Published in final edited form as:

J Mol Biol. 2014 January 9; 426(1): 84–104. doi:10.1016/j.jmb.2013.09.005.

Thermodynamic and Structural Basis for Relaxation of Specificity in Protein-DNA Recognition

Paul J. Sapienza^{1,2}, Tianyi Niu^{1,3}, Michael R. Kurpiewski¹, Arabela Grigorescu^{1,4}, and Linda Jen-Jacobson^{*,1}

¹ Department of Biological Sciences, University of Pittsburgh Pittsburgh, PA 15260 USA

Abstract

As a novel approach to the structural and functional properties that give rise to extremely stringent sequence specificity in protein-DNA interactions, we have exploited “promiscuous” mutants of EcoRI endonuclease to study the detailed mechanism by which changes in a protein can relax specificity. The A138T promiscuous mutant protein binds more tightly to the cognate GAATTC site than does wild-type EcoRI, yet displays relaxed specificity deriving from tighter binding and faster cleavage at EcoRI* sites (one incorrect base pair). AAATTC EcoRI* sites are cleaved by A138T up to 170-fold faster than by wild type enzyme if the site is abutted by a 5'-purine-pyrimidine (5'-RY) motif. When wild-type protein binds to an EcoRI* site, it forms structurally adapted complexes with thermodynamic parameters of binding that differ markedly from those of specific complexes. By contrast, we show that A138T complexes with 5'-RY-flanked AAATTC sites are virtually indistinguishable from wild-type specific complexes with respect to the heat capacity change upon binding (ΔC_p), the change in excluded macromolecular volume upon association, and contacts to the phosphate backbone. While the preference for the 5'-RY motif implicates contacts to flanking bases as important for relaxed specificity, local effects are not sufficient to explain the large differences in ΔC_p and excluded volume, as these parameters report on global features of the complex. Our findings therefore support the view that specificity does not derive from the additive effects of individual interactions, but rather from a set of cooperative events that are uniquely associated with specific recognition.

Keywords

osmolyte; heat capacity; cosolute effect; indirect readout; restriction endonuclease

Introduction

Noncovalent interactions between proteins and DNA are crucial to gene expression, chromatin structure and dynamics, DNA synthesis and repair. A deep understanding of the molecular basis of specificity in these interactions is required to devise strategies for

© 2013 Elsevier Ltd. All rights reserved.

*Corresponding author: Telephone: 412-624-4969 FAX: 412-624-4759 ljen@pitt.edu.

²Current address: Division of Chemical Biology and Medicinal Chemistry, UNC Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

³Current address: Department of Neurosurgery, UCLA School of Medicine, Los Angeles, CA. 90095, USA

⁴Current address: Keck Biophysics Facility, Northwestern University, Evanston, IL 60208, USA

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

therapeutic interventions, for designed modifications to existing DNA-binding proteins, and in the long run for rational design of novel proteins. Proteins vary considerably in their level of DNA sequence discrimination, ranging from nearly complete indifference to sequence by DNA polymerases, to extremely stringent sequence discrimination by restriction endonucleases¹⁻⁴. Our basic tenet has been that restriction endonucleases will therefore reveal the greatest number of rules and constraints governing specificity, and that other less stringent DNA-protein interactions might be understood in terms of relaxation of some of these constraints.

In site-specific DNA-protein complexes, the bound DNA recognition site is frequently distorted, to an extent ranging from subtle variations in base or backbone conformations to extreme bending and writhing. Consequently, the attainment of precise complementarity between the protein and DNA bases and backbone incurs a free-energy cost, which depends not only on the distortion in the bound state, but also on the sequence-dependent conformational propensities of the free DNA^{1; 2; 5-7}. Thus, in addition to “direct readout” of the hydrogen-bond donor-acceptor information on the base edges, proteins use “indirect readout”, meaning both the exploitation of local conformational features of the free DNA (groove widths, groove depths, etc.) and the sequence-dependent cost of optimizing protein-DNA complementarity^{2; 6; 8-10}. The fact that the cost of DNA distortion is sequence-dependent implies that “direct readout” and “indirect readout” cannot usually be considered independent.

The influence of sequence-dependent distortability, and the delicate balance among large opposing energies in recognition complexes², cause recognition interfaces to shift into “adaptive” structures in response to as little as a single base-pair change in DNA sequence, to minimize the overall binding free energy^{1; 11-13}. Such complexes have structural, functional and thermodynamic^{1-3; 14-19} properties distinct from those of the cognate complex, but are nevertheless distinct from nonspecific protein-DNA complexes. Thus, protein complexes at incorrect sites should not be viewed as correct complexes that lack only a few protein-DNA interactions; they are not isomorphous.

Restriction endonucleases require extreme sequence specificity for their *in vivo* function of selectively destroying invading DNA. The EcoRI restriction endonuclease has illuminated many of the properties that enable stringent discrimination against noncognate DNA sites. The enzyme prefers GAATTC over miscognate sites (one incorrect base-pair, so-called EcoRI* sites) by as much as 10¹¹-fold when both DNA binding and cleavage are taken into account². In the specific complex the DNA is distorted (“kinked”)²⁰ such that the scissile phosphates (one in each DNA strand) are properly aligned with the two catalytic sites of the EcoRI homodimer²¹⁻²³. The energetic cost of achieving this deviation from B-form DNA is lowest for the GAATTC site, but is greater for sites differing by even a single base pair². Further, when the free enzyme binds to the specific DNA recognition site, two “arms” protruding from the main domain of the endonuclease undergo a disorder to order transition²⁴ so as to make critical contacts to bases and phosphates. The most probable position of the arms is similar in specific and EcoRI* complexes, yet the distribution of positions is broadened in the EcoRI* complex²⁵, suggesting the possibility that the EcoRI* complex ensemble is more populated by conformers that lack contacts between the arms and DNA. Finally, the protein-DNA contacts that stabilize the specific complex persist in the transition state², such that binding energy is utilized efficiently at the catalytic step; thus, the complex with the GAATTC site has been called a “pre-transition state” complex². By contrast, binding to EcoRI* sites triggers adaptation into structures that minimize binding energy yet are off the path to the transition state, thus greatly decreasing catalytic rates². Complexes that deviate by two or more base-pairs are completely incompetent for cleavage¹. This integrated discrimination mechanism illustrates why it is so difficult to

change the specificity of a restriction endonuclease like EcoRI by rational site-directed mutagenesis²⁶⁻³².

As a novel approach to identifying the structural and functional features that give rise to extremely stringent specificity, we have studied “promiscuous” EcoRI mutants to understand the detailed mechanism by which changes in the protein can relax specificity. These mutants were isolated by Heitman and Model³³ on the basis that they caused DNA damage. Some mutant EcoRI endonuclease mutations were so lethal, even when co-expressed with the EcoRI methyltransferase, that they could only be expressed conditionally. Surprisingly, each of the relaxed specificity mutant proteins actually binds more tightly to the specific site than does the wild type protein³. However, while the mutant proteins still prefer the GAATTC site in binding and cleavage, relaxed specificity *in vivo* is attributable to tighter binding and faster cleavage of EcoRI* sites, relative to the wild type enzyme³. The crystal structure³⁴ of the most promiscuous mutant protein (A138T) in complex with the GAATTC site shows that while the structures of the protein and DNA in the complex are nearly identical to those in the wild-type complex, the longer Thr138 side chain permits the mutant enzyme to make novel water-mediated and packing interactions with the two bases immediately upstream of GAATTC; these interactions are degenerate for any 5'-flanking dinucleotide step with a purine-pyrimidine (RY) motif. Importantly, the A138T enzyme most rapidly cleaves EcoRI* sites that are flanked by 5'-RY dinucleotides^{33; 34} suggesting that these novel flanking contacts also form in A138T-EcoRI* complexes.

These observations raised the question of whether A138T-EcoRI* complexes more closely resemble the ‘specific’ binding mode of wild-type-GAATTC complexes, or the distinct ‘adaptive’ binding mode of wild-type-EcoRI* complexes. Here we use several thermodynamic and structural criteria to show that A138T complexes with RY-flanked EcoRI* sites are virtually indistinguishable from wild-type specific complexes, in contrast to the aberrant complexes formed by wild type enzyme at EcoRI* sites. Two of the criteria are “system properties”: the heat capacity change (ΔC_p) upon protein-DNA binding, and the effect of cosolutes on binding, which is related to the desolvation of macromolecular surfaces upon association. For both these metrics, formation of the wild-type-EcoRI* complex produces values intermediate between those of specific and nonspecific DNA complexes, whereas A138T-EcoRI* complexes show values more similar to the specific complex. In addition, wild-type-EcoRI* “adaptive” complexes lack many of the key protein-phosphate contacts, whereas the A138T-EcoRI* complexes have phosphate contacts more similar to those of the wild-type-GAATTC complex. For each of these properties, the A138T-EcoRI* complex behaves like the specific complex if and only if the EcoRI* site is abutted by a 5'-RY motif, implicating the novel contacts to this flanking region as critical for relaxed specificity. However, our data also imply that the A138T mutation does not undermine sequence discrimination by these local effects alone. Thus, this study supports the view that specificity does not derive from the additive effects of individual interactions but rather from a set of cooperative events that are unique to specific complex formation.

Results

Flanking context effects distinguish specific from adaptive complexes

We have shown previously³ that the three base pairs flanking the GAATTC recognition site modulate specific binding of the wild-type and A138T enzymes over a range of 4 kcal/mol (~1000-fold) (data provided for reference in Table 1, columns 2-3). Because we discuss flanking bases extensively in this work, we define the numbering for oligonucleotide constructs in scheme 1: where the hexanucleotide binding site is underlined, $R_{-3} = G$ for specific constructs, $R_{-3} = A$ for the EcoRI* constructs used in this study, $X_{-6}X_{-5}X_{-4}$ is the

variable flanking sequence, and $Z_4Z_5Z_6$ is the reverse complement of $X_{-6}X_{-5}X_{-4}$ such that flanking triplets are always symmetrical on both strands. Importantly, the crystal structures of both wild-type and A138T specific complexes show no contacts beyond 12 base-pairs³⁴ and we have shown that the identity of the fourth base-pair flanking the recognition site has no significant effect on binding (Kurpiewski, M.R. and Jen-Jacobson, L., unpublished results). Despite the dramatic effect of flanking trinucleotide on binding affinity, we show here (Table 1) that flanking context of a GAATTC site does not affect the first-order cleavage rate constants for either enzyme. It is useful to express these data in terms of the transition state interaction free energy change¹ $\Delta G^\ddagger_1$, which determines the probability of going from the free macromolecules to the activated complex where the first DNA strand is cleaved. This parameter takes into account both equilibrium DNA binding and the activation energy for the cleavage step (eq. 1):

$$\Delta G^\ddagger_1 = \Delta G^\circ_{\text{bind}} + \Delta G^{\ddagger\circ} \quad (1)$$

Discrimination between two DNA substrates in the transition state for catalysis is expressed by equation 2:

$$\Delta \Delta G^\ddagger_1 = \Delta \Delta G^\circ_{\text{bind}} + \Delta \Delta G^{\ddagger\circ} \quad (2)$$

where

$$\Delta \Delta G^{\ddagger\circ} = -RT \ln (k_{1,\text{cleave},\text{mod}}/k_{1,\text{cleave},\text{ref}}) \quad (3)$$

and the subscripts *mod* and *ref* refer to a modified DNA substrate and a reference substrate, respectively. (We use for comparison k_1 , the cleavage rate constant for the canonical GAA half-site, to register changes over the whole complex rather than local effects of a variant base-pair¹; see below.) In Figure 1 we plot $\Delta \Delta G^\circ_{\text{bind}}$ vs. $\Delta \Delta G^\ddagger_1$ for the wild-type and A138T enzymes and a series of DNA oligonucleotides containing the specific GAATTC site in different flanking contexts. The data (Table 1 and Figure 1) for both enzymes lie on the lines of unit slope, indicating that $\Delta \Delta G^\circ_{\text{bind}} \approx \Delta \Delta G^\ddagger_1$ and therefore $\Delta \Delta G^{\ddagger\circ} \approx 0$ (eq. 2). This demonstrates that altering the flanking context in a specific protein-DNA complex affects approximately equally the free energies of the ground state and the transition state for DNA cleavage. The simplest explanation is that the ground-state and transition-state complexes resemble each other so closely that context affects them equally²; this does not preclude minor differences at the catalytic center.

This is consistent with other data on the wild-type enzyme, showing that DNA context around the GAATTC site does not affect protein-base or protein phosphate contacts¹⁷, or the stoichiometry of cation release² or cosolute effects^{35; 36} upon binding. The strong similarity of the ground and transition state structures was pointed out previously by Jen-Jacobson² for the wild-type complex; Figure 1b implies that this also applies to the promiscuous mutant A138TGAATTC complex.

