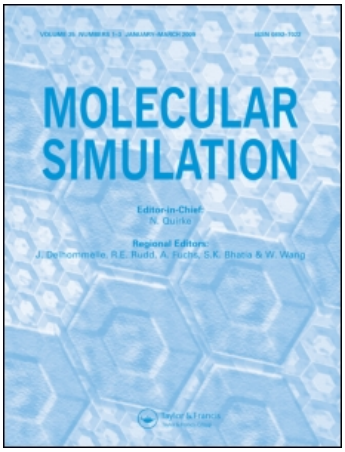

Molecular Simulation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

Simulations of biomolecule unbinding from protein using DL_POLY

P. -L. Chau[a]
[a] Bioinformatique Structurale, Institut Pasteur, CNRS URA 2185, Paris, France

To cite this Article Chau, P. -L.(2006) 'Simulations of biomolecule unbinding from protein using DL_POLY', Molecular Simulation, 32: 12, 953 — 961
To link to this Article: DOI: 10.1080/08927020600835640
URL: http://dx.doi.org/10.1080/08927020600835640

Molecular Simulation, Vol. 32, Nos. 12–13, 15 October–15 November 2006, 953–961

Taylor & Francis
Taylor & Francis Group

Simulations of biomolecule unbinding from protein using DL_POLY

P.-L. CHAU*

Bioinformatique Structurale, Institut Pasteur, CNRS URA 2185, 75724 Paris, France

(Received March 2006; in final form May 2006)

This paper briefly reviews the ideas leading to ligand–receptor interaction being a central topic of research in the biological sciences, especially in pharmacology. The simulation methods for studying ligand–receptor interaction dynamically through ligand unbinding are reviewed, together with the analysis methods devised to examine the unbinding trajectory. Examples of applying DL_POLY in these simulations are given; they include retinol unbinding from the retinol-binding protein, and of serotonin and granisetron unbinding from the 5-HT_3 receptor.

Keywords: Granisetron; Molecular recognition; Molecular dynamics; Retinol; Retinol-binding protein; Serotonin

1. Biomolecular interactions

Biomolecular interaction is the first step to many basic processes of life. Whether it be enzyme catalysis, neurotransmitter and hormone action or antibody–antigen recognition, these processes always involve the interaction of biological molecules with each other. The importance of molecular interaction has been gradually realised over history through pharmacological and chemical experiments.

Many pharmacology experiments have been performed since time immemorial, and many valuable discoveries have been made, and the results were compiled systematically in various publications. The first Chinese pharmacopeia is the *Shen Nong Ben Cao Jing* of Han Dynasty (206BC–220AD), which contains descriptions of over 300 substances of pharmacological importance [1]. Dioscorides Pedanius' *De materia medica* [2] is the first pharmacopeia to appear in Europe during the 1st century AD, and contains description of 600 plant products and 70 animal products. Nevertheless, no effort was made to construct a unifying theory that underlies the observations.

At the same time, advances in chemistry enabled pharmacologists to separate drugs into different parts. Scientists discovered that the effect of many drugs was due to an active principle. For example, Sertürner isolated morphine and meconic acid from opium in 1803, and demonstrated that the sedative property of opium was due exclusively to morphine (which he called "morphium") [3]; this was achieved by testing morphine on his friends.

Paul Ehrlich [4] formulated the *Seitenketten-Theorie* to account for the interaction between drugs and receptors: biological tissues contained side-chains, which bound toxins or dyes. The original function of the side-chains was not to bind toxins or dyes, but to carry out certain essential physiological functions. Different tissues had different side-chains, and since dyes had different affinities for these side-chains, the staining of vital dyes was specific. Toxins were harmful because they bound to the side-chains on the tissues, preventing the side-chains from carrying out their physiological functions. Antitoxins bound to the toxins to prevent the latter from coming into contact with side-chains on the tissues.

Ehrlich's conclusions spurred scientists on to uncover links between the chemical structure of drugs and their biological effects. Pharmacologists came to realise that a change in the chemical structure could bring about a change in the biological activity. The desired biological activity could be amplified and the unwanted activity diminished or abolished by appropriate chemical modifications of pre-existing ligands. A ligand is a small molecule that binds to the receptor site in question. It can be a drug, a natural hormone or neurotransmitter. Following Ehrlich's nomenclature, the molecule inside

*Corresponding author. Email: pc104@pasteur.fr

Molecular Simulation
ISSN 0892-7022 print/ISSN 1029-0435 online © 2006 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/08927020600835640

the body is called the receptor, and the introduced molecule is called the ligand. New drugs were discovered by systematic alteration of these ligands.

In the course of this work, it was realised that a better understanding of the detailed atomic structure of the receptor would facilitate our understanding. Since most biological receptors are proteins, this has to wait for the development of crystallography [5,6] and resolution of the structure of the proteins. The first protein thus studied, haemoglobin, also happens to bind a simple ligand, the oxygen molecule [7]. From then on, many soluble protein receptors have been crystallised, and their atomic structure resolved. Membrane protein receptors have to wait till the development of electron microscopy image processing techniques [8] before their structures were resolved [9–12]. In addition, nuclear magnetic resonance methods can also be used to solve the structures of proteins[13,14].

Crystallography, electron microscopy and nuclear magnetic resonance are invaluable in elucidating the structure of proteins. Nevertheless, they do not provide us with simultaneous information about the structure and dynamics of the ligand–receptor complex and the solvent molecules, crucial to understanding the binding process. Furthermore, it is difficult to isolate the factors that contribute to binding, and quantify their relative importances in each case. Simulations complement theory and experiments by making it possible to study each factor in-depth and to obtain structural and dynamical details simultaneously at picosecond resolutions.

It is known that water mediates the interaction between ligands and receptors. This could be done via steric effects [15,16], bridging hydrogen bonds [17] or entropic effects [18,19]. The effect of water is very important and it can radically change the interaction between ligands and receptors. Solvation and interaction of biological molecules is a complex many-bodied problem, and there exist no analytical methods for such investigations; only simulation methods, such as molecular dynamics (MD) and monte carlo (MC) methods, can be used.

2. Non-dissociate studies of biomolecular complexes

The earliest simulations on ligand–receptor complexes are all "non-dissociative" simulations. "Non-dissociative" means that no attempt has been made to dissociate the ligand from the receptor in these simulation studies.

