Eur Biophys J (2011) 40:181–194
DOI 10.1007/s00249-010-0638-3

ORIGINAL PAPER

Steered molecular dynamics simulations of ligand–receptor interaction in lipocalins

Janne Kalikka · Jaakko Akola

Received: 10 June 2010 / Revised: 2 October 2010 / Accepted: 13 October 2010 / Published online: 13 November 2010

Abstract Retinol binding protein (RBP) and an engineered lipocalin, DigA16, have been studied using molecular dynamics simulations. Special emphasis has been placed on explaining the ligand–receptor interaction in RBP–retinol and DigA16–digoxigenin complexes, and steered molecular dynamics simulations of 10–20 ns have been carried out for the ligand expulsion process. Digoxigenin is bound deep inside the cavity of DigA16 and forms several stable hydrogen bonds in addition to the hydrophobic van der Waals interaction with the aromatic side-chains. Four crystalline water molecules inside the ligand-binding cavity remain trapped during the simulations. The strongly hydrophobic receptor site of RBP differs considerably from DigA16, and the main source of ligand attraction comes from the phenyl side-chains. The hydrogen bonds between digoxigenin and DigA16 cause the rupture forces on ligand removal in DigA16 and RBP to differ. The mutated DigA16 residues contribute approximately one-half of the digoxigenin interaction energy with DigA16 and, of these, the energetically most important are residues His35, Arg58, Ser87, Tyr88, and Phe114. Potential "sensor loops" were found for both receptors. These are the outlier loops between residues 114–121 and 63–67 for DigA16 and RBP, respectively, and they are located near the entrance of the ligand-binding cavity. Especially, the residues Glu119 (DigA16) and Leu64 (RBP) are critical for sensing. The ligand binding energies have been estimated based on the linear response approximation of binding affinity by using a previous parametrization for retinoids and RBP.

Keywords Lipocalins · Molecular dynamics · Steered molecular dynamics · Retinol binding protein · Digoxigenin · Ligand–receptor interaction

Electronic supplementary material The online version of this article (doi:10.1007/s00249-010-0638-3) contains supplementary material, which is available to authorized users.

J. Kalikka · J. Akola (✉)
Nanoscience Center, Department of Physics,
University of Jyväskylä, P.O. Box 35, Jyväskylä 40014, Finland
e-mail: jaakko.akola@phys.jyu.fi

J. Akola
Department of Physics, Tampere University of Technology,
P.O. Box 692, Tampere 33101, Finland

Introduction

The lipocalins are a family of small, robust, secretory, and diverse extracellular proteins of 160–180 amino acids, which are able to bind small hydrophobic guest molecules (Flower 1994, 1996; Schlehuber and Skerra 2005). Lipocalins are commonly found in vertebrates and in some bacteria, and there are at least ten different lipocalins present in the human body, e.g., plasma retinol binding protein (RBP), tear lipocalin, α_1-acid glycoprotein, and apolipoprotein D (Breustedt et al. 2006). Unlike most protein families, lipocalins have little sequence similarity, but they share common short characteristic and conserved sequence motifs in tertiary structure that can be used to identify membership of the family (Flower 1996). Lipocalins have a wide variety of functions, the most remarkable of which is the ability to bind small hydrophobic ligand molecules that are of low solubility and/or chemically sensitive (e.g., vitamins, steroids, odorants, and metabolic products) (Schlehuber and Skerra 2005). Several lipocalin crystal structures have been measured, and the ligand binding site is enclosed inside an eight-stranded continuously hydrogen-bonded antiparallel β-barrel. This

is the characteristic folding motif of this protein family. Apart from the host–guest complexation that results in the capability to transport and store compounds ("transport protein"), lipocalins have also other biological functions as they bind to specific cell-surface receptors, form complexes with soluble macromolecules, and act even as enzymes in prostaglandin synthesis (Flower 1996).

The most famous lipocalin is human plasma RBP, which binds the poorly soluble and oxidation-prone vitamin A (retinol). Retinol has several important roles in biological processes, such as vision and cell growth/differentiation, and it is stored in the liver and secreted as a retinol–RBP complex for target cells. Upon complexation with RBPs (plasma or cellular), retinol is bound deep inside the RBP cavity, and its stability increases considerably in comparison with the aqueous environment. The crystal structures of human, pig, bovine, and chicken RBPs have been measured (Calderone et al. 2003; Cowan et al. 1990; Newcomer et al. 1984; Zanotti et al. 1993, 1998, 2001). Of particular interest is the work of Calderone et al. (2003), who report high-resolution crystal structures of bovine RBP at different pH values. At neutral pH, the nearly atomic resolution reveals that retinol adopts an orientation where its β-ionone ring is innermost, the polyene chain is fully extended (*trans*-retinol), and the hydroxyl end group points towards the solvent at the entrance of the cavity. Furthermore, the work by Calderone et al. (2003) provides evidence for pH-induced modifications in the RBP structure that has important biological implications: The targeted delivery of retinol requires its release in the immediate vicinity of the plasma membrane, and this can be achieved only by a significant structural modification in the RBP structure. Acidic denaturation of RBP may be the mechanism that is required in order to facilitate retinol release.

Recently, lipocalins have attracted interest as possible scaffolds in protein engineering, and a new class of artificial proteins, "anticalins," can recognize specific molecules via host–guest complexation (Schlehuber and Skerra 2005; Skerra 2007, 2008). As lipocalins naturally bind and transport molecules of low solubility or high chemical instability, their tailored modifications have potential for medical applications as transport vessels for prescribed medicine molecules. The anticalins provide an alternative to recombinant antibody fragments, and they possess several advantages, especially for binding of larger haptens and—most recently—for the CTLA-4 protein (Schönfeld et al. 2009). A prototype anticalin is DigA16 (Korndörfer et al. 2003; Schlehuber et al. 2000; Schlehuber and Skerra 2002), which was structurally reshaped from the bilin-binding protein (BBP) of *Pieris brassicae* in order to specifically form complexes with digoxigenin and digitoxigenin. These guest molecules are medically important, as they serve either as poisons or drugs with therapeutic value (depending on the dose), and the digoxigenin group is also being used in nonradioactive biolabeling of several biomolecules (Schlehuber and Skerra 2005). The strong binding activity of the DigA16 variant for digoxigenin ($K_D = 30.2 \pm 3.6\,\text{nM}$) was achieved by modifying four loops in the entrance of the BBP ligand-binding pocket by targeted random mutagenesis resulting in substitution of 17 amino acid residues and one fixed substitution prior to the randomization (Schlehuber et al. 2000). The steroid ligand is buried deep inside the pocket (by 95%) with the lactone group innermost, and the specificity for digoxigenin is maintained by hydrogen-bonding network and shape complementarity (Korndörfer et al. 2003). Interestingly, DigA16 does not recognize another steroid ouabain, which has three additional OH groups and one shifted OH group (in the steroid rings) in comparison with digoxigenin (Schlehuber et al. 2000; Schlehuber and Skerra 2002). The selective steroid binding in DigA16 differs from that in the antidigoxigenin antibodies where the crystal structures show that only the lactone ring (tip) of the guest molecule is in tight contact with the binding site (Jeffrey et al. 1993, 1995).

In this work, we have performed classical molecular dynamics (MD) simulations for two related lipocalin proteins, bovine RBP and DigA16, whose crystal structures have been reported at high resolution (Calderone et al. 2003; Korndörfer et al. 2003). Lipocalins have been studied previously by MD methods, for instance, in the works of Beaufays et al. (2008) and Chau (2004), where the former reported the binding of leukotriene B4 in the tick salivary lipocalin ("scavenger protein") and the latter paid special attention to the behavior of individual water molecules around (and inside) the ligand-binding cavity of RBP. In addition to standard MD simulations of 30 ns that probe the static structure of the protein and its dynamics (fluctuations), we have investigated ligand expulsion processes by performing steered molecular mechanics (SMD) simulations of 10–20 ns. SMD is one of the standard techniques that have been developed to study structural changes, such as ligand unbinding or protein unfolding, within timescales that are accessible to computer simulations (Gunnerson et al. 2009; Hytönen and Vogel 2008; Isralewitz et al. 2001a, b; Niu et al. 2009; Park and Schulten 2004; Peräkylä 2009; Singh et al. 2008). Our atomistic-scale simulations with explicit solvent provide detailed information of the ligand–receptor interaction in RBP and DigA16, which involves ligand sensing by particular amino acid residues and subsequent binding.

Methods

Coordinates of the DigA16 and RBP protein structures and ligands were obtained from the RCSB Protein Data Bank (Berman et al. 2000). Both entries were partial in terms of the completeness of the amino acid sequence. The holo-DigA16 entry was missing a total of 26 residues, and the simulated structure was constructed from parts as follows: Coordinates for residues 1–4 were taken from the apo-form structure of DigA16 (Protein Data Bank entry 1KXO, Korndörfer et al. 2003), and for the coordinates of residues 166–184 we used the FluA protein (Protein Data Bank entry 1T0V, Mills et al. 2009), which was recently engineered from the same BBP as DigA16. Both tails were added to the holo-form structure of DigA16 (Protein Data Bank entry 1LKE, Korndörfer et al. 2003) and aligned properly. The middle loop (residues 118–122) had been added computationally as a part of the original crystallographic refinement by using Refmac 5.2.0005 software with maximum likelihood as the refinement target (Murshudov et al. 1997). The holo-RBP entry (Protein Data Bank entry 1KT7, Calderone et al. 2003) does not include coordinates for the last eight residues in the C-terminus, and we could not find measured coordinates for any other protein with an identical sequence. Correspondingly, these amino acid residues were not included in the holo-RBP simulations.

