Molecular dynamics simulations of palmitate entry into the hydrophobic pocket of the fatty acid binding protein

Yossi Tsfadia*, Ran Friedman*, Jonathan Kadmon,
Anna Selzer, Esther Nachliel, Menachem Gutman

Department of Biochemistry, Tel Aviv University, 69978 Ramat Aviv, Tel Aviv, Israel

Abstract The entry of substrate into the active site is the first event in any enzymatic reaction. However, due to the short time interval between the encounter and the formation of the stable complex, the detailed steps are experimentally unobserved. In the present study, we report a molecular dynamics simulation of the encounter between palmitate molecule and the Toad Liver fatty acid binding protein, ending with the formation of a stable complex resemblance in structure of other proteins of this family. The forces operating on the system leading to the formation of the tight complex are discussed.
© 2007 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

Keywords: Fatty acid binding proteins; Molecular dynamics; Molecular modeling; Palmitic acid

1. Introduction

The interaction of substrate with the active site has been, for years, the 'Holy Grail' of Enzymology. As previous studies were focused on the interactions between the substrate and the active site, they were mostly concerned with the ligand–site interactions rather than with the dynamics of entry. In many enzymes, the binding site is not fully exposed to the bulk, thus rendering the entry reaction an intricate waltzing motion of the two partners that ends with the nesting of the substrate in the site. Such motions necessitate a coordinated flexing of the substrate molecule together with conformational fluctuations of the protein; both are driven by the free energy gradient along the reaction coordinate. The substrate entry is too fast to be experimentally recorded and bears no clear detectable signal. For this reason, we applied in the present study a computational approach, carrying out a molecular dynamics (MD) simulation of a medium-sized substrate molecule (palmitate), and followed its penetration into a specific cavity, identified as the binding site by published crystallographic studies. Unlike previous study, that attempted to simulate ligand exchange with external bias [1], our simulation was carried out under no external force, with explicit water and in the presence of phys-

iological ionic strength solution, thus suggesting that the scenario is compatible with the in vivo process.

Fatty acids (FAs) are a major energy source and are structural elements of the phospholipids. Yet, due to their detergent-like properties, they should be avoided in the cytoplasm. For this reason, free FA transfer is executed by a widely-distributed protein family called 'Fatty Acid Binding Proteins' [2–5]. These proteins share a 'β-clam' structure made of a β-barrel and capped by two α-helices. The β strands are linked each to the other by hydrogen bond, except for β-strands D and E, creating a flexible section proposed to function as the portal region, through which the fatty acids penetrate into the binding cavity [6].

Crystallization of these proteins, either with native substrates or with substrate analogs indicated that the fatty acids are embedded inside the protein in two main configurations. In one case, the aliphatic tail is inserted into the site, while the carboxylate is exposed to the bulk [2,3,7]. In the other configuration, the carboxylate moiety is embedded inside the protein, being stabilized by a combination of hydrogen bonds and electrostatic interactions, while its hydrophobic tail protrudes towards the bulk through the portal region [2–5,8].

The lack of direct observations as to how the substrate penetrates the protein had led to the application of MD simulations of the protein–water–substrate system [9–13]. The strategy adopted in the present study was to simulate the protein in explicit water and physiological ionic strength, and to monitor the interactions between a few palmitate molecules added to the system with the protein. The selected protein, Toad liver fatty acid binding protein (L-FABP, PDB code 1P6P), was chosen because the crystal structure of the apo-protein retains the binding site that is partially exposed to the bulk [14]. We have carried out extensive MD simulations, accumulating more than 250 ns of independent runs that lasted 8–20 ns each. In all these cases, the encounter of the fatty acids with the protein led to a rapid adsorption to the surface of the protein, mostly in the vicinity of the portal region. The adsorbed palmitate molecules were noticed to scan the surface of the protein, with many attempts of penetration. In most cases, the penetration was only partial, where either the carboxylate moiety or most of the aliphatic tail was exposed to the bulk. In this report, we describe a case in which the palmitate anion penetrated deep into the protein, reaching a configuration similar to those determined by X-ray crystallography of other members of the family; as the adipocyte lipid binding protein (PDB codes: 1LIE, 1LIC, 1LID) and proteins from other tissues (PDB codes: 1B56, 1HMS, 1FE3), where crystallographic

