Biophys Rev (2010) 2:21–27
DOI 10.1007/s12551-009-0026-3

REVIEW

Using *Situs* for the integration of multi-resolution structures

Willy Wriggers

Received: 14 October 2009 / Accepted: 19 November 2009 / Published online: 8 January 2010

Abstract *Situs* is a modular and widely used software package for the integration of biophysical data across the spatial resolution scales. It has been developed over the last decade with a focus on bridging the resolution gap between atomic structures, coarse-grained models, and volumetric data from low-resolution biophysical origins, such as electron microscopy, tomography, or small-angle scattering. Structural models can be created and refined with various flexible and rigid body docking strategies. The software consists of multiple, stand-alone programs for the format conversion, analysis, visualization, manipulation, and assembly of 3D data sets. The programs have been ported to numerous platforms in both serial and shared memory parallel architectures and can be combined in various ways for specific modeling applications. The modular design facilitates the updating of individual programs and the development of novel application workflows. This review provides an overview of the *Situs* package as it exists today with an emphasis on functionality and workflows supported by version 2.5.

Keywords Structural models · 3D data sets · Multi-platform · Modeling

Electronic supplementary material The online version of this article (doi:10.1007/s12551-009-0026-3) contains supplementary material, which is available to authorized users.

W. Wriggers (✉)
Department of Physiology and Biophysics,
Weill Medical College of Cornell University,
1300 York Ave.,
New York, NY 10065, USA
e-mail: wriggers@biomachina.org

Present Address:
W. Wriggers
D.E. Shaw Research,
120 W. 45th St.,
New York, NY 10036, USA

Introduction

Scientific computing, including modeling and simulation, is crucial for solving biophysical research problems that are beyond the reach of traditional theoretical and experimental approaches (U.S. Department of Energy 2005). Originally confined to a supporting role with respect to experimental or theoretical approaches, modeling and simulation are increasingly seen as capable of creating new evidence in their own right (Lee et al. 2009). Computer-generated hypotheses can be confirmed or refuted, like their experimental or theoretical counterparts, even though the virtual (in silico) world is at best an imperfect mirror of the physical (in vivo) world.

In the late 1990s, funding agencies in the biological sciences took notice of this opportunity. In April 1998, a special Cell Biology and Biophysics Subcommittee of the U.S. National Advisory General Medical Sciences Council examined research trends in the areas of molecular cell biology, structural biology, and biophysics. Among the needs identified by the panel were better (computational) methods for structural analysis of large macromolecular assemblies and imaging macromolecules in cells. Based in part on these recommendations, the National Institutes of Health (NIH) issued a new program announcement that altered the more traditional biological hypothesis-driven review and award criteria in favor of method development (National Institutes of Health 2000). Instead of the traditional proposal style, biophysical scientists in the U.S. could for the first time submit applications based solely on the merit of computational techniques. This paradigm shift was important for the advancement of computational

biology, because opportunities for funding computational research had hitherto existed mainly in the physical sciences (U.S. Department of Energy 2005).

Against the backdrop of the emerging research opportunities in computational biophysics, the *Situs* package was created for the modeling and simulation of large biomolecular assemblies at variable resolution scales. *Situs* was initially conceived as a platform for the dissemination of structural coarse-graining algorithms to the biophysical community. Powerful experimental techniques such as cryo-electron microscopy (Baker and Johnson 1996), tomography (Medalia et al. 2002), and small-angle scattering (Niemann et al. 2008), which routinely produced 3D structures at a reduced spatial resolution, had emerged. These methods were capable of yielding low-resolution density maps under a wide range of biochemical conditions that allow atomic structures of components to be fitted and docked (Baker and Johnson 1996; Wriggers and Chacón 2001b), and they were in need of software to help integrate the structural data.

The goal of *Situs* is to characterize the structure and functionally relevant motions of biomolecular systems by integrating experimental data across the resolution scales, using advanced algorithms from neurocomputing, image processing, and visualization. A decade has passed since the publication of the original *Situs* paper (Wriggers et al. 1999). This review will assess the software as it is used by scientists today. Naturally, the workflow has changed in many ways over the years as compatible molecular graphics programs have evolved and *Situs* tools have been enhanced, updated, or replaced.

The following sections highlight the current *Situs* workflow using published usage examples kindly provided by other laboratories. In the first section, a typical correlation-based docking approach in electron microscopy (EM) is described, using the recent model of the influenza virus ribonucleoprotein complex (Coloma et al. 2009) as an example. Next, the integration of structural data with small-angle X-ray scattering (SAXS) data is shown on models of the extracellular region of an EGF receptor family member, s-dEGFR (Alvarado et al. 2009). Finally, a flexible fitting approach is shown using coarse-grained resolution models of myosin. Personal comments and annotations by the author are provided in the electronic supplementary material.

Correlation-based docking in electron microscopy

Chacón and Wriggers (2002) introduced *colores*, a widely used registration tool that takes advantage of Fourier correlation theory to rapidly scan the six translational and rotational degrees of freedom of a probe molecule relative to a (fixed) target density map. X-ray crystallographic fitting methods, based on volumetric cross-correlation or the free R-value, are limited to resolutions <10 Å where densities exhibit internal structure. The major advantage of *colores* is that it extends the viable resolution range to ~30 Å by means of a Laplacian operator that emphasizes contour (shape) information in addition to the traditional correlation. Over the years, we have optimized the efficiency and accuracy of *colores* and ported the tool to shared memory environments that take advantage of today's multi-core architectures. The series of steps and the programs that are required to use *colores* for the docking of a probe structure to a target EM map are shown schematically in Fig. 1.

Recently, *colores* was successfully used in the modeling of a biologically active influenza virus ribonucleoprotein (RNP) complex (Coloma et al. 2009). The RNP particles of influenza A viruses are formed by the association of single-stranded RNA to multiple monomers of nucleoprotein (NP) and a single copy of the polymerase complex composed by the PB1, PB2, and PA subunits. Coloma et al. (2009) succeeded in building a 3D model of RNP by assembling 3D reconstructions from a non-symmetrical complex containing the polymerase (at 18 Å resolution) with the NP ring derived from a symmetrical volume (at 12 Å resolution). The docking of the atomic structures of NP and partial structures of PB1 and PA in this chimera map is shown in Fig. 2. The result, described in more detail by Coloma et al. (2009), is the first structural model for a functional viral RNP complex.

Visualization and modeling of small-angle scattering data

3D bead models of proteins in solution can be determined from 1D scattering data, in particular from SAXS (Chacón et al. 2000). Wriggers and Chacón (2001a) extended existing *Situs* tools to provide an atomic interpretation of SAXS-derived shapes. The workflow and the programs that are used to dock an atomic structure into low-resolution SAXS models are shown schematically in Fig. 3. The bead models can be transformed into volumetric maps for subsequent docking using convolution with a hard sphere kernel (*pdb2vol* tool). The SAXS modeler then has access to all docking strategies supported by *Situs*, including correlation-based docking (*colacor*/*colores*) and point cloud matching (*qrange*/*matchpoint*), and even flexible fitting (see below). To test the docking accuracy, we added the *pdb2saxs* tool to map atomic structures of trial proteins to hexagonal close-packed lattices with variable bead radii. The resulting models served as "simulated" low-resolution data in Wriggers and Chacón (2001a): For >100 beads typically arising in SAXS models, a rigid body docking precision can be achieved of the order of an Angstrom.

