Bubble nucleation and cooperativity in DNA melting

Yan Zeng, Awrasa Montrichok and Giovanni Zocchi

Department of Physics and Astronomy, University of California Los Angeles,
Los Angeles, CA 90095-1547

Abstract

Bubbles in DNA are related to fundamental processes such as duplication and transcription. Using a new ensemble technique to trap intermediate states, we present direct measurements of the average length of the denaturation bubble and the statistical weights of the bubble states in the temperature driven melting of DNA oligomers. For a bubble flanked by ds regions, we find a nucleation size of ~ 20 bases, and a broad distribution of bubble sizes. However for bubbles opening at the ends of the molecule there is no nucleation threshold. The measured statistical weights of different conformations agree with the predictions of the thermodynamic models in the case of unzipping from the ends; however, internal bubble states are not completely described by the models. The measurements further show that, due to end effects, the melting transition becomes a two-state process only in the limit of a molecule length L ~ 1 base pair.

Introduction.

Fundamental biological processes such as replication and transcription involve the opening and closing of the DNA double helix. In the cell, such DNA conformational changes are driven by the interaction with proteins; however, important properties of DNA which play a role in these mechanisms can be studied by thermal denaturation experiments in vitro. These include the stability of different DNA conformations, and the properties of denaturation bubbles. In addition, a quantitative understanding of the melting process for oligomers, in particular the cooperativity of the transition, is important to improve molecular biology techniques such as quantitative PCR. Finally, DNA melting is a beautiful polymer physics problem in itself, with open questions regarding both the statics and dynamics of different conformations.

DNA melting is a true phase transition [1], driven by the fact that ds and ss DNA are very different polymers: the former rather rigid (persistence length $\ell p \sim 150$ bp), the latter very flexible ($\ell p \sim 3$ bases). Thus the transition is governed by a competition between the energy cost of breaking base pairs and the entropy gain of the more flexible ss loops. As the temperature is raised, a long DNA duplex will start to develop melted ss regions ("bubbles"), before completely separating into single strands.

Theoretical studies of DNA melting start from a description in terms of helix-coil transition [2-6], or from Hamiltonian models [7-8]. Key questions include the estimation of the entropy of single-stranded loops, the role of sequence heterogeneity in determining the nature of the transition, and generally how to summarize relevant degrees of freedom into effective parameters such as stiffness and excluded volume [8-12].

Experimental melting profiles can be obtained by a variety of techniques, including spectroscopic methods [13] such as UV absorption (which measures base stacking and pairing), circular dichroism (CD) (which measures helix content), and fluorescent energy transfer (FRET) [14]; calorimetry [15-17], and electrophoretic mobility assays [18-19]. From the melting curves of opportunely chosen sequences, the free energies for pairing and stacking can be determined [20-22]; these parameters are used in the thermodynamic models (such as the nearest neighbor NN model [23] to compute the statistical weight of different conformations and predict melting profiles [24]. In this approach, one needs a model (e.g. the two-state model) to extract the thermodynamic parameters from the data.

Here we introduce a different approach, where we measure directly conformations (bubble length) and their statistical weights. These model-independent measurements validate the predictions of the nearest neighbor (NN) thermodynamic model [25] in the case of unzipping from one end. However, for the case of an "internal bubble" (bubble flanked by ds regions), we find features, such as the nucleation size of the bubble, which are not predicted by the NN model.

The experiments are based on a simple quenching method by which we can directly measure the fraction of intermediate states, and also the length of the bubble [26-27]. By designing sequences appropriately (exploiting the larger pairing energy for G-C compared to A-T) we have studied the case of a single internal bubble, and the case of a bubble opening at one (or both) ends of the molecule. For internal bubbles, we measure a nucleation size and find evidence for large fluctuations of the bubble size. For bubbles at one end, we find no nucleation threshold. An analysis of the statistical weight σ of intermediate (partially open) states vs. length of the molecule L shows that due to these end effects, the transition becomes strictly two-state only for $L \approx 1$.

Materials and Methods.

Synthetic DNA oligomers were purchased from Qiagen, some batches HPLC purified, some salt-free. Annealing of complementary single strands was performed by heating to 90°C in a water bath for 5 minutes and then slowly cooling down to room temperature overnight. All experiments were performed in phosphate buffer saline (PBS) at an ionic strength of 50 mM**.** DNA concentration was 1 μM for all experiments**.** The sequences used in the study were:

a) Sequences for "Bubble in the middle" study:

L60b36:CCGCCAGCGGCGTTATTACATTTAATTCTTAAGTATTATAAGTAATAT
GGCCGCTGCGCC

L42B18: CCGCCAGCGGCGTTAATACTTAAGTATTATGGCCGCTGCGCC

L33B9: CCGCCAGCGGCCTTTACTAAAGGCCGCTGCGCC

b) Sequences for "Bubble at the end" study:

L48AS: CATAATACTTTATATTTAATTGGCGGCGCACGGGACCCGTGCGCCGCC

L36AS: CATAATACTTTATATTGCCGCGCACGCGTGCGCGGC

L30AS: ATAAAATACTTATTGCCGCACGCGTGCGGC

L24AS: ATAATAAAATTGCCCGGTCCGGGC

L19AS_1: GCAGCGGCCTGGCCGCTGC

L19AS_2: ATAATAAAGGCGGTCCGCC

L13AS: GCCGCCAGGCGGC

L11AS: CCGCCAGGCGG

Quenching method. We developed the method described below in order to address the following problem. A melting curve obtained e.g. by UV absorption, when normalized between 0 and 1, is taken to represent the fraction of open bp, which we call *f*. However, this curve alone does not allow to pinpoint the presence or absence of intermediate states. For instance, at the midpoint of the transition (f = ½), one cannot tell whether all molecules in the sample are half-way open, or whether half are completely open and half completely closed. With the following method, we determine at each temperature the fraction of completely open molecules, and therefore the fraction of molecules in intermediate states.

We restrict ourselves to sequences that are partially self-complementary, i.e. the single strands can form hairpins. We prepare the initial state, by careful annealing, in the duplex form (which is the ground state). This is confirmed by gel electrophoresis of unheated sample aliquots, which show only a duplex (ds) band (see also Fig.1b, 40 °C lane). For each sequence we check these initial relative populations of duplexes (ds) and hairpins (hp). If the initial hp fraction is not negligible, we re-design the sequence, introducing a few mismatches in the hp (not the ds), to increase the difference in stability between hp and duplex. Upon heating to a temperature T_i within the transition region, we will obtain a mixed population of partially and completely open molecules (Fig.1a). Now the sample is rapidly quenched to ~ 0 °C. Under the diluted conditions of the experiment, the completely separated strands form hairpins, while the partially open molecules close again as duplexes. After the quench, we have a mixed population of hairpins and duplexes; the fraction of hairpins represents the fraction of completely open molecules at the temperature T_i before the quench [26-27]. This fraction, which we call p, is determined by gel electrophoresis from the relative intensities of the hairpin (hp) and duplex (ds) bands (Fig.1b). Quenching has to be fast enough and the hairpins sufficiently

stable, that hp + hp $\rightarrow$ ds recombination during and after the quench is small. We confirm this by observing that the ratio of the band intensities hp_{meas} / (hp_{meas} + ds_{meas}) saturates at a value close to 1 for high enough temperatures. The corresponding amount of recombination is taken into account by introducing a correction in the normalization of p. In the simplest model, the recombination rate is proportional to the concentration of hairpins squared (two body collisions); this leads to a correction factor γ in the normalization of p [27]:

$$p = \frac{hp_q}{hp_q + ds_q} = \frac{hp_{meas}}{hp_{meas} + ds_{meas}} \frac{1}{1 - \gamma\, hp_{meas}} \qquad (3)$$

where the subscript q means "quenched" (i.e. the value of the quantity right after the quench), and *meas* means "measured" (a time t after the quench, the same t for all aliquots in one run). The factor γ is found from the band intensities in the gel, by enforcing hp_q / (hp_q + ds_q) = 1 at high temperature.

Experiments were carried out as follows: 20 μL aliquots of the sample in PCR tubes were heated to the desired temperature T_i for 3 minutes in a water bath, and then quenched in chilled water. Gels (3% agarose) were stained with Ethidium Bromide and photographed under UV illumination. Running time for gel electrophoresis varied according to sequence length; for example, 80 minutes at 120 Volts for a 30mer sequence. The integrated intensities of the duplex (slow) and hairpin (fast) bands were determined from the digital pictures with an image processing program and used to obtain the fraction of open molecules p at each temperature. This analysis assumes that the fluorescent intensity of a band in the gel is proportional to the amount of DNA in the band, with the same proportionality constant for hp and ds bands. For the sequences of this study, we confirm this by noting that the sum of the intensities of the two bands (hp + ds) is the same (within experimental resolution) for all lanes in the gel, even though the relative intensities change according to the different temperatures before the quench. This "sum rule" can be seen in the gel Fig.1b. More generally, p can also be extracted in a manner independent of the relation between hp and ds fluorescence, by comparing hp bands (or ds bands) across lanes.

