

Scientific paper

Low-frequency Vibrations of DNA and Base Pair Opening

Franci Merzel^{1,*} and Mark R. Johnson²¹ Laboratory for Molecular Modeling, National Institute of Chemistry, Hajdrihova 19, SI-1000 Ljubljana, Slovenia² Institut Max von Laue-Paul Langevin, 6 rue Jules Horowitz, BP 156, 38042 Grenoble Cedex 9, France

* Corresponding author: E-mail: franci.merzel@ki.si

Received: 06-05-2011

Dedicated to Professor Dušan Hadži on the occasion of his 90th birthday

Abstract

DNA melts at temperatures around 70 °C (~340 K) depending on the base-pair composition, level of hydration, type of counterions and excess salt concentration. Vibrational modes, which increase the separation between base molecules, i.e. perform the base pair opening, are thought to play an important role in DNA-bubble formation but these are typically assigned frequencies around 90 cm⁻¹ (120 K). Understanding DNA melting and related biological processes requires this apparent energy gap of more than 200 K to be understood. We have performed atomistic molecular simulations which enable us to reconcile this energy difference between cause and effect. Using quasi-harmonic analysis we have analyzed the way in which vibrational modes of DNA change with the temperature.

Keywords: B-DNA, molecular dynamics simulations, quasiharmonic analysis, vibrational density of states, base-pair opening.

1. Introduction

Understanding the function of biological molecules requires knowledge of their dynamics. In the case of deoxyribonucleic acid (DNA), one of the key points of interest is base-pair opening as such dynamics play a key role in replication, transcription and melting.¹ These processes all involve the splitting of the double helix into single strands, which is initiated locally by the breaking of inter-base hydrogen bonds and the formation of bubbles spanning several base-pairs. Although these processes often involve proteins, they are thought to be driven by the dynamics of DNA itself.

Experimentally, the dynamics of DNA has attracted considerable attention. The sound velocity parallel and perpendicular to the helix axis for A and B form as a function of humidity was measured using Brillouin² and Raman³ scattering. Dynamics of DNA was also studied by coherent inelastic neutron scattering (INS)⁴ and inelastic X-ray scattering in a range of samples.⁵

Numerical and theoretical models are also essential for a detailed understanding of low frequency dynamics of

DNA. Mesoscale models have been developed, in which DNA is typically represented by a one-dimensional chain of effective atoms with a Morse potential describing the hydrogen-bonding between the base-pair molecules. The Peyrard-Bishop-Dauxois model⁶ has been parameterized to reproduce denaturation curves and describes base-pair opening frequencies at 40 cm⁻¹ and 80 cm⁻¹ for adenine-thymine (A-T) and guanine-cytosine (G-C) pairs respectively. Early atomistic models predict base-pair opening frequencies between 59 and 84 cm⁻¹ for A-T and between 95 and 124 cm⁻¹ for G-C⁷.

In our previous study⁸ we have focused on dispersion relations at low frequencies and characterization of the vibrational modes using all-atom B-DNA models and large scale phonon calculations⁹ based on the harmonic approximation. Our motivation here was to analyse the way in which the vibrational modes of DNA change with temperature. To this end we have performed a series of molecular dynamics (MD) simulations at different temperatures followed by quasi-harmonic analysis¹⁰ of the trajectories. The resulting vibrational modes are discussed in the context of base-pair opening.

2. Methods

Crystallographic coordinates of B-DNA form were taken from the Protein Data Bank (sequence AGCAGCAGAG for 1 strand). MD simulations at $T = 100$ K, 200 K, 300 K and 320 K were carried out with the CHARMM¹¹ molecular modeling package together with the all-atom force field, CHARMM22.¹² The highest temperature, $T = 320$ K, was chosen just below the melting temperature to ensure the structural stability of DNA during the MD simulation. A periodic fibril structure of DNA in the B form, 10 base-pairs per helical pitch, length 34.6 Å, was obtained during relaxation of the hexagonal simulation box ($a = 39.0$ Å, $c = 34.6$ Å) filled with 20 lithium counter ions and 1460 water molecules (Figure 1A) using periodic boundary conditions and linking the DNA strands in the primary box with their periodic images. The amount of water in the unit-cell creates a fully hydrated state of DNA fiber with the minimum fiber to unit-cell-surface distance of more than three water-shell layers. Consequently, the variation of the size of the lateral dimension of the unit cell is expected to have negligible influence on the base-pair opening characteristics of DNA in our simulations.