Another specific problem in the interpretation of SAXS data is the visualization of the beads. We found it useful to

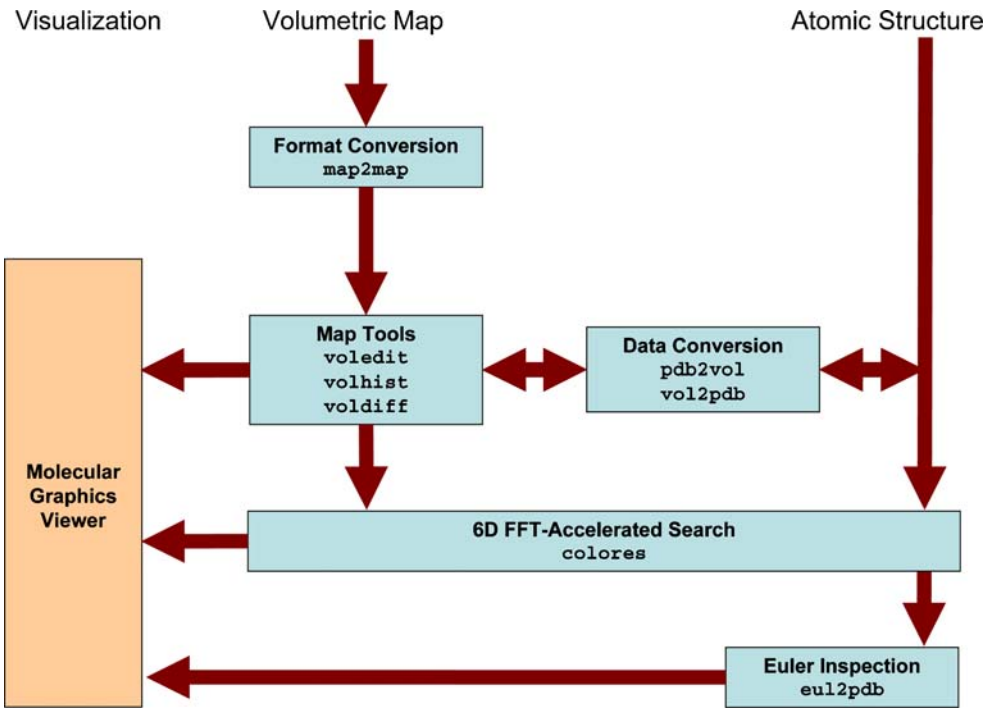


Fig. 1 Schematic diagram of *colores*-related routines in *Situs* 2.5. Major *Situs* components (*blue*) are classified by their functionality. The main workflow is indicated by brown arrows. The visualization (*orange*) for the rendering of the models requires a molecular graphics viewer such as *VMD* (Humphrey et al. 1996), *Chimera* (Pettersen et al. 2004), or *Sculptor* (http://sculptor.biomachina.org). Standard volumetric map formats are converted to cubic lattices in *Situs* format with the *map2map* utility. Subsequently, the data are inspected and, if necessary, prepared for the fitting using a variety of visualization and analysis tools. *Situs* docking tools require one volume (target) and one PDB structure (probe). Atomic coordinates in PDB format can be transformed to low-resolution maps, if necessary, and vice versa, to allow the docking of maps to maps or structures to structures. The resulting docked complex can be inspected in the graphics program. In addition, if a subset of Euler angles is chosen, these can be inspected (after conversion into PDB format) with the *eul2pdb* tool

render not the densely packed beads themselves, but rather an envelope that can be created by isocontouring a volumetric map that was created by convolution with a soft kernel such as a Gaussian (using *pdb2vol*).

Our approach to rendering and interpretation of SAXS data has been adopted by other groups (Lipfert et al. 2007; Fagan et al. 2009). Here, we highlight a recent *Nature* article on structural studies of the single epidermal growth factor receptor family member (dEGFR) in *Drosophila melanogaster*. Alvarado et al. (2009) determined the 2.7 Å X-ray crystal structure of the unliganded dEGFR extracellular region, encompassing domains I to IV (s-dEGFRΔV). A structural overlay of an active, extended, receptor tyrosine kinase sErbB2 and s-dEGFRΔV showed them to be remarkably similar, with important functional implications. One key question was whether crystal packing causes s-dEGFRΔV to be extended. This hypothesis was ruled out by SAXS studies of s-dEGFRΔV and complete s-dEGFR (Fig. 4). The *Situs*-derived models, shown in Fig. 4, indicate that s-dEGFRΔV is extended in solution (the envelope readily encompasses the crystal structure), and that domain V (orange) simply projects from the end of domain IV (red) to extend the structure further.

Flexible fitting

Rigid-body docking, as described above, laid the groundwork for the development of a flexible docking technique that brings deviating features of multi-resolution structures into

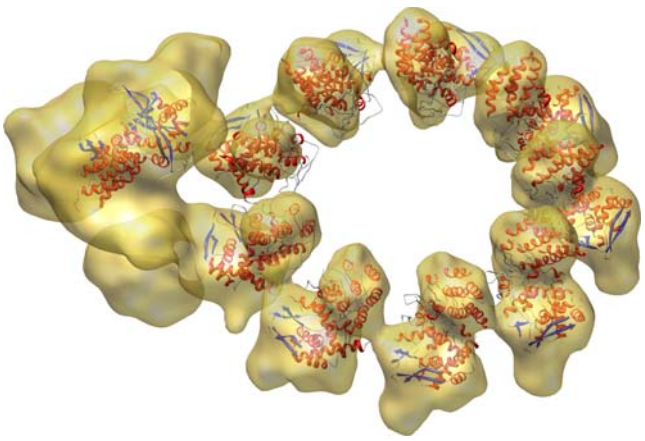


Fig. 2 Docking of the partially solved PA-PB1 complex (*left*) and NP (*ring*) monomers into the RNP structure (*transparent*) using *colores*. The graphic, kindly provided by Jaime Martín-Benito Romero, visually combines the results of Figs. 2–4 of Coloma et al. (2009)

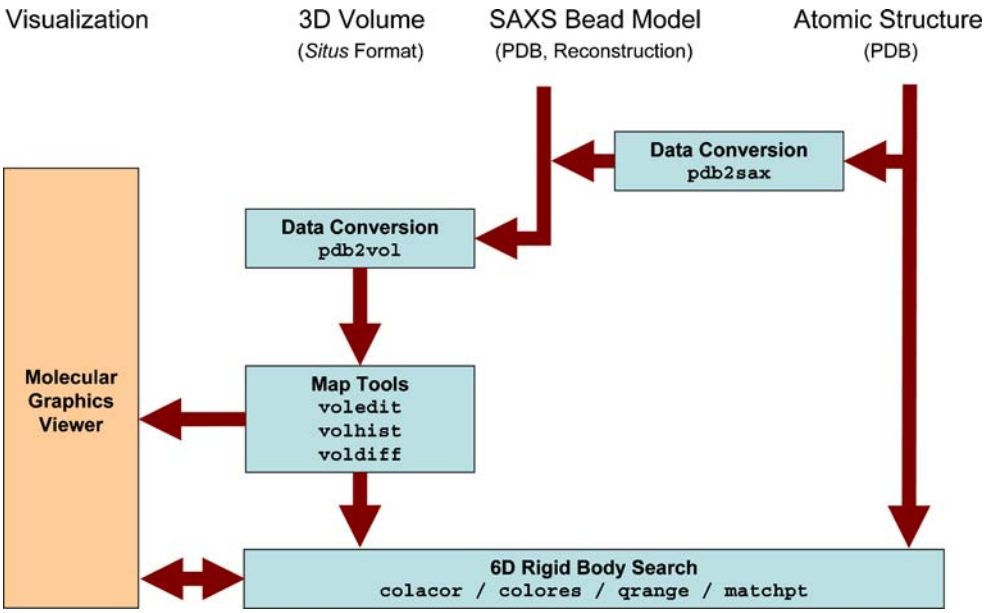


Fig. 3 Schematic diagram of SAXS-related routines in *Situs* 2.5. Visualization and modeling of atomic structures into SAXS bead models are supported through a conversion into 3D volumes representing the beads using the *pdb2vol* kernel convolution tool. Docking between atomic structures and 3D maps can be achieved through a number of approaches (see text). The data can be prepared further for the visualization using a variety of analysis and editing tools. Optional Gaussian kernel convolution with *pdb2vol* facilitates the smoothing of bead surfaces for their visualization in the form of density isocontours

register. In such situations, the atomic structure is moved towards the target density by systematically reducing the rms deviation between coarse-grained control points in a refinement of the atomic structure. One of the open questions in flexible docking is how to maintain the stereochemical quality of a fitted structure, since any over-fitting to noisy experimental data would compromise the quality of the atomic model. In an earlier review article (Wriggers et al. 2004), we described the details of a significant improvement to our flexible fitting algorithm, the Motion Capture Network (MCN). The basic idea of the workflow, depicted in Fig. 5, is that lateral connections (distance constraints) are formed between control points that reflect the connectivity of the biological polypeptide chain. This approximation of the movement can be justified by the statistics of biomolecular domain motions documented in the Protein Data Bank (PDB). In the following, a (previously unpublished) modeling of the actomyosin complex illustrates MCN-based flexible fitting.

