J|A|C|S
ARTICLES

Published on Web 08/22/2008

Residual Ligand Entropy in the Binding of *p*-Substituted Benzenesulfonamide Ligands to Bovine Carbonic Anhydrase II

Henning Stöckmann,[†,‡] Agnieszka Bronowska,[†] Neil R. Syme,[†] Gary S. Thompson,[†] Arnout P. Kalverda, Stuart L. Warriner,[†,‡] and Steve W. Homans*[,†]

Astbury Centre for Structural Molecular Biology and School of Chemistry, University of Leeds, Leeds LS2 9JT, U.K.

Received April 4, 2008; E-mail: s.w.homans@leeds.ac.uk

Abstract: In studies on the thermodynamics of ligand−protein interactions, it is often assumed that the configurational and conformational entropy of the ligand is zero in the bound state (i.e., the ligand is rigidly fixed in the binding pocket). However, there is little direct experimental evidence for this assumption, and in the case of binding of *p*-substituted benzenesulfonamide inhibitors to bovine carbonic anhydrase II (BCA II), the observed thermodynamic binding signature derived from isothermal titration calorimetry experiments leads indirectly to the conclusion that a considerable degree of residual entropy remains in the bound ligand. Specifically, the entropy of binding increases with glycine chain length *n*, and strong evidence exists that this thermodynamic signature is not driven by solvent reorganization. By use of heteronuclear ^{15}N NMR relaxation measurements in a series ($n = 1-6$) of ^{15}N-glycine-enriched ligands, we find that the observed thermodynamic binding signature cannot be explained by residual ligand dynamics in the bound state, but rather results from the indirect influence of ligand chain length on protein dynamics.

Introduction

The affinity of a given biomolecular interaction depends upon a complex interplay between the structure and dynamics of the interacting partners and solvent water. In particular, the standard free energy of binding, $\Delta G_b°$, from which affinities can be derived, is equal to the difference in the standard free energy of the solvated complex with respect to the standard free energies of the solvated uncomplexed species. Many of the factors that contribute to $\Delta G_b°$ are difficult or impossible to measure experimentally, leading to a number of assumptions when attempting to decompose binding affinities into particular contributions from protein, ligand, and solvent. In particular, it is often assumed that the configurational entropy of the ligand, composed in all practical cases of three rotational and three translational degrees of freedom, and the conformational entropy comprising internal degrees of freedom are lost on binding (i.e., the ligand is rigidly held in the binding pocket). It is particularly difficult to confirm this assumption (or not) experimentally, although a number of theoretical approaches offer some intriguing insight and suggest that this approximation is not well-founded.[1] Clearly, since $\Delta G_b°$ contains both an entropic component ($\Delta S_b°$) and an enthalpic component ($\Delta H_b°$) as expressed in the fundamental Gibbs equation $\Delta G_b° = \Delta H_b° - T\Delta S_b°$, residual entropy of the bound ligand represents a favorable contribution to binding affinities and hence is worthy of detailed study.

Here, we examine the residual ligand entropy in the binding of a series of *p*-(glycine)$_n$-substituted benzenesulfonamide ($ArGly_nO^-$) ligands to bovine carbonic anhydrase II (BCA II)[2] using isothermal titration calorimetry (ITC), heteronuclear NMR, and computational chemistry. In a recent study, Whitesides and co-workers studied these interactions for a panel of ligands ($n = 1-5$) and found almost perfect enthalpy−entropy compensation across the series: the enthalpy of binding became less favorable and the entropy more favorable with increasing chain length.[3] Changes in heat capacity were independent of chain length, suggesting that the observed changes in binding thermodynamic signatures across the series cannot be explained on the basis of the classical hydrophobic effect. To explain these data, a model was proposed on the basis of decreasing "tightness" of the protein−ligand interface as the chain length of the ligand increases, such that less favorable enthalpic interactions take place (due to fewer van der Waals contacts) together with a more favorable entropy due to the greater mobility of the chain. The increase in mobility of the ligand with chain length in this model is particularly intriguing, and it occurred to us that this can in principle be probed by ^{15}N NMR relaxation measurements of the type that are conventionally used to probe backbone dynamics in proteins.[4] The ready availability of ^{15}N glycine permits the straightforward synthesis of the respective ligands.[5] While the mobility of $ArGly_nO^-$ ligands has previously been examined in the bound state by Whitesides

[†] Astbury Centre for Structural Molecular Biology.
[‡] School of Chemistry.

(1) Chang, C. A.; Chen, W.; Gilson, M. K. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 1534–1539.
(2) Krishnamurthy, V. M.; Kaufman, G. K.; Urbach, A. R.; Gitlin, I.; Gudiksen, K. L.; Weibel, D. B.; Whitesides, G. M. *Chem. Rev.* **2008**, *108*, 946–1051.
(3) Krishnamurthy, V. M.; Bohall, B. R.; Semety, V.; Whitesides, G. M. *J. Am. Chem. Soc.* **2006**, *128*, 5802–5812.
(4) Farrow, N. A.; Muhandiram, R.; Singer, A. U.; Pascal, S. M.; Kay, C. M.; Gish, G.; Shoelson, S. E.; Pawson, T.; Formankay, J. D.; Kay, L. E. *Biochemistry* **1994**, *33*, 5984–6003.
(5) Jain, A.; Huang, S. G.; Whitesides, G. M. *J. Am. Chem. Soc.* **1994**, *116*, 5057–5062.

10.1021/ja803755m CCC: $40.75

and co-workers (reviewed in ref 2) using ^{1}H NMR spin−spin relaxation time (T_2) measurements, a reliable measure of dynamics is difficult from these data because of the influence of "external" protons that contribute substantially to ^{1}H relaxation. In contrast, ^{15}N relaxation is almost entirely dominated by ^{15}N chemical shift anisotropy and dipolar relaxation involving the amide proton, and thus ^{15}N relaxation parameters can be interpreted in terms of the dynamic motion of the amide bond vector with great confidence.

