

Small angle scattering as a tool to study the thermal denaturation of DNA

Kathleen Wood, Robert Knott, Ognyan Tonchev, Dimitar Angelov, Nikos Theodorakopoulos, Michel Peyrard

▶ **To cite this version:**

Kathleen Wood, Robert Knott, Ognyan Tonchev, Dimitar Angelov, Nikos Theodorakopoulos, et al.. Small angle scattering as a tool to study the thermal denaturation of DNA. EPL - Europhysics Letters, 2014, 108, pp.18002. 10.1209/0295-5075/108/18002 . ensl-01071136

HAL Id: ensl-01071136

https://ens-lyon.hal.science/ensl-01071136

Submitted on 3 Oct 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Small angle scattering as a tool to study the thermal denaturation of DNA

Kathleen Wood[1] , Robert Knott[1] , Ognyan Tonchev[2] , Dimitar Angelov[2] , Nikos Theodorakopoulos[3,4] and Michel Peyrard[5]

[1] *Bragg Institute, Australian Nuclear Science and Technology Organisation, New Illawarra Road, Lucas Heights NSW, Australia*
[2] *Laboratoire de Biologie Moléculaire de la Cellule, Ecole Normale Supérieure de Lyon, UMR CNRS/ENS 5239, 46 allée d'Italie, 69364 Lyon Cedex 07, France*
[3] *Theoretical and Physical Chemistry Institute, National Hellenic Research Foundation, Vasileos Constantinou 48, 116 35 Athens, Greece*
[4] *Fachbereich Physik, Universität Konstanz, 78457 Konstanz, Germany*
[5] *Laboratoire de Physique, Ecole Normale Supérieure de Lyon, 46 allée d'Italie, 69364 Lyon Cedex 07, France*

PACS 87.15.-v – Biomolecules: structure and physical properties.
PACS 61.05.cc – X-ray scattering (including small-angle scattering)
PACS 61.05.fg – Neutron scattering (including small-angle scattering)

Abstract – DNA thermal denaturation is the breaking of the base pairs, leading to a splitting of the two strands of the double helix. While it is easy to measure the fraction of open base pairs (f) versus temperature, determining the fraction (p) of fully open *molecules* is much harder. Previously, the simultaneous recording of f and p could only be achieved for special sequences. We show that small angle scattering of X-rays or neutrons allows the measurement of p for any sequence. We illustrate the method with a SAXS investigation of two sequences designed to exhibit different melting profiles and compare the SAXS data with nano-calorimetric measurements of the melting curve.

Introduction. – DNA is not the static entity shown by pictures of the double helix. It fluctuates widely. At biological temperature the base pairs may open locally, and when temperature increases, those disruptions of the helicoidal structure invade the full molecule and lead to a complete separation of the two strands. This thermal denaturation transition, also called DNA melting, has been widely studied for a long time [1], both experimentally and theoretically.

Following the opening of the base pairs in experiments is easy. The breaking of the base stacking is accompanied by a 40% increase of the UV absorbance at 260 nm which can be readily monitored. Calorimetry can directly measure the enthalpy change associated with the transition [2]. These studies measure the fraction f of open base pairs but they only give a partial view of the transition because they cannot distinguish between different strand-association states. Open pairs can either belong to partly open molecules in which open "bubbles" coexist with closed sections or to fully denatured molecules where the two strands are completely separated in the solution. To complete the measurement of f and better understand the nature of the melting transition, it is important to determine the fraction p of fully denatured *molecules*, which is not readily accessible in standard UV measurements.

Up to now the measurement of p could only be achieved for particular sequences, using a clever trick [3]. The idea is to select DNA sequences such that the single strands are partially self complementary. The DNA double helix forms because bases associate in pairs, adenine (A) with thymine (T) and guanine (G) with cytosine (C). The two strands are complementary to each other. If a segment of a single strand contains a sequence which is complementary to the sequence of another segment of the same strand however, then the strand can fold in a hairpin-like structure in which these two segments assemble in a double helix. The trick is that this hairpin formation is much faster than the assembly of two fully separated strands because the two elements do not have to find each other in the dilute solution as they belong to the same DNA

strand. Therefore when a partly denatured DNA solution is suddenly quenched to low temperature, all the molecules that had fully denatured strands lead to hairpins, while those which were only partly open re-close to their original double helix structure. This allowed the measurements of both the fraction of open base pairs f and the fraction of completely open molecules p in a landmark study [3]. However, as the method relies on self-complementary strands, it is limited to specific DNA sequences.

