Enthalpy and entropy of protein binding to DNA

The mass equation law for binding of a protein P to its DNA

$$D_{\text{free}} + P_{\text{free}} \rightleftarrows DP \qquad K_1 = \frac{D_{\text{free}} \cdot P_{\text{free}}}{DP}$$

binding of the first proteins with the dissociation constant K_1

D_{free}, concentration free DNA; P_{free}, concentration free protein

$$\text{binding constant } K_{\text{B}} = \frac{1}{\text{dissociation constant } K_{\text{D}}}$$

How fast is binding or dissociation

$$\mathrm{AB} \underset{k_{\mathrm{on}}}{\overset{k_{\mathrm{off}}}{\rightleftharpoons}} \mathrm{A+B}$$

k_{off} in s^{-1} is the reaction rate constant for dissociation

k_{on} in $M^{-1}\ s^{-1}$ is the reaction rate constant for binding

$$\frac{k_{\mathrm{off}}}{k_{\mathrm{on}}} = K_{\mathrm{d}}$$

relation to the equilibrium dissociation constant

$$\frac{1}{k_{\mathrm{off}}} = \tau$$

life time of the complex

$$\frac{d[\mathrm{AB}]}{dt} = \mathrm{k_{on}} \cdot [\mathrm{A}] \cdot [\mathrm{B}] - \mathrm{k_{off}} \cdot [\mathrm{AB}]$$

rate equation for complex formation,

can be solved but it is already difficult

k_{on} cannot be higher than 10^8 - 10^9 $M^{-1}\ s^{-1}$ for a diffusion controlled reaction

What is the meaning of the dissociation constant for binding of a single ligand to its site?

1. K_d is a concentration and has units of mol per liter

2. K_d gives the concentration of ligand that saturates 50% of the sites (when the total sit concentration is much lower than K_D)

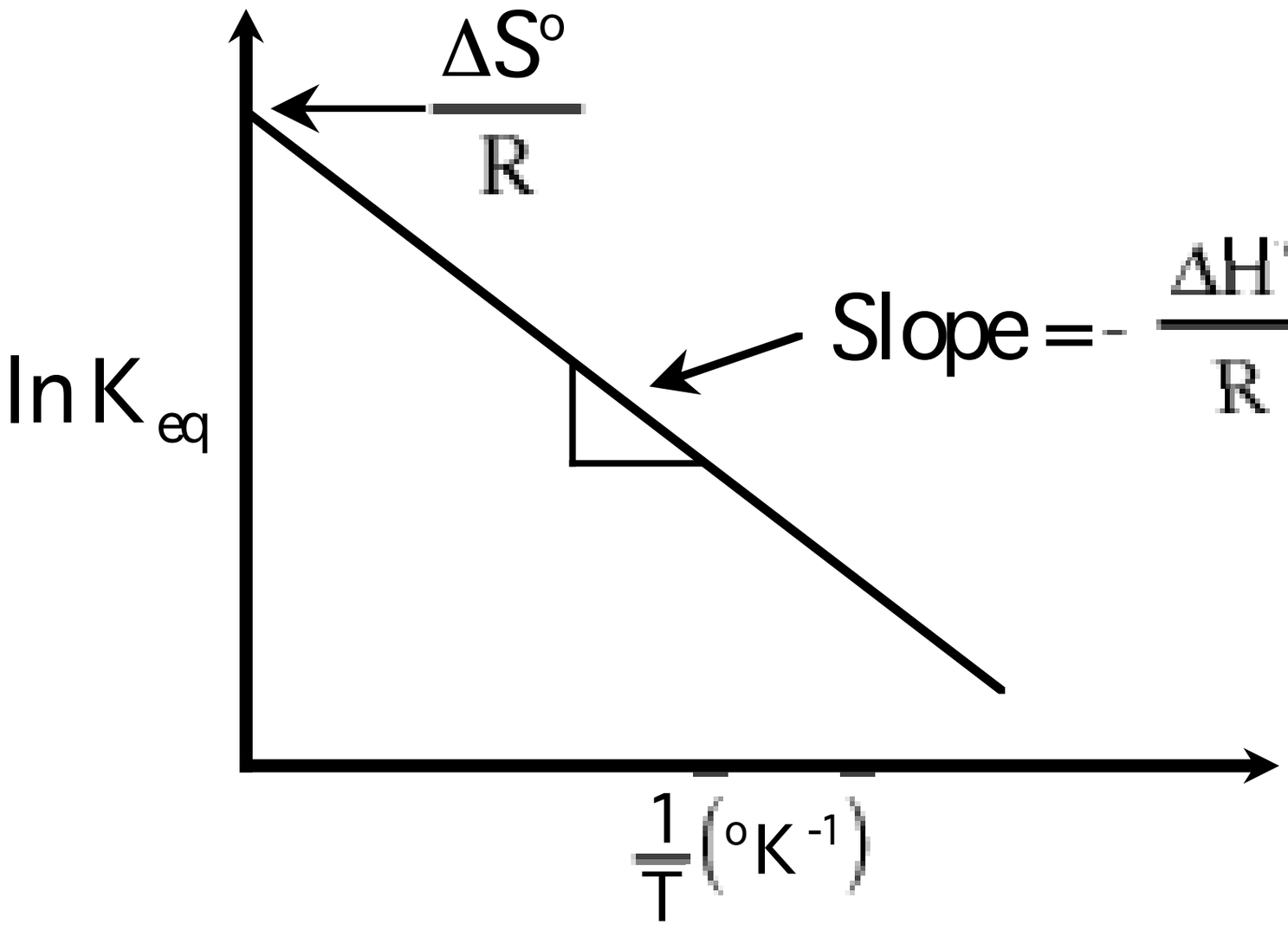
3. Almost all binding sites are saturated if the ligand concentration is 10 x K_d

4. The dissociation constant K_d is related to Gibbs free energy ΔG by the relation $\Delta G = - R\,T \ln(K_d)$

Our energy and time coordinate system

K_d (M)	**concentration scale**	**ΔG (kcal/mol)**	**k_{off} (s-1)**	**complex life time**	**Binding interaction**
			for $k_{on} = 10^5$ M^{-1} s^{-1}		
10^{-3}	1 mM	-4.1	10^{2}	10 ms	ion-DNA ion-protein
10^{-4}	0.1 mM	-5.5	10^{1}	0.1 sec	
10^{-5}	10 µM	-6.8	1	1 sec	enzyme-ligand (weak)
10^{-6}	1 uM	-8.2	10^{-1}	10 sec	protein-DNA, unspecific
10^{-7}	0.1 µM	-9.5	10^{-2}	100 sec	enzyme-ligand (strong)
10^{-8}	10 nM	-10.9	10^{-3}	16.7 min	
10^{-9}	1 nM	-12.3	10^{-4}	2.8 hours	protein-DNA specific
10^{-10}	0.1 nM	-13.6	10^{-5}	28 hours	
10^{-11}	10 pM	-15	10^{-6}	11.6 days	antibody-antigen
10^{-12}	1 pM	-16.4	10^{-7}	116 days	

