Study of the Insulin Dimerization: Binding Free Energy Calculations and Per-Residue Free Energy Decomposition

Vincent Zoete,[1,2] Markus Meuwly,[2*] and Martin Karplus[1,3*]
[1]*Laboratoire de Chimie Biophysique, ISIS/Université Louis Pasteur, Strasbourg Cedex, France*
[2]*Chemistry Department, University of Basel, Basel, Switzerland*
[3]*Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts*

ABSTRACT **A calculation of the binding free energy for the dimerization of insulin has been performed using the molecular mechanics–generalized Born surface area approach. The calculated absolute binding free energy is −11.9 kcal/mol, in approximate agreement with the experimental value of −7.2 kcal/mol. The results show that the dimerization is mainly due to nonpolar interactions. The role of the hydrogen bonds between the 2 monomers appears to give the direction of the interactions. A per-atom decomposition of the binding free energy has been performed to identify the residues contributing most to the self association free energy. Residues B24–B26 are found to make the largest favorable contributions to the dimerization. Other residues situated at the interface between the 2 monomers were found to make favorable but smaller contributions to the dimerization: Tyr B16, Val B12, and Pro B28, and to an even lesser extent, Gly B23. The energy decomposition on a per-residue basis is in agreement with experimental alanine scanning data. The results obtained from a single trajectory (i.e., the dimer trajectory is also used for the monomer analysis) and 2 trajectories (i.e., separate trajectories are used for the monomer and dimer) are similar. Proteins 2005;61:79–93.**

Key words: insulin; insulin dimerization; protein association; MM-GBSA; binding free energy; free energy decomposition; molecular dynamics; continuum model; GB-MV2

INTRODUCTION

Insulin is a 51-amino acid hormone that regulates the physiological glucose level in blood.[1] While synthesized and stored in the pancreas as a zinc-containing hexamer, which can be considered a trimer of insulin dimers, the insulin hormone interacts with its cell surface receptor as a monomer.[2] The insulin monomer contains 2 chains (Fig. 1): the A chain consists of 21 amino acids, and the B chain has 30 amino acids. When insulin monomers associate to form dimers, the extended C-terminal ends of the B chains (residues B22–B30) of the 2 molecules are brought together, forming a 2-stranded antiparallel β-sheet (Fig. 2). There are also contacts between the B-chain helices (residues B9, B12, B13, and B16). The interactions stabilizing the dimer are thought to be predominantly nonpolar,[3,4] with the β-sheet hydrogen bonds mainly replacing water hydrogen bonds and contributing to orienting the 2 monomers, rather than to their stability.

The rate-limiting step of insulin absorption by patients treated with conventional solution of native insulin appears to be the dissociation of the insulin hexamer into dimers and finally into the active monomer.[5,6] Treatment with conventional insulin is not able to mimic effectively the normal rapid insulin release following a meal, nor the basal insulin concentration between meals. To improve insulin therapy, new insulin analogues have been developed.[5,6] For example, two rapid-acting insulin analogues, insulin Lispro and insulin Aspart, have been designed with a reduced tendency to self-associate. This is obtained by the B28(Pro→Lys) and B29(Lys→Pro) reversal in the sequence compared to the native insulin for insulin Lispro,[7] and by the B28(Pro→Asp) mutation for insulin Aspart.[8] Information about energetics of the insulin dimerization at an atomistic level could be useful, since it would allow the identification of the most important residues for insulin association. This information could be used to design new insulin analogues with decreased propensity for self-association and enhanced pharmacokinetic properties.

Although free energy simulations[9,10] are the most accurate methods for studying the effect of mutations on association, they are still time-consuming to employ, despite increased computer speed. Consequently, a number of simplified approaches have been developed. One of these has been exploited by Kollman and coworkers under the name molecular mechanics-Poisson–Boltzmann surface

---

The Supplementary Materials referred to in this article can be found at http://www.interscience.wiley.com/jpages/0887-3585/suppmat/

Grant sponsor: Insulin Dependent Diabetes Trust (IDDT) (to V. Zoete and M. Meuwly in Basel). Grant sponsor: CHARMM Development Project (to V. Zoete and M. Karplus in Strasbourg). Grant sponsor: National Institutes of Health (to M. Karplus at Harvard). Grant sponsor: Schweizerischer Nationalfonds for a Förderungsprofessur (to M. Meuwly).

*Correspondence to: Martin Karplus, Department of Chemistry and Biological Chemistry, Harvard University, 12 Oxford Street, Cambridge, MA 02138. E-mail: marci@tammy.harvard.edu; and to Markus Meuwly, Chemistry Department, University of Basel, Klingelbergstrasse 80, CH-4056 Basel, Switzerland. E-mail: m.meuwly@unibas.ch

Received 6 October 2004; Revised 22 December 2004; Accepted 24 January 2005

Published online 3 August 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.20528

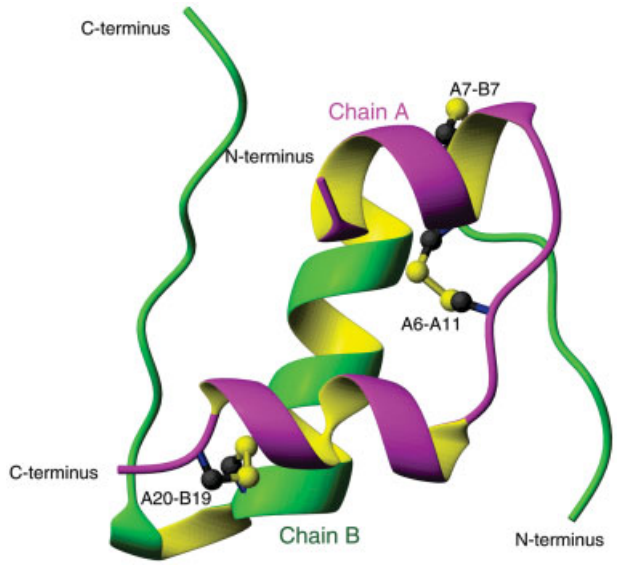

Fig. 1. Structure of insulin monomer, from 4INS (monomer 1). Chain A and chain B are indicated. Disulfide bonds are shown in ball-and-stick representation.

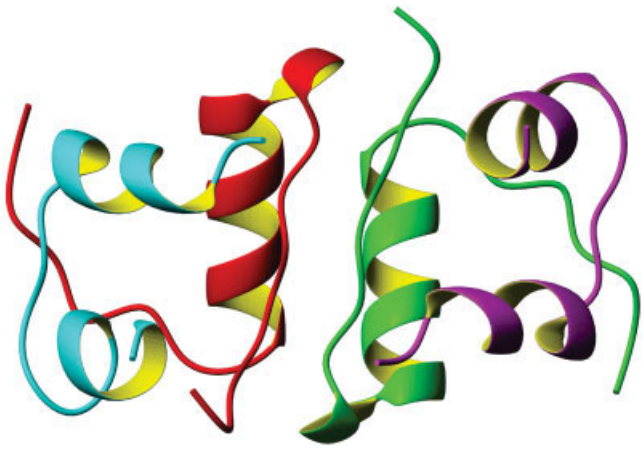

Fig. 2. Structure of insulin dimer, from 4INS (monomers 1 and 2).

area (MM-PBSA) approach;[11] similar methods have been used by others.[12] In this approach, a thermodynamic cycle is considered, and the binding free energy is estimated as the sum of the gas-phase energies, solvation free energies, and entropic contributions averaged over a series of snapshots from molecular dynamics (MD) trajectories. The gas-phase terms are obtained by molecular mechanics calculations, whereas the electrostatic contribution to the solvation term is calculated by solving the Poisson–Boltzmann (PB) equation.[13] Approaches using a free energy decomposition based on a similar thermodynamic cycle have been proposed,[14,15] with some of them introducing also an averaging of the energy terms along an MD trajectory.[16] Combinations of MD simulations with PB free energy calculations have been used previously.[17–20] The MM-PBSA approach has been applied to estimate the binding energy for protein–protein[21,22] and protein–ligand systems.[23–25] It has also been used to predict the effect of residue mutations on the binding energy of protein–protein systems with the "computational alanine scanning" approach.[26,27] Recently, a related approach, where the electrostatic contribution to the solvation energy is determined using a generalized Born (GB) model, has been introduced.[11,28] Despite its generally lower accuracy,[29] the GB model makes this variant attractive for 2 reasons. First, the GB approach is much faster than the PB approach. Second, GB allows one to decompose easily and rapidly the electrostatic solvation energy, and thus the binding free energy, into atomic contributions from only 1 calculation (see Methods section and Gohlke et al.[22]). A decomposition of the binding free energy in the context of PB calculations is also possible[12,20] but requires separate, time-consuming calculations. The decomposition gives insights into the origin of the binding at an atomistic level. The per-atom contributions can be summed over atom groups such as residues, backbones, and side-chains, to obtain their contributions to the binding free energy, offering a faster alternative to the computational alanine scanning for the detailed study of protein–protein interactions at the residue level.[22] A detailed and methodologically important analysis of the question of convergence, as well as of the dependence of the results upon the parameters and the physical models used to calculate the different energy terms in the MM-PB(GB)SA free energy estimates, has been published[30]; the study uses the Ras–Raf protein–protein complex as an example.

In the present work, we use the GB-MV2 model[31,32] to calculate the electrostatic solvation contribution to the binding free energy. This GB model was found to reproduce the solvation free energies calculated by solving the PB equation with 1% accuracy. We estimate the absolute binding free energy of insulin dimerization using MD sampling with the GB-MV2 model. A single dimer trajectory is used in most of the analysis (i.e., the monomer is sampled from the dimer trajectory). The binding free energy is decomposed on a per-atom basis to evaluate the contribution of the backbone and side-chains of each residues to the dimerization, and to identify the most important residues for insulin association. The computational methods employed are outlined in the Computational Methods section. The Results and Discussion section presents and discusses the results, and the Conclusions section follows. A comparison between using a single trajectory for the dimer and monomer and using separate trajectories is provided in the Appendix and in the Supplementary Material.

