Peter Deuflhard  Jan Hermans
Benedict Leimkuhler  Alan E. Mark
Sebastian Reich  Robert D. Skeel (Eds.)

Computational Molecular Dynamics: Challenges, Methods, Ideas


With 117 Figures and 22 Tables

Springer
Steered Molecular Dynamics

Sergei Izrailev, Sergey Stepaniants, Barry Israelewitz, Dorina Kosztin, Hui Lu, Ferenc Molnar, Willy Wriggers, and Klaus Schulten

Beckman Institute, University of Illinois at Urbana-Champaign, 405 N. Mathews, Urbana, IL 61801, USA

Abstract. Steered molecular dynamics (SMD) induces unbinding of ligands and conformational changes in biomolecules on time scales accessible to molecular dynamics simulations. Time-dependent external forces are applied to a system, and the responses of the system are analyzed. SMD has already provided important qualitative insights into biologically relevant problems, as demonstrated here for applications ranging from identification of ligand binding pathways to explanation of elastic properties of proteins. First attempts to deduce potentials of mean force by discounting irreversible work performed on the system are summarized. The non-equilibrium statistical mechanics underlying analysis of SMD data is outlined.

1 Introduction

Molecular recognition and specific ligand-receptor interactions are central to many biochemical processes. The regulation of cellular signal-transduction pathways and gene expression, activity of enzymes, cell motility, molecular immunology and the action of hormones involve the triggering of functional responses by noncovalent associations of ligands with receptors. The prediction and design of ligands (inhibitors or substrates) for a given receptor is the main goal in rational drug design, and considerable effort is spent on the development of corresponding computational methods (Cohen et al., 1990; Colman, 1994; Marrone et al., 1997). New pharmaceuticals, e.g., the HIV protease inhibitors (Thaisrivongs et al., 1996; Lebon et al., 1996; Hanessian and Devasthale, 1996), derived in part from such methods, have made a major impact on clinical medicine, and computational modeling will be of increasing importance in the future.

Despite an abundance of modeling methods for ligand-receptor interactions and protein-protein docking (Strynadka et al., 1996) little is known about processes governed by adhesive interactions such as those occurring in the binding and unbinding of ligands. Presently, the prevailing point of view concerning computer simulations describing ligand binding and determining binding affinities is to strive for the ideal of reversibility, as in umbrella sampling and free energy perturbation (McCammmon and Harvey, 1987; Ajay and Murcko, 1995; Gilson et al., 1997), with the hope that artifacts induced by the finite rate of conformational changes can be neglected. Reaching this ideal, however, requires extremely slow manipulation and, therefore, prohibitively expensive simulations. This chapter advocates a new computational method,
steered molecular dynamics (SMD), which accepts irreversibility, ceding for the present time accurate evaluation of binding affinities and potentials of mean force, but gaining access to biologically relevant information related to non-covalent bonding. We will demonstrate the wealth of such information using a broad range of examples. The concern that thermodynamic potentials cannot, even in principle, be obtained from irreversible processes has been proven unfounded by the remarkable identity derived by Jarzynski (1997a,b), 
\[ \langle \exp[-W/k_BT] \rangle = \exp[-\Delta F/k_BT] \]. This identity connects the ensemble average of an exponential of the total work \( W \) performed on the system during a non-equilibrium transition from one state to another to the free energy difference \( \Delta F \) between the two states.

Experimental techniques based on the application of mechanical forces to single molecules in small assemblies have been applied to study the binding properties of biomolecules and their response to external mechanical manipulations. Among such techniques are atomic force microscopy (AFM), optical tweezers, biomembrane force probe, and surface force apparatus experiments (Binning et al., 1986; Block and Svoboda, 1994; Evans et al., 1995; Israelachvili, 1992). These techniques have inspired us and others (see also the chapters by Eichinger et al. and by Hermans et al. in this volume) to adopt a similar approach for the study of biomolecules by means of computer simulations.

\[ \text{Fig. 1.} \]

Extraction of a ligand from the binding pocket of a protein. The force (represented by an arrow) applied to the ligand (shown in van der Waals spheres) leads to its dissociation from the binding pocket of the protein (a slice of the protein represented as a molecular surface is shown).
In SMD simulations time-dependent external forces are applied, for example, to a ligand to facilitate its unbinding from a protein, as shown in Fig. 1. The analysis of the interactions of the dissociating ligand with the binding pocket, as well as the recording (as a function of time) of applied forces and ligand position, yields important structural information about the structure-function relationships of the ligand-receptor complex, binding pathways, and mechanisms underlying the selectivity of enzymes. SMD can also be applied to investigate the molecular mechanisms that determine elastic properties exhibited by proteins subjected to deformations in AFM and optical tweezer experiments, such as stretching of titin leading to unfolding of its immunoglobulin domains (Rief et al., 1997), or stretching of tenascin which results in unfolding of its fibronectin-III domains (Oberhauser et al., 1998).

Besides yielding qualitative information, these biologically and pharmaceutically motivated applications of SMD can also yield quantitative information about the binding potential of the ligand-receptor complex. A first advance in the reconstruction of the thermodynamic potential from SMD data by discounting irreversible work was made by Balsera et al. (1997) as outlined in Sect. “Reconstruction of the potential of mean force” below.

In the following we describe the methodology of SMD, illustrate applications of SMD through key examples, present the non-equilibrium statistical mechanical theory of SMD, and describe a method of reconstruction of a potential of mean force from SMD data. The applications include studies of the dissociation of biotin from avidin and streptavidin (Izailev et al., 1997; Grubmüller et al., 1996) (see also the chapter by Eichinger et al. in this volume), the unbinding of retinal from bacteriorhodopsin (Isralewitz et al., 1997), the release of phosphate from actin (Wriggers and Schulten, 1998), the possible binding pathways of thyroid hormone to its receptor, the extraction of lipids from membranes (Stepaniants et al., 1997; Marrink et al., 1998), the unbinding of arachidonic acid from the prostaglandin H₂ synthase-1, and the force-induced unfolding of titin immunoglobulin domains (Lu et al., 1998). In the chapter by Hermans et al. in this volume another application of SMD, the extraction of bound xenon from mutant T4-lysozyme, is discussed.