UV spectroscopy. UV absorption around 260 nm arises from the π-π* electronic transition in both purine and pyrimidine bases**.** An increase in absorption represents a change in the electronic configuration of the bases due to the decrease in base stacking and pairing [22]. The absorption vs. temperature was monitored at 260 nm using a Beckmann DU-640 spectrophotometer with temperature controlled sample cell; experiments were run in steps of 0.5 °C per minute, and sample volume was 800 μL. The UV absorption curves were normalized from 0 (base line before the transition) to 1 (shoulder of the S-shaped transition curve), thus representing the fraction of open bases *f*. The "shoulder" is defined by the first point which falls below a straight line drawn through the high temperature part of the melting curve (see Fig.2a), where UV absorption increases because of unstacking in the single strands [13]. Given the resolution of our data, this point is typically defined within 0.5 °C. In cases where the slope due to unstacking is large (Fig.2c), it is delicate how to normalize based on the UV curve alone; an advantage of the quenching method is that it provides an independent information which we can use to normalize the UV melting curves, i.e. by enforcing f = 1 for p = 1.

We have done extensive modelling of the contribution of pairing and stacking to the melting curves [28]. The conclusion is that the UV absorption curve thus normalized, *f*, correctly represents the fraction of open base pairs for f < 1.

From the quenching method, we obtain the fraction of completely open molecules, *p*(T). From the UV spectroscopy, we obtain the fraction of open base pairs *f*(T). This gives a direct measure of the presence of intermediate (partially open) states. Indeed, if there are bubbles in a significant fraction of the molecules at a given temperature, then *p* < *f*. This means part of the UV absorption *f* comes from open bases in intermediate states. On the contrary, for a strictly two-state transition, *p* = *f*, since open base pairs come entirely from completely separated molecules. Moreover, we can obtain the following quantities.

$$\sigma = f - p \qquad (1)$$

is the fraction of bases in the sample which are part of a bubble. Writing the fraction of open bases as: $f(\mathrm{T}) = (1 - p(\mathrm{T}))\langle \ell \rangle + p(\mathrm{T})$, where $\langle \ell \rangle$ is the average fractional length of the bubble (averaged over the subset of the partially open molecules), we obtain the average fractional bubble length:

$$\langle \ell \rangle = (f - p) / (1 - p). \quad (2)$$

Finally, we note that the position of the melting curves along the T axis depends, as is well known, on several factors such as ionic strength and oligomer concentration. For this reason, experiments were performed under the same conditions for the spectroscopy and the quenching measurements. The consistency of the temperature calibration between the spectrometer and the water bath used for the quenching measurements was checked using a third thermistor.

Results.

First we present the case of a single bubble flanked by ds regions ("bubble in the middle"), then the case of a bubble opening at one or both ends, and finally we give an analysis of the cooperativity of the transition. The sequences used in the study are available as supporting information.

Bubble in the middle. For this case, we studied a series of three sequences, clamped at the ends by identical GC rich regions, and having AT rich middle regions of different lengths B: L60B36 (total length = 60, length of AT region = 36), L42B18, L33B9. The melting curves (Fig.2) reveal directly the presence of intermediate (bubble) states, since $p < f$ in the transition region. While the spectroscopic melting curves f(T) are rather structure-less (owing to the finite size of the molecule), the calculated average length of the bubble $\langle \ell \rangle$ shows definite structure, in particular a plateau at a value $\langle \ell \rangle_{plateau} \approx B =$ B/L, which corresponds to the fractional length of the AT rich region. This plateau is reminiscent of the isotherms in the P-V plane for a liquid-gas transition, and is the signature of a discontinuous transition for the infinite system. However the curves in Fig.2 contain additional information. In Fig.2a (B = 36), the bubble appears to open continuously, i.e. $\langle \ell \rangle = f$ before the plateau. However in Fig.2b (L = 18) the $\langle \ell \rangle$ and f curves are distinct as soon as $\langle \ell \rangle \neq 0$, i.e. before the plateau, not all open base pairs can be accounted for as originating from the bubble. In other words, as soon as $\langle \ell \rangle \neq 0$, there are also completely open molecules ($p \neq 0$). This trend is confirmed by the curves of Fig.2c, and is not an effect due to the overall length of the molecule, as we show in the next section. The effect is due to the length B of the AT rich region. We conclude that there is a minimum length for the stable bubble, of order ~ 20 bases (approximately two

turns of the double helix). This conclusion is supported by the following analysis. Some information can be obtained on the second moment of the distribution of bubble lengths, i.e. on fluctuations, by taking derivatives of the <ℓ> vs. T curve. This is easily seen e.g. starting from a partition sum of the generic form:

$$Z = \sum_{\ell} g(\ell)\, e^{-\varepsilon \ell / T} \tag{3}$$

(where ℓ is the length of the bubble, ε the energy cost of breaking one bp, and g is the number of states associated to the bubble), which leads to the following relation:

$$\frac{\partial}{\partial T} < \ell >= \frac{\varepsilon}{kT^2} [< \ell^2 > - < \ell >^2] \tag{4}$$

This is similar to the well-known relation between energy fluctuations and specific heat; in the context of protein folding, a similar van't Hoff relation is often used to assess how closely the folding transition is approximated by a two-state process. Indeed, it is easy to see that α as defined below takes, in the two extreme cases, the following values:

$$\frac{< \ell^2 > - < \ell >^2}{< \ell >^2} = \alpha = \begin{array}{ll} 1 & \text{two - state distribution} \\ 1/3 & \text{all intermediates equi - probable} \end{array} \tag{5}$$

Using (4), we have calculated the α values for the data of Fig.2 a, b, at the point where d<ℓ>/dT is maximum. The energy parameter ε was obtained by fitting the f(T) curve with the zipper model (3), with $g(\ell) = s^{\ell}$ (yielding ε ≈ 5 kT). The result is:

α = 1.37 ± 0.45 for L42B18

α = 0.38 ± 0.13 for L60B36.

The large error bars notwithstanding (which are due to taking the numerical derivative of the <ℓ> vs. T curves), we find that the α value for L60B36 is consistent with a wide distribution of bubble lengths. However the α value for L42B18 is consistent with a two-state distribution (bubble either open or closed). This supports the conclusion that there is a minimum length of the bubble, somewhere between 18 and 36 bases.

Bubble at the end. For this case we studied a set of 8 molecules of decreasing lengths (see supporting information for sequences), with an AT rich region at one end and a GC rich region at the other. Some representative melting curves are shown in Fig.3. In the case of L48AS even the spectroscopy curve alone shows that one end of the molecule melts at lower temperature than the other, as there are two kinks in the curve. For the

shorter molecules, f(T) is structure-less, but the presence of partially unzipped states is pinpointed by comparing f(T) and p(T). In all cases the bubble size grows smoothly with T, i.e. $\langle \ell \rangle = f$ for a substantial part of the transition region. Thus there is no minimum size of the bubble in this case. This is in contrast to the nucleation effect displayed by the data of Fig.2 b,c (where $\langle \ell \rangle \neq f$ throughout), and it shows that the former is not an effect of the overall length of the molecule: e.g. for L19AS the bubble opens continuously, although the molecule is much shorter than L42B18. In conclusion, there is a nucleation size for bubbles in the middle, but not for bubbles at the ends.

Statistical weight of intermediate states. If there is no nucleation size for bubbles opening at the ends, we expect that for a generic sequence the transition will never be strictly two-state. The parameter $\sigma = f - p$ represents the fraction of bases participating in a bubble state, thus it offers a quantitative measure of the incidence of intermediate states. Fig.4 shows σ vs. (T-Tm) for a series of molecules which unzip at the ends: clearly the presence of intermediate states decreases with decreasing length of the molecule. Note however that a typical sequence of length L ~ 20 is still very far from a two-state transition ($\sigma_{max} \sim 0.3$). In order to summarize in one plot the length dependence of the incidence of intermediate states, we can choose different measures, for instance the maximum value σ_{max} , or the area under the σ vs. T curve divided by the width of the curve; we call this latter quantity σ_{av} and plot it vs. L in Fig.5 (the plot of σ_{max} is basically identical). The first result is that all points fall onto the same curve, thus there is some generality to this graph (although we believe it is possible to specifically design sequences to show different behaviors, by carefully compensating end effects through the GC / AT ratio). The most interesting feature is that the data extrapolate to $\sigma_{av} = 0$ (a strictly two-state transition) for L close to 1. To make the point, we show on the graph the best linear fit to the data. Unfortunately it is not possible with the present method to explore lengths smaller than approximately 10, because of the instability of short hairpins. Thus we cannot exclude that for smaller L the plot could bend sharply and cross the abscissa at some value of L between 1 and ~ 10. But we find the extrapolation shown in Fig.5 very suggestive, and also consistent with the idea that there is no finite nucleation size for bubbles opening at the ends.

