ELSEVIER

Biophysical Chemistry 62 (1996) 121–139

Biophysical Chemistry

Thermodynamics of sequence-specific protein–DNA interactions

Torleif Härd [a,*], Thomas Lundbäck [b]

[a] *Dept. of Biochemistry and Biotechnology, Royal Institute of Technology, Center for Structural Biochemistry, NOVUM, 141 57 Huddinge, Sweden*

[b] *Dept. of Biosciences, Karolinska Institutet, Center for Structural Biochemistry, NOVUM, 141 57 Huddinge, Sweden*

Received 4 April 1996; accepted 11 June 1996

Abstract

The molecular recognition processes in sequence-specific protein–DNA interactions are complex. The only feature common to all sequence-specific protein–DNA structures is a large interaction interface, which displays a high degree of complementarity in terms of shape, polarity and electrostatics. Many molecular mechanisms act in concert to form the specific interface. These include conformational changes in DNA and protein, dehydration of surfaces, reorganization of ion atmospheres, and changes in dynamics. Here we review the current understanding of how different mechanisms contribute to the thermodynamics of the binding equilibrium and the stabilizing effect of the different types of noncovalent interactions found in protein–DNA complexes. The relation to the thermodynamics of small molecule–DNA binding and protein folding is also briefly discussed.

Keywords: Thermodynamics; Molecular recognition; DNA–protein interactions; Sequence specificity; Noncovalent interactions

1. Introduction

Transcription, replication and recombination of genes rely on noncovalent interactions between DNA and DNA-binding proteins. These interactions therefore constitute one of the key physico-chemical links between genetic information and a functional living organism. A long-term objective of structural and biophysical studies of protein–DNA interactions is to understand the molecular basis for these processes.

This topic is closely related to many other areas in which the physical basis for molecular recognition in water solution is under intense study, for instance protein–drug interactions, and interactions in host–guest complexes of organic molecules. Several issues that are continually addressed in other areas, such as the contribution of a particular hydrogen bond to the overall binding free energy, also arise in studies of protein–DNA interactions. Nevertheless, several aspects distinguish protein–DNA interactions. These include the polyelectrolyte property of DNA, the relatively large surface areas that are buried in the complexes, and the fact that a specific binding site on DNA must be recognized against a very large background of nonspecific binding sites.

A large number of structures of protein–DNA complexes have been determined in the last few years using X-ray crystallographic and nuclear magnetic resonance (NMR) methods (see, for example, Refs. [1–4], and references cited therein). These

* Corresponding author.

PII S0301-4622(96)02197-7

structures show a very large variability in terms of DNA-interacting protein motifs, DNA conformation, hydrophobic vs. hydrophilic nature of interacting surfaces, bound water, cooperativity, and conformational changes coupled to binding. In fact, the only common feature of sequence-specific complexes appears to be large interacting surfaces which display complementarity in shape and polarity (see, for example, Ref. [1]), as well as electrostatic complementarity [5,6].

Some of the various molecular events that take place on the formation of a sequence-specific protein–DNA complex are illustrated in Fig. 1, which will serve as a basis for the discussion. The collective effect of the events that accompany binding is the formation of the complementary surfaces. For instance, DNA will often undergo a conformational change in order to accommodate the protein (described in Section 6). Coupled protein folding reactions can create an additional DNA-interacting surface, and unfolding or other structural changes in DNA-binding proteins have also been observed (Section 5). Similarly, the binding of an allosteric effector to the protein may change the protein structure to better fit DNA (Section 5). The extent of these structural changes can, at least qualitatively, be detected by comparing the structure of the complex with structures of free DNA and protein components. Further, water molecules are often coordinated at the interface to fill cavities or mediate intermolecular hydrogen bonds, and such water can also be detected in high-resolution structures (Section 4).

Several additional events take place which are more difficult to detect (and understand) based on structural data alone. For instance, a large number of water molecules that hydrate the surfaces of the uncomplexed DNA and protein components need to "leave" these surfaces so that the optimal interface can be formed (Section 4). The formation of a complex also results in extensive reorganization of the counterion atmospheres (Section 3). Finally, there may be changes in dynamics in the protein backbone and side chains as well as in the DNA upon complexation (Section 7).

All these events affect the thermodynamics of the equilibrium. They also influence the sequence specificity, i.e. the difference in binding free energy for the formation of sequence-specific compared to nonspecific complexes. All events make both entropic and enthalpic contributions to the binding free energy. Of these, entropic contributions are especially difficult to estimate based on structural information.

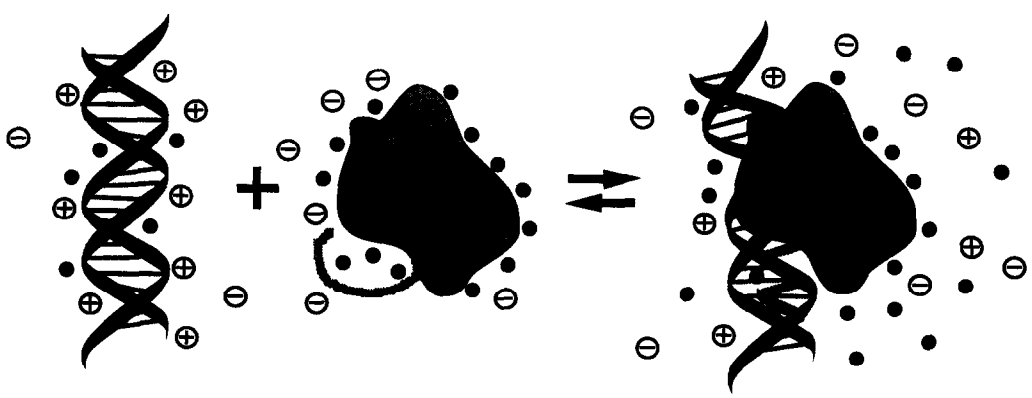

Fig. 1. Schematic illustration of some of the molecular events that determine the thermodynamics for the formation of a sequence-specific protein–DNA complex. The picture shows a (negatively-charged) DNA fragment which is hydrated and surrounded by an ion atmosphere predominantly composed of positively-charged counterions, and a (positively-charged) solvated protein with an associated ion atmosphere. The two components are in equilibrium with a complex in which the protein–DNA interface is highly complementary in terms of shape, polarity and electrostatics. Various events that contribute to the complementarity include changes in the DNA and protein conformations, a coupled protein folding reaction, dehydration of surfaces, entrapment of water molecules at the interface to mediate hydrogen bonds, reorganization of the ion atmospheres, and changes in the intramolecular dynamics and vibrational modes (not illustrated). ●, water molecules; ⊕, cations; ⊖, anions. Additional effects that may influence the thermodynamics include coupled binding equilibria involving allosteric effectors and cooperative binding of more than one protein molecule.

This fact motivates the use of thermodynamic studies to complement structural data for the elucidation of various molecular contributions to binding affinity and specificity.

The objective of the present review is to outline the thermodynamic determinants for the formation of sequence-specific protein–DNA complexes, and to give examples of how various effects are being addressed using thermodynamic methods and, in some cases, structure determinations. The literature on the subject is extensive and it is not possible to make a comprehensive survey of all contributions, or even all contributing laboratories, within the format of the present review. We have chosen to limit our discussion to brief outlines of the current status on different issues, and refer the reader to some of the previously published reviews on this subject. In order to put present studies of structure vs. thermodynamics in an appropriate historical perspective, we also include a short summary of initial ground-breaking research.

2. Historical perspective

The history of research on sequence-specific protein–DNA interactions dates back to 1961 with the formulation of the operon hypothesis [7], with the later establishment that the repressor components are proteins [8,9]. The finding that gene activity could be regulated at the level of transcription by direct interactions between proteins and DNA motivated further studies focused on the details of these interactions [10]. Before structural data became available, the lines of research developed around the questions of information content, possible mechanisms of recognition, and DNA target location and discrimination, as discussed by von Hippel [11].

The information content of binding sites on DNA and the statistical basis for recognition of a few specific sites in the presence of many nonspecific sites has been investigated (see, for example, Refs. [12,13] and work cited therein). The coding elements for the recognition of double-stranded DNA are restricted to the four Watson–Crick base pairs. A certain length of the binding site is therefore required to ensure a low probability that it occurs by a random combination of the same number of base pairs at some other nonfunctional site in the genome. If the number of proteins (repressors) is small, then this probability should, preferably, be less than unity. It was, for instance, concluded that a sequence of at least 12 base pairs is required to define a unique binding site within the *E. coli* genome.

Regarding the mechanisms of recognition, it was appreciated early that one possibility was provided by the pattern of hydrogen bond donors and acceptors displayed in the DNA grooves [14–16]. This pattern reflects the DNA sequence and may be recognized through a sufficiently precise matching pattern on the protein surface. A companion mechanism based on sequence-dependent differences in DNA structure and stability, which could be recognized by a DNA-binding protein, was also suggested [15]. Structural studies have essentially confirmed that these two mechanisms, which are known as direct and indirect readout, respectively, act in concert to form complementary interacting DNA and protein surfaces. However, subsequent structural and thermodynamic studies also illustrate the many possible variations of these basic principles and the importance of factors that were not initially recognized, such as dehydration, electrostatic complementarity, water-mediated hydrogen bonds, etc.

The mechanism by which DNA-binding proteins find their target sequences has been elucidated [17,18]. It was shown to involve two steps: the formation of an initial nonspecific complex followed by "sliding" and intersegment transfer events. The second step can be viewed as a search process within a reduced volume and leads to an accelerated apparent rate of target location. Comparisons with kinetic and equilibrium data showed that this mechanism quantitatively accounts for the high association rates between *lac* repressor and operator (see Ref. [18] and work cited therein). In addition, the sliding of *E. coli* RNA polymerase has recently been visualized directly [19].

