SUPPORTING INFORMATION

Biomolecular structure manipulation using tailored electromagnetic radiation: proof of concept on a simplified model of the active site of bacterial DNA topoisomerase

Daungruthai Jarukanont,[1] João T. S. Coimbra,[2] Bernd Bauerhenne,[1] Pedro A. Fernandes,[2] Shekhar Patel,[3,*] Maria J. Ramos,[2,§] and Martin E. Garcia[1,#]

[1]Institute of Physics, FB 10, University of Kassel, Heinrich-Plett-Str. 40, 34132 Kassel, Germany

[2]Requimte, Departamento de Química e Bioquímica, Faculdade de Ciências, Universidade do Porto, Rua do Campo Alegre s/n, 4169-007 Porto, Portugal

[3]PepsiCo Advanced Research, 3 Skyline Dr, Hawthorne NY 10532, USA

[#]corresponding author (garcia@physik.uni-kassel.de)

[§]corresponding author (mjramos@fc.up.pt)

[*]corresponding author (Shekhar.Patel@pepsico.com)

CONTENTS

A detailed description of the protocols employed in the study is presented here. The basis for designing shaped electromagnetic pulses, able to induce selective and specific bond breaking is also provided. In addition, auxiliary results mentioned throughout the original article, are given. The different sections of the SI are organized as in the original manuscript, and the contents are organized as follows:

1. METHODS AND FURTHER RESULTS

i. Determination of Electric Field threshold values for the unfolding of topoisomerase and thioredoxin's active sites. We have performed the simulations under an NPT ensemble (constant number of particles, pressure, and temperature) with explicit water molecules of the TIP3P[1] type. We have employed the ff99SB[2] force field. After an initial preparation of the system, where protonation states of aminoacids were predicted with the PROPKA web interface,[3] proteins topoisomerase III and thioredoxin were subjected to MD simulations for the determination of electric field threshold values for their respective unfolding. The simulation protocol was divided into 3 stages: phase 1,

corresponding to the systems' minimization and equilibration stages; phase 2, production MDs lasting 10 ns each for insuring that the systems were stabilized; and phase 3, where 3 different external electric fields (0.1, 0.5, and 1.0 $V.nm^{-1}$) were applied lasting 10 ns each.

For the minimization process the steepest descent algorithm of minimization was employed, with unlimited minimization steps until the maximum force, F_{max} was no greater than 1000 $kJ.mol^{-1}.nm^{-1}$. Following minimization, two equilibration MD simulations were produced: one, in a canonical ensemble (constant number of particles, volume, and temperature – also referred as NVT) for 100 ps, in order to stabilize the system to the desired temperature; another, in an isothermal-isobaric ensemble for an additional 100 ps (constant number of particles, pressure, and temperature – also referred as NPT), closely resembling experimental conditions.

For the production MD and simulations in NPT and with an external electric field, a standard protocol was conducted. The leap-frog integrator[4] was employed with an integration time step of 2 fs. This was possible since the LINCS algorithm[5] that constraints the bonds' vibration was used. Since this is the fastest movement in an MD simulation a higher integration time step could be employed. The coordinates were saved and processed every 5 ps. A 10 Å cut-off was considered for the neighbour search with the grid option. Periodic boundary conditions considering the 3 axis were employed. The Particle Mesh Ewald (PME)[6] for long-range electrostatics was employed. The modified Berendsen thermostat (V-rescale)[7] was used for the temperature coupling at 313.15 K, together with the Parrinello-Rahman pressure coupler[8] at 1 atm (isotropically). The proteins' centre of mass translation was removed.

The analysis were performed with GROMACS[9] utilities and with the Visual Molecular Dynamics (VMD) program.[10] Ramachandran plots were analysed for the last 500 ps of each MD setting: production MD and MDs with the 3 different external electric field strengths; RMSd, radius of gyration, and secondary structure analysis was performed throughout the 10 ns.

In the following, we present additional and complementary results. Specifically, root-mean-square deviation (RMSD), radius of gyration (relative to the crystallographic minimized and equilibrated structure), Ramachandran plots and secondary structure analysis are presented for the topoisomerase enzyme – Figures S1, S2 and S3, considering both the production simulation and simulations performed with the 3 different external electric fields. In addition, all the information relative to the thioredoxin protein is presented too – molecular structure representation, RMSD, radius of gyration, Ramachandran plots, and secondary structure analysis (Figure S4 – S7).

Topoisomerase:

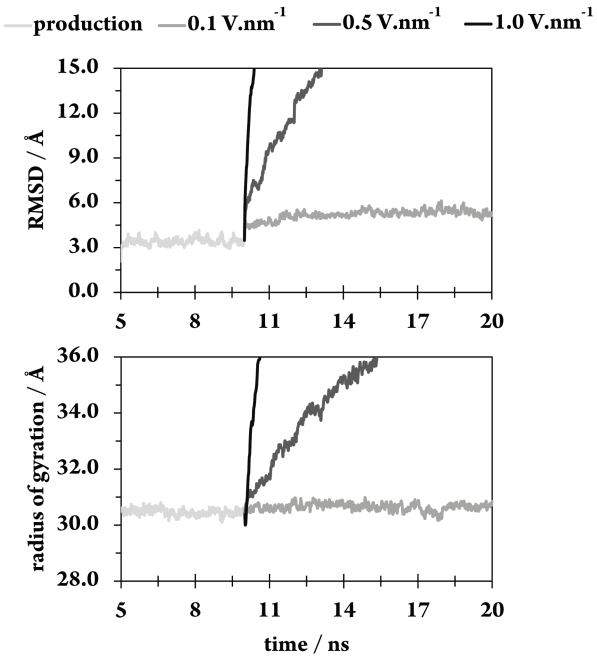

Figure S1. Root-mean-square deviation (RMSD), presented as a measure of protein stability, relative to the initial x-ray crystallography structure, minimized and equilibrated; and radius of gyration, as a measure of the proteins' compactness (if a protein is stably folded it will likely maintain a steady behaviour of the radius of gyration).