COMPUTATIONAL METHODS

The starting coordinates for the MD simulation were taken from the X-ray structure of the native porcine insulin Zn-hexamer resolved at 1.5 Å [Protein Data Bank[33,34] (PDB) entry 4INS[35]). Since we are interested in the insulin dimer, the zinc atoms, which take part in the insulin hexamerization, were removed. For disordered side-chains, with multiple conformations in the PDB file (i.e., Gln B4, Val B12, Glu B21, and Thr B27 for monomer 1, Lys B29 of monomer 2, and Arg B22 of monomers 1 and

2), only the first proposed conformation was used. The absence of sterical clashes was checked. Titratable groups were taken in their standard protonation state at neutral pH; they are Glu A4, A17, B13 and B21, Arg B22, and Lys B29. Based on visual inspection of their environment, His B10 was charged, while His B5 was set neutral and protonated at N^{δ}. The overall charge is −2 for the dimer. All calculations, except solving the PB equation, were carried out using CHARMM[36] (versions c29 and c30), and the molecules were described by the "all-atom" force field CHARMM22.[37] There also exists a cubic crystal structure with a 2-fold crystallographic symmetry axis between the 2 monomers (9INS in PDB).[38] Thus, the 2 monomers are equivalent. However, we have not used this structure for 2 reasons. First, we wanted to compare the results of the MD simulation performed for the insulin dimer[39] with a previous study that used the 4INS PDB file as a starting structure.[40] Second, a crystal structure of a mutant insulin dimer is available.[41]* This structure is not symmetric: One important difference between the 2 structures is that an ionic interaction between the C-terminus of the B chain and the N terminus of the A chain is present in one monomer but not the other, as in 4INS. Thus, that the two monomers are equivalent in the dimer of the 9INS structure may be an artifact of the crystallographic symmetry and apparently is not systematic in insulin dimer structures. This supports the use of an asymmetric insulin dimer in the present study; its use also permits a test of the effect of structural changes on the dimerization free energy analysis (see Supplementary Material). In the dimerization interface, the main conformational differences between the insulin monomers in 9INS and 4INS are found for Glu B13, Phe B25, Lys B29, and Ala B30. With the exception of Phe B25, none of these residues were found to be important for the dimerization (see discussion below). It should be noted that the structure of the insulin monomer and the insulin dimer in solution are not known experimentally. We have chosen to use the indicated structures. Since, as shown below, the binding free energies calculated from different trajectories using the single trajectory and the different trajectory methods (see Computational Methods section) agree, it is likely that the results are not sensitive to the exact structure used for the initial coordinates.

Molecular Dynamics Simulations

The insulin dimer was solvated in a rectangular 59.0 × 49.7 × 40.4 $Å^3$ box of TIP3P[42] water previously equilibrated at 300 K and 1 atm pressure. Water molecules overlapping the insulin dimer were removed, leading to a system with 3329 water molecules and a total of 11,553 atoms. The solvent was equilibrated at 300 K during 50 ps in the presence of the fixed insulin dimer. Then, the dimer structure was minimized for 500 steps of steepest descent (SD) minimization. The entire system was heated to 300 K during 6 ps and equilibrated during 100 ps at 300 K. The MD simulation was conducted with periodic boundary conditions[9] (PBC) during 5 ns. The Verlet leapfrog integrator was used for time propagation with a timestep of 0.001 ps. A 12 Å cutoff was applied to the shifted electrostatic and switched van der Waals interactions.

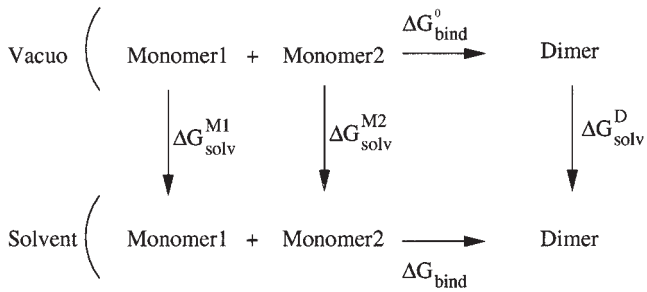

Fig. 3. Thermodynamic cycle used to calculate the binding free energies between the 2 monomers in the gas phase, ΔG^0_{bind}, and in water, ΔG_{bind}. ΔG^{M1}_{solv}, ΔG^{M2}_{solv}, and ΔG^{D}_{solv} are the solvation free energies of monomer 1, monomer 2, and of the dimer, respectively.

Binding Free Energies for Insulin Dimerization

General equation

The absolute binding free energy between the 2 insulin monomers was calculated according to the thermodynamic cycle shown in Figure 3, using the MM-GBSA approach.[11,28] In this approach, the binding free energy, ΔG_{bind}, is written as the sum of the gas phase contribution, ΔG^0_{bind}, the desolvation energy of the system upon binding, ΔG_{desolv}, and an entropic contribution, $-T\Delta S$:

$$\Delta G_{bind} = \langle \Delta G^0_{bind} \rangle + \langle \Delta G_{desolv} \rangle - \langle T\Delta S \rangle. \quad (1)$$

The brackets, ⟨⟩, indicate an average of these energy terms along the MD simulation trajectory (see below).

Gas-phase contribution

The gas-phase contribution to the binding free energy is equal to the sum of the van der Waals and electrostatic interaction energies between the two monomers, E_{vdW} and E_{elec}, and ΔE_{intra}, the difference in the internal energy of the monomers between the dimer, D, and the isolated monomers, $M1$ and $M2$:

$$\Delta G^0_{bind} = \Delta E_{intra} + E_{vdW} + E_{elec} \quad (2)$$

$$\Delta E_{intra} = E^D_{intra} - (E^{M1}_{intra} + E^{M2}_{intra}) \quad (3)$$

$$E^X_{intra} = E^X_{intra,bond} + E^X_{intra,vdW} + E^X_{intra,elec}, \quad (4)$$

$E^X_{intra,bond}$, $E^X_{intra,vdW}$ and $E^X_{intra,elec}$ are the energy of the bonded terms (bonds, angles, dihedral angles, and improper angles) within a given molecule X, and the van der Waals and electrostatic interactions between the atoms of this molecule, respectively. ΔE_{intra}, E_{elec}, and E_{vdW} are calculated according to the CHARMM22 molecular mechanics force field, with a dielectric constant of 1. As explained in the subsection on MD simulations, a 12 Å

*Since the 1B9E crystal structure corresponds to a mutant, and since the N-terminal part of the B-chain is in the open form instead of the more common closed form (see discussion in Computational Methods section of Zoete et al.[39]), we chose not to use this structure to study the insulin dimerization of wild-type insulin.

cutoff was applied to the electrostatic and van der Waals interactions during the MD simulation. However, no cutoff was used to calculate the energy terms of Eq. (2), since the postprocessing is not as computationally demanding as the MD simulation, and to be consistent with the use of the continuum solvent models described later that do not use any cutoff. The screening effect of the solvent on the electrostatic interactions is taken into account in the solvation terms calculated using these continuum solvent models (discussed later).

Entropy contribution

The entropy term is decomposed into translational, ΔS_{trans}, rotational, ΔS_{rot}, and vibrational, ΔS_{vib}, contributions:

$$-T\Delta S = -T\Delta S_{vib} - T\Delta S_{trans} - T\Delta S_{rot} \quad (5)$$

These terms are calculated using standard equations of statistical mechanics.[43,44] $T\Delta S_{trans}$ and $T\Delta S_{rot}$ are functions of the mass and moments of inertia of the molecule, respectively. The vibrational entropy term is calculated with the quantum formula from a normal mode analysis.[44] The VIBRAN module of the CHARMM program was used to calculate and diagonalize the force constant matrix to determine the normal mode vectors and frequencies. Normal modes were calculated for the fully minimized structures of the molecule (dimer or isolated monomer) in vacuo with a distance dependent dielectric, $\epsilon = 4r$. Starting from the X-ray structure, the minimization was performed using the Adopted Basis Newton–Raphson minimization algorithm, until the root-mean-square of the energy gradient reached a value of 10^{-7} kcal/mol/Å. This gradient has been shown to be satisfactory for calculating normal modes and is expected to yield real frequencies only.[44] These minimized structures were also used to calculate the moments of inertia of the molecules. The root-mean-square deviation (RMSD) for the backbone atoms between the structures before and after minimization is around 0.9–1.4 Å for the dimer and 1.2–2.0 Å for the isolated monomers (0.9–1.1 Å if residues B25 to B30, which are particularly flexible in the isolated monomer,[39] are removed from the RMSD calculation). The normal mode analysis is computationally demanding, so that $-T\Delta S$ was averaged only over 50 frames of the MD trajectory.

Contribution of the desolvation

ΔG_{desolv} is the difference between the solvation energy of the dimer, ΔG^{D}_{solv}, and those of the isolated monomers, ΔG^{M1}_{solv} and ΔG^{M2}_{solv} (Fig. 3):

$$\Delta G_{desolv} = \Delta G^{D}_{solv} - (\Delta G^{M1}_{solv} + \Delta G^{M2}_{solv}). \quad (6)$$

The solvation energy, ΔG^{X}_{solv}, where X stands for the dimer or each of the isolated monomers, is divided into the electrostatic contribution, $\Delta G^{X}_{elec,solv}$ and the nonpolar contribution $\Delta G^{X}_{np,solv}$:

$$\Delta G^{X}_{solv} = \Delta G^{X}_{elec,solv} + \Delta G^{X}_{np,solv}. \quad (7)$$

The nonpolar contribution to the solvation free energy, which can be considered as the sum of a solvent–solvent cavity term and a solute–solvent van der Waals term, is assumed to be proportional to the solvent-accessible surface area (SASA) (i.e., $\Delta G_{np,solv} = \sigma \times \text{SASA}$). This well known and often used approximation comes from the fact that the solvation energy of saturated nonpolar hydrocarbons is linearly related to the SASA.[45,46] We used a value of 0.0072 kcal/mol/Å^2 for the parameter σ.[22,47,48] The SASAs were calculated analytically with CHARMM.

Two methods can be used to calculate the electrostatic contribution to the solvation free energy, $\Delta G^{X}_{elec,solv}$. First, $\Delta G^{X}_{elec,solv}$ can be calculated by solving the Poisson equation. In that case, the method is called the molecular mechanics–Poisson surface area (MM-PSA) method. In this study, we actually solved the Poisson equation, since no ionic strength was used. The second possibility is to calculate $\Delta G_{elec,solv}$ using the analytical generalized Born GB-MV2 model implemented in CHARMM.[31,32] This model was found to reproduce the solvation free energies calculated by solving the Poisson equation with 1% accuracy. The second method is much faster than the first one (by a factor of about 20) and is therefore very useful to calculate $\Delta G_{elec,solv}$ for a large number of structures. Moreover, GB-MV2 allows one to decompose the electrostatic contribution to the binding free energy on a per-atom basis in a more straightforward manner than the Poisson equation. The GB-MV2 method was therefore used to calculate the results that are given here. Some solvation free energies were also calculated using the Poisson equation to assess the reliability of the GB-MV2 results. For this purpose, the University of Houston Brownian Dynamics (UHBD) program,[49] with an interface to CHARMM,[50] was utilized in conjunction with the CHARMM all-hydrogen parameter set (i.e., the atom partial charges and radii from CHARMM22 were employed). The electrostatic solvation free energy was calculated with a 2-step procedure. A first calculation was made using a 3.0 Å grid spacing and a solvent layer (i.e., space between the structure and the closest border of the grid) of about 20 Å around the complex. This provided the boundary potential for a second "focusing" calculation, which used a grid spacing of 0.40 Å and a 5 Å borderspace. The dielectric constant of the protein was set to 1, and a dielectric constant of 80 was used for the exterior to represent an aqueous environment. A probe sphere radius of 1.4 Å was used.

