THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Structure and Properties of DNA in Apolar Solvents

Annalisa Arcella,[†,‡] Guillem Portella,[†,‡,§] Rosana Collepardo-Guevara,[§] Debayan Chakraborty,[§] David J. Wales,[§] and Modesto Orozco*[,†,‡,∥]

[†]Institute for Research in Biomedicine (IRB Barcelona), 08028 Barcelona, Spain

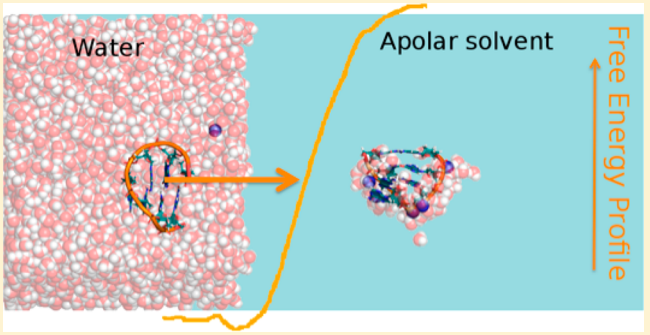
[‡]BSC-CRG-IRB Joint Program in Computational Biology, Institute for Research in Biomedicine, 08028 Barcelona, Spain

[§]Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom

[∥]Department of Biochemistry and Molecular Biology, University of Barcelona, 08007 Barcelona, Spain

S *Supporting Information*

ABSTRACT: The study of nucleic acids in low-polarity environments paves the way for novel biotechnological applications of DNA. Here, we use a repertoire of atomistic molecular simulation tools to study the nature of DNA when placed in a highly apolar environment and when transferred from aqueous to apolar solvent. Our results show that DNA becomes stiffer in apolar solvents and suggest that highly negatively charged states, which are the most prevalent in water, are strongly disfavored in apolar solvents and neutral states with conformations not far from the aqueous ones are the dominant forms. Transfer from water to an apolar solvent such as CCl_4 is unlikely to occur, but our results suggest that if forced, the DNA would migrate surrounded by a small shell of water (the higher the DNA charge, the larger the number of water molecules in this shell). Even the neutral form (predicted to be the dominant one in apolar solvents) would surround itself by a small number of highly stable water molecules when moved from water to a highly apolar environment. Neutralization of DNA charges seems a crucial requirement for transfer of DNA to apolar media, and the most likely mechanism to achieve good transfer properties.



INTRODUCTION

DNA is a highly polar molecule, evolved to be stable in high-dielectric environments, such as aqueous solution, which is the main environment considered in most experimental and theoretical studies. However, understanding the properties of DNA in low-polarity environments is crucial for many nanotechnological applications of DNA[1,2] and is instrumental in understanding how nucleic acids acting as antigene or antisense drugs pass through the highly hydrophobic cellular membrane, and how DNA can be transported in the body by means of liposome carriers.[3−7]

Water and counterions are crucial to screen the electrostatic repulsion among charged phosphates and also favor the apolar stacking of bases. Accordingly, the large impact of solvent modification on the properties of DNA is not surprising. For example, a subtle change in the neutralizing cation can lead to a dramatic conformational change[8,9] or even to a complete alteration in the sequence-stability rules of the duplex.[10] Addition of ethanol to an aqueous solution induces drastic changes in the duplex structure, forcing a B → A transition. Some organic solvents, such as methanol, formamide, pyridine, or dimethyl sulfoxide,[11−13] induce unfolding and strand separation, or formation of toroid-like conformations,[14−16] while others, such as glycerol[11] or room-temperature ionic liquids,[17] maintain the duplex structure. More surprisingly, a DNA duplex does not lose its structure completely, and the two strands remain bound when DNA is transferred from aqueous solution to the gas phase.[18−20] It is clear that changes in the balance between stabilizing and destabilizing terms controlling DNA structure due to alterations in the solvent are complex and still not well understood.

Experimental studies in apolar solvents are challenging because of the difficulty of transferring DNA from water to such media. This problem has motivated several theoretical studies, most of them using molecular dynamics (MD) simulations. For example, Cui et al.[21] have examined the structure and stability of DNA in octane, correlating the observed changes with those required for helicase activity. Lin et al.[1] explored the stability of decorated DNAs in interfaces between DNA and chloroform and have previously considered the reaction of DNA to high concentrations of pyridine.[22] Finally, Khalid et al.[23] used coarse-grained models to study the insertion of duplex DNA into a lipid bilayer, finding a significant associated free energy barrier, even for small duplexes.

Received: April 18, 2014
Revised: June 25, 2014
Published: June 26, 2014

   dx.doi.org/10.1021/jp503816r | *J. Phys. Chem. B* 2014, 118, 8540−8548

Here we use large-scale atomistic MD simulations and discrete path sampling to investigate the conformations sampled by a short DNA duplex in carbon tetrachloride (CCl_4), a medium comparable to the interior of a biological membrane. We have also explored the (free) energy and structural differences associated with the change of phase of a duplex DNA from water to CCl_4. Finally we analyzed for the first time the possibility that DNA acts as a proton donor/acceptor when it crosses from a polar to an apolar medium. Our results suggest that DNA, at least a short fragment, can maintain its duplex structure for significant periods of time in an apolar solvent and, surprisingly, appears more rigid in such media than when immersed in aqueous solution. Transfer from aqueous to apolar solvents is, obviously, strongly disfavored, especially if the DNA remains fully charged. Typically, the transferred DNA is not fully dehydrated but maintains (irrespective of its formal charge) a significant amount of water around it. Calculations suggest that a neutral DNA state with a compact structure (not too different from that in aqueous solution) is the most prevalent form of DNA in apolar solvent.

