**Supporting Information**

**A stretched conformation of DNA with a biological role?**

*Niklas Bosaeus, Anna Reymer, Tamás Beke-Somfai, Tom Brown, Masayuki Takahashi, Pernilla Wittung-Stafshede, Sandra Rocha and Bengt Nordén*

**Content:**

Methods, materials and calculations (sections 1, 2 and 3)

Tables S1-S2

Figures S1-S2

References

**Methods, materials and calculations** (in addition to what is given in figure legends)

***1. DNA oligomers sequences***

***DNA1***

5’- ATCAACAGAGCCACTTTGGCCCGCTGGTCGCCGCTATCGACAGGCTCAATGCTGGAGGGT O1

#A 12, #G 18, #C 18, #T 12 | 60-mer

5’- ACCCTCCAGCATTGAGCCTGTCGATAGCGGCGACCAGCGGGCCAAAGTGGCTCTGTTGAT O1\_Comp

#A 12, #G 18, #C 18, #T 12 | 60-mer

5'- ATCAACAGAGCCACTTTGGCCCGCTGGTCGCCGCTATCGACAGGCTCAATGCTGGAGGGT

 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

3’- TAGTTGTCTCGGTGAAACCGGGCGACCAGCGGCGATAGCTGTCCGAGTTACGACCTCCCA

***DNA2***

5'- AACGCTTCGGCGACAAACCTGGGGCTTTCCTCGCCATGTGTGACGGGACACGTAGGACCT O2

#A 12, #G 18, #C 18, #T 12 | 60-mer

5’- AGGTCCTACGTGTCCCGTCACACATGGCGAGGAAAGCCCCAGGTTTGTCGCCGAAGCGTT O2\_comp

#A 12, #G 18, #C 18, #T 12 | 60-mer

5'- AACGCTTCGGCGACAAACCTGGGGCTTTCCTCGCCATGTGTGACGGGACACGTAGGACCT

 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

3’- TTGCGAAGCCGCTGTTTGGACCCCGAAAGGAGCGGTACACACTGCCCTGTGCATCCTGGA

***L’-, R’-, L’R’-DNA1***

5’-XAATCAACAGAGCCACTTTGGCCCGCTGGTCGCCGCTATCGACAGGCTCAATGCTGGAGGGTXA O1\_with\_linker\_sites

#A 14, #G 18, #C 18, #T 12 | 64-mer

5’-YAACCCTCCAGCATTGAGCCTGTCGATAGCGGCGACCAGCGGGCCAAAGTGGCTCTGTTGATYA Complement

#A 14, #G 18, #C 18, #T 12 | 64-mer

5'-XAATCAACAGAGCCACTTTGGCCCGCTGGTCGCCGCTATCGACAGGCTCAATGCTGGAGGGTXA

 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

3’-AYTAGTTGTCTCGGTGAAACCGGGCGACCAGCGGCGATAGCTGTCCGAGTTACGACCTCCCAAY

L’R’-DNA1: X = OCTADINYL dU, Y = AZIDOHEXANOYLAMINO C6 dT.

R’DNA1: 5’-X of O1\_with\_linker\_sites and 3’-Y of Complement are replaced by T.

L’DNA1: 3’-X of O1\_with\_linker\_sites and 5’-Y of Complement are replaced by T.