Here we introduce a method which is fully general and can be applied to *any* DNA sequence. We show that p can be deduced from small angle scattering experiments, whether they use X-rays (Small Angle X-ray Scattering – SAXS) or neutrons (Small Angle Neutron Scattering – SANS), and illustrate the method by studying the thermal denaturation of two DNA sequences selected to have different melting profiles although they differ by only a few base pairs.

Method and experimental technique. –

Method. Although the results are quantitatively different for X-rays and neutrons, the principles of small angle scattering experiments are the same for both and the method that we propose to measure the fraction p of fully denatured DNA molecules applies to both. Here we describe it in the context of the SAXS experiment presented in this paper. The intensity $I_N(\mathbf{q})$, scattered with scattering vector $\mathbf{q}$ by a sample comprising N molecules in solution, is the Fourier transform of the statistical average of the spatial correlation function $\gamma_N(\mathbf{r})$ of the electronic density $\rho(\mathbf{u})$, or more precisely the difference between the electronic density of the molecules and the average electronic density of the solvent. The electronic density is actually the sum over the contributions of the N molecules $\rho(\mathbf{u}) = \sum \rho_i(\mathbf{u})$ but, in the dilute limit the molecules are far apart and uncorrelated so that only the autocorrelation function of the electronic density of single molecules contributes, i.e. $\gamma_N(\mathbf{r}) = N\gamma_1(\mathbf{r})$. In the particular case of a DNA molecule, it is interesting to introduce the electronic densities $\phi^{(1)}$, $\phi^{(2)}$ of the two strands so that $\rho_1(\mathbf{u})$ can be expressed as

$$\rho_1(\mathbf{u}) = \phi^{(1)}(\mathbf{u} + \frac{\mathbf{a}}{2}) + \phi^{(2)}(\mathbf{u} - \frac{\mathbf{a}}{2}) \qquad (1)$$

where $\mathbf{a}$ denotes the vector joining the center of mass of the electronic distributions of the two strands. The correlation function $\gamma_1(\mathbf{r})$ becomes

$$\gamma_1(\mathbf{r}) = \gamma_1^{(1,1)}(\mathbf{r}) + \gamma_1^{(2,2)}(\mathbf{r}) + 2\gamma_1^{(1,2)}(\mathbf{r}) \ , \qquad (2)$$

with

$$\gamma_1^{(1,1)}(\mathbf{r}) = \int \langle \phi^{(1)}(\mathbf{u} + \frac{\mathbf{a}}{2})\ \phi^{(1)}(\mathbf{u} + \frac{\mathbf{a}}{2} - \mathbf{r}) \rangle\ d^3u \qquad (3)$$

and similar expressions for the other terms. In the limit $q \to 0$, the intensity $I_1(q = 0)$ simply amounts to the integration of the correlation function $\gamma_1(\mathbf{r})$ over the whole space. The result depends on the association state of the DNA molecule. For a denatured DNA molecule, the distance a between the electronic distributions of the strands is much larger than the typical size of the electronic distributions, which can be estimated to be of the order of the radius of gyration of the molecule. Therefore the electronic distributions of the two strands are statistically independent and fully decorrelated so that $\gamma_1^{(1,2)}(\mathbf{r}) = 0$. Only the autocorrelations of the single strands contribute, and, owing to the meaning of the ϕ functions, each one contributes n_1^2 to the spatial integration, where n_1 is the number of electrons in a single strand minus the number of solvent electrons in the same volume. Therefore, for a denatured molecule we get $I_1(q = 0) = 2n_1^2$. If the molecule is not fully denatured, even if it is partly open, the average value of a is less than the typical size of the single-strand electronic distribution. The two strands are not statistically independent and the correlation between $\phi^{(1)}$ and $\phi^{(2)}$ does not vanish. The spatial integration of $\gamma_1^{(1,2)}(\mathbf{r})$ again gives n_1^2 and consequently for a DNA molecule which is not fully denatured $I_1(q = 0) = 4n_1^2$. Therefore the calculation shows that $I_1(q = 0)$ is halved when the DNA molecules are fully denatured. This is not a surprise because it simply expresses the familiar result that, in small angle scattering experiments, $I_1(q = 0)$ is proportional to the molecular mass of the scattering molecules, which is valid both for X-ray or neutron small-angle scattering. The decrease of I_0 is due the loss of correlations between the beams scattered by the two strands when a DNA molecule is completely denatured. When they are fully separated, the strands can be viewed as independent molecules with a molecular weight which has been cut in half. This drastic decrease of scattering intensity turns out to be of considerable interest since it allows a direct measurement of p by following the temperature dependence of $I_0 = I(q \to 0)$ in a small-angle scattering experiment, using

