



**HAL**  
open science

## Molecular simulations and visualization: introduction and overview.

Jonathan D Hirst, David R Glowacki, Marc Baaden

► **To cite this version:**

Jonathan D Hirst, David R Glowacki, Marc Baaden. Molecular simulations and visualization: introduction and overview.. Faraday Discussions, 2014, 169, pp.9-22. 10.1039/c4fd90024c . hal-01084628

**HAL Id: hal-01084628**

**<https://hal.science/hal-01084628>**

Submitted on 5 Dec 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Hirst, J. D., Glowacki, D. R., & Baaden, M. (2014). Molecular simulations and visualization: Introduction and overview. *Faraday Discussions*, 169, 9-22. <https://doi.org/10.1039/c4fd90024c>

Publisher's PDF, also known as Version of record

Link to published version (if available):  
[10.1039/c4fd90024c](https://doi.org/10.1039/c4fd90024c)

[Link to publication record in Explore Bristol Research](#)  
PDF-document

This is the final published version of the article (version of record). It first appeared online via RSC at <http://pubs.rsc.org/en/Content/ArticleLanding/2014/FD/C4FD90024C#!divAbstract>. Please refer to any applicable terms of use of the publisher.

## University of Bristol - Explore Bristol Research

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: <http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

# Molecular simulations and visualization: introduction and overview

Jonathan D. Hirst,<sup>a</sup> David R. Glowacki<sup>b,c,d,e</sup> and Marc Baaden<sup>\*f</sup>

Received 25th September 2014, Accepted 25th September 2014

DOI: 10.1039/c4fd90024c

Here we provide an introduction and overview of current progress in the field of molecular simulation and visualization, touching on the following topics: (1) virtual and augmented reality for immersive molecular simulations; (2) advanced visualization and visual analytic techniques; (3) new developments in high performance computing; and (4) applications and model building.

## Introduction

It is increasingly the case that biology, chemistry and materials science make extensive use of computational methods. The applications are diverse, spanning a range of timescales and lengthscales, from the cellular level (*e.g.* in systems biology) all the way down to detailed atomistic simulations of molecular assemblies, materials or small molecules. The 2013 Nobel Prize in Chemistry awarded to Martin Karplus, Michael Levitt and Arieh Warshel for the development of multiscale models for complex chemical systems is an indication of the achievements of such simulations. Molecular simulations and visualization offer fertile territory where research in human-computer interaction (HCI) and virtual reality may interact with and provide substantial benefit to computational molecular sciences. One of the principal challenges of this research area is that it is inherently cross-disciplinary and therefore requires deep exchanges beyond the boundaries of each discipline. In what follows, we outline progress in this emerging field, and offer a glimpse of potential directions. The discussion that follows is broken down into four interconnected topics. The first topic focuses on the use of virtual and augmented reality in the context of immersive molecular simulations. Progress here requires (a) advanced visualization and visual analytic

<sup>a</sup>School of Chemistry, University of Nottingham, Nottingham, NG7 2RD, UK<sup>b</sup>School of Chemistry, University of Bristol, Bristol, BS8 1TS, UK<sup>c</sup>Department of Computer Science, University of Bristol, BS8 1UB, UK<sup>d</sup>PULSE Institute and Dept of Chemistry, Stanford University, Stanford, CA 94305, USA<sup>e</sup>SLAC National Accelerator Laboratory, Menlo Park, California 94025, USA<sup>f</sup>Laboratoire de Biochimie Théorique, CNRS, UPR9080, Université Paris Diderot, Sorbonne Paris Cité, 13 rue Pierre et Marie Curie, 75005, Paris, France. E-mail: baaden@smplinux.de

techniques, and (b) the ability to harness new developments in high performance computing (*e.g.*, general-purpose-GPUs (GP-GPUs), clouds and more), which respectively form the second and third topics in our discussion. The final topic discusses a relatively young area where progress relies on fusing the previous three topics – namely, the development of applications and serious games: from docking to model building.

## Virtual and augmented reality and immersive molecular simulations

Setting up, running, and analyzing a molecular simulation is generally a lengthy process that requires considerable user patience and expertise. In principle, interactive simulations provide a much lower latency approach for manipulating and exploring molecular structures.<sup>1</sup> Such methods provide a virtual or augmented reality framework for immersing the user within the simulation, with the aim that molecular interaction becomes as intuitive as possible.<sup>2</sup> One of the earliest and perhaps most obvious methods for facilitating a more immersive molecular interaction experience involves the sense of touch. Indeed, the use of physical models has a long history within both chemistry and biochemistry, perhaps most famously captured by Watson and Crick's physical model of DNA. With the advent of computing and robotics, touch is a sensory channel that continues to receive a great deal of attention, mostly through the use of haptic interaction strategies.<sup>3–5</sup> Touch is also exploited in systems that utilize tangible physical models<sup>6</sup> that may be augmented virtually<sup>7</sup> and which have shown promise in enhancing student learning.<sup>8</sup> Beyond touch, a variety of sensory channels related to visual and audio feedback may be used to enhance the immersive effect, and preliminary applications of such integrated methods have occurred in the context of docking problems.<sup>9</sup> The design and utilization of efficient and effective strategies for interaction with a molecular simulation relies on careful consideration of methods developed within the field of human–computer interaction.

In general, an interactive molecular simulation framework requires several different ingredients: (1) a mechanism for representing the molecular system – usually involving a screen-based visual display and/or a physical object; (2) a mechanism for interaction with the simulation – typically through dedicated peripherals like haptic devices, cameras, a mouse, *etc.*; (3) animation of the simulation using a physics-based simulation engine; and (4) a pipeline which allows low latency coupling of these elements – often utilizing some sort of network protocol. In contrast to many fields of scientific computation, latency is an important consideration for interactive systems, and important bottlenecks need to be ameliorated – *i.e.*, the simulations need to be fast enough to run at an interactive speed, and the visualization tools need to be sufficiently elaborate and flexible so that the user can tackle both very complex and heterogeneous data. Interactive simulations are presently supported in many software packages, including NAMD,<sup>10</sup> LAMMPS,<sup>11</sup> HOOMD,<sup>12</sup> ProtoMol<sup>13</sup> and OPEP/HiRE-RNA.<sup>14</sup> As far as the simulation engines are concerned, one of the principal concerns is that of performance: for a reasonable interactive experience, animation of the molecular system needs to occur with refresh rates of at least 24 Hz.