For the AAATTC EcoRI* site, a change in the flanking context also elicits a large range in $\Delta \Delta G^\ddagger_1$ penalties, spanning 3.6 kcal/mol for the wild-type enzyme and 4.6 kcal/mol for A138T (Table 2 and Figure 1). However, in contrast to the specific GAATTC complex, differences in $\Delta \Delta G^\ddagger_1$ for EcoRI* flanking variants are dominated by the activation free energy term, $\Delta \Delta G^{\ddagger\circ}$ rather than by differences in the binding free energy, $\Delta \Delta G^\circ_{\text{bind}}$ (Figure 1). This suggests that binding free energy is not sufficient to drive these variant DNA complexes into a pre-transition state that is poised for cleavage. Rather, binding free energy is minimized in each case by adaptation into structures that are off the path to the transition state, thus preferentially affecting the activation free energy barrier $\Delta \Delta G^{\ddagger\circ}$. These data

support the idea that both the wild-type and A138T enzymes bind to EcoRI* sites by undergoing structural adaptations¹. However, the extents of adaptation are not the same, as we show below.

Notably, the interaction of the A138T protein with a subset of EcoRI* context variants has unexpected characteristics more similar to those of non-adapted “specific” complexes. First compare plots (Figure 1) of $\Delta\Delta G^{\circ}_{\text{bind}}$ vs. $\Delta\Delta G^{\circ}_{\text{I}^{\ddagger}}$ for interaction of AAATTC EcoRI* sites with wild-type and A138T proteins. The wild-type data cluster within a narrow range of ~ 0.4 kcal/mol, indicative of a common adaptation, whereas the data for A138T range over 1.6 kcal/mol in the $\Delta\Delta G^{\circ}_{\text{bind}}$ dimension. That is, in binding the miscognate AAATTC site, the A138T protein is more sensitive than wild type protein to the identity of the flanking nucleotides. Second, data points for A138T for three of the six flanking variants tested (CGC, GGT, GAT) lie just above the line of unit slope (arrows, Figure 1b), implying that for the A138T protein and these sites, the context flanking an AAATTC site affects binding affinity, but not DNA cleavage rate ($\Delta\Delta G^{\circ}_{\text{I}^{\ddagger}} \approx 0$). This resembles the behavior of the “specific” complexes. It is noteworthy that the EcoRI* sites that are preferred by the A138T enzyme are all abutted by an R₋₅Y₋₄ sequence motif. However, the GGC flank is not strongly preferred, suggesting that the R₋₅Y₋₄ motif may be necessary but not sufficient.

Global features of a subset of A138T-EcoRI* complexes resemble specific complexes

Cosolute effects on DNA binding—The effect of cosolutes (osmolytes) on binding provides another empirical probe to assess whether A138T EcoRI* complexes bear closer resemblance to the specific or adaptive (EcoRI*) binding modes. Although the detailed interpretation of cosolute effects has gone through a long and complex evolution encompassing a variety of analytical approaches³⁷⁻⁵⁷, we adopt the statistical-mechanical treatment of Schurr et al.⁵¹, based on the Kirkwood-Buff integrals^{58; 59}. In this view, the cosolute effects arise in part because cosolute molecules increase the mutually excluded volume of macromolecules. The statistical-mechanical treatments⁵¹⁻⁵⁷ also take into account the possibility that a given cosolute is found at the macromolecular surfaces at an abundance greater than (binding) or less than (exclusion) its concentration in the bulk solution. Based on the Kirkwood-Buff treatment of cosolute effects⁵¹, K_A for the protein-DNA association reaction depends on water activity a_w as follows^{35, 51}: $(\ln K_A / \ln a_w)_{T,P} = \Delta X + \Delta I$, where ΔX is a steric exclusion term that measures the change upon protein-DNA association in sterically excluded macromolecular volume in the presence of cosolute, minus the volume excluded to solvent, divided by the partial molar volume of water $\bar{V}_w$; that is, $\Delta X = \Delta\Delta V^{\text{ac}} / \bar{V}_w$ where ΔV^{ac} is the difference in volumes excluded to cosolute and to water and $\Delta\Delta V^{\text{ac}}$ is the change upon binding. ΔI (designated as ΔS by Schurr et al.⁵¹) is a term that measures the change upon macromolecular association in interaction (which may be binding or exclusion) of the cosolute with the macromolecules. For PEG200 and the EcoRI-DNA association reaction we have shown that the interaction term ΔI is small, accounting for ~ 1 to 2% of the cosolute effect³⁵. [This is also true for BamHI-DNA and EcoRV-DNA complex formation³⁵]. Thus, $(\ln K_A / \ln a_w)_{T,P} \approx \Delta X$ (after correction for sodium activity) gives the change in the excluded macromolecular volume attributable to the presence of cosolute.

In this paper, we use cosolute dependence and the parameter $(\ln K_A / \ln a_w)_{T,P} \approx \Delta X$ as an empirical index to distinguish distinct forms of protein-DNA complexes that are characterized by different changes in excluded volume. This use does not depend on the exact molecular interpretation of ΔX in terms of changes in solvation (see Discussion).

Previous studies in our lab showed that the effect of cosolute on specific EcoRI-DNA binding was much more pronounced than for nonspecific binding, from which it was

inferred that the specific protein-DNA complex is more compact than the nonspecific complex^{35; 36}. The dependence for wild-type EcoRI binding to EcoRI* DNA was intermediate between the specific and nonspecific extremes. Because the data above suggested that a subset of A138T-EcoRI* complexes more closely resemble the wild-type specific complex than wild-type EcoRI* complexes, we therefore asked if the same is true with respect to cosolute dependence of binding.

We measured the dependence of protein-DNA binding on water activity (manipulated by changing the concentration of cosolute PEG200) for a specific site (GAATTC), a nonspecific site (CTTAAG), and two RY-flanked EcoRI* sites that may form 'specific-like' complexes with the A138T enzyme. For the specific GAATTC site, the $(\ln K_A / \ln a_w)_{T,P} \approx \Delta X$ values (Figure 2a; Table 3) are the same, within experimental error, for the wild-type and A138T enzymes, consistent with the fact that the crystal structures of the two complexes are nearly identical³⁴. For both enzymes, the PEG200 dependence of binding to the nonspecific site CTTAAG is much shallower than for specific binding (Figures 2a and 2b; Table 3) and the same within error, showing that the nonspecific complexes are significantly less compacted and/or retain significantly more interfacial water than their specific counterparts. The calculations (Table 3) show that nonspecific binding involves a change in excluded volume less than half of that found for specific binding.

There are striking differences between the wild-type and the A138T enzymes in the cosolute dependence of binding to the AAATTC EcoRI* site. For the wild-type enzyme, the slope $(\ln K_A / \ln a_w)_{T,P}$ (Figure 2a) and thus the change in excluded volume ($\Delta X = 400$ for both GGT-flanked and GAT-flanked EcoRI*; Table 3) are intermediate between the values for specific and nonspecific complexes. By contrast, the change in excluded volume upon A138T binding to EcoRI* sites ($\Delta X = -520$ for both GGT-flanked and GAT-flanked EcoRI*) is the same, within error, as for mutant protein binding to the specific GAATTC site (Figure 2b and Table 3). That is, these A138T-EcoRI* complexes resemble the specific complex more than they do wild-type EcoRI* complexes.

Cosolute effects on DNA cleavage—A hallmark of the wild-type specific complex is its high structural similarity to the transition state. Consequently, the first-order rate constant for the catalytic (chemical) step is insensitive to cosolute³⁶, suggesting that no additional volume change occurs (and probably no additional desolvation occurs) in going from the metal-free recognition complex to the catalytic transition state. By contrast, for EcoRI* sites the first order rate of DNA cleavage is strongly influenced by cosolute. One interpretation is that the ground states of wild-type-EcoRI* complexes require further compaction and water release in order to become competent for cleavage³⁶, but there are other possibilities (see Discussion).

We thus also used cosolute effects to further characterize catalysis by the A138T-EcoRI* complexes. The PEG200 dependence of EcoRI* DNA cleavage by the wild-type enzyme (Figure 2c) suggests a compaction equivalent to $\Delta X = -100$ to -170 in going from the ground state to the transition state (Table 4); this value is slightly dependent on the flanking context and the DNA strand being cleaved. (Note that the scale of Figure 2c is 5 times smaller than the scale of Figure 2a.) However, cleavage of RY-flanked EcoRI* sites by the A138T enzyme exhibits much less dependence on cosolute concentration (Figure 2d), implying that after a compaction equivalent to $\Delta X \approx -500$ in binding (Table 3), little further compaction occurs in going from the metal-free ground state to the transition state (Table 4). On the other hand, when the EcoRI* site is flanked by GCA, cleavage by A138T is more sensitive to cosolute, suggesting additional compaction of about $\Delta X = -78$ for cleavage in the first (normal GAA) half-site and $\Delta X = -94$ for cleavage in the second (modified AAA) half-site (Table 4). Thus, this criterion also supports the thesis that A138T complexes with

5'-RY flanked EcoRI* sites resemble the specific complex, whereas the complex with the other EcoRI* site more closely resembles the “adaptive” EcoRI* binding mode used by the wild-type enzyme.

Change in heat capacity—The standard heat capacity change ΔC_p° upon protein-DNA binding is a useful metric for distinguishing specific from nonspecific protein-DNA binding^{14; 17; 60-63}. Specific binding is accompanied by a strongly negative ΔC_p° , derived from contributions that reflect the strong complementarity of the protein-DNA interface, including desolvation of non-polar surface, and restriction of vibrational freedom of macromolecules⁶⁴ and bound interfacial water molecules^{60; 62}. Examination of the crystal structures of the free and bound macromolecules shows that $\sim 2000 \text{ \AA}^2$ of nonpolar surface is buried in the specific complex³⁵. Semi-empirical formulas^{65; 66} based on solvent accessible surface area predict that water release accounts for less than $\sim 25\%$ of the observed ΔC_p° ⁶⁰. Much of the remainder comes from the restriction of vibrational degrees of freedom accompanying the formation of multiple polar and nonpolar protein-DNA contacts as well as the coupled folding of the protein arms^{17; 60; 67}.

By contrast, the nonspecific complex is much looser, and its formation is characterized by $\Delta C_p^\circ \approx 0$ ^{14; 16; 17; 60-63; 68}. We previously showed that the strongly negative heat capacity changes ΔC_p° were approximately the same for the wild-type and A138T enzymes when binding to the cognate GAATTC site³⁴. This is consistent with the similarity of these two specific complexes in x-ray crystal structures³⁴, implying that the complementarity of the two protein-DNA interfaces is very similar.

It is clear, however, that A138T-EcoRI* complexes differ from wild-type-EcoRI* complexes in enhanced DNA binding affinity (Table 2), higher cleavage rates (Table 2) and greater change in excluded volume (Table 3), so we asked if they also differ in ΔC_p° values. We performed van't Hoff analyses by measuring the temperature dependence of K_A for wild-type binding to the AAATTC EcoRI* site embedded in three different flanking contexts (CGC, GCA, and GAT), A138T binding to the AAATTC EcoRI* site embedded in four different flanking contexts (CGC, GCA, GGT, and GAT) and both enzymes binding to a single nonspecific (CTTAAG) DNA construct, to compare with data for binding to the GAATTC site. These binding data are shown in Figure 3. Since ΔC_p° is the second derivative of ΔG° with respect to $1/T$, curvature of the $\ln K_A$ versus $1/T$ plots is related to the magnitude of ΔC_p° . Thermodynamic parameters ΔH° , ΔS° and ΔC_p° were extracted by using two different methods to fit the temperature-dependence data (see Materials and Methods and legend to Table 5).

We focus on ΔC_p° , because interpretation of ΔH° and ΔS° in terms of their structural correlates is hazardous, since we do not have high resolution structures of the miscognate complexes, and the net observable ΔH° and ΔS° are the resultants of large and opposing contributions. Wild-type EcoRI shows a strongly negative ΔC_p° for cognate DNA binding (Figure 3a), a near-zero value for nonspecific binding (Figure 3f; note absence of curvature) and intermediate values for binding to EcoRI* sites (Figure 3b,c,e). These values are given in Table 5. Similar observations have been made for BamHI¹⁶ and EcoRV (Engler, L. E. & Jen-Jacobson, L., unpublished results) restriction endonucleases.

We further found that ΔC_p° is roughly constant ($\sim -0.7 \text{ kcal/mol}\cdot\text{K}$) for wild-type enzyme binding to the AAATTC EcoRI* site in all three flanking contexts tested (Table 5 and Figure 3b,c,e), implying that flanking context has little influence on the gross complementarity of the wild-type-EcoRI* interface. For binding of A138T endonuclease to nonspecific DNA (Table 5 and Figure 3e), we find $\Delta C_p^\circ \approx 0$, consistent with our previous report that the A138T mutation does not affect the nonspecific DNA binding mode³.