$$\begin{aligned} I_0(T) &= (1-p)\ I_0(\text{low } T) + p\ I_0(\text{low } T)/2 \\ &= (1 - p/2)\ I_0(\text{low } T)\ , \end{aligned} \qquad (4)$$

where we denote by $I_0(\text{low } T)$ the value of I_0 at a temperature sufficiently low to ensure that the fraction of fully open DNA molecules is negligible. This expression is only accurate if n_1 can be assumed to be constant, i.e. if the difference between the electronic density of the DNA molecules and the bulk solvent does not vary significantly with temperature. Actually $\phi^{(1)}$, $\phi^{(2)}$ depend on the structure of the solvent in contact with DNA. Experiments show [4] that about 30% of the surface of double-stranded DNA is covered by a layer of hydration water which is about 2.8 Å thick and has a molar volume about 10% smaller than bulk water [5] whereas such confined water is not observed around single-stranded DNA. However the variation of $\phi^{(1)}$, $\phi^{(2)}$ due to the change in the structure of the hydration shell is small with respect to the contribution of the electronic density coming from the DNA

strands themselves. As a result the variation of $\phi^{(1)}$, $\phi^{(2)}$ plays a negligible role compared to their loss of correlation when the strands split apart.

Sample and experimental methods. We studied two DNA sequences shown in Fig. 1. The reference sequence (henceforth called "sequence A") is derived from a 200 base-pair segment from λ-phage DNA completed at one end by a set of 4 GC base pairs to prevent its opening by the ends when the denaturation temperature is approached. This 200 base-pair sequence has been investigated in [6] by cyclization experiments. It was selected because it does not have a spontaneous bending. The second sequence, henceforth called "sequence B" is derived from sequence A by switching a few base pairs from GC-type to AT (marked in blue in Fig. 1) or from AT to GC (marked in red in Fig. 1) to generate a series of 32 consecutive AT pairs, clamped at both ends by short GC-rich segments. As the AT pairs are linked by only two hydrogen bonds while the GC have three, this introduces a central domain in the sequence which has a lower thermal denaturation temperature than the average of the sequence. These sequences have been obtained from Eurofins [7] and

GGCGTCTCCTTTGATGCGAGTGCCAGCGTCAGACATCATGTGCAG
ATACTCACCTGCATCCTGAACCCATTGACCTCCAACCCCGTAATA
GCGATGCGTAATGATGTCGATAGTTACTAACGGGTCTTGTTCGAT
TAACTGCCGCAGAAACTCTTCCAGGTCACCAGTGCAGTGCTTGAT
CACAGGATGCTTCCCAGGATGCCG

GGCGTCTCCTTTGATGCGAGTGCCAGCGTCAGACATCATGTGCAG
ATACTCACCTGCATCCTGAACCCATTGACCTCCAACCCCGTAATA
GCGATTAATAATAATATAAATAATTATTAATATATCGGGTTCGAT
TAACTGCCGCAGAAACTCTTCCAGGTCACCAGTGCAGTGCTTGAT
CACAGGATGCTTCCCAGGATGCCG

Fig. 1: The two DNA sequences investigated by SAXS: sequence A, upper part, is the reference sequence. Sequence B, lower part, is obtained by changing a few bases in sequence A, marked in color, to create a segment of 32 consecutive AT pairs (yellow background), limited by several GC pairs at both ends. Pairs changed from GC to AT are marked in blue, while those changed from AT to GC are marked in red.

amplified by PCR [8] to prepare about 1 mg of each DNA type. The purity of the samples has been checked by gel diffusion. The DNA was suspended in D_2O because the samples were also prepared for some SANS experiments planned in the future. The solvent contains Tris 10mM to stabilize the pH, EDTA 1 mM to inhibit nucleases that could damage DNA, and $NaCl$ 100 mM. The concentration was 946 μg/mL for both samples.