*Corresponding authors. Fax: +972 3 6409875.
E-mail addresses:* yossit@tauex.tau.ac.il (Y. Tsfadia),
r.friedman@bioc.unizh.ch (R. Friedman).

studies revealed that the negatively charged carboxylate was located inside the cavity and stabilized by salt bridges and hydrogen bonds.

The ability to follow, *in silico*, the penetration event enables the observation of the intricate motions of the reactants and to gauge the contribution of the various forces, namely the electrostatic interactions, Lennard-Jones (LJ) attraction and solvation.

A detailed report, summing many simulations will be published elsewhere (Tsfadia, manuscript in preparation).

2. Methods

2.1. Molecular dynamics simulations

The MD simulations were performed using the GROMACS 3.2.1 software [15–17], with the GROMOS 53A6 force field [18]. The calculations were carried out using the structure of the Toad basic liver fatty acid binding protein (PDB code 1P6P) [14], that was downloaded from the PDB [19]. The simulations were carried out in 100 mM NaCl solution and 4 palmitate anions to facilitate the encounter between the substrate and the protein. The palmitate molecules were randomly placed within the water box, but care was taken to ensure that the initial position of the palmitate molecules was out of the protein structure. The formal concentration of the palmitate was 22 mM, well above the CMC of this compound (5.8 mM) [20]; indeed in few of the simulations, an initiation of micelle formation were noticed. Still, as the surface provided by the protein exceeds that of the fatty acids, in most cases the adsorption of the ligands to the protein was faster than their self aggregation. The protein and the ligands were embedded in a truncated octahedral box containing SPC model water [21] that extended to

at least 1.5 nm between the protein and the edge of the box. The total number of water molecules was approximately 10^4; 16 Na^+ and 20 Cl^- ions were added to the system by replacement of water molecules in random positions, thus making the whole system neutral. A PDB file of the palmitate molecule and the partial charges of the atoms are provided in Supplementary material.

Prior to the dynamics simulation, the internal constraints were relaxed by energy minimization, followed by 40 ps equilibration under position restraints of the carbon backbone atoms through a harmonic forces constant of 1000 kJ nm^{-2}. The first 200 ps of the unrestrained simulations were not used for the analysis. During the MD runs, the LINCS algorithm [22] was used to constrain the lengths of covalent bonds; the waters were constrained using the SETTLE algorithm [23]. The time step for the simulations was 2 fs. The simulations were run under constant pressure and temperature, using Berendsen's coupling algorithm [24] ($P = 1$ bar; $\tau_P = 0.5$ ps; $T = 300$ K; $\tau_T = 0.1$ ps). VDW forces were treated using a cutoff of 1.2 nm. Long range electrostatic forces ($r > 1.2$ nm) were treated using the particle mesh Ewald method [25]. The duration of the simulations varied between 8 and 20 ns. The simulations differed in the initial location of the palmitate molecules and in the seed number, ensuring that each run was an independent calculation.

2.2. Visual presentation

The figures were created by the VMD program [26].