Materials and Methods

Ligand Synthesis. $ArGly_nO^-$ ligands were synthesized using standard Fmoc solid-phase peptide synthesis methods, which are detailed in Supporting Information. Briefly, the C-terminal glycine residue (labeled or unlabeled with ^{15}N according to the target) was attached to Wang resin, and following Fmoc deprotection subsequent amino acids were coupled using activation with HCTU. Following completion of the peptide chain, 4-carboxybenzenesulfonamide was attached to the N-terminus using DIC/HOBt activation. The target was cleaved from the resin using TFA and purified using gel filtration.

NMR Measurements. Samples of 300 μL of a solution of 600 μM BCA II (purchased from Sigma, concentration determined by UV spectrophotometry, $\varepsilon_{280} = 57\,000$ M^{-1} cm^{-1}) and 500 μM arylsulfonamide ligand (concentration determined by ^{1}H NMR) in 20 mM sodium phosphate buffer pH 7.5 (with 10 μM N_3^-) in 10% D_2O/90% H_2O were prepared for R_1, R_2, and NOE measurements. NMR data were recorded at 298 K on Varian Inova spectrometers with proton frequencies of 500 and 600 MHz. Microcell NMR tubes (Shigemi) were used to minimize condensation of vapor. ^{15}N chemical shifts were obtained from HSQC spectra. R_1, R_2, and ^{15}N{^{1}H} NOE data were obtained with pulse sequences (a) to (c) of Farrow et al.[4] The ^{15}N R_1 data were obtained with spin−lattice relaxation periods of 0.0111, 0.0555, 0.1110, 0.2220, 0.3885, 0.5550, 0.7770, 0.9990, 1.3320 s and 0.0109, 0.0543, 0.1086, 0.2173, 0.3802, 0.5432, 0.7605, 0.9778, 1.3037, 1.6296 s on 500 and 600 MHz spectrometers, respectively. The ^{15}N R_2 data were obtained by using CPMG relaxation periods of 0.0164, 0.0329, 0.0493, 0.0658, 0.0822, 0.0987, 0.1316, 0.1645, 0.1974, 0.2303, 0.2632 s and 0.0163, 0.0326, 0.0490, 0.0653, 0.0816, 0.0979, 0.1306, 0.1632, 0.1958, 0.2285, 0.2611 s on 500 and 600 MHz spectrometers, respectively. The intensities of the peaks in the 1-D spectra were determined using the VNMR package (Varian, Inc.). All spectra were processed using manual phase correction and automatic baseline correction. Two duplicate spectra per relaxation series were used to estimate the noise level for the estimation of the standard deviation of noise in the peak intensities.[6] R_1 and R_2 values were extracted by least-squares fitting of the peak intensities to a two-parameter exponential function using the Levenberg−Marquardt algorithm. Fitting was performed using in-house software employing the Scientific Python library. Bootstrap resampling using 100 simulations was performed to estimate the uncertainty of the relaxation parameters.[7] Order parameters and correlation times were extracted from models m1 to m5[8] with the *relax* NMR software suite.[9,10] The overall rotational correlation time for BCA II was set to 11.5 ns as determined by Jarvet et al. under similar experimental conditions.[11] Variation of the correlation time by ±10% resulted in squared order parameters (S^2) from the fits that differed by approximately ±10%. However, differences in S^2 between complexes involving successive members of the $ArGly_nO^-$ ligand series (see Results and Discussion) were within experimental error for correlation times within this range.

To eliminate differences in rotational correlation times due to concentration (and thus viscosity) differences between samples, all sample aliquots were prepared from the same batch of protein solution, and ligand solution concentrations were adjusted such that the same volume of ligand was added in each case. Monte Carlo resampling using 100 simulations was performed to estimate the uncertainty of the extracted parameters. Entropy changes between adjacent residues in the oligoglycine chain were calculated using the expression derived by Yang and Kay.[12] Here, the change in per-residue conformational entropy ΔS_p between two states a and b (i.e., free and bound) is given by:

$$\frac{\Delta S_p}{k} = \ln \frac{3 - \sqrt{1 + 8S_b}}{3 - \sqrt{1 + 8S_a}} \quad (1)$$

where S_b and S_a are the order parameters for a given bond vector in states b and a, respectively.

Molecular Dynamics Simulations. All simulations were carried out using the AMBER8 program[13] with the Cornell et al. force field.[14] Initial coordinates of *apo* BCA II and all complexes were based on the crystal structure of human carbonic anhydrase II complexed with 4-(aminosulfonyl)-*N*-[(2,3-2 difluorophenyl)methyl]-benzamide (PDB code: 1G52). The glycine derivatives were constructed in MOLDEN,[15] optimized using the ab initio RHF/6-31G* basis set (Gaussian 98),[16] and RESP charges[17] were generated and fitted. All ligand parameters used the parm99 force field. The models were subjected to a short energy minimization (1000 cycles) and manually docked in the conformation and orientation resembling the benzosulfonyl moiety in the crystal structure. The structures were processed by the XleaP module of AMBER, and hydrogen atoms were added to the system. A single zinc ion was also docked in the position adjacent to the benzenesulfonamide moiety, according to the crystal structure. After a short energy minimization (5000 cycles), ligand−protein models were immersed in periodic TIP3P water boxes. Approximately 7000 water molecules were added to each system. Simulations were carried out under NPT conditions ($T = 300$ K) using the particle mesh Ewald technique[18] with 12 Å nonbonded cutoff and 2-fs time step. SHAKE constraints with a tolerance of 10^{-8} Å were applied to all hydrogens during MD simulations to eliminate the fastest X−H vibrations and allow a longer simulation time step. Translational and rotational center-of-mass motion was removed every 5 ps. Equilibration was started by 25 000 cycles of energy minimization, with the atomic positions of the solute molecule restrained. This was followed by 100 ps of MD simulation, during which the system was heated to 300 K and the constraints on the solute were gradually reduced from the initial value of 100 kcal/(mol·Å^2) during a further equilibration period of 4.9 ns. The production period took 45 ns for each of the six ligand−protein complexes investigated. The coordinates were saved every 1 ps of MD simulation. MD trajectories were analyzed using the ptraj module of AMBER. The coordinates of each system were superimposed on the starting, energy-minimized structure, which was very close to the crystal

