Protein Complexation with DNA Phosphates as a Cause for DNA Duplex Destabilization: A Thermodynamic Model

MARCEL H. P. VAN GENDEREN* and HENK M. BUCK,
*Department of Organic Chemistry, Eindhoven University of
Technology, 5600 MB Eindhoven, The Netherlands*

Synopsis

Complexation of positively charged sites in a protein with the negative DNA phosphate groups shields the phosphate charges. This diminishes *inter*strand electrostatic repulsions, which stabilizes the duplex. When phosphate shielding is present in one DNA strand only, the conformation of this strand changes due to a decrease of *intra*strand phosphate–phosphate repulsions. This destabilizes the duplex since then the strands differ in conformation. A thermodynamic model is formulated to describe this stabilization/destabilization effect in terms of changed enthalpies and entropies of hybridization. It is found that protein complexation with one DNA strand can indeed lower the T_M value of a duplex. The model is applied to the action of helicases (replication), RNA polymerases (transcription), and restriction endonucleases. Mechanisms with unilateral charge shielding are proposed for their duplex-destabilizing properties.

INTRODUCTION

The ability of enzymes to unlock the genetic information in the interior of double-stranded DNA is one of the most important processes in biochemistry. For a good understanding of the way these enzymes work, it is necessary to realize that the exterior sugar–phosphate backbones are the primary recognition sites for a protein. One of the strongest and most obvious interactions between proteins and the DNA backbone is the electrostatic attraction between positively charged residues in the enzyme (e.g., lysine or arginine) and the negatively charged backbone phosphates.[1,2] At first glance, one would expect that this interaction, which shields the phosphate charges, always stabilizes the DNA duplex by decreasing interstrand electrostatic repulsions. Indeed, model studies with polycatonic proteins such as polylysine usually feature a markedly higher duplex stability.[3-5] However, in these cases the protein always associates with both DNA strands simultaneously.

In this paper we wish to present a way in which protein complexation can lead to destabilization of a DNA duplex, even though phosphate charges are shielded. This idea is derived from our work on hybrids of phosphate-methylated and natural DNA,[6-9] where a perfect charge shielding is present in *one* strand due to methylation of the phosphate groups. Stability studies on these neutral/natural hybrids have shown that unilateral charge shielding has two effects, viz. (a) diminished *inter*strand repulsions give duplex stabilization,

---

*To whom correspondence should be addressed.

and (b) diminished *intra*strand repulsions are present in the shielded DNA strand, which changes its conformation. Duplex destabilization now occurs due to structural differences between the shielded and natural strands. In a thermodynamic description, the first effect yields a larger enthalpy of hybridization, whereas the second effect is evident in an increased hybridization entropy. So the overall effect of unilateral charge shielding is determined by the balance of these two effects and is visualized in the melting temperature (T_M) of the duplex.

Using thermodynamic model calculations, it will be demonstrated that hybrids between phosphate-methylated and natural DNAs are invariably more stable than their natural counterparts, whereas complexation of a protein with one strand may result in duplex destabilization. The thermodynamic model will be applied to various types of enzymes that interact with duplex DNA.

METHODS

The duplex stability is evaluated as the duplex $\rightleftarrows$ coil transition temperature T_M, which can normally be expressed as the ratio of the enthalpy and entropy of hybridization: $T_M = (\Delta H_n^0)/(\Delta S_n^0)$, where n represents the number of base pairs. In this equation, we neglect effects of variations in DNA concentration and ionic strength, which are known to influence duplex stability. We assume that these factors are equal for all studied systems, and will evaluate only the consequences of unilateral charge shielding on the thermodynamic parameters. In our studies on hybrids between phosphate-methylated and natural DNA,[8] we found that an extra entropy term had to be added to the equation for T_M in order to account for the structural difference between the natural and the neutral strands: $\Delta S_k^{0'}$, where k represents the length of the natural DNA that has to adapt itself to the modified strand. One then obtains

