PROTEINS: Structure, Function, and Genetics 21:30–39 (1995)

Elusive Affinities

Joël Janin
Laboratoire de Biologie Structurale, UMR 9920 CNRS-Université Paris-Sud, 91198-Gif-sur-Yvette, France

ABSTRACT **The affinity of two molecules for each other and its temperature dependence are determined by the change in enthalpy, free enthalpy, entropy, and heat capacity upon dissociation. As we know the forces that stabilize protein–protein or protein–DNA association and the three-dimensional structures of the complex, we can in principle derive values for each one of these parameters. The calculation is done first in gas phase by molecular mechanics, then in solution with the help of hydration parameters calibrated on small molecules. However, estimates of enthalpy and entropy changes in gas phase have excessively large error bars even under the approximation that the components of the complex associate as rigid bodies. No reliable result can be expected at the end. The fit to experimental values derived from binding and calorimetric measurements is poor, except for the dissociation heat capacity. This parameter can be attributed mostly to the hydration step and it correlates with the size of the interface. Many protein–protein complexes have interface areas in the range 1200–2000 Å^2 and only small conformation changes, so the rigid body approximation applies. It is less generally valid in protein–DNA complexes, which have interfaces covering 2200–3100 Å^2, large dissociation heat capacities, and affect both the conformation and the dynamics of their components.** © 1995 Wiley-Liss, Inc.

Key words: protein–protein interaction, protein–DNA interaction, microcalorimetry, heat capacity changes, entropy, accessible surface area

INTRODUCTION

> Have we not already come to affinities?—We have indeed, and let us at once now get to know them in what they are and in what they do. We say of those natures which on meeting speedily connect and interact that they have an affinity for one another.
>
> Goethe, *Elective Affinities*

The ability to form noncovalent complexes of high affinity and specificity is a fundamental property of biological macromolecules. It is through specific association that proteins and nucleic acids recognize other molecules in cells and spontaneously assemble into stable structures carrying novel functions. As a result of X-ray crystallography, the structure of many complexes of major biological significance is known at the atomic level. We can give a detailed description of the interaction between an enzyme and a protein inhibitor, an antibody and an antigen, or a DNA fragment and a transcription factor. We know the physical principles that govern that noncovalent association: they are the same as for the self-assembly or folding of a single macromolecule. Protein–protein recognition usually implicates a small number (10–30) of amino acid residues on each partner.[1,2] So, it ought to be a simpler problem than folding, within the reach of present day theoretical methods.

Seen from a strictly thermodynamic point of view, the affinity of the two components of a noncovalent complex is determined by the change in energy (at constant pressure, enthalpy) and entropy of a system that contains, on the one hand, the macromolecules and surrounding solvent, and on the other hand, the complex and solvent. A number of recent papers suggest ways of calculating both the enthalpy and the entropy changes.[3–6] I want to show here that neither calculation can be done reliably and that our best effort still produces only a qualitative interpretation of recognition and self-assembly. The problem is nevertheless well worth thinking about, and on such an occasion, recalling textbook physical chemistry has its virtues: this is what is done here. Though I deal mostly with protein–protein association, the interaction of proteins with DNA is also considered. Polysacharrides and large "small molecule" ligands, peptides for instance, are not, but they may be assumed to follow the same rules with due consideration of their specific features.

ENTHALPY VS. ENTROPY IN PROTEIN FOLDING AND ASSOCIATION

The affinity of A for B is measured by the dissociation constant K_d of the noncovalent complex AB, derived by applying the law of mass action to the concentrations of each chemical species at equilib-

Received August 9, 1994; accepted September 15, 1994.
Address reprint requests to J. Janin, Laboratoire de Biologie Structurale, UMR 9920 CNRS-Université Paris-Sud, 91198-Gif-sur-Yvette, France.

rium. The law of mass action applies only to model systems, perfect gases and ideal solutions, and in real systems, chemical activities replace concentrations. However, high affinity complexes can pass for ideal and we shall ignore the correction. We could not do so if we were interested in crystallization or in the assembly of membrane components, processes that occur far from ideality.

The free enthalpy (Gibbs energy) of dissociation per mol of complex in standard state is given by the classical formula:

$$\Delta G_d = -RT \ln K_d \quad (1)$$

($R = 8.314\ \mathrm{J\ K^{-1}\ mol^{-1}} \approx 2\ \mathrm{cal\ K^{-1}\ mol^{-1}}$). Experimental values of K_d range from 10^{-6} to 10^{-14} mol liter^{-1} for the sort of systems we consider, so ΔG_d is positive in the range 8–19 kcal mol^{-1} at T = 300 K.

ΔG_d has an enthalpic component ΔH_d and an entropic component $-T\Delta S_d$. A positive ΔH_d favors association and a positive ΔS_d favors dissociation. Both components vary with temperature, and that variation is described by the heat capacity of dissociation at constant pressure:

$$\Delta C_d = \frac{d(\Delta H_d)}{dT} = T\frac{d(\Delta S_d)}{dT}. \quad (2)$$

The enthalpy of dissociation ΔH_d is equal to the heat evolved when 1 mol of complex is formed from its components. This can be measured directly in a calorimeter, or indirectly by determining K_d at several temperatures and applying the van't Hoff law. The knowledge of both ΔH_d and ΔG_d then yields the entropy of dissociation ΔS_d. As for ΔC_d, its value is obtained by measuring ΔH_d at several temperatures. Note that many authors quote thermodynamic parameters for association instead of dissociation. They have the same value with the opposite sign.

K_d can be often measured to within 10–20%. Then, we know ΔG_d with an accuracy of the order of 0.1 kcal mol^{-1}. Data on ΔH_d and ΔC_d are less abundant, and they are much less accurate when they derive from van't Hoff plots of K_d. However, isothermal mixing microcalorimetric studies are in progress on protease–inhibitor,[7,8] antigen–antibody,[9–11] protein–carbohydrate,[12,13] and protein–DNA complexes.[14,15] Data from these studies complement and correct those from van't Hoff plots. They show that ΔH_d and ΔS_d can have either sign depending on the system and on experimental conditions. Raising the temperature makes some complexes more stable (negative ΔH_d) and other less stable (positive ΔH_d). Nevertheless, high affinity complexes usually have a positive ΔH_d near 25°C, and in all cases, ΔC_d is positive: ΔH_d and ΔS_d increase with temperature.

Similar experiments have been performed on protein folding by scanning microcalorimetry.[16] They yield the enthalpy, entropy, free enthalpy, and heat capacity (ΔH_u, ΔS_u, ΔG_u, and ΔC_u) of the native-to-unfolded transition. Globular proteins unfold somewhere between 25° and 100°C with an enthalpy of unfolding ΔH_u that is positive and large compared to the free enthalpy ΔG_u. So, the existence of a native fold at 25°C requires that ΔH_u be compensated by an almost equally large $T\Delta S_u$. Under these conditions, folding can be said to be driven by a favorable enthalpy change fighting an unfavorable entropy change. However, cold denaturation is probably as basic a feature of protein thermodynamics as heat denaturation,[17] even though it may be difficult to observe as it occurs below 0°C for most proteins. It implies that at low temperature, protein folding has an unfavorable enthalpy change and a favorable entropy change, the reverse of what it is at high temperature. Yet the forces involved are exactly the same. The statement that protein folding is enthalpy driven is valid only within a narrow temperature range. Can it be of any use in describing the physical reality?