The temperature dependence of the binding constants reveals ΔH and ΔS in a van't Hoff plot if ΔH and ΔS are independent of temperature



$$\Delta G = \Delta H - T \cdot \Delta S$$

$$\Delta G = -R \cdot T \cdot \ln K_{eq}$$

$$\ln K_{eq} = -\frac{\Delta H}{RT} + \frac{\Delta S}{R}$$

$$\frac{\partial(\ln K_{eq})}{\partial(1/T)} = -\frac{\Delta H}{R}$$

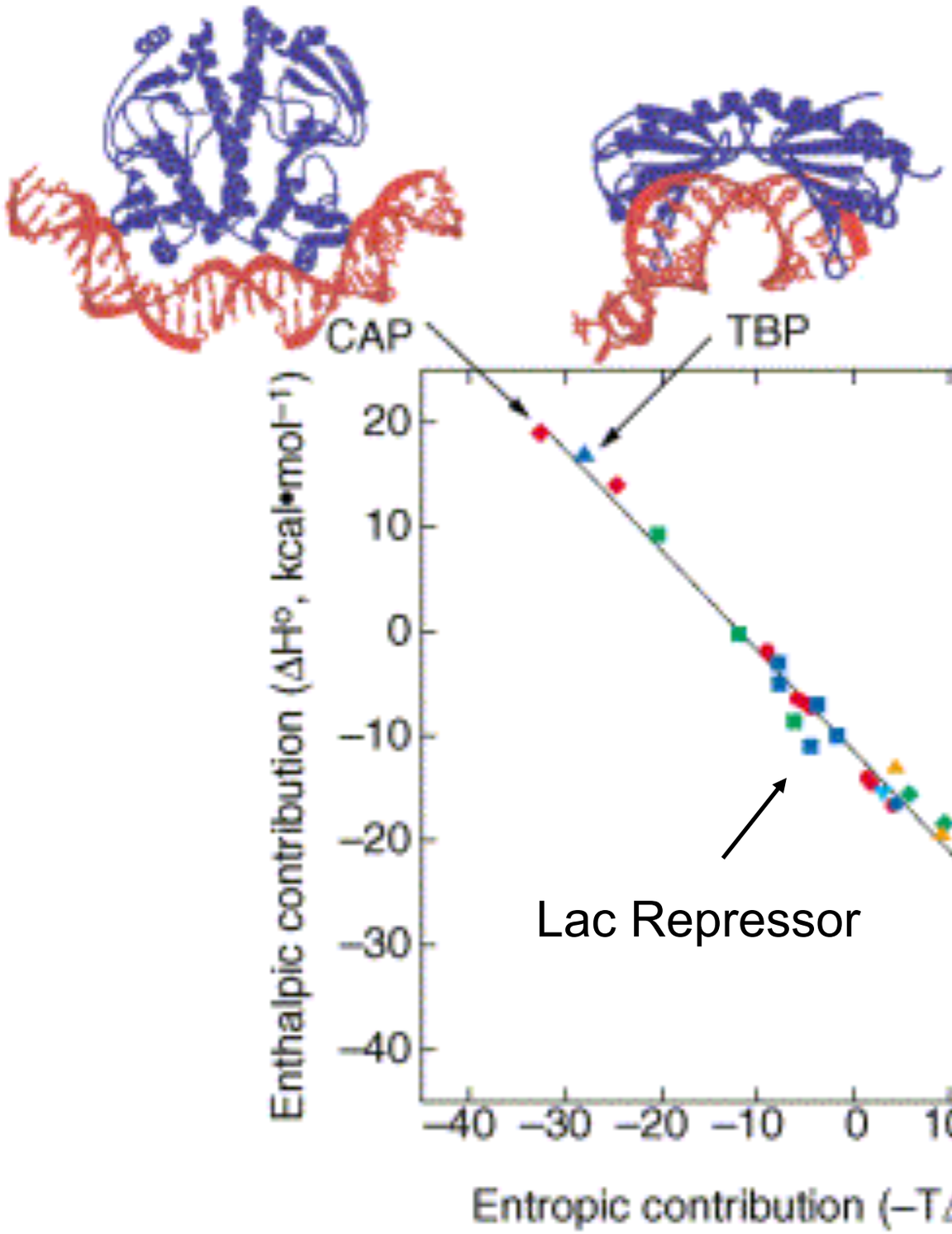
From the slope of ln K_{eq} vs. 1/T (usually from 0 to 40 °C) one can determine the ΔH and from extrapolation also ΔS . Is the van't Hoff plot curved then ΔH is temperature dependent and it can be determined from the derivative.

The unfavorable enthalpy contribution associated with DNA distortion is compensated by a favorable entropy

enthalpy

free energy⟶ $\Delta G = \Delta \mathrm{H} - T \cdot \Delta S$

entropy



Jen-Jacobson et al. (2000). Structure 8, 1015-1023

K_D and ΔG values for protein-DNA binding per site

Specific binding of a protein to DNA varies over a relatively small range of $\Delta G_{bind,sp}$ = -9 to -16 kcal/mol, with ~60 kcal/mol for ΔH and $T\Delta S$

$\Rightarrow \Delta G_{bind,sp} \approx$ const. (-11.7 ± 1.6 kcal/mol)

