PROTEINS: Structure, Function, and Bioinformatics 67:418–434 (2007)

Protein–Protein Recognition and Interaction Hot Spots in an Antigen–Antibody Complex: Free Energy Decomposition Identifies "Efficient Amino Acids"

Virginie Lafont,[1,2] Michael Schaefer,[3] Roland H. Stote,[4] Danièle Altschuh,[2] and Annick Dejaegere[1]*

[1]*Structural Biology and Genomics Department, UMR 7104, Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS/INSERM/ULP, F-67404 Illkirch Cedex, France*

[2]*UMR7175-LC1 CNRS/ULP, ESBS, Parc d'Innovation, 67412 Illkirch Cedex, France*

[3]*Laboratoire de Chimie Biophysique, Institut le Bel, Université Louis Pasteur, 67000 Strasbourg, France*

[4]*Laboratoire de Biophysicochime Moléculaire, Institut de Chimie, LC3-UMR7177, Université Louis Pasteur, 67070 Strasbourg Cedex, France*

ABSTRACT **The molecular mechanics Poisson–Boltzmann surface area (MM/PBSA) method was applied to the study of the protein–protein complex between a camelid single chain variable domain (cAb-Lys3) and hen egg white lysozyme (HEL), and between cAb-Lys3 and turkey egg white lysozyme (TEL). The electrostatic energy was estimated by solving the linear Poisson–Boltzmann equation. A free energy decomposition scheme was developed to determine binding energy hot spots of each complex. The calculations identified amino acids of the antibody that make important contributions to the interaction with lysozyme. They further showed the influence of small structural variations on the energetics of binding and they showed that the antibody amino acids that make up the hot spots are organized in such a way as to mimic the lysozyme substrate. Through further analysis of the results, we define the concept of "efficient amino acids," which can provide an assessment of the binding potential of a particular hot spot interaction. This information, in turn, can be useful in the rational design of small molecules that mimic the antibody. The implications of using free energy decomposition to identify regions of a protein–protein complex that could be targeted by small molecules inhibitors are discussed. Proteins 2007;67:418–434.** © 2007 Wiley-Liss, Inc.

Key words: MM/PBSA; molecular dynamics; macromolecular electrostatics; protein engineering; drug design; antibodies; camelids

INTRODUCTION

For a wide range of biological processes, including antibody–antigen interactions, hormone–receptor interactions, signal transduction, and regulation, the formation of stable protein–protein complexes is of fundamental importance. Interfering with or modulating the formation of these complexes therefore offers attractive opportunities for therapeutic intervention. However, the development of small molecules that modulate protein–protein interactions is difficult.[1] There are several examples of natural molecules that bind at protein interfaces[2] and a growing number of synthetic compounds that are aimed at perturbing protein–protein interactions (for a recent review, see Ref. 3). Progress in developing synthetic compounds will benefit greatly from improved understanding of the energetics and dynamics of protein–protein recognition. Our current understanding of protein–protein interactions is based on structural analysis of protein–protein complexes mostly determined by crystallography, on various biophysical characterizations of protein–protein complexes that give insight into the kinetics and thermodynamics of protein–protein association,[4] and on biochemical analyses that help identify amino acids important for recognition.

An important concept that has emerged in the last years is the notion of interaction energy "hot spots," which is the idea that a handful of amino acids at the binding interface make a dominant contribution to the binding affinity.[5,6] Mutation studies, mostly alanine scanning studies, have indeed shown[7] that only a few mutations [native → Ala] affect the free energy of binding by more than 2 kcal/mol. This has led to a distinction between the functional epitope (i.e. "hot spot residues) and the structural epitope (all residues that participate in the interface).[8]

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

Grant sponsor: French Ministry of Defense, Centre National de la Recherche Scientifique, Institut National de la Santé et de la Recherche Médicale, and Université Louis Pasteur, Strasbourg, France.

Virginie Lafont's current address is Department of Biology, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218-2685.

Michael Schafer's current address is Novartis Pharma AG, Klybeckstrasse 141, 4057 Basel, Switzerland.

*Correspondence to: Annick Dejaegere. Structural Biology and Genomics Department, UMR 7104, Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS/INSERM/ULP, BP 10142, F-67404 Illkirch Cedex, France. E-mail: annick@igbmc.u-strasbg.fr

Received 20 May 2006; Revised 6 September 2006; Accepted 15 September 2006

Published online 26 January 2007 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.21259

However, much remains to be understood of the role of residues surrounding the hot spots in creating a favorable binding environment, on the balance between electrostatic, van der Waals, and hydrophobic interactions in the formation of complexes, in entropic contributions to binding, and in the role of molecular flexibility on the binding process.

Computational and theoretical methods permit the systematic analysis of individual amino acid contributions to the formation of protein complexes.[9–13] They provide a wealth of information that are complementary to experiments and allow the determination of the structural and energetic consequences of mutations, as well as a better understanding of the physical mechanism and origin of the intermolecular interactions that drive binding. Free energy simulations methods[14–16] have been successfully used to predict the effect of point mutations on binding affinity in macromolecular complexes (for a recent review, see Ref. 17). Computational cost, however, puts practical limitations on the number of point mutations that can be analyzed by free energy simulations for a given complex. Systematic analysis of the contribution of all amino acids to the stability of macromolecular complexes have mostly relied on different empirical free energy functions[11,12] as well as on all-atom molecular mechanics methods using implicit representation of the solvent (so called MM/PBSA or MM/GBSA methods[18]). The MM/PBSA method has been used to systematically analyze the contribution of all amino acids to the formation of several macromolecular complexes. For computational efficiency, these analyses were performed on a single structure of the complex.[10,19] The generalized Born model for electrostatics has also recently been applied to the free energy decomposition in protein–protein complexes.[13,20,21] The generalized Born model is faster, but introduces additional approximations in the electrostatics treatment with respect to numerical solutions of the Poisson–Boltzmann equations. In the present work, we develop a computationally tractable, yet reliable, free energy decomposition protocol that can be used to identify amino-acids that have a dominant role in binding. In the MM/PBSA approach presented here, we use solutions to the Poisson–Boltzmann equation to obtain the electrostatic contribution to binding. This is in contrast to other recent applications of the methodology[13,20,21] where the Generalized Born model was employed. In addition, multiple conformations from molecular dynamics trajectories are used in the analysis.

The protocol is used to analyze molecular recognition in the complex between a camel single chain antibody variable domain (cAb-Lys3) and its antigen lysozyme.[22,23] The recombinant protein cAb-Lys3 is a single variable domain. Single domain antigen binding fragments can be generated from the variable domain of *Camelidae* heavy chain antibodies. Indeed, the *Camelidae* possess, in addition to classical IgG molecules, a large fraction of naturally occurring heavy chain antibodies that are devoid of light chains.[24] These single domains preserve the ability to bind specifically and with high affinity to their antigens. These small antibodies, also called VHH or nanobody™ are minimally sized, stable, and highly soluble, and therefore offer a number of benefits in medical and biotechnological applications (for a recent review, see Ref. 25). High affinity VHH can be obtained after immunization of llamas and dromedaries with specific antigens.[23,26] Alternatively, phage antibody libraries can be used.[27,28] VHHs that inhibit several enzymes[22,23,26,28–31] have been identified, as well as VHH that target antigens specific of human tumors[32,33] or specific of infectious agents.[34] VHHs against a mouse anti-DNA monoclonal antibody that are competitive binders of the original DNA antigen have also been recently identified and are shown to be useful for molecular mimicry studies.[35] VHHs against proteins of interest in human disease have potential diagnostic[33] or therapeutic[32] applications. *Camelidae* VHHs are not only able to recognize a wide range of target proteins, their high stability and small size makes them amenable to structural studies, and an increasing number of VHH–antigen structures are available (for recent examples see Refs. 36–38). The structural information from these complexes is potentially useful for the development of small molecule mimicks of the VHH.[39] For this to be realized, however, it is important to identify the important interactions in the VHH–antigen complex. From an experimental analysis alone, this is not necessarily evident. In this paper, we show, for the VHH cAb-Lys3–lysozyme complex, that an interaction energy hot spot exploited in the antibody–lysozyme complex is also used by the *N*-acetyl glucosamine (NAG) ring of saccharide substrates in subsite C of the lysozyme binding pocket. We support the results of the free energy decomposition approach by the mutational analysis of four VHH positions. We further introduce the concept of amino acid efficiency and show the potential of free energy decomposition in extracting useful information for drug design.

MATERIALS AND METHODS

We developed a protocol based on the MM/PBSA method,[13,18,40] in which conformations extracted from molecular dynamics simulations are processed by a combination of methods using a simplified description for the solvent to give an estimate of binding free energy. Individual contributions of each amino acid to the complex formation are estimated and important amino acid "hot spots" are identified.

Structures

The coordinates of the camel single-domain VH antibody–lysozyme complexes were obtained from the Protein Data Bank.[41] Two structure files containing hen egg white lysozyme as the antigen were obtained: 1MEL (solved at 2.5 Å resolution)[22] and 1JTT (solved at 2.1 Å resolution).[42] A third structure file containing the coordinates for the complex of the same antibody with turkey egg white lysozyme as the antigen was also obtained from the Protein Data Bank: 1JTP (solved at 1.9 Å resolution).[42] For the structures 1MEL and 1JTP, there are two complexes per asymmetric unit, antibody chain A with lysozyme chain L

and antibody chain B with lysozyme chain M. As there are significant structural differences between the two complexes,[42] calculations were carried out for each complex (AL and BM).

The hydrogen atoms were added to the crystal structure using the HBUILD module[43] of the CHARMM[44] program (version 29a1); the all atom parameter set of CHARMM22[45] was used for all calculations. Each structure was subjected to an energy minimization of 5000 steps using the steepest descent algorithm and 5000 steps using the adopted basis Newton Raphson algorithm. Harmonic constraints, added initially to prevent any large perturbation of structure, were gradually reduced over the course of the minimization. The final steps of the minimization were done without constraints. The minimization was stopped when the gradient of the energy was less than 10^{-4} kcal mol^{-1} Å^{-1}. A distance dependent dielectric of $4r$, where r is the distance between two atoms, and a nonbonded cutoff of 13 Å were used during minimization.

Determination of Protonation State

The pKa of titratable groups were determined for both the isolated proteins and for the complexes using continuum electrostatic as described by Schaefer et al.[46] The method assumes that electrostatic effects dominate the influence of the surroundings on the pKa of titratable groups. The change in pKa of titratable residues between the solution value (where experimental values for amino acid side chains are used) and the value in the protein interior is then calculated using continuum electrostatics. We used the UHBD program[47] to calculate the free energy of different states of protonation and a series of c-shell scripts[46] to obtain the titration curves. The protonation states of selected residues were determined at physiological pH of 7.4. Residues located at the interface of the complex are those which are the most susceptible to undergo shifts in pKa when the complex is formed. Therefore we determined the protonation state of acidic and basic residues that are buried upon complexation, i.e. residues that have a solvent accessible surface area difference between the complex and the free molecules larger than 10 Å^2. The protonation states of these residues were determined in the complex and in the two separated proteins. The protonation states of all other titratable groups were put to their standard values. A dielectric constant of 80 was used for the solvent. Two values of the internal dielectric constant for the protein were tested (4 and 20) and both gave similar results in terms of whether particular groups were protonated or not at a pH of 7.4.