$$T_M = \frac{\Delta H_n^0}{\Delta S_n^0 + \Delta S_k^{0'}}$$

According to the work of Breslauer et al.,[10] the enthalpy and free energy of hybridization for a natural duplex (ΔH_n^0 and ΔG_n^0) must be thought of as the algebraic sum of all possible dimer contributions (Δh and Δg) to account for hydrogen bonding and stacking interactions. For the free energy, it is necessary to include a start value (Δg_i). In our model calculations, we used a duplex consisting of a strand with A bases and a strand with T bases, both n nucleotides long. In such a duplex, $n - 1$ AA/TT dimers are present that contribute to the enthalpy and free energy. The Breslauer model now states that

$$\Delta H_n^0 = \Delta h(n - 1)$$

and

$$\Delta G_n^0 = \Delta g(n - 1) - \Delta g_i$$

For several natural duplexes, ΔH_n^0 and ΔG_n^0 have been determined, based on the melting behavior in $1M$ NaCl solutions. From these experimental data, the AA/TT dimer contributions Δh (9.1 kcal/mol) and Δg (1.9 kcal/mol), and the start value Δg_i (6.0 kcal/mol), were abstracted.[10] The standard entropy change can now be calculated from

$$\Delta S_n^0 = \frac{\Delta H_n^0 - \Delta G_n^0}{298} = \frac{\Delta h - \Delta g}{298}(n - 1) + \frac{\Delta g_i}{298}$$

The adaptation entropy $\Delta S_k^{0\prime}$ for hybridization between phosphate-methylated and natural DNAs has been determined[8] from hybridization experiments for several values of k: $\Delta S_{poly}^{0\prime} = 4.4$ cal/mol · K, $\Delta S_{30}^{0\prime} = 6.2$ cal/mol · K, and $\Delta S_{10}^{0\prime}$ and $\Delta S_5^{0\prime} > 8$ cal/mol · K. A continuous function was fitted to these experimental data:

$$\Delta S_k^{0\prime} = \frac{4.4k + 83.0}{k + 4.7}$$

The behavior of this function mirrors the observations that $\Delta S_k^{0\prime}$ approaches 4.4 cal/mol · K for polynucleotides (large k) and rises sharply for shorter lengths ($k < 30$). Since our experimental data represent an average entropy for hybridization in the interior and on the ends of a strand, we weighed $\Delta S_k^{0\prime}$ with a window function,

$$h(p, a, q) = \begin{cases} \dfrac{a - 1}{q}p + 1, & \text{if } 0 \leq p \leq q \\ a, & \text{if } q < p < l - q \\ \dfrac{1 - a}{q}(p - l) + 1, & \text{if } l - q \leq p \leq l \end{cases}$$

where l is the length of the duplex and p the number of nucleotides from the end of the duplex to the first shielded phosphate. The parameters a and q control the amplitude and width of the window. This window function (see Fig. 1) introduces the effect that conformational changes in the interior of a
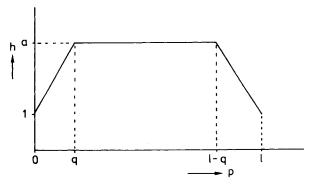

Fig. 1.   Window function $h(p, a, q)$, which is applied to the adaptation entropy $\Delta S_k^{0\prime}$.

duplex are more difficult then at the ends, so $\Delta S_k^{0\prime}$ will have a higher value in the interior. Summing all enthalpy and all entropy terms for a given duplex system and calculating the ratio of the resulting enthalpy and entropy will then give the duplex stability expressed as a T_M value.

RESULTS AND DISCUSSION

Shielded Phosphates at the End of One Strand

The situation in which a number of phosphates is shielded at the end of a duplex is depicted in Fig. 2 and gives rise to the following expression for T_M, where an asterisk indicates a base pair with one shielded and one natural phosphate:

$$T_M = \frac{\Delta H_n^0 + \Delta H_m^{0*}}{\Delta S_n^0 + \Delta S_m^{0*} + \Delta S_m^{0\prime}}$$