The rapid variation of ΔH_u and ΔS_u with temperature is a direct consequence of the large positive heat capacity of unfolding. Dissociation heat capacities are also positive and large, and this implies that temperature affects affinities. This is shown in Figure 1 for the monoclonal antibody HyHEL5 and its specific antigen hen lysozyme. The behavior of the thermodynamic parameters in this particular protein–protein complex[11] is typical of many others. Dissociation has a positive enthalpy change in the range of temperatures studied, so K_d increases and the affinity drops at higher temperatures. The free enthalpy change ΔG_d remains nearly constant over a large temperature range as the variations of ΔH_d and $T\Delta S_d$ cancel each other. ΔC_d being positive, ΔH_d and ΔS_d decrease at low temperature. As a consequence, they must change sign at some point. For ΔH_d, this occurs near −45°C, a temperature that experiments cannot reach; for ΔS_d, it is near 0°C.

So, the HyHEL5–lysozyme complex has no entropy of dissociation at 0°C. This may sound strange—except that it is true only in the standard state. We implicitly took the standard state to be 1 mol liter^{-1} when we wrote Eq. (1), and, therefore, we may now state that the entropy of an ideal (and obviously hypothetical) molar solution of the complex is equal to that of a mixture of molar lysozyme and molar antibody at 0°C. Yet, the entropy of a micromolar solution of the complex (a more realistic situation) is not equal to that of a micromolar mixture of its components, it is smaller by exactly $R \ln 10^6$. Albeit fully stable at this concentration, the complex now has a large positive dissociation entropy, equal to $R \ln 10^6$ per mol and equivalent to 7.5 kcal mol^{-1} free enthalpy (at 0°C). It is correct to say that dissociation is entropy driven at 1 μmol liter^{-1} and not at 1 mol liter^{-1}. Does that statement contain any in-

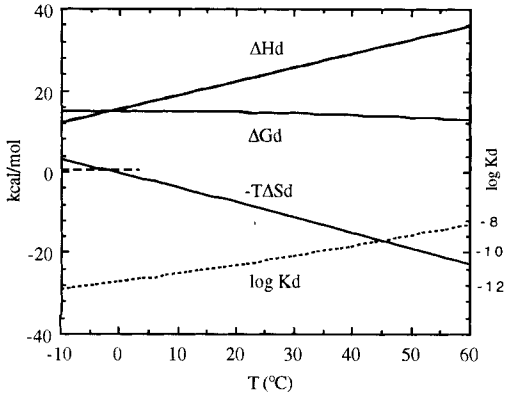

Fig. 1. Temperature dependence of thermodynamic parameters for the HyHEL5–hen lysozyme complex. The curves were extrapolated from calorimetric measurements[11] done in the range 10–37°C by applying Eq. (2) with a temperature-independent heat capacity change ΔC_d = 0.34 kcal mol^{-1} K^{-1}. K_d (dashes) varies by two orders of magnitude between 0 and 37°C, whereas ΔG_d varies by less than 1 kcal mol^{-1}.

formation on the nature of antigen–antibody interactions? Of course not. At best, it reminds us that the reaction involves a change in the number of molecules and that entropies in solution are relative to a standard state.

The *caveat* on the standard state can also be made by noting that Eq. (1) involves taking the logarithm of a quantity, K_d, that is not dimensionless. This is faulty and we ought to write

$$\Delta G_d = -RT \ln \frac{K_d}{c_\phi} \qquad (3)$$

to be in line with physical chemistry textbooks (for instance, Ch. 7.10 of Atkins[18]); c_ϕ is the concentration (or activity) of a chemical species chosen as its standard state. Equation (3) makes it obvious that ΔG_d depends on c_ϕ. ΔS_d does also, but not ΔH_d or ΔC_d. The choice of c_ϕ = 1 mol liter^{-1} is pure convention. In gas state, the convention is based on pressure, not concentration. Unitary entropies (entropies of the pure compound) and partial molal entropies are also used. For aqueous solutions, c_ϕ is then set equal to the molarity of water itself, 55.5 mol liter^{-1}. To return to the usual convention, the quantity R ln 55.5 ≈8 cal mol^{-1} K^{-1} must then be added to the entropy of all solutes.[19]

A THERMODYNAMIC CYCLE FOR THE ASSEMBLY OF PROTEIN–PROTEIN COMPLEXES

Let us now propose a way to approach ΔH_d, ΔS_d, ΔG_d, ΔC_d (noted ΔX_d for short) by considering the cycle drawn in Figure 2. There

$$\Delta X_d = \Delta X^{gas} + \Delta X^{hyd}. \qquad (4)$$

Each quantity is estimated first in gas phase, then a hydration term ΔX^{hyd} is added to it. ΔX^{hyd} is the

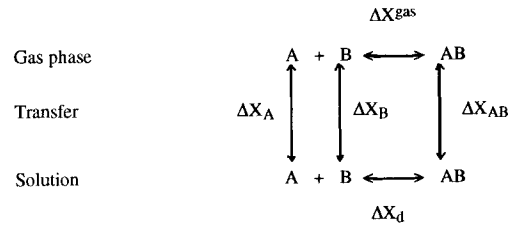

Fig. 2. A thermodynamic cycle for complex formation. ΔX stands for ΔH, ΔS, ΔG, or ΔC.

balance of hydration terms calculated independently for the complex and its isolated components:

$$\Delta X^{hyd} = \Delta X_A + \Delta X_B - \Delta X_{AB}. \qquad (5)$$

Any conformation change is assumed to occur identically in solution and in gas phase. The vertical lines represent the transfer of the complex and of its components from gas phase to the solution. They account for all protein–solvent and solvent–solvent interactions, whereas interactions between macromolecules are evaluated exclusively in the gas phase. There, the reaction involves four conceptual steps:

1. making one molecule out of two,
2. establishing electrostatic and van der Waals interactions,
3. adjusting (or changing) conformations,
4. immobilizing side chain and main chain atoms.

Note that steps (1) and (2) describe rigid body association and (3) and (4) describe departure from the rigid body approximation.

