
Thermal denaturation of double-stranded nucleic acids: prediction of temperatures critical for gradient gel electrophoresis and polymerase chain reaction

Gerhard Steger

Institut für Physikalische Biologie, Geb. 26.12.U1, Heinrich-Heine-Universität, Universitätsstraße 1, 40225 Düsseldorf, Germany

Received April 23, 1994; Revised and Accepted June 17, 1994

ABSTRACT

A program is described which calculates the thermal stability and the denaturation behaviour of double-stranded DNAs and RNAs up to a length of 1000 base pairs. The algorithm is based on recursive generation of conditional and *a priori* probabilities for base stacking. Output of the program may be compared directly to experimental results; thus the program may be used to optimize the nucleic acid fragments, the primers and the experimental conditions prior to experiments like polymerase chain reactions, temperature-gradient gel electrophoresis, denaturing-gradient gel electrophoresis and hybridizations. The program is available in three versions; the first version runs interactively on VAXstations producing graphics output directly, the second is implemented as part of the HUSAR package at GENIUSnet, the third runs on any computer producing text output which serves as input to available graphics programs.

INTRODUCTION

Several of the modern experimental methods in molecular biology need as a prerequisite knowledge about the thermal stability of double-stranded nucleic acid (dsNA): (i) for amplification of nucleic acid fragments by the polymerase chain reaction (PCR) two primers of similar stability as well as three temperatures suitable for renaturation of primer/ssDNA complexes, amplification, and complete denaturation of the synthesized dsDNA have to be chosen (1); (ii) for hybridization a temperature has to be chosen which allows formation of the wanted dsNA; and, most critically, (iii) for detection of mutations or polymorphism by temperature- or denaturing-gradient gel electrophoresis (TGGE or DGGE; 2, 3), optimal nucleic acid fragments and an optimal gradient have to be chosen. With PCR or hybridization wrong temperatures lead to amplification or detection, respectively, of wrong and unspecific nucleic acids. With TGGE or DGGE a badly chosen fragment gives only a single denaturation transition and allows only a very limited detection of mutations.

The thermal stability of double-stranded nucleic acids reflects the local base composition of the nucleic acid and is influenced

by several other parameters such as solvent conditions, especially ionic strength and concentration of denaturants, and concentration of the nucleic acid (for review of denaturation behaviour of nucleic acids and critical discussion of parameter values see 4–7). The best method available for calculation of the thermal stability of a dsNA is based on an algorithm described by Poland (8) and on an accelerated version of that algorithm described by Fixman and Freire (9). In this paper a program, called POLAND, is described which simulates transition curves of dsNA (dsDNA as well as dsRNA) based on these algorithms using thermodynamic parameters available from the recent literature. The output produced by the program is mainly graphics which visualize the thermal stability of dsNA and may help for example to design correct primers, to select optimal fragments, or to choose optimal temperatures in experiments. Two examples for use of the program are shown and discussed in comparison to experimental results, i.e. denaturation curves of dsRNA as well as of dsDNA are determined experimentally by optical detection or by TGGE and are compared to the appropriate calculated curves. Furthermore experimentally determined optimal annealing temperatures (1) of PCR are compared to calculated denaturation temperatures of primers.

SYSTEM AND METHODS

The program described in this paper is available in three different versions: the first runs interactively on VAXstations (VAXstation II and Alpha AXP 3000/800, respectively, with operating system MicroVMS V4.7 and OpenVMS AXP V1.5 (DEC Digital Equipment Corporation, Maynard, MA); it is written using VAX FORTRAN V4.0 and V6.1, respectively; graphics are implemented via VAX GKS V3.0 and V5.2, respectively. The second version is implemented as part of the HUSAR package (GENIUSnet, DKFZ, Heidelberg, Germany) using the GCG interface (10). The third version runs only non-interactively in batch mode; it is written in a slightly extended dialect of FORTRAN 77. Thus no changes of the source code were necessary for compilation with VAX FORTRAN, Microsoft FORTRAN V5.1, CONVEX fc, or NeXT FORTRAN. Source code is available on request from the author (Steger@PSTV.Biophys.Uni-Duesseldorf.DE); an executable for IBM

PC's or compatibles (with or without math coprocessor), including a short user guide, is available from Qiagen GmbH (Max-Vollmer-Straße 4, 40724 Hilden, Germany). A graphical user interface for the PC program is available on request from Dr. E. Molitor (Institut für Medizinische Mikrobiologie und Immunologie, Sigmund-Freud-Straße 25, 53127 Bonn, Germany).

ALGORITHM

The original algorithm (8) requires computing time proportional to the square of the sequence length N , i.e. is of type $O(N^2)$, but allows any type of loop entropy factor (see below, Loops). This algorithm was modified (9) by decomposition of the original loop entropy factor (8) into a sum of I exponential functions; this approximation makes computing time proportional to $O(I*N)$. With $I = 10$ coefficients, as implemented in the described program, computing time is drastically reduced; the accuracy of the results, however, is not influenced. The elementary step in both algorithms mentioned is based on the opening of a base pair; the algorithms implemented in the described program are based on base stacking (11).

INPUT

The VAX version of the program prompts the user for input alternatives with menus (GKS choice) and/or accepts character strings (GKS string). The HUSAR version uses the standard GCG interface. The input to the non-interactive program version is given via a command file, POLAND.CMD (see, for example, Fig. 2A), which may be modified by any text editor. Lines of the command file starting with an asterisk contain comments and may be removed. Lines containing tokens may be in any order; the tokens may be in upper or lower case. The command file has to be located in the same directory as the program. With all program versions the parameters described below are available or may be changed in between certain ranges.

Sequence and mismatches

Sequence. Allowed nucleotide characters are A, G, C, U, and T. For calculation of subsequences the positions of 5' and 3' end nucleotides may be given. The sequence used for calculation has to be shorter than 1001 nucleotides but longer than 5 nucleotides; the sequence file may contain a sequence of any length.

