[6] Software Introduction: Methodological advances for interacting with biomolecules using haptics

School of Computing Sciences, University of East Anglia
Georgios Iakovou g.iakovou@uea.ac.uk
Stephen D. Laycock s.laycock@uea.ac.uk
Steven Hayward steven.hayward@uea.ac.uk

Abstract

Over the past 15 years we have been developing tools for interacting with biomolecules using haptics. Interactions with biomolecules in the virtual world are made via a haptic-feedback device that is able to resist inputs from the user or even act to move the user's hand in response to molecular forces. Here we highlight the key methodological advances made in the development of these tools including Haptimol ISAS, a tool for interacting with a molecule's solvent accessible surface, Haptimol ENM, a tool for applying forces to an elastic network model of a biomolecule, DockIT (formerly Haptimol RD), for interactive rigid docking, and Haptimol FlexiDock, for interactive docking that models flexibility in the receptor molecule.

Keywords : Docking, Haptic-feedback, Elastic network model, linear response

1. Introduction

A haptic device is a force-feedback device, that enhances interactivity with the virtual world by engaging a user's sense of touch, more precisely their kinesthetic sense, i.e. their force sensation. This adds another dimension to the user's experience. Aspuru-Guzik et al. [1] states, haptics brings "*a new level of intuition to the virtual experience of the molecular world that goes far beyond its archaic and fractured perception through computer mouse and keyboard.*" There have been many attempts to incorporate haptics into biomolecular visualization. Most have approached it from the perspective of docking, where the user manipulates a ligand molecule with the haptic device, which acts as a 3D mouse, and forces (sometimes torques also) on the ligand from the receptor molecule are scaled and felt by the user on the haptic device[2-12]. Initially we took a very different approach. Our first tool, Haptimol ISAS, was born out of the question of whether one can feel the surface of a biomolecule, more precisely the van der Waals (vdW) surface. This leads to the question: what should one feel the surface with? The natural answer seemed to be a water molecule, or equivalently an oxygen atom, if ignoring electrostatic interactions. By allowing the haptic device to control the position of the oxygen atom, which acts as a spherical probe, and by feeling the hard-sphere interactions of the probe with the atoms of the biomolecule, the user could feel the solvent accessible surface; thus the acronym, ISAS, for Interactive Solvent Accessible Surface. Challenges that arose in the development of ISAS involved the navigation of the biomolecule using the haptic device when it was also being used for positioning the probe, and determination of the path of the probe when moving over a surface of overlapping spheres.

Biomolecules are flexible and conformational change is an integral part of function. Modelling conformational change accurately is something that Molecular Dynamics (MD) simulation can do very well, but incorporating haptics into MD, as in interactive MD [13], is naturally problematic, not least due to the wildly fluctuating forces that will be transmitted to the haptic device due to the stochastic nature of MD trajectories. We have developed two ways to model flexibility, one avoids MD altogether, the other separates the interactive session from the MD

simulation. The tool, Haptimol ENM[14], puts an Elastic Network Model (ENM) [15] of a biomolecule in the virtual world. A force can be applied to individual atoms via the haptic device and the conformational response seen on the screen and felt on the haptic device. It puts an investigative tool in the hands of the user, through which they can gain an understanding of the mechanical properties of a biomolecule's ENM. Haptimol FlexiDock [16] models the conformational response of a biomolecule to interaction forces from a ligand by applying the method of linear response determined from an MD trajectory of a simulation performed previously.

For graphics, the frame update rate needs to be at least 30 frames per second (fps) for the viewer to perceive a smooth continuous animation. However, due to the acute sensitivity of the human kinesthetic sense, the update rate for haptics needs to be much higher; at least 500 fps. This means that when using a haptic device, force calculations and any conformational response must be evaluated within 2 ms for the user to feel a smooth force on the haptic device. To complete computation of the conformational response within 2 ms, Haptimol ENM and Haptimol FlexiDock both employ the concept of the "important subspace". The important subspace, defined in collective coordinate space, is a relatively small subspace within which a large proportion of the total fluctuation occurs and is a feature of protein dynamics [17-19].

