Role of Hydrogen Bonding in the Interaction between a Xylan Binding Module and Xylan†

Hefang Xie,‡ David N. Bolam,‡ Tibor Nagy,‡ Lóránd Szabó,§ Alan Cooper,|| Peter J. Simpson,⊥ Jeremy H. Lakey,@ Michael P. Williamson,*,⊥ and Harry J. Gilbert‡

Department of Biological and Nutritional Sciences, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU, U.K., Gödöllô University of Agricultural Sciences, Páter Károly u. I., Gödöllô H-2101, Hungary, Department of Chemistry, Joseph Black Building, University of Glasgow, Glasgow G12 8QQ, U.K., Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield S10 2UH, U.K., and Department of Biochemistry and Genetics, University of Newcastle upon Tyne, Newcastle upon Tyne NE4 4HH, U.K.

Received January 4, 2001; Revised Manuscript Received March 7, 2001

ABSTRACT: NMR studies of the internal family 2b carbohydrate binding module (CBM2b-1) of *Cellulomonas fimi* xylanase 11A have identified six polar residues and two aromatic residues that interact with its target ligand, xylan. To investigate the importance of the various interactions, free energy and enthalpy changes have been measured for the binding of xylan to native and mutant forms of CBM2b-1. The data show that the two aromatic residues, Trp 259 and Trp 291, play a critical role in the binding, and similarly that mutants N264A and T316A have no affinity for the xylose polymer. Interestingly, mutations E257A, Q288A, N292A, E257A/Q288A, E257A/N292A, and E257A/N292A/Q288A do not significantly diminish the affinity of CBM2b-1 for the xylose polymers, but do influence the thermodynamics driving the protein−carbohydrate interactions. These thermodynamic parameters have been interpreted in light of a fresh understanding of enthalpy−entropy compensation and show the following. (1) For proteins whose ligands are bound on an exposed surface, hydrogen bonding confers little specificity or affinity. It also displays little cooperativity. Most specificity and affinity derive from binding between the face of sugar rings and aromatic rings. (2) Loss of hydrogen bonding interactions leads to a redistribution of the remaining bonding interactions such that the entropic mobility of the ligand is maximized, at the expense (if necessary) of enthalpically favorable bonds. (3) Changes in entropy and enthalpy in the binding between polysaccharide and a range of mutants can be interpreted by considering changes in binding and flexibility, without any need to consider solvent reorganization.

The plant structural polysaccharides cellulose and xylan, which comprise β-1,4-linked polymers of glucose and xylose, respectively, are the most abundant organic molecules in the biosphere (*1*). Cellulose consists mainly of crystalline fibrils, where it forms planar sheets connected by hydrogen bonds, in which the rotation between one monomer and the next is 180° (*2*). By contrast, xylan is generally thought to exist in a helical conformation with a 120° rotation between monomers (*3*). The hydrolysis of these polymers is essential for the recycling of photosynthetically fixed carbon and is therefore of fundamental biological importance. In addition, plant cell wall hydrolyzing enzymes have numerous applications in the animal feed, fruit juice, and paper pulp industries. Aerobic microbial enzymes that hydrolyze cellulose and xylan are typically modular and contain a catalytic module joined by flexible linker sequences to one or more noncatalytic carbohydrate-binding modules (CBMs)[1] (*2*). CBMs, by binding to the plant cell wall, bring these enzymes into prolonged and intimate proximity with their target substrates, thereby increasing the efficiency of catalysis (*4*−*6*). In addition, family 3c CBMs abut directly onto the substrate binding cleft of some processive cellulases, and feed the cellulose chains directly into the active site (*7*, *8*). Given the central role that CBMs play in the recycling of photosynthetically fixed carbon, the mechanism by which these modules bind to their polysaccharide ligands is therefore a topic of considerable biological and industrial importance.

There are now structural details for a number of CBMs (*9*), and in some cases, there are thermodynamic details of

† This work was suppported by a BBSRC CASE studentship (P.J.S.) and BBSRC project grants. We thank the BBSRC and the Wellcome Trust for equipment grants. The Krebs Institute is a Centre of the Biomolecular Sciences Panel of the BBSRC and EPSRC. M.P.W. and P.J.S. are members of the BBSRC-funded North of England Structural Biology Centre (NESBiC).

* To whom correspondence should be addressed: Department of Molecular Biology and Biotechnology, University of Sheffield, Firth Court, Western Bank, Sheffield S10 2UH, U.K. Telephone: +44 114 222 4224. Fax: +44 114 272 8697. E-mail: M.Williamson@Sheffield.ac.uk.

‡ Department of Biological and Nutritional Sciences, University of Newcastle upon Tyne.

§ Gödöllô University of Agricultural Sciences.

|| University of Glasgow.

⊥ University of Sheffield.

@ Department of Biochemistry and Genetics, University of Newcastle upon Tyne.

[1] Abbreviations: CBM, cellulose binding module; CBM2b-1, internal family 2b CBM of *C. fimi* xylanase 11A; GST, glutathione *S*-transferase; ITC, isothermal titration calorimetry.

10.1021/bi010034z CCC: $20.00
Published on Web 04/18/2001

their binding to polysaccharide ligands (*10*−*12*). However, the thermodynamics of protein−polysaccharide binding are by no means well understood, particularly for CBMs, which have more exposed binding sites than the lectins and monosaccharide transporters that have been the most studied representatives of this class of interaction to date. Therefore, we have undertaken a thermodynamic analysis, using site-directed mutagenesis, of a CBM for which good structural data exist. The target CBM of this study is the internal CBM2b (CBM2b-1) of *Cellulomonas fimi* Xyn11A (formerly xylanase D) (*13*). On the basis of the model for the complex between CBM2b-1 and xylohexaose (*14*), eight amino acid side chains of the protein appear to interact with its target ligand. To evaluate the importance of these amino acids in the capacity of the protein to bind carbohydrates, the residues have been mutated and the affinity of the mutant proteins for xylose polymers, and the thermodynamic forces driving these binding events, have been measured. The data were used to assess the importance of hydrogen bonding, and to develop a model for protein−polysaccharide binding that is generally applicable to a wide range of protein−ligand interactions.

EXPERIMENTAL PROCEDURES

Production and Purification of Proteins. Production of the wild-type glutathione *S*-transferase (GST)−CBM2b-1 fusion used in this study has been described previously (*14*). To produce His_{10}-tagged versions of CBM2b-1, the region of full-length *C. fimi* Xyn11A that encodes the module was amplified from pCF9 (*13*) by PCR using the following primers: 5′-CGTCATATGGACACGGGCGGAGGCGGCG-GC-3′ and 5′-GCCGGATCCTTAGCTGCCCGCGCACGT-CGC-3′, which contain *Nde*I and *Bam*HI restriction sites, respectively, at their 5′ ends. The reactions were performed using $Vent_R$ DNA polymerase (New England BioLabs) according to the manufacturer's protocol. The amplified DNA was ligated into pCR-Blunt (Invitrogen), excised using *Nde*I and *Bam*HI, and cloned into pET16b (Novagen) for expression.