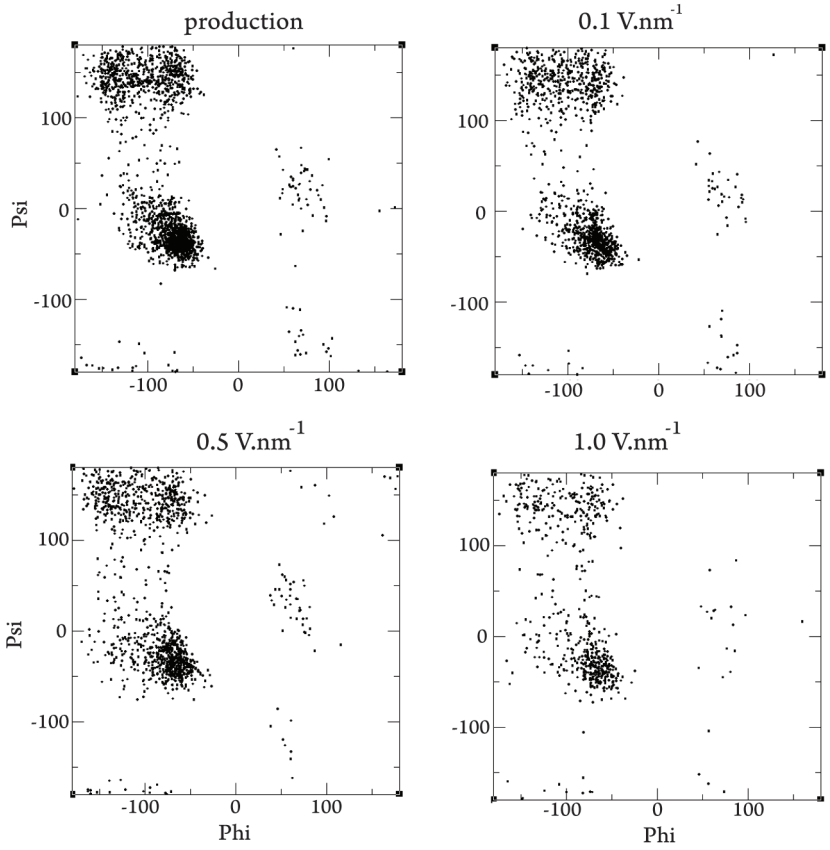

Figure S2. Ramachandran plots for the topoisomerase III protein, relative to the production MD simulation and the 3 tested external electric fields.

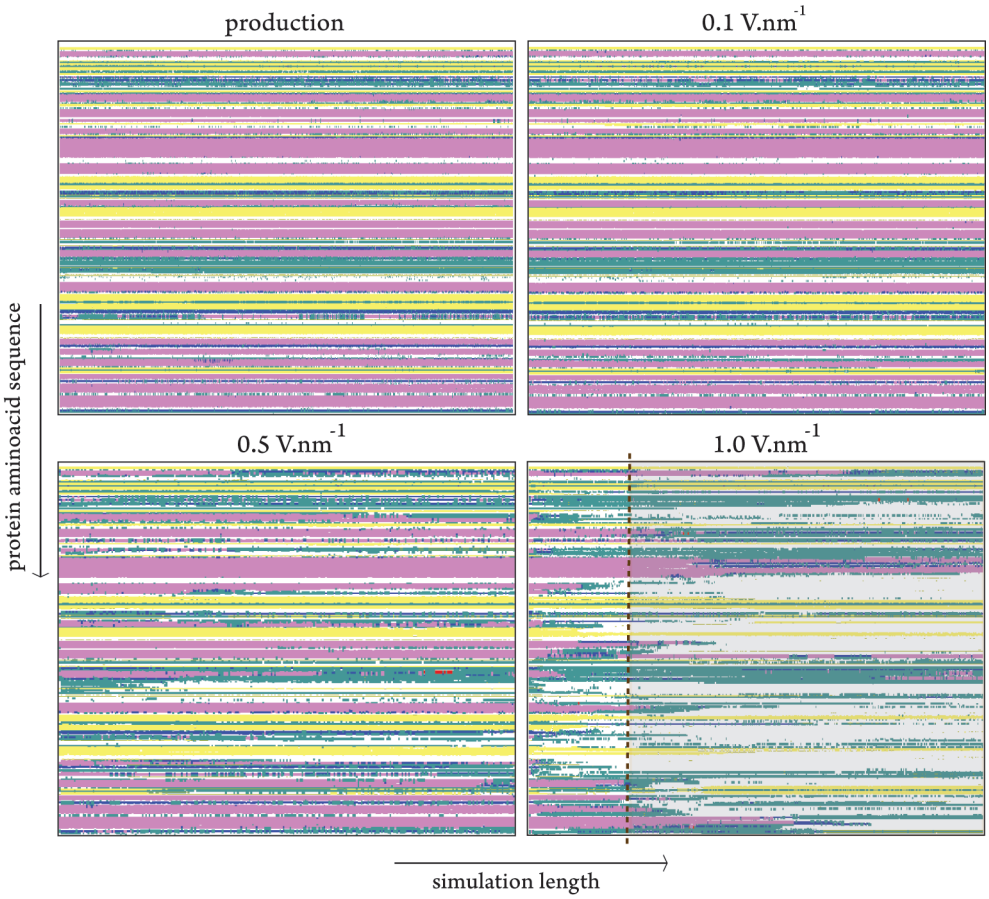

Figure S3. Graphical representation of secondary structure analysis, for the topoisomerase III protein, relative to the production MD simulation and the 3 tested external electric fields. Different colours represent the different secondary structure motifs: yellow – extended conformation (β-sheets), purple – α-helixes, green - turn, red – pi helix, blue – 3-10 helix, and white - coil. The secondary analysis should not be considered for the 1.0 $V.nm^{-1}$ past the dashed line, due to Periodic Boundary Condition (PBC) violation, since the protein extended past the periodic simulation box, and started to interact with the periodic image of itself (artefact).

As can be observed by the analysis of Figures S1 and S3, both the RMSD and radius of gyration show a relevant increase in both settings, and relevant distortion of the secondary structure elements are observed for the 0.5 $V.nm^{-1}$ and the 1.0 $V.nm^{-1}$ field strength (more pronounced for the latter).

Thioredoxin:

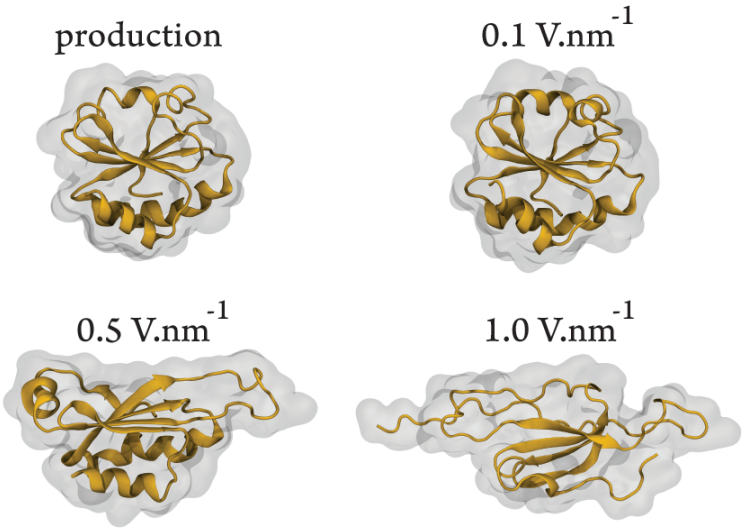

Figure S4. Molecular representations' end results for each of the MD simulations conducted for thioredoxin – at the end of 10 ns simulation in each situation. Only the protein structure and ssDNA chain are depicted. Water molecules are omitted for clarity.

