BIOINFORMATICS ORIGINAL PAPER

Vol. 30 no. 3 2014, pages 326–334
doi:10.1093/bioinformatics/btt675

Structural bioinformatics

Advance Access publication November 22, 2013

Coarse-grained versus atomistic simulations: realistic interaction free energies for real proteins

Ali May[1,2,3,†], René Pool[1,2,4,5,†], Erik van Dijk[1,2], Jochem Bijlard[1,2], Sanne Abeln[1,2], Jaap Heringa[1,2,4] and K. Anton Feenstra[1,2,4,*]

[1]Centre for Integrative Bioinformatics (IBIVU), VU University Amsterdam, [2]Amsterdam Institute for Molecules Medicines and Systems (AIMMS), VU University Amsterdam, [3]Department of Preventive Dentistry, Academic Centre for Dentistry Amsterdam (ACTA), University of Amsterdam and VU University Amsterdam, [4]Netherlands Bioinformatics Centre (NBIC), Geert Grooteplein 28 6525 GA Nijmegen, The Netherlands and [5]Department of Biological Psychology, VU University Amsterdam, 1081 HV Amsterdam, The Netherlands

Associate Editor: Anna Tramontano

ABSTRACT

Motivation: To assess whether two proteins will interact under physiological conditions, information on the interaction free energy is needed. Statistical learning techniques and docking methods for predicting protein–protein interactions cannot quantitatively estimate binding free energies. Full atomistic molecular simulation methods do have this potential, but are completely unfeasible for large-scale applications in terms of computational cost required. Here we investigate whether applying coarse-grained (CG) molecular dynamics simulations is a viable alternative for complexes of known structure.

Results: We calculate the free energy barrier with respect to the bound state based on molecular dynamics simulations using both a full atomistic and a CG force field for the TCR–pMHC complex and the MP1–p14 scaffolding complex. We find that the free energy barriers from the CG simulations are of similar accuracy as those from the full atomistic ones, while achieving a speedup of >500-fold. We also observe that extensive sampling is extremely important to obtain accurate free energy barriers, which is only within reach for the CG models. Finally, we show that the CG model preserves biological relevance of the interactions: (i) we observe a strong correlation between evolutionary likelihood of mutations and the impact on the free energy barrier with respect to the bound state; and (ii) we confirm the dominant role of the interface core in these interactions. Therefore, our results suggest that CG molecular simulations can realistically be used for the accurate prediction of protein–protein interaction strength.

Availability and implementation: The python analysis framework and data files are available for download at http://www.ibi.vu.nl/downloads/bioinformatics-2013-btt675.tgz.

Contact: k.a.feenstra@vu.nl

Supplementary information: Supplementary data are available at *Bioinformatics* online.

Received on June 17, 2013; revised on October 24, 2013; accepted on November 16, 2013

1 INTRODUCTION

Protein–protein interactions (PPI) are at the heart of all processes in life. To understand living systems beyond the genome, comprehensive knowledge of PPI is therefore essential. Experimental techniques (Ezkurdia *et al*., 2009; Kastritis and Bonvin, 2010; Sprinzak *et al*., 2003), prediction from sequence (Ezkurdia *et al*., 2009) and protein–protein docking methods (Kastritis and Bonvin, 2010; Pons *et al*., 2010) all have their specific limitations. To assess the likelihood of two proteins interacting under physiological interactions, we need to know both the concentrations and the dissociation constant (or binding free energy) of the proteins involved. Although it seems that the identification of the interface region is rather successful (Lensink and Wodak, 2010; Ofran and Rost, 2007), major open challenges are the accurate determination of interaction strength (Kastritis and Bonvin, 2010; Pons *et al*., 2010; Schueler-Furman *et al*., 2005), the incorporation of protein flexibility (Schueler-Furman *et al*., 2005; Tobi, 2010; Wollacott *et al*., 2007) and accounting for water and small solute entropic effects (Oshima *et al*., 2011; Schueler-Furman *et al*., 2005). Most importantly, Kastritis and Bonvin (2010) show that there is a poor correlation between binding affinity and scores for all nine commonly used docking algorithms they tested on 81 complexes with known binding affinity.

Molecular simulations using atomic pairwise interaction potentials are much more accurate for estimating interaction strength than docking scoring functions, though computationally much more expensive (Kastritis and Bonvin, 2010; Tuncbag *et al*., 2009). Nevertheless, an immediate bonus of molecular simulation is that it addresses all three challenges mentioned above: interaction strength, flexibility and entropic effects. For biomolecular simulation in general, the solvent (water) is the major obstacle to improve computational efficiency due to the large number of water molecules needed to solvate the protein. Many possible approaches to overcome this problem exist (for recent reviews, see Dror *et al*., 2012; Fennell and Dill, 2011). Among the fastest available are the mean-field or 'implicit solvent' methods; however, one of the main drawbacks is the lack of accurate estimation of the solute entropy, especially in combination with charged solutes (Homeyer and Gohlke, 2012). By lumping together small molecules (e.g. water molecules) or

*To whom correspondence should be addressed.
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

molecular segments into 'meta particles', coarse-grained (CG) force fields retain the explicit description of the system, including the solvent. Several CG models for water are available, each with their particular strengths and weaknesses (Hadley and McCabe, 2012). Compared with atomistic force fields, CG models provide increased computational efficiency at sufficient levels of accuracy (Tuffery and Derreumaux, 2012). Therefore, for the calculation of molecular interactions, thermodynamic integration based on atomistic simulations with explicit water is both theoretically well founded and the most accurate solution available to date (Wereszczynski and McCammon, 2012, Zhang *et al.*, 2013).

In this work, we investigate the option of addressing the high computational cost of atomistic molecular simulations by the use of a CG force field for such simulations. For this, we will use the MARTINI CG protein force field for molecular dynamics (MD) simulations (Marrink *et al.*, 2007; Monticelli *et al.*, 2008). This force field was developed for CG simulations of biological membrane–protein systems and has recently been used to simulate the spontaneous association of GPCR proteins in a lipid bilayer (Periole *et al.*, 2012). The MARTINI force field does not capture structural rearrangements, such as changes in secondary structure. We will assess its applicability and accuracy for the calculation of interaction strengths for a pair of protein structures in a water environment from constraint-force profiles.