Discussion.

Through the melting curves displayed in Fig.2 we find evidence that for a bubble flanked by ds regions there is a nucleation size, of order ~ 20 bases, i.e. roughly 2 turns of the double helix. We are not aware of any other direct measurement of bubble size.

In the case where the bubble opens at the ends, we do not see a nucleation size. This is quite plausible, since the end is already a defect, and probably "fraying" [29]. However, it is then reasonable to expect that even short sequences will not melt through a two-state process.

Let us contrast this behavior with the folding transition for single-domain proteins. In the case of small proteins the transition is cooperative in the very direct sense that often removing even a small part of the polypeptide chain prevents the molecule from folding altogether [30-32]. The opposite is the case with DNA, which can be cut into smaller and smaller pieces and still remain stable (although the melting temperature is lowered). In this sense, the melting transition of DNA is non-cooperative, as a result of end effects. On the other hand, if the ends are clamped, then an "internal" bubble opens cooperatively.

We have compared our direct measurements of conformations and statistical weights with the predictions of the NN thermodynamic model, by obtaining the free energy differences ΔG of intermediate state conformations from the program MFOLD [25], and constructing the corresponding f and p curves. Our sequences were entered in the program as longer single strands obtained by joining the original ss with its reverse complement through a short spacer (TTT), so in the simulation, the "duplex" has only one end. Complete strand dissociation ($p = 1$) in the measurements corresponds to all bases unpaired in the simulation. For sequences which unzip from the ends, there is good agreement between measurements and simulation (Fig.6a). For internal bubbles, the measurements show a minimum (nucleation) size for the bubble, which is not captured by the simulation (Fig.6 b,c).

Finally, we discuss two issues concerning the method used in this study. One question is whether the conclusions depend on the use of self-complementary sequences. We have some control over this aspect, because we do not need to use sequences which form complete hairpins, and also we can introduce mismatches in the hairpins. Experiments

using sequences with different degrees of self-complementarity show similar results. For instance, in Fig.5 the two L = 19 points were obtained with two different sequences of different degree of self-complementarity; the two points nonetheless coincide within experimental error. So we do not see dramatic effects due to self-complementarity.

A second issue is whether quenching may partially drive the transition (for the p(T) measurements). But in this case we should occasionally observe $p > f$, which is never the case. In fact, the $p = 1$ point is always consistent with the shoulder of the f curve which signals the end point of the melting transition. In short, we never observe anything indicating that quenching drives the transition for the p(T) curves.

Acknowledgements.

This research was partly supported by the US-Israel Binational Science Foundation under grant no. 2000298. We thank Kevin Plaxco for helpful discussions and the referees for valuable suggestions about the manuscript.

Figure captions

Fig.1: a) Schematic graph of the quenching method used to trap intermediate states. The lower part indicates the two kinds of sequences used in the study, and how they form bubbles.

b) Example of a gel, running from right to left. The slow (fast) band corresponds to duplexes (hairpins). The numbers on the right end of the lanes give the temperatures to which the aliquots were heated before quenching. The plot on the right shows the intensity profiles. The numbers are proportional to the areas under the peaks, and are used to calculate the fraction of open molecules p.

Fig.2: Melting curves for the three sequences L60B36 (length L = 60, length of the AT rich "bubble forming" region B = 36), L42B18 and L33B9. Clamped at the ends by identical GC rich region and having AT rich middle regions of different lengths B, the duplexes form single bubbles in the middle when the temperature is increased. The open circles represent the fraction of open base pairs f (from the UV absorption measurements); the filled circles represent the fraction of open molecules p (from the gels), and the squares represent the average relative length of the bubble $\langle \ell \rangle$ calculated from eq. (2).

a) $\langle \ell \rangle$ grows smoothly from zero and reaches a plateau for $\langle \ell \rangle \approx 0.6 = B/L$, the relative size of the AT rich middle region for this sequence. The arrow indicates the "shoulder" of the melting curve (end point of the transition).
b) With a shorter AT rich middle region (B/L = 0.42) the average bubble length reaches a plateau at a correspondingly smaller value $\langle \ell \rangle \approx 0.3$. Bubble opening is not continuous: $\langle \ell \rangle \neq f$ even before the plateau.
c) With an even shorter AT rich middle region, $\langle \ell \rangle \neq f$ as soon as $\langle \ell \rangle \neq 0$.

Fig.3: Melting curves for the two sequences L48AS (a) (L = 48, "asymmetric"), L19AS (b), a subset of the 8 sequences with an AT rich region at one end and a GC rich region at

the other. These duplexes form a bubble at one end when the temperature is increased. Symbols denote the same quantities as in Fig.2. In all cases, $<\ell> = f$ for a substantial part of the transition region, indicating that there is no minimum size of the bubble. Note that even with a short molecule (L = 19 in b) $<\ell> = f$ initially, in contrast to Fig.2 b,c.

Fig.4: The quantity $\sigma = f - p$, which gives the fraction of bases in a bubble state, plotted for a subset of the 8 sequences which form bubbles at the ends. The statistical weight of intermediate states is smaller for the shorter sequences.

Fig.5: σ_{av} is the area under the σ curve (see Fig.4) divided by the width of the peak, here plotted vs. the length of the molecule L for the 8 sequences which form bubbles at the ends. This quantity represents a measure of the frequency of intermediate states, averaged over the transition region. The two L = 19 points were obtained with two different sequences. For a strictly two-state transition, $\sigma_{av} = 0$, and extrapolation of the data indicates that this happens only for $L \approx 1$. The straight line is a linear fit through the data.

Fig.6: The melting curves f and p, and the bubble length $<\ell>$, obtained from the program MFOLD; symbols are the same as in previous figures, and the sequence is indicated on the graph. For comparison, the experiment data is shown in the inset.
a) For unzipping from the ends (L19), the simulation reproduces the experiments well (in the experiment, $\sigma = f - p$ is a factor 2 larger because the molecule can unzip from both ends).
b) and c) For internal bubbles, the agreement is not complete because the simulation always shows $<\ell> = f$ through the beginning of the transition region, whereas the experiments show $<\ell> < f$ for the shorter sequences, indicating a minimum size for the stable bubble.

References

[1] A.Y. Grosberg and A. R. Khokhlov (1994) "Statistical Physics of Macromolecules", chapter 7, AIP Press, NY.

[2] Hill, T.L. (1959) J. Chem. Phys. **30**, 383.

[3] Zimm, B.H. (1960) J. Chem. Phys. **33**, 1349.

[4] Poland, D. and Scheraga, H.A. (1966) J. Chem. Phys. **45**, 1456-63, 1464-69.

[5] Fisher, M.E. (1966) J. Chem. Phys. **45**, 1469-73.

[6] Azbel, M.Ya. (1979) Phys. Rev. A **20**, 1671-84.

[7] Peyrard, M. and A.R. Bishop (1989), Phys. Rev. Lett. **62**, 2755.

[8] Theodorakopoulos, N., Dauxois, T. and Peyrard, M. (2000) Phys. Rev. Lett. **85**, 6-9.

[9] Kafri, Y., Mukamel, D. and Peliti, L. (2000) Phys. Rev. Lett. **85**, 4988-91.

[10] Causo, M.S., Coluzzi, B. and Grassberger, P. (2000) Phys. Rev. E **62**, 3958-73.

[11] Carlon, E., Orlandini, E., and Stella, A.L., Phys. Rev. Lett. **88**, 198101 (2002).

[12] Cule, D. and Hwa, T. (2000) Phys. Rev. Lett. **79**, 2375-78.

[13] Cantor, C.R. and Schimmel, P.R. (1980) *Biophysical Chemistry* (Freeman, NY).

[14] Gelfand, C.A. et al (1999) Proc. Natl. Acad. Sci. USA **96**, 6113-18.

[15] Senior, M.M., R.A. Jones, and K.J. Breslauer, Proc. Natl. Acad. Sci. USA **85**, 6242-6 (1988).

[16] Vesnaver, G. and Breslauer, K.J. (1991) Proc. Natl. Acad. Sci. USA **88**, 3569-73.

[17] S.S. Chan et al., Biochemistry **32**, 11776 (1993).

[18] R.M. Myers et al., Nucleic Acids Res **13**, 3111 (1985).

[19] Zhu J, Wartell RM. (1997) Biochemistry **36**, 15326-35.

[20] Blake, R.D. and Delcourt, S.G. (1998) Nucleic Acids Res. **26**, 3323-32.

[21] SantaLucia, J.,Jr. (1998) Proc. Natl. Acad. Sci. USA **95**, 1460-65.