METHODS

We considered a small oligonucleotide (d(GCGAAGC)) as a model system for all our calculations. In aqueous solution, this small DNA forms a stable hairpin with a GAA triloop and a two-d(C·G)-pair stem, organized in a canonical B-type duplex, as confirmed by NMR.[24] Experiments and theory have shown that this oligonucleotide folds on a (multi)microsecond time scale[25−27] in water, and should unfold on a similar time scale in a denaturant solvent such as pyridine.[22] The small size of the system accelerates calculations, favors sampling, and reduces memory effects, which might be very important for longer oligonucleotides. Previous studies have demonstrated the ability of current force field and MD protocols to simulate this structure properly in a variety of environments.[27−29]

Equilibrium Molecular Dynamics Simulations. The 7-mer DNA hairpin was first simulated in aqueous solution as a reference for the "native trajectory". The NMR structure[24] was neutralized by adding six Na^+ ions (placed at optimal classical molecular interaction potential (CMIP) positions[30]) and solvated by 7200 water molecules, defining an octahedral box. The system was optimized, thermalized, and preequilibrated using standard procedures with triplicated simulation windows,[27,31] and equilibrated for 100 ns prior to 1 ms of production trajectory at constant temperature and pressure (T = 300 K; P = 1 atm). Long-range electrostatic corrections were represented by means of periodic boundary conditions (PBC) and the particle mesh Ewald (PME) correction with default parameters.[32]

Starting from the aqueous solution equilibrium structure, we performed unbiased MD simulations in CCl_4 considering the canonical DNA protonation state (Q = −6) as well as a fully neutralized system, in which all of the phosphates were protonated (Q = 0) following the same procedure used previously for the study of gas-phase structures.[32] The simulation cell contained one DNA molecule and around 4288 CCl_4 molecules. The optimization, thermalization, and equilibration procedures were the same as those used in the aqueous simulations. Trajectories were also collected for 1 ms at constant temperature and pressure (T = 300 K; P = 1 atm) using PBC−PME to account for long-range effects. Note that the CCl_4-simulation systems do not contain counterions, which means that Ewald's plasma is responsible for charge balance in these simulations. To discard possible Ewald artifacts in this case, we repeated the calculations for the Q = −6 system using a reaction field correction with a 1.1 nm cutoff to define the boundary between the atomistic system and the exterior dielectric continuum.

Replica Exchange Molecular Dynamics Simulations. To explore the conformational space of the hairpin in CCl_4, we performed replica exchange replica exchange molecular dynamics (RexMD)[33] simulations for the two charge states (Q = −6 and Q = 0) starting from (i) a fully extended conformation and (ii) the aqueous-equilibrated structure. As an additional test, we extended the RExMD simulations to a minimum-hydration hairpin (∼50 water molecules) obtained from our simulations of the transfer of the hairpin from water to CCl_4 (see later discussion). A total of 25 temperatures was selected to span the range of 300−440 K. The distribution of temperatures was chosen to guarantee an average exchange probability around 25%. Each of the temperature replicas was explored for 150 ns (for a total simulation time of 2 × 2 × 35 × 0.150 = 15 ms), with exchanges attempted every 1,000 MD steps.

Disconnectivity Graphs. Disconnectivity graphs provide a simple and powerful way to visualize the DNA potential or free energy landscape in terms of local minima and transition states connecting them.[34,35] This approach groups minima that are mutually accessible at a given energy threshold into disjoint sets, which are connected by vertical branches. The shorter the branches linking the minima, the smaller the activation barrier needed to move from one group to the other. The stationary points (minima and transition states) required to construct the disconnectivity graph were obtained with the discrete path sampling (DPS) method,[36,37] implemented in the programs PATHSAMPLE[38] and OPTIM.[39] DPS runs were started from an initial connected path of minima and intervening transition states between two specified end points. To identify suitable end points, we used 400,000 snapshots obtained from MD trajectories at low (298 K) and high (450 K) temperatures for the Q = 0 system in explicit CCl_4 (that for which RexMD simulations were not conclusive; see later discussion). Local minimization quenched these snapshots to 687 different local minima. The lowest energy structure was chosen as one representative of the "native form", while an extended structure without any native contacts was designated as the "unfolded form".

The connections between stationary points were calculated by approximate steepest-descent paths obtained by energy minimization following infinitesimal displacements parallel and antiparallel to the eigenvector corresponding to the unique negative eigenvalue for each transition state. We define minima and transition states as stationary points on the potential energy surface with zero and one negative Hessian eigenvalues, respectively. Local minimization of minima and transition states was carried out using a modified limited-memory Broyden−Fletcher−Golgfarb−Shano (LBFGS) algorithm.[40,41] The doubly nudged[42] elastic band (DNEB) algorithm[43] was used to identify transition-state candidates, which were accurately refined using a hybrid eigenvector-following approach.[44] Once an initial connected path was found, the database was enlarged by adding all of the minima and transition states found during successive connection-making attempts for existing pairs of minima.[45] The resulting databases contain a total of around 200,000 minima and transition states.