By applying the MARTINI CG force field, we first show that we are able to estimate the free energy barrier with respect to the bound state (ΔG^{off}) with similar accuracy compared with atomistic force field calculations and in good agreement with experimental data, but at >500-fold increased computational speed. For this we selected two test cases: a TCR–pMHC and an MP1–p14 scaffolding complex, both of which were previously studied with atomistic models (Cuendet and Michielin, 2008; Cui *et al.*, 2008, respectively). We then show that calculated contributions of surface residues to the interaction strength are sensitive to changes in the amino acid residues involved. Random mutations at the interface core yield major changes in calculated interaction strengths, whereas mutations at the partially solvated interface rim yield only minor changes. Random mutations at the rest of the surface on average hardly affect the interaction strength at all. Moreover, we find that the evolutionarily most likely mutations, as assessed by standard residue exchange propensities, at the interface core also have a negligible influence on the interaction strength, whereas evolutionarily unlikely mutations disrupt favourable PPI considerably. This behaviour with respect to mutations is consistent with what we would expect from a biological point of view. Finally, we discuss future implications of our finding that the major contributions to the interaction strength within our CG approach arise from the interface core.

2 METHODS

2.1 Software and force fields

The `mutate_model` script in Modeller (Sali and Blundell, 1993) was used to produce mutant structures. DSSP (Kabsch and Sander, 1983) and JOY (Mizuguchi *et al.*, 1998) were used to calculate the absolute and relative solvent accessibility of residues, respectively. VMD (Humphrey *et al.*, 1996) was used to visualize the structures.

We used GROMACS 4.0.5 (Hess *et al.*, 2008) for all MD simulations. Atomistic simulations were performed using the GROMOS G43a1 force field using the default time step ($\Delta t = 1$ fs) (van Gunsteren *et al.*, 1996). CG simulations were performed using the MARTINI force field with the default time step ($\Delta t = 20$ fs) (Marrink *et al.*, 2007). Coarse-graining (CG-ing) was performed as previously described for the MARTINI model (Monticelli *et al.*, 2008). All 20 amino acids were mapped into four different bead types with respect to their physicochemical properties (Supplementary Fig. S1). The non-bonded interactions between the CG solvent and solute particles were modelled by truncated and shifted Lennard–Jones pair-potential with a cut-off radius of 1.2 nm (Marrink *et al.*, 2007; Monticelli *et al.*, 2008).

2.2 The potential of mean force

We use the potential of mean force (PMF) to describe the interaction strength between two structures (Trzesniak *et al.*, 2007). The centre of mass (COM) separation r was chosen as the reaction coordinate along which the mean force is measured. Integration of the mean force along this pathway results in a free energy profile (strictly, the PMF is not a free energy profile, as it does not correct for standard conditions; this is covered in the Supplementary Information) that can be used to derive the free energy barrier with respect to the bound state ΔG^{off}. We first calculate the force F_{mean} as a function of the reaction coordinate from constrained MD simulations,

$$F_{\text{mean}}(r) = -\langle F_{\text{pull}}(r)\rangle_{NPT} = \frac{1}{2}\left\langle\left(\vec{F}_{\text{B}} - \vec{F}_{\text{A}}\right)\cdot\vec{r}_{\text{u}}\right\rangle_{NPT} \quad (1)$$

where $\langle F_{\text{pull}}(r)\rangle$ denotes the average force required to keep the interaction members at the constraint distance r, $\vec{F}_{\text{A}}$ and $\vec{F}_{\text{B}}$ the total forces acting on the first and the second interaction members, which arise from direct interactions and interactions with explicitly simulated solvent, $\vec{r}_{\text{u}} = \vec{r}/r$ the unit vector connecting the two centres of mass and angular brackets $\langle\cdots\rangle_{NPT}$ an average in the isothermal–isobaric ensemble.

We define the constraint distance r as

$$r = \left|r_{\text{COM,A}} - r_{\text{COM,B}}\right| \text{ and } r \in \{r_1, r_2, ..., r_N\} \quad (2)$$

where $r_{\text{COM,i}}$ is the COM position of interaction member i and N is the number of separation distances at which the F_{mean} values are calculated. Three arbitrary separations are illustrated in Figure 1C for MP1–p14.

We calculated the F_{mean} at 50 distances for the TCR–pMHC (where $5\,\text{nm} \leq r \leq 7.45\,\text{nm}$) and at 54 distances for MP1–p14 ($2.16\,\text{nm} \leq r \leq 4.44\,\text{nm}$). In cases where we simulated nearly identical starting conformations of a particular structure for better sampling at distance r, we included these F_{mean} values into the average of $F_{\text{mean}}(r)$. After generating the force profile $F_{\text{mean}}(r)$ for the range of separations, we calculated the PMF by numerically integrating the interpolated $F_{\text{mean}}(r)$ as

$$\text{PMF}(r) = -\int_0^r dr' F_{\text{mean}}(r') \quad (3)$$

From this profile, the free energy barrier ΔG^{off} is obtained from the difference between the minimum of PMF(r) at r_{min} and the maximum PMF value at larger distances $r > r_{\text{min}}$:

$$-\Delta G^{\text{off}} = \min[\text{PMF}(r)] - \max[\text{PMF}(r > r_{\text{min}})] \quad (4)$$

Errors in the forces are estimated from the standard deviations of the forces $\sigma_{F_{\text{mean}}}(r)$ across the set of replicate simulations at each distance r, and errors in the PMF $\sigma_{\text{PMF}}(r)$ are subsequently derived as follows:

$$\sigma_{\text{PMF}}(r) = \sqrt{\int_r^0 dr' \sigma^2_{F_{\text{mean}}}(r')} \quad (5)$$

2.3 Simulation setup

The wild-type (WT) X-ray structures of the TCR–pMHC (Garboczi *et al.*, 1996) and MP1–p14 (Kurzbauer *et al.*, 2004) complexes, resolved at 2.6 and 1.9 Å, were taken from the Protein Data Bank entries `1ao7` and `1vet`, respectively. The TCR–pMHC structure contains 707