[22] Bloomfield, V.A., D.M.Crothers and I. Tinoco, Jr. "Nucleic Acids: structures, properties, and functions" (University Science Books 1999).

[23] J. SantaLucia, Jr., Proc. Natl. Acad. Sci. USA **95**, 1460 (1998).

[24] Blake, R.D. et al (1999), Bioinformatics **15**, 370-75.

[25] M. Zuker, Nucleic Acids Res. **31 (13)**, 3406-15 (2003).

[26] A. Montrichok, G. Gruner, and G. Zocchi, Europhys. Lett. **62**, 452 (2003).

[27] Y.Zeng, A.Montrichok and G. Zocchi (2003), Phys. Rev. Lett. **91**, 148101.

[28] V. Ivanov, Y. Zeng and G. Zocchi (2003), *Statistical mechanics of base stacking and pairing in DNA melting,* submitted.

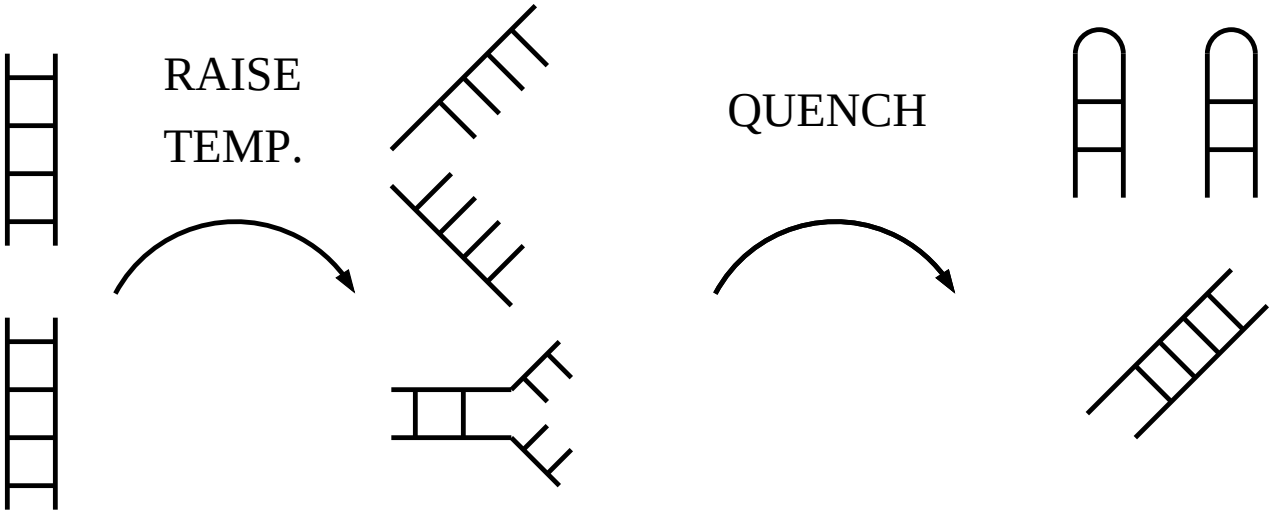
[29] Vercoutere, W.A. et al (2003)**,** Nucleic Acids Research, **31**, 1311-18.

[30] Flanagan, J.M., Kataoka, M., Shortle, D., Engelman (1992), D.M., Proc. Natl. Acad. Sci. USA **89**, 748-52.

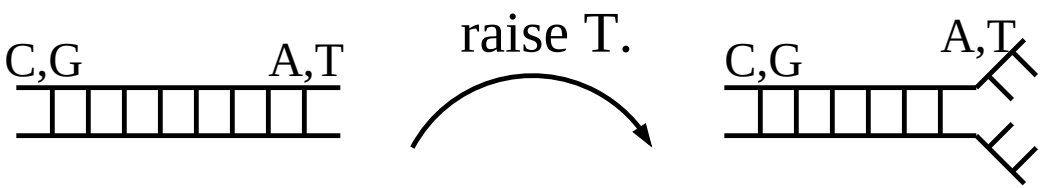
[31] A.G. Ladurner, L.S. Itzhaki, G.D. Gay and A.R. Fersht (1997), J. Mol. Biol. **273**, 317–29.

[32] Jewett AI, Pande VS, Plaxco KW (2003), J. Mol. Biol. **326**, 247-53.

Fig. 1a



Bubble at the end:



Bubble in the middle:

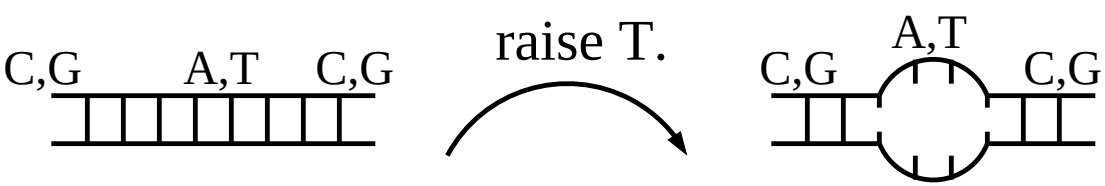


Fig. 1b

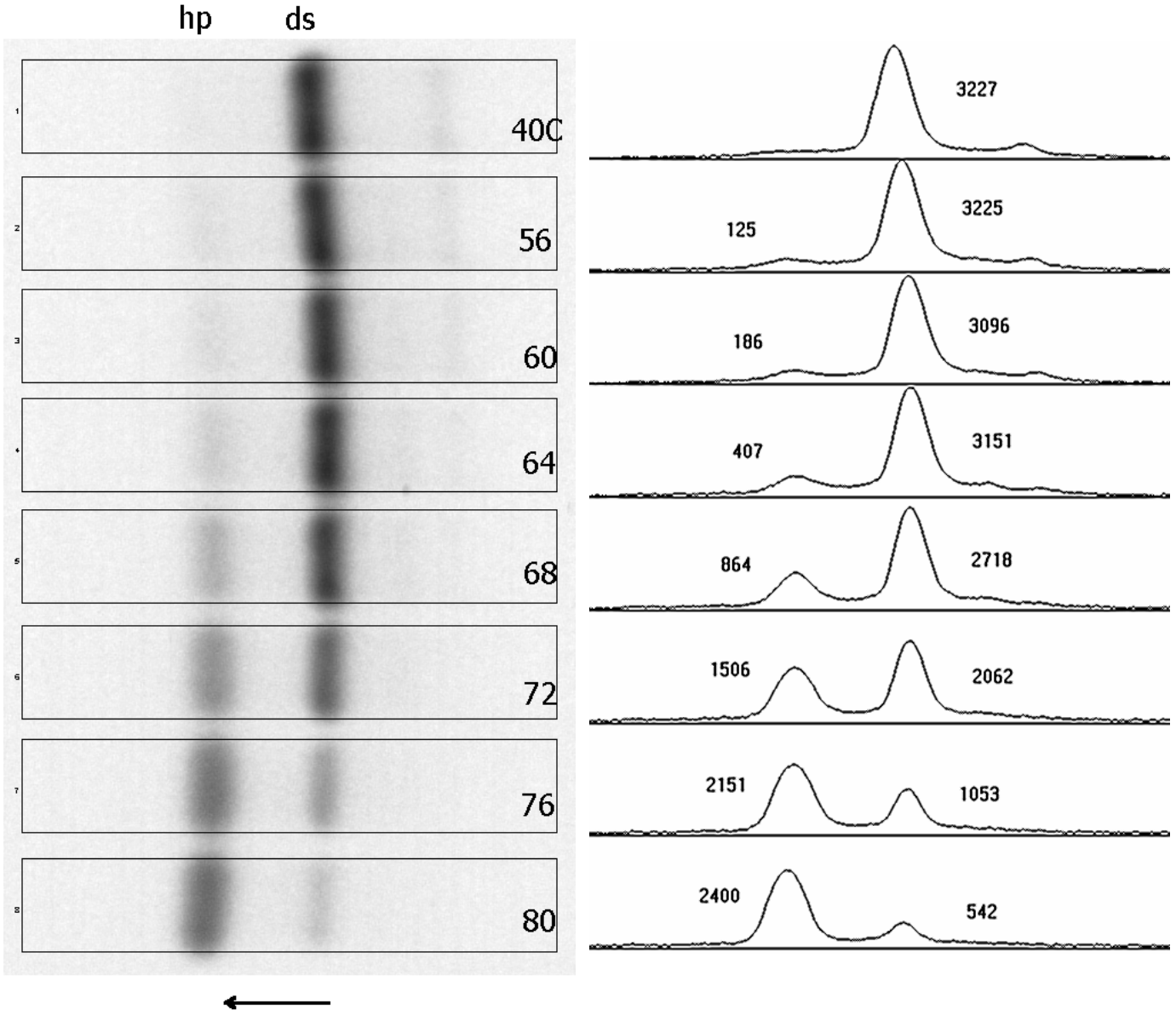

hp
ds
40C
56
60
64
68
72
76
80
3227
125
3225
186
3096
407
3151
864
2718
1506
2062
2151
1053
2400
542