dx.doi.org/10.1021/jp503816r | J. Phys. Chem. B 2014, 118, 8540−8548

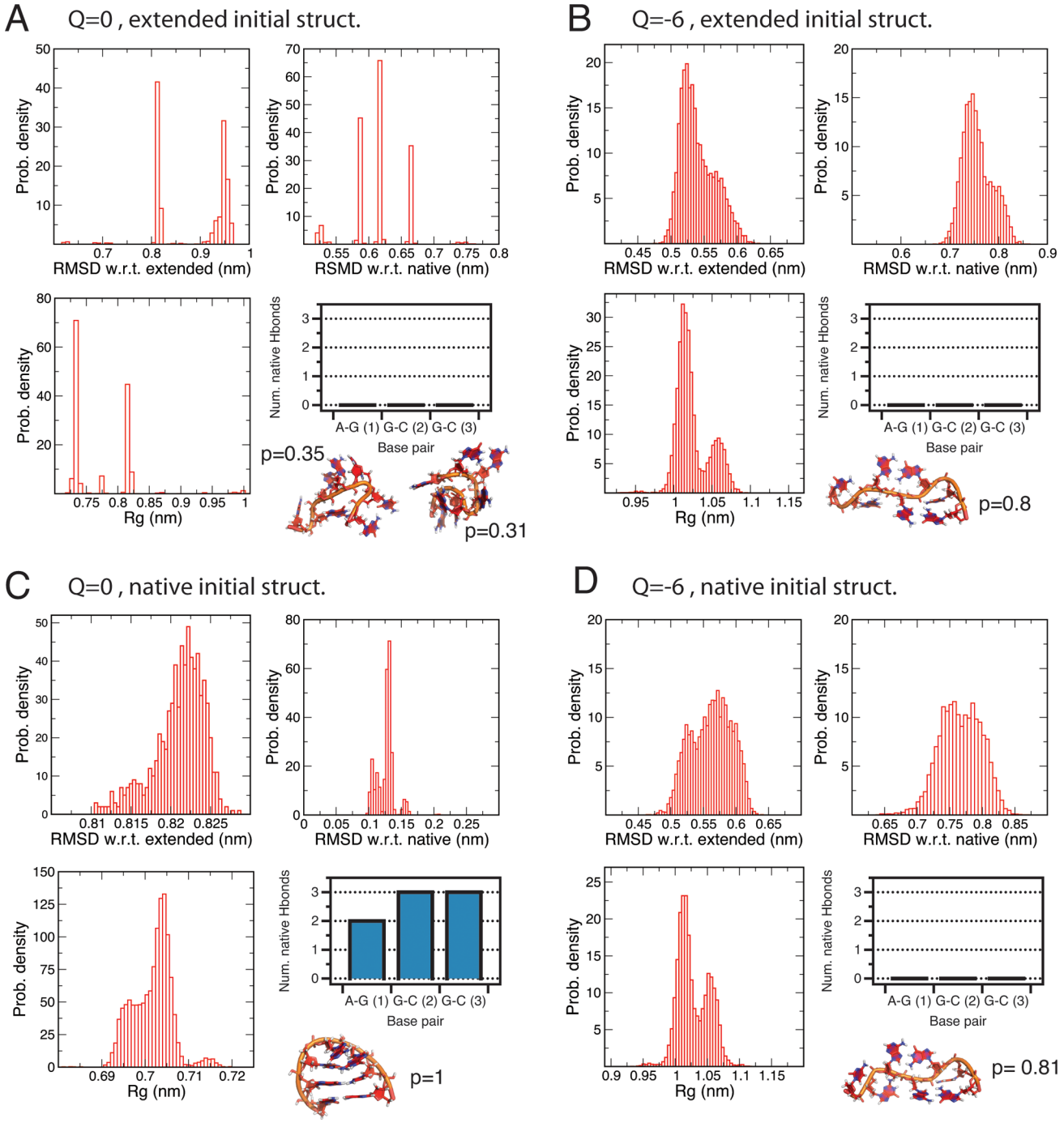


Figure 1. Summary of RexMD results for the dry hairpin in CCl_4 solution ($Q = -6$ and $Q = 0$) started from the native conformation (the folded state in water) and a fully extended structure. The results correspond to the last 50 ns of the replicas at room temperature (extracted from an aggregated simulation time of 3.75 μs; see Methods).

Finally, to construct disconnectivity graphs on the basis of free energy, we estimated the relative free energies of minima and transition states from their potential energy and normal-mode frequencies using the harmonic superposition approximation for the density of states.[46]

The energy and analytical first derivatives of the energy, required for geometry optimization, were computed through an interface for the AMBER9 package[47] available for OPTIM[39] using the parm99+parmbsc0 force field. For disconnectivity graph calculations we represented the CCl_4 solvent with the Generalized Born model implemented in AMBER[48] and without distance cutoffs for the nonbonded interactions.

Steered Molecular Dynamics and Umbrella Sampling Calculations. To obtain an atomistic representation of the transfer of the hairpin from aqueous solution to CCl_4, we performed a steered MD simulation,[49] where we moved the hairpin across the major axis of an orthorhombic box (size $5.85 \times 5.85 \times 19.4$ nm^3) containing 9,460 CCl_4 molecules (placed in the center of the long axis) and 34,935 water molecules (placed at both sides of the CCl_4 phase; see Supporting Information Figure S1). Simulations for the $Q = -6$

dx.doi.org/10.1021/jp503816r | J. Phys. Chem. B 2014, 118, 8540−8548

DNA included six Na^+ counterions (see previous discussion for neutralization protocol), while no explicit counterions were included for the $Q = 0$ simulations. The steered MD simulations either move the hairpin from one water box to the other passing through the CCl_4 phase or from the center of the CCl_4 phase to one of the neighboring water boxes. A slow steering velocity [0.05 nm/ns; with force 1,000 kJ/(mol nm^2)] was used to reduce hysteresis effects. The steered MD trajectory was then used to extract 75 starting configurations (spaced 0.2 nm along the major axis of the simulation system; see Supporting Information Figure S1) from which umbrella sampling (US) calculations were performed. Prior to the US calculations all configurations selected from the steered MD simulations were relaxed for 20 ns. For the US procedure, we used the position of the DNA hairpin in the transfer direction (z; see Supporting Information Figure S1) as the reaction coordinate, and a harmonic umbrella potential (k = 1,000 kJ/(mol nm^2)) with 75 overlapping windows, centered at consecutive values of the reaction coordinate that span the complete water→CCl_4→water transfer path. We define the central window as that centered at the center of mass (COM) of the CCl_4 box ($z = 0$ nm) and centered the first and last windows at points located at $z = -7.3$ and $z = +7.3$ nm with respect to this COM. Individual trajectories for each of the 75 US windows extend for 100 ns (for a total simulation time of 9 μs × system). The potential of mean force (PMF) was calculated using the cyclic implementation of the weighted histogram analysis method (WHAM) in GROMACS 4.5.5.[50,51] WHAM incorporates the integrated autocorrelation times (IACT) of umbrella windows. Because IACT is subject to uncertainties due to limited sampling, we have smoothed the IACT along the reaction coordinate using a Gaussian filter of width of 0.5 nm.[50] The statistical uncertainty of the PMFs was estimated using bootstrap analysis.[50]

Poisson−Boltzmann Calculations of the Phase Transfer. We used the finite difference Poisson−Boltzmann (PB) equation with a very fine grid resolution[52] as implemented in our CMIP code,[30,52] to determine the change in solvation free energy when the DNA moves from regions of low (CCl_4) and high (H_2O) dielectrics. Dielectric boundaries were defined from the placement of the interfaces between phases in the US calculations. For each of the 75 windows considered in the US calculations we selected 50 individual DNA conformations (for $Q = -6$ and $Q = 0$ states) performing then a total of 15,300 PB calculations from which an average free energy of solvation for each point of the transfer process was determined. To solve the PB equation we used a regular grid (spacing 0.05 nm) with two solvent regions (CCl_4 $\varepsilon = 2$ and H_2O $\varepsilon = 80$).

