PROTEINS: Structure, Function, and Genetics 40:185–192 (2000)

Empirical Calculation of the Relative Free Energies of Peptide Binding to the Molecular Chaperone DnaK

Patrik Kasper, Philipp Christen and Heinz Gehring*
Biochemisches Institut der Universität Zürich, Zürich, Switzerland

ABSTRACT **We describe a methodology to calculate the relative free energies of protein-peptide complex formation. The interaction energy was decomposed into nonpolar, electrostatic and entropic contributions. A free energy–surface area relationship served to calculate the nonpolar free energy term. The electrostatic free energy was calculated with the finite difference Poisson-Boltzmann method and the entropic contribution was estimated from the loss in the conformational entropy of the peptide side chains.**

We applied this methodology to a series of DnaK·peptide complexes. On the basis of the single known crystal structure of the peptide-binding domain of DnaK with a bound heptapeptide, we modeled ten other DnaK·heptapeptide complexes with experimentally measured K_d values from 0.06 μM to 11 μM, using molecular dynamics to refine the structures of the complexes. Molecular dynamic trajectories, after equilibration, were used for calculating the energies with greater accuracy. The calculated relative binding free energies were compared with the experimentally determined free energies. Linear scaling of the calculated terms was applied to fit them to the experimental values. The calculated binding free energies were between −7.1 kcal/mol and − 9.4 kcal/mol with a correlation coefficient of 0.86. The calculated nonpolar contributions are mainly due to the central hydrophobic binding pocket of DnaK for three amino acid residues. Negative electrostatic fields generated by the protein increase the binding affinity for basic residues flanking the hydrophobic core of the peptide ligand. Analysis of the individual energy contributions indicated that the nonpolar contributions are predominant compared to the other energy terms even for peptides with low affinity and that inclusion of the change in conformational entropy of the peptide side chains does not improve the discriminative power of the calculation. The method seems to be useful for predicting relative binding energies of peptide ligands of DnaK and might be applicable to other protein-peptide systems, particularly if only the structure of one protein-ligand complex is available. Proteins 2000;40:185–192. © 2000 Wiley-Liss, Inc.

Key words: DnaK; protein-peptide interactions; energy calculation of peptide binding; Poisson-Boltzmann electrostatics; molecular dynamics simulations

INTRODUCTION

DnaK is a molecular chaperone; its main function is to prevent misfolding and aggregation of polypeptide chains under stress conditions.[1,2] The chaperone consists of two domains, a peptide-binding and an ATPase domain. Binding of a peptide ligand to and release from the peptide-binding domain is controlled by binding of ATP to the ATPase domain and its hydrolysis. On the basis of the crystal structure of the peptide-binding domain of DnaK complexed with a peptide[3] and a set of DnaK-binding peptides with experimentally determined K_d values which vary in charge or hydrophobicity, we tested whether empirically calculated energy contributions, i.e., nonpolar and electrostatic interactions as well as entropic effects, account with sufficient accuracy for the observed differences in binding affinities. A variety of empirical methods for the calculation of binding free energies have been developed and are still being improved.[4,5]

For our purpose, we chose a combined approach of molecular dynamics simulation, a free energy–surface area relationship, and a continuum solvent method. The single crystal structure of a DnaK·peptide complex served for modeling complexes with 11 different peptide ligands. Molecular dynamics simulations were then applied to refine the different modeled DnaK·peptide complexes. In a similar study on the binding of peptides to major histocompatibility complex (MHC) class I proteins, crystal structures of every complex were available.[4] We performed the molecular dynamics simulations not only for refining the modeled complexes but also used the dynamic trajectories, after equilibration, for averaging the calculated energies in order to improve the accuracy. The dynamic behavior of the free peptides was also taken into account. The change in the solvent-accessible surface upon binding yielded directly a value for the nonpolar contribution according to a linear free energy-surface area relationship.[6] A continuum solvent method, i.e., solving the Poisson-Boltzmann equation,[7] provided the Coulombic and solvation energy.

The peptides used in the calculations were—with one exception— heptapeptides and related in sequence. Due to the relatively stringent binding criteria of DnaK, substitu-

Grant sponsor: Swiss National Science Foundation; Grant number: 31-45940.95.

*Correspondence to: Prof. H. Gehring, Biochemisches Institut der Universiät Zürich, Winterthurerstr. 190, CH-8057 Zürich, Switzerland. E-mail: gehring@bioc.unizh.ch

Received 12 October 1999; Accepted 2 March 2000

tion of a single amino acid residue in a peptide ligand may markedly change the binding energy. The K_d values of the peptides that were modeled into the DnaK peptide-binding site ranged from 0.06 μM to 11 μM. The comparison of the calculated binding free energies correlated surprisingly well with the experimentally determined values.

METHODS

Theory

We decomposed the Gibbs free energy of binding into nonpolar, electrostatic and conformational entropy contributions.

$$\Delta G = \Delta G_{np} + \Delta G_{el} - T\Delta S_{conf}$$

The nonpolar contribution ΔG_{np} denotes the free energy arising from hydrophobic effects, the electrostatic free energy ΔG_{el} corresponds to the electrostatic interactions, and the conformational entropy ΔS_{conf} accounts for changes in conformational entropy of the solute.