Water solvate (TIP3P force field) was added around the proteins. The structure of the added water was based on a snapshot of an equilibrated bulk water sample at room temperature (300 K). Any crystallized water molecules that were reported in the protein structure file were treated as part of the protein structure and included "as is" in the simulation structure. Counterions were added in the solvate in order to balance the net charge of the whole simulation box to neutral; 5 Na^+ ions were added in both the DigA16 and RBP systems. The DigA16 structure includes two sulfur bridges (Cys8–Cys115 and Cys42–Cys170), and the RBP structure has three sulfur bridges (Cys4–Cys160, Cys70–Cys174, and Cys120–Cys129). These were treated as covalent bonds during the simulations. The box sizes of the periodic systems were $82 \times 61 \times 62$ Å with 30,844 atoms for DigA16, and $73 \times 63 \times 68$Å with 30,814 atoms for RBP.

The MD and SMD simulations used the AMBER ff03 force field parameters [also known as the Duan et al. (2003) force field], and the atomic charges for the ligands were calculated separately using Gaussian03 revision D.02 (Frisch et al. 2004) with the HF/6-31G* method. The charges were fitted to match with the electrostatic potential by using the Merz–Singh–Kollman scheme (Besler et al. 1990; Singh and Kollman 1984). The simulations were carried out by using the NAMD2.6 program package (Phillips et al. 2005) in an *NPT*-ensemble with periodic boundary conditions. A time step of 0.5 fs was used, and covalently bonded hydrogen atoms were constrained to equilibrium positions with the ShakeH algorithm. Non-bonded interactions were evaluated once every 1 fs, and we used a 12 Å cutoff with switching function at distances over 10 Å to bring the potential smoothly to zero at the cutoff. Particle-mesh Ewald method was used for electrostatics. The table of atom pairs for nonbonded interactions was updated at 10 fs intervals with a 13.5 Å maximum distance between atoms. Temperature was kept at 300 K by Langevin dynamics (thermostat) using a damping constant of 5 ps^{-1}, and pressure was kept at 1.01325 bar by the Langevin piston Nosé–Hoover method (barostat). Atomic coordinates and system energies were recorded every 20 ps. In the beginning, the systems were initialized with 20,000 steps of conjugate gradient energy minimization before the MD simulations. In addition, the systems that were used for the SMD simulations were thermalized for 100–500 ps at 300 K.

We performed 30 ns of MD for DigA16 and RBP with the bound ligands digoxigenin (Dog) and retinol (Rtl), respectively. Next, the ligands were pulled out in several SMD simulations (10–20 ns) for each protein with different atomic restraints and spring constants. Several SMD simulations were performed to assess the response of the system to different constraints and spring constants of the artificial force, and the results are listed in Table S1 (Supplemental Information). The reference point of the artificial force moved with constant velocity of 4 Å/ns in all SMD simulations, and the more constrained simulations were used to monitor the response to an arbitrary set of constraints. One SMD simulation for each protein was selected for a more detailed analysis.

The amount of work done by the artificial force [distance constraint, $F(Q)$] can be calculated via a thermodynamical integration (Meijer and Sprik 1998; Sprik and Ciccotti 1998)

$$W(Q) = -\int_{Q_0}^{Q} \mathrm{d}Q' F(Q'). \quad (1)$$

Formally, this equals the free energy change between the two states (Jarzynski 1997a, b), but our SMD simulations are much too fast to enable complete equilibration of the system for each value of Q (for example, water solvent imposes friction on the ligand), and the resulting values which display a strong preference for ligand binding (initial state) should be regarded as overestimated upper bounds for the free energy change.

Results and discussion

DigA16

The simulated DigA16 structure after 30 ns of MD at 300 K is illustrated in Fig. 1a, b. During the MD simulation, the total energy of the system was stable, and the inevitable linear energy drift was less than 0.002 kcal/ps. The origin of this accumulating error can be traced back to the MD integrator and discretization (finite time steps) of the atomic equations of motion. The structural changes with respect to the initial (constructed) positions are rather small as seen from the root-mean-square deviation (RMSD) in Fig. 2a; it stabilizes after 5 ns of simulation, and the RMSD of the backbone C_α-atoms converges close to 3.0 Å. Another key figure, the radius of gyration (R_{gyr}) (Fig. 2b), shows a peak in the beginning of the simulation and stabilizes at 16.3 Å after 5 ns. Comparison between the full protein (184 residues) and the residues 1–174 reveals that the peak is caused by the freely floating (added) tail of the last 10 residues at the C-terminus of the protein. In the beginning, the tail is close to the bulk of the protein, and then it stretches out and folds back into another conformation. The interaction energy calculation of the tail shows that the tail interacts with the rest of the protein more strongly after the conformational change (Supplemental Information, Fig. S1).

The ligand–receptor site of DigA16 is visualized in Fig. 3 with selected interatomic contacts. Digoxigenin is bound deep inside the cavity of DigA16, and it forms several stable hydrogen bonds (H-bonds) in addition to the hydrophobic van der Waals (vdW) interaction with the phenyl side-chains. The only part of Dog that is exposed to the solvent is the uppermost OH group, which typically makes two connections with waters as an H-bond acceptor

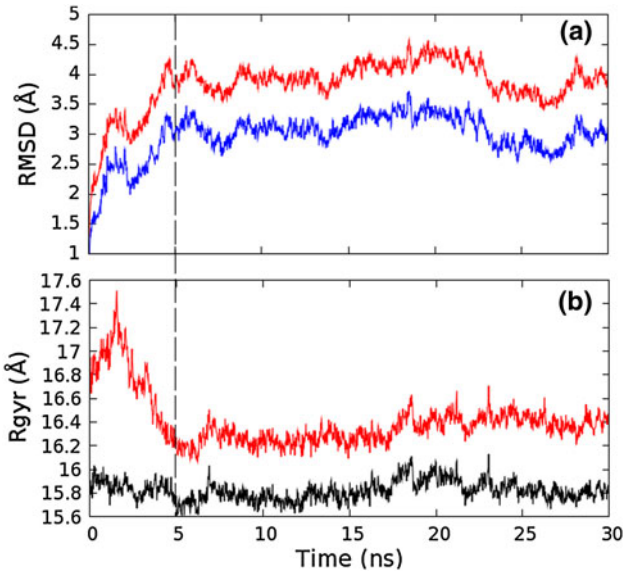

Fig. 2 Time dependence of **a** RMSD with respect to the initial structure, and **b** the radius of gyration of DigA16. Panel **a** displays RMSD for all the protein atoms in *red* and for the backbone atoms (...N-C_α-C-O...) in *blue*. Panel **b** displays the radius of gyration for all the atoms in *red* and for residues 1-174 in *black*

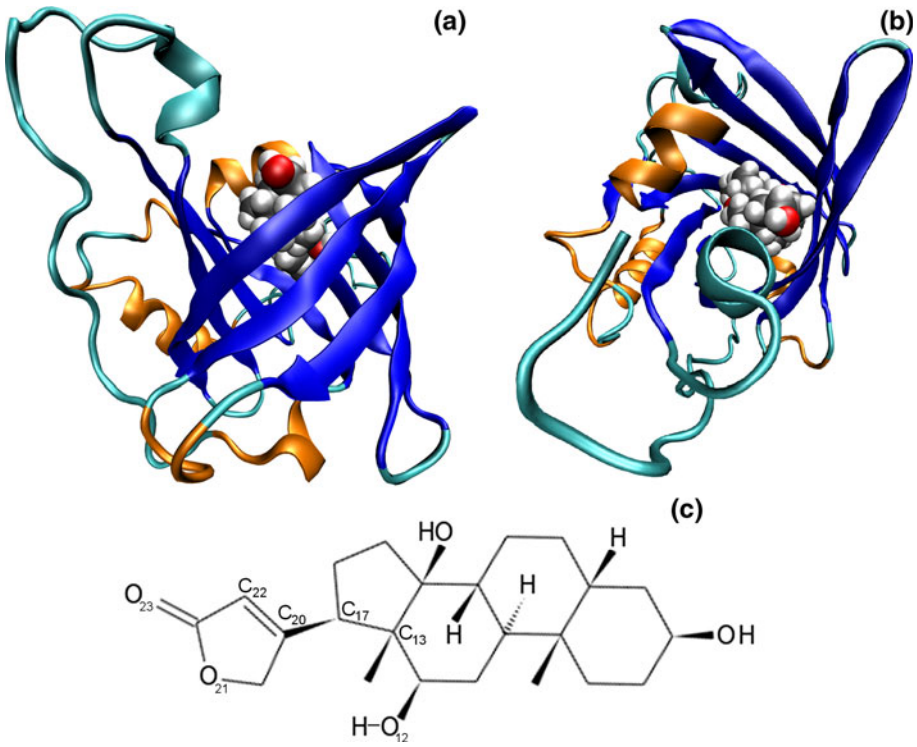

Fig. 1 **a**–**b** Ribbon schematic representation of DigA16 and digoxigenin after 30 ns of MD at 300 K from **a** the side of the β-barrel and **b** the mouth of the ligand pocket. Digoxigenin is displayed with hard spheres (carbon *gray*, oxygen *red*, hydrogen *white*), and the β-sheets are colored in *blue* and α-helices in *orange*. **c** Chemical structure of digoxigenin

and donor. The cavity opening as well as the pulling direction in the SMD simulations are upwards in Fig. 3b.