Mismatch. Mismatches are introduced into the algorithm as internal loops of size p . The position of the mismatch(es) is not

Table 1. Thermodynamic parameters for helix growth

A		Freier <i>et al.</i> (12)				Gotoh (5)			
		3' A:U	3' U:A	3' G:C	3' C:G	3' A:T	3' T:A	3' G:C	3' C:G
- ΔH^0 / kJ/mol	5' A:U/T	27.6	23.9	31.8	42.7	33.7	33.9	34.1	38.1
	5' U/T:A	33.9	27.6	44.0	55.7	31.8	33.7	33.7	37.0
	5' G:C	55.7	42.7	51.1	59.5	37.0	38.1	36.9	42.1
	5' C:G	44.0	31.8	33.5	51.1	33.7	34.1	35.5	36.9
- ΔS^0 / J/molK	5' A:U/T	77.0	64.9	80.4	109.7	102.7	102.7	102.7	102.7
	5' U/T:A	94.6	77.0	116.4	148.6	102.7	102.7	102.7	102.7
	5' G:C	148.6	109.7	124.3	146.1	102.7	102.7	102.7	102.7
	5' C:G	116.4	80.4	81.2	124.3	102.7	102.7	102.7	102.7
B		Klump (14)				Breslauer <i>et al.</i> (15)			
		3' A:T	3' T:A	3' G:C	3' C:G	3' A:T	3' T:A	3' G:C	3' C:G
- ΔH^0 / kJ/mol	5' A:T	32.3	31.2	36.0	35.8	38.1	36.0	32.7	27.2
	5' T:A	31.2	32.3	35.8	36.0	25.1	38.1	24.3	23.4
	5' G:C	36.0	35.8	35.1	39.8	23.4	27.2	46.1	46.5
	5' C:G	35.8	36.0	39.8	35.1	24.3	32.7	49.8	46.1
- ΔS^0 / J/molK	5' A:T	95.6	95.3	102.4	100.4	100.5	100.1	87.1	72.4
	5' T:A	95.3	95.6	100.4	102.4	70.8	100.5	54.0	56.5
	5' G:C	102.4	100.4	93.9	102.4	56.5	72.4	111.4	111.8
	5' C:G	100.4	102.4	102.4	93.9	54.0	87.1	116.4	111.4
		Pörschke <i>et al.</i> (13)							
		A:U	G:C	ss [#]					
- ΔH^0 / kJ/mol		45.6	62.3	33.1					
ΔS^0 / J/molK		130.1	156.4	108.8					

Free energy is calculated by: $\Delta G^0 = \Delta G^0(X:Y) - f*\Delta G^0(ss)$

In case of A-U neighbors $f = 2$ else $f = 1$.

[#]For correction of stacking in single strands after dissociation of base pairs:

$\Delta G^0(ss) = R*T*\ln(1. + \exp((-\Delta H^0(ss) + T*\Delta S^0(ss))/(R*T)))$

$R = 8.3143$ J/molK

given as odd nucleotides in the sequence file but has to be given via a menu option or on a separate line in the command file. If the mismatch is longer than a 'base pair', the position of each base has to be given separately. Calculation of asymmetric or bulge loops is not possible.

Thermodynamic parameters

Base stacking (helix growth). The program allows to choose from five different parameter sets of base stacking. These are for RNA in 1 M NaCl according to Freier *et al.* (12) (see Table 1A) or Pörschke *et al.* (13) (see Table 1B), respectively, and for DNA (see Table 1A) in 0.019 M NaCl according to Gotoh (5), in 0.1 M NaCl according to Klump (14) and in 1 M NaCl according to Breslauer *et al.* (15), respectively. In contrast to all other parameter sets the values of Pörschke *et al.* (13) differentiate only between G:C, A:U, and A:U/A:U stacks but include single-strand stacking. In all cases the program allows correction of the entropy values of base stacking by factors (see Table 4 and Discussion).

Loops (helix initiation). Loops of size p appear during denaturation of double-stranded nucleic acids by internal opening of base stacks; they may be calculated by two different methods.

(1) The original (8) loop entropy factor, $\delta(p) = \sigma * (2*p + 1)^{-1.75}$, allows the accelerated algorithm by decomposition into a sum of exponential factors (9) (see Table 2B). The loop coefficient, σ , influences the cooperativity and the half width of each transition. σ is dependent on the ionic strength and may be varied between 1 and $1.E - 7$ (16-19).

(2) Experimental values for short internal loops (see Table 2A), which are available from Gralla and Crothers (20) or from Freier *et al.* (12), may be used in combination with an extrapolation for longer loops (21, 22). Because a decomposition of the experimental values was not possible with a satisfying degree

of accuracy, these values may be used only with the original non-accelerated algorithm (8).

Strand dissociation. The product of the dissociation constant, β , and of the concentration of single strands, c_0 , influences the denaturation temperature, T_m , and the half width of the second order transition, i.e. with decreasing $\beta * c_0$ the strand separation reaction is shifted to lower temperatures. The parameter β depends on the G:C content of the base pairs involved in the dissociation transition (13) but is not known with accuracy. As input to the program, β may be given as a fixed value in the range 1 to $1.E - 5$, or calculated internally as a function (17, 23) of the number of base stacks, N , and the internal degree of denaturation, Θ_{int} :

$$\beta = K_D * N^\alpha / \Theta_{int}$$

with K_D including all factors independent of N , and α the dependence on translational and rotational degrees of freedom, $\alpha = a + b(1 - \Theta_{int})$. The values $K_D = 5E3$, $a = -2.8$ and $b = -3.2$ are used (23).

OUTPUT

The interactive program versions write on demand all input values as well as all calculated curves into output files (using, for example, Postscript, HPGL or TEK4014 syntax) which may be stored for further processing or sent to appropriate output queues or devices. The non-interactive program version writes all input values and calculated output into a text file which may be read for further processing by several different graphics programs (GLE, GnuPlot, Excel, TechPlot, etc.). In the following the different plotting options are described.

Table 2. Thermodynamic parameters for loop formation

p	- ΔS / J/molK	
	Gralla and Crothers (20)	Freier <i>et al.</i> (12)
2	1.4	10.8
3	22.5	22.9
4	35.1	33.7
5	42.1	37.8

The loop entropy function is calculated from the ΔS values assuming two A:U base pairs closing the loop.:

$$\delta(1) = 1.$$

$$\delta(p) = \sigma * \exp((\Delta S(p) - 47.8)/R).$$

For larger loops following formula is used (21, 22):

$$\delta(i) = \sigma * 0.000957 * (2*i + 1)^{-1.5} * \exp(-1.086/(2*i + 1) - (47.8 - 40.9)/R)$$

$$R = 8.3143 \text{ J/molK}$$

B

j	Fixman and Freire (9)				
	1	2	3	4	5
a(j)	0.38736802	0.16184847	0.56093484E-01	0.17068950E-01	0.47809463E-02
b(j)	0.94440088E-01	0.34619090	0.60903770	0.79331267	0.89867014
j	6	7	8	9	10
a(j)	0.12661662E-02	0.32142055E-03	0.77437566E-04	0.16545888E-04	0.22904876E-05
b(j)	0.95266253	0.97877645	0.99104792	0.99671608	0.99921376

The loop entropy function is calculated from the coefficients by: $\delta(p) = \sum_{j=1}^{10} \sigma * a(j) * \exp(-b(j) * p)$
Coefficients are determined as described by Fixman & Freire (9).

