DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD

Using Nanopores to Discriminate Between Single Molecules of DNA

Citation	Branton, Daniel and Amit Meller. 2002. Using nanopores to discriminate between single molecules of DNA. In Structure and Dynamics of Confined Polymers, ed. John J. Kasianowicz, Miklós Kellermayer, and D. W., 177-185. NATO science series, vol. 87. Dordrecht: Kluwer Academic Publishers.
Accessed	February 17, 2015 8:51:01 PM EST
Citable Link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:3109773
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

(Article begins on next page)

Using Nanopores to Discriminate
Between Single Molecules of DNA

Daniel Branton[*†] **and Amit Meller**[*†‡]

[*]Department of Molecular and Cellular Biology

Harvard University, Cambridge, MA 02138, USA

[†]The Rowland Institute for Science, Cambridge, MA 02142, USA

[‡]Center for Advanced Biotechnology, Biomedical Engineering,

Boston University, Boston, MA 02215, USA

Discrimination and characterization of unlabeled, low copy number DNA molecules may become a central requirement for many future biotechnology applications where low cost, high throughput genomic analysis is essential. To date, approaches to such analysis usually require many copies (femto-moles or more) of DNA that are amplified from a specific small region of DNA. In many cases, preparing this DNA is the rate-limiting step that significantly contributes to the overall cost of the analysis. Recently, new tools and techniques that allow detection and manipulation of single DNA molecules have been reported. These tools may eliminate the necessity for DNA amplification. One example consists of sequence-specific DNA detection using molecular beacons which offer a superior signal-to-background efficiency compared with standard DNA probes, and thus are much more suitable for single molecule detection(1). Single DNA molecules have also been sorted and sized using electric fields(2) or stretched by electrophoretic force in a specially micro-fabricated cell(3).

At this meeting we have been introduced to the energetic barriers and dynamics that may explain aspects of polymer translocation through channels in membranes (Muthukumar, Akerman, Lubensky, Parsegian, Sung). Other speakers at this conference (Kasianowicz, Deamer, Akeson), have shown that a biological channel in an insulating membrane separating two ion-containing solutions can in fact be used to detect and characterize single polynucleotide molecules. Meller *et al.*(4) have taken

advantage of these discoveries to show how several different DNA polymers can each be identified by a unique pattern in "event diagrams." These event diagrams are plots of translocation duration versus blockade current for an ensemble of events. Meller *et al.*'s (4) results are an excellent example of how the coincident recording of several independent parameters can provide a unique "finger print" that distinguishes between DNA molecules which differ from each other only by their sequence.

When a polynucleotide molecule is forced to traverse the α-hemolysin channel, it occupies, and thus blocks, much of the otherwise open pore, enabling straightforward detection of the passing molecule. Meller *et al.*(4) characterized each molecule's passage through the nanopore in an α-hemolysin channel as an event whose duration time, t_D, and its averaged normalized blockade current level, I_B, were recorded. Figure 1 displays two typical events labeled by the facing arrow pairs. I_B was calculated by averaging the blockade current during the event and dividing this average by the averaged open pore current. Although the basic apparatus Meller *et al.*(4) used was similar to the horizontal bilayer apparatus described by Akeson *et al.* (5), a special heat-conducting design and the use of a thermoelectric device made it possible to maintain the buffer solutions, bilayer, and channel at any fixed temperature between 0°C and 50°C. The effects of temperature on the movement of DNA polymers through a nanopore were found to be stronger than the

those expected due to frictional drag alone, and, as we shall see below, have pro-

vided several new insights into the translocation process.

At room temperature some polymers, such as poly(dA) and poly(dC), trans-

locate through the α-hemolysin channel at rates that differ from each other mark-

edly.   As a consequence, an event diagram (fig.2), in which each translocating

DNA molecule is characterized by the duration of the blockade it produces, t_D, and

the average blockade current, I_B, show that the events corresponding to the two

polymers each cluster in well-separated regions. Less than 1% of the poly(dA)$_{100}$

events (blue) fall in the poly(dC)$_{100}$ region (red) and vice versa. Thus, discrimina-

tion between the two polymer types is readily achieved. Strikingly, the poly(dA)$_{100}$

events separate into two groups, as do also the poly(dC)$_{100}$ events. The two sepa-

rate groups are evident as two peaks in the current histograms for each polymer

type (fig. 2b). The histograms also show that the current peaks are well fitted by

the sum of two Gaussian curves whose peak values (I_{P1} and I_{P2}) are among the

statistical translocation properties that can be measured for each polymer, e.g. for

poly(dA) $I_{P1} = 0.115$ and $I_{P2} = 0.152$.