With the exception of the entropy terms (see above), all the energy terms were calculated for 500 frames regularly separated by 10 ps along the 5 ns trajectory performed for the insulin dimer. After averaging the energy terms and grouping Eqs. (1), (2), (6), and (7), one finds:

$$\langle \Delta G_{bind} \rangle = \langle \Delta E_{intra} \rangle + \langle E_{vdW} \rangle + \langle E_{elec} \rangle + \langle \Delta G_{elec,desolv} \rangle + \langle \Delta G_{np,desolv} \rangle - \langle T\Delta S \rangle, \quad (8)$$

with $\Delta G_{elec,desolv} = [\Delta G^{D}_{elec,solv} - (\Delta G^{M1}_{elec,solv} + \Delta G^{M2}_{elec,solv})]$ and $\Delta G_{np,desolv} = \sigma \times (SASA^{D} - (SASA^{M1} + SASA^{M2}))$.

In principle, one could run separate trajectories for the complex and each of the isolated parts, and then evaluate

the binding free energy using these MD trajectories. Here this approach is called the "different trajectory method" (DTM). However, by using the same trajectory for both the dimer and the monomer (the "same trajectory method", STM), there is considerable cancellation and more stable results are expected. We use this approach, which has also been employed by others.[26,28,51] The results obtained using separate trajectories are given in the Appendix.

Binding free energy decomposition

A binding free energy decomposition at the atomic level was made to evaluate the contribution of each residue to the total binding free energy, as well as the contributions of its side-chain and backbone. For this purpose, one half of a pairwise electrostatic interaction energy between 2 atoms, each belonging to a different monomer, is attributed to both of them. Therefore, the contribution of atom i to the total electrostatic interaction energy between the 2 monomers is given by

$$E^i_{elec} = \frac{1}{2} \sum_j \frac{q_i q_j}{r_{ij}}, \quad (9)$$

where j loops over all the atoms of the monomer to which i does not belong to. r_{ij} is the distance between the 2 atoms with charges q_i and q_j, respectively. Similarly, one-half of the pairwise intermonomer van der Waals interaction energies between the 2 monomers is attributed to the atoms that are part of the interaction pairs, which avoids double-counting.

In the STM, ΔE_{intra} is equal to zero, since the internal energies of the complex and the separated parts are calculated from the same trajectory. Thus, there is no need for a per-atom decomposition. However, in the DTM, this term is important and its decomposition is described in the Appendix.

The SASA of each atom i in the dimer, $\mathrm{SASA}^{i,D}$, and in each of the 2 monomers, $\mathrm{SASA}^{i,M1}$ and $\mathrm{SASA}^{i,M2}$, are calculated by CHARMM. The contribution of this atom to the nonpolar solvation term is $\Delta G^i_{np,solv} = \sigma \times [\mathrm{SASA}^{i,d} - (\mathrm{SASA}^{i,M1} + \mathrm{SASA}^{i,M2})]$, where $\mathrm{SASA}^{i,M1}$ or $\mathrm{SASA}^{i,M2}$ is equal to zero depending on what insulin monomer the atom belongs to.

The GB-MV2 approach uses the Still et al.[48] expression to calculate the electrostatic solvation energy term:

$$\Delta G_{elec,solv} = k \sum_{i,j} \frac{q_i q_j}{\sqrt{r_{ij}^2 + \alpha_i \alpha_j \exp(-r_{ij}^2/K_s \alpha_i \alpha_j)}}, \quad (10)$$

where $k = -166.0(\epsilon^{-1}_{solute} - \epsilon^{-1}_{solvent})$, with energies expressed in kilocalories per mole and distances in angstroms. ϵ_{solute} and $\epsilon_{solvent}$ are the dielectric constant of the solute and the solvent (i.e., 1 and 80, respectively). α_i and α_j are the Born radii of atoms i and j, respectively, calculated with the GB-MV2 approach. i and j loop over all the atoms of the system (dimer or monomer). The constant K_s is equal to 8 in GB-MV2,[32] unlike the original Still equation, where it is equal to 4. This pairwise expression allows us to define the contribution of atom i to $\Delta G_{elec,solv}$ as

$$\Delta G^i_{elec,solv} = k \frac{q_i^2}{\alpha_i} + \frac{1}{2} k \sum_{j \neq i} \frac{q_i q_j}{\sqrt{r_{ij}^2 + \alpha_i \alpha_j \exp(-r_{ij}^2/K_s \alpha_i \alpha_j)}}. \quad (11)$$

By summing these atomic contributions over the atoms of a given residue, we obtain its contribution to the total binding free energy, ΔG^{res}_{bind}. This contribution can be further broken down into energy terms arising from the internal energy, the van der Waals and electrostatic interactions, and from the nonpolar or electrostatic solvation energies. It is also possible to calculate separately the contribution of the backbone or the side-chain of a given residue. It is important to note that the ΔG^{res}_{bind} provides a decomposition of $\Delta G^0_{bind} + \Delta G_{elec,desolv}$ (i.e., it does not contain the contribution of the entropy terms).

RESULTS AND DISCUSSION

In the following, CA1, CB1, CA2, and CB2 stand for chains A and B of monomer 1 (molecules labeled A and B in 4INS) and monomer 2 (molecules labeled C and D in 4INS), respectively. The detailed analysis of the MD simulation of the insulin dimer with explicit solvent and periodic boundary conditions is available elsewhere.[39] A brief description of the main results is provided in the Appendix.

GB-MV2 Benchmark

The electrostatic term of the solvation free energy contribution to the binding free energy, $\Delta G_{elec,desolv}$, was calculated by solving the Poisson equation (no ionic strength) for each of the 500 frames extracted from the MD simulation performed for the insulin dimer and regularly separated by 10 ps. These results were used as a benchmark to assess the reliability of the values calculated with the fast GB-MV2 model. We found the following correlation between the $\Delta G_{elec,desolv}$ values calculated using GB-MV2 or by solving the Poisson equation using the UHBD program:

$$\Delta G^{Poisson}_{elec,desolv} = -9.56 + 0.985 \times \Delta G^{GB=MV2}_{elec,desolv}, R = 0.9901,$$

where R is the correlation coefficient. We see that the slope of the regression line is close to 1 so that the deviation between the Poisson equation and GB-MV2 model results is nearly constant (see Fig. 4). Therefore, the fast CHARMM/GB-MV2 method appears to be in very good agreement with the UHBD/PB equation results and can be used to rapidly calculate the $\Delta G_{elec,desolv}$ term for each frame of the MD simulation, with a nearly constant difference of about -9 kcal/mol. The latter is not problematic, since we are rather interested in the relative contributions of the residues, backbones, and side-chains than in the total absolute binding free energy. GB-MV2 is about 50 times faster than solving the Poisson equation with the UHBD program for this particular system. In the following, we use the GB-MV2 model to calculate the electrostatic contribution to the solvation free energies.

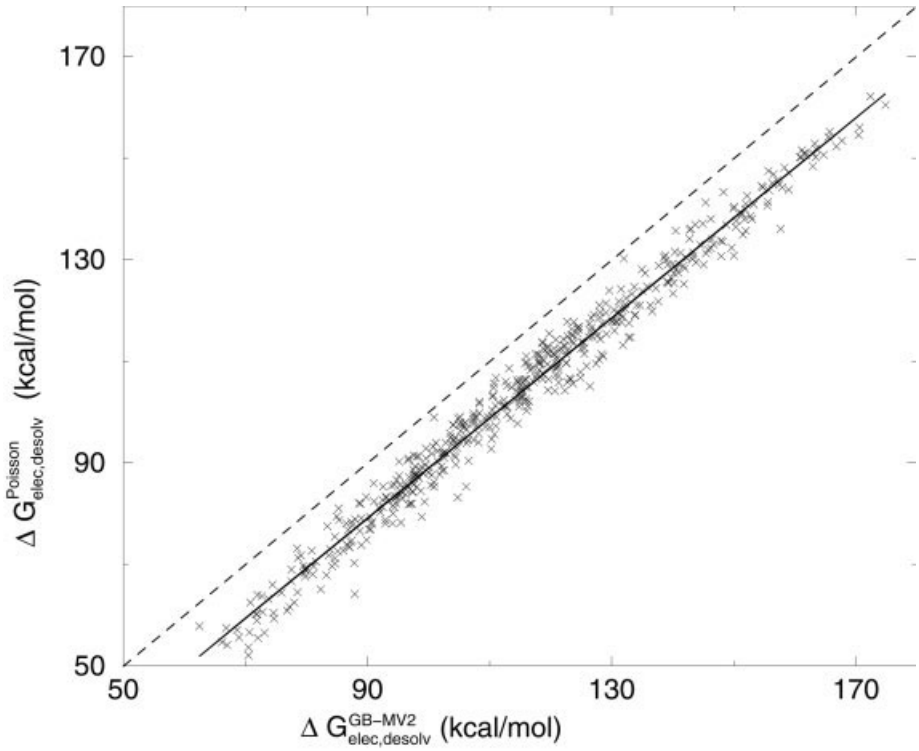

Fig. 4. Correlation between the electrostatic terms of the solvation free energy contribution to the binding free energy calculated by solving the Poisson equation, $\Delta G^{Poisson}_{elec,desolv}$, or by using the GB-MV2 model, $\Delta G^{GB-MV2}_{elec,desolv}$. The regression line (continuous line) and the diagonal (dashed line), corresponding to the ideal correlation $\Delta G^{Poisson}_{elec,desolv} = \Delta G^{GB-MV2}_{elec,desolv}$ are shown.

TABLE I. Binding Free Energy for the Dimerization of Insulin, and Contributions of Solvation, van der Waals, and Electrostatic Interactions and Entropic Terms [Eq. (1) and (8)], Using the Same Trajectory Method

	Dimer	Monomer 1	Monomer 2	$D - (M1 + M2)$
$\langle \Delta E_{intra} \rangle$	378.18 (50.0)	207.04 (41.7)	171.14 (37.8)	0.00 (0.0)
$\langle E_{vdW} \rangle$	−56.97 (3.4)	—	—	−56.97 (3.4)
$\langle E_{elec} \rangle$	−87.70 (26.2)	—	—	−87.70 (26.2)
$\langle \Delta G^{X}_{elec,solv} \rangle$	−1307.55 (43.3)	−737.03 (32.4)	−685.74 (29.8)	115.22 (23.9)
$0.0072 \times \langle SAS^{X} \rangle$	45.56 (0.9)	27.66 (0.6)	27.10 (0.5)	−9.20 (0.4)
$\langle \Delta G^{0}_{bind} \rangle + \langle \Delta G_{desolv} \rangle$				−38.65 (5.8)
$\langle -TS_{vib} \rangle$	−1194.17 (3.6)	−599.27 (3.2)	−597.11 (2.5)	2.21 (3.6)
$\langle -TS_{trans} \rangle$	−14.25[a]	−13.63[a]	−13.63[a]	13.01[a]
$\langle -TS_{rot} \rangle$	−13.55 (<0.01)	−12.53 (<0.01)	−12.54 (<0.01)	11.52 (<0.01)
$\langle \Delta G_{bind} \rangle$				−11.91 (6.7)

Energies are in kcal/mol. Figures in parentheses are standard deviations. X stands for the dimer (D) or each of the 2 monomers ($M1$ or $M2$).