Three-dimensional probability map

The main result of the algorithm is the probability, $p_i(T)$, of each base stack, i , to be in the open state as a function of temperature, T . This may be represented in a three-dimensional plot (ordinate p vs. abscissa i in 5'-3' direction vs. T ; cf. Figs 1E and 2E) which clearly shows the cooperative units of the denaturing nucleic acid. These plots do not include strand separation, thus the opening of the segment denaturing last is shown at a temperature too high.

Two-dimensional stability map

A three-dimensional probability map may be reduced to a two-dimensional stability map by showing only the temperature, $T_{pi} = 0.5(i)$, at which base stack i has a probability of 50% to be in the open state, versus i in 5'-3' direction (Fig. 2D). These plots include the influence of the second order reaction: with decreasing concentration of the nucleic acid the dissociation temperature decreases. The influence of mismatches on the stability of the nucleic acid and on the size of the cooperative units is seen easily by overlaying maps of the wild-type molecule and of molecules containing mismatches.

Optical denaturation curves

From the probability of the base stacks, after correction for strand dissociation and for hypochromicity (see Table 3), denaturation profiles are obtained. Each denaturation profile is given at two wavelengths: in case of DNA at 260 and 282 nm (24) or in case of RNA (Fig. 1C and D) at 260 and 280 nm (25), respectively. These denaturation profiles are directly comparable with experimental optical melting curves (11).

Table 3. Relative values of hypochromicity

	dsRNA (25)		dsDNA (24)	
λ	260 nm	280 nm	260 nm	282 nm
A:U/T	1.0000	0.0370	1.0000	0.0000
G:C	0.3960	0.6830	0.3664	0.7266

Gel-electrophoretic mobility

The gel-electrophoretic mobility depends on the friction of a nucleic acid molecule in the gel matrix and the solution, but is dominated by the conformation of the nucleic acid. This conformation dependence is the basis of TGGE as well as of DGGE. For calculation of the relative mobility μ/μ_0 of a completely or partially double-stranded molecule, a semi-empirical formula was introduced (3, 26); this formula is used in a modified form (2):

$$\begin{aligned} \mu(T)/\mu_0 &= \exp(-\bar{p}(T)/L_r) \text{ if } \zeta \leq 0.5 \\ SY &= \exp(-\bar{p}(T_0)/L_r) \text{ if } \zeta > 0.5 \\ \bar{p} &= \sum_{i=1}^N p_i \end{aligned}$$

μ_0 is the mobility of a completely double-stranded molecule of N base stacks; the retardation length L_r is a normalizing factor, which may be interpreted as proportional either to the length of a flexible unit of the nucleic acid or to the permeability of the gel. As long as the molecule is at least partially double-stranded (the fraction of dissociated single strands ζ is below 0.5), p is the sum of the probabilities of all base stacks to be in the open state. After strand separation ($\zeta > 0.5$) the mobility increases discontinuously. In the formula it is assumed that a change in mobility is independent of the position of the base stack; i.e. both dangling ends and internal loops have the same effect on the mobility. Examples of calculated gel-electrophoretic mobility curves are given in Fig. 1A and 2B.

DISCUSSION

Examples for the use of the program will be discussed below. Results on polymers are shown in Fig. 1 for dsRNA and in Fig. 2 for dsDNA. Table 4 summarizes combinations of values for helix growth and for loop formation used in the examples. In Table 5 experimentally determined, optimal annealing temperatures (1) of PCR are compared to calculated denaturation temperatures of primers.

Choice of parameters for prediction of nucleic acid unfolding

The goal of the described programs is the prediction of denaturation behaviour of dsNA. In order to simplify to the user

Table 4. Preferable combinations of parameters

Parameter	Double-stranded nucleic acid					
	RNA		DNA			
Ionic strength in M NaCl	1.000	1.000	0.019	0.100	1.000	
Helix growth parameters according to	Freier <i>et al.</i> (12)	Pörschke <i>et al.</i> (13)	Gotoh (5)	Klump (14)	Breslauer <i>et al.</i> (15)	
ΔS -correction*	1.021/1.000/0.961	1.000/1.040/0.970	not necessary	not necessary	*	
Loop-parameters	according to	Gralla & Crothers (20) or Freier <i>et al.</i> (12)	Fixman & Freire (9) or Poland (8)	Fixman & Freire (9) or Poland (8)	Fixman & Freire (9) or Poland (8)	*
	σ	1	1.E-6	1.E-3	1.E-3	*
Dissociation constant β	1.E-3	1.E-3	1.E-3	1.E-3	*	
Retardation Length L_r	N/2	N/2	N/2	N/2	*	

*:The ΔS -values for helix growth are corrected by the given factors; all ΔS -values are multiplied with the first number, all ΔS -values of A:U/A:U-stacks with the second and all ΔS -values of G:C/G:C-stacks with the third.
#:No satisfying fit of experimental data obtained with longer DNAs was achieved.

the choice of parameter input, optimal parameter combinations are summarized in Table 4. These combinations are used in the examples below but may be used also without alterations for other dsRNAs.

With dsRNA, helix growth parameters according to Freier *et al.* (12) are optimal after correction of all ΔS values by a factor of 1.021 and of all $\Delta S(\text{G:C})$ values by 0.961. These corrections lead to an accurate prediction of absolute T_m values as well as of T_m differences of multiple transitions in a single dsRNA. The corrections are equivalent to a shift in T_m values of A:U stacks by -7°C or -2% , respectively, and of G:C/G:C stacks by $+7^\circ\text{C}$ or $+2\%$, respectively. Thus, the corrections are below the estimated error limit (12) of 5% for the measured entropy values. For loops both experimental parameter sets (12, 20) give very similar results; both sets should not be corrected by a σ

factor. For an example using these values see Fig. 1. In case of dsRNAs containing mismatches the original loop entropy factor (8, 9) is less suited than the experimental parameter sets (12, 20) (results not shown).