(6) Palmer, A. G.; Rance, M.; Wright, P. E. *J. Am. Chem. Soc.* **1991**, *113*, 4371–4380.
(7) Efron, B. *Biometrika* **1981**, *68*, 589–599.
(8) Mandel, A. M.; Akke, M.; Palmer, A. G. *J. Mol. Biol.* **1995**, *246*, 144–163.
(9) d'Auvergne, E. J.; Gooley, P. R. *J. Biomol. NMR* **2003**, *25*, 25–39.
(10) d'Auvergne, E. J.; Gooley, P. R. *J. Biomol. NMR* **2006**, *35*, 117–135.
(11) Jarvet, J.; Olivson, A.; Mets, U.; Pooga, M.; Aguraiuja, R.; Lippmaa, E. *Eur. J. Biochem.* **1989**, *186*, 287–290.
(12) Yang, D. W.; Kay, L. E. *J. Mol. Biol.* **1996**, *263*, 369–382.
(13) Case, D. A.; et al. *AMBER 8*; University of California: San Francisco, 2004.
(14) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.
(15) Schaftenaar, G.; Noordik, J. H. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 123–134.
(16) Frisch, M. J.; et al. *Gaussian 98*, revision A.9; Gaussian, Inc.: Pittsburgh, PA, 1998.
(17) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. *J. Am. Chem. Soc.* **1993**, *115*, 9620–9631.
(18) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089–10092.

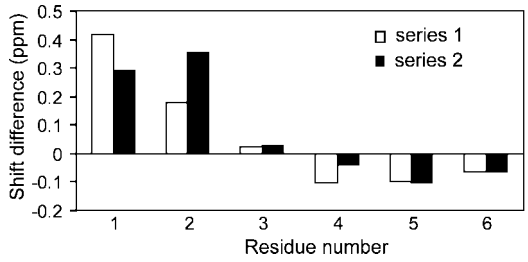

Figure 1. ^{1}H chemical shift differences (free minus bound) for series 1 and 2 ligands binding to BCA II.

structure of carbonic anhydrase II. Postprocessing of MD trajectories also included removal of water molecules and calculation of atomic fluctuations during the simulation period. The ptraj-processed trajectories were used for the calculation of generalized squared order parameters (S^2), as described by MacRaild et al.[19] Briefly, generalized order parameters were calculated from the trajectory of individual bond vectors as:[20]

$$S^2 = \frac{3}{2}[\langle x^2\rangle^2 + \langle y^2\rangle^2 + \langle z^2\rangle^2 + 2\langle xy\rangle^2 + 2\langle xz\rangle^2 + 2\langle yz\rangle^2] - \frac{1}{2} \quad (2)$$

where x, y, and z are components of a unit vector along the amide bond, and angular brackets denote the time average over the trajectory. Convergence of the dynamics of interest was tested using the approach described by Best and Vendruscolo:[20] a cumulative time function $S^2(\tau)$ is defined using eq 2, with the time averages taken from $t = 0$ to $t = \tau$. This function was evaluated for 100 equally spaced time points across the trajectory. The trajectory was deemed to have converged if the difference between the maximum and minimum values of this function over the final 50 time points (i.e., the final 22.5 ns of the trajectory) was less than 0.05. Any residue judged not to have converged was excluded from the analysis. The theoretically derived entropy was calculated from the probability density of the relevant bond vector over the MD simulation according to eq 16 of ref 12. The statistical error in the derived entropies was estimated by blocking the data into four equal parts and computation of the entropies for each part, followed by estimation of the standard error.

Results and Discussion

To examine differences in the entropic contribution to binding across the series arising from ligand degrees of freedom using ^{15}N NMR relaxation measurements, a panel of ligands was synthesized using solid-phase synthesis with ^{15}N-enriched glycine as detailed in the Materials and Methods. Two series of ligands were used. The first (series 1) comprised six ligands with different chain lengths ($n = 1-6$), isotopically ^{15}N-labeled at the COOH terminal amide, whereas the second (series 2) comprised six ligands with the same chain length ($n = 6$), but isotopically ^{15}N-labeled at a single amide at each position n.

Chemical Shift Changes on Ligand Binding. Valuable information on intermolecular interactions can be derived from an analysis of chemical shifts, and in particular the binding of a ligand to the binding pocket of BCA II alters the chemical environment around the ligand, and thus will perturb the chemical shifts in the latter. Since each ligand is selectively ^{15}N-labeled, these changes can conveniently be monitored via simple one-dimensional ^{1}H,^{15}N HSQC spectra (data not shown). Shown in Figure 1 are the ^{1}H chemical shift differences (free minus bound) for each ^{15}N-labeled amide in series 1 and 2 ligands binding to BCA II.

These data indicate that residues 1 and 2 (proximal to the aromatic ring) show substantial changes in chemical shift, which is indicative of interactions with the protein. In contrast, smaller changes in the chemical environment of residues 3−6 occur upon binding, suggesting that these residues make less substantive interactions with the protein.