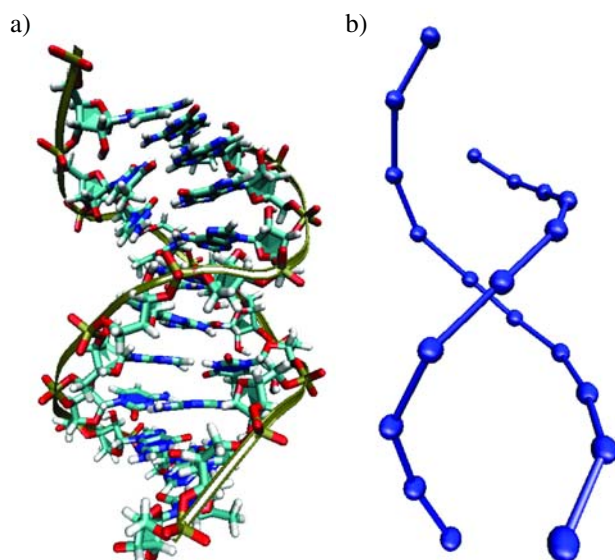


Figure 1. A) Atomistic B-DNA model used in this work (water molecules, counter ions and periodic simulation box not shown). B) Alternative representation of DNA in which the center of mass of each nucleotide is represented by a bead.

The water was treated with the TIP3P model.¹³ The MD simulations of 50 ns at constant pressure, $p = 1$ atm., and constant temperatures (CPT) using Nose-Hoover method were carried out with an integration step of 1 fs. Structures were saved every 0.1 ps to produce trajectories. The non-bonded interactions were treated using a cut-off of 12 Å.

To enable analysis of DNA dynamics at different levels of coarse graining, we introduced the projection of atomic displacements on the center of mass of a given nucleotide, or eventually, on a base molecule, as implemented in the program NMscat.⁹ Nucleotide centers of mass are represented by beads as shown in Figure 1B.

It is common to analyze molecular vibrations in terms of the vibrational density of states $g(\omega)$, which is usually extracted from the shape of the potential energy surface of the atomic system using normal-mode analysis as a frequency distribution function of vibrational modes. In this case, vibrational modes are represented as eigenvectors of the Hessian matrix, which is the matrix of mass-weighted second derivatives of potential energy with respect to atomic displacements. An alternative way to derive $g(\omega)$ is to apply the quasi-harmonic approximation (QHA) on the MD trajectory of the system. This method has shown to be very useful in studying biomolecular dynamics^{14, 15, 16} because it also probes anharmonicity of inter-atomic interactions. In QHA, the fluctuations observed during the system motions are described by a multivariate Gaussian probability distribution. Using this assumption, a temperature-dependent effective potential resulting in a Gaussian probability distribution can be defined. The quasi-harmonic modes can then be computed from the mass-weighted Hessian matrix defined in terms of its elements as $k_B T / \sigma_{ij}$, where k_B is the Boltzmann constant and T is the absolute temperature. Consequently, σ_{ij} are the elements of the covariance matrix. Those elements are defined as dynamics dependent variances (diagonal elements) and covariances (off-diagonal elements) of the fluctuations of the Cartesian atomic coordinates, x_i ,

$$\sigma_{ij} = \langle (x_i - \langle x_i \rangle)(x_j - \langle x_j \rangle) \rangle \quad (1)$$

that are obtained from MD simulation with $\langle \rangle$ denoting time average over MD trajectory. Solving the secular equation

$$\det \left(M^{1/2} \sigma M^{1/2} - \frac{k_B T}{\omega^2} I \right) = 0 \quad (2)$$

where M is the mass matrix and I is the identical matrix, one obtains the quasi-harmonic frequencies ω_i , which are used to compute vibrational density of states:

$$g(\omega) = \frac{1}{D} \sum_{i=1}^D \delta(\omega - \omega_i) \quad (3)$$

where δ is the Dirac delta function and D is the total number of degrees of freedom.