The changes in the Gibbs free energy of binding can be calculated from the atomic structures of the two molecules undergoing the binding reaction.

$$\Delta G_{binding} = \Delta G_{complex} - \Delta G_{DnaK} - \Delta G_{peptide}$$

Calculation of ΔG_{np}

Estimates of free energy changes in binding reactions have shown that contact areas in protein-protein complexes are related to binding energies in a directly proportional way.[6]

$$\Delta G_{np} = \gamma_{aw} \cdot \Delta A$$

ΔA is the change in solvent-accessible surface and γ_{aw} is the microscopic surface tension associated with the transfer from liquid alkane to water. A change of the surface (ΔA) by 1 Å^2 corresponds approximately to 25 cal/mol.[6,8] The algorithm of Lee and Richards[9] was used to calculate the solvent-accessible surface.

The change in solvent-accessible surface can be calculated as the difference of the surface area of the free and the complexed molecules. Van der Waals interactions are not being considered explicitly.

$$\Delta A = A_{complex} - A_{DnaK} - A_{peptide}$$

Calculation of ΔG_{el}

An important contribution to ligand binding in aqueous solution is the electrostatic interaction which includes Coulombic interactions and solvation effects.

$$\Delta G_{el} = \Delta G_{coul} + \Delta G_{solv}$$

Hydrogen bonds are considered to be implicitly included in this energy. In Figure 2 the entire thermodynamic process is shown which can be used to calculate the total electrostatic energy.[10] In the initial state both molecules are fully solvated and infinitely separated from each other. The first step describes the partial desolvation of the molecules by replacing the solvent molecules by a medium of dielectric constant ϵ_m which corresponds to that of the protein interior. The second step describes the interaction energy ($\Delta\Delta G_{inter}$) between the charges of both molecules and the change in solvation energy. The interaction energy was calculated with the program Delphi that solves the Poisson equation making a finite difference approximation[7,11] which assigns molecules to a three-dimensional grid. The atomic point charges are then distributed to the grid points nearest to their actual positions. A dielectric constant ϵ is assigned to each grid point, depending on whether the grid point is located in the solvent or in the solute ($\epsilon = 4$ in the solute, $\epsilon = 80$ in the solvent). The field equations are solved for the values of the electrostatic potential at each grid point by numerical iterations.

Calculation of ΔS

ΔS_{sc} is calculated using the empirical scale of Pickett and Sternberg.[12] Their model assumes that a solvent-exposed side chain with a relative accessibility (RA) larger than 60% is allowed to rotate freely.

$$RA = \frac{A_{side\ chain}}{A_{side\ chain\ in\ fully\ extended\ state}}$$

where A is the solvent-accessible surface. These authors give the entropies for all 20 amino acid side chains derived from the observed distributions of side-chain rotamers in 50 non-homologous protein crystal structures. The entropy changes were calculated for the side chains of the peptides by the calculation of the relative accessibilities in both the bound and unbound states.

The Structures Used In This Study

Three structures would be needed to calculate the free energy of binding, i.e., the structure of the DnaK·peptide complex, the unliganded DnaK and the peptide alone (Fig. 2). The only X-ray crystallographic structure available for DnaK is that of the peptide-binding domain of DnaK (Fig. 1A) complexed with the heptapeptide NRL (Table I) which was determined at a resolution of 2.0 Å (pdb1dkx.ent in the Brookhaven protein database[3]). The structures of the solvated free peptides were obtained by a molecular dynamics simulation in explicit water (see below).

The peptide-binding site is located in the β-subdomain, but it is also delimited by the α-helical subdomain (Fig. 1A). Although the α-helical subdomain has no direct contact with the peptide, it significantly influences the electrostatic field in the peptide-binding region. This effect became evident in preliminary calculations of the electrostatic field from the DnaK·peptide complex. Upon binding or release of a peptide, the two subdomains must move apart to let the peptide in or out. The movement might cause a completely different orientation or even multiple orientations of the two subdomains in the uncomplexed form. Perhaps it is for this reason that no crystal structure of the uncomplexed form is available.

Modeling and Molecular Dynamics Simulation

The calculation of the binding energies was performed on a total of 11 different peptides. We modeled the

structures of the peptide complexes on the basis of the crystal structure by replacement of the amino acid side chains in the NRL peptide using the program Insight II (Biosym/MSI) and maintaining the dihedral angles. These structures were refined subsequently by energy minimizations and molecular dynamics simulations in an explicit shell of 30 Å of water. For the molecular dynamics simulations we used the program Discover (Biosym/MSI) with the consistent valence force-field (cvff) implemented on a SGI (Origin) workstation. Instead of a cutoff for the nonbound interactions, we used the cell multipole method.

First we minimized all modeled structures for 2,000 steps while keeping the backbone of the entire protein fixed. Then the backbone was released and again 2,000 steps of minimization were applied. Subsequently, we performed a molecular dynamics simulation at 300 K for 2000 fs with the backbone kept fixed, followed by 2,000 fs in which only the atoms which were closer than 11 Å to the peptide were allowed to move. After this initialization, the molecular dynamics simulation was continued for 290 ps for all systems. Analysis of the trajectories (one structure every ps) showed in all cases little deviation from the initial structure of the protein. The backbone position of the peptides was always very close to that of peptide NRL in the crystal structure[3] but the peptide side chains relaxed into slightly different positions.