Åqvist *et al.* [20] performed the first MD simulation on a holo-protein/apo-protein pair to study the effect of ligand binding on protein dynamics. They took the experimental coordinates of holo-serum retinol-binding protein, and used the GROMOS potential for a MD simulation of 70 ps. They did not have the experimental coordinates of the apo-protein, so extracted the protein coordinates from the holo-protein experimental coordinates, and used that as the initial point of an 80-ps simulation of the apo-protein. Both simulations were performed in vacuum at 298 K using an NVT ensemble, where the number of atoms, the volume of the simulation box and the temperature were held constant, and the first 30 ps was used for equilibration. In their holo-protein simulation, they discovered that the root-mean-square deviation of the simulation coordinates from the experimental coordinates is 3.0 Å for all atoms, and 2.2 Å for C_α-atoms only.

Tilton *et al.* studied the myoglobin–xenon complex [21] using simulations with the AMBER potential. In their 96-ps simulations, they established that the r.m.s. difference between the X-ray experimental structure and their simulated structure was 1.25 Å, and that the volume of the simulated structure oscillated at 3–4% around the X-ray volume. Of the five xenon atoms placed in myoglobin, three of them moved by less than 3.5 Å. The other two moved by about 10 Å; one escaped to the surface of the protein, and the other went into the hydrophobic core.

Dahl *et al.* created a homology model of the dopamine D_2 receptor, pulled dopamine into the binding site in 80 ps using docking MD, and performed a minimisation of the ligand inside the binding site [22]. Unfortunately the details of the docking procedure were not mentioned, and the authors only reported that dopamine was in close proximity to Asp 80 and Asn 390.

Non-dissociative studies aimed to validate the simulations with experimental data, by comparing the simulated structure with the crystallographic or NMR structure. There were some attempts to locate the interactions between the ligand and the receptor. Subsequently, as computer power improved, attempts were made to simulate the unbinding of ligands from receptors.

3. Methods for unbinding ligands from receptors

3.1 Previous work—steered molecular dynamics

Steered molecular dynamics (SMD) was developed to induce conformational changes in proteins, but was later used to simulate ligand unbinding from the receptor [23]. In this method, potential-well restraints were added to part of the protein; these restraints were external forces which behaved like a spring with one end fixed to the target position, and the other end attached to the atom being "tugged". In their seminal paper, Leech *et al.* discussed the uses of the method, and also pointed out that faster and stronger "tugs" could cause disordering effects on the system, and that the only way to verify that the pathway was physically correct was to do experiments. Grubmüller *et al.* applied this method to perform simulations of the unbinding of the streptavidin–biotin complex [24]. The authors obtained the unbinding trajectory from previous experiments, and were able to reproduce experimental data using simulation methods. They used a number of unbinding speeds, ranging from 40 to 0.15 m/s, and noted that the rupture force is velocity dependent, as expected,

up to 1.5 m/s. Thereafter the rupture force saturates, though they still observed a scatter in the value of the rupture force. They attributed this to the heterogeneity of unbinding pathways, arising from the microheterogeneity of protein conformational substates. They identified conformational changes of biotin during the unbinding process, and the breaking of hydrogen bonds between the ligand, the receptor and some surrounding water molecules.

The SMD method has been applied to the avidin–biotin complex [25] in simulations of length 40–500 ps. The authors applied a harmonic force constant close to those of atomic force microscopy (AFM) magnitudes to pull biotin out from avidin. They found that the calculate rupture forces were much larger than those observed in experiments. They concluded that, in these short simulations, the rupture forces required were much stronger than those used in millisecond AFM experiments, so the simulated unbinding cannot be extrapolated to experimental results.

Subsequent work focused on the unbinding of retinoic acid from the retionic acid receptor [26]. Since the authors did not know the pathway(s) of unbinding, they applied three unbinding pathways, using an unbinding speed of 0.032 Å/ps (3.2 m/s). They concluded that pathway 2 involved large changes in protein conformation, while pathways 1 and 3 disturbed protein conformation to a lesser extent. Using structural arguments, they suggested that pathway 1 was the binding pathway, while pathway 3 was the unbinding pathway. However, no attempt was made to verify this experimentally.

A combined simulation and crystallographic study was performed in the same year [27]. The authors created a D128A mutant of streptavidin, and resolved its structure with and without biotin bound to it. They also performed ten simulations to gradually pull biotin out of the streptavidin-binding pocket. A large number of simulations were required to sample the phase space adequately. In wild-type streptavidin, the authors observed that the carboxylate side-chain of Asp 128 made a hydrogen bond with biotin. In the course of the simulated unbinding, this hydrogen bond lengthened and was replaced by a water-mediated contact between Asp 128 and biotin. In the case of the D128A mutant, a water molecule was observed between biotin and Ala 128 even in the initial stages. Thermodynamic analysis showed that the D128A mutant behaved like an intermediate in the dissociation process. This work showed that, given enough dissociation pathways, SMD simulations could provide insight into the unbinding process.

Heymann and Grubmüller [28] performed a series of SMD simulations on a complex formed by dinitrophenyl (DNP) and the monoclonal antibody AN02 F_{ab} fragment. Some of these simulations had different starting conditions, while in others, the pulling speed were assigned values from 0.1 to 50 m/s. They identified two main events leading to DNP unbinding from AN02 F_{ab}: rupture of hydrogen bond between Tyr 33 of the light chain, and opening up of the tryptophane sandwich. They were able to identify two broad classes of unbinding pathways: in the "light chain route", the tryptophane sandwich opened first, while in the "heavy chain route", the hydrogen bond ruptured first.

Bayas *et al.* [29] applied SMD to unbind CD2 from CD58. They found that at high pulling speeds of 0.01 and 5×10^{-3} m/s, the two proteins unfolded before separation. At lower unbinding speeds of 10^{-3} and 5×10^{-4} m/s, the proteins separated before the domains could unfold. This work showed the importance of using a more realistic unbinding speed in SMD so as not to perturb the protein structure.

Whilst the SMD method has seen wide use in simulation studies of ligand–protein interactions, it does suffer from the drawback that the unbinding trajectory has to be pre-defined. When this is not known, scientists have to use a number of alternative trajectories to decide which ones are more probable. To overcome this problem, a new method do not require the prior definition of the unbinding pathway has recently been developed.