Even if in docking we ignore molecular flexibility and dock molecules rigidly, the calculation of the intermolecular forces within the haptic time constraint presents a challenge. The brute-force approach of calculating all inter-atomic forces between two molecules on the CPU can only deal with molecules comprising a few hundred atoms [3,20]. A popular method to overcome this is to use pre-computed force grids [2,5,11,21]. However, such grids are memory hungry and can induce rough force transitions at cell boundaries [11]. More pertinent though, is that force-grids cannot be used if molecular flexibility is included as they must be re-computed every time there is a change in conformation. Although the CPU can only accommodate small molecules, we developed a method to calculate the force between very large molecules by exploiting parallelism on the GPU. This was implemented in DockIT (formerly Haptimol RD) [22-25] and Haptimol FlexiDock [16].

Below we detail the key methodological advances we have made including those in our latest VR version of DockIT and indicate the kind of results that can be found using these tools.

2. Methods

2.1 ISAS

The fundamental elements to Haptimol ISAS are the ability to touch the solvent accessible surface of a molecule and to navigate all around it via the use of a haptic feedback device. Two parts are required for this, firstly a haptic rendering algorithm to calculate forces to prevent the probe pushing into the surface and secondly a navigation method to handle rotations and translations of the biomolecule.

The molecule is represented in space-filling mode but for the haptic rendering algorithm the radius of the probe sphere is added to the radius of each atom in the molecule. This reduces the haptic rendering algorithm to a single point-probe approach. The method we use is a mapping of the constraint-based single point rendering technique developed for polygonal meshes [26] to spheres. The algorithm uses two points, one is used to position a sphere shown to the user constrained to the vdW surface, whilst the second, the haptic interface point, is permitted to penetrate inside. A spring force is calculated between these two points to simulate hard surface interactions between the probe and the atoms of the molecule.

To allow the user to explore larger biomolecules a *Navigation Cube* is developed. The user can touch any atoms inside the Navigation Cube, but when the user moves the probe outside of the cube to reach a section of the molecule "out of reach", the molecule will translate.

To allow for rotation the user can press a button on the haptic stylus and move to apply a rotation. The software enables a good sense of the 3D shape of the molecule to be obtained whilst exploring pockets and channels where water molecules might be able to penetrate.

2.2 Intermolecular force calculation

Within DockIT and Flexidock a key component is the GPU-accelerated calculation of the intermolecular forces between the ligand and the receptor. vdW interactions are modelled with the Lennard-Jones potential, and electrostatic interactions using point charges and Coulomb's law. The parameters for these interactions are loaded from a Gromacs [27] topology file [24]. In the calculation all atoms pairs could be used but for efficiency we only include atom pairs within a cutoff distance. The GPU force calculation approach works in five steps. An OpenCL work item is created for each atom in the larger of the structures, typically the receptor, and the atom is transformed to the local coordinates of the ligand. Using a regular grid, the atoms in the ligand that are within the cutoff distance (set to 8 Å) to the receptor atom are determined and all forces computed. The last steps then sum the forces to compute the total force to send to the haptic device. The approach can compute the interatomic forces within 2ms for molecules comprising of hundreds of thousands of atoms each.

2.3 Scaling the force

The interaction force between two molecules is the order of nano-Newtons (nN) which is obviously imperceptibly small for the human sense of touch. The haptic device we have predominantly used for our studies is the 3D Systems Touch device (formally known as the SensAble Phantom Omni; Figure 3 shows one being used) which gives a maximum force of 3 N, approximately 0.3 kg. So, the force range for the Touch device is 0-3N which is not large. A simple way to feel the force on the haptic device is simply to scale it by a constant factor, s:

$$\vec{f}_{haptic} = s\vec{f}_{atom} \quad (1)$$

For DockIT and FlexiDock, the default value for s is 1×10^9 so a force of 1 nN in the virtual world will be felt as 1 N force on the haptic device. This "fixed" factor, which is used in ENM, DockIT and FlexiDock, can be varied to increase or decrease the force that is being felt/applied. If the range of $|\vec{f}_{atom}|$ is small, for example when the ligand is far from the receptor and electrostatic interactions dominate, then force variations could be imperceptible on the haptic device. In order to overcome this, we implemented a linear scaling method where the range of force is specified by the user. This "min-max" scaling mode [24] scales the molecular force linearly between user-defined minimum and maximum forces to appear on the haptic device as a force between 0 N and 3 N (for Touch device), respectively. Molecular forces below or above the specified maximum and minimum appear as 0 N and 3 N on the haptic device, respectively. The min-max mode enables a small molecular force range to span the whole force range of the haptic device. We also implemented the non-linear "variable gain" scaling method devised by Bolopion et al. [28]. This is sensitive to small force changes when the molecular force is low and is rather insensitive to force changes when the molecular force is high. All three scaling methods have been implemented in DockIT and FlexiDock.

2.4 Calculation of conformational response

In order to model biomolecular flexibility in docking we have used the method of linear response which was first applied to proteins with considerable success by Ikeguchi et al. [29]. The approach we have taken for FlexiDock is to take the trajectory of an explicit solvent MD simulation of the ligand-free receptor molecule to determine the response of the receptor to perturbation forces from the ligand. The first step is to remove external movements from the receptor trajectory by fitting to a static structure. The second step is to calculate the average

structure and the final step is to evaluate the variance-covariance matrix, $\boldsymbol{A}$, by determining fluctuations from the average structure. Within the quasi-harmonic approach, the conformational response of the receptor in interaction with the ligand is given by:

$$\Delta \boldsymbol{r} = \frac{1}{k_B T} \boldsymbol{A} \boldsymbol{F} \quad (2)$$

where k_B is Boltzmann's constant, T the absolute temperature, $\boldsymbol{F}$, the $3N \times 1$ column vector of forces from the ligand that act on each of the N atoms of the receptor, and $\Delta \boldsymbol{r}$ is the $3N \times 1$ column vector of displacements of the receptor atoms. For Haptimol ENM, the equivalent to $\frac{1}{k_B T} \boldsymbol{A}$ was derived from a normal mode analysis (NMA) of the elastic network model [14]. A problem arises when using Equation (2) as it requires $9N^2$ multiplications which cannot be completed within the 2 ms constraint for haptic rendering for even modest size biomolecules. In order to overcome this, we diagonalize $\boldsymbol{A}$ to find its eigenvalues and eigenvectors as is done in quasi-harmonic analysis. We use the eigenvectors and eigenvalues to form an approximate expression for the displacements:

$$\Delta \boldsymbol{r} \approx \Delta \boldsymbol{r}_m = \frac{1}{K_B T} \boldsymbol{V}_m \boldsymbol{\Lambda}_m \boldsymbol{V}_m^t \boldsymbol{F} \quad (3)$$

where $\boldsymbol{V}_m$ is the $3N \times m$ matrix of eigenvectors, $\boldsymbol{V}_m^t$ is its transpose and $\boldsymbol{\Lambda}_m$ is the $m \times m$ diagonal matrix of eigenvalues sorted in descending order. The eigenvectors describe collective coordinates and the eigenvalues their corresponding mean square fluctuations. As stated in the Introduction section, a relatively small subspace, the "important subspace", of the collective coordinates – those that dominate the fluctuations – can account for a relatively large proportion of the total fluctuation. The number of multiplications in Equation (3) is $m(6N + 1)$ which means m can be adjusted so that the 2 ms constraint is achieved. The existence of the important subspace means that even for relatively small m the approximation in Equation (3) can be quite good and is also quantifiable. Equation (3) can also be used to give savings in memory. Further details are given in Matthews et al. [16].

2.5 DockIT and its VR version

DockIT provides a range of features important for interactive molecular docking including the force calculation already described and different graphical depictions like the molecular surface. Further to this we also include the rapid calculation and depiction of hydrogen bonds between receptor and ligand during the interactive docking session and enable the docking to be performed in VR using two Oculus Rift Touch Controllers to manipulate the position and orientation of the ligand and receptor. To enable the calculation of hydrogen bonds sufficiently quickly for real-time rendering rates in VR we developed a GPU-accelerated method that utilizes the same topology information used for the force calculation.

