Role of Water for Life

Citation for the original published paper (version of record):
Nordén, B. (2019). Role of Water for Life. Molecular Frontiers Journal, 3(1): 3-19.
http://dx.doi.org/10.1142/S2529732519400017

MOLECULAR FRONTIERS JOURNAL

Role of Water for Life*

Bengt Nordén[1]

The behavior of benzoic acid in polyethylene inspired me to reflect on why water is a unique molecule that all living organisms depend upon. From properties of DNA in aqueous solution a seemingly counter-intuitive conjecture emerges: water is needed for the creation of certain *dry* low-dielectric nm-size environments where hydrogen bonding exerts strong recognition power. Such environments seem to be functionally crucial, and their interactions with other hydrophobic environments, or with hydrophobic agents that modulate the chemical potential of water, can cause structural transformations via 'hydrophobic catalysis'. Possibly combined with an excluded volume osmosis effect (EVO), hydrophobic catalysis may have important biological roles, e.g., in genetic recombination. Hydrophobic agents are found to strongly accelerate spontaneous DNA strand exchange as well as certain other DNA rearrangement reactions. It is hypothesized that hydrophobic catalysis be involved in gene recognition and gene recombination mediated by bacterial RecA (one of the oldest proteins we know of) as well as in sexual recombination in higher organisms, by Rad51. Hydrophobically catalyzed unstacking fluctuations of DNA bases can favor elongated conformations, such as the recently proposed Σ-DNA, with potential regulatory roles. That living cells can survive as dormant spores, with very low water content and in principle as such travel far in space is reflected upon: a random walk model with solar photon pressure as driving force indicates our life on earth could not have originated outside our galaxy but possibly from many solar systems within it — at some place, though, where there was plenty of liquid water.

***Keywords*:** Planet Earth; Water in Biology; Origin of Life; Hydrogen Bond; Hydrophobic Interactions; Nucleic Acids; Stretched DNA.

ROLE OF WATER

Why water is special is something physical chemists have been pondering about very long, and how thermodynamic forces dictate structures and interactions of biological molecules and their complex assemblies. Experimental evidence encouraged me to propose a generalizing conjecture:

> *"a major role of water in living systems is to create certain **dry** low-dielectric nano-environments where hydrogen bonding and ion-pairs can exert strong bonding and recognition power"*

Some evidence for this statement has been around for some time and Langmuir[1,2], Kauzmann[3] and others[4–6] attracted attention to hydrophobic effects behind the folding into tertiary structures of proteins. Chandler made effort to put solvation of non-polar molecules on a solid theoretical footing[7,8], and an exhaustive review of water in cell biology has been given by Philip Ball[9].

But let me start from when I was first inspired addressing in more general terms the role of bulk water in living organisms, whether in plants, bacteria or animals. I had during a symposium for young people in India asked the audience why they thought we need to water the plants and drink so much water ourselves? It was in 2009 at a Molecular Frontiers setting (see www.molecularfrontiers.org) arranged by Professor C.N.R. Rao at his beautiful Jawaharlal Nehru Centre north of Bangalore. Probably influenced by their textbooks or authoritative teachers, high-school students age 15–18 all answered that water was important as a solvent and fluid transporting

[1]Department of Chemistry and Chemical Engineering, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden. Published online 26 February 2019; doi:10.1142/S2529732519400017

* Opening keynote speech at the Second Nordic Organ on a Chip and Micro Physiological Symposium in Oslo, 12 February 2019.

With dedication to Erik W. Thulstrup, a pioneer of stretched sheet polarized spectroscopy.

agent (which is true, in blood for example) and as a chemical agent with redox properties (which is also true, in photosynthesis). But it was another group of children, only 7–13 years old, who came up with the more imaginative suggestions. Apparently inspired by the foam she had seen wash ashore on the beach, organic substances floating up, which wind and waves had driven along the surface to concentrate and aggregate at the water front, one girl said: "I think water makes the life substances come together".

We may speculate if the origin of life, the first cell and so on, may indeed have been created in such an environment where their *insolubility* in water makes hydrophobic substances accumulate naturally on a water surface and there, exposed to sun and air, perform the first crucial chemical reactions. How lipid vesicles may dwell and develop in water, competing with and 'eating' each other, based on hydrophobic interactions, has been demonstrated elegantly by Jack Szostak[10,11].

HYDROPHOBICITY AND HYDROGEN BONDING

But let me start with an experiment I did many years ago, when I was interested in ordering molecules to study their so-called transition moments. Transition moments are microscopic antennas responsible for absorption of light. It was known, since the invention of Polaroids by Edwin Land, that polar dye molecules could be dispersed in films of poly(vinyl alcohol) and become aligned when the film was stretched in the hot air from a hair dryer. If a molecule has an elongated shape it becomes oriented with its long-axis parallel with the stretch direction, and stays so after the film has cooled. I have a film containing the dye methylene blue which I stretched 30 years ago — it still looks deep blue when looked upon through a polaroid oriented parallel with the stretch direction, but uncolored if the polarizer is turned so the light is perpendicularly polarized. This means the methylene blue molecules are still perfectly aligned!

Nonpolar molecules, by contrast, such as benzene or naphthalene which are insoluble in water, readily dissolve in nonpolar polymers such as polyethylene[12,13]. Strangely enough some polar molecules may dissolve in polyethylene too, either by forming internal hydrogen bonds to make them less polar or by forming aggregates that behave in a nonpolar way. This may appear weird since bigger aggregates ought to be generally less soluble than small molecules. But it turns out, as we shall see in other cases too, that *polarity* is much more important than *size* to determine solubility or passage of a molecule through lipid membranes or into macromolecular cavities — at least if water is around!

My discovery more than 40 year ago was that benzoic acid forms a dimer much less polar than the monomer, due to that hydrogen bonds arrange the two monomer molecules in a way so that their permanent dipole moments perfectly cancel each other so that the whole aggregate appears non-polar (Fig. 1). I found that dimer was quickly enriched into a sheet of polyethylene soaked into aqueous benzoic acid solution and was aligned (because of elongated shape) inside the film when this was stretched[14]. This allowed me to determine the directions of the transition moments in benzoic acid from the polarized spectra (the antennas indicated by double-headed arrows in Fig. 1b). Such information is useful in studies

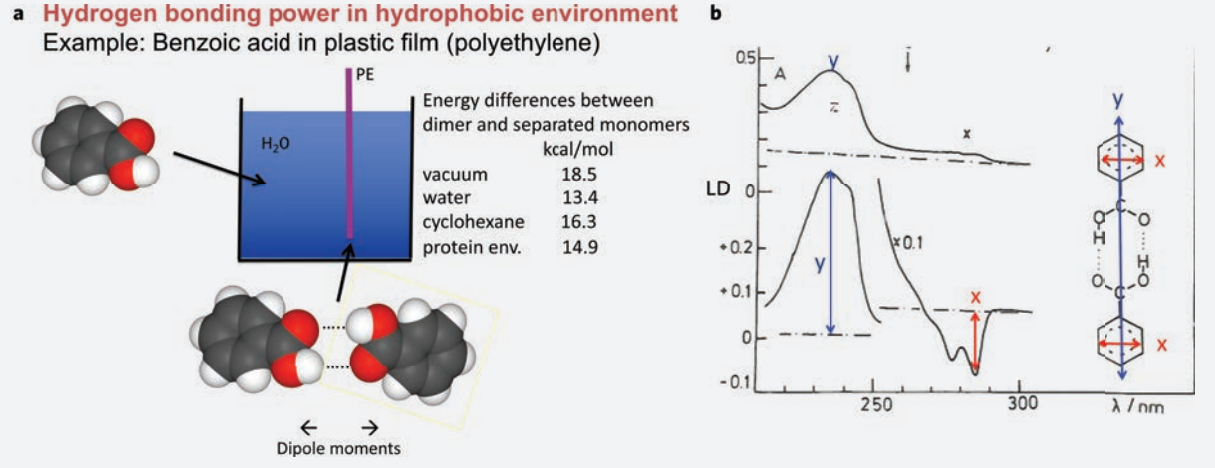


Figure 1. Hydrogen bonding makes polar benzoic acid change into a non-polar dimer (a). Theory shows that dimer is more stable than monomer in vacuum or non-polar solvent than in water. That the dimer is selectively absorbed and oriented in stretched polyethylene film was discovered 42 years ago with polarized light (Linear Dichroism, LD) spectroscopy (b)[14].

using so-called SSLD-MD (Site Selected Linear Dichroism by Molecular Replacement) a technique we developed for determination of 3-D structures of biomacromolecules in water solution, that for some reason were not amenable to structural study by standard techniques like X-ray crystallography or NMR. I shall give some relevant examples of SSLD-MD later.

I was at the time too excited by the spectroscopic opportunities of the oriented benzoic acid dimer, that I did not reflect on what the discovery of the dimer in polyethylene meant. I should have realized that it could be regarded a simple model for a hydrogen-bonded base-pair in DNA and that polyethylene represents the dry non-polar nano-environment that the stack of surrounding nucleobases provides. Fig. 1 also shows computed energy differences (kindly supplied by Tamas Beke-Somfai based on quantum mechanical calculations) between monomeric and dimeric benzoic acid in different environments: obviously the dimer is significantly less stable in water than in a non-polar solvent (cyclohexane) or vacuum. I guess you are now seeing where I am heading.

π-STACKING AND DISPERSION FORCES

Before we continue let me elaborate a little on the transition moments and their role in another context, namely so-called π-stacking. A transition moment (or electric dipole transition moment as its full name goes) is a quantum mechanical integral involving the wave-function of the electronic ground state of the molecule from where an electron is excited, multiplied by the wavefunction of the state where it lands, multiplied by unit vectors ($e_x + e_y + e_z$) for the co-ordinate axes in the molecule. This integral has a directional property (it is a vector) in the molecular frame: for example, in a planar aromatic molecule, like benzene or a DNA base, it is directed somewhere in the molecular plane (xy) if the electronic transition is between a bonding and an antibonding π orbital. It could also be perpendicular to the plane (z) in the rare case we may study a transition between a non-bonding orbital and an antibonding π orbital, but we ignore this for the moment. As mentioned, the transition moments are like antennas for light absorption and this vector property of theirs is exploited when we use polarized light to determine the orientation of molecules in structural applications: when the polarization of light is parallel with the antenna we see maximum absorption while when perpendicular there is no absorption at all. Fig. 1b shows how the two kinds of antennas with x (red) and y (blue) polarizations can be diagnosed spectroscopically on the stretched film.