[a]The standard deviation of $-TS_{trans}$ is not defined, since it is only a function of the mass of the system, which is constant.

Binding Free Energy and Influence of Each Residue: The Same Trajectory Method

Table I gives the different contributions to the calculated absolute binding free energy for the insulin dimerization. In the STM, the energy terms relative to the isolated monomers are calculated using coordinates taken from the dimer. As a consequence, the internal energy of the isolated monomers, and of every group of atoms, is identical to that in the complex, and the contribution of the difference in the internal energy upon complexation, ΔE_{intra}, is equal to zero. The calculated value of $\langle \Delta G^{0}_{bind} \rangle + \Delta G_{desolv}$ averaged over 500 frames from the MD simulation is −38.65 kcal/mol. Van der Waals and electrostatic interactions, and electrostatic and nonpolar contributions to the solvation energy contribute −56.98, −87.70, 115.24, and −9.20 kcal/mol, respectively (see Table I). We see that

TABLE II. Contributions From the Side-Chains and the Backbone Atoms in $\langle \Delta G^0_{bind} \rangle + \langle \Delta G_{desolv} \rangle$, from the STM.

	$\langle E_{vdW} \rangle$	$\langle E_{elec} \rangle$	$\langle \Delta G_{elec,desolv} \rangle$	$\langle \Delta G_{np,desolv} \rangle$	$\langle \Delta G^0_{bind} \rangle + \langle \Delta G_{desolv} \rangle$
Side-chains	−34.14	−66.55	82.86	−6.67	−24.50
Backbone	−22.83	−21.15	32.36	−2.53	−14.15

Energies are in kcal/mol.

the favorable binding free energy for the dimerization originates predominantly from the nonpolar terms $\langle E_{vdW} \rangle$ and $\langle \Delta G_{np,desolv} \rangle$. The favorable contribution of the electrostatic interactions between the 2 monomers is canceled by the electrostatic desolvation upon dimerization. These results are in agreement with other MM-PBSA and MM-GBSA studies.[21,22,26,27] According to these results, the polar interactions (like hydrogen bonds) only provide directional constraints for the dimerization (i.e., to the relative positions of the 2 monomers). A statistical analysis of the time series of the energy terms shows that $\langle \Delta G^0_{bind} \rangle + \langle \Delta G_{desolv} \rangle$ was estimated to within a standard deviation of 5.8 kcal/mol. This is very close to the standard deviation calculated for the estimation of the binding free energy of the Ras–Raf system using the MM-GBSA approach (5.4 kcal/mol).[22] The main source for uncertainties are the electrostatic terms, which are most sensitive to the atomic coordinates. $\langle \Delta G^0_{bind} \rangle + \langle \Delta G_{desolv} \rangle$ was decomposed on a per-atom basis, and the contributions of the side-chains and backbone were calculated. As can be seen in Table II, the contribution of the side-chains to the dimerization is twice that of the backbone atoms, underlining their importance.

Not surprisingly, the calculated binding free energy of the 2 monomers using Eq. (8) and neglecting the entropy term (i.e., $\langle \Delta G^0_{bind} \rangle + \langle \Delta G_{desolv} \rangle$) is larger than the experimentally measured binding free energy value of −7.2 kcal/mol.[52] Adding the calculated value of the entropy contribution, $-T\Delta S = +26.74$ kcal/mol (Table I), leads to a calculated absolute value of the binding free energy, ΔG_{bind}, of around −11.9 kcal/mol, in good agreement with the experimental value. This agreement may be partly fortuitous. However, it does support the physical relevance of the model and suggest that the decomposition analysis is meaningful. As in all such absolute binding free energy estimates, the result corresponds to a small difference between large numbers. Relatively small errors in the individual values can lead to large errors in the binding free energy. For example, the electrostatic solvation energy contribution, 115.22 ± 23.9 kcal/mol, results from the difference between the mean electrostatic solvation energy of the dimer, −1307.55 ± 43.3 kcal/mol, and the sum of the mean electrostatic solvation energy of the 2 monomers, −737.03 ± 32.4 kcal/mol and −685.74 ± 29.8 kcal/mol. In this respect, and taking account of the approximations of the model and the dependency upon parameters,[30] the constant difference between GB-MV2 and PB results, and of the 6.7 kcal/mol uncertainty on the calculated ΔG_{bind} value (see Table I), the calculated binding free energy is in satisfactory agreement with the experiment and gives us confidence in our results. Most importantly, the analysis concerns primarily the decomposition of the binding free energy into residue contributions rather than its absolute value.

The need to take account of the entropy contribution to fall into the experimental binding free energy range is well established. For instance, neglecting the entropy term, Wang and Kollman[21] calculated a value of −84.3 kcal/mol for the binding free energy of dimerization for the HIV-1 protease, using a similar equation as Eq. (8). They found that the range of the experimental value (−10 kcal/mol) could be reached theoretically by adding the entropic energy. However, without taking account of the latter, the authors were able to calculate relative dimerization free energies of homo- and heterodimerization consistent with experimental data. Also, using only the gas and solvation terms, Gohlke et al.[22] calculated binding free energies of −48.9 and −51.8 kcal/mol, respectively, for the Ras–Raf and Ras–RalGDS systems. After adding the entropy contribution, they obtained theoretical values of −15.0 and −19.5 kcal/mol, respectively, in better agreement with experimental values (−9.6 and −8.4 kcal/mol, respectively). Again, they found a fair agreement between the calculated side-chains contributions to the binding free energy and experimentally determined binding free energy differences for the corresponding alanine mutants, without taking account of the entropy terms. Therefore, while entropy terms appear to be necessary to reproduce the range of the experimental energies, they are not needed to obtain qualitative and even quantitative results concerning the role of the proteins side-chains into the binding in agreement with experience. In view of this, and since the determination of the atomic contributions to the entropy terms are time-consuming,[22] they will be neglected in the following.

Table III gives the residues (backbone and side-chain) making a significant favorable or unfavorable contribution to the binding free energy of the 2 monomers ($|\Delta G^{res}_{bind}| \geq 0.5$ kcal/mol). Tables IV and V report the most favorable or unfavorable residue side-chains or backbones to the dimerization along with their respective contributions. In all cases, the decomposition into the different energy terms is provided. A negative value for an energy term; i.e., $\langle E^X_{vdW} \rangle$, $\langle E^X_{elec} \rangle$, $\langle \Delta G^X_{elec,desolv} \rangle$, $\langle \Delta G^X_{np,desolv} \rangle$, and $\langle \Delta G^X_{bind} \rangle$, where X stands for the complete residue, *res*, the side-chain, *sc*, or the backbone, *bb*, indicates that it makes a favorable contribution to the dimerization.

As can be seen, the residues making the most favorable contributions to the binding free energy between the 2 monomers are residues Phe24, Phe25, and Tyr26 of CB1 and CB2. Their contribution to the binding free energy ranges from −2.7 to −3.9 kcal/mol. These residues are

TABLE III. Total Binding Free Energy Decomposition.

Residue	$\langle E^{res}_{vdW} \rangle$	$\langle E^{res}_{elec} \rangle$	$\langle \Delta G^{res}_{elec,solv} \rangle$	$\langle \Delta G^{res}_{np,solv} \rangle$	$\langle \Delta G^{res}_{bind} \rangle$
CB2 Phe24	−4.01	−3.49	4.07	−0.49	−3.92
CB1 Tyr26	−3.98	−3.38	4.54	−0.57	−3.40
CB1 Phe25	−3.02	−1.98	2.15	−0.51	−3.36
CB2 Tyr26	−4.00	−4.82	6.23	−0.58	−3.17
CB2 Tyr16	−3.48	−1.20	2.39	−0.68	−2.98
CB2 Phe25	−2.32	−1.16	1.00	−0.32	−2.79
CB1 Phe24	−3.70	−3.56	4.98	−0.40	−2.68
CB1 Tyr16	−3.58	−1.89	3.57	−0.72	−2.62
CB2 Pro28	−1.91	−2.15	1.97	−0.37	−2.47
CB2 Val12	−2.33	−1.59	1.96	−0.38	−2.34
CB1 Val12	−2.33	−1.98	2.42	−0.41	−2.29
CB1 Pro28	−1.91	−0.85	0.94	−0.39	−2.22
CB1 Gly23	−1.17	−1.71	1.57	−0.23	−1.54
CB2 Gly23	−1.16	−0.80	0.78	−0.20	−1.38
CB2 Thr27	−1.21	−1.58	2.05	−0.20	−0.95
CB1 Thr27	−1.01	−1.54	1.84	−0.14	−0.85
CB1 Ser9	−1.28	−1.40	2.42	−0.30	−0.56
CB1 Gly8	−0.72	−2.08	2.42	−0.14	−0.53
...					
CB2 Ala30	−0.08	18.28	−17.62	−0.00	0.57
CB1 Glu13	−0.53	7.67	−6.31	−0.10	0.72
CB2 Glu13	−0.54	7.68	−6.28	−0.09	0.76
CB1 Gly20	−0.48	−1.07	2.45	−0.10	0.80
CB1 Ala30	−0.30	30.35	−28.88	−0.11	1.07
CB2 Glu21	−1.06	9.80	−6.74	−0.31	1.68

The energy terms correspond to the contribution of the entire considered residue (backbone and side-chain). A negative value of ΔG^{res}_{bind} indicates that the corresponding residue makes a favorable contribution to the dimerization. Residues are ranked according to ΔG^{res}_{bind}. Only residues making a significant favorable or unfavorable contribution are given ($|\Delta G^{res}_{bind}| \geq 0.5$ kcal/mol). Energies are in kcal/mol.