Parallelisation therefore plays an important role in interactive simulation, because it can give substantial increases in the performance of molecular dynamics simulation engines. Stream architectures like GP-GPUs (as described below) provide one way to exploit parallelisation and achieve performance increases.<sup>12,15</sup> Compared to standard single-core environments, data transfer tends to present a significant bottleneck in massively parallel environments, resulting in a number of challenges – *e.g.*, extracting atomic positions without degrading performance in a stream parallelized software package is not trivial. More generally, MD simulations can produce a large quantity of simulation data, and transfer of this data may generate bottlenecks for coupling with visualization and interaction modules. Interactive all-atom simulations have recently been reported for systems up to 2 million atoms.<sup>16</sup> For the exploration of large systems, coarse graining offers another well-balanced alternative to speed up the physical engine driving an immersive simulation experiment.<sup>17</sup>

There are a wide variety of potential applications for interactive molecular simulation frameworks, including structural modelling, conformational searching, and interpretation of the mechanisms that drive function in complex biological models. Drug design is another active area in which new interactive simulation strategies are under investigation (including haptics interaction, virtual reality, and 3D printing<sup>18</sup>), in part motivated by the very high cost of bringing new medicines to market. Interactive simulations have been used to facilitate forms of nano-manipulation and even to prototype and design nano-robots.<sup>19,20</sup> Interactive simulation frameworks already offer considerable potential for providing microscopic insight into experiments; however, an even closer linking with experiment is likely to emerge in the near future. For example, a multitude of biophysical techniques (SAXS, cryo-EM, FRET, mass spectrometry, *etc.*) nowadays routinely generate a huge amount of biological data, opening up the possibility of combining all the data to build increasingly accurate models.<sup>21–23</sup> It will soon be possible to extend immersive approaches to explore not only simulation data, but experimental information as well (*e.g.* from NMR spectroscopy<sup>24</sup>) and to subsequently build models from such data under direct human supervision.<sup>25</sup> Another application of interactive molecular simulation involves reconstituting molecular assemblies from cryo-EM data.<sup>26</sup> These sorts of applications open a whole new range of collaborative opportunities and questions.<sup>27–29</sup>

## Advanced visualization and visual analytics

One of the cornerstones of modern molecular simulation concerns the visual representation of the structure of a molecule and its properties. Visual representation is particularly important in guiding the manner in which scientists think about atomic and molecular structure, which is partly a result of the fact that our human perception requires some form of augmentation in order to ‘see’ this world.<sup>30</sup> Attempts to conceptualize and visualize molecules reach far back in the history of chemistry. In terms of three-dimensional molecular structures, a notable milestone along this path goes back to the early 1940s when Roger Hayward depicted the arrangement of atomic assemblies in collaboration with Linus Pauling in both a scientifically accurate and aesthetically pleasing way.<sup>31</sup> Since this era (and in particular recently), technical progress has considerably improved our ability to visualize the molecular world. Nowadays, molecular

graphics are ubiquitous and every scientist can display the structure of a biomolecule on his/her personal computer<sup>32,33</sup> or tablet device.<sup>34</sup> To ensure that the enormous quantity of information contained in molecular simulation data (interactive or not) furnishes maximum insight into microscopic phenomena, the investigation of new visualization and visual analysis methods is an area of active research. One of the primary focuses of this emerging area concerns the development of new ways to understand and rapidly process the radically expanding deluge of data which molecular simulations are capable of generating. Visualization assists in grasping the complexity of these data and identifying emerging properties.

It is now possible to interactively visualise very large molecular assemblies, and new developments (including the use of GPU programming) are driving performance gains and opening up new possibilities for visualization.<sup>35</sup> Nowadays, millions of atoms and their bonds can be depicted interactively,<sup>36–38</sup> with considerable speed-ups in secondary structure representation.<sup>39</sup> Calculating molecular surfaces on-the-fly is more demanding,<sup>40–43</sup> but realistic rendering including the effects of lighting and ambient occlusion<sup>44</sup> reaches real-time refresh rates.<sup>45</sup> Impressive progress has been achieved with interactive raytracing of molecular systems on the GPU (<http://www.molecular-visualization.com/#!/home/mainPage>). Using ray-casted instancing, even whole-cell simulation data may be visualized smoothly<sup>46</sup> on stereoscopic displays,<sup>47</sup> allowing the reconstruction of 3D cellular complexes built from proteins and DNA molecules.<sup>48</sup>

The widespread availability of molecular simulation packages means that molecular visualization must have grown accustomed to depicting time-dependent dynamics, calling for on-the-fly visualization of atomic and molecular interactions (*e.g.*, hydrogen bonds),<sup>37</sup> molecular properties such as helix bending,<sup>49</sup> or the dynamics of molecular paths and cavities.<sup>50–52</sup> Visual analytics (<http://www.visual-analytics.eu>) have great potential to aid in understanding the increasing number of simulation datasets, and have been applied in a few cases.<sup>53,54</sup> Simplifying large quantities of complex data like those produced by MD simulations may be achieved by appropriate abstractions. In this context, techniques from scientific and medical illustration are helpful and have found their way into the visualization of chemical structures.<sup>55–57,112</sup> A stimulating recent example is the continuous abstraction of a molecular illustration<sup>58</sup> to yield a continuum of molecular depictions. Another challenge that arises in particular for the visualization of molecular simulations concerns the depiction of molecular flexibility.<sup>59</sup> In fact, chemical reactions themselves are difficult to render for many visualization tools. More generally, the visualization of dynamic molecular interaction networks is a very active field of research,<sup>60–62</sup> but is beyond the scope of this short introduction.