^{15}N NMR Relaxation Measurements. To obtain more quantitative information on ligand dynamics in the bound state, ^{15}N NMR relaxation parameters (R_1, R_2, NOE) were measured at two magnetic field strengths (500 and 600 MHz) for each series of ligands, and these data are shown in Figure 2.

Inspection of the R_2 data (Figure 2b,d) suggests that the Gly$_n$ chain becomes more dynamic as n increases (i.e., progressing from the aromatic ring). However, there are clear differences between series 1 and 2 ligands: series 1 ligands are much more dynamic for small n, which is intuitively not unreasonable given the labeled residues are at the COOH terminus in this series. This notion is supported by ^{15}N{^{1}H} NOE values: the NOE can vary from −3.6 for fast motions ($\omega_N\tau_m \ll 1$) to +0.82 for slow motions ($\omega_N\tau_m \gg 1$), where ω_N is the Larmor frequency of ^{15}N and τ_m is the rotational tumbling time of the molecule.[21] In the case of series 1 ligands, the NOE is negative for all values of n and is thus indicative of fast motions, whereas in the case of series 2 ligands, the NOE is positive for $n = 1-3$ and negative thereafter, indicating that the three glycine residues nearest the aromatic ring adopt slow dynamic motions, whereas the three residues distal to the aromatic ring adopt faster dynamics more characteristic of series 1 ligands.

Model-Free Analysis of Relaxation Data. To obtain more quantitative information on ligand dynamics, the above relaxation data were analyzed using the Lipari−Szabo model-free formalism[22] with the extensions of Clore and co-workers,[23] using the software package *relax*[9,10] (see Materials and Methods), to determine the amplitudes and time scales of the ligand motion. The theoretical basis of the extended model-free formalism and the selection of model-free parameters were described in detail elsewhere[8] and will not be reiterated here. Essentially, *relax* fits the relaxation data to five models of the motion, denoted m1−m5, each containing the following subsets of extended model-free parameters: S^2 (m1); S^2 and τ_e (m2); S^2 and R_{ex} (m3); S^2, τ_e and R_{ex} (m4); S_f^2, S_s^2, and τ_e (m5). Here, $S^2 = S_f^2 S_s^2$ is the square of the generalized order parameter characterizing the amplitude of the internal motions, and S_f^2 and S_s^2 are the squares of the order parameters for internal motions on fast and slow time scales, respectively, τ_e is the internal time scale parameter, and R_{ex} is a chemical exchange contribution. The best model was selected on the basis of Akaike's Information Criterion (AIC).[24] Although all models give rise to a similar tendency of generalized order parameters, m5 was selected according to the AIC criterion. The quality of fit is also confirmed by comparatively low χ^2 values for m5 (Table 1).

(19) MacRaild, C. A.; Daranas, A. H.; Bronowska, A.; Homans, S. W. *J. Mol. Biol.* **2007**, *368*, 822–832.
(20) Best, R. B.; Vendruscolo, M. *J. Am. Chem. Soc.* **2004**, *126*, 8090–8091.
(21) Kay, L. E.; Torchia, D. A.; Bax, A. *Biochemistry* **1989**, *28*, 8972–8979.
(22) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546–4559.
(23) Clore, G. M.; Driscoll, P. C.; Wingfield, P. T.; Gronenborn, A. M. *Biochemistry* **1990**, *29*, 7387–7401.
(24) Akaike, H. Information Theory and an Extension of the Maximum Likelihood Principle. In *Proceedings of the Second International Symposium on Information Theory*; Akademia Kiado: Budapest, Hungary, 1973; pp 267−281.