***(DNA1)2***

5’- ATCAACAGAGCCACTTTGGCCCGCTGGTCGCCGCTATCGACAGGCTCAATGCTGGAGGGT

ATCAACAGAGCCACTTTGGCCCGCTGGTCGCCGCTATCGACAGGCTCAATGCTGGAGGGTXATT O1+O1+link

#A 25 #G 36 #C 36 #T 26 | 124-mer | X = OCTADIYNYL dU

5’-YAACCCTCCAGCATTGAGCCTGTCGATAGCGGCGACCAGCGGGCCAAAGTGGCTCTGTTGAT

ACCCTCCAGCATTGAGCCTGTCGATAGCGGCGACCAGCGGGCCAAAGTGGCTCTGTTGAT link+O1\_Comp+O1\_Comp

#A 25 #G 36 #C 36 #T 26 | 122-mer | Y = AZIDOHEXANOYLAMINO C6 dT

***S3’3’-DNA1***

5’- AYCAACAGAGCCACTTTGGCCCGCTGGTCGCCGCTATCGACAGGCTCAATGCTGGAGGGT

TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTBBBT link+O1+biohandle

#A 12 #G 18 #C 18 #T 42 | 94-mer | Y = AZIDOHEXANOYLAMINO C6 dT, B=Biotin dT

5’- ACCCTCCAGCATTGAGCCTGTCGATAGCGGCGACCAGCGGGCCAAAGTGGCTCTGTTGAZ

TTTTTTTTTTTTTTTTTTTTTTTTTTTTTT3 O1\_Comp+link+alkhandle

#A 12 #G 18 #C 18 #T 41 | 91-mer | Z=AMINO C6 dT+BCN active ester, 3=3'mod-ALKYNE

***S5’5’-DNA1***

5’-BBBTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAYCAACAGAGCCACTTTGGCCCGCTGGTC

GCCGCTATCGACAGGCTCAATGCTGGAGGGT biohandle+link+O1

#A 12 #G 18 #C 18 #T 41 | 93-mer | Y = AZIDOHEXANOYLAMINO C6 dT, B=Biotin dT

5’-2TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTACCCTCCAGCATTGAGCCTGTCGATAGCGGC

GACCAGCGGGCCAAAGTGGCTCTGTTGAZT alkhandle+O1\_Comp+link

#A 12 #G 18 #C 18 #T 42 | 91-mer | Z=AMINO C6 dT+BCN active ester, 2=5'mod-HEXYNOL

***S5’3’-DNA1***

5’-BBBTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAYCAACAGAGCCACTTTGGCCCGCTGGTC

GCCGCTATCGACAGGCTCAATGCTGGAGGGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT3 biohandle+link+O1+alkhandle

#A 12 #G 18 #C 18 #T 71 | 123-mer | Y = AZIDOHEXANOYLAMINO C6 dT, B=Biotin dT, 3=3'mod-ALKYNE

5’- ACCCTCCAGCATTGAGCCTGTCGATAGCGGCGACCAGCGGGCCAAAGTGGCTCTGTTGAZT

#A 12 #G 18 #C 18 #T 12 | 61-mer | Z=AMINO C6 dT+BCN active ester

The 3’-ends of sequences DNA1, DNA2, L’-DNA1, R’-DNA1, L’R’-DNA1, and (DNA1)2 were extended using terminal transferase to incorporate either Biotin or Digoxigenin labelled nucleotides for attachment to Streptavidin (Spherotech) or Anti-Digoxigenin coated polystyrene beads, respectively. Sequences S3’3’-DNA1, S5’5’-DNA1, and S5’3’-DNA1 were attached to Streptavidin coated and azide-modified beads. The Anti-Digoxigenin and azide-modified beads were prepared by crosslinking either sheep polyclonal digoxigenin antibody (Roche) using dimethyl pimelimidate, or azidobutyrate NHS ester (Glen Research), to Protein G coated beads (Spherotech), respectively.

***2. Fitting a three state model***

The probability of finding a system in a given state n is given by the Boltzmann distribution:

for three non-degenerate states the partition function *q* is

The energy level for the n:th state at a given force is

In the used model we know the force and the free energy, *ΔG°* and *Δx* are fitted.

*a* and *d* contains the cooperative length for the transitions *B → I* and *B → S*, respectively. The free energy for the transitions are found in parameters *b* and *e*. We define three different transition forces: *FtrI* where *PI = PB*, *FtrII* where *PS = PI*, and *Ftr* as *max(PI)*

***3. Calculations of DNA stacking energies***

Energetics of DNA base stacking was studied by molecular dynamics (MD) simulations as well as quantum mechanics (QM) calculations (Friedman & Honig,  1995; Giudice, Várnai, & Lavery,  2003; Da̧bkowska *et al.*,  2005; Řezáč & Hobza,  2007). As summarized in **Table S1**, hydrogen bonds formed between the base pairs have a large contribution to base stability, however, accompanied with a large hydration free energy per base, which may contribute to destabilization of Watson-Crick (WC) hydrogen bonds. Entropy terms, based on QM normal mode analysis, further decrease the large stabilization by hydrogen bonds. Note, that interactions between base pair dimers will also be dependent on the stacking distance, twist and rotation, as pointed out in a detailed study by Swart *et al*. (Swart *et al.*,  2007)