RESULTS AND DISCUSSION

DNA Hairpin in Water. The d(GCGAAGC) oligonucleotide forms a very stable DNA hairpin in aqueous solution, with a well-organized and stable triloop and a short, but stable, B-GpC DNA stem. The native NMR conformation is stabilized by eight base-pairing hydrogen bonds (H-bonds)[24]—six d(G·C) bonds plus two d(G·A) bonds—that we term "canonical H-bonds". Despite backbone fluctuations, around 98% of these canonical H-bonds are preserved in the simulation. Our long MD simulations in aqueous solvent suggest that the DNA native hairpin conformation is stable at room temperature (Supporting Information Figure S1), in agreement with previous experimental (root mean square deviation (RMSD) = 0.08 ± 0.01 nm with respect to the NMR average structure) and MD studies.[25−27]

DNA Hairpin in Dry CCl_4. Unbiased microsecond MD simulations for the DNA hairpin in CCl_4 reveal that for both $Q = -6$ and $Q = 0$ charge states the aqueous conformation can survive in CCl_4 for a significant amount of time, undergoing only small structural changes (Supporting Information Figure S1). In fact, unbiased MD simulations show RMSD fluctuations smaller for trajectories in CCl_4 than for equivalent trajectories (at the same temperature) in water (see Supporting Information Figure S2). Clearly, the ability of water to establish H-bond interactions with the DNA favors flexibility, since in aqueous solution temporarily lost intra-DNA interactions (due to thermal fluctuations) are compensated with DNA−water contacts, while this is impossible in CCl_4, or any other apolar environments.[18−20] To verify that the, perhaps surprising, microsecond-scale stability of the $Q = -6$ hairpin was not related to a PME artifact, we repeated the simulation using a reaction field to correct for distant interactions, instead of the default PBC−PME procedure. As shown in Supporting Information Figure S3 and Figure S1, reaction field and Ewald simulations give (as expected[53]) nearly equivalent results, confirming that aqueous solution structure is indeed a microsecond-stable structure of the hairpin in CCl_4, even when no neutralization of the phosphates occurs. It is, however, unclear whether or not the solution structure corresponds to the real free energy minima in CCl_4 or it is simply a metastable conformation that becomes trapped due to the stiffness of the structure. To clarify this point, we performed large-scale RexMD simulations with a wide range of temperatures and an aggregated simulation time of more than 15 μs (see Methods and Figure 1a−d). As previously described, to check for memory artifacts, we started simulations from both an "extended" and a "native-like" conformation for the two charged states of the hairpin.

Irrespective of the starting conformation, both RexMD simulations of the $Q = -6$ hairpin lead to quite extended configurations at the end of the simulations (Figure 1b).

The structures sampled have a significant number of H-bonds, but few of them are native (Figure 1a,b,d). The loop and general hairpin structure is also destroyed, and it is difficult to find any signal of the original hairpin structure in the sampled structures at room temperature (Figure 1). From these RexMD and unbiased MD simulations, we conclude that the native structure is a metastable (on the microsecond time scale) conformation for a hypothetical fully charged hairpin, but it is not the most favorable conformation for this oligonucleotide in CCl_4.

The results of the RexMD simulations are less clear for the $Q = 0$ state. By comparing the independent simulations, we conclude that the conformational space is dominated by compact structures (see Figure 1), but it is not clear whether or not these compact structures are close to the native one. Thus, simulations starting from the native state remain very close to the aqueous structure, maintaining the pattern of native H-bonds (Figure 1), while simulations starting from the extended state collapse to a variety of different compact, but nonnative, structures. Clearly, replica exchange seems to be nonoptimal to explore low-lying minima in this complex conformational equilibrium.[54,55] Thus, for the hairpin ($Q = 0$) we decided to complement RexMD simulations with DPS calculations to build disconnectivity graphs (see Methods). The results in Figure 2 show that the lowest free energy structures correspond