The ubiquity of molecular images and associated visualization tools is in part a consequence of the fact that it has been benefited from other high-growth economic areas. For example, tools that are traditionally dedicated to areas such as the video game industry or the film industry may be used for molecular visualization<sup>63–65</sup> or animation.<sup>66–68</sup> The availability of such tools has enabled the use of molecular visualization in collaborative structural biology, for example using TV-based<sup>69</sup> or web-based<sup>70</sup> solutions. A similar cross-fertilization is observed for GPUs, originally used in the consumer graphics market and nowadays omnipresent in high performance computing and scientific visualization.

# Computing power revolution and new algorithms: GP-GPUs, clouds and more

The field of molecular simulation and visualization intrinsically depends on high-performance computing (HPC) to ensure the underlying calculations can be carried out in real time and on a broad range of hardware including commodity computers. In this context, general-purpose-GPUs,<sup>71</sup> cloud computing,<sup>72</sup> and many-core architectures<sup>73</sup> are finding their way into the molecular simulation community. Multi-core architectures are evolving quickly, with massive-parallelism and massive-threading available on machines like the 1.3 million thread Blue Waters supercomputer. Bespoke architecture development, like that available on the Anton machine, is similarly allowing researchers to push the boundaries of simulation<sup>74</sup> and new techniques based on cloud-based methods and ultrafast high-performance networking are just around the corner. These and likely future developments in HPC are making massively parallel computations viable, and are stimulating innovation across hardware, software, and hardware/software integration, much of which is aimed at tackling the main challenges of molecular simulation: the size of systems which can be simulated, the time-scales which it is possible to simulate, the ability to sample large regions of molecular phase space, and the rigor of the underlying physics within the models. Lane *et al.*<sup>75</sup> have touched on many of these aspects in the context of protein folding.

The advent of programmable GPUs<sup>76</sup> using high-level languages like C, in conjunction with the NVIDIA CUDA (Compute Unified Development Architecture) tools, OpenCL and other frameworks has been instrumental in porting software and developing new algorithms. The rise of GPUs offers another example of how advances in high-growth consumer markets (namely video gaming) has been exploited for the purposes of scientific simulation. As a result of the power of GPUs, and the fact that they are relatively inexpensive, much academic and commercial molecular dynamics (MD) software (*e.g.*, GROMACS, NAMD and AMBER)<sup>77,113</sup> has been GPU-accelerated. The enhancement in speed can vary, in part due to the specific algorithms used, and also as a result of the particular GPU hardware architecture on which the code is run, both of which lead to different scalability and execution time. Adding many-body terms to potentials used in classical simulations is a case where the computational cost has been mitigated through exploitation of new hardware, by the development of a shared-memory force-decomposition algorithm.<sup>78</sup> Calculations using ReaxFF, which is a reactive force field, have been accelerated using GPUs.<sup>79</sup> Aspects of reliability and reproducibility have been studied in the context of error-correcting code.<sup>80</sup>

For quantum chemistry, software adoption of GPUs has been slower than for MD simulations, but building on initial work<sup>81–83</sup> many electronic structure packages now have GPU-enabled codes, and there is significant interest in utilizing fast quantum chemical methods, for example, to investigate reaction mechanisms.<sup>84</sup> Recent work has investigated GPU acceleration in a range of contexts, for example: (1) double precision matrix multiply operations within legacy quantum chemistry codes;<sup>85</sup> (2) ONETEP, a linearly-scaling plane wave density functional theory (DFT) code;<sup>86</sup> (3) BigDFT, a hybrid DFT code based on Daubechies wavelets;<sup>87</sup> (4) VASP, GPU-accelerated electronic structure calculations;<sup>88</sup> (5) real-space DFT implementations within the Octopus code;<sup>89</sup> and (6)

semiempirical methods.<sup>90</sup> Very recently, Sisto *et al.* have outlined fragment-based quantum chemical methods which rely on both distributed and shared memory GPU parallelism to carry out very large excited state time-dependent DFT (TDDFT) calculations using the TeraChem software framework.<sup>91</sup>

Cloud computing<sup>92–95</sup> is a relatively recent approach to molecular simulation that builds on distributed computing approaches like FightAIDS@home, SETI@home, and Folding@home. Distinct from other high-performance and distributed paradigms, it provides large-scale compute infrastructure on demand. In many respects, cloud-based approaches are still in their infancy, but are attracting growing attention. For applications of molecular simulation and modelling, cloud computing can offer large-scale data and compute capability for a short 'burst' phase. Cloud computing provides another example wherein molecular simulation benefits from exploiting approaches which have applications in other sectors: for example, cloud-based computing has appeal to small and medium biotech start-ups where continuous in-house HPC facilities would be under-utilised. Embarrassingly parallel tasks, like the generation of combinatorial databases, virtual screening of millions of compounds, and the analysis of the huge genome datasets, are well suited to existing cloud provisions. A workflow system called AutoDockCloud<sup>96</sup> enables distributed screening on a cloud platform using the molecular docking program AutoDock. For applications with greater demands for inter-processor communications, scalability is a key issue. A plugin<sup>97</sup> for the popular VMD software<sup>98</sup> (a front-end for NAMD<sup>10</sup>) allows one to (1) create a cloud-compute cluster on Amazon EC2; (2) submit a parallel NAMD job; (3) transfer the results back for subsequent post-processing; and (4) shutdown and terminate the compute cluster on Amazon EC2. These and other case studies of molecular modelling using cloud computing have been reviewed by Ebejer *et al.*<sup>72</sup>

## Crowd-sourcing and serious games: from docking to protein folding

Molecular simulation, like many areas of computational science, involves a tradeoff between user control and automation. Users usually have a deeper understanding and context for the problem at hand, but limited speed and memory. Computational systems, on the other hand, excel in memory and speed, but are limited when it comes to understanding and context. Even with the tremendous advances in computation discussed in the previous section, it is likely the case that there will always be a limit to the size and accuracy of models that can be built for a particular system, and therefore some level of human understanding will always be required. It is therefore of fundamental interest to consider radically new approaches to molecular modelling – *i.e.*, utilizing paradigms that do not rely exclusively on ever-faster computational frameworks.

