have generally chosen to extract and make use of the alpha carbons for each amino acid. This decision ensures capture of significant motion, while effectively acting as a low–pass filter for the high–frequency motions of amino acid side–chains.

Given the starting point of discrete spatial coordinates provided by MD simulations we decided to represent pathlines using the piecewise–polynomial functions. Such functions, commonly known as splines, provide an excellent means for mapping from discrete to smooth continuous functions, given that they can approximate such functions on any interval arbitrarily well [13]. To produce visualizations based on these abstractions, we have made use of the open source ray–tracing engine, POV–Ray [1]. Specifically, we create cubic natural splines for each atom of interest, with each time–point representing a control point for the spline. The splines themselves, being only mathematical functions, are visualized using sphere–capped cylinders tracking the spline interpolation. The pathlines in Figure 2 correspond to the atoms from an MD simulation of the AAV–2 [2, 3, 6] capsid monomer protein. Each of the splines is colored on some gradient (in this example the gradient runs from blue to red) as an indication of increasing time along the simulation data. The next piece of the visualization consists of spheres representing atoms near the distal end of the timeline. As mentioned previously, flow–oriented lines have no inherent means for indicating time ordering. The addition of these spheres indicates an end, thus adding easily understandable orientation to the motion with respect to time.

The final useful visual cue we have added to the rendering is a set of semi–transparent ribbons that connect alternating adjacent atom pathlines with sufficient motion. These ribbons begin at preset (but customizable) fraction along the timeline of the motion flow, and provide important spatial context to the 3–D orientation of the motion. The pathlines alone in space may not provide easily understandable intuition as to the concerted motion of multiple atoms. With the addition of the ribbons, a solid surface is created between adjacent pathlines, making the 3–D orientation instantly comprehendible along the path of the atoms. Figure 2 provides a rendering of the full AAV–2 capsid protein monomer using the visualization system we have called MoFlow (Molecular Flow).

Examining the motion of the entire protein in Figure 2, the eye is immediately drawn to the three areas of significant motion. The red to blue gradient along the motion splines indicates the flow of time along the atom paths, and with the aid of the atom spheres, users are able to easily understand the flow of conformational change over time. Moreover, even though the entire protein is visualized, the motion paths are comprehendible with minimal effort.