Histograms of the translocation durations for groups 1 and 2 also exhibit

clear peak values which are defined as t_{P1} and t_{P2}. For events of short duration

(e.g., $t_D < t_{P1}$) the distributions display Gaussian behavior as shown by the fits

(solid lines in fig. 2c). But for events of long duration (e.g., $t_D > t_{P1}$), the distribu-

tion of t_D values is not Gaussian (fig. 2c) and is most reliably approximated by an exponential with time constant τ_T. As shown in figure 2c, which plots the translocation duration histograms for groups 1, this time constant was found to be much longer for the poly(dA)$_{100}$ events than for the poly(dC)$_{100}$ events. (Likewise, the group 2 values of τ_{T2} were greater for poly(dA) than for poly(dC)). Other polymer types, reported below, as well as many other polymers (unpublished data) exhibit a similar non-Gaussian time distribution of t_D. Because this non-Gaussian time distribution is observed for all the polymers that have been tested, it most likely a reflection of the basic underlying mechanism responsible for polymer translocation rather than of the particular base sequence traversing the nanopore. This non-Gaussian time distribution should be taken into account in future attempts to model polymer translocation through small diameter channels.

Although the separation into 2 groups is not as clear for poly(dC)$_{100}$ as it is for poly(dA)$_{100}$, the tendency of these DNA polymers to fall into two groups begs the obvious question: Why two groups? The same question has been raised with respect to the similar phenomena observed for many RNA polymers(6) , where it was suggested the two groups seen for many polymers could represent translocation of the same structure in either of two orientations (3' to 5' or 5' to 3'). If this were the case, one might expect that all DNA polymers, which contain the same deoxyribophosate backbone would, irrespective of their base composition, give rise

to two groups in event plots such as shown in figure 2. In fact, several polymers, such as poly(dCdT)$_{50}$, were found to produce only one group, suggesting that the grouping phenomena may be a function of the particular purines or pyrimidines that are attached to the backbone sugars.

Using the procedure described above for poly(dA) and poly(dC), we measured the translocation properties of six different polymers. By fitting the translocation duration time and blockade current distribution we obtained the characteristic parameters of each of the polymer shown in Table 1. Together, these ensemble properties can provide a unique "finger-print" that distinguishes between DNA molecules, several of which differ from each other only by their sequence. The I_P, t_P and τ_T values for the group 1 and group 2 events of six different polymers unambiguously characterizes each of the polymer types.

The differences between the translocation behavior of polymers measured at 25°C are accentuate at lower temperature. Using again poly(dA)$_{100}$ and poly(dC)$_{100}$ as an example, examination of representative data at 15.0°C, 25.0°C and 33.0°C (fig 3), make it clear that the two polymers show different trends:

1. The poly(dA) events remain as two separate groups throughout the entire temperature range, but the poly(dC) events that begin to fall into two groups at 20°C (fig. 2) merge into a single widely dispersed group above 25° (fig. 3b and 3c).

2.   The relative number of events in the two poly(dA) groups varies with tem-

perature. At 15°C, nearly 50% of the total number of events are in the second group

while at 40°C this fraction is reduced to only 20-25%.

3.   Particularly for poly(dA), the scattered events in group 2 become even more

dispersed at low temperatures (fig. 3a).

An extensive series of measurements from 15°C to 40°C with five polymer

types showed that for all of the polymers tested, the temperature dependence of t_{P1}

is best approximated by $\sim a/T^2 + b$ (figure 4, full lines) where a is a constant that

depends on the polymer type, T is the temperature in °C and b is an additive con-

stant. Exponential and $\sim T^{-1}$ temperature dependencies failed to fit this data. The

T^{-2} temperature dependence of t_{P1} cannot be accounted for by viscous drag alone

since viscosity would be expected to contribute only a factor of T^{-1}. The strong

temperature dependence of t_{P1} probably arises from a complex of factors affecting

those portions of the polymer that are in the channel and those that are outside of

the channel. We thus would expect weaker temperature dependence for shorter

polymer, especially those that are short enough to reside entirely within the pore.