The transition moments of all possible electronic transitions, spanning from transitions occurring with visible light up to light of very short wavelength, determine another interesting property of a molecule, namely its dynamic electric polarizability. The fact that the transition moments for the π electron transitions are all polarized in the molecular plane makes the polarizability values very high in the plane compared to perpendicular to the plane. This means in turn that if two aromatic molecules are stacked on top of each other, like in a coin pile, the fluctuating dipole field due to motion of electrons in one molecule will induce an oppositely directed dipole moment in the other. The two induced dipole moments will attract each other: this is the dispersion force responsible for why non-polar aromatic molecules want to stack on top of each other, so called π-stacking. If the aromatic molecule (like benzoic acid) also carries a permanent dipole moment, an additional electrostatic interaction will influence how the molecules will be oriented. Also, quadrupole moments, due to that filled π-orbitals enrich negative charge above and below the more positive aromatic plane, give rise to electrostatic interactions, but let's now only consider the dispersion forces which will favor two DNA bases stack on top of each other — in vacuum I should add, because…

THE HYDROPHOBIC FORCE AND A THEORETICAL DILEMMA

…if we include a lot of water another attractive force as strong as the dispersion force will contribute to the π-stacking as well: the hydrophobic force. This is not, in contrast to the dispersion force, a fundamental force that can be easily derived from the quantum mechanics of the electronic states of the two interacting molecules. It is a statistical thermodynamic force corresponding to the gradient of the total Gibbs free energy G of the whole molecular system which thus involves a lot of free as well as somehow bound water molecules, ions etc.

To understand the hydrophobic force contribution to the π-stacking of DNA bases, let us imagine that we force a base to swing out from its pocket within the stack and protrude into the surrounding water solvent. Of course, with the loss of contact with the adjacent surrounding base-pairs we would immediately lose the π-dispersion force, but what about the hydrophobic force? This will be determined by comparing the total free energy, G, of the system for the two situations: the base pocketed inside the base stack and outside in the water, respectively. There are several energetically unfavorable effects that will make the value of G bigger in the latter structure and thus make the base want to return to be embedded in its pocket between other hydrophobic bases. First, the empty slot if filled up with water will increase G because of both enthalpic and entropic reasons: entropic because water will be forced into more ordered structures near the surface of the non-polar base, enthalpic because compared to free water the number of water-water hydrogen bonds will decrease. However, neither the entropic nor the enthalpic contributions to the total free energy are easily determined. For example, the ordered ('ice-like') structure of water may provide a somewhat more favorable enthalpy because of absence of vigorous motion — rotations as well as distance variations, occurring typically in the ps time

regime — at higher temperatures. This effect is related to why water has its maximum density at +4°C. Likewise, the base protruding into the surrounding water layer will also contribute to increased G for similar reasons although the situation may vary from that inside the pocket since the base is free to rotate and may encounter many more water molecules that could engage in bonding to the now free hydrogen donors and acceptors of the base. An additional attractive force between the hydrophobic surfaces, increasing G for an empty gap, is due to solvent fluctuations creating vacuum or water vapor between the surfaces in the nanocavity.

I am probably more pessimistic than many theorists modelling biological macromolecules about the possibilities to very accurately assess the hydrophobic force. What we know reasonably well from experiment is the total attractive force: dispersive + hydrophobic, and that the latter is substantial and particularly important in a system balancing between aggregated and dissociated structures. Very similar situations hold for proteins and for cell membranes (briefly commented on by the end of this paper). A common feature for these systems is rather small changes in free energy ΔG during biological functions. We know that at equilibrium $\Delta G = 0$. Since $\Delta G = \Delta H - T\Delta S$ this means that $\Delta H = T\Delta S$. This is also the condition characteristic of a phase transition, for example, water-ice at 0°C (T = 273 K). A peculiar phenomenon, whose theoretical foundations are not yet fully understood, is so-called entropy-enthalpy compensation (EEC), observed empirically in both kinetic and equilibrium contexts. It appears common in biological systems involving hydrophobic effects and has been interpreted in terms of micro-phase transitions[15,16]. That ΔG is small agrees with the energy-economical conditions that most biological systems live under, reactions close to equilibrium and generally easy to drive in either direction and, therefore, seldomly strongly dissipative (heat producing). This emphasizes the problem of accurately assessing the hydrophobic force that makes the nucleobases stack. ΔG is theoretically the difference between two sums of many, both positive and negative contributions that at the end almost exactly balance each other. Obviously, if each contribution is big, a high precision is required to accurately predict their sums ΔG. We may denote this theoretical challenge as *the hydrophobic energy dilemma.* Despite my somewhat pessimistic view, advances and insights into hydration entropies contributed by Karplus, Warshel and others mentioned by the end of this paper give us some hope for better understanding and theoretical treatment of hydrophobic effects in the future.

THE DOUBLE HELIX

Now, returning to DNA we all know of the paradigm that Crick and Watson launched in 1953[17] which rendered them the Nobel Prize for unvealing the structure of the DNA double helix (Fig. 2). From observations of always identical contents of adenine (A) and thymine (T) they concluded A must be bound to T, and similarly guanine G to cytosine C — in both cases by two hydrogen bonds like benzoic acid dimer. That the exocyclic amino group of guanine could provide a third hydrogen bond, and thus make the G-C base-pair specific had escaped their notice (Fig. 2b). It was later explained by Donohue[18,19]. Crick and Watson believed that the double helix was kept together by the hydrogen bonds between the bases: they missed the fact that a lion part of the stabilization of double-stranded DNA is due the hydrophobic and dispersive π-stacking energies. Still they realized that water was somehow important and that the base organization with the planes of the bases perpendicular to the helix axis, with a separation of 3.4 Å, is particular for the B-form DNA at high humidity, while lowered water content would lead to tilt of the base planes and a different base separation — information they obtained from Rosalind Franklin. She had also observed the famous diffraction cross proving a helical structure, "Photograph 51" reproduced on the stamp celebrating the Nobel Prize to Crick, Watson and Wilkins in 1962 (Fig. 3).

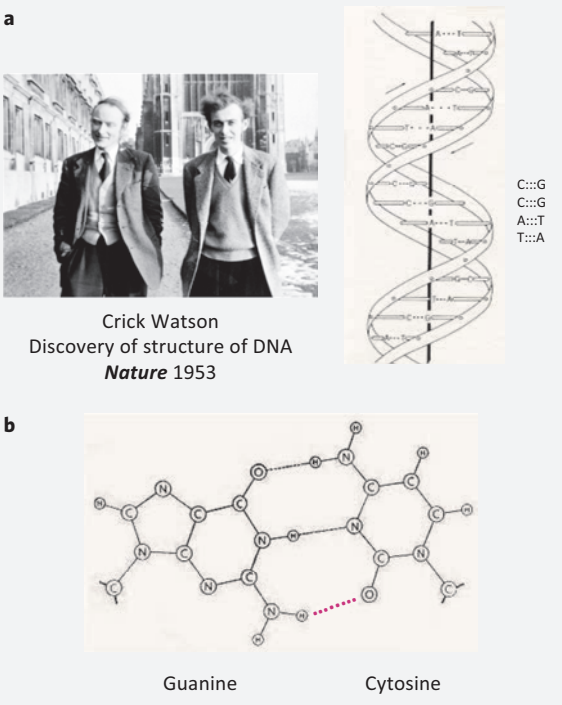


Figure 2. (a) The iconic DNA double helix proposed by Frances Crick and James Watson in 1953, with adenine bound to thymine with two hydrogen bonds and, likewise guanine to cytosine — the third hydrogen bond, explaining specificity, missing though (b). Figures adapted from illustrations in Ref. 17.

Rosalind Franklin

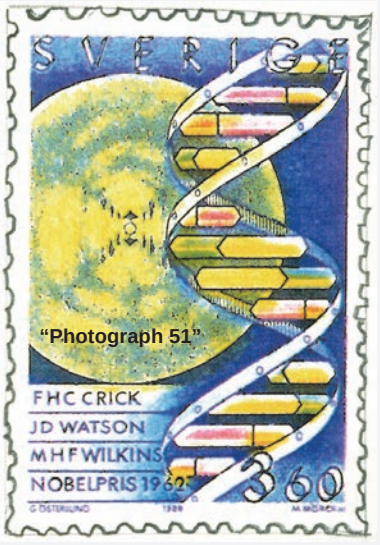


Figure 3. Rosalind Franklin's 'Photograph 51' and the diffraction cross proving DNA is a helix.

ANOTHER CROSS

A cross evidencing a helical structure was also observed when we more than 25 years ago studied DNA with recombinase protein RecA using small-angle neutron scattering (SANS) on flow-oriented solutions of DNA:RecA fibers (Fig. 4), work together with Masayuki Takahashi[20–25]. However, this cross is not due to the DNA helix but to neutrons scattered by a helical aggregate of RecA protein molecules wound around the DNA. Since genetic recombination and the function of bacterial RecA or human Rad51 is central for the water message of this paper, we shall look at those experiments in some detail. Figure 4 shows the experimental SANS pattern for a sample of RecA-DNA fibers which are aligned in a so-called Couette flow cell in the anular gap between a rotating and a static cylinder made of a foil of niobium (a metal transparent to neutrons) – a construction that rendered my technician and co-author Tore Eriksson an Honorary Doctorate at Chalmers. The helical cross is conspicuous in the color picture and this SANS pattern can be modeled theoretically from spherical RecA proteins arranged in a helical fashion as shown in the figure. The SANS anisotropy (elliptical deviation from circular pattern) could be reproduced theoretically, assuming an orientation distribution model for rigid elongated fibers in laminar flow[20].