2 Methods

One way to apply external forces to a protein-ligand complex is to restrain the ligand to a point in space (restraint point) by an external, e.g., harmonic, potential. The restraint point is then shifted in a chosen direction (Grubmüller et al., 1996; Isralewitz et al., 1997; Stepaniants et al., 1997; Marrink et al., 1998; Lu et al., 1998), forcing the ligand to move from its initial position in the protein and allowing the ligand to explore new contacts along its unbinding path. Assuming a single reaction coordinate \( x \), and an external potential \( U = K(x - x₀)^2/2 \), where \( K \) is the stiffness of the restraint, and \( x₀ \) is the initial position of the restraint point moving with a constant velocity \( v \), the
external force exerted on the system can be expressed as

$$F = K(x_0 + vt - x).$$  \hspace{1cm} (1)

This force corresponds to the ligand being pulled by a harmonic spring of stiffness $K$ with its end moving with velocity $v$. Alternatively, a fixed restraint point at a distance much larger than the length of the unbinding pathway may be chosen. In this case, the end of the spring does not move and its stiffness is linearly increased with time (Izrailev et al., 1997), i.e., $K = \alpha t$, and the force is

$$F = \alpha t(x_0 - x).$$  \hspace{1cm} (2)

Other external forces or potentials can also be used, e.g., constant forces, or torques applied to parts of a protein to induce rotational motion of its domains (Wrigners and Schulten, 1997a).

SMD simulations require selection of a path, i.e., a series of directions of the applied force. In some cases a straight line path is sufficient, e.g., for avidin-biotin (Fig. 2), actin (Fig. 4), lipids in membranes (Fig. 6), or the unfolding of titin immunoglobulin domains (Fig. 8). Other biomolecular systems involve a ligand positioned at the bottom of a convoluted binding cleft, e.g., bacteriorhodopsin (Fig. 3), prostaglandin H$_2$ synthase (Fig. 7), and nuclear hormone receptors (Fig. 5). In the latter cases the forced unbinding of the ligand requires the direction of the force to be changed during the simulation to avoid distortion of the surrounding protein. The direction of the force can be chosen randomly (Lüdemann et al., 1997) or by guessing a direction on the basis of structural information. A force is then applied to the ligand in the chosen direction, and this direction is accepted or rejected based on factors such as conservation of secondary structure of the protein, deformation of the protein, the magnitude of the force applied, the average velocity of the ligand along the unbinding pathway, etc. (Isralewitz et al., 1997; Lüdemann et al., 1997). One possible protocol for selecting force directions in SMD defines a conical region of space around a preferred direction and selects new directions randomly within this region. A small cone angle strongly biases the chosen directions to the initial guess, whereas a large cone angle leads to exploration of more directions.

An initial and desired final configuration of a system can be used by the targeted molecular dynamics (TMD) method (Schlitter et al., 1993) to establish a suitable pathway between the given configurations. The resulting pathway can then be employed during further SMD simulations for choosing the direction of the applied force. TMD imposes time-dependent holonomic constraints which drive the system from one known state to another. This method is also discussed in the chapter by Helms and McCammon in this volume.

Other methods for identifying multi-dimensional reaction paths are based on stochastic dynamics. For example, a reaction path can be found by opti-
mization of the Onsager-Machlup action between the two end points of a trajectory (Olender and Elber, 1996) (see the chapter by Elber et al. in this volume). Alternatively, using the conformational flooding method (Grubmüller, 1995), one may sample the distribution of ligand conformations through principle component analysis and use forces derived from this analysis to drive the ligand away from the current distribution, as discussed in the chapter by Eichinger et al. in this volume.

3 Applications of SMD

The ultimate criterion for the value of a method such as SMD is how much can be learned from using it. In this section we provide examples of SMD applications yielding important insights into biological processes. First, we review the study of the biotin-avidin complex which served as a test bed for the method, then discuss three examples in which SMD identified binding pathways of ligands. Next, we demonstrate how SMD elucidated two key steps in fatty acid metabolism, namely, the extraction of lipids from membranes by phospholipase A₂ and the binding of arachidonic acid by prosta glandin H₂ synthase. Finally, we show how SMD revealed the mechanism behind the stretching of titin immunoglobulin domains.

3.1 Avidin-Biotin Complex as a Test Bed for SMD

The avidin-biotin complex, known for its extremely high affinity (Green, 1975), has been studied experimentally more extensively than most other protein-ligand systems. The adhesion forces between avidin and biotin have been measured directly by AFM experiments (Florin et al., 1994; Moy et al., 1994b; Moy et al., 1994a). SMD simulations were performed on the entire tetramer of avidin with four biotins bound to investigate the microscopic detail of unbinding of biotin from avidin (Izrailev et al., 1997).

In the simulations the rupture of biotin from avidin was induced by means of a soft harmonic restraint, as described by Eq. (2) with \( K = \alpha t \) ranging from 0 to 120 pN/Å. The spatial range of thermal fluctuations of biotin associated with the restraint was \( \delta x \approx 3 \) Å, i.e., on the order of the size of the binding pocket (about 10 Å). The fluctuations of the applied force, on the other hand, were small compared to its absolute value, and the force profiles exhibited nearly linear growth with time, similar to that reported in AFM experiments (see Fig. 2). The values of the rupture forces, i.e., the maximum measured force (450–800 pN), exceeded those measured in AFM experiments (160 pN). These SMD simulations did not exhibit any particular scaling of the rupture force with the pulling rate, which covered a span of almost two orders of magnitude. In SMD simulations of a similar streptavidin-biotin complex reported by Grubmüller et al. (1996) (see the chapter by Eichinger et al. in this volume), the scheme of Eq. (1) with a stiff spring (\( K = 280 \) pN/Å) was
employed, and the rupture force was found to scale linearly with the velocity $v$ (cf. Sect. "Stochastic Modeling of SMD").