Further useful information can be extracted from QHA when analyzing the directionality and the size of atomic displacements in a given vibrational mode. For example, the quantity revealing the breathing or base-pair-opening character of the vibrational mode can be defined

in terms of projection operators of atomic/bead displacements:

$$B = \frac{1}{N_p} \sum_p \frac{N_p \left| \frac{\vec{r}}{|\vec{r}|} (\vec{d}_p[1] - \vec{d}_p[2]) \right|}{2 \max(|\vec{d}_p[1]|, |\vec{d}_p[2]|)}, \quad (4)$$

where $\vec{d}_p[1]$, $\vec{d}_p[2]$ and $\vec{d}_p[2]$ are the net displacement vectors of the beads forming an individual nucleotide pair. $\vec{r}/|\vec{r}|$ is the directional vector along the line connecting two beads of a pair, and N_p is the number of nucleotide pairs in the model DNA, i.e. 10. Plotting B as a function of vibrational frequency enables a systematic mode assignment across the entire frequency range.

3. Results

We first compare the vibrational density of states, $g(\omega)$, derived from the normal mode analysis (NMA) and QHA from the simulation at $T = 100$ K, which is expected to exhibit harmonic type of motion (Figure 2A). The overall agreement of both spectra is good. Temperature dependent $g(\omega)$ derived from QHA in the low-frequency range, i.e. < 600 cm^{-1} , is shown in Figure 2B. At higher

temperatures ($T \geq 300$ K) we observe the broadening of the first low frequency band at $\omega < 100$ cm^{-1} and increase of the density of states in the frequency range $250 < \omega < 400$ cm^{-1} . The systematic increasing of the DOS at $\omega < 30$ cm^{-1} with increasing temperature can be attributed to the softening of the DNA dynamics at higher temperatures.

For comparison we show in Figure 3 the vibrational density of states, $g(\omega)$, calculated directly from the MD trajectory at 100 K as the velocity autocorrelation function using the program nMoldyn¹⁷ which agrees remarkably well with the result obtained from the normal mode analysis (Figure 2A).

The distribution function of the base-pair opening character of the bead motion is plotted versus frequency in Figure 4. The projection of the atomic displacement vectors summed over individual nucleotides on to the corresponding inter-nucleotide vectors was calculated by equation (4) and analysis was carried out for eigenvectors obtained from NMA and QHA of dynamics at 100 K, respectively. Both distributions show a maximum at 80–130 cm^{-1} . According to the definition (4) the peak value of $B = 0.6$ means that the base-pair opening character occurs at about six nucleotide pairs, out of 10, in low frequency modes. The high level of agreement between the NMA and QHA results indicates clearly the accordance