Parameter Settings

To calculate the nonpolar binding energy we used the program Naccess[13] which calculates the solvent-accessible surface. The surface areas were calculated for the protein with and without the peptide bound. For the surfaces of the peptides, we performed separate molecular dynamics simulations with the peptides alone in a sphere of water. These simulations were performed with the same parameter setting as the simulations of the protein-peptide complexes. The solvent-accessible surfaces were then calculated for the last 50 structures of the molecular dynamics simulation (= 50 ps) and averaged. For the atomic radii we used the Van der Waals values and for the probe radius of water the commonly used value of 1.4 Å. The atomic charges used with the Delphi program were taken from the cff91 force field, implemented in the Insight II package (Biosym/MSI). They yielded results in electrostatic calculations that were comparable to those obtained with the PARSE parameter set.[14]

A dielectric constant of 80 was used for the water and 4 for the interior of the protein and peptides. Again a probe radius of 1.4 Å was used for the calculation of the surface. After calculation of a grid of 65 × 65 × 65 points, a focussing grid with the same number of points was calculated which enclosed the actual binding site. The final grid spacing was 0.52 Å.

Experimental Determination of the Dissociation Constants K_d

For determination of the dissociation equilibrium constants K_d of the peptides NRL, NRA, NRG, NDL, and RLR (Table I) we used mutant DnaK H541C (Feifel B., Schönfeld H.-J. and Christen P., in preparation). Residue 541 is located in the α-helical subdomain in close proximity of the peptide-binding site but does not interact directly with the bound peptide. The sulfhydryl group of the newly introduced cysteine residue was covalently labeled with the fluorophor MIANS (2-(4′-maleimidylanilino)naphtalene-6-sulfonic acid). Binding of a peptide to this labeled DnaK mutant resulted in a decrease of the fluorescence emitted by MIANS. Addition of increasing concentrations of peptide to the labeled DnaK mutant allowed to determine the K_d values of the peptides.

RESULTS

Choice of Peptide Ligands

The peptide-binding domain of DnaK consists of a β-sandwich and an α-helical subdomain (Fig. 1A). Peptide NRL (see Table I for amino acid sequence) interacts in its extended conformation with the loops of the β-sandwich subdomain. The α-helical subdomain forms a lid over the binding site without direct contacts to the peptide. The binding site has a large hydrophobic pocket in its middle which accommodates the central hydrophobic leucine triplet of peptide NRL (positions 3 to 5). The mainly hydrophilic amino acids at positions 1, 2 and 6, 7 of the peptide are partially in contact with the solvent. For the calculation of the binding energies, heptapeptides and the hexapeptide KWH were selected which corresponded in some cases to segments of larger peptides with known binding characteristics (Table I). For the calculations, interactions of a seven amino-acid stretch were assumed to be responsible for binding.[3,15] The heptapeptides differed either in one residue of the central hydrophobic cluster and/or in the hydrophilic and charged residues in their flanking regions. The relative binding properties of ten of the total 11 peptides had been reported previously, and covered with their K_d values a range of two orders of magnitude. Some estimated K_d values were re-determined with higher accuracy (see Methods).

The peptide SRL has the highest binding affinity of the examined peptides (Table I). It contains the binding motif RLLL, that is also present in the high affinity peptides RLR and NRL. In the case of NRA, the exchange of Leu 4 to Ala lowers as expected the affinity. Peptide NRG has a Gly instead of Ala at the Leu position 4. The shift of the positive charge in the peptide RSL from position 2 to position 1 lowers the affinity significantly, although the leucine triplet at positions 3–5 is maintained. Peptide RLQ has, similarly to RSL, the positive charge shifted from position 2 to 1. In addition, Leu at position 5 is changed to Gln, a less hydrophobic residue though of similar size. In the case of peptide KFF, two positive charges are located at positions 1 and 2 and both Leu 4 and 5 were replaced by Phe, resulting in a slightly decreased affinity. The completely changed hydrophobic core of peptide KWH to Trp-Val-His and the vacant position 1 result in a relatively high K_d value of 5 μM, despite the positive charge of Lys at position 2. The peptide with the lowest affinity used in this study is NDL, Asp instead of Arg at position 2 being the only difference to NRL. This charge reversal may be