However, the ΔC_p° for A138T binding to EcoRI* sites was strikingly influenced by the flanking context. The data for A138T binding to sites flanked by a 5'-RY dinucleotide (CGC, GGT, or GAT) show much greater curvature (Figure 3,b,c,d) than does wild-type binding to EcoRI* sites and a similar degree of curvature as A138T binding to the specific DNA site (Figure 3a, open circles). This corresponds to a much more negative heat capacity change ($\Delta C_p^\circ = -1.2$ to -1.5 kcal/mol•K) than is the case for wild-type-EcoRI* binding ($\Delta C_p^\circ = -0.7$ kcal/mol•K) (Table 5 and Figure 3). Thus, ΔC_p° for A138T binding to 5'-RY-flanked EcoRI* sites is more similar to the ΔC_p° for wild-type specific binding than to that for wild-type EcoRI* binding (Table 5 and Figure 3). Some of the additional increment of negative ΔC_p° is likely a consequence of greater surface burial during binding (see Discussion), but this is probably a minor factor, since desolvation of polar and nonpolar surfaces has opposite effects on ΔC_p° and the *total* desolvation contribution to ΔC_p° for wild-type-GAATTC binding is <25%^{17; 60}. By contrast, the temperature dependence of A138T binding to GCA-EcoRI* DNA (Figure 3e) exhibits significantly less curvature, more similar to that of the wild-type, and ΔC_p° for binding to GCAAAATTC is the same for both enzymes (Table 5).

It is important to note that the three EcoRI* sites with the large and negative ΔC_p° upon A138T binding are also those most rapidly cleaved by the A138T enzyme (Table 2). This buttresses the existing evidence^{16; 17; 60} that a large negative ΔC_p° is one diagnostic of a ground state complex that is poised to reach the transition state. (This may be a necessary property of the “specific” binding mode; we do not know if it alone is a sufficient diagnostic.) These observations support the inference that A138T carries out erroneous and potentially lethal cleavage at 5'-RY-flanked EcoRI* sites because it can form a specific-like complex rather than a structurally distinct adaptive complex.

Protein-phosphate contacts in wild-type and mutant EcoRI* complexes

EcoRI* complexes have so far resisted crystallization, so we sought another way to explore their molecular features. When wild-type EcoRI endonuclease forms adaptive complexes at EcoRI* sites, there is a pronounced shift in the ethylation-interference footprint, indicating adaptive changes in protein-phosphate contacts, relative to those in the specific complex¹. The specific footprint shows a critical role in binding (i.e., strong interference) for the symmetrically disposed phosphate contacts at pXpGAApTTC^{1; 2}. In the crystal structure of the specific complex²⁴, these phosphate contacts are extensively networked to key base recognition elements to “clamp” the latter in place, supporting their functional importance in binding. By contrast, the wild-type-EcoRI* adaptive complex shows a footprint that is weaker and markedly asymmetric¹, such that the slowly cleaved canonical GAA half-site shows a stronger footprint than does the even more slowly cleaved non-canonical half-site, where some key phosphate contacts appear to be missing¹. We thus asked whether RY-flanked A138T-EcoRI* complexes, which resemble specific complexes in cleavage rates, ΔC_p° and cosolute effects, also have footprints more like those of specific complexes.

We assessed the ethylation interference footprints made by the wild-type and A138T enzymes on the AAATTC site in three flanking contexts: two (GAT, Figure 4a,c and GGT, Figure 4b) that exhibited specific-like features with respect to multiple probes and one (GCA, Figure 4d) for which the ΔC_p° of binding to AAATTC is the same for the mutant and wild-type enzymes. The GGTAAATTC site is cleaved by A138T 140-fold faster than by wild-type (Table 2) and only 14-fold slower than the specific site with the same flanking sequence. We thus predicted that the A138T-GGTAAATTC binding ensemble is significantly populated by productive complexes, and would thus have a ‘specific-like’ character of the phosphate contacts. Indeed, the footprint of A138T on the cognate half-site of GGTGAATTT is significantly stronger than that of wild-type protein (Figure 4c, upper).

This stronger footprint is even more pronounced for the GAT-flanked site (Figures 4a and 4c, upper), which is cleaved 170-fold faster by A138T than by wild-type. Furthermore, in the non-canonical AAA half-site (Figures 4b and 4c, lower panels) the A138T footprints are even more powerfully enhanced relative to those of wild-type protein. This is most evident at the 5'-flanking phosphates pXpAAATTC. Thus, the A138T enzyme is able to make stronger and more symmetric phosphate contacts to both strands of the asymmetric EcoRI* site AAATTC when it is flanked by an RY dinucleotide. These data strongly suggest that the A138T-RYAAATTC complex has more structural resemblance to the specific wild-type-GAATTC complex than it does to the adaptive wild-type-AAATTC complex.

By contrast to RY-flanked sites, the A138T mutation does not affect footprints on the DNA strand containing the cognate GAA half-site of the pyrimidine-purine (YR)-flanked GCAAAATTC (Figure 4d, upper) and there are only small increases in the strength of the footprint on the non-cognate half-site (Figure 4d, lower). These observations are consistent with the fact that while the A138T enzyme cleaves this site 6-fold faster than the wild-type (data for the cognate half-site, Table 2), the cleavage rate is still >400-fold slower than that of the specific GAA site with the same flanking sequence (Table 2). Therefore, although the mutation may cause a shift in the A138T-GCAAAATTC binding ensemble towards cleavage competent forms, these are probably a small fraction of the total, so any differences from the wild-type complex would be difficult to detect in ensemble-averaged thermodynamic properties.

The spatial relationship between the mutation and the improved phosphate contacts is suggested by the crystal structure of the A138T-GAATTC complex. Residues Ala/Thr138, Ala139, and Ile197 are part of a hydrogen bond and packing interaction network (Figure 5a) that connects Gly196 and Arg203, the amino acids that contact phosphates at X₋₆pXp₋₅X₋₄. It has been noted previously³⁴ that Thr138 makes novel direct and water-mediated contacts to the flanking RY bases (Figures 5b,c); these protein-base contacts and protein-phosphate contacts, made by interacting sidechains (Figure 5), should form a mutually supporting network. It is also noteworthy that the I197A mutation is an intragenic suppressor of the A138T promiscuous phenotype²⁹. We conjecture that without the bulky Ile197 side chain, Thr138 may not be able to bridge Gly196 between and Arg203 (Figure 5a).

Discussion

The experiments in this paper show that the A138T mutant EcoRI endonuclease, unlike its wild-type counterpart, can interact with EcoRI* sites in something approaching the “specific” binding mode that the wild-type protein achieves uniquely at the canonical GAATTC site. Four sets of experimental criteria support this inference: First order cleavage-rate constants, the heat capacity change upon binding ΔC_p , the pattern of contacts with phosphates of the DNA backbone, and the inferred change (from cosolute effects) in excluded volume upon protein-DNA association. By all these criteria, the A138T-EcoRI* complex more nearly resembles the ‘specific’ complex at GAATTC sites than it does the ‘adaptive’ complex formed by wild type EcoRI at EcoRI* sites.

Interpretation of cosolute effects

Because the dependence on cosolute is the most unfamiliar criterion we use to distinguish among binding modes, we discuss here the conceptual underpinnings of this method. In the most general case of mixtures of macromolecules, water and cosolute, the effect of the cosolute on macromolecular association is the result of both (a) differences in cosolute abundance between bulk solution and the macromolecular surfaces and (b) changes in mutually excluded macromolecular volumes. Many early thermodynamic treatments ignored or seriously underestimated the excluded-volume effects, but it was later appreciated by

Schellman^{50; 52} and Schurr et al.⁵¹ and now is widely accepted⁵³⁻⁵⁷ that these could be at least as large as the term for interaction (positive or negative) of the cosolute with the macromolecular surfaces. The view of these phenomena that arises from the Kirkwood-Buff integrals has the advantages of exactness⁵⁵ and of providing a molecular picture of the phenomena (but see below). It has been shown^{52; 54} that development from the Kirkwood-Buff theory leads to equations of precisely the same forms as those arising from classical thermodynamic treatments (e.g., 'preferential interactions') that do not incorporate an associated molecular picture^{41;42}, and these equivalencies were further refined by Shulgin and Ruckenstein⁵⁴ and Smith and coworkers^{55; 57}. The equations become identical in the limit of a cosolute of the kind we use here.

Previous studies in our lab established that triethylene glycol (TEG) and polyethylene glycol 200 (PEG200) meet rigorous criteria of being neither bound to nor preferentially excluded from EcoRI and DNA surfaces^{35; 36}; that is, the cosolute is represented at the macromolecular surface(s) at the same abundance it has in the bulk solution. The detailed conceptual and experimental basis for our choice of PEG200 is given elsewhere^{35; 36}. For such a cosolute, there is no free energy contribution from the partitioning of cosolute between the macromolecular surfaces and bulk solution.

Such a choice of cosolute greatly simplifies interpretation because cosolute effects on macromolecular association arise almost entirely from changes in excluded volumes. The cosolute effects in such a case are most easily understood as follows: (1) Cosolute molecules at macromolecular surfaces are larger than water molecules, so they increase the mutually excluded volume of the macromolecules. This can be viewed as a reduction in entropy, producing an increase in the free-energy cost of mutual steric exclusion of the macromolecules. (2) The excluded volume of the protein-DNA complex is smaller than the sum of the excluded volumes of the free macromolecules, so the free-energy cost of steric exclusion in the presence of cosolute is less for the complex than for the free macromolecules. (3) As a result, the presence of the cosolute promotes intermolecular association.

Nevertheless, the quantitative interpretation is far from trivial, because the observed phenomena arise from changes in excluded volume, not from stoichiometric changes in solvent or cosolute molecules at the macromolecular surfaces. The parameter $(-\ln K_A / \ln a_w)_{T,P} = \Delta X$ has the units of numbers of water molecules, but does not measure exactly the stoichiometric number of water molecules released upon macromolecular association because of the difference in radius between a PEG200 ($r = 3.7 \text{ \AA}$) and a water molecule ($r = 1.4 \text{ \AA}$) (see detailed discussion³⁵). The simplest image may be that compaction upon macromolecular association is largely associated with desolvation of the interacting surfaces, but it is unclear, even if this image is valid, whether this represents the displacement of one hydration layer, multiple hydration layers, or a number of layers that varies according to local physicochemical properties of the surfaces. Furthermore, the magnitude of the effects depends on the effective radius of the cosolute^{35; 36; 51} (Jen-Jacobson, L., Cao, D.F. and Kurpiewski, M.R., unpublished results) and because cosolute molecules are inevitably larger than water molecules, it is physically impossible for the measured excluded-volume change to relate exactly to the number of water molecules released. These reservations notwithstanding, we have found^{35; 36} that the excluded-volume effects for PEG200 with EcoRI endonuclease (and EcoRV endonuclease) binding to cognate DNA are within 10% of what is computed (taking into account retained interfacial water seen in co-crystals) for the displacement of a single hydration layer at the interface between protein and DNA. It seems possible that effects from the average roughness of macromolecular surfaces partly compensate for the size difference between water and cosolute⁵⁴.

The reader should note that none of this conceptual underpinning or molecular interpretation is required to justify the use of cosolute effects as an empirical indicator of distinct modes of binding. Thus, the weaker cosolute effects on the formation of “adaptive” complexes of wild-type EcoRI at EcoRI* sites could reflect (a) A smaller degree of compaction in the “adaptive” complexes; (b) A higher retention of surface solvation in the “adaptive complexes”; (c) Both, in some measure. These ambiguities do not affect the conclusion: A138T-EcoRI* complexes more closely resemble specific complexes than they do “adaptive” wild-type-EcoRI* complexes.

Reaction rates and adaptive complexes

An EcoRI* site, e.g. AAATTC, has one normal GAA half-site and one modified AAA half-site. Protein-DNA interactions over the entire complex must contribute to the stability of ground-state and transition-state complexes, but we focus our analysis and this discussion on the cleavage rate constant in the GAA half-site (k_1), where local features of the transition state most closely resemble those of the specific GAATTC site complex.

Our data show that “adaptive” wild-type EcoRI* complexes have much reduced first-order cleavage rate constants (see k_1 , Table 2), whereas the measured rate constants for the A138T-EcoRI* complexes are substantially higher. This difference, and the difference in sensitivity of the rate constants to cosolute (Figure 2), can be given two rather distinct interpretations.

The first-order cleavage rate constant is a proportionality constant between measured rate of product formation and the concentration of enzyme-substrate complex:

$$\frac{d \text{ product}}{dt} = k_{\text{cleave}} [P - DNA] \quad (5)$$

We measure first-order rate constants with enzyme excess over DNA at high enough concentrations that all DNA is enzyme-bound. The conventional presumption in this situation is that $[P\text{-DNA}] = [\text{DNA}]_{\text{total}}$. If we suppose that the complex of wild-type endonuclease with an EcoRI* site consists of a paucidisperse population of “adaptive” complexes $P\text{-EcoRI}^*_{\text{adaptive}}$, and that these have substantial structural differences from the transition state, then we can attribute the observed lower k_{cleave} to the fact that the enzyme does not stabilize the transition-state complex with an EcoRI* site as well as it does the transition-state complex with the specific GAATTC site. That is, the two transition states differ in global structure and stability, with the EcoRI* transition state lying at a higher energy level. There is very little structural change, excluded-volume change or desolvation in the specific GAATTC complex in going from the ground state to the transition state, and thus no cosolute effect. By contrast, for the EcoRI* complex additional compaction is required to get to the transition state, and this is promoted by the cosolute because of excluded-volume effects. According to absolute reaction-rate theory⁶⁹, we can treat the measured $k_{\text{cleave}}(\text{app})$ as representing a quasi-equilibrium constant for the conversion of the ground state (metal-free enzyme-DNA complex) to the catalytic transition state (including Mg^{2+}). We can then interpret the cosolute effects as we would for a true equilibrium binding constant.