In parallel with the SANS experiment, another experiment was made with polarized UV light on the same sample under the same flow conditions in another Couette cell made of UV transparent quartz. Measuring the differential absorbance: LD ('linear dichroism') = A(parallel) – A(perpendicular), with light polarized parallel and perpendicular relative to the flow, information about the DNA base orientation could be deduced (Fig. 5). From an LD titration it was concluded that the RecA:DNA complex has a stoichiometry of 1 RecA per 3.6 bas-pairs DNA. The pitch of the RecA helix could be precisely determined from the SANS cross (the subsidiary maxima in Fig. 4) to be 9.5 nm. From this it may be deduced that the DNA is stretched by a factor ×1.50 compared to normal B-form DNA, confirming earlier electron microscopy observations. An even more important conclusion from the modeling of flow orientation distribution is that the value of the so-called orientation parameter, S, may be assessed. The UV flow LD value at 250 nm, a wavelength where the DNA bases absorb light, is related to S and the orientation of transition moments relative to the helix axis of the RecA fiber according to the simple formula:

$$LD(\lambda)/A(\lambda) = (3/2)\, S\, (3\cos^2\theta - 1) \qquad (1)$$

where $LD(\lambda)$ and $A(\lambda)$ are the LD and absorbance values measured for light of wavelength λ, and θ is the angle between the light-absorbing transition moment and the helix axis of the RecA-DNA fibre. Obviously, knowledge of S (from the SANS pattern) allows us determine the angle θ. Somewhat surprisingly at the time, values close to 90° (75°–90°) were found for all wavelengths over the absorption region of the nucleobases. This was surprising because if DNA were strongly stretched, it could no longer be in B form and the only possible base geometry is one with the bases inclined[20–25]. The result that the bases, just like in B DNA, still have their planes perpendicular to the DNA helix axis, therefore, indicated some kind of inhomogeneous arrangement, such as the one denoted Σ DNA in Fig. 6, with short stacks of perpendicular bases surrounded by bigger empty gaps. Energetic considerations indicated a structure with 3 stacked base-pairs as optimal. The alternative, a homogeneous structure for a 50% elongated DNA, called Σ′ in Fig. 6, is inconsistent with the overall negative LD spectrum in Fig. 5 and the determined angle θ near 90°. The Σ′ structure must because of the base tilt have θ near 45° for several transitions which according to the formula above would then have appeared with positive LD and in any case a ratio LD/A that would vary strongly with wavelength. Hence, from the constant negative LD/A the tilted Σ′ structure could be definitely excluded, in favor of Σ. However, as we shall argue, also Σ′ might have a biological role as an elastic energy capacitor for stretched DNA in aqueous environment. The Σ and Σ′ forms were put forward as the only two energetically possible structures after the discovery by Niklas Bosaeus of a 1.52 times elongated stable double-stranded DNA conformation for sequences with high GC content, see below[26]. The later determined structures of RecA

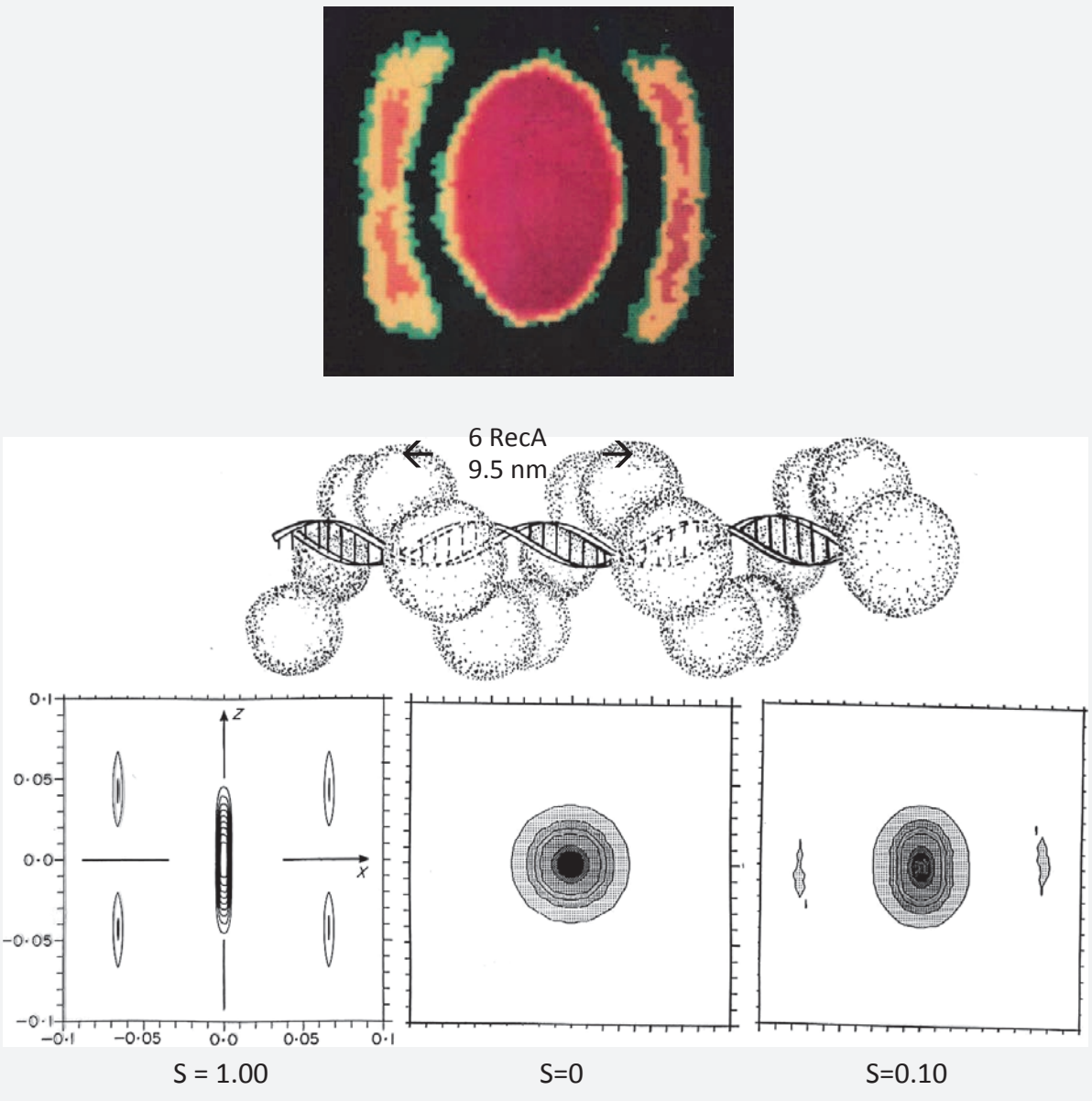


Figure 4. Structure model of RecA-DNA fiber from small-angle neutron scattering (SANS) and linear dichroism (LD) measurements of flow-aligned solutions. Experimental SANS pattern (color, top) shows the helical cross corresponding to an arrangement of RecA proteins as shown in model. Below model patterns for different values of flow orientation parameters (S). Figure adapted from results in Refs. 20–25, 27, 28.

with double as well as single-stranded DNA (Fig. 7) showed DNA structures, similar to Σ, confirming the early conclusion from LD and SANS about perpendicular base orientation.

SSLD-MR METHOD AND 3-D STRUCTURE

The mentioned SSLD-MR technique was applied to a series of RecA mutants each having one aromatic amino acid residue replaced by a 'transparent' (not UV absorbing) residue. In this way the differential LD spectrum between the mutant and the wild-type protein will specifically reflect the orientation of the replaced chromophore. With two non-parallel transition moments, two angles diagnosing the orientation in space of the particular selected residue become accessible[27,28]. Of course, it has to be checked so that the replacement does not significantly alter the protein structure. Also, the S value, if not the same, may have to be adjusted for. Doing some 10 different replacements of aromatic residues

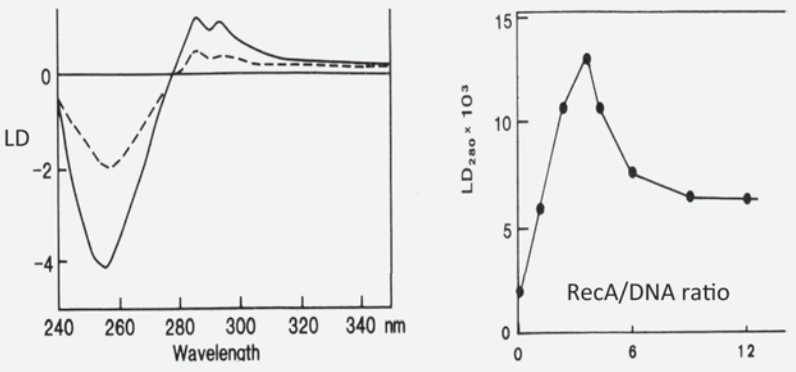


Figure 5. Linear dichroism (LD) of RecA-DNA under identical flow conditions as in Fig. 4. Titration of LD shows a stoichiometry of 3.6 base-pairs DNA per RecA protein unit. From the pitch of the RecA helix in Fig. 4, a DNA rise of 0.51 nm per DNA base is deduced, i.e. that the DNA is elongated by a factor ×1.5. From the amplitude of the negative LD peak at 255 nm and the S value deduced from Fig. 3, an angle θ = 70–90° is calculated from Eq. (1). Figure adapted from results in Refs. 20, 24.

such as various tyrosines and one tryptophan[27], replaced by phenyl alanine and threonine, respectively, enough directional data was obtained so that a 3-D model could be built for RecA complexes with both double- and single-stranded DNA[27,28] also assisted from information from a crystal structure of the more compressed structure of a pure RecA polymer from Steitz' group[29]. A later crystal structure including DNA[30] confirmed on essential points our solution structures.

Later also the protein Rad51 responsible for genetic recombination in higher organisms (including human) was studied using SSLD-MR by Karolin Frykholm and a 3-D model made by Anna Reymer[31]. The structure including a fitted-in DNA helix is shown in Fig. 8. Also shown is an electron microscopy density map into which our structure fits snugly.