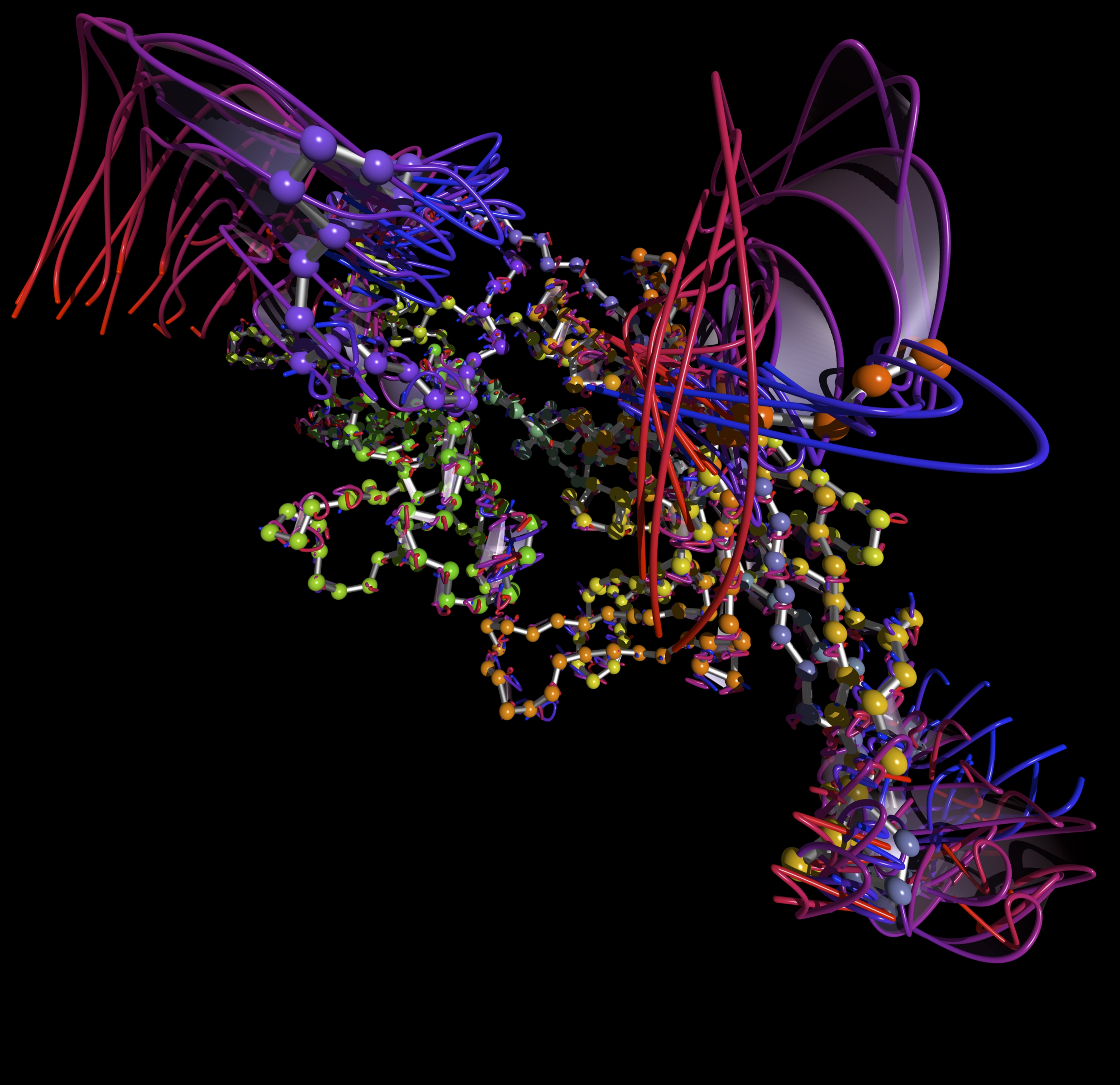

Figure 2: Backbone motion pathline-rendering of the AAV-2 capsid protein monomer transitioning from neutral to acidic pH. This rendering exhibits all of the elements of a MoFlow visualization: Atom pathlines represented by cylinders describing a spline-interpolated path punctuated by atom-locating spheres, colored on a gradient indicating time (here gradient from red to blue), and alternate pathlines connected by a semi-transparent ribbon.

Figure 3: The lid domain portion of a putative metal-dependent transcriptional regulator from Streptococcus mutans (a) Cartoon-rendered (PyMOL [5]) composite overlay view of each time-point. (b) MoFlow rendering of six alpha carbons from the middle of the protein structure. These visualizations illustrate the contrast between the seemingly simple swinging motion in the composite timeline image and the clearly more complex motion revealed by the pathline rendering.

Comparison

In order to further illustrate the utility of Moflow, we present the following three comparisons. In Figure 3 we have two renderings of the lid domain of a putative metal–dependent transcriptional regulator from Streptococcus mutans. The image in Fig. 3a is a composite cartoon rendering of each step in the hinge–like motion of the protein domain, from brick–red on the left moving clockwise to purple on the right. In this composite, the motion seems quite simple; the motile arm describes a smooth hemisphere from left to right. However, Fig. 3b clearly shows this not to be the case. The timeline representation provided by the composite obscures the more complex motion the pathline representation uncovers.

The motion in Figure 4 presents the opposite dilemma. This figure displays two alpha–helices undergoing rotational movement around a vertical axis. The composite image on the left (Figure 4a) presents a mass jumble of atoms with some possible rotational motion. In the image on the right (Figure 4b) we have extracted all the alpha carbons as well as the gamma carbons from the Lysine residues which stick out from the helices. Here the Moflow rendering of short curves and blue–to–red gradient coloring clearly informs the viewer of the simple rotational motion. In fact, this rendering includes amino acid side chains in the rendering, providing additional potentially useful information, whereas in the composite overlay, the side chains provide the majority of the obstructive clutter.

