Differences in Water Release for the Binding of EcoRI to Specific and Nonspecific DNA Sequences

Nina Y. Sidorova, Donald C. Rau

Proceedings of the National Academy of Sciences of the United States of America, Volume 93, Issue 22 (Oct. 29, 1996), 12272-12277.

Your use of the JSTOR database indicates your acceptance of JSTOR's Terms and Conditions of Use. A copy of JSTOR's Terms and Conditions of Use is available at http://www.jstor.org/about/terms.html, by contacting JSTOR at jstor-info@umich.edu, or by calling JSTOR at (888)388-3574, (734)998-9101 or (FAX) (734)998-9113. No part of a JSTOR transmission may be copied, downloaded, stored, further transmitted, transferred, distributed, altered, or otherwise used, in any form or by any means, except: (1) one stored electronic and one paper copy of any article solely for your personal, non-commercial use, or (2) with prior written permission of JSTOR and the publisher of the article or other text.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

Proceedings of the National Academy of Sciences of the United States of America is published by National Academy of Sciences. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/nas.html.

Proceedings of the National Academy of Sciences of the United States of America
©1996 National Academy of Sciences

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact jstor-info@umich.edu.

©2000 JSTOR

http://www.jstor.org/
Wed Oct 4 08:05:17 2000

Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 12272–12277, October 1996
Biochemistry

Differences in water release for the binding of *Eco*RI to specific and nonspecific DNA sequences

NINA Y. SIDOROVA*† AND DONALD C. RAU‡

*Laboratory of Structural Biology, Division of Computer Research and Technology, and ‡Office of the Director, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Building 5, Room 405, Bethesda, MD 20892

Communicated by Peter H. von Hippel, University of Oregon, Eugene, OR, July 11, 1996 (received for review March 12, 1996)

ABSTRACT **The free energy difference between complexes of the restriction nuclease *Eco*RI with nonspecific DNA and with the enzyme's recognition sequence is linearly dependent on the water chemical potential of the solution, set using several very different solutes, ranging from glycine and glycerol to triethylene glycol and sucrose. This osmotic dependence indicates that the nonspecific complex sequesters some 110 waters more than the specific complex with the recognition sequence. The insensitivity of the difference in number of waters released to the solute identity further indicates that this water is sequestered in a space that is sterically inaccessible to solutes, most likely at the protein–DNA interface of the nonspecific complex. Calculations based on the structure of the specific complex suggest that the apposing DNA and protein surfaces in the nonspecific complex retain approximately a full hydration layer of water.**

The correct functioning of transcription factors and other regulatory proteins that recognize specific DNA sequences requires not only binding with high affinity to the proper sequence but also the ability to distinguish effectively the recognition sequence from all others, including ones that differ by only 1 or 2 bp. A key problem in biophysics is understanding how such binding strength and specificity are so tightly linked in recognition reactions. As yet, there is no simple way to connect structure, derived from x-ray crystallography or NMR, with thermodynamics and with the physics of molecular interactions between individual groups on apposing surfaces. Direct measurements of forces between macromolecules in condensed arrays (reviewed in ref. 1) have been interpreted as showing the dominance of water structuring on interaction energies at close surface separation (<10–15 Å). These measurements suggest that differences in binding free energies for the association of proteins to different DNA sequences in dilute solution should correlate with differences in the number of water molecules retained in the complexes. Crystal structures of many specific DNA–protein complexes show that direct DNA–protein contacts mostly replace DNA–water and protein–water interactions with little or no water left at the interface (2, 3). Comparison of the crystal structure of the estrogen receptor-like DNA binding domain with a noncognate DNA sequence that has a binding constant about an order of magnitude smaller than the recognition sequence shows several additional waters incorporated at the interface between protein and DNA that are not seen in the specific complex (4). Just as differences in proton or salt binding accompanying macromolecular reactions can be thermodynamically probed by the sensitivity of the reaction to pH or salt activity, respectively, differences in the number of water molecules retained by different complexes can be determined from the dependence of binding equilibrium constants on the bulk water activity, controlled by the concentration of added solutes (5).

A paradigm for specific recognition is the binding of restriction nucleases to DNA. The *in vivo* action of these enzymes at incorrect sequences carries a high penalty. The restriction nuclease *Eco*RI binds to its canonical site, GAATTC, with a constant $\approx 10^{11}$ M^{-1} in 0.1 M salt but decreases by a factor of 10^3 to 10^4 when any one of the 6 bp is changed (6), to values nearly indistinguishable from nonspecific binding ($\approx 10^7$ M^{-1}). We report here that the ability of a nonspecific DNA, poly(dI-dC)·poly(dI-dC), to compete with a DNA fragment containing the specific *Eco*RI site for restriction nuclease binding is strongly dependent on bulk water activity. Competitive binding reactions between DNA sequences are not only more easily measured but also more directly relevant for probing the specificity of recognition than measurements of the binding of protein free in solution to cognate DNA sequences. With increasing concentration of neutral solutes, the ability of poly(dI-dC)·poly(dI-dC) to compete with specific fragment for *Eco*RI binding decreases significantly. The free energy difference between specific and nonspecific DNA–protein complexes increases linearly with water chemical potential, consistent with an osmotic effect of solute rather than direct binding. The observed dependence on water activity translates into some 110 water molecules that are released in forming the specific complex but are retained in the nonspecific complex. Further, this number of waters is independent of the chemical nature of the solute used to control water activity and, consequently, reflects a structurally well-defined, water-filled volume associated with the nonspecific complex that sterically excludes solute. It is reasonably expected that this sequestered water is at the protein–DNA interface.