The simulations also revealed that flapping motions of one of the loops of the avidin monomer play a crucial role in the mechanism of the unbinding of biotin. The fluctuation time for this loop as well as the relaxation time for many of the processes in proteins can be on the order of microseconds and longer (Eaton et al., 1997). The loop has enough time to fluctuate into an open state on experimental time scales (1 ms), but the fluctuation time is too long for this event to take place on the nanosecond time scale of simulations. To facilitate the exit of biotin from its binding pocket, the conformation of this loop was altered (Izrailev et al., 1997) using the interactive molecular dynamics features of MDScope (Nelson et al., 1995; Nelson et al., 1996; Humphrey et al., 1996).

During unbinding, biotin was found to move in discrete steps (see Fig. 2). Each step can be identified with the formation and rupture of a network of hydrogen bonds which stabilize biotin in the binding pocket of avidin. The strongest bonds were formed between biotin and polar residues Tyr33, Ser16 and Thr35, consistent with experimental observations. Contacts of biotin with
nonpolar residues (see Fig. 2), especially with Trp110 of an adjacent avidin monomer (in the complete tetramer), are crucial for the unbinding process (Izrailev et al., 1997). These residues prevent water molecules from entering the binding pocket. To determine the effect of water molecules on the unbinding mechanism, 50 water molecules were placed in the avidin tetramer with the program DOWSER (Zhang and Hermans, 1996) (the algorithm for placing water in proteins is discussed in the chapter by Hermans et al. in this volume). The presence of four water molecules in the outer region of the binding pocket, i.e., close to biotin's valeryl side-chain carboxylate group, did not affect the stepwise motion of biotin, but reduced the rupture (maximum) force from 600 pN to 400 pN as shown in Fig. 2. The reduction of the rupture force resulted from the participation of water molecules in breaking the hydrogen bond networks between biotin and residues located near the exit of the binding pocket. Water did not penetrate the binding pocket on the time scale of the simulations.

3.2 Binding of Retinal to Bacterio-opsin

Bacteriorhodopsin (bR) (Oesterhelt et al., 1992; Schulten et al., 1995) is a light-driven vectorial proton pump found in the membrane of Halobacterium salinarum. The protein binds a retinal molecule through a Schiff base linkage to its Lys216 side group. Formation of bR from the apoprotein and retinal has been studied experimentally (Oesterhelt and Schumann, 1974; Chang et al., 1988; Booth et al., 1996), but the pathway of initial retinal entry during bR formation was not clearly understood. Despite its extremely poor solubility in water and a considerable affinity for lipid environment, retinal was generally believed to enter the protein through the solvent-exposed loops of the protein. However, a window on the lipid-exposed surface of bR located between helices E and F (see Fig. 3) which uncovers part of retinal (its β-ionone ring) can be an entry point for retinal. Inspection of the bR structure revealed that this window, in fact, is the only opening large enough to allow retinal entry and provide access to the Lys216 binding site of retinal. SMD simulations were performed to test this hypothesis by extracting retinal from bR with an external force along a path towards and out of the putative entry window (Israelewitz et al., 1997). If such an extraction could be carried out without significantly perturbing the protein on the time scale of MD simulations, then the extraction path could also constitute the binding path of retinal to the apoprotein on the much longer time scale of bR formation.

Due to the convoluted shape of the retinal binding site, retinal cannot be extracted from bR by application of a force along a single straight line. Therefore, the unbinding path was segmented, with the direction of the force determined anew for each of the ten segments (cf. Sect. “Methods”). For each segment the applied forces were described by Eq. (1) with \( K = 10 \) \( k_B T / \AA^2 \approx 414 \) pN/Å and \( v = 0.125 \) Å/ps at \( T = 300 \) K.
(a) Extraction of retinal from bacteriorhodopsin (bR). The backbone of bR is shown in tube and ribbon representation; the seven trans-membrane helices of bR are labeled A–G. The structure presents a snapshot of the SMD simulation with retinal (dark spheres) partially extracted from the protein bR through an opening between helices E and F into the hydrophobic phase of the surrounding membrane. (b) Magnification of boxed area in (a), at completion of retinal’s exit from bR, representing the putative exit/entry window of retinal. The amino acids shown, Ala144, Met145, and Ser183, form a hydrogen bond network with retinal’s carbonyl group attracting and guiding retinal into the protein.

It was found that extraction of retinal from bR along a path through the window between helices E and F (see Fig. 3) could be accomplished during a 0.2 ns simulation without disrupting the protein structure (Israelewitz et al., 1997). The maximum force applied was about 1000 pN and accounted for breaking of a strong hydrogen bond between retinal and Lys216.

Upon exit from the interior of bR, retinal formed a stable network of hydrogen bonds with residues Ala144, Met145 and Ser183 which line the putative exit/entry window (see Fig. 3). This suggests that retinal approaches the apoprotein from the hydrophobic phase of the membrane, binds to the stated residues, subsequently moves into bR forming a hydrogen bond with Lys216 and, finally, forms the Schiff base bond.

### 3.3 Actin’s Back Door

Actin filaments are dynamic polymers whose assembly and disassembly in the cell cytoplasm drives shape changes (Small, 1989), cell locomotion (Theriot et al., 1992), and chemotactic migration (Theriot et al., 1992)Devreotes and
Zigmond, 1988). The ATP-hydrolysis that accompanies actin polymerization, ATP $\rightarrow$ ADP $+$ $P_i$, and the subsequent release of the cleaved phosphate ($P_i$) are believed to act as a clock (Pollard et al., 1992; Allen et al., 1996), altering in a time-dependent manner the mechanical properties of the filament and its propensity to depolymerize. Molecular dynamics simulations suggested a so-called back door mechanism for the hydrolysis reaction ATP $\rightarrow$ ADP $+$ $P_i$ in which ATP enters the actin from one side, ADP leaves from the same side, but $P_i$ leaves from the opposite side, the “back door” (Wriggers and Schulten, 1997b). This hypothesis can explain the effect of the toxin phalloidin which blocks the exit of the putative back door pathway and, thereby, delays $P_i$ release as observed experimentally (Dancker and Hess, 1990).