This is indeed confirmed by a preliminary set of measurements with polymers as

short as 10 bases (unpublished).

Inspection of figure 4 makes it clear that at high temperatures, the differences

between polymers are diminished. For example, the ratio of t_{P1} for poly(dA)$_{100}$ to

t_{Pl} for poly(dC)$_{100}$ (the slowest and fastest polymers in our experiments) decreases with temperature from ~3.2 at 15°C to ~2.1 at 40°C. Further experiments at higher temperatures will be needed to determine if all polymers approach a common value. If so, translocation through a nanopore could be used as a rapid measure of polymer length regardless of the polynucleotide's composition or sequence.

At low temperatures, the differences between polymers are striking. This implies that experiments at 15°C or lower should optimize the identification of individual polymers in a mixed population. For example, the discrimination between poly(dA)$_{100}$ and poly(dC)$_{100}$ at 20°C shown in figure 2c is enhanced at 15°C where there is less overlap in the distribution of t_D values of the two components. Other experiments with poly(dC)$_{100}$ and poly(dCdT)$_{50}$ show that the t_{Pl} values for these polymers differ by 50% (300μsec versus 200μsec) at 15°C even though both polymers contain only pyrimidines. Recent experiments at low temperatures demonstrate that as few as 10 substitutions of thymines spaced evenly throughout a 100 nucleotide poly(dC) polymer are readily detectable (unpublished data).

The effects of temperature on the movement of DNA polymers through a nanopore have begun to provide new insights into the origin of the two event groups observed with many polymers, such as poly(dA)$_{100}$ and poly(dC)$_{100}$ at low temperatures. Originally, the presence of two groups was interpreted as translocation of similarly structured polymers in either of two orientations, 3' to 5' or 5' to

3'. But the measurements of temperature dependent translocation parameter values have now begun to suggest that the presence of two groups may be the result of an equilibrium between polymers that contain stacked structures and polymers that are in an essentially random coil(7-10) . The stacked structure is favored at low temperature, whereas the unstacked random state is favored at high temperature. If, at low temperatures, any existing stacked structure must be broken as the DNA is translocated through the narrow α-hemolysin pore, the added time to disrupt this structure would shift t_P to longer times and broaden the distribution of translocation durations, as measured by τ_T. At the high salt concentrations used for our experiments, the time scales for unstacking the polymer bases is commensurate with the difference between the t_{P1} and t_{P2} values we observed (1).

Meller *et al.*(4) advanced four observations that implied secondary structure and base stacking could be a major explanation for the existence and the statistical properties of group 2 events:

1.   There is a particularly strong temperature dependence of t_{P2} and τ_{T2} in those polymers which contain long poly(dA) sequences. This phenomena is pronounced with poly(dA)$_{100}$, which is known to have a strong tendency for base stacking at low temperatures. In contrast, the translocation duration of poly(dAdC)$_{50}$, which cannot form strong purine-purine base stacking, is approximated by an exponential over the entire temperature range

2.   If entry into the narrow spatial environment of the α-hemolysin channel re-quires that DNA base stacking structure be broken, the energy associated with this process should yield events with a greater temporal scattering. Indeed, τ_{T2}, (which provides a direct measure of temporal dispersion in group 2) for poly(dA)$_{100}$ and poly(dA$_{50}$dC$_{50}$) diverges at low temperatures to much larger values compared with τ_{T1} for poly(dA)$_{100}$ and poly(dA$_{50}$dC$_{50}$) whereas the τ_{T2} values for poly(dAdC)$_{50}$, which cannot form strong purine-purine base stacking, did not show such diver-gence.

3.   If lower temperatures stabilize purine stacking, it is expected that the num-ber of the events associated with structured polymers will grow with decreasing temperature. This is observed. The fraction of the events in group 2 increase from about 20% at 25°C to 45% at 15° for poly(dA)$_{100}$ and poly(dA$_{50}$dC$_{50}$), while re-maining nearly constant for poly(dAdC)$_{50}$.

4.   For both poly(dC$_{50}$dT$_{50}$) and poly(dCdT)$_{50}$, in which only weak pyrimidine stacking is possible(11), the pattern of translocation events do not exhibit two groups, even at 15°C.