dx.doi.org/10.1021/jp503816r | J. Phys. Chem. B 2014, 118, 8540−8548

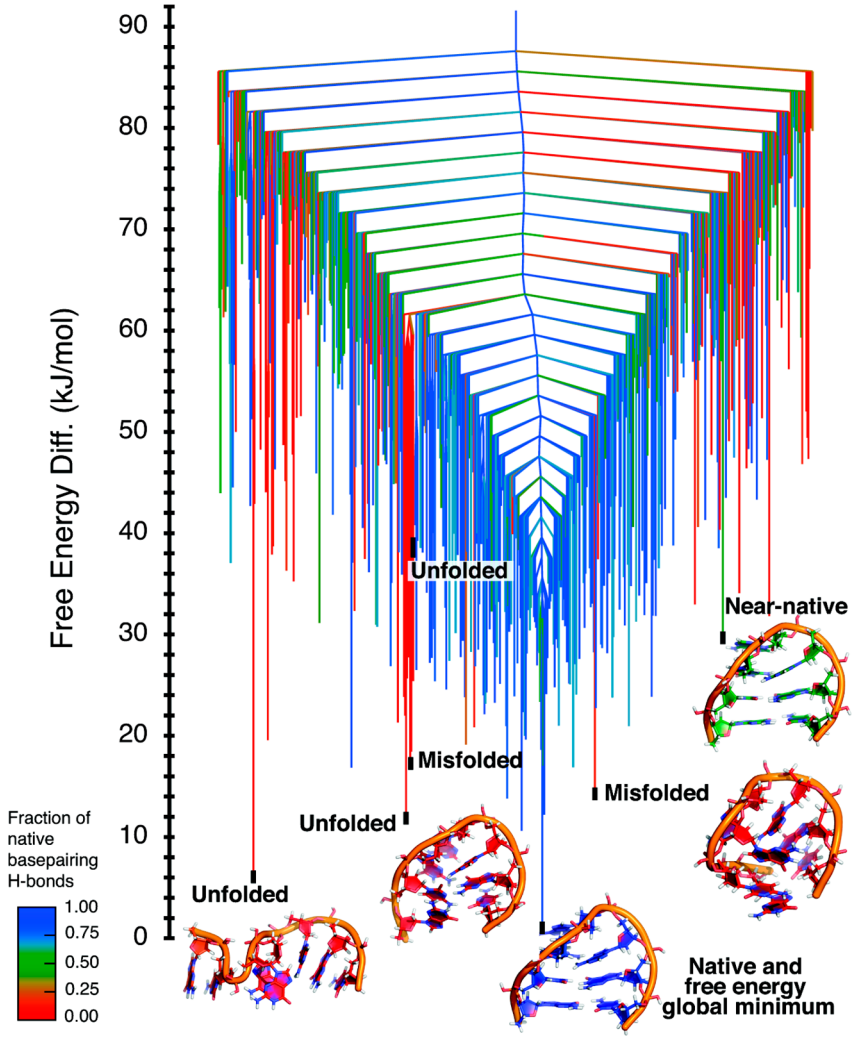


Figure 2. Disconnectivity graph for the $Q = 0$ DNA hairpin. The vertical axis represents the free energy difference with respect to the global minimum, and the spacing of the branches on the horizontal axis was chosen to reveal the structure as clearly as possible. The branches terminate at the energies defined by the individual free energy minima and are joined together at energy thresholds where the barriers separating different sets can be overcome. Branches are colored according to the fraction of native base-pairing H-bond contacts (from no contacts in red to all native contacts in blue) and highlight the diversity of conformations adopted at different energies (unfolded, misfolded, near-native, and native). Selected snapshots colored according to the same order parameter are also shown. The graph reveals that the lowest free energy funnel corresponds to folded structures with a high content of native contacts (blue), and the existence of a funnel of misfolded structures with no native contacts (red) separated from the global minimal by a high free energy barrier.

to folded hairpins with a high content of native base-pairing (native in blue and near-native in green in this Figure).

The long branches (>30 kJ/mol on the vertical axis) reflect the high-energy barriers associated with rearrangements between native and misfolded or unfolded conformations. Extended (unfolded) states (no native contacts and radius of gyration greater than 0.8 nm, red) are sparse (see Supporting Information Figure S4), and are separated by high-energy barriers from the native state (see Figure 3 and Supporting Information Figure S4). In fact, starting from the unfolded state, we observe that the pathway to the native state involves a potential energy barrier of around 130 kJ/mol (Figure 3). Overall, the disconnectivity graphs confirm that the conformational landscape of the hairpin in CCl_4 is very frustrated with several nonnative conformations separated by high-energy barriers. However, the combination of these results with RexMD suggests that compact structures, not far from the aqueous native state, are the dominant ones for the neutral state of the hairpin.

Transfer of the DNA Hairpin between Polar and Apolar Phases. Up to this point we have explored the structure of DNA in dry CCl_4, but DNA is naturally found in water, and it is therefore interesting to study the transfer process and how it affects the structure and energetics of the hairpin. To study the unlikely transfer process (see Methods), we first performed a very slow velocity steered MD simulation to obtain starting conformations for umbrella sampling free energy calculations. We first pulled the hairpin from one water box to the other, passing through the CCl_4 phase (see schematic representation in Supporting Information Figure S5). As seen in Figure 4, even for very small pulling velocities, the hairpin crosses the apolar phase hydrated and maintains the general hairpin structure (see Supporting Information Figures S5 and S6), confirming that the structure of DNA in solution is a metastable conformation in CCl_4.

dx.doi.org/10.1021/jp503816r | J. Phys. Chem. B 2014, 118, 8540−8548

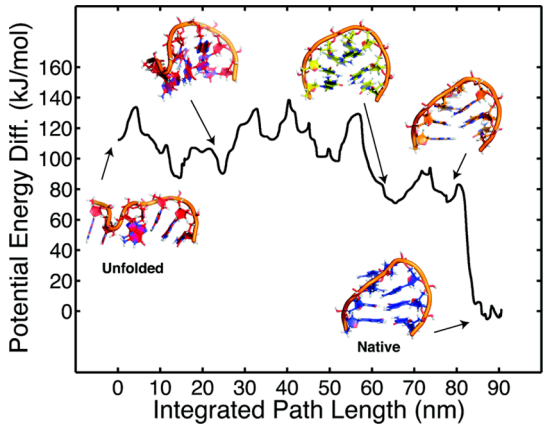


Figure 3. Potential energy difference with respect to the native state as a function of the integrated path length for a discrete path connecting an unfolded structure with the native structure for $Q = 0$. The integrated path length is the integrated Euclidean distance in Cartesian coordinates high energy. Snapshots for selected intervening structures are also shown and colored according to the fraction of native H-bond base-pairing interactions as described in Figure 2

The number of water molecules in the microhydration drop depends on the charge state of the hairpin (see Supporting Information Figure S5), but even for the neutral hairpin there is a small, but nonnegligible, number of water molecules protecting the hairpin from the apolar solvent (see Supporting Information Figure S6). Although displacements of the hairpin within the box were as slow as possible, they are still much faster than the relevant experimental time scale, and we cannot rule out potential problems in the relaxation of the solvent environment around the DNA. To avoid these potential relaxation artifacts as far as possible and to confirm that even a neutral hairpin is hydrated in CCl_4, we repeated the pulling procedure but started with a dry hairpin placed in the middle of the CCl_4 phase. These complementary simulations confirm that as soon as the hairpin moves from the center of the apolar phase, a stream of water molecules flows in (especially dramatic for the $Q = -6$ charge state), spontaneously generating a microhydration environment (Figure 4). This result indicates that the entrance of water molecules in the apolar phase is not a relaxation artifact but a real physical effect. Hence we predict that under normal conditions a hairpin (especially a fully charged one) would cross a hydrophobic phase surrounded by a small number of water molecules.