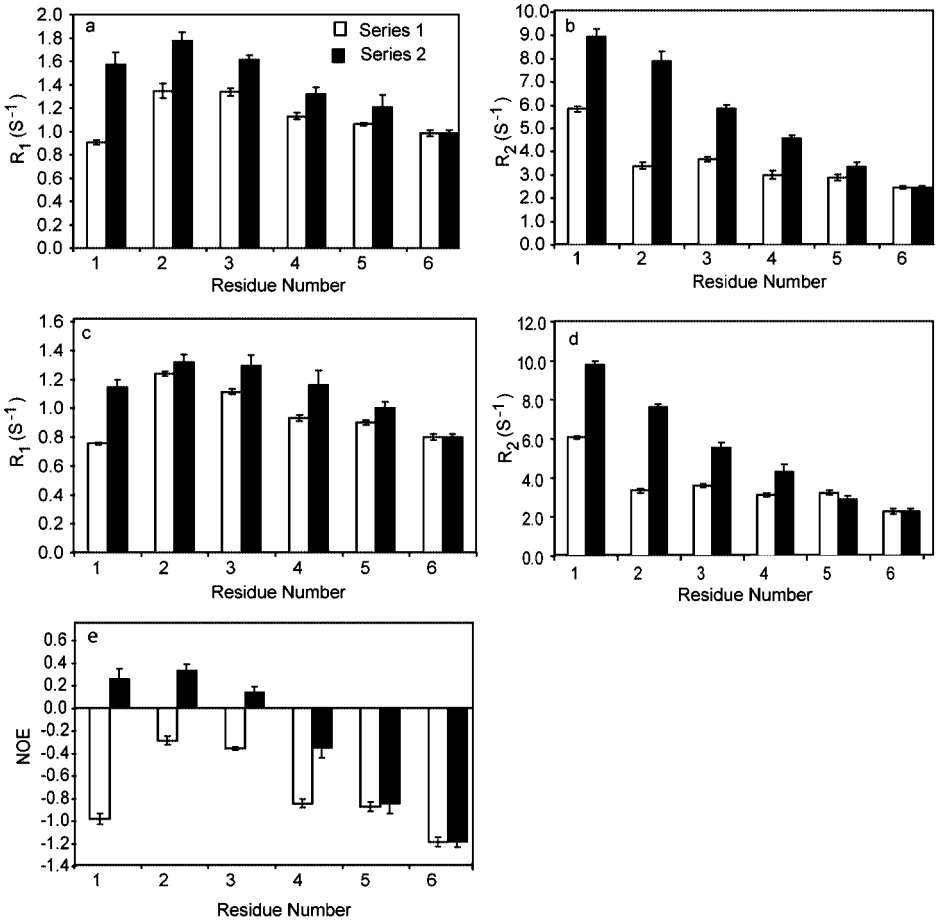

Figure 2. ^{15}N R_1 and R_2 relaxation rates at 500 MHz (a, b), 600 MHz (c, d), and ^{15}N{^{1}H} NOE values at 500 MHz for the glycine side chains of series 1 and 2 ArGly$_n$O$^-$ ligands bound to BCA II.

Table 1. Extended Model-Free Parameters for Series 1 and 2 ArGly$_n$O$^-$ Ligands Bound to BCA II

	residue	S^2	error	S_f^2	error	S_s^2	error	τ_e	error	χ^2
series 1	1	0.33	0.004	0.71	0.04	0.47	0.02	208	28	18.3
	2	0.13	0.006	0.65	0.02	0.19	0.01	1036	98	1.9
	3	0.15	0.005	0.64	0.01	0.24	0.008	949	29	21.9
	4	0.14	0.005	0.62	0.01	0.22	0.009	649	14	33.0
	5	0.14	0.006	0.61	0.008	0.23	0.01	682	18	20.6
	6	0.1	0.004	0.55	0.01	0.19	0.008	599	14	16.2
series 2	1	0.54	0.03	0.81	0.06	0.66	0.05	914	2183	3.9
	2	0.38	0.03	0.75	0.04	0.51	0.04	1668	1296	12.7
	3	0.28	0.01	0.79	0.05	0.35	0.02	1141	299	11.3
	4	0.22	0.01	0.71	0.04	0.31	0.02	877	136	2.0
	5	0.13	0.01	0.63	0.02	0.21	0.02	710	43	8.1
	6	0.1	0.004	0.55	0.01	0.19	0.008	599	14	16.2

The order parameters of all Gly residues in either series are much smaller than those of backbone residues within proteins (typically around 0.8), indicating a comparatively high mobility of the Gly$_n$ chain of the ligand. Comparing the generalized order parameters of residues within series 1, the proximal residue 1 exhibits the highest S^2 of 0.33, thus indicating a relatively low degree of internal motion. The more distal residues 2−6 appear to have an equally high degree of conformational freedom, with an average order parameter of 0.13. The relative rigidity of residue 1 can be explained on the basis of the crystal structure of a BCA II in complex with related ligands.[25,26] Glycine moiety 1 is attached to the arylsulfonamide unit which is, in turn, tightly attached to the zinc ion at the bottom of the BCA binding pocket. Moreover, nonbonded interactions with the protein serve to restrict the motion of this residue. The remaining amide units, however, do not seem to experience substantial nonbonded interactions. Inspection of the order parameters for series 2 shows that the intrachain residues possess a different dynamic behavior: Residue 1 exhibits an S^2 value of ca. 0.54, which then decreases in an approximately exponential fashion to ca. 0.1 for the C-terminal residue. This suggests that the first residue is relatively restricted and that the mobility of the following residues becomes exponentially higher. Furthermore, it is noteworthy that the order parameters of series 2 residues are higher than those of corresponding residues in series 1, suggesting a higher mobility of the terminal glycine residues in series 1. This can be explained by a mutual interaction between the *intrachain* residues in series 2, compared with *terminal* residues in series 1 that are attached only to one vicinal residue. The selection of m5 suggests that motions on two different timescales are relevant in both series of oligoglycine ligands. However, taken together with chemical shift data in Figure 1, the relaxation data for both series suggest a model whereby the Gly$_{1-2}$ residues of the ligand are relatively immobile adjacent to the aromatic ring and are engaged in nonbonded interactions with the protein, whereas the Gly$_{3-6}$

(25) Cappalonga-Bunn, A. M.; Alexander, R. S.; Christianson, D. W. *J. Am. Chem. Soc.* **1994**, *116*, 5063–5068.

(26) Boriack, P. A.; Christianson, D. W.; Kingerywood, J.; Whitesides, G. M. *J. Med. Chem.* **1995**, *38*, 2286–2291.

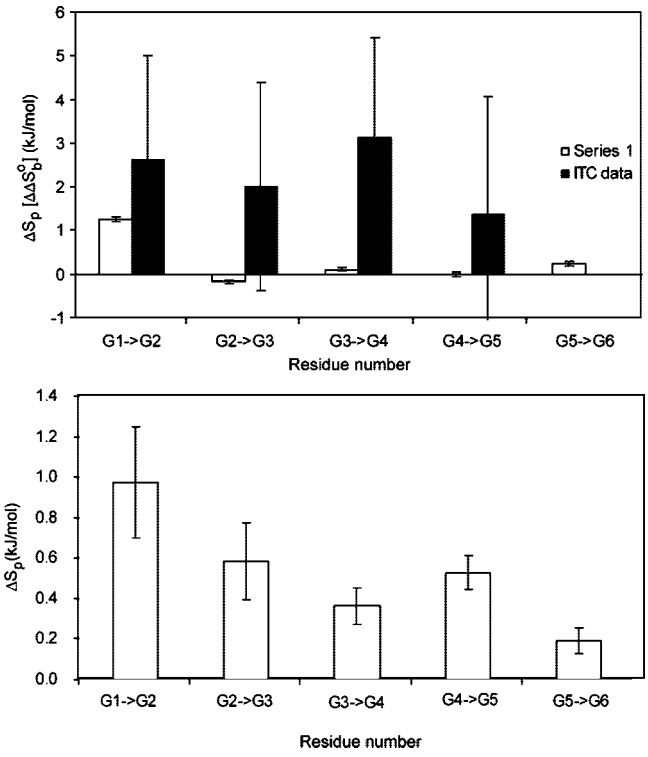

Figure 3. (Top) NMR-derived per-residue entropies for $^{15}N-^{1}H$ bond vectors and experimentally derived standard entropies of binding[3] for series 1 ligands bound to BCA II. (Bottom) NMR-derived per-residue entropies for $^{15}N-^{1}H$ bond vectors in series 2 $ArGly_nO^-$ ligands bound to BCA II.

residues are considerably more mobile, with S^2 values consistent with motion unrestricted by the protein.

