We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Practical Estimation of TCR-pMHC Binding Free-Energy Based on the Dielectric Model and the Coarse-Grained Model

Hiromichi Tsurui[1] and Takuya Takahashi[2]
[1]Juntendo University
[2]Ritsumeikan University
Japan

1. Introduction

To evaluate free energy changes of bio-molecules in a water solution, *ab initio* molecular dynamics (MD) simulations such as Quantum Mechanical Molecular Mechanics (QM/MM) and MD are the most theoretically rigorous methods (Car and Parrinello 1985; Kuhne, Krack et al. 2007), although the calculation cost is far too large for large molecular systems that contain many electrons. Therefore, all-atom MD simulations based on classical mechanics (i.e., Newton's equations) are used for the usual bio-molecular systems. As the conventional free energy perturbation (FEP) method based on all atom MD simulation is a strict method, to elucidate the molecular principles upon which the selectivity of a TCR is based, FEP simulations are used to analyse the binding free energy difference of a particular TCR (A6) for a wild-type peptide (Tax) and a mutant peptide (Tax P6A), both presented in HLA A2. The computed free energy difference is 2.9 kcal mol^{-1} and the agreement with the experimental value is good, although the calculation is very time-consuming and the simulation time is still insufficient for fully sampling the phase space. From this FEP calculation, better solvation of the mutant peptide when bound to the MHC molecule is important to the greater affinity of the TCR for the latter. This suggests that the exact and efficient evaluation of solvation is important for the affinity calculation (Michielin and Karplus 2002). Other FEP calculations of the wild-type and the variant human T cell lymphotropic virus type 1 Tax peptide presented by the MHC to the TCR have been performed using large scale massively parallel molecular dynamics simulations and the computed free energy difference using alchemical mutation-based thermodynamic integration, which agrees well with experimental data semi-quantitatively (Wan, Coveney et al. 2005). However, the conventional FEP is still very time-consuming when searching for so many unknown docking structures because all-atom MD for a large molecular system is a computationally hard task and MD simulations must be done not only in initial and final states but also in many intermediate states.

Recently, the energy representation (ER) method - where only the initial and final states of a molecular system need to be considered and the sampling cost is drastically decreased - is developed for the molecular solvation process by Matubayashi et al. (Matubayasi and Nakahara 2000; Matubayasi and Nakahara 2002) and will be applicable to the calculation of

binding free energy. Of course, this ER method can be combined with the approximate models described below. Instead of MD simulations, Monte Carlo simulations are also used for the sampling of the configurations. This type of approach which only considers initial and final states is called an endpoint method.

Most of the calculation cost in all-atom MD involves the sampling of the solvent atom configurations because the number of solvent atoms - such as water and co-solvent ions - is much larger than that of the target bio-molecules, and long-range electrostatic potential is especially time consuming although efficient algorithms such as the Fast Multi Pole method and several Ewald methods are developed for all-atom MD. To decrease the calculation cost of the long-range electrostatic term, a continuum dielectric model - which can calculate the electrostatic free energy term of the system very efficiently - is widely used in many bio-molecular systems and is described in the next section.

In the case of large molecules, the entropy term of solvation change becomes important (Asakura and Oosawa 1954), and the solvent accessible surface area (SA) based calculation method becomes insufficient because the excluded volume effect increases. Therefore, integration equation (IE) theories such as the Ornstein-Zernike equation and the closures - which are developed in molecular liquid theory - promise to evaluate entropy change, including solvation and de-solvation processes (Kinoshita 2006; Kinoshita 2009; Yasuda, Yoshidome et al. 2010). The recent MD software package AMBER also contains such an IE algorithm, 3D-RISM, which is a reference site model employing Cartesian coordinates (Luchko, Gusarov et al. 2010). In particular, the simple morphological theory obtained from this IE approach is now applied to the elucidation of protein folding (Yasuda, Yoshidome et al. 2010) and F1-ATPase mechanisms and has proven to be useful (Yoshidome, Ito et al. 2011).
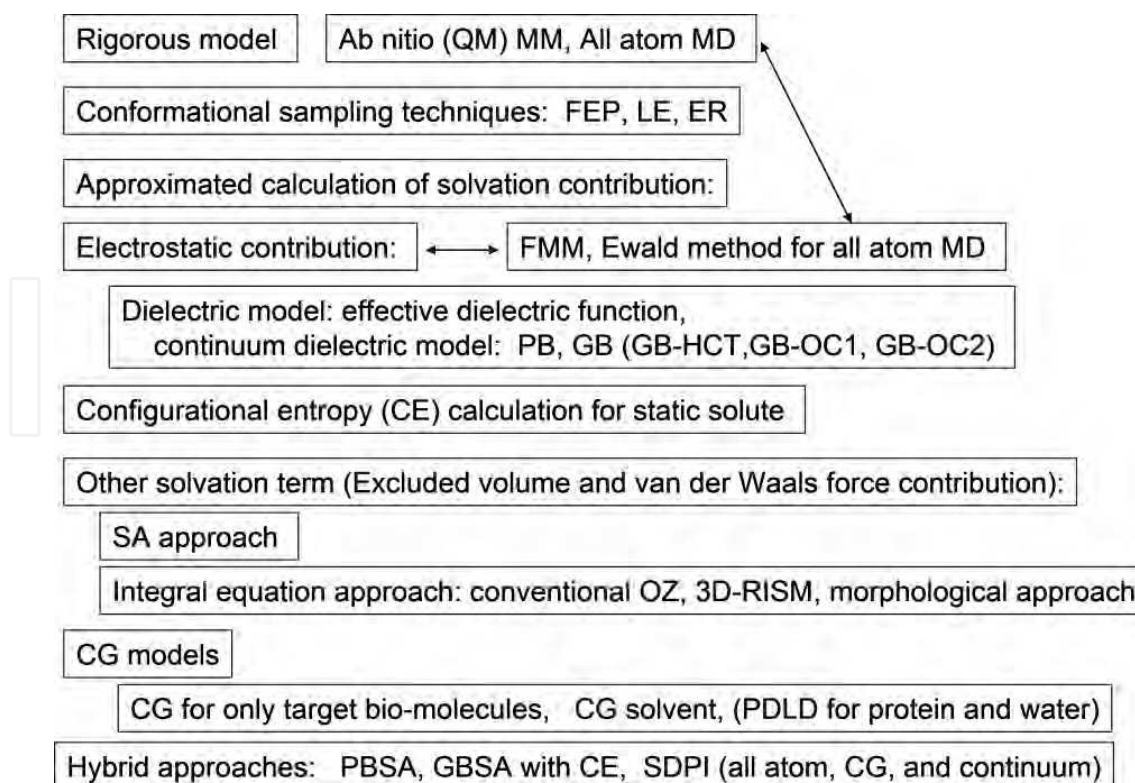

Fig. 1. The relationship among theoretical models and approaches

Practical Estimation of TCR-pMHC Binding Free-Energy
Based on the Dielectric Model and the Coarse-Grained Model
109

The other approaches for decreasing the calculation cost of the solvent molecules are coarse-grained (CG) solvent models. The protein–dipole Langevin–dipole (PDLD) model, which can efficiently calculate the electrostatic interaction among permanent dipoles and induced dipoles of proteins and solvent atoms, is one of the coarse-grained solvent models. As the PDLD model is usually used in the outer area of the all-atom region, this is a hybrid approach of CG and all-atom models (Warshel and Levitt 1976; Xu, Wang et al. 1992). Hybrid approaches of all-atom, CG and continuum solvent models are evolving. A smoothly decoupled particle interface (SDPI) model has a switching region that gradually transitions from fully interacting particles to a continuum solvent. The resulting SDPI model allows for the use of an implicit solvent model based on a simple theory that needs only to reproduce the behaviour of bulk solvent rather than the more complex features of local interactions (Wagoner and Pande 2011). Of course, CG models for solute molecules - which are described in the third section - are promising for the understanding of protein folding (Liwo, He et al. 2011) and predictions of the ligand-receptor docking structure, etc.

The relationship among the theoretical models and approaches is summarised in Fig. 1.

2. The dielectric model and the MM-PBSA (GBSA) method

In this section, we describe briefly the principles behind the methods, the differences between PBSA and GBSA and explicit and implicit treatment.

2.1 Principles of the method

A molecule has an atomic polarisability due to its electrons and an orientational polarisability when the molecule is polar and has a permanent electric dipole moment. A high value of the relative dielectric constant (ε_r=78.4 at 298K) of water is mainly due to its orientational polarisation, where the electric dipole moment is 2.95 Debye. Moreover, the solution in our body contains several co-solvent ions such as Na^+, Cl^-, K^+ and so on for the usual physiological condition. Therefore, electrostatic interactions among bio-molecules are largely decreased by water and solvent ions in a very complicated manner when compared with the *in vacuo* case (Koehl 2006).

To obtain the electrostatic contribution to free energy change, the dielectric model is a good approximation and is widely used to calculate the electrostatic potential of molecular systems in many scientific and technological fields.

First, the use of a simple function of the effective relative dielectric constant is the easiest way to reduce the calculation time required in obtaining electrostatic potentials.

A simple distance-dependent function, 4.5 r, which is proposed by Pickersgill (Pickersgill 1988), can well explain site-directed mutagenesis experiments. Warwicker showed that the simple Debye–Hückel shielding function with a uniform effective relative dielectric constant of 50 was sufficient to explain experimental results when compared with a continuum model (Warwicker 1999). Mehler et al. challenged this problem and proposed a sigmoid function considering the local hydrophobicity and hydrophilicity of protein molecules whose results were also in good agreement with pKa shift measurements (Mehler and Guarnieri 1999). These methods are simple and very fast; however, they all require parameter readjustment for each new system to be studied. Unfortunately, a universal function applicable to all macromolecular systems does not yet exist.

Empirically obtained effective dielectric functions that depend on the inter-atomic distance, r, such as linear functions (εr=r or $4r$) and the sigmoid function is simple, and low calculation-cost method is still used in recent drug design studies for the docking simulations of large molecular systems so as to save on the calculation-cost, although the calculation error is large (Takahashi, Sugiura et al. 2002).

2.1.1 PB approach

On the other hand, the typical dielectric model solves the Poisson equation and treats bio-molecules and water as continuum media which have a specific dielectric constant, although the position-dependent local dielectric constant - which is calculated from the electronic polarisation of atoms and the orientational polarisation of local dipoles - is also possible for a finite difference equation (Nakamura, Sakamoto et al. 1988; Pitera, Falta et al. 2001).

Moreover, the Poisson-Boltzmann (PB) equation, which was first proposed by Gouy in 1910 and was complemented by Chapman in 1913, is widely-used for considering the contribution of solvent ions. The Gouy-Chapman theory, which solves a simple one-dimensional nonlinear PB equation, is often used in a membrane-electrolyte system that has electrical double layers (Forsten, Kozack et al. 1994).