The first physico-chemical studies of sequence-specific protein–DNA interactions in vitro were carried out in 1968 by Riggs and co-workers [20]. Encouraged by the work of Jones and Berg on polymerase–DNA complexes [21], these authors showed that the strong binding of *lac* repressor to membrane filters could be used to quantify the binding in vitro. The membrane filter technique turned

out to be very useful because it allows various effects to be examined quantitatively, e.g. the presence of competing nonspecific DNA or inducers such as IPTG, pH conditions, salt and osmolyte concentrations, and temperature. The sensitivity to salt concentration was recognized early [21,22]. Further, in 1972 Poliakow et al. [23] found that the binding of histones to DNA was accompanied by a large positive change in entropy, and argued that this was due to changes in hydration upon complex formation. The role of water and salts on protein–DNA equilibria is now highly appreciated, as discussed below. Methods used for the quantification of protein–DNA equilibria now include filter-binding [20], DNA footprint titrations [24,25], fluorescence intensity [26], fluorescence energy transfer and anisotropy [27,28], gel mobility shifts [29,30], gel chromatography [31], pulsed-flow calorimetry [32] and isothermal titration calorimetry [33,34].

3. Electrostatic interactions

The basis for the contribution of electrostatic effects to DNA–ligand interactions lies in the physical properties of the DNA polymer: DNA is a polyelectrolyte with a high axial charge density. As already noted, this property distinguishes DNA–ligand interactions from most other biomolecular recognition processes. The high charge density results in an accumulation of counterions close to the polyion surface. The radial distribution of counterions is relatively insensitive to changes in the total salt concentration. However, the thermodynamic activity of the polyion (resulting from interactions with counterions) shows an approximately logarithmic dependence on the total salt concentration. The activity is strongly dependent on the charge density. Any process that alters the charge density (e.g. ligand binding) is therefore dependent on the total salt concentration. (An extensive review of the theoretical and experimental studies of the polyelectrolyte activity of DNA was recently presented by Anderson and Record [35].)

Hence, the contribution of electrostatic interactions to the formation of protein–DNA complexes can be elucidated based on the dependence of binding constants on the salt concentration and composition, or pH, of the reaction buffer. The effects of these solution variables were interpreted and investigated in detail in the laboratory of Record [36–38]. Record and co-workers initially interpreted the salt dependence in terms of a participation of cations, anions and water molecules in the binding reaction (reviewed in Ref. [39]). This view, which has been commonly adopted, allows for quantification of the dependence of the intrinsic equilibrium constant, K_{obs}, on the total concentration of univalent salt, [MX], as

$$SK_{obs} \equiv \frac{\partial \log K_{obs}}{\partial \log[MX]} = \Delta c + \Delta a - \frac{2m\Delta w}{[H_2O]} \qquad (1)$$

where m is the molality of univalent salt, and the effects of differential protonation have been neglected. The coefficients Δc, Δa and Δw represent the stoichiometrically weighted differences between the preferential interactions of cations, anions and water, respectively, with the macromolecular reactants (DNA, protein and protein–DNA complex). SK_{obs} is often inversely proportional to the valence of the cation [37,40], which supports the inference that SK_{obs} is dominated by the release of cations from DNA, i.e. $\Delta c \ll 0$. Anion effects on SK_{obs} are often found to be smaller ($\Delta a \approx 0$) [40,41], but may not be negligible in the case of sequence-specific binding [42]. The term representing preferential hydration is expected to be small compared to the other terms at low total salt concentration (≤ 0.2 M) [43]. The reduction of SK_{obs} with increasing salt concentration (0.25–0.9 M potassium glutamate) has been used to probe the effect of dehydration for the *lac* repressor–operator interaction [42]. Limiting law expressions based on counterion condensation theory (Ref. [44] and work cited therein) have often been applied to interpret Δc in terms of a specific number of cations released in the binding process, e.g. in Refs. [36,39]. Experimental observations that may be quantified in a similar manner include binding equilibria in the presence of divalent ions [38], the binding of oligomeric cations such as polyamines or polylysines of different lengths [36,45,46], and the pH dependence of binding equilibria [37,45].

The presence of different anions in the buffer has a much more pronounced effect on the binding constant than on its salt concentration dependence [40–

42,47]. For instance, replacing Cl^- with acetate or glutamate can increase the sequence-specific binding affinity by a factor of 40-80, a stronger effect being observed with glutamate [42]. These observations suggest that anion effects are not primarily electrostatic in character. Various molecular mechanisms for the effect of some anions on protein–DNA equilibria, e.g. differential solvent exclusion, have been discussed [42,43].

A related treatment, in which a mass balance of ions bound to a protein is considered as the protein is brought into the vicinity of DNA [48], has been applied to account for the nonlinearity of plots of $\log K_{obs}$ vs. log[MX] that is often observed at low total salt concentrations [18,49,50]. Overman and Lohman recently established a strong linkage between pH, and anion and cation effects [51]. More importantly, they demonstrated that very complex equilibria, involving contributions from differential ion interactions with both DNA and protein, can be resolved by examining the linked effects of pH and salt concentration.

The expression in Eq. (1) was originally obtained using generalized binding polynomials, and by introducing the Gibbs–Duhem relation between water and salt activities [39]. Anderson and Record [52] recently presented a more general theoretical treatment based on preferential interaction coefficients [53], in which they employed approximations, that can be justified under experimental conditions, to derive an expression for the salt dependence. Their formulation allows for comparisons between experimental data and results from, for example, Monte Carlo simulations [54]. A two-domain (dialysis) model applied in a subsequent paper [55] leads to Eq. (1) without explicit application of the Gibbs–Duhem equation. It should be noted that the formulation in terms of preferential interaction coefficients does not involve any a priori assumptions with regard to the molecular mechanism of the salt effect, i.e. cation release. The two-domain model, however, allows for a physical interpretation in terms of differences in preferential accumulations of ions in the local domains around free polymers and the complex [55].

More detailed information on the molecular origin of the salt effect may be inferred based on finite-difference solutions to the nonlinear Poisson–Boltzmann (NLPB) equation for structures where dielectric boundaries and charges are represented at atomic resolution [6,56]. These studies show that the unfavorable reorganization of ion atmospheres upon complexation provides the predominant contribution to the salt effect on ligand–DNA binding equilibria. The NLPB calculations also show that the salt-dependent electrostatic contribution to the binding free energy is unfavorable in all cases studied. Further, the free energy associated with ion reorganization (cratic entropy), which is analogous to cation release, is only slightly favorable, or even unfavorable. Comparisons of calculations on λ cI repressor complexes with different binding sites illustrate how small variations in binding site geometry may result in significant differences in the interaction free energy, due to long-range electrostatic effects [6].

The cation exchange models for the salt effect implicate the association of salt dependence with a predominant effect on binding entropy [36]. Studies of nonspecific peptide–DNA [45,57,58] and protein–DNA [59] complexes provide experimental evidence for the predominance of entropic contributions to the salt effect. These observations are also consistent with NLPB calculations which, however, indicate that the entropy change is associated with solvent reorganization (dielectric effects) as well as ion reorganization [60].

It is important to note that the research reviewed in this section concerns salt-dependent contributions to the total electrostatic free energy of binding. However, little appears to be known about the total electrostatic contribution to the free energy of binding for a protein–DNA complex. Experimental and theoretical studies indicate that the total electrostatic free energy of binding of the small cationic DAPP ligand to DNA is favorable [61], but small owing to the compensation of the Coulombic attraction by nonfavorable dehydration effects [62].

Differences in salt dependence for protein binding to different specific DNA sites [5,63], for binding of protein mutants to a single binding site [49,64], and for specific vs. nonspecific binding of a protein [38,42], suggest that electrostatic effects contribute to binding site discrimination. For instance, data on *lac* repressor binding indicate that electrostatic and nonelectrostatic interactions have to be optimized simultaneously at an adaptable protein–DNA interface [5]. Calculations on the *trp* repressor suggest that a

negative electrostatic potential on the protein surface is neutralized in the correct DNA-binding conformation, presumably induced at the cognate DNA-binding site [65].

It is generally believed that long-range Coulombic interactions (rather than, for example, dehydration effects) dominate the formation of nonspecific protein–DNA complexes [13] and the initial steps of a sequence-specific association [65,66]. A corollary to this is that electrostatics is a predominant physical driving force for a large number of biochemical processes that rely on nonspecific protein–DNA associations. Nonspecific electrostatic interactions also play a key role in the kinetics (diffusion along DNA) of binding site recognition, as discussed above.

4. Bound water molecules and dehydration effects

The role of water in biomolecular interactions may be structural owing to specific interactions at certain binding sites or entrapment of water within cavities. Water also has a dominating effect on the thermodynamics due to dehydration of interacting surfaces. The interactions and physical properties of water can be studied by a number of experimental and theoretical approaches. Below, we limit our discussion to the branches of experimental research that have provided most information on the role of water in protein–DNA interactions: the observation of bound waters in structures, studies based on osmotic stress methods, the analysis of heat capacity changes upon binding, and studies of water sites and residence times in solution based on NMR cross relaxation.

4.1. Water molecules observed in crystallographic structures

Water molecules are commonly identified at specific binding sites in crystal structures of protein–DNA complexes. They are located in positions where they mediate hydrogen bonds between the protein and the DNA phosphate backbone, and are frequently also found to play key structural roles in the interaction between protein and functional groups on the DNA bases. Some well-resolved examples of bound water molecules are found in DNA complexes of the λ repressor [67], Hin recombinase [68], Bam HI endonuclease [69], and the estrogen receptor DNA-binding domain [70].

A striking example of the extensive use of water in sequence-specific recognition is the *trp* repressor–operator complex [71,72], in which all the interactions between the protein and the three critical bases of the operator half site are mediated by water molecules. The structure of a corresponding DNA fragment shows that these three hydration sites are already occupied in the uncomplexed state. These water molecules can, therefore, be regarded as noncovalent extensions of the DNA bases which may be used as recognition elements of the DNA target sequence [73].

