The Role of Water in the EcoRI–DNA Binding

N. Sidorova, D.C. Rau

1 Introduction

In many respects, Type II restriction endonucleases are prototypical DNA-binding proteins. In order to avoid catastrophic consequences for the cell, however, these enzymes must be far more stringent in recognition of their target sequences and subsequent DNA cleavage than other specific sequence recognition proteins that regulate gene activity. In contrast to *E. coli* Lac and λ Cro repressors, for example, that show gradually decreasing binding energies as the recognition sequence is changed (Frank et al. 1997; Takeda et al. 1992), many restriction nucleases are exquisitely specific. EcoRI will bind to its recognition sequence, GAATTC, with an association equilibrium constant $K_{a,sp}$ ~10^{11} M^{-1} and to a completely nonspecific sequence with $K_{a,nonsp}$ ~10^{7} M^{-1}. A change of even a single base pair is sufficient to decrease the binding constant at least by 10^3, bringing it within a factor ~10 or less of nonspecific binding (Lesser et al. 1990).

To understand the physical basis of this specificity it is necessary both to know the structures of the complexes and to understand the energetics of molecular interactions. There has been an explosion of DNA–protein structures, in general, and of restriction nuclease–DNA complexes, in particular, solved by X-ray crystallography and NMR spectroscopy. There has not been a comparable increase in our ability to calculate binding energies from these structures. In particular, hydration energies are known to play important role in determining binding energies, but quantitating their contribution is still problematic. X-ray structures have uncovered many waters buried at protein–DNA interfaces that mediate interactions (Janin 1999), but energetic significance of these waters is unclear. The link between structure and the ener-

N. Sidorova, D.C. Rau
Laboratory of Physical and Structural Biology, National Institute of Child Health and Human Development, National Institutes of Health, Bld. 9, Room 1E108, Bethesda, Maryland 20892, USA

Nucleic Acids and Molecular Biology, Vol. 14
Alfred Pingoud (Ed.)
Restriction Endonucleases
© Springer-Verlag Berlin Heidelberg 2004

getics of intermolecular interactions is thermodynamics. The goal of this chapter is to begin correlating differences in binding energies to differences in hydration between complexes. The well-known linkage relationships that connect changes in binding energy to changes in salt concentration or pH through differences in ion binding or protonation will be extended to water activity or osmotic pressure.

We focus in this chapter on DNA complexes of the restriction endonuclease EcoRI, as a model system for delineating the role of water, in particular, in specific recognition. We compare directly specific and nonspecific binding properties, rather than simply investigating specific binding only, for two reasons. Binding specificity of protein only has meaning as a comparison of binding energies to different DNA sequences. On a more practical level, the pertinent equilibrium for specific sequence DNA-binding proteins within the DNA rich cellular environment is likely not between free and bound proteins, but between specifically and nonspecifically bound proteins. It should also be recognized that the water activity within the crowded cellular environment is not the same as in dilute aqueous solutions typically used to measure specific binding and that measuring the sensitivity of binding to water activity likely has practical applications for understanding in vivo action.

We show that nonspecific complexes of EcoRI sequester about 110 water molecules more than with the specific recognition sequence complex (Sidorova and Rau 1996). At low osmotic pressures this amount of water is seen even with complexes of EcoRI with DNA sequences that differ by a single base pair change ('star' sites) from the recognition sequence, consistent with the stringent binding specificity of this enzyme (Sidorova and Rau 1999). Much of the water sequestered by these 'star' sequence complexes, but not by other nonspecific sequence complexes, is removed at high osmotic pressures.

By combining equilibrium results with measurements of the dissociation rate of the specific sequence complex (Sidorova and Rau 2000, 2001) we are able to differentiate the effects of salt, pH, and water on the nonspecific–specific binding equilibrium and on the rate of dissociation of nonspecifically bound enzyme. The osmotic dependence of the dissociation rate constant of EcoRI from its recognition sequence is dominated by the 110 waters difference between specific and nonspecific binding modes. The dissociation of nonspecifically bound protein is accompanied by the uptake of much fewer waters. In contrast, there is very little salt and pH difference between specific and nonspecific modes of EcoRI binding. Nearly all of the dependence of the overall dissociation rate on salt and pH is coupled to the dissociation rate of the nonspecifically bound protein from the DNA.

The osmotic stress technique has now been used to measure the changes in hydration accompanying the DNA binding of several other proteins: *E. coli* Gal repressor (Garner and Rau 1995), *E. coli* CAP protein (Vossen et al. 1997), *E. coli* Lac repressor (Fried et al. 2002), Hin recombinase (Robinson and Sligar

1996), ultrabithorax and deformed homeodomains (Li and Mathews 1997), *E. coli* Tyr repressor (Poon et al. 1997), Sso7d protein (Lundback et al. 1998), TBP (Wu et al. 2001; Khrapunov and Brenowitz, pers. comm.).

2 Thermodynamics

It is by now standard practice when characterizing the binding of DNA recognition proteins to measure the sensitivity to salt concentration. The electrostatic interactions between the phosphate groups on the highly charged DNA backbone and basic amino acids of the protein in the complex are typically seen as the release of salt ions accompanying binding. The commonly used expression relating the association binding constant, K_a, salt concentration, [NaCl], and the difference in the number of associated ions between the complex and the free DNA and protein, ΔN_{NaCl}, is (e.g., Record et al. 1998),

$$\frac{d\ln(K_a)}{d\ln([NaCl])} = \Delta N_{NaCl} \tag{1}$$

Since there may be other conformational changes in the DNA or protein coupled to binding that may bind or release additional salt ions, ΔN_{NaCl} is not necessarily a direct measure of the number of DNA–protein ion pairs formed in the complex. Linkage equations just like the above can be written for any solution component such that net changes in, for example, metal ion or ligand binding or protonation accompanying the formation of a protein–DNA complex can also be measured. The ΔN values extracted from these linkage relations are the differences in the binding of these solution components between products and reactants.