Mutagenesis of both His_{10}-tagged and GST fusions of CBM2b-1 was carried out with the Transformer site-directed mutagenesis kit (Clontech) according to the manufacturer's protocol, using the following primers: E257A, 5′-GCGACGC-GCGCGGCCGAGTGGTCGGAC-3′; W259A, 5′-CGCGC-CGAGGCCTCGGACCGCTTC-3′; D261A, 5′-GCCGAG-GAGTGGAGCGCTCGCTTCAACGTC-3′; N264A, 5′-TC-GGACCGCTTCGCCGTCACGTACTCC-3′; Q288A, 5′-AG-CCAGACCATCGCGGCGTCGTGGAAC-3′; W291A, 5′-AC-CATGCAGGCGAGCGCTAACGCGAACGTC-3′; N292A, 5′-CAGGCGTCGTGGGCCGCGAACGTCACC-3′; and T316A, 5′-ACCTTCGGCGTGGCCGTCATGAAGAAC-3′.

Expression and Purification of Proteins. The *Escherichia coli* strains used in this study were JM83 (*4*) and BL21 (DE3):pLysS (*14*). Recombinant *E. coli* was cultured in Luria broth (2.4 L) supplemented with 100 μg/mL ampicillin at 30 (JM83) or 37 °C (BL21) with aeration (180 rpm). The cells were grown to mid-log phase before gene expression was induced by the addition of 0.5 (JM83) or 1 mM (BL21) isopropyl β-D-galactopyranoside, and the cells were incubated for a further 3−4 h. JM83 was used to express the GST fusion proteins (wild-type CBM2b-1 and single and double mutants), while the His-tagged CBMs (wild-type and triple mutant) were expressed in BL21 (DE3):pLysS. Cells were harvested by centrifugation (4500*g*) for 10 min at 4 °C and resuspended in $^1/_{40}$ of the volume of either PBS [140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, and 1.8 mM KH_2PO_4 (pH 7.3) for GST fusions] or 50 mM sodium phosphate buffer (pH 7.0) for His-tagged proteins before being ruptured by sonication and centrifuged at 25000*g* for 15 min at 4 °C to produce cell-free extract (supernatant) and insoluble cell material (pellet).

GST fusion proteins were purified by glutathione affinity chromatography as follows. The cell-free extract (20 mL) was passed through a 5 mL bed volume Glutathione Sepharose-4B (Pharmacia) column and washed with 30 mL of PBS, and the purified GST fusion eluted with 15 mL of 20 mM glutathione in 50 mM Tris-HCl (pH 8.0). The proteins were then dialyzed against 2 × 2 L of 50 mM sodium phosphate buffer (pH 7.0) and stored at −20 °C.

His-tagged proteins were purified from inclusion granules as follows. Insoluble cell material was solubilized by resuspending it in 50 mM sodium phosphate buffer (pH 7.0) containing 8 M urea. The resolubilized protein was centrifuged at 30000*g* for 30 min at 20 °C to remove any cell debris that was not solubilized by the urea. The supernatant (20 mL) was then applied to a 5 mL bed volume immobilized metal affinity column (TALON, Clontech). This was washed with 30 mL of 20 mM Tris-HCl (pH 8.0) containing 100 mM NaCl and 8 M urea (buffer A), and the purified protein was eluted with 15 mL of 100 mM imidazole in buffer A. The CBMs were dialyzed successively against 2 L of 50 mM sodium phosphate buffer (pH 7.0) containing 4 M, 3 M, 2 M, 1 M, and 500 mM urea and then twice against 2 L of 50 mM sodium phosphate buffer (pH 7.0) and stored at −20 °C.

The concentration of the purified protein was determined by UV absorbance at 280 nm, using a calculated molar extinction coefficient of 18 470 M^{-1} cm^{-1} for His-tagged CBMs and 59 470 M^{-1} cm^{-1} for GST fusions.

Circular Dichroism (CD) and Fluorescence Spectroscopy. CD spectra were recorded with a Jobin-Yvon CD6 spectropolarimeter. The spectra were obtained at a protein concentration of 0.5 μg/mL in 100 mM Tris-HCl buffer (pH 6.8) at 25 °C using a 0.1 cm path length quartz cuvette (Hellma). Each spectrum was accumulated from 20−30 scans between 188 and 250 nm, at a scan rate of 60 nm/min. Fluorescence spectra were collected on an SLM 8100 fluorimeter operating in the ratio mode with excitation and emission bandwidths set to 8 nm. The excitation wavelength was 295 nm to selectively excite tryptophan. Spectra were obtained at a protein concentration of 0.6 mg/mL in 100 mM Tris-HCl (pH 6.8) at 22 °C in a 5 mm × 5 mm quartz cuvette (Hellma). Excitation and emission polarizers were set at horizontal and vertical orientations, respectively, to remove scattered light interference. The small peaks on each spectrum at 390 nm are due to a Woods anomaly in the emission monochromator.

Affinity Gel Electrophoresis. The capacity for wild-type CBM2b-1 and mutants to bind to soluble oat spelt xylan was evaluated by affinity electrophoresis. Continuous native polyacrylamide gels consisting of 7.5% (w/v) acrylamide in 25 mM Tris/250 mM glycine buffer (pH 8.3) were prepared. To one of the gels was added 0.1% soluble oat spelt xylan

(Sigma) prior to polymerization. Approximately 10 μg of target proteins and GST (as a noninteracting negative control) were loaded on the gels and subjected to electrophoresis at 10 mA/gel for approximately 2 h at room temperature. Proteins were visualized by staining with Coomassie Blue.

Isothermal Titration Calorimetry (ITC). ITC measurements were taken at 25 °C using a MicroCal Omega titration calorimeter. Proteins were dialyzed extensively against 50 mM sodium phosphate buffer (pH 7.0), and the ligand was dissolved in the dialysis buffer. During a titration, the protein sample, stirred at 400 rpm in a 1.3586 mL reaction cell, was injected with 25−50 successive 10 μL aliquots of a stock solution of soluble oat spelt xylan (Sigma Chemical Co.; 14 mg/mL) at 200 s intervals. Given that the protein binds to six successive xylose residues (*14*), the ligand was deemed to comprise nonoverlapping xylohexaose units within the xylan molecules. Thus, the ligand concentration was taken to be $^1/_6$ of the xylose concentration in the xylan which equals 14.7 mM. The binding data were corrected for the heat of dilution of both proteins and ligands. Integrated heat effects were analyzed by nonlinear regression using a single-site binding model (MicroCal ORIGIN), yielding independent values for K_a and $\Delta H°$. Other thermodynamic parameters were derived from the equation $-RT \ln K_a = \Delta G° = \Delta H° - T\Delta S°$. The protein concentrations that were used were between 325 and 980 μM, and all gave *c* values (product of K_a and the total protein concentration in the cell) of >1, thus ensuring accurate deconvolution of the binding isotherm data (*15*).