TABLE IV. Decomposition of the Side-Chain Contributions to the Binding Free Energy

Residue	$\langle E^{sc}_{vdW} \rangle$	$\langle E^{sc}_{elec} \rangle$	$\langle \Delta G^{sc}_{elec,solv} \rangle$	$\langle \Delta G^{sc}_{np,solv} \rangle$	$\langle \Delta G^{sc}_{bind} \rangle$
CB2 Tyr16	−3.19	−1.01	2.26	−0.68	−2.63
CB2 Phe24	−2.79	−0.30	0.94	−0.30	−2.45
CB1 Tyr16	−3.27	−0.86	2.52	−0.72	−2.34
CB1 Tyr26	−3.09	−0.98	2.25	−0.41	−2.23
CB1 Val12	−1.85	−0.23	0.28	−0.41	−2.22
CB2 Pro28	−1.52	−6.85	6.57	−0.37	−2.17
CB1 Pro28	−1.49	−7.43	7.14	−0.38	−2.16
CB2 Val12	−1.75	−0.14	0.26	−0.37	−2.01
CB1 Phe25	−1.99	−0.13	0.66	−0.47	−1.93
CB1 Phe24	−2.49	−0.12	0.92	−0.23	−1.92
CB2 Tyr26	−2.94	−1.52	3.05	−0.42	−1.84
CB2 Phe25	−1.24	−0.04	0.24	−0.28	−1.32
CB2 Thr27	−0.58	−0.49	0.66	−0.19	−0.60
...					
CB1 Glu13	−0.25	8.94	−7.64	−0.07	0.97
CB2 Glu13	−0.24	8.57	−7.26	−0.07	0.99
CB2 Glu21	−0.39	9.25	−7.66	−0.11	1.10

The energy terms correspond to the contribution of the side chain of the considered residue. Residues are ranked according to ΔG^{sc}_{bind}. Only side chains making a significant favorable or unfavorable contribution are given ($|\Delta G^{sc}_{bind}| \geq 0.5$ kcal/mol). Energies are in kcal/mol.

situated at the interface between the 2 monomers and exchange stable hydrogen bonds between their backbone atoms and van der Waals interactions between their side-chains, as mentioned in the Appendix. They are responsible for the only 1 correlated motion seen between the 2 monomers in the dimer.[39] Table V shows that the electrostatic terms (interaction + desolvation) provided by the backbone atoms of these residues, especially Phe B25,

TABLE V. Decomposition of the Backbone Contributions to the Binding Free Energy

Residue	$\langle E^{bb}_{vdW}\rangle$	$\langle E^{bb}_{elec}\rangle$	$\langle \Delta G^{bb}_{elec,solv}\rangle$	$\langle \Delta G^{bb}_{np,solv}\rangle$	$\langle \Delta G^{bb}_{bind}\rangle$
CB1 Gly23	−1.17	−1.71	1.57	−0.23	−1.54
CB2 Phe25	−1.08	−1.11	0.76	−0.04	−1.47
CB2 Phe24	−1.22	−3.19	3.13	−0.18	−1.46
CB1 Phe25	−1.03	−1.85	1.49	−0.04	−1.44
CB2 Gly23	−1.16	−0.80	0.78	−0.20	−1.38
CB2 Tyr26	−1.06	−3.29	3.17	−0.16	−1.35
CB1 Tyr26	−0.89	−2.40	2.30	−0.16	−1.15
CB1 Phe24	−1.21	−3.44	4.06	−0.16	−0.76
CB1 Thr27	−0.57	−1.51	1.48	−0.01	−0.61
CB1 Gly8	−0.72	−2.08	2.42	−0.14	−0.53
CB1 Ser9	−0.81	−0.71	1.08	−0.08	−0.52
...					
CB1 Glu21	−0.74	−2.27	3.76	−0.23	0.52
CB2 Ala30	−0.06	18.27	−17.64	−0.00	0.56
CB2 Glu21	−0.68	0.55	0.92	−0.20	0.59
CB1 Gly20	−0.48	−1.07	2.45	−0.10	0.80
CB1 Ala30	−0.17	30.40	−29.01	−0.05	1.17

The energy terms correspond to the contribution of the backbone of the considered residue. Residues are ranked according to ΔG^{bb}_{bind}. Only residues backbones making a significant favorable or unfavorable contribution are given ($|\Delta G^{bb}_{bind}| \geq 0.5$ kcal/mol). Energies are in kcal/mol.

are generally contributing favorably to the binding. The favorable contributions made by the electrostatic terms are around −0.4 kcal/mol for the Phe B25 backbone atoms, and around −0.1 kcal/mol for Phe B24 and Tyr B26 (except for Phe 24 of CB1 that makes an unfavorable contribution of around +0.6 kcal/mol). However, the largest favorable contributions made by these residues come from nonpolar terms, $\langle E_{vdW}\rangle$ and $\langle \Delta G_{np,desolv}\rangle$, due to the interactions between their side-chains (Table IV), especially for Phe B24 and Tyr B26. The nonpolar terms arising from the side-chains of these 2 residues contribute from −3.0 to −3.4 kcal/mol each to the total binding free energy. The same contribution is around −1.5 to −2.5 kcal/mol for Phe B25 side-chains. Tyr16, Pro28, and Val12 of CB1 and CB2 are also remarkably favorable to the binding. Their contribution to the binding free energy ranges from −2.2 to −3.0 kcal/mol. Again, their favorable contribution arises essentially from the nonpolar terms due to their side-chains. Actually, upon dimerization, Val12 of the B chain of one monomer is exchanging van der Waals interactions with the Tyr16, Phe24, and Val12 side-chains of chain B of the other monomer, while the Tyr16 side-chain is also making van der Waals contact with the Gly8 residue of the B chain of the facing monomer. Residues Pro B28 are favorable to the dimerization due to exchange of nonpolar interactions by their side-chains with the Gly B23 and Arg B22 backbones of the facing monomer. Also, residues Gly B23 are favorable to the binding, but to a lesser extent, due to favorable van der Waals interactions with the backbone of the Tyr B26 and Thr B27 residues of the other monomer. For the same reason, Thr B27 is found to make small favorable contributions to the dimerization.

On the contrary, residues CB1 Ala30 and CB2 Glu21 are unfavorable to the dimerization because of an electrostatically unfavorable vicinity of the C-terminal carboxylate of residue CB1 Ala30 and the side-chain carboxylate of residue CB2 Glu21. Since the N-terminal carboxylate function of residue CB2 Ala30 is making a ionic interaction with the N-terminus of CA2, the former does not make an unfavorable contribution to the binding.

These theoretical results compare remarkably well with an experimental alanine scanning of a few residues of human insulin situated at the interface between the 2 monomers.[53] The authors mutated Asn A21, Val B12, Tyr B16, Phe B24, Phe B25, Tyr B26, and Thr B27 to alanine. They found that mutating Val B12, Tyr B16, Phe B24, and Tyr B26 to alanine prevent the dimerization. Actually, the side-chains of these residues are found theoretically to make among the most important contributions to the calculated binding free energy, essentially through nonpolar contacts (Table IV). Therefore, on the hypothesis that the mutations will not modify significantly the fold of the hormone, these nonpolar contacts are expected to be lowered if one replaces their bulky side-chains by a small methyl group, explaining the measured effect of such a mutation. Also, the authors found that mutations of Asn A21, Thr B27, and Phe B25 to alanine still allows the dimerization. As we have seen, Thr B27 makes only a small favorable calculated contribution to the binding free energy, whereas Asn A21 makes no favorable contribution, in agreement with the absence of mutation effect on the dimerization. The situation is more complicated for Phe B25. Actually, this residue was found to make an important contribution to the calculated binding free energy for the dimerization, whereas the Ala B25 mutant is found experimentally in the dimeric form. However, as we mentioned earlier, the side-chain of Phe B25 is found to contribute less to the calculated binding free energy than the side-chains of the flanking residues Phe B24 and Tyr B26, or Val B12 and Tyr B16 (Table IV). Moreover, we note

that the 2 chains B are antiparallel. Actually, for example, the Phe24 of CB1 is facing the Tyr26 of CB2, whereas the Phe24 of CB2 is facing the Tyr26 of CB1. Therefore, mutating residue Phe B24 (or residue Tyr B26) cancels the nonpolar interactions between 2 pairs of residues. The same reasoning can be done for the other residues (i.e., Val B12, Pro B28, Gly B23, etc.). On the contrary, the Phe25 of CB1 is facing the Phe25 of CB2, so that mutating Phe B25 mainly cancels the nonpolar interactions between only 1 pair of residues. Therefore, the impact of the Phe B25 mutation may be expected to be smaller than those of the other residues.

CONCLUSION

The binding free energy for the insulin dimerization was calculated by the MM-GBSA method, allowing for a decomposition of the energy at an atomic level. The calculated absolute binding free energy is −11.9 kcal/mol, in rather good agreement with the experimental value of −7.2 kcal/mol. Estimation of the contributions of the van der Waals and electrostatic interactions, as well as the contribution of the desolvation, has shown that the dimerization is mainly due to nonpolar interactions. The favorable contribution of the electrostatic interactions between the 2 monomers, provided, for example, by hydrogen bonds, is canceled by the electrostatic desolvation upon dimerization. Therefore, the role of the hydrogen bonds between the 2 monomers is, most of all, to give the direction of the interactions. Our calculations have shown that residues B24–B26 make the largest favorable contributions to the dimerization of insulin mainly through nonpolar contacts between the side-chains. This is in agreement with the correlated motions found between the 2 monomers for these residues during the MD simulation.[39] Other residues were calculated to make smaller but favorable contributions to the dimerization. They are Tyr B16, Val B12, Pro B28, and also Gly B23. All these residues are situated at the interface between the 2 monomers and contribute mainly through nonpolar contacts made by their side-chains. No residue of the A chain appears to contribute to the dimerization. The energy decomposition on a per-residue basis is in agreement with experimental alanine scanning data.[53] The present results should be helpful in the design of new insulin analogues with modified self-association properties. They also support the use of the GB-MV2 solvent model for the estimation of absolute binding free energies and free energy decomposition on a per-atom basis.

ACKNOWLEDGMENTS

We thank Roland Stote and Annick Dejeagere for helpful discussions. All calculations were performed at the Centre Informatique National de l'Enseignement Supérieur (CINES, Montpellier–France) and the Centre d'Etude du Calcul Parallèle de Strasbourg (Illkirch–France).