Other complexes are, however, devoid of bound water at the interface between the protein and DNA bases. This is the case, for instance, in a high-resolution structure of a TBP–DNA complex where the interface shows a sterically perfect complementarity with no cavities [74]. Similarly, no water molecules are found at the hydrophobic GATA-1–DNA interface, as discussed below.

General thermodynamic arguments imply that the transfer of a water molecule from bulk solution to a specific site carries an entropic cost. Carefully calibrated standard entropies of hydrated and anhydrous inorganic salts indicate that the entropy decrease is in the order of 2 kcal mol^{-1} at 300 K [75]. This entropic penalty can be expected to be balanced by favorable van der Waals and hydrogen bonding interactions at the bound state. The free energy cost (or gain) of a bound water molecule can, therefore, be expected to vary with the polarity of the environment. (It is highly unlikely that a single water molecule would be found within a small completely nonpolar cavity [76].) These general considerations are also in qualitative accordance with structural data: water molecules are bound at highly polar interfaces, e.g. between polar protein side chains and DNA phosphates, whereas hydrophobic interfaces or interfaces with very small cavities contain no bound water.

One way to dissect the energetics of a bound water molecule identified in a protein–DNA complex would be to replace it with a covalently attached analog (hydroxyl) and examine the thermodynamic consequences. We are not aware of any such

experimental study. An alternative experiment would be to create a small (presumably) water-binding cavity by removing, for instance, a thymine methyl group at a protein–DNA interface. Such an experiment was recently carried out on a complex between the glucocorticoid receptor DNA-binding domain (GR DBD) and the glucocorticoid response element (GRE) [77], and is described in more detail below.

4.2. The effect of water activity on specificity

Recently developed methodology to probe the amount of water released upon complexation is based on measurements of macromolecular equilibrium constants as a function of osmotic stress (reviewed by Parsegian et al. [78]). By adding inert osmolytes to the reaction buffer, the chemical potential of water can be changed in a direct and controlled manner. If the protein–DNA binding equilibrium is associated with water release, then the addition of osmolytes (that are excluded from the interacting surfaces) will drive the equilibrium towards its more dehydrated state. The shift in the association equilibrium constant, K_{obs}, with osmotic pressure, Π_{osm}, is directly related to the differences in the (excess) volume of water released on association [78]:

$$\left.\frac{\partial[\ln(K_{obs})]}{\partial \Pi_{osm}}\right|_{\Pi_{osm}=0} = -\frac{\Delta V}{kT} = \frac{-(V_{PD} - V_P - V_D)}{kT} \qquad (2)$$

where $\Delta V = V_{PD} - V_D - V_P$ is the difference in water volume that is inaccessible to the osmolyte around the uncomplexed DNA (D) and protein (P) components and the complex (PD), respectively. Dividing ΔV by the volume of a water molecule ($v_w \approx 30$ Å^3) gives the (net) number of water molecules released or taken up in the reaction ($\Delta N_w = \Delta V / v_w$). (Eq. (2) can be rewritten in terms of water activity (μ_w) using the relation $v_w \mathrm{d}\Pi_{osm} = -\mathrm{d}\mu_w$.) A number of considerations have to be made before a shift in equilibrium is interpreted in terms of hydration/dehydration [78]. These include the possible occurrence of pH effects, changes in dielectric constant or activity of other co-solutes (ions), protein denaturation or aggregation, and the fact that the osmolyte may not be completely excluded from the immediate surroundings of the interacting macromolecules.

Osmotic stress methods have, so far, only been applied in studies of a few protein–DNA equilibria, but more widespread use can be expected. Robinson and Sligar showed that the so-called star activity of the restriction endonuclease EcoRI, i.e. the loss of specificity of this enzyme at certain buffer conditions, is due to differential hydration effects [79,80]. Subsequent studies also provide evidence for differences in dehydration associated with enzymatic specificities of various endonucleases [81].

The interaction of the *gal* repressor with a certain operator DNA containing two specific binding sites (O_I and O_E) was found to be very sensitive to osmotic stress, with an estimated 100–180 released water molecules per binding site, depending on the particular osmolyte used [82]. Further, competition experiments suggest that 6 ± 3 more water molecules are released upon binding to the higher-affinity O_E site compared to the O_I site. The nonspecific binding of the *gal* repressor to poly(dIdC) is not accompanied by water release [82]. A similar difference between specific and nonspecific complexes was observed for the CAP–*lac* promoter DNA interaction [83]. Variation of the water activity by four different osmolytes indicates that this interaction releases 80 $\pm$ 5 water molecules. Competition experiments suggest that the corresponding number for nonspecific CAP–DNA interactions is $\approx$ 20 water molecules.

From the experiments on CAP and *gal* repressor, it appears that sequence-specific DNA-binding is associated with the directly measurable release of solute excluded water molecules, whereas a much smaller number of water molecules are released in nonspecific complexes. These observations are likely to reflect the more extensive dehydration that occurs in the formation of sterically complementary interfaces in sequence-specific complexes. The extent of dehydration measured by osmotic stress is, in these two cases, in qualitative agreement with expected changes in solvent accessible surface areas (see also the interpretations of heat capacity changes discussed below).

It is important to note that osmotic stress methods yield the *net* number of waters released or taken up in a reaction. Thus, if there is a coupled unfolding reaction in the protein, then additional surface area is

exposed in the complex. The water molecules that hydrate this additional surface contribute to the measured effect. Such a mechanism may result in an osmotic stress effect opposing that observed with *gal* repressor and CAP interactions. For instance, data on the Hin recombinase–DNA recognition [84] suggest a net uptake of a small number of water molecules. A similar observation is that the binding of a netropsin analog to random sequence DNA involves a net uptake of 50–60 solute-excluding waters [85].

4.3. The hydrophobic effect and interpretation of measured heat capacity changes

The thermodynamic signatures of the hydrophobic effect, observed for the transfer of small nonpolar solutes from a polar to a nonpolar phase, include a negative change in heat capacity, i.e. $\Delta C_p^\circ < 0$, and a convergence temperature, T_s, at which the entropy of transfer is zero [86] (see also Ref. [87] and work cited therein). Similar thermodynamics is often observed in biomolecular processes, such as protein folding [87–89]. (The interpretation of the hydrophobic effect in biochemical processes has recently been discussed by, for example, Dill [90] and Ben-Naim [91], and in work cited therein.) The magnitude of the heat capacity change is closely related to changes in solvent accessible surface areas (see, for example, Refs. [92,93]). These observations have led to the conclusion that dehydration of nonpolar surfaces contributes significantly to heat capacity changes in biomolecular processes, although other molecular mechanisms, such as changes in vibrational modes and coupled equilibria, may also contribute [94,95]. Several laboratories have used thermodynamic data for the transfer of model compounds between water and organic phases to calibrate observed thermodynamics in terms of changes in solvent accessible surface area. An important conclusion of recent studies is that changes in polar surface area to some extent oppose the effect of changes in nonpolar surface area [93,96,97].

A large negative change in heat capacity also appears to be a common thermodynamic property for sequence-specific protein–DNA interactions. Initial studies of the breakdown of the binding free energy in entropic and enthalpic contributions, and determinations of ΔC_p°, based on van't Hoff analyses include histone–DNA interactions [23], and the binding of RNA polymerase $E\sigma^{70}$ to the λP_R promoter [98], the *lac* repressor to the O^+ *lac* operator [99], the Mnt repressor to its O_{mnt}-operator site [100], the *lac* repressor headpiece to a symmetric operator sequence [101], and EcoRI endonuclease to its specific recognition site [101]. A large negative ΔC_p° was found in all these cases. Ha et al. [101] suggested that the hydrophobic effect is a dominant driving force in the formation of sequence-specific protein–DNA complexes, analogous to protein folding. Subsequent analyses based on direct determination of the binding enthalpy using titration calorimetry have been carried out on the sequence-specific binding of the λ Cro protein [32] (pulsed-flow calorimetry), *trp* repressor [34,102], methionine repressor (MetJ) [33,102], λ cI repressor [104] and GR DBD [77]. These complexes display a significant variation in measured ΔC°_p values, which are otherwise all negative.

In contrast, nonspecific protein–DNA complexes appear to involve a zero, or undetectable, ΔC°_p [32,34]. This finding indicates that nonspecific binding does not involve extensive dehydration of surfaces, in accordance with the above-discussed conclusions based on osmotic stress. Similarly, binding of smaller drugs to DNA results in little or no variation in binding enthalpy with temperature (see, for example, Refs. [105,106]). These observations are also consistent with small changes in surface area upon binding.

Thus, it appears that sequence-specific protein–DNA complexes generally conform to the concept of removal of solvent accessible surface areas, accompanied by dehydration, upon formation of complementary protein–DNA interfaces. However, in their analysis of the initial van't Hoff studies mentioned above, Ha et al. [101] noted that the measured ΔC°_p values in these cases were larger than expected based on surface area changes, and suggested that the "excess" ΔC_p° could be accounted for by conformational changes upon binding. An extension of this argument [107] suggested how changes in heat capacity and entropy can be analyzed quantitatively in terms of dehydration and coupled protein folding reactions (discussed below).

Structures of protein–DNA complexes, as well as

of free protein components, are in fact available for the systems whose precise calorimetric ΔC°_{p} values have been measured. Of these, it appears that the magnitude of the ΔC°_{p} values measured for the MetJ repressor [103] and the GR DBD [77] can be approximately accounted for by changes in accessible surface areas. This is consistent with an absence of coupled folding reactions in these complexes. The Cro complex was not analyzed in terms of surface areas [32], and a quantitative analysis of the λ cI repressor–operator complex is difficult because of complicating effects due to cooperativity or conformational changes in DNA [104].

The *trp* repressor shows excess values of ΔC°_{p} [102,34]. An analysis of surface areas in X-ray structures of all species in the equilibrium shows that conformational changes cannot account for this observation (see also the interpretation of NMR data in Section 5). It is argued that the tightening of "soft vibrational modes" [94] provides an alternative contributing mechanism [34] — the molecular interpretation being a restriction of dynamic fluctuations of polar groups and bound water molecules at the interface. Thus, although dehydration effects may dominate the measured ΔC°_{p} for some equilibria, additional experimental data are needed to quantify the relative importance of other contributing mechanisms and to ascertain their relation to, for example, the number of water molecules bound at a protein–DNA interface.