A small amount of each solution was diluted to reach a volume of 1 mL suitable for calorimetry measurements using a nano-differential scanning calorimeter N-DSC-III. The thermal denaturation curve of each sample was recorded in the $20 - 95$°C range (Figs. 3-a and 3-b).

An aliquot of 20 μL was used for the SAXS experiments with a Bruker instrument using a 2D detector. The usable range of scattering vectors is $0.011\,\text{Å}^{-1} \leq q \leq 0.35\ \text{Å}^{-1}$. Data were recorded by heating the sample between 10°C and 95°C, at steps of 5 degrees initially and then 2 degrees. After complete denaturation, data were recorded on cooling between 95°C and 55°C every 10°C. At each temperature step, temperature was equilibrated for 20 minutes and then data recorded for one hour. Subsequently the temperature of the sample was calibrated using the 80.5°C melting of naphthalene as a reference. A temperature scan with the buffer was also made to account for the known temperature dependence in the scattering cross-section for water, and hence allow a more accurate subtraction of the buffer signal.

Data analysis. The length of the 204-base-pair DNA molecules is about 675 Å and the persistence length of double stranded DNA is of the order of 500 Å. Therefore for such samples an accurate low-q limit would require data for q as low as $q \approx 0.005\ \text{Å}^{-1}$ or lower, so that we cannot expect to extract I_0 from a simple Guinier plot of our data. A method that takes advantage of the data for all recorded q values has to be used. This can be done either in q space by fitting the data with a theoretical expression of the structure factor $S(q)$ or in real space by computing the pair-distance distribution function $P(r) = 4\pi r^2 \gamma_N(r)$ from an inversion of the Fourier transform which expresses $I_N(\mathbf{q})$ as a function of the spatial correlation function of the electronic density. Then I_0 can be obtained from an integration of $P(r)$ for all r. We used both approaches.

To fit in q space, the DNA chain can be described by a Kratky-Porod (KP) model [10]. It is a discrete version of the Worm Like Chain often used to model DNA properties, which is more appropriate to describe DNA up to its denaturation transition because denatured strands have a persistence length which extends over a few base pairs only [9]. Moreover, particularly in the case of sequence B including an AT-rich region which tends to open at lower temperatures than the rest of the chain, it is necessary to use a heterogenous KP model, with a rigidity that is site-dependent. A theoretical expression for the structure factor of such a model can be calculated by taking advantage of the mapping of the KP model to a Heisenberg model [11]. However it would be unrealistic to try to fit the 202 local interaction constants required to describe our DNA molecules. Instead we use a phenomenological expression which combines three limiting cases of the KP structure factor: $S_1(q)$ corresponding to a fully closed DNA molecule, i.e. a homogeneous KP model with a persistence length L_{p1} determined by a fit of the data at the lowest temperature ($L_{p1} = 334$ Å), $S_3(q)$ corresponding to denatured DNA, i.e. a homogeneous KP model with a persistence length L_{p3} determined by a fit of the data at the highest temperature ($L_{p3} = 40$ Å), and $S_2(q)$ for an heterogeneous KP model where sites of the AT-rich domain of sample B have a coupling constant corresponding to the persistence length L_{p3} and the other sites a coupling constant corresponding to L_{p1}. It is tempting to relate the

value $L_{p1} = 334$ Å, lower than the usually accepted value of 500 Å, to the recent speculations that short DNA chains could be more flexible than originally thought [12], however, while our data appear to be sufficient for an accurate determination of I_0, the lack of low q data in the SAXS experiment does not allow us to draw definite conclusions on the value of the persistence length of our DNA samples. The recorded signal $I(q)$, obtained after subtracting the background, is fitted by

$$I(q) = A\big[c_1 S_1(q) + c_2 S_2(q) + \frac{1}{2} c_3 S_3(q)\big] \ , \qquad (5)$$