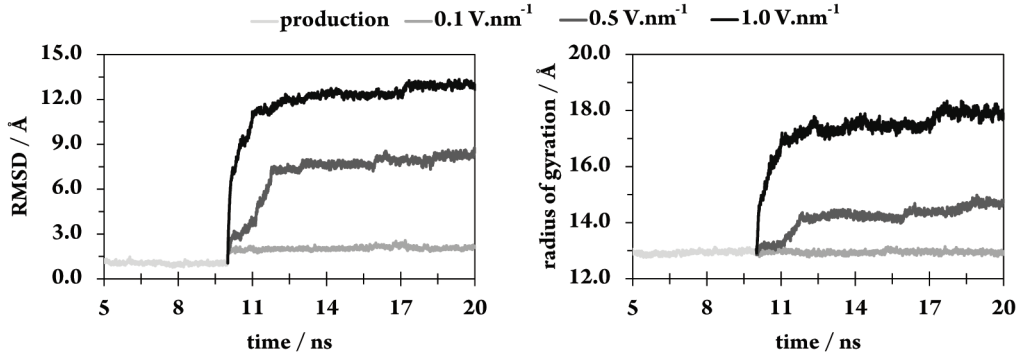

Figure S5. Root-mean-square deviation (RMSD), presented as a measure of protein stability, relative to the initial x-ray crystallography structure, minimized and equilibrated; and radius of gyration, as a measure of the proteins' compactness (if a protein is stably folded it will likely maintain a steady behaviour of the radius of gyration).

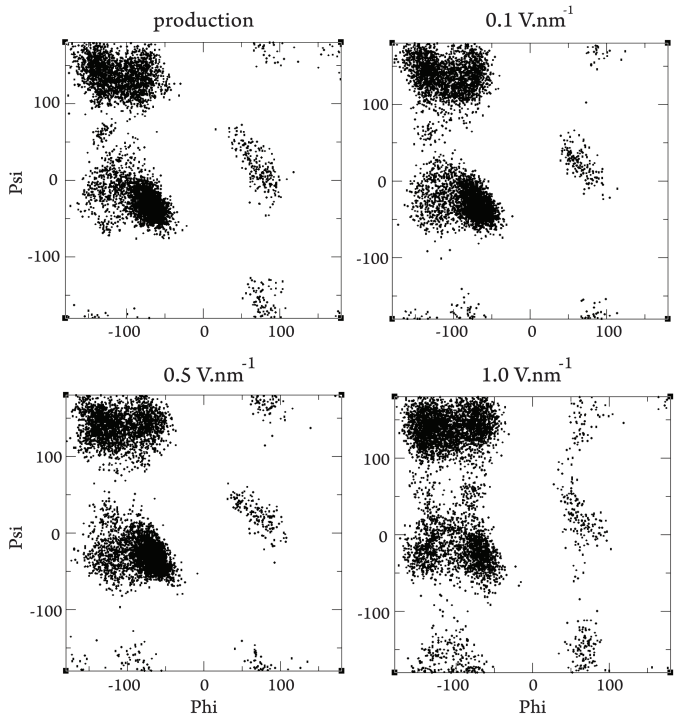

Figure S6. Ramachandran plots for the thioredoxin protein, relative to the production MD simulation and the 3 tested external electric fields.

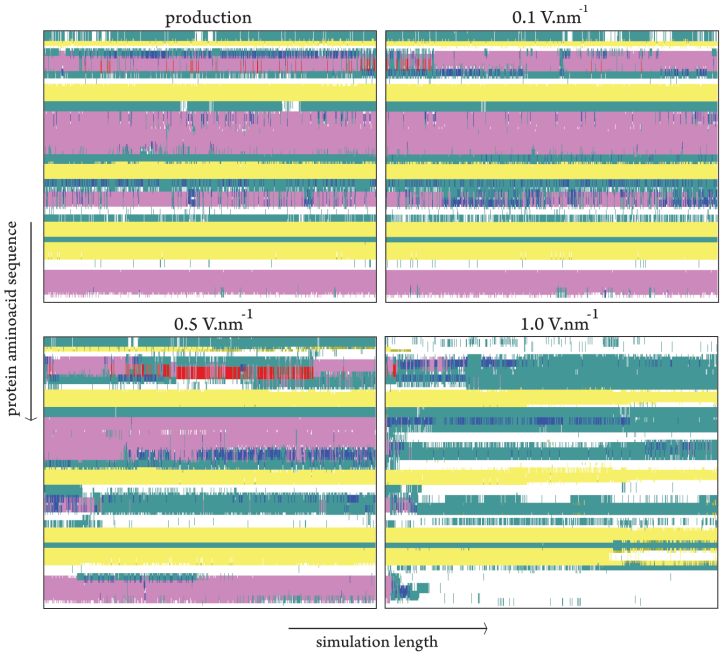

Figure S7. Graphical representation of secondary structure analysis, for the thioredoxin protein, relative to the production MD simulation and the 3 tested external electric fields. Different colours represent the different secondary structure motifs: yellow – extended conformation (β-sheets), purple – α-helixes, green - turn, red – pi helix, blue – 3-10 helix, and white - coil.

ii. The harmonic approximation.

We consider a biomolecule consisting of *N* units (which could be atoms in an atomistic model or groups of atoms in the case of a coarse grained force field). The mass of each atom is M_l (l=N). The Hamiltonian (energy) of the system is given by

$$H_B = \sum_{l=1}^{N} \frac{P_l^2}{2M_l} + V_{eff}\left(\vec{R}_1, \ldots, \vec{R}_N\right), \tag{S1}$$

where the first term refers to the kinetic energy, and the second term to the potential (usually called force field). Here, P_l and $\vec{R}_l$ refer to the momentum and position of the atom l. The potential $V_{eff}\left(\vec{R}_1, \ldots, \vec{R}_N\right)$ can be a complicated function of the positions. It includes covalent, ionic and hydrogen bonds as well as van der Waals forces. If the biomolecule is in its native state (global minimum) or in a local minimum of V_{eff} the positions of the atoms (units) acquire equilibrium values $R_l^{(0)}$, and one can denote the equilibrium configuration by $R^{(0)} = \left(\vec{R}_1^{(0)}, \ldots, \vec{R}_N^{(0)}\right)$.