4.4. Sites and residence times of bound water molecules probed by NMR

NMR offers possibilities of identifying sites of preferential hydration and residence times of bound water. The method that has been applied to study hydration in protein–DNA complexes involves comparisons of proton–proton cross relaxation properties in the laboratory frame (nuclear Overhauser enhancement, or NOE) and the rotating frame (ROE) (see Ref. [108] and work cited therein). The observation of NOEs and/or ROEs between DNA/protein and water protons can be used to identify preferential water-binding sites within a known structure. The relative sign and magnitude of the two cross-relaxation processes contain information on the dynamics, and may be interpreted in terms of residence times [108]. Studies of a 1:1 complex between the *Antp(C39S)* homeodomain and a 14-base pair DNA duplex indicate that water bound at the protein–DNA interface has residence times of more than 1 ns but less than 20 ms [109]. These properties are similar to those of water molecules buried in the interior cavities of globular proteins, where they constitute an integral part of the protein structure [108]. This observation suggests that the dynamic environments of bound waters in proteins and at protein–DNA interfaces are similar [109], at least within the time scale resolution provided by cross-relaxation techniques.

NMR studies of the specific monomeric complex between the DNA-binding domain of the GATA-1 transcription factor and DNA reveal several water-binding sites between the protein and the DNA phosphate backbone, but no water molecules at the interface between the protein and the DNA bases [110]. The exclusion of water from the protein–DNA interface in this complex is consistent with the highly hydrophobic nature of the complementary surfaces [111].

5. Coupled protein folding and conformational changes

The formation of sequence-specific protein–DNA complexes can involve conformational changes in the protein. These changes may be large scale rearrangements of domains, as with the λ Cro protein [112], as well as distinct structural rearrangements involving specific protein side chains, as in the Arc repressor [113]. The thermodynamics of structural rearrangements may be complex and involve contributions from several of the various effects discussed in this review. The fact that they are inexplicably related to the thermodynamics of the protein–DNA interactions makes the dissection of the two effects difficult to address experimentally. Calorimetric studies of several specific λ Cro–DNA complexes show the existence of two thermodynamically distinct binding modes, which differ in the extent of enthalpic and entropic contributions to the binding free energy. This effect may be attributed to the presence of conformational subclasses [32].

Alternatively, the change in structure of the free protein required to form the correct complementary

DNA-binding surface may be induced by binding of an allosteric effector, as with the *trp* [114–116] and *lac* [117] repressors. In this case, it is possible to dissect the thermodynamics of protein conformational change from the thermodynamics of DNA binding by studying the two reactions separately. Calorimetric investigations of the L-tryptophan–*trp* repressor interaction [102] shows that the binding reaction is dominated by an enthalpic contribution (at 23°C). The negative heat capacity change appears too large to be accounted for by changes in solvent accessible surface areas of the ligand and its binding site. The observed thermodynamics was therefore interpreted as resulting from differences in structure and, especially, dynamics between the apo- and holorepressor [102], in accordance with observed hydrogen exchange kinetics [118] and molecular dynamics simulations [119]. Similar conclusions could be drawn based on comparisons of the wild-type *trp* repressor with a mutant superrepressor [120].

It has been recognized for some time that sequence-specific DNA binding may be coupled with protein folding reactions, in which regions that are disordered in the uncomplexed protein assume well-defined folded structures upon binding to DNA (see, for example, Refs. [121,122]). Such reactions can be expected to have a unique effect on the binding thermodynamics [98,101,102,107]. Disorder–order transitions that are coupled to DNA binding have, for instance, been observed for the λ repressor [123], GCN4 [124,125], *lac* repressor [117], EcoRV endonuclease [126], and *Antp* homeodomain [127]. The observed folding reactions involve N-terminal folding, formation of helices and loops, or global folding. The opposite situation, i.e. order–disorder transitions, has also been observed in a few cases [69,128].

Coupled protein folding can be expected to result in a negative ΔC_p° for the association process due to a reduction in surface area, as discussed above. Spolar and Record [107] have suggested how the observed thermodynamics may be further quantified in terms of various contributions to the entropy of association. In general, measurements of ΔG°, ΔH° and ΔC°_p for the binding reaction at a given temperature allows determination of ΔS° and, assuming a temperature-independent ΔC°_p, the temperature (T_s) at which $\Delta S^\circ = 0$. A summation of entropy contributions due to the hydrophobic effect (ΔS°_{HE}), loss of rotational and translational degrees of freedom (ΔS°_{rt}), the polyelectrolyte effect (ΔS°_{PE}) and other contributions (ΔS°_{other}) at this temperature yields:

$$\Delta S^\circ(T_s) = 0 = \Delta S^\circ_{HE}(T_s) + \Delta S^\circ_{rt} + \Delta S^\circ_{PE} + \Delta S^\circ_{other} \quad (3)$$

The first of these contributions may be evaluated based on previous work on protein folding, i.e. $\Delta S^\circ_{HE}(T) = 0.32\Delta A_{np}\ln(T/386)$, where ΔA_{np} is the change in nonpolar solvent accessible surface area [87,129]. Further, analysis of the thermodynamics of protein–ligand and protein–protein interactions where coupled folding and polyelectrolyte effects are not involved [107], as well as statistical-mechanical considerations (see, for example, Ref. [130]), suggests that $\Delta S^\circ_{rt} \approx -50$ e.u. for a bimolecular association reaction. Finally, ΔS°_{PE} may be estimated based on the measured salt dependence (Eq. (1)) assuming that the polyelectrolyte effect is purely entropic in character [36,43,107]. Typical values of ΔS°_{PE} calculated for some different protein–DNA complexes fall in the range 11–54 e.u. at the salt concentration ranges used in various studies. It is suggested that the remaining term, ΔS°_{other}, contains predominant contributions from conformational changes or folding in protein and/or DNA [107]. (The fact that entropy loss due to protein folding may be compensated by entropy gained from dehydration is supported by studies of Arc repressor–DNA equilibria at different pressures [131].)

The Spolar–Record approach of breaking down heat capacity and entropy changes into various contributions may prove to be useful in understanding the complex thermodynamics of sequence-specific protein–DNA interactions. It needs to be tested on some model complexes for which (i) the (solution) structures of all participants in the equilibrium are known, (ii) a distinct disorder–order transition is observed, (iii) the polyelectrolyte effect has been measured, and iv) precise calorimetric data on the thermodynamics have been obtained. The quantitative analysis in Ref. [107] is largely based on data on the GR DBD and the *trp* holorepressor. However, the GR DBD fragment for which thermodynamic data is available does not contain completely disordered regions that become ordered in the complex

[132,133], even though more modest conformational changes appear to occur upon binding [134,135]. The analysis of the *trp* holorepressor is based on the assumption that the D helix is disordered in the uncomplexed form. However, the interpretation of NMR data leads to a picture in which the D helix, although less stable than the protein core, transiently assumes two or more different conformational states in which the helical conformation is retained [136].

6. Conformational changes in DNA

DNA conformations that deviate from the classical B-form are often observed in sequence-specific protein–DNA complexes. Examples of DNA-binding proteins that dramatically distort the DNA conformation from the B-form include the CAP protein [137], the TBP [74], and proteins within, and related to, the HMG class [4]. The DNA structure in many other complexes is more similar, but not identical, to B-form DNA. It is obvious that the DNA conformation in both cases exists because it provides an optimal steric and electrostatic complement to the specifically bound protein.

The process by which a DNA sequence is recognized by a protein from its inherent structure or flexibility rather than by direct interactions is known as indirect readout (see Section 2). The classical example of indirect readout is the structure of the phage 434 repressor and its operator [138]. The DNA conformation in this dimeric complex is distorted compared to B-form DNA. By mutating two base pairs which are not in direct contact with the bound protein, it is possible to reduce the binding affinity up to 50-fold [139], indicating that these two base pairs modulate operator affinity through their effects on the relative free energies of different DNA conformations [1].

Although unique DNA conformations are so often observed in complexes, there is little experimental information on the underlying mechanisms and thermodynamics. The thermodynamic consequences of distorting a DNA sequence upon binding should be different compared to a situation in which the DNA already exists in a non-B-form conformation or populates such a conformation in the uncomplexed state. Distortion of the DNA conformation can be expected to be costly in terms of the overall binding free energy. However, a conformational distortion will contribute to specificity if it is even more costly for a protein to distort a noncognate DNA fragment. Berg and von Hippel [140] stress the fact that, whereas conformational lability in proteins in general will be expected to *reduce* specificity in binding, conformational changes in DNA can be sequence dependent and therefore act to *increase* specificity. Erie and Bustamente [141] recently formulated this mechanism thermodynamically when considering the energetic cost of DNA bending: the total binding free energy (ΔG_i) for the formation of a specific ($i = \mathrm{S}$) or nonspecific complex ($i = \mathrm{NS}$) complex can be partitioned as $\Delta G_i = \Delta G_{i,\mathrm{NR}} + \Delta G_{i,\mathrm{B}}$, where $\Delta G_{i,\mathrm{NR}}$ is the (favorable) free energy due to protein–DNA interactions and $\Delta G_{i,\mathrm{B}}$ is the (unfavorable) free energy required to bend the DNA. The binding specificity of a protein is defined as $\Delta\Delta G_{\mathrm{Sp}} = \Delta G_{\mathrm{S}} - \Delta G_{\mathrm{NS}}$, which can be rewritten as

$$\Delta\Delta G_{\mathrm{Sp}} = \Delta G_{\mathrm{S,NR}} + \Delta G_{\mathrm{S,B}} - \Delta G_{\mathrm{NS,NR}} - \Delta G_{\mathrm{NS,B}} = \Delta\Delta G_{\mathrm{Sp,NR}} + \Delta\Delta G_{\mathrm{Sp,B}} \quad (4)$$

where $\Delta\Delta G_{\mathrm{Sp,NR}} = \Delta G_{\mathrm{S,NR}} - \Delta G_{\mathrm{NS,NR}}$ and $\Delta\Delta G_{\mathrm{Sp,B}} = \Delta G_{\mathrm{S,B}} - \Delta G_{\mathrm{NS,B}}$. Eq. (4) shows that sequence-specific differences in bending (or deformation) free energies will also determine the specificity in a hypothetical case where the interfacial protein–DNA interactions present in sequence-specific and nonspecific complexes are energetically equivalent. Experimental data supporting the concept that DNA bending deformations observed in sequence-specific complexes also occur in nonspecific complexes come from the scanning force microscopy of λ Cro with a 1 kb DNA fragment containing specific binding sites at known locations [142]. Calorimetric data on the interaction between λ cI repressor and DNA containing various combinations of the three operator sites demonstrate that sequence-dependent conformational changes in DNA play a significant role in the mechanisms of binding and cooperativity to the sites within the operator region [104]. Studies by Merabet and Ackers [104] also give an indication of the thermodynamic consequences (on heat capacity and enthalpy of binding) of these conformational changes.