where A, c_1, c_2, c_3 are parameters with the condition $c_1 + c_2 + c_3 = 1$. This is therefore a three-parameter fit once L_{p1} and L_{p3} have been fixed. The coefficient $\frac{1}{2}$ in front of c_3 takes into account our analysis showing that the scattered intensity is divided by 2 when the two strands are fully separated. From the fit we get $I_0 = A(c_1 + c_2 + \frac{1}{2}c_3) = \frac{1}{2}A(1 + c_1 + c_2)$. Tests show that I_0 is only weakly sensitive to random errors in the signal or small shifts in the background, while the parameters c_i are more sensitive to such errors. Therefore, although c_3 plays the same role as p in Eq. (4), the evaluation of the fraction of fully denatured molecules from Eq. (4) is more reliable than the fit in q space. Figure 2 shows an example of the experimental data and the corresponding fit, at temperature $T = 81°$C, slightly below the denaturation temperature.

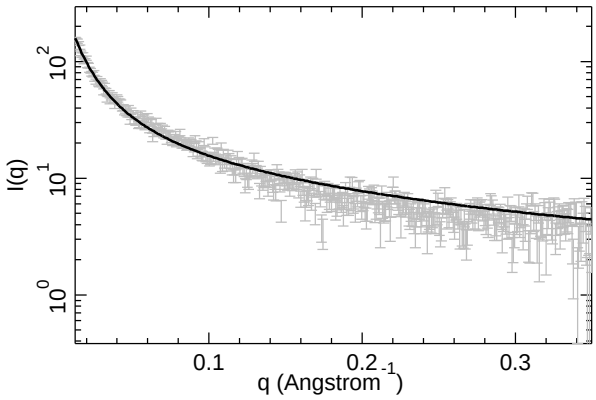


Fig. 2: Example of experimental result $I(q)$ for sequence B at $T = 81°$C, (logarithmic scale, counts, showing the experimental error bars) (grey) and the corresponding fit by Eq. (5) with $c_1 = 0.39$, $c_2 = 0.17$ (full black line).

The analysis in real space computes $P(r)$ by inversion of the Fourier transform that relates $I_N(\mathbf{q})$ to the correlation function $\gamma_N(\mathbf{r})$ of the electronic density and then integrates over the whole range of r to obtain I_0. The inversion is however not fully straightforward because it is an ill-posed problem. Small errors in $I(q)$ can lead to large errors in $P(r)$. Moreover, with the truncation of the q-range for the recorded data, it is prone to instabilities, often leading to spurious oscillations in $P(r)$ [13, 14]. A method using a regularization parameter has been introduced by [14]. It is implemented in the GNOM program available from EMBL [15]. We used this program to calculate $P(r)$. The boundary conditions set $P(r) = 0$ at both ends of the selected range $0.0 \leq r \leq 620$ Å. The value of $r_{\max} = 620$ Å was chosen because the length of the DNA molecules, if they were perfectly straight, would be 675 Å but, even at low temperature, for which the persistence length of DNA is of the order of 500 Å, we do not expect the molecules to behave as perfect rigid rods.