We now define the dimer contributions for shielded phosphates in one strand as $\Delta h^* = \beta \cdot \Delta h$, and $\Delta g^* = \delta \cdot \Delta g$. Using $l = m + n$, and counting only one duplex initiation, one can write

$$\tau = \frac{T_M}{298} = \frac{\Delta h(l - m - 1) + \Delta h^* m}{(\Delta h - \Delta g)(l - m - 1) + \Delta g_i + (\Delta h^* - \Delta g^*)m + 298\dfrac{4.4m + 83.0}{m + 4.7}}$$

Note that in this case $h(p, a, q) = 1$ always, since $p = 0$. Substituting all numerical values, it then follows that

$$\tau = \frac{9.1(l - m - 1) + 9.1\beta m}{7.2(l - m - 1) + 6 + (9.1\beta - 1.9\delta)m + \dfrac{1.31m + 24.73}{m + 4.7}}$$

We have evaluated τ as a function of m, the number of shielded phosphates, with β, δ, and l as variable parameters, using an IBM PC-AT personal computer. Note that $m \geq 2$ must hold, since dinucleotide contributions are counted. The general result of these calculations is shown in Fig. 3. First, we varied the thermodynamic parameters β and δ. Increasing the value of δ at a fixed value of β [Fig. 3(a)] and decreasing the value of β at a fixed value of δ [Fig. 3(b)] give the same result, viz., a continuously rising function $\tau(m)$. At low values of δ or high values of β, a constantly decreasing graph is found. For intermediate values, the curves show a maximum for τ. These three domains
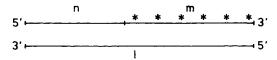

Fig. 2. Schematic representation of a DNA duplex with m shielded phosphates (*) at the end of one strand.
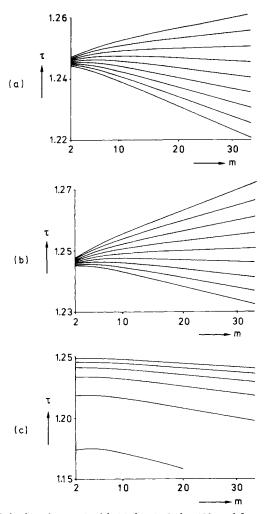
Fig. 3.   Graphs of the function $\tau(m)$ with (a) $\beta = 1.10$, $l = 100$, and δ varying from 0.80 (lower trace) to 1.20 (upper trace); (b) $\delta = 1.10$, $l = 100$, and β varying from 0.90 (upper trace) to 1.30 (lower trace); (c) $\beta = 1.10$, $\delta = 0.95$, and l varying from 20 (lower trace) to 120 (upper trace).

can be characterized roughly with the ratio δ/β. A value of $\delta/\beta \geq 1$ will always give a rising τ function, while values below ca. 0.85 give a decreasing stability for more shielded phosphates. The shape of the curves is virtually unaffected by variation of the duplex length l [Fig. 3(c)], although higher τ values are of course found for longer duplexes.

From the experimental data[6,8] on the phosphate-methylated/natural system d[Ap(OMe)Ap(OMe)A] · poly(dT), we can calculate $\beta = 1.29$ and $\delta = 1.93$. It is now evident, that since $\delta/\beta = 1.50$ duplex stabilization will always occur for this type of hybrid. In accordance with our earlier experimental work,[8] the T_M value goes up in the model when more phosphate-methylated trimers associate with poly(dT), giving a cooperative effect.

As we have remarked before, shielding of the phosphate charges can also occur in the case of protein complexation with DNA phosphates. *The present*

results indicate that protein complexation with DNA phosphates at the end of one strand can give destabilization of the duplex, as long as $\delta/\beta < 1$. This may be the mechanism that DNA unwinding enzymes (helicases) use during the replication of DNA. In this process, the DNA strands are separated, and their genetic information is copied to new strands in a structure known as a "replication fork" (Ref. 11, chap. 27). Several structural features are known for helicases that support our argument. In general, helicases cover about 8 base pairs at the end of a DNA duplex in the replication fork[12,13] and cause the duplex to dissociate into single strands. More specifically, the unwinding enzyme of phage fd (gene 5 protein) is known to bind electrostatically with 5 phosphates via lysine and arginine residues,[14,15] and has a high affinity for a single strand. In fact, in model studies it has been found that complexation of gene 5 protein with the duplex of poly(dAT) decreases the T_M value from 54 to 11.5°C.[16] Our model calculations now provide a possible mechanism for these protein-induced unwinding processes. Binding of a helicase enzyme to one of the strands in a duplex will deform this strand and can therefore destabilize the duplex when the deformation is not outweighed by diminished interstrand repulsions.