Fig. 2a

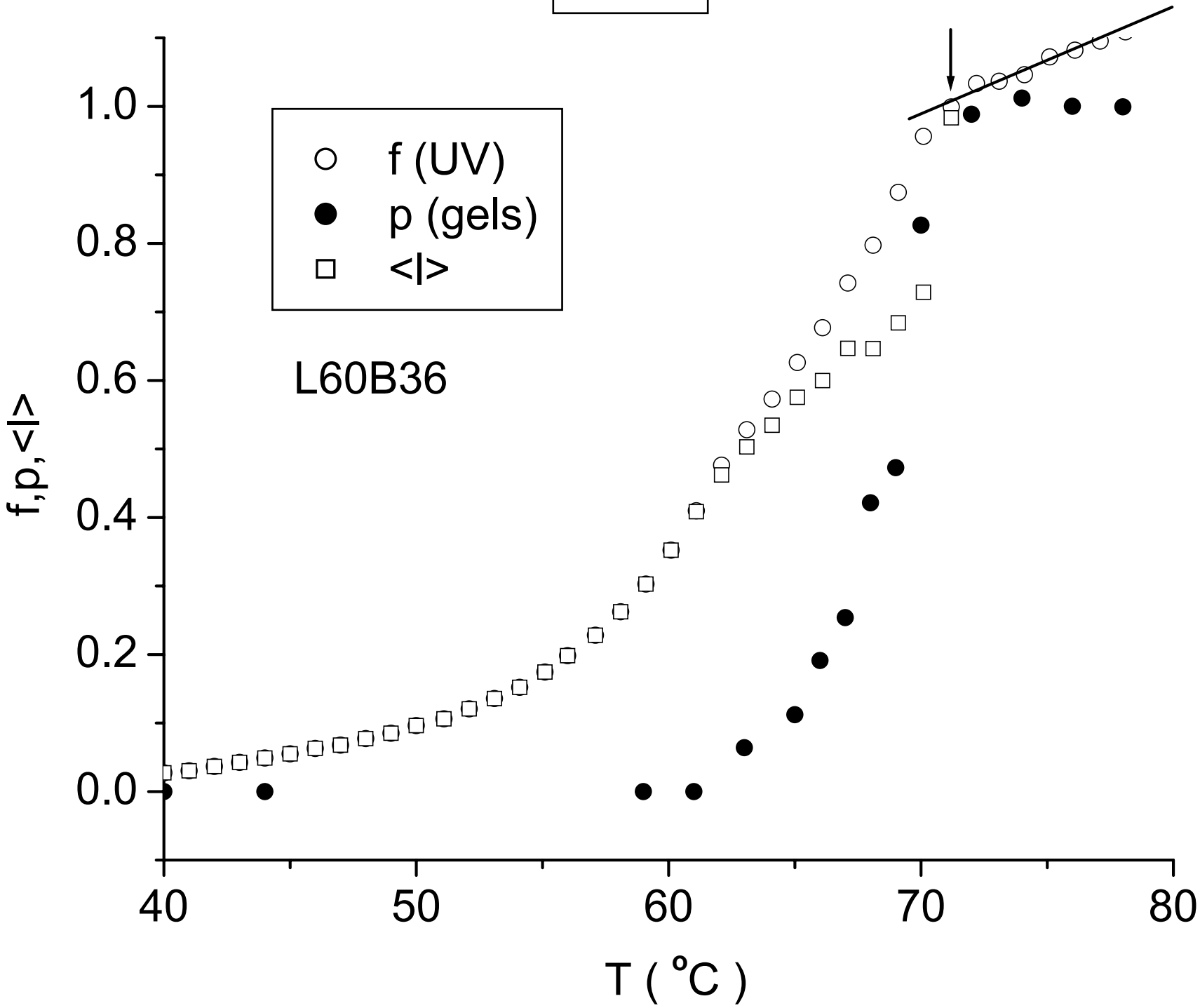


Fig. 2b

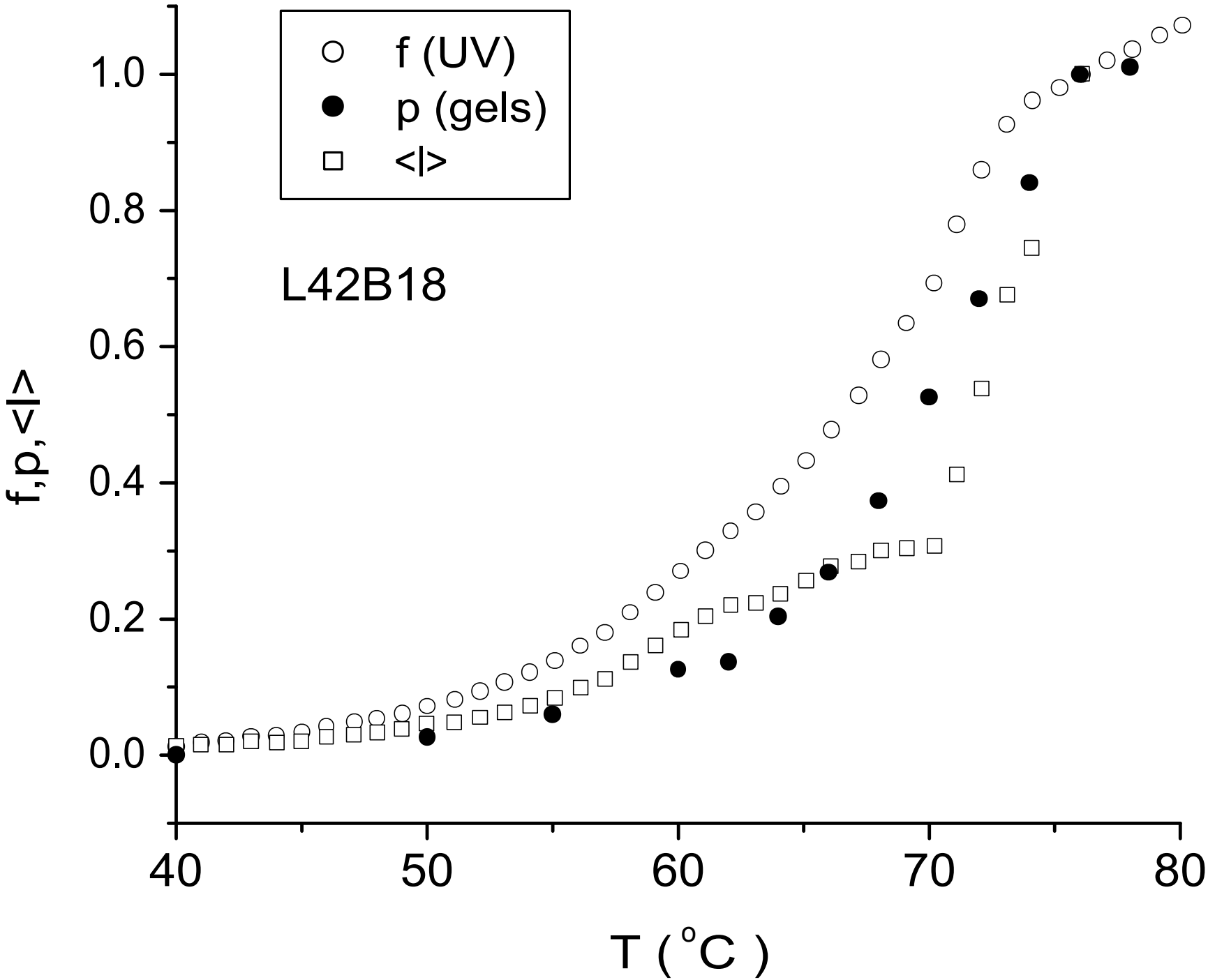


Fig. 2c

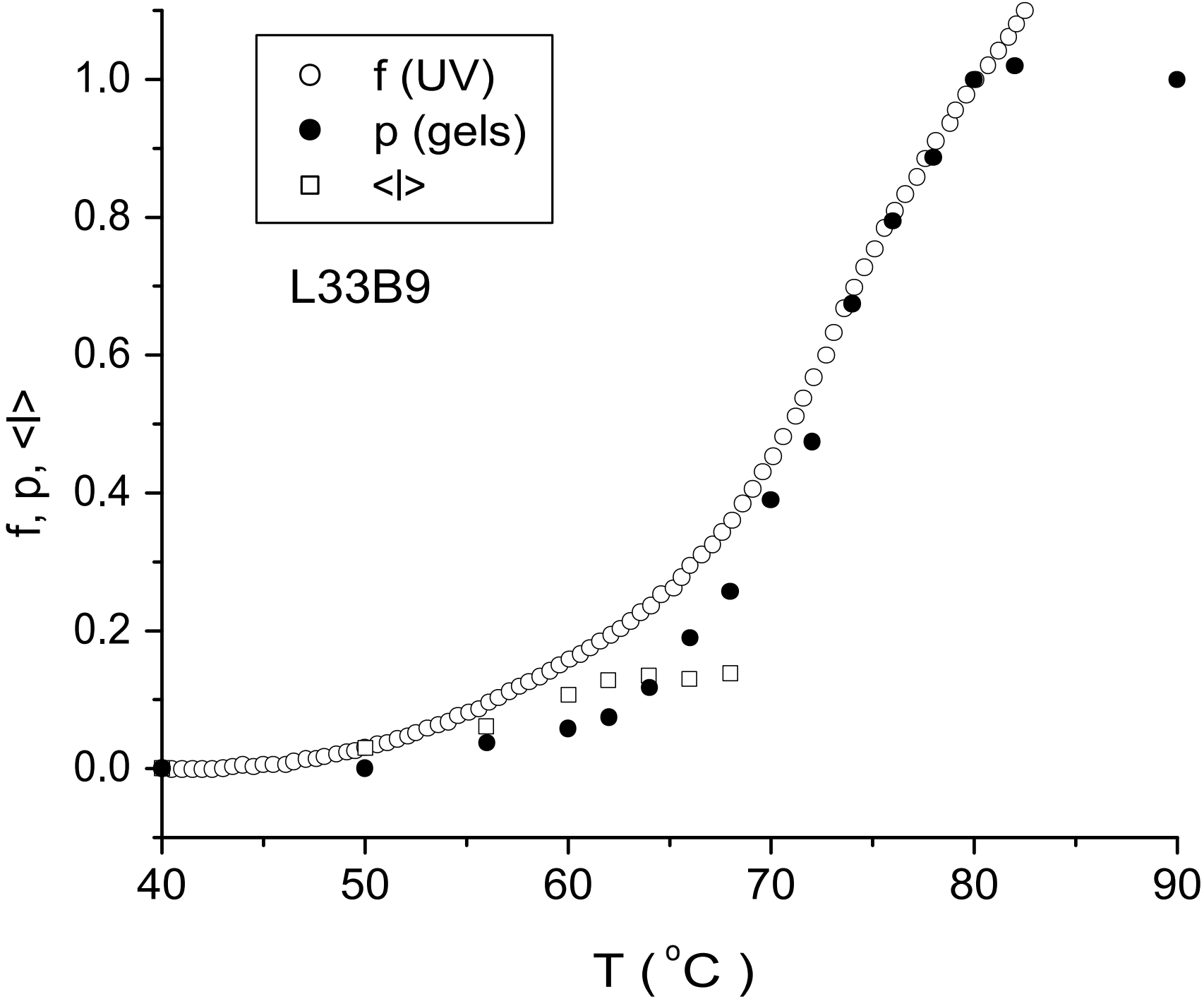


Fig. 3a

