Protein–Ligand Complexes

DOI: 10.1002/anie.201208561

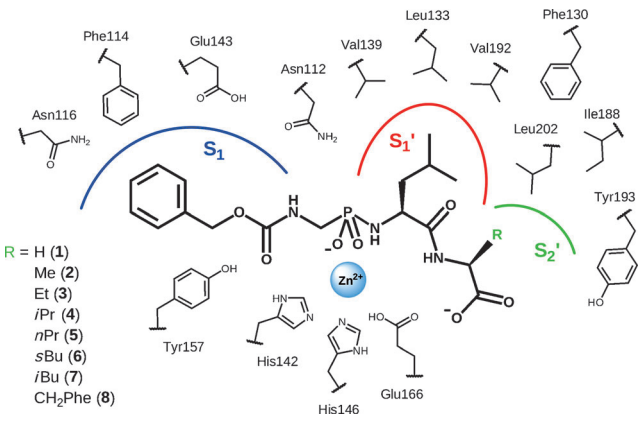
Dissecting the Hydrophobic Effect on the Molecular Level: The Role of Water, Enthalpy, and Entropy in Ligand Binding to Thermolysin**

Adam Biela, Nader N. Nasief, Michael Betz, Andreas Heine, David Hangauer, and Gerhard Klebe**

Dedicated to Professor Jack D. Dunitz on the occasion of his 90th birthday

The hydrophobic effect is viewed as the driving force for the aggregation of nonpolar substances with extended lipophilic molecular surfaces in aqueous solution through the exclusion of water molecules from the formed interfaces.[1,2] It is usually quoted to explain why an oil/water mixture spontaneously separates, why soluble proteins fold with a hydrophobic core and a hydrophilic outer surface,[3,4] why membrane components assemble as lipid bilayers and micelles, why membrane proteins are accommodated in membrane segments, and why small molecules associate in protein binding pockets with mutual burial of hydrophobic surfaces.[5] In the latter instance, it is a general strategy in medicinal chemistry to improve protein–ligand binding by increasing the ligand's hydrophobic surface which becomes buried in hydrophobic pockets of the target protein. In all cases, the hydrophobic effect is considered to be the major force of association. On the molecular level, this phenomenon is commonly attributed to the displacement of water molecules arranged around the hydrophobic surfaces, and entropic effects are made responsible to drive this association. The entropic profile is related to changes in the degree of ordering and the dynamic properties of the water molecules, which are assumed to be more disordered in the bulk water phase relative to where they were located prior to being displaced upon hydrophobic association. Recent studies have demonstrated, however, that hydrophobic interactions can originate either from enthalpy- or entropy-driven binding, making simple explanations often presented for the hydrophobic effect insufficient.[6–12] Also in computational design tools the handling of explicit water molecules has received increasing recognition. Tools such as WaterMap and Szmap[13,14] try to take into account water structures in drug design and the properties of individual water molecules are discussed in terms of enthalpy and entropy.

In order to obtain a better understanding of the hydrophobic effect on the molecular level and its role in protein–ligand binding, we embarked on a systematic study using thermolysin (TLN) as a model system.[6] This thermostable bacterial zinc metalloprotease from *Bacillus thermoproteolyticus* exhibits three specificity pockets of predominantly hydrophobic nature (Scheme 1). It has been considered



Scheme 1. A schematic view of the binding pocket of thermolysin with the bound ligand and the substitution pattern of the studied ligands **1**–**8**.

a prototype for the entire class of enzymes[15] owing to its highly conserved active-site architecture, despite remarkable sequence differences to other zinc proteases. Potent TLN inhibitors are often designed as transition-state analogues.[16–18] The enzyme has been frequently used as a surrogate[19–23] for other metalloenzymes against which new drugs are developed, and served as a model system to test ideas[24,25] and new methodological concepts.[26] TLN was one of the first crystallographically investigated metalloproteases[27,28] and its catalytic zinc ion is coordinated by His142, His146, and Glu166. The adjacent S_1 subsite is rather nonspecific and accommodates hydrophobic ligand portions.[16] In contrast, the S_1' pocket is a deep and well-defined cavity which hosts

[*] Dr. A. Biela, M. Betz, Dr. A. Heine, Prof. Dr. G. Klebe
Department of Pharmaceutical Chemistry
Philipps-University Marburg
Marbacher Weg 6, 35032 Marburg (Germany)
E-mail: klebe@mailer.uni-marburg.de
N. N. Nasief, Prof. Dr. D. Hangauer
Department of Chemistry, University at Buffalo
The State University of New York
Buffalo, NY 14260 (USA)
E-mail: hangauer@buffalo.edu

[**] G.K. is grateful to the EU for providing an ERC Advanced Grant (DrugProfilBind 268145). We thank the beamline support staff at the Helmholtz-Zentrum Berlin, Bessy II for radiation time and their advice during data collection. We thank the HZB for financial support of travel costs.