MATERIALS AND METHODS

Materials. The plasmid pUC19 and restriction enzymes *Eco*RI and *Pvu*II were purchased from New England Biolabs and used without further purification. The 322-bp fragment DNA carrying the *Eco*RI binding site, GAATTC, was isolated from *Pvu*II digestion of pUC19 using standard agarose gel techniques. The synthetic polynucleotide poly(dI-dC)·poly(dI-dC) was purchased from Pharmacia. To avoid interference with visualization of the 322-bp fragment, the size of the poly(dI-dC)·poly(dI-dC) was reduced to less than 200 bp by extensive hydrolysis (several hours) at 94°C, followed by slow cooling and concentration by ethanol precipitation. Plasmid DNA and polynucleotide concentrations were determined spectrophotometrically using molar (base pair) extinction coefficients of 1.33×10^4 and 1.48×10^4, respectively. Absorption spectra were obtained with a Shimadzu UV-2101PC spectrophotometer.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

†On leave from: V. A. Engelhardt Institute of Molecular Biology Russian Academy of Sciences, Vavilov Street 32, 117984 Moscow B-334, Russia.

Betaine glycine was purchased from United States Biochemical, triethylene glycol and methyl glucoside were from Fluka, sucrose and glycerol were from Mallinckrodt, and glycine was from Bio-Rad. All were used without further purification. Water chemical potentials are linearly proportional to solute osmolal concentrations ($d\mu_w = -d[\text{osmolal}]/55.6$). Osmolal concentration of solutes were determined by direct measurement using a vapor pressure osmometer (Wescor, Logan, UT; model 5100). The solution pH and Na^+ ion activity (using a sodium ion-selective electrode) were measured with added solute and adjusted if necessary.

Experimental difficulties were encountered with several solutes. DNA and protein will phase separate from polyethylene glycols with molecular weights greater than about 400 at modest concentrations, ≈0.5 osmolal. Gel mobility-shift experiments with added glucose were unsatisfactory. Instead of clear, sharp bands corresponding to free and *Eco*RI-bound DNA fragments, only unquantifiable smears were observed. No problems were encountered with methyl glucoside. There is probably a sufficient concentration of the free aldehyde form of glucose to react significantly with protein amines.

Gel Mobility-Shift Experiments. Mixtures of *Eco*RI, specific site fragment, and varying concentrations of poly(dI-dC)·poly(dI-dC) were incubated at 25°C for 30 min in 25 mM Tris·Cl, pH 7.5/0.1 M NaCl/2.5 mM EDTA/1 mM DTT/0.1 mg/ml BSA/2.5% Ficoll. Under these conditions, we observed no measurable cleavage of the DNA in the absence of Mg^{2+}. Free specific DNA fragment and *Eco*RI-bound complex were separated on 1.5% agarose gels by electrophoresis for 1.5–2 hr at 100 V. The DNA fragment concentration was about 0.75 μM bp (≈2.3 nM in *Eco*RI sites), in 25-μl volumes. Sufficient *Eco*RI restriction nuclease was added to give about 40–60% stoichiometrically bound fragment.

All gel mobility-shift experiments were performed with 1.5% agarose gels rather than polyacrylamide as is standard for this type of experiment (7, 8). Well-separated and easily quantitated bands for DNA–protein complex and free DNA were seen using agarose gels with *Eco*RI binding. The stability of protein–DNA complexes during the gel experiment was probed previously (9) by monitoring the fraction of DNA in complex as a function of time of electrophoresis as described by Fried (8). No dependence of the fraction complex on the time of gel electrophoresis was observed.

Titration both of the protein with the specific DNA fragment and of the DNA with protein verified that the binding of active protein to the specific sequence was stoichiometric under the experimental conditions used. Specifically, at constant protein concentration and with no added competitor DNA, the total amount of specific *Eco*RI–DNA complex is constant over the *Eco*RI sites concentration range of 1–10 nM, corresponding to a 10–90% range of fraction-bound DNA. The error in the measurement of complex only allows a lower bound estimate of the association binding constant of at least $3 \times 10^{10}\ M^{-1}$. The amount of active *Eco*RI protein in the stock solution increases somewhat, depending on the osmolyte nature, with increasing osmolyte concentration. Protein binding is still stoichiometric under the experimental conditions used.