3. Results and discussion

3.1. Dynamic observations of palmitate entry into the protein

Fig. 1 presents two sets of snapshots taken during the entry of a single palmitate ion to the protein. The upper frames is a
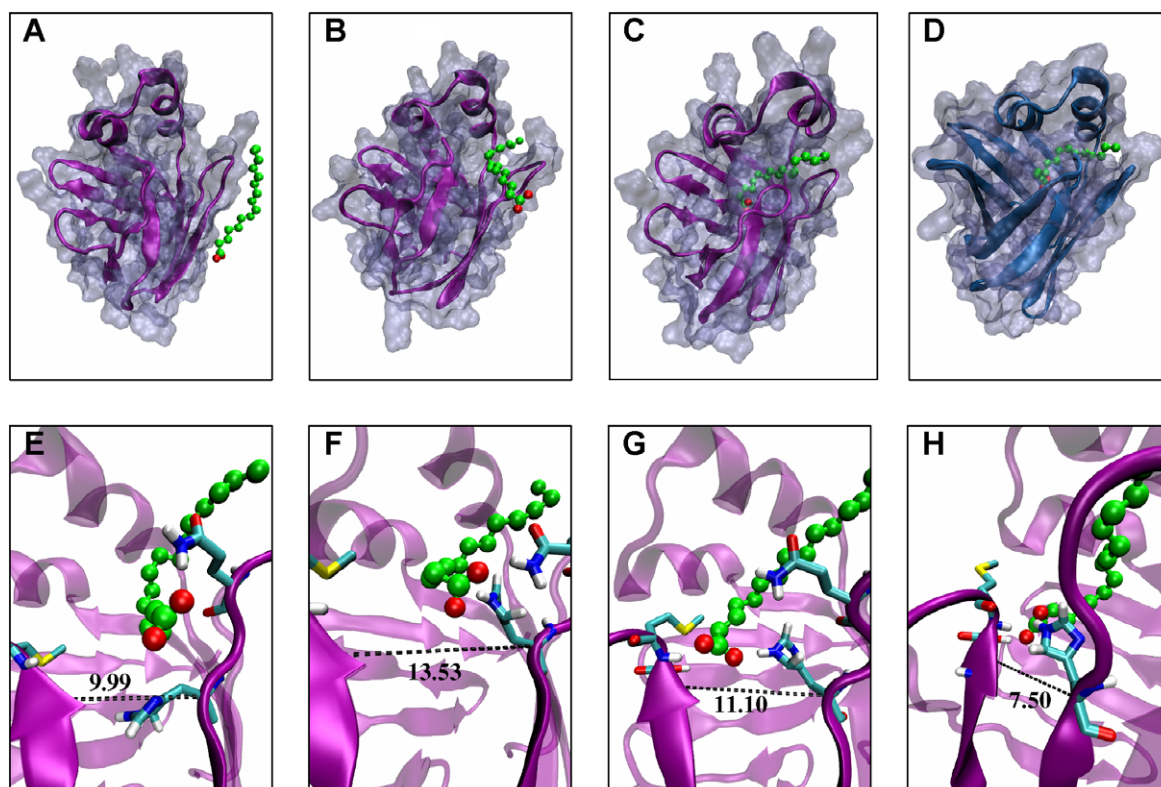

Fig. 1. Snapshots taken during the entry of a palmitate anion to the cavity of the Toad-liver FABP. Upper panel: the frames A–C were taken from a MD simulation at 200, 450, and 8000 ps. Frame D is the crystal structure of the adipocyte lipid binding protein (ALBP) that was crystallized with palmitate anion. In this structure, the α carbon trace is in blue. Lower panel: a closer look at the interaction between the carboxylate with the amino acids at the portal region. Frames E–H were taken, respectively, at 600, 730, 850 and 1550 ps. The β strands D and E are colored with opaque purple; the distance between them is measured in Å.

sequence of snapshots, where four palmitate molecules were added, at random locations in the bulk some 7 Å from the protein's surface (for sake of clarity all other palmitate molecules are not presented). Within the first 200 ps of the simulation (A), one of the fatty acid molecules diffused and adsorbed onto the protein's surface at the vicinity of the portal region, represents the beginning of the binding process. At $t = 450$ ps (B), the palmitate is already wedging itself into a cleft between the β strands D and E and the helix that caps them. The aliphatic section seems to pry the cleft open, while the carboxylate moiety retains a full contact with the water matrix. The penetration process lasted for about 1 ns, and at $t \sim 1600$ ps the complex between the protein and the palmitate molecule reached a stable conformation, which remains almost invariable until the end of the simulation at $t = 8$ ns (C). This final configuration, is comparable to the crystal structure (1LIE) of the Holo-protein of the adipocyte lipid protein (D) [27]. During the simulation of the penetration process, we have noticed an impressive decrease of the number of water molecules present inside the cavity, from ~ 25 molecules at the starting point down to 12–14 molecules at the final state with the bound ligand, in accordance with experimentally estimations [3].