There are four (crystalline) water molecules inside the ligand cavity. Two of these (Wat213 and Wat193) are trapped at the bottom of the cavity, interacting with the Dog lactone ring (polar oxygen, O_{23}) and the surrounding protein residues. Wat213 acts as a bridge between Dog and Trp26 (see Fig. 3a, panels A and B) and rotates around its symmetry axis frequently. Most of the time, Wat193 acts as a similar bridge between the Wat213 oxygen and Gln28, but it rotates occasionally to a parallel orientation with Wat213 and bridges Dog (lactone ring, O_{21}) to Trp44. A typical H-bond lifetime is less that 0.5 ns for these molecules. The other two trapped water molecules are Wat212 and Wat192, and they are located on the side of Dog and are interacting with one Dog:OH group (O_{12}). Wat212 and Wat192 are more mobile than Wat213 and Wat193, because the former pair can switch places: One water connects Dog to the protein, while the other makes H-bonds with the first one (via oxygen) and the protein. The graph for the connection between Dog:H_{12} (hydroxyl) and Ser87:O (panel C) shows that the H-bond is stable for the first 18 ns of the simulation, but during the last 12 ns the H atom alternates between Ser87:O and His86:N. The orientation of the Dog lactone group varies slightly as the corresponding torsion angle with respect to the steroid rings (C_{13}-C_{17}-C_{20}-C_{22}) oscillates around 80-100°, but the ligand does not change its conformation during the simulation.

Simulation energetics were studied using the program VMD and NAMDEnergy plugin (Humphrey et al. 1996). Simulations of digoxigenin in bulk water (863 molecules, 20 ns) and apo-DigA16 (9,400 molecules, 10 ns) were performed for comparison. The energy components of digoxigenin and their differences for holo-DigA16 and the ligand in water are given in Table 1. The total energy difference of the various components provides the Dog binding energy, $\Delta E^{\mathrm{lig}}_{\mathrm{bind}}$, *without* the entropy contribution. Furthermore, the relaxation effects in the protein–protein and protein–solvent interactions are not included, as these would require simulations considerably longer than 10 ns for the apo-form in order to ensure full relaxation. In the present case, the energy difference for the ligand is −17.42 kcal/mol in DigA16, and it overestimates the experimental binding affinity of $\Delta G = -10.25$ kcal/mol (30.2 ± 3.6 nM) (Schlehuber et al. 2000).

An estimate for the ligand binding energy can also be calculated by using a linear response approximation for

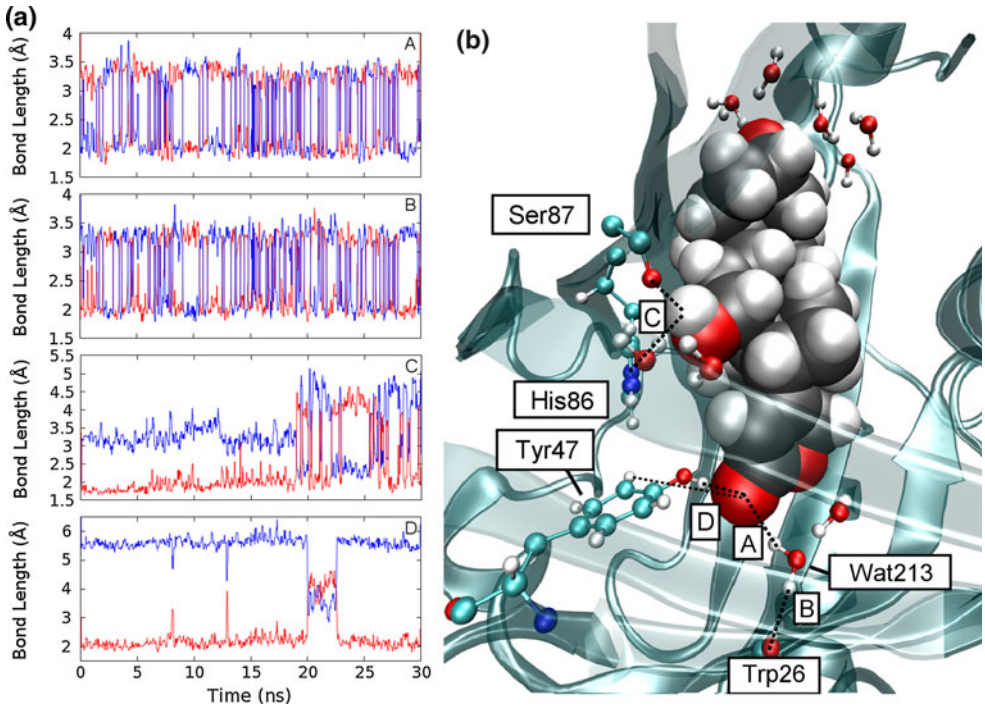

Fig. 3 **a** Selected connections between DigA16, digoxigenin, and water molecules in the bound state during the MD simulation, and **b** a snapshot of digoxigenin in the ligand pocket of DigA16. Bonds A–D in the snapshot correspond to curves in (**a**). The connections are Dog:O_{23}-Wat213:H_1 in *red* and Dog:O_{23}-Wat213:H_2 in *blue* (A), Trp26:O-Wat213:H_1 in *red* and Trp26:O-Wat213:H_2 in *blue* (B), Dog:H_{12}-Ser87:O in *red* and Dog:H_{12}-His86:N in *blue* (C), and Dog:O_{23}-Tyr47:H (hydroxyl) in *red* and Dog:O_{23}-Tyr47:H (aromatic) in *blue* (D). Only one connection is displayed for each Wat213 hydrogen in (**b**) but panels A and B show data for both Wat213:H atoms as the molecule fluctuates between two identical orientations (note the correlations between A and B). Color key for the ligand as in Fig. 1; DigA16 carbon in *cyan* and nitrogen in *blue*. The protein residues in the foreground are *translucent* in order improve the visibility of the ligand and residues (waters) inside

Table 1 Potential energy in MD simulations

	Energy component [kcal/mol]			
	vdW	Coul.	Ligand (internal)	Total
DigA16+Dog	−57.73	−52.25	102.90	−7.08
Dog in water	−24.65	−68.27	103.25	10.33
$\Delta E_{\text{bind}}^{\text{lig}}$	−33.08	16.02	−0.35	−17.42
RBP+Rtl	−44.73	−16.47	55.61	−5.59
Rtl in water	−27.08	−20.96	54.78	6.74
$\Delta E_{\text{bind}}^{\text{lig}}$	−17.65	4.49	0.83	−12.33

Interaction energy components (averages) between the ligand and its surroundings are included in the vdW and Coulomb terms. The ligand internal energy contains vdW, Coulomb, and covalent components inside the ligand

binding affinities (also known as the linear interaction energy, LIE, approach) (Åqvist et al. 1994; Åqvist and Marelius 2001; Åqvist et al. 2002), which is a semiempirical method and requires only two simulations of the ligand in different environments. The corresponding equation is

$$\Delta G_{\text{bind}} = \alpha\left(\left\langle V_{\text{lig-prot}}^{\text{vdW}}\right\rangle - \left\langle V_{\text{lig-sol}}^{\text{vdW}}\right\rangle\right) + \beta\left(\left\langle V_{\text{lig-prot}}^{\text{el}}\right\rangle - \left\langle V_{\text{lig-sol}}^{\text{el}}\right\rangle\right) + \gamma, \quad (2)$$

where $\langle V_{\text{lig-prot}}\rangle$ is the (average) potential energy between the ligand and the rest of the simulation box in a complex with the binding protein and $\langle V_{\text{lig-solv}}\rangle$ is the same for a ligand in water solvent. Superscripts denote van der Waals interaction ("vdW") and Coulombic interaction ("el"). Different values have been proposed in literature for the parameters α, β, and γ (energy), depending on the system in question, but formally β should be 0.5 in order to satisfy the Zwanzig expression (Zwanzig 1954) of free energy differences and dielectric theory (Åqvist et al. 2002). It has been shown that the value of α depends on the details of the binding site, and especially on its hydrophobicity (Wang et al. 1999). For the last term, $\gamma = 0$ kcal/mol is used in most cases.