In order to be able to discuss the role of water in DNA and genetic recombination, let us briefly review the function of RecA. The main steps consist of assembling RecA protein units to a single-stranded DNA in a

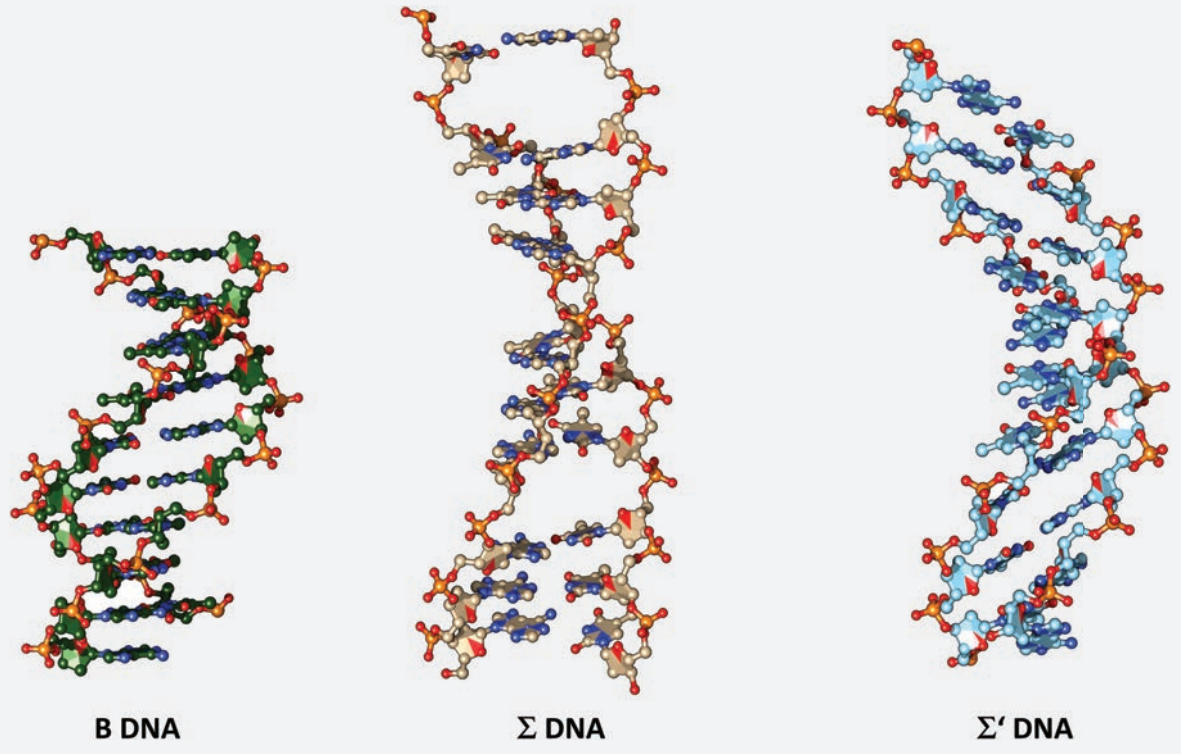


Figure 6. Two energetically plausible models for 50% elongated double-stranded DNA. Right: a homogeneously stretched model with tilted bases (Σ′). Middle: an inhomogeneous model with perpendicular bases (Σ). Left: normal B DNA. From the perpendicular base orientation evidenced from Figs. 4, 5, the Σ′ structure was excluded in favor of Σ. Structures kindly provided by Dr Anna Reymer[26].

Figure 7. Answer is correct! Crystal structure of RecA-DNA (kindly adapted from results in Ref. 30 by Dr Anna Reymer) shows similarity to Σ DNA.

highly cooperative manner to form a helical scaffold. Inside this scaffold the single strand of DNA is supposed to meet and be compared to a double-stranded DNA sequence, both being elongated by about 50% so that short triplet stacks with B like structure are surrounded by gaps that may be empty or partly filled by protein chain. If the base sequence is homologous, or near homologous, a strand replacement occurs. If the sequence is wrong, the RecA-single-strand DNA scaffold does not accept the double-stranded DNA but promotes dissociation. Despite exhaustive studies over several decades, however, the mechanistic details of genetic recombination are still enigmatic and how the recombinase proteins interact with DNA to mediate strand exchange is not understood at an atomistic level.

HYDROPHOBIC CATALYSIS OF STRAND EXCHANGE

Two pieces of information seem to be significant in this context: the observation that the RecA binding to DNA is base-unspecific and a discovery, by Bobo Feng, that spontaneous DNA strand-exchange occurs in solutions containing poly(ethylene glycol) (PEG) via a mechanism we call 'hydrophobic catalysis'[32,33]. Figure 9 shows the principles of Bobo's experiment. A double-stranded sequence of DNA carrying a fluorescent group (F) by the end of one strand and a quenching moiety (T) on the other, is mixed with a single strand with a sequence

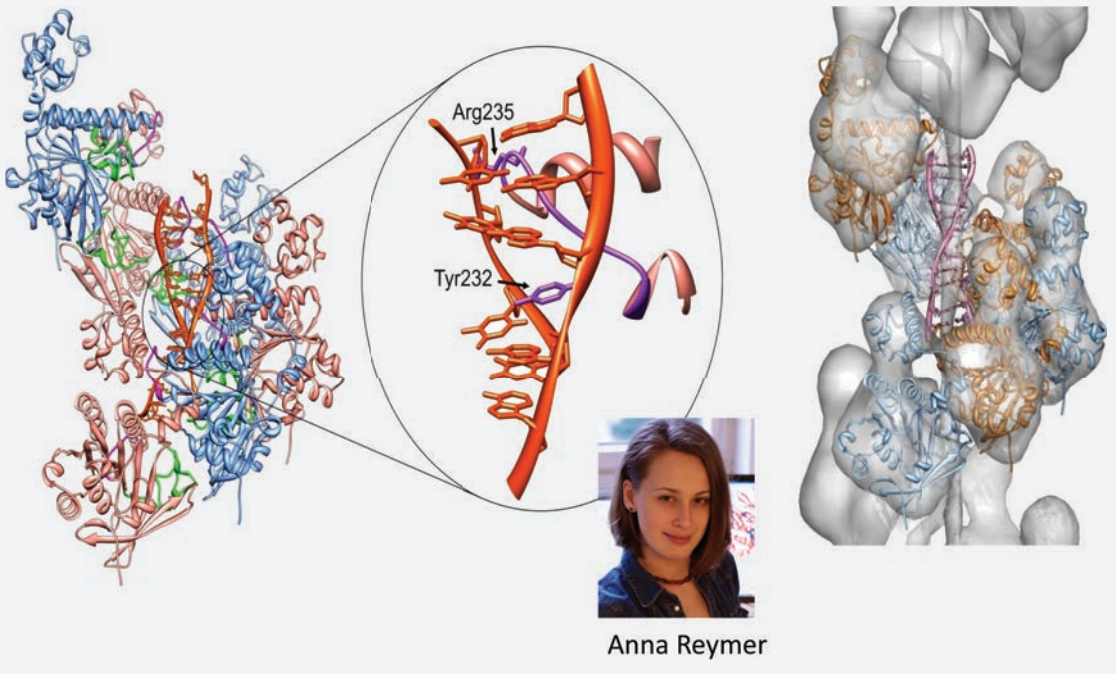
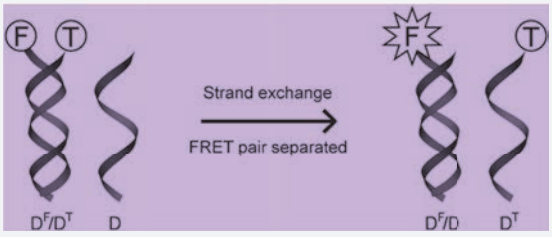


Figure 8. Structure of human Rad51-DNA complex determined from Site Specific Linear Dichroism by Molecular Replacement (SSLD-MR) spectroscopy and molecular modeling[31]. To the right an electron density map from electron microscopy showing agreement of contours with model. Figure adapted from results kindly provided by Dr Anna Reymer.



%PEG w/w	t/s
0	13 000
40	9 600
45	8 600
47	2 200
50	500
52	34

Bobo Feng

Figure 9. Spontaneous DNA strand exchange catalyzed by poly(ethylene glycol) (PEG)[33]. Figure adapted from results in Ref. 33 by Dr Bobo Feng.

identical to the T strand and thus exactly matching the F strand. By monitoring the fluorescence increase one could follow the spontaneous strand exchange. Not unexpectedly it is very slow in absence of catalyst, but speeds up remarkably in presence of PEG. Another interesting observation that Bobo Feng made (B. Feng, Ph.D. Thesis) was that the introduction of a mismatch into the double-stranded DNA construct significantly speeded up the strand exchange rate and the catalytic effect of PEG. The effect is somewhat stronger for mismatches near the end of DNA and one may argue that fraying ends could provide a toehold for an incoming strand. However, the PEG-catalyzed dissociation of a strand having a mismatch in middle regions is also significant, thus possibly paralleling the function of RecA rejecting non-homologous DNA sequences. The mechanism of rejection of non-homologous DNA by RecA is still obscure, whether it is a single matching/non-matching step or whether the mismatched assembly is tested both at a pairing step and before strand exchange and when dissociated, the required dissociation energy could be somehow recruited from the stretched DNA being in a higher energetic state. We shall return to the latter hypothesis in connection with studying stretching of single DNA molecules, where we may speculate if stretch energy elastically stored could provide the required dissociation energy.

Another, remarkable discovery by Bobo Feng is his finding that the threading of bulky complexes through the DNA stack of bases is also catalyzed by the presence of PEG as well as by smaller hydrophobic molecules (Feng *et al.* to be published). In Fig. 10 raw data for the rate of intercalation of the depicted bis-ruthenium complex shows how the presence of PEG accelerates the entering of the complex. The two bulky ruthenium moieties, each carrying hydrophobic phenanthroline wings, for steric reasons require a substantial opening in the DNA base stack to be created. Such pores may be transiently formed by thermal fluctuations, but normally with very low probability so the rate of thread-intercalation at room temperature is measured in hours or days.

A fore-shadowing of the discovery of hydrophobic catalysis was a surprising observation by Per Lincoln and our students of the kinetics of DNA binding and dissociation of certain thread-intercalating bis-ruthenium complexes[34–36]. The thread-intercalation and dissociation rates were found to depend sensitively on the size and nature of the auxiliary ligands of the Ru complex. While the Ru coordination with just two bipyridines as ligand wings had an extremely slow threading rate, the bulkier complex (shown in Fig. 10a) with the bigger phenanthroline passed more easily through the DNA base stack. This strongly counter-intuitive finding (that a bigger object passes more easily through a small hole than a small object does) is in Fig. 10b explained in terms of the hydrophobicity of the phenanthroline wings which may lower the activation barrier to passing through the DNA due to the formation of a transition-state structure in which one base is stacked to a phenanthroline surface (envisaged by Per Lincoln). If this proposed transition state is correct, which a recent crystal structure suggests[37], it would represent a clear case of hydrophobic catalysis.

BREATHING OR CROWDING?