![Fig. 4.](image)

Phosphate release from actin. (a) Monomeric actin with ADP and $P_i$ bound. The protein backbone (tube), ADP (grey spheres), and Ca$^{2+}$-$P_i$ (black spheres) are shown. The orientation of the spring indicates the pulling direction during $P_i$ unbinding. (b) Force exerted on the deprotonated (solid line) and protonated (dashed line) phosphate during the SMD simulations.

To reveal the microscopic processes underlying the unbinding of $P_i$, SMD simulations were carried out in which $P_i$ was pulled along the back door pathway and the adhesion forces were measured (Wriggers and Schulten, 1998). The simulations revealed that the dissociation of $P_i$ is likely to be controlled by its protonation. $P_i$, which is singly protonated (HPO$_4^{2-}$) after cleavage from ADP, needs to overcome a strong Coulomb energy barrier due to the presence of a Ca$^{2+}$ ion associated with ADP. The resulting forces reached 3,000 pN as shown in Fig. 4; in case of protonated $P_i$ (H$_2$PO$_4^-$) the maximal forces measured 2,400 pN. This suggests that protonation of $P_i$ is required for unbinding from actin, consistent with kinetic measurements (Allen et al., 1996).
The P$_i$-coordinating amino acid residues and solvent molecules in the dissociation pathway exhibited a remarkable functional diversity (Wriggers and Schulten, 1998). A methylated histidine, highly conserved among actin species and believed to be functionally relevant, stabilized bound P$_i$ through a rotation of its side chain relative to the crystal structure. The long side chain of an arginine remained attached to P$_i$ for the most part of the unbinding, guiding the ligand to the protein surface. Other P$_i$-coordinating side chains were replaced by water molecules in the solvated back door channel. This hydration of P$_i$ gave rise to a velocity-dependent unbinding force that reflects the mobility of the water molecules relative to the displaced P$_i$. A hydration step during unbinding has also been observed in other SMD simulations, e.g., in case of the extraction of retinal from bR (Isralewitz et al., 1997).

### 3.4 Binding of Hormones to Nuclear Hormone Receptors

Hormone binding to the thyroid hormone receptor initiates a series of molecular events culminating in the activation or repression of transcription of target genes. The transition between the bound and unbound form of the thyroid receptor is accompanied by a conformational change that enables the hormone–receptor complex to bind to specific sequences of DNA and other transcriptional coactivators or repressors (Brent et al., 1989; Damm et al., 1989; Andersson et al., 1992). SMD can determine likely pathways of hormone binding and unbinding, reveal components of the receptor involved in the unbinding, and thus contribute to the design of new ligands for hormone therapy.

An examination of the crystal structure of the rat α$_1$ thyroid hormone receptor (TR) ligand binding domain bound with a thyroid hormone agonist (Wagner et al., 1995) suggests three entry/exit points for the hormone as shown in Fig. 5a. By applying an external force to the ligand to facilitate its unbinding from the protein, the three possible pathways were explored. In the simulations, the protein-ligand system was surrounded by a water bath. One atom of the hormone was harmonically restrained ($K = 10 \text{ kcal/molÅ}^2 \approx 695 \text{ pN/Å}$) to a point moving with a constant velocity $v = 0.08 \text{ Å/ps}$ in a chosen direction. The investigation is still ongoing and presently only preliminary conclusions can be drawn from the SMD data.

During the unbinding process the force exerted on the hormone varied according to the interaction of the hormone with surrounding protein residues. The hormone was found to leave the binding pocket along all three pathways, but exerted the least effect on the protein conformation when pulled along path 1. The carboxylate group of the hormone formed direct hydrogen bonds with the guanidium of Arg228 and the amino nitrogen of Ser277. The external force exerted on the hormone increased until the Arg228 residue was dislodged and the associated hydrogen bond was broken. These events corresponded to the maximum values of the force shown in Fig. 5b. It should also be mentioned that path 1 was lined with flexible amino acid side groups,
Steered Molecular Dynamics

Fig. 5.

(a) Possible unbinding pathways for the dinit hormone from the thyroid hormone receptor: (1) the hormone leaves through the only discernible opening in the molecular surface; (2) the hormone moves underneath the last two helices of the protein (helices 11 and 12); (3) the hormone moves between helices 11 and 12. (b) Force profiles for the three unbinding pathways shown in (a) (top to bottom: path 1, path 2, path 3)

the flexibility of which was reflected by high experimentally observed temperature factors. Along paths 2 and 3, the hormone encountered Phe residues that needed to be moved out of the way. In both cases the force increased (see Fig. 5b) until the Phe residues changed their positions.

3.5 Extraction of Lipids from Membranes

SMD simulations were performed to investigate the extraction of a lipid from the dilauroyl-phosphatidyl-ethanolamin (DLPE) monolayer into the aqueous phase (see Fig. 6a) (Stepanians et al., 1997). External forces described by Eq. (1) with $K = 10 \text{kcal/molÅ}^2 \approx 695 \text{pNÅ}$ and $v = 0.014 \text{Å/ps}$ were applied to the head group of the lipid, pulling it out from the membrane. The forces required to extract the lipid measured about 200 pN and remained constant within the range of fluctuations as shown in Fig. 6c. Analogous simulations were carried out by Marrink et al. (1998), extracting lipids from a dipalmitoyl-phosphatidyl-choline (DPPC) bilayer with pulling velocities of $v = 0.01$–0.5 Å/ps and resulting forces of 200–800 pN. In agreement with the results of SMD simulations of unbinding of biotin from streptavidin (Grubmüller et al., 1996) and avidin (Grubmüller et al., 1996) Izrailev et al.,
1997), performed with a range of pulling rates, the rupture force was found to decrease with decreasing \( v \). The applied force as a function of distance increased up to the point of rupture, and then gradually decreased as the lipid proceeded into the solvent.