Molecular Dynamic Simulations

Molecular dynamic simulations were done using the CHARMM[44] program and the all atom parameter set of CHARMM22.[45] The minimized protein–protein complexes were solvated in a $93 \times 63 \times 63$ Å box of TIP3P water. Molecular dynamic simulations were performed at constant temperature using the Nose–Hoover method[48,49] with periodic boundary conditions. Bond lengths involving bonds between heavy atoms and hydrogen atoms were constrained using SHAKE.[50] With the protein–protein complex fixed, the water was equilibrated for 6 ps at 300 K. The constraints were removed and the entire system was again equilibrated for 6 ps. A time step of 1 fs was used in the simulation and structures were collected every 10 fs. Van der Waals interactions were truncated at a cutoff distance of 12.5 Å using a switch function. Electrostatic interactions were truncated at 12.5 Å using a shift function. All simulations contain between 35,500 and 35,700 atoms, which includes about 3800 atoms for the protein complex. Six trajectories ranging from 180 to 300 ps in length were generated using different starting structures and initial conditions. The different simulations performed are listed in Table I. The purpose of the simulations was to generate multiple structures to improve the statistical sampling of the binding energy calculations.

TABLE I. MD Simulations Done for Conformational Sampling

Complex	PDB ID	Chains	HIS 111	Trajectory {ps}
cAb-Lys3/HEL	MEL	AL	HSD	180
cAb-Lys3/HEL	MEL	AL	HSE	180
cAb-Lys3/HEL	MEL	BM	HSE	240
cAb-Lys3/HEL	JTT	AL	HSE	250
cAb-Lys3/TEL	JTP	AL	HSD	180
cAb-Lys3/TEL	JTP	BM	HSE	300

Details for four simulations of the cAb-Lys3/Hen egg-white lysozyme complex, two simulations of the cAb-Lys3/Turkey egg-white lysozyme complex. For each PDB crystallographic structure, both complexes of chains A+L and of chains B+M were used. The protonation state for the residue His 111 was determined, see text (HSD/HSE: protonation on Nδ/Nε, respectively). For the complex of chains A+L in PDB entry MEL (cAb-Lys3/HEL), trajectories were run for each of the two protonation states of His 111 (HSE and HSD).

Free Energy Calculations and Free Energy Decomposition

Free energy was estimated according to a MM/PBSA approach,[18] in which the free energy is calculated according to the following equation:

$$G_{\text{molecule}} = E_{\text{MM}} + G_{\text{PBSA}} - TS_{\text{MM}} \quad (1)$$

where E_{MM} is the total molecular mechanical energy; G_{PBSA}, the solvation free energy; and TS_{MM}, the entropy. The molecular mechanical energy contains an electrostatic term, a van der Waals term, and an internal energy term. The solvation free energy is divided into two contributions: the electrostatic contribution and the nonpolar (or hydrophobic) contribution.

$$G_{\text{PBSA}} = G_{\text{PB}} + G_{\text{np}} \quad (2)$$

For the purpose of the present study, we calculated only a subset of terms entering the MM/PBSA Eq. (1), namely the total electrostatic contribution, the van der Waals contribution, and the nonpolar, or hydrophobic, contribution

to the free energy. As we neglect conformational change upon complexation (see later), there are no changes to the internal energy terms. The binding free energy was then estimated according to the following equation:

$$\Delta G = \Delta E^{\text{elec}} + \Delta E^{\text{vdw}} + \Delta G^{\text{elec}}_{\text{solv}} + \Delta G^{\text{np}}_{\text{solv}} \quad (3)$$

where ΔE^{elec} and ΔE^{vdw} are the electrostatic and van der Waals contributions to the energy of the complex formation, respectively; $\Delta G^{\text{elec}}_{\text{solv}}$ and $\Delta G^{\text{np}}_{\text{solv}}$ are the electrostatic and nonpolar contributions related to solvation, respectively.

Equation 3 neglects conformational entropy contributions to the thermodynamics of binding. We did not estimate this term because our main goal was to apply this method to identify residues that play a dominant role in stabilizing the bound complex rather than obtaining a quantitative account of the binding thermodynamics. Although methods have been proposed to analyze individual contributions to conformational entropies,[51,52] their generalization in the present context would entail a significant and unnecessary computational cost. Another simplification introduced concerns the neglect of conformation changes between the bound and unbound proteins. In the absence of an experimental structure for the unbound antibody, the structures of the unbound proteins were taken from that of the complex. There are experimental indications that the antibody–antigen association is dominated by enthalpy and occurs without large structural rearrangements.[53,54] These simplifications preclude calculations of absolute values of the binding free energy; however, they are expected to be satisfactory in the context of identifying interaction energy "hot spots" in the protein–protein complex, which is the goal of the present study. Similar simplifications have been adopted by other authors.[10,19,55,56] Each contribution to the binding free energy entering Eq. (3) was decomposed as a sum of individual contributions per residue, as described later.

van der Waals contributions

The van der Waals energy was calculated by the 6-12 Lennard-Jones potential of the CHARMM22 force field. The individual contributions to the van der Waals interaction energy of each residue of the antibody was estimated as one half the interaction with the set of lysozyme residues located in a 12.5 Å cutoff distance, and vice versa for lysozyme. The division by two ensures that the sum of individual contributions gives the total van der Waals interaction energy of the complex.

Electrostatic contributions

The total electrostatic energy of a macromolecular system can be partitioned into contributions from interactions between pairs of atomic partial charges of the macromolecule and from the interactions of individual partial charges with the solvent.

In a process such as protein association, there is always interplay between favorable and unfavorable contributions. For example, there is always a destabilizing contribution due to the loss of favorable interactions between charged and polar residues and the solvent molecules (desolvation). In addition, there is a change in the number, and the environment, of pair-wise interactions resulting from complex formation, which counterbalances the unfavorable desolvation term.

We used the UHBD program[47] to solve the linearized Poisson–Boltzmann equation using a finite-difference method (FDPB). A shell script, which includes interfaces to UHBD, was used to compute the electrostatic binding free energy of binding of the two molecules. The boundary between the solute and solvent regions with different dielectric constants is determined by the molecular surface. van der Waals radii and charges of atoms are obtained from parameters of the force field CHARMM22. A dielectric constant of 80 was used for the solvent and 1 for the protein (the value of 1 for the protein dielectric constant is used when conformational averaging is performed; for the single calculation on the X-ray structure of the cAbLys3-HEL, a value of 4 was used for the protein dielectric constant[10]). The final mesh grid on which charges are located and on which the electrostatic potential is calculated has a grid spacing of 0.35 Å.

The electrostatic contribution was decomposed by residues. This decomposition is possible because the linear Poisson–Boltzmann equation is used, which permits a superposition of the electrostatic potential at point i by the sum of potentials created by each individual charge.

$$\Delta G^{\text{elec}} = \sum_{i\in C}\frac{1}{2}q_i\phi_i^{\text{complex}} - \left(\sum_{i\in A}\frac{1}{2}q_i\phi_i^{\text{antibody}} + \sum_{i\in L}\frac{1}{2}q_i\phi_i^{\text{lysozyme}}\right) \quad (4)$$

This equation shows that the total electrostatic energy of a complex C (antibody + lysozyme) is calculated by summing over all charges, the product of the charge q_i and the potential ϕ_i at the position of the charge. The potential ϕ_i is obtained by solving the linear Poisson–Boltzmann (PB) equation using the program UHBD. The contribution of one residue j to the electrostatic free energy of complex formation is estimated by summing over all charges q_i using the potentials ϕ_i created at the position i by the atomic charges of residue j:

$$\Delta G^{\text{elec}}_j = \sum_{i\in C}\frac{1}{2}q_i\phi_i^{j\in C} - \left(\sum_{i\in A}\frac{1}{2}q_i\phi_i^{j\in A} + \sum_{i\in L}\frac{1}{2}q_i\phi_i^{j\in L}\right) \quad (5)$$

where j belongs either to A (antibody) or to L (lysozyme).

Such decompositions were first used to interpret continuum electrostatic calculations for solvation of phosphoranes in water,[57] and a decomposition scheme similar to equation (4) was first applied to proteins by Hendsch and Tidor.[10] However, in earlier free energy decomposition work based on the Poisson–Boltzmann method, only a sin-

gle conformation of the protein–protein complex was usually used,[10,19,58] which is not the case here.

Nonpolar contributions

The nonpolar contributions to the free energy of binding are taken to be proportional to the loss in solvent accessible surface area.

$$\Delta G^{\mathrm{np}} = \gamma\left(SAS^{\mathrm{complex}} - \left(SAS^{\mathrm{antibody}}_{\mathrm{unbound}} + SAS^{\mathrm{lysozyme}}_{\mathrm{unbound}}\right)\right) \quad (6)$$

The radii of protein atoms are obtained from the parameters of the force field CHARMM22. We used CHARMM for calculating the accessible surface with a probe radius of 1.4 Å. We determined the accessible surface of each residue in the free molecules and in the complex and then we calculated the accessible surface area difference between the free molecule and the complex for each residue. A value of 5 cal mol^{-1} $\mathrm{Å}^{-2}$ was used[13,20,21] for the constant γ.

Selection of Structures for the Free Energy Decomposition

Past free energy decomposition analysis based on the MM/PBSA or MM/GBSA methods have been performed using either a single experimental structure[9,10] or an ensemble of conformations extracted from molecular dynamic simulations.[13,56] It is known that the different energy terms used in the MM/PBSA method fluctuate significantly with small changes in structure linked to thermal fluctuations.[59] It is therefore necessary to use an ensemble of structures rather than one experimental structure in free energy decompositions based on the MM/PBSA scheme. Obtaining converged values of binding free energies for quantitative assessment of binding thermodynamics entails significant computational effort.[13] The intention of the present work was to reliably identify those amino acids that play a dominant role in binding, and to obtain a semi-quantitative estimate of their contributions to the binding thermodynamics. We, therefore, developed a protocol that uses only a limited number of conformations extracted from molecular dynamics trajectories for the free energy decomposition, but that allows sufficient sampling to reduce artifacts linked to the use of a single conformation. For the system studied, several crystal structures were available (cf. *Structures* Section). These structures were used as starting points for short molecular dynamics trajectories. Thus, significant conformational differences would be taken into account by using the different initial structures and the MD simulations would give small thermal fluctuations around the crystal structure. Ten conformations were extracted from the MD trajectories for analysis. The structures were not extracted at regular time intervals, but selected using a protocol that was designed to cover a large, representative range of individual MM/PBSA energy terms. The component of the binding free energy that shows the largest fluctuations with small changes in structures is the electrostatic Coulomb term. We therefore computed the Coulomb interaction energy *in vacuo* between the protein cAb-Lys3 and lysozyme for all conformations saved from the MD trajectory (coordinates were collected every 0.1 ps from the MD). The Coulomb electrostatic interaction energy *in vacuo* was calculated using the CHARMM program. A dielectric constant of 1 and a nonbonded cutoff of 12.5 Å was used, with a shift truncation function for electrostatics. The conformations were clustered in ten groups that had similar values of the electrostatic interaction energy, and that were distributed evenly between the lowest and highest values recorded for the electrostatic interactions over the course of the trajectory. One conformation was extracted from each cluster (with vacuum interaction energy closest to the cluster average value). Each extracted conformation was then postprocessed using the MM/PBSA procedure. The total free energy and its individual components were then obtained for each trajectory as a weighted average over the conformations extracted, with higher weight given to conformations extracted from more populated clusters (i.e. the weight of a conformation is proportional to the ratio of the number of conformations with that energy to the total number of conformations extracted from the MD trajectory). This procedure reduces the computational cost of performing MM/PBSA calculations while allowing an estimate of the effect of conformational variation on the decomposition analysis. While the energy used for selecting the conformations was the vacuum electrostatic term, which differs from the solvent term computed by continuum electrostatics, the two terms show a significant statistical correlation, so that conformations that show large variation in vacuum energy also show significant variation in solution energy (data not shown,[60]), hence the interest of the simplified protocol.