RESULTS

Identification and Characterization of Ligand-Binding Residues of CBM2b-1. The solution structure of CBM2b-1 (*14*) has identified several residues that contribute to the ligand binding site of the protein. Titration of the CBM with xylohexaose resulted in significant changes in the chemical shifts of Asn 292, Gln 288, and Glu 257 (*14*), strongly indicating that they participate in hydrogen bonds with the ligand. Thus, the ^{15}N and ^{1}H chemical shifts in the side chain amide groups change on binding of xylohexaose by 1.16 and 0.10 ppm, respectively, for Asn 292 and by 2.63 and 0.26 ppm, respectively, for Gln 288. For Glu 257, the backbone ^{15}N and ^{1}H chemical shift changes are 1.16 and 0.33 ppm, respectively. A model of CBM2b-1 bound to xylohexaose has been constructed on the basis of these results (Figure 1), in which the following amino acids appear to interact with the ligand: Glu 257, Trp 259, Asp 261, Asn 264, Gln 288, Trp 291, Asn 292, and Thr 316. To investigate the importance of these residues in binding xylose polymers, each of these amino acids was mutated to alanine. In addition, the double mutants Q288A/N292A, Q288A/E257A, and N292A/E257A and the triple mutant E257A/Q288A/N292A were also made. The 12 mutant proteins were expressed as GST fusion proteins (wild-type protein and CBM2b derivatives containing single or double mutations) or as discrete modules containing an N-terminal His_{10} tag (wild-type protein and CBM2b with the triple mutation). Proteins were all produced in reasonable amounts except for D261A and E257A/N292A. A small amount of the double mutant could be obtained, but all attempts to produce D261A, either as a His_{10}-tagged protein or as a fusion with GST, were unsuccessful. All the proteins were purified by affinity chromatography to apparent homogeneity as judged by SDS−PAGE (Figure 2).

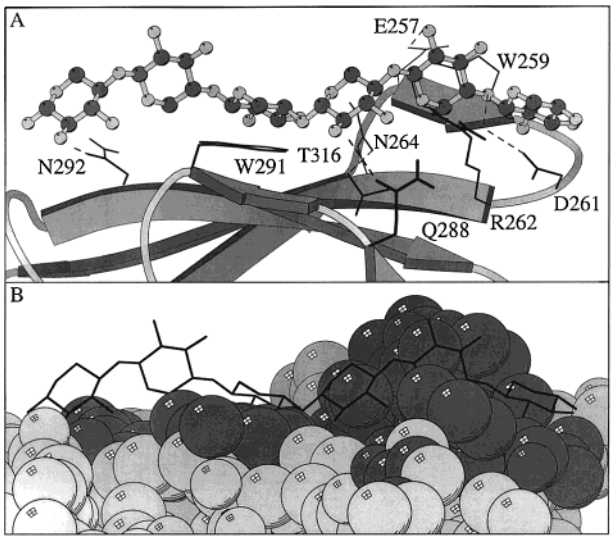

FIGURE 1: Main interactions expected between CBM2b-1 and xylan (*14*). (A) Side chains implicated in the recognition of xylooligosaccharides. (B) A space-filling representation of the complex, showing the exposed surface available to much of the xylan chain.

The binding of wild-type and mutant proteins was analyzed qualitatively using gel affinity electrophoresis. Native gels were run in the presence of 0.1% oat spelt xylan. Mutants W259A, N264A, W291A, and T316A showed no detectable binding, and mutant D261A could not be analyzed due to the very small amounts of this protein that were produced. Thus, only mutants E257A, Q288A, and N292A had significant binding affinity, as did the double mutants and the triple mutant involving these residues (Figure 3). Approximate values for the binding affinities could be derived by running the gels using different concentrations of xylan (data not shown), which suggested that the affinities of all of this group of mutants were close to that of the wild type.

Circular dichroism (CD) and fluorescence spectroscopy were used to evaluate the structural integrity of the CBM2b-1 mutants in which polar residues were substituted with alanine and that did not bind to xylan, namely, N264A and T316A. The CD spectra of N264A and T316A were similar to that of the wild type (Figure 4a), indicating that these amino acid substitutions did not significantly alter the secondary structure of CBM2b-1. However, their fluorescence spectra were different from that of native CBM2b-1 (Figure 4b). N264A and T316A exhibited higher and lower fluorescent intensities, respectively, than the wild-type protein, suggesting that there is some perturbation of a tryptophan side chain(s) in each mutant.

Isothermal Titration Calorimetry. Thermodynamic parameters were obtained for wild-type CBM2b-1, the three mutants (Q288A, N292A, and E257A) that exhibited measurable affinity for xylan, the two double mutants that could be expressed at reasonable levels (Q288A/E257A and Q288A/N292A), and the triple mutant (Figure 5). The interpretation of the ITC data is based on the premise that CBM2b-1 binds to six successive xylose units within xylan, and thus, the ligand binding site comprises nonoverlapping xylohexaose units within the xylan chains (*14*). This premise is based on three main observations. (a) NMR spectroscopic

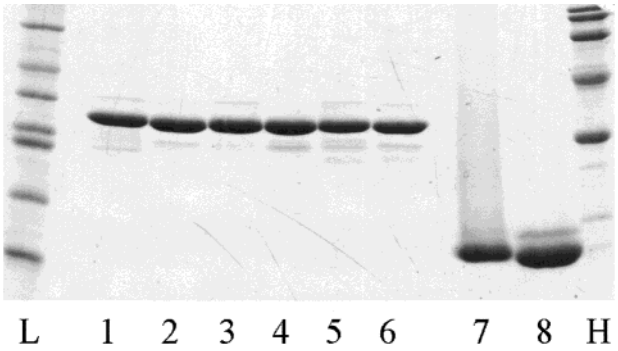

FIGURE 2: SDS−PAGE of wild-type and mutant forms of CBM2b-1. Proteins were electrophoresed in a 10% (w/v) polyacrylamide gel. Samples are GST fusions of the wild type (lane 1), E257A (lane 2), N292A (lane 3), Q288A (lane 4), Q288A/N292A (lane 5), E257A/Q288A (lane 6), the His-tagged wild type (lane 7), and E257A/Q288A/N292A (lane 8). Lanes L and H contain Sigma low- and high-molecular weight markers, respectively.