3. Results

3.1 ISAS

Although electrostatic interactions play an important role in the mechanism of water transport through aquaporins, the simplest way to allow or prevent passage of water through a channel is through steric interaction. Here we illustrate this with a plant aquaporin for which open (PDB: 2B5F) and closed (PDB: 1Z98) structures have been deposited in the Protein Data Bank (PDB) [30]. The default probe radius is 1.52 Å, the vdW radius of an oxygen atom, but it was found that a probe of this size could not pass through either structure. This is likely due to the structures being rigid. A simple method to model the effect of atomic fluctuations on the passage of water is to reduce the size of the oxygen atom. This was done gradually until the probe could pass through one of the channels. This was achieved for the open-channel structure for a probe radius of 1.03 Å but not for the closed-channel structure (see Figure 1). The main residue blocking the passage was found to be Leu197 which has

been identified as a key conserved residue for creating a barrier in the closed channel structure [30].

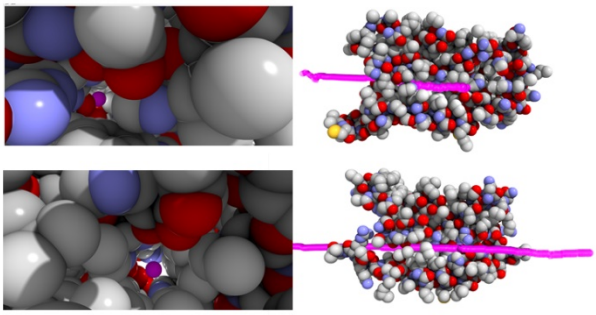

Figure 1. *Top row: Closed conformation of aquaporin (single subunit). Bottom row: Open conformation of aquaporin (single subunit). Left column: Looking down aquaporin channel with probe sphere (reduced size oxygen atom) in magenta. Right column: "Probecast", showing trail of probe through aquaporin.*

3.2 ENM

Aspartate transcarbamoylase (ATCase) is a complex enzyme that exhibits an allosteric mechanism. An NMA of the R-start structure has been performed previously [14] and is available from the Haptimol website to load into Haptimol ENM. Even though this is a relatively large protein all the modes could be used on a laptop equipped with an Intel(R) HD Graphics 520 card, i.e. the 2 ms constraint was satisfied. A force applied to cause movement of a regulatory dimer can produce the counter rotation of the catalytic trimers seen in the experimentally observed R to T transition. Figure 2 illustrates Haptimol ENM being used on ATCase.

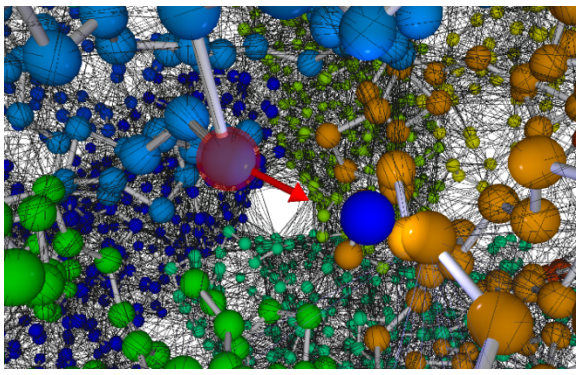

Figure 2. *Screenshot from Haptimol ENM showing ATCase. The dark blue sphere in the foreground is the "probe" used to select an atom (surrounded by a red halo) to which the force is applied. The force vector is represented by the red arrow. Spheres (C^{α} atoms) of the same colour are within the same subunit. Thin black lines indicate an elastic bond and grey cylinders "bonds" between consecutive C^{α} atoms.*

3.3 DockIT

DockIT [25] is a tool for the rigid docking of molecules. Figure 3 shows it being used to redock (separate a receptor and ligand in a solved structure and bring them back to their original binding conformation) the receptor-binding domain of the spike protein of SARS-CoV-2 to the antibody CC12.1 (PDB: 6XC2) [31]. Redocking showed that for some complexes there was a "sucking effect" on the haptic device as the ligand is drawn into the correct binding pose, whereas for others one cannot achieve the correct binding pose [32]. This led to the concept of locked interfaces. Access to the binding site in these cases can be achieved by "ghosting" (see but not feel) regions that overlap upon docking.