The PB equation is a differential equation and it describes electrostatic interactions between molecules in ionic solutions by using a mean-filed approximation where the correlations among the solvent ions are neglected. The equation in SI units can be written as:

$$\bar{\nabla}\left[\varepsilon(\vec{r})\bar{\nabla}\psi(\vec{r})\right]=-\rho(\vec{r})-\sum_i c_i^{\infty}z_iq\lambda(\vec{r})\exp\left[\frac{-z_iq\bar{\nabla}\psi(\vec{r})}{k_BT}\right] \tag{1}$$

where $\vec{\nabla}\cdot$ is the divergence operator and $\varepsilon(\vec{r})$ is the position-dependent dielectric, which is set to be constant in the solvent, the bio-molecule and the boundary regions in continuum dielectric models. $\bar{\nabla}\psi(\vec{r})$ is the gradient of the electrostatic potential, $\rho(\vec{r})$ represents the charge density of the solute (i.e., the fixed charges of the bio-molecule), c_i^{∞} represents the concentration of the ion i at a distance of infinity from the solute, z_i is the charge of the solvent ion, q is the charge of a proton, k_B is the Boltzmann constant, T is the temperature and $\lambda(\vec{r})$ is a factor for the position-dependent accessibility of position $\vec{r}$ to the ions in the solution. If the potential is small and the electrostatic energy is negligible compared to the thermal fluctuation, k_BT, the equation can be linearised and solved more efficiently.

$$\bar{\nabla}^2\psi(\vec{r})=\kappa^2\psi(\vec{r}) \tag{2}$$

Here, κ is the Debye shielding parameter, defined as follows:

$$\kappa^2=\sum_i^m\frac{z_i^2q^2c_i^{\infty}}{\varepsilon k_BT} \tag{3}$$

This weak field limit approach is called the Debye–Hückel approximation (Fogolari, Brigo et al. 2002).

To solve the PB equation, there are typically three numerical methods: a finite difference (FD) method is relatively time consuming, but simple and applicable to a complex system which has a position-dependent local dielectric constant. Therefore, the FD method is firstly applied to calculate the electrostatic potential in a protein-solvent system, and the pKa shift of the protein ionisable residues are well-explained (Gilson and Honig 1987) and the effect of the salt concentration on the pKa are also reproduced (Takahashi, Nakamura et al. 1992). The finite element method (FEM) and the boundary element method (BEM) are more powerful and the calculation cost is smaller than the FD method, although only a uniform dielectric constant must be set in each region (Lu, Zhou et al. 2008).

2.1.2 GB approach

One other powerful way to obtain the electrostatic potential based on the dielectric model is the Generalised Born (GB) model, which solves the linearised PB equation by approximating such bio-molecules as proteins and nucleic acids as a set of spheres whose internal dielectric constant differs from the external solvent (Koehl 2006). The functional form of the model is written as:

$$Gs = \frac{1}{8\pi}\left(\frac{1}{\varepsilon_0} - \frac{1}{\varepsilon}\right)\sum_{i,j}^{N}\frac{q_i q_j}{f_{ij}} \tag{4}$$

where

$$f_{ij} = \sqrt{r_{ij}^2 + a_{ij}^2 e^{-D}} \tag{5}$$

and

$$D = \left(\frac{r_{ij}}{2a_{ij}}\right)^2 , a_{ij} = \sqrt{a_i a_j} \tag{6}$$

where ε_0 is the dielectric constant *in vacuo*, ε is the dielectric constant of the solvent, q_i is the electrostatic charge on the particle i, r_{ij} is the distance between particles i and j, and a_i is a length defined as the effective Born radius (Still, Tempczyk et al. 1990).

The effective Born radius of an atom represents its degree of burial inside the solute and corresponds to the distance from the atom to the molecular surface. The exact evaluation of the effective Born radii is the central issue for the GB model (Onufriev, Bashford et al. 2004).

To consider the electrostatic shielding effect due to the solvent ions, a simplified function based on the Debye-Hückel approximation is added to the function Gs in the AMBER software package (Case, Cheatham et al. 2005), which is one of the most used packages in the world of bio-molecular simulations, as follows by (Srinivasan, Trevathan et al. 1999):

$$Gs = \frac{1}{8\pi}\left(\frac{1}{\varepsilon_0} - \frac{\exp(-\kappa f_{ij})}{\varepsilon}\right)\sum_{i,j}^{N}\frac{q_i q_j}{f_{ij}}. \tag{7}$$

They calculated the solvation free energies, G_s, with this GB model for proteins and nucleic acids, which agreed very well with those of the PB model. The salt-dependence of the electrostatic binding free energy based on the Debye-Hückel approximation is still under investigation (Harris, Bredenberg et al. 2011).

2.1.3 The GBSA (PBSA) approach

GBSA (PBSA) is simply a GB (PB) model with the hydrophobic solvent accessible surface area (SA) term. This is the most commonly used implicit solvent model combination and is widely used in MD simulations for large bio-molecules. This approach is known as MM/GBSA in the context of molecular mechanics. This formulation can well identify the native states of short peptides with a precise stereoscopic structure (Ho and Dill 2006), although the conformational ensembles produced by GBSA models in other studies differ significantly from those produced by an explicit solvent and do not identify the protein's native state (Zhou 2003). In particular, strong charge-charge interaction such as salt bridges are overstabilised due to insufficient electrostatic screening, and the alpha helix population became higher than the native one. These problems are common in PBSA. Variants of the GB model have also been developed to approximate the electrostatic environment of membranes, which have had some success in folding the transmembrane helices of integral membrane proteins (Im, Feig et al. 2003).

There are several kinds of software containing the GB algorithm. For example, the AMBER software package has three types of GBSA models as has as the PBSA model.

The MM-PBSA and GBSA approaches are the endpoint methods and usually only consider the initial unbound state and the final bound state. The binding free energy change, dG_{bind}, is written as:

$$dG_{bind} = dG_{gas} + dG_{solv} = (dH_{gas} + dH_{tr/ro} - TdS) + (dG_{elsolv} + dG_{npsolv}). \qquad (8)$$

The term dG_{gas} refers to total free energy change and the term dH_{gas} contains the van der Waals and electrostatic interaction energies as well as internal energy variation, such as bond, angle and torsional angle energies *in vacuo* (i.e., gas phase). The terms $dH_{tr/ro}$ denote the energy difference due to translational and rotational degrees of freedom, and becomes 3 RT in the classical limit (i.e., thermal energy is large enough). The term dS refers to the conformational entropy change (Tidor and Karplus 1994; Ben-Tal, Honig et al. 2000). The term dG_{solv} is the difference between the initial and final solvation free energies and is divided into the electrostatic contribution, dG_{elsolv}, and the nonpolar contribution, dG_{npsolv}. The term dG_{npsolv}, which is the sum of a cavity term and a solute-solvent van der Waals term, is calculated from the SA as follows:

$$dG_{npsolv} = \gamma SA + b. \qquad (9)$$

The surface tension γ and the constant *b* are 0.00542 kcal mol^{-1} Å^{-2} and 0.92 kcal mol^{-1} respectively, for the MM-PBSA model (Sitkoff, D., K. Sharp and B. Honig. 1994). For GB models, 0.0072 kcal mol^{-1} Å^{-2} and 0 kcal mol^{-1} (Jayaram, Sprous et al. 1998), or else 0.005 kcal mol^{-1} Å^{-2} and 0 kcal mol^{-1} (Gohlke, Kuhn et al. 2004) are used. The SA in AMBER is calculated by using the LCPO algorithm (Weiser, Shenkin et al. 1999; Still, Tempczyk et al. 1990) to compute an analytical approximation to the solvent accessible area of the molecule.

Practical Estimation of TCR-pMHC Binding Free-Energy
Based on the Dielectric Model and the Coarse-Grained Model
113

The several types of the GBSA models are not only applied to many protein folding simulations (Zhou 2003), but also to nucleic acid conformational dynamics from massively parallel stochastic simulations, where the ubiquitous helical hairpin conformation is reproduced and folding pathway is investigated (Sorin, Rhee et al. 2003).

2.2 Review of recent work

As mentioned in the previous section 2.1, the dielectric models and the hybrid approaches are widely used in many scientific and technological fields, such as protein folding, molecular docking and drug design, etc. In particular, the binding free energy (BFE) calculation and the prediction of the binding affinity and binding structure between ligands and proteins is the most important aim (Gilson and Zhou 2007) because the major purpose of molecular docking (Zacharias and Fiorucci 2010; Leis and Zacharias 2011) is to predict the experimentally-obtained BFE and the binding site of a receptor to a specific ligand molecule, and drug design is usually supported by suitable molecular docking methods.

For example, the linear interaction energy (Rastelli, Rio et al. 2010) method - which combines two different continuum solvent models - is applied to calculate protein-ligand BFEs for a set of inhibitors against the malarial aspartic protease plasmepsin II, and the explicit solvent LIE calculations and LIE-PB reproduce absolute experimental BFEs with an average unsigned error of 0.5 and 0.7 kcal mol^{-1} respectively (Carlsson, Ander et al. 2006). Moreover, the ligand-water interaction energies - which are calculated from both PB and GB models using snapshots from explicit solvent MD simulations of the ligand and protein-ligand complex - are compared with the explicit solvent MD results. The obtained energy from the explicit water MD agrees well with those from the PB model, although the GB model overestimates the change in solvation energy, which overestimation is caused by consistent underestimation of the effective Born radii in the protein-ligand complex.

Xu and Wang applied the MM-PBSA method to FK506-binding proteins (Xu and Wang 2006) - which are important targets of pharmaceutical interests - and calculated the binding of a set of 12 non-immunosuppressive small-molecule inhibitors to FKBP12 through MD simulations, where each complex is subjected to 1-ns MD simulation conducted in an explicit solvent environment under constant temperature and pressure. The BFE of each complex is then calculated with the MM-PBSA method in the AMBER program and the MM-PBSA computation agrees very well with the experimentally determined BFEs, with a correlation coefficient (R^2) of 0.93 and a standard deviation as low as 0.30 kcal mol^{-1}. The vibrational entropy term given by the normal mode analysis is necessary for achieving this correlation. Moreover, an adjustment to one weight factor in the PBSA model is essential to correct the absolute values of the final binding free energies to a reasonable range, which suggests that the very good correlation is due to the similar properties of ligand molecules and that this artificial weight factor is not universal. A comparison of the MM-PBSA model with a Linear Response Approximation model suggests that the MM-PBSA method seems to be robust in binding affinity prediction for this class of compounds (Lamb, Tirado-Rives et al. 1999).

To systematically evaluate the performance of MM-PBSA and several versions of the MM-GBSA models, extensive calculations of BFEs are done for 59 ligands interacting with six different proteins with the AMBER 9.0 software (Hou, Wang et al. 2011). First, the effects of

the length of the MD simulation are explored, ranging from 400 to 4800 ps, and the simulation length has an obvious impact on the predictions. Interestingly, longer MD simulation is not always necessary for achieving better predictions. Second, the effect of a solute dielectric constant (1, 2, or 4) on the BFEs of MM-PBSA is also checked and the predictions are quite sensitive to the solute dielectric constant. Therefore, this parameter should be carefully determined according to the characteristics of the protein/ligand binding interface. Third, conformational entropy often shows large fluctuations in MD trajectories, and a large number of snapshots are necessary to achieve stable predictions. Next, the comparison of the accuracy of the BFEs of three GB models: (1) GB-HCT, the pair wise model by Hawkins et al. (Hawkins, Cramer et al. 1996) parameterised by Tsui and Case (Tsui and Case 2000); (2) GB-OC1 and (Case, Cheatham et al.) GB-OC2, the parameters of which are modified by Onufriev et al. (Onufriev, Bashford et al. 2004) and the GB-OC1 model which gives better results compared to the other two GB models in ranking the binding affinities of the studied inhibitors. This may be explained by the better agreement of GB-OC1 with PBSA. The better performance of MM-PBSA when compared with MM-GBSA in calculating absolute - but not necessarily relative - BFEs is confirmed, which is not surprising because the GBSA is the approximation of PBSA, but it suggests the reliability of the dielectric continuum model itself. Considering its computational efficiency, MM-GBSA gives good relative BFEs and is much faster than MM-PBSA, and can serve as a powerful tool in drug design where the correct ranking of inhibitors is often emphasised and the obtaining of the absolute value of BFEs is not so important.