An atomic model of F-actin (Holmes et al. 2003) was fitted to the 14 Å resolution actomyosin map (data kindly provided by Rasmus R. Schröder, now at University of Heidelberg, during his visit to Houston in 2003). The F-actin structure allowed us to create a mask for a single myosin S1 unit by low-pass filtering from the docked atomic structure using *pdb2vol*. As described by Wriggers and Chacón (2001b), the mask was needed by the tools *voledit* and *voldiff* to segment and subtract densities from actin and neighboring symmetry-related subunits and to obtain the density of a single myosin S1 from the helical 3D map. This single myosin S1 map was then compared to the atomic structure.

We first attempted rigid-body fitting of the atomic model, taken from the supplementary structure "motor domain.pdb" (Holmes et al. 2003), into the 3D map with *colores*, as described above. Rigid-body docking was not satisfactory with respect to the position of the upper 50K domain and the lever arm, even when performed independently for each structural subunit. Therefore, we subjected the predicted atomic model to flexible docking (Fig. 6) to characterize the observed changes. The flexible docking procedure was based on a connected MCN of identified features within the atomic model (Wriggers et al. 2004). The atomic model was allowed to move according to

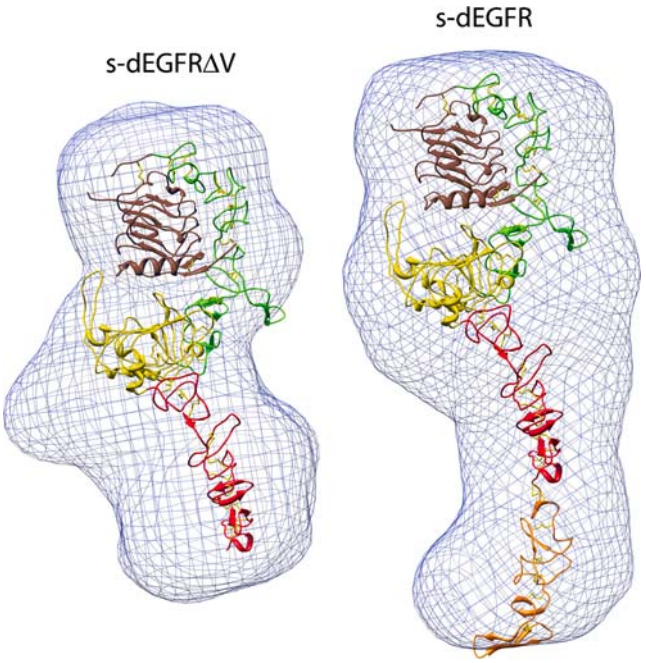


Fig. 4 Low-resolution molecular envelopes from SAXS studies of s-dEGFRΔV (*left*) and s-dEGFR (*right*). The envelopes (*blue*), created after kernel convolution with *pdb2vol*, readily accommodate the *Situs*-docked crystallographic models (see text). The graphic, kindly provided by Diego Alvarado and Mark Lemmon, emphasizes interior domain details; see also Fig. 2b in Alvarado et al. (2009)

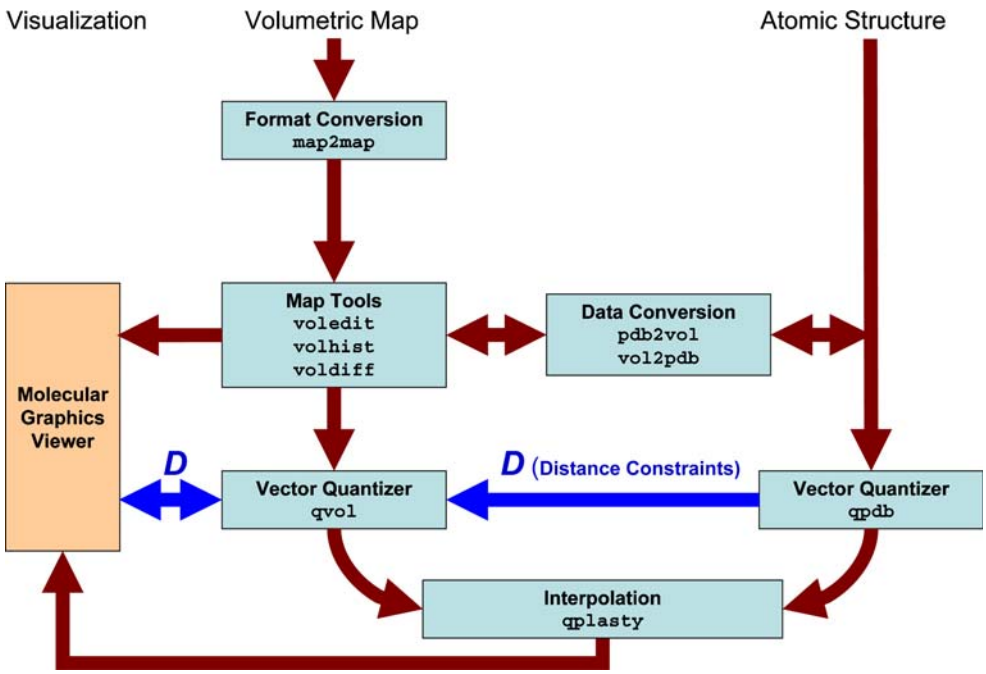


Fig. 5 Schematic diagram of flexible fitting with *Situs* 2.5. The modeling of distance constraints for the MCN is shown in *dark blue*. Standard volumetric map formats are converted with the *map2map* utility and the data can be prepared for coarse graining by vector quantization using a variety of map visualization and analysis tools. Atomic coordinates in PDB format can be transformed to low-resolution maps, if necessary, and vice versa. During vector quantization of the high-resolution structure, distances can be learnt that are sent to the vector quantizer of the low-resolution structure to enable MCN-based fitting. After the vector quantization, the high-resolution structure is flexibly docked by the *qplasty* tool (Rusu et al. 2008). As an alternative to spatial interpolation with *qplasty*, a molecular dynamics refinement is also supported

displacements tracked by 10 control points defined by the network, to find the best match to the cryo-EM map. The number of control points was judged to be sufficient for capturing the shape details of the single S1 map that occupies a volume of 185,000 $Å^3$ at the isocontour level shown (Fig. 6). The number of independent pieces of information contained in the 14 Å resolution map is then $185{,}000/14^3 \approx 67$. This number comprises an upper bound for the number of recognizable features in this particular volume. The conservative choice of 10 points (corresponding to a spatial resolution of 26 Å in the reduced network) was significantly below this upper bound to avoid an over-fitting of the data (Wriggers and Chacón 2001b). This level of detail, however, was quite sufficient for the flexing.

The longitudinal distance constraints in the MCN were assigned manually, as described by Wriggers et al. (2004), by following the connectivity of the polypeptide chain and to ensure robustness of the control points during the shape change. We found by trial and error that motion capture was best achieved through allocating more flexibility to the 50K regions (effectively allowing cleft closure) by eliminating all constraints on the motion of control points in this region. The final network used for the automated flexing is shown in Fig. 6.