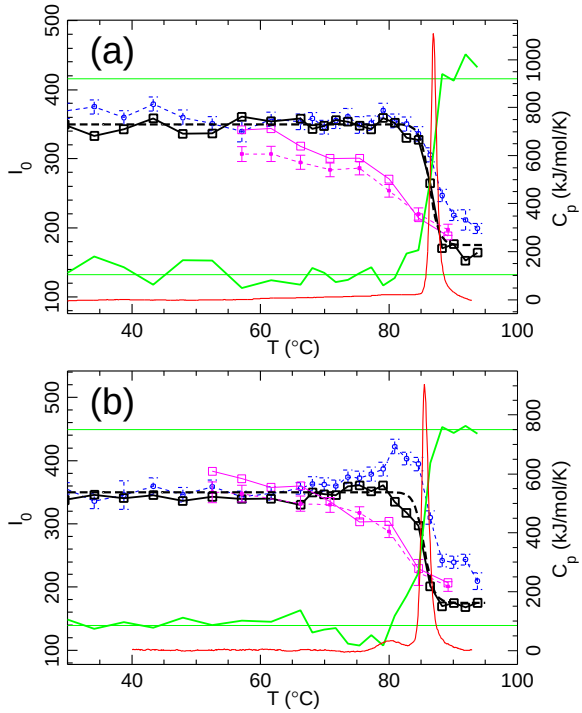


Fig. 3: Results of the calorimetry and analysis of the SAXS data for sequence A (a) and sequence B (b). The red lines show the specific heat measured by calorimetry. The squares and black lines show the variation of $I_0(T)$ on heating, deduced from the fit in q space. The black dashed lines show a fit of those data with the sigmoidal curve $I_{\rm fit}(T)$ (Eq. 6). The magenta squares and full line show $I_0(T)$ from the measurements made on cooling from the highest temperature. The blue circles and dotted lines show $I_0(T)$ deduced from the analysis in real space for the measurements on heating, and the magenta dots and dashed lines show $I_0(T)$ deduced from the analysis in real space for the measurements on cooling. The green full lines show the variation of the fraction of fully open molecules, on a scale going from 0 to 1 (marked by the two green horizontal lines), obtained from Eq. 4, computed with the results of the analysis in q space.

Results and discussion. – Figure 3 summarizes our results. The calorimetric curves (full red lines) show the difference in the denaturing transitions of the two sequences. Sequence A exhibits a one-step melting with a sharp peak at 87°C while for sequence B the large AT-rich domain leads to a two-step melting, with a small premelting maximum at 80° followed by full melting (sharp peak)

at 85.5°.

The variation of $I_0(T)$, also plotted of Fig. 3 can tell us more about the strand-association state of the DNA molecules when premelting occurs. First let us notice that both methods to determine I_0 are in good agreement in a large temperature range, as shown by Fig. 3, however some systematic discrepancies can be observed. They are noticeable for sample B in the temperature range corresponding to the premelting at around 80°C, and for both samples, after the full melting. The determination of I_0 by inversion of a Fourier transform to get $P(r)$ systematically overestimates I_0 when the persistence length of the molecules gets small. This is due to the lack of low-q data in our SAXS experiment. This can be tested by using the GNOM program [15] with data provided by the theoretical expressions $S_1(q)$, $S_2(q)$, $S_3(q)$ of the KP-model structure factor. If the inversion is made by providing data in the q range $0.002\,\text{Å}^{-1} \leq q \leq 0.35\,\text{Å}^{-1}$, the value of I_0 computed from $P(r)$ is accurate within a 0.1% error. If we truncate the q range and perform the inversion with $0.012\,\text{Å}^{-1} \leq q \leq 0.35\,\text{Å}^{-1}$, which corresponds to the range experimentally accessible, the computed I_0 is accurate within 2% for $S_1(q)$, which corresponds to double-helical DNA, but overestimated by 15% to 20% for $S_2(q)$ or $S_3(q)$, independently of the choice of the parameters used to control the accuracy of the calculation in the GNOM program. This error can be traced to an overestimation of $P(r)$ for large r by comparing with its value deduced from MD simulations of the KP model. This shows that the technical limitations of our SAXS measurements should not be ignored. The data obtained from the fit in q space appear to be much less affected and this is why we use them to calculate the fraction of open molecules. The calculation of $P(r)$ is nevertheless useful as a control and to help in the understanding of the results because the tests that we made show that $P(r)$ stays accurate for $r \leq r_{\max}/2 = 310$ Å.

Figure 3 shows that, for both samples, $I_0(T)$ determined by the fit in q space (black full line and squares) drops sharply when DNA denatures, as expected from our analysis of small-angle scattering experiments. This corresponds to the separation of DNA in two independent strands. As shown by the dashed black lines on Fig. 3 the decrease of $I_0(T)$ is well described by a function

$$I_{\text{fit}}(T) = \frac{1}{4}\, I_0(\text{low } T)\, \Big[3 - \tanh[(T - T_m)/\Delta T]\Big] \qquad (6)$$

which decays by a factor 2 around the melting temperature T_m in the temperature range ΔT. The fit gives $\Delta T = 1.5$°C, comparable to the width at half maximum of the calorimetry peak. The values of $T_m = 86.3$°C for sequence A and $T_m = 85.5$°C for sequence B, close to the denaturation temperatures observed by calorimetry, show that the measurement of $I_0(T)$ is also able to detect subtle effects associated to local changes in the sequence in spite of the temperature step of 2°C of the SAXS measurements in the transition region, which limits the accuracy in the determination of T_m.