A standard deviation of each amino acid contribution is also calculated. The statistical weight of the conformation is taken into account in computing the standard deviation.[61] When results from different simulations using different starting structures are averaged, the total number of conformations in all trajectories is used to normalize the weight of each individual conformation in the average.

Amino Acid Efficiency

We define the amino acid efficiency as the free energy contribution of each amino acid divided by its number of nonhydrogen atoms (i.e. 4 for Gly, 5 for Ala, 14 for Trp etc.). This definition is similar in spirit to the definition of ligand efficiency[62] which is defined as the binding affinity per heavy atom. When extending the notion to amino acids within a macromolecule, we also considered that some amino acids may be partially or even fully buried in unbound proteins. We therefore tested an alternative definition of the number of atoms as the solvent accessible atoms rather than all heavy atoms. The effective number of solvent accessible atoms was computed in the following way: for each of the 20 amino acid, AA, a tripeptide Gly-AA-Gly was constructed in extended conformation and the solvent accessible surface of the amino acid in the tri-

peptide was computed. The radii of amino acids atoms were obtained from the CHARMM force field and CHARMM program was used for calculating the accessible surface with a probe radius of 1.4 Å. The solvent accessible surface of each amino acid of the unbound protein was computed using the same procedure. Then an effective number of atoms (N_{effect}) for each amino acid within the protein was computed as

$$N_{\text{effect}} = \left(N_{\text{heavy}} \cdot \left[\frac{S_{\text{prot}}}{S_{\text{peptide}}} \right] \right) + 1 \qquad (7)$$

In this equation, N_{heavy} is the number of heavy atoms in the amino acid; S_{prot}, the solvent accessible surface of the amino acid in the unbound protein; S_{peptide}, the solvent accessible surface of the amino acid in the tripeptide. If the amino acid is completely buried in the unbound protein ($S_{\text{prot}} = 0$), the effective number of atoms is set to one.

The amino acid efficiency was calculated in two ways: first by dividing the free energy contribution of each amino acid by the effective number of atoms (N_{effect}), and second, by dividing the free energy contribution by the number of heavy atom (N_{heavy}). The results were compared and they were not found to be significantly different between the two procedures. The data using N_{effect} are presented in the Supplementary Material and the amino acid efficiency calculated by dividing the free energy by the number of heavy atoms was used in the analysis of the results presented.

Mutations and Binding Affinity Measurements

Production of VHH mutants

The cAb-Lys3-pHEN6 vector[29] containing the gene for the expression of the recombinant cAb-Lys3 with a C-terminal His_6 was a kind gift from Drs. M. Lauwereys and S. Muyldermans (Laboratorium voor Ultrastructuur, Brussels). The mutations T28A, T28D, T28K, D73S, D73R, D73K, Y118F, and I29T were introduced by standard site directed mutagenesis techniques. All mutant genes were confirmed by nucleotide sequencing. The proteins were expressed and purified as described by Lauwereys et al.[29]

Affinity measurements

The interaction of the wild type or mutant cAb-Lys3 with either TEL or HEL was analyzed using a Biacore® 2000 instrument (Biacore AB, Uppsala, Sweden) as described previously.[63] Briefly, all experiments were carried out at 25°C in HBS-EP buffer (10 m*M* HEPES pH 7.4, 150 m*M* NaCl, 3.4 m*M* EDTA with 0.005% Surfactant P20). TEL or HEL were immobilized on a CM5 sensor surface using the amine coupling chemistry as described.[63] The active concentrations of the antibody fragments were measured by injecting the samples on a surface containing high levels of lysozyme in conditions of total mass transport.[64] The reaction rate was recorded 15 s after injection, and used to calculate the active cAb-Lys3 concentration based on a previously established calibration curve. For kinetic measurements, the WT and mutant cAb-Lys3 molecules were injected at different concentrations on a control surface and on a surface containing low levels of immobilized lysozyme. The injection and postinjection phases lasted between 3 and 5 min. The surfaces were regenerated with 50 m*M* HCl. Data were fitted globally with the BIAevaluation 3.1 software (Biacore AB), using the Langmuir binding model.

RESULTS

Free Energy Decomposition for the Antibody cAb-Lys3

Here, we report the results concerning the contribution of the camelid single chain variable domain (cAb-Lys3) to total binding free energy of the complex with hen egg white lysozyme (HEL) and turkey egg white lysozyme (TEL). The cAb-Lys3 single variable domain fragment binds to HEL and TEL with nanomolar affinity.[65] On average, the solvent accessible surface of the antibody that is buried on binding is 998 $Å^2$ for the HEL and 945 $Å^2$ for the TEL complex. To identify those amino acids that make significant interactions in the bound complex, the free energy decomposition scheme was applied to the cAb-Lys3/HEL and cAb-Lys3/TEL complexes. In all, 6 simulations were run of the different complexes and the MM/PBSA analysis was applied to all trajectories following the procedure outlined in Methods. The results presented here in detail are average results over the different antigens, HEL and TEL, respectively. The decomposition results, presented in Figs. 1–3, are for the cAb-Lys3/HEL complex. A comparison to the results obtained for the cAb-Lys3/TEL complex is given in Supplementary Material.

In cAb-Lys3, the residues that contact lysozyme belong to the Complementary Determining Regions (CDRs) (as defined in Ref. 67; see also Ref. 22). The results show that amino acids that constitute the CDRs make the most important contributions to the binding interactions, while the framework region of the antibody (defined as the antibody without the CDRs) makes mostly negligible contributions (see Fig. 1 and Table II). The residues of the CDRs do not, however, contribute uniformly to the binding interactions (see Fig. 1). The third Complementary Determining Region (CDR 3), which is particularly long in the cAb-Lys3 and inserts in the lysozyme active site,[22] makes the dominant contribution to binding (see Fig. 1 and Table II); in particular, a contiguous stretch of seven residues (Thr 101 - Ile 102 - Tyr 103 - Ala 104 - Ser 105 - Tyr 106 - Tyr 107) is shown to contribute 75% of the total binding free energy, while it contributes only 57% of the buried cAb-Lys3 surface.

An analysis of the binding in terms of electrostatic, van der Waals and hydrophobic (based on surface area) contributions (see Fig. 2 and Table II) shows that for cAb-Lys3, binding is dominated by the van der Waals interactions, while the hydrophobic (surface area) contribution is smaller. The relative contributions of the three CDRs and of the framework region to the hydrophobic term are roughly correlated with their respective van der Waals contributions.

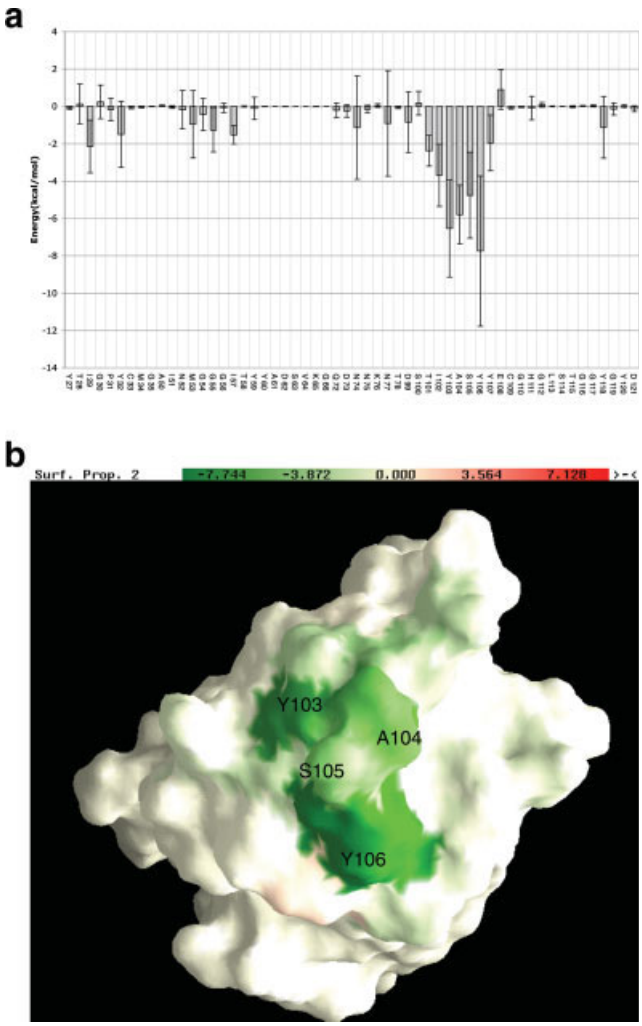

Fig. 1. (a) Net binding energy contribution by residue of the antibody for the cAb-Lys3/hen egg white lysozyme complex. The values reported are averages of the total contribution ΔG, as estimated from Eq. (3). The results are presented for residues belonging to complementary determining region (CDR1 from Pro 31 to Gly 35, CDR2 from Ala 50 to Gly 66, and CDR3 from Asp 99 to Asp 121) and for residues of the framework region with a contribution different from zero. Averages are calculated over all structures extracted from the trajectories (MEL AL HSD, MEL AL HSE, MEL BM, and JTT). Units: kcal/mol. (b) Antibody surface colored according to the total energetic contribution of each residue. Coloring from green (favorable) to red (unfavorable) with contributions in the range from -7.74 to $+7.13$ kcal/mol. Figure generated using the program GRASP.[66] The residues with the most favorable or unfavorable contributions are labeled.

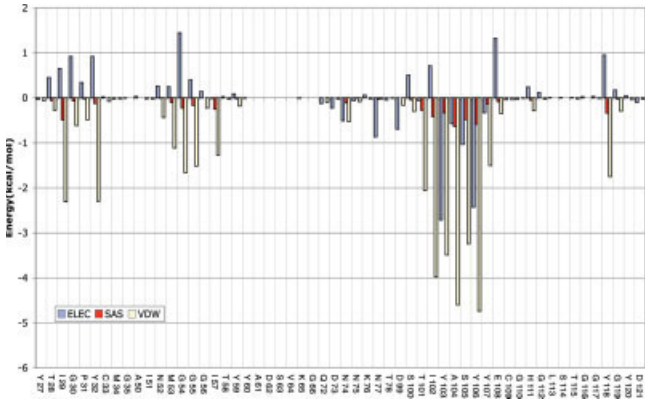

Fig. 2. Decomposition of the total energetic contribution of antibody residues presented in Figure 1a. The electrostatic (ELEC in blue), the nonpolar (SAS in red) and the van der Waals (VDW in yellow) contributions, as defined in Methods, are displayed separately for selected residues (see Figure 1a for selected residue). Units: kcal/mol.