3.2 Mutual repulsion

In the method of mutual repulsion [30], the centres of mass of the ligand and the receptor are assigned what can be called "pseudo-charges", g, that increase linearly with time. A potential $\Psi(\mathbf{r})$ is defined:

$$\Psi(\mathbf{r}_i) = \frac{g^2}{|\mathbf{R}_1 - \mathbf{R}_2|} \tag{1}$$

where $\mathbf{r}_i$ is the position vector of the atom i, g is the magnitude of the pseudo-charge at that time, $\mathbf{R}_1$ is the position vector of the centre of mass of the receptor, and $\mathbf{R}_2$ is that of the centre of mass of the ligand.

The pseudo-charges interact with each other, but do not affect the normal electronic partial charges assigned to each atom. They repel or attract each other under rules similar to those for normal electronic partial charges. This method approaches the problem of unbinding as a rare event with large energy difference between the bound and the unbound states. The $\Psi(\mathbf{r}_i)$ potential artificially reduces this energy difference so that the transition from one state to another is facilitated.

There are two advantages of the method. Firstly, since the force due to the pseudo-charges is calculated with respect to the centre of mass, no torque is generated on the molecules. The molecules will be allowed to explore the unbinding path with fewer artificial forces. Secondly, since the pseudo-charge increases slowly, the potential $\Psi(\mathbf{r}_i)$ can be exploited for the calculation of the Helmholtz free energy ΔA by the adiabatic switching method, as explained later.

To demonstrate the validity of this method, I applied it to unbind retinol from the bovine serum retinol-binding protein. Atomic coordinates of the holo-sRBP [31] were obtained from the Protein Databank (PDB code: 1HBP);

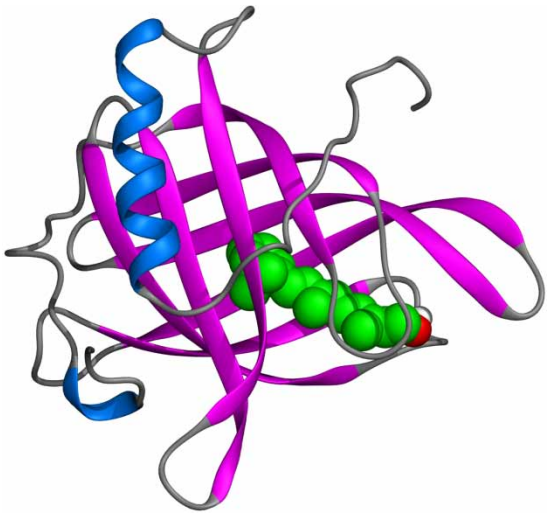

Figure 1. Diagram showing the structure of the bovine serum retinol-binding protein (1HBP), at the beginning of the unbinding simulation.

the structure of the protein is shown in figure 1. The model for the protein and retinol was parameterised using the GROMOS 87 potential [32]. With conservation of crystallographically determined water molecules, the protein was solvated in truncated octahedral boxes filled with 5537 SPC/E water [33] molecules using the water add utility in DL_POLY. Four sodium ions were added to ensure total electronic neutrality.

All minimisations and simulations were performed using version 2 of DL_POLY [34]. The solvated structure was minimised using the zero-K minimization option, and simulations were started with an NVT run at 30 K, taking initial velocities from a Maxwellian distribution. The temperature was gradually increased in steps to 310 K over a period of 100 ps, using a Nosé-Hoover thermostat [35,36]. SHAKE was used to constrain bond lengths with a tolerance of 10^{-8}. The time step was 1 fs. A cutoff of 10 Å was used, with simple truncation for Coulombic and van der Waals interactions. The simulations were then continued for 400 ps in an NPT ensemble using the Nosé-Hoover thermostat and the Melchionna barostat [37], all thermostats and barostats being options integral to DL_POLY. The last 200 ps of this NPT simulation was used to determine the optimal box size for a pressure of 1.01×10^5 Pa (1 atm). This box size was then used for a subsequent 100-ps NVT simulation to equilibrate the system, followed by a mutual repulsion run of 1 ns, again in the NVT ensemble. Configurational data were output every 20 fs during the mutual repulsion run.

Pseudo-charges were given dimensionless units. The force due to the pseudo-charges between the centres of mass was:

$$F = \frac{kg^2}{|\boldsymbol{R}_1 - \boldsymbol{R}_2|^2} \tag{2}$$

where F was the magnitude of the force between the two pseudo-charges ($k = 1.6605402 \times 10^{-33}$ kg m^3 s^{-2} and is a constant inserted to ensure that the force due to the pseudo-charges is adjusted to the internal units of DL_POLY, and that it is also of comparable magnitude to the other forces). The pseudo-charges were implemented by inserting the commands in the subroutine extnfld.f; extra forces were assigned to the centres of mass of the ligand and the receptor, respectively, and then equally distributed to all atoms.

The mutual repulsion was coupled with the method of adiabatic switching to give an estimate of the free energy change of unbinding. In the NVT ensemble, the NVE Hamiltonian of the physical system, H_0, is added to the thermostat part of the Nosé-Hoover thermostat, H', to give H_n, a "Nosé conserved quantity" or a "Nosé pseudo-Hamiltonian":

$$H_n = H_0 + H' \tag{3}$$

The partition function for this extended Nosé pseudo-Hamiltonian can be written as:

$$Z = \frac{1}{3N+1}\sqrt{\frac{2\pi Q}{kT}}\mathrm{e}^{E/kT}Z_c^o \tag{4}$$

where Q is a parameter that behaves like a mass in the Nosé-Hoover thermostat and acts as a heat source/sink, and Z_c^o is the partition function of the NVE ensemble for the physical system $\mathbf{H}_0$. Under adiabatic change of $\mathbf{H}_0$, an equivalent expression exists for the new Hamiltonian $\mathbf{H}_f$ and extended Nosé pseudo-Hamiltonian $\mathbf{H}_{n,f}$:

$$Z_f = \frac{1}{3N+1}\sqrt{\frac{2\pi Q}{kT}}\mathrm{e}^{E'/kT}Z_c^f \tag{5}$$

When the $\Psi(\mathbf{r}_i)$ potential is applied in addition to the thermostat, the gradually increasing mutual repulsion can be viewed as a switching function of the type described by Watanabe and Reinhardt [38]. These authors also showed that the Helmholtz free energy difference, ΔA, is equal to the total energy difference of the extended Nosé conserved quantity (the NVE Hamiltonian plus the thermostat part) under the adiabatic transformation:

$$Z_c^f = Z_c^o \mathrm{e}^{-\Delta A/kT} \tag{6}$$

It is imperative for the mutual repulsive force to increase very slowly, otherwise this method is no longer valid.