The catalytic power of PEG, in particular at concentrations close to its phase transition of precipitation, is remarkable! One can envisage three possible mechanisms, either of which or combined together, may provide an explanation of the effect of the polymer:

1. PEG may exert an excluded volume osmosis effect (EVO),
2. PEG may act as a denaturant breaking the base-pair hydrogen bonds, or
3. PEG may weaken the hydrophobic force holding the base-pair stack together due to that it locally modulates the chemical potential of water

The first of these effects, EVO, among biologists popularly called ‘crowding’ is an effect first described by Oozawa’s school in the 1950-ies[38]. In an ideal system it does not involve any interactions but is a purely entropic effect due to that an added macromolecule by occupying a volume in the solution will reduce the volume available to solvent and thus increase the solute concentration of smaller molecules and ions. The osmosis part of EVO is due to that the added macromolecule does not have access to interstitial cavities where the solvent and other small molecules or ions may go. This may lead to pressure effects, generally when the added macromolecule is in excess tending to close the cavities by making more compact conformations preferred. The catalytic effect we observe is concluded not to be a typical EVO effect for several reasons.

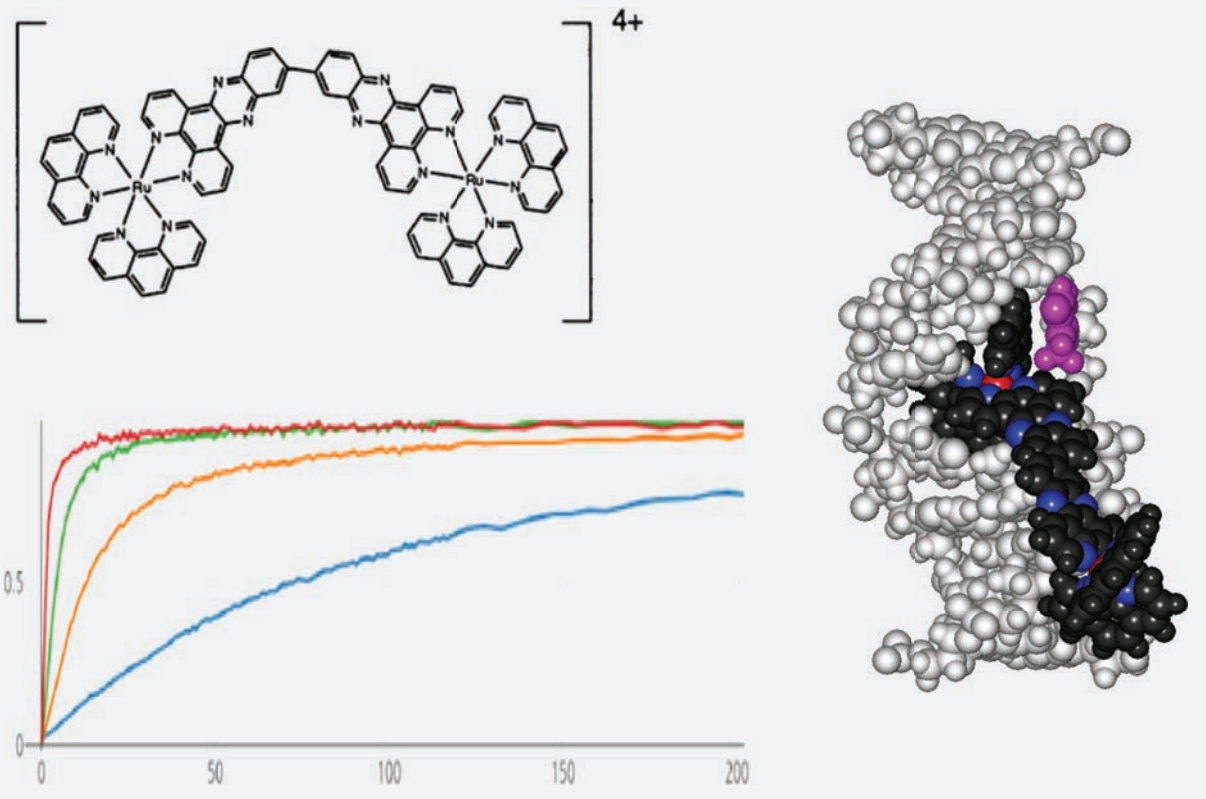


Figure 10. Left: Thread-intercalation of bulky bimetallic complex is catalyzed by PEG (A). Curves showing intercalation rates: 0 (blue), 20 (orange), 30 (green) and 40 % PEG (red). Figure adapted from results kindly provided by Dr Bobo Feng (Feng et al, to be published). Right: Hydrophobic phenanthroline on complex self-catalyzes passage through DNA by forming a transition state by stacking to a DNA base (thymine). Figure kindly provided by Prof Per Lincoln.

PEG, in contrast to e.g. dextran, may adopt amphiphilic (thus partly hydrophobic) conformations and also involves enthalpic effects. Therefore, we tested instead dextran, a typical crowding agent: no effect on the exchange rate was then to be seen. Another argument against EVO is the observation by Fredrik Westerlund and his group that the presence of PEG, studied for DNA in micro-channels, reduces the apparent persistence length in contrast to the increase expected if crowding were significant. A reduction of the persistence length is also observed under unbounded conditions in flow LD experiments, manifested by a decreased effective orientation due to an increased flexibility. Finally, the fact that small molecules, modeling monomers of PEG, exhibit effects very similar to PEG strongly indicates that crowding (EVO) is not an explanation for the catalysis.

BREATHING: BASE-PAIR OPENING OR DE-STACKING OF BASE-PAIRS?

The second mechanism, denaturation, is more difficult to exclude, since a natural opening (breathing) of the double-stranded DNA would lead to faster kinetics of the exchange of DNA sequences, and also explain the faster thread-intercalation. Flow LD, as well as circular dichroism (CD) spectra and experiments showing no significant reaction with glyoxal, however, all indicate that the structure remains as double-stranded B DNA. Still though, a low-probability fluctuating local base-pair opening cannot be excluded and is even probable.

The third option, that PEG reduces the hydrophobic stacking energy because of modulated chemical potential of water, we consider a very likely scenario, and is expected to lead to openings of slots (like in Σ DNA). However, associated strain can also lead to that bases let go of their hydrogen bonded partners and swing out into the now less polar aqueous surrounding. Obviously, we cannot definitely exclude the occurrence of some local denaturation bubbles. What we can exclude, though, is a more significant degree of denaturation, like thermal denaturation, since LD would in such a case drop to zero, and the CD would change significantly. In such a case the DNA strands would not keep in register and would be unable to quickly return to a perfectly native double-stranded state. A crude estimate of the density of transient holes or

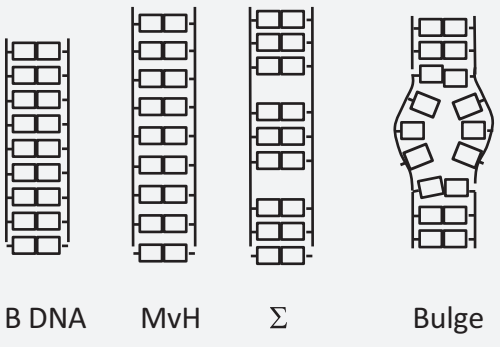


Figure 11. Various hypothetical DNA opening and elongation mechanism models. From left to right: B DNA structure, MvH (McConnel — von Hippel) homogeneous longitudinal elongation, Σ DNA inhomogeneous elongation, bulge base-pair opening. Sketch kindly made by Dr Bobo Feng.

bubbles is obtained by assuming a linear model and that each opening gives rise to a highly flexible hinge ('universal joint') — the Kuhn chain model. The modest reduction of persistence length observed (10%), by this approach suggests a very low frequency of additional breathing, of the order of 1% (1 opening per 100 base-pairs) still however much higher than the natural breathing of B DNA: 0.01%[39,40]. Such an opening scenario could explain the catalytic effect of PEG. Whether this is due to opening of base pairs (natural breathing) or to base-paired de-stacking, however, remains to be determined. Fig. 11 shows schematically the possible ways openings in the DNA may be created. The untwisted uniformly unstacked, but still hydrogen-bonded structure was suggested long ago as a hypothetical possibility by McConnel and von Hippel[41,42]. However, the energy cost for the unstacking of all base-pairs is judged too high and we anticipate that such a stretched structure would quickly re-equilibrate into either the Σ or Σ′ forms in Fig. 6[26]. As proven by flow LD experiments, Σ′ is impossible in the DNA complexes with recombinases. However, this does not exclude it from being a feasible conformation in pure stretched DNA (see below).

DISCOVERY OF STRETCHED DNA CONFORMATION

The remarkable discovery of a well-defined conformation of elongated double-stranded DNA was made by my student Niklas Bosaeus[43,44] using single-molecule laser-tweezers force spectroscopy as developed by Bustamante's school. From the observed DNA extension by 51% and energetic arguments[26], it was concluded to have either of the Σ and Σ′ structures in Fig. 6, the Σ structure possibly related to the stretched DNA in the recombination protein complexes. At a distinct force, corresponding to 1.57 kcal (mol base-pair)$^{-1}$, short (60) as well as longer (122) sequences of synthetic GC-rich DNA, when pulled by their 3′-3′ strands, undergo a sharp transition to a 1.51 times longer Σ (or Σ′) DNA.

Intriguingly, this is within experimental accuracy the same degree of extension of DNA that was also found in the recombinase complexes mentioned earlier. In Fig. 12 the result of a typical single-molecule experiment is shown and how at a certain precisely defined force the DNA molecule shows bistability and oscillates between B form and a 51% longer conformation. The process is amazingly reproducible and allows the population of ground and excited states to be probed with high precision. The high quality of force-distance data and their almost perfect fit to a two-state thermodynamic model demonstrates both accuracy of the experimental setting as well as that the conformational transition is well defined. The displacements and forces may be displayed in an energy diagram (Fig. 13) for the completely reversible process showing how the elastic distortion of the double helix is related to two conformations — a physical property thus inherent to the DNA structure. Since the experiment measures transition energy, the diagram may be regarded as an activation energy landscape. It is found that the result depends on whether the force is applied to opposite strands (3′-3′ or 5′-5′) or to the same strand (3′-5′) and, quite surprisingly, the 3′-3′case differs significantly from 5′-5′, implying a diastereomeric effect — i.e. that the sense of helicity and the local chirality interact with one another[26].

It is interesting to compare the energy of the stretch deformation with energies for unstacking of nucleobases from thermal melting experiments.