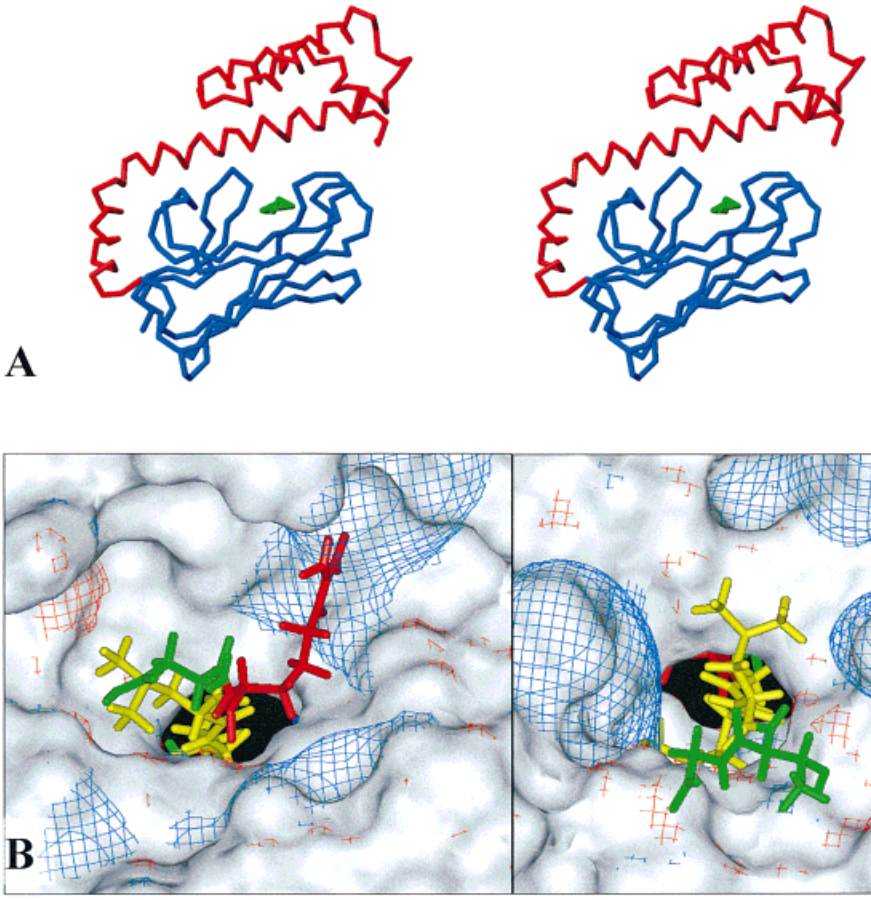

Fig. 1. **A:** Stereo-view of the Cα backbone of DnaK[3], projected along the peptide ligand backbone. The peptide (in green) which is bound by the β-subdomain (in blue) is blocked by the α-subdomain (in red) forming a lid atop of the binding groove. For binding or release of the peptide, the α-subdomain has to move away. **B:** The electrostatic potential of the peptide-binding site with peptide NRL[3], displayed as 3D contour over the solvent-accessible surface. Negative potentials are colored in blue, positive in red. The left panel shows the view of the binding site from the NH_2-terminus of the peptide; the right panel from its COOH-terminus. The positively charged Arg (red) is located on the surface of DnaK in a strong negative field (blue). A strong negative field is also present on the other side. The central Leu-triplet is shown in yellow. The NH_2-terminal Asn and the COOH-terminal Thr and Gly are displayed in green.

expected to result in a substantial change in Coulombic interactions. Peptide RLR was designed on the basis of preliminary results with the intention to synthesize a ligand with even higher affinity than SRL or NRL. In addition to the Leu triplet it has positive charges both at positions 2 and 6. Peptide YQK has a positive charge at position 6 and Gln-Leu-Ala as the hydrophobic core.

Calculations

To compute the free energy change of the binding reaction, the energy value for the full thermodynamic process would have to be calculated. However, for comparing the binding energies of different peptides bound to the same acceptor only the relative free energy changes which arise directly from the peptides were calculated. Contributions of the unliganded DnaK, the structure of which is the same in all cases, cancel out and were neglected.

Nonpolar binding free energy

The contribution of ΔG_{np} is proportional to the gain in buried surface of the molecules participating in the binding process (see Methods). The value of ΔG_{np} was calculated from the solvent-accessible surface areas of the peptides (ΔA) that become buried upon binding to DnaK (Table II). A microscopic surface tension of 25 cal mol^{-1} $Å^{-2}$ was used.[6] In order to obtain statistically representative surface differences, the free peptides and the modeled DnaK·peptide complexes (see Methods) were subjected to molecular dynamics simulations in water. The surfaces were then calculated for each of the last 50 frames of the simulation of the free peptide (50 ps of total 150 ps) or of the DnaK·peptide complex (50 ps of total 290 ps) and averaged. ΔA of the unliganded protein was not calculated because no structure was available.

A tendency for a significantly larger ΔA is evident in the case of those peptides, which have a completely buried Leu at position 4 in the binding site. An interesting exception is peptide NDL that has a relatively low ΔA despite its three central leucine residues. Most likely a negative charge effect in a negative field (see below) counteracts the hydrophobic interaction.

Electrostatic energy

The electrostatic energy contributions ΔG_{coul} and ΔG_{solv} were calculated according to the thermodynamic process (Fig. 2) as described under Methods, neglecting, however, the contributions of the first step. It turned out that $\Delta\Delta G_{inter}$, which includes Coulombic and solvation energies, was sufficient to describe the relative binding energies because contributions from the unknown unliganded structure of DnaK will cancel out. From each of the last 15 structures (15 ps) of the simulated liganded structure the

TABLE I. Amino Acid Sequences and Experimentally Determined K_d Values of the Peptides Used in the Study†

Peptide	Sequence[a]	K_d (exp) (μM)	ΔG_{exp} (kcal/mol)
SRL	. . LQ**SRLLLSA**PR . .	0.06[b] (±0.02)	−9.8
RLR	**NRLLLRG**	0.1[c]	−9.5
NRL	**NRLLLTG**	0.2[d] (±0.02)	−9.1
NRA	**NRLALTG**	0.5[e] (±0.1)	−8.6
NRG	**NRLGLTG**	0.8[c] (±0.5)	−8.3
RSL	CA**RSLLLSS**	0.9[d] (±0.3)	−8.2
YQK	**FYQLAKT**CPV	1.0[f]	−8.2
RLQ	**RALLQSC**	1.4[b] (±0.4)	−8.0
KFF	. . AQ**RKLFFNL**RK . .	3.8[f]	−7.4
KWH	**KWVHLF**G	5.0[e] (±3.0)	−7.2
NDL	**NDLLLTG**	11.0[c]	−6.8