In the beginning of the simulation, the pseudo-charge was held at zero for 50 ps. At the end of the simulation, when retinol could not go any further without colliding with the other side of the protein through periodic boundary conditions, the pseudo-charge was held at 290.5, and the system equilibrated for 50 ps, followed by a data collection period of 50 ps. The final and initial 50-ps periods were used for evaluating the change in Helmholtz free energy when the ligand was unbound from the receptor. Let the initial kinetic energy, potential energy and pseudo-charge potential energy be, respectively, K_i, V_i and Ψ_i, and their final values be, respectively, K_f, V_f and Ψ_f. The kinetic energy and potential energy values are

derived from the Nosé conserved quantity. The Helmholtz free energy change ΔA is given by:

$$\Delta A = K_f + V_f + \Psi_f - K_i - V_i - \Psi_i. \qquad (7)$$

These quantities could all be read off the OUTPUT file of DL_POLY, so it was quite straight-forward to compute ΔA.

4. Results from retinol unbinding

Figure 2 shows retinol and retinol-binding protein at the end stage of the unbinding. It can be seen that retinol cleared the binding site at the end, and that the protein structure was largely unchanged. A quantitative analysis [30] showed that the r.m.s. deviation of the simulated structure from the crystal structure was always below 2.0 Å.

4.1 Free energy changes

Both my simulation and previous experiments showed that the free energy change was purely entropic. The experiment was performed in the NPT ensemble [39], and the entropy change $\Delta S = -364$ J/mol K. This value compares favourably with my NVT simulation result of $\Delta S = -580 \pm 30$ J/mol K.

4.2 Changes in binding site

I was able to examine the trajectory to define the movement of water and protein atoms in ligand–receptor interactions. I located the atoms of the binding site, and found that they formed a quasi-convex hull. Triangles were constructed from the polyhedron vertices using a convex hull programme [40]. The binding site volume was found to be about 5900 Å^3 in the initial 500 ps. It then decreased to below 5700 Å^3 at about 700 ps. During the last 300 ps, the volume increased to about 5800 Å^3 and then decreased to about 5700 Å^3 (figure 3). So there was a net decrease of about 200 Å^3.

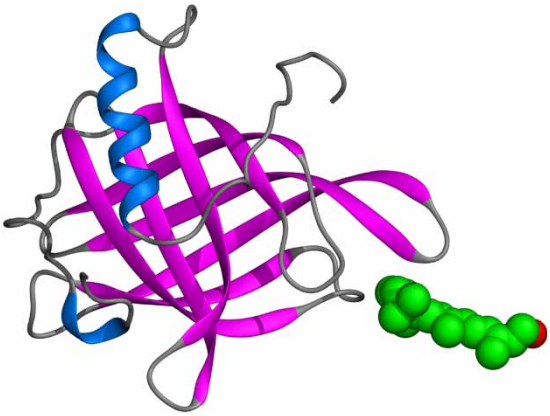

Figure 2. Diagram showing the structure of the bovine serum retinol-binding protein (1HBP), towards the end of the simulation.

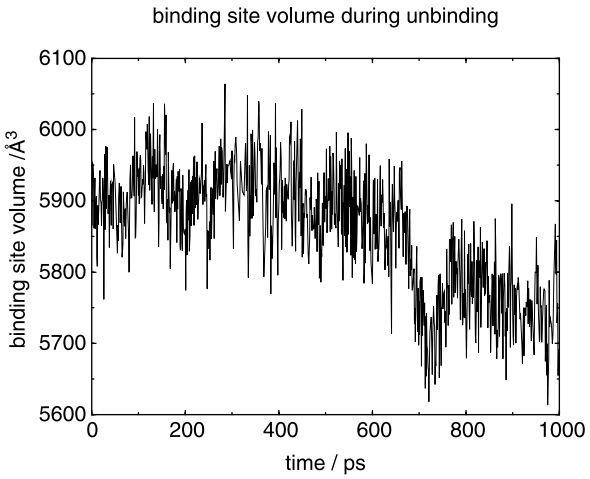

Figure 3. Diagram showing the volume of the binding site during the course of unbinding.

Using a previously developed method [41], I calculated the volume occupied by retinol, which was about 270 Å^3. Thus with the departure of retinol from the binding site, and a decrease in binding volume, there is a net increase of about 70 Å^3 in the space of the binding site. This is about the volume occupied by 2 or 3 water molecules. Would water from outside the binding site enter it, to occupy the space vacated?

4.3 Water movements

The number of water molecules inside the binding site during the unbinding simulation is shown in figure 4. During the first 200 ps, there were about 77 water molecules inside the binding site. Over time, the number of water molecules decreased to about 73. Thus, at least within the timescale of the simulation, there was net loss of water from the binding site, which is surprising [42]. How would the water molecules re-arrange themselves to accommodate the changes?

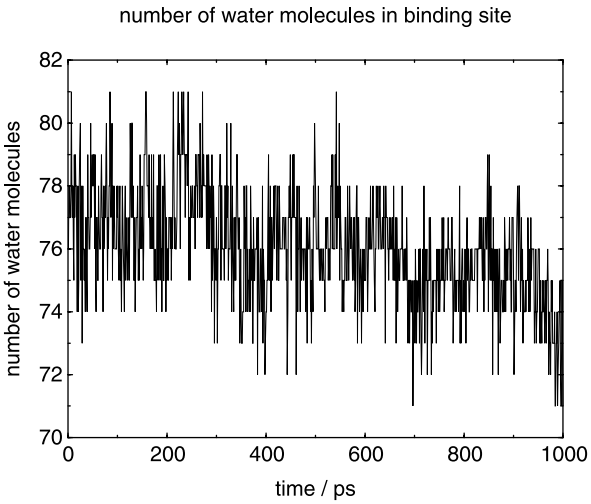

Figure 4. Diagram showing the number of water molecules inside the binding site during the course of unbinding.