The lower row of frames (E–H) presents a detailed view of the penetration process, revealing how a sequence of electrostatic interactions coupled with the random motion of the residues drives the entry of the carboxylate headgroup into the cavity. At $t = 610$ ps (E), the carboxylate head group of the palmitate interacts with Q56 at the surface of the protein through a hydrogen bonding, in accordance with the path deduced from experimental results [2]. Some 90 ps later ($t \sim 700$ ps, F), the carboxylate is pulled deeper into the cavity and stabilized both by Q56 and H58. At ~ 850 ps (G), new

hydrogen bonds are formed; one with the backbone NH group of M73 (2.82 Å) and the second one with the OH moiety of S72 (2.63 Å), both on β-strand E. Please note that these steps are associated with a widening of the cleft between strands D and E. The last structure, detailed in frame H (t-1600 ps), is of an apparent stable configuration, where negatively charged oxygens are stabilized by hydrogen bonds with H58 on strand D and with S72 and M73 on strand E. The distance between the strands forming the portal region had been closed to 7.5 Å. The O1 atom of the carboxylate also forms a hydrogen bond through a water molecule ($O \rightarrow H_2O$ 2.77 Å) to NH_2 of R120 ($H_2O \rightarrow NH_2$ 3.04 Å, not shown).

This shape set of hydrogen bond interactions is comparable to the interactions recorded for the crystal structure of the palmitate-ALBP (PDB code: 1LIE), where the carboxylate interacts with the hydroxyl of Y128 (2.69 Å), the NH_2 of R126 (2.98 Å) and through a water molecule ($O \rightarrow H_2O$ 2.92 Å) with the NH_2 atom of R106 ($H_2O \rightarrow N$ 2.81 Å). It must be recalled that the two proteins are not identical in sequence, yet one can find similarity between the interactions in the case of the complex at 8 ns of the simulation and those of the Holo-protein ALBP in its crystalline form. On the basis of this similarity, we concluded that, at the end of the simulation time, the binding of the palmitate to Toad L-FABP resembles the crystal structure of Holo ALBP.

3.2. The driving forces for the binding

The penetration of the substrate protein was followed up by calculating the electrostatic and LJ potentials for the substrate–protein interactions, as presented in Fig. 2A. A short time after the initiation of the simulation, the system gains stability due a rapid decrease of the LJ potential. Only later (~ 500 ps after initiation of the simulation), when the LJ
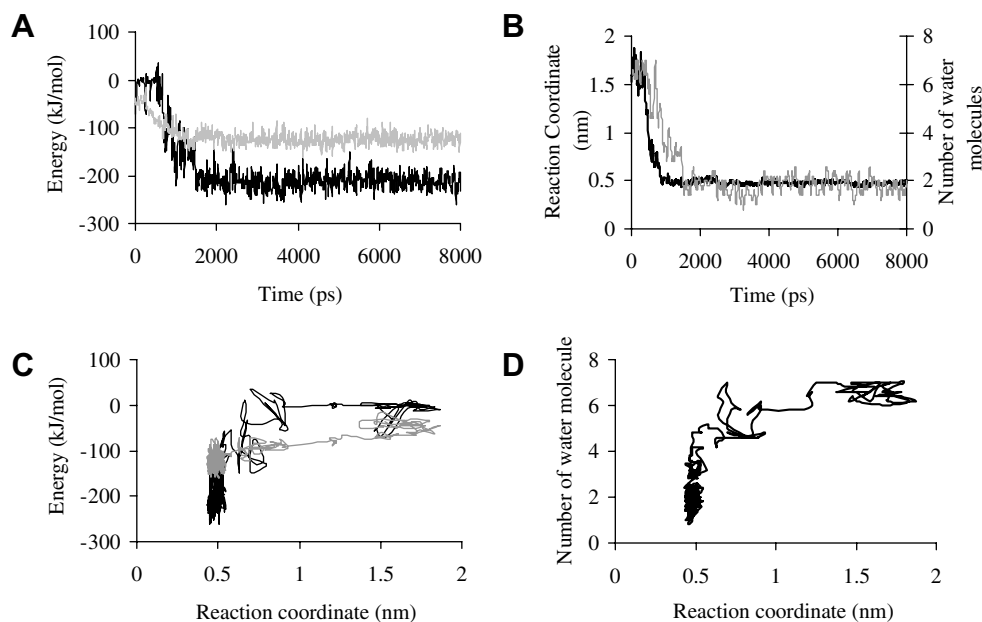