Hansson et al. (1998) have reported that uncharged ligands with certain dipolar groups may require deviations from the electrostatic linear response, and $\beta < 0.5$ should be considered. A series of MD simulations with the GROMOS force field (van Gunsteren and Berendsen 1987) for 18 ligand–receptor complexes (endothiapepsin, HIV proteinase, glucose binding protein, and trypsin) concluded that the LIE values $\alpha = 0.186$ and $\beta = 0.320$ provided the best fit, with root-mean-square (RMS) error of 0.97 kcal/mol ($\gamma = 0$ kcal/mol). However, considerable deviations were observed later with the entirely hydrophobic binding of retinoids to RBP, and this required a correction term $\gamma = -7$ kcal/mol (Åqvist and Marelius 2001). Another way of incorporating the hydrophobic interactions is to scale α according to the fine details of the binding site (nonpolar desolvation ratios) (Wang et al. 1999). In the case of DigA16, we observe that the LIE estimate with $\alpha = 0.186$, $\beta = 0.320$, and $\gamma = 0$ kcal/mol results in $\Delta G_{\text{bind}} = -1.03$ kcal/mol, which severely underestimates the measured digoxigenin binding. This indicates that a similar correction term should be applied as reported for retinoids and RBP, and $\gamma = -7$ kcal/mol results in $\Delta G_{\text{bind}} = -8.03$ kcal/mol, which is significantly closer to measurement (−10.25 kcal/mol). There is no rigorous basis for this correction in DigA16, and in order to achieve a more accurate value one needs to perform simulations with a training set of relevant molecules (steroids) inside DigA16. The above set of parameters turns out to be valid for RBP (see below, Sect. 3.2).

SMD simulation of DigA16

Several SMD simulations of Dog removal were performed in order to see how the system responds to different constraints and spring constants of the artificial force, and the results are collected in Supplemental Information (Table S1). We have selected a simulation with $k = 1$ kcal/mol/Å^2 and $W = 98$ kcal/mol ($F_{max} = 800$ pN) for further inspection, and the course of the expulsion process is demonstrated in Figs. 4 and 5. Upon Dog removal, the artificial stretching force and center of mass (CoM) of the ligand show a significant jump at 3.2 ns. Here, several contacts between Dog and DigA16 that are present in the bound state (e.g., the H-bond with Tyr47) break as the ligand suddenly moves 4.8 Å outwards (Fig. 5a, b), and we see a rapid change in the interaction energy between Dog and DigA16 (Fig. 5c). Also the other SMD simulations showed characteristic phases during the ligand expulsion process, which were separated by sudden movements of the ligand (2–10 Å), followed by stabilization at a new location for 0.1–1.0 ns.

During the removal, different residues make connections ("bonds") with the ligand (Fig. 5d). As the ligand is pulled out, the first residues to break H-bonds with the ligand are the ones at the side of the β-barrel and the α-helix (residues 33–40) on the opposing side above the barrel. At the rupture point (Fig. 4), the residues closest to the ligand are close to the cavity mouth in the β-barrel and in the aforementioned α-helix. The force required to pull Dog out after the initial twitch is low compared with the maximum force but comparable to the rupture force of Rtl removal from RBP (see below). A total of 20 residues are responsible for almost all of the interaction energy between Dog and DigA16 (Fig. 5c), and these residues are listed in Table 2.

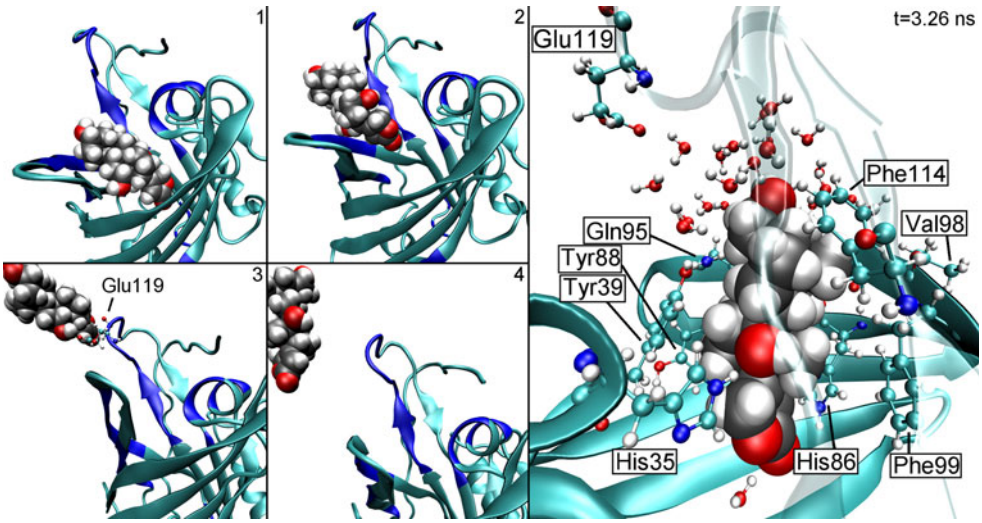

Fig. 4 Snapshots of DigA16 during the selected SMD simulation at times corresponding to the *numbered lines* in Fig. 5, and a closeup of Dog before the first twitch (rupture point) at 3.26 ns. Times (1–4) are: (1) 3.0 ns, (2) 3.5 ns, (3) 5.6 ns, and (4) 6.4 ns. Residues highlighted in *blue* in (1–4) are the residues listed in Table 2. Other *colors* and representation as in Fig. 3

Of these, eight were mutated during the engineered mutagenesis of DigA16 from wild-type BBP. In addition, residues Tyr47 (from BBP) and Arg58 (mutated) contribute significantly in the bound state, with interaction energies of −5.68 and −5.99 kcal/mol (ligand accessible area 15 and 21 $Å^2$), respectively, but these contacts are broken at an early stage of the ligand stretching.

After 5 ns, ligand–receptor interaction is completely due to residues 114–121. These residues are located in the outlier loop of the protein, and they are the last residues to make contact with the ligand as it is being pulled away. Although the ligand was removed and not inserted in the SMD simulation, we suggest that this loop could function as an initial sensor that recognizes a potential molecule for the protein binding site. A particularly interesting residue is Glu119, which resides in the suspected sensor loop. It does not have interactions with the fully buried ligand (Fig. 4), but during the expulsion process it makes a significant contribution to the interaction energy: From 4 ns to 6 ns almost 50% of the protein-ligand interaction is associated with Glu119, and after 6 ns Glu119 is the only residue interacting with Dog. Glu119 binds to the Dog lactone ring via its polar oxygen and amine hydrogens, and especially in the end of the removal, the contacts are very well lined up. This appears to be a mechanism for keeping the incoming ligand in a preferable orientation with respect to the receptor site.

As digoxigenin leaves the cavity, Wat193 and Wat213 initially follow it towards the cavity center and remain there. They compensate for the absence of digoxigenin OH-groups and form H-bonds with the protein. External water molecules protrude into the cavity and bond with the aforementioned crystalline waters, but they do not completely fill the free space left by the ligand. Despite this, the cavity does not shrink upon Dog removal, and the gyration radius is at the same level (16.3 Å, Supplemental Information, Fig. S2). At the end of the simulation, there are a total of eight water molecules inside the ligand cavity. A reference simulation for apo-DigA16 results in a gyration radius of 16.2 Å with 16–21 water molecules inside the cavity after 10 ns. This indicates that the receptor site can host more water than present at the end of the SMD simulation, and that the changes in water occupation will not be reflected in R_{gyr}.

Interaction energy calculations of potential salt bridge residues indicate that two pairs could react to ligand expulsion: The pairs Asp126–Lys159 and Asp1–Lys38 show sharp energy increases after 3.2 ns, and they decrease back to the initial level at 3.8 ns. This coincides with the rupture point of the ligand removal. Asp126 and Lys159 are located on the side of the β-barrel and in the beginning of the C-terminus tail, respectively, whereas Asp1 and Lys38 reside in the N-terminus and α-helix near the entrance of the ligand-binding cavity, respectively. Presumably, the changes in interaction energies reflect the moderate deformations of the β-barrel during the ligand removal process.