Evidence that a particular DNA fragment may be pre-disposed to adopt the non-B-form conformation

present in a sequence-specific complex comes from comparisons of the structure of the uncomplexed *trp*-operator DNA with that in a *trp* repressor–operator complex [73]. In this case, it was concluded that the inherent structure of the DNA sequence as well as its inherent flexibility are major determinants in the recognition process. An example of a protein that relies almost exclusively on the intrinsic and induced DNA conformation for sequence recognition is the human serum response factor [143].

Different DNA conformations have, in fact, also been observed in complexes involving the same DNA-binding protein. Comparisons of complexes between GCN4-bZIP fragments, and AP-1 and ATF/CREB target DNA sites, respectively, reveal that AP-1 DNA remains close to the B-form, whereas ATF/CREB DNA is slightly bent. The high degree of similarity in protein–DNA interactions and binding free energies between the two complexes suggests that the structural forms of DNA observed in these crystals are of comparable energy [144].

7. Changes in dynamics

Changes in the conformational flexibility and dynamics of DNA and protein can, in general, be expected to influence the thermodynamics of the interaction. The internal dynamics of biomolecules is complex and involves a wide variety of dynamic processes occurring at time scales spanning about 15 orders of magnitude (see, for example, Ref. [145]). Changes in protein backbone dynamics with an associated entropy loss are inherent in the local protein folding reactions discussed above. However, changes in dynamics will also occur when the interacting species retain the equilibrium conformations of the uncomplexed states. Energetically unfavorable contributions to the binding can, for instance, be expected owing to the restriction of protein side chain motions at the DNA-interacting surface or restriction of torsional motions in DNA. Changes in dynamics may reflect the fact that a more limited number of approximately iso-energetic states are accessible in the complex. This mechanism is a direct (conformational) entropy effect, and analyses of, for example, side chain conformations in proteins show that it may be extensive in protein folding reactions [146]. Changes in dynamics may also reflect a tightening of potentials [94], in which case the primary thermodynamic consequence is a decrease in heat capacity. However, vibrational entropy can also be gained in the complex owing to an increase in the number of low-frequency vibrational modes, as calculated for the dimerization of insulin [147]. This latter scenario corresponds to an increased motional freedom at low characteristic frequencies of the interacting molecules in the complex.

Only a few experimental or theoretical investigations of protein–DNA interactions in solution have, so far, specifically addressed the thermodynamic consequences of changes in dynamics. However, this issue can be expected to be subject to more intense future research. Ladbury et al. [34] support their calorimetric data on the *trp* repressor–operator interaction with an analysis of crystallographic temperature factors for the free and complexed repressors. Changes in temperature factors are notably large within the DNA-interacting fragments of the repressor. These observations may therefore reflect reduced amplitudes of vibrational motions at the intermolecular interface.

Berglund et al. used ^{15}N NMR relaxation measurements to study the changes in dynamics of a single arginine side chain in the Sso7d protein, as this arginine forms part of the protein–DNA interface [148]. The motions of the arginine become restricted in the complex, but not as restricted as those of an arginine buried in a protein core. The amplitude of internal motions on the pico- to nanosecond time scale can be related to the free energy loss due to restriction of these motions upon binding [149]. A comparison of NMR order parameters measured for arginine side chains at a protein surface, at a protein–DNA interface and in a protein core indicates that the concerted effect of restricting the flexibility of many side chains at an interface may be comparable to other determinants of binding thermodynamics [148].

The inference that some of the inherent flexibility of the interacting molecules needs to be retained in the complex, to increase affinity and specificity, is supported by ^{31}P NMR studies of *lac* repressor headpiece–operator complexes [150,151]. The extent of perturbation of ^{31}P NMR chemical shifts in different operator sequences upon binding of wild-type

and mutant repressor fragments was found to correlate with measured dissociation constants. These observations suggest that specific high-affinity protein–operator complexes retain the inherent flexibility of the free operator, whereas the phosphate esters are conformationally restricted in the lower-affinity complexes [150].

8. Thermodynamic dissection of specific interactions

Protein–DNA complexes involve a large number of noncovalent interactions. It follows that the various interactions contribute only a small amount to the total stabilization energy or, alternatively, that only a few specific interactions make a stabilizing contribution. The latter scenario implies that many interactions, although crucial for the distinction between a sequence-specific and a nonspecific complex, make a zero or perhaps even negative contribution to the total stability. An important consequence of this is that part of the discrimination can be achieved by interfacial groups finding the least destabilizing counterpart only at the correct binding site [13]. The majority of specific interactions observed in complexes are hydrogen bonds (see Refs. [3,152] and references cited therein). These are complemented by ion pairs involving DNA phosphates, and van der Waals contacts with nonpolar surfaces of sugars and bases. The DNA major and minor grooves provide a number of hydrogen-bonding possibilities, but only a smaller number of potential sites for nonpolar interactions. Most nonpolar interactions involve a thymine methyl group that protrudes into the major groove. In this section we review studies aimed at dissecting the energetic contributions of various specific interactions.

A number of studies have focused on the thermodynamic consequences (in terms of the free energy of binding) of systematic base-pair substitutions in cognate DNA sequences and/or specific amino acid mutations on DNA-interacting protein surfaces. The studied systems include the *lac* repressor [5,153–155], the *trp* repressor [156,157], the λ repressor [64,158], the Cro repressor [32,159], the Mnt repressor [160], the EcoRI [49,161–164], EcoRV [165] and RsrI [163] endonucleases, the CAP protein [166,167] and the GR DBD [76,168,169]. These studies, along with a large number of studies employing methods that provide qualitative rather than quantitative data, show that the exchange of single amino acids or base pairs at the contacting surfaces can result in very different contributions to the free energy of binding. At a first glance, this may seem reasonable, because only some of the replaced building blocks are involved in direct interactions between the molecules. However, a closer inspection reveals that many observations cannot be explained in this way. Part of the problem, which is detected as a nonadditivity of different changes, arises because most substitution experiments remove or alter the functional groups at only one of the interacting surfaces, leaving the formerly complementary surface intact. Consequently, these experiments do not measure the energetic contribution of a particular set of identified interactions, but rather a difference compared to the (maybe very unfavorable) interactions that replace the removed ones.

Another part of the problem is due to the fact that the "wild-type" interaction interface is formed by conformational adaptation, as discussed above. Conformational adaptation occurs because the forces stabilizing a protein–DNA complex are the same ones, both in terms of nature and magnitude, that shape the macromolecules. Thus, small adjustments, such as DNA bending and major groove widening or amino acid rearrangements, will be part of the recognition process enabling the molecules to form a complementary surface [5]. For instance, by making a base pair substitution in DNA with the purpose of removing an observed hydrogen bond, the conformational adaptability of DNA, which is strongly sequence dependent, will also be changed. As a consequence, one or more favorable contacts may be lost which require an appropriate conformation of the phosphate backbone. The matter is further complicated by the presence of small water molecules that extend the reach of hydrogen-bonding groups or fill cavities in places where it is not otherwise possible to achieve good complementarity. In order to dissect the role of single protein–DNA contact energies, it is therefore important to make only small perturbations of the contacting surfaces and to compare the observed thermodynamics with complementary structural data.

A particularly thorough investigation is provided

by Lesser et al., who studied the ability of EcoRI endonuclease to discriminate between the correct recognition site and systematically altered DNA sites [162]. This study separates the contribution of identifiable protein–DNA contacts from contributions due to macromolecular structure reorganization. The discrimination is achieved by comparing substitutions that preserve the sequence-dependent conformation of the DNA, i.e. isosteric base analogs, to those that do not, i.e. another natural base pair. Relatively small free energy penalties (≈ 1.5 kcal mol^{-1}) are observed for the loss of a single hydrogen bond or nonpolar interaction with a thymine methyl in isosteric sites. Further, these penalties are found to be independent of the position in the DNA sequence. In contrast, the introduction of an incorrect base pair results in position-dependent free energy penalties that are too large to be attributed to changes in protein–DNA base pair contacts alone. The authors conclude that the extra free energy loss is due to changes in protein–DNA backbone interactions and differences in the energetic cost of distorting the DNA. This observation, as well as the figures associated with removal of single hydrogen bonds, are confirmed in a later study [164].

The energetic effect of small changes at a protein–DNA surface is also given by Smith et al., who studied *trp* repressor binding to operator sequences containing isosteric base analogs [157]. The base analogs differed from the natural bases by a single atom, the N^7-purine nitrogen, which was replaced by a carbon. The results show that the water-mediated hydrogen bonds with N^7-purine nitrogens observed in the crystal structure are critical for the formation of a high-affinity complex. From observed free energy differences, an average value of approximately 1 kcal mol^{-1} was attributed to the loss of a water-mediated hydrogen bond.

