This article was downloaded by:
On: *29 January 2011*
Access details: *Access Details: Free Access*
Publisher *Taylor & Francis*
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK

Supramolecular Chemistry
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713649759

Thermodynamic Parameters of Specific and Nonspecific Protein-DNA Binding

Linda Jen-jacobson[a]; Lisa E. Engler[a]; Jennifer T. Ames[a]; Michael R. Kurpiewski[a]; Arabela Grigorescu[a]
[a] Department of Biological Sciences, University of Pittsburgh, Pittsburgh, PA, USA

To cite this Article Jen-jacobson, Linda , Engler, Lisa E. , Ames, Jennifer T. , Kurpiewski, Michael R. and Grigorescu, Arabela(2000) 'Thermodynamic Parameters of Specific and Nonspecific Protein-DNA Binding', Supramolecular Chemistry, 12: 2, 143 — 160
To link to this Article: DOI: 10.1080/10610270008027446
URL: http://dx.doi.org/10.1080/10610270008027446

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.informaworld.com/terms-and-conditions-of-access.pdf

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

SUPRAMOLECULAR CHEMISTRY, Vol. 12, pp. 143-160
Reprints available directly from the publisher
Photocopying permitted by license only

© 2000 OPA (Overseas Publishers Association) N.V.
Published by license under
the Harwood Academic Publishers imprint,
part of the Gordon and Breach Publishing Group.
Printed in Malaysia

Thermodynamic Parameters of Specific and Nonspecific Protein-DNA Binding

LINDA JEN-JACOBSON*, LISA E. ENGLER, JENNIFER T. AMES, MICHAEL R. KURPIEWSKI and ARABELA GRIGORESCU

Department of Biological Sciences, University of Pittsburgh, Pittsburgh PA 15260, USA

Downloaded At: 15:35 29 January 2011

Proteins that bind preferentially to specific recognition sites on DNA also bind more weakly to nonspecific DNA. We have studied both specific and non-specific binding of the *Eco*RI and *Bam*HI restriction endonucleases, and determined enthalpic and entropic contributions to binding free energy (ΔG°_{bind}) using both the van't Hoff method and isothermal titration calorimetry. Specific binding is characterized by a strongly negative ΔC°_p and can be either enthalpy-driven or entropy-driven, depending on temperature. Nonspecific binding has $\Delta C^\circ_p \approx 0$ and is enthalpy-driven. A strongly negative ΔC°_p is the "thermodynamic signature" of site-specific binding, because it reflects the characteristics of a tight complementary recognition interface: the burial of previously hydrated nonpolar surface and restriction of configurational-vibrational freedoms of protein, DNA, and water molecules trapped at the protein-DNA interface. These factors are absent in nonspecific complexes. We probed the contributions to ΔC°_p by varying the sequence context surrounding the recognition site. As ΔG°_{bind} improves, ΔC°_p, ΔH° and ΔS° all become more negative, and there is a linear correlation between ΔH° and ΔS° (enthalpy-entropy compensation). Because these context variations do not change the protein-base or protein-phosphate contacts, the hydrophobic contribution or the number of trapped water molecules at the interface, we conclude that a better sequence context improves the "goodness of fit" in the interface and and thus increases the magnitude of the negative configurational-vibrational contribution to ΔC°_p.

Keywords: specific protein-DNA complexes, nonspecific protein-DNA complexes, heat capacity change, thermodynamics, enthalpy-entropy compensation

INTRODUCTION

Sequence-specific DNA-binding proteins, including restriction endonucleases and methylases, transcriptional repressors and activators, bind to their recognition sites on DNA with extremely high specificity [1]. High specificity is achieved using both *direct* readout of sequence information (direct point-to-point interactions, e.g., hydrogen bonds with individual functional groups on the bases) and *indirect* readout of sequence information [2–4]. The term "indirect readout" has been used in different ways [2–4] but we use it here to denote the use of sequence-dependent conformational information, including the precise orientations of base-pairs, sugar and phosphate geometries, DNA groove widths [5] and DNA distortion [6].

Our previous work has shown that the overall discrimination by the *Eco*RI [4] and *Bam*HI [7] restriction endonucleases against any site with even one incorrect base-pair is extremely high because it includes both the direct and indirect readouts. There is discrimination in DNA binding, and also in the first-order rate constants for DNA cleavage. Taking both into account, we saw orderly hierarchies of discrimination energies of about 6.6 to 13 kcal/mol or 70,000 to

* Correspondence to: Linda Jen-Jacobson, Tel: 412–624–4969, Fax: 412–624–4759, e-mail: LJEN@pitt.edu

10^9-fold [4]. We have pointed out elsewhere [1] that discrimination in the cleavage step is a consequence of the fact that the specific recognition (binding) complex bears a close resemblance to the transition-state complex, such that very tight binding to the recognition site on the DNA substrate does not inhibit catalysis, but instead provides energy that is efficiently utilized along the path to the transition state. Structural changes in the recognition complex, such as those that inevitably occur in binding to an incorrect recognition site, therefore adversely impact catalytic rates as well as binding.

High specificity requires a structurally complementary protein-DNA interface. Crystal structures of protein-DNA complexes (see [8] for a comprehensive list of references for 63 protein-DNA complexes) have shown the formation of tightly apposed surfaces that form both polar and nonpolar contacts between protein and the DNA bases and phosphates in the recognition sites.

It is often true that the free protein and DNA cannot form such extensive complementary interfaces without changes in conformation. To form their recognition interfaces, both *Eco*RI [9] and *Bam*HI [10] endonucleases must undergo large conformational changes and the DNA in each complex must suffer distortion, which is mild for *Bam*HI [10] and dramatic for *Eco*RI [5]. The deformed DNA in the specific EcoRI complex has a distinct torsional kink at the central base-pairs, which unwinds the DNA and widens the major groove so that the protein can insert its recognition elements to contact all available functional groups on the six base-pairs (GAATTC) in the recognition site.

The formation of complementary protein-DNA interfaces leads to strong binding, but the observed negative binding free energy (ΔG°_{bind}) is the result of large opposing contributions. There are favorable contributions from protein-base and protein-phosphate interactions (the latter subsuming cation release [11]) and from the hydrophobic effect, that is, the release of water from nonpolar surfaces [12]. Set against these are a number of unfavorable contributions, arising from the losses of translational, rotational, conformational and vibrational freedom, from the desolvation of polar surfaces, and from molecular strain (e.g., bond distortion and repulsive interactions) of both protein and DNA in the complex. It should always be borne in mind that this division into components, while conceptually useful, may have little relation to the physical reality; in fact, these components are completely interdependent during the highly cooperative formation of a protein-DNA interface.

To account for the experimentally observed binding free energy (ΔG°) of –15 kcal/mol, we have made estimates [1] based on experimental data that the aggregate of the favorable free energy contributions can be as large as –180 kcal/mol, paying for the cost of unfavorable contributions as large as +165 kcal/mol. Beveridge and coworkers [13] have used molecular mechanics combined with molecular dynamic analyses to compute theoretically that the aggregate of favorable free energy contributions is –302 kcal/mol, counterbalanced by unfavorable factors adding up to ~+290 kcal/mol.

Whether net favorable or unfavorable, each contribution to ΔG°_{bind} is itself the resultant of enthalpic and entropic factors, which generally oppose each other. In this paper, we consider how the signs and magnitudes of some of these enthalpic and entropic factors can be assessed, and what additional insights this provides beyond those available from considering free-energy changes. In particular, we show that a strongly negative heat capacity change (ΔC°_p) is the "thermodynamic signature" that distinguishes site-specific from nonspecific DNA binding. Finally, we show that the negative ΔC°_p associated with specific binding is the result not only of burial of nonpolar surface, but also of the loss of configurational-vibrational freedom in the tight complementary interface of the specific complex.