RBP

The secondary structure of RBP after 30 ns of MD is visualized in Fig. 6, and RMSD and R_{gyr} of the holo-form are displayed in Fig. 7a, b. RMSD and R_{gyr} show little deviation from the initial values during the MD simulation: RMSD has a small, approximately linear upward trend, and the RMSD of C_α-atoms increases by 0.1 Å over 5.5 ns but shows also signs of stabilization towards the end of the simulation. The changes with respect to the initial structure are considerably smaller than for DigA16, and this is presumably due to the fact that the RBP simulation was started

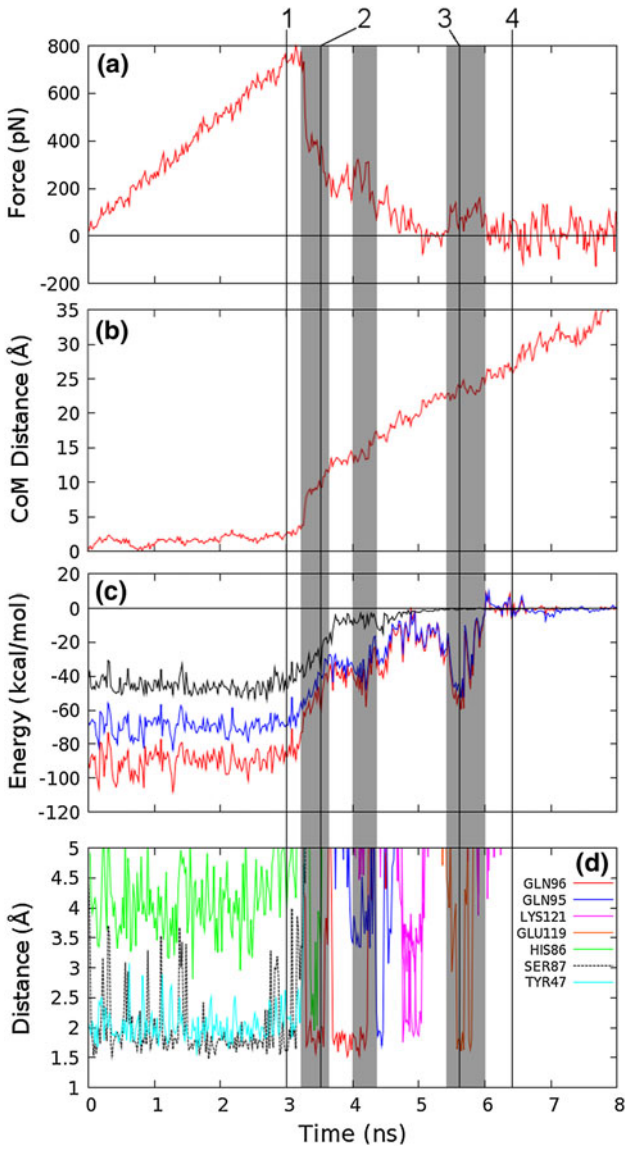

Fig. 5 Expulsion process of digoxigenin. **a** Time dependence of the artificial force exerted on Dog, **b** CoM distance from the initial position, **c** interaction energy between Dog and DigA16 in *red*, between Dog and residues listed in Table 2 in *blue*, and between Dog and the mutated residues in *black*, and **d** selected connections between Dog and DigA16 residues during the SMD simulation. The *shaded areas* correspond to transitions in the Dog position or in the artificial force, and *black lines* (1–4) correspond to the snapshots in Fig. 4 (3.0 ns, 3.5 ns, 5.6 ns, and 6.4 ns)

from a uniform, partial set of experimental coordinates (residues 1–175), and there is no floating C-terminus tail.

Retinol, being a hydrophobic molecule with a single polar oxygen atom in one end, makes a hydrogen bond to the protein via the sole hydroxyl group (Fig. 7c) in contrast to the three polar OH groups and lactone ring in digoxigenin. The ligand–receptor site of RBP is visualized in Fig. 7d, and it shows that RBP has several residues with a phenyl side-chain on the inner surface of the ligand cavity.

Table 2 DigA16 residues that have relevance in the Dog expulsion process

Residue	LAA	SAB[a]	E_{inter}	Residue	LAA	SAB[a]	E_{inter}
Lys25	–	–	−3.18	**Gly97 (Asn)**	**4**	**13**	**−2.43**
Asp34 (Asn)	–	–	**−1.47**	Phe99	11	11	−2.27
His35 (Ser)	**13**	**11**	**−7.22**	**Phe114 (Tyr)**	**22**	**28**	**−4.53**
Lys38	–	–	2.23	**Ser116 (Lys)**	–	–	**−0.02**
Tyr39	32	35	−6.12	Tyr117	–	–	0.00
His86	38	32	−7.47	Asp118	–	–	−1.26
Ser87 (Lys)	**6**	**2**	**−6.63**	Glu119	–	–	−3.06
Tyr88 (Leu)	**20**	**36**	**−6.59**	Asp120	–	–	−0.75
Gln95 (Lys)	**7**	**19**	**−3.41**	Lys121	–	–	0.95
Glu96	6	15	−4.13	Trp129	31	34	−11.53

The values for the ligand accessible area (LAA, in Å^2) are calculated from the final structure of the MD simulation (after 30 ns) using a 1.5 Å probe radius. Interaction energy in the bound state between the ligand and the residue, E_{inter} (in kcal/mol), is calculated as an average over the 30 ns MD simulation. Residues mutated during the DigA16 design are printed in boldface with the original BBP residue in parenthesis

[a] Surface area of the protein residue buried (SAB, in Å^2), values from the work by Korndörfer et al. (2003) for reference

The interaction energy analysis between Rtl and RBP (Table 1) confirms that the phenyl rings close to the ligand are the main cause of attraction for the bound state. Retinol typically forms one hydrogen bond with an external water molecule via its terminal OH group (hydrogen). The hydroxyl oxygen is mostly bound to the residue Gln98 with a H-bond lifetime of several nanoseconds (Fig. 7c, d), but the OH group can occasionally point towards the protein (e.g., after 12 ns), breaking this connection and exposing the oxygen to the solvent. This results in temporary hydrogen bonds with two external water molecules.

The retinol starting configuration is an all-*trans* isomer, but two of the torsion angles can rotate to a *cis* conformation (Supplemental Information, Fig. S3): first, the torsion around the C_8-C_9 bond, which is in a *cis* conformation on two occasions for a total of 11 ns; second, the torsion around the C_{12}-C_{13} bond, which switches to a *cis* conformation once for 3 ns during the 30 ns simulation. The changes in the polyene chain torsion are coupled to the contacts with Leu97 and Gln98 and their side-group orientations, and these are reflected in the corresponding residue–ligand interaction energies (Figs. S3 and S4).

The potential energy components of Rtl and RBP are listed in Table 1. The effect of the hydrophobic receptor

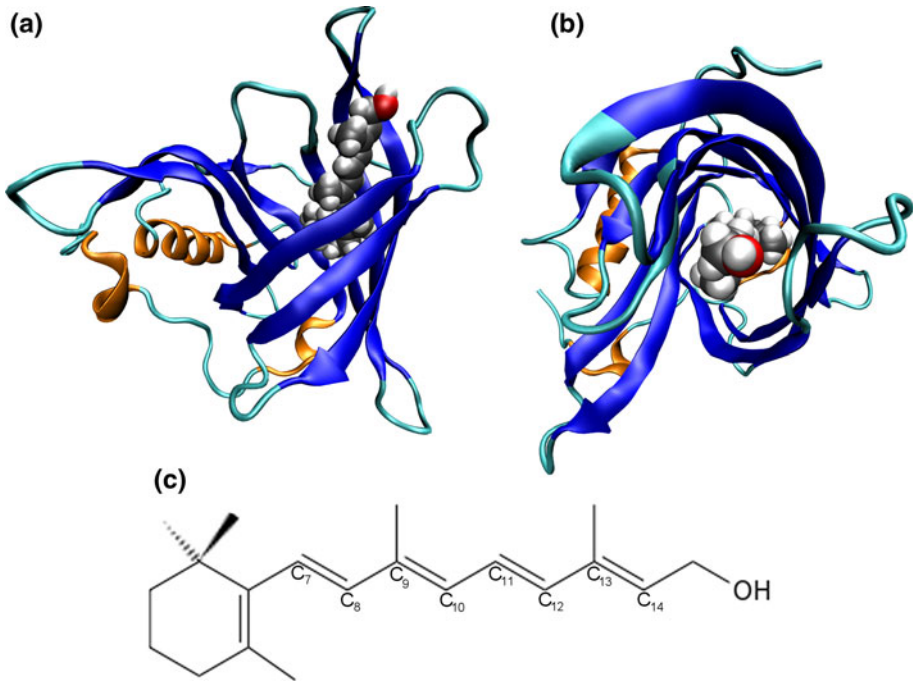

Fig. 6 **a**, **b** Ribbon schematic representation of RBP and retinol after 30 ns of MD at 300 K from **a** the side of the β-barrel and **b** the mouth of the ligand pocket. Retinol is displayed with hard spheres. *Colors* as in Fig. 1. **c** Chemical structure of retinol