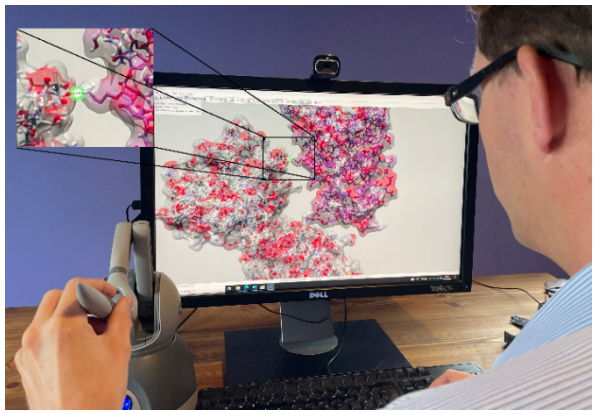

Figure 3. *Using DockIT to redock the receptor-binding domain of the spike protein SARS-CoV-2 with an antibody. Inset: enlarged image of an intermolecular hydrogen bond indicated by a green broken line with a green halo.*

3.4 FlexiDock

The reason one cannot access the true binding

conformation for those with “locked” interfaces is due to conformational change upon binding. It is obvious that there will be some degree of shape change upon binding another molecule and for some this is dramatic and has functional purpose. An example is maltose binding protein (MPB), which undergoes a 36° hinge bending movement on binding maltose. We applied FlexiDock to MBP. We first performed a 100 ns explicit solvent MD simulation on MBP alone and then performed quasi-harmonic analysis on the trajectory to evaluate the matrices in Equation (3). In order to get under the 2 ms constraint in the interactive session, only 3% of the total number of the eigenvectors (17,205), could be used but they accounted for 87% of the fluctuation that occurred in the MD simulation. Figure 4 shows the closure of the domains that occurs when maltose is maneuvered into the interdomain cleft. It was shown that this movement approximates the experimentally determined movement very well [16].

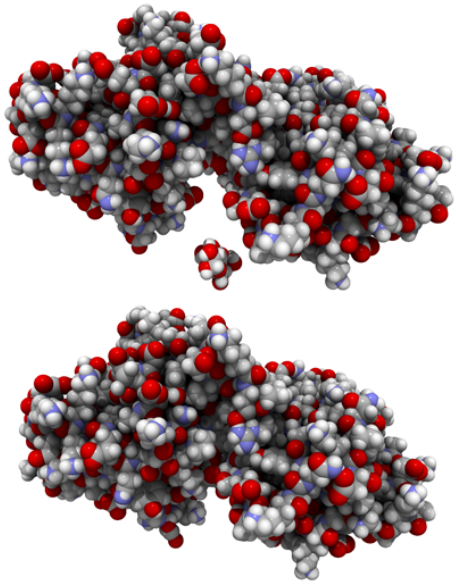

Figure 4*. Screenshots from Haptimol FlexiDock session of MBP and maltose. Top: Maltose approaching the open domain conformation of MBP. Bottom: MBP closed upon maltose after maltose was docked inside the interdomain region. A video of the process can be viewed at https://pubs.acs.org/doi/abs/10.1021/acs.jcim.9b00112.*

4. Conclusions

We have reviewed the key methodological advances that were required to develop our biomolecular haptics tools and shown how the resulting tools can be used to explore biomolecules and discover new things about them. Haptics draws the user deeper into the virtual world immersing them in an environment that nurtures ideas and fosters exploration. Our new VR version of DockIT deepens this feeling of immersion considerably.

Acknowledgements

We thank Matthew Stocks who helped develop Haptimol ISAS and ENM, and Nick Matthews who helped develop FlexiDock. G.I. acknowledges John Lusted, Dave Crooks, Gemma Bernal and the eSec/DIG team at Aviva Plc for their continuous support. We gratefully acknowledge support from NVIDIA with the donation of the Quadro P6000 GPU used for this research.

Availability

All the tools referred to can be downloaded from www.haptimol.co.uk.

References

[1] A. Aspuru-Guzik, R. Lindh, and M. Reiher, ACS Central Science **4**, 144 (2018).