Since many waters hydrating both the protein and DNA surfaces are typically displaced in forming the recognition interface of a specific complex (cf. Fig. 1), water itself should also be considered an important solution component. As with salt, the numbers of waters coupled to a reaction can be determined from the sensitivity of the equilibrium constant to water 'concentration' or activity. Although it is generally thought that 'water concentrations' (~55.6 M in dilute solutions) or activities change too little in ordinary solutions to affect reactions, the numbers of waters associated with macromolecular binding reactions are typically large enough, however, to cause significant effects. In order to change water concentration or activity a solute (osmolyte) must be added to the solution. This other component, of course, must necessarily not itself bind to the DNA or protein. Rather than speaking of a 'water concentration' ($[H_2O]$), however, it is more appropriate to use osmotic pressures or osmolal concentrations of solutes as measured by a vapor pressure osmometer, for example. For many commonly used solutes, osmolal concentrations are nearly the same (to within ~20 %) as osmolyte

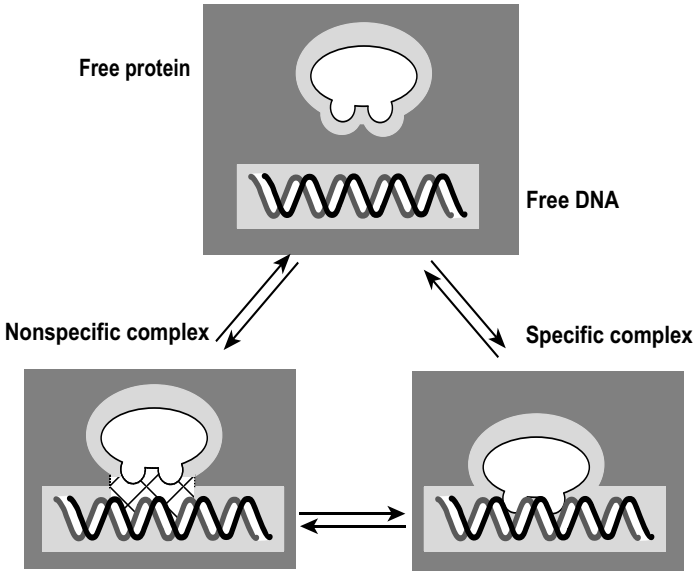

Fig. 1. Exclusion of water from protein, DNA, and their complexes. A schematic representation of a protein–DNA recognition reaction is shown to illustrate the two classes of protein- and DNA-associated water that exclude solutes and can affect binding constants as water chemical potential (osmotic stress) is changed. The protein is shown as globular with two lobes that represent helices that, e.g. specifically interact with DNA bases and are responsible for sequence recognition. Bulk solution (water and solute) is shown in *dark gray*. The free protein and DNA primarily exclude solutes from exposed surfaces through preferential hydration or crowding mechanisms. The *light gray* regions surrounding the protein and DNA surfaces represent a zone of osmolyte exclusion. The extent of solute exclusion (or water inclusion) from this zone will depend on the size and nature of the osmolyte probing the surface. In the specific complex (*lower right*), the DNA and protein come into direct contact, decreasing the amount of water that excludes solutes by preferential hydration. In addition to an exclusion by preferential hydration, the nonspecific EcoRI–DNA complex can also have a volume of water presumably in a cavity at the interface between surfaces, depicted by the *cross-hatched area* that sterically excludes solutes. Exclusion from this water will not depend on either solute size (after some minimum) or chemical nature

molal (mol/1 kg water) concentrations through ~1 molal. In terms of osmotic pressures expressed as osmolal concentrations,

$$\frac{d\ln(K_a)}{d[\text{osmolal}]} = -\frac{\Delta N_w}{55.6} \tag{2}$$

A more detailed discussion of linkage equations applied to water can be found in Parsegian et al. (1995, 2000).

There are generally two classes of waters illustrated for a DNA-binding protein (Fig. 1) that can be probed by adding osmolytes and changing water

activity. Water sequestered in pockets, grooves, or cavities are often sterically inaccessible to solutes (the cross-hatched area of the nonspecific complex in Fig. 1). In this case, the size or chemical nature of the added solute does not matter. There is simply osmotic pressure acting on a volume of water.

For water hydrating protein and DNA surface areas that are exposed to the bulk solution (the light gray areas in Fig. 1), however, the number of waters probed depends on the competition between water and solute for interaction with the macromolecules and so varies with the size and chemical nature of the osmolyte. Two mechanisms are commonly considered. Crowding (Minton 1998) recognizes that there is a steric exclusion of large solutes from surfaces; osmolytes simply cannot approach as closely as water. Preferential hydration (Timasheff 1993, 1998) further recognizes that the interaction of water with groups on protein and nucleic acids surface may be energetically more favorable than with osmolyte. Many experiments measuring both exclusion of solutes from protein and DNA surfaces and the effect of this exclusion on macromolecular reactions show that the apparent number or change in the number of hydrating waters is constant over a wide range of solute concentrations for each osmolyte, but that this number is dependent on the particular solute probing surface hydration (e.g., Timasheff 1993; Courtenay et al. 2000; Davis-Searles et al., 2001). For the osmolytes we typically use, a total range of about three to five fold difference in exclusion is commonly observed. If a wide variety of osmolytes are examined, then the sensitivity to the solute nature can be used to distinguish changes in numbers of waters sequestered in pockets and cavities from changes in exposed surface area accompanying binding reactions.

Although we initially considered a simple binding reaction, linkage expressions such as Eqs. (1) and (2) can be written for any reaction and can even be applied to reaction rates. The competition reaction between specific and nonspecific DNA sequences for protein binding is $DNA_{nonsp} \cdot Protein + DNA_{sp} \Leftrightarrow DNA_{nonsp} + DNA_{sp} \cdot Protein$, with an equilibrium constant $K_{nonsp\text{-}sp}$. The osmotic pressure or salt concentration sensitivity, $d\ln(K_{nonsp\text{-}sp})]/d[osmolal]$ or $d[\ln(K_{nonsp\text{-}sp})]/d[\ln[NaCl]]$, gives the difference in 'bound' waters or salts: $(N_{nonspDNA} + N_{spComplex}) - (N_{spDNA} + N_{nonspComplex})$, where N is the total number of water molecules or salt ions associated with the reaction components. Since salt binding and solute exclusion from DNA is dominated by the sugar–phosphate backbone with very little contribution from the particular sequence, $N_{nonspDNA} \sim N_{spDNA}$ for both salt and water, leaving only the difference between specific and nonspecific complexes. It should be noted that free protein does not contribute to $K_{nonsp\text{-}sp}$. Competition experiments can either be done by measuring the change in specific binding constant as specific sequence DNA is titrated with protein with and without added competitor DNA or by measuring the loss of specific binding as specific DNA–protein complexes are titrated with competitor DNA.