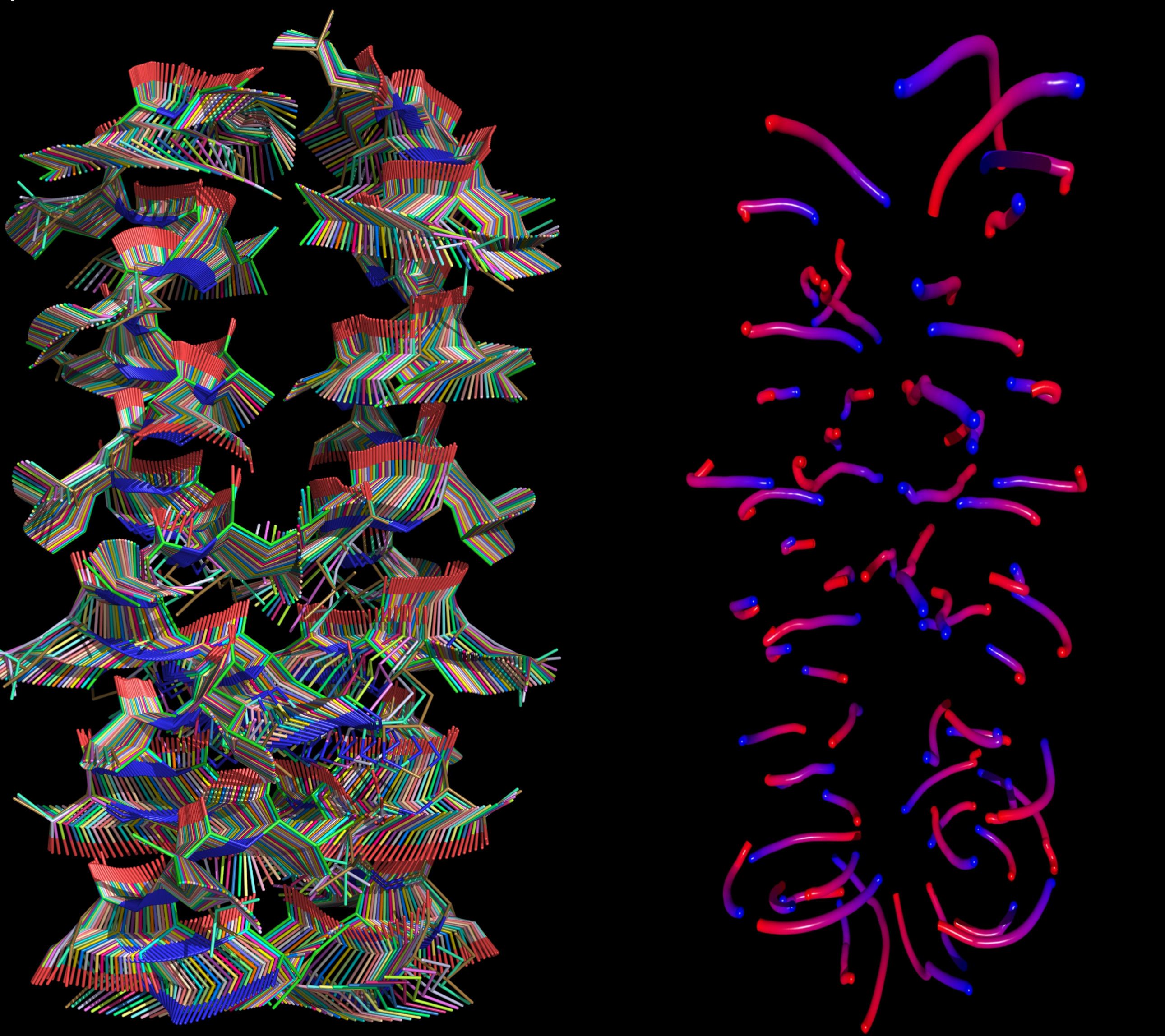

Figure 4: This motion example is of two non-connected alpha helices completing a half-turn rotation around their respective vertical axes. (a) The simplicity of the motion is hidden by the composite overlay of the structures at each time-point. (b) Rendering of backbone alpha carbons along the helices as well as the gamma carbons from each of the Lysine residues. The short curves and blue-to-red gradient clearly and efficiently describe the simple half-turn rotations. Data from Yale morph server [8]. MorphID: m155101-32532.