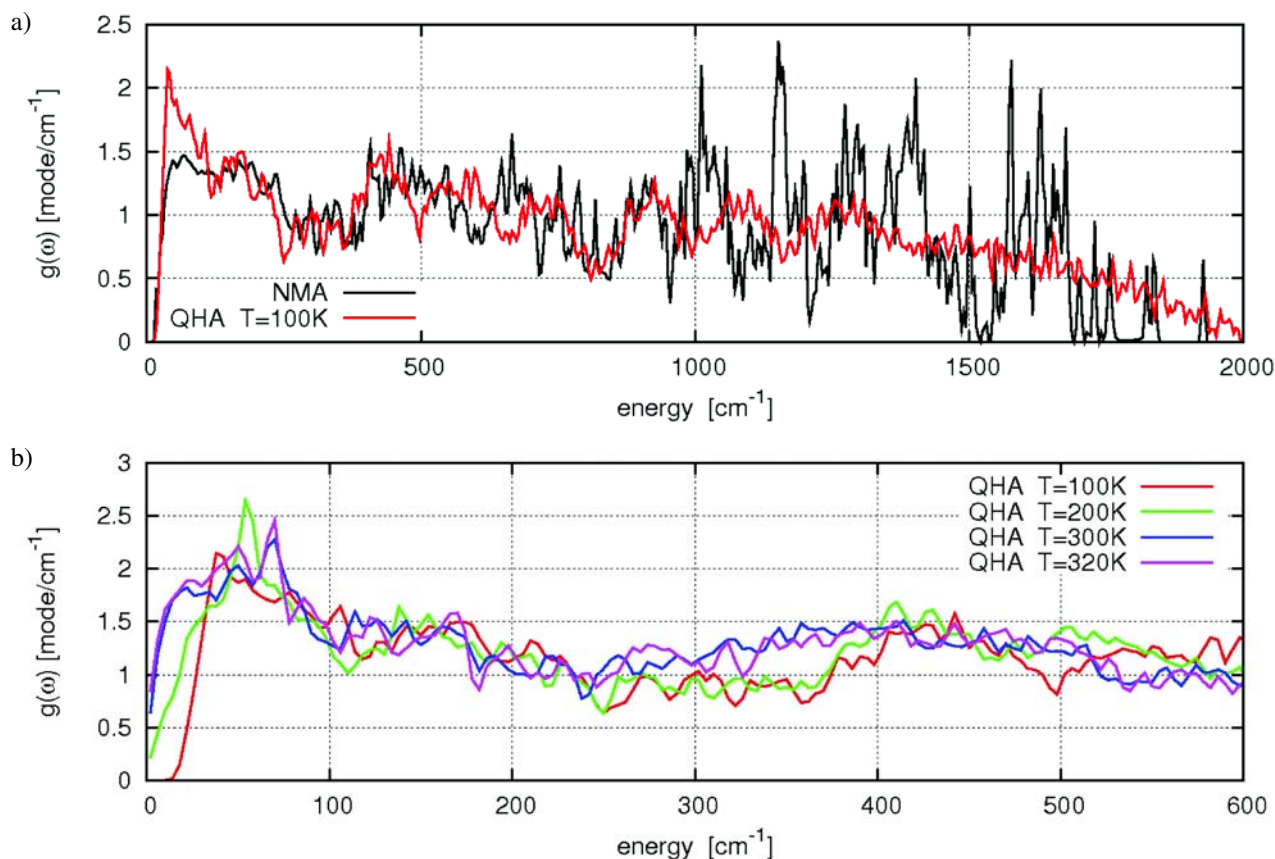


Figure 2. Vibrational density of states (vDOS) of B-DNA. A) Comparison of vDOS obtained by the normal mode analysis (NMA) and by the quasi-harmonic analysis (QHA) of the MD simulation at 100K. B) Temperature dependent low-frequency vDOS derived from QHA.