Although extraction of lipids from membranes can be induced in atomic force apparatus (Leckband et al., 1994) and biomembrane force probe (Evans et al., 1991) experiments, spontaneous dissociation of a lipid from a membrane occurs very rarely because it involves an energy barrier of about 20 kcal/mol (Cevc and Marsh, 1987). However, lipids are known to be extracted from membranes by various enzymes. One such enzyme is phospholipase A\(_2\) (PLA\(_2\)), which complexes with membrane surfaces, destabilizes a phospholipid, extracts it from the membrane, and catalyzes the hydrolysis reaction of the sn-2-acyl chain of the lipid, producing lysophospholipids and fatty acids (Slotboom et al., 1982; Dennis, 1983; Jain et al., 1995). SMD simulations were employed to investigate the extraction of a lipid molecule from a DLPE monolayer by human synovial PLA\(_2\) (see Fig. 6b), and to compare this process to the extraction of a lipid from a lipid monolayer into the aqueous phase (Stepaniants et al., 1997).

Due to the selection of a stiff restraint, the head group of the lipid was not allowed to fluctuate substantially and its motion essentially followed that of the restraint point. The forces measured during the extraction of the lipid exhibited large fluctuations on the order of 300 pN, as expected when a stiff restraint is employed (Izrailev et al., 1997; Balsera et al., 1997). The forces required to displace the lipid from the membrane into the binding pocket of PLA\(_2\), shown in Fig. 6d, were larger than those required to displace the lipid from the membrane into the aqueous phase. This difference in the measured forces was due in part to the fact that the steric hindrance experienced by the lipid on its way out of the membrane into the active site of PLA\(_2\) was larger than that for its movement into the aqueous phase; repositioning of PLA\(_2\) could have reduced this hindrance. The results do not agree with the hypothesis of destabilization of the lipids by PLA\(_2\) facilitating lipid extraction by the enzyme. The disagreement may have resulted from the steric effects mentioned above, an imperfect choice of the pulling direction for the lipid extraction into the enzyme, or insufficient sampling due to the short (500 ps) simulation time.

### 3.6 Binding of Arachidonic Acid to Prostaglandin H\(_2\) Synthase-1

The enzyme prostaglandin H\(_2\) synthase-1 (PGHS-1) catalyzes the transformation of the essential fatty acid, arachidonic acid (AA), to prostaglandin H\(_2\) (Smith and DeWitt, 1996). This is the first committed step in the biosynthesis of prostanoids which modulate physiological processes such as platelet aggregation and inflammation. Aspirin, flurbiprofen, and other non-steroidal anti-inflammatory drugs directly target PGHS-1 by preventing the access of AA to its cyclooxygenase active site. This site involves a hydrophobic chan-
Fig. 6.
(a) Extraction of a lipid (black spheres) from the DLPE monolayer (lines) into the aqueous phase. (b) Extraction of a lipid (black spheres) from the DLPE monolayer (lines) into protein phospholipase A$_2$ (tube) solvated in water. (c) Force applied to the head of the lipid along the pulling direction during the extraction of the lipid into the aqueous phase. (d) Force applied to the head of the lipid along the pulling direction during the extraction of the lipid into the binding pocket of PLA$_2$.

nel, approximately 25 Å deep and 8 Å wide, which bends the fatty acid into a U-shape (Fig. 7). This shape is required for the catalyzed cyclooxygenation reaction.

Based on the crystal structure of PGHS-1, with flurbiprofen bound at the active site, a model for AA embedded in the enzyme was suggested, in which AA replaces the inhibitor (Picot et al., 1994). The aim of the investigation was to identify how AA folds itself into the required U-shape in the narrow binding channel, rather than entering the channel in a straight conformation. The simulations also sought to identify key residues guiding AA binding.

One monomer of the PGHS-1 homo-dimer (Picot et al., 1994), shown in Fig. 7 with AA bound in its putative cyclooxygenation site was used as a starting point for simulations enforcing the unbinding of AA. Pulling on the methyl end-group of the fatty acid with a harmonic spring as described by Eq. (1) and a range of force constants ($K = 200$–$400$ pN/Å) and pulling velocities ($v = 10$–$0.1$ Å/ps) led to the exit of the ligand from its narrow hydrophobic binding channel. During the forced unbinding a series of concerted torsional motions was observed (Fig. 7). AA contains four rigid cis double
Fig. 7.
Binding site of arachidonic acid (AA) in prostaglandin H₂ synthase-1. An intermediate configuration observed during the unbinding event is depicted. AA contains four rigid cis double bonds connected to each other by a pair of conformationally flexible single bonds. The counter-rotation of atoms C⁵ and C⁶ around an axis defined by the C⁵-C⁶ bond was monitored through the change in dihedral angles defined by the C-atoms abcd and cdef. This rotation proceeded in a way which left the “backbone” of the ligand made of the conformationally rigid cis double bonds relatively unaffected (dihedral angle between C-atoms acdf).

bonds connected to each other by a pair of conformationally flexible single bonds. The unbinding mechanism can be described as a series of rotations around these single bonds which leave the “rigid backbone” of the fatty acid, formed by the conformationally inflexible cis double bonds, relatively unaffected (Fig. 7). This type of concerted motion was shown to be specific for the chemical structure of AA. If the all-cis-isomer was altered to an isomer with a trans double bond, the concerted motion became hindered and the unbinding of the molecule rendered energetically unfavorable.

Another set of simulations was carried out with the targeted molecular dynamics (TMD) method (Schlitter et al., 1993). The initial and final structures of an SMD simulation were used as input for the TMD simulations as discussed in “Methods”. TMD trajectories were calculated in “both directions” between the input structures, simulating both the binding and the
unbinding events. A comparison of the SMD and TMD simulations revealed that the pathways generated by both methods show very similar modes of concerted rotations around single bonds during the unbinding of AA.

Both methods suggest that the chemical structure of AA (cis double bonds connected by two single bonds) allows the fatty acid to access the cyclooxygenase active site of PGHS-1 through a narrow hydrophobic channel and to bind in a shape favorable for the cyclooxygenation reaction.