$\Rightarrow \Delta H = -T \cdot \Delta S$ - 11.7 kcal/mol

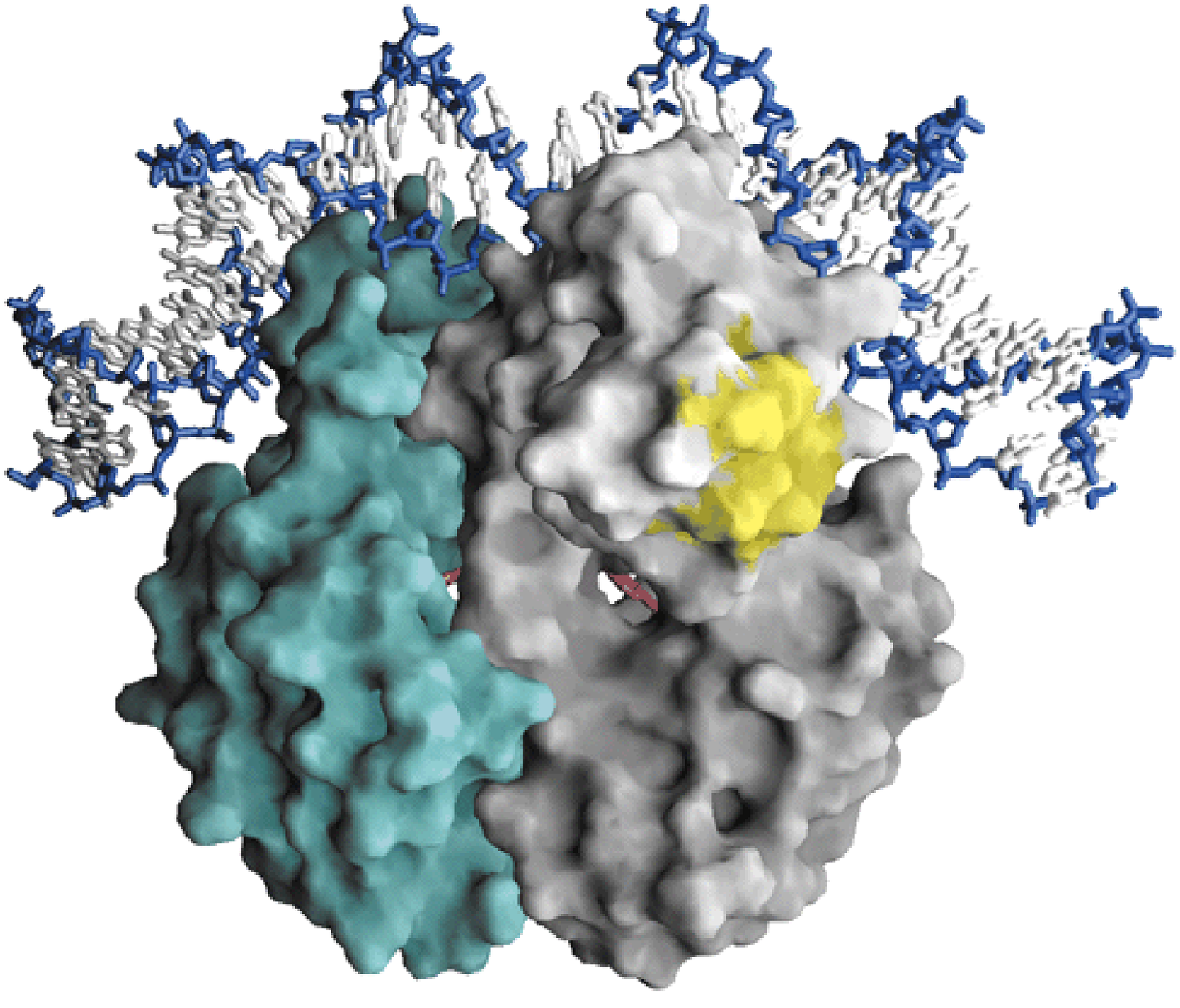
Protein needs to select specific binding site from unspecific sites

$\Rightarrow \Delta\Delta G$(specific - unspecific) ~ -5 to -9 kcal/mol

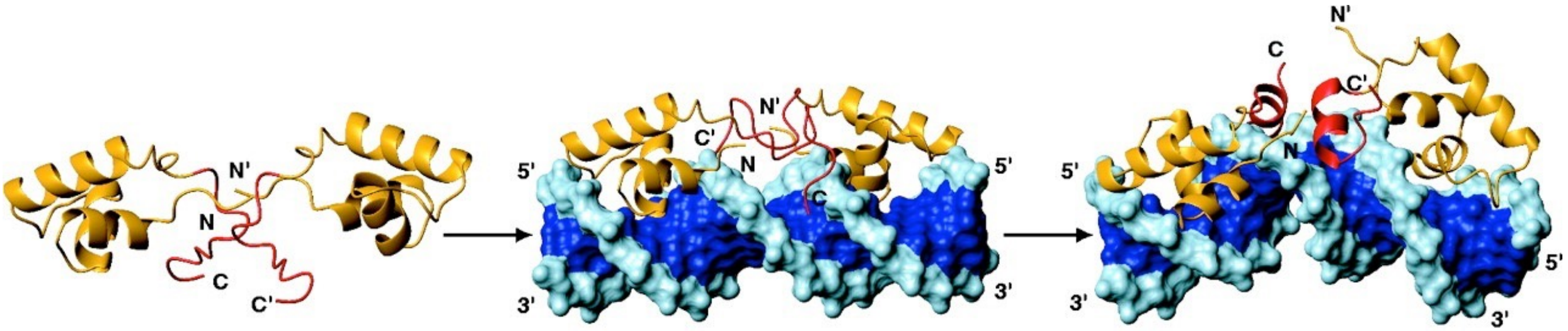
Protein binding must be reversible on the cell's time scale

$\Rightarrow \Delta G_{bind,sp} \leq$ -16 kcal/mol

Molecular structure of E. coli CRP (also called CAP for catabolite gene activator protein)



The hinge region (50-62 in red) of Lac-DBD is folded only in the specific complex with DNA



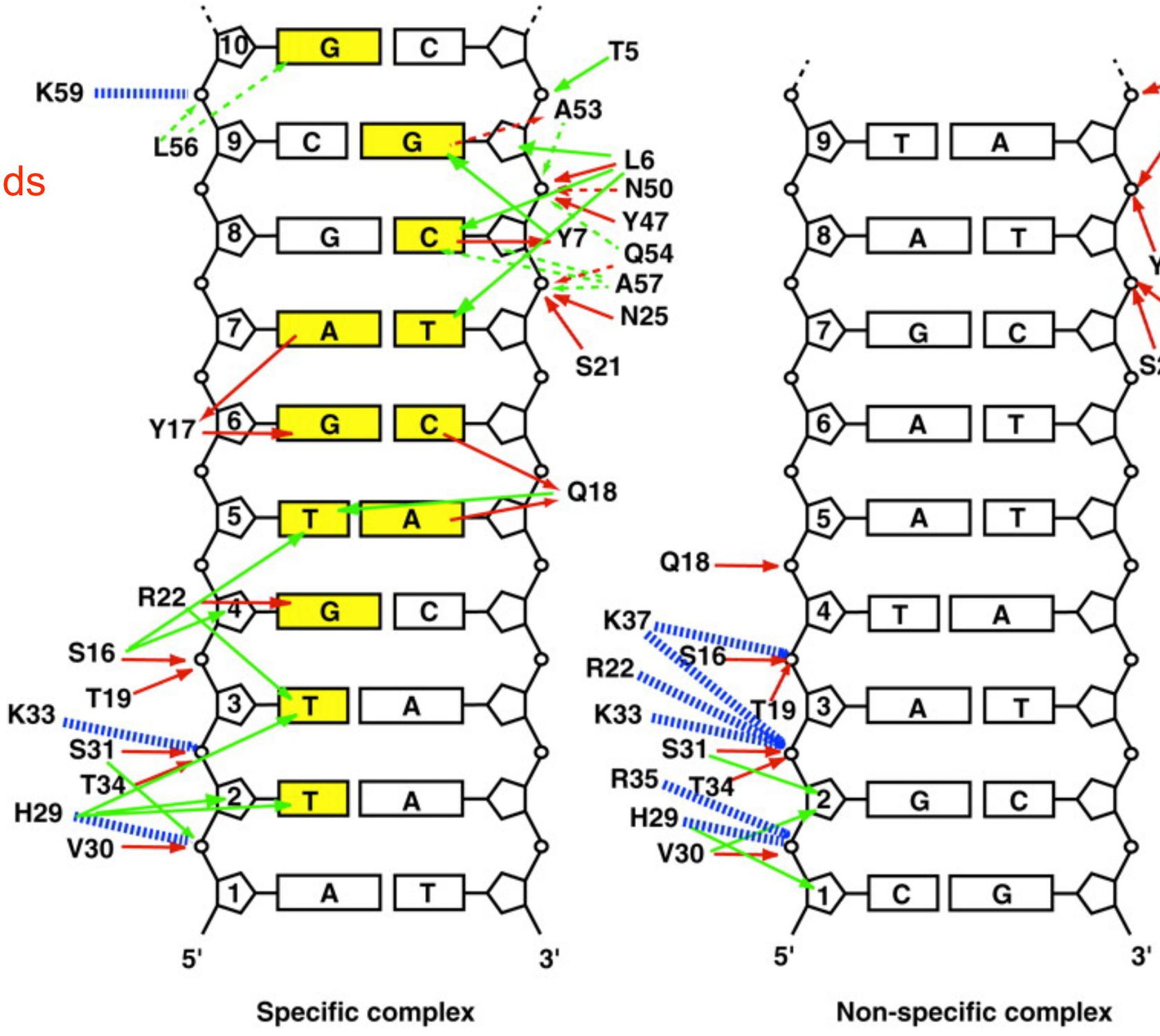
- folding of hinge region with specific contacts in minor groove
- specific interactions major groove
- less electrostatic interactions
- curvature of DNA