Very recently, there has been a great deal of interest directed at investigating whether human intuition and problem solving skills can be effectively mobilized (usually *via* the internet) as a new resource for solving research questions.<sup>99</sup> The interest in these solutions is such that participants may be stimulated by the prospect of being remunerated.<sup>100</sup> Success in this area requires that the research

approach or proposition is cast in a way that is sufficiently engaging, entertaining, or educational. Along these lines, the Defense Advanced Research Projects Agency (DARPA) recently developed a challenge to see how quickly it is possible to involve a large number of people to fulfil a particular task.<sup>101</sup> Such 'crowd-sourced' research approaches<sup>102</sup> have received increasing attention. For example, one particularly successful example is the Galaxy Zoo<sup>103</sup> project, which transforms a potentially mundane, but difficult computer vision task (classifying images of galaxies) into an attractive challenge. When it comes to solving scientific research problems, collective and intrinsic motivation can marshal large communities of volunteers. This requires a high-level of visibility, which social media and modern communication technologies can successfully facilitate. Once a volunteer community is established, strategies and structures must be in place to maintain the ongoing engagement of the community. In many cases, crowd-sourced scientific computing paradigms raise interesting questions related to data analysis and data integrity.

Crowd-sourced approaches to research can generate useful insight owing to user intuition as a solution to cope with complex data and unveil emerging properties: a striking example is the game Foldit.<sup>104,105</sup> This project presents protein folding as a sort of three-dimensional jigsaw puzzle, where players are invited to shake and wiggle the three-dimensional structure of proteins in order to find the most stable conformations. Since May 2008, when the first beta version of this game was released, the project has gathered a large community. In some cases, Foldit players have been able to find optimal structures that automated search strategies failed to sample. Players do not necessarily require significant knowledge of biology to play the game and to find stable protein configurations. It is more a matter of spatial representation in three dimensions, as well as collaboration between players. The first 'levels' of the Foldit game are designed to train the players in order to accomplish increasingly complicated tasks. Interactions among players have led to remarkable results from a biological point of view<sup>106–108</sup> and also led players to collaboratively develop new algorithms to solve a particular problem.<sup>109</sup>

Not only do interactive and video game interfaces offer the potential for crowd-sourced research studies, they also offer an engaging medium for scientific education, helping students of all ages learn scientific principles and knowledge. As a consequence, educational games are flourishing. For example, the Spectral Game<sup>110</sup> seeks to teach quite advanced concepts in spectroscopy, specifically proton nuclear magnetic resonance (NMR). In addition to meeting specific targets, educational games and interactive molecular simulation platforms (like the distributed computing projects discussed above<sup>111</sup>) have a more general effect – *i.e.*, they engage the public and thereby increase public awareness and understanding of scientific problems. New channels for engaging the public with scientific ideas are also emerging in less traditional venues – *i.e.*, on the frontiers of aesthetic imagination and scientific visualization. As art moves increasingly toward digital mediums, artists have become fascinated with the glimpse into the invisible atomic world provided by molecular simulations and visualizations, to the extent that it has inspired new forms of artistic expression and aesthetic content.<sup>30</sup>

## Conclusion

Molecular simulation and visualization represent a vibrant melting pot of many scientific disciplines that both benefits from and drives significant progress across a range of fields. New hardware architectures, new software algorithms, and new technological developments inspire this evolution and herald an exciting era of increasingly sophisticated and perhaps unconventional molecular simulations. The potential for these new simulation frameworks is extremely exciting: they will allow us to obtain unprecedented new research insights, develop new ways for interacting with and imagining the microscopic world, drive progress in HCI and computer science, and ultimately have profound effects beyond the scientific realm within the broader culture.

## Acknowledgements

The authors thank Drs Matthieu Chavent, Antoine Taly and Alex Tek for stimulating discussions and exchanges that have nourished parts of this manuscript. M.B. acknowledges support by the “Initiative d’Excellence” program from the French State (Grant “DYNAMO”, ANR-11-LABX-0011). J.D.H. thanks EPSRC for support under grant EP/I006559/1. D.R.G. acknowledges support as a Royal Society Research Fellow.

## References

- 1 A. Tek, B. Laurent, M. Piuze, Z. Lu, M. Baaden, O. Delalande, M. Chavent, N. Ferey, C. Martin, L. Piccinalli, B. Katz, P. Bourdot, and L. Autin, *Advances in Human-Protein Interaction – Interactive And Immersive Molecular Simulations*, in *Protein Interactions – Computational and Experimental Tools/Book 2*, ed. W. Cai and H. Hong, Intech, Croatia, 2012.
- 2 J. E. Stone, A. Kohlmeyer, K. L. Vandivort and K. Schulten, Immersive molecular visualization and interactive modeling with commodity hardware, *Lect. Notes Comput. Sci.*, 2010, **6454**, 382–393.
- 3 A. M. Wollacott and K. M. Merz Jr, Haptic applications for molecular structure manipulation, *J. Mol. Graphics Modell.*, 2007, **25**, 801–805.
- 4 N. Zonta, I. J. Grimstead, N. J. Avis and A. Brancale, Accessible haptic technology for drug design applications, *J. Mol. Model.*, 2009, **15**, 193–196.
- 5 A. Ricci, A. Anthopoulos, A. Massarotti, I. Grimstead and A. Brancale, Haptic-driven applications to molecular modeling: state-of-the-art and perspectives, *Future Med. Chem.*, 2012, **4**, 1219–1228.
- 6 M. Francl, Tangible assets, *Nat. Chem.*, 2013, **5**, 147–148.
- 7 A. Gillet, M. Sanner, D. Stoffler and A. Olson, Tangible Interfaces for Structural Molecular Biology, *Structure*, 2005, **13**, 483–491.
- 8 G. E. Höst, C. Larsson, A. Olson and L. A. Tibell, Student learning about biomolecular self-assembly using two different external representations, *CBE Life Sci. Educ.*, 2013, **12**, 471–482.
- 9 N. Ferey, J. Nelson, C. Martin, L. Piccinalli, G. Bouyer, A. Tek, P. Bourdot, J. M. Burkhardt, B. F. G. Katz, M. Ammi, C. Etchebest and L. Autin, Multisensory VR interaction for protein-docking in the CoRSAIRE project, *Virtual Reality*, 2009, **13**, 273–293.