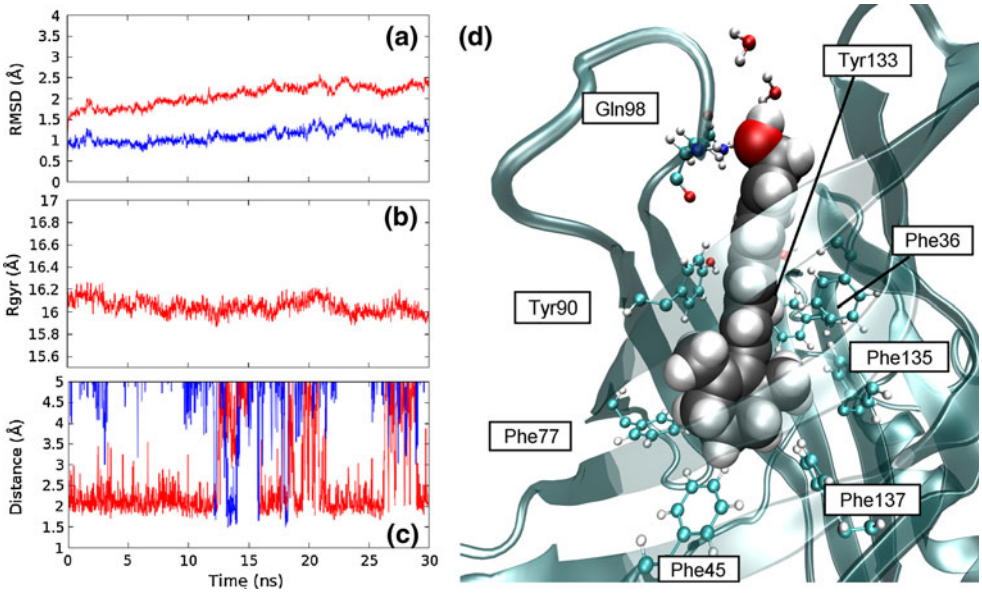

Fig. 7 **a** RMSD with respect to the initial RBP structure during the MD simulation, **b** radius of gyration, **c** H-bond distances between the retinol OH group and Gln98 of RBP, and **d** a snapshot of retinol in the RBP receptor site. Panel **a** contains RMSD for all the atoms in *red* and for the backbone atoms (...N-C_α-C-O...) in *blue*, and (c) displays bond distances Gln98:H-Rtl:O in *red* and Gln98:O-Rtl:H (hydroxyl) in *blue*. Colors and representation in panel (**d**) are as in Fig. 3

site is evident as a considerable change of the vdW interaction, while the Coulombic term remains rather similar. The resulting energy difference, i.e., the ligand binding energy without entropy contribution and protein relaxation effects, is -12.33 kcal/mol. Taking into account the hydrophobic nature of retinol, it is clear that association with the RBP receptor site increases entropy (of the water solvent) and decreases the free energy. The first LIE estimate for the retinol binding energy ΔG_{bind} has been calculated with Eq. (2) using the parameters $\alpha = 0.186$ and $\beta = 0.320$ by Hansson et al (1998). Reference simulations of retinol in bulk water (886 molecules, 20 ns) and apo-RBP (9,325 waters, 10 ns) were performed for this purpose, and the LIE energy is $\Delta G_{\text{bind}} = -1.85$ kcal/mol. This value underestimates significantly the real free energy change of -9.15 kcal/mol $\left(1.9 \times 10\,\text{M}^{-7}\right)$ (Cogan et al. 1976), and a correction term γ has to be used. Remarkably, the value $\gamma = -7$ kcal/mol, originally reported by Åqvist and Marelius (2001) for retinoids and RBP, results in $\Delta G_{\text{bind}} = -8.85$ kcal/mol, which differs by only 0.30 kcal/mol from the experimental value. This lends further support to the conclusion that differences in classical force fields and MD simulations are not significantly reflected in the LIE parameters (Wang et al. 1999; Åqvist et al. 2002).

SMD simulation of RBP

The SMD simulation with $k = 0.2\,\text{kcal/mol/Å}^2$ and $W = 66\,\text{kcal/mol}$ ($F_{\max} = 291\,\text{pN}$) is selected for further inspection. The course of the process is visualized in Fig. 8, and related properties are shown in Fig. 9. RBP displays similar behavior as DigA16 with a linearly increasing pulling force. However, there are several maxima in the force profile, and Rtl requires much less force for removal compared with Dog. This is due to the absence of H-bonds between the ligand and protein. The sole polar OH group in retinol is partially immersed into water already in the bound state. This means that most of the interaction energy must come from the vdW contributions of the phenyl side-chains close to Rtl, and the interaction is spread out over groups of atoms (aromatic rings). The reduced directionality lowers the rupture force compared with that of Dog expulsion.

The time evolution of the artificial force (Fig. 9a) shows three peaks. Several maxima are typical for the SMD simulations of RBP as opposed to a single peak for DigA16. This must be related to the differences in ligand–receptor interaction (hydrophobic versus hydrophilic) in these systems. The retinol CoM distance with respect to the bound state (Fig. 9b) shows small up-and-down movement that was not observed for Dog. This suggests that RBP adjusts to the ligand stretching and tries to restore the ligand back into the cavity on several occasions. Interaction energy calculations show that a selection of 28 residues almost entirely covered the ligand–receptor interaction during the expulsion process (Fig. 9c). The residues are mainly located in the outermost β-sheet and three loops surrounding the mouth of the ligand-binding cavity (Fig. 8). At the rupture point, Rtl is more immersed into water than Dog in the same situation, and this is even more pronounced at the last force maximum (Fig. 8, $t = 6.8$ ns). There are several possible reasons for this behavior. The most obvious explanation is that retinol is hydrophobic, and thus, water resists its entry into the solvent. Another possible cause is the bulky head group of Rtl, which interacts with RBP more strongly than the narrow polyene tail. Ligand stretching requires opening of the protein cavity in order to adjust to the sterically hindered release of the β-ionone ring, and this occurs at a later stage (4–7 ns). After the ligand is almost completely unbound, the Gln98 side-chain turns from its perpendicular position with respect to the ligand and adopts a parallel orientation with the retinol polyene chain covering one side of the ligand and offering a hydrophobic contact surface. Leu64 is the last residue to have contact with retinol (Fig. 8) in clear analogy with that of Glu119 in DigA16, indicating that the outlier loop of residues 63–67 could serve as a sensor loop for ligand recognition in RBP.

During the selected SMD simulation, retinol starts in an all-*trans* conformation, switches to a *cis* conformation around the C_8-C_9 bond almost immediately (0.5 ns) as a response to the pulling, and changes back to the all-*trans* conformation at 4 ns as the initial contact with the Gln98 side-chain finally breaks. Retinol remains in the all-*trans* conformation after the final removal. It must be emphasized that the fluctuations in torsion angles depend sensitively on the retinol force field parameters.

RMSD analysis of the SMD simulation (Supplemental Information, Fig. S5) shows modulation after 4 ns upon retinol removal, and a small overall increase at the end. The radius of gyration displays a negligible decrease of 0.1 Å upon retinol detachment, which is in agreement with the recent experimental finding for bovine RBP and retinoic acid (Inoue et al. 2009). In contrast, Inoue et al.

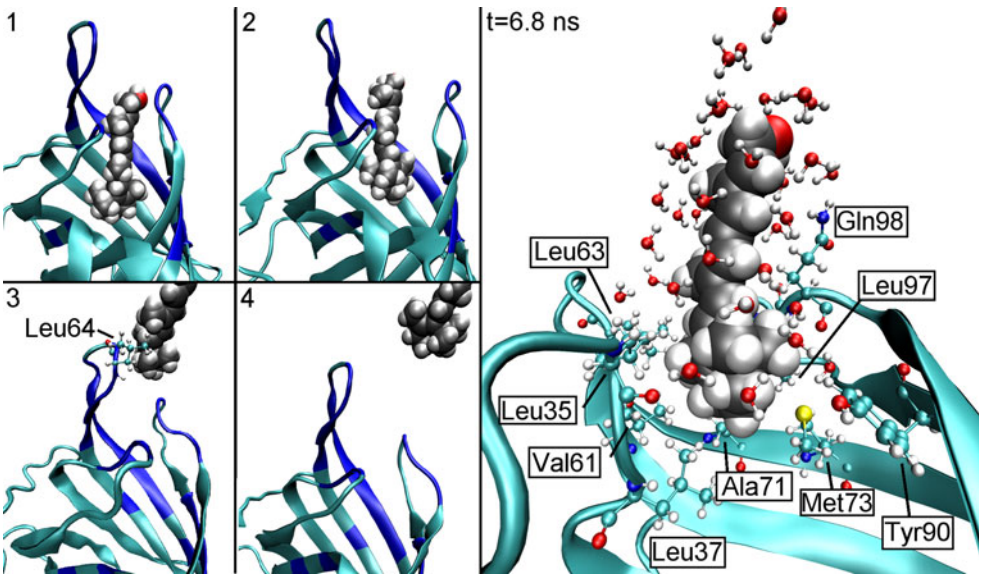

Fig. 8 Snapshots of RBP during the selected SMD simulation at times corresponding to *vertical lines* in Fig. 9, and a closeup of the ligand environment before the last twitch at 6.8 ns. Snapshots (1–4) are: (1) 4.5 ns, (2) 6.0 ns, (3) 7.0 ns, and (4) 7.3 ns. Residues highlighted in *blue* in (1–4) are listed in Table 3. Other colors and representation as in Fig. 3