We performed the flexing by adding a constraint energy function to the Hamiltonian of a molecular dynamics simulation that penalizes global shape differences between the data sets (Wriggers et al. 2004). In the molecular dynamics run, we added water molecules predicted by DOWSER (Zhang and Hermans 1996) to the system, which resulted in a total system size of 12,008 atoms.

One can expect that at 14 Å resolution the flexing faithfully reproduces conformational differences with a precision of 2 Å if atomic structures are locally conserved (Wriggers et al. 2004). Side chains are rearranged automatically to accommodate global conformational changes. Otherwise, the algorithm leaves the initial structure intact at the local level. Whether this assumption holds depends on the nature of the conformational difference between the two isoforms, which is not known a priori. However, it has been shown that only about 7% of protein domain rearrangements documented in the PDB are irregular motions where the tertiary structure is significantly perturbed (Gerstein and Krebs 1998). Therefore, it is plausible, at least for the predominantly hinge-type domain motions exhibited by myosin, that the low-resolution flexible fitting approach visualizes conformational changes with a precision of single amino acid residues. The final flexing-induced rms deviation in the atomic model was 5.3 Å.

To validate the precision and probe for systematic errors, we also performed a control flexing calculation on the structure of myosin 5 (Coureux et al. 2003). Myosin 5 is

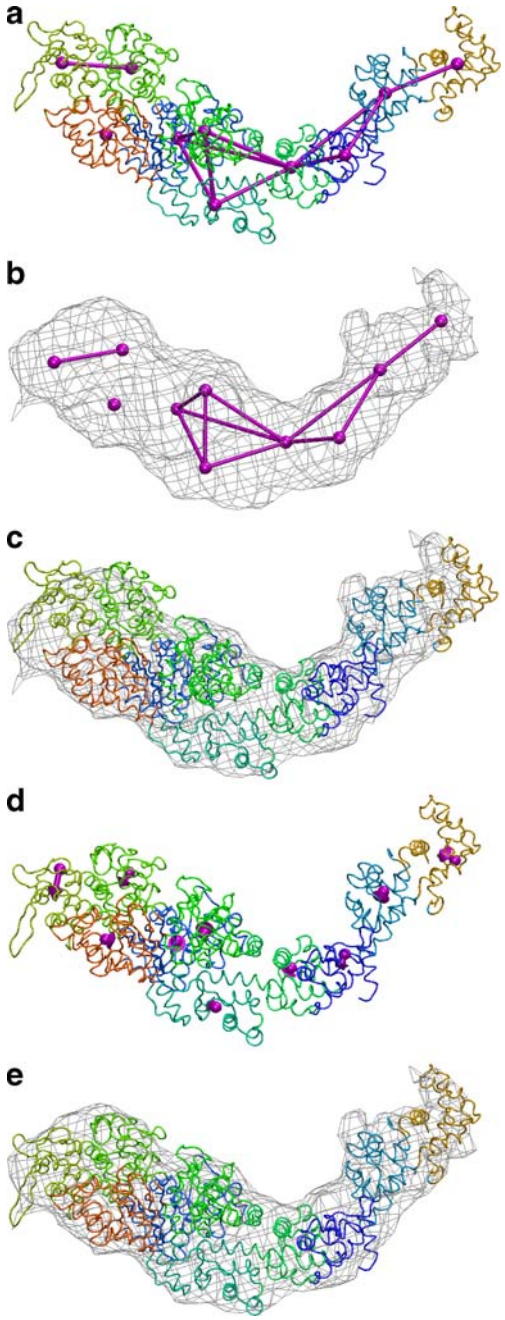


Fig. 6 Flexible fitting of myosin 2 subfragment 1. **a** MCN (see text) and Voronoi tessellation of the atomic structure. **b** MCN fitted to the segmented EM density. **c** Before flexing. **d** Displacements sampled at the control points. **e** After flexing

deemed to be in closer agreement with the 3D map of S1 in actomyosin (Holmes et al. 2004). Following the above protocols, we created a model of myosin 5, resulting in a total system size of 11,150 atoms. The observed flexing-induced rms deviation in the atomic model was 3.8 Å, which was indeed much lower than that observed in the myosin 2 case.

The above tests validate the *Situs*-based flexible fitting approach with a real EM data set. More detailed and systematic tests of flexible fitting were published in Rusu et al. (2008). In addition, the myosin fitting was recently extended to full thick filaments of tarantula muscle in collaboration with the group of Raúl Padrón in Venezuela (Alamo et al. 2008).

Conclusion

One key to the success of *Situs* over the years has been that the programs were ported to multiple platforms and their source code was freely available on the Internet (http://situs.biomachina.org). While we strive to teach at workshops and symposia, it seems that many researchers prefer to explore software in their own laboratories.

Our web-based tutorials have helped hundreds of electron microscopists and small- angle scattering experts to learn the use of the programs. For their dissemination, we obtained our own web domain (http://biomachina.org) and web server.

Another helpful aspect was the modular design of the programs. As mentioned above, the software consists of multiple, stand-alone tools that can be combined in various creative ways. The modular design allowed us to update individual programs over time (inevitably, it becomes necessary to update algorithms and to implement bug fixes for problems reported to us). We are managing an e-mail list to communicate with the more than 2,000 registered users who opted to receive information. Readers should feel free to send comments to situs@biomachina.org.

This brief review primarily focused on the scientific use of the software, but the development of *Situs* was also a personal journey for the author, with many memorable encounters along the way. A personal history of this work, as well as annotated references, can be found in the electronic supplementary material.

Acknowledgements I would like to thank Jaime Martín-Benito Romero, Diego Alvarado, and Mark Lemmon for providing Figs. 2 and 4, and for critical reading of this manuscript. In addition, I thank Rasmus R. Schröder and Kenneth C. Holmes for discussions. This work was supported in part by grants from National Institutes of Health (R01GM62968), the Alfred P. Sloan Foundation (BR-4297), and the Human Frontier Science Program (RGP0026/2003).

References

Alamo L, Wriggers W, Pinto A, Bártoli F, Salazar L, Zhao F-Q, Craig R, Padrón R (2008) Three-dimensional reconstruction of tarantula myosin filaments suggests how phosphorylation may regulate myosin activity. J Mol Biol 384:780–797

Alvarado D, Klein DE, Lemmon MA (2009) ErbB2 resembles an autoinhibited invertebrate epidermal growth factor receptor. Nature 461:287–291

Baker TS, Johnson JE (1996) Low resolution meets high: towards a resolution continuum from cells to atoms. Curr Opin Struct Biol 6:585–594

Chacón P, Wriggers W (2002) Multi-resolution contour-based fitting of macromolecular structures. J Mol Biol 317:375–384

Chacón P, Díaz JF, Morán F, Andreu JM (2000) Reconstruction of protein form with X-ray solution scattering and a genetic algorithm. J Mol Biol 299:1289–1302

Coloma R, Valpuesta JM, Arranz R, Carrascosa JL, Ortín J, Martín-Benito J (2009) The structure of a biologically active influenza virus ribonucleoprotein complex. PLoS Pathog 5:e1000491

Coureux PD, Wells AL, Menetrey J, Yengo CM, Morris CA, Sweeney HL, Houdusse A (2003) A structural state of the myosin V motor without bound nucleotide. Nature 425:419–423

Fagan RP, Albesa-Jové D, Qazi O, Svergun DI, Brown KA, Fairweather NF (2009) Structural insights into the molecular organization of the S-layer from Clostridium difficile. Mol Microbiol 71:1308–1322

Gerstein M, Krebs W (1998) A database of macromolecular motions. Nucl Acids Res 26:4280–4290

Holmes KC, Angert I, Kull FJ, Jahn W, Schröer RR (2003) Electron cryo- microscopy shows how strong binding of myosin to actin releases nucleotide. Nature 425:423–427

Holmes KC, Schröder RR, Sweeney HL, Houdusse A (2004) The structure of the rigor complex and its implications for the power stroke. Philos Trans R Soc Lond B 359:1819–1828

Humphrey WF, Dalke A, Schulten K (1996) VMD - Visual Molecular Dynamics. J Mol Graph 14:33–38