Shielded Phosphates in the Interior of One Strand

In the situation where a protein associates with the DNA phosphates in the interior of a duplex (see Fig. 4), the effect of the window function $h(p, a, q)$ becomes apparent. It now holds for the T_M value that

$$T_M = \frac{\Delta H_p^0 + \Delta H_m^{0*} + \Delta H_n^0}{\Delta S_p^0 + \Delta S_m^{0*} + \Delta S_n^0 + \Delta S_m^{0\prime} \cdot h(p, a, q)}$$

With $l = p + m + n$, this can be written as

$$\tau = \frac{T_M}{298} = \frac{\Delta h(p-1) + \Delta h^* m + \Delta h(l - m - p)}{(\Delta h - \Delta g)(p-1) + \Delta g_i + (\Delta h^* - \Delta g^*)m} \\ + (\Delta h - \Delta g)(l - m - p) + 298\frac{4.4m + 83.0}{m + 4.7}h(p, a, q)$$

Inserting numerical values then yields

$$\tau = \frac{9.1(l - m - 1) + 9.1\beta m}{7.2(l - m - 1) + 6 + (9.1\beta - 1.9\delta)m + \dfrac{1.31m + 24.73}{m + 4.7}h(p, a, q)}$$

Fig. 4. Schematic representation of a DNA duplex with m shielded phosphates (*) in the interior of one strand.
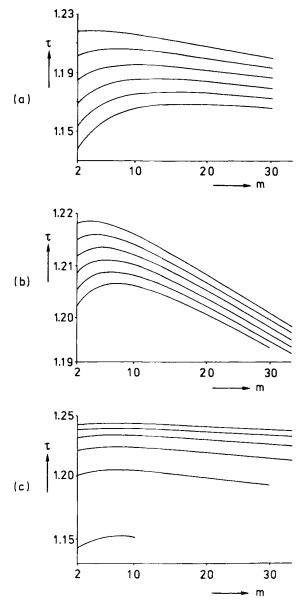
Fig. 5.   Graphs of the function $\tau(m)$ with $\beta = 1.10$, $\delta = 0.95$, and (a) $l = 40$, $p = q = 10$, and a varying from 1 (upper trace) to 6 (lower trace); (b) $l = 40$, $q = 10$, $a = 2$, and p varying from 0 (upper trace) to 10 (lower trace); (c) $p = q = 10$, $a = 2$, and l varying from 20 (lower trace) to 120 (upper trace).

Evaluation of $\tau(m)$ now depends on the parameters β, δ, and l, and the window parameters a and q. The results are shown in Figs. 5 and 6. The window parameter q defines the length of the end effect, and is only relevant in comparison with the value of p. An appropriate length should be the Debye–Hückel screening length κ^{-1}, which is the relevant distance for electrostatic interactions in solution. However, in these model calculations we neglect
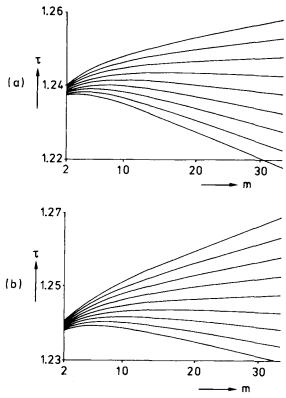
Fig. 6. Graphs of the function $\tau(m)$ with $l = 100$, $p = q = 10$, $a = 2$, and (a) $\beta = 1.10$ and δ varying from 0.80 (lower trace) to 1.20 (upper trace); (b) $\delta = 1.10$ and β varying from 0.90 (upper trace) to 1.30 (lower trace).

concentration effects, which are implicitly present in the screening length κ^{-1}. Therefore a fixed value of 10 is used throughout, which corresponds with one complete turn of the helix. Increasing the window height (a) results in a decrease for the overall τ value, and especially disfavors short stretches of shielded phosphates [Fig. 5(a)]. Normally, the value of parameter a was taken as 2. Variation of the distance from the duplex end (p) indeed gives an extra destabilization for charge shielding in the interior of the duplex [Fig. 5(b)], although no extra destabilization is present for $p > q$. Again, variation of the duplex length (l) does not alter the essential features of the curves [Fig. 5(c)], whereas the ratio δ/β determines the characteristics of the curves [Fig. 6(a, b)].