As a function of the distance $d = \left|\vec{R}_l - \vec{R}_n\right|$ between two atoms l and n connected by any type of chemical bond, either of covalent, ionic, hydrogen- or van der Waals type, the potential V_{eff} will be just the potential energy of the chemical bond V_{bond} and will mostly have a form as that sketched in Fig 2(a). For very short distances the potential is repulsive, then it becomes negative and shows a minimum at the equilibrium distance $d^{(0)} = \left|R_l^{(0)} - R_n^{(0)}\right|$, and finally it goes to zero for large distances (dissociation limit). The dissociation limit corresponds to the energy at which the motion of the atoms involved in the bond is no longer oscillatory but unbound. Therefore, bond breaking (dissociation) occurs if the bond, initially at the equilibrium distance $d^{(0)}$, absorbs an amount of energy equal to V_0 [see Fig. S8(a)]. It is important to point out that near the equilibrium distance the potential can generally be fitted by a parabola [see Fig. S8(b)]. This means that the bond between the atoms l and n will oscillate with a well-defined frequency near equilibrium. However, this so-called harmonic approximation[11] is only valid for small oscillations. If, for some reasons, the bond performs large amplitude oscillations, then the harmonic potential can considerably deviate from V_{bond} and the harmonic approximation breaks down [see Fig. S8(b)].

Since, as mentioned before, V_{eff} is a complicated function of the positions of all atoms, the oscillations of the distance between l and n will depend on the oscillations of other bonds. This means that, even for small oscillations the bonds will be coupled.

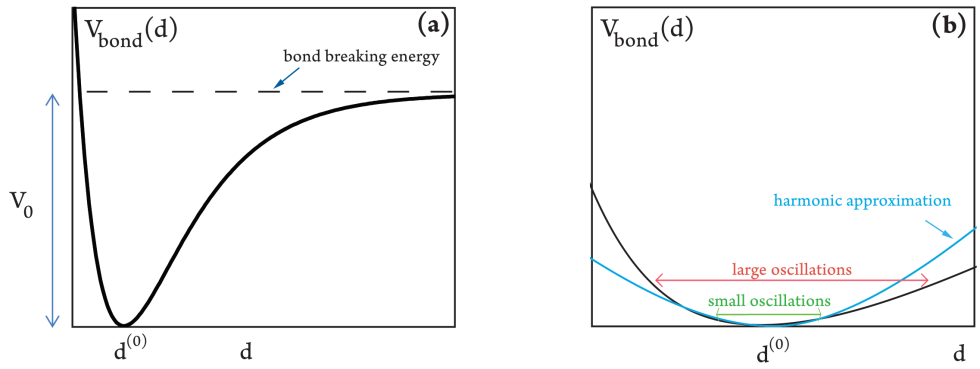

Figure S8. (a) schematic plot of the behavior of the potential energy of a chemical bond (covalent, ionic, hydrogen or van der Waals) as a function of the distance d between the bonded atoms. The dashed line indicates the dissociation limit leading to bond breaking. If the bond absorbs an amount of energy equal to V_0, bond breaking occurs. (b) Comparison between the real potential curve and the parabolic potential calculated using the harmonic approximation. For small oscillations the harmonic approximation works well. For large oscillations the harmonic potential deviates from the real potential.

In the harmonic approximation, one can formally map the biomolecule to a system of 3*N* coupled harmonic oscillators by expanding the potential up to a quadratic order in the atomic displacements. If we write $\vec{R}_l = \vec{R}_l^{(0)} + \vec{u}_l$, where $\vec{u}_l$ is a vector representing the displacement of atom l from its equilibrium position, then V_{eff} reads

$$V_{eff}(\vec{R}_1, \cdots, \vec{R}_N) = V_{eff}\left(\vec{R}_1^{(0)}, \cdots, \vec{R}_N^{(0)}\right) + \frac{1}{2}\sum_{l,n,\alpha,\beta} \frac{\partial^2 V_{eff}}{\partial R_{l\alpha}\partial R_{n\beta}}\bigg|_{R^{(0)}} u_{l\alpha}u_{n\beta} + \cdots, \qquad \text{(S2)}$$

where the first term is just $-V_0$. α and β label the three projections of the vectors in the x,y and z directions in space (α,β=1,2,3). The force constants

$$k_{l\alpha n\beta} = \frac{\partial^2 V_{eff}}{\partial R_{l\alpha}\partial R_{n\beta}}\bigg|_{R^{(0)}}, \qquad \text{(S3)}$$

constitute the elements of the 3N x 3N Hessian matrix $\boldsymbol{K}$.

The equations of motion for the displacements can be written in matrix form as

$$\boldsymbol{M}\cdot\ddot{\vec{U}}(t) = -\boldsymbol{K}\cdot\vec{U}(t) \qquad (S4)$$

where **M** is a diagonal 3*N* x 3*N* matrix containing the atomic masses, given by $M_{ln\alpha\beta} = \delta_{ln}\delta_{\alpha\beta}M_l$, and the vector $\vec{U} = (u_{1x}, \ldots, u_{l\alpha}, \ldots, u_{Nz})$ contains the displacements of all atoms. Equation (S4) can be easily rewritten as

$$\boldsymbol{M}^{\frac{1}{2}}\cdot\ddot{\vec{U}}(t) = -\boldsymbol{M}^{-\frac{1}{2}}\cdot\boldsymbol{K}\cdot\boldsymbol{M}^{-\frac{1}{2}}\cdot\boldsymbol{M}^{\frac{1}{2}}\cdot\vec{U}(t) = \boldsymbol{D}\cdot\boldsymbol{M}^{\frac{1}{2}}\cdot\vec{U}(t). \qquad \text{(S5)}$$

Since the mass weighted Hessian (or dynamical) matrix $\boldsymbol{D}$ is real, symmetric and positive definite, and therefore it has only real and positive eigenvalues $\{\omega_1^2, \ldots, \omega_{3N}^2\}$, which are the squares of the eigenfrequencies of the system. Diagonalization is achieved by applying a unitary matrix $\boldsymbol{C}$ such that

$$\boldsymbol{C}\cdot\boldsymbol{D}\cdot\boldsymbol{C}^{+}=\boldsymbol{\Omega}=\begin{pmatrix}\omega_1^2 & 0 & \cdots & 0\\ 0 & \omega_2^2 & 0 & \vdots\\ \vdots & 0 & \ddots & 0\\ 0 & \cdots & 0 & \omega_{3N}^2\end{pmatrix}. \qquad \text{(S6)}$$

Three of the eigenvalues are zero, since they correspond to translations of the molecule in space. If we place our coordinate system at the center of mass of the molecule, both translations and rotations of the molecule as a whole are no longer considered and the total number of degrees of freedom reduces to 3*N*-6.