Interestingly, the successive study of MM-PBSA and MM- GBSA-OC1 using 98 protein-ligand complexes to develop an excellent scoring function by Hou et al. shows that MM-GBSA (success rate 69.4%) outperformed MM-PBSA (45.5%) and many popular scoring functions in identifying the correct binding conformations, and the best prediction of the MM-GBSA model with an internal dielectric constant of 2.0 produced a Spearman correlation coefficient of 0.66, which is better than MM/PBSA (0.49) and almost all the scoring functions used in molecular docking (Hou, Wang et al. 2011). However, the reason why the PBSA underperforms the GBSA is not clear. One possibility is the difference of the SA term and the other possibility is the insufficiency of the conformational sampling of proteins, as the authors are also emphasising the importance of MD calculation time. In any case, MM-GBSA performs well, for both binding pose predictions and binding free-energy estimations and it is efficient at re-scoring the top-hit poses produced by other less-accurate scoring functions.

As AMBER and other software packages - including the PB and GB models - are widely used and drug design is the important issue, many studies concerning ligand-protein docking based on the dielectric model have been done (Rastelli, Rio et al. 2010). The above calculation results of the GB and PB dielectric models are limited to relatively small ligand molecules and receptor proteins, and the size of the complex is not so large compared to so-called super-molecules, such as the immune complex and membrane proteins, etc. To calculate and analyse the BFE of a large, complex T-cell receptor (TCR) and immunogenic peptides (p) presented by class I major histocompatibility complexes (MHC), binding free energy decomposition (BFED) calculations based on the MM–GBSA approach including entropic terms were done on the 2C TCR/SIYR/H-2Kb system and provided a detailed description of the energetics of the interaction (Zoete and Michielin 2007), since this BFED method can detect the important individual   side chains for the stability of a protein fold

with computational alanine scanning of the insulin monomer (Zoete and Meuwly 2006). A correlation between the decomposition results and experimentally-determined activity differences for alanine mutants of the TCR-pMHC complex is 0.67 when the conformational entropy is neglected, and 0.72 when the entropy is considered. Similarly, a comparison of experimental activities with variations in the BFEs determined by computational alanine scanning yields correlations of 0.72 and 0.74 when the entropy is neglected or taken into account, respectively. In addition, a comparison of the two theoretical approaches for estimating the role of each side chain in the complex formation is given, and a new *ad hoc* approach for decomposing the vibrational entropy term into atomic contributions - the linear decomposition of vibrational entropy (LDVE) - is introduced. The latter allows the rapid calculation of the entropic contribution of interesting side chains to the binding. This approach is justified by the idea that the most important contributions to the vibrational entropy of a molecule originate from residues that contribute most to the vibrational amplitude of the normal modes. The results of the LDVE are very similar to those of the exact but highly computationally demanding method. The BFED approach is also applicable to the design of rational TCR by calculating each amino acid contribution in mutated TCR. As melanoma patients frequently show unusually positive clinical outcomes, it represents an interesting target for adoptive transfer with modified TCR. Sequence modifications of TCR which potentially increase the affinity for this epitope have been proposed and tested *in vitro*. T-cells expressing some of the proposed TCR mutants showed better T-cell functionality, with the improved killing of peptide-loaded T2 cells and better proliferative capacity compared to the wild type TCR expressing cells (Zoete, Irving et al. 2010).

As there are still not many applications for massive simulations with dielectric models to large bio-molecules like the TCR-pMHC complex, more extensive studies are necessary to evaluate the validity of the method and improve its accuracy and performance because the excluded volume effect due to water entropy change in binding will become larger in the larger systems.

2.3 The correlation between calculation-cost and accuracy

It is not easy to state the calculation cost and accuracy exactly because the method is only now developing and the accuracy depends on the system size.

Previous studies have shown a very good correlation between PB and GB results because the GB parameter is modified to achieve better agreement with that of PB (Gohlke, Kuhn et al. 2004; Onufriev, Bashford et al. 2004). Moreover, GB and PB methods also enable the rapid scoring of protein structures when they are combined with physics-based energy functions. The direct comparison of these two approaches on large protein data sets is done with a scoring function based on a GB and PB solvation model and short MD simulations. Against seven publicly available decoy sets, the results of the MM-PBSA approach are comparable to the GB-based scoring function (Lee, Yang et al. 2005).

We also compared the MM-PBSA and MM-GBSA methods. Table 1 shows the comparison of the binding electrostatic free energies of the PB and GB methods for two TCR-pMHC complexes (2gj6 and 3pwp), a complex of A6 and Tax peptide-HLA A2, and A6 with Hud-A2 respectively. Constant regions of TCR were removed (Gregoire, Lin et al. 1996), hydrogen was added and the complexes were neutralised and solvated with TIP3P. The

numbers of atoms involved in the systems were 130,545 for 2GJ6 and 127,023 for 3PWP. Calculations were performed with Sander of AMBER 11 for 5 ns. The G_{elsolv}, which is always largely negative in each case, represents the electrostatic energy contribution due to solvents. The G_{np} is the hydrophobic and van der Waals contributions were calculated from the solvent accessible surface area (SA). The difference of the PB and GB results of each case is 3-4%, although the total binding free energy, dG_{bind}, differs by almost 20% because the binding energies *in vacuo*, dG_{gas}, and the contribution of the solvent, $dG_{solv}=dG_{elsolv}+dG_{npsolv}$, have a different sign and cancel each other. We must note that the ratio of the SA contributions between GBSA and PBSA is larger than the E_{elsov}, although the absolute contribution is $1/10^{th}$ of the G_{elsolv}.

	2gj6			3pwp		
	GBSA	PBSA	GB/PB	GBSA	PBSA	GB/PB
Complex						
G_{elsolv}	-7502.3	-7233.6	1.037	-7367.2	-7112.6	1.036
G_{npsolv}	193.3691	152.9235	1.2645	188.3416	151.8496	1.2403
Receptor protein						
G_{elsolv}	-2512	-2421.1	1.038	-2358.79	-2259.25	1.044
G_{npsolv}	73.4152	58.0731	1.2642	71.2015	57.4698	1.2389
Ligand						
G_{elsolv}	-5340.8	-5151.5	1.037	-5340.08	-5162.4	1.034
G_{npsolv}	133.8713	105.4387	1.2697	131.2283	104.4543	1.2563
Difference = Complex-(Receptor + Ligand)						
dG_{elsolv}	350.479	338.916	1.034	331.6509	309.0868	1.073
dG_{npsolv}	-13.9175	-10.5883	1.314	-14.0882	-10.0745	1.398
$dE_{el\_solv}+dE_{SA}$	336.561	328.328	1.025	317.563	299.012	1.062
dG_{gas}	-388.76	-388.76	1	-380.55	-380.55	1
dG_{solv}	336.561	328.328	1.025	317.563	299.012	1.062
dG_{bind}	-52.199	-60.432	0.864	-62.99	-81.54	0.773

Table 1. A comparison of the binding electrostatic free energies of the PBSA and GB methods for two TCR-pMHC complexes (PDB ID: 2gj6 and 2pwp). The G_{elsolv}, which is always largely negative in each case, represents the electrostatic energy contribution due to solvents. The G_{npsolv} is the hydrophobic and van der Waals contributions are calculated from the solvent accessible surface area (SA). The total binding free energy, dG_{bind}, is the sum of the binding energy *in vacuo*, dG_{gas}, and the contribution of the solvent, $dG_{solv}=dG_{elsolv}+dE_{SA}$. All energies in the table are given in kcal mol^{-1}.

3. Coarse-grained (CG) simulation

3.1 The limits of all-atom simulations

Even though all-atom simulations provide the most detailed information about the system of interest, its calculation costs are quite high. A system containing a large protein molecule such as several 10^5 to 10^6 Dalton comes up to several 10^5 atoms when solvated in explicit water molecules, and expands to nm^3 in size; hence, the calculation time of less than μsec

Practical Estimation of TCR-pMHC Binding Free-Energy
Based on the Dielectric Model and the Coarse-Grained Model
117

even using a recent multi-core PC. These figures are too short and still too small to reproduce such biologically interesting phenomena as protein folding, protein-assembly and enzymatic reaction, etc. Therefore, the increase of calculation efficiency is quite an urgent requirement. The calculation cost increases approximately in proportion to the square of the number of atoms, and the time for one step is approximately proportional to the order of the square-root of the mean mass of elements. The number of atoms constituting an amino acid (AA), when polymerized in a peptide, is 7 (Gly) to 24 (Trp), and the mass is between 57 (Gly) and 186 (Trp) - about 5 to 15 times of a C, N or O. When an AA is coarse-grained to 2 to 4 pseudo-atoms, the calculation cost decreased by 2 to 3 orders of magnitude, and the time for a step increases by 2 or 3 orders. In most CG models, the interaction between pseudo-atoms through bonds of less than 5 is described as follows:

Extension potential between two beads

$$U^{bond} = \frac{1}{2}k(r-r_0)^2 \tag{10}$$

Angle potential between three beads

$$U^{angle} = \frac{1}{2}k(\theta-\theta_0)^2 \tag{11}$$

Dihedral angle between four beads

$$U^{dihedral} = \frac{1}{2}k\left[1-\cos(n\phi-\phi_0)\right] \tag{12}$$

And, the unbound potential between two beads can be expressed as

$$U^{LJ} = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6\right] \tag{13}$$

The whole energy of the system is described as the combination of these elemental potentials. For example, the Head-Gordon et al. model (Brown, Fawzi et al. 2003) is described as:

$$H = \sum_\theta \frac{1}{2}K_\theta(\theta-\theta_0)^2 + \sum_\phi\left[A(1+\cos\phi) + B(1-\cos\phi) + C(1+\cos 3\phi) + D\left(1+\cos\left[\phi+\frac{\pi}{4}\right]\right)\right]$$
$$+ \sum_{i,j\geq i+3} 4\varepsilon_H S_1\left[\left(\frac{\sigma}{r_{ij}}\right)^{12} - S_s\left(\frac{\sigma}{r_{ij}}\right)^6\right] \tag{14}$$

where θ, φ, and i, j are summed for all the AAs contained in the peptide. The interaction between non-bonded pseudo-atoms is usually described as the Lennard-Jones potential. The methods for configuration sampling, usually MD (Shih, Arkhipov et al. 2006) and the Monte Carlo simulation (Levy, Karplus et al. 1980; Horejs, Mitra et al. 2011), are the same as those used in all-atom simulations. The equation of motions for MD is principally the same as that used in all-atom simulations, i.e.,

$$m\frac{d^2\vec{r}}{dt^2} = \vec{F} - \Gamma\frac{d\vec{r}}{dt} + \overrightarrow{W} \tag{15}$$

where $\vec{F}$, Γ, and $\overrightarrow{W}$ are external force, friction and thermal noise, respectively. Any modification is made according to the kind of ensemble adopted.