Overall, the electrostatic contribution of the cAb-Lys3 to binding is about zero, which results from the compensation between (+4 kcal/mol) unfavorable electrostatic contributions of CDR's 1 and 2 and favorable contributions of a few amino acids in CDR3. The standard deviation on the electrostatic contribution is comparatively large (see Table II), which is expected, given the magnitude of the Coulomb and the solvation free energy terms that are involved. In the small seven residue sequence of CDR 3 that dominates the binding energy, some of the residues have favorable electrostatic contributions to binding, particularly in the cAb-Lys3/HEL complex. Although more extensive conformational sampling would be required to get a quantitative assessment of the electrostatic contribution for these residues, it is striking (cf. Fig. 2) that this sequence of residues emerges as the only region where favorable electrostatic contributions are found. This favorable region of cAb-Lys3 is shown in Fig. 1b.

Free Energy Decomposition for Lysozyme

The analysis of the contributions of lysozyme to the binding energetics reveals important qualitative differences with respect to the cAb-Lys3 (cf. Figs. 3 and 4 and Table II). Lysozyme buries on average a solvent accessible surface of 900 Å^2 for HEL and 822 Å^2 for TEL upon complex formation.

The amino acids that make a dominant contribution to binding are not contiguous, but form two discontinuous patches around amino acids 58-63 and 106-109 (see Fig. 3b). The residues that make the most favorable contributions to binding contribute only 42% of the buried lysozyme (HEL) surface (50% of TEL). As for cAb-Lys3, the favorable interactions of these residues are dominated by the van der Waals contribution (cf. Table II and Fig. 4). These residues line the binding pocket into which the CDR3 of cAb-Lys3 inserts (cf. Fig. 3b and discussion given below).

Lysozyme (HEL) contains several charged residues at or near the binding interface that become buried upon binding. An analysis of the binding in terms of electrostatic, van der Waals, and hydrophobic contributions (see Fig. 4 and Table II) shows that several charged residues (D52, R73, and R112) have large unfavorable electrostatic terms. Although some of them make important interactions with cAb-Lys3 amino acids (see later), the interaction term is not important enough to compensate for their desolvation, and as a result, the total electrostatic contribution of these amino acids is positive (see Figure 4). While small local conformational differences between the

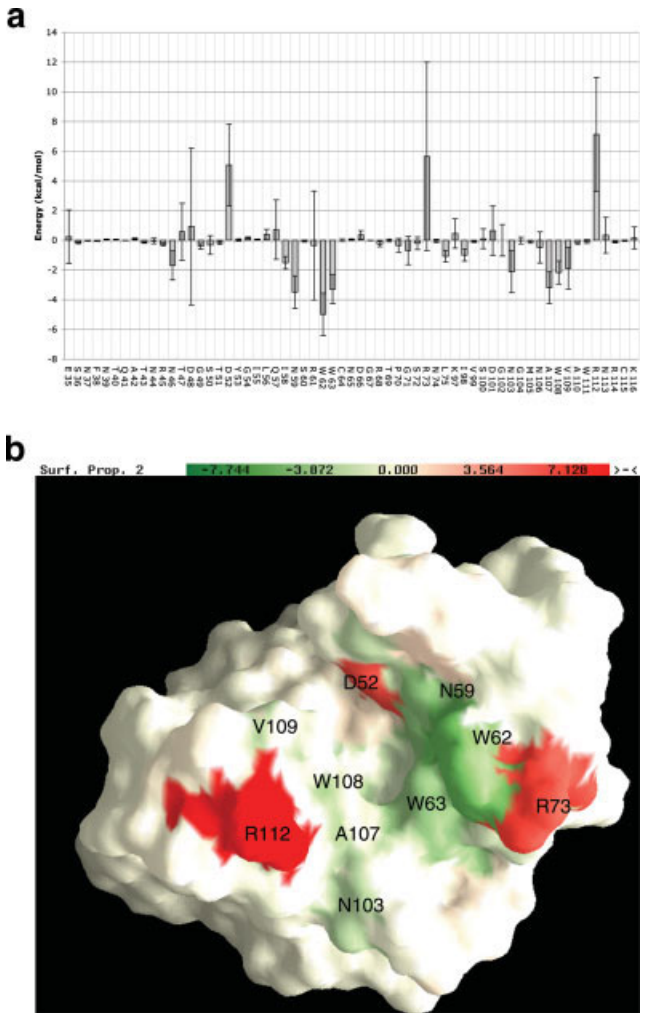

Fig. 3. (a) Net binding energy contribution by residue of the HEL lysozyme for the cAb-Lys3/hen egg white lysozyme complex. The values reported are averages of the total contribution ΔG [cf. Eq. (3)]. The results are presented for residues with a contribution different from zero. The averages were calculated over all structures extracted from the trajectories of the different initial structures (MEL AL HSD, MEL AL HSE, MEL BM, and JTT). Units: kcal/mol. (b) Projection of the total energetic contribution of each residue onto the lysozyme surface. The residues are colored from green (favorable contributions) to red (unfavorable contributions). The values range from -7.74 to $+7.13$ kcal/mol. The picture was prepared using the program GRASP[66]. The residues with the most favorable or unfavorable contributions are noted. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

complexed and uncomplexed lysozyme may modulate the magnitude of this term, the desolvation effect explains, in large part, why the electrostatic contribution is large and positive for lysozyme (about $+36 \pm 6$ kcal/mol, average of HEL and TEL; see Table II).

This is an important difference with the cAb-Lys3, where such large positive contributions were not observed. As a result, although unfavorable electrostatic interactions are compensated by the favorable nonpolar ones, the contribution of lysozyme to the binding energetics is small when compared with that of cAb-Lys3. The relative contribution of lysozyme to the total binding free energy is only about 15% (cf. Table II, -7 kcal/mol from a total of -52 kcal/mol for HEL, and -6 kcal/mol out of a total of -38 kcal/mol for TEL). The observation that the difference in total contribution to binding between the two partners is directly linked to the difference in electrostatic interaction also results from the fact that, by definition (see Methods), the van der Waals interactions are shared equally between the two proteins, and that the difference in buried surface area between two proteins in a complex is very small.[68] Electrostatic desolvation, on the other hand, is strongly dependent on the individual structure and charge distribution of each partner, and large differences can be observed for two proteins in a complex as demonstrated here. Although the absolute values of the antibody and lysozyme contributions to binding differ between the HEL and TEL complexes (see Table II), the relative magnitudes of the individual contributions are very similar.

Comparison of the Different Structures

We studied three experimental structures of the cAb-Lys3 in complex with HEL, and two of the complex with TEL. The three HEL complex structures differ significantly by a rigid body rotation of one protein with respect to the other[42] and both the sequence and the internal structure of the lysozyme monomer are different in the TEL and in the HEL complex structures. The RMS differences between the different protein–protein complexes (in both the crystal structures and structures obtained from the MD trajectory) are given in Table 1 of Supplementary Material.

With these differences, it is possible to test the robustness of the free energy decomposition scheme with respect to changes in sequence and in conformation space. The short MD simulations that were run for each structure were used to generate ensembles of conformations and thus to reduce statistical error. We are then able to obtain a more general understanding of the recognition process in that our results are less dependent on the specific three-dimensional structure.

A comparison of the results for the two lysozymes (see Fig. 5) shows that the interaction energy hot spots are the same in the two complexes; similar results are obtained for the antibody (see Fig. 1 of Supplementary Material). However, some charged amino acids of the interface, such as Asp 48 and Arg 73 (HEL) or Lys 73 (TEL), show significantly different binding energy contributions for the five structures of the complex (three structures for HEL and two structures for TEL). These differences can be explained either structurally or based on the lysozyme sequences. For example, Asp 48 of lysozyme can form favorable interactions with Tyr 106 of cAb-Lys3, but it is situated in a loop and shows larger structural fluctuations than other amino acids. Hence its contribution varies considerably between the different structures. Likewise, Arg 73 (HEL) or Lys 73 (TEL) are in a region of the protein–protein interface that shows significant structural

TABLE II. Individual Contributions to the Binding Energy for the cAb-Lys3/HEL Complex (top part) and the cAb-Lys3/TEL Complex (lower part)

	ELEC	SAS	VDW	TOTAL
CHICKEN				
CDR1 {P31-G35}	1.3 ± 1.2	-0.1 ± 0.04	-2.9 ± 0.9	-1.8 ± 2.2
CDR2 {A50-G66}	2.6 ± 2.6	-0.7 ± 0.1	-6.4 ± 1.2	-4.6 ± 3.9
CDR3 {D99-D121}	-3.8 ± 5.5	-3.4 ± 0.2	-26.9 ± 2.0	-34.0 ± 7.7
HOT SPOT {T101-Y107}	-6.4 ± 4.6	-2.8 ± 0.2	-23.6 ± 1.9	-32.8 ± 6.7
FRAME {all without CDRs}	0.3 ± 5.3	-0.7 ± 0.3	-4.0 ± 1.9	-4.4 ± 7.4
cAb-Lys3 {all antibody}	0.4 ± 4.5	-5.0 ± 0.2	-40.2 ± 3.6	-44.8 ± 8.3
Lysozyme: N46, I58, N59, W62, W63, N103, A107, W108, V109	-1.8 ± 2.1	-1.9 ± 0.1	-20.6 ± 1.1	-24.3 ± 3.3
Lysozyme: D48, D52, R61, R73, R112	29.5 ± 6.3	-1.5 ± 0.1	-9.6 ± 1.7	18.4 ± 8.1
HEL {all lysozyme}	37.4 ± 7.0	-4.5 ± 0.3	-40.2 ± 3.6	-7.2 ± 10.9
Total	37.8 ± 8.6	-9.5 ± 0.5	-80.4 ± 7.2	-52.0 ± 16.2
TURKEY				
CDR1 {P31-G35}	2.6 ± 0.8	-0.2 ± 0.03	-2.1 ± 0.5	0.3 ± 1.3
CDR2 {A50-G66}	3.7 ± 1.2	-0.7 ± 0.04	-5.0 ± 0.4	-2.1 ± 1.7
CDR3 {D99-D121}	1.8 ± 3.9	-3.4 ± 0.2	-26.9 ± 1.7	-28.5 ± 5.8
HOT SPOT {T101-Y107}	1.1 ± 2.9	-2.8 ± 0.1	-23.8 ± 1.7	-25.5 ± 4.7
FRAME {all without CDR's}	0.4 ± 0.6	-0.4 ± 0.1	-2.0 ± 0.3	-2.0 ± 1.0
cAb-Lys3 {all antibody}	8.5 ± 4.7	-4.7 ± 0.2	-36.0 ± 2.0	-32.3 ± 6.8
Lysozyme: N46, I58, N59, W62, W63, N103, A107, W108, V109	0.02 ± 1.9	-2.1 ± 0.1	-20.0 ± 1.2	-22.08 ± 3.2
Lysozyme: D48, D52, R61, R73, R112	30.1 ± 3.3	-1.3 ± 0.1	-8.8 ± 0.7	20.0 ± 4.1
TEL {all lysozyme}	34.3 ± 3.3	-4.2 ± 0.2	-36.0 ± 2.0	-5.8 ± 5.5
Total	42.8 ± 7.4	-8.9 ± 0.3	-72.1 ± 3.9	-38.1 ± 11.6

All contributions are averages over the structures extracted from several molecular dynamics trajectories, as described in the Methods section. The contributions for CDRs (Complementary Determining Regions) 1 to 3, for the hot spot residues (T101 to Y107 for cAb-Lys3), for the framework (residues not in the CDR regions of cAb-Lys3, FRAME), for the entire antibody (cAb-Lys3), for lysozyme (HEL or TEL), for the hydrophobic patches in lysozyme, and for charged residues in lysozyme are obtained by summing over the specified residues. All energies are in kcal/mol.