Per-Residue Entropic Contributions to Binding of Gly_n Chains. In an effort to correlate the motional properties of the Gly_n chains in each ligand series with ligand binding thermodynamics, per-residue conformational entropies were estimated from the S^2 values listed in Table 1 by use of the relation derived by Yang and Kay (eq 1).[12] The entropic contribution to binding from each glycine residue in ligand series 1 or 2 can thus be calculated from eq 1 using the relevant S^2 values from Table 1, resulting in the data shown in Table 1s and Figure 3. Also included with these data are the global standard entropies of binding for the interaction of the $ArGly_nO^-$ ligand series 1 ($n = 1-5$) determined from ITC measurements at 298 K reported by Krishnamurthy et al.[3]

From these data, it is clear that a poor correlation exists between the change in ligand conformational entropy determined from NMR relaxation measurements above versus the global entropies of binding derived from ITC. This is not entirely unexpected since the global entropies derived from ITC data include contributions not only from the ligand, but from protein and solvent.

Ligand Dynamics Does Not Explain the Observed Thermodynamic Binding Signature. A rigorous theoretical description of the binding process can be formulated by use of conventional Born−Haber thermodynamic cycles,[27−29] leading to the following expression for the standard entropy of binding:

$$\Delta S_{obs}^\circ = \Delta S_i^\circ + [\Delta S_{sb}^\circ - \Delta S_{su}^\circ] \quad (3)$$

where ΔS_{obs}° is the observed global entropy of binding (as determined by ITC), ΔS_i° is the change in "intrinsic" (solute−solute) degrees of freedom on binding, and ΔS_{sb}° and ΔS_{su}° are the

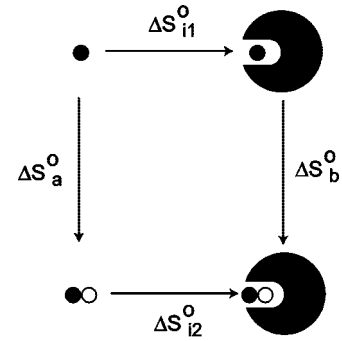

Figure 4. Thermodynamic cycle for the binding of two adjacent ligands in the $ArGly_nO^-$ series. In this example, $ArGly_2O^-$ and $ArGly_1O^-$ are illustrated, where the unlabeled and labeled glycine units are shown diagrammatically by open and filled circles, respectively.

standard entropies of solvation of the bound and unbound species, respectively. In the present context where we consider *differences* in binding thermodynamics between two related ligands, eq 4 can be derived from eq 3:[30]

$$\Delta S_{obs2}^\circ - \Delta S_{obs1}^\circ = [\Delta S_{i2}^\circ - \Delta S_{i1}^\circ] + \{[\Delta S_{sb2}^\circ - \Delta S_{sb1}^\circ] - [\Delta S_{sl2}^\circ - \Delta S_{sl1}^\circ]\} \quad (4)$$

In this equation, the subscripts 1 and 2 refer to adjacent ligands in the series, and the terms ΔS_{sl1}° and ΔS_{sl2}° are the standard hydration entropies of the free ligands; the standard solvation entropy of the free protein exactly cancels. Ignoring for the moment the solvation contribution, the solute−solute contribution $[\Delta S_{i2}^\circ - \Delta S_{i1}^\circ]$ can conveniently be understood in terms of the thermodynamic cycle shown in Figure 4.

If ligand dynamics are responsible for the observed binding signature, the *change* in entropy of a $Gly_{(n+1)}$-bearing ligand on binding would need to be greater (more positive) than the change in entropy of a $Gly_{(n)}$ ligand on binding, that is, $[\Delta S_{i2}^\circ - \Delta S_{i1}^\circ] = \Delta S_b^\circ - \Delta S_a^\circ$ must be greater than zero. The only mechanism by which this can occur requires that the addition of a Gly unit in the bound state destabilizes preceding Gly units, as was recognized by Krishnamurthy et al.[3] However, it is clear from comparison of the bound-state dynamics of series 1 and 2 ligands, which can most readily be appreciated by considering S^2 values (Figure 5), that the addition of Gly units to the bound-state ligand *reduces* the dynamics of preceding residues. Consequently, a model based on increased dynamics of the ligand in the bound state with respect to Gly chain length is not a plausible explanation for the observed thermodynamic binding signature.

Possible Role of Solvation. A second possible source of the observed thermodynamic signature arises from the solvation terms in eq 3. However, two observations suggest that these do not contribute significantly to the observed increase in entropy with respect to chain length. First, ΔC_p values for the interaction determined by Krishnamurthy et al. are uncharacteristically small (~80 J/mol/K) and are essentially independent of Gly chain length.[3] Second, while Gly residues proximal to the benzene

(27) Chervenak, M. C.; Toone, E. J. *J. Am. Chem. Soc.* **1994**, *116*, 10533–10539.
(28) Daranas, A. H.; Shimizu, H.; Homans, S. W. *J. Am. Chem. Soc.* **2004**, *126*, 11870–11876.
(29) Barratt, E.; Bingham, R.; Warner, D. J.; Laughton, C. A.; Phillips, S. E. V.; Homans, S. W. *J. Am. Chem. Soc.* **2005**, *127*, 11827–11834.
(30) Malham, R.; Johnstone, S.; Bingham, R. J.; Barratt, E.; Phillips, S. E. V.; Laughton, C. A.; Homans, S. W. *J. Am. Chem. Soc.* **2005**, *127*, 17061–17067.