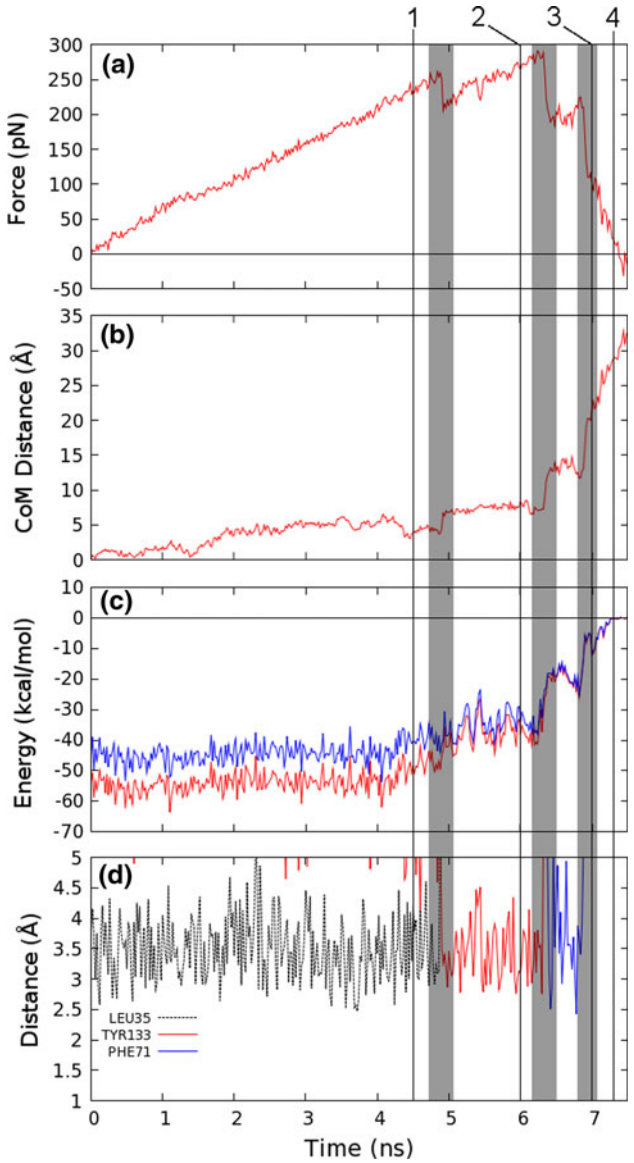

Fig. 9 Retinol expulsion process. **a** Force exerted on retinol, **b** CoM distance from the initial position, **c** interaction energy between Rtl and RBP in *red*, and between Rtl and the residues listed in Table 3 in *blue*, and **d** selected distances between retinol and RBP residues during the SMD simulation. *Shaded areas* correspond to sudden changes in the retinol position or in the artificial force, and *black lines* (1–4) correspond to the snapshots in Fig. 8 (4.5, 6.0, 7.0, and 7.3 ns)

(2009) reported significant cavity expansion upon ligand removal for another lipocalin, prostaglandin D synthase, in complex with retinoid acid, bilirubin, and biliverdin. The ligand pocket of RBP remains empty for over 2.5 ns after Rtl removal, but after that two water molecules enter the pocket. Previously, Chau (2004) reported no water molecules entering the ligand pocket of bovine serum RBP after retinol removal by using 1 ns simulations. Our reference simulation for apo-RBP has four water molecules inside the cavity in the beginning, but these move out during the

Table 3 RBP residues relevant to the retinol expulsion process

Residue	LAA	E_{inter}	Residue	LAA	E_{inter}	residue	LAA	E_{inter}
Leu35	16	−1.86	Asp68	–	−0.46	Lys89	–	−0.65
Phe36	11	−2.74	Val69	–	−0.07	Tyr90	28	−3.99
Leu37	45	−3.97	Cys70	–	−0.01	Phe96	–	−0.63
Ala57	17	−1.98	Ala71	–	−0.41	Leu97	26	−3.55
Val61	11	−1.36	Asp72	–	0.15	Gln98	15	−4.71
Arg62	2	0.08	Met73	11	−3.50	Lys99	–	−0.91
Leu63	4	−0.27	Val74	6	−0.84	Tyr133	9	−1.82
Leu64	–	0.03	Gly75	5	−0.94	Phe135	14	−1.77
Asn65	–	0.03	Phe77	1	−0.95			
Trp67	–	−0.14	Met88	40	−2.60			

The ligand accessible area (LAA, in Å^2) and the interaction energy (E_{inter}, in kcal/mol) are calculated as in Table 2

10 ns simulation without affecting the gyration radius. According to these results, water molecules can occasionally visit the hydrophobic receptor site, but most of the time water is absent.

The interaction energy analysis of the potential salt bridges (Calderone et al. 2003) reveals three residue pairs that respond to the ligand removal: Lys17–Asp79, Arg155–Glu158, and Glu157–Arg163. The first pair indicates a bond breaking at 9 ns with a re-emerging binding after that (strengthening of 10 kcal/mol). The residue pair is located at the bottom end of the β-barrel and stabilizes its closed conformation. The second pair (Arg155–Glu158) shows a bond breaking at an early stage of the ligand removal (before significant retinol movements), and it binds again after the removal has been finished. This pair is located in two adjacent α-helix turns and is unlikely to affect the protein overall conformation considerably. The third connection (Glu157–Arg163) is broken after the first ligand twitch at 4.8 ns and is re-formed 0.5 ns after the breaking. This bond is between the aforementioned α-helix and the following C-terminus sequence (tail).

Conclusions

The structure and dynamics of the lipocalin proteins DigA16 and RBP have been studied using classical MD simulations at 300 K. It is characteristic of lipocalins that they can bind small hydrophobic ligand molecules inside their β-barrel, and we have focused on the ligand–receptor interactions in each system. DigA16 is an engineered protein ("anticalin") with enhanced specificity for the steroid digoxigenin and potential applications in medicine, and RBP is a biologically important protein as the carrier of vitamin A (retinol) in the human body. The starting structures of the holo-forms of both proteins were based on

high-resolution x-ray diffraction measurements (Calderone et al. 2003; Korndörfer et al. 2003), where the DigA16 structure was missing residues in both its N- and C-terminus tails and in the middle loop. These were added by using computer modeling.

During the MD simulations, DigA16 experiences a conformational change in its free-floating C-terminus tail. The ten last residues (175–184) of the sequence start from a conformation in which they are folded close to the bulk of the protein, and then stretch out and fold back in a different conformation. This occurs during the first 5 ns of the MD simulation, and the protein does not show further conformational changes (Fig. 2). Therefore, our final structure of DigA16 (after 30 ns) provides a model for the secondary structure of the whole protein, including the segments that are absent in the original crystal structure (coordinates of the optimized final structure are available in Supplemental Information). Analysis of the ligand-binding site reveals that Dog is rather rigid, the hydrogen-bonding network is stable, and that the four crystalline water molecules are trapped in between the ligand and protein.

The RBP structure shows only small changes during MD, as it is based solely on the experimental (crystalline) starting geometry. The ligand-binding cavity of RBP differs from DigA16 because it is completely hydrophobic, and the main source of attraction (of vdW type) is caused by the ligand and aromatic side-chains. The retinol OH end group points towards the solvent and forms H-bonds with external water and Gln98 in the entrance of the cavity. Retinol experiences some conformational changes as it starts in an all-*trans* conformation but switches occasionally to other conformations during the simulation, and these are coupled to the interaction between the OH end group and residues Leu97 and Gln98.

Ligand binding energies have been estimated based on the LIE method, which requires only two MD simulations of the ligand (a) inside the receptor and (b) in the solvent. Previously, it has been reported that the LIE parameters $\alpha = 0.186$ and $\beta = 0.320$ in Eq. 2 are valid for a wide variety of ligand–receptor complexes (Hansson et al. 1998). However, the hydrophobic binding of retinoids to RBP requires a special correction term $\gamma = -7$ kcal/mol (Åqvist and Marelius 2001). By using these parameters, our estimate for the retinol binding energy is $\Delta G_{\text{bind}} = -8.85$ kcal/mol, which differs only 0.30 kcal/mol from the measured value -9.15 kcal/mol. In the case of DigA16 and Dog, we observe that the same set of parameters results in a value $\Delta G_{\text{bind}} = -8.03$ kcal/mol, which underestimates the measured digoxigenin binding by 2.22 kcal/mol. Obviously, a correction term of $|\gamma| = 7 - 10$ kcal/mol is needed in this case as well.

SMD simulations of the ligand expulsion process were performed for both proteins with various constraints and spring constants in order to gain information of the reverse process, ligand recognition. Analysis of the artificial force, interaction energies, and connections per residue enables a detailed description of the ligand–receptor interaction. DigA16 and RBP display different rupture forces, and this difference can be attributed to the H-bonds between Dog and DigA16. Once the initial digoxigenin H-bonds have been broken, the force required to pull the ligand out is close to that of retinol (in RBP), whose binding is completely hydrophobic. The mutated DigA16 residues (8 of 18 are important during the expulsion) contribute approximately one-half of the Dog interaction energy with DigA16 (Table 2 and Fig. 5c), and the mutated residues His35, Arg58, Ser87, Tyr88, and Phe114 are of particular importance because of their ligand accessible area and/or interaction energy. The ligand removal does not have a significant effect on the protein radius of gyration for DigA16 and RBP.

A potential "sensor loop" was found for both receptors. These are the outlier loops 114–121 and 63–67 for DigA16 and RBP, respectively, and they are located at the immediate vicinity of the mouth of the ligand-binding cavity. The outlier loops are the last parts of the protein to have interaction with the ligand as the ligand is being pulled out, and especially, the residues Glu119 (DigA16) and Leu64 (RBP) are critical for sensing. The former forms hydrogen bonds with the Dog lactone group, whereas the latter makes a hydrophobic contact with the retinol ring.