The alternative view is that the “adaptive” protein-DNA complex is intrinsically incapable of progressing directly to the transition state, but must first be converted to an “active” form that closely resembles the “specific” ground-state complex:

$$\frac{d \text{ product}}{dt} = k_{\text{cleave}} [P - \text{EcoRI}^*_{\text{active}}] \quad (7)$$

For wild-type endonuclease the equilibrium for EcoRI* complexes lies far toward P-EcoRI*_{adaptive}, such that the amount of the active conformer P-EcoRI*_{active} is small. This causes the measured value of $k_{\text{cleave}}(\text{app})$ to appear small, even though its “intrinsic” value (per molecule of P-EcoRI*_{active}) may be the same as for a GAATTC site.

This implies that the transition states for the GAATTC reaction and the EcoRI* reaction, both of which accomplish the same hydrolytic chemistry at the same GpA phosphodiester bond, may differ little in global structure and stability and lie at very nearly the same energy levels. In this model, cosolute affects only the equilibrium between active and inactive conformers of ground-state protein-DNA complex, promoting the active conformer because it is more compact. Cosolute would thus raise the measured $k_{\text{cleave}}(\text{app})$ even if it had no effect on the “intrinsic” rate constant k_{cleave} . This model thus avoids conjecture about cosolute effects on a catalytic rate constant or on a transition-state structure that cannot be studied directly.

In this model, the A138T mutation shifts the conformer equilibrium toward P-EcoRI*_{active}, thus increasing $k_{\text{cleave}}(\text{app})$. Note that the ground-state ethylation-interference footprint data (Figure 4) and the heat capacity data (Table 5, Figure 3) also represent ensemble-averages, and could be interpreted according to this view.

Either of these models is consistent with essentially all our observations on cleavage of EcoRI* sites by wild-type and A138T endonucleases. One intriguing aspect of the data is that the cosolute effects on binding and cleavage of EcoRI* sites indicate an overall degree of excluded-volume change that is quite similar to that for the GAATTC site. For example, for wild-type enzyme at GAATTC, regardless of flanking context, $\Delta X_{\text{bind}} = -540$ at the binding step and $\Delta X_{\text{cleave}} \approx 0$ at the cleavage step. For wild-type enzyme and EcoRI* sites, the GGT flank shows $\Delta X_{\text{bind}} (-400) + \Delta X_{\text{cleave}} (-160) = -560$; for the GAT flank $\Delta X_{\text{bind}} (-400) + \Delta X_{\text{cleave}} (-120) = -520$. For A138T-EcoRI* $\Delta X_{\text{bind}} (-520) + \Delta X_{\text{cleave}} (-30) = -550$.

The sum $\Delta X_{\text{bind}} + \Delta X_{\text{cleave}}$ is a measure of the overall change in excluded volume in going from the free protein and DNA molecules to the transition state for cleavage in a GAA half-site. The approximate constancy of this sum therefore suggests that the overall degree of compaction is similar for the transition state complexes with GAATTC or with EcoRI*. The detailed interpretation differs, however, according to the two models given above. In the first model (paucidisperse P-EcoRI*_{adaptive} complex), ΔX_{bind} measures the compaction in forming the P-EcoRI*_{adaptive} complex, and ΔX_{cleave} measures the additional compaction required to reach the transition state. In the second model (ground-state conformer equilibrium), ΔX_{bind} measures the average compaction in forming the equilibrium mixture of P-EcoRI*_{adaptive} and P-EcoRI*_{active}, whereas ΔX_{cleave} measures the average compaction involved in converting P-EcoRI*_{adaptive} to P-EcoRI*_{active}. We see no present basis for distinguishing between these two models.

How specificity goes awry

The fundamental rationale for studying a restriction endonuclease was to learn how a protein achieves extremely high specificity for a DNA site. EcoRI endonuclease has evolved to make a set of highly discriminatory interactions with a particular DNA base sequence, exploiting both detailed features of the molecular surfaces and conformational predispositions^{1, 2}. The number of these interactions, and the networking among them, is such that recognition of the canonical GAATTC site is over-specified⁷⁰. If wild-type EcoRI endonuclease were to attempt to form a “specific” complex with a closely related but incorrect DNA site (an EcoRI* site), there would be both a mismatch of the hydrogen-bonding groups (notably at the incorrect base-pair) and an increased cost of deforming the DNA into the ‘kinked’ conformation that produces maximum protein-DNA complementarity. The energetic penalties in this situation are too large to overcome, so the protein instead minimizes binding free energy by forming a structurally distinct ‘adaptive’ complex^{1, 2}. This EcoRI* complex is not only off the direct path to the catalytic transition state (consequently cleaves EcoRI* sites very slowly), but also differs from the specific complex in less mutual complementarity of the protein-DNA surfaces, reflected in a less negative ΔC_p (Table 5) and less decrease in excluded volume (Table 3). Furthermore, some phosphates that made crucial contacts in the specific complex may show weaker or no ethylation interference [¹; Figure 4], indicating that the protein can no longer make close contacts to those phosphates in the adaptive complexes.

In this and earlier papers^{3, 34} on ‘promiscuous’ mutants of EcoRI endonuclease, we have been seeking to learn how a single amino acid change can cause a pronounced (indeed, lethal) relaxation of the normally stringent specificity. The naive expectation was that relaxation of stringent specificity would entail the loss of some subset of the required protein-base recognition contacts. That expectation turned out to be incorrect.

Our crystallographic and computational studies of the A138T-GAATTC complex³⁴ showed that the A138T mutant protein unexpectedly gained novel interactions with DNA, including some to the flanking bases on the 5’-side of the 6 base-pair recognition motif. Because asymmetric EcoRI* complexes have proven recalcitrant to crystallization, we can only surmise how the mutant protein might interact with EcoRI* sites by extrapolation from the structure of the A138T-GAATTC complex³⁴. Such extrapolation is given some justification by our experimental demonstration here that the A138T-EcoRI* complex more closely resembles the ‘specific’ complex by several criteria, including catalytic rates, the pattern of protein-phosphate contacts and the ΔC_p and volume compaction that attend complex formation. The preference for a flanking R₋₅Y₋₄ motif suggests that A138T-EcoRI* complexes might contain some of the novel contacts to flanking bases that were observed in the A138T-GAATTC complex³⁴. It is noteworthy that the three EcoRI* sites that are preferentially cleaved by the promiscuous mutants *in vivo*³³ are all flanked by a T₋₄.

However, it is also notable that A138T ‘specific-like’ binding to EcoRI* sites most frequently shows less favorable ΔH° , but more favorable ΔS° than does ‘adaptive’ binding of wild-type protein to EcoRI* sites (Table 5). If the A138T mutation simply augmented the wild-type interface with novel hydrogen-bond or van der Waals contacts, this in isolation should have resulted in improved ΔH° , contrary to our observation. In forming an intricate protein-DNA interface whose formation is also strongly influenced by release of bulk water and retention of specific water molecules, many factors affect ΔH° and ΔS° and these often have compensating effects⁶⁰. It should also be emphasized that differences in parameters ΔH° and ΔS° for protein-DNA binding can originate from differences in the complexes, in the free protein, or both. It is thus extremely hazardous to associate changes in net observable ΔH° or ΔS° with particular “structural” differences in a complex. These considerations make

it unlikely that the loss of specificity (increased ability to cleave EcoRI* sites) can be easily understood from novel recognition contacts alone.

An integrative account of promiscuous specificity

We can provide the following general explanation of why this mutant enzyme shows 'promiscuous' behavior (less stringent specificity):

1. We propose that the A138T protein makes novel contacts to flanking RY bases of EcoRI* sites, as it does to GAATTC sites (Figures 5b and 5c). These provide a relatively small increment of binding free energy.
2. This increment taken alone would likely be insufficient to overcome the cost of distorting an AAATTC site, but there is also a contribution from the better phosphate contacts made by A138T. The additional protein-phosphate contacts, particularly in the flanking regions at pXpAAATTC, not only provide more free energy to drive the DNA distortion, but also help position protein sidechains precisely for both recognition and catalysis^{1; 71}. Whether the DNA distortion permits these phosphate contacts to be formed, or the phosphate contacts help stabilize the DNA distortion, is a merely semantic distinction.
3. When a complex more nearly approximates the "specific" binding mode, a more intimate protein-DNA interface is formed. This is attended by greater reduction in excluded volume and likely greater desolvation of the interfacial regions (Table 3), contributing to more favorable ΔS° and $\Delta G^\circ_{\text{bind}}$ ('hydrophobic effect'). On the other hand, the more intimate intermolecular interface likely provides a partially compensating unfavorable contribution to ΔS° via decreased vibrational freedom.
4. The observed effects on thermodynamic parameters ΔH° , ΔS° and ΔG° are the nets of all these factors. Both greater solvent release and greater vibrational restriction contribute to a more negative ΔC_p , giving A138T-EcoRI* binding a value approaching those of the wild-type-GAATTC and A138T-GAATTC binding processes (Table 5).
5. The non-adapted, "specific-like" interface of the A138T-EcoRI* complex is not identical to that of the 'specific' complex with GAATTC, because cleavage rate constants are still at least 10-fold lower (Table 2). Nevertheless, these rates are 100-200-fold higher than the wild-type enzyme can achieve at the same EcoRI* sites (Table 2), suggesting that the non-adapted A138T-EcoRI* complex lies more nearly along the path to the catalytic transition state. That is, the A138TEcoRI* complex can utilize binding energy toward the transition state, albeit not quite as efficiently as the GAATTC complex.

Conclusions

Some of the simple structural ideas about specificity determination in protein-DNA interactions underrate the complexity of high-stringency recognition systems. As much as high-resolution crystal structures have helped us understand recognition interfaces, they promote an unduly enthalpy-centric view of these interfaces, and tempt us to focus exclusively on the complexes while ignoring both the free macromolecules and the dominant role of water in the energetics. The case of EcoRI endonuclease and its promiscuous mutants shows that we cannot think of specificity change in terms of addition or subtraction of contact points in a recognition complex, but must anticipate the possibility of significant structural and thermodynamic differences between complexes. There is reason for hope that spectroscopic methods may be an efficient approach to elucidating the

differences among the various adaptive states, as well as the roles in specificity for dynamic motions of protein and DNA^{25; 72-76}.

Materials and Methods

Enzyme expression and purification

Wild-type and A138T EcoRI endonucleases were expressed and purified as described³. Enzyme preparations were greater than 99% pure as analyzed on Coomassie Blue stained SDS gels. Enzyme concentrations were determined by comparing intensities of bands on SDS gels with a standard curve of intensities derived from an EcoRI endonuclease preparation whose concentration was determined using direct amino acid analysis⁷⁷.

Oligonucleotide substrates—Oligodeoxynucleotides were synthesized on an Expedite 8909 synthesizer at the University of Pittsburgh DNA Synthesis Facility and purified using denaturing PAGE. Concentrations of single stranded oligonucleotides were determined by UV spectroscopy using extinction coefficients calculated from the nearest neighbor method⁷⁸ as elaborated by Senior et al⁷⁹. Duplex DNAs were formed by mixing equimolar amounts of single stranded species and annealing by slow cooling in a water bath. Concentrations of duplex DNAs were confirmed by ethidium visualization of native gels loaded with a concentration standard of the same length. Concentrations determined from the UV and PAGE methods were within 5%. Oligonucleotides were 5'-end labeled with ³²P using T4 polynucleotide kinase as described⁸⁰. The particular oligonucleotide sequences used for various kinds of experiments are specified in the legends to Figures and Tables. The lengths were varied from 13 to 24-bp to accommodate the needs of each method; however, sequences beyond the flanking trinucleotides have no effect on binding or cleavage rates.

Equilibrium binding and cleavage kinetics

Equilibrium association constants (K_A) for specific DNA sites were determined by either direct binding (specific sites) or competition methods (EcoRI* and nonspecific sites) on nitrocellulose filters as described^{15; 80}. Direct equilibrium binding titrations included 10 different enzyme concentrations spanning a two order of magnitude range. Enzyme and DNA concentrations were chosen such that $[E]_{\text{total}} \approx [E]_{\text{free}}$ and to ensure that plateaus, representing conditions for which 100% of the DNA is bound, were well defined in each isotherm. Equilibrium competition binding isotherms included eight different competitor DNA concentrations with a concentration range of up to 300-fold. Means and standard deviations are derived from replicate (2) determinations of K_A . Reactions were carried out at 22°C (unless otherwise noted) in binding buffer (10 mM bis-Tris-propane, 100 μ M dithiothreitol, 1 mM EDTA, 100 μ g/ml bovine serum albumin 0.01% NaN₃, pH 7.3) plus 0.22 M KCl unless indicated otherwise.