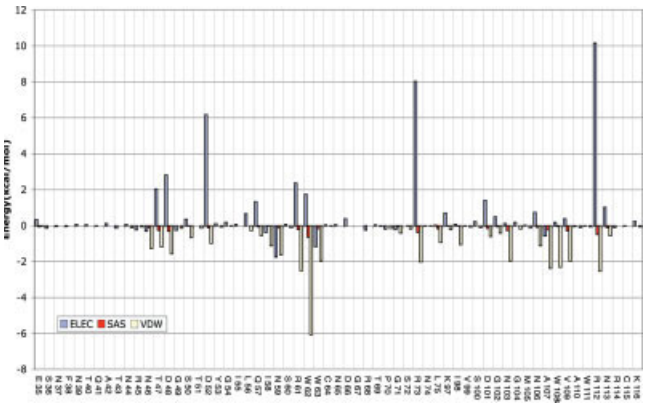

Fig. 4. Decomposition of the total energetic contribution of lysozyme residues presented in Figure 3a. The electrostatic (ELEC in blue), the nonpolar (SAS in red) and the van der Waals (VDW in yellow) contributions are displayed separately for selected residues (see Fig. 3a for selected residue). Units: kcal/mol. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

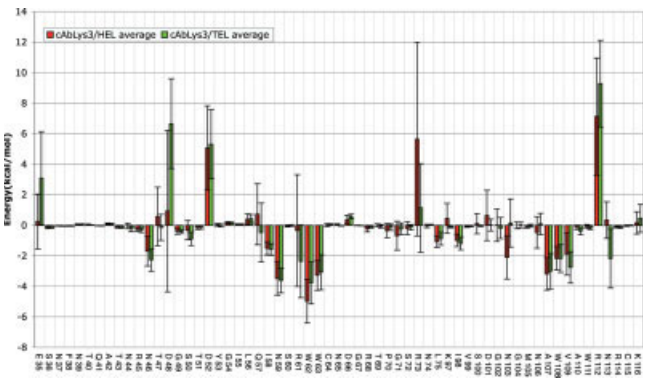

Fig. 5. Comparison of results obtained for the two complexes cAb-Lys3/hen egg white lysozyme (red) and the cAb-Lys3/turkey egg white lysozyme (green). The contributions plotted are averages of total contributions of lysozyme residues. The two lysozymes differ by only seven residues. The sequence used in this plot is the one for the chicken egg-white lysozyme. Amongst the residues shown in the graph, the following sequence differences in the turkey egg-white lysozyme are found: Q41H, R73K, V99A, and D101G. The averages were calculated over all trajectories generated for each complex. The standard deviation for each residue is shown. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

variations. These results underline the need to exercise caution when interpreting the results from a free energy decomposition for a single crystallographic structure. The conformational sampling applied in this study yields a mean value with a standard deviation for each contribution; the standard deviation is a useful first indication of the dependence of the results on small structural changes.

For example, in the cAb-Lys3/HEL complex, the conformational sampling clearly helps discriminating the true hot spot residues from amino acids that have favorable but highly fluctuating contributions (see Fig. 1 of Supplementary Material). These results suggest that, whenever

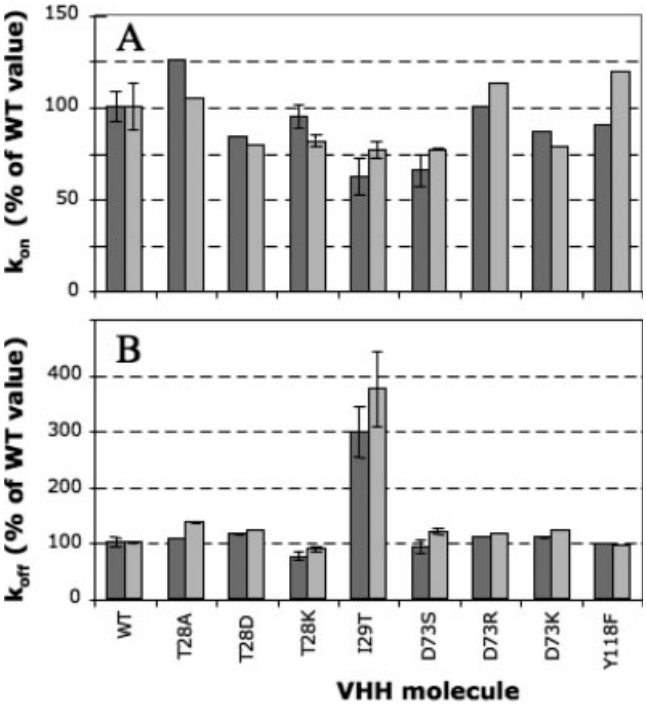

Fig. 6. Comparison of kinetic on-rate (a) and off-rate (b) parameters for the interaction between the WT or mutant cAb-Lys3 with HEL (dark grey) or TEL (light grey). Rate parameters are expressed as percent of the wild type value. Measurements were repeated 2–3 times for the WT cAb-Lys3 and mutants T28K, I29T, and D73S, for which error bars are shown.

available, averaging over several experimental structures is likely to increase the significance of the results obtained from the decomposition scheme.

Mutational Analysis of Four cAb-Lys3 Positions

To further investigate the importance of using conformational averaging when calculating a free energy decomposition, 4 amino acids of the antibody, identified as making nonzero contributions to binding free energy by a free energy decomposition calculation on a single crystal structure, were mutated as specified in Methods. These were Thr 28, Asp 73, and Tyr 118, which, in the cAbLys3-HEL crystal structure (BM complex), have total contributions to the binding free energy of −1.1, −0.5 and −2.3 kcal/mol respectively. Thr 28 and Asp 73 were replaced by residues that lead to important modifications of size and/or charge (T28A, T28D, T28K, D73S, D73R, and D73K), while a more conservative change was made at position 118 (Y118F), to avoid large structural effects of the mutation. Residue Ile 29, which was identified as important for binding, was mutated to Thr as a positive control. The kinetics of the interaction of the wild type and eight cAbLys3 mutants with either HEL or TEL were measured by Biacore. Kinetic parameters, expressed as percentage of those measured for the interaction of HEL or TEL with the unmodified cAb-Lys3, are compared in Figure 6. Except for the I29T change, which produced a 3- to 4-fold increase in k_{off}, other changes had no measurable effect on interaction kinetics. From the computational analysis, contributions from Thr 28, Asp 73, and Tyr 118 showed large fluctuations with conformational averaging and yielded small binding averages (see Fig. 1a and Fig. S1 in Supplementary material). However, contributions from individual structures can be important, as is shown by the large error bars.

DISCUSSION

Structural Analysis of the Binding Energy Hot Spots

A useful aspect of the free energy decomposition scheme presented here is that it not only identifies the binding energy hot spots, but also gives insight into the nature of the key interactions.

The amino acids from the cAb-Lys3 (see Figs. 1 and 2 and Table II) that make the most favorable contributions to binding (between −2 and −8 kcal/mol) form a contiguous stretch of residues in the CDR3: Thr 101 - Ile 102 - Tyr 103 - Ala 104 - Ser 105 - Tyr 106 - Tyr 107 (one additional important amino acid is Ile 29 from CDR1), while for lysozyme, the most favorable interactions are made by Asn 46- Ile 58 - Asn 59 - Trp 62 - Trp 63 - Asn 103 - Ala 107 - Trp 108 - Val 109 (see Figs. 3 and 4 and Table II), which are grouped together spatially.

For a more detailed analysis, we determined whether the favorable interactions of the hot spot residues are predominantly van der Waals or electrostatic in nature and whether they face each other on the two partners of the complex. Concerning cAb-Lys3, the hot spot interactions are mainly nonpolar. The important amino acids from CDR3 make favorable contributions to the interaction energy via the solvent accessible surface, or hydrophobic term and via the van der Waals interactions. Some of these important van der Waals interactions are made with amino acids on lysozyme that are also, correspondingly calculated as making important contributions. For example, the side chain of Ala 104 of cAb-Lys3 fills a hydrophobic pocket lined by the side chains of Ile 58, Ala 107, and Trp 108 of lysozyme. Particularly favorable van der Waals interactions are also made by Trp 62 of lysozyme. This amino acid not only makes an important contact with a hot spot residue from cAb-Lys3 (Ile 102 from CDR3), but it also makes contacts with several amino acids from CDR1 (Ile 29 and Tyr 32) and CDR2 (Gly 54), which have smaller van der Waals contributions to the binding energy. The large contribution of Trp 62 (lysozyme) arises from its central position at the protein–protein interface that allows a large number of van der Waals contacts.

Thus the size of the amino acid and the shape of the contact surface plays an important role and in some cases a van der Waals hot spot is located on one of the proteins, whereas there is no dominating van der Waals contribution arising from the reciprocal residues on the other protein. Thus, even though van der Waals interactions are shared equally between the two proteins, at the residue level, the hot spots are not necessarily symmetrical.

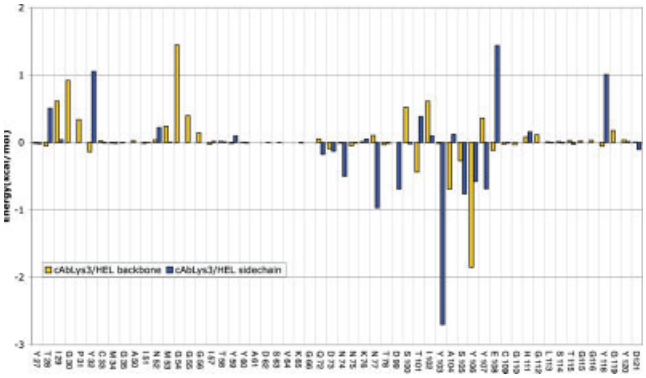

Fig. 7. Decomposition of the electrostatic contributions into side chain and main chain components for the antibody in the cAb-Lys3/hen egg white lysozyme complex. The values represented are averages of the electrostatic contribution of backbone (yellow) and side chain (blue) of antibody residues. (see Fig. 1a for selected residues). The average was calculated over all the trajectories done with the different initial structures (MEL AL HSD, MEL AL HSE, MEL BM, and JTT). Units: kcal/mol. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Even though van der Waals interactions have an important contribution to hot spots, some important electrostatic interactions can also be identified. The electrostatic contribution is favorable for cAb-Lys3 residues Tyr 103 - Ala 104 - Ser 105 - Tyr 106, which are involved in a hydrogen bond network with several residues from lysozyme. The most favorable electrostatic interactions are made by the two tyrosine residues; they can be further decomposed into side chain and main chain contributions. As can be seen from Figure 7, the dominant contributions are made by the side chain for Tyr 103 (the main chain also makes a favorable, but less important, contribution) and by the main chain for Tyr 106. Examination of the structures indicates that the Tyr 103 side chain hydroxyl group makes an H-bond with Arg 112 from lysozyme, while the Tyr 106 main chain NH makes an H-bond with the carboxyl group of Asp 52 from lysozyme. The favorable contribution of these H-bonds is directly linked to the strong electrostatic potential created by the charged amino acids on lysozyme. Remarkably, the free energy decomposition for lysozyme (see Figs. 3 and 4) yields unfavorable electrostatic contributions to the binding free energy for Arg 112 and Asp 52. This is due to the large desolvation penalty upon binding for the charged side chains of these residues. The fact that electrostatic interactions are not always favorable to binding has been noted in other free energy decomposition studies.[13,19,21,58] In this particular example, it can be seen that Arg 112 and Asp 52, although not favorable themselves, create a favorable environment for Tyr 103 and Tyr 106 on the cAb-Lys3, and thus have an important role in the formation of the complex. The other charged amino acid on lysozyme that has a large unfavorable desolvation contribution is Arg 73. Arg 73 is not involved in hydrogen bonds with amino acids on cAb-Lys3, and its binding energy contribution shows larger fluctuations than those of Arg 112 and Asp 52 (see Fig. 3a).