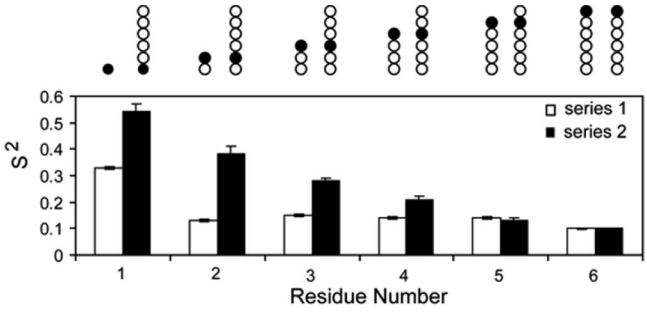

Figure 5. NMR-derived order parameters for $^{15}N-^{1}H$ bond vectors in series 1 and 2 $ArGly_nO^-$ ligands bound to BCA II. Shown above the figure in diagrammatic form are the corresponding ligand Gly_n chains, where the unlabeled and labeled glycine units are shown diagrammatically by open and filled circles, respectively.

ring clearly interact with the protein in crystal structures[25,26] and might be supposed to give rise to desolvation of the protein binding pocket and the ligand, more distal residues that extend beyond the binding pocket can hardly do so, and the approximately linear increase in entropy with respect to chain length thus strongly argues against a significant solvation contribution.

Protein Dynamics Are Influenced by Ligand Chain Length. Given that the source of the observed thermodynamic signature cannot be explained by ligand dynamics or solvation, the conclusion that the observed binding thermodynamics must involve favorable changes in protein entropy with increasing ligand chain length appears to be inescapable. Possible destabilization of the protein by bound ligand was acknowledged by Krishnamurthy et al., but these authors found no evidence for or against this model in their recent study. In principle, such changes can be probed experimentally using NMR relaxation measurements[12,31] on the various complexes. However, these changes are likely to be quite small and will be manifest principally in protein side chains that interact with the ligands. Robust methods have been developed for the estimation of protein methyl-containing side-chain entropies by measurement and cross-validation of ^{2}H relaxation parameters for protein side chains,[32,33] and recent developments permit such methods to be applied to Val, Leu, and Ile residues in complexes in excess of 30 kDa.[34] However, the dearth of methyl-containing residues in the binding pocket of BCA II (one Val and one Leu) renders this approach untenable. In principle, NMR methods could be developed to probe ^{13}C relaxation for a more complete set of binding pocket side chains, but as described by Muhandiram et al.,[32] the interpretation of such data is fraught with difficulty, particularly in the study of complexes where the relaxation of ^{13}C may be influenced substantially by "external" ligand protons.

Thus, to establish whether favorable changes in protein dynamics correlate with ligand chain length as suggested by experimental observations, we resorted to all-atom molecular dynamics simulations on series 1 and 2 ligands in complex with BCA II with explicit inclusion of solvent water. As a test of convergence, squared order parameters for amide vectors in each ligand were calculated from these trajectories (see Materials and Methods),[20] and these are reported in Table 2, where for convenience the experimental S^2 values for the respective vectors are also shown (taken from Table 1).

Table 2. Squared Generalized Order Parameters for Series 1 and 2 Ligands in Complex with BCA II, Derived from All Atom MD Simulations with Explicit Inclusion of Solvent Water

		bound	
	residue	S^2 (theor)	S^2 (exptl)
series 1	1	0.38	0.33 (0.04)
	1	0.38	
	2	0.24	0.13 (0.02)
	1	0.40	
	2	0.19	
	3	0.10	0.15 (0.01)
	1	0.44	
	2	0.26	
	3	0.16	
	4	0.10	0.14 (0.01)
	1	0.48	
	2	0.27	
	3	0.17	
	4	0.11	
	5	0.08	0.14 (0.01)
	1	0.54	
	2	0.32	
	3	0.22	
	4	0.16	
	5	0.09	
	6	0.03	0.1 (0.01)
series 2	1	0.54	0.54 (0.06)
	2	0.32	0.38 (0.04)
	3	0.22	0.28 (0.05)
	4	0.16	0.22 (0.04)
	5	0.09	0.13 (0.02)
	6	0.03	0.1 (0.01)

From these data, it is immediately apparent that the theoretically derived S^2 values for series 2 ligands in the complex are in very close agreement with the experimental values. However, this is not the case for series 1 ligands in the complex, where the computed S^2 values for residues 1−2 are significantly higher than experimental values. Indeed, the theoretical S^2 values for these residues have not converged after 45 ns (based on the test for convergence described in the Materials and Methods). Nonetheless, the qualitative trend in S^2 in the simulations of both ligand series in the complex mirrors that observed experimentally.

These MD data were used to probe the influence of ligand binding on protein dynamics. Specifically, S^2 values were computed for each terminal heavy-atom bond vector over the whole protein and for each side chain in the binding pocket for each complex with series 1 ligands, from which conformational entropies were computed,[12] and the results are collated in Table 3.