A detailed analysis was carried out to examine the trajectory of each individual water molecule. I calculated the actual and net distances traversed by each molecule and its hydrogen-bonding pattern. The water molecules could be divided into different groups, according to their relative mobility. Some of the high-mobility water molecules were found to move as quickly as in the bulk phase, and they moved in to occupy the space vacated by the leaving retinol. The low-mobility water molecules, on the other hand, usually exhibited non-directional movement, and some of them were effectively exhibiting random motion [42].

4.4 Hydrophobic interaction

The hydration properties of retinol as it emerged from its binding site were examined [30]. The first hydrophobic hydration shell limit distance was 5.5 Å. When the ligand was bound, the number of water molecules within the first hydration shell of retinol was 44.7 ± 0.5, but this number increased to 76.2 ± 1.4 for the last 50 ps, when the ligand was unbound.

Moreover, I identified a group of water molecules that were in the bulk region when retinol was bound, but came to coat retinol when the ligand had unbound. I examined the properties of these water molecules, and showed that they were indeed in the bulk phase in the beginning of the simulation, but came to adopt a hydrophobic-hydration configuration with respect to retinol towards the end of the simulation [30].

5. Modified mutual repulsion

5.1 Methods

The mutual repulsion method was applied to unbind 5-hydroxytryptamine (5-HT) from a homology model of the 5-HT_{3A} receptor. The structure of the receptor was obtained from previous work [43], and its structure is shown in figure 5. Two different ligands, 5-HT and granisetron, whose chemical diagrams are shown in figure 6, are docked into one of the five binding sites of the receptor. 5-HT is an agonist of the receptor, and granisetron is an antagonist. Figure 7 shows one 5-HT molecule docked to the 5-HT_{3A} receptor.

The CHARMm22 potential was used throughout [44]. Two series of simulations were carried out. The first one consists of 5-HT docked into the binding site [45], and the other with granisetron docked into the same site [46]. All simulations were carried out in vacuum.

In each simulation, the structure of the protein–ligand complex was minimized for 20,000 steps using the zero-K minimisation method. All non-hydrogen atoms of the ligand, and all C_α atoms of the protein were tethered to their original positions, and the structure heated up to 310 K over 250 ps, followed by an equilibration period of 100 ps. Data production was carried out after

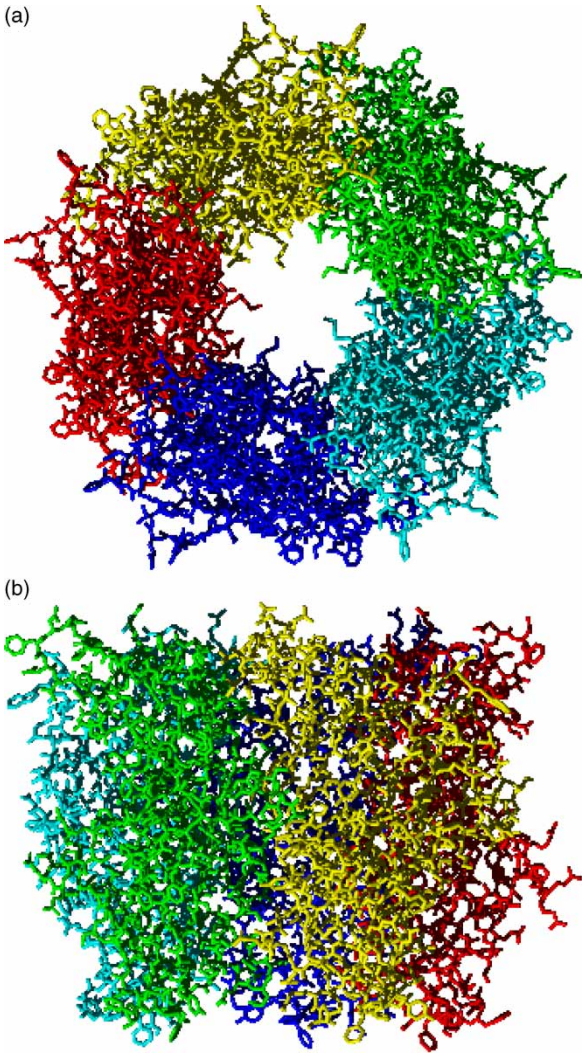

Figure 5. Diagrams showing the structure of the homology model of the extracellular domain of the 5-HT_{3A} receptor. (a) view from the extracellular space towards the membrane. The central pore leads to the ion channel; (b) side view of the extracellular domain. The membrane is at the bottom, and the extracellular space at the top. Each subunit is coloured differently, and the binding site is at the interface of two subunits. See online version for colour.

equilibration, with configuration dumping taking place every 1 ps. The time-step used was always 1 fs. A Nosé-Hoover thermostat was applied with a thermostat constant of 0.1 ps.

The unbinding required special care. Indeed, these simulations revealed a question with this unbinding method: mutual repulsion acted on the centre of mass of the two molecules, so if the unbinding trajectory was not around the line joining the two centres of mass, would the ligand leave the binding site? Initially, I implemented the unmodified mutual repulsion method, with the line of force acting almost perpendicular to the membrane-wards direction. This forced the ligand out without any problems; there were no serious distortions of the protein structure. However, there was good evidence that the unbinding pathway faces membrane-wards (negative z-direction). The proposed trajectory was almost perpendicular to

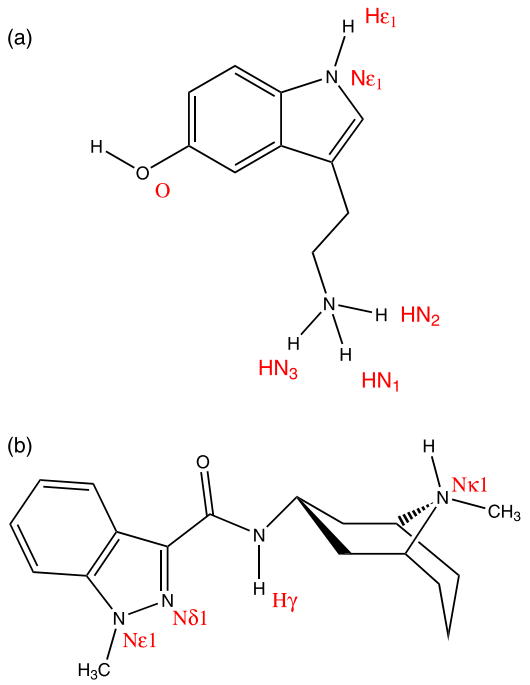

Figure 6. Chemical structures of (a) 5-HT and (b) granisetron. The hydrogen bonding groups are in red in online version.