3.2 The difference of CG models between proteins and other molecules

As might be guessed by Fig.2A and B, it is easier to treat a homopolymer by the CG model than to treat a polypeptide or a protein. A homopolymer can be described with rather a few parameters and, under certain circumstances, several components can be coarse-grained as a pseudo-atom (4 styrenes in a dotted circle are treated as one bead). Rheological features such as phase-transition, diffusion coefficient, compressibility, ductility, elasticity and viscosity have been reproduced fairly well (Yaoita, Isaki et al. 2008; Harmandaris and Kremer 2009; Kalra and Joo 2009; Posel, Lísal et al. 2009). On the other hand, peptides and proteins consist of diversified 20 AAs and the particular functions of proteins such as specific binding and enzymatic functions are based on a unique configuration of those characteristic AAs. Therefore, to evaluate the interaction on the CG model is especially difficult due to the effect of averaging specific properties and the anisotropicity of components. Notwithstanding this state of affairs, some CG models have come to predict the docking and binding of proteins fairly well. In this section, representative protein CG models are reviewed and the application of the CG model to the evaluation of TCR-pMHC interaction is foreseen.

3.3 Representative CG models

3.3.1 The one-bead model

Many one-bead models (Taketomi, Ueda et al. 1975; Brown, Fawzi et al. 2003; Jang, Hall et al. 2004) can be deemed as descendants of the Go-model. Go-like models, even though extremely simplified in their format, principally succeeded in reproducing several aspects of protein folding. This is presumably due to the finding that the protein-folding rate and mechanism are largely determined by a protein's topology rather than its inter-atomic interaction (Baker 2000). Those descendant models have equipped their own features, but still have a tendency towards a reference configuration. This might be due to the difficulty of incorporating the geometric and physicochemical aspects of all the AAs in only a few parameters. Recently, the finding that the underlying physicochemical principles of the interaction between the domains in protein folding are similar to those between the binding sites of protein assembly has been accepted (Haliloglu, Keskin et al. 2005; Levy, Cho et al. 2005; Turjanski, Gutkind et al. 2008; Baxter, Jennings et al. 2011). This fact will probably provide another aspect of the application of the CG model to issues of protein-binding.

Miyazawa and Jernigan (MJ) extracted inter-residue potentials from the crystallography of 1168 proteins (Miyazawa and Jernigan 1996). The principle adopted in this method is that the number of residue-residue contacts observed in a large number of protein crystals will represent the actual intrinsic inter-residue interactions. Namely, to regard the effect (contacts in the observed structure) in the same light as the cause (interaction energy) based on "the principle of structural consistency" or "the principle of minimal frustration".
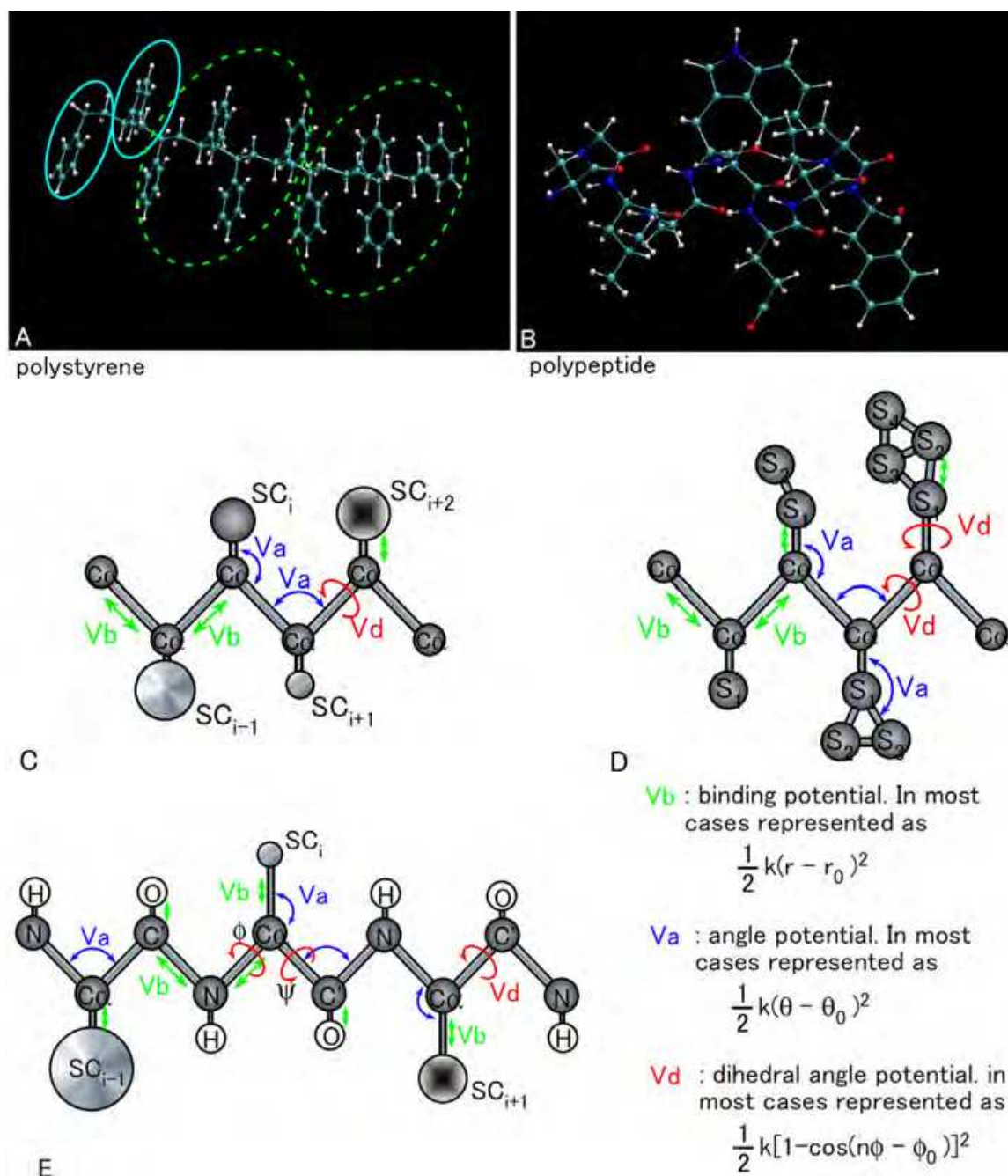
A
polystyrene

B
polypeptide

C

D

Vb : binding potential. In most
cases represented as

$$\frac{1}{2}k(r - r_0)^2$$

Va : angle potential. In most
cases represented as

$$\frac{1}{2}k(\theta - \theta_0)^2$$

Vd : dihedral angle potential. in
most cases represented as

$$\frac{1}{2}k[1-\cos(n\phi - \phi_0)]^2$$

E

Fig. 2. Coarse-grained models for protein
Homopolymer such as polystytene can be described with rather a few parameters, and in
some cases, several units are mapped to one bead (A). Protein consists of heterogenious
components, hence more detailed and complicated description (B). Main chain is
represented by Cα and each side chain is mapped to one bead, which retains its original
geometiric and physico-chemical features (Liwo, Pincus et al. 1993) (C). MARTINI force field
maps more beads to a side chain (Marrink, Monticelli et al. 2008), enabled to simulate the
release of inner water molecules through stress-sensitive channel enbedded in vesicle
membrane (Louhivouri, Lisselada et al. 2008) (D). OPEP model all the atoms of main chain
and maps one bead for side chain. As can be guessed, this models is suitable for dealing
with the issues where backbone structure such as α-helix and β-sheet play essential roles
(Chebaro et al. 2009, Laghaei et al. 2011, Nasica-Labouze et al. 2011) (E).

Adopting this model for the parameters of LJ potentials, Kim and Hummer constructed a one-bead model combined with the Debye-Hückel type potential and performed configure-sampling on replica exchange MC - applied to ubiquitin binding - and obtained good agreement with other experiments (Kim and Hummer 2008). Chakraborty's group applied an MJ matrix to estimate TCR-pMHC and explained the effect of the HLA class I haplotype on TCR repertoire-formation (Kosmrlj, Read et al. 2010). The above mentioned CG models are tabulated in Table 2.

3.3.2 UNRES

Scheraga's group described a CG model which consists of a Cα, side chain centroid (SC) and one dihedral angle (Liwo, Pincus et al. 1993). They searched the conformation space on this model with compactness of the protein as an indicator. The obtained structure was then decoded into an all-atom-backbone with the SC model and then searched further for the lowest-energy structure. Finally, an all-atom model was reconstructed from the obtained structure and searched for the lowest-energy structure on an electrostatically driven Monte Carlo (EDMC) simulation based on the ECEPP/2 potential. They succeeded in predicting *ab initio* the moderate size of proteins (53-235 residues) (Oldziej, Czaplewski et al. 2005). This hybrid method - the sampling of a configuration on a CG model and the estimation of binding energy on an atomistic model - presents quite a reasonable combination of efficiency and accuracy. Their recent accomplishment was a 1 msec simulation of more than 500 AA proteins through massive parallelisation (Scheraga, Maisuradze et al. 2010).

3.3.3 ATTRACT

Zacharias described a docking method of protein-protein or protein-ligand using a reduced protein model and docking algorithm, ATTRACT (Zacharias 2003). An AA is represented with 2 to 3 (Zacharias 2003) or 2 to 4 (Zacharias and Fiorucci 2010) pseudo-atoms and the interactions of specific pseudo-atom pairs, including their size and physicochemical characters, are interweaved into the parameters of the Lennard-Jones potential. ATTRACT assumes that both interacting molecules are rigid, smaller molecules is tried to dock from thousands of sites with 6 degrees of freedom, 3-translational and 3-rotational. Docking includes the minimisation of side chains described as rotamer, hence total minimisation is performed. They applied this CG model and ATTRACT to the Critical Assessment of Prediction of Interest (CAPRI) (Janin 2002) and showed two acceptable bindings out of 6 targets (May and Zacharias 2007) or else obtained better (4 out of 6 targets) prediction by improving the scoring function and docking method (Zacharias and Fiorucci 2010). The estimation of TCR and pMHC binding not only deals with the binding energy of a pre-determined configuration, but also deals with the determination of the binding-configuration, because the TCR-pMHC complex has several binding modes (Wucherpfennig, Call et al. 2009). They showed that it is possible to uncover a binding site by using an electrostatic desolvation profile (Zacharias and Fiorucci 2010) based on ODA method (Fernandez-Recio, Totrov et al. 2005).