Using thermodynamic parameters for helix growth of Pörschke *et al.* (13) melting curves are obtained which are similar but less accurate compared to those obtained with parameters of Freier *et al.* (12). $\Delta S(\text{A:U})$ values might be corrected by a factor of 1.040 and $\Delta S(\text{G:C})$ values by 0.970. The Fixman–Freire algorithm with a loop coefficient of $\sigma = 1.E-6$ might be used (results not shown; for an example see 27).

With dsDNA, helix growth parameters of Gotoh (5) and of Klump (14) are equally well suited for predictions. Because parameters are valid for different ionic strength conditions, there is only a shift in the temperature scale. In neither case a ΔS

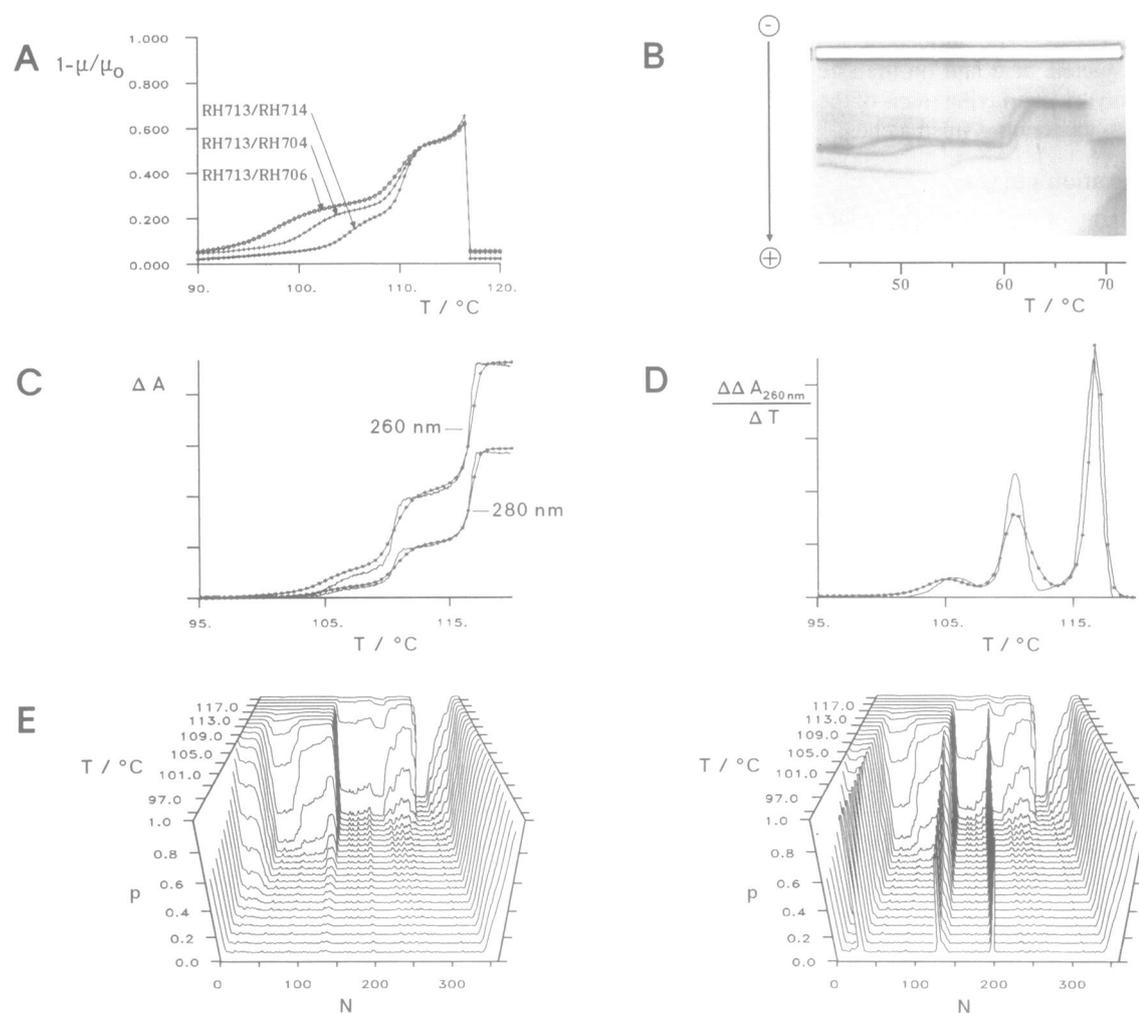


Figure 1. Analysis of dsRNA synthesized from T7-transcription plasmids pRH713 and pRH714 (homologous dsRNA), pRH713 and pRH704 (heterologous dsRNA with mismatches at nucleotide positions 29, 30, 130, and 199), pRH713 and pRH706 (heterologous dsRNA with mismatches at nucleotide positions 28, 34, 124, and 125). For experimental details see Zimmat *et al.* (28). (A) Calculated gel mobility curve of homologous dsRNA RH713/RH714 (*) and of heterologous dsRNAs RH713/RH704 (+) and RH713/RH706 (o) using a coefficient $L_r = 200$. (B) TGGE of homologous dsRNA RH713/RH714 and of heterologous dsRNAs RH713/RH704 and RH713/RH706. The gel contained 5% acrylamide, 0.12% bisacrylamide, 0.08% TEMED, 89 mM Tris, 89 mM boric acid, 0.24 mM EDTA, 9 M urea, and 0.06% ammonium peroxydisulphate; electrophoresis in the presence of the temperature gradient was carried out at 300 V for 90 min; the gel was stained with silver. (C, D) Experimental (—) and calculated (*) thermal transition curves of homologous dsRNA RH713/RH714 in integrated (C) and differentiated (D) form. The measurement was carried out with 0.25 A260 dsRNA/ml in 10 mM NaCl, 0.1 mM EDTA, 1 mM Na cacodylate, pH 6.8; the resulting transition curve was shifted by 28.8°C in order to correspond to the ionic-strength condition of the thermodynamic parameters. (E) $p(N,T)$ plot (probability of each stack to be in the denatured state vs. nucleotide position in 5'–3' direction) of homologous dsRNA RH713/RH714 (left) and of heterologous dsRNA RH713/RH704 (right). This type of plot does not include the second order reaction of strand separation.