Except for (1), protein folding goes through the very same steps. Thus, we should be able to use theoretical methods established for protein folding when evaluating thermodynamic parameters for the assembly of complexes. ΔH^{gas}, the enthalpy change associated with steps (2) and (3), is the potential energy of interaction between A and B in vacuo. It can be calculated by molecular mechanics and conventional energy refinement. Results of such calculations are quoted in Table I for three complexes in which the protein components are known to associate more or less like rigid bodies. The relatively small enthalpic cost of conformation changes is not separated from the van der Waals contribution E^{vdw}, 60–80 kcal mol^{-1}, and the electrostatic potential energy of interaction E^{elec}. The latter arises mostly from the 10–15 hydrogen bonds and salt bridges formed in each complex, and it is less in the subtilisin–eglin complex, which has no salt bridge, than in the other two complexes, which have two or three. Two sets of numbers are listed for E^{elec}. Those calculated with a dielectric permittivity ε = 1 are excessively large. With a more realistic (gas phase) ε = 3, which takes into account atomic polarizabilities and the dielectric relaxation within molecules, E^{elec} is in the range 30–100 kcal mol^{-1}, still a large and definitely not a reliable number. More sophisti-

cated techniques (recently reviewed[21]) may be applied to handle dielectric relaxation effects, but they would not help with the fact that E^{elec} is very sensitive to small displacements of electric charges.

ENTROPY CHANGES

Formally, neither step (2) nor (3) changes the system entropy. ΔS^{gas}, the entropy of dissociation in gas phase, is attributable to steps (1) and (4), which impose restraints on atomic movements as molecules come into contact. Though entropy is notoriously difficult to calculate, this can be done for the vibrational component by normal mode analysis within the harmonic approximation. The entropy of free side chain is approximately known from statistics on rotamer libraries or empirical calibrations.[22,23] Only a fraction of that entropy is lost by side chains at interfaces, for there are few side chains that are really free to rotate in the dissociated components, and not all of them become completely immobilized in the complex, but it can also be estimated.

The entropy change associated with step (1), making one molecule out of two in gas phase, is

$$\Delta S^{rt} = S^{rt} - \Delta S^{vib}. \quad (6)$$

S^{rt} is the entropy associated with the six degrees of rotational/translational freedom lost, and ΔS^{vib} the vibrational entropy change associated with the new low frequency modes that appear in the complex. The Sackur–Tetrode equation yields a value of S^{rt} of the order of 100 cal mol^{-1} K^{-1} for a small globular protein in 1 mol liter^{-1} standard state at 300 K. This number depends on the shape of the molecule and it increases slowly (logarithmically) with its size. We estimated ΔS^{vib}, the compensating vibrational entropy change, to be $\approx$50 cal mol^{-1} K^{-1} based on the average amplitude of atomic movements observed in crystalline proteins.[24] The balance then contributes $T\Delta S^{rt} \approx 15$ kcal mol^{-1} in favor of dissociation at 300 K.

Our estimate—I should probably say our guess—of $T\Delta S^{rt}$ has been used in several theoretical studies. Nevertheless, it is not accepted by all authors and other estimates range between 0 and 30 kcal mol^{-1}. Some calculations ignore ΔS^{rt} altogether.[4,6] Then, the entropy and free enthalpy of dissociation are no longer dependent on the standard state, and they may not be compared to experimental values, which are. Others consider that the Sackur–Tetrode equation does not apply to macromolecules.[5] However, it has a sound theoretical basis, it correctly predicts the entropy of rare gases and that of small molecules both in gas phase and in solution.[25] There is no reason to exclude proteins from its field of application. Authors who use the Sackur–Tetrode equation for S^{rt} still differ on the size of the compensating term ΔS^{vib}. Some allow for more atomic mo-

TABLE I. Thermodynamic Parameters for Complex Dissociation

Complex	Fab HyHEL5 –lysozyme	Subtilisin–leech eglin	Barnase –barstar
Interface area*			
Total (Å^2)	1680	1480	1590
Aliphatic (%)	43	56	40
Aromatic (%)	13	6	14
Polar (%)	44	36	46
Gas phase parameters			
ΔH^{gas} (kcal mol^{-1})†			
E^{vdw}	55 59	62 64	75 82
E^{elec}	235 66	100 31	312 100
$-T\Delta S^{rt}$ (kcal mol^{-1})‡	−15	−15	−15
$-T\Delta S^{sc}$ (kcal mol^{-1})§	−24	−15	−18
$-T\Delta S^{vib}$ (kcal mol^{-1})**	−25		
Hydration parameters††			
ΔG^{hyd} (kcal mol^{-1})			
Nonpolar	16	18	12
Polar	−69	−50	−63
ΔH^{hyd} (kcal mol^{-1})			
Nonpolar	−15	−13	−15
Polar	−91	−66	−85
ΔC^{hyd} (kcal mol^{-1} K^{-1})	0.26	0.29	0.21
Overall‡‡			
$\Delta G^{gas} + \Delta G^{hyd}$	8		
$\Delta H^{gas} + \Delta H^{hyd}$	19	16	82
Experimental§§			
ΔG_d (kcal mol^{-1})	14.5	13.1	18.0
ΔH_d (kcal mol^{-1})	22.6		
ΔC_d (kcal mol^{-1} K^{-1})	0.34		

*The interface area of a complex is $A_1 + A_2 - A_{12}$, where A_{12} is the accessible surface area of the complex, A_1 and A_2, those of its components taken from the same set of coordinates (Protein Data Bank files 2HFL, 1CSE, 1BGS, refs. 37,38,39). Accessible surface areas[20] were calculated with program ASA (Pr. A. Lesk, Cambridge) and a probe size of 1.4 Å. Polar atoms are N, O, S, and C in guanidinium, carbonyl, or carboxylate groups.

†Energies of van der Waals and electrostatic interaction between the two components were calculated with X-PLOR[29] after 300 cycles of conjugate gradient refinement.[30] Atoms far from the interface were kept fixed; near the interface, the main chain was constrained by an harmonic potential; side chains were free to move. Two separate minimizations were carried out with the dielectric permittivity set to $\varepsilon = 1$ (left) and $\varepsilon = 3$ (right).

‡Rotational/translational entropy change ΔS^{rt} estimated.[24] $T = 300$ K.

§Eighteen to 24 side chains become immobilized in the complexes. For each one, ΔS^{sc} is estimated from its change in accessibility and from the entropy of the free side chain.[31]

**The vibrational entropy change ΔS^{vib} was calculated by assuming that ΔC^{vib} is the difference between ΔC_d and ΔC^{hyd} and applying the empirical relationship $\Delta S^{vib} = 1.05\ \Delta C^{vib}$ proposed by Sturtevant.[28]

††Estimated at 300 K from the interface areas using proportionality coefficients.[32]

‡‡Calculated from numbers above with $\varepsilon = 3$.

§§Data at 25°C for the hen egg lysozyme–HyHEL5 antibody complex.[11] Dissociation constants at 25°C taken for the subtilisin–eglin complex[33] and for the barnase–barstar complex.[34]

bility in the complex than we do, making ΔS^{vib} larger and ΔS^{rt} smaller.[26] Computer simulations support the opposite point of view. ΔS^{vib} was recently derived by applying molecular mechanics to insulin dimerization,[27] yielding a value equal to about half our estimate of 50 cal mol^{-1} K^{-1}. Since only the harmonic part of the vibrations was taken into account, the actual entropy change should be larger. In any case, this calculation suggests that our guess[24] was not too far off.