REFERENCES

1. Voet D, Voet JG. Biochemistry. New York: Wiley; 1995. Chapter 18-2, p 523–524.
2. Dodson G, Steiner D. The role of assembly in insulin's biosynthesis. Curr Opin Struct Biol 1998;8:189–194.
3. Jorgensen AM, Olsen HB, Balschmidt P, Led JJ. Solution structure of the superactive monomeric des-[Phe(B25)] human insulin mutant: elucidation of the structural basis for the monomerization of des-[Phe(B25)] insulin and the dimerization of native insulin. J Mol Biol 1996;257:684–699.
4. Brems DN, Alter LA, Beckage MJ, Chance RE, DiMarchi RD, Green LK, Long HB, Pekar AH, Shields JE, Frank B. Altering the association properties of insulin by amino acid replacement. Prot Eng 1992;5:527–533.
5. Brange J, Volund A. Insulin analogs with improved pharmacokinetic profiles. Adv Drug Deliv Rev 1999;35:307–335.
6. Madsbad S. Insulin analogues: have they changed insulin treatment and improved glycaemic control? Diabetes Metab Res Rev 2002;18(Suppl 1):S21–S28.
7. DiMarchi RD, Chance RE, Long HB, Shields JE, Slieker LJ. Preparation of an insulin with improved pharmacokinetics relative to human insulin through consideration of structural homology with insulin-like growth factor I. Horm Res 1994;41(Suppl 2):93–96.
8. Brange J, Ribel U, Hansen JF, Dodson G, Hansen MT, Havelund S, Melberg SG, Norris F, Norris K, Snel L, Sørenson AR, Voigt HO. Monomeric insulins obtained by protein engineering and their medical implications. Nature 1988;333:679–682.
9. Brooks CL III, Karplus M, Pettitt BM. Protein: a theoretical perspective of dynamics, structure, and thermodynamics. In: Prigogine I, Rice SA, editors. Advances in chemical physics. Vol. 71. New York: Wiley-Interscience; 1988.
10. Simonson T, Archontis G, Karplus M. Free energy simulations come of age: protein–ligand recognition. Acc Chem Res 2002;35:430–437.
11. Srinivasan J, Cheatham TE III, Cieplak P, Kollman PA, Case DA. Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate-DNA helices. J Am Chem Soc 1998;120:9401–9409.
12. Archontis G, Simonson T, Karplus M. Binding free energies and free energy components from molecular dynamics and Poisson–Boltzmann calculations: application to amino acid recognition by aspartyl-tRNA synthetase. J Mol Biol 2001;306:307–327.
13. Gilson MK, Honig BH. Energetics of charge–charge interactions in proteins. Proteins 1988;3:32–52.
14. Ajay, Murcko MA. Computational methods to predict binding free energy in ligand–receptor complexes. J Med Chem 1995;38:4953–4967.
15. Noskov SY, Lim C. Free energy decomposition of protein–protein interactions. Biophys J 2001;81:737–750.
16. Lazaridis T, Masunov A, Gandolfo F. Contributions to the binding free energy of ligands to avidin and streptavidin. Proteins 2002;47:194–208.
17. Gilson MK, Honig BH. Calculation of the total electrostatic energy of a macromolecular system: solvation energies, binding energies, and conformational analysis. Proteins 1988,4:7–18.
18. Warwicker J, Watson HC. Calculation of the electric potential in the active site cleft due to alpha-helix dipoles. J Mol Biol 1982,157:671–679.
19. Rogers NK. The modeling of electrostatic interactions in the function of globular proteins. Prog Biophys Mol Biol 1986;48:37–66.
20. Hendsch ZS, Tidor B. Electrostatic interactions in the GCN4 leucine zipper: substantial contributions arise from intramolecular interactions enhanced on binding. Protein Sci 1999,8:1381–1392.
21. Wang W, Kollman PA. Free energy calculations on dimer stability of the HIV protease using molecular dynamics and a continuum solvent model. J Mol Biol 2000;303:567–582.
22. Gohlke H, Kiel C, Case DA. Insights into protein–protein binding by binding free energy calculation and free energy decomposition for the Ras–Raf and Ras–RalGDS complexes. J Mol Biol 2003;330:891–913.
23. Kuhn B, Kollman PA. A ligand that is predicted to bind better to avidin than biotin: insights from computational fluorine scanning. J Am Chem Soc 2000;122:3909–3916.
24. Lee TS, Kollman PA. Theoretical studies suggest a new antifolate as a more potent inhibitor of thymidylate synthase. J Am Chem Soc 2000;122:4385–4393.
25. Wang J, Morin P, Wang W, Kollman PA. Use of MM-PBSA in

reproducing the binding free energies to HIV-1 RT of TIBO derivatives and predicting the binding mode to HIV-1 RT of efavirenz by docking and MM-PBSA. J Am Chem Soc 2001;123: 5221–5230.
26. Massova I, Kollman PA. Computational alanine scanning to probe protein–protein interactions: a novel approach to evaluate binding free energies. J Am Chem Soc 1999;121:8133–8143.
27. Huo S, Massova I, Kollman PA. Computational alanine scanning of the 1:1 human growth hormone–receptor complex. J Comput Chem 2002;23:15–27.
28. Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, Lee M, Lee T, Duan Y, Wang W, Donini O, Cieplak P, Srinivasan J, Case DA, Cheatham TE III. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. Acc Chem Res 2000;33:889–897.
29. Massova I, Kollman PA. Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding. Persp Drug Disc Des 2000;18:113–135.
30. Gohlke H, Case DA. Converging free energy estimates: MM-PB(GB)SA studies on the protein–protein complex Ras–Raf. J Comput Chem 2004;25:238–250.
31. Lee MS, Salsbury Jr. FR, Brooks CL III. Novel generalized Born methods. J Chem Phys 2002;116:10606–10614.
32. Lee MS, Feig M, Salsbury Jr. FR, Brooks CL III. New analytic approximation to the standard molecular volume definition and its application to generalized Born calculations. J Comput Chem 2003;24:1348–1356.
33. Bernstein FC, Koetzle TF, Williams GJ, Meyer EE Jr, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M. The Protein Data Bank: a computer-based archival file for macromolecular structures. J Mol Biol 1977;112:535–542.
34. Berman HM, Battistuz T, Bhat TN, Bluhm WF, Bourne PE, Burkhardt K, Feng Z, Gilliland GL, Iype L, Jain S, Fagan P, Marvin J, Padilla D, Ravichandran V, Schneider B, Thanki N, Weissig H, Westbrook JD, Zardecki C. The Protein Data Bank. Acta Crystallogr D Biol Crystallogr 2002;58:899–907.
35. Baker EN, Blundell TL, Cutfield JF, Cutfield SM, Dodson EJ, Dodson GG, Hodgkin DM, Hubbard RE, Isaacs NW, Reynolds CD, et al. The structure of 2Zn pig insulin crystals at 1.5 Å resolution. Philos Trans R Soc Lond B Biol Sci 1988;319:369–456.
36. Brooks BR, Bruccoleri R, Olafson BD, States DJ, Swaminathan S, Karplus M. CHARMM: a program for macromolecular energy, minimization and dynamics calculations. J Comput Chem 1983; 4:187–217.
37. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE, Roux B, Schlenkrich M, Smith, JC, Stote R, Straub J, Watanabe M, Wiorkiewicz-Kuczera J, Yin D, Karplus M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B 1998;102:3586–3616.
38. Badger J, Harris MR, Reynolds CD, Evans AC, Dodson EJ, Dodson GG, North AC. Structure of the pig insulin dimer in the cubic crystal. Acta Crystallogr B 1991;47:127–136.
39. Zoete V, Meuwly M, Karplus M. A comparison of the dynamic behavior of monomeric and dimeric insulin shows structural rearrangements in the active monomer. J Mol Biol 2004;342:913–929.
40. Falconi M, Cambria MT, Cambria A, Desideri A. Structure and stability of the insulin dimer investigated by molecular dynamics simulation. J Biomol Struct Dyn 2001;18:761–772.
41. Yao ZP, Zeng ZH, Li HM, Zhang Y, Feng YM, Wang DC. Structure of an insulin dimer in an orthorhombic crystal: the structure analysis of a human insulin mutant (B9 Ser→Glu). Acta Crystallogr D Biol Crystallogr 1999;55:1524–1532.
42. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. J Chem Phys 1983;79:926–935.
43. McQuarrie DA. Statistical mechanics. New York: Harper & Row; 1976.
44. Tidor B, Karplus M. The contribution of vibrational entropy to molecular association. J Mol Biol 1994;238:405–414.
45. Hermann RB. Theory of hydrophobic bonding: II. Correlation of hydrocarbon solubility in water with solvent cavity surface area. J Phys Chem 1972;76:2754–2759.
46. Amidon GL, Yalkowsky SH, Anik ST, Valvani SC. Solubility of nonelectrolytes in polar solvents: V. Estimation of the solubility of aliphatic monofunctional compounds in water using a molecular surface area approach. J Phys Chem 1975;79:2239–2246.
47. Hasel W, Hendrikson TF, Still WC. A rapid approximation to the solvent accessible surface areas of atoms. Tetrahedron Comput Methodol 1988;1:103–116.
48. Still WC, Tempczyk A, Hawley RC, Hendrickson T. Semianalytical treatment of solvation for molecular mechanics and dynamics. J Am Chem Soc 1990;112:6127–6129.
49. Madura JD, Briggs JM, Wade RC, Davis ME, Luty BA, Ilin A, Antosiewicz J, Gilson MK, Bagheri B, Scott LR, McCammon JA. Electrostatics and diffusion in solution: simulations with the University of Houston Brownian Dynamics program. Comp Phys Comm 1995;91:57–95.
50. Schaefer M. Private communication.
51. Kuhn B, Kollman PA. Binding of a diverse set of ligands to avidin and streptavidin: an accurate quantitative prediction of their relative affinities by a combination of molecular mechanics and continuum solvent models. J Med Chem 2000;43:3786–3791.
52. Strazza S, Hunter R, Walker E, Darnall DW. The thermodynamics of bovine and porcine insulin and proinsulin association determined by concentration difference spectroscopy. Arch Biochem Biophys 1985,238:30–42.
53. Chen H, Shi M, Guo ZY, Tang YH, Qiao ZS, Liang ZH, Feng YM. Four new monomeric insulins obtained by alanine scanning the dimer-forming surface of the insulin molecule. Protein Eng 2000; 13:779–782.
54. Hua QX, Shoelson SE, Kochoyan M, Weiss MA. Receptor binding redefined by a structural switch in a mutant human insulin. Nature 1991;354:238–241.
55. Olsen HB, Ludvigsen S, Kaarsholm NC. Solution structure of an engineered insulin monomer at neutral pH. Biochemistry 1996;35: 8836–8845.
56. Olsen HB, Ludvigsen S, Kaarsholm NC. The relationship between insulin bioactivity and structure in the NH2-terminal A-chain helix. J Mol Biol 1998;284:477–488.
57. Hua QX, Hu SQ, Frank BH, Jia W, Chu YC, Wang SH, Burke GT, Katsoyannis PG, Weiss, MA. Mapping the functional surface of insulin by design: structure and function of a novel A-chain analogue. J Mol Biol 1996;264:390–403.
58. Hua QX, Hu SQ, Jia W, Chu YC, Burke GT, Wang SH, Wang RY, Katsoyannis PG, Weiss MA. Mini-proinsulin and mini-IGF-I: homologous protein sequences encoding non-homologous structures. J Mol Biol 1998;277:103–118.
59. Xu B, Hua QX, Nakagawa SH, Jia W, Chu YC, Katsoyannis PG, Weiss MA. Chiral mutagenesis of insulin's hidden receptor-binding surface: structure of an allo-isoleucine(A2) analogue. J Mol Biol 2002;316:435–441.
60. Xu B, Hua QX, Nakagawa SH, Jia W, Chu YC, Katsoyannis PG, Weiss MA. A cavity-forming mutation in insulin induces segmental unfolding of a surrounding alpha-helix. Protein Sci 2002;11: 104–116.
61. Hua QX, Chu YC, Jia W, Phillips NF, Wang RY, Katsoyannis PG, Weiss MA. Mechanism of insulin chain combination: asymmetric roles of A-chain alpha-helices in disulfide pairing. J Biol Chem 2002;277:43443–43453.
62. Ludvigsen S, Roy M, Thogersen H, Kaarsholm NC. High-resolution structure of an engineered biologically potent insulin monomer, B16 Tyr→His, as determined by nuclear magnetic resonance spectroscopy. Biochemistry 1994;33:7998–8006.
63. Keller D, Clausen R, Josefsen K, Led JJ. Flexibility and bioactivity of insulin: an NMR investigation of the solution structure and folding of an unusually flexible human insulin mutant with increased biological activity. Biochemistry 2001;40: 10732–10740.