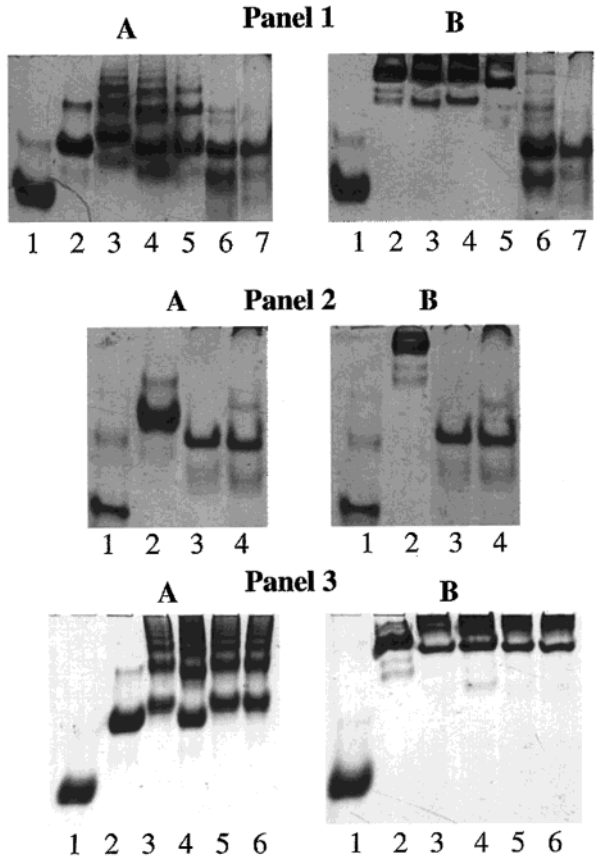

FIGURE 3: Affinity gel electrophoresis of wild-type CBM2b-1 and mutants. GST fusions were electrophoresed in nondenaturing polyacrylamide gels containing no polysaccharide (gel A) or 0.1% soluble oat spelt xylan (gel B). Panel 1: GST (lane 1), wild type (lane 2), E257A (lane 3), N292A (lane 4), Q288A (lane 5), N264A (lane 6), and T316A (lane 7). Panel 2: GST (lane 1), wild type (lane 2), W259A (lane 3), and W291A (lane 4). Panel 3: GST (lane 1), wild type (lane 2), E257A/N292A (lane 3), Q288A/N292A (lane 4), E257A/Q288A (lane 5), and E257A/Q288A/N292A (lane 6).

studies of CBM2b-1 titrated with xylohexaoase showed that the protein has only one binding site for the oligosaccharide (*14*). In addition, the three-dimensional structure of the protein showed that the binding site (which comprised the full length of the protein) could accommodate up to six xylose moieties (*14*). (b) The ITC data show a stoichiometry

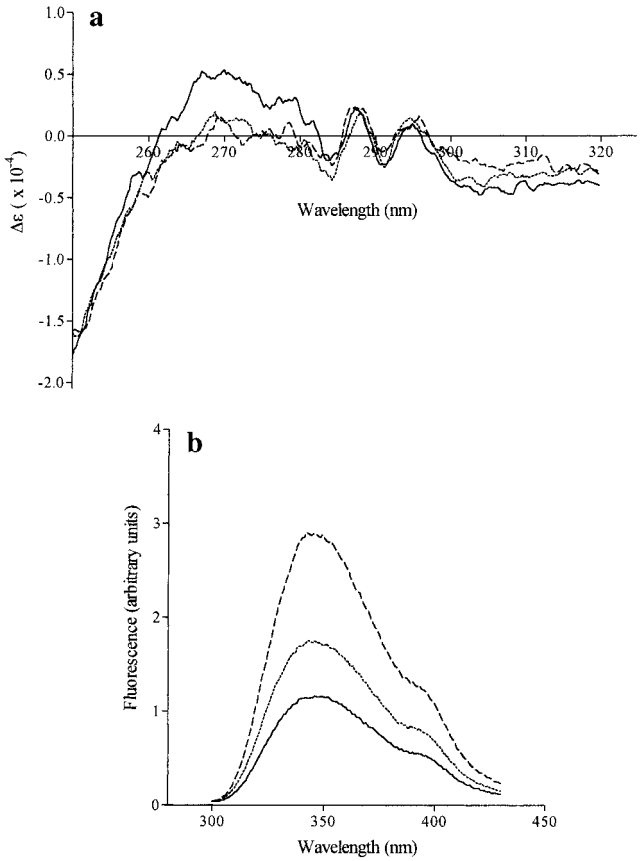

FIGURE 4: Spectra of wild-type and mutant CBMs. (a) Far-UV CD and (b) fluorescence spectra of wild-type CBM2b-1 (- - -) and mutants N264A (− − −) and T319A (—).

for the binding of CBM2b-1 to xylan that is very close to 1 protein per 6 xyloses (i.e., very close to 1 protein per 1 xylohexaose unit). (c) The thermodynamics, affinity, and stoichiometry of the binding of the wild type and single mutants to xylohexaose were very similar to those for binding to xylan, if it is assumed that the binding site comprises nonoverlapping xylohexaose units (data not shown).

The ITC data yield independent values for the free energies and enthalpies of binding, from which entropies can also be calculated. The results are described in detail in Table 1. The most interesting aspect of the data is that while the overall affinities of the mutants for the ligands were similar to that of the wild-type protein, the mutations had a bigger impact on the enthalpy and entropy driving the carbohydrate−protein interaction. This is particularly evident in the triple mutant, in which the free energy of binding was similar to that of native CBM2b-1, but the three mutations caused a large decrease in the magnitude of the enthalpy of binding, which was largely offset by a significantly more favorable change in entropy. The other striking aspect of the data is that the free energies, enthalpies, and entropies of the double and triple mutants are obtained in an approximately additive manner from those of the corresponding single mutants. These results imply some wide-ranging consequences for the nature of the binding, as discussed below.