Linkages relations can also be written for rate processes (Lohman 1985). If we consider the dissociation of a specifically bound protein, the osmotic or salt dependence of the rate, $d[\ln(k_d)]/d[\text{osmolal}]$ or $d[\ln(k_d)]/d[\ln[\text{NaCl}]]$, gives the difference in bound water or salt between the transition state and the specific complex. The transition state is defined for the rate-limiting kinetic step in dissociation as the high-energy structure from which it is about equally probable to rebind as to dissociate. Rate constants may also include a contribution from the osmolyte to the solution viscosity depending on whether the energy dissipation of the rate-limiting kinetic step is dominated by solvent friction or by the internal friction of the complex.

Many restriction nucleases, as EcoRI, are able to slide quite efficiently along DNA bound nonspecifically to facilitate finding recognition sequences (Wright et al. 1999; Stanford et al. 2000; Pingoud and Jeltsch 2001). The rate-limiting step for dissociation of EcoRI from DNA is the dissociation of the nonspecific complex. To a first order approximation, the overall dissociation rate, k_d, is the product of the dissociation rate of the nonspecific complex, $k_{d,nonsp}$, and the equilibrium constant for the specific to nonspecific complex reaction, $K_{nonsp\text{-}sp}$. The osmotic sensitivity (or salt or pH) of the dissociation can consequently be divided into these two steps,

$$\frac{d\ln(k_d)}{d[\text{Osmolal}]} = \frac{d\ln(k_{d,nonsp})}{d[\text{Osmolal}]} - \frac{d\ln(K_{nonsp\text{–}sp})}{d[\text{Osmolal}]} = -\frac{\Delta N_{w,d,nonsp}}{55.6} + \frac{\Delta N_{w,nonsp?sp}}{55.6} \qquad (3)$$

If solution friction limits the dissociation rate, k_d can be corrected for viscosity. A number of waters linked to the dissociation rate of a nonspecifically bound protein, $k_{d,nonsp}$, can be calculated from the osmotic dependence of the specific-nonspecific binding equilibrium and of the overall dissociation rate, k_d.

Finally, it must be recognized that with added osmolytes sequestered waters are under a pressure. If there are alternate conformations of nonspecific complexes, for example, that sequester less water, then these states will be stabilized by osmotic stress. Loss of water in nonspecific complexes will be observed as a decrease in ΔN_w between specific and nonspecific complexes as the osmotic pressure increases. The energy of removing water from these nonspecific complexes can be calculated from pressure–volume ($\Pi\Delta V$) work accompanying this loss of water. It is important, however, that proper controls are performed to ensure that an apparent change in ΔN_w is actually due to water loss and not to the many other possible effects of osmolytes that may occur.

3 Experimental Applications

3.1 Equilibrium Competition

3.1.1 Osmotic Stress Dependence or Knonsp-sp

Figure 2a shows a typical gel mobility-shift assay illustrating the competition for EcoRI binding between nonspecific poly(dI-dC) · poly(dI-dC) and a 322 bp DNA fragment containing the specific EcoRI recognition sequence, GAATTC, in the presence of different concentrations of triethylene glycol. In each series, EcoRI protein concentration is held constant and the amount of competitor DNA increased. The fraction of specific complex decreases with increasing competitor DNA concentration. Significantly more competitor DNA is required at 1 molal triethylene glycol to reach the same level of competition than in the absence of solute.

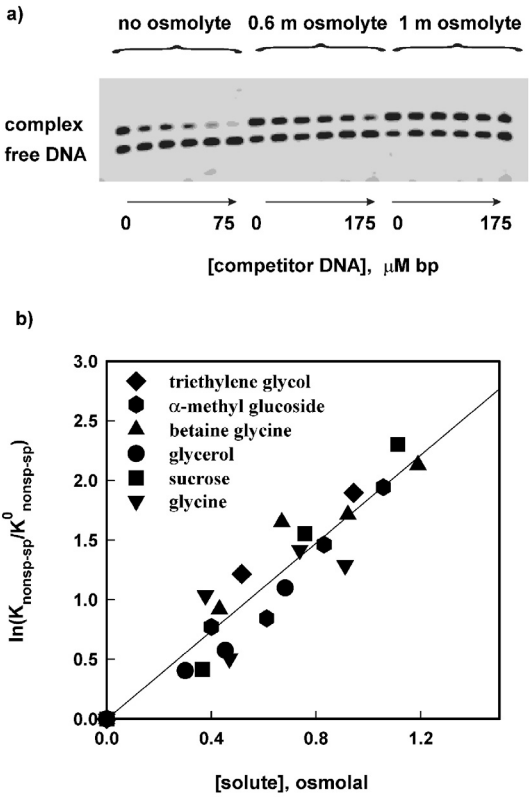

Fig. 2. Osmotic pressures favor specific sequence binding of EcoRI. **a** Poly(dI-dC)Epoly(dI-dC) competes with a DNA fragment containing the specific recognition sequence for EcoRI binding. With increasing concentrations of neutral solutes, the ability of nonspecific polynucleotide to compete is significantly diminished. Competition experiments at three osmotic pressures are shown: no osmolyte added, 0.6 molal triethyleneglycol (0.62 osmolal), and 1 molal triethyleneglycol (1.05 osmolal). No divalent ion is present in order to avoid the enzymatic cleavage reaction. **b** The dependence of free energy difference between specific and nonspecific EcoRI–DNA complexes, $RT\ln(K_{nonsp\text{-}sp}/K^0_{nonsp\text{-}sp})$, on solute osmolal concentration is shown for several solutes. Competitive binding free energies scale linearly with osmotic pressure. The slope of the best fitting line translates into a difference of about 110±15 water molecules between specific and nonspecific EcoRI–DNA complexes

The ability of nonspecific polynucleotide to compete with specific DNA site for EcoRI binding depends of course on the ratio of specific and nonspecific binding constants ($K_{nonsp\text{-}sp}$). In the absence of osmolyte, $K^0_{nonsp\text{-}sp}$ was determined as $(2.8\pm0.4)\times10^4$. Compared with no added osmolyte, $K_{nonsp\text{-}sp}$ is about 4.5 times larger in 0.6 molal triethylene glycol and about 11 times larger in 1 molal triethylene glycol. The difference in the binding free energy of the EcoRI to specific and nonspecific DNA sequences is $RT\ln(K_{nonsp\text{-}sp})$. The dependence of free energy difference (divided by factor RT=0.6 kcal/mole) on concentration of added solute is shown in Fig. 2b as a function of osmolal concentration for several chemically distinct solutes. Two most important features in this graph to note are: (1) changes in competitive binding free energies scale linearly with changes in osmolal concentration or, equivalently, water chemical potential; (2) there is practically no difference among the several neutral solutes: betaine glycine, sucrose, glycerol, triethylene glycol, glycine, and α-methyl glucoside.