Quantitation and Data Analysis. DNA bands on agarose gels were visualized and quantitated with SYBR Green I (Molecular Probes) staining. The gels were photographed with a Panasonic BD 400 videocamera (averaging 128 frames) connected to a Macintosh IIfx microcomputer, and two Spectroline (Spectronics, Westbury, NY) X-Series UV 254-nm epi-illuminating lamps were used for excitation of the dye fluorescence. Band intensities were quantitated using IMAGE (version 1.45) and specialized macros developed for us. The linearity of the system response was confirmed from the linearity of bands intensities versus DNA size for pBR322 DNA fragments generated by *Msp*I digestion.

The ability of poly(dI-dC)·poly(dI-dC) to compete with specific fragment for the *Eco*RI binding depends on the ratio between specific (sp) and non-specific (nonsp) DNA constants ($K_{comp} = K_{sp}/K_{nonsp}$) (10, 11). At equilibrium, the binding equations for specific and nonspecific reactions to form the DNA–protein complex must be simultaneously satisfied,

$$K_{comp} = \frac{K_{sp}}{K_{nonsp}} = \frac{[(DNA_{sp}\cdot Eco\text{RI})][DNA_{nonsp}]_{free}}{[(DNA_{nonsp}\cdot Eco\text{RI})][DNA_{sp}]_{free}}. \quad [1]$$

The fraction, $f = [(DNA_{sp}\cdot Eco\text{RI})]/[DNA_{sp}]_{total}$, of specific DNA fragment in the complex is determined directly from the gel mobility-shift assay. For weak nonspecific binding relative to specific binding ($K_{nonsp} \ll K_{sp}$), the free and total nonspecific DNA concentrations are essentially identical ($[DNA_{nonsp}]_{free} \approx [DNA_{nonsp}]_{total}$). Lastly, under conditions of virtually stoichiometric binding of *Eco*RI and constant total specific site DNA and protein concentrations, the concentration of nonspecific DNA sites complexed with protein is identical to the loss of specific DNA complex compared with no added competitor,

$$[(DNA_{nonsp}\cdot Eco\text{RI})] = (f_0 - f)[DNA_{sp}]_{total}, \quad [2]$$

where f_0 is the experimentally determined fraction of specific fragment with bound *Eco*RI at [poly(dI-dC)·poly(dI-dC)] = 0.

With these simplifications,

$$F = (1 - f)\left(\frac{f_0}{f} - 1\right) = \frac{1}{K_{comp}}\frac{[DNA_{nonsp}]_{total}}{[DNA_{sp}]_{total}}, \quad [3]$$

K_{comp} can be extracted from the slope of F versus $[DNA_{nonsp}]_{total}$, while maintaining constant specific site DNA and protein concentrations. The total concentration of nonspecific sites is essentially the base pair concentration of poly(dI-dC)·poly(dI-dC).

RESULTS

Fig. 1 shows a typical gel mobility-shift assay illustrating the competition for *Eco*RI binding between nonspecific poly(dI-dC)·poly(dI-dC) and a 322-bp DNA fragment containing the specific *Eco*RI recognition sequence and the effect on the competition of added triethylene glycol. The fraction of specific site DNA fragment complexed with *Eco*RI decreases with increasing competitor polynucleotide concentration. With increasing triethylene glycol concentration, significantly more poly(dI-dC)·poly(dI-dC) is necessary to achieve the same level of competition as that achieved without added solute.

The ability of poly(dI-dC)·poly(dI-dC) to compete with specific site DNA for *Eco*RI binding depends on the ratio of specific and nonspecific binding constants. The competitive binding constant, K_{comp} ($K_{comp} = K_{sp}/K_{nonsp}$), can be extracted from the change in the fraction, f, of specific complex with increasing nonspecific competitor DNA concentration using the expression shown in Eq. **3**. The upper curve in Fig. 2 shows the variation of the parameter F, $F = [f_0/(f - 1)]\ (1 - f)$ with poly(dI-dC)·poly(dI-dC) concentration. The ratio of binding constants in the absence of osmolyte, K^0_{comp}, is determined from the slope as $(2.8 \pm 0.4) \times 10^4$. This ratio is somewhat larger than the $\approx 1 \times 10^4$ value given by Lesser and coworkers (6), measuring *Eco*RI binding constants to various oligonucleotides sequences by a filter binding assay. The effect of triethylene glycol is to increase K_{comp}. Compared with no added osmolyte, K_{comp} is about 4.5 times larger in 0.6 molal triethylene glycol (0.62 osmolal) and about 11 times larger in 1 molal triethylene glycol (1.05 osmolal).