The discussion in the previous paragraph suggests that if transferred directly from water, the DNA in CCl_4 will probably be surrounded by several water molecules. Thus, we repeated the RexMD calculations but used microhydrated species ($Q = 0$ and $Q = -6$) obtained from the pulling experiments when the center of mass of the hairpin is placed at the middle of the apolar phase. We found that the impact of the microhydration environment helps to stabilize the folded state (see Supporting Information Figures S7 and S8), shielding the electrostatic repulsion between the charged phosphate groups. As a result, a microhydrated DNA placed in the middle of a CCl_4 phase maintains a conformation that is not far away for the native structure in water, even when we consider that phase transfer does not imply change in the ionization state of the oligonucleotide. In summary, our simulations strongly suggest that the neutral hairpin maintains well the solution structure when transferred to either dry or wet CCl_4. On the contrary, the fully charged hairpin, while remaining close to the aqueous structure in wet CCl_4, is completely unfolded in dry CCl_4.

We used WHAM and umbrella sampling (from steered-MD snapshots; see Methods) to compute the free energy change associated with the transfer of the hairpin across polar and apolar phases. The results displayed in Figure 5 (see Methods and Supporting Information Figure S9 for details) clearly illustrate the large free energy barrier associated with crossing a hydrophobic phase: around 650 kJ/mol for a fully charged hairpin and 150−200 kJ/mol for the neutral. These numbers are in qualitative agreement with those obtained from a Poisson−Boltzmann calculation (see Figure 6) which demonstrated that transfer of a hairpin from water to an apolar solvent, even if microhydrated, is largely disfavored. These results reinforce the notion that decoration of the DNA with apolar moieties[3−7] is desirable to improve permeation.

Finally, we studied the relative weight of neutral and charged states of a hairpin in CCl_4. Using standard thermodynamic cycles (see Supporting Information Figure S10), we computed

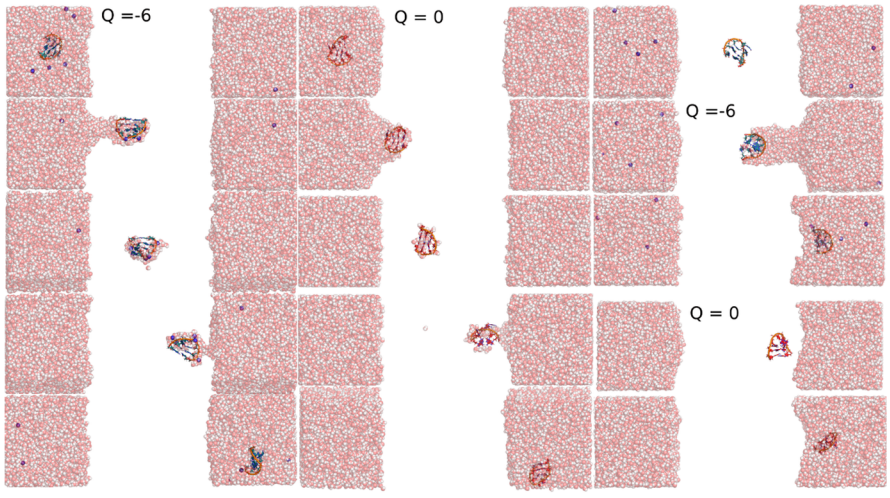


Figure 4. Representation of different snapshots obtained in steered MD simulations forcing the phase transfer of DNA hairpin: (A) from water to CCl4 and back to water for $Q = -6$ DNA hairpin; (B) from water to CCl_4 and back to water for $Q = 0$ DNA hairpin; (C) from CCl_4 to water for the $Q = -6$ DNA hairpin; (D) from CCl4 to water for the $Q = 0$ DNA hairpin.

dx.doi.org/10.1021/jp503816r | J. Phys. Chem. B 2014, 118, 8540−8548

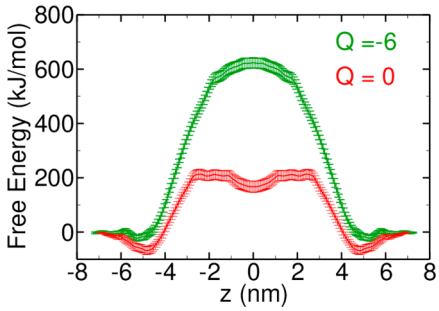


Figure 5. Potential of mean force associated with the transfer of the hairpin ($Q = -6$ and $Q = 0$) across a polar/apolar/polar simulation box (see Methods and Supporting Information Figure S1).

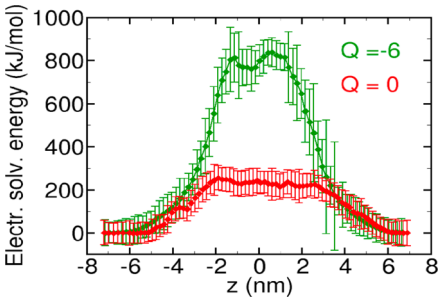


Figure 6. Change in the electrostatic solvation energy associated with the transfer water→CCl_4→water of a fully charged ($Q = -6$) and neutral ($Q = 0$) hairpin.

the free energy cost of transforming a neutral to a charged state in CCl_4, the cost of removing the hairpin charge in water, and the cost associated with the phase transfer of the neutral and charged hairpins. The free energy change associated with the annihilation of the hairpin charge can be determined at neutral pH from the experimental pK_a in water of a DNA phosphate (pK_a around 2; ΔG around 160 kJ/mol assuming that all phosphate groups protonate independently of each other). Thus, transfer of a hairpin ($Q = -6$) in water into a ($Q = 0$) hairpin in CCl_4 is overall disfavored by around 350 kJ/mol, while the same transfer assuming $Q = -6$ in both solvents has an associated free energy change around 450−500 kJ/mol higher. In other words, our simulations strongly suggest that DNA in apolar media should have a reduced charge and could even be completely neutral. Under these conditions DNA seems able to maintain its native structure in solution.