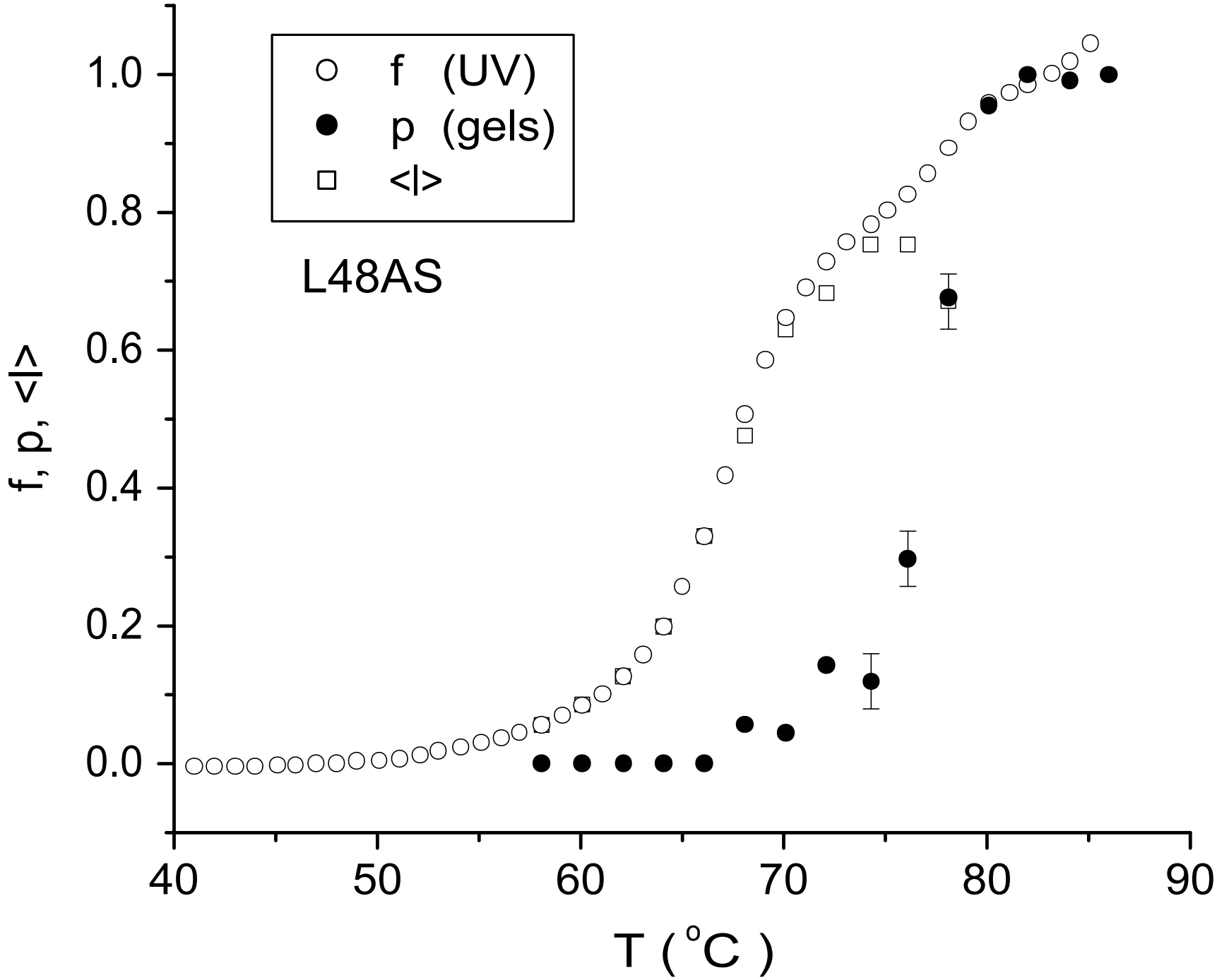


Fig. 3b

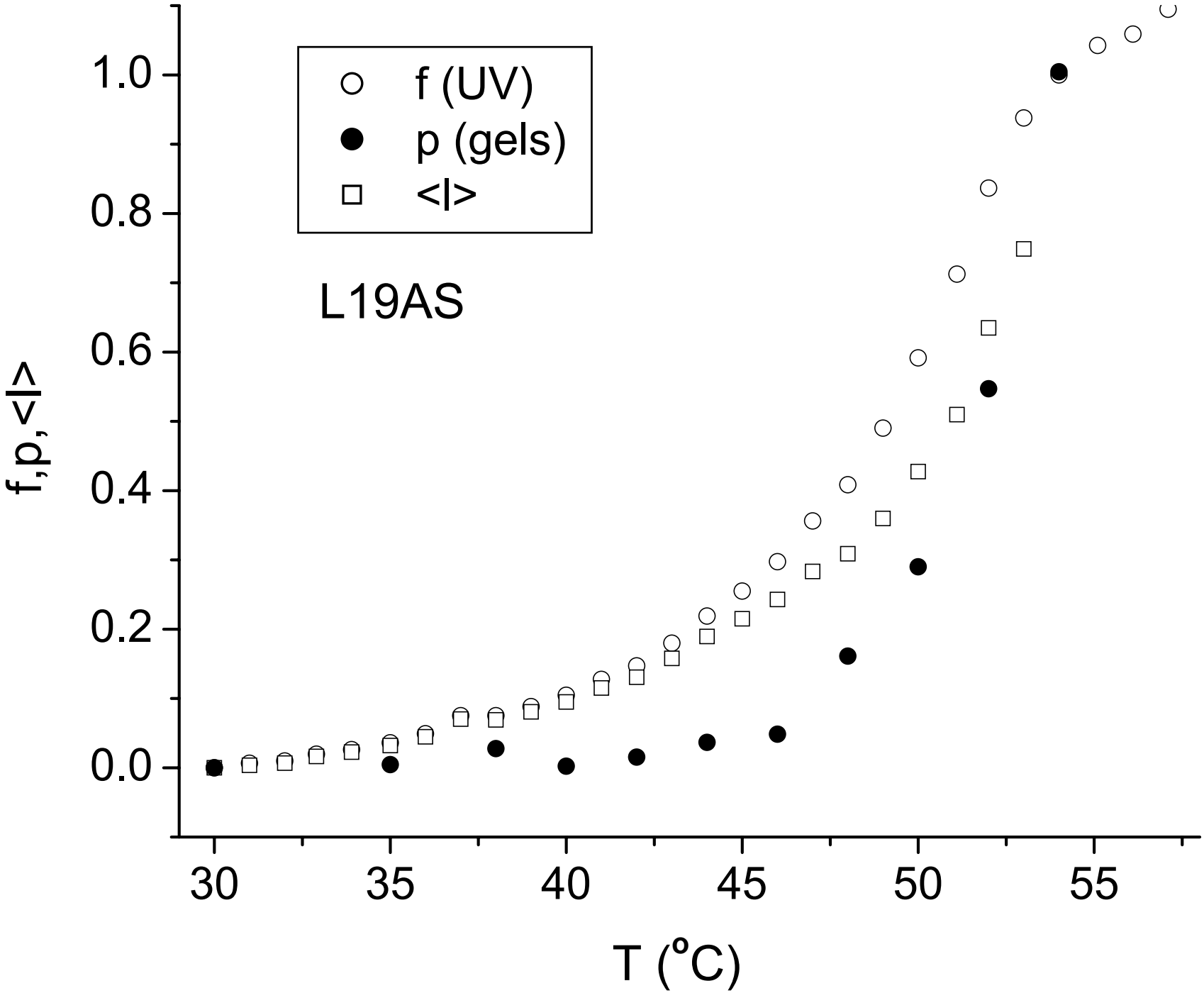


Fig. 4

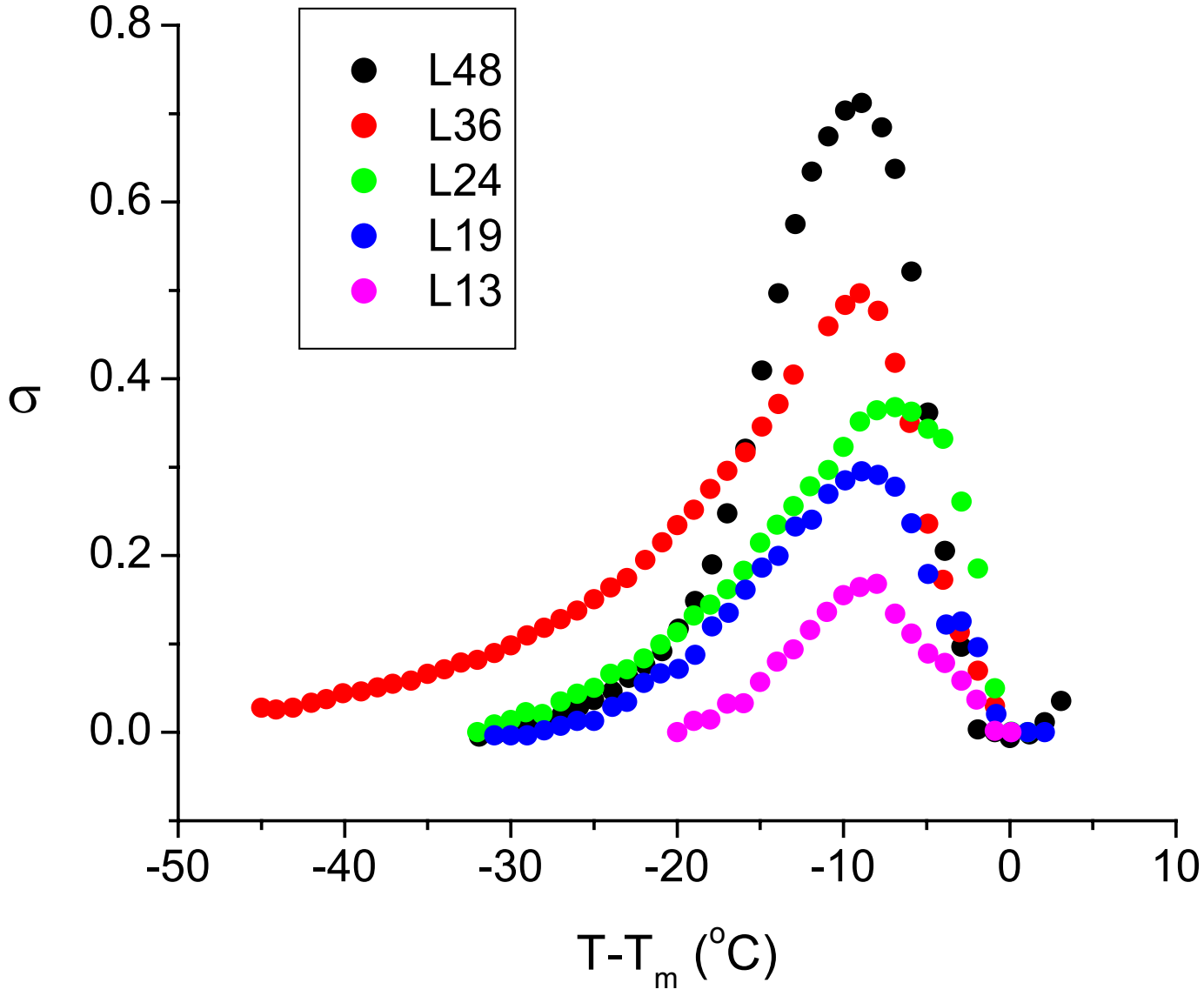


Fig. 5