### 3.7 Force-Induced Unfolding of Titin

The giant muscle protein titin, also known as connectin, is a roughly 30,000 amino acid long filament which plays a number of important roles in muscle contraction and elasticity (Labeit et al., 1997; Maruyama, 1997; Wang et al., 1993). The I-band region of titin, largely composed of immunoglobulin-like (Ig) domains, is believed to be responsible for the molecule’s extensibility and passive elasticity. Recent AFM (Rief et al., 1997) and optical tweezer (Rief et al., 1997) Kellermayer et al., 1997; Tskhovrebova et al., 1997) experiments directly measured the force-extension profile of single titin molecules. In the AFM experiment, cloned sections of titin composed of adjacent I-band Ig domains were stretched at constant speed. The force-extension profile showed a pattern of sawtooth-shaped peaks, spaced 250–280 Å apart, with each force peak corresponding to the unfolding of a single Ig domain. The Ig domains were thus observed to unfold one by one under the influence of an applied external force. To examine in atomic detail the dynamics and structure-function relationships of this behavior, SMD simulations of force-induced titin Ig domain unfolding were performed (Lu et al., 1998).

The SMD simulations were based on an NMR structure of the Ig domain I27 of the cardiac titin I-band (Improna et al., 1996). The Ig domains consist of two β-sheets packed against each other, with each sheet containing four strands, as shown in Fig. 8b. After I27 was solvated and equilibrated, SMD simulations were carried out by fixing one terminus of the domain and applying a force to the other in the direction from the fixed terminus to the other terminus. Simulations were performed as described by Eq. (1) with \( v = 0.5 \, \text{Å/ps} \) and \( K \approx 10 \, k_B T / \text{Å}^2 \approx 414 \, \text{pN/Å} \). The force-extension profile from the SMD trajectory showed a single force peak as presented in Fig. 8a. This feature agrees well with the sawtooth-shaped force profile exhibited in AFM experiments.

The simulation trajectory shown in Fig. 8b provides an explanation of how the force profile in Fig. 8a arises. During extension from 0 to 10 Å the two β-sheets slid away from each other, each maintaining a stable structure and its intra-sheet backbone hydrogen bonds. As the extension of the domain reached 14 Å, the structure within each sheet began to break: in one sheet, strands A' and G slid past each other, while in the other sheet, strands A and B slid past each other. The A'–G and A–B backbone hydrogen bonds broke nearly simultaneously, producing the large initial force peak seen in Fig. 8a.
(a) Force extension profile from SMD simulations of titin I27 domain with a pulling velocity $v = 0.5$ Å/ps. The extension domain is divided into four sections: I. pre-burst, II. major burst, III. post-burst, IV. pulling of fully extended chain. (b) Intermediate stages of the force-induced unfolding. All I27 domains are drawn in the cartoon representation of the folded domain; solvating water molecules are not shown. The four figures at extensions 10 Å, 17 Å, 150 Å, and 285 Å correspond, respectively, to regions I to IV defined in (a).

These events marked the beginning of the Ig domain unfolding, after which the strands unraveled one at a time, accompanied by a large reduction in the recorded force. After an extension of 260 Å, the domain was completely unfolded; further stretching of the already extended polypeptide chain caused the force to increase dramatically.

The simulation (Lu et al., 1998) suggested how Ig domains achieve their chief design requirement of bursting one by one when subjected to external forces. At small extensions, the hydrogen bonds between strands A and B and between strands A' and G prevent significant extension of a domain, i.e.,
the domain maintains its $\beta$-sandwich structure. After these bonds break, resistance to unfolding becomes much smaller, and the domain unfolds rapidly. Only when a domain is fully extended does the force increase enough to begin the unfolding process in another domain.

4 Stochastic Modeling of SMD

In AFM experiments as well as in SMD simulations, the ligand extracted from the binding pocket is subjected to a time-dependent external force. The rate of change and the functional form of the applied force critically affect the behavior of the ligand and the information one can obtain from experiment and simulation (Evans and Ritchie, 1997; Izrailev et al., 1997; Balsera et al., 1997). To better understand the results of SMD simulations and how they compare to experimental measurements, it is helpful to consider an idealized one-dimensional stochastic model that captures the essence of unbinding phenomena and reveals the limit of the information that can be gained from SMD simulations about the binding potentials. The model assumes that the motion of the ligand proceeds in the strong friction limit along a reaction coordinate $x$, and is governed by a one-dimensional Langevin equation

$$\gamma \dot{x} = -\frac{dU}{dx} + F(x, t) + \sigma N(t),$$

(3)

where $\gamma$ is the time-independent coefficient of friction of the ligand in the binding pocket of the protein, $U(x)$ and $F(x, t)$ are respectively the potential surface governing dissociation and the external force applied to the ligand, and $\sigma N(t)$ is a stochastic force of amplitude $\sigma$ and zero mean.

We assume that the unbinding reaction takes place on a time scale long compared to the relaxation times of all other degrees of freedom of the system, so that the friction coefficient can be considered independent of time. This condition is difficult to satisfy on the time scales achievable in MD simulations. It is, however, the most favorable case for the reconstruction of energy landscapes without the assumption of thermodynamic reversibility, which is central in the majority of established methods for calculating free energies from simulations (McCammon and Harvey, 1987; Elber, 1996) (for applications and discussion of free energy calculation methods see also the chapters by Helms and McCammon, Hermans et al., and Mark et al. in this volume).