3.3.4 The MARTINI force field

The MARTINI force field was originally devised for describing lipids or surfactants, such as dipalmitoylphosphatidylcholine (DPPC), dicaemployl-PC (DCPC), dodecylphosphocholine

Number of beads	Authors	force field or residual potential	sampling	purpose or reproduc
1	Go and coworkers	Go model	MC	protein fol
1	Jang H, Hall CK and Zhou Y	Go like model	MD	protein fol
1	Miyazawa & Yernigan	inter residue potential	static model	calculation of
1	Kim and Hummer	MJ matrix	MC	Ubiquitin
1	Chakraborty and coworkers	MJ matrix	static model	T cell peperto
1	Head-Goadon and coworkers	Go like model	MD	protein foldi
1	Turjanski et al.	Go like model	MD	protein bind
2	Liwo et al.	UNRES and ECEP/2	MC	protein foldi
2	Bahar and Jernigan	distance dependent statistical potential	static method	peptide-prot
2	Scheraga and coworkers	UNRES and ECEP/3	MD	structure, fol
2	Scheraga and coworkers	UNRES and ECEP/3	REMD by parallelism	millisecond
3	Zacharias	reduced protein	ATTRACT	protein-prote
3	Marrink and coworkers	MARTINI	MD	protein behavior in
4	Fiorucci and Zacharias	reduce protein model	modified ATTRACTA	docking, predictio
5	Louhivouri	MARTINI	MD	water release fr
6	Derreumaux	OPEP	MC	structure and fo
6	Derreumaux and Mousseau	OPEP	MD	aggregation mec
6	Derreumaux and coworkers	OPEP	REMD	formation of
6	Nasica-Labouze et al	OPEP+all atom	REMD	detailed process o
6	Laghaei et al	OPEP	HT-REMD	structure and thermodyna

1 (Taketomi, Ueda et al. 1975); 2 (Jang, Hall et al. 2004);3 (Miyazawa and Jernigan 1996); 4 (Hummer and Kim 20
6 (Brown, Fawzi et al. 2003); 7 (Turjanski, Gutkind et al. 2008); 8(Liwo, Pincus et al. 1993); 9 (Bahar and Jernigan
2005); 11 (Liwo, Oldziej et al. 2010); 12 (Zacharias 2003); 13 (Marrink, Monticelli et al. 2008); 14 (Fiorucci and Zac
Risselada et al. 2010); 16 (Derreumaux and Forcellino 2001); 17 (Derreumaux and Mousseau 2007); 18 (Derreuma
(Nasica-Labouze, Meli et al. 2011); 20 (Laghaei, Mousseau et al. 2011)

Table 2. Representative coarse-grained models for protein and noteworthy results based on them

(DPC) and cholesterol (Marrink, de Vries et al. 2004; Marrink, Risselada et al. 2007). The adoption of a very limited atom type and short range potentials provided very efficient computation, hence the micrometer length in scales and milliseconds in time, and succeeded in the simulation of the spontaneous aggregation of DPPC lipids into a bilayer and the formation of DPC in water. The hydrogen atom is neglected in this model. Heavy four atoms on average are represented as one pseudo-atom (four-to-one mapping) with an exception for ringlike molecules. Ringlike molecules are mapped with higher resolution (up to two-to-one mapping). Interaction sites are classified into 4 types: polar (P), nonpolar (N), apolar (C) and charged (Q). Within a main type, subtypes are distinguished either by a letter denoting the hydrogen-bonding capabilities (d = donor, a = acceptor, da= both, 0 = none) or by a number indicating the degree of polarity (from 1 = lower polarity to 5 = higher polarity). The interaction of each atom-type was parameterised at five levels: attractive (e = 5 kJ/mol), semi-attractive (e = 4.2 kJ/mol), intermediate (e = 3.4 kJ/mol), semi-repulsive (e = 2.6 kJ/mol) and repulsive (e = 1.8 kJ/mol). Non-bonded interactions between the interaction sites i and j are described by the Lennard-Jones potential:

$$U_{LJ}(r) = 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r}\right)^{12} - \left(\frac{\sigma_{ij}}{r}\right)^{6}\right] \tag{16}$$

with σ_{ij} representing the effective minimum distance of approach between two particles and ε_{ij} representing the strength of their interaction. This model was extended to deal with proteins (Marrink, Monticelli et al. 2008). The basic parameters are the same as used in the lipid model. Bonded interaction is described by the following set of potential energy functions acting between the bonded sites i, j, k, and l with an equilibrium distance d_b, an angle ϕ_a, and a dihedral angle ψ_i and ψ_{id}:

$$V_b = \frac{1}{2}K_b\left(d_{ij} - d_b\right)^2 \tag{17}$$

$$V_a = \frac{1}{2}K_a\left[\cos\left(\varphi_{ijk}\right) - \cos\left(\varphi_a\right)\right]^2 \tag{18}$$

$$V_d = K_d\left[1 + \cos\left(n\psi_{ijkl} - \psi_d\right)\right] \tag{19}$$

$$V_{id} = K_d\left(\psi_{ijkl} - \psi_{id}\right)^2 \tag{20}$$

where V_b, V_a, V_d and V_{id} represent potential sites for bonding, stiffness, dihedral angle and improper dihedral angle, respectively. The total energy of the system is obtained by summing (17) to (20). The mapping of all AAs is mapped into 4 types of beads or a combination of them. In this mapping, Leu, Pro, Ile, Val, Cys and Met are classified as apolar (C-type), where as Thr, Ser, Asn and Gln are polar (P-type). Glu and Asp are charged (Q-type), and Arg and Lys are modelled by a combination of a Q and an uncharged particle (N-type). The bulky ring-based side chains are modelled by three (His, Phe, and Tyr) or four (Trp) beads. Gly and Ala residues are only represented by the backbone particle. The type of

Practical Estimation of TCR-pMHC Binding Free-Energy
Based on the Dielectric Model and the Coarse-Grained Model
123

the backbone particle depends on the protein secondary structure; free in solution or in a coil or bend, the backbone has a strong polar character (P-type); as part of α helix or β strand, the interbackbone hydrogen bonds reduce the polar character significantly (N-type). Proline is less polar due to the lack of hydrogen-donor capabilities. More detailed geometrical representation is given in Fig.2 D, illustrating the binding distance, angle, dihedral angle, improper angle and bead configuration. This CG protein model contains directional specificity and heterogeneity in side chains to some extent, hence a feature of a secondary structure (α-helix and β-strand) and the gross physicochemical property, such as being charged, hydrophilic and hydrophobic. They succeeded in the partitioning of AAs in the DOPC bilayer, keeping the AA association (Leu-Leu, Lys-Glu) constant in water, the portioning and orientation of pentapeptides at the border of the water and cyclohexane. The tilt and orientation of hexapeptides in the DOPC bilayer is also reproduced after sub-μ sec to μsec MD simulation on GROMACS software (van Der Spoel, Lindahl et al. 2005). They recently accomplished the simulation of the rapid release of content from a pressurised liposome through a particular mechano-sensitive protein channel, MscL, embedded in the liposomal membrane (Louhivuori, Risselada et al. 2010). The behaviour of this tiny functional organelle, which consists of 5 MscL molecules, 2108 DOPC lipids, 5,444 water beads with an additional 54,649 water beads forming a 4 μm layer around the vesicle, was described in almost atomistic detail. In response to the increase of internal pressure, this vesicle released water molecules by opening the Mscl channel. MD was performed for 40 μs, which corresponds to 160 μs in an all-atom model. This model demonstrated that CG-MD provides for the computer-aided design of super-molecules and organelles of a practically usable size.

3.3.5 The optimised potential for efficient peptide-structure representation (OPEP) model

OPEP is, as shown in Fig.2 E, a CG protein model that uses a detailed representation of all backbone atoms (N, H, Cα, C and O) and reduces each side chain to one single bead with appropriate geometrical parameters and physicochemical properties (Derreumaux and Forcellino 2001). The OPEP energy function, which includes the implicit effects of an aqueous solution - expressed as (21) - is formulated as a sum of local potentials (E_{local}), non-bonded potential ($E_{nonbonded}$), and hydrogen-bonding potential ($E_{H\text{-}bond}$):

$$E = E_{local} + E_{nonbonded} + E_{H-bond} \tag{21}$$

Local potentials are expressed by:

$$E_{local} = w_b \sum_{bonds} K_b(r - r_{eq})^2 + w_a \sum_{angle} K_\alpha(\alpha - \alpha_{eq})^2 + w_\Omega \sum_{imp-torsions} k_\Omega(\Omega - \Omega_{eq})^2 + w_{\phi,\psi}(\sum_\phi E_\phi + \sum_\psi E_\psi) \tag{22}$$

K_b, K_a, and K_Ω represent force constants associated with changes in bond length, the bond angles of all particles and force constants related to changes in improper torsions of the side chains. The dihedral potentials associated with N-Cα associated are expressed as (23) and Cα-C expressed as as (24), respectively:

$$E_\phi = k_{\phi\psi}(\phi - \phi_0)^2 \tag{23}$$

$$E_\psi = k_{\phi\psi}(\psi - \psi_0)^2 \tag{24}$$

The non-bonded functions are expressed by:

$$E_{nonbonded} = w_{1,4} \sum_{1,4} E_{VdW} + w_{C\alpha,C\alpha} \sum_{C\alpha,C\alpha} E_{VdW} + w_{1>4} \sum_{M',M'} E_{VdW} +$$
$$w_{1>4} \sum_{M',C\alpha} E_{VdW} + w_{1>4} \sum_{M,SC} E_{VdW} + \sum_{Sc,Sc} w_{Sc,Sc} E_{VdW} \tag{25}$$

which includes all the interaction works through more than 3-bonds, and all these functions are expressed as Van der Waals potentials, as shown in (11):

$$E_{VdW} = \varepsilon_{ij}\left(\left(\frac{r_{ij}^0}{r_{ij}}\right)^{12} - 2\left(\frac{r_{ij}^0}{r_{ij}}\right)^6\right)H(\varepsilon_{ij}) - \varepsilon_{ij}\left(\frac{r_{ij}^0}{r_{ij}}\right)^6 H(-\varepsilon_{ij}) \tag{26}$$

Here, the Heavyside function $H(x) = 1$ if $x >= 0$ and 0 of $x < 0$, r_{ij} is the distance between particles i and j, $r_{ij}^0 = (r_i^0 + r_j^0)/2$ with r_i^0 as the Van der Waals radius of particle i.

The hydrogen-bonding potential ($E_{H\text{-bond}}$) consists of two-body and three-body terms (Derreumaux, Maupetit et al. 2007).

This model was originally devised for predicting the structure and folding of proteins (Derreumaux 1999; Derreumaux and Forcellino 2001) and, by combining a Monte Carlo simulation, fairly succeeded in prefiguring basic supersecondary structures. This model, containing all the protein-backbone components, excels in issues where secondary-structure features play an essential role. They combined this potential with MD, which resulted in reproducing the aggregation of Alzheimer's $A\beta_{16\text{-}22}$, (Derreumaux and Mousseau 2007; Wei, Song et al. 2008). In adopting the sampling of Replica Exchange MD (REMD), they obtained an accurate structural description of Alzheimer's Amyloid-β, β-hairpin and Trp-cage peptides (Derreumaux, Chebaro et al. 2009; Derreumaux, Chebaro et al. 2009). A detailed atomic characterisation of oligomer-formation was obtained by combining OPEP, the atomistic model and REMD (Nasica-Labouze, Meli et al. 2011) Their reduced model on REMD enabled the calculation of several tens of μsec in 40 replicas and the full assessment of convergence to the equilibrium ensemble, demonstrating the probability of determining the thermodynamic features of large proteins and assemblies (Laghaei, Mousseau et al. 2011).