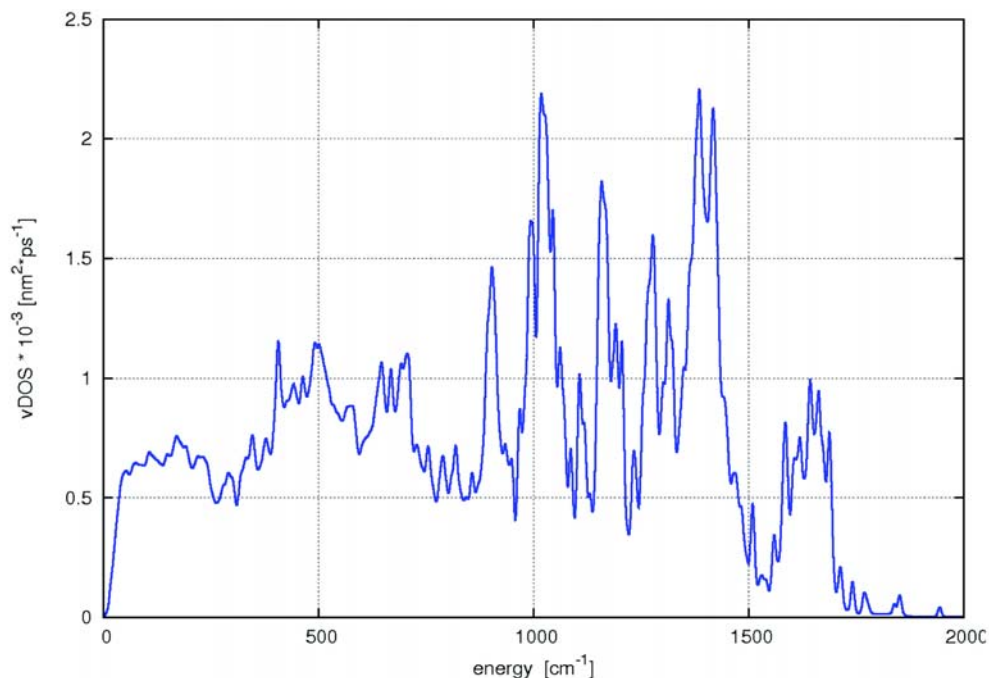


Figure 3. Vibrational density of states derived from the velocity autocorrelation function using the program nMoldyn¹⁷.

of static and dynamic representations of low-frequency vibrations.

In Figure 5 we show distribution functions of the base-pair opening character of bead dynamics at two levels of coarse-graining, as revealed by simulations of DNA at different temperatures. First we define beads as

representing nucleotides (Figure 5A) and secondly as representing only the base molecules (Figure 5B). Apart from the modest shift of the position of the first maxima towards lower frequencies with increasing temperatures from 100 K to 300 K, the distribution functions in Figure 5A do not differ much from each other. However, clo-

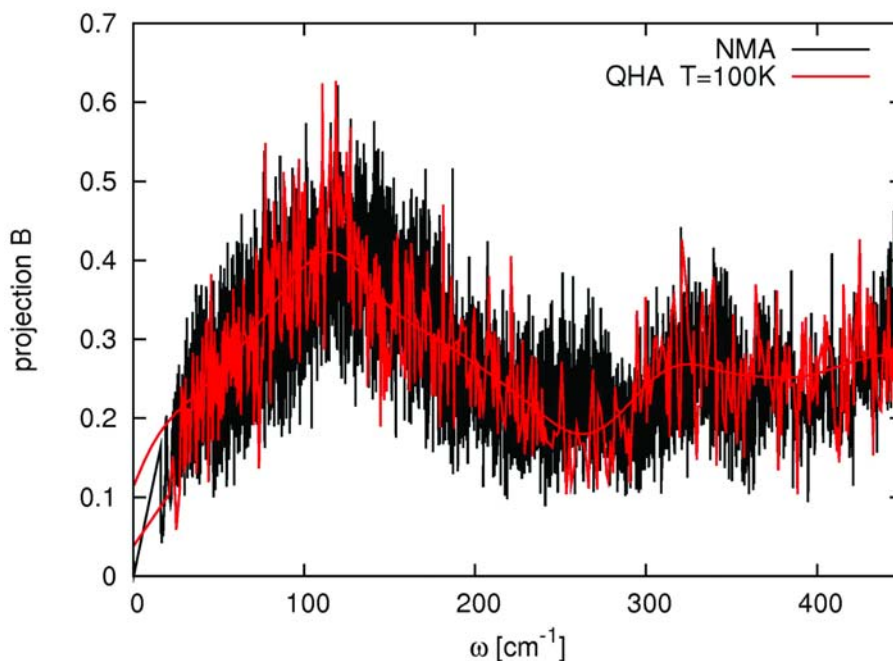


Figure 4. “Breathing” or base-pair-opening character of individual vibrational modes as a function of mode frequency. Results for the normal modes is shown in black and for the quasi-harmonic modes in red. (Smoothed red curve is obtained by a Gaussian convolution of the QHA values.)