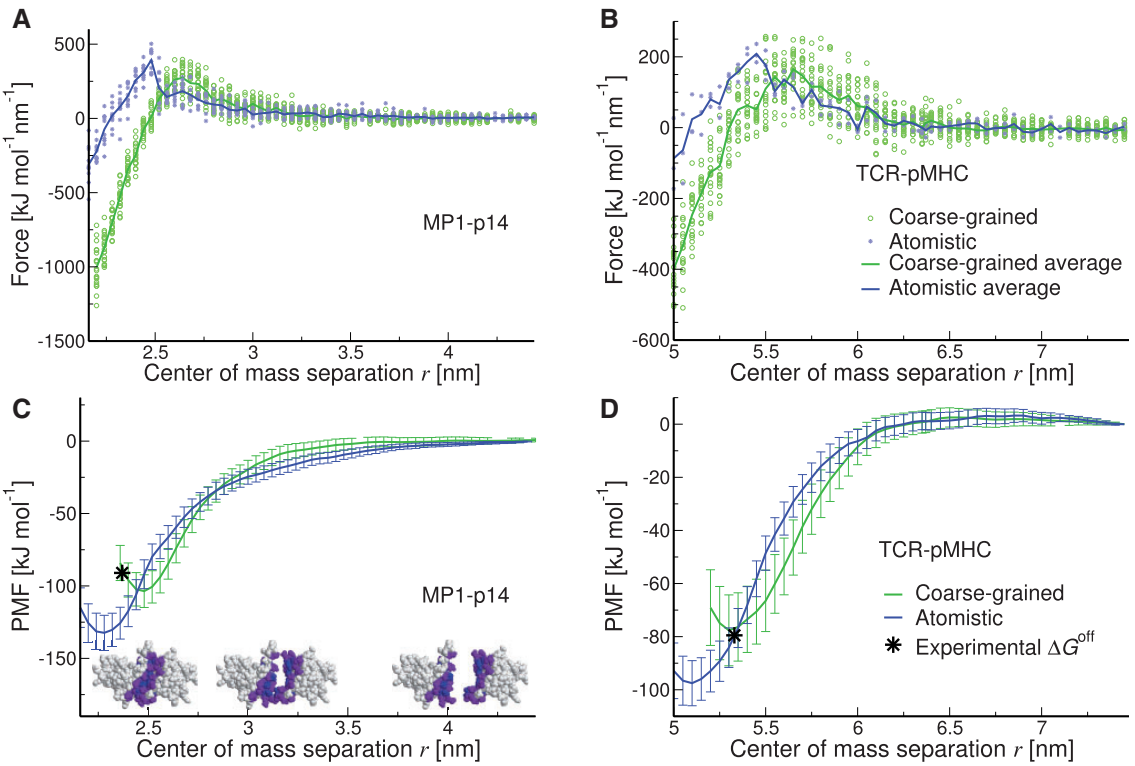

Fig. 1. The mean force profile and the PMF from atomistic and CG calculations. (Top) Identical symbols at a given distance are calculations of the mean force from simulations that differ slightly in their starting conformations. Blue and green lines describe the means of these values at each distance, respectively for atomistic and CG simulations. (Bottom) PMF describing the free energy of dissociation. Distances >1 Å below the PMF minimum are not shown. Blue and green lines describe the PMF obtained by atomistic and CG simulations, respectively. The corresponding free energy profiles are shown in Supplementary Figure S3. Stars show experimentally determined ΔG^{off} and equilibrium distances in the native crystal structures. Error bars are based on standard deviations in the mean force, cf. Equation (5). (**A** and **C**) MP1–p14 and (**B** and **D**) TCR–pMHC complex

residues, of the human A6 TCR in complex with the MHC-bound Tax nanopeptide. The MP1–p14 complex contains 240 residues, of two structurally very similar chains of low sequence similarity, with a large and shallow interface.

2.3.1 Equilibration procedure

For the atomistic and CG simulations, a general equilibration scheme was performed, which was identical for atomistic and CG except where noted below. First, the structure was energy-minimized in vacuum, followed by the separation of the interacting proteins to the constraint distance r along the reaction coordinate. These structures, for each r, were solvated in a periodic cubic box with a size ensuring a minimum distance between the proteins and the box edges to avoid self-interactions with periodic images. For the atomistic simulations, the SPC water model and a minimum distance of 1.2 nm were used. For the CG simulations, MARTINI water was used with a 1.0 nm minimum distance. Energy minimization on the solvated system was performed first with, and then without, position restraints. In the atomistic energy minimization, position restraints were first put on non-H atoms, followed by restraining only C_α atoms. The system was neutralized by adding as many Na^+ or Cl^- counter ions as needed for a total charge of 0. Then, another unrestrained energy minimization was performed. The neutralized and solvated structure was simulated for 20 ps (atomistic) or 30 ps (CG) with position restraints on the C_α atoms (atomistic) or whole structure (CG), to allow the solvent to equilibrate around the solute. The temperature was set to $T = 303$ K using a Berendsen (Berendsen *et al.*, 1984) (CG) or a Nosé–Hoover (Cheng and Merz, 1996) (atomistic) thermostat with $\tau_T = 0.1$ ps. The pressure in atomistic simulations was set to $P = 1$ bar using a Parrinello–Rahman barostat (1981) with $\tau_P = 0.5$ ps. In CG simulations, pressure was set to $P = 300$ bar, using a Berendsen barostat (Berendsen *et al.*, 1984) with $\tau_P = 0.5$ ps. The high pressure in the CG simulations was applied to ensure the bulk MARTINI CG water is in the fluid region of the phase diagram. We determined the MARTINI CG water phase diagram from separate Gibbs ensemble simulations (Frenkel and Smit, 2002). Temperature and pressure were equilibrated for 0.1 ns (atomistic) or 0.2 ns (CG). The resulting conformations were used in the production simulations without position restraints.

2.3.2 Atomistic production simulations Atomistic production MD simulations were performed for the WT TCR–pMHC (3 replicates) and MP1–p14 (10 replicates). Equilibration procedure and production runs were repeated for each distance, cf. Equation (2). Each production simulation was run for 2 ns.

2.3.3 CG production simulations CG production MD simulations were performed for the WTs and *in-silico* mutants (20 replicates). Before starting the CG equilibration procedure, the atomistic structure was first energy minimized and then coarse-grained using `atom2cg` (http://md.chem.rug.nl/cgmartini). The tertiary structure of the CG complex was stabilized by generating distance restraints on the backbone atoms (Marrink *et al.*, 2007). This equilibration procedure and the production runs were repeated for each distance in Equation (2). Production simulations were run for 2 ns or 2 μs as indicated.

2.4 Definition of different residue classes

The relative and absolute solvent accessible surface areas (SASAs) of the residues were calculated in the monomer (e.g. TCR) and dimer

forms (e.g. TCR–pMHC). Subsequently, residues that have <7% of their side chain accessible to the solvent in both forms were defined as protein core. Residues that have >7% of their side chain exposed to the solvent in the monomer and <7% in the dimer were defined as interface core. Residues that have >7% of their side chain accessible to the solvent in both forms were defined as interface rim if there was a difference of at least 1 Å^2 in their absolute SASA between two forms. Residues that do not fall into one of the above groups were called surface residues. The outer interface rim class is composed of the closest surface residue neighbours of interface rim residues located on the same monomer. The number of outer rim residues identified in a complex is the same as the number of rim residues.

2.5 *In silico* mutations and statistical analysis

To probe the biological relevance of the CG calculations, we introduce mutations guided by the BLOSUM62 substitution matrix (Henikoff and Henikoff, 1992) in the MP1–p14 complex. We mutated all 23 interface core residues to the amino acid with the highest substitution score (other than itself) to obtain the evolutionarily most likely mutant (see Supplementary Table S6). In a separate set of mutations, we mutated the interface core to the lowest-scoring substitutions to obtain the least likely mutant (see Supplementary Table S6). We performed 20 replicate simulations for both mutants, each starting from slightly different starting conformations. Next, we also made all 23×19 possible interface core mutants and simulated each 10 times.