Downloaded At: 15:35 29 January 2011

Downloaded At: 15:35 29 January 2011

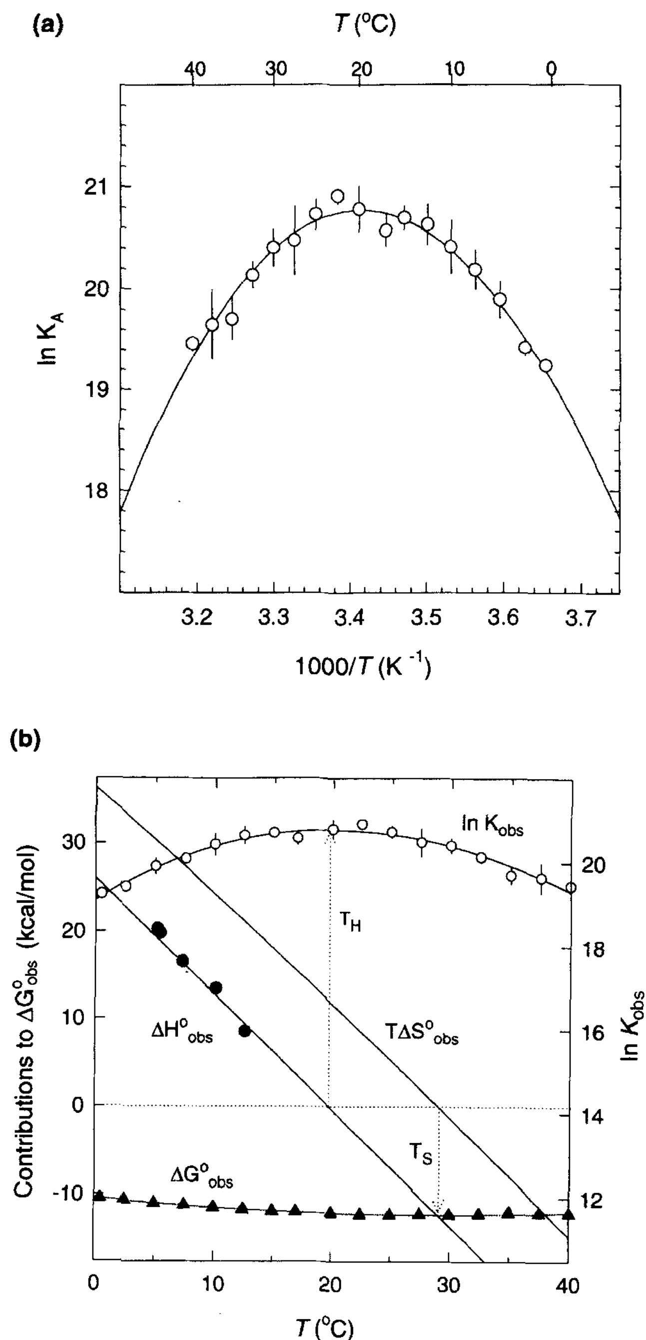

FIGURE 1 (a) Representative van't Hoff plot for the site-specific binding of *Bam*HI endonuclease. Values of ln K_A are the means ± S.D. of 4 independent determinations. The curve is the non-linear least squares best-fit of *all* data points (not means ± SD) to ln $K_{obs} = (\Delta C°_p/R)[(T_H/T)-\ln(T_s/T)-1]$ assuming a temperature independent $\Delta C°_p$. For this sequence (CGCGGGCGGCggatccG-GGCGGGC), T_H (where $\Delta H°_{obs} = 0$) is 292.8 ± 0.3 K and T_s (where $\Delta S° = 0$) is 302.1 ± 0.6 K. (b) Thermodynamic profiles derived from data presented in Fig. 1 (a). K_A and $\Delta G°$ values are experimentally determined. The predicted $\Delta H°$ and $\Delta S°$ contributions to $\Delta G°$ as a function of temperature (lines) were calculated from $\Delta H° = \Delta C°_p(T-T_H)$ and $\Delta S° = \Delta C°_p \ln(T/T_s)$; $\Delta C°_p = -1.3 \pm 0.1$ kcal/mol·K. Experimental values (•) of $\Delta H°_{obs}$ are from ITC measurements

TABLE I Thermodynamic Signatures of Specific and Nonspecific *Eco*RI and *Bam*HI Complexes

	EcoRI Endonuclease[a]				*BamHI Endonuclease*[b]			
	"Best"	*"Medium"*	*"Worst"*	*Nonspecific*[c]	*"Very Good"*	*"Medium"*	*"Worst"*	*Nonspecific*[c]
ΔG° bind[d] kcal/mol	−14.6 ± 0.1	−12.3 ± 0.3	−11.1 ± 0.4	− 6.8 ± 0.3	−12.7 ± 0.2	−12.4 ± 0.1	−10.9 ± 0.1	−7.5 ± 0.1
ΔC°_p [e] kcal/mol·K	− 2.5 ± 0.3	−1.5 ± 0.2	− 1.2 ± 0.3	0	−1.5 ± 0.1	− 1.3 ± 0.1	−1.2 ± 0.1	0
ΔH° [e] kcal/mol	− 9.7 ± 0.4	−0.3 ± 2.5	+ 9.3 ± 3.5	−14.2 ± 1.5	−16.6 ± 1.7	− 6.5 ± 1.0	−1.9 ± 0.9	− 7.4 ± 0.3
$T\Delta S^\circ$ [e] kcal/mol	+ 4.9 ± 1.0	+12.0 ± 4.1	+20.4 ± 5.8	− 7.4 ± 1.5	− 4.0 ± 0.7	+ 5.2 ± 0.7	+9.0 ± 3.0	+0.1 ± 0.2

a. Equilibrium association constant values (K_A) determined in binding buffer A. Sequences for specific sites embedded in different contexts and K_A values are shown in Table II. "Best"- seq. no. 1, "medium" – sequence no. 4, "worst" – seq. no. 6.
b. K_A values (determined in binding buffer B) for specific *Bam*HI sequences vary 21-fold; "very good" – GGGATGGGTGggatccCACCCAC, "medium" – CGCGGGCGGCggatccGGGCGGGC, "worst" – GGGATGGTGGggatccCCACCAC.
c. Nonspecific site (CTTAAG) for *Eco*RI embedded in "best" context; nonspecific site (CCTAGG) for *Bam*HI embedded in "medium" context.
d. Calculated from $\Delta G^\circ = -RT \ln KA$ for T = 298 K.
e. Obtained from fits to equations given in legend, Fig. 1.

Downloaded At: 15:35 29 January 2011

RESULTS AND DISCUSSION

The thermodynamic parameters of specific binding

To extract the enthalpic (ΔH°) and entropic (ΔS°) contributions to the free energy of binding for both the *Eco*RI and *Bam*HI-DNA interactions, we have taken two experimental approaches. First, we measured the equilibrium association constant K_A as a function of temperature by nitrocellulose filter-binding methods [14, 15]. Fig. 1(a) shows a representative van't Hoff plot ($\ln K_A$ vs $1/T$) for *Bam*HI binding to a 24 base-pair oligodeoxynucleotide containing its specific recognition site GGATCC. The highly nonlinear plot (exhibiting a maximum for K_A) indicates that $\Delta H^\circ_{obs} (= -R(\partial \ln K_{obs}/\partial 1/T)_P$ depends strongly on temperature.

There is a potential problem of extracting numerical values of ΔH° and ΔS° from a three-parameter fit to the van't Hoff equation (see legend, Fig. 1), namely that the computed values might be influenced by strong covariance. We have therefore also used isothermal titration calorimetry (ITC) to obtain independent verification of ΔH° values. The lines in Fig 1(b) show the computations from van't Hoff data of the enthalpic (ΔH°) and entropic ($T\Delta S^\circ$) contributions to the standard free energy change as a function of temperature and the plotted points are the calorimetrically determined ΔH° values. There is close agreement of the values of ΔH° and its temperature dependence between the indirect van't Hoff and the direct calorimetric methods. The key features of this plot are:

1) Both ΔH° and $T\Delta S^\circ$ decrease greatly with increasing temperature, deriving from a large negative value of ΔC°_p for the association process.
2) Binding is entropy-driven below a temperature T_H (where $\Delta H^\circ = 0$), enthalpy-driven above a temperature T_S (where $\Delta S^\circ = 0$) and has both ΔH° and ΔS° favorable between T_H and T_S. Thus, it is not meaningful to speak of "entropy-driven" or "enthalpy-driven" binding unless the temperature is specified.
3) The strongly compensating temperature dependences of ΔH° and $T\Delta S^\circ$ result in a ΔG°_{bind} that is relatively temperature-invariant.

We also observe a large negative ΔC°_p for the interaction of *Eco*RI endonuclease with its specific site GAATTC embedded in a 23 bp sub-

strate. Table I presents values of the thermodynamic parameters ($\Delta G°$, $\Delta C°_p$, $\Delta H°$ and $T\Delta S°$) for *Bam*HI and *Eco*RI endonucleases. Although the $\Delta C°_p$ values for the site-specific association of *Bam*HI and *Eco*RI endonucleases are always large and negative (-1.2 to –2.5 kcal/mol·K), the magnitude of the values are dependent on the context in which the specific sites are embedded (see below). Large negative values of $\Delta C°_p$ resulting in strong but compensating temperature dependences of $\Delta H°_{obs}$ and $\Delta S°$ have also been reported for the formation of a number of other site-specific protein-nucleic acid complexes [16–22].

The thermodynamic parameters of nonspecific binding

All known site-specific DNA-binding proteins also bind to nonspecific DNA, albeit 10^3- to 10^7-fold more weakly [1]. There are only 2 available crystal structures of *nonspecific* complexes: *Eco*RV endonuclease [23] and the glucocorticoid receptor DNA-binding domain [24]. In both, the association between protein and DNA appears to be loose; the total surface area buried in the complex is much smaller than in a specific complex. In contrast to the multiple direct protein-base hydrogen bond and van der Waals contacts in the specific complexes, there are only one or two protein-base contacts in the nonspecific complexes. However, the number of protein-phosphate contacts is similar. In striking contrast to the dramatic distortions seen in the specific *Eco*RV complex (and many other specific protein-DNA complexes), there appears to be little or no DNA distortion in the nonspecific *Eco*RV complex; the structural parameters for the DNA in the nonspecific complex appear to be in the range typically observed for B-form DNA fragments [23].