Single-molecule stretch force experiments[26]

- 600 pN nm = 1.57 kcal/mol base-pair

Thermal melting experiments[45]

- End-stacking adenine 2.0 kcal/mol
- End-stacking thymine 1.1 kcal/mol
- End-stacking pyrene 3.4 kcal/mol

Obviously these ΔG vales are not very different from each other and also fit well with estimates[26] made for the two energetically feasible Σ conformations in Fig. 6. A plausible scenario is first the occurrence of a fast highly cooperative process leading to the Σ′ structure with tilted bases, then followed by what we have called a 'disproportionation'[26] into the inhomogeneous structure Σ, which is thus strikingly similar to the DNA conformation observed in complex with RecA in a crystal study by Pavletich (Fig. 7)[30].

A connection between the triplet base conformation and the three-letter genetic code has been speculated on based on the expected similarity with RNA and our observed broken symmetry for 3′-3′ and 5′-5′ stretch experiments[26]. The origin of the genetic code is part of the question of origin of

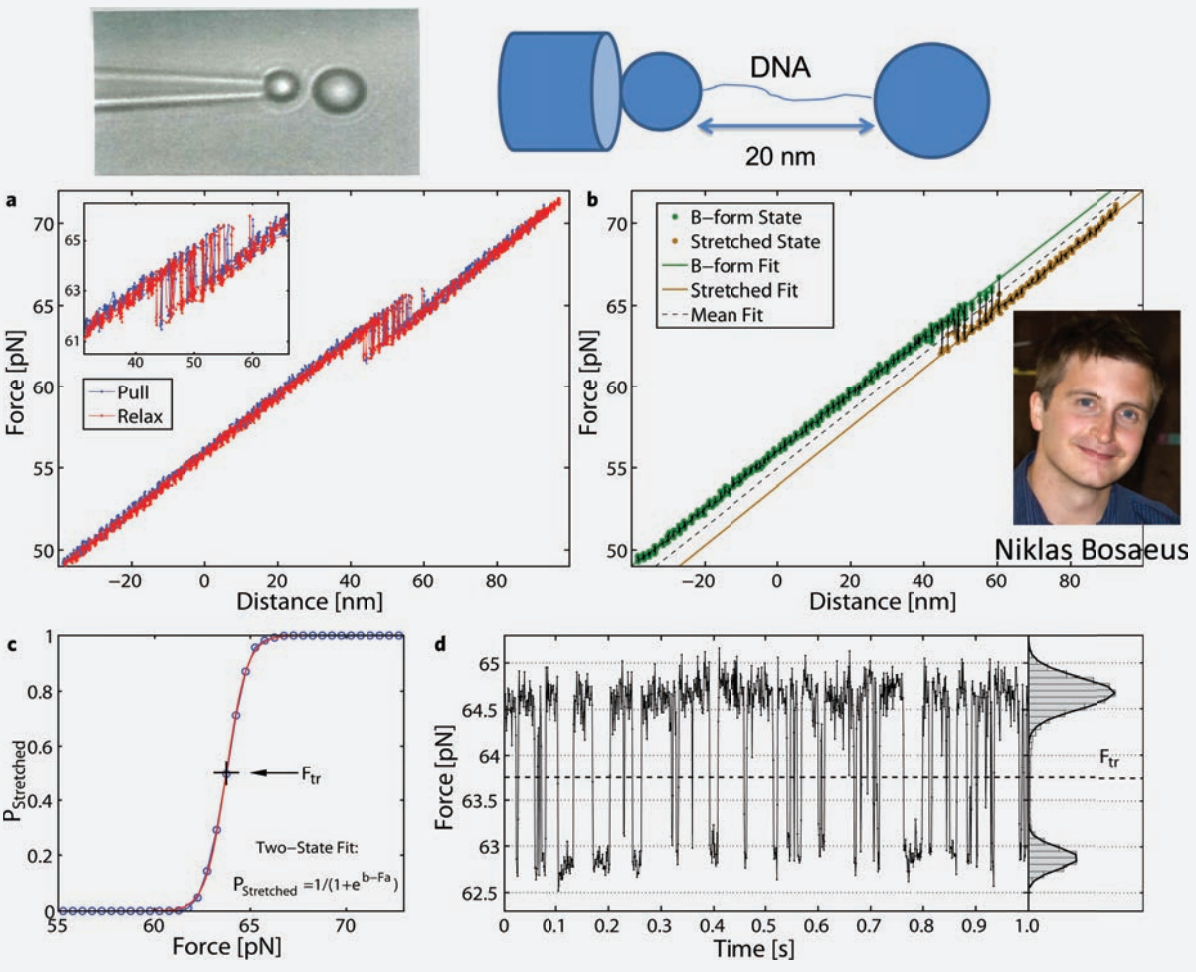


Figure 12. Discovery by Dr Niklas Bosaeus of a 51% elongated DNA conformation, by laser-tweezers stretching. Bistability noticed at 63 pN applied force along the 3'-3' strands. A 60 base-pair long duplex DNA applied between latex beads, one bead in laser trap and one at a pipette. No reaction with glyoxal in stretched state proves that base-pairs remain intact. Figure adapted from results by Bosaeus et al.[26,43,44]

life[46,47], so any progress in this direction could be disruptively important.

LIPID MEMBRANES

Let us finally see how our postulate of water-induced dry environments applies to membrane structure and function. The major components of cell membranes are the phospholipids, amphiphilic molecules which in water spontaneously form a double layer, typically 3–5 nm thick, with the lipids' polar heads facing the water phase and the hydrophobic hydrocarbon tails packing themselves into a more or less fluid 'oily' interior. Immediately we see the parallel with benzoic acid in polyethylene — in fact benzoic acid we have found can penetrate a lipid membrane, as a dimer — in experiments on lipid vesicles, cell-like artificial water-containing bubbles surrounded by a lipid bilayer that people use as models of cells. The lipid membrane is semi-permeable but salts and other ionic or highly polar species are normally more or less efficiently barred from penetration. It has been found that certain peptides, 'cell-penetrating peptides' (CPPs) are easily taken up and transferred through membranes, and may be used in cell-biological and therapeutic contexts to deliver 'cargo' into the cells. Not unlike the benzoic acid dimer, several CPP molecules show ability to fold in such a way that internal hydrogen bonds are formed, for example in arginine, the protonated planar guanidinium ion itself is poorly hydrated, i.e. hydrophobic, and can pass through hydrophobic lipid barrier[48–52]. Likewise, an ion can be carried through the membrane as an ion-pair if it is 'neutralized' by a counterion, preferably one that is partly hydrophobic[52]. Such phenomena, enabling a molecule to transfer through the membrane by reshaping into a nonpolar conformation,

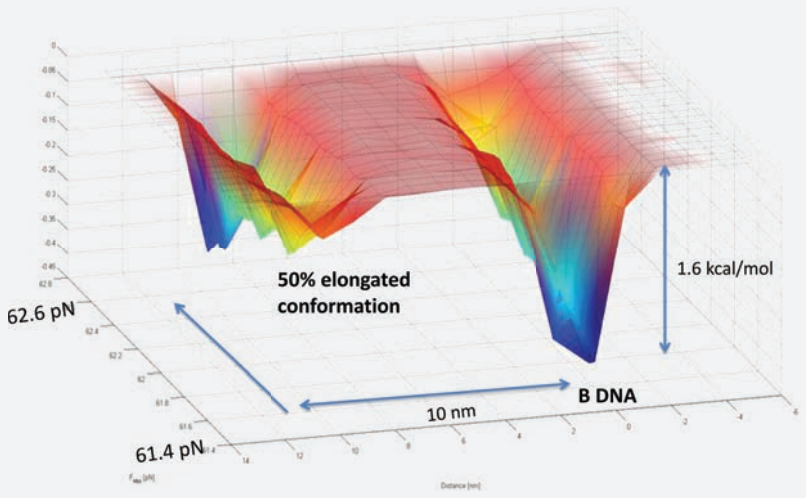
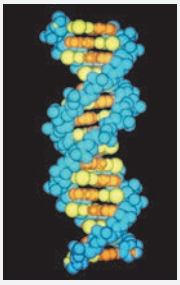
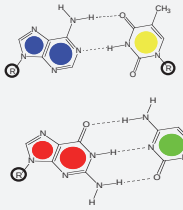


Figure 13. Activation energy landscape adapted from Bosaeus et al.[26]



How DNA recognizes itself



A=T

G≡C

Problem: Recognition works only in "dry" environment
Hydrophobic effect makes DNA bases stack
Bulk water is needed for hydrophobic effect
Conclusion: **water is needed to keep DNA inside dry**

!

Figure 14. Synopsis of role of water in nucleic acids.

can be considered examples of hydrophobic catalysis. This includes also the process of certain peptides able of forming pores through the membrane. We have then of course ignored the main actors of the biological membranes, the proteins, where the membrane provides the precious hydrophobic scene where hydrophobic catalysis may play an important role. Here attention should be drawn to the potential roles of fluctuations, where the membrane can be a collector and mediator of large-amplitude waves which could serve proteins with activation energy for driving chemical processes. A mechanism in such contexts could be the constructive and destructive interference of coherent and incoherent fluctuations, or elastic pressure waves in the lipid membrane. We have postulated that adjacent catalytic centra may collaborate by piggybacking fluctuations (phonons) and by destructive interference this way bring down activation potentials (so called 'Phase Shift Catalysis')[53].

HYDROPHOBIC CATALYSIS — A GENERALIZATION

If we return to the cases where hydrophobic agents may lower the barrier to passage into or through the hydrophobic stack of DNA bases, we are very much in favor of the 'third option': PEG weakening the hydrophobic force holding the base-pair stack together due to that it locally modulates the chemical potential of water. This however, is just a thermodynamic formulation and one would like to see and understand some mechanistic details, i.e. how the hydrophobic agent may stabilize certain open transition states, for example by partially entering the opening in the base stack.

The RecA-DNA complex may be an example where the transition state structure could be 'caught in the act' — for example, in the crystal structure the slot in the base-stack is not empty but a methionine residue appears to be partly inserted[30]. The question whether some hydrophobic aromatic residue might be intercalated to explain the elongation of DNA was already asked long ago and from spectroscopic studies we could early exclude natural aromatic candidates such as cofactor ATP[54], as well as tryptophan and tyrosine from RecA[27,28]. In analogy with the partly amphiphilic partly lipophilic methionine, we may thus envisage that PEG may act as a stabilizer of an unstacked DNA structure such as Σ DNA.