Supporting information for this article (experimental procedures for the crystallization and soaking of the ligands, data collection and processing, crystal structure determination and refinement, kinetic assay, isothermal titration calorimetry, and the synthesis of ligands **1**–**8**) is available on the WWW under http://dx.doi.org/10.1002/anie.201208561.

preferentially hydrophobic residues (e.g. the side chains of Val, Leu, or Phe), and determines substrate specificity.[29] The neighboring, more shallow and bowl-shaped S_2' pocket is also of hydrophobic nature; however, it is more easily accessible to water molecules from the bulk phase. Nevertheless, with respect to the binding of ligand side chains, S_2' can host groups very similar to those preferentially hosted by the S_1' pocket. Interestingly, across ligand series with P_1' or P_2' side chains modified from Gly to Leu the inhibitory potency increases strongly by 800-fold for the S_1' pocket whereas for S_2' only a 50-fold increase is experienced.[6] Furthermore, increasing the hydrophobic interactions in the S_1' pocket is strikingly enthalpy-driven and not, as assumed for the classical hydrophobic effect, entropically beneficial.[6] The favorable enthalpic signal, observed for growing hydrophobicity of the P_1' substituent in the congeneric series, was attributed to "poor solvation" of the S_1' pocket. The latter property most likely does not relate to a water-free enzyme pocket in the unbound state but to the accommodation of several highly mobile water molecules which are scattered over multiple positions and thus are hardly detectable by crystal structure analysis. The situation is very different for the S_2' pocket where, based on recently determined high-resolution crystal structures, a sophisticated arrangement of a complex water network is observed next to the S_2' pocket that exerts a dominant influence on the binding properties of the accommodated ligands.[30,31]

Homans et al.[7,8] reported a similar enthalpy-driven hydrophobic interaction which has also been attributed to suboptimal solvation of a protein pocket. The desolvation enthalpy of the protein binding pocket is greatly decreased so that overall the thermodynamic signature shifts to an overwhelmingly enthalpy-driven Gibbs free energy of binding. Snyder et al.[9] also reported an enthalpy-driven thermodynamic profile for heterocyclic aromatic sulfonamides with increasing hydrophobic properties against carbonic anhydrase. They explain the observed differences with changes in the number and organization of well-ordered water molecules in the binding site. In a study regarding the displacement of several well-ordered water molecules from the $S_{3/4}$ pocket of thrombin by increasingly hydrophobic P_3 substituents of peptidomimetic inhibitors, we observed an entropy-driven signal.[10] In the field of host–guest chemistry, several examples of complex formation have been reported to be driven either by enthalpy or entropy improvement.[11,12] These studies indicate that the thermodynamic signature of hydrophobic binding is determined by changes in the water structure, and will depend on the properties of the water molecules being reorganized during the binding process.

Since the S_1' and S_2' pockets of TLN exhibit opposite features with respect to the observed solvation patterns, but can host chemically similar ligand side chains, we embarked on the study of a congeneric series of peptidomimetics which contain a step-by-step series of modifications in the P_2' side chain. High-resolution crystal structures of the protein–ligand complexes were analyzed in order to investigate the changes in the water structure, as it is modulated by the interactions of the P_2' side chains. Isothermal titration calorimetry (ITC) data were recorded to complement crystallographic findings and to provide insights into the driving forces associated with hydrophobic binding to the S_2' pocket of TLN.

The crystal structures of eight TLN inhibitors containing the Cbz-Gly-$(PO_2)^-$-L-Leu-L-X scaffold (Cbz = carboxybenzyl, X = Gly **1**, Ala **2**, Et-Gly **3**, Val **4**, nPr-Gly **5**, Ile **6**, Leu **7**, Phe **8**; see Scheme 1) in complex with thermolysin have been determined at high resolution (1.28–1.66 Å). Crystal structures with ligands **1**, **2**, and **7** were studied previously.[6,30] For TLN-**7**, the original data have been newly refined so that the same protocol could be applied to all complexes. As the binding mode of the parent scaffold has been already described,[30] we will briefly discuss the predominant interactions of the ligand to TLN and focus only on novel structural features.

The electron density of the scaffold is well defined for all the studied inhibitors (see **3** in Figure S1 in the Supporting Information as a representative of the series). No significant changes in the binding mode are observed among all ligands (Figure S2). The Cbz moiety binds to the unspecific S_1 pocket and the central phosphonamidate group coordinates in monodentate fashion through one of its oxygen atoms to the zinc ion (2.0 Å), whereas the other is hydrogen-bonded to OE1 of Glu143. The leucyl P_1' and the structurally varied P_2' ligand side chains interact with the hydrophobic environment of the S_1' and S_2' pockets, respectively (for further details see the Supporting Information).

Major differences among the complexes are evident in the water network adjacent to the S_2' pocket. This network is perturbed and modulated by the size of the P_2' substituent. Unfortunately, not all complexes could be determined at the same resolution; the complex of the valyl derivative **4** even shows some disorder in this crucial region. For the related *n*-propylglycyl **5**, the residual difference density indicates that some disorder of the side chain might occur. We however decided to describe the final density with one model. Local disorder makes it difficult to reliably detect water molecules with increasing distance from the protein surface, or from polar ligand functional groups, particularly if chains of contiguously connected water molecules are analyzed. Therefore, the diffraction properties of the crucial water molecules have been thoroughly inspected by difference electron density maps (F_o-F_c) to examine the accuracy and reliability of the hydration properties of the S_2' pocket. Particularly, the B-factors and occupancies which are highly correlated have been regarded with care. These limitations complicate a straight-forward comparison of the absolute numbers of water molecules across the ligand series, especially considering the relative inventory of released or picked-up water molecules. The distances along the water network vary and may even correlate with the strength of formed hydrogen bonds. However, the determined spatial accuracy of individual water positions can be affected by, for example, residual mobility, disorder, and partial occupancy which limits positional accuracy. Therefore, we refrained from any detailed analysis of the variation in H-bond length.

Virtually the same solvation pattern is observed for all complexes next to the Cbz carbonyl and the negatively charged terminal carboxylate group (Figure S2 right, upper part). A network of at least seven mutually connected water

molecules mediates interactions between the latter two ligand functional groups and Asp226, Asn227, and Asn112. In two complexes water molecules corresponding to the second solvation shell are indicated (TLN-**4**, TLN-**5**).