APPENDIX

Separate MD Simulation of the Insulin Dimer and Insulin Monomers

The insulin dimer is stable during the 5 ns and stays close to the starting structure. The RMSD from the X-ray structure is approximately 1.4 Å for the backbone and 1.9 Å for all heavy atoms. The difference in the fold between the 2 monomers in the dimer found in the X-ray structure

is maintained throughout the simulation. The RMS fluctuations per residue of the insulin dimer calculated from the MD trajectory are in good agreement with those calculated from the experimental B-factors of the hexamer crystal, with the exception of the N-terminal part of chains B, which showed larger fluctuations during the MD simulation of the dimer. This is not surprising, because in the hexamer the N-terminus of the B chain is involved in contacts between dimers. For the N-terminus, our results agree with NMR solution structures of insulin monomers showing a larger flexibility of the N-terminus of the B chain of the insulin monomer, as compared to insulin hexamer.[54–63] The N-terminal part of the B chains are around 16 Å away from the dimerization interface, and are thus expected to have little influence on the insulin dimerization. Correlated motions between the 2 insulin monomers were found for residues B24–B26. These residues are situated at the interface between the 2 monomers. The correlations are explained by the hydrogen bonds between the residues 24 and 26 of CB1 and CB2, and also by the nonpolar contacts between the side-chains of CB2 Tyr26 and CB1 Phe24, CB1 Phe25 and CB2 Phe25, and CB1 Tyr26 and CB2 Phe24.

Several MD simulations of isolated monomers with explicit water and periodic boundary conditions have been done previously and have revealed a high flexibility of the C-terminal part of the B chain.[39] This flexibility makes difficult a complete sampling of the conformational space of the isolated monomer on the time scale of the MD simulations. Starting from the X-ray conformation of the monomer in the structure of the complex, it has been observed that the high flexibility of the C-terminal part of the B chain is reached before 3 ns, both for monomer 1 and monomer 2. The ionic interaction between the N-terminus of the A chain and the C-terminus of the B chain that exists in the complex in monomer 2, but not in monomer 1, and which is responsible for the different fold of the 2 monomers in the dimer, is broken. After that (3 ns), the 2 monomers show similar dynamical behavior. Since the 2 components of the complex are actually identical monomers, only 2 MD simulations are necessary: 1 for the complex and 1 for an isolated monomer. The MD simulation of the latter was performed for 10 ns. Only the last 5 ns were used to calculate the energy terms of Eq. (8), to get rid of the conformational transition between the low and high flexibility regimes of the C-terminal part of the B chain, as described above.

The MD simulation for the monomeric insulin was performed using a similar procedure as for the dimer, except that the system contained 1977 water molecules and a total of 6714 atoms, in a 46.6 × 43.5 × 34.1 Å^3 cell. The MD simulation was started from the coordinates of monomer 1 in the crystal structure of the dimer. The solvent was equilibrated during 30 ps in the presence of the fixed insulin, and after heating to 300 K, the entire system was equilibrated for 80 ps. The MD simulation was conducted for 10 ns.

The decomposition of the binding free energy at the atomic level is similar to that described for that using a single trajectory (i.e., the single trajectory or STM). However, with separate trajectories for the double trajectory DTM, a per-atom decomposition of the internal energy terms is required. The nonbonded internal energy terms of each monomer, $E^X_{intra,vdW}$ and $E^X_{intra,elec}$, is performed in the same way as for E_{vdW} and E_{elec}. The $E^X_{intra,bond}$ energy term was decomposed into the contribution of the side-chain and the contribution of the backbone of each residue. $E^X_{intra,bond}$ contains the energies of the bonds, angles, dihedral angles and improper angles of the system. For a given subset of atoms, like a side-chain, for example, the energy arising only from atoms of the subset is entirely attributed to it. However, only half of the energy arising from an internal coordinate involving 2 atoms of the subset and atoms out of it is attributed to the subset. In CHARMM, the energy of an angle is allocated to the central atom, the energy of a dihedral is defined between the central 2 atoms, and the energy of an improper angle is defined by the first atom. This decomposition prevents double counting of interactions.

Results and Discussion of Data Obtained With Two Trajectories

In principle, the use of separate trajectories for the monomer and dimer can take account of conformational changes that take place on binding. However, it does not have the benefit of the cancellations of terms that occurs when the same trajectory is used, so that the results are

TABLE VI. Same as Table I, But Using the Different Trajectory Method

	Dimer	Monomer	$D-(M+M)$
$\langle \Delta E_{intra} \rangle$	378.18 (50.0)	207.06 (40.3)	−35.94
$\langle E_{vdW} \rangle$	−56.97 (3.4)	—	−56.97
$\langle E_{elec} \rangle$	−87.70 (26.2)	—	−87.70
$\langle \Delta G^X_{elec,solv} \rangle$	−1307.55 (43.3)	−737.25 (29.1)	166.95
$0.0072 \times \langle SAS^X \rangle$	45.56 (0.9)	28.84 (0.8)	−12.12
$\langle \Delta G^0_{bind} \rangle + \langle \Delta G_{desolv} \rangle$			−25.78
$\langle -TS_{vib} \rangle$	−1194.17 (3.6)	−597.42 (2.8)	0.67
$\langle -TS_{trans} \rangle$	−14.25[a]	−13.63[a]	13.01
$\langle -TS_{rot} \rangle$	−13.55 (<0.01)	−12.53 (<0.01)	11.51
$\langle \Delta G_{bind} \rangle$			−0.59

[a]The standard deviation of $-TS_{trans}$ is not defined, since it is only a function of the mass of the system, which is constant.

TABLE VII. Same as Table II, But Using the Different Trajectory Method

	$\langle \Delta E_{intra} \rangle$	$\langle E_{vdW} \rangle$	$\langle E_{elec} \rangle$	$\langle \Delta G_{elec,desolv} \rangle$	$\langle \Delta G_{np,desolv} \rangle$	$\langle \Delta G^0_{bind} \rangle + \langle \Delta G_{desolv} \rangle$
Side-chains	24.22	−34.14	−66.55	62.70	−8.31	−22.08
Backbone	−60.16	−22.83	−21.15	104.25	−3.81	−3.70

TABLE VIII. Same as Table III, But Using the Different Trajectory Method

Residue	$\langle \Delta E^{res}_{intra} \rangle$	$\langle \Delta E^{res}_{vdW} \rangle$	$\langle \Delta E^{res}_{elec} \rangle$	$\langle \Delta G^{res}_{elec,solv} \rangle$	$\langle \Delta G^{res}_{np,solv} \rangle$	$\langle \Delta G^{res}_{bind} \rangle$
CB1 Tyr26	−2.96	−3.98	−3.38	5.20	−0.50	−5.63
CB2 Phe24	−1.76	−4.01	−3.49	4.78	−0.23	−4.71
CB2 Tyr26	−2.99	−4.00	−4.82	7.90	−0.52	−4.44
CB2 Pro28	0.25	−1.91	−2.15	1.11	−0.47	−3.18
CB2 Tyr16	1.39	−3.48	−1.20	0.74	−0.57	−3.13
CB1 Pro28	0.47	−1.91	−0.85	−0.29	−0.41	−3.00
CB1 Phe24	−0.33	−3.70	−3.56	5.18	−0.23	−2.65
CB2 Val12	−0.05	−2.33	−1.59	1.59	−0.16	−2.54
CB1 Val12	−0.05	−2.33	−1.98	2.37	−0.17	−2.15
CB1 Tyr16	1.83	−3.58	−1.89	2.07	−0.57	−2.14
CB1 Gly23	−0.54	−1.17	−1.71	1.96	−0.27	−1.73
CB2 Leu11	−1.21	−0.23	−1.16	0.94	−0.00	−1.67
CB1 Lys29	−19.92	−0.22	−17.61	36.64	−0.39	−1.50
CB1 Ser9	−2.83	−1.28	−1.40	4.35	−0.31	−1.47
CA1 Gln5	−2.47	−0.01	0.03	0.97	0.03	−1.45
CB1 Leu11	−1.56	−0.20	−1.42	1.75	−0.00	−1.43
CA1 Glu4	−0.53	−0.02	16.11	−16.63	−0.30	−1.37
CB2 Arg22	−6.53	−0.85	−8.32	14.41	0.05	−1.25
CB2 Leu15	−1.13	−0.33	−0.64	0.88	−0.00	−1.22
CB1 Phe1	−2.02	−0.01	−3.68	4.85	−0.34	−1.19
CB2 Phe25	−0.19	−2.32	−1.16	3.27	−0.75	−1.14
CB2 Gly23	−0.80	−1.16	−0.80	1.93	−0.29	−1.12
CB1 Gly8	−1.03	−0.72	−2.08	3.07	−0.18	−0.95
...						
CA2 Glu17	2.63	−0.02	7.75	−9.38	−0.05	0.93
CA2 Gln15	0.58	−0.01	0.07	0.50	−0.03	1.11
CB1 Ala30	−2.78	−0.30	30.35	−25.94	−0.14	1.20
CB1 Glu13	7.13	−0.53	7.67	−12.48	−0.07	1.71
CA2 Gly1	−48.26	−0.02	−17.95	68.38	−0.36	1.79
CB2 Glu21	4.67	−1.06	9.80	−10.84	−0.25	2.32
CB2 Thr27	0.42	−1.21	−1.58	5.62	−0.61	2.63
CB2 Glu13	11.08	−0.54	7.68	−15.43	−0.13	2.65
CA1 Ile2	2.73	−0.03	−0.04	−0.11	0.21	2.76
CB2 Ala30	−19.96	−0.08	18.28	6.03	−0.48	3.78
CA2 Glu4	17.47	−0.03	17.79	−30.87	−0.12	4.24

sensitive to the detailed structures. The contribution of this rearrangement on the binding free energy is included in the difference in the internal energy of the components between their isolated and complexed forms. Therefore, in the DTM, it is very important to include the internal energy term. For instance, in the present case, an ionic interaction internal to monomer 1 between the side-chains of CB1 Lys29 and CA1 Glu4 exists in the complex but not in the isolated monomer. This results in a larger electrostatic desolvation energy of the side-chains when calculated from the DTM compared to the STM, but also in a large decrease of the internal electrostatic energy of monomer 1 upon complexation, which is included in $\langle \Delta E_{intra} \rangle$.