Fig. 2. Quantitative description of the reaction between the palmitate anion and the Toad-liver fatty acid binding protein. (A) The interaction-energy between the substrate and the protein, is presented by the short range Coulomb potential (black) and the LJ interactions (grey), as a function of the simulation time. (B) The progression of the reaction with time, depicted as the variation of reaction coordinate (RC_t) versus the simulation time (black line with the scale on the left ordinate). The grey line (right ordinate) represents the number of water molecules 0.4 nm from the carboxylate. (C) The stabilization of the system as a function of the reaction coordinates. The colors of the lines are as in frame A. (D) The variation of the number of water molecules that form a hydrogen bond with the carboxylate moiety, presented with respect to the reaction coordinate.

attraction progressed almost halfway, the electrostatic attraction comes into play.

Relation of the observed events with the time axis, as in Fig. 2A, emphasizes the temporal sequentiality of the process. The reaction was also expressed by an internal frame of reference (the reaction coordinate), by which the progression is related with the relative location of the ligand with respect to the binding site. The reaction coordinate (RC_t) is defined as the average distance between the carboxylate moiety of the palmitate and the center of mass of the three residues that tightly interact with the carboxylate at the later phase of the simulation, namely; H58, S72 and M73. At the early phase (Fig. 2B), the fatty acid is still at its initial random location, $RC_t \sim 1.7$ nm, while in the final one the value $RC_t \sim 0.5$ nm remains stable for $\sim$6 ns.

Frame C presents the gained stability (electrostatic and LJ interactions) as a function of the progression along the reaction coordinate. The wobbling lines represent the values of the electrostatic (black) and LJ (grey) potentials as they vary when the carboxylate is penetrating into the cavity. The two traces, appearing in frame C, are very tortuous, thus indicating that the progression into the protein is a continuous search in the conformation space. In the early phase, $RC_t \sim 1.5$ nm, the palmitate molecule is still out of the protein and samples the space by a Brownian motion that drives the substrate towards contact with the protein. During this phase, there is hardly a change in the system's energy. The next phase of the reaction is steady progression towards $RC_t \sim 0.8$ nm, which is mostly driven by the steady decrease of the LJ potential. At this phase the aliphatic tail penetrates the hydrophobic cavity, while the carboxylate retains a full state of solvation (see also Fig. 1B). The reaction then proceeds by another wobbling phase, where RC_t varies around 0.7 nm, that ends by a significant increase in the electrostatic potential, corresponding to the entry of the carboxylate into the protein, along with a subsequent reduction of the number of water molecules solvating the carboxylate from 6 to 4 (Fig. 2B and 2D, grey). Finally, there is a convergence of both electrostatic and LJ potentials into a time invariant value that lasts until the end of the simulation time, with no progression along the reaction coordinate ($RC_t = 0.5 \pm 0.05$ nm). At that phase of the reaction, the carboxylate had lost most of its solvation water. This phase is associated with a very minor molecular motion, thus representing optimization of the stabilizing energy through local rearrangements of the atoms near the carboxylate's binding site.