The difference in the number of solute excluding water molecules between specific and nonspecific EcoRI–DNA complexes, ΔN_w, can be determined from the slope of the best-fitting line to the data in Fig. 2b as specified by Eq. (2). The best fit to all data gives $\Delta N_w=-110\pm15$, indicating that the nonspecific complex sequesters about 110 waters more than the specific one. The observed insensitivity of ΔN_w to solute size and nature strongly suggests that observed difference in exclusion of solutes from 110 water molecules is strictly steric. Since the crystal structure of the specific EcoRI–DNA complex (McClarin et al. 1986; Kim et al. 1990) shows essentially anhydrous contact between DNA and protein surfaces, these 110 waters are likely sequestered at the protein–DNA interface of the nonspecific complex (see cartoon in Fig. 1).

3.1.2 pH and Salt Dependence of Knonsp-sp

In contrast to the very strong dependence of specific binding of free EcoRI on salt concentration corresponding to the release of 10–12 ions (Jen-Jacobson 1997), only a small sensitivity to salt is seen for the equilibrium between specific and nonspecific binding of the EcoRI using the competitive binding assay. Between 90 and 160 mM NaCl the competitive binding constant of the EcoRI to specific versus nonspecific DNA sequences increases by only 50 % (Sidorova and Rau 2001). The salt dependence of $K_{nonsp\text{-}sp}$ on salt concentration can be analyzed either as a difference in ion binding reflecting a difference in DNA–protein charge interactions between specific and nonspecific complexes or as an indirect, osmotic effect of salt. The data over this limited salt concentration range are insufficiently precise to distinguish between these alternatives. If analyzed as a change in direct salt binding, the data are consistent with a release of an additional 0.6 ion in forming the nonspecific complex from the specific one. This small increase could be due to closer interactions of nonspecifically bound protein charge with DNA phosphate

groups on the backbone. This dependence is, however, more likely due to the osmotic contribution of salt. As long as salt is excluded from the same water-filled cavities as the neutral solutes, an osmotic effect is required. If salt acts osmotically on the equilibrium between specific versus nonspecific binding with the same ΔN_w=–110 molecules as for the neutral solutes, then the osmotic contribution would be equivalent to the release of ~ 0.5 ion over the salt concentration range examined, very close to the experimentally measured value of 0.6 ions. Several other DNA-binding proteins also show slightly more ions coupled to nonspecific binding than to specific binding (Lohman 1985; Record and Spolar 1990; Jen-Jacobson 1997). It seems probable that this general behavior is a reflection of a general osmotic action of salt on differences in water sequestered between specific and nonspecific complexes rather than an electrostatic effect as is commonly assumed.

Similarly, even though the equilibrium constant for specific binding of free EcoRI shows a strong dependence on pH (Jen-Jacobsen et al. 1983), there is practically no effect of pH sensitivity on the competitive specific and nonspecific sequence binding equilibrium of EcoRI.

3.2 Dissociation Kinetics of EcoRI from its Specific Site

Although most thermodynamic work has focused on equilibrium constants for DNA–protein complex formation, studies of dissociation rates (k_d) of DNA–protein complexes are also extensive. Dissociation kinetics are important not only for understanding reaction rates of nucleases, ligases, polymerases, and repair enzymes, for example, but also because the binding of regulatory proteins to their target DNA sequences within a cell may be kinetically controlled and not an equilibrium reaction. Often the equilibrium constants and dissociation rates show similar dependences on salt concentration and pH (Lohman 1985). Differences, however, can lead to a better formulation of the detailed binding scheme. In particular, the dissociation of many DNA–protein complexes, including the specific EcoRI–DNA complex, occurs in two steps (Lohman 1985; Pingoud and Jeltsch 1997, 2001). There is a steady-state reaction between specific and nonspecific binding of the protein to the DNA. Nonspecifically bound protein can linearly diffuse along the DNA and can either return to the specific binding site or eventually dissociate from the DNA. Differences in the sensitivities of relative specific–nonspecific equilibrium binding constants and dissociation rates on solution conditions (salt concentration, pH, water activity, etc.) can distinguish between factors that are in common and that are different for specific and nonspecific binding. The total number of water molecules, salt ions or protons seen in the dissociation reaction of the EcoRI from its specific sequence can be separated into the contributions from the differences in the numbers of associated water molecules, ions or protons between the nonspecifically bound complex and the transi-

tion state for protein dissociation, and between EcoRI specifically and non-specifically bound complexes, respectively, as given in Eq. (3).

3.2.1 Osmotic Dependence of k_d

The dissociation of EcoRI from its specific site on DNA is well described by a single exponential and the rate of dissociation is sensitive to osmolyte concentration. Just as the dependence of an equilibrium binding constant on salt or water activity gives the difference in the number of ions or water molecules associated with products and reactants, the sensitivity of a rate constant to salt or water activities is determined from the difference in the number of ions or water molecules associated with the initial state and the transition state of the rate-limiting kinetic step. The sensitivity of the dissociation rate on water activity, for example, is determined by the difference in the numbers of water molecules associated with the initial specific complex of the enzyme with DNA and the transition state of the rate-limiting kinetic step of dissociation. Figure 3 shows plots of $\ln(k_d)$ versus solute osmolal concentration for a wide variety of solutes. Plots are linear for each solute (including the no osmolyte limit). All of the solutes are closely similar in their ability to slow down dissociation. There is, however, more solute-specific variation than we observed previously for $K_{nonsp\text{-}sp}$. The number of waters linked to the dissociation of the EcoRI from its specific site varies between 100 water molecules for sucrose and 155 for triethylene glycol and TMAO. These variations (50 % at the most) are still much less than the factor differences of two to five expected for an osmotic effect based solely on exclusion of solutes from exposed surface