†The experimental binding energies ΔG_{exp} were calculated using the relation $\Delta G = RT \ln K_d$, where K_d is the dissociation constant. For peptide SRL, the experimental K_d value had been measured with a larger, 22-residue peptide representing the prepeptide of mitochondrial aspartate aminotransferase.[16] The K_d value of peptide NRG was determined using a fluorescence-labeled DnaK mutant (see Methods). For comparison, the same method was applied also to peptide NRA, the K_d value of which was known from literature. It turned out that peptide NRA binds to the mutant DnaK with half the affinity of wild-type DnaK. For this reason we corrected the measured apparent K_d value for NRG from 1.6 μM to 0.8 μM. The measured dissociation constant of the peptide NDL was found to be comparable to estimations based on previous reports.[18] The K_d value of peptide RLR was also determined with the fluorescence-labeled DnaK mutant, as was, for comparison, the K_d value of the very similar peptide NRL. Peptide RLR was found to bind indeed 2-fold better (apparent K_d 0.4 μM) than NRL (apparent K_d 0.8 μM). Since wild-type DnaK has a 4-fold higher affinity towards the high-affinity ligand NRL, we estimated for peptide RLR a K_d value of 0.1 μM. The dissociation constant for the peptide KFF had been determined with a much larger 24-residue peptide (J. Reinstein, personal communication).

[a]The K_d values were measured with the complete peptides. For the computational analysis, only the residues in bold were considered.

[b]Measured with fluorescence-labeled peptide.[16]

[c]Measured with fluorescence-labeled DnaK mutant (see Methods).

[d]Measured with fluorescence-labeled peptide.[17]

[e]From Gragerov et al.[18]

[f]J. Reinstein, personal communication.

TABLE II. Calculated Solvent-Accessible Surface Changes ΔA_{pep} and the Unscaled Nonpolar, Electrostatic and Entropic Energy Contributions†

			ΔG_{np}	ΔG_{coul}	ΔG_{solv}	$-T\Delta S_{sc}$
Peptide	Sequence	ΔA_{pep} (Å^2)	(kcal/mol)			
SRL	**SRLLLSA**	715 (60)	−17.9 (1.5)	−28.9 (2.7)	2.8 (2.4)	−6.1
RLR	**NRLLLRG**	766 (64)	−19.2 (1.6)	−117.6 (3.8)	59.8 (2.8)	−8.0
NRL	**NRLLLTG**	704 (44)	−17.6 (1.1)	−29.5 (2.9)	2.7 (2.4)	−7.6
NRA	**NRLALTG**	689 (27)	−17.2 (0.7)	−28.2 (5.1)	8.0 (3.7)	−5.2
NRG	**NRLGLTG**	653 (34)	−16.3 (0.9)	−29.2 (3.1)	5.2 (2.8)	−6.8
RSL	**RSLLLSS**	694 (35)	−17.3 (0.9)	−20.7 (6.2)	2.4 (3.9)	−7.8
YQK	**FYQLAKT**	795 (71)	−19.9 (1.8)	−96.5 (5.3)	48.2 (4.0)	−8.0
RLQ	**RALLQSC**	702 (43)	−17.5 (1.1)	−75.2 (4.3)	35.5 (2.7)	−7.4
KFF	**RKLFFNL**	647 (32)	−16.2 (0.8)	−141.3 (3.4)	73.9 (2.6)	−7.5
KWH	**.KWVHLF**	633 (35)	−15.8 (0.9)	−18.1 (4.7)	1.1 (1.6)	−3.8
NDL	**NDLLLTG**	676 (29)	−16.9 (0.7)	78.7 (3.2)	−54.0 (2.7)	−5.2

†ΔG_{np} is the unscaled nonpolar binding free energy corresponding to ΔA_{pep}, averaged over the last 50 structures of the dynamic trajectory (50 ps). The standard deviations are given in parentheses. ΔG_{coul} and ΔG_{solv} are the electrostatic energy contributions. They have been calculated and averaged over the last 15 structures of the dynamics trajectory. $T\Delta S_{SC}$ is the entropic energy contribution from the peptide side chains. It has been averaged over the last 50 structures as in the case of ΔG_{np}.

energies were calculated with the program Delphi which numerically solved the Poisson-Boltzmann equation.[7] Subsequently, the calculated energies were averaged to eliminate possible artifacts arising from grid effects.

A listing of the averaged Coulombic and solvation values of the electrostatic interaction energies (Table II) shows that the most negative Coulombic contributions resulted for the peptides RLR and KFF, which both have two positive charges, RLR one on both sides, KFF two on the N-terminal side. The most positive Coulombic contribution was found with peptide NDL with its negative charge at the N-terminus. All other peptides with just one positive