In contrast to the unfavorable electrostatic contributions calculated for Arg 112 and Asp 52 of lysozyme, a small favorable contribution is found for Trp 63 of lysozyme, which interacts with Tyr 103 main chain from cAb-Lys3. Thus, even though the electrostatic interaction from the main chain of Tyr 103 of cAb-Lys3 is smaller than the interaction made by its side chain (cf. Fig. 7), the main chain H-bond yields a favorable binding energy contribution both on cAb-Lys3 and on lysozyme. Two additional amino acids of lysozyme make favorable (although smaller than those found in cAb-Lys3) electrostatic contributions: Asn 59 (HEL: −1.7 kcal/mol) and Ala 107 (HEL: −0.6 kcal/mol), see Figure 4. An analysis of the structures shows that Asn 59 and Ala 107 are engaged in hydrogen bonds with Ala 104 of cAb-Lys3 (cf. Figure 9), and this buried network of H-bonds is particularly favorable.

Comparison of Experimental and Computational Data

The conclusions of the calculations were challenged with both existing and new mutational studies of cAb-Lys3. While the existing studies mainly focus on residues identified here as "hot spot," we chose to focus on four positions with a calculated small contribution to binding energy.

Existing experimental studies include replacements of residues Thr 101 (by L,S,P) and Ser 105 (by N,H,Q,A, G,P,M) from CDR3 of cAb-Lys3 for a quantitative structure-activity study,[53] and replacements of residues from CDR1 and CDR2 to revert to the germ line sequence.[54] The replacements at positions 101 and 105 had varied effects, the largest losses in binding affinity being 120-fold for the T100P change, and 100-fold for the S105Q change.[53] In particular, the mutation S105A resulted in a 10-fold loss in experimental binding affinity (i.e. a change of 1.4 kcal/mol in binding affinity), which is in reasonable agreement with the computed binding free energy contribution of -4.8 ± 2.3 kcal/mol for the serine as a whole (i.e. including main chain contributions that are not lost upon mutation to Ala). Multiple mutations were introduced in CDR1 and CDR2 so that the influence of individual changes is not known.[54] Thermodynamic data however show a role of residues from CDR1 and CDR2 for enthalpic stabilization of the bound complex during maturation,[54] which is in agreement with the results of the present calculations that show that CDR1 and CDR2 make favorable van der Waals interactions with lysozyme and thus play a role in stabilizing the complex.

To investigate whether positions with small, but nonzero, computed contributions can nevertheless influence binding, we targeted residues Thr 28, Asp 73, and Tyr 118. The residue Ile 29, which is identified as important in CDR1, was taken as a positive control. None of the mutations at positions 28, 73, and 118 had a measurable effect on binding kinetics (see Figure 6), while the mutation at position I29 (I29T) increased k_{off} 3- to 4-fold, thus resulting in a loss of binding affinity. These experimental results support the validity of our computational ap-

proach. In particular, the mutations data support our simplified conformational averaging scheme, and show that the three amino acids tested, which showed large fluctuations in the computed contribution, do not influence binding, while their contribution was non negligible if only one experimental X-ray structure is used.

Experimental vs. Computational Hot Spots

The comparison of the experimental data to the computational results raises some important issues. Amino acids that are found to play an important role in binding are often referred to in the computational literature as "binding energy hot spots," irrespective of whether the quantity calculated is the change in binding free energy upon mutating the residue to alanine, or a related but different quantity. In this paper, we used the term hot spot for amino acids that make significant individual contributions (<-1.5 kcal/mol) to the total binding free energy, as described in Material and Methods. It is therefore useful to clarify the relationship between "hot spots" as discussed in the present computations and "hot spots" obtained by alanine scanning experiments.

Experimentally, binding energy hot spots are often defined as residues that, when mutated to alanine, give rise to a significant drop in the binding affinity (typically, changes in binding free energy larger than 1.5 kcal/mol).[7] As discussed in Ref. 69, mutations to alanine can affect the binding affinity by both changing the structure and/or free energy of the unbound partners as well as those of the bound complex. This combination of effects complicates the interpretation of alanine scanning experiments, and consequently their exploitation in, for example, designing small molecule inhibitors that mimic important interactions. An advantage of computational approaches is that the energetic contributions can be separated in a manner that is not attainable experimentally. For example, as done in the present study, it is possible to analyze the energetic contribution of an amino acid without resorting to explicit mutation to alanine.

This is an important aspect of the present calculations, which is to identify binding energy hot spots and, thus, free energy decomposition is used rather than explicit mutation to alanine. Individual amino acid contributions to binding presented in Figs. 1, 2, 3, 4, 7 are therefore not the same quantity as the change in free energy upon mutating the residue to alanine. If a direct comparison to Ala-scanning experiments was desired, it would be possible to calculate free energy changes upon mutating a residue to alanine by constructing the mutant and reevaluating the free energy of binding.[11,70] A full estimate of the free energy change should also consider the structural consequences of mutating a specific amino acid to alanine, both in the complex and in the unbound partners.

Another difference with experimental alanine scanning data is that single amino acid free energy contributions to binding, as computed here, include contributions to the binding energy from main chain atoms. They are obtained by writing the total electrostatic, van der Waals and surface area (hydrophobic solvation) terms as a sum over individual amino acid contributions (as described in Material and Methods). The decomposition is performed in such a way that the sum of all individual contributions corresponds to the total free energy (cf. Eq. 1). Decomposing into individual amino acid contributions implies that pairwise interactions (such as van der Waals and electrostatic) are divided equally between the interacting amino acids i.e. if for example the van der Waals interaction energy between amino acids A and B amounts to -4 kcal/mol, a contribution of -2 kcal/mol will be attributed to A and -2 kcal/mol to B, and likewise for electrostatic interactions. It must be noted that in the continuum description of the solvent that is used here, desolvation (whether electrostatic or hydrophobic) is a self-energy term and is thus naturally decomposed in individual contributions.[57,71]

The earlier-mentioned differences between computational free energy decomposition and experimental alanine scanning data should be kept in mind when comparing computational and experimental hot spot residues.

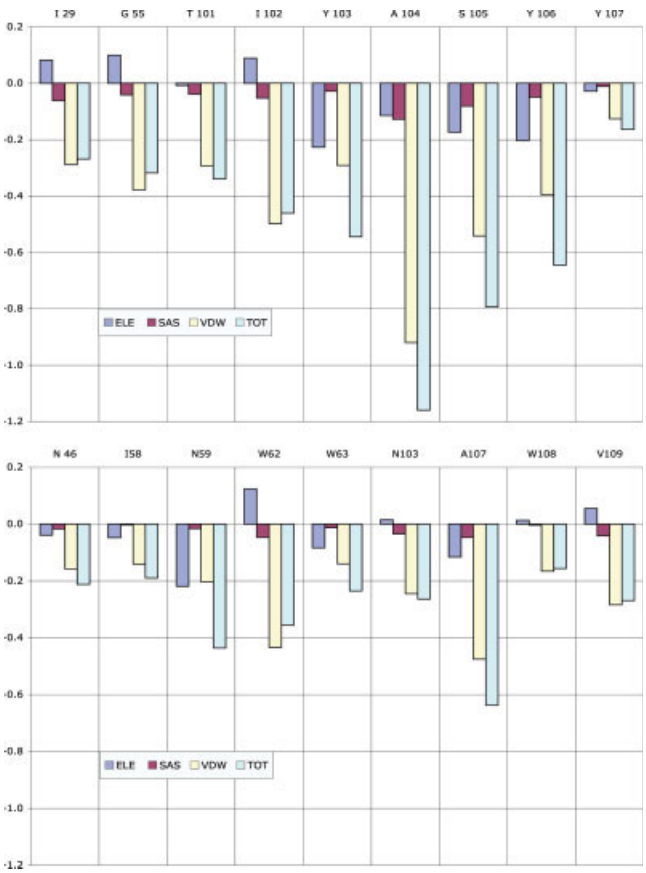

Fig. 8. Bar graph of the amino acid efficiency. Each free energy contribution (Electrostatic (ELE), nonpolar solvation (SAS), van der Waals energy (VDW), and total ΔG (TOT)) of amino acids is divided by the number of nonhydrogen atoms. The results are presented for the residues having the most efficient amino-acids in the interaction for the antibody (top graph) and for the lysozyme (bottom graph). See Methods for details. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