To further probe biological relevance, we compare the effects of random mutations at different locations at and around the interface for both the MP1–p14 and the TCR–pMHC complexes. We created 20 compositionally different interface core mutants where all n interface core residues (see Supplementary Table S2) were substituted with randomly chosen amino acids. Then, we created 20 different interface rim mutants by randomly substituting n interface rim residues with random amino acids. Next, we obtain 20 outer interface rim mutants by randomly substituting the n non-interface neighbours of the interface rim. We calculated the PPI for each of these mutants only once rather than multiple replicates. Differences in the free energy barrier $\Delta\Delta G$ were statistically tested with a two-sided Mann–Whitney test separately for every mutant and for the WT.

Finally, we made specific interface mutations in a homologue of the TCR–pMHC complex described by Wu *et al*. (2002). We chose the seven mutants with the largest measured $\Delta\Delta G$s: in the beta chain Q64A, E69A and A73G, and in the peptide K99R, T102N, T102S and the double mutant Y97F/T102S.

3 RESULTS

3.1 Calculation of free energy barrier with respect to the bound state

To obtain the free energy barriers to compare atomistic and CG results, we calculate the PMF (Fig. 1C and D). This can be derived from the mean force F_{mean} cf. Equation (1) required to constrain interaction members (e.g. TCR and pMHC) at a number of COM separation distances (Fig. 1A and B).

Several immediate observations can be drawn from Figure 1. First, force profiles obtained by the atomistic and the CG model are in reasonable agreement for both complexes (Fig. 1A and B). Second, at a given separation distance r, simulations of nearly identical starting conformations (identical symbols in Fig. 1A and B on the distance r) yield a distribution of force values in both atomistic and CG simulations rather than converging to some value. Finally, these distributions overlap closely for distances larger than ~2.75 nm (Fig. 1A) and ~5.5 nm (Fig. 1B), but diverge at shorter distances.

Next we calculated the average mean force at each distance to obtain the force profiles $F^{\text{atom}}_{\text{mean}}(r)$, in blue, and $F^{\text{CG}}_{\text{mean}}(r)$, in green in Figure 1A and B. We calculated the PMF by numerically integrating the interpolated $F_{\text{mean}}(r)$ cf. Equation 3 (Fig. 1C and D). The free energy barrier can now simply be obtained from the well-depth of the PMF, cf. Equation (4). The resulting PMFs shown in Figure 1C and D for both complexes provide a comparison between the free energy minima calculated from the atomistic and the CG simulations, as well as with the experimentally determined interaction strengths.

For the MP1–p14 complex, atomistic and CG simulations yielded free energy barriers $\Delta G^{\text{off}} = 132$ and 104 kJ mol^{-1}, respectively, both overestimating the reported experimental $\Delta G^{\text{off}} = 91$ kJ mol^{-1} (Kurzbauer *et al*., 2004). For the TCR–pMHC complex, atomistic and CG simulations yielded free energy barriers $\Delta G^{\text{off}} = 101$ and 80 kJ mol^{-1}, respectively, close to the experimental values $\Delta G^{\text{off}} = 79.5$ kJ mol^{-1} (Ding *et al*., 1999) and $\Delta G^{\text{off}} = 78.6$ kJ mol^{-1} (Davis-Harrison *et al*., 2007 and see Supplementary Table S1 for further details). As noted above, a very small correction must be made to the free energy barrier (Gilson *et al*.,1997), which amounts to ~4 kJ mol^{-1} (see Supplementary Data for a detailed calculation). To probe the effect of sampling of the rotational degrees of freedom, for the MP1–p14 complex, we also calculated the PMF based on a triplicate set of long 2 μs CG simulations. Supplementary Figure S4 shows ordering is present in the short (2 ns) simulations, as well as at close distances in the long (2 μs) simulations; however, this ordering disappears at farther distances (>4 nm) in the long simulations, indicating a strongly improved rotational sampling. We found a negligible change in the ΔG^{off} as a result of this improved sampling.

Besides interaction strength, Figure 1C and D presents a comparison between the experimentally reported PPI equilibrium distance (the separation distance in X-ray structure) and the corresponding value from simulations (the distance at the PMF minimum). We see that the experimental distances were slightly underestimated by the atomistic PMF in both cases, whereas the CG PMF appears to yield distances closer to the crystal structure, especially in the case of TCR–pMHC interaction.

3.2 Computational speed-up

The running time of atomistic MD simulations for MP1–p14 at the longest COM separation distance $r = 4.44$ nm was 284 CPU-hours (CPUh). CG-ing reduced the running time at this distance to 0.5 CPUh, yielding a 568-fold speed-up. Shorter separations are progressively faster, but show similar speed-ups. In the case of the TCR–pMHC, this was 1333 and 2.5 CPUh, respectively, for a speed-up of 533-fold. In total, we invested well over 300 000 CPUh in the atomistic simulations, over the 3 and 10 replicate calculations for the TCR–pMHC and MP1–p14 complexes, respectively. In contrast, the CG simulations for the 456 different mutants (10 or 20 replicates each) only required ~100 000 CPUh. The 2 μs CG simulations (three replicates) required an additional 64 000 CPUh.

3.3 Evolutionary likelihood of mutations and the free energy

To obtain a simple, although biologically relevant test case, we substituted interface core residues in MP1–p14 according to BLOSUM62 (Henikoff and Henikoff, 1992). We define two sets of mutations; one where each residue is mutated into the evolutionarily most distant residue, and one where it is mutated into the closest residue type (detailed substitutions in Supplementary Table S6). One would expect minimal distortion of the PMF from the closest substitutions, and maximal distortion for the evolutionarily most distant substitutions.

It is clear from the results shown in Figure 2A that when the interface core composition was altered with the most similar amino acids, change in the interaction strength remains insignificant ($\Delta\Delta G^{\text{off}} = 7\,\text{kJ mol}^{-1}$, $P < 0.38$; see Supplementary Table S5 for detailed significance values), whereas substitutions with the most dissimilar amino acids have a strongly disruptive effect on the PPI that is highly significant ($\Delta\Delta G^{\text{off}} = 71.5\,\text{kJ mol}^{-1}$, $P < 10^{-11}$; see Supplementary Table S5). However, we still observe a favourable interaction ($\Delta G^{\text{off}} = 32 \pm 13\,\text{kJ mol}^{-1}$) with the most dissimilar mutant structure. Note that the force field alone has no such predictive value (see Supplementary Fig. S7).