The energetic importance of hydrogen-bonding contacts with the central backbone phosphate of the DNA recognition site for EcoRI has been examined by systematically replacing each phosphate oxygen by sulfur to produce stereospecific phosphorothioates [170]. The substitution results in a more localized charge on the sulfur. For phosphates where the protein makes a hydrogen-bonding interaction to only one of the phosphoryl oxygens, one of the substitutions is expected to improve binding and the other to attenuate binding. This is the case for the central phosphate in the binding site and the most favorable energetic contribution of this contact is estimated at -1.5 kcal mol^{-1}.

A stabilizing contribution of about 1–2 kcal mol^{-1} for a hydrogen bond is also consistent with measurements on several other protein–DNA equilibria [156,163,165]. The value is also comparable to those estimated from studies of protein stability for the extreme case of burying hydrogen-bonding groups without any partner, where an unpaired uncharged hydrogen bond contributes approximately 0.5–1.5 kcal mol^{-1} and an unpaired charged hydrogen bond contributes an additional 3 kcal mol^{-1} [171].

The thymine methyl group projects into the DNA major groove where it can make van der Waals contacts with a bound protein, thereby contributing to the sequence discrimination. Several studies have focused on the role of this methyl group by substituting thymine with uracil. In 1979, Fisher and Caruthers [155] showed that the thymine methyl could account for the reduced affinity of a *lac* repressor mutant for *lac* operators where AT base pairs were replaced successively by TA, AU, GC and GC^{Me5}. Only the last substitution involving a 5-methyl cytosine restored full binding affinity. The removal of a contacted methyl group has been found to be energetically unfavorable by approximately 0.6–2.0 kcal mol^{-1} [32,153,154,156,159,162,165,167]. These studies indicate that the thymine methyl is an important, and sometimes even the only, determinant of specificity for AT base pairs. The estimated figures for this interaction are also in accordance with a theoretical calculation on the dehydration of a thymine methyl [172]. (Note, however, the potential pitfall associated with analyses of system and microscopic bond stabilities [173].)

Additional insight into the molecular basis of individual interactions is provided by determinations of the enthalpy and entropy of DNA binding. Takeda et al. applied pulsed-flow microcalorimetry to investigate the thermodynamics of Cro protein–DNA interactions [32]. The studied reactions included both single and multiple base pair alterations as well as single amino acid mutants. The authors were able to correlate measured thermodynamic parameters with expected changes in interactions based on structural

data. For example, the removal of an identifiable van der Waals interaction with a thymine methyl group (T to U substitution) results in a free energy penalty of 1.6 kcal mol^{-1}. This effect is mainly due to an enthalpy change of 1.3 kcal mol^{-1} with an approximately zero change in entropy, which is expected for the loss of this interaction. However, a possible complication in this kind of experiment is the small potentially water-filled cavity created when the methyl group is removed from the protein–DNA interface. This issue was recently addressed using isothermal titration calorimetry [77] for the sequence-specific binding of the GR DBD to the GRE, a complex for which structural data are available [134]. The study focuses on the removal of a thymine methyl group, which is not involved in any critical sequence-specific interaction. The affinities for the mutated sites (AT to AU, GC and GC^{Me5}) were all found to be similar to that of the original site. However, the removal of the methyl is associated with an unfavorable entropy contribution of about 1 kcal mol^{-1} which is compensated by a corresponding favorable enthalpic component. The differences in thermodynamics were therefore interpreted as a replacement of the thymine methyl by an ordered water molecule, which is supported, although not confirmed, by structural data.

9. Concluding remarks

The common denominator for sequence-specific protein–DNA complexes is a large complementary interaction surface. The need for large interacting surfaces can be understood by the requirement of a sufficiently long DNA sequence to define a unique specific binding site. This property distinguishes protein–DNA interactions from small molecule–DNA interactions. (The strongest DNA-binding drugs appear to gain more binding free energy per interacting surface area than sequence-specific proteins. [1]) The number of interactions that actually stabilize a protein–DNA complex is either small or, if it is large, the average interaction contributes only a small part of the binding energy. This situation is reminiscent of protein folding thermodynamics. Protein folding involves the removal of large surfaces from exposure to solvent, and the interior of a folded protein displays optimum steric (packing) and polar/nonpolar complementarity. Nevertheless, the free energy of folding is surprisingly small compared to the large number of noncovalent interactions. Another important thermodynamic consequence of the large interacting surfaces is a significant negative change in heat capacity upon complex formation. The heat capacity change, which is also found in protein folding, results in large temperature dependencies in binding enthalpies and entropies. This is in contrast to DNA–small ligand interactions, for which measurable heat capacity changes are seldom observed, although binding affinities are high (see, for examle, Refs. [105,106]). Despite these similarities, protein–DNA interactions differ from protein folding in many ways, e.g. owing to the polyelectrolyte nature of DNA.

We have outlined the various molecular events that contribute to the formation of the complementary interface, and summarized the current knowledge on how these affect binding thermodynamics. Substantial research has led to significant progress in several areas. These include the extent and thermodynamic effects of dehydration; the combined effects of cations, anions, and pH on binding equilibria; and a realization of the effects of coupled equilibria and protein folding interactions. Further, the contribution to the binding free energy due to various specific interactions at the interface has been determined for a few well-studied model systems. Areas that, in our opinion, need more attention are changes in dynamics, the details of specific electrostatic interactions and desolvation effects, and the thermodynamic effect of conformational changes in DNA. Future research may also consider evolutionary aspects, as these may lead to important clues regarding the most crucial structural and thermodynamic requirements for sequence-specific binding.

This survey illustrates the complexity of the thermodynamics and the way in which various effects make different complexes thermodynamically unique,

[1] T. Härd and T. Lundbäck, unpublished compilation of structural and thermodynamic data for the binding of Hoechst 33258, netropsin, the GR DBD and the *trp* repressor. Literature values for binding free energies (corrected for salt concentration dependence) were compared to buried surface areas calculated from known structures.

just as they are structurally unique. This conclusion may seem rather obvious, but an implication is that it is easier to reach an understanding of the contributing thermodynamic determinants for a single protein–DNA complex than to formulate a general set of rules. This has also been the case for some well-studied drug–DNA complexes (see, for example, Refs. [62,174,175]). Thus, in our opinion, the study of the model systems mentioned here should be continued in order to build a data base of observations. There is much to learn from combinations of experiments addressing structure, dynamics and thermodynamics complemented by computer simulations and statistical-mechanical considerations. It is worth noting that many of the most conclusive studies mentioned in this review involve model systems for which there is an abundance of data on structure and dynamics, and on which only very small systematic changes in chemical structure, buffer conditions, etc., were made.

Acknowledgements

This work was supported by grants from the Swedish Natural Sciences Research Council.

References

[1] S.C. Harrison and A.K. Aggarwal, Annu. Rev. Biochem., 59 (1990) 933–969.
[2] P.S. Freemont, A.N. Lane and M.R. Sanderson, Biochem. J., 278 (1991) 1–23.
[3] C.O. Pabo and R.T. Sauer, Annu. Rev. Biochem., 61 (1992) 1053–1095.
[4] M. Werner, A.M. Gronenborn and G.M. Clore, Science, 271 (1996) 778–784.
[5] M.C. Mossing and M.T. Record, J. Mol. Biol., 186 (1985) 295–305.
[6] V.K. Misra, J.L. Hecht, K.A. Sharp, R.A. Friedman and B. Honig, J. Mol. Biol., 238 (1994) 264–280.
[7] F. Jacob and J. Monod, J. Mol. Biol., 3 (1961) 318–356.
[8] W. Gilbert and B. Müller-Hill, Proc. Natl. Acad. Sci. USA, 56 (1966) 1891–1898.
[9] M. Ptashne, Nature, 214 (1967) 232–234.
[10] C. Hélène and G. Lancelot, Prog. Biophys. Mol. Biol., 39 (1982) 1–68.
[11] P.H. von Hippel, Science, 263 (1994) 769–770.
[12] P.H. von Hippel, in R.F. Goldberger (Ed.), Biological Regulation and Development, Plenum, New York, 1979, pp. 279–347.
[13] P.H. von Hippel and O.G. Berg, Proc. Natl. Acad. Sci. USA, 83 (1986) 1608–1612.
[14] M. Yarus, Annu. Rev. Biochem., 38 (1969) 841–880.
[15] P.H. von Hippel and J.D. McGhee, Annu. Rev. Biochem., 41 (1972) 231–294.
[16] N.C. Seeman, J.M. Rosenberg and A. Rich, Proc. Natl. Acad. Sci. USA, 73 (1976) 804–808.
[17] P.H. Richter and M. Eigen, Biophys. Chem., 2 (1974) 255–263.
[18] R.B. Winter, O.G. Berg and P.H. von Hippel, Biochemistry, 20 (1981) 6961–6977.
[19] H. Kabata, O. Kurosawa, I. Arai, M. Washizu, S.A. Margarson, R.E. Glass and N. Shimamoto, Science, 262 (1993) 1561–1563.
[20] A.D. Riggs, S. Bourgeois, R.F. Newby and M. Cohn, J. Mol. Biol., 34 (1968) 365–368.
[21] O.W. Jones and P. Berg, J. Mol. Biol., 22 (1966) 199–209.
[22] A.D. Riggs, S. Suzuki and S. Bourgeois, J. Mol. Biol., 48 (1970) 67–83.
[23] M.C. Poliakow, M.H. Champagne and M.P. Daune, Eur. J. Biochem., 26 (1972) 212–219.
[24] D.J. Galas and A. Smith, Nucl. Acids Res., 5 (1978) 3157–3170.
[25] V. Petri, M. Hsieh and M. Brenowitz, Biochemistry, 34 (1995) 9977–9984.
[26] D.E. Draper and L. Gold, Biochemistry, 19 (1980) 1774–1781.
[27] T. Heyduk and J.C. Lee, Proc. Natl. Acad. Sci. USA, 87 (1990) 1744–1748.
[28] C.A. Royer, A.E. Chakerian and K.S. Matthews, Biochemistry, 29 (1990) 4959–4966.
[29] M.G. Fried and D.M. Crothers, Nucl. Acids. Res., 9 (1981) 6505–6525.
[30] M.M. Garner and A. Revzin, Nucl. Acids Res., 9 (1981) 3047–3060.
[31] A.D. Frankel, G.K. Ackers and H.O. Smith, Biochemistry, 24 (1985) 3049–3054.
[32] Y. Takeda, P.D. Ross and C.P. Mudd, Proc. Natl. Acad. Sci. USA, 89 (1992) 8180–8184.
[33] A. Cooper, A. McAlpine and P.G. Stockley, FEBS Lett., 348 (1994) 41–45.
[34] J.E. Ladbury, J.G. Wright, J.M. Sturtevant and P.B. Sigler, J. Mol. Biol., 238 (1994) 669–681.
[35] C.F. Anderson and M.T. Record, Annu. Rev. Phys. Chem., 46 (1995) 657–700.
[36] M.T. Record, T.M. Lohman and P. deHaseth, J. Mol. Biol., 107 (1976) 145–158.
[37] P.L. deHaseth, T.M. Lohman and M.T. Record, Biochemistry, 16 (1977) 4783–4790.
[38] M.T. Record, P.L. deHaseth and T.M. Lohman, Biochemistry, 16 (1977) 4791–4796.
[39] M.T. Record, C.F. Anderson and T.M. Lohman, Quart. Rev. Biophys., 11 (1978) 103–178.
[40] M.D. Barkley, P.A. Lewis and G.E. Sullivan, Biochemistry, 20 (1981) 3842–3851.
[41] S. Leirmo, C. Harrison, D.S. Cayley, R.R. Burgess and M.T. Record, Biochemistry, 26 (1987) 2095–2101.