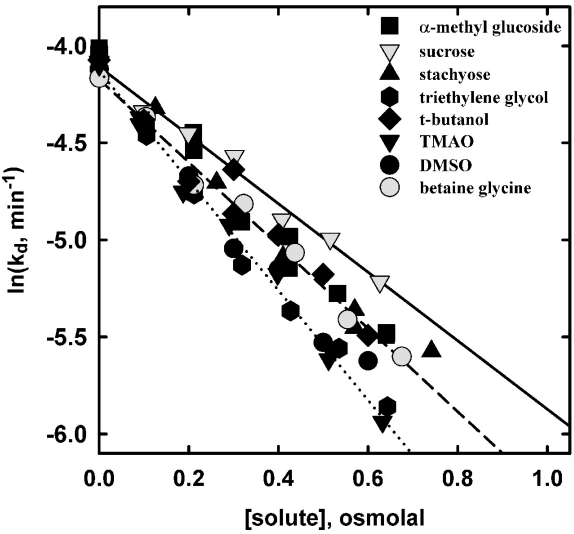

Fig. 3. Dependence of the dissociation rate on osmotic pressure. The dependence of $\ln(k_d)$ on solute osmolal concentration for the EcoRI for a wide variety of osmolytes. The slopes of the lines translate into 100±6 waters for sucrose, 120±8 for stachyose, 125±6 for a-methyl glucoside, 120±10 for *t*-butanol, 120±6 for betaine glycine, 155±8 for triethylene glycol, 145±8 for dimethyl sulfoxide (DMSO), and 155±4 for trimethylamine N-oxide (TMAO)

areas. Of these total waters coupled to EcoRI dissociation, 110 water molecules accompany the specific – nonspecific complex equilibrium. For the seven solutes other than sucrose the remaining number of waters accompanying the dissociation of nonspecifically bound protein is positive and varies from 10 to 45 extra water molecules. Given this dependence on solute identity, this reaction step is very likely characterized by a change in solute-accessible surface area. For sucrose this number was slightly negative –10±20) water molecules. This represents either no exclusion from or a small preferential inclusion of sucrose with the newly exposed surface area. Adjusting dissociation rates for the solution viscosity gave physically unrealistic results (Sidorova and Rau 2001), suggesting that the rate of dissociation of nonspecifically bound protein is limited by an internal friction rather than solution viscosity.

3.2.2 pH and Salt Dependence of k_d

The dissociation rate is coupled to a pH titration. Over the range of pH between 6.2 and 9.2, k_d increases by about 40-fold (Sidorova and Rau 2001). Since we observe virtually no pH-dependence of the equilibrium between specific and nonspecific binding measured by our competition assay, pH affects only the dissociation of the nonspecifically bound protein from the DNA. The pH-dependence can be fit adequately by assuming two titrating groups with identical pK values. The number of water molecules coupled to the rate of dissociation is insensitive to pH. Water activity and pH are acting independently to affect binding and dissociation.

Similarly, the dissociation rate is strongly dependent on salt concentration (Jen-Jacobson et al. 1986; Lesser et al. 1993; Sidorova and Rau 2001). We have measured with binding of 5.8±0.5 ions linked to the EcoRI dissociation from its specific site. Since the salt concentration dependence of the specific to nonspecific binding reaction is equivalent to the release of ~0.6 ions (see Sect. 3.1.2), ~6.4 ions are linked to the dissociation of nonspecifically bound EcoRI. Not surprisingly, pH and salt are coupled. The net binding of two extra Na^+ ions linked to the dissociation of nonspecifically bound enzyme accompanies the protonation of the two extra protein groups between pH 9.2 and 6.2. In contrast to the osmotic dependence, both salt and pH predominantly affect the dissociation of nonspecifically bound protein, not the equilibrium between specific and nonspecific binding.

These results clearly demonstrate the importance of water compared with salt and pH in distinguishing specific and nonspecific EcoRI binding. The osmotic dependence is primarily due to differences between specific and nonspecific binding of the EcoRI. Both salt and pH sensitivities are mostly due to the nonspecific binding of free protein and do not distinguish between specific and nonspecific binding. Our results also suggest that osmotic stress might be a convenient way to increase stability and lifetime of weak com-

plexes for separation, chemical modification, or measurement of physical properties. Since the numbers of water molecules coupled with the dissociation of nonspecifically bound EcoRI are quite solute dependent then the ability to stabilize weak nonspecific complexes will also differ among osmolytes. Among all solutes used triethylene glycol, TMAO, and DMSO would be the most effective for stabilizing nonspecific EcoRI complexes. For example, half-life of a nonspecific complex should increase about 3 times between 0 and 1 molal triethylene glycol. In contrast, sucrose either would not influence the stability of a nonspecific complex or might even slightly destabilize it.

3.3 Removing Water from an EcoRI–Noncognate DNA Complex with Osmotic Stress

The equilibrium competition experiments discussed in Section 3.1.1 were restricted to comparatively low osmotic stresses and, therefore, small energy perturbations that are less than about 2RT or 1.2 kcal/mol. The linear dependence of the free energy difference on osmolality (Fig. 2b) shows that difference in the number of water molecules sequestered by nonspecific versus specific complex remains the same at least up to 1.2 osmolal. At high enough osmotic pressures, however, the pressure–volume or $\Pi\Delta V$ work gained by removing waters from a nonspecific complex will be comparable to the unfavorable interaction energy incurred in forcing closer contact between non-complementary surfaces. In principle, any sequestered water can be removed by applying high enough osmotic stress, but the work necessary to dehydrate complexes will obviously depend on the resulting DNA–protein contacts and complex structure.

It has long been known that EcoRI restriction endonuclease as well as many other Type II restriction endonucleases are capable of cleaving sequences that are similar to but not identical with the canonical recognition sequence, 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 ‘star’ activity. It has been shown by Robinson and Sligar (1993) that the increased digestion by EcoRI at ‘star’ sequences caused by neutral solutes is strictly correlated with water activity. One possible explanation for the effect of solutes on the ‘star’ activity is that osmotic stress modulates an equilibrium between a predominating nonspecific (water-mediated contact) and an energetically unfavorable, but enzymatically active, specific-like (direct protein–DNA contacts) modes of EcoRI binding to ‘star’ sequences.