A more complex pattern is observed next to the area of the S_2' pocket where the growing P_2' side chain extends and perturbs the water network (Figure S2, right, lower part). In a previous study the crystal structures of TLN-**1** and TLN-**2** were compared.[30] The glycine derivative **1** shows two water molecules hydrogen-bonded to the backbone carbonyl group of Asn111 which are displaced from the TLN-**2** complex (Figure 1 a, circled in cyan) due to steric conflicts with the attached methyl group in the alanyl derivative **2**. In contrast, the latter recruits two additional water molecules (Figure 1 b, circled in yellow and green) which are picked up and form favorable van der Waals contacts with the terminal methyl group. In TLN–**2** the water network establishes a contiguously connected water chain from the ligand's carboxylate group to Asn111(C=O), whereas in TLN–**1** the water network is disrupted (Figure 1 a,b).

A comparison of the water networks in TLN–**2** and TLN–**3** (Figure 1 b,c) suggests nearly identical hydration patterns, whereas those in TLN–**4** and TLN–**5** (Figure 1 d,e) seem to deviate and are disconnected at the lower left rim of the pocket. Nonetheless, the valyl **4** and *n*-propyl **5** derivatives display very similar water network patterns. In TLN–**4**, the isopropyl side chain is scattered over at least two conformations. This partial disorder is translated to the neighboring Leu202 residue as its isobutyl side chain adopts two conformations which refine to 55 % and 45 % occupancy (Figure S4). Both conformations occur in a correlated manner due to mutual steric interference. When one considers the water molecules picked up by TLN–**2** and –**3** compared to TLN–**1**, the complexes with **4** and **5** show the water molecule capping the position of the carboxylate group (Figure 1 d,e, circled in green).

The crystal structure of the complex with the Ile derivative **6** (Figure 1 f) suggests again a contiguously connected water network which wraps around the terminal hydrophobic group as is also seen in TLN–**2** and –**3**. The network takes a more extended detour around the butyl group than in TLN–**2**. TLN–**7**, reported in a previous study (PDB code 3FWD[6]), was re-refined in the present study in order to apply exactly the same refinement protocol and program suite. Even though the indicated network is not as complete as that for TLN–**6**, a related pattern is indicated for TLN–**7** (Figure 1 g). For both complexes the water molecule at the position capping the carboxylate group is no longer observed, in contrast to the complexes with **2**, **3**, **4**, and **5** (Figure 1 b–e, circled in green). In the latter crystal structures the shortest distance between the capping water and the alkyl side chain amounts to 3.80–3.85 Å, in TLN–**4** with the branched valyl group, present with disordered geometry, the water is slightly shifted (Figure S4a). If one takes the mean