[42] J.-H. Ha, M.W. Capp, M.D. Hohenwalter, M. Baskerville and M.T. Record, J. Mol. Biol., 228 (1992) 252–264.
[43] M.T. Record, J.-H. Ha and M.A. Fisher, Methods Enzymol., 208 (1991) 291–343.
[44] G.S. Manning, Quart. Rev. Biophys., 11 (1978) 179–246.
[45] T.M. Lohman, P.L. deHaseth and M.T. Record, Biochemistry, 19 (1980) 3522–3530.
[46] W.H. Braunlin, T.J. Strick and M.T. Record, Biopolymers, 21 (1982) 1301–1314.
[47] L.B. Overman, W. Bujalowski and T.M. Lohman, Biochemistry, 27 (1988) 456–471.
[48] M.G. Fried and D.F. Stickle, Eur. J. Biochem., 218 (1993) 469–475.
[49] L. Jen-Jacobson, M. Kurpiewski, D. Lesser, J. Grable, H.W. Boyer, J.M. Rosenberg and P.J. Greene, J. Biol. Chem., 258 (1983) 14638–14646.
[50] B.J. Terry, W.E. Jack, R.A. Rubin and P. Modrich, J. Biol. Chem., 158 (1983) 9820–9825.
[51] L.B. Overman and T.M. Lohman, J. Mol. Biol., 236 (1994) 165–178.
[52] C.F. Anderson and M.T. Record, J. Phys. Chem., 68 (1993) 786–794.
[53] J. Wyman, Adv. Prot. Chem., 19 (1964) 223–286.
[54] M.C. Olmstedt, J.P. Bond, C.F. Anderson and M.T. Record, Biophys. J., 68 (1995) 634–647.
[55] M.T. Record and C.F. Anderson, Biophys. J., 68 (1995) 786–794.
[56] V.K. Misra, K.A. Sharp, R.A. Friedman and B. Honig, J. Mol. Biol., 238 (1994) 245–263.
[57] D.P. Mascotti and T.M. Lohman, Biochemistry, 31 (1992) 8932–8946.
[58] D.P. Mascotti and T.M. Lohman, Biochemistry, 32 (1993) 10568–10579.
[59] T. Lundbäck and T. Härd, J. Phys. Chem., in press.
[60] K.A. Sharp, R.A. Friedman, V. Misra, J. Hecht and B. Honig, Biopolymers, 36 (1995) 245–262.
[61] R.L. Jones and W.D. Wilson, Biopolymers, 20 (1981) 141–154.
[62] V.K. Misra and B. Honig, Proc. Natl. Acad. Sci. USA, 92 (1995) 4691–4695.
[63] D.F. Senear and R. Batey, Biochemistry, 30 (1991) 6677–6688.
[64] H.C.M. Nelson and R.T. Sauer, Cell, 42 (1985) 549–558.
[65] J. Guenot, R.J. Flatterick and P.A. Kollman, Prot. Sci., 3 (1994) 1276–1285.
[66] M. Zacharias, B.A. Luty, M.E. Davis and J.A. McCammon, Biophys. J., 63 (1992) 1280–1285.
[67] L.J. Beamer and C.O. Pabo, J. Mol. Biol., 227 (1992) 177–196.
[68] J.-A. Feng, R.C. Johnson and R.E. Dickerson, Science, 263 (1994) 348–355.
[69] M. Newman, T. Strzelecka, L.F. Dorner, I. Schildkraut and A.K. Aggarwal, Science, 269 (1995) 656–663.
[70] J.W.R. Schwabe, L. Chapman, J.T. Finch and D. Rhodes, Cell, 75 (1993) 567–578.
[71] Z. Otwinowski, R.W. Schevitz, R.-G. Zhang, C.L. Lawson, A. Joachimiak, R.Q. Marmorstein, B.F. Luisi and P.B. Sigler, Nature, 335 (1988) 321–329.
[72] C.L. Lawson and J. Carey, Nature, 366 (1993) 178–182.
[73] Z. Shakked, G. Guzikevich-Guerstein, F. Frolow, D. Rabinovich, A. Joachimiak and P.B. Sigler, Nature, 368 (1994) 469–473.
[74] J.L. Kim and S.K. Burley, Nat. Struct. Biol., 1 (1994) 638–653.
[75] J.D. Dunitz, Science, 264 (1994) 670.
[76] R. Wolfenden and A. Radzicka, Science, 265 (1994) 936–937.
[77] T. Lundbäck and T. Härd, Proc. Natl. Acad. Sci. USA, 93 (1996) 4754–4759.
[78] V.A. Parsegian, R.P. Rand and D.C. Rau, Methods Enzymol., 259 (1995) 43–94.
[79] C.R. Robinson and S.G. Sligar, J. Mol. Biol., 234 (1993) 302–306.
[80] C.R. Robinson and S.G. Sligar, J. Biol. Chem., in press.
[81] C.R. Robinson and S.G. Sligar, Proc. Natl. Acad. Sci. USA, 92 (1995) 3444–3448.
[82] M.M. Garner and D.C. Rau, EMBO J., 14 (1995) 1257–1263.
[83] K. Vossen, Ph.D. thesis, Pennsylvania State University 1996.
[84] C.R. Robinson, personal communication, 1996.
[85] N. Sidorova and D.C. Rau, Biopolymers, 35 (1995) 377–384.
[86] J.T. Edsall, J. Am. Chem. Soc., 57 (1935) 1506–1507.
[87] R.L. Baldwin, Proc. Natl. Acad. Sci. USA, 83 (1986) 8069–8073.
[88] W. Kauzmann, Adv. Prot. Chem., 14 (1959) 1–63.
[89] P.L. Privalov, Adv. Prot. Chem., 33 (1979) 167–240.
[90] K.A. Dill, Biochemistry, 29 (1990) 7133–7155.
[91] A. Ben-Naim, Curr. Opin. Struct. Biol., 4 (1994) 264–268.
[92] S.J. Gill and I. Wadsö, Proc. Natl. Acad. Sci. USA, 73 (1976) 2955–2958.
[93] K.P. Murphy and E. Freire, Adv. Prot. Chem., 43 (1992) 313–361.
[94] J.M. Sturtevant, Proc. Natl. Acad. Sci. USA, 74 (1977) 2236–2240.
[95] J. Gómez, V.J. Hilser, D. Xie and E. Freire, Proteins, 22 (1995) 404–412.
[96] R.S. Spolar, J.R. Livingstone and M.T. Record, Biochemistry, 31 (1992) 3947–3955.
[97] G.I. Makhatadze and P.L. Privalov, J. Mol. Biol., 232 (1993) 639–659.
[98] J.H. Roe, R.R. Burgess and M.T. Record, J. Mol. Biol., 184 (1985) 441–453.
[99] P.A. Whitson, J.S. Olson and K.S. Matthews, Biochemistry, 25 (1986) 3852–3858.
[100] A.K. Vershon, S.-M. Liao, W.R. McClure and R.T. Sauer, J. Mol. Biol., 195 (1987) 311–322.
[101] J.-H. Ha, R.S. Spolar and M.T. Record, J. Mol. Biol., 209 (1989) 801–816.
[102] L. Jin, J. Yang and J. Carey, Biochemistry, 32 (1993) 7302–7309.