3.3.1 Competitive Equilibrium at High Osmotic Stress

The general strategy is the same as described in Section 3.1.1. Comparative binding constants of EcoRI to different DNA sequences are measured by a competition assay. The loss of EcoRI binding to a DNA fragment containing its specific recognition sequence as the concentration of a competing oligonucleotide increases is measured using the gel mobility shift assay. Figure 4a shows the dependence on the osmolal concentration of betaine glycine of the relative binding free energies of EcoRI to two 30 bp oligonucleotides, differing only in that one contains a central 'star' sequence TAATTC and the other nonspecific oligonucleotide contains the inverted recognition sequence CTTAAG instead. The 'star' sequence oligonucleotide binds EcoRI only about twofold more strongly than the inverted sequence, nonspecific oligonucleotide. As seen from the slopes of the curves in Fig. 4a at low pressures, both oligonucleotide complexes sequester some 110 more waters than the specific sequence one. Even a single base pair change from the recognition sequence is sufficient to trigger the nonspecific-binding mode of EcoRI. Above ~2 osmolal the slope of the plot for the TAATTC 'star' sequence oligonucleotide complex is clearly smaller than at low pressures suggesting a loss of

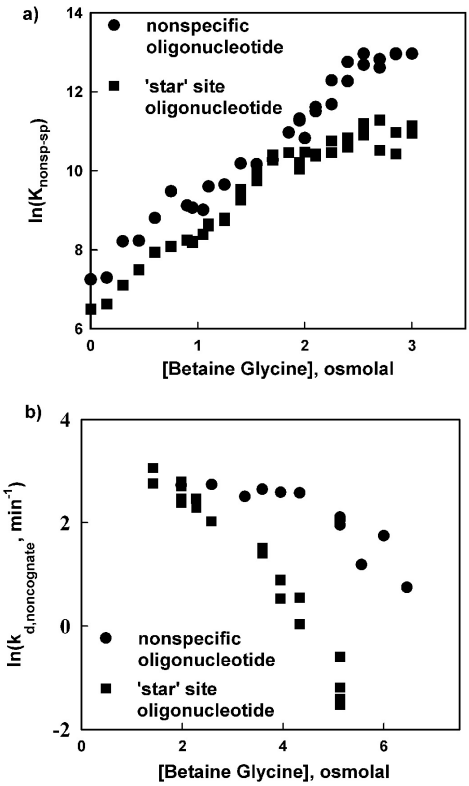

Fig. 4. Removing water at high osmotic pressures. **a** The binding energies, $\ln(K_{nonsp\text{-}sp})$, for two oligonucleotides relative to the specific complex is shown as a function of betaine glycine osmolal concentration. $K_{nonsp\text{-}sp}$ is the ratio between EcoRI binding constants to specific and competitor noncognate sequences measured in the experiments analogous to the one shown in Fig. 2a. The 'star' sequence competitor contains the sequence TAATTC. The nonspecific, inverted sequence oligonucleotide has CTTAAG instead. The two oligonucleotides are otherwise identical. The 'star' sequence complex seems to lose water at high osmotic pressures. **b** The dependence of $\ln(k_d)$ for the EcoRI dissociation from the two noncognate oligonucleotides, 'star' sequence and nonspecific, on betaine glycine osmolal pressure is shown. Significantly more water can be removed from the 'star' sequence complex compared with the nonspecific, inverted sequence one

water from the 'star' sequence complex with increasing osmotic stress. The average slope in the 2–3 osmolal range corresponds to 70 waters. No loss of water is seen for the complex with the nonspecific, inverted sequence oligonucleotide.

3.3.2 Dissociation Kinetics of the EcoRI from Noncognate Sites

The equilibrium experiments at high osmotic stress are difficult since dissociation rates are so very slow and the competitor DNA concentrations necessary to observe competition are very high. Alternatively, the osmotic dependence of the dissociation rate of 'star' and nonspecific DNA sequence complexes can be conveniently measured to very high osmotic stresses. The dissociation rate from noncognate DNA sequence complexes can be determined by adding EcoRI to a mixture of specific sequence DNA fragment and noncognate oligonucleotide and measuring the time course for specific binding. Under conditions of fast association, the appearance of specific complex depends on three factors: (1) the ratio of specific and nonspecific association rate constants, (2) the concentrations of specific sequence fragment and noncognate oligonucleotide, and (3) the nonspecific complex dissociation rate. The osmotic dependences of the EcoRI dissociation rate constant from nonspecific and 'star' sequence oligonucleotide complexes are shown in Fig. 4b.

At relatively low osmotic pressures there is only a small difference between two oligonucleotides as expected from equilibrium competition experiments (Fig. 4a). The slope at osmotic pressures <5 osm for the nonspecific sequence complex corresponds to ~10 waters, very close to the value inferred for the dissociation of nonspecifically bound protein from the specific complex dissociation rate measurements (Sect. 3.2.1). EcoRI dissociation from the 'star' site (TAATTC) oligonucleotide is obviously much more sensitive to osmotic stress. In the 2–3 osmolal range, the slope corresponds to ~40 waters or only ~80 waters left in the 'star' sequence complex, in reasonable agreement with the estimate from equilibrium experiments in the same pressure range (Sect. 3.3.1). At even higher pressures between 4 and 5 osmolal the osmotic sensitivity of $k_{d,nonsp}$ translates into uptake of about 100 waters for this noncognate complex, leaving only ~20 waters in the complex. Over the same high range of pressures the nonspecific, inverted sequence complex may also be losing some water, the slope giving ~25 waters or about 95 waters remaining in the complex. The precision of experiments, both equilibrium and kinetics, is not high enough to distinguish between a two state model, i.e., between a 'star' sequence complex with ~110 sequestered waters and a second discrete binding mode with much less associated water, or a continuum of states, i.e., a gradual loss of water from the 'star' complex. The approximate $\Pi\Delta V$ work for removing ~90 waters from the 'star' sequence complex is about 4 kcal/mol.

3.4 Other Applications of Osmotic Stress to Restriction Nucleases

Robinson and Sligar (1998) have measured the osmotic sensitivities of nonspecific and specific EcoRI binding separately using a single osmolyte, ethylene glycol. Since these reactions are characterized by a large change in solution accessible surface area, apparent hydration numbers should depend on the solute size and nature. A difference of only ~70 waters (a release of 146 waters accompanying specific binding and 76 waters coupled to nonspecific complex formation) was observed rather than the 110 found using the competition assay and many more solutes (Sect. 3.1.1). Additionally, Lynch and Sligar (2000) report that the specific binding of BamHI was accompanied by the release of only about 20 waters. No osmotic dependence for nonspecific binding was reported.

The osmotic stress approach has been used to measure a number of waters coupled to enzymatic rates, k_{cat} or V_{max}. Robinson and Sligar (1998) find an uptake of ~28 coupled to the turnover rate of EcoRI; Lynch and Sligar (2000) report almost +90 waters linked to the BamHI turnover; and Wenner and Bloomfield (1999) measure 28 waters for EcoRV. Since these reaction rates were measured under conditions of multiple turnover it is not clear what step in the reaction osmotic stress is probing, cleavage or subsequent dissociation.

Robinson and Sligar (1993, 1995b) have reported that the 'star' activities not only of EcoRI (as already noted in Sect. 3.3), but also of PvuII and BamHI correlate well with osmotic stress for many different solutes. The 'star' activity of EcoRV, however, is not sensitive to osmolytes.