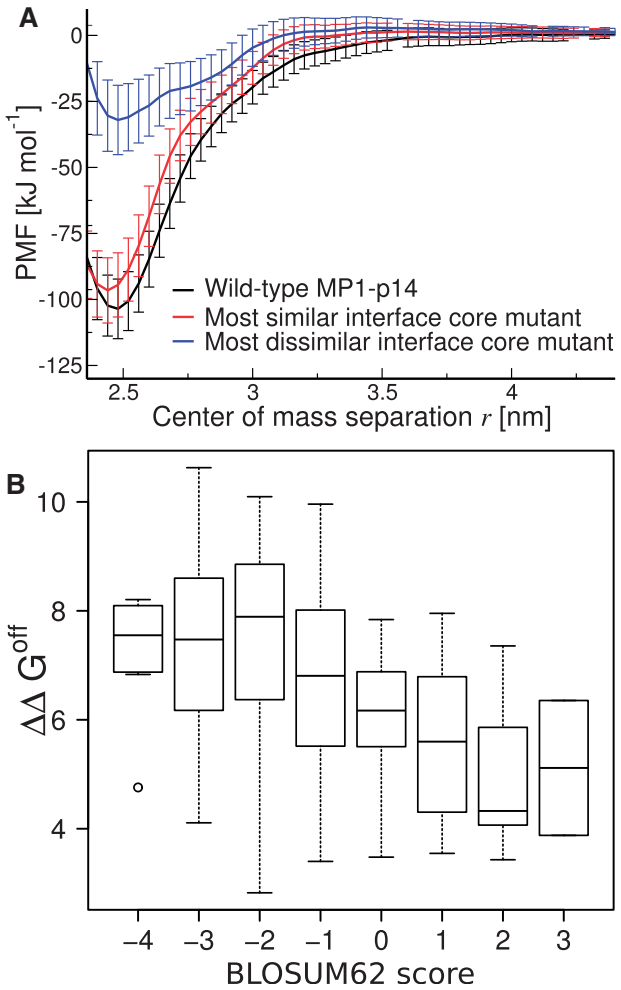

Fig. 2. The PPI effects of different mutations at the MP1–p14 interface. (**A**) PMF curves of the WT, the most similar and dissimilar mutants in black, red and blue, respectively. Error bars are based on standard deviations in the mean force, cf. Equation (5). (**B**) Box-plot of average $\Delta\Delta G^{\text{off}}$ per substitution type, based on pairwise differences between single mutations at the interface core, versus the corresponding BLOSUM62 substitution score, with a correlation $R = -0.39$ and significance $P \leq 2.2 \cdot 10^{-8}$

We have in addition considered all single mutations for each of the 23 interface core residues for MP1–p14. Of all $\Delta\Delta G^{\text{off}}$s with respect to the WT, 16 are significant ($P \leq 0.05$ using Student's t-test with Hommel multiple testing correction; 121 at $P \leq 0.05$ with just t-test). This corresponds to a detection limit of about $\Delta\Delta G^{\text{off}} \geq 8\,\text{kJ mol}^{-1}$. Figure 2B compares $\Delta\Delta G^{\text{off}}$ between all mutants (excluding the WT as the rest of the protein structure is biased towards the WT amino acid) with the corresponding BLOSUM62 scores, and shows that there is a strong correlation between them, as expected. In a homologue of the TCR–pMHC complex, we furthermore show that we can reproduce experimental ΔG^{off}s for interface mutations to within the accuracy of our calculations (see Supplementary Fig. S7).

3.4 The interface core dominates the interaction

Our final *in silico* experiment with the CG model was aimed at investigating the effective role of interface residues in the interaction between two proteins. We first defined classes of residues based on SASA and distance (Fig. 3, for definitions see Section 2, a summary is provided in Supplementary Table S2).

We substituted the same number of residues from each class with randomly chosen amino acids, and calculated the interaction strength in the resulting mutant complexes. The resulting PMFs shown in Figure 4 indicate significant disruptive effects of mutations at the interface core in both the MP1–p14 ($P < 10^{-8}$; see Supplementary Table S3 for detailed significance values) and TCR–pMHC ($P < 2 \cdot 10^{-4}$; see Supplementary Table S4), whereas mutations in the interface rim appear to have very little influence on the PMF compared with the WT. Note that in both complexes, mutations of outer rim sometimes yielded enhanced interaction strengths.

4 DISCUSSION AND CONCLUSION

In this article, we have set out to answer the following questions. First, can we use CG MD simulations to get the free energy barrier with respect to the bound state for PPI that are of similar accuracy to those from atomistic MD simulations? Second, what is the gain in speed and overall sampling for CG versus atomistic? Finally, can we get biologically relevant results from the CG model, similar to what we expect to get from the atomistic model (at much higher computational cost)?

4.1 Biological relevance

We have shown that by using the CG MARTINI force field, we obtain free energy barriers that are at least as close to the experimental values as those obtained using an atomistic force field. Moreover, evolutionarily least likely mutations, according to BLOSUM62 substitution propensities, at the interface core disrupt binding, whereas the most likely mutations hardly influence