DISCUSSION

Importance of Aromatic Residues in CBM2 Proteins. Previous studies have shown that aromatic residues play a key role in the binding of CBMs, including CBM2a proteins,

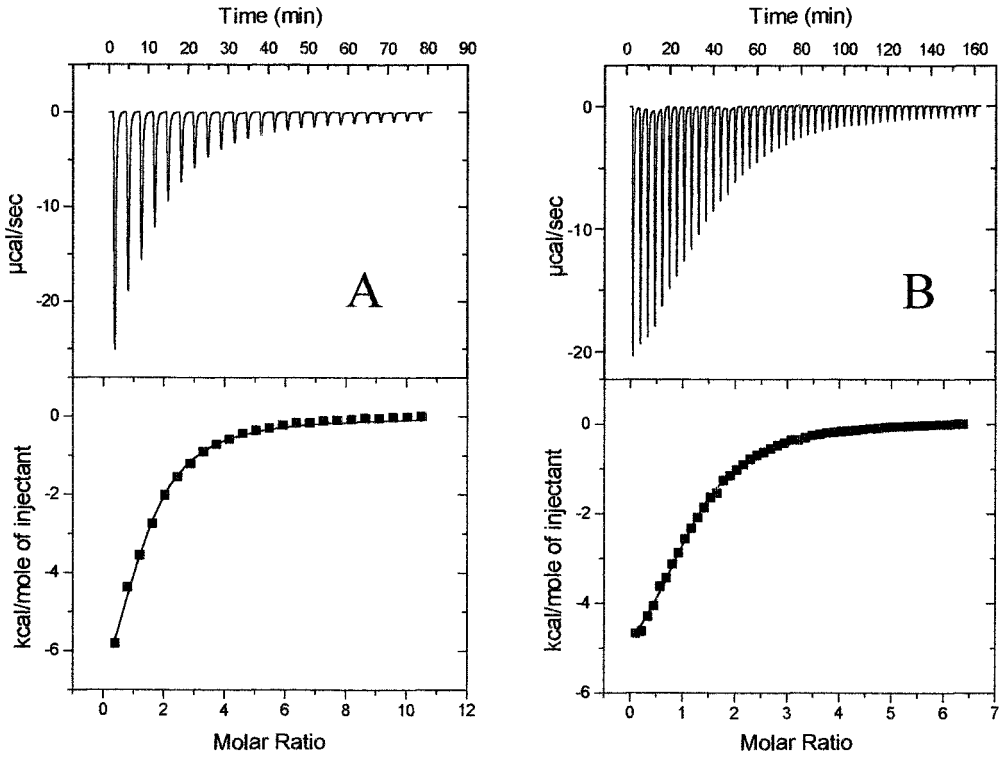

FIGURE 5: ITC results for the binding of wild-type CBM2b-1 (A) and the triple mutant (B) to soluble oat spelt xylan. The top half of each panel shows the calorimetric titration of protein with ligand, and the lower half shows the integrated heats from the upper panel, corrected for control dilution heats. The solid line is the best fit curve that was used to derive parameters K_a and $\Delta H°$.

Table 1: Thermodynamic Parameters for the Binding of Wild-Type CBM2b-1 and Mutants to Xylan[a]

protein	K_a ($\times 10^3$ M^{-1})	K_d (μM)	$\Delta G°$	$\Delta H°$	$T\Delta S°$	$\Delta\Delta G°$ (kcal mol^{-1})	$\Delta\Delta H°$ (kcal mol^{-1})	$T\Delta S°$ (kcal mol^{-1})	n
wild-type	6.42 ± 0.90	156	-5.2 ± 0.06	-9.3 ± 0.48	-4.1 ± 0.51				1.20 ± 0.08
Q288A	4.04 ± 0.55	247	-4.9 ± 0.05	-8.6 ± 0.30	-3.7 ± 0.63	0.3	0.7	0.4	1.18 ± 0.10
E257A	4.72 ± 0.14	212	-5.0 ± 0.05	-9.0 ± 0.21	-4.0 ± 0.48	0.2	0.3	0.1	1.14 ± 0.16
N292A	5.02 ± 0.22	199	-5.0 ± 0.07	-7.9 ± 0.50	-2.9 ± 0.33	0.2	1.4	1.2	1.22 ± 0.15
E257A/Q288A	3.08 ± 0.53	325	-4.8 ± 0.05	-7.8 ± 0.31	-3.0 ± 0.37	0.4	1.5	1.1	1.04 ± 0.11
Q288A/N292A	4.49 ± 0.75	223	-5.0 ± 0.10	-6.9 ± 0.25	-1.9 ± 0.45	0.2	2.4	2.2	1.13 ± 0.17
E257A/Q288A/N292A	2.77 ± 0.40	361	-4.7 ± 0.09	-5.4 ± 0.49	-0.7 ± 0.53	0.5	3.9	3.4	1.17 ± 0.15

[a] Wild-type protein data were from CBM2b fused at the N-terminus to either GST or a His tag sequence. The biochemical properties of the two forms of wild-type CBM2b were indistinguishable. The CBM2b derivatives Q288A, E257A, N292A, E257A/Q288A, and Q288A/N292A were fused, at the N-terminus, to GST, while E257A/Q288A/N292A contained an N-terminal His tag. *n* is the number of binding sites on the protein for the polysaccharide.

to crystalline cellulose by stacking against the pyranose rings of the glucose polymer (*16*−*19*). Two of the aromatic residues that are involved in cellulose binding in family 2a CBMs (Trp 17 and Trp 54 in *C. fimi* Xyn10A CBM2a) are conserved in CBM2b proteins (Trp 259 and Trp 291 in *C. fimi* Xyn11A CBM2b-1), implying that they play a critical role in ligand binding in both CBM2 subfamilies. Results presented in this report show that the removal of either of these residues abolishes ligand binding, and thus, the results strongly support the model proposed by Simpson et al. (*14*), which suggests that the surface aromatic residues in CBM2b-1 play a pivotal role in the interaction of the protein with xylan. These results are also consistent with the data presented in ref *20*, which indicate that the orientation of the surface tryptophan residues in CBM2a and CBM2b proteins plays a key role in ligand recognition.

Functional Importance of Asp 261, Asn 264, and Thr 316. Of the eight sites chosen for mutation, one produced no protein (D261A) and four produced mutants that did not bind to xylan (W259A, N264A, W291A, and T316A), leaving only three that produced mutants capable of binding to xylan (E257A, Q288A, and N292A). The lack of protein production for D261A presumably indicates that the mutant protein was misfolded and digested by proteases. The side chains of Asp 261 and Arg 262 are both somewhat unusual in that they have restricted rotation about the Cα−Cβ bond, as evidenced by differential $^{15}N-^{1}H\beta$ coupling constants and HN−Hβ NOEs in the NMR spectrum. This puts the side chains in the appropriate orientation to hydrogen bond with each other (Figure 1). In view of the central role that Arg 262 plays in determining the orientation of the surface residue Trp 259 (*20*), these results imply that Asp 261 has an important function in maintaining the geometry of key residues on the binding face.

The side chains of Asn 264 and Thr 316 hydrogen bond to each other on the surface of the protein, forming part of the binding surface for the ligand, and sitting between the two exposed tryptophan side chains (Figure 1). Mutation of either of these residues causes intensity changes to the fluorescence spectrum (Figure 4). These two mutants are

therefore presumed to have a perturbed environment around one or both tryptophan side chains, explaining their loss of binding. Thus, we propose that the primary role of Thr 316 and Asn 264 is to maintain the correct structure of the ligand binding region of CBM2b-1.