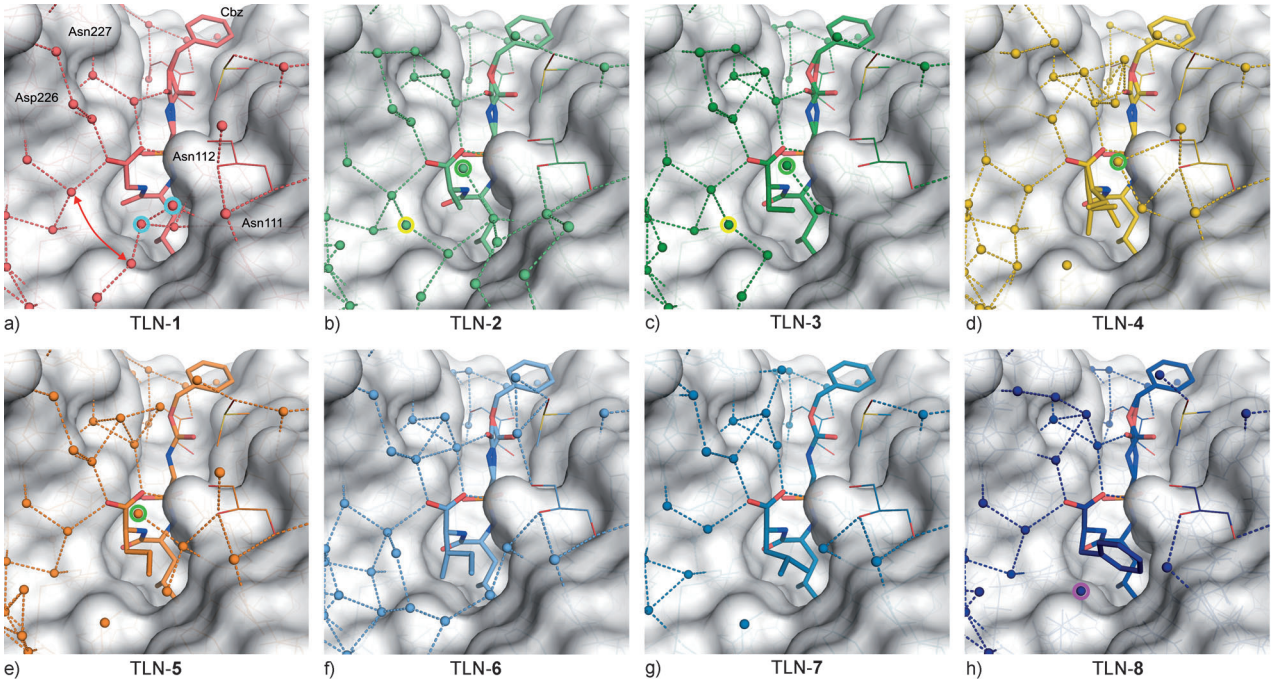


Figure 1. Binding modes of the ligands **1**–**8**. Each complex is shown with a different color, heteroatoms in atom-type color coding, water molecules as spheres with the same color as the parent structure. In TLN–**1** two water molecules (circled in cyan) are present that are replaced in the other complexes as a result of the steric requirement of the growing P_2' substituent. TLN–**1** shows a break in the contiguously connected water network (red arrow) which is closed in TLN–**2** and TLN–**3** by the pick-up of an additional water molecule (circled in yellow) and which is stabilized by both the favorable van der Waals contacts with the P_2' methyl or ethyl group in **2** or **3**, and the H-bonds with other water molecules. Similar favorable van der Waals interactions help accommodate a water molecule at a position capping the ligand's carboxylate group (circled in green) in complexes with **2**, **3**, **4**, and **5**. In the complexes with **6**, **7**, and **8** this water molecule is repelled, whereas **8** picks up a water molecule (circled in magenta) next to the benzyl moiety of the ligand's P_2' substituent.

position of the capping water in these complexes to calculate putative distances to the more bulky butyl side chains in TLN–**6** and –**7**, an 0.3 Å shorter contact would be created. Supposedly, this contact distance is too short and sterically unfavorable so that the capping water is no longer found in TLN–**6** and –**7**.

Finally, TLN–**8** with a terminal benzyl moiety shows the least amount of ordered water molecules next to the S_2' pocket. It seems that nearly all water molecules observed in other complexes close to the lower rim of the S_2' pocket are either repelled or not sufficiently well ordered. Most likely this correlates with increasing steric requirements of the benzyl group which fills the S_2' pocket quite significantly. Remarkably, however, one water molecule, which is also found in TLN–**2** and –**3** (Figure 1h, circled in magenta) and occupies a site very close to the ligand, can be observed in the TLN–**8** complex. This site is clearly not accommodated in TLN–**6** and –**7**, whereas in TLN–**8** a water molecule is found at this site stabilizing interactions (3.2 Å) with the π-system of the ligand's neighboring phenyl ring (Figure S5). Interestingly, the water molecule at the position capping the ligand's carboxylate group is also missing in TLN–**8**, apparently because of steric interference with the ligand's terminal P_2' substituent.

The bulky benzyl group in TLN–**8** has an impact on the neighboring protein molecule. It interferes with the carbonyl group of Asn111, which is pushed to a different position, giving rise to a second conformation (Figure S5). This perturbation is accompanied by a partial loss of the planarity of the peptide bond between Asn111 and Asn112 which is not observed in the other complexes of the series (ω angle deviates from planarity by 10.4° and −17.3°; more details in Table S2 in the Supporting Information). Apparently, the carbonyl oxygen evades in two directions to create enough space for the large benzyl side chain of **8**.

Changes in the thermodynamic parameters for binding across the inhibitor series were determined by ITC. Absolute thermodynamic values of the inhibitor binding could not be determined because ligand binding is superimposed by the displacement of the cleavage product Val-Lys resulting from autoprotolysis at high TLN concentrations. In this respect, all the measured thermodynamic values include a constant contribution resulting from the displacement of Val-Lys. This constant contribution of the dipeptide displacement is cancelled out in the relative values. Enzyme kinetic inhibition data of all inhibitors **1**–**8** were obtained, and demonstrated binding free energy changes similar in magnitude to the changes recorded by ITC, despite a constant offset (see Table S3 in the Supporting Information). Kinetic inhibition data, therefore, confirm the data obtained in ITC experiments.

Furthermore, a buffer dependence of our recorded ITC data was measured which shows that all complexes pick up one proton per formed protein–ligand complex. The data indicate that Glu143 next to the catalytic zinc ion changes its protonation state upon inhibitor binding (see the Supporting Information). As this residue is not directly involved in the binding of the P_2' portion, the thermodynamic data will be affected for all complexes in the same way, thus in a relative comparison also this contribution cancels out.

Figure 2 illustrates that, apart from TLN–**1**, binding becomes increasingly more entropic as the size of the attached hydrophobic P_2' substituent increases.[32] Simultaneously, significant enthalpy–entropy compensation is observed. As a consequence, across the series the net changes in the free energy are much smaller than those observed for either enthalpy or entropy.

A remarkable relative gain in the potency ($\Delta\Delta G_{1/2} = -5.7$ kJ mol^{-1}) is obtained going from **1** to **2**. This gain is mainly due to an improvement in enthalpy ($\Delta\Delta H_{1/2} =$