Specific (left) and nonspecific (right) protein-DNA contacts of Lac-DBD repressor with DNA

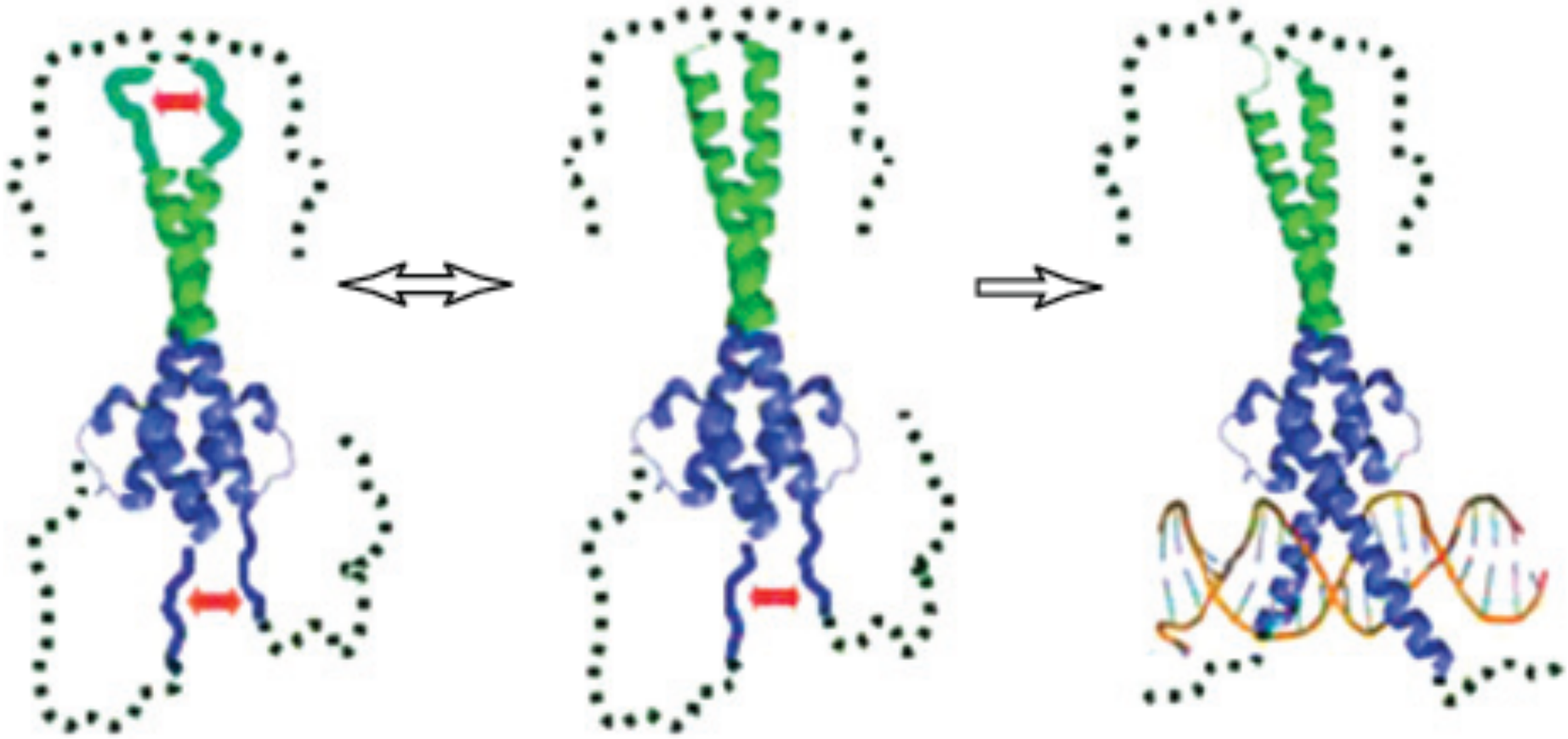
red:
hydrogen bonds

green:
hydrophobic
contacts

dashed blue:
electrostatic
contacts



Local folding of the Max transcription factor upon dimerization and binding



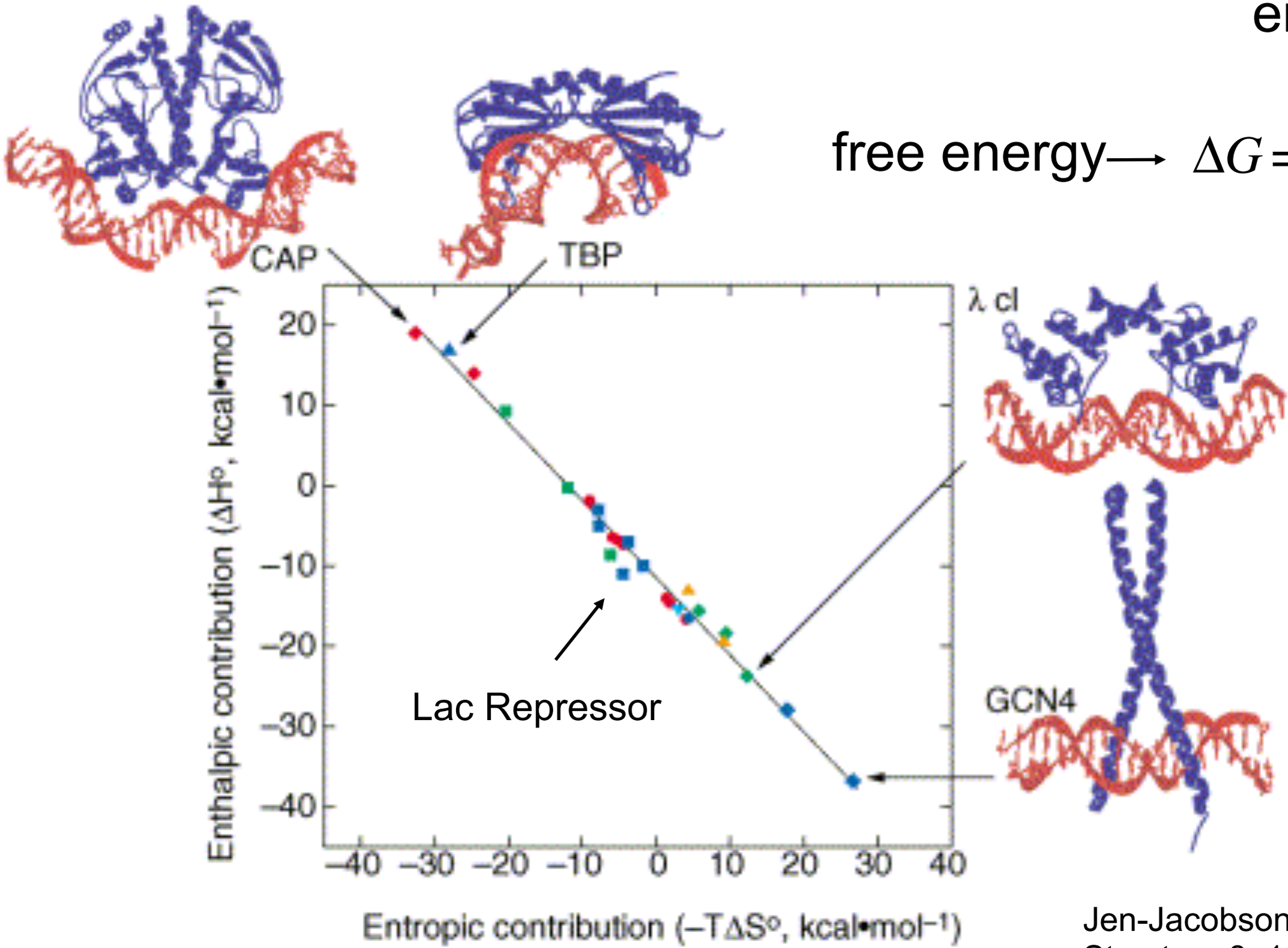
The Max transcription factor (PDBcode: 1NKP) binds DNA as a dimer. The disordered N-terminal region (upper dotted line) reduces the electrostatic repulsion (red arrows) between the two monomers, and increases the population of the folded state at the flanking leucine zipper (green). This also stabilizes the bHLH region (blue) and thus improves binding affinity for DNA.

Contribution of enthalpy and entropy to binding energies of different protein-DNA complexes

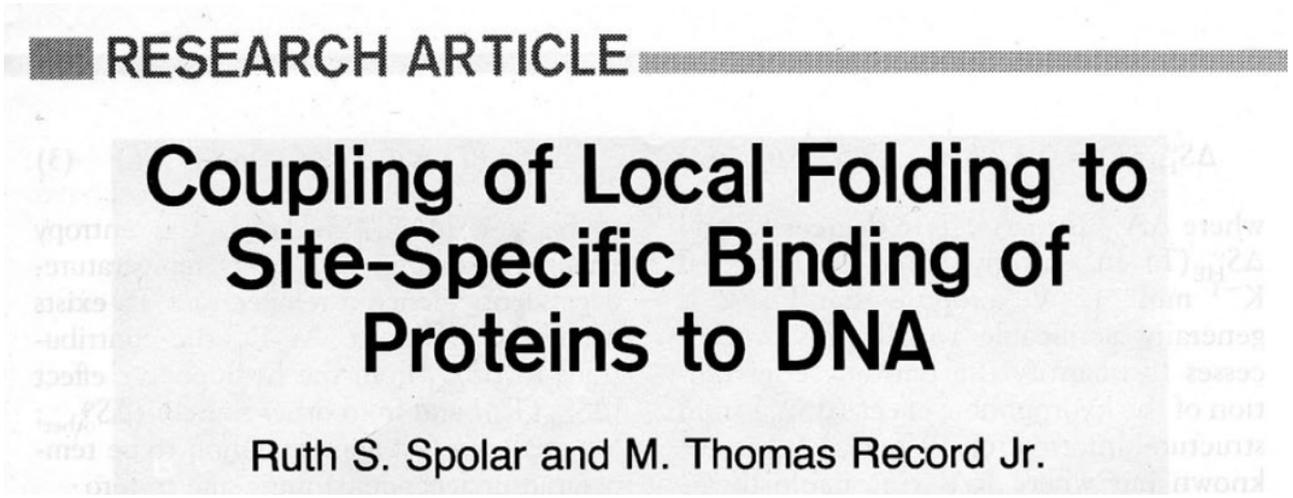
enthalpy

free energy⟶ $\Delta G = \Delta \mathrm{H} - T \cdot \Delta S$

entropy



Jen-Jacobson et al. (2000). Structure 8, 1015-1023


RESEARCH ARTICLE

Coupling of Local Folding to Site-Specific Binding of Proteins to DNA

Ruth S. Spolar and M. Thomas Record Jr.

Science 263, 777-784, 1994

The analysis of ΔS for protein binding to DNA is conducted at the characteristic temperature T_s where $\Delta S_{bin} = 0$ so that:

$$\Delta S_{bin}(T_S) = 0 = \Delta S_{HE}(T_S) + \Delta S_{rt} + \Delta S_{PE} + \Delta S_{other}$$

The term ΔS_{other} arises primarily from folding/conformational changes in the protein and/or the DNA upon specific DNA binding.

Protein folding includes two dominant and opposing contributions to the entropy:

a) One positive from the hydrophobic effect or the “release” of water on burial of nonpolar surfaces

b) One negative from the reduction in conformational entropy

$$\Delta S_{fold}(T_S) = 0 = \Delta S_{HE}(T_S) + \Delta S_{conf}$$

S&R, Table 1: Protein folding

325 cal $K^{-1}mol^{-1}$/56 residues = 5.8 cal $K^{-1}mol^{-1}$:

Table 1. Entropic contribution to protein folding from the hydrophobic effect.