First order rates of DNA cleavage were measured in single turnover experiments as described³. Solution conditions for the cleavage experiments were 1 μ M EcoRI, 0.5 μ M DNA, 10 mM BTP, 0.1 M NaCl, 100 μ M DTT, 100 μ g/ml BSA, 8 mM MgCl₂, pH 7.3, 22°C. At least 10 time points were included in each rate determination. Means and standard deviations are derived from replicate (2) determinations of k_{cleave} . Note that the parallel sequential model for double stranded DNA cleavage yields four rate constants¹. In this work, we focus on k_1 , the rate for the first single stranded nicking event, which for EcoRI* sites, always occurs in the cognate, GAA half site¹.

For van't Hoff analyses, the temperature dependence of the equilibrium association constant, K_A was measured over a temperature range of 1 $\pm$ 0.2 to 40 $\pm$ 0.2° C in an environmentally controlled room. These binding reactions were performed in 10 mM cacodylate, 0.21 M

KCl, 1 mM EDTA, 100 μ M DTT, 100 μ g/ml BSA and 0.01% NaN_3 (pH 7.3). We used cacodylate buffer, which has a very small heat of ionization⁸¹ (-0.47 kcal/mol) to minimize contributions from coupled equilibria to ΔH° and ΔC°_p upon protein-DNA binding⁸²⁻⁸⁴. Cacodylate also has the advantage that the total number of protons and the pH of the solutions were constant over the temperature range, so that no pH adjustment is required at each temperature. At least three determinations of K_A were made at each temperature. Weak binding of EcoRI* sites requires the equilibrium competition method in which increasing concentrations of an unlabeled noncognate DNA competes against a fixed amount of radiolabeled specific DNA for enzyme binding. In order to obtain accurate binding constants, the K_A of the radiolabeled specific DNA must be known. Thus as a pre-requisite to EcoRI* van't Hoff studies, we measured the temperature dependence of binding of two different specific probes (GACGAATTC, and GGCGAATTC). In the specific binding assay where $[E] \gg [DNA]$, enzyme instability was an issue at temperature extremes (above 30°C and below 0°C). To obtain accurate binding constants at the extremes, we performed “converse” titrations in which increasing amounts of radiolabeled DNA is titrated against a fixed enzyme concentration ($<10\%$ of the K_D) while varying the radiolabeled DNA concentration. In this scenario, excess DNA protects the enzyme. Enzyme stability under these conditions was confirmed by incubating the binding reactions at the temperature extremes, then transferring to 21°C and filtering. The binding constants determined from these reactions were the same as those determined from reactions incubated at 21°C only. Instability is not a problem in equilibrium competition assays because a high concentration of the specific, radiolabeled probe (6-40-fold over the K_D) was used to keep the enzyme active at elevated temperatures. This was confirmed by incubating the binding reactions at the temperature extrema, then transferring back to 21°C and filtering. The binding constants determined from these reactions were the same as those determined from reactions incubated at 21°C only. Because the K_D of the specific probe and its concentration in the reaction are parameters in fitting competition curves, we performed two additional controls to demonstrate the robustness of the method. We show in Figure S2 that fitted competitor DNA K_A is independent of both the identity and concentration of the specific probe.

Determination of thermodynamic parameters

ΔC°_p , ΔH° , and ΔS° values were calculated from the experimentally determined values of K_A (and $\Delta G^\circ_{\text{bind}}$) as a function of temperature, using two methods, as described in detail previously³⁴. The first method, which assumes that ΔC°_p is constant over the experimental temperature range, uses a multiparametric-fit to an integrated form of the van't Hoff equation^{85,86,34}. Alternatively, the method of Clarke and Glew^{87, 88} assumes only that the equilibrium association constant and the thermodynamic parameters are continuous functions of temperature. Because the Clarke and Glew analysis does not assume a constant ΔC°_p , any change in ΔC°_p with temperature can be evaluated. As shown here and in our previous work³⁴ there is very good agreement between the two methods. It is also important to note that we showed with carefully controlled conditions, that the van't Hoff method and isothermal titration calorimetry yield values of ΔH° and its temperature dependence ΔC°_p in very close agreement^{60; 84}. We detect no temperature-dependence of ΔC°_p in the range examined.

Cosolute effects on DNA binding

The conceptual framework for using the effect of increasing concentrations of polyethylene glycol 200 (PEG200) on EcoRI-DNA complex formation to obtain an experimental measure of the excluded-volume change upon DNA binding has been discussed in detail³⁵ and is summarized in Results. Equilibrium constants (K_A) for the association of wild type EcoRI and A138T with specific and miscognate DNA sites were measured as a function of water activity (a_w), varied by increasing concentrations (0.0 to 0.75 M) of PEG200 in binding

buffer (see above) plus 0.22M NaCl. Representative binding data are shown in Figure S1. Because a cosolute creates some nonideality in the reaction buffer, the actual sodium activity was measured at each cosolute concentration, using a sodium ion selective electrode (Radiometer Analytical, Inc). Based on the slope of a $\log K_A$ vs $\log [\text{NaCl}]$ plot, values of K_A were corrected to a constant sodium activity of 0.22M. Osmolalities of the PEG200 solutions were measured using a Wescor 5520 Vapor Pressure osmometer (Logan, UT). The output of the osmometer, expressed in units of mmol/kg (milli-osmolality, mOs), is related to the a_w as follows: $m\text{Osm} \equiv -(10^6 \ln a_w) M_1$, where M_1 is the molecular weight of water⁸⁹. Slopes were obtained by least squares regression and errors in slopes were estimated by Monte Carlo methods. Briefly, 150 $\ln(K_A/K_0)$ vs $\ln(a_w)$ data sets were created such that the spread of individual points for a given PEG200 concentration was a normal distribution with the mean and standard deviation of the actual measurements. The 150 “curves” were then fit and the mean and standard deviation of the fitted slopes were reported in Table 3.

Cosolute effects on first order cleavage reactions

First-order cleavage rate constants were determined (see above) in the presence of various concentrations of PEG200, as indicated in the preceding section. Slopes and the associated errors were determined as described in the preceding section.

Ethylation interference footprinting

Footprints of wild type EcoRI and A138T proteins on miscognate DNAs were obtained using the ethylation interference method⁹⁰ as described^{1, 15}. Each DNA strand was ethylated separately with N-ethyl-N-nitrosourea, then 5'-end labeled and annealed to its unlabeled, unethylated complement. An ethylation reaction time of 90 minutes produced the desired one ethylation event (on average) per DNA. The footprinting procedure was modified as follows: Protein complexes with the ethylated DNA constructs were formed in binding buffer (pH 7.5) plus 3 mM CaCl_2 and 50 mM NaCl (for wild type protein) or 90 mM NaCl (for A138T protein). The presence of Ca^{2+} reduces partitioning of EcoRI* complexes into nonspecific binding mode. Concentrations of the DNA and enzyme were adjusted such that 30% of the DNA was enzyme bound at equilibrium. Enzyme-bound DNA was separated from free DNA by electrophoresis on nondenaturing 10% polyacrylamide gels (20 mM Tris-acetate, 0.5mM EDTA plus 3% DMSO, pH 7.5) in running buffer (8mM Tris-acetate, 0.2 mM EDTA). Gel slices of bands for free and bound DNA were excised and the eluted DNA was purified over Alltech C-18 columns. Alkaline cleavage of the phosphotriester bonds at the points of ethylation, followed by electrophoresis on a 22% denaturing (8M urea) polyacrylamide gel (pH 8.5), results in two 5'-radiolabeled cleavage products terminating in 3'-OH or 3'-OPOEt. The identity of one band of the doublet was determined by alignment with a set of 5'-labeled synthetic oligonucleotides terminating in 3'-OH. Interference data obtained from the autoradiograms were analyzed as described¹.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Lewis Jacobson for valuable discussions. This work was supported by National Institutes of Health MERIT award 5R37-GM29207 to L.J.-J.

⁺Supported by NIH MERIT award 5R37-GM29207

References

1. Lesser DR, Kurpiewski MR, Jen-Jacobson L. The energetic basis of specificity in the Eco RI endonuclease--DNA interaction. *Science*. 1990; 250:776–86. [PubMed: 2237428]
2. Jen-Jacobson L. Protein-DNA recognition complexes: conservation of structure and binding energy in the transition state. *Biopolymers*. 1997; 44:153–80. [PubMed: 9354759]
3. Sapienza PJ, Dela Torre CA, McCoy WH III, Jana SV, Jen-Jacobson L. Thermodynamic and kinetic basis for the relaxed DNA sequence specificity of “promiscuous” mutant EcoRI endonucleases. *J Mol Biol*. 2005; 348:307–24. [PubMed: 15811370]
4. Pingoud A, Fuxreiter M, Pingoud V, Wende W. Type II restriction endonucleases: structure and mechanism. *Cell Mol Life Sci*. 2005; 62:685–707. [PubMed: 15770420]
5. Rohs R, Sklenar H, Shakked Z. Structural and energetic origins of sequence-specific DNA bending: Monte Carlo simulations of papillomavirus E2-DNA binding sites. *Structure*. 2005; 13:1499–509. [PubMed: 16216581]
6. Lesser DR, Kurpiewski MR, Waters T, Connolly BA, Jen-Jacobson L. Facilitated distortion of the DNA site enhances EcoRI endonuclease-DNA recognition. *Proc Natl Acad Sci U S A*. 1993; 90:7548–52. [PubMed: 8356054]
7. Hancock SP, Hiller DA, Perona JJ, Jen-Jacobson L. The energetic contribution of induced electrostatic asymmetry to DNA bending by a site-specific protein. *J Mol Biol*. 2011; 406:285–312. [PubMed: 21167173]
8. Otwinowski Z, Schevitz RW, Zhang RG, Lawson CL, Joachimiak A, Marmorstein RQ, Luisi BF, Sigler PB. Crystal structure of trp repressor/operator complex at atomic resolution. *Nature*. 1988; 335:321–9. [PubMed: 3419502]
9. Rohs R, Jin X, West SM, Joshi R, Honig B, Mann RS. Origins of specificity in protein-DNA recognition. *Annu Rev Biochem*. 2010; 79:233–69. [PubMed: 20334529]
10. Lawson, CL.; Berman, HM. Indirect readout of DNA sequence by proteins.. In: Rice, PA.; Correll, C., editors. *Structural Biology of Protein-Nucleic Acid Interactions*. Royal Society of Chemistry Publishing; Cambridge: 2008. p. 66-90.
11. Mossing MC, Record MT Jr. Thermodynamic origins of specificity in the lac repressor-operator interaction. Adaptability in the recognition of mutant operator sites. *J Mol Biol*. 1985; 186:295–305. [PubMed: 4087296]
12. Jen-Jacobson L. Structural-perturbation approaches to thermodynamics of site- specific protein-DNA interactions. *Methods Enzymol*. 1995; 259:305–44. [PubMed: 8538460]
13. Hiller DA, Rodriguez AM, Perona JJ. Non-cognate enzyme-DNA complex: structural and kinetic analysis of EcoRV endonuclease bound to the EcoRI recognition site GAATTC. *J Mol Biol*. 2005; 354:121–36. [PubMed: 16236314]
14. Frank DE, Saecker RM, Bond JP, Capp MW, Tsodikov OV, Melcher SE, Levandoski MM, Record MT Jr. Thermodynamics of the interactions of lac repressor with variants of the symmetric lac operator: effects of converting a consensus site to a non-specific site. *J Mol Biol*. 1997; 267:1186–206. [PubMed: 9150406]
15. Engler LE, Welch KK, Jen-Jacobson L. Specific binding by EcoRV endonuclease to its DNA recognition site GATATC. *J Mol Biol*. 1997; 269:82–101. [PubMed: 9193002]
16. Engler, LE. Specificity determinants in the BamHI-DNA interaction. Ph.D. Thesis. University of Pittsburgh; 1998.
17. Grigorescu, A. Structural and Energetic Determinants of the Binding Specificity of EcoRI Endonuclease. Ph.D. Thesis. University of Pittsburgh; 2003.
18. Kalodimos CG, Biris N, Bonvin AM, Levandoski MM, Guennuegues M, Boelens R, Kaptein R. Structure and flexibility adaptation in nonspecific and specific protein-DNA complexes. *Science*. 2004; 305:386–9. [PubMed: 15256668]
19. Romanuka J, Folkers GE, Biris N, Tishchenko E, Wienk H, Bonvin AM, Kaptein R, Boelens R. Specificity and affinity of Lac repressor for the auxiliary operators O2 and O3 are explained by the structures of their protein-DNA complexes. *J Mol Biol*. 2009; 390:478–89. [PubMed: 19450607]