Thus, $T\Delta S^{rt}$, the free enthalpy cost of making one molecule out of two, is ~15 kcal mol^{-1}. In the complexes listed in Table I, $T\Delta S^{sc}$, the cost of side chain immobilization, is 15–24 kcal mol^{-1} depending on the number and nature of the side chains forming the interface. Though we included a vibrational component in ΔS^{rt}, main chain immobilization has not been explicitely accounted for. The calculation done on insulin found the vibrational entropy to increase upon dimerization.[27] However, evidence from crystallographic temperature factors suggests that the mobility of main chain atoms at interfaces may be less than in the free molecules, and, therefore, association may lower their entropy. It is not strictly correct to split the entropy change into the main chain and side chains, since their movements are tightly coupled. Nevertheless, we did so in the case of the HyHEL5–lysozyme complex, just to show that estimates of the entropic contribution to the free enthalpy in gas phase can vary from 30 to 64 kcal mol^{-1}.

HYDRATION AND THE HEAT CAPACITY OF DISSOCIATION

Semiempirical estimates of free enthalpy components, of which $T\Delta S^{rt}$ is an example, are less satisfying to the physical chemist than analytical calculations, but they sometimes have smaller error bars. This remark certainly applies to the transfer step. In spite of the great complexity of the physical phenomenon, calibrations based on small molecule solubility and calorimetry yield useful values for thermodynamic parameters ΔG^{hyd}, ΔH^{hyd}, and ΔC^{hyd}. Transferring a protein molecule from gas phase to the solution involves hydrating all polar and non polar chemical groups on its surface. In a complex, we see from Figure 2 and Eq. (5) that only groups located at the interface are relevant. Other chemical groups are buried or make the same interactions with the solvent in the free components and in the complex. Numbers listed in Table I are based on measurements of interface areas and on proportionality coefficients[32] that distinguish between aliphatic and aromatic carbons among nonpolar groups, and between nitrogen and oxygen containing polar groups. Other coefficients[35,36] yield numbers that differ by 10–30%, which sets a range of confidence for hydration parameters.

At 300 K, the hydration free enthalpy is strongly negative for polar groups, which are highly soluble in water, and positive for the nonpolar ones, which are insoluble. The enthalpies are nevertheless all negative, as they should be since they represent the energy of interactions atoms at the interface make with water when the complex dissociate. We may now compare these interactions to those formed in gas phase between protein atoms. We see in Table I that the nonpolar component to ΔH^{hyd} is smaller in absolute value than the fraction of E^{vdw} (at least one-half) that can be attributed to nonpolar groups. This means that van der Waals interactions between atoms in a well-packed protein–protein interface are stronger on average than with water. The polar component to ΔH^{hyd} is much smaller in absolute value than the value of E^{elec} calculated with $\varepsilon = 1$, but, with $\varepsilon = 3$, the two are of comparable size. Equation (4) implies that the sum of ΔH^{gas} and ΔH^{hyd} should be equal to the experimental value of ΔH_d. The sum comes out approximately right in the case of the HyHEL5–lysozyme complex with $\varepsilon = 3$, but this can be no more than a coincidence given the error bar on each term, and the free enthalpies do not add up.

Hydration has a large heat capacity change, and therefore both ΔH^{hyd} and ΔS^{hyd} vary rapidly with temperature. As pointed out by Sturtevant,[28] hydration dominates the heat capacity change in processes involving proteins. Degrees of freedom lost in step (1) are worth no more than $3R = 6$ cal mol^{-1} K^{-1}, which is small. The enthalpy of step (2), forming interactions, and step (3), conformation changes in gas phase, is formally temperature independent; steps (2) and (3) should therefore contribute little to the heat capacity. The heat capacity change in gas phase derives mostly from step (4), main chain and side chain immobilization. In other terms, ΔC^{gas} is zero in the rigid body approximation and then all of ΔC_d comes from the transfer step. In real molecules, vibrational changes may also contribute to ΔC^{gas}. We can write with Sturtevant[28]:

$$\Delta C_d \approx \Delta C^{hyd} + \Delta C^{vib} \qquad (7)$$

Nonpolar groups contribute positively to ΔC^{hyd} and polar groups contribute negatively, but much less. As in protein folding and many other processes involving biological macromolecules, the heat capacity change derives mostly from nonpolar groups being removed from contact with water. Rather than the contribution of nonpolar groups to ΔG^{hyd}, which is strongly temperature dependent, it is their contribution to ΔC^{hyd} that is the trademark of the hydrophobic effect. For the three complexes tested, ΔC^{hyd} is in the range 210–290 cal mol^{-1} K^{-1}. It is larger for the subtilisin–eglin than for the HyHEL5–lysozyme and barnase–barstar complexes, due to the high proportion of aliphatic groups in this particular interface. Experimental values of ΔC_d obtained by calorimetry on several protease–inhibitor

complexes[7,8] are near 250 kcal mol^{-1} K^{-1}, that for HyHEL5–lysozyme[11] is somewhat more, and in excess over ΔC^{hyd}. The difference (80 cal mol^{-1} K^{-1}) includes all the errors in the calculation, but as it does fit with Sturtevant's interpretation of the heat capacity change, we used it in Table I to estimate the vibrational contribution to the heat capacity and entropy of dissociation of the HyHEL5–lysozyme complex.

INTERFACE AREAS IN PROTEIN–PROTEIN AND PROTEIN–DNA COMPLEXES

The three complexes mentioned in Table I are different from almost every point of view: structure, function of the proteins, amino acid residues, and chemical groups forming the interface. They nevertheless ressemble each other in the size of the interface. It has been shown[2] that protease–inhibitor and antibody–antigen complexes form interfaces with areas in the 1200–2000 Å^2 range. Table II and the histogram in Figure 3 confirm this property and suggest that it is shared by other types of protein–protein complexes. Of 15 protease–inhibitor complexes, 13 have interface areas that are within 18% (range) or 9% (standard deviation) of an average 1470 Å^2. In five antibody–antigen complexes, the average interface areas is 1550 Å^2, in six other protein–protein complexes, it is 1500 Å^2. These values are not significantly different from that obtained for the average protease–inhibitor complex.

The smallest interface (1150 Å^2) is between cytochrome *c* peroxydase and cytochrome *c* which form a redox complex.[58] The largest by far (3300 Å^2) is between thrombin and hirudin, a potent anticlotting agent from medicinal leech. Half of the interface area derives from the N-terminal domain of hirudin binding at the active site much like other protein inhibitors do, and the other half from 17 residues in the C-terminal tail of hirudin filling the fibrinogen exo site of the protease.[53] The tail is disordered in the free protein and it undergoes a disorder-to-order transition upon binding. An engineered version of hirudin called hirulog, in which much of the structured N-terminal domain is deleted, still forms with thrombin a stable complex with a large interface.[52] The rigid-body association approximation is clearly not valid in the case of the hirudin tail, and we may assume that the larger than usual interface area is needed to compensate for the entropic cost of main chain immobilization. Trypsinogen undergoes a disorder-to-order transition of comparable importance when it binds the trypsin pancreatic inhibitor. This is reflected not in the interface area, which is the same as in the complex with trypsin,[43,44] but in the affinity: ΔG_d drops by at least 8 kcal mol^{-1}.