Comparing the graphs in Fig. 6(a) with those in Fig. 3(a), it is evident that for an equal value of δ/β (e.g., $\beta = 1.10$ and $\delta = 0.90$), a longer stretch of shielded phosphates [ca. 10 in Fig. 6(a) vs ca. 5 in Fig. 3(a)] is needed in the interior of the duplex to achieve the same destabilization. This corresponds nicely with the characteristics of RNA polymerases, the enzymes that locally unwind two DNA strands in the interior of a duplex to start transcription (i.e., copying of the DNA information to an mRNA molecule—see Ref. 11, chap. 29). The RNA polymerase enzyme is always affixed to one strand of the duplex at the promoter site.[17-19] Electrostatic attractions are thought to be

the main protein–DNA interactions in this case,[20] and the length of duplex DNA covered by the protein is 35–70 base pairs in the initiation phase,[21] and at least 17 base pairs in the elongation phase.[22] *The present model can adequately explain why at least twice the length of duplex must be covered for RNA polymerases in comparison with helicase enzymes (vide supra).* Recently Schinkel et al.[23] have provided experimental support for the proposed conformational changes in a DNA duplex after protein complexation in the interior and at the end (see previous section). They observed in gel retardation experiments that binding of yeast mitochondrial RNA polymerase to a DNA fragment containing the promoter either in the interior or near the end induces a bending of the helix. This bending must be due to a difference in conformation between the two DNA strands in the complex, which causes a local melting of the DNA duplex according to our model calculations. It follows therefore that protein complexation in the interior can produce the open complex, which serves as a starting point for the transcription,[21] while formation of a replication fork can be induced by complexation at the end of a duplex.

Shielded Phosphates in Both Strands Symmetrically

The situation in which phosphate charges in both strand ends are symmetrically shielded by a protein (see Fig. 7) has two modes with different mathematical expressions for T_M. For mode A, one can easily derive with $l = m + n$

$$T_{M,A} = \frac{2\,\Delta H_m^{0*} + \Delta H_{l-2m}^0}{2\,\Delta S_m^{0*} + \Delta S_{l-2m}^0 + 2\,\Delta S_m^{0\prime}}$$

[Note that $h(p, a, q) = 1$.]
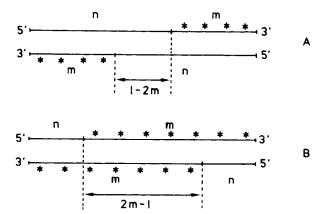

Fig. 7. Schematic representation of a DNA duplex with symmetrical regions of m shielded phosphates (*) in both strands.

This means for τ,

$$\tau_A = \frac{2\,\Delta h^* m + \Delta h(l - 2m - 1)}{2(\Delta h^* - \Delta g^*)m + (\Delta h - \Delta g)(l - 2m - 1) + \Delta g_i + 298\dfrac{8.8m + 166.0}{m + 4.7}}$$

which becomes with numerical values inserted,

$$\tau_A = \frac{9.1(l - 2m - 1) + 18.2\beta m}{7.2(l - 2m - 1) + 6 + (18.2\beta - 3.8\delta)m + \dfrac{2.26m + 49.46}{m + 4.7}}$$