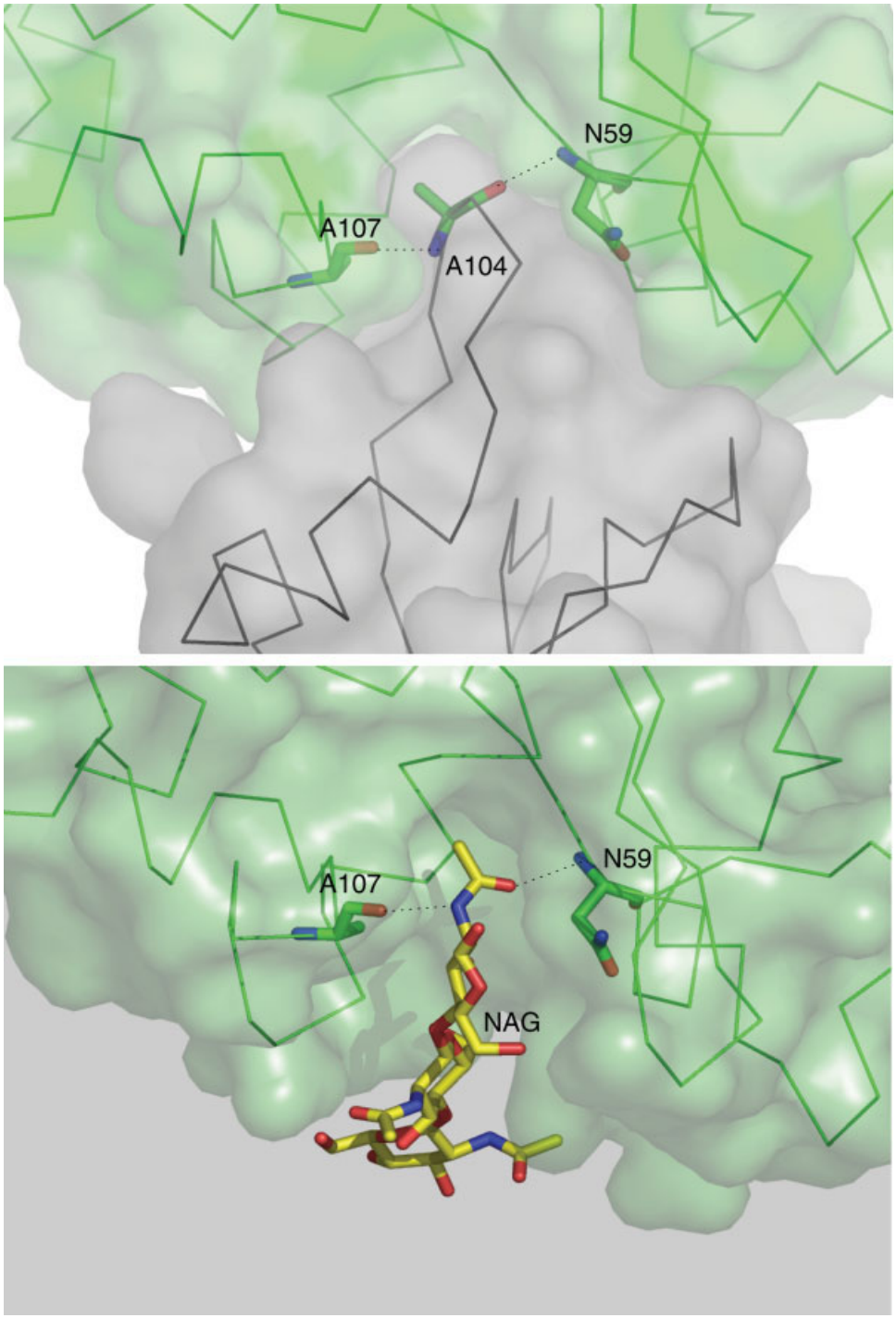

Fig. 9. Structural comparison of the lysozyme–cAb-Lys 3 complex and lysozyme-*N*-acetyl glucosamine (NAG) ring complex. The similarity between the hydrogen bond network formed between A107 and N59 of lysozyme and A104 of cAb-Lys3 (upper picture) and the network formed between A107 and N59 of lysozyme and NAG is shown. This network involves efficient amino acids of the lysozyme-cAb-Lys3 complex, and shows the energetic mimicry of the substrate (see text for details). [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

The two approaches nevertheless give convergent information i.e. when the free energy decomposition identifies a residue as making a large favorable contribution to binding, one can expect that mutating this residue to alanine will be detrimental to the binding affinity (unless the contribution is made by main chain atoms). Significant correlations between experimental free energy changes upon mutating a residue to alanine and free energy

decomposition based on MM/GBSA have indeed been reported.[13] In contrast, a computed unfavorable contribution to binding does not necessarily mean that mutating the residue to alanine will increase the binding affinity. As discussed earlier, charged amino acids such as Asp 52 and Arg 112 from lysozyme have a positive (unfavorable) contribution, yet most likely they have a structural role in stabilizing important residues from cAb-Lys3.

Efficient Amino Acids and Ligand Binding

A different, instructive way to identify amino acids that have an important role in binding is to compute the "amino acid efficiency," a quantity that we define here in analogy to ligand efficiency. Ligand efficiency is defined as the ratio of affinity to molecular size, a common measure of size being the number of nonhydrogen atoms. The notion of ligand efficiency has recently attracted a considerable interest in drug design studies,[62,72] particularly in the context of fragment based drug design. In fragment based drug design, the goal is to find small molecular building blocks that bind to a medicinal target. Subsequently, several building blocks can be linked or iteratively increased in size while keeping the molecular weight in the "druggable" range.[73] Small ligands with high binding efficiency are thus sought as a good starting point for lead identification and drug design.

By analogy with this concept, we identified "efficient amino acids" of the binding interface between cAb-Lys3 and lysozyme by dividing the free energy contribution of each amino acid by its number of nonhydrogen atoms (i.e. 4 for Gly, 5 for Ala, 14 for Trp etc.). An alternative definition of amino acid efficiency, based on the solvent accessible surface of the amino acids in the unbound proteins was also tested. The results are not significantly different from the simple definition mentioned earlier. (See Methods and Supplementary Material for details).

Amino acid efficiency can provide an assessment of the binding potential of a particular hot spot interaction. In drug design strategies aimed at protein–protein interfaces, the identification of efficient amino acids could help identify key interactions and prioritize the region of the protein–protein interface that could be targeted by small molecules. To gain physical insight in the origin of the amino acid efficiency, we divided each individual component of the binding free energy, as well as the total per residue free energy of binding, by the corresponding number of heavy atoms (see Fig. 8). The most efficient amino acids are Ala 104 and Ser 105 on cAb-Lys3 and Asn 59, Trp 62 and Ala 107 on lysozyme. In general, the efficiency of the amino acids on the antibody is better than on lysozyme, and the van der Waals contribution is the dominant energy term for efficient binding. A notable exception is Asn 59 from lysozyme, where electrostatic and van der Waals interactions contribute equally. Structurally, the ligand efficiency of Trp 62 from lysozyme is linked to its position as a central residue anchoring amino acids from CDR 1, 2, and 3 of cAb-Lys3. The three amino acids Ala 104 (cAb-Lys3), Asn 59, and Ala 107 (lysozyme) make a network of hydrogen bonds and hydrophobic contacts (see Figure 9) that make them particularly efficient. These interactions closely resemble[65] those made by lysozyme substrates, in particular the *N*-acetyl glucosamine ring (NAG) in subsite C of the binding pocket. The acetamido substituent of NAG in site C exhibits low temperature factor and a well defined electron density[74] and the saccharide in site C has been shown to make energetic contributions to binding that are more favorable than the saccharides bound to other sites within the lysozyme binding pocket.[75] Thus the cAb-Lys3 not only is a structural mimic of the lysozyme substrate, but it also exploits binding energy hot spots in a similar way as the substrate. This further shows that identifying binding hot spots or conversely efficient amino acids can pinpoint regions of the protein–protein interface that can be targeted by synthetic ligands.

CONCLUSIONS

In this work, we analyzed the binding interface between the single chain variable antibody domain (VHH) cAb-Lys3 and its antigen lysozyme, using a free energy decomposition based on Poisson–Boltzmann electrostatics, van der Waals interactions, and a hydrophobic surface area term. We considered conformational flexibility of the complex explicitly by using several experimental structures as starting points and performing short molecular dynamics simulations on each structure. We selected ten conformations from each molecular dynamics simulation that sample representative values of the Coulomb interaction energies between the proteins and performed the free energy decomposition on each conformation; the results represent averages over each individual calculation. This protocol was shown to be able to discriminate between consistently favorable interactions that give rise to binding hot spots on one side, and interactions that vary significantly with small changes in structures on the other side. The reliability of the free energy decomposition with respect to using a single crystal structure is improved, while avoiding the large computational cost of a free energy simulation method. The predictive ability of the procedure was subsequently tested by point mutations of amino acids of the interface; and it was shown that amino acids that have a contribution to binding when only one crystal structure is used, but that show large fluctuations when using the proposed conformational sampling protocol, can be mutated without influencing binding affinity.

The analysis of the complex has shown that electrostatic interactions do not systematically favor binding because of the compensation between protein desolvation and protein–protein interactions when forming macromolecular complexes. van der Waals interactions, as expected, systematically favor complex formation. This is in agreement with studies using similar methods on different complexes.[10,13,19,21,58] To identify amino acids that play a particularly important role in the protein–protein interface, we defined and computed the amino acid efficiency, i.e., the free energy contribution divided by the

number of heavy atoms in the amino acid. The analogous concept of ligand efficiency is commonly used in drug design, but to our knowledge, it has not been applied to protein–protein interactions. The efficient amino acids correspond to binding energy hot spots, and they originate mainly from efficient van der Waals packing of key residues on the interface. Electrostatic interactions make an important contribution to a binding energy hot spot of three neutral amino acids on the interface that form a network of hydrogen bonds, Asn 59 and Ala 107 of lysozyme, and Ala 104 of cAb-Lys3. The formation of networks of hydrogen bonds that stabilize protein–protein interactions has been noticed previously.[19] It shows that continuum electrostatics using point charges from a classical force field, despite its simplified description of the hydrogen bond,[76,77] can identify favorable electrostatic interactions in a context where, on average, protein–protein hydrogen bonds are not contributing or even unfavorable to binding.

Interestingly, the amino acids from cAb-Lys3 identified as efficient are structural mimics of the lysozyme saccharide substrate, and they interact with amino acids of lysozyme that are also identified as efficient. Our analysis thus establishes that the antibody not only is a structural mimic of the substrate,[65] but that it exploits key interactions that are consistent with those of a small molecule substrate. This indicates that the free energy decomposition method presented in this work can be used to guide antibody engineering, and that it can also serve as a tool for designing small molecule inhibitors of protein–protein interactions with potential therapeutic impact.

NOTE ADDED IN PROOF

In a recent study (Thanos, CD, DeLano WL, Wells, JA. Hot spot mimicry of a cytokine receptor by a small molecule Proc Natl Acad Sci USA 2006;103:15422–15427) the ligand efficiency of a protein-small molecule and protein-protein complex were compared. However, the ligand efficiency was defined as the total binding free energy divided by the number of contact atoms, which differs from the definition of amino acid efficiency as described here.

ACKNOWLEDGMENTS

We thank IDRIS (Institut du Développement et des Ressources en Informatique Scientifique) and CINES (Centre Informatique National de l'Enseignement Supérieur) for generous allowance of computer time, Drs Serge Muyldermans and M. Lauwereys for the kind gift of the cAbLys3-pHEN6 vector and Julien Gras and Elyette Martin for programming assistance.