The free energy change for transferring *Eco*RI from nonspecific competitor to the specific recognition site is $RT\ \ln(K_{comp})$, where RT is thermal energy. The changes in transfer free energy with added solute, $RT\ \ln(K_{comp}/K^0_{comp})$, are shown

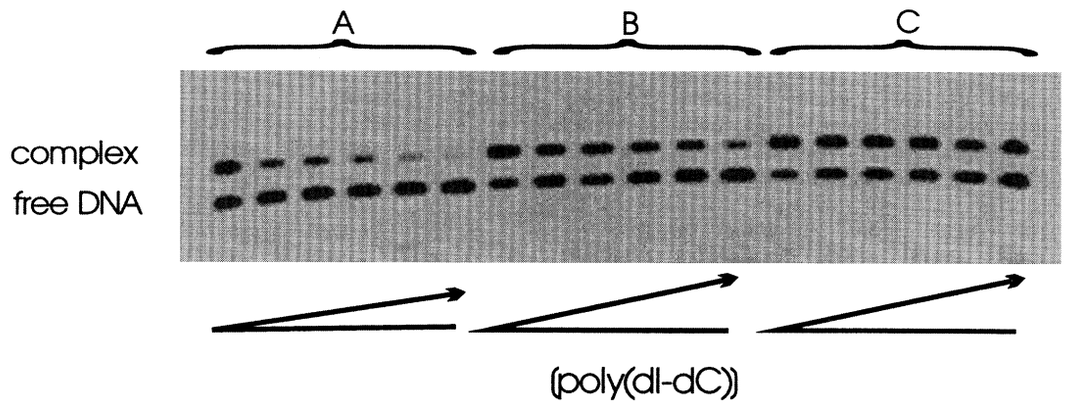

FIG. 1. poly(dI-dC)·poly(dI-dC) competes with a DNA fragment containing the specific recognition sequence for *Eco*RI binding. With increasing concentrations of neutral solutes, the ability of polynucleotide to compete decreases significantly. The decrease in binding of the restriction nuclease *Eco*RI to a 322-bp fragment containing its recognition site with increasing nonspecific poly(dI-dC)·poly(dI-dC) concentration is monitored by the gel mobility-shift assay as a function of triethylene glycol concentration. Mixtures of *Eco*RI, 322-bp fragment, and varying concentrations of poly(dI-dC)·poly(dI-dC) were incubated at 25°C for 30 min in the absence of Mg^{2+}, which is required for cleavage. Competition experiments at three osmotic pressures are shown: no osmolyte added (*A*), 0.6 molal triethylene glycol (0.62 osmolal), and 1 molal triethylene glycol (1.05 osmolal) (*C*). The leftmost lane in each series corresponds to no poly(dI-dC)·poly(dI-dC) added. Competitor polynucleotide concentration increases from 0 to 70 μM bp in *A*, and from 0 to 175 μM bp in *B* and *C*.

in Fig. 3 as a function of osmolal concentration for several chemically distinct solutes. There are two important features to note. Changes in competitive binding free energies scale linearly with changes in water chemical potential. This is important for distinguishing an indirect osmotic effect from direct solute binding. Second, there is practically no difference among the several neutral solutes: betaine glycine, sucrose, glycerol, triethylene glycol, glycine, and methyl glucoside.

The change in the number of solute excluding water molecules associated with transfer of bound *Eco*RI from poly(dI-dC)·poly(dI-dC) to its recognition sequence, Δn_w, can be determined from the slope of the best-fitting line to the data in Fig. 3 as described in Parsegian and coworkers (5),

$$\frac{RT\, d\ln(K_{comp}/K^0_{comp})}{d[\text{osmolal}]} = -\frac{RT\Delta n_w}{55.6}. \qquad [4]$$

The best fit to all the data gives $\Delta n_w = -110 \pm 15$. The negative sign means that the nonspecific complex sequesters more water than the specific one. The variation among the individual

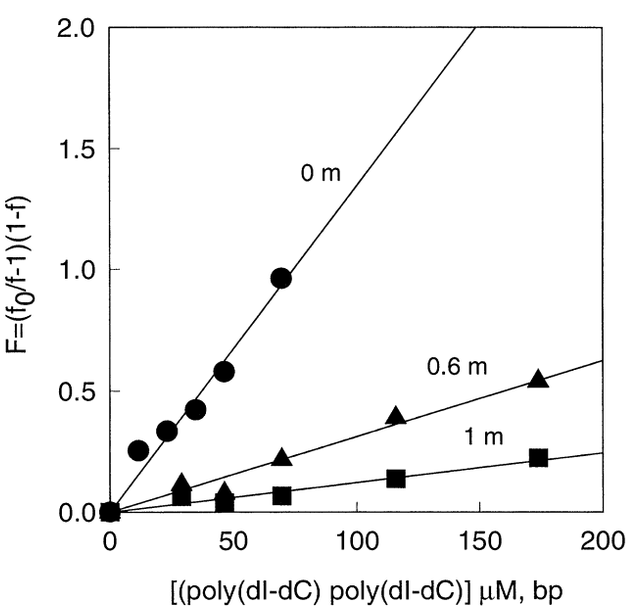

FIG. 2. The ratio of specific and nonspecific DNA binding constants ($K_{comp} = K_{sp}/K_{nonsp}$) is extracted from the ability of poly(dI-dC)·poly(dI-dC) to compete with the specific fragment for *Eco*RI binding. The parameter $F = [f_0/(f-1)](1-f)$ is shown plotted against polynucleotide concentration, where f is the fraction of specific fragment with bound *Eco*RI in the presence of nonspecific competitor and f_0 is this fraction with [poly(dI-dC)·poly(dI-dC)] = 0. The slope of the best-fitting straight line is $1/(K_{comp}\cdot[DNA_{sp}]_{total})$. For 0 (●), 0.6 (▲), and 1.0 molal (■) triethylene glycol, $K_{comp} = (2.8 \pm 0.4) \times 10^4$, $(1.2 \pm 0.2) \times 10^5$, and $(3.1 \pm 0.2) \times 10^5$, respectively.