To better understand energetics of DNA base stacking, we performed QM calculations using the X-D functional calculations of stacking stabilization for a planar benzoic-acid hydrogen-bonded dimer, which we view as a primitive symmetric model for a DNA base pair. The obtained values are in close agreement with previous stacking energy calculations on benzene dimers (Swart *et al.*,  2007), which resulted in -1.70 kcal/mol and -1.58 kcal/mol for CCSD(T) and KT1 potentials, respectively. These values lie close to the -2.1 kcal/mol obtained here for benzoic acid dimers. The small difference is probably due to the carboxyl groups present in the latter (our) system rendering a shorter distance ( ~3.4 Å) between the benzoic acid planes as compared to in the benzene dimers (~3.8-3.9 Å) (Swart *et al.*,  2007).

We also considered effects of surrounding environment, water and non-polar cyclohexane, on hydrogen bond strength. This was calculated using the integral equation formalism for polarizable continuum solvent model (Cancès, Mennucci, & Tomasi,  1997). As detailed in the main text, hydrogen bonds are strengthened by a non-polar matrix (**Table S2**), whereas for intra-strand stacking, the stacking of base pairs is more favoured in water by 1.2 kcal/mol.

Further, we investigated how relative energy gains of adding an additional dimer of benzoic acid to a stacked construct may determine where it is most economic to separate them under external force. Note that the available X-ray structures on extended DNA conformations harbour triplets (Chen, Yang, & Pavletich,  2008) that closely resemble B-DNA both in terms of DNA base pair twists and distance between the stacked pairs. Consequently, a simplified approximation appeared reasonable. Using the obtained energy values on WC pairing, intra-strand as well as inter-strand stacking, we calculated how the relative energy gain per BAbase pair changes when the number of pairs is gradually increased (**Figure S2**). Here we considered the effect of environment on base pair stacking and WC hydrogen bonding: as there is a disruption of base pair stacking the end base pairs were estimated to be 50% hydrated, and thus energy values were modified accordingly. The resulting curve indicates that the relative change in energy gain per benzoic acid pair of dimers gets significantly lower after the third base pair (**Figure S2**). As already discussed in the main text, the energy gain between the 2nd and 3rd benzoic acid pair is nearly double the gain for the 3rd to 4th pair. Note, that the identification of the exact values has so far not been achieved by modelling studies. Nevertheless, this characteristic relative energy difference among the stacking of the 2nd, 3rd, and 4th base pairs is the result of the consideration that relative hydration of the end base pairs will be different from those being in the core of a stacked conformation. Consequently, the decay of the energy curve on **Figure S2** will likely be similar irrespective of the exact stacking energy values and exact percentage of end base pair hydration.

**Table S1.** Energetic contributions to DNA base pair stacking interactions

|  |  |  |  |
| --- | --- | --- | --- |
| ***Interaction*** | ***Base(s)/base pair(s) involved*** | ***Energy term*** | ***Value(s) (kcal/mol)*** |
| Base pairing | AT, GC | *T**ΔS*(normal mode analysis) | 11.1-12.2 a |
| Base hydration, | A, T, G, C | Gibbs energies | -12.0; -12.4; -22.4; -18.4 b |
| Watson-Crick hydrogen bonds (WC) | AT, GC | Relative energies | -16.4; -35.8 c |
| Intrastrand stacking (S) | AT, GC | Relative energies | -8.1; -7.9 c |
| AA, GG | Relative energies | -0.7; -4.5 c  |
| TT, CC | Relative energies | 1.0; 1.4 c  |
| Base stacking | CC, GG | Free energies | -4.36; -7.79 d |
| Base stacking | CC, GG | Non-polar contributions to base stacking | -5.26; -9.11 d  |

a(Hobza & Šponer,  1996); b(Miller & Kollman,  1996); c(Da̧bkowska *et al.*,  2005); d(Friedman & Honig,  1995)

**Table S2.** Relative energies: dispersion and stacking energies of benzoic acid dimerization as function of the surrounding environment\*.