REFERENCES

1. Arkin MR, Wells JA. Small-molecule inhibitors of protein-protein interactions: progressing towards the dream. Nat Rev Drug Discov 2004;3:301–317.
2. Pommier Y, Cherfils J. Interfacial inhibition of macromolecular interactions: nature's paradigm for drug discovery. Trends Pharmacol Sci 2005;26:138–145.
3. Fletcher S, Hamilton AD. Protein surface recognition and proteomimetics: mimics of protein surface structure and function. Curr Opin Chem Biol 2005;9:632–638.
4. Janin J. Principles of protein-protein recognition from structure to thermodynamics. Biochimie 1995;77:497–505.
5. Wells JA. Systematic mutational analyses of protein–protein interfaces. Methods Enzymol 1991;202:390–411.
6. Clackson T, Wells JA. A hot spot of binding energy in a hormone–receptor interface. Science 1995;267:383–386.
7. Bogan AA, Thorn KS. Anatomy of hot spots in protein interfaces. J Mol Biol 1998;280:1–9.
8. Cunningham BC, Wells JA. Comparison of a structural and a functional epitope. J Mol Biol 1993;234:554–563.
9. Honig B, Nicholls A. Classical electrostatics in biology and chemistry. Science 1995;268:1144–1149.
10. 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.
11. Kortemme T, Baker D. A simple physical model for binding energy hot spots in protein–protein complexes. Proc Natl Acad Sci USA 2002;99:14116–14121.
12. Guerois R, Nielsen JE, Serrano L. Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. J Mol Biol 2002;320:369–387.
13. 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.
14. Wong C, McCammon JA. Dynamics and design of enzymes and inhibitors. J Am Chem Soc 1986;108:3830–3832.
15. Kollman P. Free energy calculations: application to chemical and biochemical phenomena. Chem Rev 1993;93:2395–2417.
16. Simonson T, Archontis G, Karplus M. Free energy simulations come of age: protein-ligand recognition. Acc Chem Res 2002;35:430–437.
17. Rodinger T, Pomes R. Enhancing the accuracy, the efficiency and the scope of free energy simulations. Curr Opin Struct Biol 2005;15:164–170.
18. 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. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. Acc Chem Res 2000;33:889–897.
19. Sheinerman FB, Honig B. On the role of electrostatic interactions in the design of protein-protein interfaces. J Mol Biol 2002;318:161–177.
20. 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.
21. Zoete V, Meuwly M, Karplus M. Study of the insulin dimerization: binding free energy calculations and per-residue free energy decomposition. Proteins 2005;61:79–93.
22. Desmyter A, Transue TR, Ghahroudi MA, Thi MH, Poortmans F, Hamers R, Muyldermans S, Wyns L. Crystal structure of a camel single-domain VH antibody fragment in complex with lysozyme. Nat Struct Biol 1996;3:803–811.
23. Arbabi Ghahroudi M, Desmyter A, Wyns L, Hamers R, Muyldermans S. Selection and identification of single domain antibody fragments from camel heavy-chain antibodies. FEBS Lett 1997;414:521–526.
24. Hamers-Casterman C, Atarhouch T, Muyldermans S, Robinson G, Hamers C, Songa EB, Bendahman N, Hamers R. Naturally occurring antibodies devoid of light chains. Nature 1993;363:446–448.
25. De Genst E, Saerens D, Muyldermans S, Conrath K. Antibody repertoire development in camelids. Dev Comp Immunol 2006;30:187–198.
26. Conrath KE, Lauwereys M, Galleni M, Matagne A, Frere JM, Kinne J, Wyns L, Muyldermans S. β-Lactamase inhibitors derived from single-domain antibody fragments elicited in the camelidae. Antimicrob Agents Chemother 2001;45:2807–2812.
27. Tanha J, Xu P, Chen Z, Ni F, Kaplan H, Narang SA, MacKenzie CR. Optimal design features of camelized human single-domain antibody libraries. J Biol Chem 2001;276:24774–24780.
28. Jobling SA, Jarman C, Teh MM, Holmberg N, Blake C, Verhoeyen ME. Immunomodulation of enzyme function in plants by single-domain antibody fragments. Nat Biotechnol 2003;21:77–80.

29. Lauwereys M, Arbabi Ghahroudi M, Desmyter A, Kinne J, Holzer W, De Genst E, Wyns L, Muyldermans S. Potent enzyme inhibitors derived from dromedary heavy-chain antibodies. EMBO J 1998;17:3512–3520.
30. Decanniere K, Desmyter A, Lauwereys M, Ghahroudi MA, Muyldermans S, Wyns L. A single-domain antibody fragment in complex with RNase A: non-canonical loop structures and nanomolar affinity using two CDR loops. Structure Fold Des 1999;7:361–370.
31. Desmyter A, Spinelli S, Payan F, Lauwereys M, Wyns L, Muyldermans S, Cambillau C. Three camelid VHH domains in complex with porcine pancreatic α-amylase. Inhibition and versatility of binding topology. J Biol Chem 2002;277:23645–23650.
32. Cortez-Retamozo V, Backmann N, Senter PD, Wernery U, De Baetselier P, Muyldermans S, Revets H. Efficient cancer therapy with a nanobody-based conjugate. Cancer Res 2004;64:2853–2857.
33. Saerens D, Kinne J, Bosmans E, Wernery U, Muyldermans S, Conrath K. Single domain antibodies derived from dromedary lymph node and peripheral blood lymphocytes sensing conformational variants of prostate-specific antigen. J Biol Chem 2004;279: 51965–51972.
34. Stijlemans B, Conrath K, Cortez-Retamozo V, Van Xong H, Wyns L, Senter P, Revets H, De Baetselier P, Muyldermans S, Magez S. Efficient targeting of conserved cryptic epitopes of infectious agents by single domain antibodies. African trypanosomes as paradigm. J Biol Chem 2004;279:1256–1261.
35. Zarebski LM, Urrutia M, Goldbaum FA. Llama single domain antibodies as a tool for molecular mimicry. J Mol Biol 2005;349:814–824.
36. Loris R, Marianovsky I, Lah J, Laeremans T, Engelberg-Kulka H, Glaser G, Muyldermans S, Wyns L. Crystal structure of the intrinsically flexible addiction antidote MazE. J Biol Chem 2003; 278:28252–28257.
37. De Genst E, Silence K, Ghahroudi MA, Decanniere K, Loris R, Kinne J, Wyns L, Muyldermans S. Strong in vivo maturation compensates for structurally restricted H3 loops in antibody repertoires. J Biol Chem 2005;280:14114–14121.
38. De Genst E, Silence K, Decanniere K, Conrath K, Loris R, Kinne J, Muyldermans S, Wyns L. Molecular basis for the preferential cleft recognition by dromedary heavy-chain antibodies Proc Natl Acad Sci 2006;103:4586–4591.
39. Marquardt A, Muyldermans S, Przybylski M. A synthetic camel anti-lysozyme peptide antibody (Peptibody) with flexible loop structure identified by high-resolution affinity mass spectrometry. Chemistry 2006;12:1915–1923.
40. Vorobjev YN, Almagro JC, Hermans J. Discrimination between native and intentionally misfolded conformations of proteins: ES/IS, a new method for calculating conformational free energy that uses both dynamics simulations with an explicit solvent and an implicit solvent continuum model. Proteins 1998;32: 399–413.
41. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. Nucleic Acids Res 2000;28:235–242.
42. Decanniere K, Transue TR, Desmyter A, Maes D, Muyldermans S, Wyns L. Degenerate interfaces in antigen-antibody complexes. J Mol Biol 2001;313:473–478.
43. Brunger AT, Karplus M. Polar hydrogen positions in proteins: empirical energy placement and neutron diffraction comparison. Proteins 1988;4:148–156.
44. Brooks BR, Bruccoleri RE, 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.
45. 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.
46. Schaefer M, Vlijmen HWv, Karplus M. Electrostatic contributions to molecular free energies in solution. Adv Protein Chem 1998;51: 1–57.
47. Davis ME, Madura JD, Luty BA, McCammon JA. Electrostatics and diffusion of molecules in solution: simulations with the University of Houston Brownian Dynamics program. Comput Phys Commun 1991;62:187–197.
48. Nose S. A unified formulation of the constant temperature molecular dynamics methods. J Chem Phys 1984;81:511.
49. Hoover WG. Canonical dynamics: equilibrium phase-space distributions. Phys Rev A 1985;31:1695.
50. Ryckaert J-P, Ciccotti G, Berendsen HJC. Numerical integration of the cartesian equations of motion of a system with constraints. Molecular Dynamics of n-Alkanes. J Comp Phys 1977;23:327–341.
51. Fischer S, Smith JC, Verma CS. Dissecting the vibrational entropy change on protein/ligand binding: burial of a water molecule in BPTI. J Phys Chem B 2001;105:8050–8055.
52. Fischer S, Verma CS. Binding of buried structural water increases the flexibility of proteins. Proc Natl Acad Sci USA 1999; 96:9613–9615.
53. De Genst E, Areskoug D, Decanniere K, Muyldermans S, Andersson K. Kinetic and affinity predictions of a protein–protein interaction using multivariate experimental design. J Biol Chem 2002;277:29897–29907.
54. De Genst E, Handelberg F, Van Meirhaeghe A, Vynck S, Loris R, Wyns L, Muyldermans S. Chemical basis for the affinity maturation of a camel single domain antibody. J Biol Chem 2004;279: 53593–53601.
55. Gouda H, Kuntz ID, Case DA, Kollman PA. Free energy calculations for theophylline binding to an RNA aptamer: MM-PBSA and comparison of thermodynamic integration methods. Biopolymers 2003;68:16–34.
56. 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.
57. Dejaegere A, Liang X, Karplus M. Phosphate ester hydrolysis: calculation of gas-phase reaction paths and solvation effects. J Chem Soc Faraday Trans 1994;90:1763.
58. Green DF, Tidor B. Design of improved protein inhibitors of HIV-1 cell entry: optimization of electrostatic interactions at the binding interface. Proteins 2005;60:644–657.
59. Tsui V, Case DA. Calculations of the absolute free energies of binding between RNA and metal ions using molecular dynamics simulations and continuum electrostatics. J Phys Chem B 2001;105:11314–11325.
60. Lafont V. Analyse de la reconnaissance antigène-anticorps par modélisation moléculaire et mesures biophysiques en vue du développement de méthodes d'ingénierie rationnelle des protéines. Strasbourg: Louis Pasteur; 2004.
61. Spiegel MR. In:Schaum , editor. Probability and statistics. New York: McGraw Hill; 1975.
62. Hopkins AL, Groom CR, Alex A. Ligand efficiency: a useful metric for lead selection. Drug Discov Today 2004;9:430, 431.
63. Dejaegere A, Choulier L, Lafont V, De Genst E, Altschuh D. Variations in antigen-antibody association kinetics as a function of pH and salt concentration: a QSAR and molecular modeling study. Biochemistry 2005;44:14409–14418.
64. Karlsson R, Fagerstam L, Nilshans H, Persson B. Analysis of active antibody concentration. Separation of affinity and concentration parameters. J Immunol Methods 1993;166:75–84.
65. Transue TR, De Genst E, Ghahroudi MA, Wyns L, Muyldermans S. Camel single-domain antibody inhibits enzyme by mimicking carbohydrate substrate. Proteins 1998;32:515–522.
66. Nicholls A, Sharp KA, Honig B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins 1991;11:281–296.
67. Kabat EA, Wu TT, Perry H, Gottesman K, Foeller C. Sequences of proteins of immunological interest, 5th ed. 1991 NIH Publication No. 91–3242.
68. Janin J, Wodak SJ. Protein modules and protein–protein interaction. Introduction. Adv Protein Chem 2002;61:1–8.
69. DeLano WL. Unraveling hot spots in binding interfaces: progress and challenges. Curr Opin Struct Biol 2002;12:14–20.
70. 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.

71. Bashford D, Karplus M. pKa's of ionizable groups in proteins: atomic detail from a continuum electrostatic model. Biochemistry 1990;29:10219–10225.
72. Andrews PR, Craik DJ, Martin JL. Functional group contributions to drug–receptor interactions. J Med Chem 1984;27:1648–1657.
73. Lipinski CA. Drug-like properties and the causes of poor solubility and poor permeability. J Pharmacol Toxicol Methods 2000;44: 235–249.
74. Strynadka NC, James MN. Lysozyme revisited: crystallographic evidence for distortion of an *N*-acetylmuramic acid residue bound in site D. J Mol Biol 1991;220:401–424.
75. Chipman DM, Sharon N. Mechanism of lysozyme action. Science 1969;167:454–469.
76. Kortemme T, Morozov AV, Baker D. An orientation-dependent hydrogen bonding potential improves prediction of specificity and structure for proteins and protein–protein complexes. J Mol Biol 2003;326:1239–1259.
77. Morozov AV, Kortemme T, Tsemekhman K, Baker D. Close agreement between the orientation dependence of hydrogen bonds observed in protein structures and quantum mechanical calculations. Proc Natl Acad Sci USA 2004;101:6946–6951.