20. Frederick CA, Grable J, Melia M, Samudzi C, Jen-Jacobson L, Wang BC, Greene P, Boyer HW, Rosenberg JM. Kinked DNA in crystalline complex with EcoRI endonuclease. *Nature*. 1984; 309:327–31. [PubMed: 6328307]
21. Kostrewa D, Winkler FK. Mg²⁺ binding to the active site of EcoRV endonuclease: a crystallographic study of complexes with substrate and product DNA at 2 Å resolution. *Biochemistry*. 1995; 34:683–96. [PubMed: 7819264]
22. Newman M, Strzelecka T, Dorner LF, Schildkraut I, Aggarwal AK. Structure of Bam HI endonuclease bound to DNA: partial folding and unfolding on DNA binding. *Science*. 1995; 269:656–63. [PubMed: 7624794]
23. Yang Z, Kurpiewski MR, Ji M, Townsend JE, Mehta P, Jen-Jacobson L, Saxena S. ESR spectroscopy identifies inhibitory Cu²⁺ sites in a DNA-modifying enzyme to reveal determinants of catalytic specificity. *Proc Natl Acad Sci U S A*. 2012; 109:E993–1000. [PubMed: 22493217]
24. Grigorescu, A.; Horvath, M.; Wilkosz, PA.; Chandrasekhar, K.; Rosenberg, JM. The integration of recognition and cleavage: X-ray structures of pre-transition state complex, post reactive complex and the DNA-free endonuclease.. In: Pingoud, AM., editor. *Restriction Endonucleases*. Springer; Berlin: 2003. p. 137-173.
25. Stone KM, Townsend JE, Sarver J, Sapienza PJ, Saxena S, Jen-Jacobson L. Electron Spin Resonance Shows Common Structural Features for Different Classes of EcoRI-DNA Complexes. *Angew Chem Int Ed Engl*. 2008; 47:10192–4. [PubMed: 19021169]
26. Alves J, Ruter T, Geiger R, Fliess A, Maass G, Pingoud A. Changing the hydrogen-bonding potential in the DNA binding site of EcoRI by site-directed mutagenesis drastically reduces the enzymatic activity, not, however, the preference of this restriction endonuclease for cleavage within the site-GAATTC. *Biochemistry*. 1989; 28:2678–84. [PubMed: 2659077]
27. Dorner LF, Bitinaite J, Whitaker RD, Schildkraut I. Genetic analysis of the base-specific contacts of BamHI restriction endonuclease. *J Mol Biol*. 1999; 285:1515–23. [PubMed: 9917393]
28. Fritz A, Kuster W, Alves J. Asn141 is essential for DNA recognition by EcoRI restriction endonuclease. *FEBS Lett*. 1998; 438:66–70. [PubMed: 9821960]
29. Ivanenko T, Heitman J, Kiss A. Mutational analysis of the function of Met137 and Ile197, two amino acids implicated in sequence-specific DNA recognition by the EcoRI endonuclease. *Biol Chem*. 1998; 379:459–65. [PubMed: 9628338]
30. Needels MC, Fried SR, Love R, Rosenberg JM, Boyer HW, Greene PJ. Determinants of EcoRI endonuclease sequence discrimination. *Proc Natl Acad Sci U S A*. 1989; 86:3579–83. [PubMed: 2657723]
31. Osuna J, Flores H, Soberon X. Combinatorial mutagenesis of three major groove-contacting residues of EcoRI: single and double amino acid replacements retaining methyltransferase-sensitive activities. *Gene*. 1991; 106:7–12. [PubMed: 1937043]
32. Zhu Z, Zhou J, Friedman AM, Xu SY. Isolation of BsoBI restriction endonuclease variants with altered substrate specificity. *J Mol Biol*. 2003; 330:359–72. [PubMed: 12823974]
33. Heitman J, Model P. Mutants of the EcoRI endonuclease with promiscuous substrate specificity implicate residues involved in substrate recognition. *Embo J*. 1990; 9:3369–78. [PubMed: 2209548]
34. Sapienza PJ, Rosenberg JM, Jen-Jacobson L. Structural and thermodynamic basis for enhanced DNA binding by a promiscuous mutant EcoRI endonuclease. *Structure*. 2007; 15:1368–82. [PubMed: 17997963]
35. Jen-Jacobson, L.; Jacobson, LA. The role of water and the effects of small ions in site-specific protein-DNA interactions.. In: Rice, PA.; Correll, C., editors. *Structural Biology of Protein-Nucleic Acid Interactions*. Royal Society of Chemistry Publishing; Cambridge: 2008.
36. Cao, DF. Co-solute effects as probes of the role of water release in DNA binding and catalysis by three restriction endonucleases. Ph.D. Thesis. University of Pittsburgh; 2001.
37. Schellman JA. A Simple-Model for Solvation in Mixed-Solvents - Applications to the Stabilization and Destabilization of Macromolecular Structures. *Biophysical Chemistry*. 1990; 37:121–140. [PubMed: 2285775]

38. Eisenberg H. Protein and Nucleic-Acid Hydration and Cosolvent Interactions - Establishment of Reliable Base-Line Values at High Cosolvent Concentrations. *Biophysical Chemistry*. 1994; 53:57–68. [PubMed: 7841332]
39. Record MT, Anderson CF. Interpretation of Preferential Interaction Coefficients of Nonelectrolytes and of Electrolyte Ions in Terms of a 2-Domain Model. *Biophysical Journal*. 1995; 68:786–794. [PubMed: 7756545]
40. Windsor, DJ.; Wills, PR. Thermodynamic nonideality and protein solvation.. In: Gregory, RB.; Dekker, M., editors. *Protein-Solvation Interactions*. New York, NY: 1995. p. 483
41. Record MT, Zhang WT, Anderson CF. Analysis of effects of salts and uncharged solutes on protein and nucleic acid equilibria and processes: A practical guide to recognizing and interpreting polyelectrolyte effects, Hofmeister effects, and osmotic effects of salts. *Advances in Protein Chemistry*. 1998; 51:281–353. 51. [PubMed: 9615173]
42. Timasheff SN. Control of protein stability and reactions by weakly interacting cosolvents: The simplicity of the complicated. *Advances in Protein Chemistry*. 1998; 51:355–432. 51. [PubMed: 9615174]
43. Neal BL, Asthagiri D, Lenhoff AM. Molecular origins of osmotic second virial coefficients of proteins. *Biophysical Journal*. 1998; 75:2469–2477. [PubMed: 9788942]
44. Minton AP. Molecular crowding: analysis of effects of high concentrations of inert cosolutes on biochemical equilibria and rates in terms of volume exclusion. *Methods Enzymol*. 1998; 295:127–49. [PubMed: 9750217]
45. Parsegian VA, Rand RP, Rau DC. Osmotic stress, crowding, preferential hydration, and binding: A comparison of perspectives. *Proceedings of the National Academy of Sciences of the United States of America*. 2000; 97:3987–3992. [PubMed: 10760270]
46. Chitra R, Smith PE. Preferential interactions of cosolvents with hydrophobic solutes. *Journal of Physical Chemistry B*. 2001; 105:11513–11522.
47. Davis-Searles PR, Saunders AJ, Erie DA, Winzor DJ, Pielak GJ. Interpreting the effects of small uncharged solutes on protein-folding equilibria. *Annu Rev Biophys Biomol Struct*. 2001; 30:271–306. [PubMed: 11340061]
48. Timasheff SN. Protein-solvent preferential interactions, protein hydration, and the modulation of biochemical reactions by solvent components. *Proceedings of the National Academy of Sciences of the United States of America*. 2002; 99:9721–9726. [PubMed: 12097640]
49. Anderson CF, Felitsky DJ, Hong J, Record MT. Generalized derivation of an exact relationship linking different coefficients that characterize thermodynamic effects of preferential interactions. *Biophysical Chemistry*. 2002; 101:497–511. [PubMed: 12488023]
50. Schellman JA. Protein stability in mixed solvents: A balance of contact interaction and excluded volume. *Biophysical Journal*. 2003; 85:108–125. [PubMed: 12829469]
51. Schurr JM, Rangel DP, Aragon SR. A contribution to the theory of preferential interaction coefficients. *Biophys J*. 2005; 89:2258–76. [PubMed: 16055532]
52. Schellman JA. Destabilization and stabilization of proteins. *Quarterly Reviews of Biophysics*. 2005; 38:351–361. [PubMed: 16526966]
53. Shimizu S, Matubayasi N. Preferential hydration of proteins: A Kirkwood-Buff approach. *Chemical Physics Letters*. 2006; 420:518–522.
54. Shulgin IL, Ruckenstein E. A protein molecule in a mixed solvent: The preferential binding parameter via the Kirkwood-Buff theory. *Biophysical Journal*. 2006; 90:704–707. [PubMed: 16272443]
55. Smith PE. Chemical potential derivatives and preferential interaction parameters in biological systems from Kirkwood-Buff theory. *Biophysical Journal*. 2006; 91:849–856. [PubMed: 16679363]
56. Rosgen J, Pettitt BM, Bolen DW. An analysis of the molecular origin of osmolyte-dependent protein stability. *Protein Science*. 2007; 16:733–743. [PubMed: 17327389]
57. Pierce V, Kang M, Aburi M, Weerasinghe S, Smith PE. Recent applications of Kirkwood-Buff theory to biological systems. *Cell Biochemistry and Biophysics*. 2008; 50:1–22. [PubMed: 18043873]

58. Kirkwood JG, Buff FP. The Statistical Mechanical Theory of Solutions .1. Journal of Chemical Physics. 1951; 19:774–777.
59. Ben-Naim, A. Statistical Thermodynamics for Chemists and Biochemists. Plenum; NY: 1992. p. 372-382.p. 448-457.
60. Jen-Jacobson L, Engler LE, Ames JT, Kurpiewski MR, Grigorescu A. Thermodynamic parameters of specific and nonspecific protein-DNA binding. Supramolecular Chemistry. 2000; 12:143–160.
61. Takeda Y, Ross PD, Mudd CP. Thermodynamics of Cro protein-DNA interactions. Proc Natl Acad Sci U S A. 1992; 89:8180–4. [PubMed: 1518844]
62. Ladbury JE, Wright JG, Sturtevant JM, Sigler PB. A thermodynamic study of the trp repressor-operator interaction. J Mol Biol. 1994; 238:669–81. [PubMed: 8182742]
63. Berger C, Jelesarov I, Bosshard HR. Coupled folding and site-specific binding of the GCN4-bZIP transcription factor to the AP-1 and ATF/CREB DNA sites studied by microcalorimetry. Biochemistry. 1996; 35:14984–91. [PubMed: 8942664]
64. Sturtevant JM. Heat capacity and entropy changes in processes involving proteins. Proc Natl Acad Sci U S A. 1977; 74:2236–40. [PubMed: 196283]
65. Livingstone JR, Spolar RS, Record MT Jr. Contribution to the thermodynamics of protein folding from the reduction in water-accessible nonpolar surface area. Biochemistry. 1991; 30:4237–44. [PubMed: 2021617]
66. Spolar RS, Livingstone JR, Record MT Jr. Use of liquid hydrocarbon and amide transfer data to estimate contributions to thermodynamic functions of protein folding from the removal of nonpolar and polar surface from water. Biochemistry. 1992; 31:3947–55. [PubMed: 1567847]
67. Spolar RS, Record MT. Coupling of Local Folding to Site-Specific Binding of Proteins to DNA. Science. 1994; 263:777–784. [PubMed: 8303294]
68. Haq I, O'Brien R, Lagunavicius A, Siksnys V, Ladbury JE. Specific DNA recognition by the type II restriction endonuclease MunI: the effect of pH. Biochemistry. 2001; 40:14960–7. [PubMed: 11732916]
69. Laidler KJ, King MC. The Development of Transition-State Theory. Journal of Physical Chemistry. 1983; 87:2657–2664.
70. Rosenberg JM. Structure and function of restriction endonucleases. Current Opinion in Structural Biology. 1991; 1:104–113.
71. Kurpiewski MR, Engler LE, Wozniak LA, Kobylanska A, Koziolkiewicz M, Stec WJ, Jen-Jacobson L. Mechanisms of coupling between DNA recognition specificity and catalysis in EcoRI endonuclease. Structure (Camb). 2004; 12:1775–88. [PubMed: 15458627]
72. Sapienza PJ, Lee AL. Using NMR to study fast dynamics in proteins: methods and applications. Curr Opin Pharmacol. 2010; 10:723–30. [PubMed: 20933469]
73. Nucci NV, Pometun MS, Wand AJ. Site-resolved measurement of water-protein interactions by solution NMR. Nat Struct Mol Biol. 2011; 18:245–9. [PubMed: 21196937]
74. Tzeng SR, Kalodimos CG. Protein activity regulation by conformational entropy. Nature. 2012; 488:236–40. [PubMed: 22801505]
75. Wand AJ. The dark energy of proteins comes to light: conformational entropy and its role in protein function revealed by NMR relaxation. Curr Opin Struct Biol. 2013; 23:75–81. [PubMed: 23246280]
76. Boehr DD, Nussinov R, Wright PE. The role of dynamic conformational ensembles in biomolecular recognition. Nat Chem Biol. 2009; 5:789–96. [PubMed: 19841628]
77. Jen-Jacobson L, Kurpiewski M, Lesser D, Grable J, Boyer HW, Rosenberg JM, Greene PJ. Coordinate ion pair formation between EcoRI endonuclease and DNA. J Biol Chem. 1983; 258:14638–46. [PubMed: 6315732]
78. Warshaw MM, Cantor CR. Oligonucleotide interactions. IV. Conformational differences between deoxy- and ribonucleoside phosphates. Biopolymers. 1970; 9:1079–103. [PubMed: 5449436]
79. Senior M, Jones RA, Breslauer KJ. Influence of dangling thymidine residues on the stability and structure of two DNA duplexes. Biochemistry. 1988; 27:3879–85. [PubMed: 3408733]