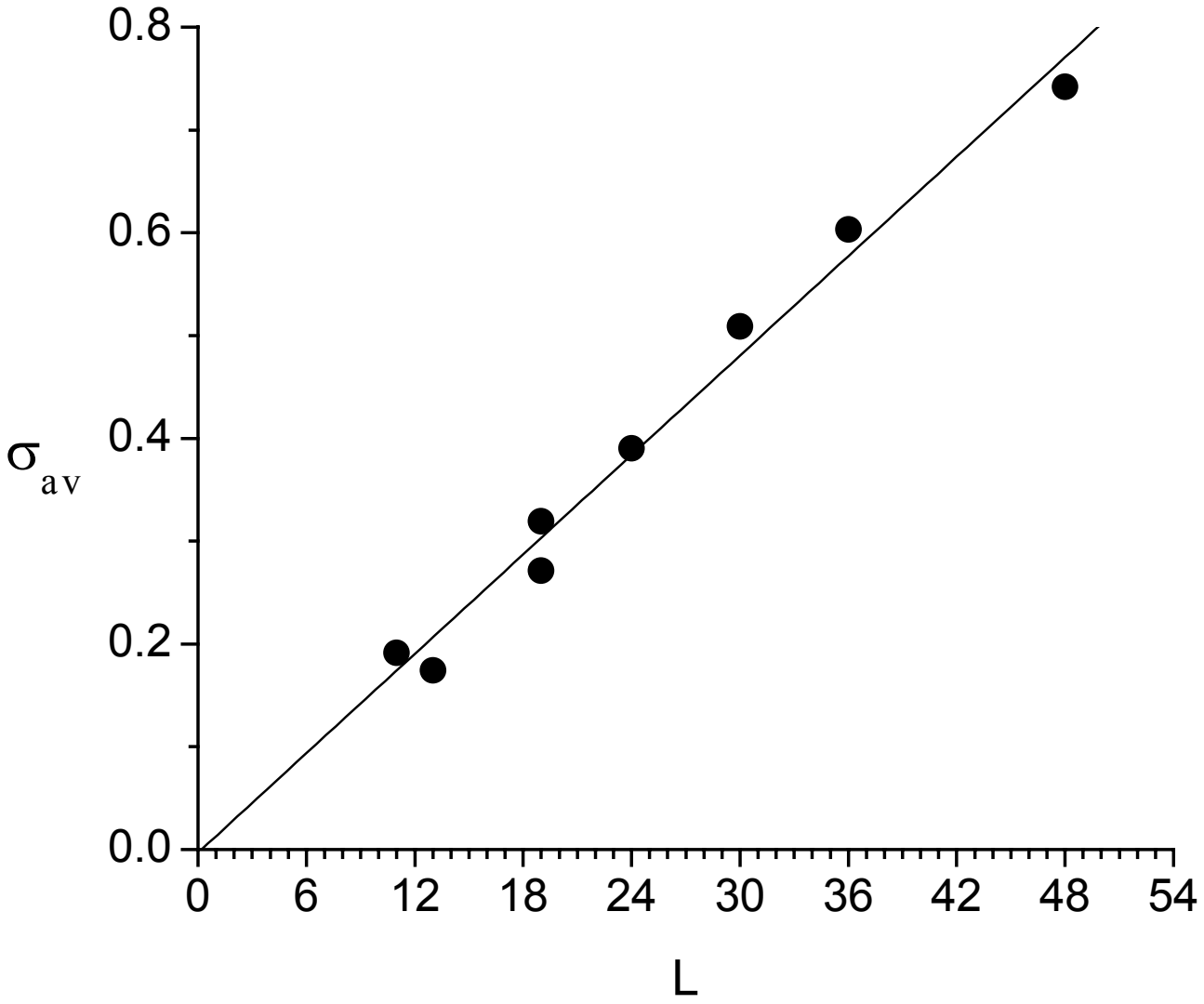


Fig. 6a

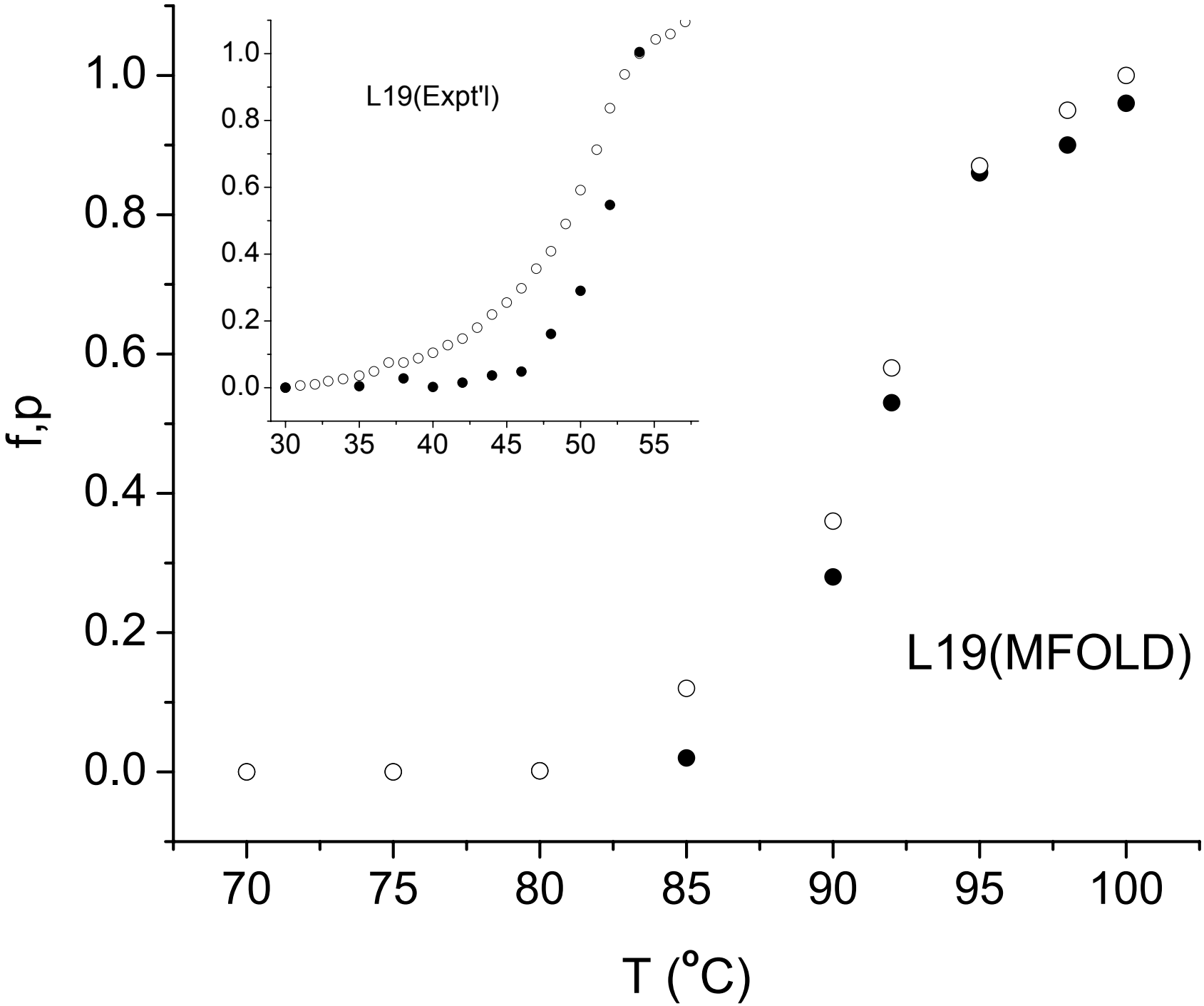

f,p
T (°C)
L19(MFOLD)
L19(Expt'l)
1.0
0.8
0.6
0.4
0.2
0.0
70
75
80
85
90
95
100
30