CONCLUSION

DNA is a highly polar molecule, which has evolved over millions of years to be stable and functional in aqueous solution. DNA will not spontaneously transfer to an apolar phase, such as a biological membrane, which represents a clear advantage for confinement in a given cellular compartment. However, biotechnological and biomedical applications often require the DNA to cross biological membranes, be encapsulated in a hydrophobic, environment, or just act as a chemical reactant in a nonaqueous solvent. It is then necessary to study the behavior of nucleic acids in apolar solvents as a preliminary step to design modifications that can favor transfer from water. For this purpose, theoretical approaches are ideal, since they allow us to study DNA in such hostile environments. Our large-scale simulation effort, described in the present report, illustrates the significant changes occurring in the conformational landscape of a small model system of DNA when transferred from water to CCl_4. The DNA becomes stiffer and the landscape is more frustrated in the apolar phases, which might generate kinetic traps. Compact states, probably not far from the aqueous structure, are sampled if the DNA is neutralized, while more extended conformations are expected if the DNA maintains its ionic state in water.

The transfer of DNA from aqueous solution to CCl_4 is energetically disfavored, especially if the DNA maintains its aqueous ionic state. The transfer always occurs with the DNA inside a microdroplet of water. The size of the water droplet increases with the net charge of the DNA, but even for a neutral DNA a significant number of water molecules appears in our simulations for the apolar phase. A combination of the umbrella sampling potential of mean force simulations (supported by Poisson−Boltzmann calculations) with experimental pK_a estimates demonstrates that the neutral state is favored, suggesting that transfer of DNA could generate a net proton flux across a membrane. Overall, neutralization of DNA charges seems a key mechanism to improve transfer properties of DNA and either changing phosphoric by another acid with higher pK_as or neutralizing decorations neutralizing DNA charge seems a promising approach to improve DNA bioavailability.

ASSOCIATED CONTENT

Supporting Information

Figures illustrating multiple structural descriptors from our calculations (RMSD, hydrogen bond content, radius of gyration), a disconectivity graph, the setup for the steering MD simulation protocol, evolution of the number of waters bound to DNA, structural descriptors of the DNA hairpin during phase transfer, RexMD results for the microhydrated hairpin in CCl_4 solution, the distribution of histograms from umbrella sampling simulations, thermodynamic cycles and transport pathway through a low polarity medium. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: modesto.orozco@irbbarcelona.org. Tel.: +34 93 403 7155.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Adam Hospital for help with the Poisson−Boltzmann calculations. Our work has been supported by the Spanish MINECO (BIO2012-32868), the Spanish National Institute of Bioinformatics (INB), and the European Research Council (ERC 291433). M.O. is an ICREA-Academia fellow, and G.P. is a Sara Borrell fellow.

REFERENCES

(1) Lin, J.; Seeman, N. C.; Vaidehi, N. Molecular-Dynamics Simulations of Insertion of Chemically Modified DNA Nanostructures into a Water-Chloroform Interface. *Biophys. J.* **2008**, *95*, 1099−107.

(2) Mann, S. Self-Assembly and Transformation of Hybrid Nano-Objects and Nanostructures under Equilibrium and Non-Equilibrium Conditions. *Nat. Mater.* **2009**, *8*, 781−92.

(3) Bailey, A. L.; Sullivan, S. M. Efficient Encapsulation of DNA Plasmids in Small Neutral Liposomes Induced by Ethanol and Calcium. *Biochim. Biophys. Acta, Biomembr.* **2000**, *1468*, 239−52.