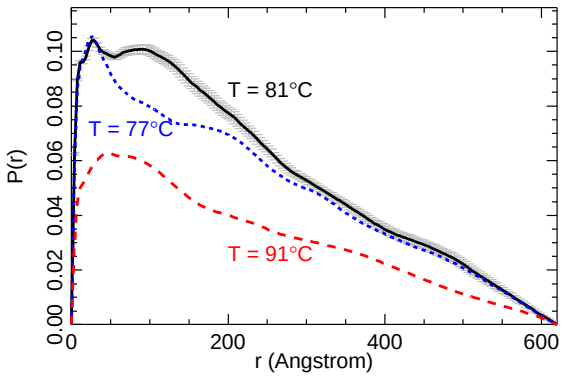


Fig. 4: The pair-distance distribution function $P(r)$ for sequence B at three temperatures: blue dotted line, $T = 77$°C, black full line $T = 81$°C, red dash line $T = 91$°C. The error bars for the case $T = 81$°C show the error computed by the GNOM program.

In spite of the dispersion due to the uncertainties in the determination of I_0, the separation of the strands shows up in the value of p deduced from Eq. (4), which increases from $p \approx 0$ at low temperature to $p \approx 1$ when the denaturation is complete. Interestingly, for sample B, Fig. 3-b shows that, slightly below T_m and for two consecutive temperatures, the measured $I_0(T)$ decays under the values of the fitting sigmoidal curve which corresponds to a single-step transition. Although the noise in $I_0(T)$ does not allow a definite conclusion, it suggests that, for this sample, the complete separation of the strands starts in the premelting stage.

The calculation of $P(r)$ shown in Fig. 4 supports the analysis. At $T = 77$°C, i.e. below the premelting maximum in the calorimetric results, $P(r)$ shows a distinct maximum for $r \leq 20$Å, which corresponds to the diameter of the DNA molecule. This is what we expect for the double helix structure: all pair distances up to the maximum size of the molecule must be represented, but, owing to the repetition of the motif of the base pairs, the typical size of this motif has to be heavily represented in $P(r)$. The backbone phosphates, lying on the outside of the helix, are the strongest scatterers, therefore it is natural to find a strong contribution for the value of r corresponding to the distance between the two phosphate groups of each base pair. The figure also shows a slight hump for very low r values ($r \approx 3$ Å) which is the distance between successive phosphate groups along the strands. At $T = 91$°C the fraction of denatured molecules is close to 1 according to Fig. 3 but nevertheless still below its maximum reached a few degrees above. The plot of $P(r)$ shows that the maximum for $r \leq 20$Å has almost fully disappeared although a small hump can still be detected at that position. The most significant contribution to $P(r)$ is around $r \approx 100$ Å however. This is what the KP model gives when we test it with an interaction constant corresponding to denatured strands.

The strong decrease of $P(r)$ at $T = 91°C$ corresponds to the expected decrease of I_0 when DNA denatures. At the intermediate temperature $T = 81°C$, slightly higher than the temperature of the premelting maximum in the calorimetry curve, $P(r)$ exhibits two maxima, one corresponding to the double helical structure and a second one for $r \approx 100$ Å which corresponds to denatured strands. The figure of $P(r)$ alone cannot distinguish unambiguously between the two strand-association states, one with a large open domain in the middle of the molecule, in the AT-rich region, and two closed regions on the two sides, and another situation where fully open molecules coexist with closed ones (or partly open ones). The decrease of I_0 observed on Fig. 3 at temperatures below the large drop due to the main denaturation transition suggests that the second option is the most probable.

The variation of $I_0(T)$ determined for the data recorded on cooling, plotted in Fig. 3, gives some information on the kinetics of renaturation. After complete denaturation the samples were equilibrated for 20 minutes at each temperature followed by a one-hour acquisition time, as they were on heating. A large temperature step of 10°C was used so that the cooling was too fast to allow a re-assembly of the DNA strands along the a pathway reverse to the one on heating. Therefore it is not surprising to find a large temperature lag as observed in Fig. 3.