- 10 J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé and K. Schulten, Scalable molecular dynamics with NAMD, *J. Comput. Chem.*, 2005, **26**, 1781–1802.
- 11 S. Plimpton, Fast Parallel Algorithms for Short-Range Molecular Dynamics, *J. Comput. Phys.*, 1995, **117**, 1–19.
- 12 J. A. Anderson, C. D. Lorenz and A. Travesset, General purpose molecular dynamics simulations fully implemented on graphics processing units, *J. Comput. Phys.*, 2008, **227**, 5342–5359.
- 13 T. Matthey, and J. A. Izaguirre, ProtoMol: A Molecular Dynamics Framework with Incremental Parallelization, in *Proc Tenth SIAM Conf Parallel Processing for Scientific Computing (PP01), Proceedings in Applied Mathematics*, 2001.
- 14 F. Sterpone, S. Melchionna, P. Tuffery, S. Pasquali, N. Mousseau, T. Cragolini, Y. Chebaro, J. F. Saint-Pierre, M. Kalimeri, A. Barducci, Y. Laurin, A. Tek, M. Baaden, P. Hoang Nguyen and P. Derreumaux, The OPEP protein model: from single molecules, amyloid formation, crowding and hydrodynamics to DNA/RNA systems, *Chem. Soc. Rev.*, 2014, **43**, 4871–4893.
- 15 J. E. Stone, D. J. Hardy, I. S. Ufimtsev and K. Schulten, GPU-accelerated molecular modeling coming of age, *J. Mol. Graphics Modell.*, 2010, **29**, 116–125.
- 16 M. Dreher, M. Piuze, A. Turki, M. Chavent, M. Baaden, N. Ferey, S. Limet, B. Raffin and S. Robert, Interactive Molecular Dynamics: Scaling up to Large Systems, *Procedia Comput. Sci.*, 2013, **18**, 20–29.
- 17 O. Delalande, N. Ferey, G. Grasseau and M. Baaden, Complex Molecular Assemblies at hand via Interactive Simulations, *J. Comput. Chem.*, 2009, **30**, 2375–2387.
- 18 G. A. Dalkas, D. Vlachakis, D. Tsagkrasoulis, A. Kastania and S. Kossida, State-of-the-art technology in modern computer-aided drug design, *Briefings Bioinf.*, 2013, **14**, 745–752.
- 19 M. Hamdi, A. Ferreira, G. Sharma and C. Mavroidis, Prototyping bio-nanorobots using molecular dynamics simulation and virtual reality, *Microelectron. J.*, 2008, **39**, 190–201.
- 20 C. T. Chen, S. Y. Chen, C. H. Liao and S. C. Zeng, An interactive nanomanipulation visualization based on molecular dynamics simulation and virtual reality, *Trans. Can. Soc. Mech. Eng.*, 2013, **37**, 991–1000.
- 21 T. Schwede, Protein Modeling: What Happened to the “Protein Structure Gap”?, *Structure*, 2013, **21**, 1531–1540.
- 22 E. Karaca and A. M. J. J. Bonvin, Advances in integrative modeling of biomolecular complexes, *Methods*, 2013, **59**, 372–381.
- 23 D. Schneidman-Duhovny, R. Pellarin and A. Sali, Uncertainty in Integrative Structural Modeling, *Curr. Opin. Struct. Biol.*, 2014, **28**, 96–104.
- 24 J. N. Block, D. J. Zielinski, V. B. Chen, I. W. Davis, E. C. Vinson, R. Brady, J. S. Richardson and D. C. Richardson, KinImmerse: Macromolecular VR for NMR ensembles, *Source Code Biol. Med.*, 2009, **4**, 3.
- 25 G. Da Costa, A. Bondon, O. Delalande, L. Mouret and J. P. Monti, Elucidation by NMR solution of neurotensin in small unilamellar vesicle environment: molecular surveys for neurotensin receptor recognition, *J. Biomol. Struct. Dyn.*, 2013, **31**, 809–817.