Protein	$\Re$	$-\Delta C^{\circ}_{p,fold}$ (cal mol^{-1} K^{-1})	T_S* (K)	$-\Delta A_{np}$† (Å^2)	$\Delta S^{\circ}_{HE}(T_S)$‡ (e.u.)	$-\Delta s^{\circ}_{other}$§ (e.u.)
Streptococcal protein G, domain B1	56	620 (*68*)	272	2900	325	5.8
BPTI	58	720 (*24*)	306	2640	196	3.4
		400 (*69*)	221		471	8.1
Parvalbumin b	108	1100 (*70*)	268	5485	640	5.9
Ribonuclease A	124	1230 (*25*)	255	5815	771	6.2
Lysozyme (hen egg white)	129	1540 (*25*)	270	6870	786	6.1
Ferricytochrome c	104	1730 (*25*)	294	5540	483	4.6
Staphylococcal nuclease	141	1820 (*25*)	288	7880	738	5.2
Holo‖ myoglobin	153	2770 (*25*)	301	9710	773	5.1
β trypsin	223	2850 (*25*)	281	11830	1200	5.4
Papain	212	2920 (*25*)	290	12755	1167	5.5
α chymotrypsin	245	3020 (*25*)	280	14770	1517	6.2
Carbonic anhydrase	256	3820 (*25*)	290	15760	1442	5.6
Pepsinogen	370	6090 (*25*)	297	23730	1990	5.4
					Average¶	5.6 ± 0.5

*Values of T_S were calculated from values of $\Delta C^{\circ}_{p,fold}$ and ΔS°_{fold} cited in the reference indicated in column 3. Reported uncertainties in $\Delta C^{\circ}_{p,fold}$ range from 5 to 20 percent. Corresponding uncertainties in T_S range from 1 to 7 K degrees. †Calculations of ΔA_{np} model the denatured state as an extended β chain (*26*, *27*). The value of ΔA_{np} for folding the B1 domain of streptococcal protein G was calculated as described in (*27*) from Brookhaven Protein Database (*67*) file 2GB1. All other values of ΔA_{np} are from (*26*). ‡Eq. 2. §$\Delta s^{\circ}_{other} = \Delta S^{\circ}_{other}/\Re$, calculated from Eq. 4. ‖In this and subsequent tables, holo refers to the protein associated with its cofactor. ¶Not including BPTI.

$T\Delta S_{conf}$ = 5.6 cal K^{-1} mol^{-1}·298 K = 1.7 kcal mol^{-1}

Conclusion 1 from entropy analysis:

The unfavorable conformational entropy of folding per residue is

$\Delta S = -5.6$ cal mol^{-1} K^{-1}
or $T\Delta S = -1.7$ kcal mol^{-1}

ΔS_{rt}: Unfavorable entropy due to loss of movements of protein upon binding

Loss of rigid body rotational and translational entropy ΔS_{rt}

$$\Delta S_{bin}(T_S) = 0 = \Delta S_{HE}(T_S) + \Delta S_{rt}$$

Estimated from studies of entropic changes arising from rigid body protein-protein association.

S&R, Fig. 1: Rigid body association for subtilisin binding to its inhibitor protein

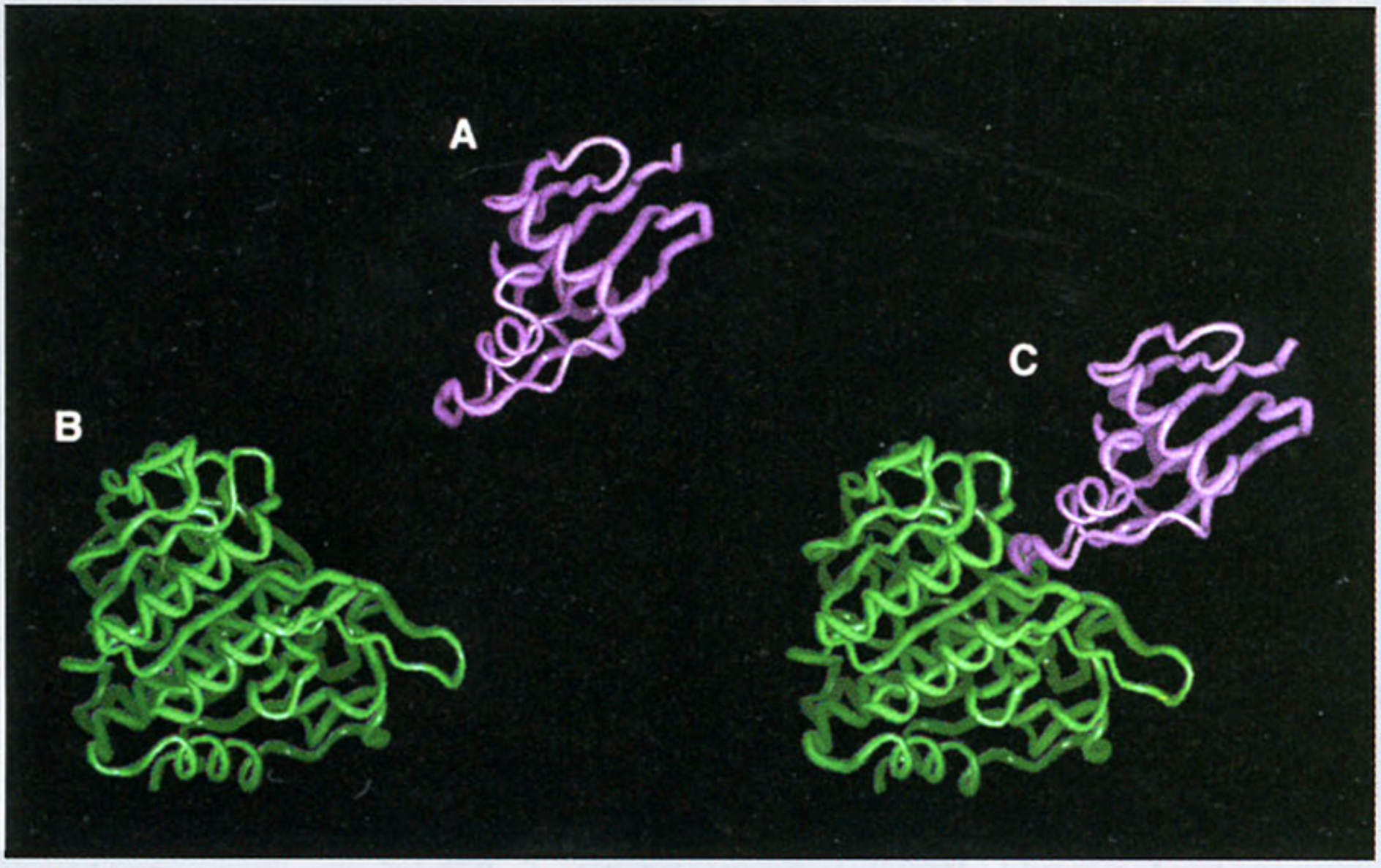


Fig. 1. Ribbon model of a "rigid-body" association. X-ray crystallographic structures of (**A**) subtilisin inhibitor monomer (purple, PDB file 2SSI) and (**B**) uncomplexed subtilisin (green, PDB file 2ST1), shown in the same orientation as in the complex. (**C**) Enzyme-inhibitor complex (PDB file 2SIC), same colors as in (A).