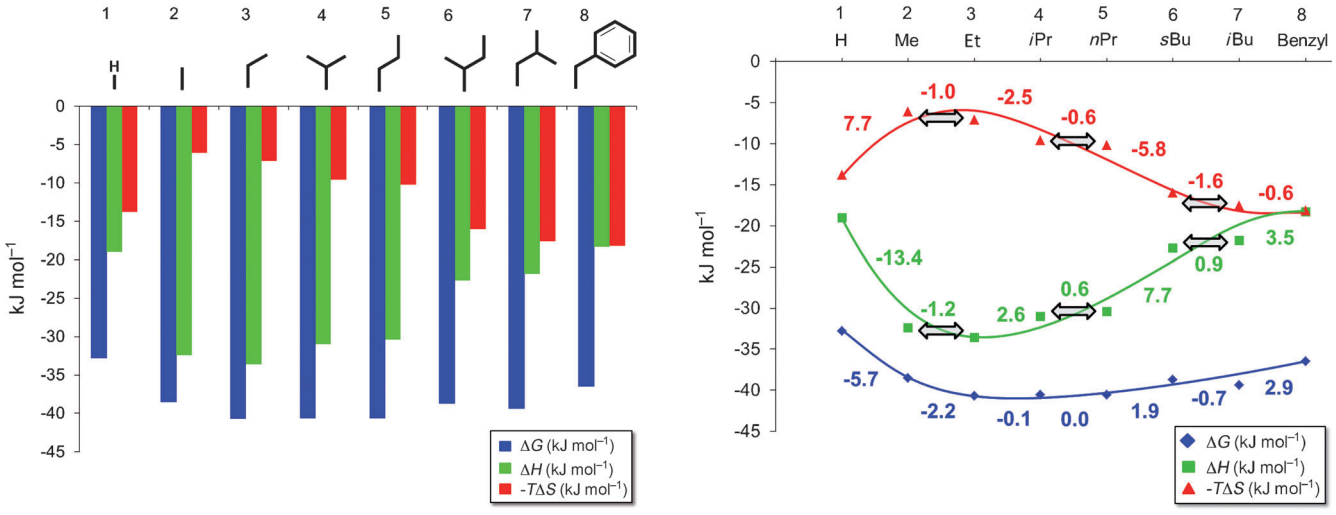


Figure 2. The diagram shows the observed thermodynamic results for ΔG (blue), ΔH (green), and $-T\Delta S$ (red) as obtained by ITC. The experiments were performed in HEPES buffer and not corrected for superimposed protonation steps and replacement of the autocleavage product Val-Lys. Right: To show the relative differences, mutual enthalpy–entropy compensation leading to minor changes in free energy and stepwise changes in terms of related pairs. The thermodynamic data are depicted in an alternative way and the numbers shown give the relative differences between neighboring ligands in the diagram.

$-13.4\ kJ\,mol^{-1}$) and the effect is partly compensated by the smaller entropic signal of **2** relative to **1** ($-T\Delta\Delta S_{1/2} = 7.7\ kJ\,mol^{-1}$) (Table S5). Across the series, the Gibbs free energy improves from **1** to **5**, whereby **3**, and **5** show the same values, within experimental accuracy. The remaining ligands **6**, **7**, and **8** decrease slightly in affinity (Figure 2).

If one considers closely related ligand pairs, some systematic changes are observed. For example, individual ligands in the pairs **2**/**3** ($\Delta\Delta H_{2/3} = -1.2\ kJ\,mol^{-1}$, $-T\Delta\Delta S_{2/3} = -1.0\ kJ\,mol^{-1}$), **4**/**5** ($\Delta\Delta H_{4/5} = -0.6\ kJ\,mol^{-1}$, $-T\Delta\Delta S_{4/5} = 0.6\ kJ\,mol^{-1}$), and **6**/**7** ($\Delta\Delta H_{6/7} = 0.9\ kJ\,mol^{-1}$, $-T\Delta\Delta S_{6/7} = -1.6\ kJ\,mol^{-1}$) exhibit very similar changes in their properties relative to each other, whereas going from **3** to **4** ($\Delta\Delta H_{3/4} = 2.6\ kJ\,mol^{-1}$, $-T\Delta\Delta S_{3/4} = -2.5\ kJ\,mol^{-1}$), and from **5** to **6** ($\Delta\Delta H_{5/6} = 7.7\ kJ\,mol^{-1}$ and $-T\Delta\Delta S_{5/6} = -5.8\ kJ\,mol^{-1}$) involves larger changes. Finally, the Phe derivative deviates from the pair **6**/**7** and shows a thermodynamic signature with balanced enthalpic and entropic portions.

In the series of the peptidomimetic transition-state-analogue inhibitors reported herein, the terminal hydrophobic substituent gradually penetrates the S_2' pocket of thermolysin. This shallow bowl-shaped pocket is open to the bulk solvent. It can host substituents up to the size of a benzyl moiety. This group fills the pocket quite substantially. It even pushes to the limit so that the backbone carbonyl group of the adjacent peptide bond in the protein has to move out of the position it occupies in the other complexes for steric reasons. This carbonyl group evades in two directions and thereby produces two alternative geometries which are likely to be energetically disfavored.

Even though the hydrophobic surface increases continuously in the series (**1**–**8**) from hydrogen in the Gly derivative to phenyl in the Phe derivative by about 130 $Å^2$, the overall Gibbs free energy improves only by $-3.7\ kJ\,mol^{-1}$. This is a minor contribution considering the rough estimate for the free energy of dehydration of about -2 to $-3\ kJ\,mol^{-1}$ per methyl group that becomes buried upon protein binding.[33] Purely based on surface patch considerations, we would expect a much larger value as the hydrophobic effect for this change. Interestingly, the affinity trend shows an optimum with an ethyl (**3**), isopropyl (**4**), or *n*-propyl (**5**) substituent, even though, in terms of size, these groups do not yet fill the S_2' pocket completely. This indicates even more that simple considerations based on hydrophobic surface patches buried upon complex formation break down in the current analysis. For the Phe derivative **8**, the crystal structure indicates disfavored conformations for Asn111 which will also influence the decrease in the affinity of this ligand. More remarkable is the trend in enthalpy/entropy partitioning (Figure 2) which is largest with the most potent inhibitors. Changes are not consistently observed across the series, but the structurally closely related pairs **2**/**3**, **4**/**5**, and **6**/**7** exhibit very similar thermodynamic profiles (Figure 2). This suggests for each pair similarities in the structural solvation patterns of the individual complexes.

In the analysis of the binding modes of **1**–**8**, the parent scaffold remains virtually unchanged across the entire series. Also the hydration pattern next to the ligand's Cbz group and the terminal carboxylate group, which involves at least seven conserved water molecules and mediates a complex network between ligand and protein functional groups, does not show any significant changes across the series (Figure S2, right upper part). In addition, no change can be detected for the contacts to the glycerol and DMSO molecules picked up from the cryo buffer. Thus, the only differences occur next to the hydrophobic P_2' substituents which increase in size (Figure S2, right, lower part). Here, the network of the adjacent water molecules is highly perturbed.