As was mentioned above, the main CG models are tabulated in Table 2.

3.4 The trial for the TCR-pMHC and larger systems

At the starting point of the whole immunological synapse (IS) simulation, Wan, Flower and Coveny constructed a ternary complex of TCR-pMHC-CD4 between opposite membranes - which consists of 329,265 atoms - and performed molecular dynamics for 10 ns on 128 processors of SGI Altix (Wan, Flower et al. 2008). It took 23 hours for one ns simulation. This run was not enough to calculate the binding free-energy by MM/PBSA due to the shortness of the simulation time and the lack of entropy evaluation. They intended to simulate a system consisting of four sets of the TCR-pMHC-CD4 complex, made up of about one

Practical Estimation of TCR-pMHC Binding Free-Energy
Based on the Dielectric Model and the Coarse-Grained Model
125

million atoms. They pointed out the difficulty of the whole IS simulation on the all-atom model due to the too heavy load imposed upon the computer, and pointed out the feasability of adopting the hybrid atomistic/CG simulation for accomplishing the project (Diestler, Zhou et al. 2006).

At present, there have been only very limited trials of evaluated TCR-pMHC binding energy by the CG model. The evaluation of TCR-pMHC binding consists of at least three steps: 1) to determine the binding site, 2) to determine the binding configuration, and 3) to calculate the binding energy. Several works have provided not only the method to determine the binding configuration but also to detect the binding site from the surface nature of its own (Fernandez-Recio, Totrov et al. 2005; Burgoyne and Jackson 2006; Fiorucci and Zacharias 2010). The factors that concern the evaluation of TCR-pMHC binding are: 1) the evaluation of energy from a particular configuration, and 2) the sampling of independent configurations. In most CG models, the calculation of binding energy as the function of the configuration is based on their own parameters (Liwo, Pincus et al. 1993; Miyazawa and Jernigan 1996; Derreumaux 1999; Zacharias 2003; Buchete, Straub et al. 2004; Oldziej, Czaplewski et al. 2005; Zhou, Thorpe et al. 2007; Kim and Hummer 2008; Marrink, Monticelli et al. 2008). The sampling of independent configurations is most time-consuming but critically important process. If the sampling on the CG model reflects the distribution of the atomistic model with reasonable fidelity, it is quite a smart way to sample configurations on a CG-model (Chebaro, Dong et al. 2009), to reconstruct to the atom-scale the structure and then calculate the binding energy on these reconstructed atomistic structures using MM/PBSA. From this point of view, a general method to reconstruct the all-atom from the Cα atom position, RACOGS, was devised and the energy landscapes of both the CG- and the all-atom-model were shown to be quite similar, suggesting the validity of this principle (Heath, Kavraki et al. 2007).

4. Application of GPGPU in molecular dynamics

As mentioned above, all-atom simulation is very expensive, and hence is restricted scope in both time and scale. There have been attempts to breakthrough these circumstances, not only by improving the algorithm but also by devising novel hardware. Special purpose machines for MD have been developed (Susukita, Ebisuzaki et al. 2003; Shaw, Deneroff et al. 2008) and showed fairly good performance (Kikugawa, Apostolov et al. 2009). However, such purpose-specific machines are very expensive and their continuous development is difficult. The recent development of the general purpose graphic processor unit (GPGPU) has had much influence on high performance computing (Giupponi, Harvey et al. 2008). In 2011, three of the top 5 super-computers are constructed mainly on NVIDIA's GPGPU (http://www.top500.org/). Many applications are now being preparing to respond to this momentum, and representative molecular dynamics software such as Amber, CHARM, GROMACS and NAMDA are now being prepared to equip programs working on GPGPUs. Recent representative GPGPUs, such as Tesla C2075, have a performance of 1.03 T Flops on single precision. We calculated the binding energy of two TCR-pMHC complexes, 2GJ6 and 3PWP, on C2075 and compared the results calculated on a Xeon processor. After heating, density-equilibration and equilibration, product runs were performed for 10 runs, corresponding to 5 nsec in total. The results are shown in Table 3. As can be seen from the table, the performance of a Tesla

C2075 is about 40 to 50 cores of a present Xeon CPU. The obtained values are 8% ~ 30% larger by GPGPU than that by CPU. This may be due to the difference of detailed algorithm adopted for the calculations using CPU and GPGPU.

	PDB ID	Product run /run	ΔH	
Xeon E5620 (4 core)	2GJ6	35.7 hours	GB -50.3 ± 8.7	
			PB -54.1 ± 11.8	
	3PWP	34.9 hours	GB -51.3 ± 8.9	
			PB -60.1 ± 11.2	
Tesla C2075	2GJ6	3.09 hour	GB -53.8 ± 7.64	
			PB -63.4 ± 12.5	
	3PWP	3.03 hours	GB -61.5± 6.8	
			PB -78.0 ± 10.9	

Table 3. Comparison of CPU and GPGPU All energies in the table are given in kcal mol^{-1}.

5. Conclusion

Physically meaningful models are rapidly advancing and are being applied to large macro-molecular systems with the rapid evolution of parallel computation and hardware, such as multi-core processors and GPGPUs. Although the exact models become realistic for calculations of large bio-molecules, continuum dielectric models are still useful for the binding free energy calculation and bound complex structure prediction as well as the structure prediction tasks of bio-molecules such as proteins and nucleic acids, etc., because of the high cost performance and fairly good accuracy. In future, hybrid approaches will become promising, where QM model, the all-atom model, the CG model and continuum models are combined with a good conformational sampling technique such as the ER method, and we can choose the optimal hybrid approach according to purpose and the system size.

It has been clear that the calculation of TCR-pMHC binding energy with reasonable efficiency and accuracy is feasible. MMPBSA/GBSA seems quite promising. The sampling method affects both the efficiency and accuracy of the calculation. The combination of sampling on the CG model and energy-calculation on the atomistic model is very reasonable approaches. GPGPUs will be quite important facilities. A combination of those factors will provide for the valid simulation of biologically interesting phenomena for an adequately long time.

6. Acknowledgment

7. References

Asakura, S. & Oosawa, F. (1954). "On interaction between two bodies immersed in a solution of macromolecules." *Journal of Chemical Physics* 22: 1255-1256.
Baker, D. (2000). "A surprising simplicity to protein folding." *Nature* 405(6782): 39-42.

Baxter, E. L.; Jennings, P. A. & Onuchic, J. N. (2011). "Interdomain communication revealed in the diabetes drug target mitoNEET." *Proc Natl Acad Sci U S A* 108(13): 5266-5271.

Ben-Tal, N.; Honig, B.; Bagdassarian, C. K. & Ben-Shaul, A. (2000). "Association entropy in adsorption processes." *Biophys J* 79(3): 1180-1187.

Brown, S.; Fawzi, N. J. & Head-Gordon, T. (2003). "Coarse-grained sequences for protein folding and design." *Proc Natl Acad Sci U S A* 100(19): 10712-10717.

Buchete, N. V.; Straub, J. E. & Thirumalai, D. (2004). "Development of novel statistical potentials for protein fold recognition." *Curr Opin Struct Biol* 14(2): 225-232.

Burgoyne, N. J. & Jackson, R. M. (2006). "Predicting protein interaction sites: binding hot-spots in protein-protein and protein-ligand interfaces." *Bioinformatics* 22(11): 1335-1342.

Car, R. & Parrinello, M. (1985). "Unified approach for molecular dynamics and density-functional theory." *Physical Review Letters* 55(22): 2471-2474.

Carlsson, J.; Ander, M.; Nervall, M. & Aqvist, J. (2006). "Continuum solvation models in the linear interaction energy method." *Journal of Physical Chemistry B* 110(24): 12034-12041.

Case, D. A.; Cheatham, T. E., 3rd; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M. Jr.; Onufriev, A.; Simmerling, C.; Wang, B. & Woods, R. J. (2005). "The Amber biomolecular simulation programs." *J Comput Chem* 26(16): 1668-1688.

Chebaro, Y.; Dong, X.; Laghaei, R.; Derreumaux, P. & Mousseau, N. (2009). "Replica exchange molecular dynamics simulations of coarse-grained proteins in implicit solvent." *J Phys Chem B* 113(1): 267-274.

Derreumaux, P. (1999). "From polypeptide sequences to structures using Monte Carlo simulations and an optimized potential." *Journal of Chemical Physics* 111(5): 2301-2310.

Derreumaux, P.; Chebaro, Y.; Dong, X.; Laghaei, R. & Mousseau, N. (2009). "Replica Exchange Molecular Dynamics Simulations of Coarse-grained Proteins in Implicit Solvent." *Journal of Physical Chemistry B* 113(1): 267-274.

Derreumaux, P.; Chebaro, Y. & Mousseau, N. (2009). "Structures and Thermodynamics of Alzheimer's Amyloid-beta A beta(16-35) Monomer and Dimer by Replica Exchange Molecular Dynamics Simulations: Implication for Full-Length A beta Fibrillation." *Journal of Physical Chemistry B* 113(21): 7668-7675.

Derreumaux, P. & Forcellino, F. (2001). "Computer simulations aimed at structure prediction of supersecondary motifs in proteins." *Proteins-Structure Function and Genetics* 45(2): 159-166.

Derreumaux, P.; Maupetit, J. & Tuffery, P. (2007). "A coarse-grained protein force field for folding and structure prediction." *Proteins-Structure Function and Bioinformatics* 69(2): 394-408.

Derreumaux, P. & Mousseau, N. (2007). "Coarse-grained protein molecular dynamics simulations." *J Chem Phys* 126(2): 025101.

Diestler, D. J.; Zhou, H.; Feng, R. & Zeng, X. C. (2006). "Hybrid atomistic-coarse-grained treatment of multiscale processes in heterogeneous materials: a self-consistent-field approach." *J Chem Phys* 125(6): 64705.

Fernandez-Recio, J.; Totrov, M.; Skorodumov, C. & Abagyan, R. (2005). "Optimal docking area: a new method for predicting protein-protein interaction sites." *Proteins* 58(1): 134-143.

Fiorucci, S. & Zacharias, M. (2010). "Prediction of protein-protein interaction sites using electrostatic desolvation profiles." *Biophys J* 98(9): 1921-1930.

Fogolari, F.; Brigo, A. & Molinari, H. (2002). "The Poisson-Boltzmann equation for biomolecular electrostatics: a tool for structural biology." *Journal of Molecular Recognition* 15(6): 377-392.

Forsten, K. E.; Kozack, R. E.; Lauffenburger, D. A. & Subramaniam, S. (1994). "Numerical solution of the nonlinear Poisson-Boltzmann equation for a membrane-electrolyte system." *J Phys Chem B* 98(21): 5580-5586.