Conclusion

We have described here a new method for rendering and visual exploration of molecular motion that specifically takes into account the natural human intuition of movement over time. The core of this system is representing the motion of atoms as pathlines, as opposed to the orthogonal timeline representation of traditional molecular motion imagery and animation. The advantage is that the eye is naturally drawn along the pathlines, thereby providing immediate understanding of the overall motion. As we have demonstrated with Figures 2, 3, and 4, gaining an intuitive grasp of molecular motion using MoFlow rendering is simple at any scale. Indeed, while providing useful images of larger scale motion is certainly within the capabilities of MoFlow, the latter two images reflect the value of investigation at smaller scales as well.

Digitally, the MoFlow renderings are initially stored as a POV–Ray scene file. This allows users to make quick parameter changes such as the location of the motion–orientation spheres, the image size and quality, and even produce simple animations. This brings us to another advantage of MoFlow rendering, namely that (as discussed in the introduction) traditional motion animations suffer from the same timeline–related perception issues. With the pathline representations, we can, for example, drag the motion–orientation spheres and concomitant surface–delineating ribbons along the length of the pathlines. Such animations can be quite useful in investigations involving especially convoluted motions. Thus in both domains: static and animated, the MoFlow method of enhanced perceptual rendering provides quite clear advantages over traditional motion visualization methods.

Moving forward, we are working to expand the usefulness and application of the MoFlow rendering system. We are currently investigating the visualization of the inherent uncertainty and randomness in molecular systems. Brownian motion and other forces cause MD simulations as well as NMR investigation to produce varying motion profiles in every experiment. An example of such a rendering is given in Figure 5, which was generated from a smaller section of the AAV–2 capsid protein monomer. The same basic elements carry over from the previously seen renderings, however here the data from five separate MD simulations are displayed as thinner, semitransparent pathlines containing atom spheres colored to match the larger distal atom spheres to aid in visual identification. These pathlines convey the range of motion over multiple simulation runs, while a much larger opaque set of pathlines indicate the average path taken by each atom. We are also pursuing effective means for visualizing the pace–variant motion produced by difference in magnitude such as that seen in comparing side chain motion to that undergone by the peptide backbone.