Free Energy Changes on Binding. Removal of either of the two tryptophan side chains (Trp 259 or Trp 291) leads to loss of any detectable binding, implying a loss of at least 2.5 kcal mol^{-1} in binding affinity. Many cellulose binding modules contain three tryptophan residues, and the loss of any one of these causes a reduction in binding energy of ~2 kcal mol^{-1} (see ref *18*). Merely on the basis of the hydrophobic binding energy, one tryptophan−xylan interaction reduces the exposed hydrophobic surface area by approximately 35 $Å^2$, which corresponds to a free energy of ~2.5 kcal mol^{-1} (using the relationship in which hydrophobic burial contributes ~50−100 cal mol^{-1} $Å^{-2}$ buried; *21*−*23*). It is therefore not surprising that removal of a tryptophan side chain causes such a fundamental change in binding energy.

Removal of the side chains of hydrogen bonding residues Asn 292, Gln 288, or Glu 257 has only a very small effect on the free energy of binding (Table 1). The free energy change for a hydrogen bond in water is generally taken to be somewhere between 0.5 and 4.5 kcal mol^{-1} (*21*, *24*, *25*), with strong hydrogen bonds contributing in excess of 1.5 kcal mol^{-1} (*26*). In sugar−protein complexes, the loss of a single hydrogen bond has been observed to lead to drastic changes in binding energy. Thus, in several complexes of lectins with their oligosaccharide ligands, loss of a single hydrogen bond (using deoxy sugar derivatives) leads to a complete loss of observed binding (*27*). It is therefore somewhat surprising that the loss of a hydrogen bond in CBM2b-1 causes changes in the free energy of binding of ≤0.3 kcal mol^{-1}, and that the loss of three hydrogen bonds in the triple mutant only weakens the binding by 0.5 kcal mol^{-1} (Table 1). In broad terms, the rather small change in binding energy from the loss of a hydrogen bonding interaction can be rationalized in that the unfavorable change in enthalpy is almost exactly compensated by a favorable change in entropy, arising largely from the fact that the ligand has more translational and rotational freedom in the absence of the hydrogen bond. This brings up the thorny question of enthalpy−entropy compensation, which is discussed in more detail below. We should however note in support of this general statement that the loss of the hydrogen bond from Asn 292 to the terminal sugar of the xylose polymer occupying the binding site (Figure 1) causes a smaller free energy change and a more favorable entropy change than loss of either of the other two hydrogen bonds, which are to internal sugars (cf. Table 1, row 4 vs rows 2 and 3). The loss of a hydrogen bond to a terminal sugar might be expected to permit a greater degree of motional freedom to the bound ligand than loss of a hydrogen bond to an internal sugar.

Enthalpy−Entropy Compensation. Although the free energy of binding of CBM2b-1 to xylose polymers is only changed to a small degree by the loss of hydrogen bonding interactions, there are much larger changes to the enthalpy and entropy of binding. An increase in entropy associated with a decrease in enthalpy, when either the ligand or protein has been modified, is generally termed enthalpy−entropy compensation. This phenomenon has been observed in a wide variety of molecular associations (*28*−*30*), including protein−saccharide interactions (*31*−*33*), and the relative contribution of enthalpy and entropy varies widely between different binding interactions. In particular, the loss of a single hydrogen bond between the ligand and sugar-binding proteins that have deep binding pockets and high affinites for their target saccharide (such as periplasmic sugar transport proteins and certain lectins) has a drastic effect on $\Delta G°$. This implies that the free energy of binding is more sensitive to changes in enthalpy than entropy, although some compensation does occur. This view is supported by the results of Swaminathan et al. (*33*), who showed that the linear relationship between $\Delta H°$ and $T\Delta S°$ is greater than unity for concanavalin A−mannooligosaccharide binding.

The data presented in this report show that CBM−polysaccharide interactions also involve enthalpy−entropy compensation, but in this instance, the decrease in enthalpy through the loss of hydrogen bonds can be almost completely compensated by an increase in entropy. One reason for the different behavior of CBMs on one hand, and periplasmic transport proteins and lectins on the other hand, lies in the different structure of their ligand binding sites. CBM binding sites comprise shallow clefts or planar surfaces, while the sugar binding sites in many lectins (and much more so in transport proteins) are deeper clefts or pockets (*34*, *35*). This means that in lectins and sugar transport proteins the loss of a saccharide−protein hydrogen bond cannot be compensated by increased motional freedom of the ligand, due to steric hindrance imposed by the binding site. In contrast, by extrapolation of the data presented in this report, we propose that when polysaccharides interact with CBMs on exposed binding sites, the loss of a hydrogen bond can be compensated by a more positive entropy value through increased motional freedom of the ligand, because it is not confined by a sterically restricted binding site. Thus, the importance of hydrogen bonds in the binding of ligands to CBMs with surface binding sites might be less important than in other protein−saccharide interactions, and instead it is the complementarity of the surface of the ligand and binding site that determines the specificity and affinity of these macromolecular associations.

A key feature of enthalpy−entropy compensation is that enthalpy and entropy changes are of almost exactly the same magnitude, leading to a slope of $\Delta H°/\Delta S°$ of close to 1.0, and consequently only a small change in the overall free energy (Figure 6). The phenomenon has generally been attributed to "solvent reorganization", which is normally taken to mean that liberation of water molecules from a hydrophobic surface produces a less favorable enthalpy of binding (because water molecules surrounding a hydrophobic surface are forced to form more and/or stronger hydrogen bonds among themselves than they do in bulk solvent) but a more favorable entropy of binding (because the water molecules are now free; but see ref *36*, which presents experimental evidence that bound water has an entropy similar to that of bulk water). Although this explanation is attractive, it hinders understanding at least as much as it helps: first, because further understanding of the thermodynamics is hampered by our lack of understanding at a molecular level of the nature of the solvent reorganization (*32*); and second, because enthalpy−entropy compensation

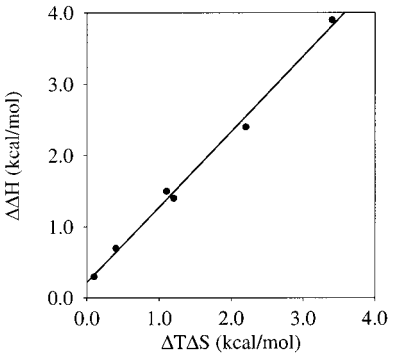

FIGURE 6: Enthalpy−entropy compensation plot for the mutants described in this work, plotted as differences from the wild-type interaction. The line is the least-squares fit, with a gradient of 1.06.