This holds as long as $m \leq n$, since for longer shielded stretches an overlap will occur (mode B). Then, there arises the situation where two strands are simultaneously shielded, which is indicated with a double asterisk. The dimer contributions in this case are written as $\Delta h^{**} = \gamma \cdot \Delta h$ and $\Delta g^{**} = \epsilon \cdot \Delta g$. For mode B, one can write with $l = m + n$,

$$T_{M,B} = \frac{2\,\Delta H^0_{l-m}{}^* + \Delta H^0_{2m-l}{}^{**}}{2\,\Delta S^0_{l-m}{}^* + \Delta S^0_{2m-l}{}^{**} + 2\,\Delta S^{0\prime}_{l-m}}$$

or

$$\tau_B = \frac{2\,\Delta h^*(l - m - 1) + \Delta h^{**}(2m - l)}{\begin{aligned}&2(\Delta h^* - \Delta g^*)(l - m - 1) + \Delta g_i\\ &+ (\Delta h^{**} - \Delta g^{**})(2m - l) + 298\dfrac{8.8(l - m) + 166.0}{l - m - 4.7}\end{aligned}}$$

In the case of protein complexation, the values of γ and ϵ could be abstracted from the thermodynamic studies of Fujioka et al.[24] on DNA–polylysine systems: $\gamma = 0.83$ and $\epsilon = 1.38$. With all numerical values substituted, it follows that

$$\tau_B = \frac{18.2\beta(l - m - 1) + 7.55(2m - l)}{(18.2\beta - 3.8\delta)(l - m - 1) + 6 + 4.93(2m - l) + \dfrac{2.62(l - m) + 49.46}{l - m + 4.7}}$$

which holds for $m > n$. The transition from mode A to mode B is discontinuous in this description. A more detailed model will have to incorporate a better description for the transition area ($m \approx n$), thus yielding a smooth τ function. Inspection of $\tau(m)$ as a function of β, δ, and l shows that l only influences the halfway transition point between mode A and B, but not the shape of the curves [Fig. 8(a)]. The ratio δ/β again determines the function's behavior in mode A [Fig. 8(b, c)]. For certain values of δ/β, destabilization occurs initially, but for $m > n$ one always finds a strong stabilization. Thus, a
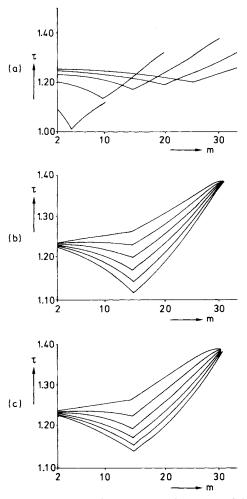
Fig. 8. Graphs of the function $\tau(m)$ with (a) $\beta = 1.10$, $\delta = 1.00$, and l varying from 10 (lower trace) to 50 (upper trace); (b) $\beta = 1.10$, $l = 30$, and δ varying from 0.80 (lower trace) to 1.30 (upper trace); (c) $\delta = 1.00$, $l = 30$, and β varying from 0.80 (upper trace) to 1.30 (lower trace).

minimum in duplex stability can be present in this symmetrical shielding situation.

This model is especially pertinent for restriction endonucleases, enzymes that can recognize and cut palindromic sequences in duplex DNA. It is known for the enzyme EcoRI that first an electrostatic complexation with DNA phosphates takes place via protonated amino acid residues and polarized α-helices.[25] This complexation locally unwinds the DNA duplex 25° by the formation of a so-called neokink, after which recognition of the base sequence occurs with a set of specific hydrogen bonds. Since EcoRI is a symmetrical dimeric protein, shielding of the phosphates will always be present on two DNA strands symmetrically.[25] *Therefore, the present model may account for the generation of neokinks in duplex DNA by endonucleases, which is an essential mechanism for the recognition process.*

CONCLUDING REMARKS

The thermodynamic model presented above is clearly able to give a mechanism for the enzymatic DNA duplex destabilization. Moreover, with relatively few parameters it can account for certain salient features of enzymes that interact with DNA. In the present version, effects of ionic strength on the duplex stability are neglected. It cannot be excluded that inter- and intrastrand electrostatic repulsions are influenced in a different way by addition of salts, so in future studies this will have to be taken into consideration. From the present model, the difference in shielded length of DNA duplex between helicases and RNA polymerases is readily explained. Also, a better understanding of the relevance of symmetrical complexation in the case of endonucleases can be obtained. We therefore feel that unilateral shielding of the phosphate charges in DNA may be an important mechanism in protein–DNA interactions. Further experimental studies will be performed to establish the ratio δ/β for protein complexation with one strand from simple model compounds and to verify the effect on duplex stability.