- 26 J. Heyd and S. Birmanns, Immersive structural biology: a new approach to hybrid modeling of macromolecular assemblies, *Virtual Reality*, 2009, **13**, 245–255.
- 27 J. Simard, M. Ammi and M. Auvray, Collaborative strategies for the search of 3D targets in molecular environments, *IEEE Trans. Syst. Man Cybern. C Appl. Rev.*, 2012, **42**, 1555–1565.
- 28 J. Simard and M. Ammi, Haptic communication tools for collaborative deformation of molecules, in *Proc. of EuroHaptics*, Tampere, Finland, 2012, pp. 517–527.
- 29 P. Romano, R. Giugno and A. Pulvirenti, Tools and collaborative environments for bioinformatics research, *Bioinformatics*, 2011, **12**, 549–561.
- 30 ZKM Center for Art and Media Karlsruhe, *Molecular Aesthetics*, ed. Peter Weibel and Ljiljana Fruk, The MIT Press, Cambridge, MA, U.S.A./London, England, Germany, 2013, p. 400.
- 31 L. C. Pauling and R. Hayward, *Architecture of Molecules*, W.H. Freeman & Co Ltd, 1970.
- 32 S. I. O'Donoghue, D. S. Goodsell, A. S. Frangakis, F. Jossinet, R. A. Laskowski, M. Nilges, H. R. Saibil, A. Schafferhans, R. C. Wade, E. Westhof and A. J. Olson, Visualization of macromolecular structures, *Nat. Methods*, 2010, **7**, S42–S55.
- 33 C. Mura, C. M. McCrimmon, J. Vertrees and M. R. Sawaya, An Introduction to Biomolecular Graphics, *PLoS Comput. Biol.*, 2010, **6**, e1000918.
- 34 E. Krieger and G. Vriend, YASARA View—molecular graphics for all devices—from smartphones to workstations, *Bioinformatics*, 2014, DOI: 10.1093/bioinformatics/btu426.
- 35 M. Chavent, B. Levy, M. Krone, K. Bidmon, J. P. Nomine, T. Ertl and M. Baaden, GPU-powered tools boost molecular visualization, *Briefings Bioinf.*, 2011, **12**, 689–701.
- 36 S. Grottel, G. Reina, C. Dachsbacher and T. Ertl, Coherent Culling and Shading for Large Molecular Dynamics Visualization, *Comput. Graph. Forum*, 2010, **29**, 953–962.
- 37 M. Chavent, A. Vanel, A. Tek, B. Levy, S. Robert, B. Raffin and M. Baaden, GPU-accelerated atom and dynamic bond visualization using hyperballs: A unified algorithm for balls, sticks, and hyperboloids, *J. Comput. Chem.*, 2011, **32**, 2924–2935.
- 38 N. Lindow, D. Baum and H. C. Hege, Interactive Rendering of Materials and Biological Structures on Atomic and Nanoscopic Scale, *Comput. Graph. Forum*, 2012, **31**, 1325–1334.
- 39 M. Wahle and S. Birmanns, GPU-Accelerated Visualization of Protein Dynamics in Ribbon Mode, in *Proceedings of SPIE, Visualization And Data Analysis 2011*, ed. P. C. Wong, J. Park, M. C. Hao, C. Chen, K. Borner, D. L. Kao and J. C. Roberts, 2011, vol. 7868.
- 40 M. Chavent, B. Levy and B. Maigret, MetaMol: High-quality visualization of molecular skin surface, *J. Mol. Graphics Modell.*, 2008, **27**, 209–216.
- 41 M. Krone, K. Bidmon and T. Ertl, Interactive Visualization of Molecular Surface Dynamics, *IEEE Trans. Visual. Comput. Graph.*, 2009, **15**, 1391–1398.
- 42 M. Krone, J. E. Stone, T. Ertl and K. Schulten, Fast Visualization of Gaussian Density Surfaces for Molecular Dynamics and Particle System Trajectories, in *EuroVis 2012 Short Papers*, 2012.

- 43 J. Parulek and I. Viola, Implicit Representation of Molecular Surfaces, in *Proceedings of the IEEE Pacific Visualization Symposium (PacificVis 2012)*, 2012, pp. 217–224.
- 44 M. Tarini, P. Cignoni and C. Montani, Ambient occlusion and edge cueing to enhance real time molecular visualization, *IEEE Trans. Visual. Comput. Graph.*, 2006, **12**, 1237–1244.
- 45 S. Grottel, M. Krone, K. Scharnowski and T. Ertl, Object-Space Ambient Occlusion for Molecular Dynamics, *IEEE Pacific Visualization Symposium 2012*, ed. H. Hauser, S. Kobourov and H. Qu, 2012, pp. 209–216.
- 46 M. Falk, M. Krone and T. Ertl, Atomistic Visualization of Mesoscopic Whole-Cell Simulations Using Ray-Casted Instancing, *Comput. Graph. Forum*, 2013, **32**, 195–206.
- 47 B. Sommer, C. Bender, T. Hoppe, C. Gamroth and L. Jelonek, Stereoscopic cell visualization: from mesoscopic to molecular scale, *J. Electron. Imag.*, 2014, **23**, 011007.
- 48 S. Hornus, B. Levy, D. Lariviere and E. Fourmentin, Easy DNA Modeling and More with GraphiteLifeExplorer, *PLoS One*, 2013, **8**, e53609.
- 49 A. C. E. Dahl, M. Chavent and M. S. P. Sansom, Bendix: intuitive helix geometry analysis and abstraction, *Bioinformatics*, 2012, **28**, 2193–2194.
- 50 N. Lindow, D. Baum, A. N. Bondar and H. C. Hege, Exploring cavity dynamics in biomolecular systems, *BMC Bioinf.*, 2013, **14**(Suppl 19):S5.
- 51 J. Parulek, C. Turkay, N. Reuter and I. Viola, Visual cavity analysis in molecular simulations, *BMC Bioinf.*, 2013, **14**, S4.
- 52 M. Krone, G. Reina, C. Schulz, T. Kulschewski, J. Pleiss and T. Ertl, Interactive Extraction and Tracking of Biomolecular Surface Features, *Comput. Graph. Forum*, 2013, **32**, 331–340.
- 53 D. Bromley and V. Daggett, Analyzing disease-associated protein structures with visual analytics, *AMIA Joint Summits Translat Sci Proc*, 2013, vol. 2013, p. 33.
- 54 S. Thakur and M. A. Pasquinelli, Adapting Visual-Analytical Tools for the Exploration of Structural and Dynamical Features of Polymer Conformations, *Macromol. Theory Simul.*, 2011, **20**, 286–298.
- 55 J. Weber, ProteinShader: illustrative rendering of macromolecules, *BMC Struct. Biol.*, 2009, **9**, 19.
- 56 S. Bruckner and M. E. Gröller, Style transfer functions for illustrative volume rendering, *Comput. Graph. Forum*, 2007, 715–724.
- 57 G. Cipriano, G. Wesenberg, T. Grim, G. N. Phillips and M. Gleicher, GRAPE: GRaphical Abstracted Protein Explorer, *Nucleic Acids Res.*, 2010, **38**, W595–W601.
- 58 M. van der Zwan, W. Lueks, H. Bekker and T. Isenberg, Illustrative Molecular Visualization with Continuous Abstraction, *Comput. Graph. Forum*, 2011, **30**, 683–690.
- 59 A. Bryden, G. N. Phillips and M. Gleicher, Automated Illustration of Molecular Flexibility, *IEEE Trans. Visual. Comput. Graph.*, 2012, **18**, 132–145.
- 60 D. C. Y. Fung, S. S. Li, A. Goel, S. H. Hong and M. R. Wilkins, Visualization of the interactome: What are we looking at?, *Proteomics*, 2012, **12**, 1669–1686.
- 61 T. Praneenararat, T. Takagi and W. Iwasaki, Integration of interactive, multi-scale network navigation approach with Cytoscape for functional genomics in the big data era, *BMC Genomics*, 2012, **13**, S24.