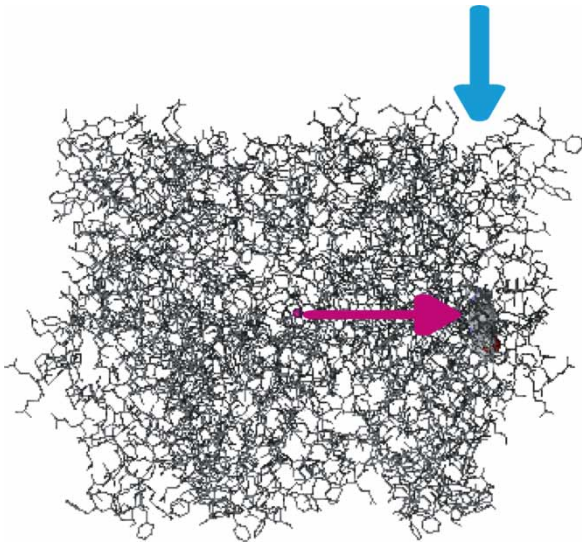

Figure 8. Diagram showing the relative disposition of the centre of the receptor (magenta circle) and the ligand (shown in space-filling model). The protein receptor is shown in grey wireframe. A mutually repulsive force would have forced the ligand out along the magenta arrow. However, there is evidence to show that the ligand most probably comes out membrane-wards, along the direction of the arrow. A dummy push-point was thus placed on the extracellular side of the protein to push the ligand out in the membrane-wards direction.

the line joining the centre of mass of the protein and the ligand (figure 8). It seemed more reasonable to force the ligand out along this direction. So instead of creating a force between the two centres of mass, I defined a dummy centre of mass for the receptor. This was placed in the positive z-direction of the ligand, so the mutual repulsive force would force the ligand out membrane wards (blue arrow of figure 8). This, unfortunately would create a torque on the receptor, so the C_α atoms of every five amino acids were tethered throughout the unbinding simulation.

In all simulations using the modified method, unbinding forces were increased to achieve an unbinding speed of about 20 m/s. In half the simulations, tethering remained on the C_α atoms of every five amino acids, during the whole unbinding process (symmetric tethering). The amino acids within 5 Å of the unbinding ligand were identified; these amino acids are called the "binding path amino acids". In the other half of simulations, the tethering was lifted from the C_α atoms of these binding path amino acids (asymmetric tethering). In order to examine the effect of initial conditions on the unbinding trajectory, two starting conditions were used after equilibration. Either the unbinding simulation continued from equilibration (no velocity re-scaling), or the velocity of all particles were re-scaled to 310 K once, at the beginning of the data production run (velocity re-scaling) [47].

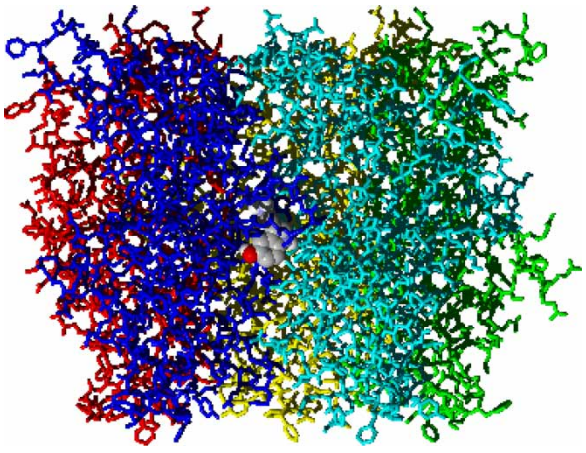

Figure 7. Diagram showing 5-HT docked to its receptor, binding to the interface between two subunits. The protein is colour-coded according to subunit. The drug is displayed in CPK model. See online version for colour.

5.2 Results from 5-HT unbinding

By examining the eight trajectories of unbinding, I identified key amino acids in the pathway [47]. A comparison with experimental data showed that many of the pathway amino acids were important for binding and function, and that mutation of those amino acids would reduce the binding and/or render the receptor non-functional [47]. This work showed that even in the absence of aqueous solvent, MD simulations were capable of giving us a qualitative understanding of intermolecular interaction in biological systems.

5.3 Future directions

This work shows that, under certain circumstances, the mutual repulsion method can allow the ligand to follow an unbinding trajectory very different from the line joining

the centres of mass of the ligand and of the receptor. Whether this is generally applicable is not certain. A more useful method would be to unbind a ligand using random forces [48].

A random force method has been developed by Lüdemann *et al.* [49] In this method, a force is applied to the ligand in a random direction. If the ligand cannot travel a certain distance $r_{\min}$ within a defined time period, the path is abandoned and another direction chosen, until the ligand exits the binding site. The random force $\mathbf{F}_{\text{rand}}$ is expressed thus:

$$\boldsymbol{F}_{\text{rand}} = k\boldsymbol{r}_o \tag{8}$$

where k is a force constant which is kept constant throughout the whole simulation, and $\mathbf{r}_o$ is the unit vector whose magnitude is also kept constant, but whose direction is changed randomly if the ligand velocity drops below a certain value.

In their implementation, Lüdemann *et al.* used a time-step of 2 fs, and for each random direction, between 10 and 100 steps were executed before deciding if a new random direction was required. Obviously the longer the trial period, the greater the danger of imposing conformational changes on the protein. They also maintained the unbinding speed of the ligand between 0.025 and 0.1 m/s.

As test systems, they produced 14 trajectories on the cytochrome P450camcamphor complex, 4 trajectories on the same protein complexed with 5-hydroxycamphor, and also 4 trajectories using the ligand endo-borneol allyl ether. They discovered three main exit pathways, which they named pathways 1–3. Camphor showed a small preference in the order of pathway 2 > pathway 1 > pathway 3, whilst 5-hydroxy-camphor and endo-borneol allyl ether only existed via pathway 1 or 2, with endo-borneol allyl ether showing a preference for pathway 1, and 5-hydroxy-camphor having equal preference for the two possibilities. Interestingly, the pathway with the lowest B-factor is the one least commonly used (pathway 3).