In this section we describe the behavior of a ligand subjected to three types of external forces: a constant force, forces exerted by a moving stiff harmonic spring, and forces exerted by a soft harmonic spring. We then present a method of reconstruction of the potential of mean force from SMD force measurements employing a stiff spring (Izrailev et al., 1997; Balsera et al., 1997).
4.1 Unbinding Induced by a Constant Force

Unbinding processes can be viewed as taking place in several qualitatively different regimes (Izrailev et al., 1997; Marrink et al., 1998). These regimes can be illustrated by considering the simplest binding potential

\[
U(x) = \begin{cases} 
+\infty & \text{for } x < a, \\
\Delta U \frac{x-a}{b-a} & \text{for } a \leq x \leq b, \\
\Delta U & \text{for } x > b.
\end{cases}
\]

We assume in the following that the ligand is bound in a binding pocket of depth \(b-a = 7 \text{ Å}\) involving a potential barrier \(\Delta U = 25 \text{ kcal/mol}\), similar to that of streptavidin (Chicotti et al., 1995). We also assume that the diffusion coefficient of the ligand is similar to the diffusion coefficient of the heme group in myoglobin (\(D = 1 \text{ Å}^2/\text{ns}\)) as determined from Mössbauer spectra (Nadler and Schulten, 1984).

To unbind from a protein the ligand has to move from \(a\), the minimum of the potential \(U(x)\), to \(b\), the maximum of \(U(x)\). The mean first passage time \(\tau(F)\) of such motion is (Izrailev et al., 1997)

\[
\tau(F) = 2\tau_d \delta(F)^{-2} \left[ e^{\delta(F)} - \delta(F) - 1 \right],
\]

where \(\tau_d = (b-a)^2/2D\) and

\[
\delta(F) = \frac{\Delta U - F(b-a)}{k_BT}.
\]

**Activated Regime** For a small applied force \(F\) corresponding to positive \(\delta\) and \(e^{\delta(F)} \gg \max(1, \delta)\), the mean time of unbinding is

\[
\tau_{act} \approx 2\tau_d \left[ \delta(F) \right]^{-2} e^{\delta(F)}.
\]

This result reflects the Kramers' relation (Gardiner, 1985). A millisecond time of unbinding, i.e., \(\tau_{act} \approx 1 \text{ ms}\), corresponds in this case to a rupture force of 155 pN. For such a force the potential barrier \(\Delta U\) is not abolished completely; in fact, a residual barrier of 9 kcal/mol is left for the ligand to overcome. The AFM experiments with an unbinding time of 1 ms are apparently functioning in the thermally activated regime.

**Diffusive Regime** In the case of a stronger force, such that \(F \approx \Delta U/(b-a)\) and \(\delta(F) \approx 0\), one obtains from (5)

\[
\tau_{diff} = \tau_d \left( 1 + \frac{\delta(F)}{3} \right).
\]

In this regime the applied force completely overwhelms the binding potential and the ligand is subject to free diffusion. The mean free passage time in this regime is equal to \(\tau_d\) and is on the order of 25 ns.
Drift Regime For still stronger forces corresponding to $\delta(F) < 0$ and $e^{\delta(F)} \ll 1 \ll |\delta(F)|$, using (5), one obtains

$$\tau_{\text{drift}} \approx 2\tau_d |\delta(F)|^{-1}. \quad (9)$$

This regime involves forces which are so strong that the ligand undergoes a drift motion governed by (3) in the limit that the fluctuating force $\sigma N(t)$ is negligible compared to the applied force. In this case a force of about 800 pN would lead to rupture within 500 ps.

These examples illustrate that SMD simulations operate in a different regime than existing micromanipulation experiments. Considerably larger forces (800 pN vs. 155 pN) are required to induce rupture, and the scaling behavior of the drift regime, characterized by (9), differs qualitatively from the activated regime as characterized by (7). Hence, SMD simulations of rupture processes can not be scaled towards the experimental force and time scales.

4.2 Unbinding Induced by Harmonic Springs

Assume now that the ligand is pulled by a harmonic spring, that is, subjected to an external force $F(x, t)$ of the form given by (1). The position of the ligand in the binding pocket fluctuates; according to the Boltzmann distribution of a harmonically bound particle, the position fluctuations associated with the spring are characterized by a variance $\delta x \sim (k_B T/K)^{1/2}$, and the corresponding variance in the applied force is related to $K$ through $\delta F \sim (K k_B T)^{1/2}$. A stiff spring confines the ligand to fluctuate in a small region of the binding pocket, so that only local properties of the binding potential are sampled, while the fluctuations of the force are large. For a soft spring, on the other hand, the ligand is able to fluctuate in a large region of the binding pocket, and the fluctuations of the force are small.

Stiff Spring For a stiff spring, satisfying $K \gg |d^2U/dx^2|$, under the overdamped condition assumed in (3) the average force measured by the spring can be expressed as

$$\bar{F} = \frac{dU}{dx} + \gamma v, \quad (10)$$

where $\bar{F} = K(\nu t - \bar{x})$, and where $\bar{F}$ and $\bar{x}$ denote a running time average of $F$ and $x$, respectively (Balsera et al., 1997). Equation (10) implies that for a stiff restraint the average applied force measures the local slope of the binding potential plus a frictional contribution that depends linearly on the pulling velocity. This dependence was observed in the MD simulations of the biotin-streptavidin complex (Grubmüller et al., 1996).
**Soft Spring** For a soft spring, no linear scaling of the rupture force with the pulling velocity should result (Izrailev et al., 1997). For millisecond unbinding times and soft springs employed in AFM experiments \((K \approx 6 \text{ pN/Å})\) (Florin et al., 1994), thermal fluctuations facilitate the exit of the ligand from the binding pocket of the protein. This means that the unbinding is thermally activated and the unbinding time \(\tau_R(F)\), according to Bell's relation, depends exponentially on the height of the energy barrier \(\Delta U^\dagger\) reduced by the applied force \(F\) (Bell, 1978). The rupture force \(F_{AFM}\), in this case, satisfies (Izrailev et al., 1997)

\[
F_{AFM} = \frac{\Delta U^\dagger}{L} - \frac{k_B T}{L} \ln \frac{\tau_R(F)}{\tau_d}.
\]  
(11)

The rupture force measured in AFM experiments is given, therefore, by the average slope of the energy profile minus a correction related to the effects of thermal fluctuations. Equation (11) demonstrates that the rupture force measured in AFM experiments grows linearly with the activation energy of the system (Chilcotti et al., 1995). A comparison of (10) and (11) shows that the unbinding induced by stiff springs in SMD simulations, and that induced by AFM differ drastically, and that the forces measured by both techniques cannot be readily related.