3.5 Application of Hydrostatic Pressure to Restriction Nucleases

Hydrostatic pressure is another probe of water structuring that has been applied to restriction endonuclease binding and cleavage kinetics, primarily by Sligar and coworkers (e.g., Lynch and Sligar 2002). Whereas osmotic pressure favors those species that have fewer numbers of waters that exclude solutes, hydrostatic pressure stabilizes those species that have larger densities or, equivalently, smaller volumes (e.g., Silva et al. 2001). If, for example, the hydrating water of DNA and protein that is released in forming a complex is more dense than bulk, then, all else remaining constant, hydrostatic pressure promotes complex dissociation, as observed for specific BamHI binding (Lynch et al. 2002). The stabilization energy is $P\Delta V$, where ΔV is the solution volume change. If the average molecular volume of bulk water is $\bar{v}_w^0$ and of the N_w released hydrating waters $\bar{v}_w$, then $DV = N_w (\bar{v}_w - \bar{v}_w^0)$. Of course, protein and DNA conformational changes may also have accompanying density changes that will be sensitive to hydrostatic pressure. We have already noted that Sligar and coworkers found that the 'star' activity of a number of restriction nucleases (EcoRI, PvuII, and BamHI) is enhanced by osmotic stress.

Hydrostatic pressure acts to reverse the effect (Robinson and Sligar 1995a, b). Either hydrostatic pressure simply inhibits the turnover reaction or, more interestingly, alters the energetics of removing water from 'star' sequence complexes. The density of water in the DNA–protein cavity could be much different from bulk water.

4 Summary

No structure is available for the nonspecific EcoRI complex to compare with the specific sequence complex structure (McClarin et al. 1986; Kim et al. 1990) to confirm a water cavity at the protein–DNA interface. Structures of specific and nonspecific complexes of a closely similar restriction nuclease BamHI, however, have been reported (Newman et al. 1995, Viadiu and Aggarval 2000). The cavity at the protein–DNA interface of the nonspecific complex has a volume of 4763 $Å^3$, compared with only 282 $Å^3$ for the specific sequence complex. Assuming a typical volume of 30 $Å^3$ per water molecule, the difference in cavity sizes corresponds to 150 waters comparable to the 110 waters we find for EcoRI. This 'loose', 'water-lubricated' association of DNA within a binding cleft of the protein is a conformation that would easily allow linear diffusion of the protein along the DNA.

We see that ion release and proton binding are simply a result of nonspecific EcoRI binding. There is no further sensitivity to pH or salt (except for an apparent osmotic contribution from salt) between specific and nonspecific modes of binding. There is, however, a large difference in water. If indeed this water is at the DNA–protein interface of the nonspecific complex as suggested by the structure of the nonspecific BamHI complex, then 110 waters corresponds to ~1.5 hydration layers. The noncognate DNA and protein surfaces prefer to keep their hydration interactions, suggesting that water does play an important role in recognition. This is also consistent with heat capacity measurements. The formation of many specific DNA–protein complexes, including EcoRI (Ha et al. 1989), is accompanied by a large change in heat capacity. A large portion of this change seems to come from the release of hydration waters that are structured differently from bulk water (Spolar and Record 1994). In contrast, only very small heat capacity changes accompany the formation of several nonspecific complexes, consistent with retention of hydrating waters. The waters at the noncognate DNA–protein interface can only be removed with great difficulty. Indeed, we have only been able to remove water from EcoRI complexes with DNA sequences that differ by only one base pair from the recognition and even then the energy required is quite significant, ~4 kcal/mol.

Direct measurements of forces between macromolecules in condensed arrays indicates that water structuring forces do seem to dominate interactions at close spacings, the last 10–15 Å of separation between surfaces (Leikin

et al. 1993). The key to understanding the binding strength and specificity of restriction endonuclease–DNA interactions is in understanding the energetics of the hydration interactions that must be replaced in forming a complex.

References

Courtenay ES, Capp MW, Anderson CF, Record MT Jr (2000) Vapor pressure osmometry studies of osmolyte-protein interactions: implications for the action of osmoprotectants in vivo and for the interpretation of "osmotic stress" experiments in vitro. Biochemistry 39:4455–4471

Davis-Searles PR, Saunders AJ, Erie DA, Winzor DJ, Pielak GJ (2001) Interpreting the effects of small uncharged solutes on protein-folding equilibria. Annu Rev Biophys Biomol Struct 30:271–306

Frank DE, Saecker RM, Bond JP, Capp MW, Tsodikov OV, Melcher SE, Levandovski MM, Record MT Jr (1997) Thermodynamic of the interactions of lac repressor with variants of the symmetric lac operator: effects of converting a consensus site to a nonspecific site. J Mol Biol 267:1186–1206

Fried MG, Stickle DF, Smirnakis KV, Adams C, MacDonald D, Lu P (2002) Role of hydration in the binding of lac repressor to DNA. J Biol Chem 277:50676–50682

Garner MM, Rau DC (1995) Water release associated with specific binding of gal repressor. EMBO J 14:1257–1263

Ha J-H, Spolar RS, Record MT Jr (1989) Role of the hydrophobic effect in stability of site-specific protein-DNA complexes. J Mol Biol 209:801–816

Janin J (1999) Wet and dry interfaces: the role of solvent in protein–protein and protein–DNA recognition. Structure 7:R277–R279

Jen-Jacobson L (1997) Protein–DNA recognition complexes: conservation of structure and binding energy in the transition state. Biopolymers 44:153–180

Jen-Jacobson L, Kurpiewski M, Lesser D, Grable J, Boyer HW, Rosenberg JM, Greene PJ (1983) Coordinate ion pair formation between EcoRI endonuclease and DNA. J Biol Chem 258:14638–14646

Jen-Jacobson L, Lesser D, Kupriewski M (1986) The enfolding arms of EcoRI endonuclease: role in DNA binding and cleavage. Cell 45:619–629

Kim YC, Grable JC, Love R, Greene PJ, Rosenberg JM (1990) Refinement of EcoRI endonuclease crystal structure: a revised protein change tracing. Science 249:1307–1309

Leikin S, Parsegian VA, Rau DC (1993) Hydration Forces. Annu Rev Phys Chem 44:369–395

Lesser DR, Kurpiewski MR, Jen-Jacobson L (1990) The energetic basis of specificity in the EcoRI endonuclease-DNA interaction. Science 250:776–786

Lesser D, Kurpiewski MR, Waters T, Connolly BA, Jen-Jacobson L (1993) Facilitated distortion of the DNA site enhances EcoRI endonuclease-DNA recognition. Proc Natl Acad Sci USA 90:7548–7552