Unfortunately, the random force method developed is not linked to any free energy estimation routine, so would be able to predict an unbinding pathway only. Work is in progress to couple the two methods.

6. Conclusions

Since the advent of simulations methods in the 1950s, they have been gradually taken up by biologists to study ligand–receptor interactions. This method has proved to be another useful tool in the armamentarium of scientists. Used properly, it is capable of giving us information which complements experimental data.

DL_POLY has been used for some of the simulations described. It has many advantages. It is a well-written and transparent piece of software based on rigorous physics (for example, the Nosé-Hoover thermostat [35,36] available gives an NVT ensemble, unlike some temperature-stabilising regimes which makes the temperature constant but does not give an NVT ensemble [50]). The intrinsic options available caters for a large variety of potentials and simulation conditions, it is parallelisable, and the programme package allows the user to alter the source code easily; this is important in research where new potentials, simulation conditions and/or simulation methods need to be incorporated.

The main problem related to using DL_POLY is to generate the CONFIG and FIELD files used by DL_POLY from the coordinates of a biomolecule, usually in the Protein DataBank (PDB) format. A related piece of software, DL_PROTEIN, contains a suite of programmes called build-topology to perform this. Unfortunately, these programmes are rather unwieldy, and, at the time of my carrying out the retinol simulation work, ridden with bugs. I had to locate and correct them. I note that DL_POLY is not yet widely used by biological scientists, which is a pity because it is a more versatile and scientifically-rigorous piece of software than many others. A better and more user-friendly interface between the PDB format and the DL_POLY CONFIG and FIELD files should encourage more biological scientists to switch to this package.

Acknowledgements

I am grateful to Ole Olsen for the original idea of mutual repulsion, Søren Toxværd, Andrew Hardwick, Ruth Lynden-Bell, Chan Shek Ling, Owen Saxton, Alex Selby, André Juffer, Peter Howe, Wolfgang Rieping and Michael Witty for mathematical discussions, Bill Smith for advice on implementing mutual repulsion on DL_POLY, Simone Melchionne for suggesting the method of adiabatic switching, and Greg McMullan, Jenny Barna, Nick Maclaren, Guy Cormier and Leung Hin-Tak for technical help. I thank the UK Biotechnology and Biological Sciences Research Council for the award of a David Phillips Fellowship, New Hall, Cambridge, for a Research Fellowship, and the Royal Society for the award of a European Science Exchange Programme travel grant, for different parts of this work. Provision of computational resources on the Hitachi SR2201 at the High Performance Computing Facility of the University of Cambridge is gratefully acknowledged.

References

[1] English translation by S.-Z. Yang, *Divine Farmer's Materia Medica*, Blue Poppy Press, Boulder, Colorado, USA (1998).
[2] Dioscorides Pedanius of Anazarbos, *De materia medica*, 1518 edition published in Venice by Aldus Manutius.
[3] F.W.A. Sertürner. Über das Morphium, eine neue salzfähige Grundlage, und die Mekonsäure, als Hauptbestandtheile des Opium. *Annalen der Physik*, **55**, 56 (1817).
[4] P. Ehrlich. Über den jetzigen Stand der Chemotherapie. *Berichte der Deutschen Chemischen Gesellschaft*, **42**, 17 (1909).

[5] W.H. Bragg. The reflection of X-ray by crystals. *Proc. Royal Soc., Lond.*, **89A**, 246 (1913).

[6] W.L. Bragg. The structure of some crystals as indicated by their diffraction of X-rays. *Proc. Royal Soc., Lond.*, **89A**, 248 (1913).

[7] M.F. Perutz. Crystal structure of oxyhaemoglobin. *Nature*, **150**, 324 (1942).

[8] W.O. Saxton. *Computer Techniques for Image Processing in Electron Microscopy*, supplement to volume 10 of *Advances in Electronics and Electron Physics* series, Academic Press, New York (1978).

[9] N. Unwin. Nicotinic acetylcholine receptor at 9A resolution. *J. Mol. Biol.*, **229**, 1101 (1993).

[10] N. Unwin. Acetylcholine receptor channel imaged in the open state. *Nature*, **373**, 37 (1995).

[11] A. Miyazawa, Y. Fujiyoshi, N. Unwin. Structure and gating mechanism of the acetylcholine receptor pore. *Nature*, **423**, 949 (2003).

[12] N. Unwin. Refined structure of the nicotinic acetylcholine receptor at 4A resolution. *J. Mol. Biol.*, **346**, 967 (2005).

[13] R. Kaptein, E.R.P. Zuiderweg, R.M. Scheek, R. Boelens, W.F. van Gunsteren. A protein structure from NMR data. *lac* repressor headpiece. *J. Mol. Biol.*, **182**, 179 (1985).

[14] M.P. Williamson, T.F. Havel, K. Wüthrich. Solution conformation of proteinase inhibitor IIA from bull seminal plasma by ^{1}H nuclear magnetic resonance and distance geometry. *J. Mol. Biol.*, **182**, 295 (1985).

[15] C.S. Poornima, P.M. Dean. Hydration in drug design. 1. Multiple hydrogenbonding features of water molecules in mediating protein-ligand interactions. *J. Comput.-Aided Mol. Des.*, **9**, 500 (1995).

[16] D.S. Wilson, B. Guenther, C. Desplan, J. Kuriyan. High resolution crystal structure of a paired (Pax) class cooperative homeodomain dimer on DNA. *Cell*, **82**, 709 (1995).

[17] B.C. Braden, B.A. Fields, R.J. Poljak. Conservation of water molecules in an antibody-antigen interaction. *J. Mol. Recogn.*, **8**, 317 (1995).

[18] R. de Cristofaro, M. Picozzi, E. de Candia, B. Rocca, R. Landolfi. Thrombinthrombomodulin interaction—energetics and potential role of water as an allosteric effector. *Biochem. J.*, **310**, 49 (1995).

[19] D.N. Dubins, R. Filfil, R.B. MacGregor, T.V. Chalikian. Role of water in protein-ligand interactions: volumetric characterisation of the binding of 2′-CMP and 3′-CMP to ribonuclease A. *J. Phys. Chem. B*, **104**, 390 (2000).