Lee EH, Hsin J, Sotomayor M, Comellas G, Schulten K (2009) Discovery through the computational microscope. Structure 17:1295–1306

Lipfert J, Das R, Chu VB, Kudaravalli M, Boyd N, Herschlag D, Doniach S (2007) Structural transitions and thermodynamics of a glycine-dependent riboswitch from Vibrio cholerae. J Mol Biol 365:1393–1406

Medalia O, Weber I, Frangakis AS, Nicastro D, Gerisch G, Baumeister W (2002) Macromolecular architecture in eukaryotic cells visualized by cryoelectron tomography. Science 298:1209–1213

National Institutes of Health. Technology development for high-resolution electron microscopy (PA-00-084). URL http://grants.nih.gov/grants/guide/pa-files/pa-00-084.html (accessed 10/15/2009). 2000

Niemann HH, Petoukhov MV, Härtlein M, Moulin M, Gherardi E, Timmins P, Heinz DW, Svergun DI (2008) X-ray and neutron small-angle scattering analysis of the complex formed by the Met receptor and the Listeria monocytogenes invasion protein InlB. J Mol Biol 377:489–500

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF Chimera - A visualization system for exploratory research and analysis. J Comp Chem 25:1605–1612

Rusu M, Birmanns S, Wriggers W (2008) Biomolecular pleiomorphism probed by spatial interpolation of coarse models. Bioinformatics 24:2460–2466

U.S. Department of Energy (2005) Energy department requests proposals for advanced scientific computing research. URL http://www.energy.gov/news/2823.htm (accessed 10/15/2009)

Wriggers W, Chacón P (2001a) Using *Situs* for the registration of protein structures with low-resolution bead models from X-ray solution scattering. J Appl Cryst 34:773–776

Wriggers W, Chacón P (2001b) Modeling tricks and fitting techniques for multi-resolution structures. Structure 9:779–788

Wriggers W, Milligan RA, McCammon JA (1999) Situs: A package for docking crystal structures into low-resolution maps from electron microscopy. J Struct Biol 125:185–195

Wriggers W, Chacón P, Kovacs J, Tama F, Birmanns S (2004) Topology representing neural networks reconcile biomolecular shape, structure, and dynamics. Neurocomputing 56:365–379

Zhang L, Hermans J (1996) Hydrophilicity of cavities in proteins. Proteins Struct Func Genet 24:433–438

Supplementary Material for

"Using *Situs* for the Integration of Multi-Resolution Structures" by Willy Wriggers

A. A Personal History of the Development of Situs

At the time of this writing, the *Situs* package has just passed the ten-year anniversary of its 1999 journal publication. Whereas the main article documents how *Situs* continues to be useful to the biophysical community, it may be worthwhile on this occasion to reflect on how the package came into existence.

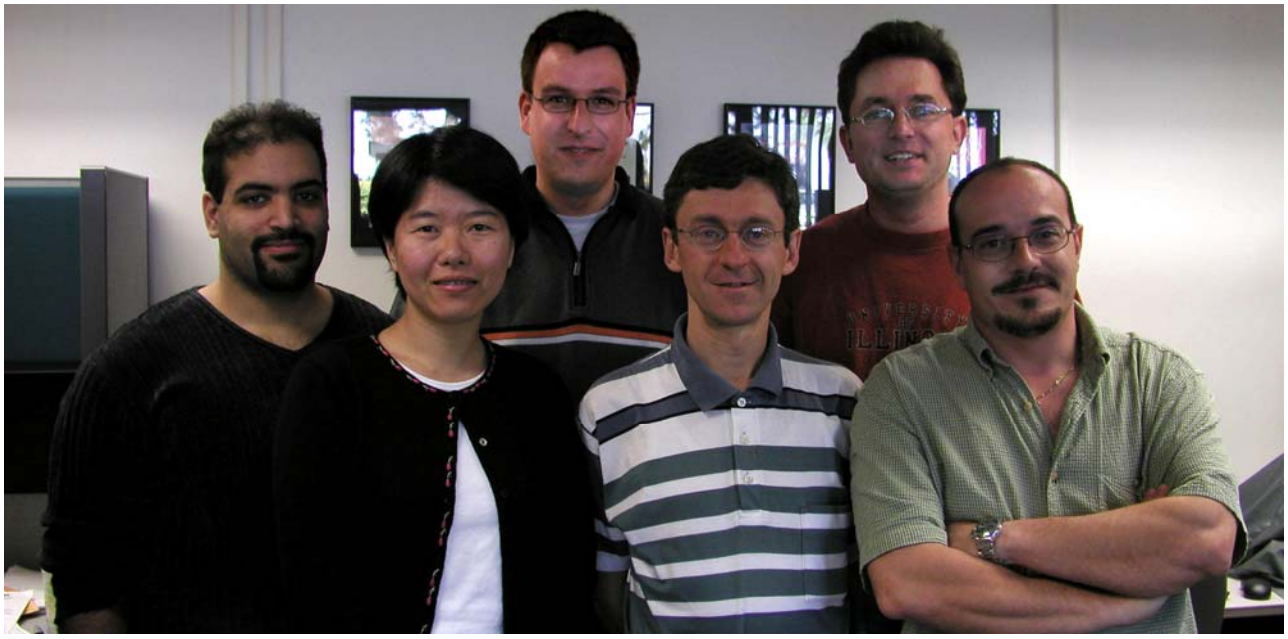
As a physics graduate student in Klaus Schulten's group at the University of Illinois at Urbana-Champaign (UIUC), I received rigorous training in biomolecular modeling and simulation. In the early 1990s, Schulten was still leading a formidable computational neuroscience program that he had built at Technical University Munich, Germany, and moved to Illinois in 1988. In particular, Schulten and visiting professor Annette Zippelius taught a course on the "Physics of Neural Networks" in the spring of 1993, which I attended as a graduate student at UIUC. For the term paper for this course, I studied the geometric properties of the Kohonen self-organizing map [1] and the "neural gas" algorithm that had been developed by Thomas Martinetz in the group [2,3]. The term paper and the corresponding programs proved valuable during the initial design of *Situs*. In the meantime, by the mid-1990s, Schulten decided to focus exclusively on biophysics. Consequently, I shifted my interests to molecular dynamics simulations of cell motility proteins and methods for simulating and mathematically characterizing large-scale conformational changes. In one resulting article, I predicted a conformational change of the kinesin motor when superimposed with a 2D projection of the microtubule density solved by the Milligan laboratory (Figure 12A in [4]). Although only in 2D, the figure proved inspirational during a visit to San Diego in the summer of 1997, and both Milligan and I saw a potential for extending this type of fitting to the third dimension.

In November 1997, I joined the McCammon group at UCSD for a postdoctoral tenure. J. Andrew McCammon, a power house in physical chemistry and biomolecular simulation, actively encouraged my growing interactions with Milligan, whose laboratory at The Scripps Research Institute (TSRI) was within walking distance. Given the freedom to chose suitable methodologies, I decided to continue the work on neural networks started with Schulten five years earlier. In January 1998, it became clear that the Martinetz neural gas algorithm would be highly valuable for biomolecular docking in that it could provide a coarse-grained representation of 3D data. McCammon suggested that I bring in my own funding for a joint postdoctoral research project and, together with Milligan, we were successful in securing a *La Jolla Interfaces in Science* (LJIS) postdoctoral fellowship. LJIS was an interdisciplinary and multi-institutional training program supported by the Burroughs Wellcome Fund that encouraged dual mentorship of postdocs in the La Jolla community. The fellows gained wide visibility in San Diego, which later proved important. In

August 1998, I also became formally affiliated with the Milligan laboratory at TSRI. This arrangement proved very productive; the underlying methods were tested in the McCammon group at UCSD and applied to experimental data at TSRI. To this end, I received instruction in electron microscopy (EM) theory and practice from Milligan and his friendly (and patient) coworkers. The first article resulting from the LJIS collaboration was published the same year [5].