It is inappropriate to interpret the MD data in Table 3 quantitatively, since while the estimated statistical errors are quite low, additional errors may exist in the force-field parametrization. However, qualitatitively it is clear from the data in Table 3 that the arylsulfonamide moiety becomes correspondingly more rigid with respect to series 1 ligand chain length, as expected given that the addition of successive glycine residues decreases the dynamics of preceding units as shown from the NMR measurements above. Moreover, there exists a qualitative trend of increased dynamics of residue side chains across the protein, much of which is manifest in the binding-pocket residues including the coordinating histidines, and increased

(31) Li, Z. G.; Raychaudhuri, S.; Wand, A. J. *Protein Sci.* **1996**, *5*, 2647–2650.
(32) Muhandiram, D. R.; Yamazaki, T.; Sykes, B. D.; Kay, L. E. *J. Am. Chem. Soc.* **1995**, *117*, 11536–11544.
(33) Millet, O.; Muhandiram, D. R.; Skrynnikov, N. R.; Kay, L. E. *J. Am. Chem. Soc.* **2002**, *124*, 6439–6448.
(34) Tugarinov, V.; Ollerenshaw, J. E.; Kay, L. E. *J. Am. Chem. Soc.* **2005**, *127*, 8214–8225.

Table 3. Differences in Per-Residue Entropies ($ArGly_{n+1}O^-$ Complex Minus $ArGly_nO^-$ Complex) of Terminal Heavy-Atom Bond Vectors Measured as $T\Delta S_p$ (in Kilojoules Per Mole, or Atomic Fluctuations in Angstroms in the Case of the Zinc Atom) for All Residues and for Residue Side Chains in the Binding Pocket of BCA II Series 1 Ligand Complexes

	series 1 ligand chain length ($n + 1 - n$)				
residue	2−1	3−2	4−3	5−4	6−5
Zn	0.27	0.39	0.52	0.70	0.82
Ar-sulf	−2.57	−0.80	−0.85	−5.59	−3.74
His92	1.00	0.94	0.86	0.32	0.17
His94	0.20	0.31	0.34	1.18	1.25
His117	0.08	1.52	0.56	0.17	0.70
Glu104	1.38	0.69	0.33	0.65	0.57
His105	0.38	0.21	0.09	0.13	2.08
Trp3	0.17	0.77	0.13	0.69	0.20
Asn60	0.01	0.42	0.1	0.08	0.26
Asn65	0.66	0.43	0.13	0.48	1.33
Gln90	1.64	0.63	0.50	1.13	1.36
Leu195	0.56	−0.04	0.08	1.05	0.33
Thr196	0.08	0.03	0.42	0.78	0.42
Thr197	0.63	0.08	0.35	0.96	0.22
Val204	0.17	0.09	1.26	1.08	0.88
total	4.37 ± 1.1[a]	5.28 ± 1.2	4.33 ± 1.0	3.11 ± 1.0	6.04 ± 1.3
all residues	14.9 ± 1.7	4.6 ± 1.8	5.5 ± 2.2	9.9 ± 2.4	8.4 ± 2.5

[a] Statistical errors in computed parameters determined as described in Materials and Methods.

atomic displacement of the zinc ion, with respect to ligand chain length. The observation that ligand binding increases protein dynamics is counterintuitive but has nevertheless been observed in a number of systems.[19,35−38] The small size of the changes with respect to ligand chain length for individual residues underpins the difficulty in obtaining quantitative data from the simulations or indeed from experimental NMR relaxation measurements.

In summary, our investigations suggest that the observed thermodynamic signature for binding of $ArGly_nO^-$ ligands to BCA II derives principally from an increase in protein dynamics, rather than ligand dynamics, with respect to Gly_n chain length. One possible caveat with this observation is that while our NMR measurements and dynamics simulations are probing picosecond−nanosecond motions, it is not impossible that slower time scale motions are superimposed on the fast motions, but which are "invisible" to the techniques we employed. However, given that the S^2 values for residues distal to the aromatic ring of the ligand are small, this is very unlikely.

Krishnamurthy et al. showed that enthalpy−entropy compensation was observed for a range of arylsulfonamides with oligoglycine and oligosarcosine and oligoethylene glycol chains.[3] That structurally distinct chain types give a similar thermodynamic signature suggests that a common process is underway that is unlikely to be related to specific interactions between the chain and the protein. In this work, we demonstrated a decrease in the dynamics of the proximal parts of the ligand upon increasing oligomer length. This increase in order results from motion of the proximal part of the ligand being constrained by the motions of the other residues in the oligomeric chain and thus it is likely that a similar effect is observed in the oligosarcosine and oligo PEG-substituted sulfonamides studied by Krishnamurthy et al. Increased dynamics of the protein results from decreased mobility of the proximal portion of the ligand, providing a pathway for the observed enthalpy−entropy compensation across these structurally distinct molecules.

Acknowledgment. This work was supported by BBSRC, Grant No. BB/E014844/1 to S.W.H., The Wellcome Trust, Grant No. 075520 to S.W.H., and Astra Zeneca (scholarship to H.S.).

Supporting Information Available: Complete refs 13 and 16, Table 1s, and details of ligand synthesis. This material is available free of charge via the Internet at http://pubs.acs.org.

JA803755M

(35) Popovych, N.; Sun, S.; Ebright, R. H.; Kaladimos, C. G. *Nat. Struct. Mol. Biol.* **2006**, *13*, 831–838.
(36) Yu, L. P.; Zhu, C. X.; TseDinh, Y. C.; Fesik, S. W. *Biochemistry* **1996**, *35*, 9661–9666.
(37) Arumugam, S.; Gao, G.; Patton, B. L.; Semenchenko, V.; Brew, K.; Van Doren, S. R. *J. Mol. Biol.* **2003**, *327*, 719–734.
(38) Yun, S.; Jang, D. S.; Kim, D. H.; Choi, K. Y.; Lee, H. C. *Biochemistry* **2001**, *40*, 3967–3973.