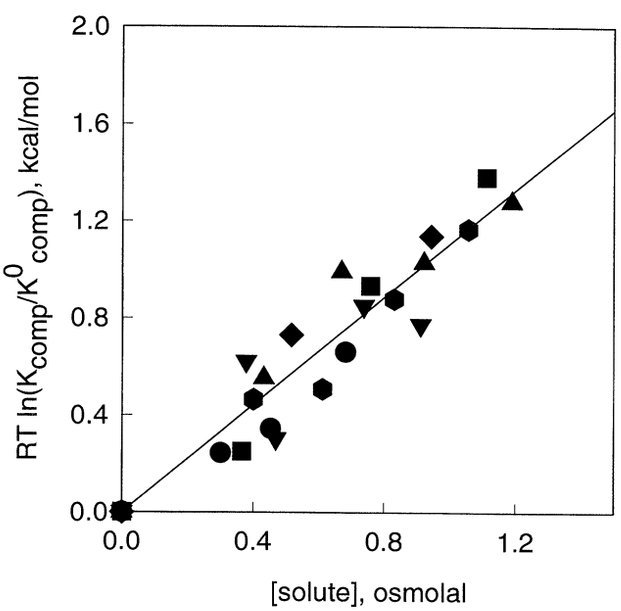

FIG. 3. The dependence of *Eco*RI transfer free energies, $RT\ln(K_{comp}/K^0_{comp})$, on solute osmolal concentration is shown for several solutes. ●, Glycerol; ▲, betaine; ■, sucrose; ◆, triethylene glycol; ●, methyl glucoside; ▼, glycine. Each point is the average of 2–4 separate experiments. The average errors for each point did not exceed 15%. Competitive binding free energies scale linearly with water chemical potential. The slope of the lines translates into a release of about 110 ± 15 water molecules for the transfer of *Eco*RI from poly(dI-dC)·poly(dI-dC) to the specific site. The slight dependence of Δn_w on the solute identity indicates a steric exclusion of these solutes from a well-defined, water-filled space, most probably associated with the nonspecific complex.

solutes is from $\Delta n_w = -90 \pm 15$ for glycerol to -120 ± 20 for triethylene glycol. These solutes span a wide range of chemical and physical properties and sizes. Glycine and betaine glycine are zwitterionic and increase the dielectric constant of the bulk solution. All the rest are uncharged and lower the bulk dielectric constant by varying degrees. Molecular weights and volumes range from 75 for glycine to 342 for sucrose.

DISCUSSION

An osmotic action of solutes on DNA–protein competition reactions requires a difference in the exclusion of solute from the water associated with specific and nonspecific DNA–protein complexes. Osmotic stress drives reactions toward those conformations or structures that exclude the least amount of water from solutes. An exclusion can result either from sterics (the solutes are too large to enter water-filled spaces) or from a net "preferential hydration" (reviewed in refs. 12–14) (exposed protein and DNA surfaces on average prefer their interactions with water over those with solutes). A difference in steric exclusion due to changes in the sizes of water-filled cavities inaccessible to solutes between two states will not depend either on solute size (once large enough to be excluded) or on its chemical nature, as seen, for example, for the opening and closing reactions of membrane incorporated channels (15, 16) and the oxygenation of hemoglobin (17).

Preferential interactions of solutes and water with proteins and nucleic acids have been extensively investigated (12–14). Direct measurement of solute–macromolecule interactions has shown both exclusion of solute from some regions of surfaces through a preferential hydration and inclusion of solute to other areas through apparent weak binding. The magnitude of the inclusion/exclusion effect on macromolecular energetics depends on both the size and the chemical nature of both the solute and the macromolecule. The effect of preferential interactions on conformational transitions or binding reactions depends only on the differences in solute inclusion and exclusion between the two states. Any preferential interaction that does not change between the two states of a macromolecular reaction does not contribute to a change in the equilibrium constant. Energy changes due to differences between two states in solute exclusion and in the weak binding of solutes can still scale linearly with osmotic pressure or, equivalently at low concentrations, solute activity. In this case, the slope of free energy change versus osmotic pressure can still be used to define a change in the effective number of solute-excluding waters between states. For preferential interactions, this strictly operational definition, however, can be strongly dependent on solute type and should not be taken to mean that only changes in solute exclusion are occurring.