The idea of developing *Situs* as a software package first took shape in the summer of 1998, mainly through interactions with professor Gina Sosinsky at UCSD. In late August, Sosinsky alerted me to an upcoming issue on molecular visualization software in the *Journal of Structural Biology*. These special issues are widely read in the EM community. A first version of the package was written, and the article [6] submitted less than eight weeks later, in retrospect an incredibly short time. *Situs* would be the first designated fitting package for EM, but I soon became aware of a competing article in the same journal issue. I realized that I could legally distribute my code freely on the web in open source form. Eager to be the first to roll out a package publicly, I created a tutorial and user guide and used McCammon's web server for their dissemination. Some early adopters heard about the web site and the first documented download of *Situs* before its official release occurred on November 10, 1998. Version 1.0 was officially announced on six mailing lists on February 1, 1999. The *Situs* paper [6] made the journal cover and appeared in April 1999. At around the same time, I received a phone call from Charles L. Brooks III (TSRI). Brooks was about to renew his NIH Resource grant and indicated that he could utilize some of the ideas; he asked if I would like to join TSRI as an assistant professor in his domain. At the time I was only 17 months into my postdoctoral tenure, but I saw an opportunity to continue my work seamlessly at a prestigious institution, and accepted. I remain indebted to Andy McCammon, who allowed me to move *Situs* to TSRI. A few weeks later I attended the Gordon Research Conference on 3D EM in Henniker, New Hampshire, June 20-25, 1999. The special journal issue with our cover figure was widely distributed at that meeting and I met many leading electron microscopists for the first time. The conference chair, Bridget Carragher, invited me to give a talk, and I gave the first public demonstration of *Situs* in a session organized by Gina Sosinsky and Michael Schatz (Image Science, Berlin, Germany). The events of 1999 were a highlight of my early career, and I enjoyed the sudden attention my work received.

What happened in the decade after the first distribution of the code? The original manuscript [6] was later recognized with the *Journal of Structural Biology* Paper of the Year Award [7]. Among several grants, I secured an NIH grant in 2001 (R01GM62968, renewed in 2006) that was directly based on the development of *Situs* and supported a growing group of coworkers. Supplementary Figure 1 shows the TSRI group in 2002.

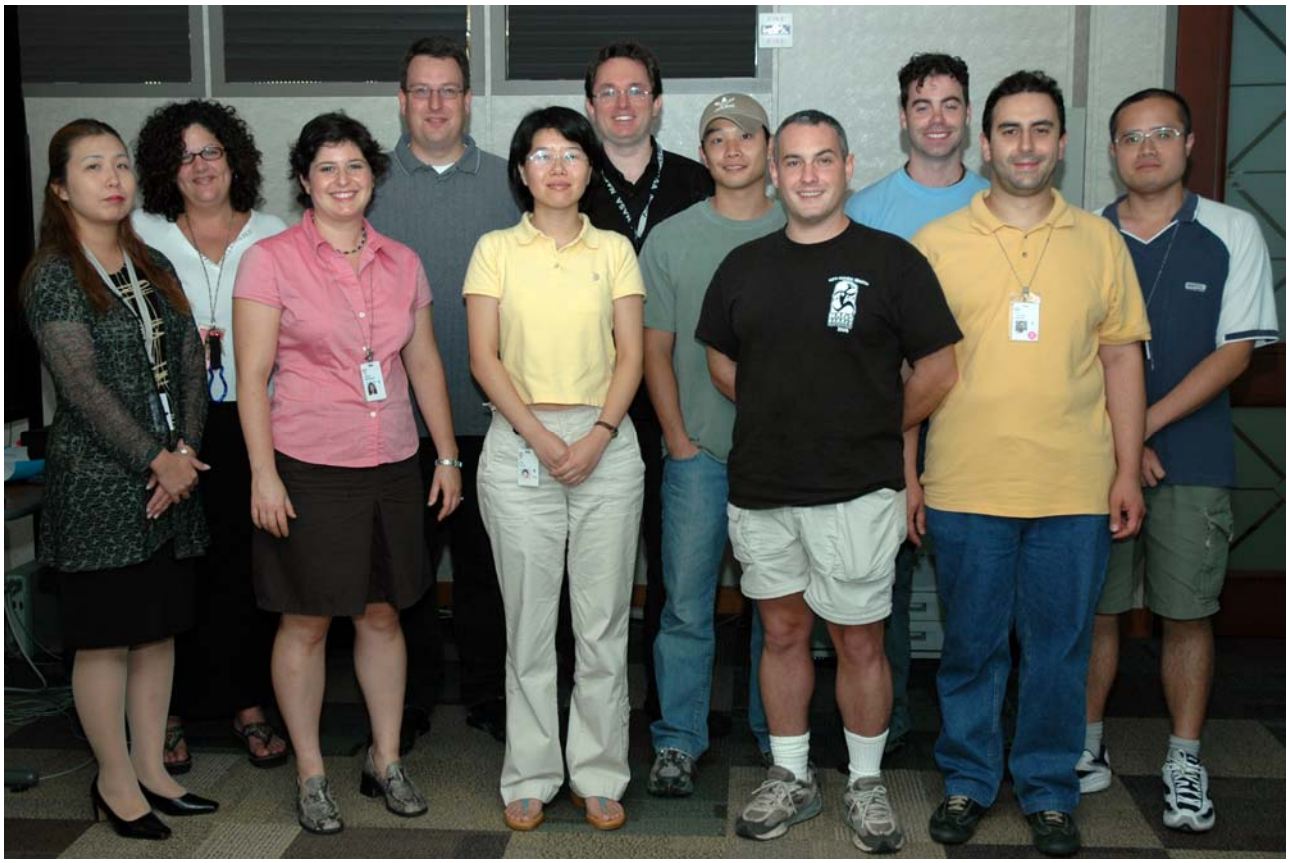


Supplementary Figure 1. Computational Structural Biology Group (TSRI), 2002. From Left: Essam Metwally, Yao Cong, Stefan Birmanns, Julio Kovacs, Willy Wriggers, and Pablo Chacón.

I cannot comment here on every one of my 30 coworkers over the years (for more information, see our alumni pages at URL http://wriggers.biomachina.org/people/alumni.html), but I'd like to state that attracting a good postdoc right away makes a big difference in productivity. I was fortunate to have Pablo Chacón contact me in 2000, shortly after I entered my position at TSRI. Pablo was well trained in biophysics and his experience fit the project perfectly. Brooks kindly provided the initial funding for Pablo's position until fellowships and grants were secured. At the time, my small team was still closely interacting with Ron Milligan's laboratory at TSRI. An amusing fact is that we mostly met to discuss science during Milligan's cigarette breaks (both Pablo and Milligan were smokers). Consequently, I almost ended up smoking myself (it was for a good cause; our collaborator was otherwise very busy!). Inspired in part by Pablo's background in scattering, we broadened the application of the *Situs* tools to other areas of biophysics such as small angle X-ray scattering [8]. The coarse graining that we initially used for rigid-body docking [9] also laid the groundwork for the later development of a flexible docking technique and elastic network models [10-12]. In a series of papers, we took advantage of Fourier correlation theory to rapidly scan the six translational and rotational degrees of freedom of a probe molecule relative to a (fixed) target density map [13-15].

After our move to Houston in 2003, my group at the University of Texas Health Science Center (UTHSCH) doubled in size (Supplementary Figure 2), and I found myself busy teaching courses. A trusted postdoctoral teaching assistant in my “Computational Structural Biology” course was Sugoto Chakravarty, who co-authored a review article on biomolecular linkers based in part on *Situs* calculations [16]. We were also fortunate to have Stefan Birmanns, a computer scientist who had collaborated with the group over the years while completing his Ph.D. in Jülich, Germany, join us as a postdoc to coordinate the group's hardware and software development. Although Stefan primarily developed his own visualization software, *Sculptor* (http://sculptor.biomachina.org), he made substantial contributions to *Situs* tools with next-generation algorithms [9,12]. Although

many more of my group members deserve praise, the above three postdoctoral coworkers (Chacón, Chakravarty, and Birmanns) stand out because they each moved on to academic faculty positions.