80. Lesser DR, Grajkowski A, Kurpiewski MR, Koziolkiewicz M, Stec WJ, Jen-Jacobson L. Stereoselective interaction with chiral phosphorothioates at the central DNA kink of the EcoRI endonuclease-GAATTC complex. *J Biol Chem.* 1992; 267:24810–8. [PubMed: 1447218]
81. Fukada H, Takahashi K. Enthalpy and heat capacity changes for the proton dissociation of various buffer components in 0.1 M potassium chloride. *Proteins.* 1998; 33:159–66. [PubMed: 9779785]
82. Baker BM, Murphy KP. Evaluation of linked protonation effects in protein binding reactions using isothermal titration calorimetry. *Biophys J.* 1996; 71:2049–55. [PubMed: 8889179]
83. Kozlov AG, Lohman TM. Large contributions of coupled protonation equilibria to the observed enthalpy and heat capacity changes for ssDNA binding to Escherichia coli SSB protein. *Proteins Suppl.* 2000; 4:8–22.
84. Horn JR, Brandts JF, Murphy KP. van't Hoff and calorimetric enthalpies II: effects of linked equilibria. *Biochemistry.* 2002; 41:7501–7. [PubMed: 12044184]
85. Bectel WJ, Schellman JA. Protein stability curves. *Biopolymers.* 1987; 26:1859–77. [PubMed: 3689874]
86. Ha JH, Spolar RS, Record MT Jr. Role of the hydrophobic effect in stability of site-specific protein-DNA complexes. *J Mol Biol.* 1989; 209:801–16. [PubMed: 2585510]
87. Dec SF, Gill SJ. Temperature dependence of errors in parameters derived from van't Hoff studies. *J. Chem. Educ.* 1985; 62:879–881.
88. Clark ECW, Glew DN. Evaluation of thermodynamic functions from equilibrium constants. *Trans. Faraday Soc.* 1966; 62:539–547.
89. Zhang W, Capp MW, Bond JP, Anderson CF, Record MT Jr. Thermodynamic characterization of interactions of native bovine serum albumin with highly excluded (glycine betaine) and moderately accumulated (urea) solutes by a novel application of vapor pressure osmometry. *Biochemistry.* 1996; 35:10506–16. [PubMed: 8756707]
90. Siebenlist U, Gilbert W. Contacts between Escherichia coli RNA polymerase and an early promoter of phage T7. *Proc Natl Acad Sci U S A.* 1980; 77:122–6. [PubMed: 6987648]

Highlights

- DNA site specificity is relaxed in promiscuous A138T mutants of EcoRI endonuclease
- ΔC°_p and cosolute effects for A138T-EcoRI* binding indicate a “specific” binding mode
- Phosphate contacts to EcoRI* DNA by A138T resemble the wild-type specific complex
- Novel A138T contacts may confer preference for 5'-RY flanked EcoRI* site
- Global changes in the mutant complexes produce relaxed DNA site specificity

scheme 1.

Figure 1.

Effect of flanking context on specific and EcoRI* binding and cleavage. The transition state interaction free energy difference $\Delta\Delta G^{\ddagger}_I$ (y-axis) is plotted against binding free energy difference $\Delta\Delta G^{\circ}_{bind}$ (x-axis). Each parameter (see eqs. 1-3) is set at zero for the GAATTC site with the highest binding affinity. Circles show data (Table 1) for GAATTC sites with various flanking trinucleotides while inverted triangles show data for AAATTC EcoRI* sites (Table 2). Lines show unit slope ($\Delta\Delta G^{\ddagger}_I = \Delta\Delta G^{\circ}_{bind}$). For mutant A138T (right panel), a subset of data (arrows) for $R_{-5}Y_{-4}AAATTC$ lie just above the line of unit slope.

GAATTC / gcaAAATTC / gatAAATTC / ggtAAATTC / NS

Figure 2.

Cosolute effects distinguish specific and miscognate modes of interaction of DNA sites with different flanking contexts. Panels (a) and (b) show effects of PEG200 on equilibrium association constants (representative raw data are shown in Figure S1); panels (c) and (d) show effects on k_1 , the first-order cleavage rate constant in the GAA half-site (see Materials and Methods). Data points and associated error bars represent the mean and standard deviation from two and three independent measurements of the equilibrium binding constant and first-order cleavage rate constant, respectively. In the absence of cosolute, K_0 is the equilibrium association constant and $k_{1,0}$ is the cleavage rate constant. Water activity a_w was determined by vapor pressure osmometry ($m\text{Osm} = (10^6 \ln a_w)/18$) as described in Materials & Methods. The slope of each line for binding [panels (a),(b)] the change in excluded volume (in units of numbers of water molecules) that differ between reactants and products

(see text), so negative slopes indicate compaction upon binding. Data are from Table 3 (binding) and Table 4 (catalysis). The slight tendency to decreasing slope [panel (c) or negative slope (panel (d))] at high PEG200 concentrations may reflect cosolute-driven aggregation of free enzyme, so that free EcoRI dimer is no longer in large excess to DNA. The color scheme for the oligonucleotides with various flanking triplets is shown below the plots. We have shown previously that cosolute does not affect the rate of *specific* DNA cleavage³⁶.

Figure 3.

A138T binding to RY-flanked-EcoRI* DNA sites has the thermodynamic signature of the specific binding mode. The temperature dependence of binding to DNA constructs containing sites (underlined) as follows: (a), specific GGCGAATTC (b), CGCAAATTC; (c) GATAAATTC; (d) GGTAAATTC (A138T only) (e), GCAAAATTC (f) nonspecific CTTAAG. See Table 5 for the contexts in which these sites are embedded. Note that the axes in all panels have the same scale interval, but the Y-axis windows are shifted; thus, curvatures in the panels are directly comparable. Each symbol is an independent binding experiment. Data are shown for the wild-type EcoRI endonuclease (●) and mutant A138T (○). The change in heat capacity (ΔC_p) upon binding was determined by multiparametric

fits (lines in this Figure) to the van't Hoff equation. Since ΔC_p° is the second derivative of ΔG° with respect to $1/T$, curvature of the plots indicates the magnitude of ΔC_p° . Fitted thermodynamic parameters and associated errors are presented in Table 5.

Figure 4

. Specific-like phosphate contacts in A138T complexes with EcoRI* sites. (a) Sample ethylation interference footprints of wild-type and A138T proteins on the EcoRI* AAATTC site, with flanking triplet 5'-GAT. The duplex 17-nt substrate used for footprinting is shown at the bottom of figure. The EcoRI* site is shown in bold and the flanking triplet is shown in red. Top strand contains the unmodified GAA half-site; bottom strand contains modified AAA-half site. Lanes show ethylated DNA alone (Co), enzyme-bound (B) and free (F) DNA. Note the stronger footprints made by the A138T enzyme on both strands relative to the wild-type. (b)–(d) Quantitation of interference footprints. Values on the ordinate (free DNA/bound DNA) were obtained by densitometric quantitation of bands from footprints like that in panel (a). Data show means $\pm$ SD of at least 3 independent footprinting experiments. A value >1 (i.e., band enriched in free over bound DNA) indicates that ethylation at that particular phosphate interferes with protein binding. Phosphates outside the indicated region gave no interference. The AAATTC site is underlined for each DNA strand. Upper half of each panel shows the strand containing the unmodified GAA half-site; lower half shows strand containing the AAA half-site. (b) GGT flank. (c) GAT flank, (d) GCA flank. The A138T mutant protein makes stronger phosphate contacts to both DNA strands, with the most dramatic differences in the AAA half-site at **GpApTAAApTTC** (cf. panel A) and at **GpGpTAAApTTC**. There is little or no difference between wild-type and A138T proteins for the GCAAAATTC site.

Figure 5.

Phosphate contacts and flanking base contacts in the x-ray structure (PDBID 2OXV) of the A138T-GAATTC complex. (A) The hydrogen bond network contacting the phosphates at X₋₆pX₋₅X₋₄ (highlighted by asterisks). X₋₆pX₋₅ is recognized by a main-chain hydrogen bond from Gly196 and X₋₅pX₋₄ is by a side-chain hydrogen bond from Arg203, which also recognizes G₋₃ by a water-mediated hydrogen bond. Arg200 donates a direct hydrogen bond to recognize G₋₃. (B) The same network region in the wild-type-GAATTC complex (coordinates from a highly refined version of PDBID 1CKQ²⁴), showing that the wild-type enzyme relies on water molecules to achieve shape complementarity with the two bases upstream of GAATTC. (C) The increased size of the Thr138 side-chain permits the mutant enzyme to make contacts to G₋₅ and C₋₄ with concomitant release of an ordered water molecule.

Table 1
Effect of flanking context on cognate site binding and cleavage by EcoRI endonuclease

Sequence ^a	$K_A(M^{-1})$ ^{b,c}	$\Delta\Delta G_{bind}^o$ (kcal/mol) ^d	$k_{1,cleave}$ (s ⁻¹) ^e	$\Delta\Delta G^{o,\ddagger}$ (kcal/mol) ^f	$\Delta\Delta G^{o,\ddagger}$ (kcal/mol) ^g
wild-type					
CGCgaattcGCG	$9.6 \pm 0.3 \times 10^9$	0	0.3 ± 0.1	0	0
GGCgaattcGCC	$1.3 \pm 0.4 \times 10^9$	1.2 ± 0.2	0.6 ± 0.1	-0.4 ± 0.2	0.8 ± 0.3
GGTgaattcACC	$7.9 \pm 0.9 \times 10^8$	1.5 ± 0.1	0.3 ± 0.1	0.0 ± 0.3	1.5 ± 0.3
GATgaattcATC	$1.7 \pm 0.1 \times 10^8$	2.4 ± 0.1	0.7 ± 0.1	-0.5 ± 0.2	1.9 ± 0.2
GCAgaattcTGC	$8.1 \pm 0.2 \times 10^7$	2.8 ± 0.1	0.5 ± 0.1	-0.3 ± 0.2	2.5 ± 0.2
CCCgaattcGGG	$4.2 \pm 0.5 \times 10^7$	3.2 ± 0.1	0.5 ± 0.1	-0.3 ± 0.2	2.9 ± 0.2
AAAAGAattcTTTT	$1.1 \pm 0.1 \times 10^7$	4.0 ± 0.1	0.2 ± 0.1	0.2 ± 0.4	4.2 ± 0.4
A138T					
CGCgaattcGCG	$1.3 \pm 0.1 \times 10^{11}$	0	0.4 ± 0.1	0	0
GGTgaattcACC	$1.2 \pm 0.1 \times 10^{11}$	0.1 ± 0.1	0.6 ± 0.1	-0.2 ± 0.2	-0.1 ± 0.2
GATgaattcATC	$3.6 \pm 0.5 \times 10^{10}$	0.7 ± 0.1	0.7 ± 0.1	-0.3 ± 0.2	0.4 ± 0.2
CCCgaattcGGG	$2.2 \pm 0.1 \times 10^{10}$	1.0 ± 0.1	0.5 ± 0.1	-0.1 ± 0.2	0.9 ± 0.2
GGCgaattcGCC	$1.8 \pm 0.1 \times 10^{10}$	1.1 ± 0.1	0.6 ± 0.1	-0.2 ± 0.2	0.9 ± 0.2
GCAgaattcTGC	$1.1 \pm 0.1 \times 10^9$	2.8 ± 0.1	0.5 ± 0.1	-0.1 ± 0.2	2.7 ± 0.2
AAAAGAattcTTTT	$1.1 \pm 0.2 \times 10^8$	4.1 ± 0.1	0.2 ± 0.1	0.4 ± 0.3	4.5 ± 0.3

^aSequence of double stranded 24 base-pair oligonucleotide is 5'-GGGGCGGGXXXgaattcZZZGGCGGC; segment shown in Table is underlined.

^bEquilibrium association constants were measured by direct filter binding (See Methods for details) in bis-Tris-propane binding buffer (pH 7.3) plus 0.22M KCl. Values represent means $\pm$ SD of 3 determinations.