In light of this structural view, we asked whether the nonspecific complex was also "non-intimate" from a thermodynamic point of view, by determining $\Delta G°_{bind}$, $\Delta H°$, $\Delta S°$ and $\Delta C°_p$ for nonspecific binding of *Eco*RI and *Bam*HI endonucleases (Table I). As "nonspecific" DNA for each enzyme we used a 23-bp oligonucleotide containing an inverted six base-pair recognition site, such that the protein could form none of the normal hydrogen-bonds or nonpolar interactions with the bases [25].

A key result (Table I) is that specific and nonspecific complexes differ markedly in $\Delta C°_p$. In contrast to the strongly negative $\Delta C°_p$ for site-specific binding, $\Delta C°_p \approx 0$ for *Eco*RI and *Bam*HI binding to nonspecific DNA, as has also been reported for the nonspecific binding of cro [26], trp [17], lac [22] and λ_{CI} [19] repressors, and of GCN4 transcriptional activator [21]. We have also analyzed the dependence of equilibrium affinity on the concentration of cosolvents such as triethylene glycol (TEG), betaine or trimethylamine-N-oxide (TMAO) and found that release of water upon specific binding by both *Bam*HI and *Eco*RI is dramatically larger (~14-fold) than that upon nonspecific binding [27]. Importantly, results from molecular dynamic studies [28] indicate that the magnitudes of the motions for protein and DNA in a nonspecific *Eco*RV complex are greatly increased relative to those in the specific complex; furthermore, mobile solvent channels are present in the less compact nonspecific complex, but are not in the tight specific complex. Thus we infer that in the nonspecific complexes, there is little hydrophobic effect, and no significant restriction of the dynamic fluctuations of protein or DNA or water.

For both *Eco*RI and *Bam*HI, nonspecific binding appears to be enthalpy-driven, like the binding of lac repressor to nonspecific DNA [22] and binding of the nonspecific SSB protein [29], in contrast to previous proposals [17, 19, 26] that the entropically favorable release of counterions from DNA would be the driving force for nonspecific interaction. We find for nonspecific binding that the entropic contribution ($-T\Delta S°$) to free energy is either negligible or unfavorable, implying that loss of the rotational-translational degrees of freedom are not completely paid for by the entropically favorable release of counterions.

Downloaded At: 15:35 29 January 2011

Downloaded At: 15:35 29 January 2011

Comparison of specific and nonspecific binding

For specific binding, the absolute values of $\Delta H°$ and $\Delta S°$ are of course temperature-dependent (because $\Delta C°_p \neq 0$), but moreover they are dependent on the nucleotide sequence surrounding the recognition site (Table I). Crystal structures show that these surrounding base-pairs make no direct contacts with the proteins, yet they have profound effects on binding, as will be detailed below. This complicates a quantitative comparison of the $\Delta H°$ and $\Delta S°$ for specific and nonspecific binding, but several useful generalizations can be extracted from the data.

(1) For the specific *Eco*RI complexes, the entropic contribution ($-T\Delta S°$) to standard free energy is net favorable for all sequence contexts, whereas for the specific *Bam*HI complex it is less favorable or even unfavorable (in the context where $\Delta G°_{bind}$ is best). The amount of nonpolar surface buried in the *Eco*RI (~2100 Å^2) complex is much larger than that for *Bam*HI (1620Å^2) [7] so that the more favorable $-T\Delta S°$ for EcoRI derives primarily from a larger contribution attributable to the hydrophobic effect. (These calculations include induced folding of the protein upon binding.) Desolvation of polar surfaces cannot be the major source of favorable $-T\Delta S°$ for *Eco*RI, since 25% more polar surface is buried in the specific *Bam*HI complex.

(2) For both proteins, $\Delta H°$ is less favorable for specific (most sequence contexts) than for nonspecific binding. This contradicts the simplistic view that $\Delta H°$ should be more favorable for the specific complexes because of enthalpically favorable protein-base recognition contacts, which are absent in the nonspecific complex. It follows that specific binding must include unfavorable contributions to $\Delta H°$ which are absent in the nonspecific complex. These unfavorable components are desolvation (polar groups contribute most of the unfavorable $\Delta H°$; nonpolar groups do so to a lesser extent) and molecular strain. Thermodynamic data do not permit deconvolution of these components, but some insight may be gained by comparing different proteins. The *Bam*HI and *Eco*RI interfaces have roughly similar numbers and kinds of protein-base and protein-phosphate contacts, but the specific *Eco*RI complex has dramatic DNA distortion, whereas in the *Bam*HI complex there is little. Thus we expect strain energy to contribute more unfavorable $\Delta H°$ for *Eco*RI, and indeed $\Delta H°_{bind}$ is generally less favorable than for *Bam*HI, although the ranges overlap. This comparison may actually underestimate the differential due to strain energy, because the amount of polar surface buried in the *Bam*HI complex is about 25% greater than for EcoRI. In other words, *Bam*HI would have been expected to have *less* favorable $\Delta H°$ on the basis of desolvation alone. Instead, it has *more* favorable $\Delta H°$ because it is less strained.

The most informative and general point is thus that the absence of a negative $\Delta C°_p$ for nonspecific binding supports the inference from the specific complexes: The formation of a tight complementary interface produces the negative $\Delta C°_p$.

Factors that contribute to a negative $\Delta C°_p$

What are the origins of the large negative $\Delta C°_p$ observed for site specific protein-nucleic acid associations? Sturtevant [30] considered six possible sources of the large heat capacity and entropy changes frequently observed for processes involving proteins in aqueous solution and suggested that the primary contributors were the hydrophobic effect and changes in vibrational frequencies. Subsequently, it was found [31–33] that calorimetrically determined $\Delta C°_{p,fold}$ values for the folding of a number of proteins could be accounted for quantitatively using empirical correlations in which the burial of nonpolar surface (hydrophobic effect) made a dominant negative

contribution to $\Delta C^\circ_{p,\,fold}$ and the burial of polar surface made a partially compensating positive contribution.

The success of this approach for protein folding prompted Record and coworkers [12] to suggest that the hydrophobic effect was primarily responsible for the large negative ΔC°_p values observed for proteins binding to specific DNA sites. For many DNA-binding proteins, however, use of the empirical relationship between ΔC°_p and the change in solvent accessible surfaces gives nonpolar areas much larger than those calculated from crystal structures [16, 17, 20–22, 34, 35]. Put another way, the nonpolar surface area buried on complex formation, as computed from crystal structures, cannot account for all of the observed negative ΔC°_p. Searching for more surface area to contribute to ΔC°_p, Spolar and Record [12] suggested that induced folding of local regions of protein accompanying binding caused the burial of large amounts of additional nonpolar surface. However, for *Bam*HI and *Eco*RI endonucleases, even after including contributions from induced folding (comparing disorder-order transitions in the free and bound enzymes) we find [7] burial of previously solvated surfaces (polar and nonpolar) cannot account for the observed negative ΔC°_p values. For example, compare predicted values of $\Delta C^\circ_p \approx -0.3$ kcal/mol·K and $\Delta C^\circ_p \approx -0.04$ kcal/mol·K for *Eco*RI and *Bam*HI, respectively, with the observed ΔC°_p values of –1.2 kcal/mol·K to –2.5 kcal/mol·K (Table I). It may be that these calculations over-correct for burial of polar surface because these surfaces remain in relatively polar environments in protein-DNA complexes, but even the burial of nonpolar surface without correction can account for only 50% or less of the observed negative ΔC°_p values. Similar discrepancies (after correction for folding transitions) have beeen noted for other DNA-binding proteins, including trp and MetJ repressors [17, 35], the glucocorticoid receptor [34], the TATA binding protein [20] and interaction of the transcription factor GCN4 with its ATF/CREB site [21].

Based on calorimetric results for the trp repressor, it was suggested [17] that in addition to the hydrophobic effect, restriction of the degrees of freedom of the complementary polar, hydrated surfaces at specific protein-DNA interfaces must make a significant contribution to the negative ΔC°_p. High resolution structural data indicate that specific protein-DNA interfaces contain many "trapped" water molecules that are hydrogen-bonded to protein, DNA, or other water molecules (e.g., Fig. 2). Many of these water molecules mediate protein-base and protein-phosphate interactions that supplement the direct protein-DNA contacts. In this respect protein-DNA association is fundamentally unlike protein folding, where few water molecules become trapped in the interior.

The following argument allows us to place bounds on the contribution of trapped water molecules to ΔC°_p: Connelly [36] used crystalline salt hydrate data to estimate the heat capacity change for the incorporation of water molecules from bulk solvent into the interface of a protein-ligand complex as $\Delta C^\circ_p = -8.2$ cal/K·mol of water. We can take this value as an upper limit. At the *Eco*RI-DNA interface, there are 50 water molecules [37], that are either within hydrogen bonding distance of both the protein and DNA, or are involved in highly ordered water networks at the interface. Of these, 22 mediate direct interactions between protein and DNA functional groups at the interface. (Although total bound water obviously cannot be taken as equivalent to water visible in a crystal structure, there is little or no room remaining for additional waters to be "trapped" between protein and DNA.) Thus fully immobilized water might contribute an upper limit of ~–0.4 kcal/mol·K to ΔC°_p for specific recognition by *Eco*RI. A similar argument for *Bam*HI also gives an upper limit of –0.4 kcal/mol·K.