A *** POLAND.CMD *****
 *** Full filename of sequence (including directory) ***
 Sequence = \$DiskV:[STEGER.SEQ]HumHbb_18GC.SEQ
 Begin = 1
 End = 125
 **** Mismatch positions (none = -1) ****
 Mismatch = 89
 **** Full filename of output file ***
 Output = Hybrid.TXT
 **** Plots ****
 3DPlot = N
 GelPlot = n
 MeltPlot = no
 TempPlot = yes
 **** Temperatures (Tmin, Tmax, TStep) ****
 Temperature = 35., 85., 0.5
 **** Thermodynamic parameters (Turner, Pörschke, Gotoh, Klump, or Breslauer) ****
 Thermo = G
 **** DeltaS - Factors (all, A:U/T-, G:C-stacks) ****
 DScorr = 1.0, 1.0, 1.0
 **** Concentration of single strands ****
 C0 = 1.e-8
 **** Dissociation constant, β (Benight & Wartell = -1) ****
 Beta = -1.
 **** Loop parameters (Gralla & Crothers, Turner, Poland) ****
 Loop = p
 Sigma = 1.e-3
 **** Algorithm (Poland or Fixman) ****
 Algorithm = fixman
 **** Stiffness of nucleic acid ****
 Lr = 40, 90, 200

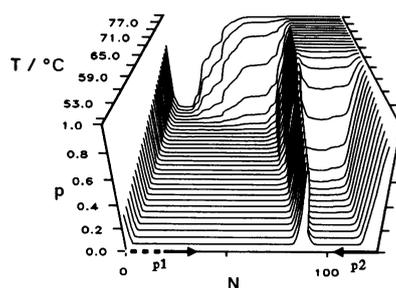
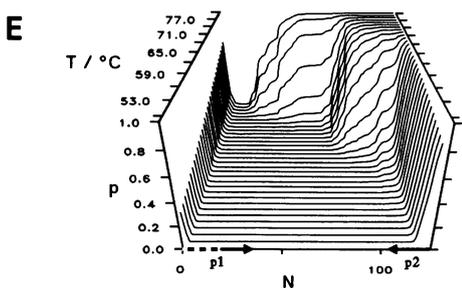
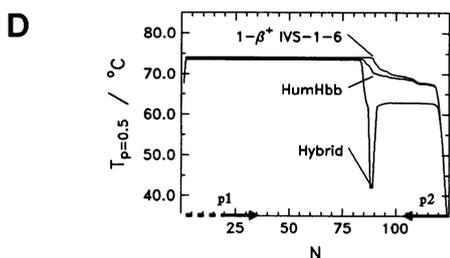
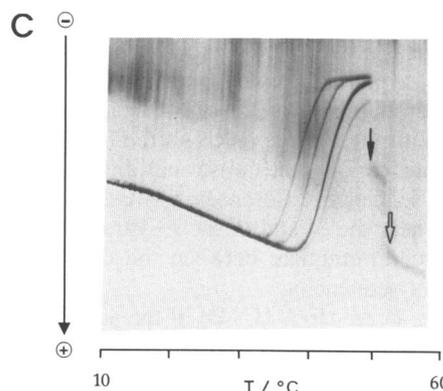
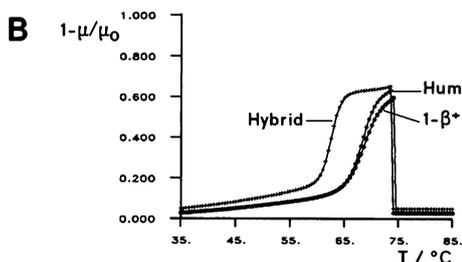


Figure 2. Analysis of dsDNA synthesized by PCR-amplification of the human β -globin locus (30). In (D) and (E) the 5' (p1) and 3' (p2) primers used for amplification are depicted as arrows; the 5' primer contains a 5' GC clamp shown as a dotted line. For experimental details see text and (2). (A) Command file used for calculation. (B) Calculated gel mobility curves of dsDNA with wild type sequence (HumHbb), mutated dsDNA ($1-\beta^+$ IVS-1-6), and hybrid with a single mismatch at position 89 of the wild-type sequence. A normalizing factor $L_r = 40$ was used. (C) TGGE-analysis of a mixture of wild type segment and IVS-1-6 mutant (in slight excess) after a denaturation/renaturation cycle. The gel contained 8 % polyacrylamide, 0.13 % bisacrylamide, 8.9 mM Tris, 8.9 mM boric acid, 0.2 mM EDTA, 4 M urea; TGGE was carried out for 1.5 h at 500 V. The retardation transitions belong (from left to right) to the hybrids with C·A and T·G mismatch, respectively, and to the dsDNAs with T:A and C:G base pair, respectively. For arrows see text. (D) $T_{p=0.5}(N)$ plot (dissociation temperature of each stack vs. nucleotide position in 5'–3' direction) of dsDNAs with wild type sequence (HumHbb), with mutated sequence ($1-\beta^+$ IVS-1-6), and of hybrid with a single mismatch at position 89 of the wild type sequence. (E) $p(N,T)$ plot (probability of each stack to be in the denatured state vs. nucleotide position in 5'–3' direction vs. temperature) for dsDNA with wild type sequence (left) and for the hybrid with a single mismatch at position 89 of the wild type sequence (right).

correction is necessary for reproduction of experimental results. Use of the Fixman–Freire algorithm with a loop coefficient of $\sigma = 1.E-3$ is recommended. For an example using Gotoh parameters see Fig. 2.

With values of Breslauer *et al.* (15) neither the example described below nor several others were simulated with the necessary degree of accuracy. Predicted transition temperatures are extremely high resulting in dependencies above 30°C per factor of 10 in ionic strength. With some dsDNAs additional transitions are predicted; this may be due to the extraordinary stability of C:G/A:T stacks whereas G:C/C:G is the most stable stack with parameters of Gotoh (5) and Klump (14).

For both dsDNA and dsRNA, a dissociation constant β of $1.E-3$ is adequate. Calculation of β as a function (17, 23) results in a sharpening of the second order transition which might be desirable for prediction of mobility curves. The retardation length L_r is depended on polymer and crosslinker concentrations; appropriate values are equal to about half of the length of the nucleic acid ($L_r \approx N/2$).