Gilson, M. K. & Honig, B. H. (1987). "Calculation of electrostatic potentials in an enzyme active site." *Nature* 330(6143): 84-86.

Gilson, M. K. & Zhou, H. X. (2007). "Calculation of protein-ligand binding affinities." *Annual Review of Biophysics and Biomolecular Structure* 36: 21-42.

Giupponi, G.; Harvey, M. J. & De Fabritiis, G. (2008). "The impact of accelerator processors for high-throughput molecular modeling and simulation." *Drug Discov Today* 13(23-24): 1052-1058.

Gohlke, H.; Kuhn, L. A. & Case, D. A. (2004). "Change in protein flexibility upon complex formation: Analysis of Ras]Raf using molecular dynamics and a molecular framework approach." *PROTEINS: Structure, Function, and Bioinformatics* 56(2): 322-337.

Gregoire, C.; Lin, S. Y.; Mazza, G.; Rebai, N.; Luescher, I. F. & Malissen, B. (1996). "Covalent assembly of a soluble T cell receptor-peptide-major histocompatibility class I complex." *Proc Natl Acad Sci U S A* 93(14): 7184-7189.

Haliloglu, T.; Keskin, O.; Ma, B. & Nussinov, R. (2005). "How similar are protein folding and protein binding nuclei? Examination of vibrational motions of energy hot spots and conserved residues." *Biophys J* 88(3): 1552-1559.

Harmandaris, V. A. & Kremer, K. (2009). "Dynamics of Polystyrene Melts through Hierarchical Multiscale Simulations." *Macromolecules* 42(3): 791-802.

Harris, R. C.; Bredenberg, J. H.; Silalahi, A. R.; Boschitsch, A. H. & Fenley, M. O. (2011). "Understanding the physical basis of the salt dependence of the electrostatic binding free energy of mutated charged ligand-nucleic acid complexes." *Biophys Chem* 156(1): 79-87.

Hawkins, G. D.; Cramer, C. J. & Truhlar, D. G. (1996). "Parametrized models of aqueous free energies of solvation based on pairwise descreening of solute atomic charges from a dielectric medium." *J Phys Chem B* 100(51): 19824-19839.

Heath, A. P.; Kavraki, L. E. & Clementi, C. (2007). "From coarse-grain to all-atom: toward multiscale analysis of protein landscapes." *Proteins* 68(3): 646-661.

Ho, B. K. & Dill, K. A. (2006). "Folding very short peptides using molecular dynamics." *PLoS Comput Biol* 2(4): e27.

Horejs, C.; Mitra, M. K.; Pum, D.; Sleytr, U. B. & Muthukumar, M. (2011). "Monte Carlo study of the molecular mechanisms of surface-layer protein self-assembly." *J Chem Phys* 134(12): 125103.

Practical Estimation of TCR-pMHC Binding Free-Energy
Based on the Dielectric Model and the Coarse-Grained Model
129

Hou, T.; Wang, J.; Li, Y. & Wang, W. (2011). "Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations." *J Chem Inf Model* 51(1): 69-82.

Hou, T.; Wang, J.; Li, Y. & Wang, W. (2011). "Assessing the performance of the molecular mechanics/Poisson Boltzmann surface area and molecular mechanics/generalized Born surface area methods. II. The accuracy of ranking poses generated from docking." *J Comput Chem* 32(5): 866-877.

Im, W.; Feig, M. & Brooks, C. L. (2003). "An implicit membrane generalized born theory for the study of structure, stability, and interactions of membrane proteins." *Biophys J* 85(5): 2900-2918.

Jang, H.; Hall, C. K. & Zhou, Y. (2004). "Assembly and kinetic folding pathways of a tetrameric beta-sheet complex: molecular dynamics simulations on simplified off-lattice protein models." *Biophys J* 86(1 Pt 1): 31-49.

Janin, J. (2002). "Welcome to CAPRI: A Critical Assessment of Predicted Interactions." *Proteins-Structure Function and Genetics* 47(3): 257-257.

Jayaram, B.; Sprous, D. & Beveridge, D. (1998). "Solvation free energy of biomacromolecules: Parameters for a modified generalized Born model consistent with the AMBER force field." *The Journal of Physical Chemistry B* 102(47): 9571-9576.

Kalra, V. & Joo, Y. L. (2009). "Coarse-grained molecular dynamics study of block copolymer/nanoparticle composites under elongational flow." *J Chem Phys* 131(21): 214904.

Kikugawa, G.; Apostolov, R.; Kamiya, N.; Taiji, M.; Himeno, R.; Nakamura, H. & Yonezawa, Y. (2009). "Application of MDGRAPE-3, a special purpose board for molecular dynamics simulations, to periodic biomolecular systems." *J Comput Chem* 30(1): 110-118.

Kim, Y. C. & Hummer, G. (2008). "Coarse-grained models for simulations of multiprotein complexes: application to ubiquitin binding." *J Mol Biol* 375(5): 1416-1433.

Kinoshita, M. (2006). "Roles of entropic excluded-volume effects in colloidal and biological systems: Analyses using the three-dimensional integral equation theory." *Chemical Engineering Science* 61(7): 2150-2160.

Kinoshita, M. (2009). "Importance of Translational Entropy of Water in Biological Self-Assembly Processes like Protein Folding." *International Journal of Molecular Sciences* 10(3): 1064-1080.

Koehl, P. (2006). "Electrostatics calculations: latest methodological advances." *Current opinion in structural biology* 16(2): 142-151.

Kosmrlj, A.; Read, E. L.; Qi, Y.; Allen, T. M.; Altfeld, M.; Deeks, S. G.; Pereyra, F.; Carrington, M.; Walker, B. D. & Chakraborty, A. K. (2010). "Effects of thymic selection of the T-cell repertoire on HLA class I-associated control of HIV infection." *Nature* 465(7296): 350-354.

Kuhne, T. D.; Krack, M.; Mohamed, F. R. & Parrinello, M. (2007). "Efficient and accurate Car-Parrinello-like approach to Born-Oppenheimer molecular dynamics." *Physical Review Letters* 98(6): 066401.

Laghaei, R.; Mousseau, N. & Wei, G. (2011). "Structure and thermodynamics of amylin dimer studied by Hamiltonian-temperature replica exchange molecular dynamics simulations." *J Phys Chem B* 115(12): 3146-3154.

Lamb, M. L.; Tirado-Rives, J. & Jorgensen, W. L. (1999). "Estimation of the binding affinities of FKBP12 inhibitors using a linear response method." *Bioorganic & medicinal chemistry* 7(5): 851-860.

Lee, M. C.; Yang, R. & Duan, Y. (2005). "Comparison between Generalized-Born and Poisson-Boltzmann methods in physics-based scoring functions for protein structure prediction." *Journal of molecular modeling* 12(1): 101-110.

Leis, S. & Zacharias, M. (2011). "Efficient inclusion of receptor flexibility in grid-based protein-ligand docking*." *J Comput Chem* 32(16): 3433-3439.

Levy, R. M.; Karplus, M. & McCammon, J. A. (1980). "Molecular dynamics studies of NMR relaxation in proteins." *Biophys J* 32(1): 628-630.

Levy, Y.; Cho, S. S.; Onuchic, J. N. & Wolynes, P. G. (2005). "A survey of flexible protein binding mechanisms and their transition states using native topology based energy landscapes." *J Mol Biol* 346(4): 1121-1145.

Liwo, A.; He, Y. & Scheraga, H. A. (2011). "Coarse-grained force field: general folding theory." *Phys Chem Chem Phys* 13(38): 16890-16901.

Liwo, A.; Pincus, M. R.; Wawak, R. J.; Rackovsky, S. & Scheraga, H. A. (1993). "Prediction of protein conformation on the basis of a search for compact structures: test on avian pancreatic polypeptide." *Protein Sci* 2(10): 1715-1731.

Louhivuori, M.; Risselada, H. J.; van der Giessen, E. & Marrink, S. J. (2010). "Release of content through mechano-sensitive gates in pressurized liposomes." *Proc Natl Acad Sci U S A* 107(46): 19856-19860.

Lu, B. Z.; Zhou, Y. C.; Holst, M. J. & McCammon, J. A. (2008). "Recent progress in numerical methods for the Poisson-Boltzmann equation in biophysical applications." *Communications in Computational Physics* 3(5): 973-1009.

Luchko, T.; Gusarov, S.; Roe, D. R.; Simmerling, C.; Case, D. A.; Tuszynski, J. & Kovalenko, A. (2010). "Three-dimensional molecular theory of solvation coupled with molecular dynamics in Amber." *Journal of Chemical Theory and Computation* 6(3): 607-624.

Marrink, S. J.; de Vries, A. H. & Mark, A. E. (2004). "Coarse grained model for semiquantitative lipid simulations." *Journal of Physical Chemistry B* 108(2): 750-760.

Marrink, S. J.; Monticelli, L.; Kandasamy, S. K.; Periole, X.; Larson, R. G. & Tieleman, D. P. (2008). "The MARTINI coarse-grained force field: Extension to proteins." *Journal of Chemical Theory and Computation* 4(5): 819-834.

Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P. & de Vries, A. H. (2007). "The MARTINI force field: coarse grained model for biomolecular simulations." *J Phys Chem B* 111(27): 7812-7824.

Matubayasi, N. & Nakahara, M. (2000). "Theory of solutions in the energetic representation. I. Formulation." *Journal of Chemical Physics* 113(15): 6070-6081.

Matubayasi, N. & Nakahara, M. (2002). "Theory of solutions in the energy representation. II. Functional for the chemical potential." *Journal of Chemical Physics* 117(8): 3605-3616.

May, A. & Zacharias, M. (2007). "Protein-protein docking in CAPRI using ATTRACT to account for global and local flexibility." *Proteins* 69(4): 774-780.

Practical Estimation of TCR-pMHC Binding Free-Energy
Based on the Dielectric Model and the Coarse-Grained Model
131

Mehler, E. L. & Guarnieri, F. (1999). "A self-consistent, microenvironment modulated screened coulomb potential approximation to calculate pH-dependent electrostatic effects in proteins." *Biophys J* 77(1): 3-22.

Michielin, O. & Karplus, M. (2002). "Binding free energy differences in a TCR-peptide-MHC complex induced by a peptide mutation: a simulation analysis." *Journal of Molecular Biology* 324(3): 547-569.

Miyazawa, S. & Jernigan, R. L. (1996). "Residue-residue potentials with a favorable contact pair term and an unfavorable high packing density term, for simulation and threading." *J Mol Biol* 256(3): 623-644.

Nakamura, H.; Sakamoto, T. & Wada, A. (1988). "A theoretical study of the dielectric constant of protein." *Protein engineering* 2(3): 177.

Nasica-Labouze, J.; Meli, M.; Derreumaux, P.; Colombo, G. & Mousseau, N. (2011). "A multiscale approach to characterize the early aggregation steps of the amyloid-forming peptide GNNQQNY from the yeast prion sup-35." *PLoS Comput Biol* 7(5): e1002051.