Multiplying Eq. (S6) by $\boldsymbol{C}$ and using the property $\boldsymbol{C}\cdot\boldsymbol{C}^{+}=\boldsymbol{C}^{+}\cdot\boldsymbol{C}=\mathbf{1}$, where $\mathbf{1}$ is the identity matrix, one can completely decouple the equations of motions as

$$\boldsymbol{C}\cdot\boldsymbol{M}^{\frac{1}{2}}\cdot\ddot{\vec{U}}(t)=\ddot{\vec{A}}(t)=-\boldsymbol{\Omega}\,\vec{A}(t), \qquad \text{(S7)}$$

where the vector $\vec{A}(t)=\left(a_1(t),a_2(t),\ldots,a_{3N-6}(t)\right)$ contains the displacements written on the basis of the eigenvectors $\vec{G}_j$ of the dynamical matrix. Note that we now consider only 3*N*-6 degrees of freedom.

Thus, in the harmonic approximation one can describe the biomolecule as a system of *3N* coupled harmonic oscillators (see Figure S9). This means that the equations of motion for the oscillators will be coupled, independently of whether one treats the oscillators classically or quantum mechanically.

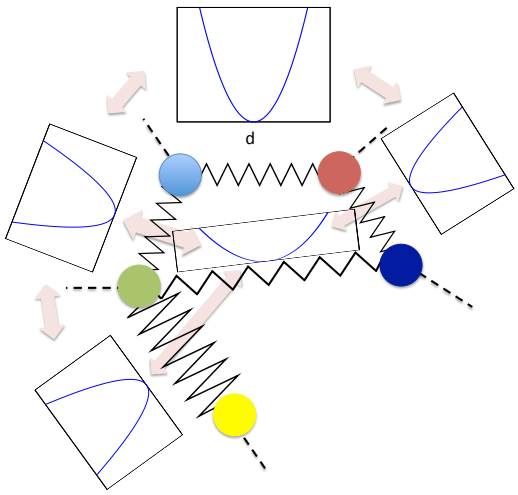

Figure S9: sketch of the mapping of the force field of a biomolecule onto a system of coupled harmonic oscillators. Only a part of an ideal molecule is shown. Dashed lines indicate where bonds between the selected part and the rest of the biomolecule might be present. The colored circles represent atoms within the biomolecule. Through the mapping the different bonds are replaced by springs. The harmonic potentials corresponding to the different springs are also shown**.** The couplings between oscillators are represented by arrows. The magnitudes of the spring constants might be different, since the springs connect atoms of different types, and the second derivative of the potential with respect to the atomic coordinates depends on the properties of the specific atoms (see Equation S3).

iii. Frequency analyses for topoisomerase. A detailed description of the normal modes and determined frequencies is discussed subsequently for Model 1 and Model 2. Only gas-phase results will be discussed.

Model 1. The first model (Figure S10) was truncated at the C_β carbon and at the 2nd heavy atom counting from the C_α carbon of the catalytic tyrosine residue (in each direction of the backbone) – comprising the carbonyl oxygen atom and the amide nitrogen atom. The determined frequencies (presented in wavenumbers) and normal modes are within the expected ranges, and a detailed description is given below.

Frequencies above 2500 cm^{-1}: This region is associated with the stretching modes for OH, NH, and CH (in this descending order); and also with the characteristic amide A vibration. For our study, these vibrational modes are not of a relevant nature, since we are interested in breaking more relevant peptide bonds (such as CN, CO, or CC). The highest frequency for this region is associated with the antisymmetric bond stretching between the N12-H14 and the N12-H15 atoms; the lowest frequency with the C1-H13 stretching. These values are comprised between 2885-3525 cm^{-1}.

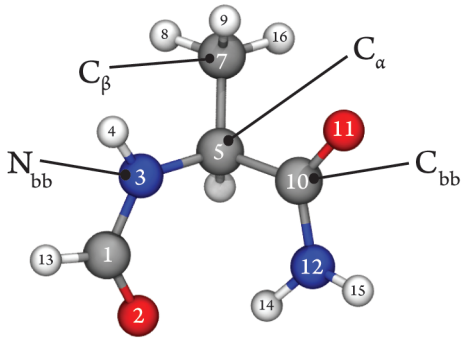

Figure S10. Model 1 representation. Numbered in white are the heavy atoms in which the analysis of vibrational frequencies was focused – 1, 5, 7 and 10 are carbon atoms (colored grey); 2 and 11 are oxygen atoms (colored red); 3 and 12 are nitrogen atoms (colored blue).

Frequencies 1200-1800 cm^{-1}: This is our main region of interest, since relevant peptide group vibrations are within this range of values: CO stretch, CN stretch, and CH_3 in plane bending. This region is also associated with the NH_3^+ and CH deformations. The higher frequencies observed – 1741, 1709 cm^{-1}, are associated with CO stretching events, of both CO groups; and concomitantly the antisymmetrical CN stretching (Figure S13 - panel 1). The other frequencies (ranging from 1260-1592 cm^{-1}) are associated with CC and CN stretching modes (Figure S11 – panel 2), with C_α-C_β stretching – 1386 cm^{-1}, and with some bending contaminations associated with hydrogen atoms bound to heavy atoms.

Frequencies below 1200 cm^{-1}: in this region the categorisation of vibrational modes is tentative due to the wide range of skeletal vibrations that takes place in this region for each normal mode of vibration. But relevant stretching is found for the Cα-C_β bond at 1074 cm^{-1}.

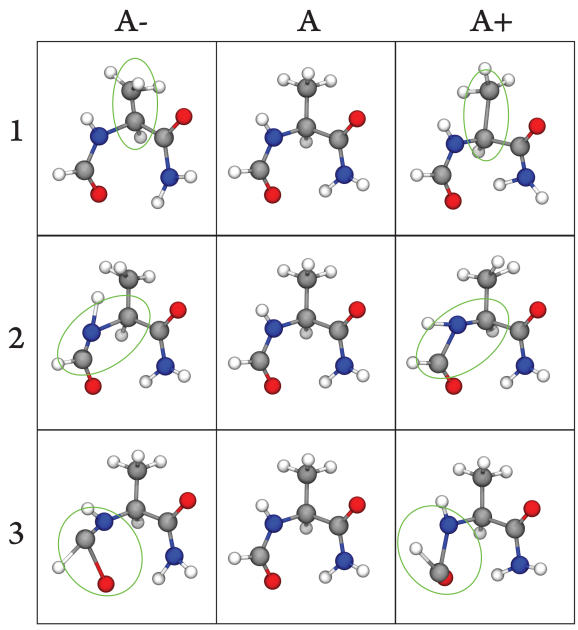

Figure S11. Model 1 selected frequencies. Panels A-, A, and A+, represent the displacement (ranging from the maximum to the minimum amplitude) of the selected vibration mode; panel 1 represents the vibration at 1074 cm^{-1}, panel 2 at 1481 cm^{-1}, and panel 3 at 1709 cm^{-1}. The principal vibrational modes are surrounded by ellipses.