These upper limits are almost certainly not approached, because (a) water molecules that hydrate the polar surfaces of *unbound* protein and DNA are more immobilized than in bulk

Downloaded At: 15:35 29 January 2011

Downloaded At: 15:35 29 January 2011

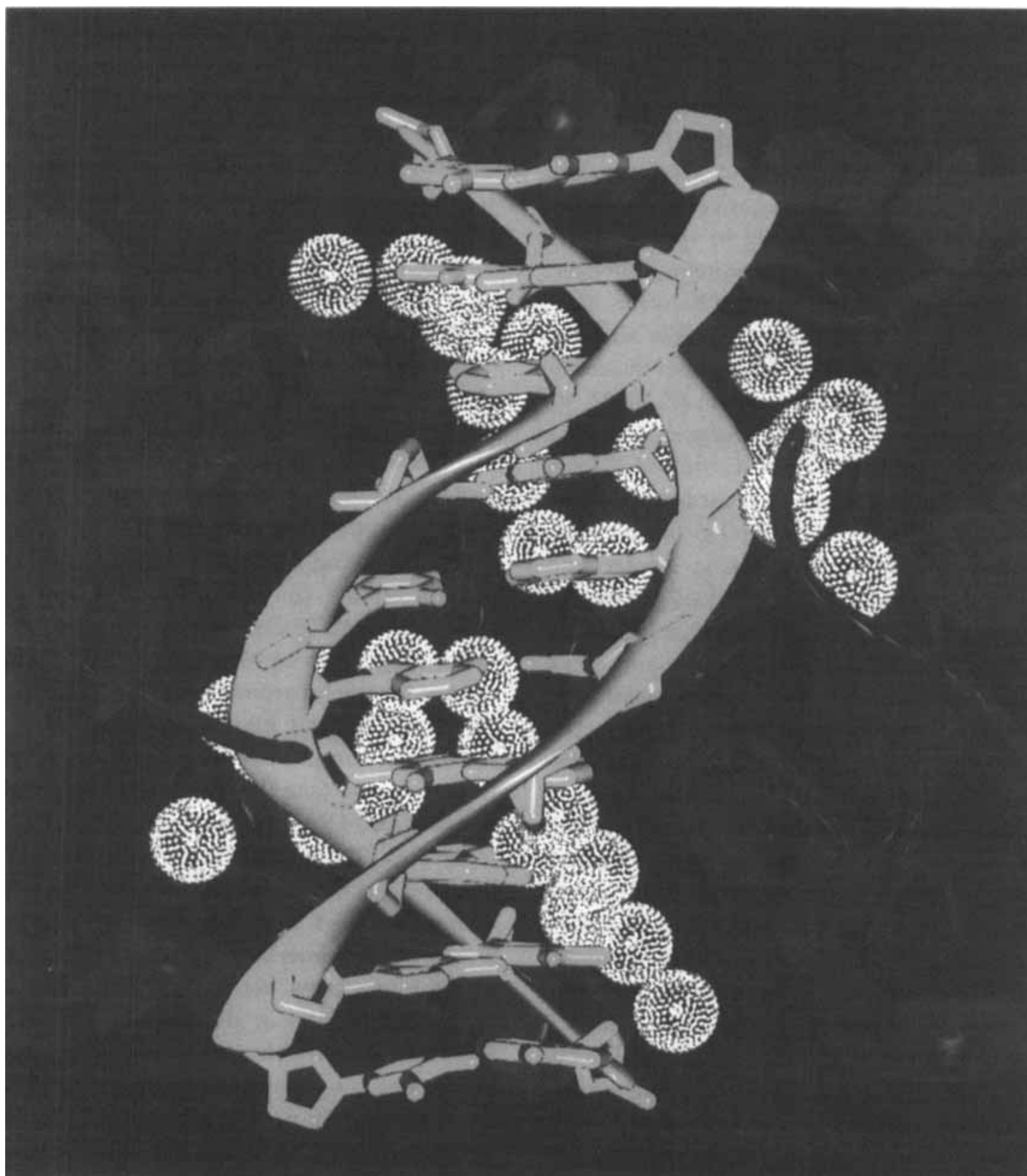

FIGURE 2 Tightly bound water molecules at the *Bam*HI endonuclease – DNA (major groove) recognition interface. View into the DNA (yellow) minor groove. Water molecules are displayed as white van der Waals spheres. Image generated with Midas-plus software using atomic coordinates (PDB accession number 1BHM) (See Color Plate I at the back of this issue)

solvent [36, 38, 39] and (b) water trapped in the protein-DNA interface in solution is likely not as immobilized as that in a crystalline salt hydrate. Thus, the difference in immobilization of water between free and bound states in solution is not as great as the difference between bulk solvent and a crystalline hydrate, so trapped interfacial water likely contributes less than 25% of the negative $\Delta C°_p$.

TABLE II Effect of flanking sequence context on binding and cleavage by *Eco*RI endonuclease

No.	Sequence[a]	K_A $(M^{-1})^b$	Relative $K_A{}^c$ Fold decrease	$\Delta\Delta G^\circ_{bind}{}^d$ (kcal/mol)	k_1 $(sec^{-1})^e$	k_2 $(sec^{-1})^e$
1	GGGCGGG**CGC** gaattc **GCG**GGCGG CCCGCCC**GCG** cttaag **CGC**CCGCC	5.1 (± 0.4) × 10^{10}	1	0	0.5 ± 0.1	0.5 ± 0.1
2	GGGCGGG**TGT** gaattc **ACA**GGCGG CCCGCCC**ACA** cttaag **TGT**CCGCC	3.1 (± 0.2) × 10^9	17	1.7 ± 0.1	0.6 ± 0.1	0.6 ± 0.1
3	GGGCGGG**GCA** gaattc **TGC**CGCGG CCCGCCC**CGT** cttaag **ACG**GCGCC	1.2 (± 0.2) × 10^9	43	2.2 ± 0.1	0.6 ± 0.1	0.6 ± 0.1
4	GGGCGGG**GTG** gaattc **CAC**CGCGG CCCGCCC**CAC** cttaag **GTG**GCGCC	7.3 (± 0.6) × 10^8	70	2.5 ± 0.1	0.5 ± 0.1	0.5 ± 0.1
5	GATGGA**CCCC** gaattc **GGGG**TGTA CTACCT**GGGG** cttaag **CCCC**ACAT	1.8 (± 0.2) × 10^8	283	3.3 ± 0.1	0.5 ± 0.1	0.5 ± 0.1
6	GCGCGG**AAAA** gaattc **TTTT**GCGG CGCGCC**TTTT** cttaag **AAAA**CGCC	1.3 (± 0.1) × 10^8	392	3.5 ± 0.1	0.2 ± 0.1	0.2 ± 0.1

a. Oligodeoxynucleotides (24 bp) containing the *Eco*RI recognition/cleavage site (gaattc, lower case) embedded in different flanking sequence contexts. For sequences no. 1 through 4, only the symmetrically disposed six base-pairs (boldface) surrounding the site vary; sequences 5 and 6 have C-tracts and A-tracts.
b. Equilibrium association constant (K_A) values determined in binding buffer A (22°C) by the nitrocellulose filter binding assay.
c. Calculated relative to the "best" sequence no. 1; $\frac{K_A(\text{"best"})}{K_A(\text{"variant"})}$.
d. Binding free energy penalty relative to "best" sequence no. 1; $\Delta\Delta G^\circ_{bind} = -RT \ln \frac{K_A(\text{"variant"})}{K_A(\text{"best"})}$.
e. First order cleavage rate constants: k_1, cleavage in top strand; k_2, cleavage in bottom strand.

Downloaded At: 15:35 29 January 2011

Immobilization in specific recognition complexes is not limited to the polar hydrated regions of the interface. Molecular dynamics simulations [28] showed that the dynamic fluctuations of DNA in specific *Eco*RI and *Eco*RV complexes are severely restricted relative to those of the free DNA. Both the computed "theoretical" and crystallographically determined B-factors for free and bound protein and DNA suggest that the restriction of configurational-vibrational degrees of freedom is global, not just localized to the DNA-protein interface.

Thus for *Eco*RI and *Bam*HI, the hydrophobic effect and the immobilization of water, protein sidechains, DNA bases and backbone elements all contribute to the negative ΔC°_p. All these factors relate to protein-DNA complementarity at the interface, and so we might expect that a large negative ΔC°_p is an invariant feature of forming an intimate recognition complex, a "thermodynamic signature" of specific recognition. Nonspecific complexes lack all of the above factors that contribute to ΔC°_p.