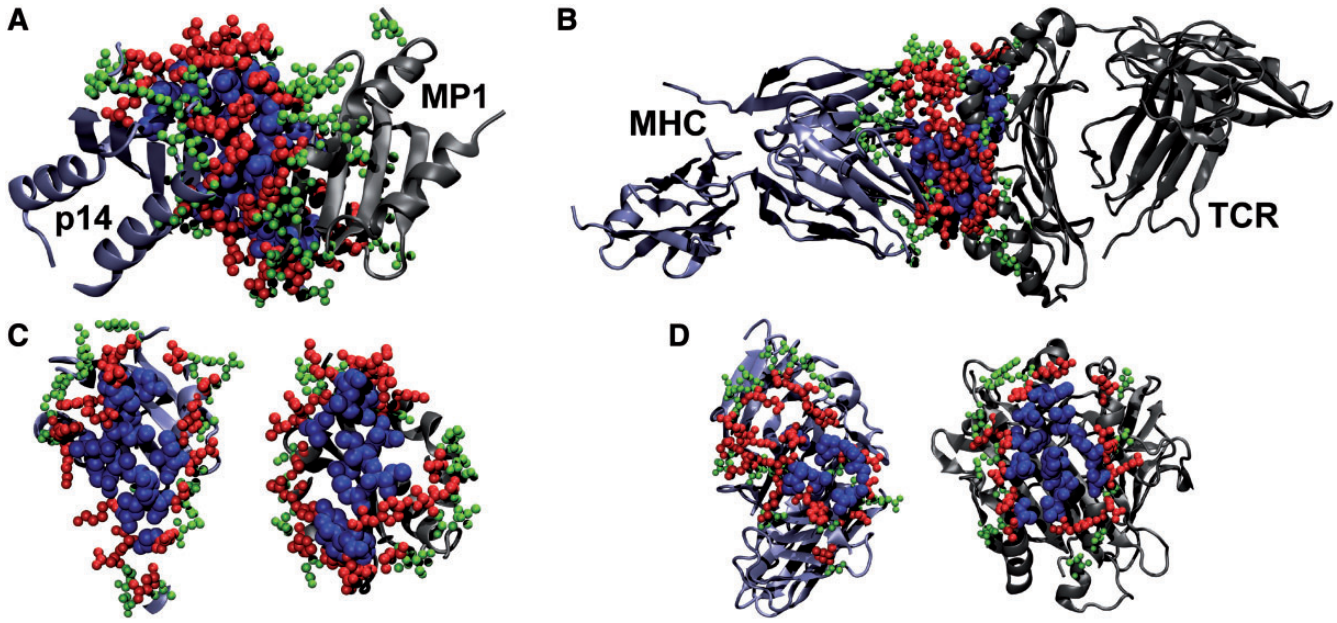

Fig. 3. Detailed view of the interface regions of interacting proteins. (**A**) MP1–p14 and (**B**) TCR–pMHC showing the interface core residues in blue, interface rim in red and outer interface rim in green in VDW representation. The remaining proteins are shown as cartoons. (**C**) and (**D**) show the interfaces 'opened up' by rotation outward by 90° around the vertical axes to expose the interacting 'faces' using the same colour as in (**A**) and (**B**)

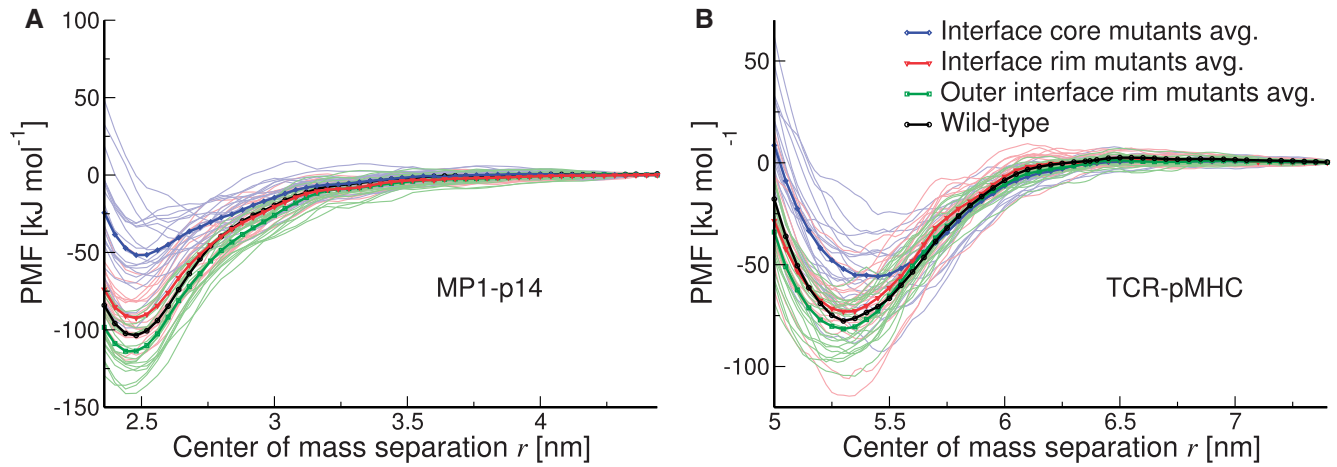

Fig. 4. The effect of mutations in different regions of the interacting pairs. The interface core mutants are in blue, interface rim in red, outer interface rim in green and the WT in black. Pale colours on the background are for each individual mutant and darker colours indicate the mean PMF for each mutant class. (**A**) MP1–p14 and (**B**) TCR–pMHC complex. Supplementary Figure S5 shows corresponding plots with only one of three residues mutated

binding. It is important to note here that the residue similarities from large-scale sequence comparisons, as captured in the BLOSUM62 matrix, are independent of the MARTINI CG force field.

Likewise, mutations on the interface core have a much more disruptive effect on the binding than those at the interface rim. Mutations farther away from the interface have negligible effects. This is consistent with the findings that residues at the interface core in general behave much like protein core, whereas interface rim residues are more similar to those at the surface (e.g. Tuncbag *et al*., 2009).

These results show that the CG MARTINI force field is sensitive to changes in the shape and physicochemical properties of the interface, and suggest it is suitable for studying the effects of biologically relevant structural changes in PPIs.

Of note, we observe that even the evolutionarily least likely mutations in the interface core still result in a minimum well in the PPI PMF, which means that some (weak) affinity is retained. This may be due to retained shape compatibility of the mutant interface or due to favourable 'supporting' interactions of residues in the interface rim. This could also explain our finding that least likely mutations that are performed only on one side of the interface (see Supplementary Table S5) have much smaller disruptive effects on the binding.

We also noticed that some outer interface rim mutations significantly lower the well depth; note that these residues do

not participate in the interaction in the WT complex. This increased binding strength could be caused by additional interactions that are introduced by these mutations. We found that the most significantly increased binding strength occurred in mutants with a higher net charge. This could mean that either additional salt-bridges contribute to the binding, or that possibly (additional) counter-ions interact at the interface to increase the binding strength. Alternatively, the interface rim is known to infer specificity through complementarity of shape and binding properties (Guharoy and Chakrabarti, 2005), and is therefore not tuned for optimal binding strength. Likewise, the interface as a whole may not be evolutionarily optimized for binding affinity, as other functional aspects are also likely to give selective advantages. This means that there may in fact be ample room for optimization of the binding affinity in any particular naturally occurring protein–protein interface.

Finally, we do see qualitative differences between the atomistic and CG PMFs, although our set of two complexes does not allow us to decide which one is in better agreement with experimental dissociation constants. Underlying reasons for these differences are likely related to the different parameterization of short-range interactions in these force fields. The difference between the experimental and computed equilibrium distances can have a number of explanations. The crystal structures represent vacuum systems in which the protein structure is very rigid, whereas we simulated systems of solvated 'breathing' proteins. Furthermore, the MARTINI force field includes restraints on the tertiary structure of the protein, whereas the unrestrained atomistic simulations will allow larger deviations in the interface on binding. This may explain why equilibrium binding distance in the CG calculations is closer to the crystal structure than the atomistic simulations. In any case, a comparison of the experimental equilibrium COM distances to the ones we computed is expected to yield differences.

4.2 Sampling is crucial

Surprisingly, for both complexes, the CG model appears to be at least as accurate in approximating the experimental value as the atomistic model. Part of the reason for that could be better sampling of the mean force for the CG model where the faster simulations enabled us to perform 20 independent simulations for each separation distance, whereas only 3 (TCR–pMHC) or 10 (MP1–p14) simulations could realistically be performed for the atomistic model.