Oldziej, S.; Czaplewski, C.; Liwo, A.; Chinchio, M.; Nanias, M.; Vila, J. A.; Khalili, M.; Arnautova, Y. A.; Jagielska, A.; Makowski, M.; Schafroth, H. D.; Kazmierkiewicz, R.; Ripoll, D. R.; Pillardy, J.; Saunders, J. A.; Kang, Y. K.; Gibson, K. D. & Scheraga, H. A. (2005). "Physics-based protein-structure prediction using a hierarchical protocol based on the UNRES force field: assessment in two blind tests." *Proc Natl Acad Sci U S A* 102(21): 7547-7552.

Onufriev, A.; Bashford, D. & Case, D. A. (2004). "Exploring protein native states and large-scale conformational changes with a modified generalized born model." *Proteins* 55(2): 383-394.

Pickersgill, R. W. (1988). "A rapid method of calculating charge-charge interaction energies in proteins." *Protein Eng* 2(3): 247-248.

Pitera, J. W.; Falta, M. & Van Gunsteren, W. F. (2001). "Dielectric properties of proteins from simulation: The effects of solvent, ligands, pH, and temperature." *Biophys J* 80(6): 2546-2555.

Posel, Z.; Lísal, M. & Brennan, J. K. (2009). "Interplay between microscopic and macroscopic phase separations in ternary polymer melts: Insight from mesoscale modelling." *Fluid Phase Equilibria* 283(1-2): 38-48.

Rastelli, G.; Rio, A. D.; Degliesposti, G. & Sgobba, M. (2010). "Fast and accurate predictions of binding free energies using MM]PBSA and MM]GBSA." *J Comput Chem* 31(4): 797-810.

Scheraga, H. A.; Maisuradze, G. G.; Senet, P.; Czaplewski, C. & Liwo, A. (2010). "Investigation of Protein Folding by Coarse-Grained Molecular Dynamics with the UNRES Force Field." *Journal of Physical Chemistry A* 114(13): 4471-4485.

Shaw, D. E.; Deneroff, M. M.; Dror, R. O.; Kuskin, J. S.; Larson, R. H.; Salmon, J. K.; Young, C.; Batson, B.; Bowers, K. J.; Chao, J. C.; Eastwood, M. P.; Gagliardo, J.; Grossman, J. P.; Ho, C. R.; Ierardi, D. J.; Kolossvary, I.; Klepeis, J. L.; Layman, T.; Mcleavey, C.; Moraes, M. A.; Mueller, R.; Priest, E. C.; Shan, Y. B.; Spengler, J.; Theobald, M.; Towles, B. & Wang, S. C. (2008). "Anton, a special-purpose machine for molecular dynamics simulation." *Communications of the Acm* 51(7): 91-97.

Shih, A. Y.; Arkhipov, A.; Freddolino, P. L. & Schulten, K. (2006). "Coarse grained protein-lipid model with application to lipoprotein particles." *J Phys Chem B* 110(8): 3674-3684.

Sorin, E. J.; Rhee, Y. M.; Nakatani, B. J. & Pande, V. S. (2003). "Insights into nucleic acid conformational dynamics from massively parallel stochastic simulations." *Biophys J* 85(2): 790-803.

Srinivasan, J.; Trevathan, M. W.; Beroza, P. & Case, D. A. (1999). "Application of a pairwise generalized Born model to proteins and nucleic acids: inclusion of salt effects." *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)* 101(6): 426-434.

Still, W. C.; Tempczyk, A.; Hawley, R. C. & Hendrickson, T. (1990). "Semianalytical treatment of solvation for molecular mechanics and dynamics." *Journal of the American Chemical Society* 112(16): 6127-6129.

Susukita, R.; Ebisuzaki, T.; Elmegreen, B. G.; Furusawa, H.; Kato, K.; Kawai, A.; Kobayashi, Y.; Koishi, T.; McNiven, G. D.; Narumi, T. & Yasuoka, K. (2003). "Hardware accelerator for molecular dynamics: MDGRAPE-2." *Computer Physics Communications* 155(2): 115-131.

Takahashi, T.; Nakamura, H. & Wada, A. (1992). "Electrostatic forces in two lysozymes: calculations and measurements of histidine pKa values." *Biopolymers* 32(8): 897-909.

Takahashi, T.; Sugiura, J. & Nagayama, K. (2002). "Comparison of all atom, continuum, and linear fitting empirical models for charge screening effect of aqueous medium surrounding a protein molecule." *Journal of Chemical Physics* 116(18): 8232-8237.

Taketomi, H.; Ueda, Y. & Go, N. (1975). "Studies on protein folding, unfolding and fluctuations by computer simulation. I. The effect of specific amino acid sequence represented by specific inter-unit interactions." *Int J Pept Protein Res* 7(6): 445-459.

Tidor, B. & Karplus, M. (1994). "The Contribution of Vibrational Entropy to Molecular Association:: The Dimerization of Insulin." *Journal of Molecular Biology* 238(3): 405-414.

Tsui, V. & Case, D. A. (2000). "Theory and applications of the generalized Born solvation model in macromolecular simulations." *Biopolymers* 56(4): 275-291.

Turjanski, A. G.; Gutkind, J. S.; Best, R. B. & Hummer, G. (2008). "Binding-induced folding of a natively unstructured transcription factor." *PLoS Comput Biol* 4(4): e1000060.

Van Der Spoel, D.; Lindah, E.; Hess, B.; Groenhof, G.; Mark, A.E. & Berendsen, H.J.C. (2005) "GROMACS: fast, flexible and free." J Comput Chem 26(16): 1701-1718

Wagoner, J. A. & Pande, V. S. (2011). "A smoothly decoupled particle interface: New methods for coupling explicit and implicit solvent." *J Chem Phys* 134: 214103.

Wan, S.; Coveney, P. V. & Flower, D. R. (2005). "Molecular basis of peptide recognition by the TCR: affinity differences calculated using large scale computing." *J Immunol* 175(3): 1715-1723.

Wan, S.; Flower, D. R. & Coveney, P. V. (2008). "Toward an atomistic understanding of the immune synapse: large-scale molecular dynamics simulation of a membrane-embedded TCR-pMHC-CD4 complex." *Mol Immunol* 45(5): 1221-1230.

Warshel, A. & Levitt, M. (1976). "Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme." *Journal of Molecular Biology* 103(2): 227-249.

Warwicker, J. (1999). "Simplified methods for pKa and acid pH-dependent stability estimation in proteins: removing dielectric and counterion boundaries." *Protein science* 8(02): 418-425.

Wei, G. H.; Song, W.; Derreumaux, P. & Mousseau, N. (2008). "Self-assembly of amyloid-forming peptides by molecular dynamics simulations." *Frontiers in Bioscience* 13: 5681-5692.

Weiser, J.; Shenkin, P. S. & Still, W. C. (1999). "Approximate atomic surfaces from linear combinations of pairwise overlaps (LCPO)." *J Comput Chem* 20(2): 217-230.

Wucherpfennig, K. W.; Call, M. J.; Deng, L. & Mariuzza, R. (2009). "Structural alterations in peptide-MHC recognition by self-reactive T cell receptors." *Curr Opin Immunol* 21(6): 590-595.

Xu, Y.; Wang, C. & Shi, Y. (1992). "Improvements on the protein-dipole Langevin-dipole model." *J Comput Chem* 13(9): 1109-1113.

Xu, Y. & Wang, R. (2006). "A computational analysis of the binding affinities of FKBP12 inhibitors using the MM]PB/SA method." *PROTEINS: Structure, Function, and Bioinformatics* 64(4): 1058-1068.

Yaoita, T.; Isaki, T.; Masubuchi, Y.; Watanabe, H.; Ianniruberto, G.; Greco, F. & Marrucci, G. (2008). "Statics, linear, and nonlinear dynamics of entangled polystyrene melts simulated through the primitive chain network model." *J Chem Phys* 128(15): 154901.

Yasuda, S.; Yoshidome, T.; Oshima, H.; Kodama, R.; Harano, Y. & Kinoshita, M. (2010). "Effects of side chain packing on the formation of secondary structures in protein folding." *J Chem Phys* 132(6): 065105(065101-065110).

Yoshidome, T.; Ito, Y.; Ikeguchi, M. & Kinoshita, M. (2011). "Rotation Mechanism of F1-ATPase: Crucial Importance of the Water Entropy Effect." *Journal of the American Chemical Society* 133: 4030-4039.

Zacharias, M. (2003). "Protein-protein docking with a reduced protein model accounting for side chain flexibility." *Protein Sci* 12(6): 1271-1282.

Zacharias, M. & Fiorucci, S. (2010). "Binding site prediction and improved scoring during flexible protein-protein docking with ATTRACT." *Proteins-Structure Function and Bioinformatics* 78(15): 3131-3139.

Zhou, J.; Thorpe, I. F.; Izvekov, S. & Voth, G. A. (2007). "Coarse-grained peptide modeling using a systematic multiscale approach." *Biophys J* 92(12): 4289-4303.

Zhou, R. H. (2003). "Free energy landscape of protein folding in water: Explicit vs. implicit solvent." *Proteins-Structure Function and Genetics* 53(2): 148-161.

Zoete, V.; Irving, M. & Michielin, O. (2010). "MM-GBSA binding free energy decomposition and T cell receptor engineering." *Journal of Molecular Recognition* 23(2): 142-152.

Zoete, V. & Meuwly, M. (2006). "Importance of individual   side chains for the stability of a protein fold: computational alanine scanning of the insulin monomer." *J Comput Chem* 27(15): 1843-1857.

Zoete, V. & Michielin, O. (2007). "Comparison between computational alanine scanning and per-residue binding free energy decomposition for protein-protein association using MM-GBSA: application to the TCR-p-MHC complex." *Proteins* 67(4): 1026-1047.
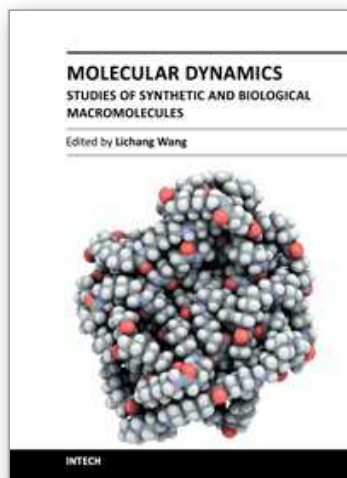
Molecular Dynamics - Studies of Synthetic and Biological Macromolecules

Edited by Prof. Lichang Wang

ISBN 978-953-51-0444-5

Hard cover, 432 pages

Publisher InTech

Published online 11, April, 2012

Published in print edition April, 2012

Molecular Dynamics is a two-volume compendium of the ever-growing applications of molecular dynamics simulations to solve a wider range of scientific and engineering challenges. The contents illustrate the rapid progress on molecular dynamics simulations in many fields of science and technology, such as nanotechnology, energy research, and biology, due to the advances of new dynamics theories and the extraordinary power of today's computers. This second book begins with an introduction of molecular dynamics simulations to macromolecules and then illustrates the computer experiments using molecular dynamics simulations in the studies of synthetic and biological macromolecules, plasmas, and nanomachines. Coverage of this book includes: Complex formation and dynamics of polymers Dynamics of lipid bilayers, peptides, DNA, RNA, and proteins Complex liquids and plasmas Dynamics of molecules on surfaces Nanofluidics and nanomachines

INTECH

open science | open minds