Context variants as probes of specific recognition interfaces

In order to probe the quantitative contributions of various factors that contribute to ΔC°_p for the specific complexes, we sought a situation in which we could hold the desolvation contributions (negative $\Delta C^\circ_{p,\ nonpolar}$ and positive $\Delta C^\circ_{p\ polar}$) constant while other factors varied. We found that we could do this by holding the six-base-pair recognition site constant while systematically varying the surrounding sequence context; that is, by changing the noncontacted base sequences flanking the recognition site. We have studied this most systematically for the *Eco*RI endonuclease-DNA interaction (measure-

ments have been made for 40 of the 64 symmetrical possibilities for 3 base-pairs surrounding the recognition site [1, 40] but the same kinds of phenomena are observed for *Eco*RV [41] and *Bam*HI [7]. Data for six representative 22-bp oligonucleotides (Table II) demonstrate that changing the noncontacted bases outside the *Eco*RI recognition site (GAATTC) causes the equilibrium affinities to vary by about 400-fold, equivalent to a range of differences in binding free energy ($\Delta\Delta G^{\circ}_{bind}$) of +3.5 kcal/mol.

The following lines of evidence indicate that modulation of binding by different contexts flanking the recognition site does not derive from alteration of protein-base or protein-phosphate contacts at the recognition interface:

(a) Crystal structures of *Eco*RI complexes with its recognition site embedded in two different contexts (identical to sequences 1 and 4, Table II, for the 3 bp on either side of the recognition site) show identical protein-base and protein-phosphate contacts [5, 42]. There are no direct protein contacts to base pairs outside the GAATTC recognition site.

(b) Ethylation-interference footprints [40, 43] are indistinguishable for the sequences shown in Table II, showing that there are no major changes in protein-phosphate contacts within or outside the site. All these sequences have identical salt-dependences of equilibrium binding (i.e., $d\ \log K_{obs}/d\log[\mathrm{NaCl}]$ is invariant), indicating equal stoichiometries of cation release upon binding [40].

(c) For all sequences except the flanking A-tract (known to confer unusual rigidity) [44, 45] there is no change in cleavage rate constant (Table II). Cleavage rate constants are extremely sensitive indicators of conformation in the complex, and are altered [46] by displacements of even a fraction of an Ångstrom in a flanking region two base-pairs away from the point of reaction. This sensitivity reflects the fact that specific protein-DNA binding complexes must be very precisely similar in structure to their transition-state complexes in order to efficiently utilize binding free energy for catalysis [1]. The absence of changes in catalytic rate for different flanking contexts therefore implies that any structural changes must be extremely subtle.

In addition, water release upon binding is not affected by flanking sequence context. For all six sequences, analysis of the dependence of the equilibrium affinities on the concentration of cosolvents (triethylene glycol, betaine, TMAO) indicates that the stoichiometry of water release upon complex formation is the same within experimental error [27]. Although the preferential exclusion-interaction behavior differs for each cosolvent, there are no differences in behavior amongst the six sequences with respect to any one cosolvent.

It is crucial that the protein-DNA recognition interfaces for the context variants are the same. That is, the burial of polar and nonpolar surfaces is presumably the same, and the number of water molecules trapped at the interface must be the same or very nearly so. Thus, the differences between the complexes must derive primarily from differences in strain and the vibrational factors related to immobilization of the protein and DNA, as we shall show below.

Sequence context affects the thermodynamics of binding

We next asked how sequence context affects the detailed thermodynamic parameters of the site-specific protein-DNA interaction. Fig. 3 (a) shows the van't Hoff plots for the *Bam*HI specific complex in different sequence contexts and Fig. 3 (b) the predicted ΔH° dependences on temperature. We have verified the van't Hoff ΔH° values at particular temperatures for all the sequences by calorimetric (ITC) measurements (data not shown). Similar van't Hoff analyses were obtained for the *Eco*RI interaction with its

Downloaded At: 15:35 29 January 2011

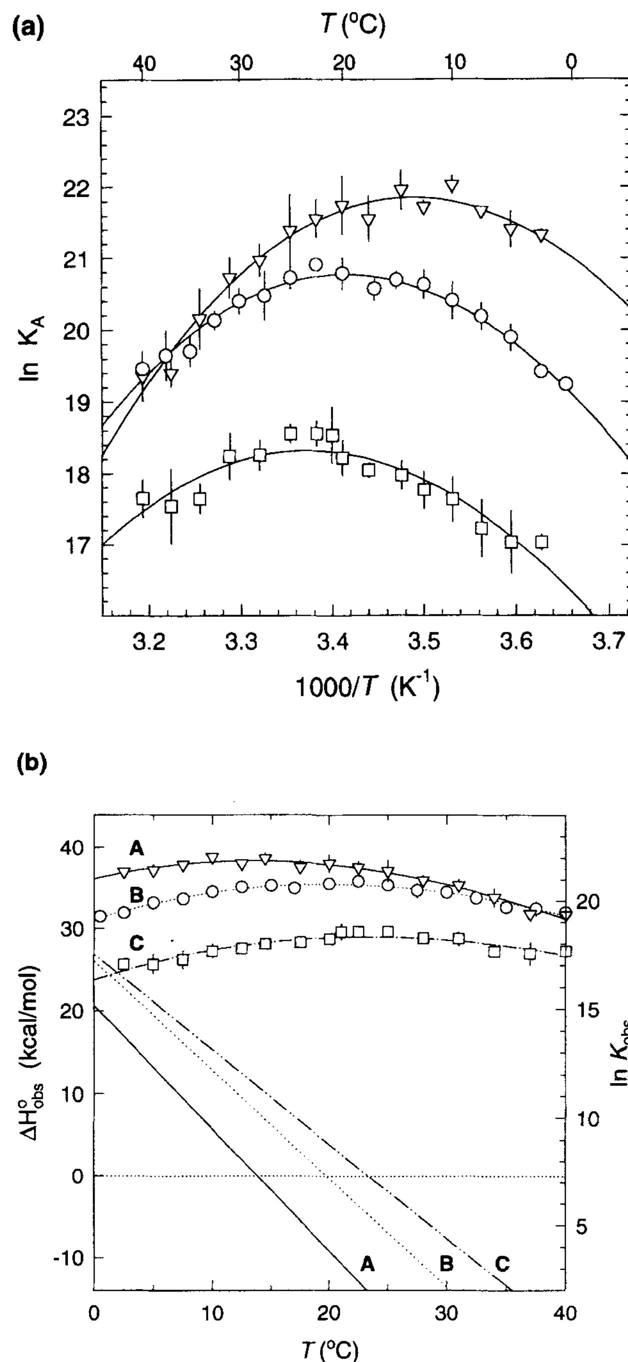

FIGURE 3 (a) Temperature dependence of *Bam*HI binding to the specific site embedded in different flanking contexts. Values of ln K_A (means ± SD for at least 3 determinations) are plotted as a function of 1/T for sequences indicated in Table I as "very good" (▽), "medium" (∘), and "worst" (□). (b) Enthalpic contribution to binding free energy as a function of temperature for *Bam*HI flanking context variants. Experimentally determined K_A values are for sequences designated in Table I as "very good" (▽), "medium" (∘) and "worst"(□). Lines (A, B, C) are enthalpic contributions to ΔG°, calculated from van't Hoff plots (Fig. 3a) as described in legend, Fig. 1(b)

Downloaded At: 15:35 29 January 2011

cognate site embedded in different contexts (Table I). The range of $\Delta G°_{bind}$ values for *Eco*RI may be greater only because we have tried more sequence contexts -- we may not yet have found the "best" or "worst" contexts for *Bam*HI.

It is evident that the relatively small changes in $\Delta G°_{bind}$ conceal much larger compensating changes in $\Delta H°$ and $\Delta S°$ (Table I). As $\Delta G°_{bind}$ improves for better sequence contexts, $\Delta H°$ becomes more favorable and $\Delta S°$ becomes less favorable. Since these relatively small changes in nucleotide sequence cannot have large effects on the enthalpies or entropies of the free DNA molecules, they must affect primarily the enthalpies and entropies of the protein-DNA complexes. Any molecular interpretation of how sequence context affects the complexes must therefore be consistent with the signs and magnitudes of changes in $\Delta H°$ and $\Delta S°$. In particular, the more favorable $\Delta H°$ does not reflect the emergence of any new protein-DNA contacts, so must be explicable in terms of decreased strain energy and/or optimization (bond lengths, angles) of a constant set of protein-DNA contacts.

Downloaded At: 15:35 29 January 2011

A striking finding that emerges from this analysis (Table I) is that as flanking context improves, $\Delta C°_p$ values become more negative, showing consistent trends for both enzymes. Fig. 4 presents these changes in $\Delta C°_p$ in graphical form. The slopes of the plots of $\Delta H°$ versus $T\text{-}T_H$ (where T_H is the temperature at which $\Delta H°$ is zero for the particular sequence) give the $\Delta C°_p$ values for the cognate sites embedded in different flanking contexts. For *Eco*RI, the decrease in $\Delta C°_p$ is dramatic: the $\Delta C°_p$ value for the cognate site embedded in the best flanking context (seq 1, table II) is 1.3 kcal/mol·K more negative than (i.e., nearly double) that of a cognate site in a different flanking context (Seq. X, Table II) that binds more weakly ($\Delta\Delta G°_{bind}$ = +2.5 kcal/mol).