A huge change in the thermodynamic profile is experienced when a methyl group is added to the glycine derivative **1** resulting in the alanine substituent in **2**. One difference is that the Gly derivative has in solution, prior to protein binding, access to a larger conformational space than the Ala analogue. Consequently, TLN–**2** will experience a smaller loss in entropy than TLN–**1** upon binding. However, since the thermodynamic profile of TLN–**2** shows a larger, rather than a smaller loss in entropy, additional effects are in operation. It might be well possible that the conformational differences are of minor entropic importance as ligands in a solvent cage will hardly experience full flexibility. They could be as restricted in their degrees of freedom as they would be at a binding site which opens to the bulk solvent. As shown in our previous study,[30] which also involves the non-carboxylated analogues of **1** and **2**, binding includes the rupture of the contiguously connected water network which wraps around the terminal methyl group in **2** (Figure 1 a,b). It is apparent that the methyl group provides favorable interaction sites for the two additional water molecules which are further stabilized in their binding positions through van der Waals contacts. On the other hand, two water molecules (Figure 1 a, circled in cyan) H-bonded in TLN-**1** to Asn111(C=O) are repelled from the complex as a result of steric conflicts with the additional methyl group in **2**. The rupture of the contiguously connected H-bonding network disfavors the exothermic binding of **1**, whereas the binding of **2** is entropically less favorable owing to a stronger fixation of the water network. Accordingly, going from **1** to the more hydrophobic **2** is enthalpy-driven and could be classified—in formal terms—as a "nonclassical hydrophobic effect".

The thermodynamic signature of the ethyl derivative **3** is nearly identical to that of **2**. As both enthalpy and entropy become more favorable relative to **2**, $\Delta\Delta G$ improves by $-2.2\ kJ\,mol^{-1}$, a value found in the typical range for favorably placed additional methyl groups. It is in good agreement with the estimated free energy of desolvation of a methyl group. Again the additional degree of conformational freedom in **3** seems to be of minor importance. With respect to the water network, TLN–**3** is nearly identical with TLN–**2**. The isopropyl and *n*-propyl derivatives **4** and **5** share again very similar thermodynamic properties, but their enthalpy/entropy values are much different from those of the previous pair **2**/**3**. In TLN–**4** the branched and more rigid valyl ligand side chain is distributed over two conformations and also a disorder of the Leu202 side chain is detected. In TLN–**5**, the even more flexible P_2' *n*-propyl group seems to be ordered; nonetheless some disorder cannot be fully excluded by crystallography though it is less evident. At the far end of the pocket two water molecules mediating the water network in TLN–**2** and

–**3** cannot be detected and the network appears incomplete in TLN–**4** and –**5**. This should result, as seen for TLN–**1**, in an enthalpic loss and an entropic gain. Both disorder and water release support the entropic advantage of TLN–**4** and –**5**; however, with respect to the free energy, the latter advantage is nearly canceled out. Nonetheless, it is remarkable that **4** and **5** not only share similar thermodynamic signatures, but also the details of their local water networks show significant resemblance. The hydrophobic effect related to the change from **2**/**3** to **4**/**5** can be viewed as a "classical entropy-driven hydrophobic effect".

All four complexes TLN–**2**, –**3**, –**4**, and –**5** host one water molecule at a position capping the ligand's carboxylate group. They all show a nearly 10 kJ mol^{-1} more enthalpic signal than TLN–**1**, –**6, --7**, and –**8** which lack this capping water. Its spatial position should be energetically favored for electrostatic reasons. Frequent occupancy of such carboxylate–water contacts was highlighted by Paulini et al.[34] in protein structures and a clear preference for this geometry can be found in the compilation of water–carboxylate contacts as assembled in IsoStar.[35]

TLN–**6** and –**7** place a *sec*-butyl and isobutyl group, respectively, into the S_2' pocket in an ordered fashion. Again both complexes experience a very similar enthalpy/entropy profile with strong enthalpic loss and entropic gain relative to the pair **4**/**5**. With respect to their water solvation pattern quite similar networks are observed that wrap around the terminal hydrophobic ligand side chain, and orient along the rim of the S_2' pocket. Compared to TLN–**2** and TLN–**3**, TLN–**6** has two water molecules shifted in position to more remote sites in order to create an expanded network. The capping water above the carboxylate group is no longer observed, supposedly as a consequence of steric repulsion with the larger butyl substituent. Apparently, the release of this special water is one of the causes of the observed enthalpy loss and entropy gain, which we would relate to the "classical hydrophobic effect".

Finally, the benzyl derivative **8** loses 2.9 kJ mol^{-1} in $\Delta\Delta G$ with respect to **7**. This price is paid in enthalpy, and correlates with the steric clash of the phenyl group with the backbone carbonyl group of Asn111. Furthermore, in this complex the capping water is also repelled from the site above the carboxylate group. The solvation structure at the far end of the S_2' pocket appears to be thinned out; however, one water site that is close to the ligand and has already been accommodated in TLN–**2** and –**3** is newly populated (Figure S5a). This water molecule (Figure 1 h, circled in magenta) is most likely stabilized at this pivotal position perpendicular to the terminal aromatic benzyl moiety by the interactions with the phenyl group. Thus in this final case, compared to TLN–**6** and –**7**, the hydrophobic effect is paid in enthalpy and even a new water binding site is populated.