Model 2. The frequencies of Model 1 were then evaluated for Model 2 (Figure S12). Model 2 represents 3 amino acid residues, comprising the catalytically relevant tyrosine residue and the two immediate residues.

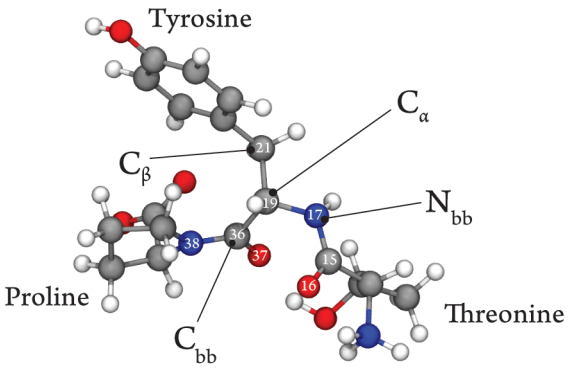

Figure S12. Representation of Model 2, 27 heavy atoms altogether. Numbered in white are the heavy atoms in which the analysis of vibrational frequencies was focused.

Frequencies above 2500 cm^{-1}: In this model, the highest frequencies are associated with OH and NH stretching events - 3319-3604 cm^{-1}. The normal modes for CH stretching are followed ranging from 2907-3088 cm^{-1}.

Frequencies 1200-1800 cm^{-1}: The higher frequencies observed are associated with the antisymmetric COO^- stretching and two C(N)=O antisymmetric stretching (1708, 1700, and 1625 cm^{-1}), and the angular deformation of the NH_3^+ group (1658 and 1512 cm^{-1}). The other frequencies are associated with the C-C and C-N stretching modes, with some angle bending contaminations associated with hydrogen atoms bound to heavy atoms, and with the

deformation of tyrosine's aromatic ring. One of the most promising results is the frequency difference of two OCN antisymmetric vibrational modes – O16C15N17 and O37C36N38 (Figure S13 – panel 1 and 2). The same is observed for other vibrational frequencies, *e.g.* the CN stretching – N17C15 and N38C36.

Frequencies below 1200 cm^{-1}: again, categorisation of this region's vibrational modes is tentative due to the wide range of skeletal vibrations that takes place in this region in each normal mode of vibration. However, two possible frequencies of interest arise from this region: one at 1040 cm^{-1} (C_{α}-N_{bb} stretching – C19N17) and another at the 1007 cm^{-1} (relevant C_{α}-C_{β} stretching – C19C21) (Figure S13 – panel 5).

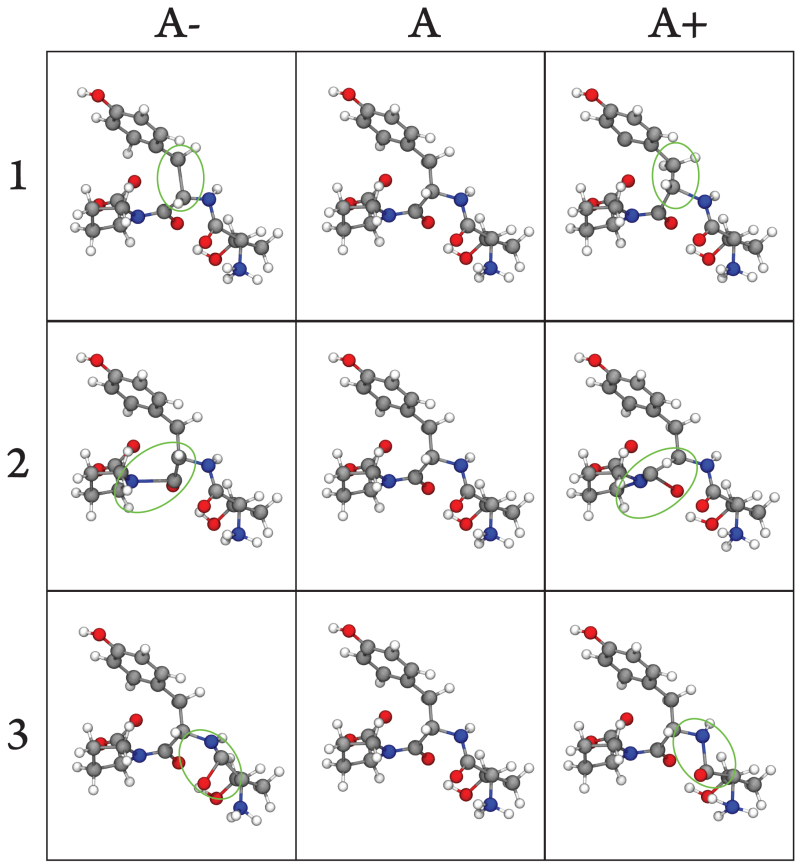

Figure S13. Model 2 selected frequencies. Panels A-, A, and A+, represent the range from the maximum to the minimum displacement of the selected vibrational mode; panel 1 represents the vibration at 1007 cm^{-1}, panel 2 at 1625 cm^{-1}, and panel 3 at 1700 cm^{-1}. The principal vibrational modes are surrounded by ellipses.

iv Application of optimal control based on analytical and numerical calculations to induce energy concentration on particular bonds of model 1 of topoisomerase

Further Results on the Analytical Optimization

Below we show results for the analytical optimization of pulse trains 1 and 2 (see Figure 5 of the article). In Fig. S14 we show the target structure for the action of pulse train 1. In Fig. S15 we show results for the tailored pulse train 1 and its effect on all bond lengths of Model 1 of topoisomerase (see article). The tailored pulse train is represented by the vertical lines (amplitudes are not shown). Remarkably, the pulse induces a huge elongation of the bond between atoms 3 and 5 at the desired time. Since the elongation reaches 100%, the amount of energy concentrated on the bond is enough to produce bond breaking. Finally, Figures S16 and S17 show results of the analytical optimization for the action of pulse train 2 on Model 1 of topoisomerase.