Acknowledgments The parallel simulations were performed on the SGI/Altix platform at the Nanoscience Center of the University of Jyväskylä. The charge calculations with Gaussian03 were performed on the HP CP4000 BL platform at CSC in Espoo, Finland. All molecular visualizations were done with Visual Molecular Dynamics (Humphrey et al. 1996). We thank A. Skerra for discussions and providing the DigA16 structure (1LKE) with an added middle loop (residues 118–122), O. Pentikäinen for instructions with the simulation setups, and R.O. Jones for critical reading of the manuscript.

References

Åqvist J, Marelius J (2001) The linear interaction energy method for predicting ligand binding free energies. Comb Chem High Throughput Screen 4:613–626

Åqvist J, Medina C, Samuelsson J-E (1994) New method for predicting binding-affinity in computer-aided drug design. Protein Eng 7:385–391

Åqvist J, Luzhkov VB, Brandsdal BO (2002) Ligand binding affinities from MD simulations. Acc Chem Res 35:358–365. doi:10.1021/ar010014p

Beaufays J et al (2008) Ir-LBP, an Ixodes ricinus tick salivary LTB4-binding lipocalin, interferes with host neutrophil function. PLoS ONE. doi:10.1371/journal.pone.0003987

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The Protein Data Bank. Nucleic Acids Research 28:235–242. (http://www.pdb.org)

Besler BH, Merz KM Jr, Kollman PA (1990) Atomic charges derived from semiempirical methods. J Comp Chem 11:431–439

Breustedt DA, Schönfeld DL, Skerra A (2006) Comparative ligand-binding analysis of ten human lipocalins. Biochim Biophys Acta 1764:161–173. doi:10.1016/j.bbapap.2005.12.006

Calderone V, Berni R, Zanotti G (2003) High-resolution structures of retinol-binding protein in complex with retinol: pH-induced protein structural changes in the crystal state. J Mol Biol 329:841–850. doi:10.1016/S0022-2836(03)00468-6

Chau P-L (2004) Water movement during ligand unbinding from receptor site. Biophys J 87:121–128. doi:10.1529/biophysj.103.036467

Cogan U, Kopelman M, Mokady S, Shinitzky M (1976) Binding affinities of retinol and related compounds to retinol binding proteins. Eur J Biochem 65:71–78

Cowan SW, Newcomer ME, Jones TA (1990) Crystallographic refinement of human serum retinol binding protein at 2 Å resolution. Proteins Struct Funct Genet 8:44–61. doi:10.1002/prot.340080108

Duan Y et al (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. J Comput Chem 24:1999–2012. doi:10.1002/jcc.10349

Flower DR (1994) The lipocalin protein family: a role in cell regulation. FEBS Lett 354:7–11

Flower DR (1996) The lipocalin protein family: structure and function. Biochem J 318:1–14

Frisch MJ et al (2004) Gaussian03, Revision D. 02. Gaussian Inc., Wallingford CT

Gunnerson KN, Pereverzev YV, Prezhdo OV (2009) Atomistic simulation combined with analytic theory to study the response of the P-selectin/PSGL-1 complex to an external force. J Phys Chem B 113:2090–2100. doi:10.1021/jp803955u

Hansson T, Marelius J, Åqvist J (1998) Ligand binding affinity prediction by linear interaction energy methods. J Comput Aided Mol Design 12:27–35

Humphrey W, Dalke A, Schulten K (1996) VMD—Visual Molecular Dynamics. J Mol Graphics 14:33–38

Hytönen VP, Vogel V (2008) How force might activate talin's vinculin binding sites: SMD reveals a structural mechanism. PLoS Comput Biol. doi:10.1371/journal.pcbi.0040024

Inoue K, Yagi N, Urade Y, Inui T (2009) Compact Packing of Lipocalin-type Prostaglandin D Synthase Induced by Binding of Lipophilic Ligands. J Biochem 145:169–175

Isralewitz B, Baudry J, Gullingsrud J, Kosztin D, Schulten K (2001) Steered molecular dynamics investigations of protein function. J Mol Graphics Modell 19:13–25

Isralewitz B, Gao M, Schulten K (2001) Steered molecular dynamics and mechanical functions of proteins. Curr Opin Struc Biol 11:224–230

Jarzynski C (1997) Nonequilibrium equality for free energy differences. Phys Rev Lett 78:2690–2693

Jarzynski C (1997) Equilibrium free-energy differences from non-equilibrium measurements: A master-equation approach. Phys Rev E 56:5018–5035

Jeffrey PD, Strong RK, Sieker LC, Chang CY, Campbell RL, Petsko GA, Haber E, Margolies MN, Sheriff S (1993) 26-10 Fab-digoxin complex: affinity and specificity due to surface complementarity. Proc Natl Acad Sci USA 90:10310–10314

Jeffrey PD, Schildbach JF, Chang CY, Kussie PH, Margolies MN, Sheriff S (1995) Structure and specificity of the anti-digoxin antibody 40–50. J Mol Biol 248:344–360

Korndörfer IP, Schlehuber S, Skerra A (2003) Structural mechanism of specific ligand recognition by a lipocalin tailored for the complexation of digoxigenin. J Mol Biol 330:385–396. doi:10.1016/S0022-2836(03)00573-4

Meijer EJ, Sprik M (1998) Ab initio molecular dynamics study of the reaction of water with formaldehyde in sulfuric acid solution. J Am Chem Soc 120:6345–6355

Mills JL, Liu G, Skerra A, Szyperski T (2009) NMR Structure and Dynamics of the Engineered Fluorescein-Binding Lipocalin FluA Reveal Rigidification of β-Barrel and Variable Loops upon Enthalpy-Driven Ligand Binding. Biochemistry 48:7411–7419. doi:10.1021/bi900535j

Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of Macromolecular Structures by the Maximum-Likelihood Method. Acta Cryst D 53:240–255

Newcomer ME, Jones TA, Åqvist J, Sundelin J, Eriksson U, Rask L, Peterson PA (1984) The three-dimensional structure of retinol-binding protein. EMBO J 3:1451–1454

Niu C et al (2009) Dynamic mechanism of E2020 binding to acetylcholinesterase: a steered molecular dynamics simulation. J Phys Chem B 109:23730–23738. doi:10.1021/jp0552877

Park S, Schulten K (2004) Calculating potentials of mean force from steered molecular dynamics simulations. J Chem Phys 120:5946–5961. doi:10.1063/1.1651473

Peräkylä M (2009) Ligand unbinding pathways from the vitamin D receptor studied by molecular dynamics simulations. Eur Biophys J 38:185–198. doi:10.1007/s00249-008-0369-x

Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kalé L, Schulten K (2005) Scalable molecular dynamics with NAMD. J Comp Chem 26:1781–1802. doi:10.1002/jcc.20289

Schlehuber S, Skerra A (2005) Lipocalins in drug discovery: from natural ligand-binding proteins to 'anticalins'. Drug Discovery Today 10:23–33

Schlehuber S, Beste B, Skerra A (2000) A novel type of receptor protein, based on the lipocalin scaffold, with specificity for digoxigenin. J Mol Biol 297:1105–1120

Schlehuber S, Skerra A (2002) Tuning ligand affinity, specificity, and folding stability of an engineered lipocalin variant—a so-called 'anticalin'—using a molecular random approach. Biophys Chem 96:213–228

Schönfeld D et al (2009) An engineered lipocalin specific for CTLA-4 reveals a combining site with structural and conformational features similar to antibodies. Proc Natl Acad Sci USA 106:8198–8203

Singh RP, Brooks BR, Klauda JB (2008) Binding and release of cholesterol in the Osh4 protein of yeast. Proteins 75:468–477. doi:10.1002/prot.22263

Singh UC, Kollman PA (1984) An approach to computing electrostatic charges for molecules. J Comp Chem 5:129–145

Skerra A (2007) Alternative non-antibody scaffolds for molecular recognition. Curr Opin Biotechnol 18:295–304. doi:10.1016/j.copbio.2007.04.010

Skerra A (2008) Alternative binding proteins: Anticalins—harnessing the structural plasticity of the lipocalin pocket to engineer novel binding activities. FEBS J 275:2677–2683. doi:10.1111/j.1742-4658.2008.06439.x

Sprik M, Ciccotti G (1998) Free energy from constrained molecular dynamics. J Chem Phys 109:7737–7744

van Gunsteren WF, Berendsen HJC (1987) Groningen molecular simulation (GROMOS) library manual. Biomos, Groningen

Wang W, Wang J, Kollman PA (1999) What determines the van der Waals coefficient b in the LIE (linear interaction energy) method to estimate binding free energies using molecular dynamics simulations? Proteins Struct Funct Genet 34:395–402

Zanotti G, Berni R, Monaco HL (1993) Crystal structure of liganded and unliganded forms of bovine plasma retinol-binding protein. J Biol Chem 268:10728–10738

Zanotti G, Panzalorto M, Marcato A, Malpeli G, Folli C, Berni R (1998) Structure of pig plasma retinol-binding protein at 1.65 Å resolution. Acta Crystallog D 54:1049–1052

Zanotti G, Calderone V, Beda M, Malpeli G, Folli C, Berni R (2001) Structure of chicken plasma retinol-binding protein. Biochim Biophys Acta 1550:64–69

Zwanzig RW (1954) High-temperature equation of state by a perturbation method. 1. nonpolar gases. J Chem Phys 22:1420–1426