References

1. Saenger, W. (1984) *Principles of Nucleic Acid Structure*, Springer Verlag, New York, chap. 18.
2. Ollis, D. L. & White, S. W. (1987) *Chem. Rev.* **87**, 981–995.
3. Tsuboi, M. (1967) in *Conformation of Biopolymers*, Vol. 2, Academic Press, New York, pp. 689–702.
4. Von Hippel, P. H. & McGhee, J. D. (1972) *Ann. Rev. Biochem.* **41**, 231–300.
5. Hélène, C. & Maurizot, J.-C. (1981) *CRC Crit. Rev. Biochem.* **10**, 213–250.
6. Koole, L. H., van Genderen, M. H. P., Reiniers, R. G. & Buck, H. M. (1987) *Proc. K. Ned. Akad. Wet. Ser. B* **90**, 41–46 (communicated by H. M. Buck at the meeting of Dec. 15, 1986).
7. Van Genderen, M. H. P., Koole, L. H., Merck, K. B., Meijer, E. M., Sluyterman, L. A. AE. & Buck, H. M. (1987) *Proc. K. Ned. Akad. Wet. Ser. B* **90**, 155–159 (communicated by H. M. Buck at the meeting of Feb. 23, 1987).
8. Van Genderen, M. H. P., Koole, L. H., Moody, H. M. & Buck, H. M. (1988) *Proc. K. Ned. Akad. Wet. Ser. B* **91**, 59–62 (communicated by H. M. Buck at the meeting of Nov. 30, 1987).
9. Van Genderen, M. H. P., Koole, L. H. & Buck, H. M. (1989) *Recl. Trav. Chim. Pays-Bas* **108**, 28–35.
10. Breslauer, K. J., Frank, R., Blöcker, H. & Marky, L. A. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 3746–3750.
11. Stryer, L. (1988) *Biochemistry*, 3rd ed., Freeman and Co., New York.
12. Szekely, M. (1980) *From DNA to Protein*, McMillan Press Ltd., London, chap. 2.
13. Chase, J. W. & William, K. R. (1986) *Ann. Rev. Biochem.* **55**, 103–136.
14. Brayer, G. D. & McPherson, A. (1983) *J. Mol. Biol.* **69**, 565–596.
15. Coleman, J. E., Anderson, R. A., Ratcliffe, R. G. & Armitage, I. M. (1976) *Biochemistry* **15**, 5419–5430.
16. Alberts, B., Frey, L. & Delius, H. (1972) *J. Mol. Biol.* **68**, 139–152.
17. Hinkle, D. C. & Chamberlin, M. (1972) *J. Mol. Biol.* **70**, 157–185.
18. Mangel, W. F. & Chamberlin, M. (1974) *J. Biol. Chem.* **249**, 2995–3001.
19. Chamberlin, M. (1976) in *RNA Polymerase*, Losick, R. & Chamberlin, M., Eds., Cold Spring Harbor Press, New York, pp. 159–191.
20. Shaner, S. L., Melançon, P., Lee, K. S., Burgess, R. R. & Record, M. T., Jr. (1983) *Cold Spring Harbor Symp. Quant. Biol.* **47**, 463–472.

21. Von Hippel, P. H., Bear, D. G., Morgan, W. D. & McSwiggen, J. A. (1984) *Ann. Rev. Biochem.* **53**, 389–446.

22. Schmitz, A. & Galas, D. J. (1979) *Nucleic Acids Res.* **6**, 111–137.

23. Schinkel, A. H., Groot-Koerkamp, M. J. A., Teunissen, A. W. R. H. & Tabak, H. F. (1988) *Nucleic Acids Res.* **16**, 9147–9163.

24. Fujioka, K., Baba, Y. & Kagemoto, A. (1979) *Polymer J.* **11**, 509–513.

25. McClarin, J. A., Frederick, C. A., Wang, B.-C., Greene, P., Boyer, H. W., Grable, J. & Rosenberg, J. M. (1986) *Science* **234**, 1526–1541.