The described structural features next to the S_2' pocket show that simple models[36,37] based on buried ligands and protein surface patches to describe beneficial desolvation contributions arising from size differences of the ligands are not sufficient to explain the observed thermodynamic signature across the series. Differences in the varying conformational properties of the P_2' side chains exhibiting different numbers of rotatable bonds also seem to be of minor importance. The bound ligand side chains contribute together with the protein to the newly formed complex surfaces. Water molecules with their strong structure-determining properties arrange along the new surfaces; they can even impact the ligand's bound geometry to find a best compromise with respect to the formed water network. In total, all these contributions starting from the separately solvated binding partners to the formed complex describe quantitatively the observed thermodynamic profiles. As thermolysin is a very rigid protein, influences resulting from changes of the residual mobility of protein residues or induced-fit adaptations, apart from TLN–**8**, will be of minor importance. Even though our discussion only intends to provide a qualitative correlation, it is remarkable that complexes with side chains of comparable size and number of rotatable bonds show very similar thermodynamic signature.

The presented series can be used to correlate the details of how the first solvation layer around a binding pocket impacts ligand-binding affinity. It also shows that water networks can have significant influence on modulating the structure–activity relationships. Increasing the hydrophobicity of ligand functional groups that bind in hydrophobic pockets is usually discussed in terms of the hydrophobic effect. This effect has been attributed to either an enthalpic or an entropic signature.[6,7,11,12] Crystal structure analyses along with the ITC data of the presented complex series underscore that both, enthalpy and entropy, are involved in the hydrophobic effect, and that many detailed structural phenomena determine the final overall signature. For example, if a contiguously connected water network ruptures like in TLN–**1**, an enthalpic loss and an entropic gain are experienced relative to TLN–**2**.[26] The loss of the hydrogen bonds caused by this rupture allows the system to activate and thus distribute energy over more degrees of freedom. The displacement of ordered water molecules can be entropically favorable. On the other hand, the release of largely disordered waters can also reveal a predominately enthalpic signal.[6–9] If parts of the ligand are accommodated in pockets that are open to the bulk solvent, and parts of the ligand are exposed to the water phase, new binding sites for water molecules can be generated, for example, in our study the position capping the carboxylate group or the site found on top of the phenyl ring. The capping water seems to provide a significant contribution. The four complexes TLN–**1**, –**6**, –**7**, and –**8** do not show this water molecule supposedly because of steric interference and they lose in the Gibbs free energy predominantly for enthalpic reasons. This loss in enthalpy is partly compensated by entropy as these complexes do not capture the water molecule, a process which would be entropically unfavorable if it were to occur. All these phenomena contribute on the molecular level to the finally determined hydrophobic effect. In summary, there are no universally valid reasons why the hydrophobic effect should be predominantly "entropic" or "enthalpic"; small structural changes in the binding features of water molecules on the molecular level determine whether hydrophobic binding is enthalpically or entropically driven.

Admittedly, this study reaches the limits of experimental accuracy accomplishable in contemporary protein–ligand

structural work. Assignment of water molecules to the difference electron density is crucial and depends on both resolution and the local ordering phenomena. In the present series, eight crystal structures indicate similar and internally consistent solvation patterns which give confidence about the reliability necessary to underscore their relevance. Our structural data originate from crystal environments of flash-cooled crystals which should capture a frozen image of the structure at ambient temperature.[38] The thermodynamic data were recorded in solution. Nevertheless, a very consistent picture emerges. Surprising pairwise systematic changes in the thermodynamic data are experienced for complexes of related ligands, and they are convincingly well reflected by the structural properties. The present study unravels small but important details. Computational methods simulate molecular properties at the atomic level, and are usually determined by the summation of many small details.[39] However, details such as those observed here are usually not regarded by these computational methods as relevant, simply because we are not fully aware of their importance for protein–ligand binding, structure–activity relationships, and rational drug design in general. The study presented here sheds light on the role these details play in the binding process. Taking such details into account will make computational simulations not necessarily simpler, but hopefully better in predicting ligand-binding affinity, as well as the other molecular recognition aspects.

Experimental Section

Coordinates and structure factors have been deposited in the Protein Data Bank with the following accession codes: TLN–**3** complex 3T87; TLN–**4** complex 3T8H; TLN–**5** complex 3T8C; TLN–**6** complex 3T8D; and TLN–**8** complex 4D9W. TLN–**1** and TLN–**2** were previously reported as 3T8G and 3T74, respectively. TLN–**7** (4H57) has been newly refined and deposited replacing the former refinement results.

Received: October 24, 2012
Published online: January 2, 2013

Keywords: crystal structure analysis · enthalpy–entropy compensation · hydrophobic effect · protein–ligand interactions · water solvation