^cData in this column were taken from ref. 3.

^dFor each enzyme, the difference in the standard binding free energy between a reference site and other GAATTC sites in different flanking contexts is calculated at 294K as:
$$\Delta\Delta G_{bind}^o = -RT \ln \frac{K_A^{flank}}{K_A^{ref}}$$

^eFirst-order rate constants for DNA cleavage were measured as described in Methods. k_1 reports on cleavage of one DNA strand; in all cases for the GAATTC site, cleavage in each DNA strand proceeded at the same rate (data not shown).

^fFor each enzyme, the difference in the standard activation free energy for DNA cleavage between a reference site and other GAATTC sites in different flanking contexts is calculated at 294K as:

$$\Delta\Delta G^{\circ}_{bind} = -RT \ln \frac{k_{cleave}^{flank}}{k_{cleave}^{ref}}$$

^gFor each enzyme, the difference in the transition state interaction free energy between a reference site and other GAATTC sites in different flanking contexts is calculated at 294K as: $\Delta\Delta G^{\circ}[\ddagger] = \Delta\Delta G^{\circ}_{bind} + \Delta\Delta G^{\circ}[\ddagger]$

Table 2

Effect of flanking context on EcoRI* DNA binding and cleavage

Sequence ^a	K _a (M ⁻¹) ^b	ΔΔG ^c _{bind} (kcal/mol)	k _{1,cleave} (S ⁻¹) ^{b,c,d}	ΔΔG ^c (kcal/mol) ^b	ΔΔG ^c ₁ (kcal/mol) ^b
wild-type					
CGCgaattcGCG	9.6 ± 0.3 × 10 ⁹	0	0.3 ± 0.1	0	0
CGCaattcGCG	1.1 ± 0.1 × 10 ⁵	6.7 ± 0.1	3.8 ± 0.4 × 10 ⁻³	2.6 ± 0.2	9.3 ± 0.2
GGTaaattcACC	5.3 ± 0.2 × 10 ⁴	7.1 ± 0.1	3.0 ± 0.1 × 10 ⁻⁴	4.0 ± 0.2	11.1 ± 0.2
GATTaaattcATC	6.0 ± 0.9 × 10 ⁴	7.0 ± 0.1	1.1 ± 0.1 × 10 ⁻⁴	4.6 ± 0.2	11.6 ± 0.2
GGCaaattcGCC	6.1 ± 0.7 × 10 ⁴	7.0 ± 0.1	6.9 ± 0.7 × 10 ⁻⁴	3.6 ± 0.2	10.6 ± 0.2
GC AaaattcTGC	6.1 ± 0.3 × 10 ⁴	7.0 ± 0.1	1.9 ± 0.3 × 10 ⁻⁴	4.3 ± 0.2	11.3 ± 0.2
AAA AaaattcTTTT	8.0 ± 1.2 × 10 ⁴	6.8 ± 0.1	1.3 ± 0.6 × 10 ⁻⁵	5.9 ± 0.3	12.7 ± 0.3
AI38T					
CGCgaattcGCG	1.3 ± 0.1 × 10 ¹¹	0	0.4 ± 0.1	0	0
CGCaattcGCG	2.3 ± 0.5 × 10 ⁶	6.4 ± 0.1	2.1 ± 0.5 × 10 ⁻²	1.7 ± 0.2	8.1 ± 0.2
GGTaaattcACC	4.6 ± 0.7 × 10 ⁵	7.3 ± 0.1	4.2 ± 0.7 × 10 ⁻²	1.3 ± 0.2	8.6 ± 0.2
GATTaaattcATC	3.9 ± 0.8 × 10 ⁵	7.4 ± 0.1	1.9 ± 0.3 × 10 ⁻²	1.8 ± 0.2	9.2 ± 0.2
GGCaaattcGCC	1.6 ± 0.4 × 10 ⁵	8.0 ± 0.2	3.9 ± 0.1 × 10 ⁻³	2.7 ± 0.1	10.7 ± 0.2
GC AaaattcTGC	4.6 ± 1.7 × 10 ⁵	7.3 ± 0.2	1.2 ± 0.1 × 10 ⁻³	3.4 ± 0.2	10.7 ± 0.2
AAA Aaaattd TTT	1.4 ± 0.1 × 10 ⁵	8.0 ± 0.1	1.2 ± 0.5 × 10 ⁻⁴	4.7 ± 0.2	12.7 ± 0.2

^aSequence of double stranded 24 base-pair oligonucleotide (segment in Table underlined) was
5'-GGCGGGGXX~~XXaaattcZZZ~~GGCGC
3'-CCCCCCCZZZttaagXXCCGCG

^bAll parameters and conditions are defined in the legend to Table 1.

^cData in this column were given graphically in Fig. 5 of ref. 34.

^dFor EcoRI* sites, k1 is the rate of cleavage of the lower DNA strand (footnote a) containing the canonical GAA half-site.

Table 3

PEG200 effects on specific, EcoRI*, and nonspecific DNA binding

Sequence ^a	wild-type ^b $\left(-\frac{\ln K_A}{\ln a_w} \right)_{T,P}$	A138T ^c $\left(-\frac{\ln K_A}{\ln a_w} \right)_{T,P}$
GAC Specific ^c	-540 ± 8.9	-530 ± 14
GGT EcoRI*	-400 ± 12	-520 ± 22
GAT EcoRI*	-400 ± 9.0	-520 ± 5.1
Nonspecific	-220 ± 42	-260 ± 21

^a DNA sequences (recognition site in lowercase) were as follows: GAC Specific: 5'-TGCCAGGACgaattcGTCCAGCGA GGT EcoRI*: 5'-TGGTaaattcACC GAT EcoRI*: 5'-TCGATAaattcATCC Nonspecific: 5'-GTGCcttaagCGCG

^b Equilibrium association constants K_A were measured in bis-Tris-propane binding buffer (pH 7.3) plus 0.22 M NaCl plus various concentrations of PEG200 from 0 to 0.75 M. Data for the specific GAATTC site are by direct binding, whereas those for AAATTC EcoRI* and nonspecific DNA are by equilibrium competition (cf. Figure S1). Slopes [see Figures 2 (a) and (b)] represent least squares fits of all data points with at least two determinations per PEG200 concentration. Errors are estimated based on Monte Carlo methods (see Materials and Methods).

^c The GAC flank for the specific site was chosen for direct equilibrium binding for its reduced binding affinity ($\Delta\Delta G^\circ$ relative to the reference (Table 1) CGCGAATTC of +2.6 and +2.8 kcal/mol for wild-type and A138T respectively in the absence of PEG200). This avoids prohibitively low DNA concentrations while measuring the extremely tight binding at high [PEG]. Previous work in our lab has shown that flanking sequence does not affect the cosolute dependence of specific DNA binding³⁶.

Table 4

Effect of PEG200 on first order rate constants for EcoRI* DNA cleavage

Sequence ^a	wild-type	wild-type	A138T	A138T
	$\left(\frac{\ln k_1}{\ln a_w} \right)_{T,P}^b$	$\left(\frac{\ln k_2}{\ln a_w} \right)_{T,P}^c$	$\left(\frac{\ln k_1}{\ln a_w} \right)_{T,P}$	$\left(\frac{\ln k_2}{\ln a_w} \right)_{T,P}$
GCA EcoRI*	-100 ± 1.8	-160 ± 3.1	-79 ± 15	-93 ± 8.1
GGT EcoRI*	-160 ± 4.4	-170 ± 4.3	-35 ± 13	-40 ± 24
GAT EcoRI*	-120 ± 4.0	-130 ± 3.8	-30 ± 6.9	-23 ± 9.3

^a Only the 5'-flanking trinucleotide is shown. Full base sequence of the double stranded 24 base-pair oligonucleotide is:
 5' -GGGCGGGXXXaaattcZZZGGCGC
 3' -CCCGCCZZZttaagXXXCCGCG

^b In EcoRI* sites, k_1 is the rate of cleavage of the canonical GAA half-site (lower DNA strand). Errors in the slopes (see Figures 2c,d) were estimated by Monte Carlo methods (See Materials and Methods).

^c In EcoRI* sites, k_2 is the rate of cleavage of the non-canonical AAA half-site (upper DNA strand).

Table 5
Thermodynamic Parameters for Specific, EcoRI*, and Non-specific DNA binding^a

Sequence ^b	wild-type multiparametric Fit ^d				A138T multiparametric Fit ^d			
	ΔG°_{bind} (kcal/mol)	ΔH° (kcal/mol)	$T\Delta S^{\circ}$ (kcal/mol)	ΔC°_P (kcal/mol•K)	ΔG°_{bind} (kcal/mol)	ΔH° (kcal/mol)	$T\Delta S^{\circ}$ (kcal/mol)	ΔC°_P (kcal/mol•K)
GGC Specific	-13.9 ± 0.1	-11.3 ± 0.5	2.6 ± 0.6	-1.6 ± 0.1	-15.3 ± 0.1	-12.7 ± 0.9	2.5 ± 1.0	-1.8 ± 0.1
CGC EcoRI*	-7.9 ± 1.4	-10.8 ± 1.3	-2.9 ± 0.3	-0.7 ± 0.1	-9.3 ± 1.4	-11.4 ± 1.1	-2.1 ± 0.8	-1.2 ± 0.1
GCA EcoRI*	-6.9 ± 1.6	-8.0 ± 1.3	-1.2 ± 0.9	-0.8 ± 0.1	-7.5 ± 1.6	-2.8 ± 0.6	4.7 ± 1.4	-0.7 ± 0.1
GGT EcoRI*	ND	ND	ND	ND	-8.9 ± 0.9	-6.2 ± 0.5	2.8 ± 0.7	-1.4 ± 0.1
GAT EcoRI*	-7.6 ± 1.2	-7.9 ± 1.0	-0.3 ± 0.6	-0.7 ± 0.1	-9.1 ± 1.0	-6.6 ± 0.6	2.6 ± 0.8	-1.5 ± 0.1
Nonspecific	-6.7 ± na	-16.8 ± na	-10.1 ± 7.2	-0.1 ± 0.1	-6.6 ± na	-18.3 ± na	-11.7 ± 8.3	-0.2 ± 0.2

wild-type Clarke and Glew Fit ^d				A138T Clarke and Glew Fit ^d			
GGC Specific	-13.9 ± 0.1	-11.6 ± 0.6	2.3 ± 0.6	-1.4 ± 0.1	-15.3 ± 0.1	-11.6 ± 0.8	3.7 ± 0.7
CGC EcoRI*	-7.9 ± 0.1	-11.2 ± 0.8	-3.4 ± 0.1	-0.5 ± 0.2	-9.4 ± 0.1	-11.1 ± 0.9	-1.7 ± 0.1
GCA EcoRI*	-6.9 ± 0.1	-8.0 ± 1.6	-1.1 ± 0.1	-0.7 ± 0.4	-7.5 ± 0.1	-3.6 ± 1.3	3.9 ± 0.1
GGT EcoRI*	ND	ND	ND	nd	-9.0 ± 0.1	-6.1 ± 0.8	2.9 ± 0.1
GAT EcoRI*	-7.6 ± 0.1	-8.1 ± 1.0	-0.5 ± 0.1	-0.5 ± 0.3	-9.2 ± 0.1	-6.3 ± 0.9	2.8 ± 0.1
Nonspecific	-6.6 ± 0.1	-17.8 ± 1.9	-11.1 ± 1.9	-0.06 ± 0.1	-6.7 ± 0.1	-16.9 ± 1.8	-10.3 ± 1.8

^aEquilibrium association constants (K_A) were measured as a function of temperature by filter binding in a binding buffer containing 0.21 M KCl, 10 mM cacodylate, 1 mM EDTA, 0.01% NaN₃, 100 µg/ml BSA, 50 µM DTT, pH 7.3. Data for the specific GAATTC EcoRI* site are by direct binding, whereas those for AAATTC EcoRI* and nonspecific DNA are by equilibrium competition (see Methods). Values of ΔG° , ΔH° , $T\Delta S^{\circ}$ are fitted values at 298K.

^bDue to the effect of flanking context on DNA melting temperature, DNA lengths were varied to ensure 100% duplex over the entire experimental temperature range. Sequences (recognition site lowercase) were as follows: GGC Specific 5'-GGCGGGGGCGgaattcGCCGGCGC CGC EcoRI*, 5'- TCGCaattcGCG GCA EcoRI*, 5'- GCGCAaattcTGCCG GGT EcoRI*, 5'- GCGGTaattcACCCG GAT EcoRI*, 5'- GCGATaattcATCCG Nonspecific, 5'- GTGCcttaagCGCG

^c ΔG°_{bind} values are not directly comparable to the experiments in Tables 1 and 2, because of differences in temperature (298K here vs. 294K), the buffer ion (cacodylate here vs. bis-Tris-propane) and salt concentration (0.21M here vs. 0.22M).

^dSee Materials and Methods.