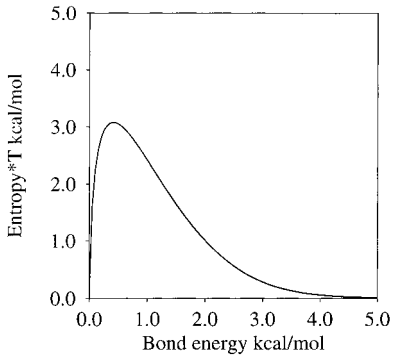

FIGURE 7: Dependence of the entropy of bond distributions on the enthalpy of the bond. The entropy is calculated from eq 10 of ref *40*: $S = R \ln[\sum_{j=0}^{M} \binom{M}{j}^2 (1-p)^j p^{M-J}]$ where $p = \exp(-E/RT)$, using $M = 10$.

happens to a similar degree in solvents other than water (*21*, *37*). In particular, it is observed in carbon tetrachloride, a solvent not capable of forming strong intermolecular interactions. The explanation for the large compensatory changes in enthalpy and entropy therefore cannot be attributed to any special properties of water (as proposed for example by refs *38* and *39*), but must be more general.

An alternative explanation of enthalpy−entropy compensation has been proposed (*40*), which is more helpful for understanding the processes occurring here, and in many other binding interactions. This explanation suggests that a weakening in bonding energy (for example, a loss in enthalpy caused by the loss of a hydrogen bond) is necessarily followed by a redistribution of the energetics of the remaining weak intermolecular bonds. A key element of this explanation is that while *strong* bonds, for example, covalent bonds, tend to have favorable enthalpy but unfavorable entropy, *weak* bonds achieve cooperatively a minimum free energy when their strengths are reorganized such that they have relatively unfavorable enthalpy but high entropy, essentially by allowing the ligand a greater motional flexibility. This is to some extent a rewording of the second law of thermodynamics, and also of the law of equipartition of energy, and its strength is that it stresses that redistribution of entropy and enthalpy within a system is an inherent feature of the system, and does not require any theorizing about the nature of solvent reorganization. Indeed, to quote Weber (*40*), *the energetic participation of the solvent [is limited] to that of providing a suitably matching enthalpy contribution*. This treatment focuses on weak bonds, because the energetic gain from redistributing binding energy into entropy is maximized when the individual bonds have an energy of only $RT \ln 2$ (0.4 kcal mol^{-1}) (Figure 7). Similar conclusions have also been reached (but based on experimental rather than theoretical results) by the Williams group (*21*, *37*, *41*−*43*), where the emphasis has been that residual motions in bound complexes contribute strongly to the overall stability of the complexes, and can therefore form the basis of an explanation of enthalpy−entropy compensation.

On the basis of the foregoing discussion, a valid explanation for the enthalpy and entropy changes observed here is as follows. Removal of a hydrogen bond (which has strong directional character) allows the complex greater motional freedom, which provides an increase in entropy that almost offsets the loss of enthalpy caused by the loss of the hydrogen bond. Remaining hydrogen bonds may well also weaken at the same time, thus allowing the complex further motional freedom, and further shifting the balance between enthalpy and entropy.

Several conclusions arise from this analysis.

(1) Redistribution of energies is only possible where the bound ligand has the possibility of increased motion. Therefore, this analysis only applies to cases, exemplified by family 2 CBMs, where the binding site is exposed on the surface. Relatively buried binding sites, such as those for saccharides in lectins, or completely buried, such as those in bacterial periplasmic sugar binding proteins (*35*, *44*), have no possibility of internal motion. In such cases, the loss of a hydrogen bond can have a very strong effect on overall binding energy, as noted above, as the loss in enthalpy cannot be compensated by an increase in entropy. CBM4s occupy an intermediate position, in that the polysaccharide binds in a restricted cleft. In agreement with this intermediate position, hydrogen bonds have been shown to contribute an intermediate value of ~1 kcal mol^{-1} (*12*).

(2) The corollary to this is that when ligands bind at an exposed surface, hydrogen bonds are relatively unimportant, and the complementarity of binding surfaces is much more important. This conclusion is clearly relevant to protein−protein association as well as protein−polysaccharide association.

(3) Hydrogen bonds are normally described as being highly cooperative, essentially because the presence of a strongly directional hydrogen bond markedly increases the effective concentration for the interaction between a neighboring hydrogen-bonded pair (*45*). However, the cooperativity between the hydrogen bonds mutated here is weak. Thus, although the triple mutant shows the greatest loss in free energy on binding compared to the wild type, and also the greatest changes in enthalpy and entropy (Table 1), the changes are close to additive. This lack of cooperativity is seen as another manifestation of the fact that in this complex, between a protein surface and a polymeric sugar, the hydrogen bonds have little effect on the strength of binding, and a rather small effect on the specificity of binding.

CONCLUSIONS

In summary, we have shown that hydrogen bonding between the protein and the surface-bound polysaccharide makes a modest contribution to the overall binding, because although hydrogen bonds have favorable enthalpy, they restrict the residual motion of the ligand and thus reduce

the entropy of the system. Most of the affinity and specificity of the binding derives from interactions between the faces of sugar rings and aromatic residues on the protein (*20*). This result can be contrasted to the situation for binding of buried saccharides, exemplified by bacterial periplasmic sugar binding proteins and many lectins, where hydrogen bonding has a much stronger effect on both specificity and affinity (*27*, *31*, *46*, *47*).