35
40
45
50
55

Fig. 6b

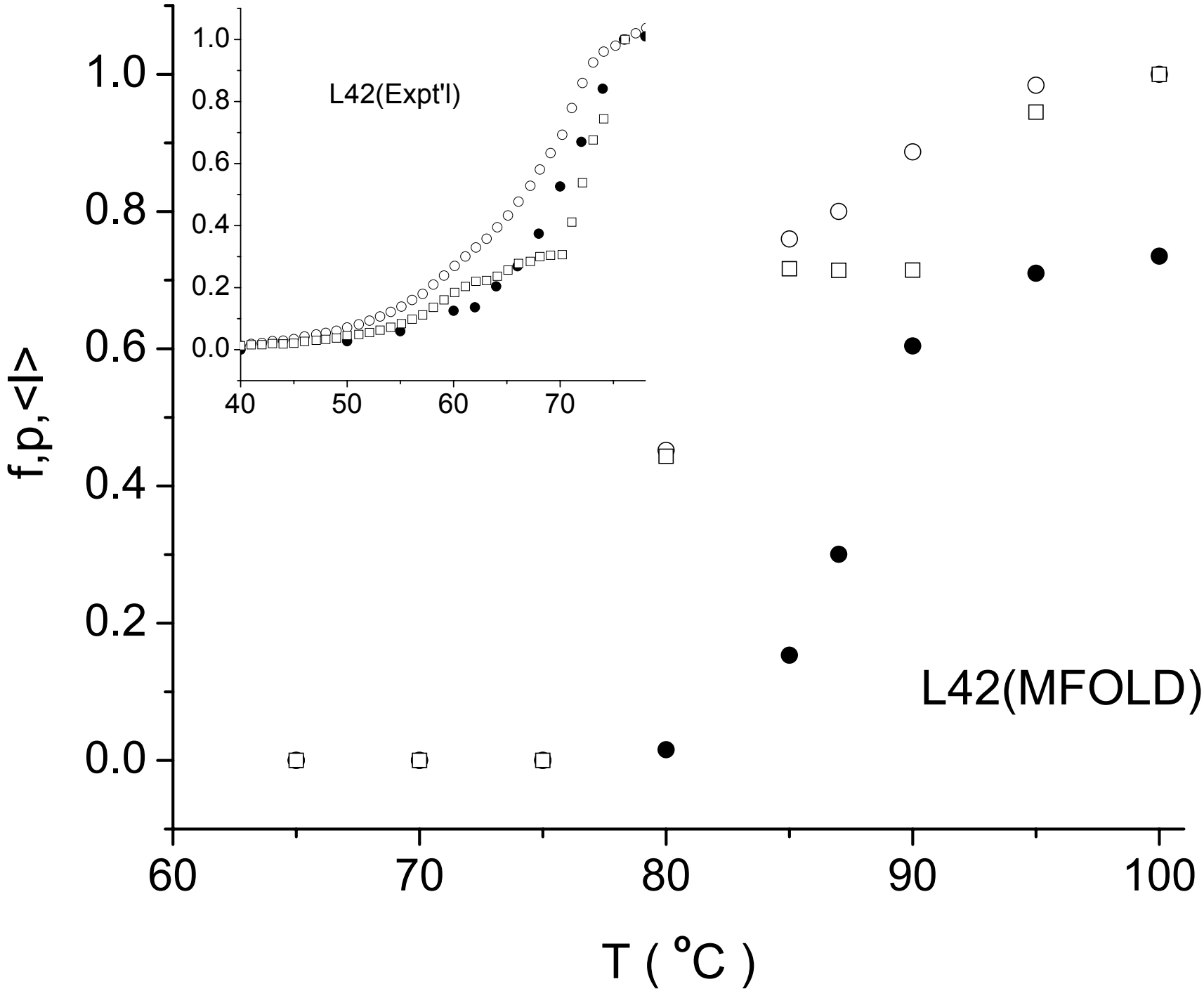

L42(Expt'l)
L42(MFOLD)
f,p,<|>
T (°C)
1.0
0.8
0.6
0.4
0.2
0.0
60
70
80
90
100
40
50
60
70

Fig. 6c