There is no correlation between ΔG_d and the interface area in the trypsin and trypsinogen complexes, and there ought be none in general. Yet, Horton and Lewis[78] found that they could produce a linear correlation for 15 protein–protein complexes by breaking interface areas into nonpolar and polar components and adjusting weights. Their analysis failed with the trypsinogen–trypsin inhibitor complex, and it would fail each time two structurally similar complexes have different stabilities, as happens after a point mutation for instance.[79] The correlation that they observe is nevertheless real and significant. It suggests that for proteins that have undergone natural selection to form optimal contacts, the energy of van der Waals and electrostatic interactions scales up with the size of the interface; the number of immobilized side chains also, and, since hydration parameters do so anyhow, a linear relationship between free enthalpy and interface area may well be maintained as long as the rigid body approximation applies.

The presence of packing defects in suboptimal interfaces, disorder-to-order transitions and enthalpy-costing conformation changes will destroy such a relationship. These phenomena are poorly represented in the protein–protein complexes listed in Table II. For many of them, the structure of protein components is independently known and it justifies the rigid-body association approximation. In the present data base of X-ray structures, there is nevertheless a class of protein–protein interactions where disorder-to-order transitions cannot be ignored. This is in oligomeric proteins. The smallest subunit interfaces in dimers are similar in size to those listed here,[80] and they are probably equivalent in energetic and thermodynamic terms. But most dimers and all larger oligomers have subunit interfaces that are much larger, covering 2000–4000 Å^2. These large interfaces compensate for the entropy lost by the polypeptide chains which fold as they associate. This interpretation fits with the many experiments that show strong coupling between folding and assembly in oligomeric proteins.

The hydrophobic contribution of large interfaces and the vibrational contribution associated with order-to-disorder transitions add up to make the heat capacity of dissociation in such systems much larger than for the HyHEL5–lysozyme complex. This also appears to be a general property of specific protein–DNA association,[15,81,82] whereas nonspecific complexes may have a small heat capacity.[14] We have included several protein–DNA complexes in Table II for comparison with protein–protein complexes. The smallest protein–DNA interface in Table II covers 1600 Å^2, comparable to a protein–protein interface, but most values are in the range 2200–3100 Å^2 and the average is 2800 Å^2, almost twice the average protein–protein interface. These large protein–DNA interfaces occur in assemblies where the protein component is a dimer, or a succession of three zinc fingers for Zif268. With the proteases and their inhibitors, contacts made by preformed complemen-

TABLE II. Interface Areas in Protein–Protein and Protein–DNA Complexes*

File	Complex	B (Å^2)	Ref.
Protease–inhibitor			
3SGB	*Streptomyces* protease B-ovomucoid	1250	40
1PPF	Elastase–ovomucoid	1300	41
4CPA	Carboxypeptidase A–potato inhibitor	1350	42
2PTC	Trypsin–pancreatic trypsin inhibitor	1400	43
2TGP	Trypsinogen–pancreatic trypsin inhibitor	1400	44
2KAI	Kallikrein–pancreatic trypsin inhibitor	1400	45
1CHO	Chymotrypsin–ovomucoid	1450	46
5SIC	Subtilisin–subtilisin inhibitor	1500	47
1CSE	Subtilisin–leech eglin	1500	38
1MCT	Trypsin–bitter gourd inhibitor	1550	48
1ACB	Chymotrypsin–leech eglin	1550	49
2SNI	Subtilisin–chymotrypsin inhibitor 2	1600	50
1TGS	Trypsinogen–pancreatic secretory inhibitor	1750	51
1ABI	Thrombin–hirulog	2500	52
4HTC	Thrombin–hirudin	3300	53
Antibody–antigen			
1JHL	Fv D11.15–lysozyme	1250	54
1VFB	Fv D1.3–lysozyme	1350	55
3HFM	Fab HyHEL10–lysozyme	1600	56
2HFL	Fab HyHEL5–lysozyme	1700	37
1NCD	Fab NC41–Flu virus neuraminidase	1900	57
Other protein–protein complexes			
2PCC	Cytochrome c peroxydase–yeast cytochrome *c*	1150	58
1GLC	Glycerol kinase–IIIG1c	1300	59
1FC2	Fc fragment–protein A fragment	1300	60
1BGS	Barnase–barstar	1600	39
1ATN	Actin–DNase I	1800	61
2BTF	Actin–profilin	2050	62
Protein–DNA			
2DNJ	DNase I–DNA	1600	63
1CMA	Met J repressor dimer–operator	1750	64
2BPF	Polymerase β-DNA	1900	65
1YSA	GCN4 dimer–operator	2200	66
1D66	Gal4 dimer–operator	2450	67
1ZAA	ZIF268 Zn fingers–operator	2550	68
1HDD	Engrailed homeodomain dimer–operator	2700	69
1HCR	Hin recombinase site	2750	70
1PER	Phage 434 repressor dimer–operator	2800	71
3CRO	Cro repressor dimer–operator	2900	72
1CGP	cAMP binding protein dimer–operator	2900	73
2BOP	Papilloma E2 dimer–operator	2900	74
1LMB	Lambda phage repressor dimer–operator	2900	75
1TRO	Trp repressor dimer–operator	3100	76
4RVE	*Eco*RV dimer–DNA	4300	77

*Interface areas were estimated from atomic coordinates in Protein Data Bank files as described for Table I.

tary surfaces covering about 1500 Å^2 in total are sufficient to form very stable complexes. Preformed complementary surfaces on the DNA double helix and a protein monomer cover the same area and should have the same capacity. Yet, we find that these contacts are almost always repeated by symmetry in dimers or iteration in the zinc fingers. This may be given several interpretations. One is that enough nucleotide bases must be involved in the contact for sequence-specific recognition to occur. This is probably one reason why DNase I, which is not sequence specific, makes do with a 1600 Å^2 interface, whereas the strictly specific *Eco*RV restriction enzyme uses one almost three times larger. Another reason is more fundamental. A rigid-body protein–DNA interaction may be unable to perform the required functions. Transcription regulation, which is what most of the proteins in Table II do, works by changing the conformation of DNA and possibly also its dynamics, the rate of base pair opening for instance. A high affinity is needed, and the two together can be achieved only if the protein–

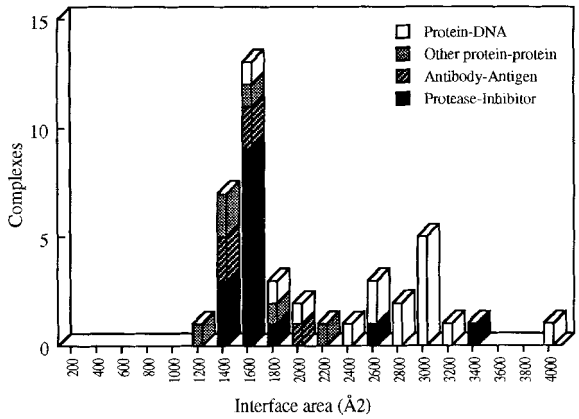

Fig. 3. Histogram of interface areas (data from Table II).