|  |  |  |  |
| --- | --- | --- | --- |
| ***Model*** | ***vacuum*** | ***water*** | ***C6H12(cyclohexane)*** |
|  | ***ΔE* (kcal/mol)** |
| WC | -18.5 | -13.46 | -16.44 |
| S | -2.1 | -3.6 | -2.4 |
| IS1 | 2.1 | 1.9 | 2.0 |
| IS2 | -0.3 | 0.5 | 0.2 |
| S2x2BA | -8.4 | -8.3 | -8.6 |

*\*All values were obtained at the B97X-D/6-311++G(d,p) level of theory, using the IEFPCM model for solvent effects. Values are relative to the corresponding number of monomeric benzoic acids obtained at the same level of theory. The dispersion terms were obtained from the empirical atomic-pairwise dispersion correction of B97X-D*(Chai & Head-Gordon,  2008) *functional as implemented in the Gaussian 09 software package*(Frisch et al.,  2009)*.*

***WC****: Benzoic acid dimer with 2 H-bonds;* ***S****:-stacked benzoic acid dimer,* ***IS****: interstrand stacking in a dimer opposite to each other as shown on* ***Figure S1****.* ***S2x2BA****: Intrastrand stacking energy obtained from a doublet* ***BA****dimers,-stacked (****Figure S1****).*



**Figure S1.** Schematic representation of -stacked benzoic acid (**BA**) dimers as symmetric models for DNA base pair stacking. The pairwise interactions labelled WC, S, and IS were calculated in vacuum, water and in an apolar environment (cyclohexane) representing the inner core of a  stacked filament.

**Figure S2**. The relative energy gain when increasingly longer stacked **BA** pairs are considered based on calculations using **Table S2**. The number of **BA** pairs stacked together are displayed on the X-axis, while the relative energies per base pair in kcal/mol are given on the Y-axis. To estimate environmental effects, the end **BA** pairs in each stacked fragment are considered as 50 % hydrated.

**References**

Cancès, E., Mennucci, B. & Tomasi, J. (1997) A new integral equation formalism for the polarizable continuum model: Theoretical background and applications to isotropic and anisotropic dielectrics. *The Journal of Chemical Physics* **107**, 3032–3041. American Institute of Physics.

Chai, J.-D. & Head-Gordon, M. (2008) Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections. *Physical Chemistry Chemical Physics* **10**, 6615–6620. The Royal Society of Chemistry.

Chen, Z., Yang, H. & Pavletich, N.P. (2008) Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures. *Nature* **453**, 489–494. Nature Publishing Group.

Da̧bkowska, I., Gonzalez, H.V., Jurečka, P. & Hobza, P. (2005) Stabilization Energies of the Hydrogen-Bonded and Stacked Structures of Nucleic Acid Base Pairs in the Crystal Geometries of CG, AT, and AC DNA Steps and in the NMR Geometry of the 5‘-d(GCGAAGC)-3‘ Hairpin:  Complete Basis Set Calculations at the MP2 and . *The Journal of Physical Chemistry A* **109**, 1131–1136. American Chemical Society.

Friedman, R.A. & Honig, B. (1995) A free energy analysis of nucleic acid base stacking in aqueous solution. *Biophysical journal* **69**, 1528–1535.

Frisch et al., M.J. (2009) Gaussian09, Revision A.1. *Gaussian, Inc., Wallingford CT*.

Giudice, E., Várnai, P. & Lavery, R. (2003) Base pair opening within B-DNA: free energy pathways for GC and AT pairs from umbrella sampling simulations. *Nucleic Acids Research* **31**, 1434–1443. Oxford University Press, Oxford, UK.

Hobza, P. & Šponer, J. (1996) Thermodynamic characteristics for the formation of H-bonded DNA base pairs. *Chemical Physics Letters* **261**, 379–384.

Miller, J.L. & Kollman, P.A. (1996) Solvation Free Energies of the Nucleic Acid Bases. *The Journal of Physical Chemistry* **100**, 8587–8594. American Chemical Society.

Řezáč, J. & Hobza, P. (2007) On the Nature of DNA-Duplex Stability. *Chemistry - A European Journal* **13**, 2983–2989. WILEY‐VCH Verlag.

Swart, M., van der Wijst, T., Fonseca Guerra, C. & Bickelhaupt, F.M. (2007) π-π stacking tackled with density functional theory. *Journal of Molecular Modeling* **13**, 1245–1257.