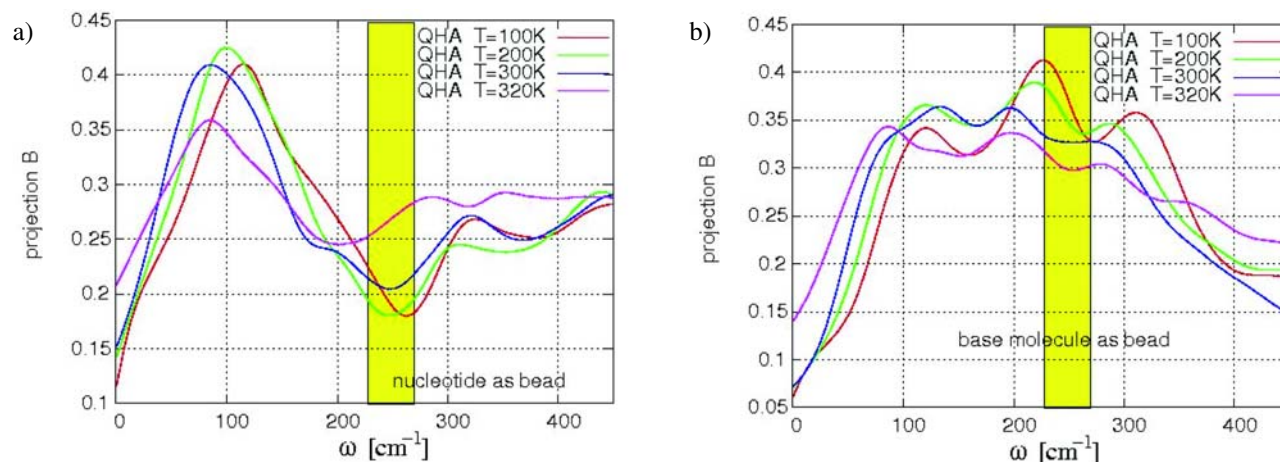


Figure 5. Base-pair-opening character of QHA vibrational modes of B-DNA at different temperatures as a function of frequency analyzed in terms of A) nucleotides and B) base-molecules. The yellow band indicates the energy range corresponding to the melting temperature.

se to the melting point (320 K), the distribution becomes flatter: the maximum at ~ 100 cm⁻¹ is lower while modes at $\omega > 200$ cm⁻¹ exhibit stronger base-pair opening character. The energy range $230 < \omega < 260$ cm⁻¹ depicted by the yellow boxes in Figure 5 corresponds to the thermal energy of DNA melting. For $T \leq 300$ K base-pair opening occurs over about four nucleotide pairs in low frequency modes and only about two nucleotide pairs at higher frequencies. The increase of the breathing character of modes at $\omega > 200$ cm⁻¹ demonstrated for the simulation at $T = 320$ K is relevant to melting. Also, the bubble formation preceding melting seems to be enhanced by more localized modes close to the melting point.

Redefining beads as base (Figure 5B), sugar and phosphate molecules (not shown) in our analysis gives a distribution functions with a continuous plateau up to ~ 320 cm⁻¹, demonstrating the presence of the breathing modes of base molecules up to frequencies that match the melting temperature of DNA. Comparing these distributions with those in which beads representing whole nucleotides (Figure 5A) leads to a conclusions that the average motion of nucleotide beads does actually hide the breathing-like character of the inter-base vibrations. This is due to the out-of-phase motion of nucleotide units (sugar, phosphate and base molecules) which reduce the net contribution to the displacement of the bead as a whole nucleotide. Overall, distribution functions become broader and lower with an increasing temperature. For the base-molecules, the projection B shows an opposite (decreasing) trend with increasing temperature as for whole nucleotides (Figure 5A) at frequency range $200 < \omega < 300$ cm⁻¹. While this suggests more irregularity in the breathing character of the base molecules motion themselves with increasing temperature, it is clear that this effect must be compensated by the presence of increasingly higher level of the in-phase motion of nucleotide units within the nucleotide.