DNA interface is much larger than for blocking the active site of a protease.

CONCLUSION

DNA recognition by proteins emphasizes a point that we might have missed by looking only at protein–protein complexes. There is more in recognition than just tight binding at thermodynamic equilibrium. We have concentrated our analysis on the affinity; we should now recall that recognition also implies the capacity to discriminate between partners and alter their properties. Specificity—the discrimination between closely similar molecules—means both that A has a high affinity for B, and that it has a low affinity for C, D, E . . . , some of which ressemble B and others not. In this paper, we have only looked at how A associates with B. Discrimination against an almost endless list of chemical species has yet to be given a physicochemical basis. Moreover, the kinetic and dynamic aspects of recognition are of great functional importance. Many of the complexes considered here should not be viewed as static structures. Transcription must be turned on, then off, as well as enzymic activity. The pancreatic inhibitor, which blocks the active site of trypsin, and barstar, which does the same on a bacterial ribonuclease, bind with such high affinities that no spontaneous dissociation can take place in the pancreas or in the bacterial cell. The inhibitors suppress the hydrolytic activity of the enzymes where it would be harmful. They must nevertheless let them go free after excretion to perform their function. Thus, the story is more involved that we have assumed to now. Just as in Goethe's novel *Elective Affinities,* bonds established between pairs at Chapter One must break at some later stage and leave way to new associations—or it would be a very dull plot.

ACKNOWLEDGMENTS

My interest in protein–protein recognition owes much to long-standing interactions with Dr. C. Chothia (Cambridge, UK) and my more recent interest in calorimetry, to discussions with Dr. N. N. Khechinashvili (Pushchino, Russia). I am grateful to F. Zimmermann (Gif-sur-Yvette, France) for performing some of the calculations reported here.

REFERENCES

1. Chothia, C., Janin, J. Principles of protein-protein recognition. Nature (London) 256:705–708, 1975.
2. Janin, J., Chothia, C. The structure of protein-protein recognition sites. J. Biol. Chem. 265:16027–16030, 1990.
3. Novotny, J., Bruccoleri, R.E., Saul, F.A. On the attribution of energy in antigen-antibody complexes McPC 603, D1.3 and HyHEL5. Biochemistry 28:4735–4749, 1989.
4. Murphy, K.P., Xie, D., Garcia, K.C., Amzel, L.M., Freire, E. Structural energetics of peptide recognition: Angiotensin II/Antibody binding. Proteins 15:113–120, 1993.
5. Murphy, K.P., Xie, D., Thompson, K.S., Amzel, L.M., Freire, E. Entropy in biological binding processes: Estimation of translational entropy loss. Proteins 18:63–67, 1994.
6. Miyamoto, S., Kollman, P.A. Absolute and relative binding free energy calculations of the interaction of biotin and its analogs with streptavidin using molecular dynamics/free energy perturbation approaches. Proteins 16:226–245, 1993.
7. Takahashi, K., Fukada, H. Calorimetric studies of the binding of Streptomyces subtilisin inhibitor to subtilisin of Bacillus subtilis strain N'. Biochemistry 24:297–300, 1985.
8. Fukada, H, Takahashi, K., Sturtevant, J.M. Thermodynamics of the binding of Streptomyces subtilisin inhibitor to α-chymotrypsin. Biochemistry 24:5109–5115, 1985.
9. Tello, D., Goldbaum, F.A., Mariuzza, R.A., Ysern, X., Schwarz, F.P., Poljack, R.J. Three-dimensional structure and thermodynamics of antigen binding by anti-lysozyme antibodies. Biochem. Soc. Trans. 21:943–946, 1993.
10. Bhat, T.N., Bentley, G.A., Boulot, G., Greene, M.I., Tello, D., Dall'Acqua, W., Souchon, H., Schwarz, F.P., Mariuzza, R.A., Poljak, R.J. Bound water molecules and conformational stabilization help mediate an antigen-antibody association. Proc. Natl. Acad. Sci. U.S.A. 91:1089–1093, 1994.
11. Hibbits, K.A., Gill, D.S., Willson, R.C. Isothermal titration calorimetric study of the association of hen egg lysozyme and the anti-lysozyme antibody HyHEL5. Biochemistry 33:3584–3590, 1994.
12. Schwarz, F.P., Puri, K.D., Bhat, R.G., Surolia, A. Thermodynamics of monosaccharide binding to concanavalin A, a pea (Pisum sativum) lectin, and lentil (Lens culinaris) lectin. J. Biol. Chem. 268:7668–7677, 1993.
13. Mandal, D.K., Kishore, N., Brewer, C.F. Thermodynamics of lectin-carbohydrate interactions. Titration calorimetry measurements of the binding of N-linked carbohydrates and ovalbumin to concanavalin A. Biochemistry 33:1149–1156, 1994.
14. Takeda, Y., Ross, P.D., Mudd, C.P. Thermodynamics of Cro protein-DNA interactions. Proc. Natl. Acad. Sci. U.S.A. 89:8180–8184, 1992.
15. Ladbury, J.E., Wright, J.G., Sturtevant, J.M., Sigler, P.B. A thermodynamic study of the trp repressor-operator interaction. J. Mol. Biol. 238:669–681, 1994.
16. Privalov, P.L., Potekhin, S.A. Scanning microcalorimetry in studying temperature-induced changes in proteins. Methods Enzymol. 131:4–51, 1986.
17. Privalov, P.L., Gill, S.J. Stability of protein structure and hydrophobic interaction. Adv. Prot. Chem. 39:193–234, 1988.
18. Atkins, P.W. "Physical Chemistry," 4th ed. Oxford: Oxford University Press, 1990.
19. Kauzmann, W. Some factors in the interpretation of protein denaturation. Adv. Protein Chem. 14:1–63, 1959.
20. Lee, B.K., Richards, F.M. Solvent accessibility of groups in proteins. J. Mol. Biol. 55:379–400, 1971.
21. Sharp, K.A. Electrostatic interactions in macromolecules. Curr. Op. Struct. Biol. 4:234–239, 1994.
22. Creamer, T.P., Rose, G.D. Side-chain entropy opposes α-helix formation but rationalizes experimentally determined