HYDROPHOBIC INTERACTIONS — SOME RECENT THEORETICAL PROGRESS

As indicated, a correct treatment of the statistical thermodynamics of the DNA or any bio-macromolecule requires a global inclusion of all particles in the system and the system could not be too small. In principle, current Molecular Dynamics simulations include all water molecules and ions and should, therefore, give a realistic result. The entropy may be the challenge but recent progress is encouraging[55–59], both generally and regarding base-stacking and other non-covalent interactions in nucleic-acids contexts in particular[60–66].

CONCLUSIONS

We recently demonstrated, using single-molecule force spectroscopy on synthetic DNAs, the existence of a 51% extended stable conformation of double-stranded (base-paired) DNA. The same degree of extension is also found in DNA (single- as well as doube-stranded) complexes with recombinase enzymes, bacterial RecA and human Rad51 — is this a coincidence? As a general conjecture for the role of water in all living systems we have proposed that hydrophobic effects create basis for cohesive forces leading to nano-environments where molecular recognition and precise structure, needed for function, are enhanced. This seems a general rule holding for:

- DNA, RNA
- Proteins
- Membranes

Particularly with DNA, a homogeneously stretched DNA (Σ'), created by high-amplitude thermal fluctuations may undergo 'disproportionation' into the inhomogeneous Σ-form consisting of locally B-like perpendicularly stacked bases. Energy considerations suggest that preferentially three bases stack (3×3.4 Å $= 10$ Å) and that the openings in between are maximally 5 Å. Intriguingly, such a 50% extended structure is found in DNA complexes with recombinase proteins, RecA and Rad51. The triplet base stack structure may ensure improved fidelity of base-pair recognition and promote rejection in case of mismatch during homologous recombination reaction. Triplet is also length of gene codon and we speculate that the structural physics of nucleic acids may have biased the evolution of recombinase proteins to exploit triplet base stacks and also the genetic code.

I started by mentioning hydrogen bonding in non-polar environment and lipid vesicles in context of the origin and evolution of life. In addition to work by Szostak[10,11] also studies by Rebek[67,68] on hydrogen-bonded capsules and by Parikh[69,70] on self-regulating lipid vesicles should be noticed providing important theoretical insight. Related topics are questions about the water state in bacterial spores[71], and whether spores could be somehow shielded to survive cosmic radiation for longer times (or DNA fragments may survive)[72]. Included as Supplementary Information is a slide which I provocatively presented when chairing a session at an Alfred Nobel Symposium in Sånga Säby outside Stockholm in 2005. At the 100th anniversary of Albert Einstein's Ph.D. Thesis on diffusion, I used his formula to estimate from how far away a bacterium may have migrated in space if created outside our planet. If the photon pressure was assumed to be the migrative driving force a true random walk would be reasonable to assume (directed radially from each sun). Further assuming that the first organism was created at a time after the first planets had cooled would give us a 'travel time' of at least some 7 billion years. The result of a (literally) 'back of an envelope calculation' indicated that life could hardly have come from anywhere outside our galaxy, but possibly from any of some billion solar systems within it and, if travel is 'only' a billion years and rate reduced an order of magnitude, the origin could have been still millions of solar systems (and planets). Of course, this does not help us understand *how* life originated, but if including the condition that liquid water has to be present, the number of exo-planet candidates may be reduced in the future when we know more about our cosmic surrounding. It seems reasonable that impacting meteorites could stir up dust particles to reach low gravity altitudes and there be caught by the photon pressure or solar wind (as Arrhenius once suggested) to migrate far from a planet. Also comets, with liquid water on their 'sunny side' ought to be included as potential cradles of life!

OUTLOOK

From benzoic acid in polyethylene to the conjecture that water is needed for life for the creation of *dry* patches where hydrogen bonding exerts strong recognition power, one may further hypothesize how water, and 'modulated chemical potential' of water in the presence of hydrophobic (or amphiphilic) agents may be exploited in a Darwinianly evolved biochemistry. Possibly combined with an excluded volume osmosis effect (EVO), hydrophobic catalysis may have important biological roles, e.g., in genetic recombination. Hydrophobic agents are found to strongly accelerate spontaneous DNA strand exchange as well as certain other DNA rearrangement reactions. Thus, it is speculated that hydrophobic catalysis may be involved in gene recognition and gene recombination mediated by bacterial RecA (one of the oldest proteins we know of) as well as in sexual recombination in higher organisms, by Rad51. Hydrophobically catalyzed unstacking fluctuations of DNA bases can favor elongated conformations, such as the recently proposed Σ-DNA, with potential regulatory roles. Large-amplitude thermal fluctuations can be anticipated to occur anywhere where loops of DNA are free to move and pick up and transfer angular and linear momentum (the so called whip-lash effect). Similar dynamic effects may explain collective phenomena in membranes where also phase-shift and hydrophobic catalysis may kick in.

REFERENCES

1. Langmuir, I. & Waugh, D.F. The adsorption of proteins at oil-water interfaces and artificial protein-lipoid membranes. *J. Gen. Physiol.* **21**, 745–755 (1938).
2. Langmuir, I. Overturning and anchoring of monolayers. *Science* **87**, 493–500 (1938).
3. Kauzmann, W. Some forces in the interpretation of protein denaturation. *Adv. Prot. Chem.* **14**, 1–63 (1959).

4. Stillinger, F.H. Structure in aqueous solutions of nonpolar solutes from the standpoint of scaled-particle theory. *J. Solution Chem.* **2**, 141 (1973).
5. Parsegian, V.A., Rand, R.P. & Rau, D.C. Osmotic stress, crowding, preferential hydration and binding: a comparison of perspectives. *Proc. Natl. Acad. Sci. USA* **97**, 3987–3992 (2000).
6. Dill, K.A. & MacCallum, J.L. The protein folding problem, 50 years on. *Science* **338**, 1042–1046 (2012). DOI: 10.1126/science.1219021
7. Chandler, D. Interfaces and the driving force of hydrophobic assembly. *Nature* **437**, 640–647 (2005). DOI: 10.1038/nature04162.PMID 16193038.
8. Lum, K., Chandler, D. & Weeks, J.D. Hydrophobicity at small and large length scales. *J. Phys. Chem. B.* **103**, 4570–4577 (1999).
9. Ball, P. Water as an active constituent in cell biology. *Chem. Rev.* **108**, 74–108 (2008).
10. Matsushita-Ishiodori, Y., Hanczyc, M.M., Wang, A., Szostak, J.W. & Yomo, T. Using imaging flow cytometry to quantify and optimize giant vesicle production by water-in-oil emulsion transfer methods. *Langmuir* **35**, 2375–2382 (2019). DOI: 10.1021/acs.langmuir.8b03635 Publication Date (Web): January 15, 2019
11. Adamala, K. & Szostak, J.W. Competition between model protocells driven by an encapsulated catalyst. *Nat. Chem.* **5**, 495–501 (2013). DOI: 10.1038/nchem.1650
12. Thulstrup, E.W., Michl, J. & Eggers, J.H. Polarization spectra in stretched polymer sheets. II. Separation of pi->pi* absorption of symmetrical molecules into components. *J. Phys. Chem.* **74**, 3868–3878 (1970).
13. Nordén, B. Applications of linear dichroism spectroscopy. *Appl. Spectrosc. Rev.* **14**, 157–248 (1978).
14. Nordén, B. General aspects on linear dichroism spectroscopy and its application. Proceedings of the Nobel Workshop in Lund, October 25–27, 1976 on Molecular Optical Dichroism and Chemical Applications of Polarized Spectroscopy. *Spectrosc. Lett.* **10**, 381–400 (1977).
15. Starikow, E.B. & Nordén, B. Enthalpy-entropy compensation: a phantom or something useful? *J. Phys. Chem.* **111**, 14431–14435 (2007).
16. Starikov, E.B. & Nordén, B. Entropy-enthalpy compensation as a fundamental concept and analysis tool for systematical experimental data. *Chem. Phys. Lett.* **538**, 118–120 (2007).
17. Watson, J.D. & Crick, F.H.C. Molecular structure of nucleic acid. *Nature* **171**, 737–738 (1953).
18. Donohue, J. The third bond. *Proc. Natl. Acad. Sci. USA* **42**, 60–65 (1956).
19. Wain-Hobson, S. The third Bond. *Nature* **439**, 539 (2006).
20. Nordén, B., Elvingson, C., Kubista, M., Sjöberg, B., Ryberg, H., Ryberg, M., Mortensen, K. & Takahashi, M. Structure of RecA-DNA complexes studied by combination of linear dichroism and small angle neutron scattering measurements in flow-oriented samples. *J. Mol. Biol.* **226**, 1175–1192 (1992).
21. Takahashi, M., Kubista, M. & Nordén, B. Linear dichroism study of RecA-DNA complexes: structural evidence and binding stoichiometries. *J. Biol. Chem.* **292**, 8109–8111 (1987).
22. Takahashi, M., Kubista, M. & Nordén, B. Binding stoichiometry and structure of RecA-DNA complexes studied by flow linear dichroism and fluorescence spectroscopy: evidence for multiple heterogeneous-DNA coordination. *J. Mol. Biol.* **205**, 137–147 (1989).
23. Nordén, B., Elvingson, C., Eriksson, T., Kubista, M., Sjöberg, B., Takahashi, M. & Mortensen, K. Structure of a RecA-DNA complex from linear dichroism and small-angle neutron-scattering in flow-oriented solution. *J. Mol. Biol.* **216**, 223–228 (1990).
24. Takahashi, M. & Nordén, B. Structure of RecA-DNA complexes and mechanism of DNA strand exchange reaction in homologous recombination. *Adv. Biophys.* **30**, 1–35 (1994).
25. Nordén, B., Wittung-Stafshede, P., Ellouze, C., Kim, H.-K., Mortensen, K. & Takahashi, M. Base orientation of second DNA in RecA·DNA fi laments: analysis by combination of linear dichroism and small angle neutron scattering in flow-oriented solution. *J. Biol. Chem.* **273**, 15682–15686 (1998).
26. Bosaeus, N., Reymer, A., Beke-Somfai, T., Brown, T., Takahashi, M. & Wittung-Stafshede, P. A stretched conformation of DNA with a biological role? *Quart. Rev. Biophys.* **50**, e11 (2017).
27. Hagmar, P., Norden, B., Baty, D., Chartier, M., Takahashi, M., Nordén, B., Baty, D., Chartier, M. & Takahashi, M. Structure of DNA-RecA complexes studied by residue differential linear dichroism and fluorescence spectroscopy for a genetically engineered RecA protein. *Journal of Molecular Biology* **226**, 1193–1205 (1992).
28. Morimatsu, K., Takahashi, M. & Nordén, B. Arrangement of RecA protein in its active filament determined by polarized-light spectroscopy. *Proc. Natl. Acad. Sci. USA* **99**, 11688–11693 (2002).
29. Story, R.M.M., Weber, I.T.T. & Steitz, T.A.T.A.A. The structure of the *E. coli* RecA protein monomer and polymer. *Nature* **355**, 318–325 (1992).
30. Chen, Z., Yang, H. & Pavletich, N.P. Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures. *Nature* **453**, 489–494 (2008).
31. Reymer, A., Frykholm, K., Morimatsu, K., Takahashi, M. & Nordén, B. Structure of human Rad51 protein filament from molecular modeling and site-specific linear dichroism spectroscopy. *Proc. Natl. Acad. Sci.* **106**, 13248–13253 (2009).