- 62 G. Agapito, P. H. Guzzi and M. Cannataro, Visualization of protein interaction networks: problems and solutions, *BMC Bioinf.*, 2013, **14**, S1.
- 63 G. T. Johnson, L. Autin, D. S. Goodsell, M. F. Sanner and A. J. Olson, ePMV Embeds Molecular Modeling into Professional Animation Software Environments, *Structure*, 2011, **19**, 293–303.
- 64 R. M. Andrei, M. Callieri, M. F. Zini, T. Loni, G. Maraziti, M. C. Pan and M. Zoppe, Intuitive representation of surface properties of biomolecules using BioBlender, *BMC Bioinf.*, 2012, **13**, S16.
- 65 Z. Lv, A. Tek, F. Da Silva, C. Empereur-mot, M. Chavent and M. Baaden, Game on, Science – how video game technology may help biologists tackle visualization challenges, *PLoS One*, 2013, **8**, e57990.
- 66 G. McGill, Molecular movies... Coming to a lecture near you, *Cell*, 2008, **133**, 1127–1132.
- 67 S. Bromberg, W. Chiu and T. E. Ferrin, Workshop on Molecular Animation, *Structure*, 2010, **18**, 1261–1265.
- 68 J. H. Iwasa, Animating the model figure, *Trends Cell Biol.*, 2010, **20**, 699–704.
- 69 R. Yennamalli, R. Arangarasan, A. Bryden, M. Gleicher and G. N. Phillips, Using a commodity high-definition television for collaborative structural biology, *J. Appl. Crystallogr.*, 2014, **47**, 1153–1157.
- 70 C. Marion, J. Poudroux, J. Jomier, S. Jourdain, M. Hanwell and U. Ayachit, A Hybrid Visualization System for Molecular Models, *WEB3D 2013: 18th International Conference On 3d Web Technology*, ed. S. N. Spencer, 2013, pp. 117–120.
- 71 J. A. Baker and J. D. Hirst, Molecular dynamics simulations using graphics processing units, *Mol. Inf.*, 2011, **30**, 498–504.
- 72 J. P. Ebejer, S. Fulle, G. M. Morris and P. W. Finn, The emerging role of cloud computing in molecular modelling, *J. Mol. Graphics Modell.*, 2013, **44**, 177–187.
- 73 R. M. Farber, Topical perspective on massive threading and parallelism, *J. Mol. Graphics Modell.*, 2011, **30**, 82–89.
- 74 D. E. Shaw, P. Maragakis, K. Lindorff-Larsen, S. Piana, R. O. Dror, M. P. Eastwood, J. A. Bank, J. M. Jumper, J. K. Salmon, Y. Shan and W. Wriggers, Atomic-level characterization of the structural dynamics of proteins, *Science*, 2010, **330**, 341–346.
- 75 T. J. Lane, D. Shukla, K. A. Beauchamp and V. S. Pande, To milliseconds and beyond: challenges in the simulation of protein folding, *Curr. Opin. Struct. Biol.*, 2013, **23**, 58–65.
- 76 J. E. Stone, D. J. Hardy, I. S. Ufimtsev and K. Schulten, GPU-accelerated molecular modeling coming of age, *J. Mol. Graphics Modell.*, 2010, **29**, 116–125.
- 77 R. Salomon-Ferrer, A. W. Gotz, D. Poole, S. Le Grand and R. C. Walker, Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald, *J. Chem. Theory Comput.*, 2013, **9**, 3878–3888.
- 78 W. M. Brown and M. Yamada, Implementing molecular dynamics on hybrid high performance computers-Three-body potentials, *Comput. Phys. Commun.*, 2013, **184**, 2785–2793.
- 79 M. Zheng, X. X. Li and L. Guo, Algorithms of GPU-enabled reactive force field (ReaxFF) molecular dynamics, *J. Mol. Graphics Modell.*, 2013, **41**, 1–11.