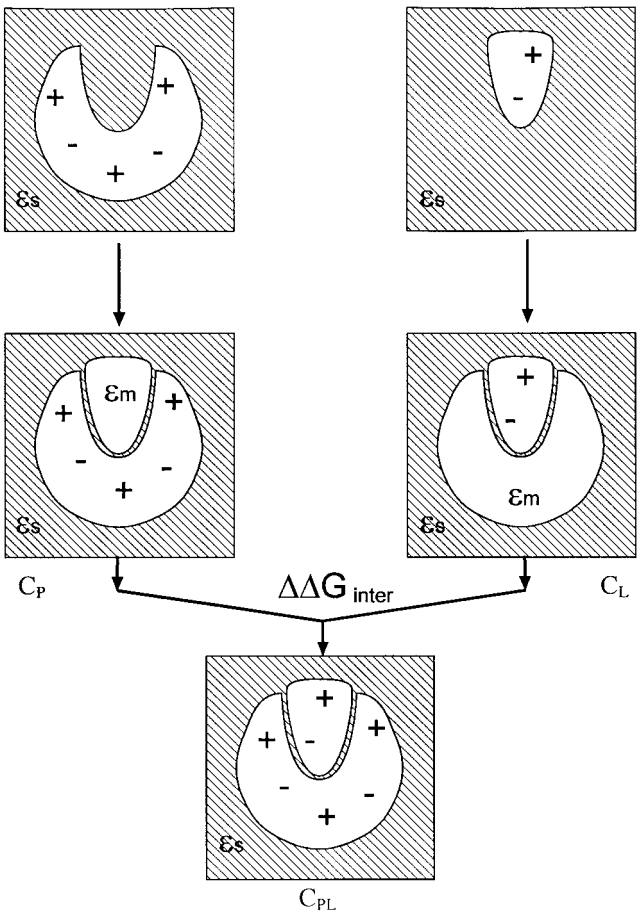

Fig. 2. Thermodynamic process for calculating the change in total electrostatic energy ΔG_{el} upon intermolecular binding. The scheme is adapted from Gilson and Honig.[10] The first step describes the change in the charge-solvent interaction energy of the protein and the ligand upon binding. ϵ_m is the dielectric constant of the molecular interior, ϵ_s is the dielectric constant of the solvent water. The second step describes the interaction energy between the charges of the two molecules : $\Delta\Delta G_{inter} = \Delta G_{el}(C_{PL}) - \Delta G_{el}(C_P) - \Delta G_{el}(C_L)$.

charge exhibit intermediary ΔG_{coul} values. The correlation of the Coulombic contribution with the positive charge of the peptide reflects the existence of a protein-generated negative field at both sides of the central hydrophobic binding pocket (Fig. 1B).

Entropic contributions

As in the calculations of ΔG_{np} and ΔG_{el}, only the entropy changes of the peptides were calculated, assuming that the differences in entropy between free DnaK and liganded DnaK is essentially the same for all peptides. The entropy change of the peptide upon binding to DnaK can be calculated from the change in degrees of freedom of the side chains (Table II). The term $-T\Delta S_{sc}$ was calculated from the relative solvent accessibility of the side chains as described under Methods. For the calculation of the solvent-accessible surfaces, the side chain accessibilities of the last 50 structures of the molecular dynamics simulations were averaged. The entropic contribution was found to depend strongly on the behavior of the free peptide in solution.

Scaling

Because only the relative energies but not the absolute free energies were calculated, the calculated energy terms had to be scaled to account for neglected constant contributions in the thermodynamic process (see Methods). In a first linear fitting, the individual energy contributions were scaled applying the leave-one-out method on the set of all 11 peptides. All scaling calculations were performed with the program Excel using the Solver module. Initially, we used five scaling parameters for which the following values were obtained: 0.46 ± 0.11 (ΔG_{np}), 0.17 ± 0.02 (ΔG_{coul}), 0.28 ± 0.03 (ΔG_{solv}), 0.13 ± 0.09 ($-T\Delta S_{sc}$) and 0.20 ± 0.17 (constant term). The large fluctuations of the scaling parameters of the nonpolar, entropic and constant terms indicated an overfitting. The correlation coefficient achieved was 0.88 with an average absolute error of 0.38 kcal mol^{-1}. A second fitting was executed with only four parameters leaving out the constant term. The resulting values were 0.35 ± 0.04 (ΔG_{np}), 0.16 ± 0.02 (ΔG_{coul}), 0.27 ± 0.03 (ΔG_{solv}) 0.10 ± 0.09 ($-T\Delta S_{sc}$). The correlation coefficient was virtually the same (0.87) as in the 5-parameter fitting with an average absolute error of 0.39 kcal mol^{-1}. The fluctuations of the nonpolar scaling parameter decreased to 10% whereas the fluctuations of the entropic term were increased to 88%. The very high fluctuations in the entropic scaling term and the fact that the entropic term after scaling contributed only little to the total binding energy suggested a further reduction of the number of fitting parameters to only three by leaving out the entropic term $-T\Delta S_{sc}$. An almost equally good correlation coefficient of 0.86 and an average absolute error of 0.40 kcal mol^{-1} resulted from this fitting with three parameters (Table III and Fig. 3). The final scaling parameters were 0.32 ± 0.02 (ΔG_{np}), 0.15 ± 0.02 (ΔG_{coul}) and 0.26 ± 0.03 (ΔG_{solv}). The low fluctuations and the still good correlation coefficient indicated that the peptide binding energy could be adequately described by the nonpolar (ΔG_{np}) and electrostatic (ΔG_{coul} and ΔG_{solv}) contributions.

DISCUSSION

In this study on the binding of peptides to a protein, we used a physical model which attributes energies to structures on the atomic level. Surface area calculations for the nonpolar contributions, continuum solvent calculations, and estimates of conformational entropy were applied. The goal was to probe the strengths and limits of the model in calculating relative binding free energies rather than to calculate absolute free binding energies. We applied the calculations to complexes of the molecular chaperone DnaK with peptide ligands. The choice of this system was due to inherent interest but also to the challenge of calculating the binding energies of different peptides with known binding affinities to a specific binding site without having available, however, the experimentally determined 3D structure of each individual complex.