In general, the insertion of a negative charge (carboxylate) to the low dielectric constant matrix is associated with a penalty of the Born energy and the loss of the solvation shell of the charge. The penetration can be made feasible by compensating interactions, such as electrostatic interactions, with charged residues inside the protein or by retaining some of the solvation water molecules, as in the present case. To facilitate the insertion, the carboxylate moiety retains its solvation shell as it progresses along the reaction coordinate, lowering it to only two water molecules, when the system appears to reach its apparent stable configuration. This feature is depicted in Fig. 2 frame B, grey line, and frame D, where the number of water molecules that are in hydrogen bond with the carboxylate is related to the propagation in time and along the RC_t. The two water molecules that solvate the carboxylate in its final stage of penetration differ in their interaction with the carboxylate; one of them is bridging, by hydrogen bond, the carboxylate with the positive charge of R120. This water molecule exhibits a very slow exchange with other water molecules, having a residence time longer than 2000 ps. The other water molecule is less dedicated to the site and is routinely replaced by other water molecule present inside the protein.

The interaction of the ligand with the protein is much more complex than that expressed by the electrostatic and LJ potentials. The actual ΔG is affected by many other terms like $\Delta G_{solvation}$, and the $T\Delta S$ terms for the protein, the ligand and the solvent. For these reasons, the energy terms appearing in Fig. 2 reflect the trends in the operating forces, rather than the sole contributors of the free energy of the reaction. The interplay between the two potentials reveals how the charged and aliphatic domains of the ligand contribute, each at a time to the penetration into the binding cavity.

This study reports, for the first time, how a large fully-solvated substrate enters a specific binding site located inside a protein molecule, driven by intricate random walk and steering forces that operate almost simultaneously. From the present simulation, and others that will be reported elsewhere, it appears that the reaction consists of three sequential steps; the first is a diffusive search in the solution, ending by an encounter the protein's surface. Following the adsorption, the ligand is probing the surface of the protein, and the third step is a fruitful penetration. The rate-limiting step of the reaction under physiological conditions, where the concentration of the free fatty acids is low, is probably the encounter step. Under conditions in which the adsorption prevails, the search for the open portal state and getting to a proper orientation of the ligand with respect to the site can lead to a series of abortive entry attempts. However, once the conditions are proper, the entry itself is a very rapid event.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.febslet.2007.02.033.

References

[1] Zanotti, G., Feltre, L. and Spadon, P. (1994) A possible route for the release of fatty-acid from fatty-acid-binding protein. Biochem. J. 301, 459–463.

[2] Stewart, J. (2000) The cytoplasmic fatty-acid-binding proteins: thirty years and counting. Cell. Mol. Life Sci. 57, 1345–1359.

[3] Zimmerman, A.W. and Veerkamp, J.H. (2002) New insights into the structure and function of fatty acid-binding proteins. Cell. Mol. Life Sci. 59, 1096–1116.

[4] Wolfrum, C. and Spener, F. (2000) Fatty acids as regulators of lipid metabolism. Eur. J. Lipid Sci. Technol. 102, 746–762.

[5] Haunerland, N.H. and Spener, F. (2004) Fatty acid-binding proteins – insights from genetic manipulations. Prog. Lipid Res. 43, 328–349.

[6] Richieri, G.V., Ogata, R.T. and Kleinfeld, A.M. (1996) Kinetics of fatty acid interactions with fatty acid binding proteins from adipocyte, heart, and intestine. J. Biol. Chem. 271, 11291–11300.

[7] Thompson, J., Winter, N., Terwey, D., Bratt, J. and Banaszak, L. (1997) The crystal structure of the liver fatty acid-binding protein. A complex with two bound oleates. J. Biol. Chem. 272, 7140–7150.

[8] Sacchettini, J.C., Gordon, J.I. and Banaszak, L.J. (1989) Crystal-structure of rat intestinal fatty-acid-binding protein – refinement and analysis of the *Escherichia-coli*-derived protein with bound palmitate. J. Mol. Biol. 208, 327–339.