Thus, the group 2 events may represent the mostly structured (base-stacked) polymers and the group 1 events the more unstructured, randomly conformed poly-mers. While Meller *et al.*'s (4) data does not exclude specific polymer-pore inter-

actions that could explain the two groups as two states of the channel, such interac-

tions would not readily explain the strong temperature effects.

References

1.   Bonnet, G., Krichevsky, O. & Libchaber, A. (1998) *Proc. Natl. Acad. Sci.* **95,** 8602-06.

2.   Chou, H., Spence, C., Scherer, A. & Quake, S. (1999) *Proc. Natl. Acad. Sci.* **96,** 11-13.

3.   Bakajin, O. B., Duke, T. A. J., Chou, C. F., S., C. S., Austin, R. H. & Cox, E. C. (1998) *Phys. Rev. Lett.* **80,** 2737-2740.

4.   Meller, A., Nivon, L., Brandin, E., Golovchenko, J. & Branton, D. (2000) *Proc. Natl. Acad. Sci.* **in press**.

5.   Akeson, M., Branton, D., Kasianowicz, J. J., Brandin, E. & Deamer, W. D. (1999) *Biophys. J.* **77,** 3227-3233.

6.   Kasianowicz, J. in *this meeting*.

7.   Luzzati, V., Mathis, A., Mason, F. & Witz, J. (1964) *J. Mol. Biol.* **10,** 28-41.

8.   van Holde, K. E., Brahms, J. & Michelson, A. M. (1965) *J. Mol. Biol.* **12,** 726-739.

9.   Holcomb, D. N. & Tinoco, I., Jr. (1965) *Biopolymers* **3,** 121-133.

10.   Dewey, T. G. & Turner, D. H. (1979) *Biochemistry* **18,** 5757-5762.

11.   Solie, T. N. & Schellman, J. A. (1968) *J. Mol. Biol.* **33,** 61-77.

TABLE 1

Polymer	I_{P1}	t_{P1} (μsec)	τ_{T1} (μsec)	I_{P2}	t_{P2} (μsec)	τ_{T2} (μsec)
$(dA)_{100}$	0.126±0.012	192±10	55±3	0.178±0.013	291±20	111±7
$(dC)_{100}$	0.134±0.010	76±4	15±1	0.170±0.013	64±4	10±1
$(dA)_{50}(dC)_{50}$	0.128±0.010	136±7	32±2	0.168±0.014	231±16	176±12
$(dAdC)_{50}$	0.141±0.011	177±9	38±2	0.182±0.011	163±11	41±3
$(dC)_{50}(dT)_{50}$	0.140±0.011	137±7	25±1	No group 2	--	--
$(dCdT)_{50}$	0.144±0.012	82±4	91±5	No group 2	--	--

Summary of the statistical translocation properties of six different polymers characterized at 25.0°C. The standard error of the mean is shown for at least 5 groups of measurements of the same polymer. From the work of Meller *et al.* (4).

Figure Legends

Fig. 1 Definition of the translocation duration time $,t_D,$ and the normalized block-ade level, I_B, for two typical events.

Fig. 2 (a) Event diagram showing translocation duration versus blockade level for poly(dA)$_{100}$ (blue) and poly(dC)$_{100}$ (red) at 20.0°C. Each point on this diagram rep-resent the translocation of a single molecule that was characterized by its translo-cation duration, t_D, and blockade current, I_B. (b) Current histogram projected from the above event diagram, same color codes. The two peaks corresponding to the two groups of events are denoted by I_{P1} and I_{P2}. The solid lines are fits of the data to a sum of two Gaussians. (c) Duration histogram projected from (a) for the first group of events. The solid lines are fits, see text. From the work of Meller *et al.* (4).

Fig. 3 Event diagrams for poly(dA)$_{100}$ (blue) and poly(dC)$_{100}$ (red) at: (a) 15°C, (b) 25°C and (c) 33°C. The insets are the corresponding translocation current (top) and duration time histograms, same color codes. The solid lines are fits similar to those shown in figure 2. From the work of Meller *et al.* (4).