We also find that, for a given separation distance, simulations from slightly different starting conditions do not converge to the same value of the constraint force (see Fig. 1A and B, $*$ and $\circ$ symbols) for both the CG and the atomistic models, and this effect is strongest at short separation distances. We attribute this observation to the complex potential energy surface at small separation distances. Simulations of biomolecules starting from nearly identical initial conditions are known to get trapped in local minima (Luo *et al.*, 2006). In general, this can be overcome as well with sufficient sampling.

The reason we could perform so many more simulations for the CG model is of course its increased computational efficiency, 500–600 times faster than the atomistic simulations. The primary reason for this large computational gain is the reduction in particle density, and the ensuing quadratic decrease in numbers of pairwise interactions calculated. Moreover, the larger heavier particles with softer interaction potentials in the CG force field allow much longer integration time-steps without loss of accuracy than is possible with the atomistic force field. Therefore, we can conclude that, for all but very small-scale analyses, the amount of sampling required for accurate determination of the ΔG^{off} barrier can only realistically be achieved using CG models.

Some further efficiency could possibly be gained by optimizing simulation parameters, particularly the integration time-step. For the atomistic simulations, time-steps of up to $\Delta t = 5$ or 6 fs may be used with negligible loss of accuracy (Feenstra *et al.*, 1999). A similar speed-up may likewise be achieved in the CG simulations.

One final challenge regarding the accurate calculation of the free energy of binding in particular remains largely open: that of the accurate estimation of loss of rotational entropy on formation of a protein complex (Chang *et al.*, 2008; Grunberg *et al.*, 2006; Tamura and Privalov, 1997; Yu *et al.*, 2001). In the results presented here, at 2 ns sampling, the rotational entropy is not fully sampled (see Supplementary Fig. S4). For the free energy difference between the bound state and the unbound state, this would result in a discrepancy between the simulation and the experiment. However, at the top of the barrier, this error is expected to be much smaller (Cuendet and Michielin, 2008); our 2 μs simulations for the MP1–p14 complex confirm that this effect can be extremely small. Furthermore, the correction is expected to be independent of the details of the force field used, and therefore we can directly compare the results between the atomistic and CG simulations.

It is interesting to note that we could use the increased efficiency in the CG simulations to directly sample rotational degrees of freedom, as indicated by the complete loss of orientational ordering at 2 μs and large separation distance (see Supplementary Fig. S4). It is beyond the scope of this work, but once sufficient sampling is established, these results can be used to estimate directly the changes in entropy during the bound to unbound transition in complexes like this one.

4.3 Limitations of the approach

In this article, we compared two force fields on two complexes of known structure. We have shown that for these two complexes, the interaction profiles are highly similar when comparing the atomistic force field to the CG MARTINI force field. To claim generality of these findings, a larger test set may be required. However, the computational requirement for the reference simulations using the atomistic force field is prohibitive. Additionally none of the current methods in docking or simulation can predict binding affinities without knowledge of the bound structure (Kastritis and Bonvin, 2010).

The comparison has been made by assuming that conformations in the free state as well as in the bound state are relatively stable. For other complexes, this may be different, e.g. when the PPI involves 'induced fit' effects, especially for highly flexible binding partners. It is unlikely that the methods proposed here will be directly applicable to such cases. It should be emphasized that the atomistic approach is equally unfeasible here, but in that

case due to limits in computational sampling. The *in silico* mutational study presented here did not account for the possibility that mutations might disrupt protein secondary structure, as these kinds of structural rearrangements are not possible in the MARTINI model.

4.4 Future implications

Our results confirm the dominance of the interface region in determining the PPI for the two complexes. This opens the possibility to restrict the simulated system to the interface area and intervening water only. However, when only a limited volume at the interface region is simulated, the number of water molecules in that volume cannot be assumed constant during force calculation. Rather, this system should be considered at a constant chemical potential (μ) with fluctuating numbers of particles (N), i.e. the grand-canonical (GC) or μVT ensemble.

Traditionally, MD simulations are performed in the microcanonical (NPE), canonical (NVT) or isothermal/isobaric (NPT) ensemble, as this simplifies calculations and the complexity of the software required (Frenkel and Smit, 2002). We have recently published a python library interface to the GROMACS simulation engine (Pool *et al*., 2012) that enables simulation of a GC μVT ensemble through a hybrid MD/Monte Carlo integration scheme.

The combined speed-up achieved by coarse-graining and volume restriction would be sufficient to incorporate the calculation of binding free energies into a three-stage approach to PPI calculation for genomic-scale application. First, non-interacting protein pairs would be filtered out using cheap sequence-based methods (e.g. Ezkurdia *et al*., 2009). Second, docking will be used to find the most likely binding interfaces (Lensink and Wodak, 2010; Pons *et al*., 2010). In a final step, binding energy calculation will then be used to select the stable complexes (Pool and Bolhuis, 2010). Although this goal, for now, remains in the future, it does seem only a few small steps away.

ACKNOWLEDGEMENTS

The authors thank Qingzhen Hou for help with the MARTINI simulations. They also thank the anonymous reviewers for helpful suggestions.

Funding: BioRange programme of the Netherlands Bioinformatics Centre (NBIC), which is supported by a BSIK grant through the Netherlands Genomics Initiative (NGI). SA has been supported by a Veni grant on the project 'Understanding Toxic Protein Oligomers through Ensemble Characteristics' from the Netherlands Organisation for Scientific Research (NWO) SURFsara (www.surfsara.nl) by way of use of the Lisa Compute Cluster.

Conflict of Interest: none declared

REFERENCES

Berendsen,H.J.C. *et al.* (1984) Molecular dynamics with coupling to an external bath. *J. Chem. Phys.*, **81**, 3684.

Chang,C.E. *et al.* (2008) Entropic contributions and the influence of the hydrophobic environment in promiscuous protein-protein association. *Proc. Natl Acad. Sci. USA*, **105**, 7456–7461.

Cheng,A. and Merz,K.M. (1996) Application of the noséhoover chain algorithm to the study of protein dynamics. *J. Phys. Chem.*, **100**, 1927–1937.

Cuendet,M.A. and Michielin,O. (2008) Protein-protein interaction investigated by steered molecular dynamics: the TCR- pMHC complex. *Biophys. J.*, **95**, 3575–3590.

Cui,Q. *et al.* (2008) Molecular dynamics-solvated interaction energy studies of protein-protein interactions: the MP1-p14 scaffolding complex. *J. Mol. Biol.*, **379**, 787–802.

Davis-Harrison,R. *et al.* (2007) T cell receptor binding transition states and recognition of peptide/MHC. *Biochemistry*, **46**, 1840–1850.

Ding,Y. *et al.* (1999) Four A6-TCR/peptide/HLA-A2 structures that generate very different T cell signals are nearly identical. *Immunity*, **11**, 45–56.

Dror,R.O. *et al.* (2012) Biomolecular simulation: a computational microscope for molecular biology. *Annu. Rev. Biophys.*, **41**, 429–452.