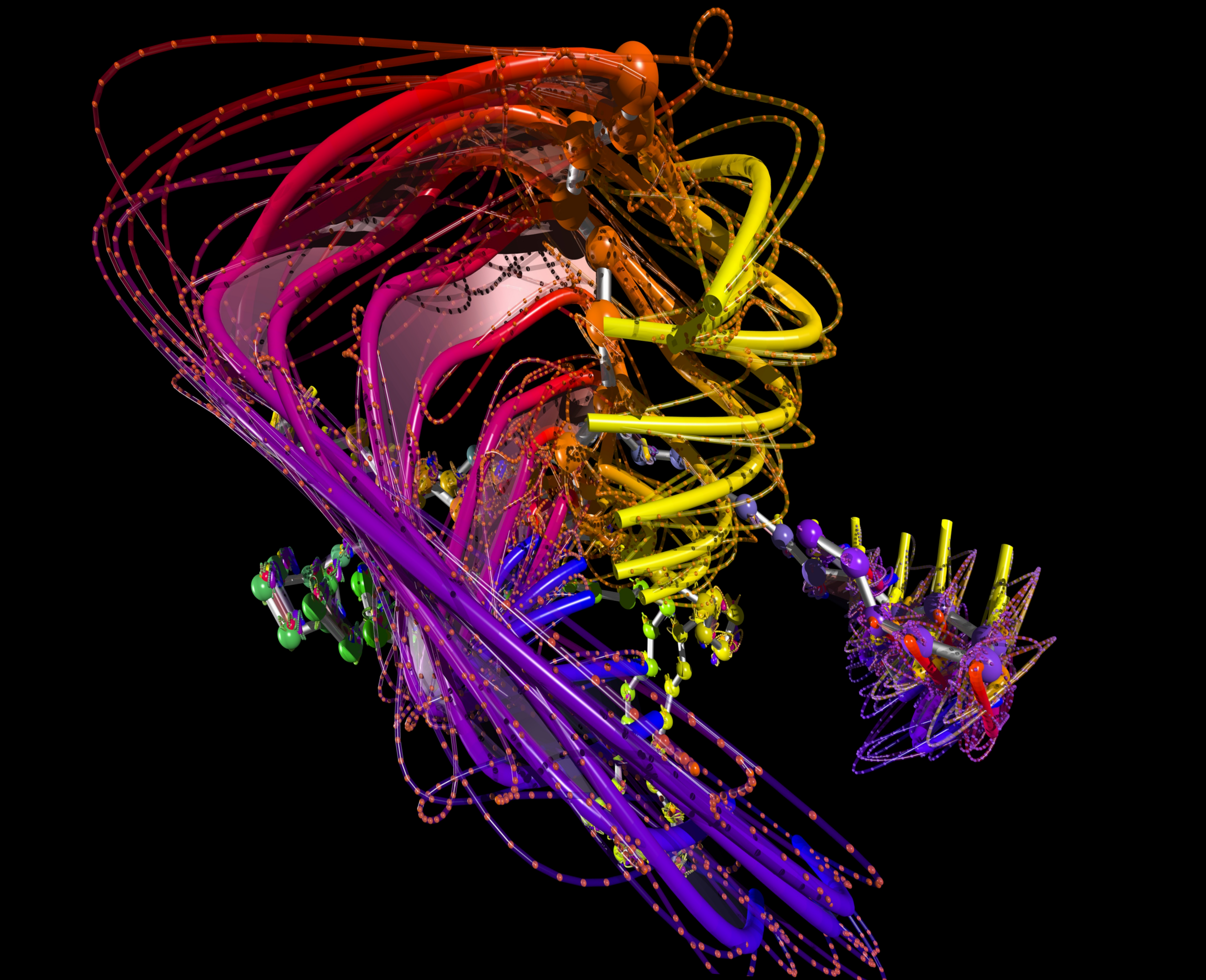

Figure 5: Visualizing the uncertainty inherent in molecular systems. Motion over pH change by a subsection of the AAV-2 capsid protein monomer was simulated five times. The thicker pathlines follow the familiar visualization paradigm, and represent the average of the five simulations. The thinner, transparent pathlines trace the individual simulations. The transparency allows the viewer to see the atom spheres within that are colored to match the distal atom spheres of the average pathlines, allowing the user to easily identify the individual-simulation paths with the atom path they describe.

References

[1] Persistence of vision raytracer (version 3.6). http://www.povray.org/, 2004.

[2] J. Bartlett, et.al. Infectious entry pathway of AAV and AAV vectors. J. Virology, 74(6), 2000.

[3] B. J. Carter. AAV and the development of AAV vectors: A historical perspective. Mol Ther, 10(6), Dec. 2004.

[4] D. A. Case, et.al. The amber biomolecular simulation programs. J.Comp.Chem, 26(16), 2005.

[5] W. DeLano. The PyMOL molecular graphics system. http://www.pymol.org, 2002.

[6] J. Grieger and R. Samulski. AAV as a gene therapy vector: Vector development, production and clinical applications. Gene Thpy & Gene Delivery Sys, 2005.

[7] V. Interrante and C. Grosch. Visualizing 3D flow. IEEE Computer Graphics and Applications, 18(4):49–53, 1998.

[8] W. G. Krebs and M. Gerstein. Survey and Summary: The Morph Server NAR, 28(8), 2000.

[9] D. A. Lane. Visualization of time–dependent flow fields. 4th Viz Conference, pages 32–38, 1993. IEEE Computer Society.

[10] R. S. Laramee, et.al. The state of the art in flow visualization: Dense and Texture–Based techniques. Comp. Graphics Forum, 23(2), 2004.

[11] D.Morikis, et.al. Solution structure of compstatin, a potent complement inhibitor. Protein Science, 7(3), 1998.

[12] F. H. Post, et.al. Feature extraction and visualisation of flow fields. SAR. Eurographics, 2002.

[13] L. Schumaker. Spline Functions: Basic Theory. Cambridge University Press, 3 edition, Sept. 2007.

[14] D. J. Selkoe. Folding proteins in fatal ways. Nature, 426(6968):900–904, 2003.

[15] R. G. Smock and L. M. Gierasch. Sending signals dynamically. Science, 324(5924):198–203, Apr. 2009.

Acknowledgements

The authors wish to thank Abhilash Mohan and Dr. Jeffery Bartlett for assistance with the experimental and biological data.