[2] F. P. Brooks Jr, M. Ouh-Young, J. J. Batter, and P. J. Kilpatrick, in *Proceedings of the 17th annual conference on Computer graphics and interactive techniques*1990), pp. 177

[3] H. Nagata, H. Mizushima, and H. Tanaka, Bioinformatics **18**, 140 (2002).

[4] Y. G. Lee and K. W. Lyons, Computer-Aided Design **36**, 75 (2004).

[5] E. Subasi and C. Basdogan, Presence-Teleoperators and Virtual Environments **17**, 73 (2008).

[6] B. Daunay, A. Micaelli, S. Regnier, and IEEE, in *Proceedings of the 2007 Ieee International Conference on Robotics and Automation, Vols 1-10*2007), pp. 840.

[7] N. Zonta, I. J. Grimstead, N. J. Avis, and A. Brancale, Journal of Molecular Modeling **15**, 193 (2009).

[8] N. Ferey *et al.*, Virtual Reality **13**, 273 (2009).

[9] A. Anthopoulos, I. Grimstead, and A. Brancale, J.Comput.Chem. **34**, 2249 (2013).

[10] P. Bivall, S. Ainsworth, and L. A. E. Tibell, Science Education **95**, 700 (2011).

[11] A. M. Wollacott and K. M. Merz, J. Mol. Graph. Modell. **25**, 801 (2007).

[12] A. Bolopion, B. Cagneau, S. Redon, and S. Regnier, Journal of Molecular Graphics & Modelling **29**, 280 (2010).

[13] J. Stone, J. Gullingsrud, P. Grayson, and K. Schulten, in *ACM SIGGRAPH*2001), pp. 191.

[14] M. B. Stocks, S. D. Laycock, and S. Hayward, Journal of Computer-Aided Molecular Design **25**, 203 (2011).

[15] M. M. Tirion, Physical Review Letters **77**, 1905 (1996).

[16] N. Matthews, A. Kitao, S. Laycock, and S. Hayward, Journal of Chemical Information and Modeling **59**, 2900 (2019).

[17] N. Go, Biophysical Chemistry **35**, 105 (1990).

[18] S. Hayward and N. Go, Annual Review of Physical Chemistry **46**, 223 (1995).

[19] A. Kitao and N. Go, Current Opinion in Structural Biology **9**, 164 (1999).

[20] O. Sourina, J. Torres, and J. Wang, in *Transactions on Edutainment Ii*, edited by Z. Pan *et al.*2009), pp. 105.

[21] N. Pattabiraman, M. Levitt, T. E. Ferrin, and R. Langridge, J.Comput.Chem. **6**, 432 (1985).

[22] G. Iakovou, S. Hayward, and S. Laycock, Faraday Discussions **169**, 359 (2014).

[23] G. Iakovou, S. Hayward, and S. D. Laycock, Journal of Molecular Graphics & Modelling **61**, 1 (2015).

[24] G. Iakovou, S. Hayward, and S. D. Laycock, Journal of Chemical Information and Modeling **57**, 1142 (2017).

[25] G. Iakovou, M. Alhazzazi, S. Hayward, and S. Laycock, Bioinformatics **36**, 5698 (2020).

[26] C.-H. Ho, C. Basdogan, and M. A. Srinivasan, Presence: Teleoperators and Virtual Environments **8**, 477 (1999).

[27] D. Van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. C. Berendsen, J.Comput.Chem. **26**, 1701 (2005).

[28] A. Bolopion, B. Cagneau, S. Redon, and S. Régnier, in *2011 IEEE World Haptics Conference*2011), pp. 469.

[29] M. Ikeguchi, J. Ueno, M. Sato, and A. Kidera, Physical Review Letters **94** (2005).

[30] S. Törnroth-Horsefield, Y. Wang, K. Hedfalk, U. Johanson, M. Karlsson, E. Tajkhorshid, R. Neutze, and P. Kjellbom, Nature **439**, 688 (2006).

[31] M. Yuan *et al.*, Science (New York, N.Y.) **369**, 1119 (2020).

[32] G. Iakovou, S. Laycock, and S. Hayward, Biophysics and Physicobiology **13**, 97 (2016).