Previously, Garner and Rau (11) measured the release of water associated with binding of free *Escherichia coli* galactose operon repressor from the bulk solution to its operator sequences. The osmotic sensitivity of the operator binding of repressor from solution showed a significant dependence on the chemical nature of the solute. The number of waters released in the binding of *gal* repressor from bulk solution to operator sequence varied from ≈100 for betaine glycine and 130 for sucrose to 180 for triethylene glycol. At least part of the osmotic stress seems due to preferential hydration exclusion of solutes from the exposed, complementary surfaces of free protein and DNA. An even stronger dependence on solute size and chemical nature is seen in the apparent osmotic sensitivity of the B → Z transition of poly(dG-m^{5}dC) with polyols and alcohols (18). The apparent Δn_w for sucrose, for example, is about 3–4 times larger than for glycerol.

The insensitivity of Δn_w for the transfer of *Eco*RI from poly(dI-dC)·poly(dI-dC) to its recognition sequence to solute size and nature is consistent with a difference in exclusion of solutes from 110 water molecules that is strictly steric. Since the crystal structure of the specific *Eco*RI–DNA complex (19, 20) shows essentially anhydrous contact between the two surfaces, these 110 waters are sequestered in the nonspecific complex, sterically inaccessible to solute, probably at the protein–DNA interface.

Fig. 4 shows a schematic representation illustrating the different classes of water that exclude solute and are coupled with DNA–protein complex formation. The surfaces of free protein and DNA are well-exposed to the bulk solution. Net exclusion of solute is likely dominated by preferential hydration and, therefore, sensitive to solute size and chemical nature. As seen in many crystal structures, the interface between protein and DNA in specific complexes has little or no water remaining. The remaining exposed surface area, however, still excludes solute by a preferential hydration. The effect of solutes on the specific binding of protein free in solution is, therefore, expected to be dominated by changes in the exposed surface area that preferentially excludes solute. This is consistent, for example, with the results for the binding of free *gal* repressor to operator (11). The results for *Eco*RI–DNA complexes suggest that the contact area between protein and DNA in a nonspecific complex delimits a structurally well-defined volume of water sterically inaccessible to the several different solutes used here and that changes in solute exposed surface area between nonspecific and specific complexes are small in comparison. The difference in solute-inaccessible water observed between the specific and nonspecific complexes is a thermodynamic parameter that measures a structural difference in the closeness of contact between protein and DNA surfaces. If there are no changes in DNA or protein conformation of surfaces still exposed to bulk solution between the specific and nonspecific complexes, then preferential interactions will not contribute to changes in the competitive binding constant. The slight dependence of Δn_w on solute identity remaining could reflect these preferential interactions and small difference in solute-accessible surface area due either to a small difference in protein–DNA overlap areas or to a small change in protein or DNA conformation between the specific and nonspecific complexes.

The close contact area between *Eco*RI and a DNA oligomer containing its recognition sequence can be estimated as 750–800 Å^2 from the crystal structure (19, 20). If the area of a water molecule is taken as ≈10 Å^2, then a first hydration layer on these buried protein and DNA surfaces corresponds to ≈150-160 waters. If the nonspecific complex has about the same protein–DNA overlap area as the specific complex, then, to a first order approximation, the 110 more waters sequestered by the *Eco*RI–poly(dI-dC)·poly(dI-dC) complex than in the specific complex corresponds to almost a full complement of hydrating water at the interface, in contrast to the anhydrous interface of the specific complex.

The results of Garner and Rau (11) also suggested that the nonspecific complex of *gal* repressor retains the full hydration of the free protein and DNA. The difference in *gal* repressor binding to nonspecific poly(dI-dC)·poly(dI-dC) and specific operator sequences had a sensitivity to added sucrose consistent with Δn_w ≈-140, compared with −130 for the binding of free repressor also using sucrose. Ha and coworkers (21) suggested that the difference in curvature of log(K) versus log[salt] plots at high salt concentrations for the nonspecific and specific operator binding of *lac* repressor is due to differences in water release and an osmotic effect of excluded salt. Nonlinear fits to the data, incorporating changes in both salt and water binding, gave $\Delta n_w \approx -200$ for specific operator binding and ≈0 for nonspecific binding. A sensitivity to solute nature was not investigated.

A retention of primary hydration water in the nonspecific complex as a general feature of DNA–protein interactions would also be consistent with the absence of a significant change in heat capacity that generally accompanies nonspecific