### 4.3 Reconstruction of the Potential of Mean Force

Measurement of the unbinding force should not be the only goal of SMD simulations. Even if the value of that force corresponded to experimental observations it would still not yield sufficient information to understand the dynamics of association/dissociation. Knowledge of the free energy profile of the system along the unbinding coordinate is required. Balsera et al. (1997) showed that it is possible, under the idealized conditions of (3), to reconstruct a one-dimensional potential of mean force from SMD simulation data. For a stiff spring the frictional contribution to the applied force can be explicitly discounted. One can simply integrate (10) to obtain an estimate of the potential \(U(x)\),

\[
\overline{U}(x) - U(0) = \int_0^x dx' (F - \gamma v).
\]  
(12)

Dissipation, however, imposes limits on how precisely the potential can be reconstructed. With the introduction of the work performed by the frictional force \(W_{fr}\), the uncertainty \(\sigma_U^2\) in the reconstructed potential \(U(x)\) can be presented as (Balsera et al., 1997)

\[
\sigma_U^2 = 2 k_B T W_{fr} \equiv 2 k_B T \gamma v x.
\]  
(13)

Thus, the uncertainty in the potential \(U\) is determined by the irreversible work done on the system. This irreversible work is proportional to the pulling...
velocity \( v \) and can be reduced with an increase in simulation time. In the avidin-biotin system, for example, the size of the avidin binding pocket is \( x \sim 10 \text{ Å} \). Assuming again a diffusion coefficient \( D = 1 \text{ Å}^2/\text{ns} \), simulation periods of 1 ns and 10 ns, corresponding to pulling velocities of the order of \( 10^{-2} \) and \( 10^{-3} \text{ Å}/\text{ps} \), yield \( \sigma_U \sim 8 \) and 3 kcal/mol, respectively. By contrast, an attempt to reconstruct the potential of mean force using a soft spring clearly fails.

5 Discussion

SMD is a novel approach to the study of the dynamics of binding/unbinding events in biomolecular systems and of their elastic properties. The simulations reveal the details of molecular interactions in the course of unbinding, thereby providing important information about binding mechanisms. The advantage of SMD over conventional molecular dynamics is the possibility of inducing relatively large conformational changes in molecules on nanosecond time scales. Other methods, such as umbrella sampling, free energy perturbation (McCammon and Harvey, 1987), and weighted histogram analysis (Kumar et al., 1992), aiming at the determination of the energy landscapes, typically involve small conformational changes and require extensive computations to achieve accuracy.

In cases where irreversible work done during unbinding can be attributed to a non-dispersive frictional force \( \gamma v \), a quantitative description of the thermodynamic potentials governing the binding and unbinding processes can be achieved by discounting the irreversible work from the calculated potential of mean force. However, the error in the reconstructed potentials is related to irreversible work done on the system and, therefore, may be unacceptably large. The estimate of the friction also presents a challenge, since it can be highly dispersive and may exhibit memory effects (Balsera et al., 1997).

Irreversibility of the unbinding process can also be accounted for by averaging over an ensemble of SMD trajectories according to the non-equilibrium equality for free energy differences (Jarzynski, 1997a; Jarzynski, 1997b) as described in the Introduction. This approach, however, requires averaging over multiple trajectories, and may be extremely sensitive to insufficient sampling of reaction pathways.

Irreversible work might also be discounted by forcing a conformational change in the system followed by the reverse conformational change, i.e., inducing hysteresis. Such an approach may yield a “model free” estimate of the irreversible work component from the hysteresis (Baljon and Robbins, 1996; Xu et al., 1996). Finally, lengthening the simulation time decreases the amount of irreversible work and the simulated process could, ideally, reach quasi-equilibrium in the limit of very long simulation times.

The simulations of the avidin-biotin complex (Izrailev et al., 1997) showed that a major difficulty involved in studies of the binding and flexibility of
macromolecules is the long time scale of motions such as the fluctuations of the avidin loop discussed above. These fluctuation times, ranging from several nanoseconds to seconds, are beyond the reach of SMD simulations that can presently be realized in a feasible amount of time (Balsera et al., 1996).

Solvation is likely to influence protein-ligand binding and, hence, the forces measured in SMD simulations. During the extraction of retinal from bacteriorhodopsin (see Fig. 3) water facilitated the breaking of a hydrogen bond between retinal and Lys216 (Isralevitz et al., 1997). In the simulations of the avidin-biotin complex, placement of several water molecules near the exit from the binding pocket reduced the measured binding force (see Fig. 2). In the simulations of the streptavidin-biotin complex (Grubmüller et al., 1996) the binding pocket was exposed to solvent due to exclusion of the adjacent streptavidin monomer. This allowed water molecules to enter the binding pocket and participate in breaking hydrogen bond networks between the ligand and the protein during the unbinding. The issue of how water molecules participate in the process of protein-ligand dissociation remains unclear and should be further investigated. For example, one may add water molecules to the binding pocket in the course of SMD simulations, as the retracting ligand frees up space in the binding pocket (Resat et al., 1996).

Binding and unbinding of non-covalently attached biomolecules are at the heart of many important processes and are the target of experimental investigations. SMD may serve to interpret measurements and suggest new experiments. The rapidly growing computer power available for simulations and increasing time resolution of experimental techniques will provide the basis for further advances in the method and will help bridge the gap in time scales between computer simulation and experiment.

6 Acknowledgments

The authors thank Y. Oomo and M. Balsera for invaluable contributions to the joint development of SMD and J. Gullingsrud for many suggestions in the preparation of the manuscript. Images of molecular systems were produced with the program VMD (Humphrey et al., 1996). This work was supported by grants from the National Institute of Health (PHS 5 P41 RR05969-04), the National Science Foundation (BIR-9318159, BIR 94-23827 (EQ)), and the Roy J. Carver Charitable Trust.

References