The DTM introduces additional errors compared to the STM for several reasons. First, it is necessary to include the contribution of $\langle \Delta E_{intra} \rangle$, which is an additional source for errors. Second, in the STM approach, the energies relative to 1 given term are calculated from the frames of the same trajectory for the complex and its isolated components. As a consequence, their variations are correlated; thus, the standard deviation on the averaged difference, D − (M1 + M2), is lower than the standard deviation on each of the energies. For instance, in the STM approach (Table I), the standard deviations on $\langle \Delta G^D_{elec,solv} \rangle$, $\langle \Delta G^{M1}_{elec,solv} \rangle$, and $\langle \Delta G^{M2}_{elec,solv} \rangle$ are 43.32, 32.4, and 29.8 kcal/mol, respectively, but the standard deviation on $\langle \Delta G_{elec,desolv} \rangle = \langle \Delta G^D_{elec,solv} - (\Delta G^{M1}_{elec,solv} + \Delta G^{M2}_{elec,solv}) \rangle$ is only 23.9 kcal/mol. This is due to the fact that the variations of $\Delta G^D_{elec,solv}$ and $(\Delta G^{M1}_{elec,solv} + \Delta G^{M2}_{elec,solv})$ are correlated; the correlation coefficient between the 2 sets of values is 0.84. This cancellation of errors does not exist in the DTM approach, since $\Delta G^D_{elec,solv}$ and $(\Delta G^{M1}_{elec,solv} + \Delta G^{M2}_{elec,solv})$ are calculated from different trajectories. Also, in our case, the high flexibility of the C-terminal part of the

TABLE IX. Same as Table IV, But Using the Different Trajectory Method

Residue	$\langle \Delta E^{res}_{intra} \rangle$	$\langle \Delta E^{res}_{vdW} \rangle$	$\langle \Delta E^{res}_{elec} \rangle$	$\langle \Delta G^{res}_{elec,solv} \rangle$	$\langle \Delta G^{res}_{np,solv} \rangle$	$\langle \Delta G^{res}_{bind} \rangle$
CB1 Tyr26	−1.25	−3.09	−0.98	2.60	−0.39	−3.12
CB2 Tyr16	1.67	−3.19	−1.01	0.31	−0.52	−2.75
CB2 Tyr26	−0.59	−2.94	−1.52	2.83	−0.41	−2.64
CB2 Pro28	4.14	−1.52	−6.85	2.10	−0.40	−2.53
CB1 Tyr16	1.62	−3.27	−0.86	0.71	−0.52	−2.33
CB1 Pro28	1.31	−1.49	−7.43	5.75	−0.38	−2.24
CB2 Phe25	−1.30	−1.24	−0.04	1.22	−0.66	−2.02
CB2 Val12	0.12	−1.75	−0.14	0.18	−0.16	−1.76
CB1 Phe1	−1.29	−0.01	−0.01	−0.01	−0.34	−1.66
CB1 Lys29	−18.83	−0.04	−19.27	37.04	−0.36	−1.46
CB1 Val12	0.60	−1.85	−0.23	0.20	−0.17	−1.46
CB1 Phe25	0.52	−1.99	−0.13	0.83	−0.45	−1.22
CB2 Leu11	−0.86	−0.12	−0.14	0.18	−0.00	−0.93
CB2 Phe24	1.76	−2.79	−0.30	0.51	−0.08	−0.90
CB2 Thr27	−3.84	−0.58	−0.49	4.54	−0.52	−0.89
CB1 Phe24	0.98	−2.49	−0.12	0.84	−0.08	−0.87
CB1 Leu11	−0.89	−0.09	−0.14	0.34	−0.00	−0.78
CA1 Val3	−0.50	−0.07	−0.17	0.26	−0.28	−0.75
...						
CB2 Hsp10	4.07	−0.05	−12.70	9.58	−0.02	0.88
CA1 Tyr19	0.41	−0.13	0.18	0.85	−0.08	1.23
CA1 Ile2	1.46	−0.01	0.06	0.12	0.20	1.82
CB1 Glu13	7.88	−0.25	8.94	−14.57	−0.05	1.95
CB2 Glu13	11.40	−0.24	8.57	−17.23	−0.11	2.39
CA2 Glu4	19.17	−0.01	18.20	−32.56	−0.12	4.67

B chain of the isolated monomer makes it difficult to sample completely its conformational space in the time of the MD simulation.

Table VI gives the different contributions to the calculated absolute binding free energy for the insulin dimerization according to the DTM. Table VII decomposes this in terms of backbone and side-chain contributions. Since the terms relative to the monomers and the dimers are not calculated from the same trajectory, the standard deviations on the averaged differences, D − (M + M), are not defined. However, given the above considerations, we can expect the errors on $\langle \Delta G^0_{bind} \rangle + \langle \Delta G_{desolv} \rangle$ and $\langle \Delta G_{bind} \rangle$ to be larger than in the STM. The absolute binding free energy calculated using the DTM is −0.6 kcal/mol, significantly less favorable than the −11.9 kcal/mol value calculated using the STM. The difference is mainly due to a less favorable contribution from the backbone atoms (i.e., −3.7 kcal/mol with the DTM instead of the −14.2 kcal/mol according to the STM). This is explained by the larger desolvation energy of the backbone atoms according to the DTM, which is not entirely compensated by the −60.2 kcal/mol change in the internal energy of the backbone upon complexation. The large favorable contribution of the change in the internal energy upon complexation is due most of all to the formation of ionic interactions within the monomers upon complexation. Tables VIII, IX, and X give the contribution of the residues, and each side-chain or backbone of the residues, on $\langle \Delta G^0_{bind} \rangle + \langle \Delta G_{desolv} \rangle$. We can see that the most important residues for the dimerization of insulin according to the DTM are Val B12, Tyr B16, Phe B24, Tyr B26, and Pro B28, and to a less extent Gly B8, Ser B9, and Gly B23, in very good agreement with the STM. However, we see that, according to the DTM, the Phe B25 makes a less favorable contribution to the dimerization than according to the STM, while Thr B27 no longer makes a favorable contribution. On the contrary, several other residues were found to make a small favorable contribution to the dimerization of insulin. They are CB1 and CB2 Leu11, CB1 Lys29, CA1 Glu4 and Gln5, CB2 Leu15 and Arg22, and CB1 Phe1. Explanations for these changes are presented in the Supplementary Material.

Conclusion

The detailed comparison between the STM and DTM underlines the importance of taking account of the internal energy change in the latter. The two approaches provide similar results concerning the relative importance of the residues for the insulin dimerization. The most important residues according to the two methods are the same, with the exception of Phe B25, whose importance is decreased according to the different trajectory method. Some other differences were found concerning less important residues. In some cases, they are explained by a modification of the insulin conformation upon binding, while in other cases it is difficult to conclude whether the change is meaningful or if it is a consequence of the larger noise in the different trajectory method.

TABLE X. Same as Table V, But Using the Different Trajectory Method

Residue	$\langle \Delta E^{res}_{intra} \rangle$	$\langle \Delta E^{res}_{vdW} \rangle$	$\langle \Delta E^{res}_{elec} \rangle$	$\langle \Delta G^{res}_{elec,solv} \rangle$	$\langle \Delta G^{res}_{np,solv} \rangle$	$\langle \Delta G^{res}_{bind} \rangle$
CB2 Phe24	−3.52	−1.22	−3.19	4.27	−0.15	−3.81
CB1 Tyr26	−1.71	−0.89	−2.40	2.60	−0.11	−2.51
CB2 Tyr26	−2.40	−1.06	−3.29	5.06	−0.11	−1.81
CB1 Phe24	−1.31	−1.21	−3.44	4.34	−0.15	−1.78
CB1 Gly23	−0.54	−1.17	−1.71	1.96	−0.27	−1.73
CA1 Glu4	0.34	−0.01	−0.12	−1.38	−0.02	−1.18
CB2 Gly23	−0.80	−1.16	−0.80	1.93	−0.29	−1.12
CB1 Ser9	−0.61	−0.81	−0.71	1.14	−0.06	−1.05
CA1 Gln5	−0.82	−0.01	−0.04	−0.12	−0.01	−0.99
CB1 Gly8	−1.03	−0.72	−2.08	3.07	−0.18	−0.95
CB2 Leu15	−0.99	−0.14	−0.53	0.70	0.00	−0.95
CB2 Arg22	−0.29	−0.72	0.33	−0.30	0.04	−0.94
CA2 Tyr19	−3.31	−0.07	0.14	2.49	−0.06	−0.80
CB1 Hsd5	−0.53	−0.01	−0.14	−0.09	−0.02	−0.79
CB2 Val12	−0.17	−0.57	−1.45	1.41	0.00	−0.78
CB1 Pro28	−0.84	−0.42	6.58	−6.05	−0.03	−0.77
CB2 Gly8	−0.43	−0.72	−1.55	2.11	−0.17	−0.76
CB2 Leu11	−0.35	−0.12	−1.02	0.76	0.00	−0.73
...						
CA1 Gly1	22.46	−0.01	−13.75	−7.99	0.13	0.84
CA2 Glu17	1.44	−0.01	0.13	−0.72	0.01	0.85
CB2 Phe25	1.11	−1.08	−1.11	2.05	−0.09	0.88
CB1 Gly20	0.73	−0.48	−1.07	1.82	−0.10	0.90
CA1 Ile2	1.27	−0.01	−0.10	−0.25	0.01	0.92
CA2 Gln15	0.83	−0.01	0.19	0.08	−0.02	1.08
CB2 Lys29	0.72	−0.13	−1.05	1.81	−0.11	1.23
CB1 Ala30	−2.73	−0.17	30.40	−26.10	−0.05	1.35
CB1 Glu21	2.12	−0.74	−2.27	2.51	−0.24	1.38
CA1 Val3	2.51	−0.02	−0.33	−0.71	0.00	1.46
CA2 Gly1	−48.26	−0.02	−17.95	68.38	−0.36	1.79
CB2 Glu21	3.31	−0.68	0.55	−0.90	−0.22	2.06
CB2 Thr27	4.26	−0.64	−1.09	1.08	−0.09	3.52
CB2 Ala30	−19.73	−0.06	18.27	5.85	−0.39	3.93