Examples

In Fig. 1 the theoretical as well as the experimental analysis of a dsRNA with 359 base pairs is shown together with the analysis of two different heteroduplexes of this RNA which have four mismatches each. With these three dsRNAs the (+)strand has always the same sequence, only the (–)strands differ from each other. Most of the sequence differences between these three mutants are located near the 3' end of the (–)strand; thus, the analysis was used to differentiate between the (–)strands in TGGE easily without sequencing.

The optical melting curve (Fig. 1C, D) of the homoduplex is determined experimentally in 10 mM NaCl and shifted by 28.8°C in order to meet the 1 M NaCl conditions of the thermodynamic values (see Ionic Strength Dependence). The theoretical curves (Fig. 1A, C–E) are calculated using the thermodynamic parameters for helix growth of Freier *et al.* (12); all ΔS values are corrected by a factor of 1.021 and all $\Delta S(G:C)$ values by 0.961. The loop parameters of Gralla and Crothers (20) are used and $\beta * c_0$ is $1.E-10$. The theoretical calculation gives a fit to the experimental optical denaturation curve (Fig. 1C, D) which is nearly identical or at least optimal concerning T_m values, half widths and relative hypochromicities of the three transitions. Thus

Table 5. Comparison of optimal (1) and calculated annealing temperatures of PCR

T_a OPT	$T_m, p1$	$T_m, p2$	p1	p2
50.5 ± 1.3	51.9	56.2	e440	e323
54.5 ± 0.7	55.8	60.3	e218	e10
44.0 ± 1.5	43.0	50.0	e1287	e1520
55.3 ± 0.8	55.1	56.2	e695	e323
50.9 ± 0.7	51.9	60.3	e440	e10
60.0 ± 0.7	60.0	62.4	λ7131	λ7606
48.4 ± 0.6	50.0	52.0	e1520	e885
56.0 ± 0.4	55.1	60.3	e695	e10
59.8 ± 1.2	57.9	61.0	λ30042	λ31017
59.8 ± 1.7	68.7	69.6	λ24182	λ20161
56.6 ± 0.5	61.0	69.6	λ31017	λ20161

T_a OPT: Experimentally determined, optimal annealing temperature for primers p1 and p2 (reproduced from Table II of (1)).

$T_m, p1, T_m, p2$: Calculated denaturation temperatures of primers p1 and p2 using thermodynamic parameters (14) and loop parameters (9) with $\sigma = 1E-3$ and $\beta * c_0 = 0.25E-9$.

P_1, P_2 : For sequence of primers and amplified fragments see (1).

the theory is able to predict accurately the denaturation behaviour of a dsRNA.

The 5' terminal base pairs (positions 1–~40) of the homoduplex RH713/RH714 (see $p(N,T)$ plot in Fig. 1E left) form the most unstable cooperative unit which gives rise to the first transition at 105.6°C in the calculated mobility curve (Fig. 1A) and in the optical transition curve (Fig. 1C, D). The second transition at 110.5°C is due to the cooperative denaturation of base pairs at positions ~40–~140. The final transition at 116.7°C is due to the cooperative denaturation of the 3' terminal base pairs (positions ~140–359) including a G:C-rich region (positions 288–315 with 24 G:C base pairs). Because this transition includes the strand separation it is irreversible in the low ionic strength conditions of the TGGE and may be seen as a discontinuous acceleration of mobility in Fig. 1A and B.

The heteroduplex RH714/RH704 contains four mismatches (see $p(N,T)$ plot in Fig. 1E right). Two of them are located near the 5' terminus (positions 29 and 30) and destabilize the first transition, i.e. they shift the denaturation temperature from 105.6°C to 102°C. The third mismatch at position 130 destabilizes only slightly the second transition (see Fig. 1A). The fourth mismatch at position 199 is located in the most stable region but has no visible effect on the strand separation in TGGE.

The heteroduplex RH714/RH706 also contains four mismatches. Two of them are at positions 28 and 34. Even at low temperature the base pairs in between these positions are denatured, thus giving rise to a first transition at the lowest temperature of the examples shown (Fig. 1A and B). The two remaining mismatches (positions 124 and 125) destabilize only slightly the second transition.

The theoretical results for the three mutants shown in the mobility plot of Fig. 1A may be compared to the experimental temperature-gradient gel in Fig. 1B. At the lowest temperature shown the distance of migration is highest for the homoduplex in the theoretical plot as well as in the gel whereas the molecules containing mismatches are significantly retarded. This effect is experimentally reproducible and may be used to separate mutant sequences even in non-denaturing gel electrophoresis (ndPAGE; 28, 29). In the temperature-gradient gel the migration velocity of all molecules outside of the transitions is increased with increasing temperature; this is even more pronounced in the gel shown in Fig. 2B. This effect may be due to decreasing viscosity of the gel solution but is not corrected for in the mobility plot.

In Fig. 2 the analysis of a dsDNA is shown which was synthesized by PCR-amplification of the human β -globin locus (30) from nucleotide 62,251 to 62,315 using two primers complementary to nucleotides 62,233–62,250 (with an additional 5' GC clamp of 19 nucleotides) and 62,316–62,338. A T → C mutation at nucleotide 62,302 inhibits the correct splicing between exon 1 and 2; this IVS-1-6 mutation is the basis for one type of β -globin thalassemia.

In Fig. 2A the command file of the PC version of the program is shown which was used to produce the data for one of the curves of the $T_p = 0.5(N)$ plot in Fig. 2D. The 3' terminal base pairs (positions ~90–125, see $p(N,T)$ plot in Fig. 2E left) of the dsDNA with wild type sequence denature first, the 5' base pairs (positions 1–~90) are stabilized by the 5' terminal GC-clamp and denature in the second order transition. The length of the GC-clamp was optimized (2) in order to get a large difference between the T_m values of the 3' and the 5' segment. The mismatch at position 89 is at the border between the first and

the second cooperative unit (see $p(N,T)$ plot in Fig. 2E right). Therefore the first transition is shifted by about 6°C to lower temperatures whereas the second transition is not influenced visibly.

In Fig. 2B ($T_p = 0.5(N)$ plot) and D (mobility plot) the analysis of three dsDNAs is shown. The first has the wild type sequence, the second has the sequence of the IVS-1-6 mutation; i.e. the base pair at position 89 is changed from T:A to C:G. This change leads to a slight stabilization (about 1°C) of the first cooperative unit. The third dsDNA has a mismatch at position 89; the mismatch leads to a drastic destabilization of about 6°C.