[20] J. Åqvist, P. Sandblom, T.A. Jones, M.E. Newcomer, W.F. van Gunsteren, O. Tapia. Molecular dynamics simulations of the holo and apo forms of retinol binding protein. *J. Mol. Biol.*, **192**, 593 (1986).

[21] R.F. Tilton, U.C. Singh, I.D. Kuntz, P.A. Kollman. Protein-ligand dynamics: a 96-ps simulation of a myoglobin-xenon complex. *J. Mol. Biol.*, **199**, 195 (1988).

[22] S.G. Dahl, Ø. Edvardsen, I. Sylte. Molecular dynamics of dopamine at the D_2-receptor. *Proc. Natl Acad Sci, USA*, **88**, 8111 (1991).

[23] J. Leech, J. Prins, J. Hermans. SMD: visual steering of molecular dynamics for protein design. *IEEE Comput. Sci. Eng.*, **3**, 38 (1996).

[24] H. Grubmüller, B. Heymann, P. Tavan. Ligand binding: molecular mechanics calculation of the streptavidin-biotin rupture force. *Science*, **271**, 997 (1996).

[25] S. Izrailev, S. Stepaniants, M. Balsera, Y. Oono, K. Schulten. Molecular dynamics study of unbinding of the avidin-biotin complex. *Biophys. J.*, **72**, 1568 (1997).

[26] D. Kosztin, S. Izrailev, K. Schulten. Unbinding of retinoic acid from its receptor studied by steered molecular dynamics. *Biophys. J.*, **76**, 188 (1999).

[27] S. Freitag, V. Chu, J.E. Penzotti, L.A. Klumb, R. To, D. Hyre, I. le Trong, T.P. Lybrand, R.E. Stenkamp, P.S. Stayton. A structural snapshot of an intermediate on the streptavidin-biotin dissociation pathway. *Proc. Natl Acad. Sci., USA*, **96**, 8384 (1999).

[28] B. Heymann, H. Grubmüller. Molecular dynamics force probe simulations of antibody/antigen unbinding: entropic control and non-additivity of unbinding forces. *Biophys. J.*, **81**, 1295 (2001).

[29] M.V. Bayas, K. Schulten, D. Leckband. Forced detachment of the CD2-CD58 complex. *Biophys. J.*, **84**, 2223 (2003).

[30] P.-L. Chau. Process and thermodynamics of ligand-receptor interaction studied using a novel simulation method. *Chem. Phys. Lett.*, **334**, 343 (2001).

[31] G. Zanotti, R. Berni, H.L. Monaco. Crystal-structure of liganded and unliganded forms of bovine plasma retinol-binding protein. *J. Biol. Chem.*, **268**, 10728 (1993).

[32] W.F. van Gunsteren, H.J.C. Berendsen. *Groningen Molecular Simulation Library Manual*, Biomos B.V, Groningen, Netherlands (1987).

[33] H.J.C. Berendsen, J.R. Grigera, T.P. Straatsma. The missing term in effective pair potentials. *J. Phys. Chem.*, **91**, 6269 (1987).

[34] T.R. Forester, W. Smith. DL_POLY 2.0—a general-purpose parallel molecular dynamics package. *J. Mol. Graph.*, **14**, 136 (1996).

[35] S. Nosé. A unified formulation of the constant temperature molecular dynamics methods. *J. Chem. Phys.*, **81**, 511 (1984).

[36] W.G. Hoover. Canonical dynamics—equilibrium phase space distributions. *Phys. Rev. A*, **31**, 1695 (1985).

[37] S. Melchionna, G. Ciccotti, B.L. Holian. Hoover NPT dynamics for systems varying in size and shape. *Mol. Phys.*, **78**, 533 (1993).

[38] M. Watanabe, W.P. Reinhardt. Direct dynamical calculation of entropy and free energy by adiabatic switching. *Phys. Rev. Lett.*, **65**, 3301 (1990).

[39] N. Noy, Z.J. Xu. Thermodynamic parameters of the binding of retinol to binding proteins and to membranes. *Biochemistry*, **29**, 3888 (1990).

[40] K.L. Clarkson, K. Mehlhorn, R. Seidel. Four results on randomized incremental constructions. *Comput. Geom.—Theory Appl.*, **3**, 185 (1993).

[41] S.L. Chan, E.O. Purisima. Molecular surface generation using marching tetrahedra. *J. Comput. Chem.*, **19**, 1268 (1998).

[42] P.-L. Chau. Water movement during ligand unbinding from receptor site. *Biophys. J.*, **87**, 121 (2004).

[43] D.C. Reeves, M.F.R. Sayed, P.-L. Chau, K.L. Price, S.C.R. Lummis. Identification of 5-HT_3 receptor agonist-binding residues using homology modelling. *Biophys. J.*, **84**, 2338 (2003).

[44] A.D. MacKerell, et al. All-atom empirical potential for molecular modeling and dynamics Studies of proteins. *J. Phys. Chem. B*, **102**, 3586 (1998).

[45] Structure 4 from D.C. Reeves, M.F.R. Sayed, P.-L. Chau, K.L. Price, S.C.R. Lummis. Identification of 5-HT_3 receptor agonist-binding residues using homology modelling. *Biophys. J.*, **84**, 2338 (2003).

[46] Structure 2 from A.J. Thompson, K.L. Price, D.C. Reeves, S.L. Chan, P.-L. Chau, S.C.R. Lummis. Locating an antagonist in the 5-HT_3 receptor binding site: a modeling and radioligand binding study. *J. Biol. Chem.*, **280**, 20476 (2005).

[47] A.J. Thompson, P.-L. Chau, S.L. Chan, S.C.R. Lummis. Unbinding pathways of an agonist and an antagonist from the 5-HT_3 receptor. *Biophys. J.*, **90**, 1979 (2006).

[48] Private communication with Søren Toxværd in 1996.

[49] S.K. Lüdemann, V. Lounnas, R.C. Wade. How do substrates enter and products exit the buried active site of cytochrome P450cam? 1. Random expulsion molecular dynamics investigation of ligand access channels and mechanisms. *J. Mol. Biol.*, **303**, 797 (2000).

[50] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola, J.R. Haak. Molecular dynamics with coupling to an external bath. *J. Chem. Physics*, **81**, 3684 (1984).