Supplementary Figure 2. Laboratories for Biocomputing and Imaging (UTHSCH), 2004. From Left: Namiko Burleson, Joan Zuñiga, Hilary Wriggers, Stefan Birmanns, Yao Cong, Willy Wriggers, Char Hu, Chance Coble, Paul Boyle, Valerio Mariani, and Zhiyong Zhang.

Here, I would also like to honor the memory of UTHSCH graduate student Paul C. Boyle from Killibegs, Ireland, who died in a tragic car accident on November 29, 2004. Paul was a bright and cheerful coworker, and made many contributions in my group in Houston (Supplementary Figure 2). Paul was working on a new release of *Situs* and carried out molecular dynamics simulations to simulate Epidermal Growth Factor [17]. He also contributed a software program (now in *Sculptor*) that allows for the inspection of volumetric maps. Paul was posthumously awarded an M.S. degree, and UTHSCH has established an annual student achievement award in his name.

Like most scientists, software developers are judged by the quality of scientific work in peer-reviewed publications. On the other hand, to gain acceptance in the user community, it is important to offer support and hands-on outreach programs. Consequently, a balance is needed between the two modes of communication. Although we participate in meetings and symposia by other groups, we have also found it useful to organize our own seminars and workshops to connect more directly with our user community (programs, lecture materials, and video recordings of the following events are available at URL http://wriggers.biomachina.org/events):

In 2001 we organized an afternoon workshop, “Docking Do's and Don'ts,” that was attended by more than 100 participants at the Gordon Research Conference on 3D Electron Microscopy, Bristol, RI, June 24-29. In San Diego in 2003, we organized a “*Situs* EM/X-tal Fitting and Modeling Workshop,” February 3-5, with 29 participants from the US, Europe, and Japan. In addition to a lectureship series of external speakers that took place in Houston approximately every second month, we also organized a complementary Houston town meeting every other month for local speakers. Overall, we hosted 27 visiting scientists on site and trained them in the usage of our software.

Our greatest venture to date was a six-day international workshop in late April 2006 on “Innovations in Nanoscale Modeling and Imaging of Biological Systems” in Houston (Supplementary Figure 3). The workshop featured 21 invited speakers and 50 invited participants from as far away as Germany, Israel, Singapore, South Africa, Spain, and the United Kingdom. In addition, about 30 Houston-area scientists attended the public lectures. Hands-on computer sessions trained participants in the use of software developed by selected speakers and by our group. The speakers were mainly early-career software developers at other institutions that I knew from my research and from international conferences. Many new collaborations were fostered during this time, and we were able to connect with the users of our software, who presented their own work on posters. As an organizer, I found it to be an incredibly motivating experience to meet with the enthusiastic users in person.

Supplementary Figure 3. “Innovations in Nanoscale Modeling and Imaging of Biological Systems” in Houston, April 17-23, 2006. (Left) Opening speakers (from left): Robert Glaeser, Willy Wriggers, Jack Smith, and Zhiyong Zhang. (Right) Group picture of late night participants at Cabo MixMex Grill, downtown Houston.

Ron Milligan once commented that I was “clever or lucky, it doesn't much matter which” in getting into computational biophysics at this particular time. In 2006 I was awarded tenure at UTHSCH, in part based on the success of the *Situs* package and the organizational and grantsmanship milestones we achieved. I went from student to associate professor in nine years and had my share of tragedy

and triumph along the way. What would be the next challenge in this personal journey? After some soul searching I found I was happiest when I had the opportunity to be personally involved in science, and I wanted to free myself form administrative responsibilities as much as possible so that I could focus more on research and learning. I knew that such a rejuvenation would require either a longer sabbatical or a permanent relocation. I was fortunate in being selected to join the privately funded laboratory, D. E. Shaw Research, in Manhattan (http://www.DEShawResearch.com). Stefan Birmanns took the helm in Houston, so I was able to leave my coworkers in good hands while exploring a new phase in my career.

Although my research interests have diversified, *Situs* continues to be useful to many researchers in biophysics. Over the years, more than 2,000 users have registered the download, and a growing number of research groups use the software (about 100 research projects and more than 1,000 citations refer to our work). With the assistance of collaborators, alumni, and enthusiasts in the community, I hope to continue the development of the package in the next decade and beyond.

References:

[1] *Teuvo Kohonen*. Automatic Formation of Topological Maps of Patterns in a Self-Organizing System. 1981. In Oja, E. and Simula, O., editors, *Proceedings of 2SCIA, Scand. Conference on Image Analysis*, pp. 214-220, Helsinki, Finland.

[2] *Thomas* M. *Martinetz*, Stanislav G. *Berkovich, and Klaus Schulten*. "Neural Gas" for Vector Quantization and its Application to Time-Series Prediction. 1993. *IEEE Trans. Neural Networks Vol.* 4, pp. 558–569.

[3] *Thomas Martinetz and Klaus Schulten*. Topology Representing Networks. *Neural Networks*, 1994, Vol. 7, pp. 507-522.

[4] *Willy Wriggers and Klaus Schulten.* Nucleotide-Dependent Movements of the Kinesin Motor Domain Predicted by Simulated Annealing. *Biophysical Journal*, 1998, Vol. 75, pp. 646-661.

[5] *Willy Wriggers, Ronald A. Milligan, Klaus Schulten, and J. Andrew McCammon.* Self-Organizing Neural Networks Bridge the Biomolecular Resolution Gap. *J. Molecular Biology*, 1998, Vol. 284, pp. 1247-1254.

[6] *Willy Wriggers, Ronald A. Milligan, and J. Andrew McCammon. Situs*: A Package for Docking Crystal Structures into Low-Resolution Maps from Electron Microscopy. *J. Structural Biology*, 1999, Vol. 125, pp. 185-195.

[7] Announcement: Paper of the Year Award. *J. Struct. Biol.* 2001, Vol. 135, pp. 79-81.

[8] *Willy Wriggers and Pablo Chacón.* Using *Situs* for the Registration of Protein Structures with Low-Resolution Bead Models from X-ray Solution Scattering. *J. Appl. Cryst.*, 2001, Vol. 34, pp. 773-776.

[9] *Stefan Birmanns and Willy Wriggers.* Multi-Resolution Anchor-Point Registration of Biomolecular Assemblies and Their Components. *J. Struct. Biol.* 2007, Vol. 157, pp. 271-280.

[10] *Pablo Chacón, Florence Tama and Willy Wriggers.* Mega-Dalton Biomolecular Motion Captured from Electron Microscopy Reconstructions. *J. Mol. Biol.*, 2003, Vol. 326, pp. 485-492.

[11] *Willy Wriggers, Pablo Chacón, Julio Kovacs, Florence Tama and Stefan Birmanns.* Topology Representing Neural Networks Reconcile Biomolecular Shape, Structure, and Dynamics. *Neurocomputing*, 2004, Vol. 56, pp. 365-379.

[12] *Mirabela Rusu, Stefan Birmanns, and Willy Wriggers*. Biomolecular Pleiomorphism Probed by Spatial Interpolation of Coarse Models. *Bioinformatics* 2008, Vol. 24, pp. 2460-2466.