The theoretical results shown are verified by TGGE (Fig. 2C). DNA of wild type sequence as well as DNA of the IVS-1-6 sequence were amplified separately. The obtained DNA fragments were mixed, denatured and renatured resulting in four different dsDNAs. Two of the DNAs are homologous double strands of wild type and of IVS-1-6 sequence, respectively; each of the two others has a mismatch at position 89, either a T·G- or a C·A-mismatch, respectively. Thus, in the gel four curves are visible. The retardation transitions of the homologous dsDNAs are at higher temperature separated by about 1°C; those of the heterologous dsDNAs are at lower temperature separated by about 2°C. The difference in T_m values of the heterologous dsDNAs is not predicted by theory because each mismatch is calculated as the same internal loop independent of the sequence of the loop. Because only the C·A mismatch is a real loop the calculated curve of the heterologous DNA corresponds to the curve with lowest T_m of the gel. In contrast to the predicted mobility curve the increase in mobility in the gel is observed as a two step process; probably the denaturation of base pairs ~20–~90, resulting in a molecule in which the strands are connected only by the GC-clamp, leads to an additional denaturation step (see thin arrow in Fig. 2B) followed by the strand separation into the single strands (see thick arrow in Fig. 2B).

With PCR one of the most critical parameters for purity and yield of the reaction is the temperature for annealing of oligomeric primers on the template to be amplified. Using standard PCR conditions Rychlik *et al.* (1) determined optimal annealing temperatures (T_a OPT; see column 1 of Table 5) for eleven primer–primer–template triplets and developed by rule of thumb a method for calculation of T_a OPT as a function of both the T_m of the less stable primer–template pair and the T_m of the product. In spite of all that, calculation of T_m of the less stable primer–template pair alone (using more suited parameters) is sufficient for determination of T_a OPT (see column 2 of Table 5). Only with primer λ 20161 annealing temperatures too high are calculated (see two bottom rows of Table 5); this is due to a high thermal stability (~71°C) but low specificity of the 5' part of the primer combined with a low thermal stability ($\leq 61^\circ\text{C}$) but high specificity of the 3' part from which the polymerase has to elongate. Notice that parameters for calculation (14) of T_m , p_1 are for 100 mM NaCl conditions which seem to be equivalent to the PCR conditions (0.8 mM dNTPs, 1.5 mM MgCl₂, 50 mM KCl, 10 mM Tris–HCl, pH 8.3).

Ionic strength dependence

An extrapolation of the calculated results to ionic strengths different of those implied by the thermodynamic parameters of helix growth is not included in the program. The following formula and values may be used for the extrapolation of calculated

T_m values (index 1) to the ionic strengths and solvent conditions of experiments (index 2):

$$(T_{m,2} - T_{m,1})/\log(c_2/c_1) = f_{G:C} * I_{G:C} + (1 - f_{G:C}) * I_{A:(U \text{ or } T)}$$

with T_m = transition (midpoint, melting) temperature of nucleic acid

c = ionic strength of buffer (= concentration of Na ions)

$f_{G:C}$ = G:C content of nucleic acid

$I_{X:Y}$ = ionic strength dependence of base pair type X:Y

$$\text{for DNA: } I_{A:T} = 18.30^\circ\text{C} \quad (31)$$

$$I_{G:C} = 11.30^\circ\text{C} \quad (32)$$

$$\text{for RNA: } I_{A:U} = 20.00^\circ\text{C}$$

$$I_{G:C} = 8.40^\circ\text{C} \quad (33)$$

In case of buffers containing Tris, formamide or urea, or of oligonucleotides instead of polynucleotides, correcting formulas are also available from the literature (33–38). The use of these formulas, especially in combination with each other, can only be seen as a first empirical approximation; therefore an incorporation into the thermodynamics of the program is avoided.

In Fig. 1C and D the optical melting curve was shifted by 28.8°C in order to correct for the theoretical condition of 1 M ionic strength instead of the experimental ionic strength of 0.01 M NaCl; a similar shift of 26.5°C is predicted by the above formula using the mean G:C content of the dsRNA, $f_{G:C} = 0.58$ (209 G:C pairs in 359 base pairs). Between the transitions of the calculated mobility curves in Fig. 1A and those of the temperature-gradient gel in Fig. 1B there is a difference in T_m values of about 48°C; the gel buffer contained 89 mM Tris–borate and 9 M urea whereas calculations were performed for 1 M ionic strength. Using a ionic strength value of $c_1 = 0.02$ M for the Tris buffer and $f_{G:C} = 0.58$ the formula gives a $\Delta T_m = 22.5^\circ\text{C}$. Thus a shift of 2.8°C/1 M urea remains, which is a reasonable value.

Between the T_m values of the mobility curves in Fig. 2A and of the temperature gradient gel in Fig. 2B is a difference of about 23°C; the gel buffer contained 8.9 mM Tris–borate and 4 M urea whereas calculations were performed for 0.019 M ionic strength. Using $f_{G:C} = 0.61$ (76 G:C pairs in 125 base pairs) and $c_1 = 0.003$ M for the Tris buffer the formula gives a $\Delta T_m = 11.3^\circ\text{C}$. Thus a shift of 2.9°C/1 M urea remains, which is nearly identical to the value determined above.

Design of an optimum probe for detection of mutations in TGGE and DGGE

The detection of mutations in genomic DNA by PCR amplification of a specific fragment followed by analysis with TGGE or DGGE is the problem for which the described program is most helpful. Thus, a protocol for the design of an optimum fragment using the presented program is described shortly in the following.

(i) Calculate the denaturation behaviour of a genomic DNA fragment of about 600 base pairs with the critical base pair(s) in the center of the fragment. Use the parameters given in the third column of Table 4 with large temperature steps of at least 2°C.

(ii) By examination of the $p(N,T)$ plot select a shorter fragment which denatures in a two-step process with the cooperative unit containing the critical base pair(s) denaturing first.

(iii) Vary borders of the fragment until the T_m difference between the first cooperative unit with and without the mismatch and the T_m difference between the first and the second cooperative unit are optimally large. Thus, decrease the temperature steps for calculation to 0.5 or 1°C and examine the $T_p = 0.5(N)$ and the mobility plots. In case there is no fragment which fulfills the conditions of steps (ii) and (iii), add a GC-clamp of appropriate size to the unit of higher stability or an AT-clamp to the unit of lower stability.