- 80 R. M. Betz, N. A. DeBardleben and R. C. Walker, An investigation of the effects of hard and soft errors on graphics processing unit-accelerated molecular dynamics simulations, *Concurrency Comput. Pract. Ex.*, 2014, **26**, 2134–2140.
- 81 I. Ufimtsev and T. Martínez, Graphical Processing Units for Quantum Chemistry, *Comput. Sci. Eng.*, 2008, **10**, 26–34.
- 82 P. Brown, C. Woods, S. McIntosh-Smith and F. R. Manby, Massively Multicore Parallelization of Kohn–Sham Theory, *J. Chem. Theory Comput.*, 2008, **4**, 1620–1626.
- 83 K. Yasuda, Accelerating Density Functional Calculations with Graphics Processing Unit, *J. Chem. Theory Comput.*, 2008, **4**, 1230–1236.
- 84 M. P. Haag and M. Reiher, Real-time quantum chemistry, *Int. J. Quantum Chem.*, 2013, **113**, 8–20.
- 85 S. S. Leang, A. P. Rendell and M. S. Gordon, Quantum Chemical Calculations Using Accelerators: Migrating Matrix Operations to the NVIDIA Kepler GPU and the Intel Xeon Phi, *J. Chem. Theory Comput.*, 2014, **10**, 908–912.
- 86 K. Wilkinson and C. K. Skylaris, Porting ONETEP to Graphical Processing Unit-Based Coprocessors. 1. FFT Box Operations, *J. Comput. Chem.*, 2013, **34**, 2446–2459.
- 87 L. Genovese, M. Ospici, T. Deutsch, J. F. Mehaut, A. Neelov and S. Goedecker, Density functional theory calculation on many-cores hybrid central processing unit-graphic processing unit architectures, *J. Chem. Phys.*, 2009, **131**, 034103.
- 88 M. Hacene, A. Anciaux-Sedrakian, X. Rozanska, D. Klahr, T. Guignon and P. Fleurat-Lessard, Accelerating VASP electronic structure calculations using graphic processing units, *J. Comput. Chem.*, 2012, **33**, 2581–2589.
- 89 X. Andrade and A. Aspuru-Guzik, Real-Space Density Functional Theory on Graphical Processing Units: Computational Approach and Comparison to Gaussian Basis Set Methods, *J. Chem. Theory Comput.*, 2013, **9**, 4360–4373.
- 90 X. Wu, A. Koslowski and W. Thiel, Semiempirical Quantum Chemical Calculations Accelerated on a Hybrid Multicore CPU–GPU Computing Platform, *J. Chem. Theory Comput.*, 2012, **8**, 2272–2281.
- 91 A. Sisto, D. R. Glowacki and T. J. Martinez, Ab Initio Nonadiabatic Dynamics of Multichromophore Complexes: A Scalable Graphical-Processing-Unit-Accelerated Exciton Framework, *Acc. Chem. Res.*, 2014, **47**, 2857–2866.
- 92 R. Buyya, C. S. Yeo, S. Venugopal, J. Broberg and I. Brandic, Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility, *Future Generat. Comput. Syst.*, 2009, **25**, 599–616.
- 93 S. Zhou, R. Liao and J. Guan, When cloud computing meets bioinformatics: a review, *J. Bioinf. Comput. Biol.*, 2013, **11**, 1330002.
- 94 O. Korb, P. W. Finn and G. Jones, The cloud and other new computational methods to improve molecular modelling, *Expert Opin. Drug Discovery*, 2014, **22**, 1–11.
- 95 P. M. Kasson, Computational biology in the cloud: methods and new insights from computing at scale, *Pac. Symp. Biocomput.*, 2013, 451.
- 96 S. R. Ellingson and J. Baudry, High-throughput virtual molecular docking with AutoDockCloud, *Concurrency Comput. Pract. Ex.*, 2014, **26**, 907–916.
- 97 A. K. L. Wong and A. M. Goscinski, A VMD Plugin for NAMD Simulations on Amazon EC2, *Procedia Computer Science*, 2012, **9**, 136–145.

- 98 W. Humphrey, A. Dalke and K. Schulten, VMD: visual molecular dynamics, *J. Mol. Graphics*, 1996, **14**, 33–38.
- 99 B. M. Good and A. I. Su, Crowdsourcing for bioinformatics, *Bioinformatics*, 2013, **29**, 1925–1933.
- 100 K. R. Lakhani, K. J. Boudreau, P.-R. Loh, L. Backstrom, C. Baldwin, E. Lonstein, M. Lydon, A. MacCormack, R. A. Arnaout and E. C. Guinan, Prize-based contests can provide solutions to computational biology problems, *Nat. Biotechnol.*, 2013, **31**, 108–111.
- 101 G. Pickard, W. Pan, I. Rahwan, M. Cebrian, R. Crane, A. Madan and A. Pentland, Time-Critical Social Mobilization, *Science*, 2011, **334**, 509–512.
- 102 J. Surowiecki, *The Wisdom of Crowds: Why the Many Are Smarter than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations*, Doubleday Books, New York, 2004, p. 336.
- 103 C. Lintott, K. Schawinski, S. Bamford, A. Slosar, K. Land, D. Thomas, E. Edmondson, K. Masters, R. C. Nichol, M. J. Raddick, A. Szalay, D. Andreescu, P. Murray and J. Vandenberg, Galaxy Zoo 1: data release of morphological classifications for nearly 900 000 galaxies, *Mon. Not. R. Astron. Soc.*, 2011, **410**, 166–178.
- 104 S. Cooper, F. Khatib, A. Treuille, J. Barbero, J. Lee, M. Beenen, A. Leaver-Fay, D. Baker, Z. Popovic and F. Players, Predicting protein structures with a multiplayer online game, *Nature*, 2010, **466**, 756–760.
- 105 B. M. Good and A. I. Su, Games with a scientific purpose, *Genome Biol.*, 2011, **12**, 135.
- 106 C. B. Eiben, J. B. Siegel, J. B. Bale, S. Cooper, F. Khatib, B. W. Shen, F. Players, B. L. Stoddard, Z. Popovic and D. Baker, Increased Diels-Alderase activity through backbone remodeling guided by foldit players, *Nat. Biotechnol.*, 2012, **30**, 190–192.
- 107 S. Cooper, F. Khatib, A. Treuille, J. Barbero, J. Lee, M. Beenen, A. Leaver-Fay, D. Baker, Z. Popovic and F. Players, Predicting protein structures with a multiplayer online game, *Nature*, 2010, **466**, 756–760.
- 108 F. Khatib, F. DiMaio, Foldit Contenders Group, *et al.*, Crystal structure of a monomeric retroviral protease solved by protein folding game players, *Nat. Struct. Mol. Biol.*, 2011, **18**, 1175–1177.
- 109 F. Khatib, S. Cooper, M. D. Tyka, *et al.*, Algorithm discovery by protein folding game players, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 18949–18953.
- 110 J. C. Bradley, R. J. Lancashire, A. S. I. D. Lang and A. J. Williams, The Spectral Game: leveraging Open Data and crowdsourcing for education, *J. Cheminf.*, 2009, **1**, 9.
- 111 J. Bohannon, Distributed computing. Grassroots supercomputing, *Science*, 2005, **308**, 810.
- 112 G. T. Johnson and S. Hertig, A guide to the visual analysis and communication of biomolecular structural data, *Nat. Rev. Mol. Cell Biol.*, 2014, **15**, 690–698.
- 113 P. Larsson, B. Hess and E. Lindahl, Algorithm improvements for molecular dynamics simulations, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2011, **1**, 93–108.