[1] "hydrophobic interaction:" M. Nic, J. Jirat, B. Kosata, *IUPAC Compendium of Chemical Terminology* (Online ed.), **2006**, DOI: 10.1351/goldbook.H02907.
[2] D. Chandler, *Nature* **2005**, *437*, 640 – 647.
[3] W. Kauzmann, *Adv. Protein Chem.* **1959**, *14*, 1 – 63.
[4] M. Charton, B. I. Charton, *J. Theor. Biol.* **1982**, *99*, 629 – 644.
[5] G. M. Whitesides, V. M. Krishnamurthy, *Q. Rev. Biophys.* **2005**, *38*, 385 – 395.
[6] L. Englert, A. Biela, M. Zayed, A. Heine, D. Hangauer, G. Klebe, *Biochim. Biophys. Acta Gen. Subj.* **2010**, *1800*, 1192 – 1202.
[7] R. J. Bingham, J. B. Findlay, S. Y. Hsieh, A. P. Kalverda, A. Kjellberg, C. Perazzolo, S. E. Phillips, K. Seshadri, C. H. Trinh, W. B. Turnbull, G. Bodenhausen, S. W. Homans, *J. Am. Chem. Soc.* **2004**, *126*, 1675 – 1681.
[8] S. W. Homans, *Drug Discovery Today* **2007**, *12*, 534 – 539.
[9] P. W. Snyder, J. Mecinovic, D. T. Moustakas, S. W. Thomas, 3rd, M. Harder, E. T. Mack, M. R. Lockett, A. Heroux, W. Sherman, G. M. Whitesides, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 17889 – 17894.
[10] A. Biela, F. Sielaff, F. Terwesten, A. Heine, T. Steinmetzer, G. Klebe, *J. Med. Chem.* **2012**, *55*, 6094 – 6110.
[11] D. B. Smithrud, T. B. Wyman, F. Diederich, *J. Am. Chem. Soc.* **1991**, *113*, 5420 – 5426.
[12] K. J. Breslauer, D. P. Remeta, W. Y. Chou, R. Ferrante, J. Curry, D. Zaunczkowski, J. G. Snyder, L. A. Marky, *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 8922 – 8926.
[13] T. Young, R. Abel, B. Kim, B. J. Berne, R. A. Friesner, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 808 – 813.
[14] J. M. Word, M. T. Geballe, A. Nicholls, *Abstr. Pap. Am. Chem. Soc.* **2010**, *240* (Meeting Abstract 280-Comp).
[15] J. T. Bukrinsky, M. J. Bjerrum, A. Kadziola, *Biochemistry* **1998**, *37*, 16555 – 16564.
[16] H. M. Holden, D. E. Tronrud, A. F. Monzingo, L. H. Weaver, B. W. Matthews, *Biochemistry* **1987**, *26*, 8542 – 8553.
[17] D. E. Tronrud, A. F. Monzingo, B. W. Matthews, *Eur. J. Biochem.* **1986**, *157*, 261 – 268.
[18] P. A. Bartlett, C. K. Marlowe, *Biochemistry* **1983**, *22*, 4618 – 4624.
[19] M. Selkti, A. Tomas, J. F. Gaucher, T. Prange, M. C. Fournie-Zaluski, H. Chen, B. P. Roques, *Acta Crystallogr. Sect. D* **2003**, *59*, 1200 – 1205.
[20] J. F. Gaucher, M. Selkti, G. Tiraboschi, T. Prange, B. P. Roques, A. Tomas, M. C. Fournie-Zaluski, *Biochemistry* **1999**, *38*, 12569 – 12576.
[21] T. Takeuchi, A. Bottcher, C. M. Quezada, T. J. Meade, H. B. Gray, *Bioorg. Med. Chem.* **1999**, *7*, 815 – 819.
[22] K. J. Lee, D. H. Kim, *Bioorg. Med. Chem. Lett.* **1998**, *8*, 323 – 326.
[23] D. G. Hangauer, A. F. Monzingo, B. W. Matthews, *Biochemistry* **1984**, *23*, 5730 – 5741.
[24] A. C. English, C. R. Groom, R. E. Hubbard, *Protein Eng.* **2001**, *14*, 47 – 59.
[25] J. Behnen, H. Koster, G. Neudert, T. Craan, A. Heine, G. Klebe, *ChemMedChem* **2012**, *7*, 248 – 261.
[26] L. Englert, K. Silber, H. Steuber, S. Brass, B. Over, H. D. Gerber, A. Heine, W. E. Diederich, G. Klebe, *ChemMedChem* **2010**, *5*, 930 – 940.
[27] B. W. Matthews, L. H. Weaver, W. R. Kester, *J. Biol. Chem.* **1974**, *249*, 8030 – 8044.
[28] M. A. Holmes, B. W. Matthews, *J. Mol. Biol.* **1982**, *160*, 623 – 639.
[29] K. Morihara, *Biochem. Biophys. Res. Commun.* **1967**, *26*, 656 – 661.
[30] A. Biela, M. Betz, A. Heine, G. Klebe, *ChemMedChem* **2012**, *7*, 1423 – 1434.
[31] N. N. Nasief, H. Tan, J. Kong, D. G. Hangauer, *J. Med. Chem.* **2012**, *55*, 8283 – 8302.
[32] N. N. Nasief, A. Biela, G. Klebe, D. Hangauer in *242nd ACS National Meeting & Exposition*, Denver, CO, United States, **2011**.
[33] C. S. Leung, S. S. Leung, J. Tirado-Rives, W. L. Jorgensen, *J. Med. Chem.* **2012**, *55*, 4489 – 4500.
[34] R. Paulini, K. Müller, F. Diederich, *Angew. Chem.* **2005**, *117*, 1820 – 1839; *Angew. Chem. Int. Ed.* **2005**, *44*, 1788 – 1805.
[35] I. J. Bruno, J. C. Cole, J. P. Lommerse, R. S. Rowland, R. Taylor, M. L. Verdonk, *J. Comput.-Aided Mol. Des.* **1997**, *11*, 525 – 537.
[36] M. H. Abraham, G. S. Whiting, R. Fuchs, E. J. Chambers, *J. Chem. Soc. Perkin Trans. 2* **1990**, 291 – 300.
[37] W. L. Jorgensen, J. P. Ulmschneider, J. Tirado-Rives, *J. Phys. Chem. B* **2004**, *108*, 16264 – 16270.
[38] M. Nakasako, *Philos. Trans. R. Soc. London Ser. B* **2004**, *359*, 1191 – 1204.
[39] D. J. Huggins, W. Sherman, B. Tidor, *J. Med. Chem.* **2012**, *55*, 1424 – 1444.