(iv) Select primers from the borders of the fragment. Avoid primers containing regions of complementarity between a primer and either itself or its partner. If the primers are not taken from the exact 5' or 3' end of the fragment, repeat steps (ii) and (iii).

(v) Select optimal temperatures for PCR and gradient-gel electrophoresis by extrapolation of the calculated T_m values to the conditions of the electrophoresis.

(vi) Synthesize selected primers, amplify the fragment from the wild type and from the mutated DNA, and perform the electrophoresis.

Steps (i)–(iv) will take up to an hour but increases enormously the rate of success in the experimental, most time-consuming step (vi).

ACKNOWLEDGEMENTS

I am indebted to Prof. Dr D. Riesner for stimulating discussions and support throughout the course of this work. I thank all the people using the programs in their work and thus giving reason for improvements. I thank especially M. Schmitz, Dr H. Werntges and U. Wiese for helpful comments. K.-H. Glatting and G. Raether adapted the program to the GCG/HUSAR conventions. Part of the work was supported by grants from the Deutsche Forschungsgemeinschaft.

REFERENCES

- Rychlik, W., Spencer, W.J. and Rhoads, R.E. (1990) *Nucleic Acids Res.* **18**, 6409–6412.
- Riesner, D., Henco, K. and Steger, G. (1991) In Chrambach, A., Dunn, M.J. and Radola, B.J. (eds.) *Advances in Electrophoresis* Vol 4. VCH Verlagsgesellschaft, Weinheim, pp. 169–250.
- Lerman, L.S., Fischer, S.G., Hurley, I., Silverstein, K. and Lumelsky, N. (1984) *Ann. Rev. Biophys. Bioeng.* **13**, 399–423.
- Wada, A., Yabuki, S. and Husimi, Y. (1980) *CRC Crit. Rev. Biochem.* **9**, 87–144.
- Gotoh, O. (1983) *Adv. Biophys.* **16**, 1–52.
- Wartell, R.M. and Benight, A.S. (1985) *Phys. Rep.*, **126**, 67–107.
- Wada, A. and Suyama, A. (1986) *Prog. Biophys. Molec. Biol.* **47**, 113–157.
- Poland, D. (1974) *Biopolymers* **13**, 1859–1871.
- Fixman, M. and Freire, J.J. (1977) *Biopolymers* **16**, 2693–2704.
- Devereux, J., Haerberli, P. and Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387–395.
- Werntges, H., Steger, G., Riesner, D. and Fritz, H.J. (1986) *Nucleic Acids Res.* **14**, 3773–3790.
- Freier, S.M., Kierzek, R., Jaeger, J.A., Sugimoto, N., Caruthers, M.H., Neilson, T. and Turner, D.H. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 9373–9377.
- Pörschke, D., Uhlenbeck, O.C. and Martin, F.H. (1973) *Biopolymers* **12**, 1313–1335.
- Klump, H.H. (1990) In W. Saenger (ed.), *Landolt-Börnstein, New Series, VII Biophysics, Vol. 1 Nucleic Acids, Subvol. c Spectroscopic and Kinetic Data, Physical Data I*. Springer-Verlag, Berlin, pp. 241–256.
- Breslauer, K.J., Frank, R., Blöcker, H. and Marky, L.A. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 3746–3750.
- Ueno, S., Tachibana, H., Husimi, Y. and Wada, A. (1978) *J. Biochem.* **84**, 917–924.
- Benight, A.S. and Wartell, R.M. (1983) *Biopolymers* **22**, 1409–1425.
- Kozyavkin, S.A., Mirkin, S.M. and Amirikyan, B.R. (1987) *J. Biomol. Struct. Dyn.* **5**, 119–126.
- Blake, R.D. (1987) *Biopolymers* **26**, 1063–1074.
- Gralla, J. and Crothers, D.M. (1973) *J. Mol. Biol.* **73**, 497–511.
- Jacobson, H. and Stockmayer, W.H. (1950) *J. Chem. Phys.* **18**, 1600–1606.
- Steger, G., Hofmann, H., Förtsch, J., Gross, H.J., Randles, J.W., Sängler, H.L. and Riesner, D. (1984) *J. Biomol. Struct. Dyn.* **2**, 543–571.
- Benight, A.S., Wartell, R.M. and Howell, D.K. (1981) *Nature* **289**, 203–205.
- Blake, R.D. and Haydock, P.V. (1979) *Biopolymers* **18**, 3089–3109.
- Coutts, S.M. (1971) *Biochim. Biophys. Acta* **232**, 94–106.
- Fischer, S.G. and Lerman, L.S. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 1579–1583.
- Steger, G., Tien, P., Kaper, J. and Riesner, D. (1987) *Nucleic Acids Res.* **15**, 5085–5103.
- Zimmat, R., Gruner, R., Hecker, R., Steger, G. and Riesner, D. (1990) In Sarma, R.H. and Sarma, M.H. (eds.) *Proceedings of the 6th Conversation in Biomolecular Stereodynamics*. Adenine Press, Schenectady, pp. 339–357.
- Qu, F., Heinrich, C., Loss, P., Steger, G., Tien, P. and Riesner, D. (1993) *EMBO J.* **12**, 2129–2139.
- Orkin, S.H. and Kazazian, H.H. (1984) *Annu. Rev. Genet.* **18**, 131–171.
- Owen, R.J., Hill, L.R. and Lapage, S.P. (1969) *Biopolymers* **7**, 503–516.
- Frank-Kamenetskii, M.D. (1971) *Biopolymers* **10**, 2623–2624.
- Steger, G., Müller, H. and Riesner, D. (1980) *Biochim. Biophys. Acta* **606**, 274–284.
- Michov, B.M. (1986) *Electrophoresis* **7**, 150–151.
- Klump, H. and Burkhart, W. (1977) *Biochim. Biophys. Acta* **475**, 601–604.
- McConaughy, B.L., Laird, C.D. and McCarthy, B.J. (1969) *Biochemistry* **8**, 3289–3295.
- Riesner, D. and Steger, G. (1990) In W. Saenger (ed.), *Landolt-Börnstein, New Series, VII Biophysics, Vol. 1 Nucleic Acids, Subvol. d, Physical Data II, Theoretical Investigations*. Springer-Verlag, Berlin, pp. 194–243.
- Record, M.T. and Lohman, T.M. (1978) *Biopolymers* **17**, 159–166.