[103] D.E. Hyre and L.D. Spicer, Biochemistry, 34 (1995) 3212–3221.
[104] E. Merabet and G.K. Ackers, Biochemistry, 34 (1995) 8554–8563.
[105] F.G. Loontiens, P. Regenfuss, A. Zechel, L. Dumortier and R.M. Clegg, Biochemistry, 29 (1990) 9029–9039.
[106] D. Rentzeperis, L.A. Marky, T.J. Dwyer, B.H. Geierstanger, J.G. Pelton and D.E. Wemmer, Biochemistry, 34 (1995) 2937–2945.
[107] R.S. Spolar and M.T. Record, Science, 263 (1994) 777–784.
[108] G. Otting, E. Liepinsh and K. Wüthrich, Science, 254 (1991) 974–980.
[109] Y.Q. Qian, G. Otting and K. Wüthrich, J. Am. Chem. Soc., 115 (1993) 1189–1190.
[110] G.M. Clore, A. Bax, J.G. Omichinski and A.M. Gronenborn, Structure, 15 (1994) 89–94.
[111] J.G. Omichinski, G.M. Clore, O. Schaad, G. Felsenfeld, C. Trainor, E. Appella, S.J. Stahl and A.M. Gronenborn, Science, 261 (1993) 438–446.
[112] R.G. Brennan, S.L. Roderick, Y. Takeda and B.W. Matthews, Proc. Natl. Acad. Sci. USA, 87 (1990) 8165–8169.
[113] B.E. Raumann, M.A. Rould, C.O. Pabo and R.T. Sauer, Nature, 367 (1994) 754–757.
[114] R.-G. Zhang, A. Joachimiak, C.L. Lawson, R.W. Schevitz, Z. Otwinowski and P.B. Sigler, Nature, 327 (1987) 591–597.
[115] D. Zhao, C.H. Arrowsmith, X. Jia and O. Jardetzky, J. Mol. Biol., 229 (1993) 735–746.
[116] H. Zhang, D. Zhao, M. Revington, W. Lee, X. Jia, C. Arrowsmith and O. Jardetzky, J. Mol. Biol., 238 (1994) 592–614.
[117] M. Lewis, G. Chang, N.C. Horton, M.A. Kercher, H.C. Pace, M.A. Schumacher, R.G. Brennan and P. Lu, Science, 271 (1996) 1247–1254.
[118] C.H. Arrowsmith, J. Czaplicki, S.B. Iyer and O. Jardetzky, J. Am. Chem. Soc., 113 (1991) 4020–4022.
[119] A.E. Howard and P.A. Kollman, Protein Sci., 1 (1992) 1173–1184.
[120] R.J. Reedstrom and C.A. Royer, J. Mol. Biol., 253 (1995) 266–276.
[121] L. Walters and E.T. Kaiser, J. Am. Chem. Soc., 197 (1985) 6422–6424.
[122] A.D. Frankel and P.S. Kim, Cell, 65 (1991) 717–719.
[123] N.D. Clarke, L.J. Beamer, H.R. Goldberg, C. Berkower and C.O. Pabo, Science, 254 (1991) 267–270.
[124] K.T. O'Neil, R.H. Hoess and W.F. DeGrado, Science, 249 (1990) 774–778.
[125] M.A. Weiss, T. Ellenberger, C.R. Wobbe, J.P. Lee, S.C. Harrison and K. Struhl, Nature, 347 (1990) 575–578.
[126] F.K. Winkler, D.W. Banner, C. Oefner, D. Tsernoglou, R.S. Brown, S.P. Heathman, R.K. Bryan, P.D. Martin, K. Petratos and K.S. Wilson, EMBO J., 12 (1993) 1781–1795.
[127] G. Otting, Y.Q. Qian, M. Billeter, M. Müller, M. Affolter, W.J. Gehring and K. Wüthrich, EMBO J., 9 (1990) 3085–3092.
[128] J.M. Petersen, J.J. Skalicky, L.W. Donaldson, L.P. McIntosh, T. Alber and B.J. Graves, Science, 269 (1995) 1866–1869.
[129] J.R. Livingstone, R.S. Spolar and M.T. Record, Biochemistry, 30 (1991) 4237–4244.
[130] A.V. Finkelstein and J. Janin, Protein Eng., 3 (1989) 1–3.
[131] D. Foguel and J.L. Silva, Proc. Natl. Acad. Sci. USA, 91 (1994) 8244–8247.
[132] H. Berglund, H. Kovács, K. Dahlman-Wright, J.-Å. Gustafsson and T. Härd, Biochemistry, 31 (1992) 12001–12011.
[133] M.A.L. Eriksson, H. Berglund, T. Härd and L. Nilsson, Proteins, 17 (1993) 375–390.
[134] B.F. Luisi, W.X. Xu, Z. Otwinowski, L.P. Freedman, K.R. Yamamoto and P.B. Sigler, Nature, 352 (1991) 497–505.
[135] H. Baumann, K. Paulsen, H. Kovács, H. Berglund, A.P.H. Wright, J.-Å. Gustafsson and T. Härd, Biochemistry, 32 (1993) 13463–13471.
[136] M.R. Gryk, M.D. Finucane, Z. Zheng and O. Jardetzky, J. Mol. Biol., 246 (1995) 618–627.
[137] S.C. Schultz, G.C. Shields and T.A. Steitz, Science, 253 (1991) 1001–1007.
[138] A.K. Aggarwal, D.W. Rodgers, M. Drottar, M. Ptashne and S.C. Harrison, Science, 242 (1988) 899–907.
[139] G.M. Koudelka, S.C. Harrison and M. Ptashne, Nature, 326 (1987) 886–888.
[140] O.G. Berg and P.H. von Hippel, TIBS, 13 (1988) 207–211.
[141] A. Schepartz, D.A. Erie and C. Bustamente, Science, 269 (1995) 989–990.
[142] D.A. Erie, G. Yang, H.C. Schultz and C. Bustamente, Science, 266 (1994) 1562–1566.
[143] L. Pellegrini, S. Tan and T.J. Richmond, Nature, 376 (1995) 490–498.
[144] P. König and T.J. Richmond, J. Mol. Biol., 233 (1993) 139–154.
[145] J.A. McCammon and S.C. Harvey, Dynamics of Proteins and Nucleic Acids, Cambridge University Press, 1987.
[146] A.J. Doig and M.J.E. Sternberg, Protein Sci., 4 (1995) 2247–2251.
[147] B. Tidor and M. Karplus, J. Mol. Biol., 238 (1994) 405–414.
[148] H. Berglund, H. Baumann, S. Knapp, R. Ladenstein and T. Härd, J. Am. Chem. Soc., 117 (1995) 12883–12884.
[149] M. Akke, R. Brüschweiler and A.G. Palmer, J. Am. Chem. Soc., 115 (1993) 9832–9833.
[150] C. Karslake, M.V. Botuyan and D.G. Gorenstein, Biochemistry, 31 (1992) 1849–1858.
[151] M.V. Botuyan, D.A. Keire, C. Kroen and D.G. Gorenstein, Biochemistry, 32 (1993) 6863–6874.
[152] Y. Mandel-Gutfreund, O. Schueler and H. Margalit, J. Mol. Biol., 253 (1995) 370–382.
[153] D.V. Goeddel, D.G. Yansura and M.H. Caruthers, Nucl. Acids Res., 9 (1977) 3039–3052.
[154] D.V. Goeddel, D.G. Yansura and M.H. Caruthers, Proc. Natl. Acad. Sci. USA, 75 (1978) 3578–3582.
[155] E.F. Fisher and M.H. Caruthers, Nucl. Acids Res., 7 (1979) 401–416.
[156] J.M. Mazzarelli, S.B. Rajur, P.L. Iadarola and L.W. McLaughlin, Biochemistry, 31 (1992) 5925–5936.

[157] S.A. Smith, S.B. Rajur and L.W. McLaughlin, Nature Struct. Biol., 1 (1994) 18–22.
[158] H.C.M. Nelson and R.T. Sauer, J. Mol. Biol., 192 (1986) 27–38.
[159] Y. Takeda, A. Sarai and V.M. Rivera, Proc. Natl. Acad. Sci. USA, 86 (1989) 439–443.
[160] P. Youderian, A. Vershon, S. Bouvier, R.T. Sauer and M.M. Susskind, Cell, 35 (1983) 777–783.
[161] L. Jen-Jacobson, D. Lesser and M. Kurpiewski, Cell, 45 (1986) 619–629.
[162] D.R. Lesser, M.R. Kurpiewski and L. Jen-Jacboson, Science, 250 (1990) 776–786.
[163] C.R. Aiken, L.W. McLaughlin and R.I. Gumport, J. Biol. Chem., 266 (1991) 19070–19078.
[164] D.R. Lesser, M.R. Kurpiewski, T. Waters, B.A. Connolly and L. Jen-Jacobson, Proc. Natl. Acad. Sci. USA, 90 (1993) 7548–7552.
[165] P.C. Newman, V.U. Nwosu, D.M. Williams, R. Cosstick, F. Scela and B.A. Connolly, Biochemistry, 29 (1990) 9891–9901.
[166] R.H. Ebright, A. Kolb, H. Buc, T.A. Kunkel, J.S. Krakow and J. Beckwith, Proc. Natl. Acad. Sci. USA, 84 (1987) 6083–6087.
[167] A. Gunasekera, Y.W. Ebright and R.H. Ebright, J. Biol. Chem., 267 (1992) 14713–14720.
[168] T. Lundbäck, C. Cairns, J.-Å. Gustafsson, J. Carlstedt-Duke and T. Härd, Biochemistry, 32 (1993) 5074–5082.
[169] T. Lundbäck, J. Zilliacus, J.-Å. Gustafsson, J. Carlstedt-Duke and T. Härd, Biochemistry, 33 (1994) 5955–5965.
[170] D.R. Lesser, A. Grajkowski, M.R. Kurpiewski, M. Koziolkiewski, W.J. Stec and L. Jen-Jacobson, J. Biol. Chem., 267 (1992) 24810–24818.
[171] A.R. Fersht, J.-P. Shi, J. Knill-Jones, D.M. Lowe, A.J. Wilkinson, D.M. Blow, P. Brick, P. Carter, M.M.Y. Waye and G. Winter, Nature, 314 (1985) 235–238.
[172] K.W. Plaxco and W.A. Goddard, Biochemistry, 33 (1994) 3050–3054.
[173] K.A. Sharp and S.W. Englander, TIBS, 19 (1994) 526–529.
[174] L.A. Marky and K.J. Breslauer, Proc. Natl. Acad. Sci. USA, 84 (1987) 4359–4363.
[175] J.B. Chaires, S. Satyanarayana, D. Suh, I. Fokt, T. Przewloka and W. Priebe, Biochemistry, 35 (1996) 2047–2053.