[13] *Pablo Chacón and Willy Wriggers.* Multi-Resolution Contour-Based Fitting of Macromolecular Structures. *J. Mol. Biol.*, 2002, Vol 317, pp. 375-384.
[14] *Julio A. Kovacs, Pablo Chacón, Yao Cong, Essam Metwally, and Willy Wriggers*. Fast Rotational Matching of Rigid Bodies by Fast Fourier Transform Acceleration of Five Degrees of Freedom. *Acta Cryst. D*, 2003, Vol. 59, pp. 1371-1376.
[15] *Andreas G. Martin, Frank Depoix, Michael Stohr, Ulrich Meissner, Silke Hagner-Holler, Kada Hammouti, Thorsten Burmester, Jochen Heyd, Willy Wriggers, and Juergen Markl*. Limulus polyphemus hemocyanin: 10 A structure, sequence analysis, molecular modelling and rigidbody fitting reveals the interfaces between the eight hexamers. *J. Mol. Biol.* 2007, Vol. 366, pp. 1332-1350.
[16] *Willy Wriggers, Sugoto Chakravarty, and Patricia A. Jennings*. Control of Protein Functional Dynamics by Peptide Linkers. *Peptide Science*, 2005, Vol. 80, pp. 736-746.
[17] *Zhiyong Zhang, Paul C. Boyle, Bao-Yuan Lu, Jui-Yoa Chang, and Willy Wriggers*. Entropic Folding Pathway of Human Epidermal Growth Factor Explored by Disulfide Scrambling and Amplified Collective Motion Simulations. *Biochemistry.* 2006, Vol. 45, pp. 15269-15278.

B. Personally Annotated List of References for the Main Article

L. Alamo, W. Wriggers, A. Pinto, F. Bártoli, L. Salazar, F.-Q. Zhao, R. Craig, and R. Padrón. Three-dimensional reconstruction of tarantula myosin filaments suggests how phosphorylation may regulate myosin activity. J. Mol. Biol., 384:780–797, 2008.

Application of *Situs* flexible fitting to the thick filament of tarantula myosin.

D. Alvarado, D. E. Klein, and M. A. Lemmon. ErbB2 resembles an autoinhibited invertebrate epidermal growth factor receptor. Nature, 461:287–291, 2009.

Application of *Situs* to the docking and visualizing of SAXS data.

T. S. Baker and J. E. Johnson. Low resolution meets high: towards a resolution continuum from cells to atoms. Curr. Opinion Struct. Biol., 6:585–594, 1996.

A good early review of docking challenges in EM.

P. Chacón and W. Wriggers. Multi-resolution contour-based fitting of macromolecular structures. J. Mol. Biol., 317:375–384, 2002.

The development of the Laplacian correlation criterion for intermediate-resolution docking.

P. Chacón, J.F.Díaz, F. Morán, and J. M. Andreu. Reconstruction of protein form with X-ray solution scattering and a genetic algorithm. J. Mol. Biol., 299:1289–1302, 2000.

3D bead model reconstruction from 1D SAXS data.

R. Coloma, J. M. Valpuesta, R. Arranz, J. L. Carrascosa, Juan Ortín, and J. Martín-Benito. The structure of a biologically active influenza virus ribonucleoprotein complex. PLoS Pathog., 5:e1000491, 2009.

Application of *Situs* to the modeling and atomic interpretation of a large complex solved with EM.

P. D. Coureux, A. L. Wells, J. Menetrey, C. M. Yengo, C. A. Morris, H. L. Sweeney, and A. Houdusse. A structural state of the myosin V motor without bound nucleotide. Nature, 425:419–423, 2003.

Crystal structure of myosin 5 used in section 3 of the main text.

R. P. Fagan, D. Albesa-Jové, O. Qazi, D.I. Svergun, K. A. Brown, and N.F. Fairweather. Structural insights into the molecular organization of the S-layer from Clostridium difficile. Mol. Microbiol., 71:1308–1322, 2009.

Application of *Situs* to docking and visualizing SAXS data.

M. Gerstein and W. Krebs. A database of macromolecular motions. Nucl. Acids Res., 26:4280–4290, 1998.

Classic survey of the PDB with observed frequencies of hinge, shear, and irregular motions.

K. C. Holmes, I. Angert, F. J. Kull, W. Jahn, and R. R. Schröder. Electron cryomicroscopy shows how strong binding of myosin to actin releases nucleotide. Nature, 425:423–427, 2003.

Manual rigid-body docking of myosin 2 into actomyosin.

K. C. Holmes, R. R. Schröder, H. L. Sweeney, and A. Houdusse. The structure of the rigor complex and its implications for the power stroke. Philos. Trans. R. Soc. Lond. B, 359:1819–1828, 2004.

Manual rigid-body docking of myosin 5 into actomyosin.

W. F. Humphrey, A. Dalke, and K. Schulten. VMD – Visual Molecular Dynamics. J. Mol. Graphics, 14:33–38, 1996.

Standard reference of the VMD graphics package (compatible with *Situs*).

E. H. Lee, J. Hsin, M. Sotomayor, G. Comellas, and K. Schulten. Discovery through the computational microscope. Structure, 17:1295–1306, 2009.

A good review on the use of computing to drive scientific discoveries, focusing on molecular dynamics simulations.

J. Lipfert, R. Das, V. B. Chu, M. Kudaravalli, N. Boyd, D. Herschlag, and S. Doniach. Structural transitions and thermodynamics of a glycine-dependent riboswitch from Vibrio cholerae. J. Mol. Biol., 365:1393–1406, 2007.

Application of *Situs* to docking and visualizing SAXS data.

O. Medalia, I. Weber, A. S. Frangakis, D. Nicastro, G. Gerisch, and W. Baumeister. Macromolecular architecture in eukaryotic cells visualized by cryoelectron tomography. Science, 298:1209–1213, 2002.

A major breakthrough in electron tomography of individual subcellular components.

National Institutes of Health. Technology development for high-resolution electron microscopy (PA-00-084). URL http://grants.nih.gov/grants/guide/pa-files/pa00-084.html (accessed 10/15/2009). 2000.

NIH program announcement implementing new methodological award criteria.

H. H. Niemann, M. V. Petoukhov, M. Härtlein, M. Moulin, E. Gherardi, P. Timmins, D. W. Heinz, and D. I. Svergun. X-ray and neutron small-angle scattering analysis of the complex formed by the Met receptor and the Listeria monocytogenes invasion protein InlB. J. Mol. Biol., 377:489–500, 2008.

A recent application of both neutron and X-ray solution scattering to determine low-resolution 3D models.

E. F. Pettersen, T. D.Goddard,C. C. Huang,G.S.Couch, D. M. Greenblatt, E. C. Meng, and T. E. Ferrin. UCSF Chimera -A visualization system for exploratory research and analysis. J. Comp. Chem., 25:1605–1612, 2004.

Standard reference of the Chimera graphics package (compatible with *Situs*).

M. Rusu, S. Birmanns, and W. Wriggers. Biomolecular pleiomorphism probed by spatial interpolation of coarse models. Bioinformatics, 24:2460–2466, 2008.

Simplified flexible fitting with *Situs* using spatial interpolation.

U. S. Department of Energy. Energy department requests proposals for advanced scientific computing research. URL http://www.energy.gov/news/2823.htm (accessed 10/15/2009). 2005.

DOE program announcement describing the utility of scientific computing in the physical sciences.

W. Wriggers and P. Chacón. Using *Situs* for the registration of protein structures with low-resolution bead models from X-ray solution scattering. J. Appl. Cryst., 34: 773–776, 2001a.

First paper describing the use of *Situs* for SAXS modeling.

W. Wriggers and P. Chacón. Modeling tricks and fitting techniques for multi-resolution structures. Structure, 9:779–788, 2001b.

An early review article of integrative modeling techniques in biophysics.

W. Wriggers, R. A. Milligan, and J. A. McCammon. *Situs*: A package for docking crystal structures into low-resolution maps from electron microscopy. J. Struct. Biol., 125:185–195, 1999.

The original *Situs* article.

W. Wriggers, P.Chacón, J. Kovacs, F. Tama, and S. Birmanns. Topology representing neural networks reconcile biomolecular shape, structure, and dynamics. Neurocomputing, 56:365–379, 2004.

An early review of the use of coarse graining in flexible fitting and haptic rendering.

L. Zhang and J. Hermans. Hydrophilicity of cavities in proteins. Proteins: Struc. Func. and Genetics, 24:433–438, 1996.

Reference for the DOWSER package for the placement of water into buried cavities for a more accurate molecular dynamics refinement of atomic structures.