[9] Bakowies, D. and van Gunsteren, W.F. (2002) Simulations of apo and holo-fatty acid binding protein: structure and dynamics of protein, ligand and internal water. J. Mol. Biol. 315, 713–736.

[10] Friedman, R., Nachliel, E. and Gutman, M. (2005) Molecular dynamics simulations of the adipocyte lipid binding protein reveal a novel entry site for the ligand. Biochemistry 44, 4275–4283.

[11] Friedman, R., Nachliel, E. and Gutman, M. (2006) Fatty acid binding proteins: same structure but different binding mechanisms? Molecular dynamics simulations of intestinal fatty acid binding protein. Biophys. J. 90, 1535–1545.

[12] Woolf, T.B. (1998) Simulations of fatty acid-binding proteins suggest sites important for function. I. Stearic acid. Biophys. J. 74, 681–693.

[13] Woolf, T.B. and Tychko, M. (1998) Simulations of fatty acid-binding proteins. II. Sites for discrimination of monounsaturated ligands. Biophys. J. 74, 694–707.

[14] Di Pietro, S.M., Corsico, B., Perduca, M., Monaco, H.L. and Santome, J.A. (2003) Structural and biochemical characterization of toad liver fatty acid-binding protein. Biochemistry 42, 8192–8203.

[15] Berendsen, H.J.C., Vanderspoel, D. and Vandrunen, R. (1995) Gromacs – a message-passing parallel molecular-dynamics implementation. Comput. Phys. Commun. 91, 43–56.

[16] Lindahl, E., Hess, B. and van der Spoel, D. (2001) Gromacs 3.0: a package for molecular simulation and trajectory analysis. J. Mol. Mod. 7, 306–317.

[17] Van der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A.E. and Berendsen, H.J.C. (2005) GROMACS: fast, flexible, and free. J. Comput. Chem. 26, 1701–1718.

[18] Oostenbrink, C., Villa, A., Mark, A.E. and Van Gunsteren, W.F. (2004) A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. J. Comput. Chem. 25, 1656–1676.

[19] Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N. and Bourne, P.E. (2000) The protein data bank. Nucleic Acids Res. 28, 235–242.

[20] Mehrotra, K.N. and Upadhyaya, S.K. (1989) Ultrasonic measurements and other allied parameters of praseodymium and neodymium palmitates in mixed organic solvents. Colloid Polym. Sci. 267, 741–747.

[21] Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F. and Hermans, J. (1969) Interaction models for water in relation to protein hydration. Nature 224, 175–177.

[22] Hess, B., Bekker, H., Berendsen, H.J.C. and Fraaije, J.G.E.M. (1997) LINCS: a linear constraint solver for molecular simulations. J. Comp. Chem. 18, 1463–1472.

[23] Miyamoto, S. and Kollman, P.A. (1992) SETTLE: an analytical version of the SHAKE and RATTLE algorithms for rigid water models. J. Comp. Chem. 13, 952–962.

[24] Berendsen, H.J.C., Postma, J.P.M., DiNola, A. and Haak, J.R. (1984) Molecular dynamics with coupling to an external bath. J. Chem. Phys. 81, 3684–3690.

[25] Darden, T., York, D. and Pedersen, L. (1993) Particle mesh Ewald: An $N - ;\log(N)$ method for Ewald sums in large systems. J. Chem. Phys. 98, 10089–10092.

[26] Humphrey, W., Dalke, A. and Schulten, K. (1996) VMD: visual molecular dynamics. J. Mol. Graph. 14, 33–38.

[27] LaLonde, J.M., Bernlohr, D.A. and Banaszak, L.J. (1994) X-ray crystallographic structures of adipocyte lipid-binding protein complexed with palmitate and hexadecanesulfonic acid. Properties of cavity binding sites. Biochemistry 33, 4885–4895.