32. Feng, B., Frykholm, K., Nordén, B. & Westerlund, F. DNA strand exchange catalyzed by molecular crowding in PEG solutions. *Chem. Commun.* **46**, 8231–8233 (2010).
33. Feng, B., Westerlund, F. & Nordén, B. Evidence for hydrophobic catalysis of DNA strand exchange. *Chem. Comm.* **34**, 7390–7392 (2015).
34. Westerlund, F., Nordell P., Norden B. & Lincoln, P. Kinetic characterization of an extremely slow DNA binding equilibrium. *J. Phys. Chem.* **111**, 9132–9137 (2007).
35. Nordell, P., Westerlund, F., Wilhelmsson, L.M., Nordén, B. & Lincoln, P. Kinetic recognition of ATRich DNA by ruthenium complexes. *Angew. Chem. Int.* **46**, 2203–2206 (2007).
36. Nordell, P., Westerlund, F., Reymer, A., El-Sagheer, A.H., Brown, T., Norden, B. & Lincoln, P. DNA polymorphism as an origin of adenine-thymine tract length-dependent threading intercalation rate. *J. Am. Chem. Soc.* **130**, 14651–14658.
37. Boer, D.R., Wu, L., Lincoln, P. & Coll, M. Thread insertion of a bis(dipyridophenazine) diruthenium complex into the DNA double helix by the extrusion of AT base pairs and cross-linking of DNA duplexes. *Angew. Chem.* **126**, 1980–1983 (2014).
38. Asakura, Sho & Oosawa, F. Interaction between particles suspended in solutions of macromolecules. *J. Polym. Sci.* **33**, 183–192 (1958).
39. Frank-Kamenetskii, M.D. & Prakash, S. Fluctuations in the DNA double helix: a critical review. *Phys. Life Rev.* **11**, 153–170 (2014). DOI: 10.1016/j.plrev.2014.01.005. Epub 2014.
40. Phelps, C., Lee, W., Jose, D., von Hippel, P.H. & Marcus, A.H. Single molecule FRET and linear dichroism studies of DNA breathing and helicase binding at replication fork junctions. *Proc. Natl. Acad. Sci. USA* **110**, 17320–17325 (2013).
41. Von Hippel, P.H., Johnson, N.P. & Marcus, A.H. 50 years of DNA 'breathing': reflections on old and new approaches. *Biopolymers* **99**, 923–954 (2013).
42. McConnel, B. & von Hippel, P.H. Hydrogen exchange as a probe of the dynamic structure of DNA. I. General acid-base catalysis. *J. Mol. Biol.* **14**, 297–316 (1970).
43. Bosaeus, N., El-Sagheer, A.H., Brown, T., Smith, S.B., Akerman, B., Bustamante, C. & Norden, B. Tension induces a base-paired overstretched DNA conformation. *Proc. Natl. Acad. Sci. USA* **109**, 15179–15184 (2012).
44. Bosaeus, N., El-Sagheer, A.H., Brown, T., Åkerman, B. & Nordén, B. Force-induced melting of DNA — evidence for peeling and internal melting from force spectra on short synthetic duplex sequences. *Nucleic Acids Res.* **42**, 8083–8091 (2014).
45. Kool, E.T.T. Hydrogen bonding, base stacking, and steric effects in DNA replication. *Annu. Rev. Biophys. Biomol. Struct.* **30**, 1–22 (2001).
46. Koonin, E.V.V. & Novozhilov, A.S.S. Origin and evolution of the genetic code: the universal enigma. *IUBMB Life* **61**, 99–111 (2009).
47. Lagerkvist, U. 'Two out of three': an alternative method for codon reading. *Proc. Natl. Acad. Sci. USA* **75**, 1759–1762 (1978).
48. Thorén, P.E.G., Persson, D., Karlsson, M. & Nordén, B. The antennapedia penetratin peptide translocates across lipid bilayers: the first direct observation. *FEBS Lett.* **482**, 265–268 (2000).
49. Mason, P.E., Neilson, G.W., Dempsey, C.E., Barnes, A.C. & Cruickshank, J.M. The hydration structure of guanidinium and thiocyanate ions: implications for protein stability in aqueous solution. *Proc. Natl. Acad. Sci. USA* **100**, 4557–4561 (2003).
50. Rydberg, H., Matson, M., Åmand, H. & Nordén, B. Effects of Tryptophan content and backbone spacing on the uptake efficiency of cell-penetrating peptides. *Biochemistry* **51**, 5531–5539 (2012).
51. Ardhammar, M., Lincoln, P. & Nordén, B. Invisible liposomes: refractive index matching with sucrose enables flow dichroism assessment of peptide orientation in lipid vesicle membranes. *Proc. Natl. Acad. Sci. USA* **99**, 15313–15317 (2002).
52. Esbjörner, E.K., Lincoln, P. & Nordén, B. Counterion-mediated membrane penetration: cationic cell-penetrating peptides overcome Born energy barrier by ion-pairing with phospholipids. *Biochim. Biophys. Acta* **1768**, 1550–1558 (2007).
53. Beke-Somfai, T., Feng, B. & Nordén, B. Energy phase shift as mechanism for catalysis. *Chem. Phys. Lett.* **535**, 169–172 (2012).
54. Takahashi, M. & Nordén, B. The cofactor ATP in DNA-RecA complexes is not intercalated between DNA bases. *J. Mol. Recog.* **7**, 221–226 (1994).
55. Florian, J. & Warshel, A. Calculations of hydration entropies of hydrophobic, polar and ionic solutes in the framework of the Langevin dipoles solvation model. *J. Phys. Chem.* **103**, 10282–10288 (1999).
56. Singh, N. & Warshel, A. A comprehensive examination of the contributions to the binding entropy of protein–ligand complexes. *Proteins* **78**, 1724–1735 (2010).
57. Yoon, H., Kolev, V. & Warshel, A. Validating the water flooding approach by comparing it to grand canonical Monte Carlo simulations. *J. Phys. Chem. B.* **121**, 9358–9365 (2017).
58. El Hage, K., Hedin, F., Gupta, P.K., Meuwly, M. & Karplus, M. Valid molecular dynamics simulations of human hemoglobin require a surprisingly large box size. *eLife* **7**, e35560 (2018). DOI:10.7554/eLife.35560.
59. Amaro, R.A., Baudry, J., Chodera, J., Demir, Ö., McCammon, J.A., Miao, Y. & Smith, J.C. Ensemble docking in drug discovery. *Biophys. J.* **114**, 2271–2278 (2018).

60. Hobza, P. Calculations on noncovalent interactions and databases of benchmark interaction energies. *Acc. Chem. Res.* **45**, 663–672 (2012).
61. Cooper, V.R., Thonhauser, T., Puzder, A., Schröder, E., Lundqvist, B.I. & Langreth, D.C. Stacking interactions and the twist of DNA. *J. Am. Chem. Soc.* **130**, 1304–1308 (2008).
62. Sturm, A., Perczel, A., Ivics, Z. & Vellai, T. The Piwi-piRNA pathway: road to immortality. *Aging Cell* **16**, 1–6 (2017).
63. Dobson, J.F. & Gould, T. Calculation of dispersion forces. *J. Phys. Condens. Matter* **24**, 40–50 (2012).
64. Giudic, E., Varnai, P. & Lavery, R. Base-pair opening within B-DNA: free energy pathways for GC and AT pairs from umbrella sampling simulations. *Nucleic Acids Res.* **31**, 1434–1443 (2003).
65. Chapman, C.D., Gorczyca, S. & Robertson-Anderson, R.M. Crowding induces complex ergodic diffusion and dynamic elongation of large DNA molecules. *Biophys. J.* **108**, 1220–1228 (2015).
66. Marcus, R.A. Micelle-enhanced dissociation of a Ru cation/DNA complex. *J. Phys. Chem. B.* **109**, 21419–21424 (2005). DOI: 10.1021/jp051988s.
67. Rebek, J. *Hydrogen-Bonded Capsules: Molecular Behavior in Small Spaces.* World Scientific — London 2016.
68. Jiang, W., Tiefenbacher, K., Ajami, D. & Rebek, J. Complexes within complexes: hydrogen bonding in capsules. *J. Chem. Sci.* **3**, 3022–302 (2012).
69. Ho, J.C.S., Rangamani, P., Liedberg, B. & Parikh, A.N. Mixing water, transducing energy, and shaping membranes: autonomously self-regulating giant vesicles. *Langmuir* **32**, 2151–2163 (2016).
70. Oglecka, K., Rangamani, P., Liedberg, B., Kraut, R.S. & Parikh, A.N. Oscillatory phase separation in giant lipid vesicles induced by trans-membrane osmotic differentials. *eLife* **3**, e03695 (2014). DOI: 10.7554/eLife.03695.
71. Sunde, E.P., Stelow, P., Hederstedt, L. & Halle, B. The physical state of water in bacterial spores. *Proc. Natl. Acad. Sci. USA* **106**, 19334–19339 (2009). DOI: 10.1073/pnas.0908712106.
72. Horneck, G., Klaus, D.M. & Mancinelli, R.L. Space microbiology. *Microbiol. Mol. Biol. Rev.* **74**, 121–156 (2010). DOI: 10.1128/MMBR.00016-09.