4. Conclusions

Using MD simulations and QHA we investigated dynamics of DNA over different temperatures ranging from 100 K to 320 K, i.e. close to the DNA melting temperature, in terms of the base pair opening, a precursor event of bubble formation. While the prediction of the maximum of the base-pair opening character of DNA dynamics of the present work is consistent with previous works, and the frequencies of these vibrations in bead models, the explanation of the vibration-melting energy gap was still missing. By analyzing the way in which vibrational modes change with temperature we were able to show that most significant changes of modes take place in the frequency range corresponding to the thermal energy of melting when approaching the melting temperature. By redefining beads as base molecules in our analysis gives distribution functions of the base-pair opening with a continuous plateau from 100 cm⁻¹ up to ~ 320 cm⁻¹, indicating that base-pair opening modes exist up to frequencies that match the melting temperature of DNA. We have also shown that static (NMA) and dynamic (QHA) representation of low-frequency vibrations of DNA are in good agreement.

5. Acknowledgement

F. M. acknowledges the financial support from the grant P1-0002 of the Slovenian Research Agency.

6. References

1. M. Peyrard, *Nature Physics* **2006**, 2, 13–14.
2. S. A. Lee, S.M. Lindsay, J.W. Powell, T. Weidlich, N. J. Tao, G. D. Lewen, A. Rupprecht, *Biopolymers* **1987**, 26, 637–1665.

3. Y. Tominaga, M. Shida, K. Kubota, H. Urabe, Y. Nishimura, M. Tsuboi *J. Chem. Phys.* **1985**, *83*, 5972–5975.
4. H. Grimm, H. Stiller, C. F. Majkrzak, A. Rupprecht, U. Dahlborg, *Phys. Rev. Lett.* **1987**, *59*, 1780–1783.
5. M. Krisch, A. Mermet, H. Grimm, V. T. Forsyth, A. Rupprecht, *Phys. Rev. E* **2006**, *73*, 061909 1–10.
6. M. Peyrard, A. R. Bishop, *Phys. Rev. Lett.* **1989**, *62*, 2755–2759.
7. R. Berger, Y. Feng, E. W. Prohofsky, *Biophys. J.* **1990**, *58*, 437–445.
8. F. Merzel, F. Fontaine-Vive, M. R. Johnson, G. J. Kearley, *Phys. Rev. E* **2007**, *76*, 031917 1–5.
9. F. Merzel, F. Fontaine-Vive, M.R. Johnson, *Comp. Phys. Comm.*, **2007**, *177*, 530.
10. B. R. Brooks, D. Janežič, M. Karplus, *J. Comput. Chem.* **1995**, *16*, 1522–1542.
11. B. R. Brooks, R. E. Brucoleri, B. D. Olafson, D. J. States, S. Swaminathan, M. Karplus, *J. Comput. Chem.* **1983**, *4*, 187–217.
12. A. D. J. MacKerrel, D. Bashford, M. Bellott, et al., *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
13. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926–935.
14. D. Janežič, R. M. Venable, B. R. Brooks, *J. Comp. Chem.* **1995**, *16*, 1554–1566.
15. M. A. Balsera, W. Wriggers, Y. Oono, K. Schulten, *J. Phys. Chem. B* **1996**, *100*, 2567–2572.
16. J. Zidar, F. Merzel, *J. Phys. Chem. B* **2011**, *115*, 2075–2081.
17. T. Rog, K. Murzyn, K. Hinsen, G. R. Kneller, *J. Comp. Chem.* **2003**, *24*, 657–667.

Povzetek

Molekula DNK ima tališče okrog 70 °C (~340 K) odvisno od zaporedja baznih parov, stopnje hidracije, tipa protiionov in presežne koncentracije soli. Nihajnim načinom DNK, pri katerih se spreminja razmik med baznimi molekulami in ki imajo pomembno vlogo pri nastanku mehurčkov v dvojni vijačnici, ustrezajo frekvence okrog 90 cm⁻¹ (120 K). Za zvezo med nihajnimi načini, ki razpirajo bazne pare in taljenjem DNK je torej potrebno razumeti obstoj energijske reže velikosti ~200 K. V ta namen smo izvedli atomistične simulacije, ki nam pojasnijo energijsko razliko v luči povezave med vzrokom in učinkom. Z uporabo kvaziharmonske analize smo analizirali tudi temperaturno odvisnost narave nihajnih načinov DNK.