Fig. 4 Dependence of t_P for group 1 events for poly(dA)$_{100}$ (blue), poly(dC)$_{100}$ (red), poly(dA$_{50}$dC$_{50}$) (orange), poly(dAdC)$_{50}$ (green), and poly(dCdT)$_{50}$ (purple). The error bars indicate the standard error of the mean of more than 5 groups of measurements. With rising temperature between 15°C and 40°C there is a 12 fold

decrease of t_{Pl} for the slowest polymer poly(dA), and an 8 fold decrease of t_{Pl} for the fastest poly(dC). The dotted black line that matches closely to the poly(dA$_{50}$dC$_{50}$) data is the algebraic average between t_{Pl} of poly(dA)$_{100}$ and t_{Pl} of poly(dC)$_{100}$. Note that the temperature dependence is not exponential; rather, $\sim T^{-2}$ scaling (solid lines) yielded the best fit to the data. From the work of Meller *et al.* (4).
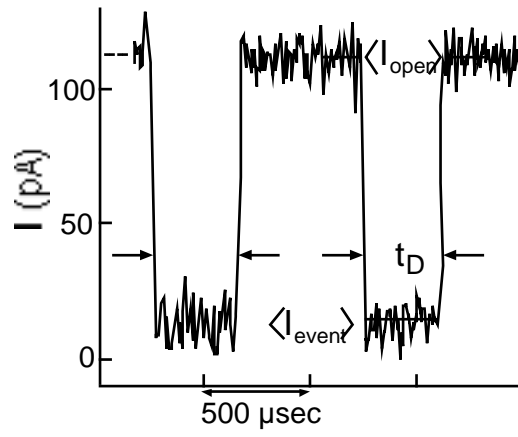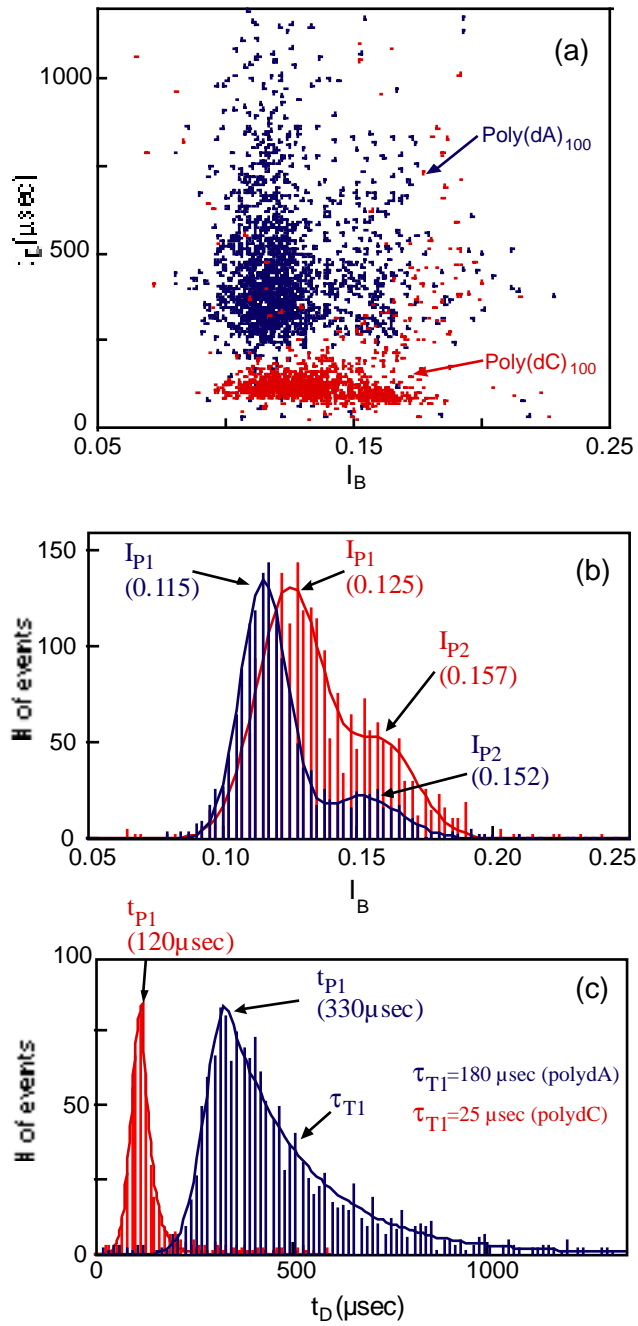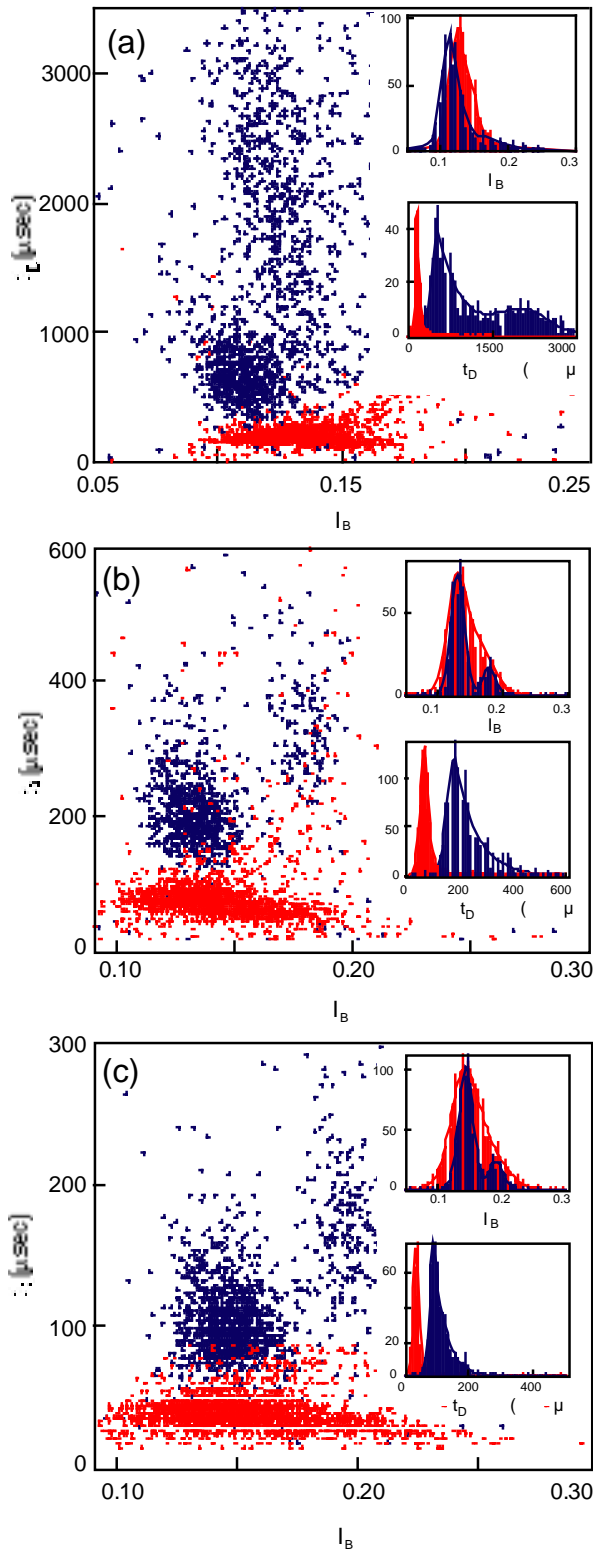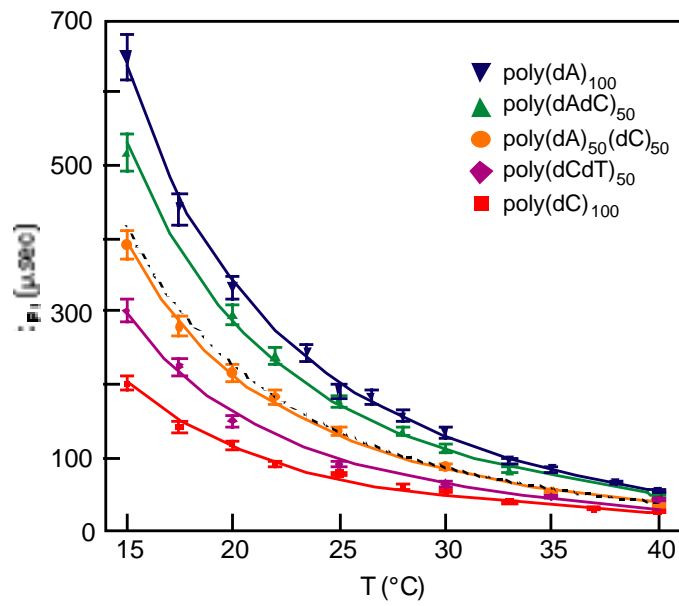
Figure 1

Figure 2

Figure 3

Figure 4