REFERENCES

1. Brett, C., and Walden, K. (1990) *Physiology and Biochemistry of Plant Cell Walls*, Unwin and Hyman, London.
2. Tomme, P., Warren, R. A. J., and Gilkes, N. R. (1995) *Adv. Microb. Physiol. 37*, 1−81.
3. Atkins, E. D. T. (1992) in *Xylan and xylanases: Progress in biotechnology* (Visser, J., Beldman, G., van Kusters, S., and Voragen, A. G. L., Eds.) Vol. 7, pp 39−50, Elsevier, Amsterdam.
4. Bolam, D. N., Ciruela, A., McQueen-Mason, S., Simpson, P. J., Williamson, M. P., Rixon, J. E., Boraston, A., Hazlewood, G. P., and Gilbert, H. J. (1998) *Biochem. J. 331*, 775−781.
5. Gill, J., Rixon, J. R., Bolam, D. N., McQueen-Mason, S., Simpson, P. J., Williamson, M. P., Hazlewood, G. P., and Gilbert, H. J. (1999) *Biochem. J. 342*, 473−480.
6. Black, G. W., Rixon, J. R., Hazlewood, G. P., Theodorou, M. K., Morris, P., and Gilbert, H. J. (1996) *Biochem. J. 319*, 515−520.
7. Sakon, J., Irwin, D., Wilson, D. B., and Karplus, P. A. (1997) *Nat. Struct. Biol. 4*, 810−817.
8. Irwin, D., Shin, D.-H., Zhang, S., Barr, B. K., Sakon, J., Karplus, P. A., and Wilson, D. B. (1998) *J. Bacteriol. 180*, 1709−1714.
9. Boraston, A. B., McLean, B. W., Kormos, J. M., Alam, M., Gilkes, N. R., Haynes, C. A., Tomme, P., Kilburn, D. G., and Warren, R. A. J. (1999) in *Recent Advances in Carbohydrate Engineering* (Gilbert, H. J., Davies, G. J., Svensson, B., and Henrissat, B., Eds.) pp 202−211, Royal Society of Chemistry, Cambridge, U.K.
10. Tomme, P., Creagh, A. L., Kilburn, D. G., and Haynes, C. A. (1996) *Biochemistry 35*, 13885−13894.
11. Creagh, A. L., Ong, E., Jervis, E., Kilburn, D. G., and Haynes, C. A. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 12229−12234.
12. Brun, E., Johnson, P. E., Creagh, L. A., Tomme, P., Webster, P., Haynes, C. A., and McIntosh, L. P. (2000) *Biochemistry 39*, 2445−2458.
13. Millward-Sadler, S. J., Poole, D. M., Henrissat, B., Hazlewood, G. P., Clarke, J. H., and Gilbert, H. J. (1994) *Mol. Microbiol. 11*, 375−382.
14. Simpson, P. J., Bolam, D. N., Cooper, A., Ciruela, A., Hazlewood, G. P., Gilbert, H. J., and Williamson, M. P. (1999) *Structure 7*, 853−864.
15. Wiseman, T., Williston, S., Brands, J. F., and Lin, L.-N. (1989) *Anal. Biochem. 179*, 131−137.
16. Din, N., Forsythe, I. J., Burntnick, L. D., Gilkes, N. R., Miller, R. C., Jr., Warren, R. A. J., and Kilburn, D. G. (1994) *Mol. Microbiol. 11*, 747−755.
17. Bray, M. R., Johnson, P. E., Gilkes, N. R., McIntosh, L. P., Kilburn, D. G., and Warren, R. A. J. (1996) *Protein Sci. 5*, 2311−2318.
18. Nagy, T., Simpson, P. J., Williamson, M. P., Hazlewood, G. P., Gilbert, H. J., and Orosz, L. (1998) *FEBS Lett. 429*, 312−316.
19. Ponyi, T., Szabó, L., Nagy, T., Orosz, L., Simpson, P. J., Williamson, M. P., and Gilbert, H. J. (2000) *Biochemistry 39*, 985−991.
20. Simpson, P. J., Xie, H., Bolam, D. N., Gilbert, H. J., and Williamson, M. P. (2000), *J. Biol. Chem. 275*, 41137−41142.
21. Searle, M. S., Williams, D. H., and Gerhard, U. (1992) *J. Am. Chem. Soc. 114*, 10697−10704.
22. Kellis, J. T., Nyberg, K., and Fersht, A. R. (1989) *Biochemistry 28*, 4914−4922.
23. Fersht, A. R. (1985) *Enzyme structure and mechanism*, 2nd ed., p 301, W. H. Freeman, New York.
24. Brummell, D. A., Sharma, V. P., Anand, N. N., Bilous, D., Dubuc, G., Michniewicz, J., MacKenzie, C. R., Sadowska, J., Sigurskjold, B. W., Sinnott, B., Young, N. M., Bundle, D. R., and Narang, S. A. (1993) *Biochemistry 32*, 1180−1187.
25. Fersht, A. R., Shi, J. P., Knill-Jones, J., Lowe, D. M., Wilkinson, A. J., Blow, D. M., Brick, P., Carter, P., Waye, M. M. Y., and Winter, G. (1985) *Nature 314*, 235−238.
26. Street, I. P., Armstrong, C. R., and Withers, S. G. (1986) *Biochemistry 25*, 6021−6027.
27. Lemieux, R. U. (1989) *Chem. Soc. Rev. 18*, 347−374.
28. Lumry, R., and Rajender, S. (1970) *Biopolymers 9*, 1125−1227.
29. Jencks, W. P. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 4046−4050.
30. Lemieux, R. U., Delbaere, L. T. J., Beierbeck, H., and Spohr, U. (1991) *Host−guest molecular interactions from chemistry to biology*, pp 231−248, Ciba Foundation Symposium, Wiley, Chichester, U.K.
31. Toone, E. J. (1994) *Curr. Opin. Struct. Biol. 4*, 719−728.
32. Lemieux, R. U. (1996) *Acc. Chem. Res. 29*, 373−380.
33. Swaminathan, C. P., Surolia, N., and Surolia, A. (1998) *J. Am. Chem. Soc. 120*, 5153−5159.
34. Weis, W. I., and Drickamer, K. (1996) *Annu. Rev. Biochem. 65*, 441−473.
35. Vyas, N. K. (1991) *Curr. Opin. Struct. Biol. 1*, 732−740.
36. Denisov, V. P., Venu, K., Peters, J., Horlein, H. D., and Halle, B. (1997) *J. Phys. Chem. B 101*, 9380−9389.
37. Searle, M. S., and Williams, D. H. (1992) *J. Am. Chem. Soc. 114*, 10690−10697.
38. Williams, B. A., Chervenak, M. C., and Toone, E. J. (1992) *J. Biol. Chem. 267*, 22907−22911.
39. Ramkumar, R., Surolia, A., and Podder, S. K. (1995) *Biochem. J. 308*, 237−241.
40. Weber, G. (1993) *J. Phys. Chem. 97*, 7108−7115.
41. Williamson, M. P., and Williams, D. H. (1984) *Eur. J. Biochem. 138*, 345−348.
42. Searle, M. S., Sharman, G. J., Groves, P., Benhamu, B., Beauregard, D. A., Westwell, M. S., Dancer, R. J., Maguire, A. J., Try, A. C., and Williams, D. H. (1996) *J. Chem. Soc., Perkin Trans. 1*, 2781−2786.
43. Williams, D. H., Bardsley, B., Tsuzuki, W., and Maguire, A. J. (1997) *Chem. Biol. 4*, 507−512.
44. Quiocho, F. A. (1991) *Curr. Opin. Struct. Biol. 1*, 922−933.
45. Creighton, T. E. (1993) *Proteins*, 2nd ed., pp 165−167, W. H. Freeman, New York.
46. Quiocho, F. A. (1989) *Pure Appl. Chem. 61*, 1293−1306.
47. Sharon, N. (1993) *Trends Biochem. Sci. 18*, 221−226.

BI010034Z