S&R, Table 3: Rigid body association

Table 3. Entropic contributions to "rigid body" associations.

Process	$-\Delta C^{\circ}_{assoc}$ (cal mol^{-1}K^{-1})	T_S* (K)	$\Delta S^{\circ}_{HE}(T_S)$† (e.u.)
Soybean inhibitor + trypsin → complex	440 (*83*)	349	(*60*)
Subtilisin inhibitor + subtilisin monomer → complex	240 (*71*)	339	41
Subtilisin inhibitor + α chymotrypsin monomer → complex	270 (*84*)	343	(*43*)
FK506 + FKBP-12 → complex	260 (*73*)	289	60
			Average: 50 ± 10

*Values of T_S calculated from values of ΔC°_{assoc}, ΔS°_{assoc} in references cited in column 2. †$\Delta S^{\circ}_{HE}(T_S)$ calculated from Eq. 2 with values for ΔA_{np} from Table 2. Values of $\Delta S^{\circ}_{HE}(T_S)$ in parentheses are calculated from Eq. 3 for systems lacking structural data to evaluate ΔA_{np}.

$$\Delta S_{bin}(T_S) = 0 = \Delta S_{HE}(T_S) + \Delta S_{rt}$$

Conclusion 2 from entropy analysis:

The unfavorable entropy for rigid body association of two macromolecules is

$\Delta S = -50$ cal mol^{-1} K^{-1}
or $T\Delta S = -14.9$ kcal mol^{-1}

S&R, Fig. 2: Induced folding of an avidin monomer upon binding to biotin

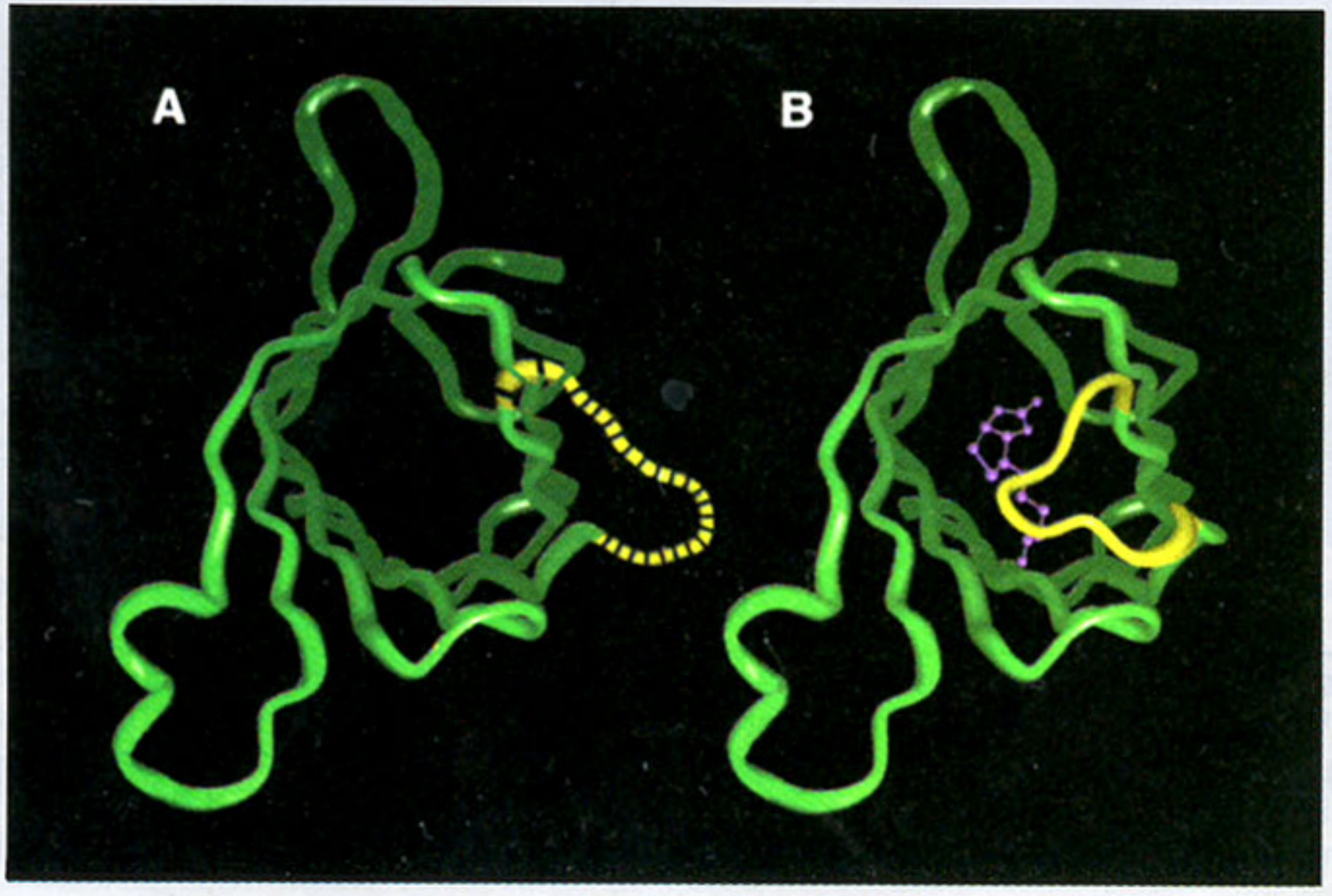


Fig. 2. Ribbon model of avidin-biotin "induced fit" interaction. (**A**) Model of the uncomplexed avidin monomer in solution (green). Residues (36–44) (dashed loop in yellow) are disordered in the free crystal structure (*49*) and are inferred to be in a flexible coil state of high conformational entropy in solution. (**B**) Avidin-biotin complex. Ordering of the looped region (yellow) upon binding encloses biotin (in purple) in a "hydrophobic box" (*49*).

S&R, Table 4: Coupled folding in protein-protein association

R (number of residues involved in folding transition)
= ΔS_{other} /-5.6 cal K^{-1} mol^{-1}

Table 4. Entropic contributions where folding is coupled to association: predictions of the number of residues participating in the folding transition.