The results indicate that it suffices to calculate the energy contributions of the DnaK·peptide complex and of the free peptide. The other energy terms, i.e., the terms of the unliganded DnaK, appear to be independent of the

TABLE III. Experimental ΔG_{exp} and Calculated ΔG_{calc} Binding Free Energies and the Individual Energy Contributions†

Peptide	Sequence	ΔG_{np}	ΔG_{coul}	ΔG_{solv}	ΔG_{el}	ΔG_{calc}	ΔG_{exp}
		(kcal/mol)					
SRL	**SRLLLSA**	−5.7	−4.5	0.7	−3.7	−9.4	−9.8
RLR	**NRLLLRG**	−6.1	−18.1	15.4	−2.7	−8.8	−9.5
NRL	**NRLLLTG**	−5.6	−4.5	0.7	−3.8	−9.4	−9.1
NRA	**NRLALTG**	−5.5	−4.4	2.1	−2.3	−7.8	−8.6
NRG	**NRLGLTG**	−5.2	−4.5	1.3	−3.2	−8.4	−8.3
RSL	**RSLLLSS**	−5.5	−3.2	0.6	−2.6	−8.1	−8.2
YQK	**FYQLAKT**	−6.3	−14.9	12.4	−2.5	−8.8	−8.2
RLQ	**RALLQSC**	−5.6	−11.6	9.2	−2.5	−8.0	−8.0
KFF	**RKLFFNL**	−5.2	−21.8	19.0	−2.8	−7.9	−7.4
KWH	**.KWVHLF**	−5.0	−2.8	0.3	−2.5	−7.5	−7.2
NDL	**NDLLLTG**	−5.4	12.1	−13.9	−1.8	−7.1	−6.8

†Each energy term is scaled using the factors obtained from the linear fitting. The scaling factors are 0.32 for ΔG_{np}, 0.15 for ΔG_{coul} and 0.26 for ΔG_{solv}. ΔG_{el} is the sum of the scaled ΔG_{coul} and ΔG_{solv}. ΔG_{calc} is the sum of ΔG_{np} and ΔG_{el}. ΔG_{exp} is the experimentally determined binding free energy.

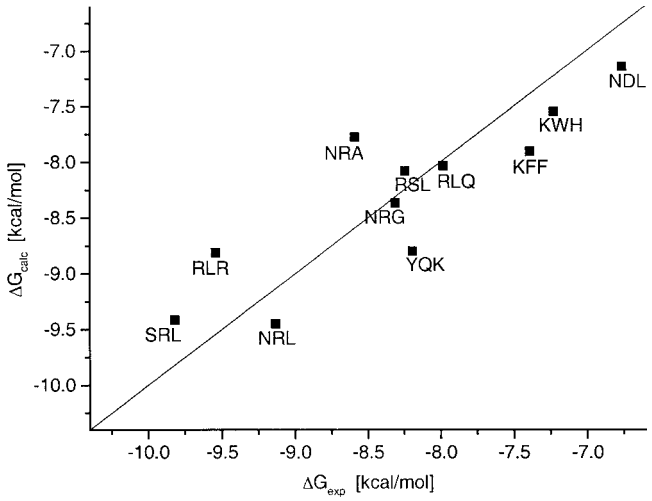

Fig. 3. Correlation of the calculated ΔG_{calc} and the experimental binding free energies ΔG_{exp}. Three fitting parameters were used and the correlation coefficient was 0.86 with an average absolute error of 0.40 kcal mol^{-1}. The values are taken from Table III.

peptide and cancel out for the calculation of the relative energies. Furthermore, the contributions from the loss in translational, rotational, and backbone entropy were assumed to cancel out and were also neglected. Also not accounted for is the strain in the modeled structures. An indication for low strain in the structures are the small deviations from the starting structures upon energy minimization and molecular dynamics simulation. We found only minor rearrangements of some side chains. The deviations would have been significantly larger if the peptides had been modeled in an unsuitable way. The energies were calculated from the last frames of the dynamic trajectory of each modeled complex. These energy values were then averaged in order to diminish grid artifacts in the electrostatic calculations and to include the dynamic behavior of the complexes in the calculations.

In the calculations we have restricted the length of the peptides to seven amino acid residues which corresponds with the size of the peptide-binding site.[3,15] The lack of one N-terminal residue in peptide KWH (Table II) can be assumed to contribute only negligibly to the loss in backbone entropy; the contribution from both peripheral residues was estimated to be small because their mobility is much higher than that of the more centrally located residues. Potential contributions to the binding energy from residues outside the peptide-binding site that encompasses seven amino acid residues could not be considered and perhaps have prevented an even better correlation (see below). An example of an electrostatic contribution by residues outside the binding site is provided by the two peptides with the sequences **FYQLAKT**CPV and **FYQLAKT**CRV, the experimentally determined dissociation constants of which are 1.0 μM and 5.4 μM, respectively (J. Reinstein, personal communication). For the calculation of the binding energy only the putative seven-residue binding sequence **FYQLAKT** was considered. The additional C-terminal positive charge of the arginine residue substituting proline obviously diminishes the binding affinity.