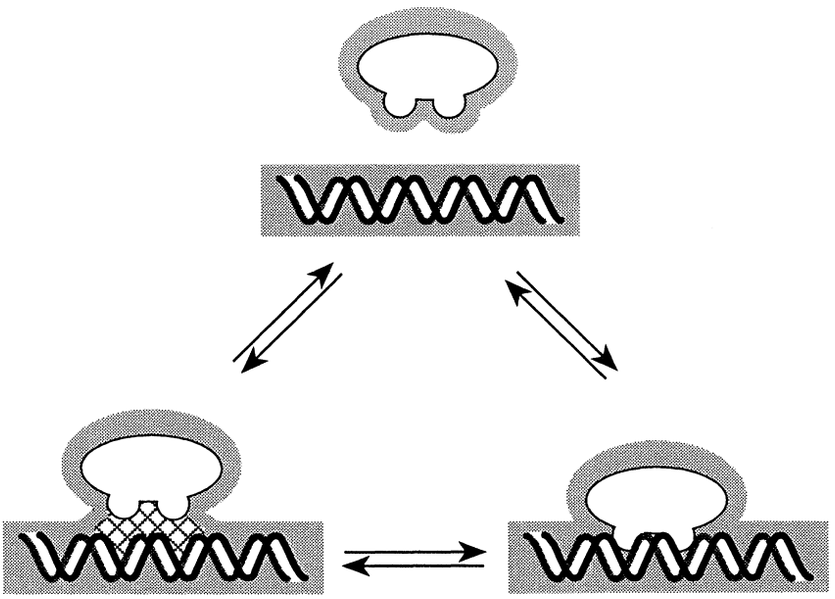

FIG. 4. A schematic representation of a protein–DNA recognition reaction is shown to illustrate the two classes of protein- and DNA-associated water that exclude solutes and can affect binding constants as water chemical potential or, equivalently, osmotic stress is changed. The protein is simply shown as globular with two lobes that represent, for example, helices that specifically interact with DNA and are responsible for sequence recognition. The free protein and DNA (*Upper*) primarily exclude solutes through a preferential hydration shown by the solid gray regions surrounding the protein and DNA surfaces. The extent of exclusion depends both on solute size and chemical nature and on the nature of the macromolecular surfaces. In the specific complex (*Lower Right*), the DNA and protein come into direct contact, decreasing the amount of water that excludes solute by preferential hydration. In addition to an exclusion by preferential hydration, the nonspecific *Eco*RI–DNA complex (*Lower Left*) also seems to have a volume of water presumably at the interface between surfaces, depicted by the crosshatched area, that sterically excludes solutes. Exclusion from this water will not depend on either solute size (once beyond some minimum) or chemical nature. The amount of sterically sequestered water is a measure of the closeness or directness of contact between protein and DNA surfaces.

binding, in contrast to the large changes seen with specific associations seen for *trp* (22) and λ cro (23) repressors. Large ΔC_p values have been postulated as reflecting the release of water structured at the protein and DNA surfaces to the bulk solution (23, 24).

It has long been known that, under nonstandard reaction conditions, *Eco*RI restriction nuclease is capable of cleaving sequences that are similar to but not identical with the canonical recognition site, termed "star" activity sites. The presence of neutral solutes, such as glycerol, dimethyl sulfoxide, ethanol, ethylene glycol, and sucrose, are among those solution conditions that promote this star activity. Recently, Robinson and Sligar (25, 26) showed that the increased digestion by *Eco*RI at star activity sequences caused by neutral solutes is correlated with water activity. The results seem to indicate that osmotic stress causes a loss in specificity of the enzyme activity, in contrast to the increase in binding specificity seen here. One interpretation of the star activity data consistent with the present experiments is that osmotic stress is modulating an equilibrium between a predominating nonspecific (water-mediated contact) and an energetically unfavorable, but enzymatically active, specific-like (direct protein–DNA contacts) modes of *Eco*RI binding to star activity sites.

We have restricted our experiments to comparatively low osmotic stresses and, therefore, comparatively small energy perturbations $[RT \ln(K_{comp}/K^0_{comp})]$ or a $\Pi\Delta V$ work $\leq 2RT \approx$ 1.2 kcal/mol. At high enough osmotic pressures, however, the energy gained by removing waters from the nonspecific complex will be of the same order as the unfavorable interaction energy between surfaces. The nonspecific complex will eventually dehydrate with increasing applied stress. This might be a potential problem for crystal structures of complexes with water incorporated at the interface. Typically, agents as polyethylene glycol or methylpentanediol are added at fairly high concentrations to enhance crystallization. Both polyethylene glycol and methylpentanediol are strongly excluded from polar protein and DNA surfaces and will apply an osmotic stress on the complex. The waters seen in the crystal structures of nonspecific complexes may only reflect a fraction of the water actually present under *in vivo* conditions.

Waters have been seen at the DNA–protein interface in crystal structures of several specific complexes, as, for example, the *trp* repressor/operator complex (27) and homeodomain–DNA complexes (28–30). Garner and Rau (11) additionally inferred from osmotic stress experiments that six water molecules are released in the transfer of *gal* repressor from a complex with one operator sequence to a second that has a factor of two larger binding constant. Comparing crystal structures of complex of steroid receptor-like DNA-binding domains with cognate and "noncognate" DNA sequences that differ by a factor of about 10 in binding constant shows that the noncognate complex has several more waters at the interface than the specific complex (4, 31). It is axiomatic that these waters are retained because the water-mediated protein–DNA interactions in these noncognate complexes are energetically more favorable than direct protein–DNA contacts. The retention of an approximately full hydration layer on the *Eco*RI and DNA surfaces in the nonspecific complex represents the extreme limit that few or no direct DNA–protein contacts are energetically preferred over the interactions with water. This suggests that there might be a correlation between the number of waters released in forming a complex and the binding free energy. In fact, a direct connection between molecular interactions and water structuring has been inferred from osmotic stress experiments on macroscopic ordered arrays of macromolecules. The measurement of forces (1) between many biopolymers (DNA, collagen, lipid bilayers, and several polysaccharides) demonstrate that classical double-layer electrostatics and van der Waals interactions, conventionally considered important, are simply not significant at close spacing (≤10-15 Å between surfaces) compared with the energies experimentally observed. The observed forces, both attractive

and repulsive, between a wide variety of polar surfaces (charged, net neutral, or even totally uncharged) are strikingly similar. The common characteristics led to the conclusion that the interaction between surfaces at close spacing is dominated by energies associated with the structuring of the intervening water. The water release parameter, Δn_w, provides a convenient link between binding thermodynamics and structural changes that is necessary for further analysis of the molecular forces underlying sequence specific recognition.