A key question is: What is the source of this $\Delta\Delta C°_p$(that is, a difference in $\Delta C°_p$) among context variants? As enumerated earlier, a more negative $\Delta C°_p$ might derive from an increased hydrophobic effect (i.e., burial of more solvent accessible surface), from further immobilization of the trapped water molecules at the protein-DNA interface, from greater reduction of the vibrational-conformational freedom of the protein and DNA, or in some measure from all of these.

These complexes cannot differ only in the hydrophobic contribution to $\Delta C°_p$, for if that were the case the $\Delta S°$ values should converge at 386 K, where the entropic contribution from the hydrophobic effect ($\Delta S°_{HE}$) disappears because water-hydrocarbon mixtures show ideal entropy of mixing [47]. In fact there is strong divergence of the $\Delta S°$ values at 386 K (data not shown).

An increased hydrophobic effect in the more favored contexts can also be rejected. If the hydrophobic contribution increased for those sequences that bind better, then $\Delta S°$ should become more favorable. On the contrary, it becomes less favorable (Table I).

We also reject the possibility that the complexes with different sequence contexts differ primarily in the *number* of water molecules trapped at the the interface The magnitudes of $\Delta\Delta H°$ and $T\Delta\Delta S°$ as sequence context changes (for *Eco*RI, 19 and 15 kcal/mol, respectively) are much too large to reflect a difference of binding one or two water molecules on each side of the recognition site in the flanking regions. [See Dunitz [39] for a *maximum* estimate of 2 kcal/mol at 298 K for the entropic contribution to the free energy cost of immobilizing water in a crystalline salt hydrate]. In the high resolution (1.85 Å) cryogenic structure of the *Eco*RI-DNA complex [37], only 5 waters are bound in the immediately flanking regions on each side of the recognition site. In the implausible case that *all* of these are absent from the complex with another flanking sequence, this would scarcely account for the observed $\Delta\Delta H°$ and $T\Delta\Delta S°$,

The signs of $\Delta\Delta H°$ and $\Delta\Delta S°$ as $\Delta G°$ improves are consistent with further immobilization of protein and DNA and/or interfacial water molecules. That is, the marked $\Delta\Delta C°_p$ is not dominated by changes in the *numbers* of solvent

Downloaded At: 15:35 29 January 2011

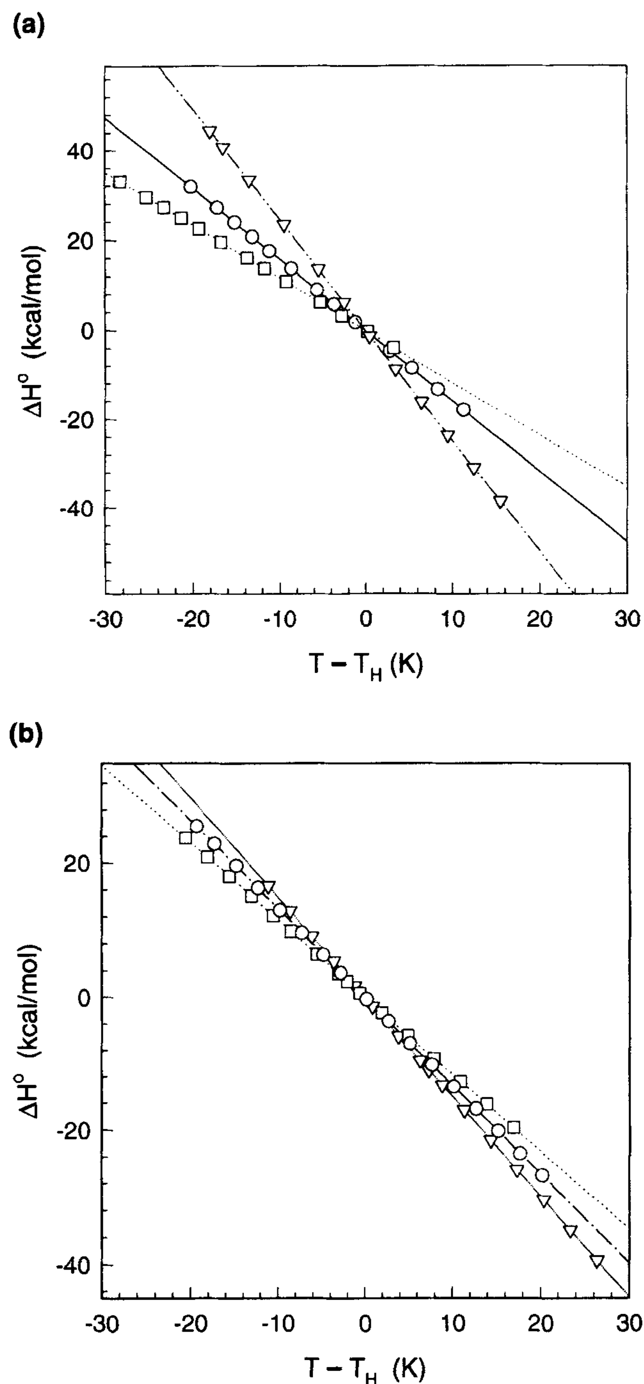

FIGURE 4 Heat capacity changes ($\Delta C°_p$) for formation of specific *Eco*RI (a) and *Bam*HI (b) complexes with different flanking context variants. The slope of each line is the $\Delta Cp°$ value for the particular sequence. Values of $\Delta H°$ at 25°C and of $\Delta C°_p$ are tabulated in Table I. T_H is the temperature where $\Delta H° = 0$. (a) EcoRI sequences: "best" (▽), $T_H = 294.5 \pm 0.3$ K; "medium" (◦), $T_H = 297.8 \pm 1.6$ K; "worst" (□), $T_H = 305.9 \pm 4.3$ K. (b) *Bam*HI sequences: "very good" (▽), $T_H = 287.0 \pm 0.7$ K; "medium" (◦), $T_H = 292.8 \pm 0.3$ K; "worst" (□), $T_H = 296.5 \pm 0.8$ K

Downloaded At: 15:35 29 January 2011

molecules released or trapped, but by further restriction of vibrational-configurational degrees of freedom in protein-DNA complexes. Given the range of changes shown in Table I for variations of only flanking context (without varying the recognition site), we can see that these factors might be large enough in magnitude to contribute significantly to the negative ΔC°_p that distinguishes specific from nonspecific complexes.

But which is the dominant contribution to the difference in ΔC°_p (i.e. $\Delta\Delta C^\circ_p$) – further immobilization of water or of the protein-DNA surfaces? Taking the 50 interfacial waters in the *Eco*RI complex as an example, much of their immobilization must have been accomplished in the "worst" sequence context. Considering that the total increment of heat capacity change for 50 waters going from bulk solvent to crystal hydrate is only −0.4 kcal/mol·K, additional immobilization of a constant number of interfacial waters with improving sequence context cannot possibly account for the observed $\Delta\Delta C^\circ_p$ of −1.3 kcal/mol·K.

Thus, vibrational-configurational factors at the protein-DNA interface must be major contributors to ΔC°_p for specific protein-DNA recognition. The "baseline" ΔC°_p for specific recognition in a "poor" context includes contributions from the hydrophobic effect, immobilized trapped water and the immobilization of protein and DNA. (It follows that previous calculations [12] of the hydrophobic driving force from ΔC°_p should be reconsidered.) Further decrease in ΔC°_p for "better" contexts represents primarily additional immobilization of the DNA-protein interface.

Entropy-enthalpy compensation

The data in Table I show a correlation between ΔH° and ΔS° for the various contexts, and this is shown graphically in Fig. 5 for the interaction of *Bam*HI endonuclease with its specific site embedded in 7 different sequence contexts. The data plainly show that as the context gets better (that is, ΔG°_{bind} becomes more negative) ΔH° gets better and ΔS° gets worse. Such a relationship is called "enthalpy-entropy compensation" (also known as an "isoequilibrium" relationship or "isokinetic" for $\Delta H^\ddagger$ and $\Delta S^\ddagger$) and has been discussed extensively [39, 48–55] with reference to a variety of "host-guest" molecular recognition processes both chemical and biological, especially those occurring in aqueous solution.

Many cases of enthalpy-entropy compensation have been reported for processes in which ΔH° and ΔS° are linearly correlated as the chemical structure of the ligand is changed (e.g. a homologous series) under fixed experimental conditons. Enthalpy-entropy compensation in protein-DNA interactions has been discussed with respect to solvent reorganization effects (ΔH° and ΔS° as a function of temperature) [16, 22, 56] but here we consider compensation over a "sequence space" analogous to a homologous series of ligands; that is, a set of DNA molecules in which a constant recognition site is embedded in surrounding DNA of constant length but varying nucleotide sequence.