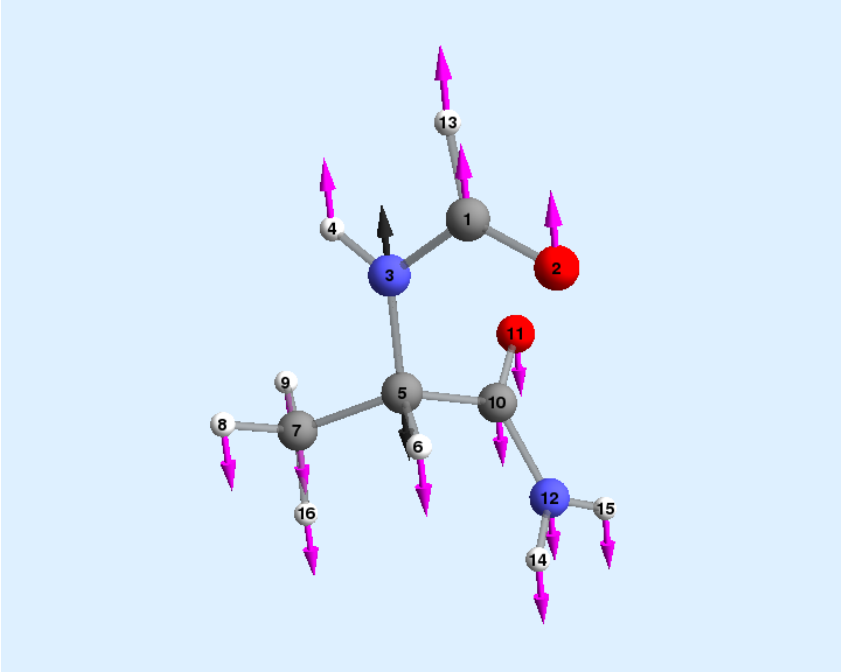

Figure S14: representation of the vector $\vec{U}_{target}$ to be reached by model 1 after excitation through the tailored pulse train 1. Arrows indicate direction of atomic displacements. Note that we set the displacements of atoms 3 and 5 opposite to each other and large. The displacements of the other atoms are chosen in such a way that the molecule can effectively be broken at the desired position. This means, the displacements of atoms 1,2,4 and 13 are parallel to that of atom 3, while the other atoms move in direction of atom 5. In this way, only the bond between atoms 3 and 5 will be strongly stretched, while the other bonds will not suffer considerable stress.

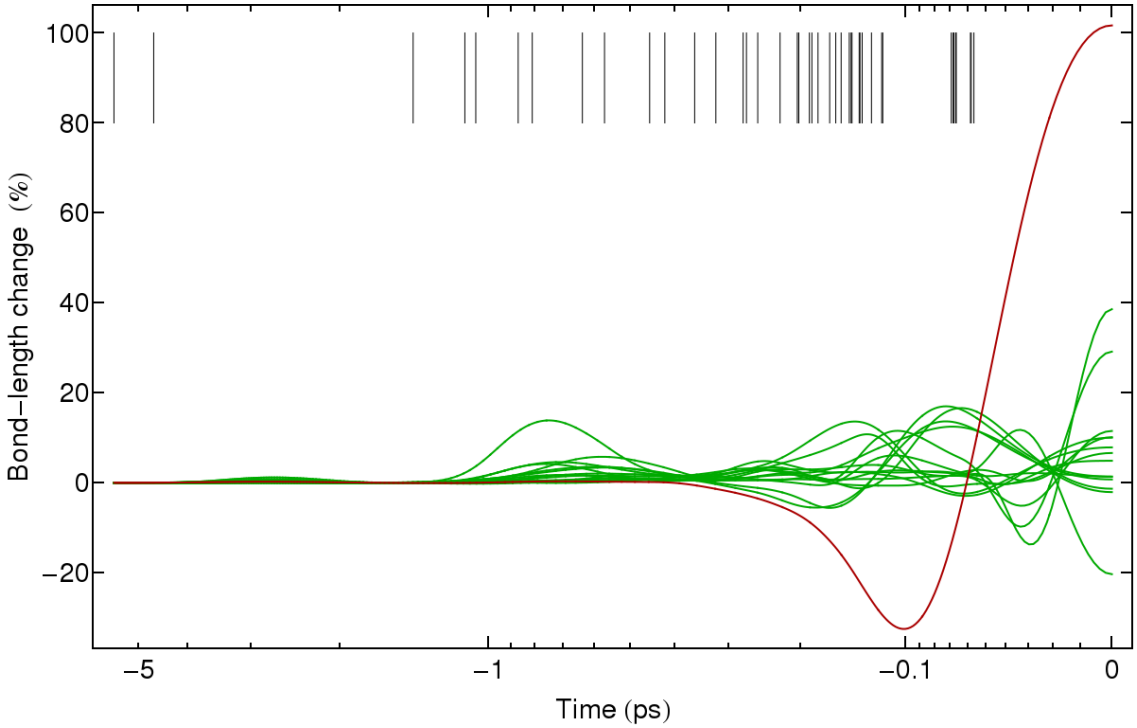

Figure S15: effect of the tailored pulse train 1 on the time dependence of the 15 bonds of model 1 of topoisomerase. The black vertical lines represent the times at which the different pulses of the designed train are switched on. The red line shows the behavior of the bond length between atoms 3 and 5. Note that this bond is elongated by 100% at the final time, whereas all other bonds (green lines) are at most compressed/elongated by less than 40%.

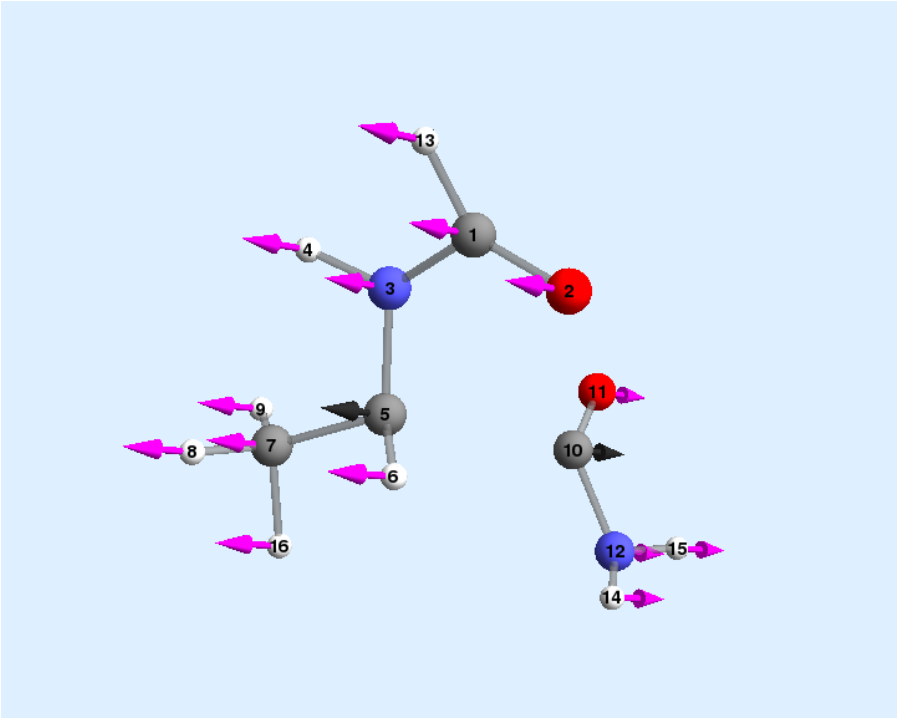

Figure S16: representation of the vector $\vec{U}_{target}$ to be reached by model 1 after excitation through the tailored pulse train 2. Arrows indicate direction of atomic displacements. Note that we set the displacements of atoms 5 and 10 opposite to each other and large. The displacements of the other atoms are chosen in such a way that the molecule can effectively be broken at the desired position. In this way, only the bond between atoms 5 and 10 will be strongly stretched, while the other bonds will not suffer considerable stress.