Conclusion. – In small-angle scattering experiments from proteins the value of the intensity I_0 in the limit of a vanishing scattering vector is often used to determine the molecular weight of the scattering particle [16,17] and therefore as an indication of biomolecular oligomerisation or major changes due to drug/ligand binding for example. Interpretation of I_0 can be used as much more than an indication of sample quality [18] and here we have shown that actually $I_0(T)$ is an interesting quantity, which can be used to follow the fraction p of completely denatured molecules in a DNA denaturation experiment. This provides a tool to measure, for any DNA sequence, this quantity which could only be obtained for specific sequences and through a fairly complex process. Our experiments required less than 20 μg of DNA so that the method can be applied even when only small amounts are available. It is suitable for sequences not longer than a few persistence lengths. For sequences longer than several thousand base pairs, segments of the two single strands could become statistically independent in the presence of very large open domains, without a complete separation of the stands. This would affect the accuracy of the measurement of p. The method is general and can be applied to all processes involving the dissociation or association of molecules or molecular complexes.

Ideally I_0 could be readily obtained by extrapolation from a Guinier plot. Our experiments did not allow us to do this due to the lack of data at sufficiently small scattering vector. More elaborate methods to determine I_0 had to be used. What could appear as a drawback can also be viewed as a strength because it shows that one can nevertheless obtain useful results, albeit with higher uncertainties, without having to rely on highly sophisticated X-ray apparatus or small-angle neutron scattering. We showed that a laboratory SAXS experiment can be used to monitor the fraction of completely open molecules in a DNA denaturation and distinguish between different behaviors for two related DNA sequences, without imposing any particular condition on the sequence.

* * *

This work has been supported by ANSTO (Australian Nuclear Science and Technology Organisation) Bragg Institute as part of proposal 3203. O.T. and D.A. acknowledge support from European Community's Seventh Framework Programme FP7/2007-2013 under Grant agreement number 289611 "HEM_ID". We would like to thank A. Wildes and L. Porcar, ILL, Grenoble (France), with whom a related SANS project is under way, for helpful discussions.

REFERENCES

[1] Wartell R. M. and Benight A. S., *Physics Reports*, **126** (1985) 67.
[2] Breslauer K. J., Frank R., Blöcker H. and Marky L.A., *Proc. Natl. Acad. Sci. USA*, **83** (1986) 3746.
[3] Montrichok A., Gruner G. and Zocchi G., *Europhys. Lett.*, **62** (2003) 452.
[4] Bastos M., Castro V., Mrevlishvili G. and Texeira J., *Biophys. J.*, **86** (2004) 3822.
[5] Son I., Shek Y. L., Dubins D. N. and Chalikian V., *J. Am. Chem. Soc.*, **136** (2014) 4040
[6] Vologodskaia M. and Vologodskii A., *J. Mol. Biol.*, **317** (2002) 205.
[7] Eurofins Scientific, `http://www.eurofins.com`
[8] The primers used for PCR are: forward 5′-ggc gtc tcc ttt gat gcg ag-3′ and reverse 5′-cgg cat cct ggg aag cat cc-3′. The annealing temperature was 53°C. We used "OneTaq" Hot Start DNA Polymerase from BioLabs (`http://www.neb.com`) and "OneTaq" standard reaction buffer.
[9] Theodorakopoulos N. and Peyrard M., *PRL*, **108** (2012) 078104
[10] Kratky O. and Porod G., *Recl. Trav. Chim Pays Bas*, **68** (1949) 1106.
[11] Theodorakopoulos N., *Calculation of the structure factor of an inhomogeneous Kratky-Porod model.*, unpublished
[12] Kahn, J. D., *Biophys. J.*, **107** (2014) 282
[13] Glatter 0., *J. Appl. Cryst.*, **14** (1981) 101.
[14] Svergun D. I., *J. Appl. Cryst.*, **25** (1992) 495.
[15] `http://www.embl-hamburg.de/biosaxs/gnom.html`
[16] Glatter O. and Kratky O., *Small Angle X-ray Scattering* (Academic Press, London) 1982
[17] Jacques D. A. and Trewhella J., *Protein Science*, **19** (2010) 642.
[18] Gabel F., Jensen M. R., Zaccai G., and Blackledge M., *J. Am. Chem. Soc.*, **131** (2009) 8769.