We thank Drs. V. A. Parsegian and H. Strey for helpful and enlightening conversations and discussions. We also thank Drs. M. Hodoscek and R. Podgornik for calculating the contact area between *Eco*RI and DNA from the crystal structure.

1. Leikin, S., Parsegian, V. A. & Rau, D. C. (1993) *Annu. Rev. Phys. Chem.* **44,** 369–395.
2. Steitz, T. A. (1990) *Q. Rev. Biophys.* **23,** 205–280.
3. Pabo, C. O. & Sauer, R. T. (1992) *Annu. Rev. Biochem.* **61,** 1053–1095.
4. Gewirth, D. T. & Sigler, P. B. (1995) *Nat. Struct. Biol.* **2,** 386–394.
5. Parsegian, V. A., Rand, R. P. & Rau, D. C. (1995) *Methods Enzymol.* **259,** 43–94.
6. Lesser, D. R., Kurpiewski, M. R. & Jen-Jacobson, L. (1990) *Science* **250,** 776–786.
7. Garner, M. (1992) *Adv. Electrophor.* **5,** 137–157.
8. Fried, M. G. (1989) *Electrophoresis* **10,** 366–376.
9. Sidorova, N. Y., Gazoni, P. & Rau, D. C. (1995) *J. Biomol. Struct. Dyn.* **13,** 367–385.
10. Lin, S.-Y. & Riggs, A. D. (1972) *J. Mol. Biol.* **72,** 671–690.
11. Garner, M. M. & Rau, D. C. (1995) *EMBO J.* **14,** 1257–1263.
12. Timasheff, S. N. (1992) *Biochemistry* **31,** 9857–9864.
13. Timasheff, S. N. (1993) *Annu. Rev. Biophys. Biomol. Struct.* **22,** 67–97.
14. Eisenberg, H. (1994) *Biophys. Chem.* **53,** 57–68.
15. Zimmerberg, J. & Parsegian, V. A. (1986) *Nature (London)* **323,** 36–39.
16. Bezrukov, S. M., Vodyanoy, I. & Parsegian, V. A. (1994) *Nature (London)* **370,** 279–281.
17. Colombo, M. F., Rau, D. C. & Parsegian, V. A. (1992) *Science* **256,** 655–659.
18. Preisler, R. S., Chen, H.-H., Colombo, M. C., Choe, Y., Short, B. J. & Rau, D. C. (1995) *Biochemistry* **34,** 14400–14407.
19. McClarin, J. A., Frederick, C. A., Wang, B.-C., Greene, P., Boyer, H. W., Grable, J. & Rosenberg, J. M. (1986) *Science* **234,** 1526–1541.
20. Rosenberg, J. M. (1991) *Curr. Opin. Struct. Biol.* **1,** 104–113.
21. Ha, J.-H., Capp, M. W., Hohenwalter, M. D., Baskerville, M. & Record, M. T., Jr. (1992) *J. Mol. Biol.* **228,** 252–264.
22. Ladbury, J. E., Wright, J. G., Sturtevant, J. M. & Sigler, P. B. (1994) *J. Mol. Biol.* **238,** 669–681.
23. Takeda, Y., Ross, P. D. & Mudd, C. P. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 8180–8184.
24. Spolar, R. S. & Record, T. M., Jr. (1994) *Science* **263,** 777–784.
25. Robinson, C. R. & Sligar, S. G.(1993) *J. Mol. Biol.* **24,** 302–306.
26. Robinson, C. R. & Sligar, S. G. (1995) *Methods Enzymol.* **259,** 395–427.
27. Otwinowski, Z., Schevitz, R. W., Zhang, R.-G., Lawson, C. L., Joachimiak, A., Marmorstein, R. Q., Luisi, B. & Sigler, P. B. (1988) *Nature (London)* **335,** 321–329.
28. Wilson, D. S., Guenther, B., Desplan, C. & Kuriyan, J. (1995) *Cell* **82,** 709–719.
29. Li, T., Stark, M. R., Johnson, A. D. & Wolberger, C. (1995) *Science* **270,** 262–269.
30. Hirsh, J. A. & Aggarwal, A. K. (1995) *EMBO J.* **14,** 6280–6291.
31. Arbuckle, N. D. & Luisi, B. (1995) *Nat. Struct. Biol.* **2,** 341–346.