We have applied several critical criteria [52, 57, 58] to confirm that the compensation effects arise from physico-chemical factors rather than covariance in the errors for determining ΔH° and ΔS°: (1) the compensation temperature (obtained from the slope of the ΔH° versus ΔS° plot) is markedly different from the experimental temperature. (2) a ΔG° vs. ΔH° plot shows good monotonic relationship. That is, correlated errors in ΔH° and ΔS° will give a linear ΔH° vs ΔS° plot with $T_c = T_{exp}$, but will show up as a scattered ΔG° versus ΔH° plot.

The directions of change in ΔH° and ΔS° as well as the compensation between them suggest a general explanation for the sequence-context effects. In these cases, we can neglect the entropic and enthalpic effects of solvent (e.g., water release or bound water in the interface) because these are almost certainly invariant as sequence context changes. Complexes containing the "better" sequence contexts must immobilize more

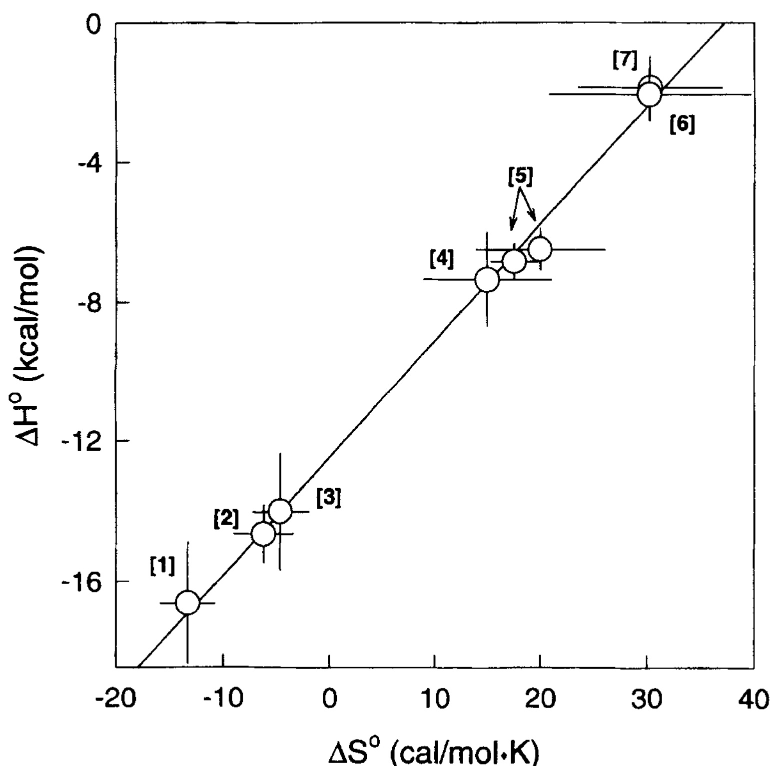

FIGURE 5 Enthalpy-entropy compensation for *Bam*HI binding across "sequence space" of varying flanking context. Values of $\Delta H°$ and $\Delta S°$ (298 K) determined as in Fig. 1. General form of compensation equation is $\Delta H° = b\,(\Delta S°) + a$, where b is the "compensation temperature" and a is $\Delta G°$ at the compensation temperature. Regression equation for this plot is $\Delta H° = 335\,(\Delta S°) - 12.5$ ($r^2 = 0.99$). Note that the compensation temperature (335 ± 8.7 K) is significantly different from the experimental temperature (298 K), so that the observed correlation between $\Delta H°$and $\Delta S°$ is due to physico-chemical and not statistical factors (see reference [57] for discussion). Sequences [1], [5] and [7] are those designated "very good", "medium", and "worst" in Table I. Other sequences (*Bam*HI recognition site in lowercase): CGCGGGTTATggatccATAAGGGC [2], TGGGTGggatccCACCCAC [3], GGGATGGGTGggatccCCACCAC [4] and GGGATGGGGGggatccCCACCAC [6]

completely the protein and DNA, because $\Delta S°$ is less favorable. At the same time, $\Delta H°$ becomes more favorable because the energy required to strain the protein and DNA in the complexes becomes less unfavorable. In other words, a sequence context that improves binding free energy does so because the complexes "fit better".

Thermodynamics and specificity

We can now see that the highly specific recognition of a DNA site by a protein has predictable thermodynamic consequences. By the term "specific", we mean not only that binding to the correct recognition site is much stronger than binding to other sites, but also – for catalytic proteins like restriction endonucleases – that the recognition complex lies along the path to the transition state and thus facilitates the efficient use of binding energy for catalysis [1].

The formation of such a complex requires that an extensive complementary interface be formed between protein and DNA. As a result, entropy decreases not only by the restricted translation and rotation of the binding partners, but also by the loss of vibrational freedom and the entrapment of water molecules in the interface. This

Downloaded At: 15:35 29 January 2011

loss of entropy is compensated, or even overcompensated, by the release of water and cations.

The very same components that contribute to binding entropy also contribute to the loss of heat capacity of the system. The loss of vibrational freedom in protein, DNA and interfacial trapped water, as well as the release of water from nonpolar surfaces, all contribute to a large negative ΔC°_{p}, while the release of water from polar surfaces makes an opposing positive contribution to ΔC°_{p}. It follows that the formation of a highly specific, complementary interface must invariably entail a negative ΔC°_{p}, and this is what we and others observe as the "thermodynamic signature" of specific binding. In the nonspecific protein-DNA interfaces, these factors are absent or greatly diminished, so ΔC°_{p} is near-zero.

If we make some relatively minor adjustment to the specific complex, say, a change in sequence surrounding the recognition site, the binding free energy can be improved somewhat. Our data indicate that the "fit" in the complex improves without forming any new protein-DNA interactions. The interface becomes more immobilized, so ΔS° becomes less favorable and ΔC°_{p} becomes more negative. However, ΔH° improves *even without any new interactions* by the reduction of strain.

It is striking that these changes in ΔH° and ΔS° are in opposite directions and tend to compensate each other. It is equally striking that compensation must be imperfect, because ΔG°_{bind} does change. One of the major challenges for the immediate future is to determine which aspects of these molecular adjustments produce perfect entropy-enthalpy compensation and which do not.

MATERIALS AND METHODS

*Eco*RI endonuclease was prepared as described [59]. *Bam*HI, overexpressed and purified as described previously [60], was the generous gift of R. Kucera and I. Schildkraut (New England Biolabs). All oligodeoxynucleotide sequences were synthesized on a Perseptive Bioystems Expedite 8909 synthesizer at the Pittsburgh DNA Synthesis Facility, Dept. of Biological Sciences, University of Pittsburgh. Purification, annealing and end-labeling of oligonucleotides were as described previously [59, 61].

For the van't Hoff analyses, enzyme-DNA equilibrium association constants (K_A) for specific sites were determined by the direct binding method [14, 59] and for nonspecific sites by an equilibrium competition method [15, 41] using the nitrocellulose filter binding assay. Binding buffer A is 10 mM bis-tris-propane, 180 mM NaCl, 100 mg/mol bovine serum albumin, 5 mM dithiothreitol. Binding buffer B differs only in that it contains 180 mM NaF instead of the NaCl. Binding buffers were titrated to pH 7.3 at each experimental temperature. Binding reactions were incubated and filtered at the appropriate temperature in an environmentally controlled room (temperature range 2°±0.2 to 42°C ± 0.2°C). From the experimentally determined values of K_A and ΔG° as a function of temperature, values of ΔC°_{p}, ΔH° and ΔS° were calculated as described [16]. First order cleavage rate constants were determined at pH 7.3 (22°C) as described [4, 61]. Isothermal titration calorimetry (ITC) experiments were performed using a MCS ITC (Microcal Inc., Northampton, MA). Enzyme and DNA stocks were exhaustively dialyzed against the calorimetric binding buffer containing 20 mM potassium phosphate, 380 mM NaF, 1 mM disodium (ethylenedinitrilo) tetraacetate, 5% glycerol (pH 7.3 at T = 4°C). Binding isotherms were obtained by injecting a series of 5-μL increments of DNA solution (50 to 120 μM) into a sample cell (volume =1.34 mL) containing enzyme (3 to 5 μM) at appropriate temperatures until complete saturation where there was no further change in heat of reaction. Data were analyzed using Origin, a software package provided by Microcal [62]. The net heat of binding

Downloaded At: 15:35 29 January 2011

was obtained by correcting the heat of reaction (integration of the peak obtained after each injection) for the heat of dilution. The Origin software applies nonlinear least squares to the binding isotherm to determine best-fit values for the binding constant, the observed binding enthalpy and overall reaction stoichiometry. The fitted stoichiometry of n=1 (moles DNA per mole of enzyme dimer) is that seen in the X-ray structures of the co-crystalline complexes and also that determined by other biochemical methods [7, 59]; this provides a further validation of the fitted parameters.

Acknowledgements

This work was supported by a grant (GM-29207) from the U.S. National Institutes of Health.