Li L, Matthews KS (1997) Differences in water release with DNA binding by ultrabithorax and deformed homeodomains. Biochemistry 36:7003–7011

Lohman TM (1985) Kinetics of protein–nucleic acid interactions: use of salt effects to probe mechanisms of interaction. CRC Crit Rev Biochem 19:191–245

Lundback T, Hansson H, Knapp S, Ladenstein R, Hard T (1998) Thermodynamic characterization of non-sequence-specific DNA-binding by the Sso7d protein from *Sulfolobus solfataricus*. J Mol Biol 276:775–786

Lynch TW, Sligar SG (2000) Macromolecular hydratoion changes associated with BamHI binding and catalysis. J Biol Chem 275:30561–30565

Lynch TW, Sligar SG (2002) Experimental and theoretical high pressure strategies for investigating protein–nucleic acid assemblies. Biochim Biophys Acta 1595:277–282

Lynch TW, Kosztin D, McLean MA, Schulten K, Sligar SG (2002) Dissecting the molecular origins of specific protein–nucleic acid recognition: hydrostatic pressure and molecular dynamics. Biophys J 82:93–98

McClarin JA, Frederick CA, Wang BC, Greene P, Boyer HW, Grable J, Rosenberg JM (1986) Structure of the DNA–EcoRI endonuclease recognition complex at 3 Å resolution. Science 234 (4783):1526–1541

Minton AP (1998) Molecular crowding: analysis of effects of high concentrations of inert cosolutes on biochemical equilibria and rates in terms of volume exclusion. Methods Enzymol 295:127–149

Newman M., Strzelecka T., Dorner LF, Schildkraut I, Aggarwal AK (1995) Structure of BamHI endonuclease bound to DNA: partial folding and unfolding on DNA binding. Science 269:656–663

Parsegian VA, Rand RP, Rau DC (1995) Macromolecules and water: probing with osmotic stress. Methods Enzymol 259:43–94

Parsegian VA, Rand RP, Rau DC (2000) Osmotic stress, crowding, preferential hydration, and binding: a comparison of perspectives. Proc Natl Acad Sci USA 97:3987–3992

Pingoud A, Jeltsch A (1997) Recognition and cleavage of DNA by Type II restriction endonucleases. Eur J Biochem 246:1–22

Pingoud A, Jeltsch A (2001) Structure and function of Type II restriction endonucleases. Nucleic Acids Res 15:3705–3727

Poon J, Bailey M, Winzor DJ, Davidson BE, Sawyer WH (1997) Effects of molecular crowding on the interaction between DNA and the *Escherichia coli* regulatory protein TyrR. Biophys J 73:3257–3264

Record MT Jr, Spolar RS (1990) Some thermodynamic principles of nonspecific and site-specific protein–DNA interactions. In: Revzin A (ed) The biology of nonspecific DNA–protein interactions. CRC Press, Boca Raton, pp 33–69

Record MT Jr, Zhang W, Anderson CF (1998) 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. Adv Prot Chem 51:281–353

Robinson CR, Sligar SG (1993) Molecular recognition mediated by bound water. A mechanism for star activity of the restriction endonuclease EcoRI. J Mol Biol 234:301–306

Robinson CR, Sligar SG (1995a) Hydrostatic pressure reverses osmotic pressure effects on the specificity of EcoRI–DNA interactions. Method Enzymol 259:395–427

Robinson CR, Sligar SG (1995b) Heterogenity in molecular recognition by restriction endonucleases: osmotic and hydrostatic pressure effects on BamHI, PvuII, and EcoRV specificity. Proc Natl Acad Sci USA 92:3444–3448

Robinson CR, Sligar SG (1996) Participation of water in Hin recombinase–DNA recognition. Protein Sci 5:2119–2124

Robinson CR, Sligar SG (1998) Changes in solvation during DNA binding and cleavage are critical to altered specificity of the EcoRI endonuclease. Proc Natl Acad Sci USA 95:2186–2191

Sidorova N, Rau DC (1996) Differences in water release for the binding of EcoRI to specific and nonspecific DNA sequences. Proc Natl Acad Sci USA 93:12272–12277

Sidorova N, Rau DC (1999) Removing water from and EcoRI–noncognate DNA complex with osmotic stress. J Biomol Struct Dynam 17:19–31

Sidorova N, Rau DC (2000) The dissociation rate of the EcoRI–DNA specific complex is linked to water activity. Biopolymers 53:363–368

Sidorova N, Rau DC (2001) Linkage of EcoRI dissociation from its specific DNA recognition site to water activity, salt concentration, and pH: separating their roles in specific and non-specific binding. J Mol Biol 310:801–816

Silva JL, Fougel D, Royer CA (2001) Presure provides new insights into protein folding, dynamics and structure. Trends Biochem Sci 26:612–618

Spolar RS, Record MT, Jr. (1994) Coupling of local folding to site-specific binding of proteins to DNA. Science 263:777–784

Stanford NP, Szczelkun MD, Marko JF, Halford SE (2000) One- and three-dimensional pathways for proteins to reach specific DNA sites. EMBO J 19:6546–6557

Takeda Y, Ross PD, Mudd CP (1992) Thermodynamics of Cro protein–DNA interactions. Proc Natl Acad Sci USA 89:8180–8184

Timasheff SN (1993) The control of protein stability and association by weak interactions with water: how do solvents affect these processes. Annu Rev Biophys Biomol Struct 22: 27–65

Timasheff SN (1998) Control of protein stability and reactions by weakly interacting cosolvents: the simplicity of the complicated. Adv Protein Chem 51:355–432

Viadiu H, Aggarwal AK (2000) Structure of BamHI bound to nonspecific DNA: a model for DNA sliding. Mol Cell 5:889–895

Vossen KM, Wolz R, Daugherty MA, Fried MG (1997) Role of macromolecular hydration in the binding of the *Escherichia coli* cyclic AMP receptor to DNA. Biochemistry 36:11640–647

Wenner JR, Bloomfield VA (1999) Osmotic pressure effects on EcoRV cleavage and binding. J Biomol Struct Dynam 17:461–471

Wright DJ, Jack WE, Modrich P (1999) The kinetic mechanism of EcoRI endonuclease. J Biol Chem 274:31896–31902

Wu J, Parhurst KM, Powell RM, Brenowitz M, Parkhurst LJ (2001) DNA bends in TATA-binding protein–TATA complexes in solution are DNA sequence-dependent.