Process (structural references)	T_S* (K)	$\Delta S^\circ_{HE}(T_S)$† (e.u.)	ΔS°_{rt}‡ (e.u.)	ΔS°_{other}§ (e.u.)	$\Re^{th}$‖	$\Re^{str}$¶
Angiotensin II (*48*) + antibody Fab 131 (*85*) → complex (*85*)	312	68	−50	−18	3	8
Avidin (*49*) + biotin → complex (*49*)	291	85	−50	−35	6	9
S-peptide (*47*) + S-protein (*47*) → ribonuclease S (*86*)	253#	145	−50	−95	17	15
L-tryptophan + apo Trp R monomer (*11*) → complex (*12*)	263	127	−50	−77	14	17**
Holo Trp R dimer (*11*) + *trp* operator DNA → complex (*12*)	319	147	−50	−97	17	16
2 GR DBD (*13*) + DNA → complex (*14*)	308	285	−100	−185	33	40
3 glucagon (*81*) → trimer (*87*)	271	364	−100	−264	47	48–72
4 melittin (*82*) → tetramer (*88*)	313	477	−150	−327	58	104
2 arc repressor (*78*) → dimer (*77*)	289	525	−50	−475	85	80–92
2 λ cro repressor (*80*) → dimer (*79*)	287	620	−50	−570	102	120

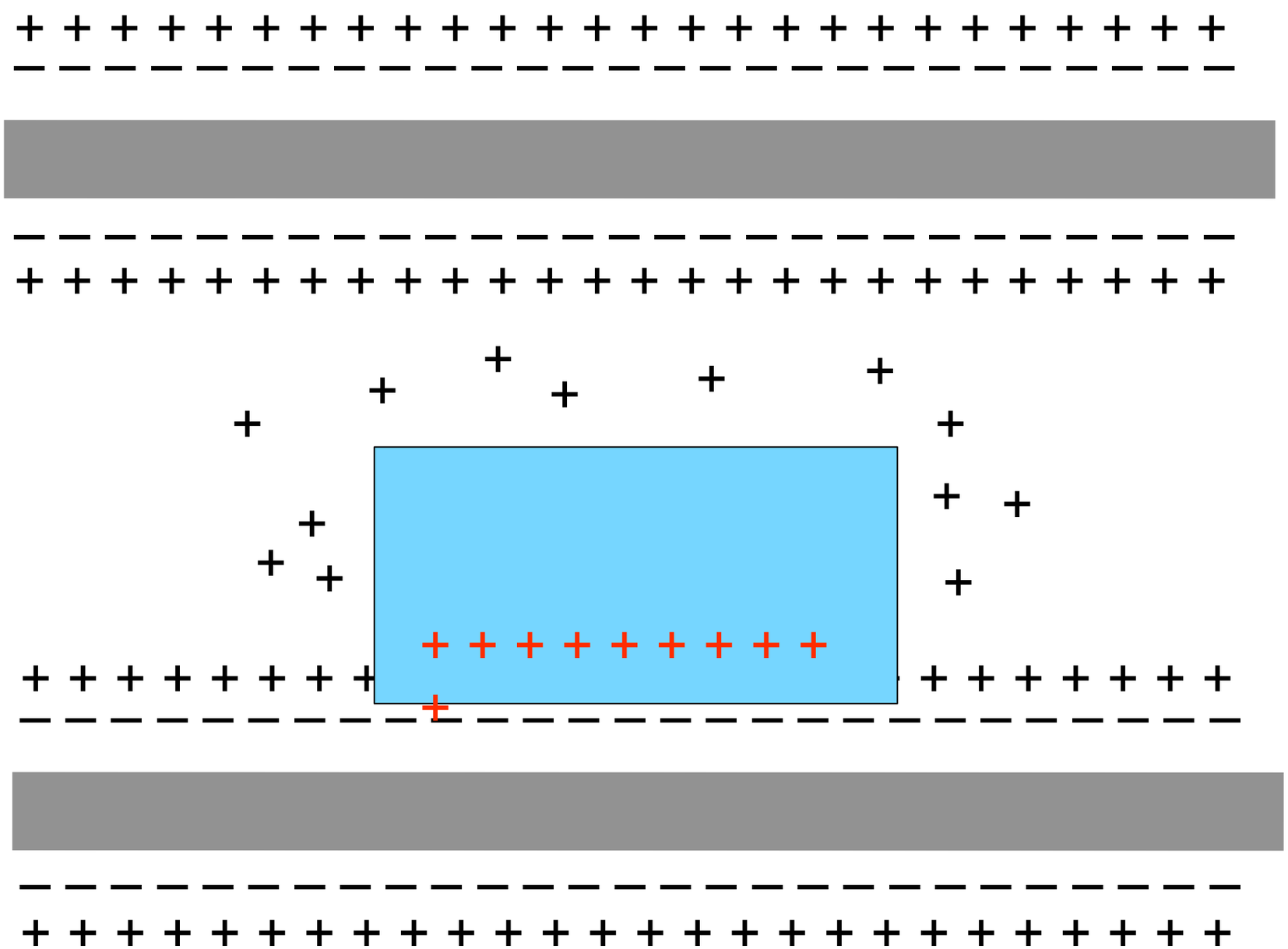
*References for data used to calculate T_S are the same as those for ΔC°_{assoc} in Table 2. †$\Delta S^\circ_{HE}(T_S)$ evaluated from Eq. 2 with values for ΔA_{np} from Table 2. ‡Table 3. §Eq. 5. ‖Eq. 6. Propagated uncertainties in $\Re^{th}$ increase from ±15 percent for λ cro repressor to ±50 percent for angiotensin II, and are typically ±25 percent. ¶$\Re^{str}$ represents the difference between the number of residues folded in the crystal structure of the complex and the number of residues observed to be folded in the free species by NMR, x-ray, or CD as referenced in column 1. #T_S estimated from values of ΔC°_{assoc} (273) and ΔS°_{assoc} (273) obtained from the temperature dependence of ΔC°_{assoc} given in (*47*), based on the assumption that S protein is completely native at 273 K. **Number of residues folded in the complex based on the NMR structure.

$$\Delta S_{bin}(T_S) = 0 = \Delta S_{HE}(T_S) + \Delta S_{rt} + \Delta S_{other}$$

Conclusion 3 from entropy analysis:

The number of residues involved in the folding transition can be calculated from the ΔS_{other} term and the value of ΔS = -5.6 cal mol^{-1} K^{-1} derived from the entropy analysis of protein folding.

ΔS_{PE}: Favorable displacement of ions from the DNA



$-T\Delta S_{PE}$ = 6 to 18 kcal/mol

Summary of protein and DNA energy contributions

- very different distributions of ΔH and ΔS for protein-DNA binding but similar ΔGs

- Unfavorable conformational entropy of folding per residue or 1.7 kcal mol^{-1}

- can be compensated by the hydrophobic effect or the “release” of water on burial of nonpolar surfaces

- for rigid body association 15 kcal mol^{-1} rotational/translation entropy loss

- additional folding of the protein can occur during binding to very different degrees

- 6-18 kcal mol^{-1} depending on interaction surface or the displacements of counter-ions upon protein binding to DNA which drives binding