Ezkurdia,I. *et al.* (2009) Progress and challenges in predicting protein-protein interaction sites. *Brief Bioinformatics*, **10**, 233–246.

Feenstra,K.A. *et al.* (1999) Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-rich systems. *J. Comput. Chem.*, **20**, 786–798.

Fennell,C.J. and Dill,K.A. (2011) Physical modeling of aqueous solvation. *J. Stat. Phys.*, **145**, 209–226.

Frenkel,D. and Smit,B. (2002) *Understanding Molecular Simulation: From Algorithms to Applications*. 2nd edn. Academic Press, San Diego, CA.

Garboczi,D.N. *et al.* (1996) Structure of the complex between human t-cell receptor, viral peptide and HLA-A2. *Nature*, **384**, 134–141.

Gilson,M.K. *et al.* (1997) The statistical-thermodynamic basis for computation of binding affinities: a critical review. *Biophys. J.*, **72**, 1047–1069.

Grunberg,R. *et al.* (2006) Flexibility and conformational entropy in protein-protein binding. *Structure*, **14**, 683–693.

Guharoy,M. and Chakrabarti,P. (2005) Conservation and relative importance of residues across protein-protein interfaces. *Proc. Natl Acad. Sci. USA*, **102**, 15447–15452.

Hadley,K.R. and McCabe,C. (2012) Coarse-grained molecular models of water: a review. *Mol. Simul.*, **38**, 671–681.

Henikoff,S. and Henikoff,J.G. (1992) Amino acid substitution matrices from protein blocks. *Proc. Natl Acad. Sci. USA*, **89**, 10915–10919.

Hess,B. *et al.* (2008) Gromacs 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.*, **4**, 435–447.

Homeyer,N. and Gohlke,H. (2012) Free energy calculations by the molecular mechanics Poisson–Boltzmann surface area method. *Mol. Inform.*, **31**, 114–122.

Humphrey,W. *et al.* (1996) VMD – visual molecular dynamics. *J. Mol. Graph.*, **14**, 33–38.

Kabsch,W. and Sander,C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, **22**, 2577–2637.

Kastritis,P. and Bonvin,A. (2010) Are scoring functions in protein-protein docking ready to predict interactomes? Clues from a novel binding affinity benchmark. *J. Proteome Res.*, **9**, 2216–2225.

Kurzbauer,R. *et al.* (2004) Crystal structure of the p14/mp1 scaffolding complex: How a twin couple attaches mitogen-activated protein kinase signaling to late endosomes. *Proc. Natl Acad. Sci. USA*, **101**, 10984–10989.

Lensink,M. and Wodak,S. (2010) Blind predictions of protein interfaces by docking calculations in capri. *Proteins*, **78**, 3085–3095.

Luo,G. *et al.* (2006) Dynamic distance disorder in proteins is caused by trapping. *J. Phys. Chem. B*, **110**, 9363–9367.

Marrink,S.J. *et al.* (2007) The martini force field: coarse grained model for biomolecular simulations. *J. Phys. Chem. B*, **111**, 7812–7824.

Mizuguchi,K. *et al.* (1998) Joy: protein sequence-structure representation and analysis. *Bioinformatics*, **14**, 617–623.

Monticelli,L. *et al.* (2008) The martini coarse-grained force field: extension to proteins. *J. Chem. Theory Comput.*, **4**, 819–834.

Ofran,Y. and Rost,B. (2007) Protein-protein interaction hotspots carved into sequences. *PLoS Comput. Biol.*, **3**, e119.

Oshima,H. *et al.* (2011) Crucial importance of the water-entropy effect in predicting hot spots in protein-protein complexes. *Phys. Chem. Chem. Phys.*, **13**, 16236–16246.

Parrinello,M. and Rahman,A. (1981) Polymorphic transitions in single crystals: a new molecular dynamics method. *J. Appl. Phys.*, **52**, 7182–7190.

Periole,X. *et al.* (2012) Structural determinants of the supramolecular organization of g protein-coupled receptors in bilayers. *J. Am. Chem. Soc.*, **134**, 10959–10965.

Pons,C. *et al.* (2010) Present and future challenges and limitations in protein-protein docking. *Proteins*, **78**, 95–108.
Pool,R. and Bolhuis,P.G. (2010) The influence of micelle formation on the stability of colloid surfactant mixtures. *Phys. Chem. Chem. Phys.*, **12**, 14789–14797.
Pool,R. *et al.* (2012) Enabling grand-canonical monte carlo: extending the flexibility of gromacs through the grompy python interface module. *J. Comput. Chem.*, **33**, 1207–1214.
Sali,A. and Blundell,T.L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.*, **234**, 779–815.
Schueler-Furman,O. *et al.* (2005) Progress in modeling of protein structures and interactions. *Science*, **310**, 638–642.
Sprinzak,E. *et al.* (2003) How reliable are experimental protein-protein interaction data? *J. Mol. Biol.*, **327**, 919–923.
Tamura,A. and Privalov,P.L. (1997) The entropy cost of protein association. *J. Mol. Biol.*, **273**, 1048–1060.
Tobi,D. (2010) Designing coarse grained-and atom based-potentials for protein-protein docking. *BMC Struct. Biol.*, **10**, 40.
Trzesniak,D. *et al.* (2007) A comparison of methods to compute the potential of mean force. *ChemPhysChem.*, **8**, 162–169.
Tuffery,P. and Derreumaux,P. (2012) Flexibility and binding affinity in protein–ligand, protein–protein and multi-component protein interactions: limitations of current computational approaches. *J. R. Soc. Interface*, **9**, 20–33.
Tuncbag,N. *et al.* (2009) A survey of available tools and web servers for analysis of protein-protein interactions and interfaces. *Brief Bioinformatics*, **10**, 217–232.
van Gunsteren,W.F. *et al.* (1996) *Biomolecular Simulation: The GROMOS96 Manual and User Guide*. Vdf Hochschulverlag, AG an der ETH Zürich, Zürich, Switzerland.
Wereszczynski,J. and McCammon,J.A. (2012) Statistical mechanics and molecular dynamics in evaluating thermodynamic properties of biomolecular recognition. *Q. Rev. Biophys.*, **45**, 1–25.
Wollacott,A. *et al.* (2007) Prediction of structures of multidomain proteins from structures of the individual domains. *Protein Sci.*, **16**, 165–175.
Wu,L.C. *et al.* (2002) Two-step binding mechanism for T-cell receptor recognition of peptide MHC. *Nature*, **418**, 552–556.
Yu,Y.B. *et al.* (2001) Contribution of translational and rotational motions to molecular association in aqueous solution. *Biophys. J.*, **81**, 1632–1642.
Zhang,H. *et al.* (2013) Quantification of solvent contribution to the stability of noncovalent complexes. *J. Chem. Theory Comput.*, **9**, 4542–4551.