After scaling, the calculated binding energies ΔG_{calc} in the range from −7.1 kcal mol^{-1} to −9.4 kcal mol^{-1} (Table III) correlated well with the experimental data (Fig. 3). The ΔG_{calc} values of all eleven peptides were fitted using the leave-one-out method with three fitting parameters, not considering the constant and the entropic term (see Results). The Coulombic and solvation energy contributions had to be scaled down to 15% and 26%, respectively. The nonpolar contribution had also to be scaled down to 32% if using a microscopic surface tension γ_{aw} of 25 cal mol^{-1} Å^{-2}.[6] The overestimations suggest the existence of neglected terms as for example entropic terms, or terms from the thermodynamic cycle caused by the missing unliganded structure. Another possible explanation is that the assumed value of the dielectric constant of 4 at the interface and in the interior of the solute might be inappropriate. In a similar study, linear fitting (without applying molecular dynamics simulations and the leave-one-out method) has been used to scale the calculated free binding

energy contributions (nine mutant variants of a protein, after excluding one data point, were considered) with a scaling factor of 0.6 for the nonpolar, 0.4 for the electrostatic and 1.3 for the entropic terms and a correlation coefficient of 0.8.[5]

The nonpolar energy ΔG_{np} contributed most to the binding energy of the peptides as found for other peptide-binding systems.[4] However, one should keep in mind that ΔG_{np} and ΔG_{el} values are not independent from each other as for example an electrostatic repulsion might reduce the contact area and thus ΔG_{np}. The nonpolar contribution ΔG_{np} is mostly dependent on the three central hydrophobic residues. The lowest contribution is achieved in the case of the peptides NRG, KFF, and KWH. They all have small (NRG) or not optimally fitting bulky (KWH and KFF) residues in the central positions 3 to 5 that reduce the hydrophobic contacts (Table II). The Coulombic energy ΔG_{coul} and the solvation energy ΔG_{solv} are strongly dependent on the charges of the amino acid residues in positions 1,2 and 6,7 (Fig. 1B, Table III). The large effects are opposite and mostly compensate each other. Although it is not easy to interpret the values in detail, the single negatively charged peptide NDL has indeed the least ΔG_{el} value (the sum of the scaled ΔG_{coul} and ΔG_{solv} contributions, Table III). The scaled conformational side chain entropies are relatively small compared to the other terms (scaling factor 0.1, see Results). They do not improve the predictive power of the present method as is evident from the only small increase of the correlation coefficient in the 4-parameter fitting and high fluctuations of this parameter in the leave-one-out method (see Results).

In conclusion, the present procedure for calculating the relative binding free energy is relatively simple, provides reasonable results for the DnaK·peptide system, and is capable of treating both specific charge effects and changes in the buried hydrophobic surface. The method might be applicable to other peptide-binding systems with only one liganded crystal structure of the protein available if the ligands in question are similar among themselves and if the complex can be modeled without extensive changes in conformation of the binding site.

ACKNOWLEDGMENTS

We thank Amedeo Caflisch for helpful discussion and Jochen Reinstein for providing the dissociation constants of the two peptides YQK and KFF.

REFERENCES

1. Ellis RJ, van der Vies SM. Molecular chaperones. Annu Rev Biochem 1991;60:321–347.
2. Hartl FU. Molecular chaperones in cellular protein folding. Nature 1996;381: 571–579.
3. Zhu X, Zhao X, Burkholder WF, et al. Structural analysis of substrate binding by the molecular chaperone DnaK. Science 1996;272:1606–1614
4. Froloff N, Windemuth A, Honig B. On the calculation of binding free energies using continuum methods: application to MHC class I protein-peptide interactions. Protein Sci 1997;6:1293–1301.
5. Novotny J, Bruccoleri RE, Davis M, Sharp KA. Empirical free energy calculations: a blind test and further improvements to the method. J Mol Biol 1997;268:401–411.
6. Chothia C. Hydrophobic bonding and accessible surface area in proteins. Nature 1974;248:338–339.
7. Nicholls A, Honig B. A rapid finite-difference algorithm, utilizing successive over-relaxation to solve the Poisson-Boltzmann equation. J Comput Chem 1991;12:435–445.
8. Novotny J, Bruccoleri RE, Saul FA. On the attribution of binding energy in antigen-antibody complexes McPC 603, D1.3, and HyHEL-5. Biochemistry 1989; 28:4735–4749.
9. Lee B, Richards FM. The interpretation of protein structures: estimation of static accessibility. J Mol Biol 1971; 55:379–400.
10. Gilson MK, Honig B. Calculation of the total electrostatic energy of a macromolecular system: solvation energies, binding energies, and conformational analysis. Proteins 1988;4:7–18.
11. Harvey SC. Treatment of electrostatic effects in macromolecular modeling. Proteins 1989;5:78–92.
12. Pickett SD, Sternberg MJ. Empirical scale of side-chain conformational entropy in protein folding. J Mol Biol 1993;231:825–839.
13. Hubbard SJ, Thornton JM. 'NACCESS', Computer Program. Department of Biochemistry and Molecular Biology, University College London.1993.
14. Sitkoff D, Sharp KA, Honig B. Accurate calculation of hydration free-energies using macroscopic solvent models. J Phys Chem 1994;98:1978–1988.
15. Flynn G, Pohl J, Flocco T, Rothman J. Peptide-binding specificity of the molecular chaperone BiP. Nature 1991;353:726–730.
16. Schmid D, Baici A, Gehring H, Christen P. Kinetics of molecular chaperone action. Science 1994;263:971–973.
17. Pierpaoli EV, Gisler SM, Christen P. Sequence-specific rates of interaction of target peptides with the molecular chaperones DnaK and DnaJ. Biochemistry 1998;37:16741–16748.
18. Gragerov A, Zeng L, Zhao X, Burkholder W, Gottesman ME. Specificity of DnaK-peptide binding. J Mol Biol 1994;235:848–854.