helix-forming propensities. Proc. Natl. Acad. Sci. U.S.A. 89:5937–5941, 1992.
23. Sternberg, M.J.E., Chickos, J.S. Protein side-chain conformational entropy derived from fusion data-comparison with other empirical scales. Protein Eng. 7:149–155, 1994.
24. Finkelstein, A.V., Janin, J. The price of lost freedom: Entropy *of bimolecular* complex formation. Protein Eng. 3:1–3, 1989.
25. Page, M.I., Jencks, W.P. Entropic contributions to rate accelerations in enzymic and intramolecular reactions and the chelate effect. Proc. Natl. Acad. Sci. U.S.A. 68:1678–1683, 1971.
26. Erickson, H.P. Cooperativity in protein-protein association. The structure and stability of the actin filament. J. Mol. Biol. 206:465–474, 1989.
27. Tidor, B., Karplus, M. The contribution of vibrational entropy to molecular association. J. Mol. Biol. 238:405–414, 1994.
28. Sturtevant, J.M. Heat capacity and entropy changes in processes involving proteins. Proc. Natl. Acad. Sci. U.S.A. 74:2236–2240, 1977.
29. Brünger, A.T. X-PLOR Manual, Yale University, 1990.
30. Cherfils, J., Duquerroy, S., Janin, J. Protein-protein recognition analyzed by docking simulation. Proteins 11:271–280, 1991.
31. Pickett, S.D., Sternberg, M.J.E. Empirical scale of side chain conformational entropy in protein folding. J. Mol. Biol. 231:825–839, 19
32. Ooi, T., Oobatake, M., Némethy, G., Scheraga, H.A. Accessible surface areas as a measure of the thermodynamic parameters of hydration of peptides. Proc. Natl. Acad. Sci. U.S.A. 84:3086–3090, 1987.
33. Ascenzi, A., Amiconi, G., Menagatti, E., Guarneri, M., Bolognesi, M., Schnaebli, H.P. Binding of the recombinant proteinase inhibitor Eglin C from leech Hirudo medicinalis to human leukocyte elastase, bovine α-chymotrypsin and subtilisin Carlsberg: Thermodynamic study. J. Enzyme Inhib. 2:167–172, 1988.
34. Hartley, R.W. Directed mutagenesis of barnase-barstar recognition. Biochemistry 32:5978–5984, 1993.
35. Privalov, P.L., Makhatadze, G.I. Contribution of hydration and non-covalent interactions to the heat capacity effect on protein unfolding. J. Mol. Biol. 224:715–723, 1992.
36. Spolar, R.S., Livingstone, J.R., Record, M.T. Jr. Use of liquide hydrocarbon and amide transfer data to estimate contributions to thermodynamic functions of protein folding from the removal of non polar and polar surface from water. Biochemistry 31:3947–3955, 1992.
37. Sheriff, S., Silverton, E. W., Padlan, E. A., Cohen, G. H., Smith-Gill, S. J., Finzel, B. C., Davies D. R. Three-dimensional structure of an antibody-antigen complex. Proc. Natl. Acad. Sci. U.S.A. 84:8075–8079, 1987.
38. Bode, W., Papamokos, E., Musil, D. The high-resolution X-ray crystal structure of the complex formed between subtilisin Carlsberg and Eglin C, an elastase inhibitor from the leech Hirudo medicinalis. Structural analysis, subtilisin structure and interface geometry. Eur. J. Biochem. 166:673–692, 1987.
39. Guillet, V., Lapthorn, A., Hartley, R.W., Mauguen, Y. Recognition between a bacterial ribonuclease, barnase, and its natural inhibitor, barstar. Structure 1:165–177, 1993.
40. Read, R.J., Fujinaga, M., Sielecki, R., James, M.N.J. Structure of the complex of Streptomyces griseus protease B and the third domain of the turkey ovomucoid inhibitor at 1.8 Å resolution. Biochemistry 22:4420–4433, 1983.
41. Bode, W., Wei, A. Z., Huber, R., Meyer, E., Travis, J., Neumann, S. X-Ray crystal structure of the complex of human leukocyte elastase (Pmn elastase) and the third domain of the turkey ovomucoid inhibitor. EMBO J. 5:2453–2458, 1986.
42. Rees, D.C., Lipscomb, W.N. Refined crystal structure of potato inhibitor complex of carboxypeptidase A at 2.5 Å resolution. J. Mol. Biol. 160:475–498, 1982.
43. Huber, R., Kukla, D., Bode, W., Schwager, P., Bartels, K., Deisenhofer, J., Steigemann, W. Structure of the complex formed by bovine trypsin and bovine pancreatic trypsin inhibitor. II Crystallographic refinement at 1.9 Å resolution. J. Mol. Biol. 89:73–101, 1974.
44. Marquart, M., Walter, J., Deisenhofer, J., Bode, W., Huber, R. The geometry of the reactive site and of the peptide groups in trypsin, trypsinogen and its complexes with inhibitors. Acta Crystallogr. B39:480–492, 1983.
45. Chen, Z., Bode, W. Refined 2.5 Å X-ray crystal structure of the complex formed by porcine kallikrein A and the bovine pancreatic trypsin inhibitor. Crystallization, Patterson search, structure determination, refinement, structure and comparison with its components and with the bovine trpsin-bovine trypsin inhibitor complex. J. Mol. Biol. 164: 283–311, 1983.
46. Fujinaga, M., Sielecki, R., Read, R.J., Ardelt, W., Laskowski, M., Jr., James, M.N.J. Crystal and molecular structures of α-chymotrypsin with its inhibitor turkey ovomucoid third domain at 1.8 Å resolution. J. Mol. Biol. 195: 397–418, 1987.
47. Takeuchi, Y., Noguchi, S., Satow, Y., Kojima, S., Kumagai, I., Miura, K.I., Nakamura, K.T., Mitsui, Y. Molecular recognition at the active site of subtilisin BPN′: Crystallographic studies using genetically engineered proteinaceous inhibitor SSI (Streptomyces subtilisin inhibitor). Protein Eng. 4:501–508, 1991.
48. Huang, Q., Liu, S., Tang, Y. The refined 1.6 angstroms resolution crystal structure of the complex formed between porcine beta-trypsin and MCTI-A, a trypsin inhibitor of squash family. J. Mol. Biol. 229:1022–1036, 1993.
49. Frigerio, F., Coda, A., Pugliese, L., Lionetti, C., Menegatti, E., Amiconi, G., Schnebli, H. P., Ascenzi, P., Bolognesi M. Crystal and molecular structure of the bovine alpha-chymotrypsin-Eglin C complex at 2.0 Å resolution. J. Mol. Biol. 225:107–123, 1992.
50. McPhalen, C.A., James, M.N.J. Structural comparison of two serine proteinase-protein inhibitor complexes. Eglin C-subtilisin Carlsberg and CI2-subtilisin novo. Biochemistry 27:6582–6598, 1988.
51. Bolognesi, M., Gatti, G., Menegatti, E., Guarneri, M., Marquart, M., Papamokos, E., Huber, R. Three-dimensional structure of the complex between pancreatic secretory inhibitor (Kazal type) and trypsinogen at 1.8 Å resolution. Structure solution, crystallographic refinement and preliminary structural interpretation. J. Mol. Biol. 162:839–868, 1982.
52. Qiu, X., Pabmanabhan, K.P., Carperos, V.E., Tulinsky, A., Kline, T., Maraganore, J.M., Fenton, J.W. II. Structure of the hirulog 3-thrombin complex and nature of the S′ subsites of substrates and inhibitors. Biochemistry 31:11689–11697, 1992.
53. Rydel, T.J., Tulinsky, A., Bode, W., Huber, R. The refined structure of the hirudin-thrombin complex. J. Mol. Biol. 221:583–601, 1991.
54. Chitarra, V., Alzari, P.M., Bentley, G.A., Bhat, T.N., Eisele, J. L., Houdusse, A., Lescar, J., Souchon, H., Poljak, R. J. Three-dimensional structure of a heteroclitic antigen-antibody cross-reaction complex. Proc. Natl. Acad. Sci. U.S.A. 90:7711–7715, 1993.
55. Amit, A.G., Mariuzza, R.A., Phillips, S.E.V., Poljak, R.J. Three-dimensional structure of an antigen-antibody complex at 2.8 Å resolution. Science 233:747–753, 1986.
56. Padlan, E.A., Silverton, E.W., Sheriff, S., Cohen, G.H., Smith-Gill, S.J., Davies, D.R. Structure of an antibody-antigen complex. Crystal structure of the HyHEL-10 Fab-lysozyme complex. Proc. Natl. Acad. Sci. U.S.A. 86:5938–5942, 1989.
57. Tulip, W.R., Varghese, J.N., Laver, W.G., Webster, R.G., Colman, P.M. Refined crystal structure of the influenza virus N9 neuraminidase-NC41 Fab complex. J. Mol. Biol. 227:122–148, 1992.
58. Pelletier, H., Kraut, J. Crystal Structure of a complex between electron transfer partners, cytochrome c peroxidase and cytochrome c. Science 258:1748–1755, 1992.
59. Hurley, J.H., Faber, H.R., Worthylake, D., Meadow, N.D., Roseman, S., Pettigrew, D.W., Remington, S.J. Structure of the regulatory complex of *E. coli* III^{Glc} with glycerol kinase. Science 259:673–677, 1993.
60. Deisenhofer, J. Crystallographic refinement and atomic models of a human Fc fragment and its complex with fragment B of protein A from Staphylococcus aureus at 2.9 and 2.8 Å resolution. Biochemistry 20:2361–2370, 1981.
61. Kabsch, W., Mannherz, H.G., Suck, D., Pai, E.F., Holmes K.C. Atomic structure of the actin:DNAse I complex. Nature (London) 347:37–44, 1990.
62. Schutt, C.E., Myslik, J.C., Rozycki, M.D., Goonesekere,