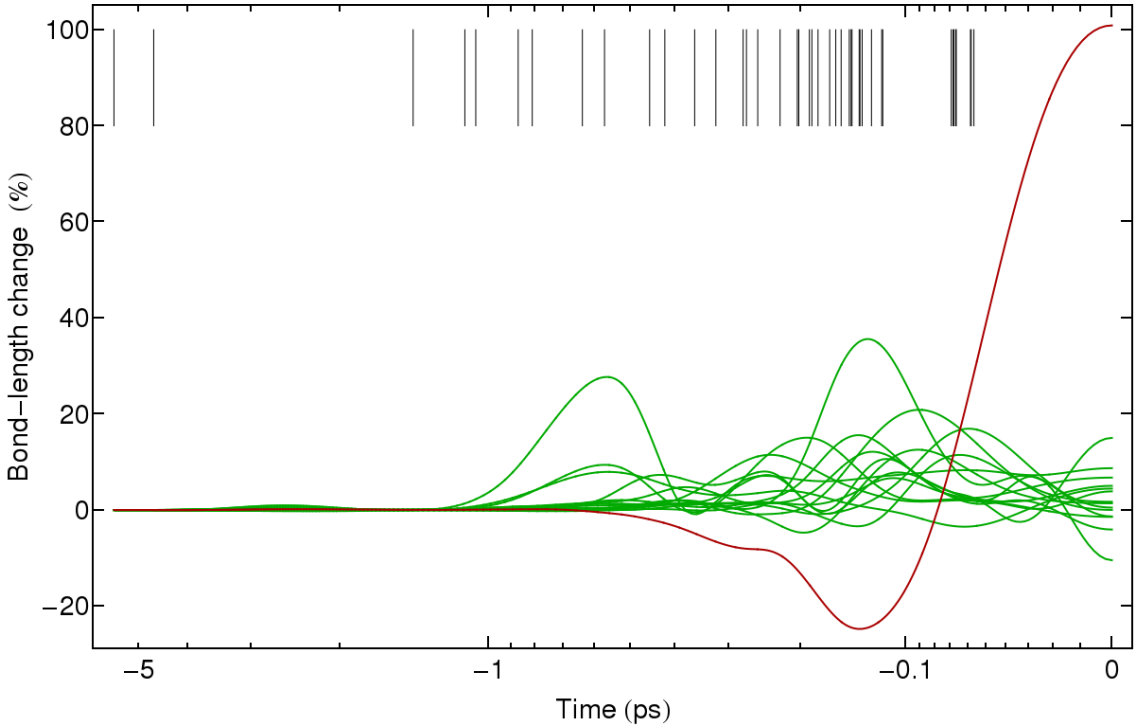

Figure S17: effect of the tailored pulse train 1 on the time dependence of the 15 bonds of model 1 of topoisomerase. The black vertical lines represent the times at which the different pulses of the designed train are switched on. The red line shows the behavior of the bond length between atoms 5 and 10. Note that this bond is elongated by 100% at the final time, whereas all other bonds (green lines) are at most compressed/elongated by less than 35%.

3. REFERENCES

[1] Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **1983,** *79* (2), 926-935.

[2] Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C., Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Structure, Function, and Bioinformatics* **2006,** *65* (3), 712-725.

[3] (a) Bas, D. C.; Rogers, D. M.; Jensen, J. H., Very fast prediction and rationalization of pKa values for protein–ligand complexes. *Proteins: Structure, Function, and Bioinformatics* **2008,** *73* (3), 765-783; (b) Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H., PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. *Journal of Chemical Theory and Computation* **2011,** *7* (2), 525-537; (c) Li, H.; Robertson, A. D.; Jensen, J. H., Very fast empirical

prediction and rationalization of protein pKa values. *Proteins: Structure, Function, and Bioinformatics* **2005,** *61* (4), 704-721; (d) Søndergaard, C. R.; Olsson, M. H. M.; Rostkowski, M.; Jensen, J. H., Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization of pKa Values. *Journal of Chemical Theory and Computation* **2011,** *7* (7), 2284-2295.

[4] Hockney, R. W.; Goel, S. P.; Eastwood, J. W., Quiet high-resolution computer models of a plasma. *Journal of Computational Physics* **1974,** *14* (2), 148-158.

[5] Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M., LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997,** *18* (12), 1463-1472.

[6] (a) Darden, T.; York, D.; Pedersen, L., Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics* **1993,** *98* (12), 10089-10092; (b) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G., A smooth particle mesh Ewald method. *The Journal of Chemical Physics* **1995,** *103* (19), 8577-8593.

[7] Bussi, G.; Donadio, D.; Parrinello, M., Canonical sampling through velocity rescaling. *The Journal of Chemical Physics* **2007,** *126* (1).

[8] (a) Parrinello, M.; Rahman, A., Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **1981,** *52* (12), 7182-7190; (b) Nosé, S.; Klein, M. L., Constant pressure molecular dynamics for molecular systems. *Mol. Phys.* **1983,** *50* (5), 1055-1076.

[9] (a) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R., GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* **1995,** *91* (1–3), 43-56; (b) D. van der Spoel; E. Lindahl; B. Hess; team, a. t. G. d., GROMACS User Manual version 4.6.1. www.gromacs.org, 2013; (c) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C., GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **2005,** *26* (16), 1701-1718.

[10] Humphrey, W.; Dalke, A.; Schulten, K., VMD: Visual molecular dynamics. *J. Mol. Graphics* **1996,** *14* (1), 33-38.

[11] Maradudin; AA, Montroll, EW, Weiss, GH and Ipatova, IP, Theory of lattice dynamics in the harmonic approximation, in "Solid State Physics ",**1971** Academic Press, *Second Edition,* New York.