Downloaded At: 15:35 29 January 2011

References

1. Jen-Jacobson, L. (1997). *Biopoly.* **44**, 153.
2. Otwinowski, Z., Schevitz, R.W., Zhang, R.G., Lawson, C.L., Joachimiak, A., Marmostein, R.Q., Luisi, B.F., and Sigler, P. (1988). *Nature.* **335**, 321.
3. Harrison, S.C. and Aggarwal, A.K. (1990). *Annu. Rev. Biochem*, **59**, 933.
4. Lesser, D.R., Kurpiewski, M.R., and Jen-Jacobson, L. (1990). *Science.* **250**, 776.
5. Kim, Y., Choi, J., Grable, J.C., Greene, P., Hager, P., and Rosenberg, J.M. (1994). In: *Structural Biology: The State of the Art. Proc. of the Eighth Conversation, SUNY, Albany*, R. Sarma and M. Sarma, (Eds). Adenine Press: New York. p. 225.
6. Dickerson, R.E. (1998). *Nucl. Acids. Res.* **26**, 1906.
7. Engler, L.E. (1998). *Ph.D. Thesis, University of Pittsburgh, Pittsburgh, PA.*
8. Dickerson, R.E. (1998). In: *Structure, Motion, Interaction and Expression of Biological Macromolecules, Proceedings of Tenth Conversation in Biomolecular Stereodynamics*, R.H. Sarma and M.H. Sarma, (Eds). Academic Press: Schenectady, N.Y.
9. Rosenberg, J.M. (1991). *Curr. Op. Struct. Biol.* **1**, 104.
10. Newman, M., Strzelecka, T., Dorner, L.F., Schildkraut, I., and Aggarwal, A.K. (1995). *Science.* **269**, 656.
11. Record, M.T., Jr, Ha, J.H., and Fisher, M.A. (1991). *Meth Enzymol.* **208**, 291.
12. Spolar, R.S. and Record, M.T., Jr. (1994). *Science*, **263**, 777.
13. Jayaram, B., McConnell, K.J., Dixit, S.B., and Beveridge, D.L. (1999). *J. Comp. Phys.*, 151, 333.
14. Riggs, A.D., Suzuki, H., and Cohn, M. (1970). *J. Mol. Biol.* **48**, 67.
15. Lin, S.-Y. and Riggs, A.D. (1972). *J. Mol. Biol.* **72**, 671.
16. Ha, J.H., Spolar, R.S., and Record, M.T., Jr. (1989). *J Mol Biol.* **209**, 801.
17. Ladbury, J.E., Wright, J.G., Sturtevant, J.M., and Sigler, P.B. (1994). *J. Mol. Biol.* **218**, 669.
18. Hyre, D.E. and Spicer, L.D. (1995). *Biochemistry.* **34**, 3212.
19. Merabet, E. and Ackers, G.K. (1995). *Biochemistry.* **34**, 8554.
20. Petri, V., Hsieh, M., and Brenowitz, M. (1995). *Biochemistry.* **34**, 9977.
21. Berger, C., Jelesarov, I., and Bosshard, H.R. (1996). *Biochemistry.* **35**, 14984.
22. Frank, D.E., Saecker, R.M., Bond, J.P., Capp, M.W., Tsodikov, O.V., Melcher, S.E., Levandoski, M.M., and Record, M.T., Jr. (1997). *J. Mol. Biol.* **267**, 1186.
23. Winkler, F.K., Banner, D.W., Oefner, C., Tsernoglou, D., Brown, R.S., Heathman, S.P., Bryan, R.K., Martin, P.D., Petratos, K., and Wilson, K.S. (1993). *The EMBO Journal.* **12**, 1781.
24. Luisi, B., Xu, W.X., Otwinowski, Z., Freedman, L.P., Yamamoto, K.R., and Sigler, P.B. (1991). *Nature.* **352**, 497.
25. Jen-Jacobson, L. (1995). *Meth. Enzymol.* **259**, 305.
26. Takeda, Y., Ross, P.D., and Mudd, C.P. (1992). *Proc. Natl. Acad. Sci. USA.* **89**, 8180.
27. Cao, D., Engler, L.E., and Jen-Jacobson, L. unpublished results.
28. Duan, Y., Wilkosz, P., and Rosenberg, J.M. (1996). *J. Mol. Biol.* **264**, 546.
29. Kozlov, A.G. and Lohman, T.M. (1998). *J. Mol. Biol.* **278**, 999.
30. Sturtevant, J.M. (1977). *Proc. Natl. Acad. Sci. USA.* **74**, 2236.
31. Spolar, R.S., Livingstone, J.R., and Record, M.T., Jr (1992). *Biochemistry.* **31**, 3947.
32. Murphy, K.P. and Freire, E. (1992). *Adv. Prot. Chem.* **43**, 313.
33. Makhatadze, G.I. and Privalov, P.L. (1995). *Adv. Prot. Chem.* **47**, 307.
34. Lundbäck, T., Cairns, C., Gustafsson, J.-A., Carlstedt-Duke, J., and Härd, T. (1993). *Biochemistry*, **32**, 5074.
35. Morton, C.J. and Ladbury, J.E. (1996). *Protein Science.* **5**, 2115.
36. Connelly, P.R. (1997). In: *Thermodynamics and structure based drug design.* R.G. Landes: Austin, TX.
37. Horvath, M., Choi, J., Kim, Y., Wilkosz, P.A., Chandresekhar, K., Grable, J.C., Greene, P., Hager, P., and Rosenberg, J.M. *Proteins. Struct. Funct. Genet. submitted.*
38. Karplus, P.A. and Faerman, C. (1994). *Curr. Opin. Struct. Biol.* **4**, 770.
39. Dunitz, J.D. (1995). *Chemistry & Biology.* **2**, 709.
40. Kurpiewski, M.R., Chi, D.J., and Jen-Jacobson, L., unpublished results.
41. Engler, L.E., Welch, K.K., and Jen-Jacobson, L. (1997). *J. Mol. Biol.* 269, 82.
42. Grable, J.C. (1990). *Ph.D. Thesis, University of Pittsburgh, Pittsburgh, PA.*
43. Becker, M.M., Lesser, D., Kurpiewski, M., Baranger, A., and Jen-Jacobson, L. (1988). *Proc. Natl. Acad. Sci.* **85**.
44. Leroy, J.L., Charretier, E., Kochoyan, M., and Guéron, M. (1988). *Biochemistry.* **29**, 4227.
45. Zhurkin, V.B., Ulyanov, N.B., Gorin, A.A., and Jernigan, R.L. (1991). *Proc. Natl. Acad. Sci. USA.* **88**, 7046.
46. Kurpiewski, M.R., Koziolkiewicz, M., Wilk, A., Stec, W.J., and Jen-Jacobson, L. (1996). *Biochemistry.* **35**, 8846.
47. Baldwin, R.L. (1986). *Proc. Natl. Acad. Sci.* **83**, 8069.

48. Leffler, J.E. and Grunwald, E., Rates and Equilibria of Organic Reactions, Wiley, New York, 1963, reprinted by Dover, 1989.
49. Lumry, R. and Rajender, S. (1970). *Biopolymers*. **9**, 1125.
50. Exner, O. (1973). *Progr. Phys. Org. Chem.* **10**, 414.
51. Krug, R.R., Hunter, W.G., and Grieger, R.A. (1976). *J. Phys. Chem.* **80**, 2335.
52. Tomlinson, E. (1983). *Int. J. Pharmaceutics*. **13**, 115.
53. Eftink, M.R., Anusiem, A.C., and Biltonen, R.L. (1983). *Biochemistry*. **22**, 3884.
54. Searle, M.S., Westwell, M.S., and Williams, D.H. (1995). *J. Chem. Soc. Perkin Trans.* **2**, 141.
55. Grunwald, E. and Comeford, L.L In: *Protein-Solvent Interactions*, R.B. Gregory, (Ed.), Marcel Dekker, Inc, New York, 1995, p. 421.
56. Lundback, T. and Hard, T. (1996). *Proc. Natl. Acad. Sci. USA*. **93**, 4754.
57. Krug, R.R., Hunter, W.G., and Grieger, R.A. (1976). *J. Phys. Chem.* **80**, 2341.
58. Linert, W. (1994). *Chem. Soc. Rev.* **23**, 429.
59. Jen-Jacobson, L., Kurpiewski, M., Lesser, D., Grable, J.C., Boyer, H.W., Rosenberg, J.M., and Greene, P.J. (1983). *J. Biol. Chem.* **258**, 14638.
60. Jack, W.E., Greenough, L., Dorner, L.F., Xu, S.-Y., Strzelecka, T., Aggarwal, A.K., and Schildkraut, I. (1991). *Nucl. Acids Res.* **19**, 1825.
61. Lesser, D.R., Grajkowski, A., Kurpiewski, M.R., Koziolkiewicz, M., Stec, W.J., and Jen-Jacobson, L. (1992). *J. Biol. Chem.* **267**, 24810.
62. Wiseman, T., Williston, S., Brandts, J.F., and Lin, L.N. (1989). *Anal. Biochem.* **179**, 131.

Downloaded At: 15:35 29 January 2011