N.C.W., Lindberg U. The structure of crystalline profilin-beta-actin. Nature (London) 365:810–816, 1993.
63. Lahm, A., Suck D. DNase I-induced DNA conformation. 2 Å structure of a DNAse I-octamer complex. J. Mol. Biol. 222:645–667, 1991.
64. Somers, W.S., Phillips, S.E.V. Crystal structure of the met repressor-operator complex at 2.8 Å resolution reveals DNA recognition by β-strands. Nature (London) 359:387–393, 1992.
65. Pelletier, H., Sawaya, M.R., Wilson, S.H., Kraut, J. Structures of ternary complexes of rat DNA polymerase β, a DNA template-primer, and ddCTP. Science 264:1891–1903, 1994.
66. Ellenberger, T.E., Brandl, C.J., Struhl, K., Harrison, S.C. The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted alpha helices: Crystal structure of the protein-DNA complex. Cell 71:1223–1237, 1992.
67. Marmorstein, R., Carey, M., Ptashne, M., Harrison, S.C. DNA recognition by Gal4: Structure of a protein/DNA complex. Nature (London) 356:408–414, 1992.
68. Pavletich, N.P., Pabo, C.O. Zinc finger-DNA recognition: Crystal structure of a Zif268-DNA complex at 2.1 Å. Science 252:809–817, 1991.
69. Kissinger, C.R., Liu, B., Martin-Blanco, E. Kornberg, T.B., Pabo, C.O. Crystal structure of an engrailed homeodomain-DNA complex at 2.8 Å remsolution: A framework for understanding homeodomain-DNA interactions. Cell 63:579–590, 1990.
70. Feng, J.A., Johnson, R.C., Dickerson, R.E. Hin recombinase bound to DNA: The origin of specificity in major and minor groove interactions. Science 263:348–355, 1994.
71. Rodgers, D.W., Harrison, S.C. The complex between phage 434 repressor DNA-binding domain and operator site OR3: Structural differences between consensus and non-consensus half-sites. Structure 1:227–240, 1993.
72. Mondragon, A., Harrison, S.C. The phage 434 Cro/OR1 complex at 2.5 Å resolution. J. Mol. Biol. 219:321–334, 1991.
73. Schultz, S.C., Shields, G.C., Steitz, T.A. Crystal structure of a CAP-DNA complex: The DNA is bent by 90 degrees. Science 253:1001–1007, 1991.
74. Hegde, R.S., Grossman, S.R., Laimins, L.A., Sigler, P.B. Crystal structure at 1.7 Å of the bovine papillomavirus-1 E2 DNA-binding domain bound to its DNA target. Nature (London) 359:505–512, 1992.
75. Beamer, L.J., Pabo, C.O. Refined 1.8 Å crystal structure of the lambda repressor-operator complex. J. Mol. Biol. 227: 177–196, 1992.
76. Otwinowski, Z., Schevitz, R.W., Zhang, R.G., Lawson, C.L., Joachimiak, A.J., Marmorstein, R., Luisi, B.F., Sigler, P.B. Crystal structure of Trp repressor operator complex at atomic resolution. Nature (London) 335:321–329, 1988.
77. Winkler, F.K., Banner, D.W., Oefner, C., Tsernoglou, D., Brown, R.S., Heathman, S.P., Bryan, R.K., Martin, P.D., Petratos, K., Wilson, K.S. The crystal structure of Eco RV endonuclease and of its complexes with cognate and non-cognate DNA fragments. EMBO J. 12:1781–1795, 1993.
78. Horton, N., Lewis, M. Calculation of the free energy of association for protein complexes. Protein Sci. 1:169–181, 1992.
79. Ysern, X., Fields, B.A., Bhat, T.N., Goldbaum, F.A., Dall'Acqua, W., Schwarz, F.P., Poljak, R.J., Mariuzza, R.A. Solvent rearrangement in an antigen-antibody interface introduced by site-directed mutagenesis of the antibody combining site. J. Mol. Biol. 238:496–500, 1994.
80. Janin, J., Miller, S., Chothia, C. Surface, subunit interfaces and interior of oligomeric proteins. J. Mol. Biol. 204: 155–164, 1988.
81. Ha, J.H., Spolar, R.S., Record, M.T., Jr. Role of the hydrophobic effect in stability of site-specific protein-DNA complexes. J. Mol. Biol. 209:801–816, 1989.
82. Spolar, R.S., Record, M.T. Jr. Coupling of local folding to site specific binding of proteins to DNA. Science 263:777–784, 1994.