(4) Hayes, M. E.; Drummond, D. C.; Kirpotin, D. B.; Zheng, W. W.; Noble, C. O. Genospheres: Self-Assembling Nucleic Acid-Lipid Nanoparticles Suitable for Targeted Gene Delivery. *Gene Ther.* **2006**, *13*, 646−51.
(5) Jeffs, L. B.; Palmer, L. R.; Ambegia, E. G.; Giesbrecht, C.; Ewanick, S. A Scalable, Extrusion-Free Method for Efficient Liposomal Encapsulation of Plasmid DNA. *Pharm. Res.* **2005**, *22*, 362−72.
(6) Maurer, N.; Wong, K. F.; Stark, H.; Louie, L.; McIntosh, D. Spontaneous Entrapment of Polynucleotides Upon Electrostatic Interaction with Ethanoldestabilized Cationic Liposomes. *Biophys. J.* **2001**, *80*, 2310−26.
(7) Skjorringe, T.; Gjetting, T.; Jensen, T. G. A Modified Protocol for Efficient DNA Encapsulation into Pegylated Immunoliposomes (Pils). *J. Controlled Release* **2009**, *139*, 140−5.
(8) Gao, Y. G.; Robinson, H.; van Boom, J. H.; Wang, A. H. Influence of Counter-Ions on the Crystal Structures of DNA Decamers: Binding of $[Co(NH_3)_6]^{3+}$ and Ba^{2+} to a-DNA. *Biophys. J.* **1995**, *69*, 559−68.
(9) Varnai, P.; Zakrzewska, K. DNA and Its Counterions: A Molecular Dynamics Study. *Nucleic Acids Res.* **2004**, *32*, 4269−80.
(10) Portella, G.; Germann, M. W.; Hud, N. V.; Orozco, M. MD and NMR Analyses of Choline and Tma Binding to Duplex DNA: On the Origins of Aberrant Sequence-Dependent Stability by Alkyl Cations in Aqueous and Water-Free Solvents. *J. Am. Chem. Soc.* **2014**, *136*, 3075−86.
(11) Bonner, G.; Klibanov, A. M. Structural Stability of DNA in Nonaqueous Solvents. *Biotechnol. Bioeng.* **2000**, *68*, 339−44.
(12) Herskovits, T. T.; Harrington, J. P. Solution Studies of the Nucleic Acid Bases and Related Model Compounds. Solubility in Aqueous Alcohol and Glycol Solutions. *Biochemistry* **1972**, *11*, 4800−11.
(13) Ke, F.; Luu, Y. K.; Hadjiargyrou, M.; Liang, D. Characterizing DNA Condensation and Conformational Changes in Organic Solvents. *PLoS One* **2010**, *11*, No. e13308.
(14) Feng, Y.; Spisz, T. S.; Hoh, J. H. Ethanol-Induced Structural Transitions of DNA on Mica. *Nucleic Acids Res.* **1999**, *27*, 1943−9.
(15) Montesi, A.; Pasquali, M.; MacKintosh, F. C. Collapse of a Semiflexible Polymer in Poor Solvent. *Phys. Rev. E* **2004**, *69*, No. 021916.
(16) Pereira, G. G.; Williams, D. R. M. Toroidal Condensates of Semiflexible Polymers in Poor Solvents: Adsorption, Stretching, and Compression. *Biophys. J.* **2001**, *80*, 161−8.
(17) Mamajanov, I.; Engelhart, A. E.; Bean, H. D.; Hud, N. V. DNA and RNA in Anhydrous Media: Duplex, Triplex, and G-Quadruplex Secondary Structures in a Deep Eutectic Solvent. *Angew. Chem.* **2010**, *49*, 6454−8.
(18) Arcella, A.; P, G.; Ruiz, M. L.; Eritja, R.; Vilaseca, M.; Gabelica, V.; Orozco, Modesto Structure of Triplex DNA in the Gas Phase. *J. Am. Chem. Soc.* **2012**, *134*, 6596−606.
(19) Rueda, M.; Luque, F. J.; Orozco, M. G-Quadruplexes Can Maintain Their Structure in the Gas Phase. *J. Am. Chem. Soc.* **2006**, *128*, 3608−19.
(20) Rueda, M. K.; S, G.; Luque, F. J.; Orozco, M. The Structure and Dynamics of DNA in Gas Phase. *J. Am. Chem. Soc.* **2003**, *125*, 8007−14.
(21) Cui, S.; Yu, J.; Kühner, F.; Schulten, K.; Gaub, H. E. Double-Stranded DNA Dissociates into Single Strands When Dragged into a Poor Solvent. *J. Am. Chem. Soc.* **2007**, *129*, 14710−6.
(22) Perez, A.; Orozco, M. Real-Time Atomistic Description of DNA Unfolding. *Angew. Chem., Int. Ed.* **2010**, *49*, 4805−8.
(23) Khalid, S.; Bond, J. P.; Holyoake, J.; Hawtin, W. R.; Sansom, M. DNA and Lipid Bilayers: Self-Assembly and Insertion. *J. R. Soc., Interface* **2008**, *5*, S241−56.
(24) Padrta, P.; Stefl, R.; Králík, L.; Zídek, L.; Sklenár, V. Refinement of D(Gcgaagc) Hairpin Structure Using One- and Two-Bond Residual Dipolar Couplings. *J. Biomol. NMR* **2002**, *24*, 1−4.
(25) Ansari, A.; Kuznetsov, S. V.; Shen, Y. Configurational Diffusion Down a Folding Funnel Describes the Dynamics of DNA Hairpins. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 7771−6.
(26) Ma, H.; Wan, C.; Wu, A.; Zewail, A. H. DNA Folding and Melting Observed in Real Time Redefine the Energy Landscape. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 712−6.
(27) Portella, G.; Orozco, M. Multiple Routes to Characterize the Folding of a Small DNA Hairpin. *Angew. Chem., Int. Ed.* **2010**, *49*, 7673−6.
(28) Orozco, M.; Perez, A.; Noy, A.; Luque, F. J. Theoretical Methods for the Simulation of Nucleic Acids. *Chem. Soc. Rev.* **2003**, *32*, 350−64.
(29) Perez, A.; Luque, F. J.; Orozco, M. Frontiers in Molecular Dynamics Simulations of DNA. *Acc. Chem. Res.* **2012**, *45*, 196−205.
(30) Gelpi, J. L.; Kalko, S. G.; Barril, X.; Cirera, J.; de La Cruz, X.; Luque, F. J.; Orozco, M. Classical Molecular Interaction Potentials: Improved Setup Procedure in Molecular Dynamics Simulations of Proteins. *Proteins* **2001**, *45*, 428−37.
(31) Shields, G. C. L.; C, A.; Orozco, M. Molecular Dynamics Simulations of the d(T·A·T) Triple Helix. *J. Am. Chem. Soc.* **1997**, *119*, 7463−9.
(32) Darden, T.; Y, D.; Pedersen, L. Particle Mesh Ewald—An N.Log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089−92.
(33) Sugita, Y.; Okamoto, Y. Replica-Exchange Molecular Dynamics Method for Protein Folding. *Chem. Phys. Lett.* **1999**, *314*, 141−51.
(34) Becker, O. M.; Karplus, M. The Topology of Multidimensional Potential Energy Surfaces: Theory and Application to Peptide Structure and Kinetics. *J. Chem. Phys.* **1997**, *106*, 1495−517.
(35) Wales, D. J.; Miller, M. A.; Walsh, T. R. Archetypal Energy Landscapes. *Nature* **1998**, *394*, 758−760.
(36) Wales, D. J. *Energy Landscapes: With Applications to Clusters, Biomolecules, and Glasses*; Cambridge University Press: Cambridge, U.K., 2004.
(37) Wales, D. J. Discrete Path Sampling. *Mol. Phys.* **2002**, *100*, 3285−305.
(38) Wales, D. J. Pathsample: A Program for Generating Connected Stationary Point Databases and Extracting Global Kinetics. http://www-wales.ch.cam.ac.uk/software.html.
(39) Wales, D. J. Optim: A Program for Optimising Geometries and Calculating Pathways. http://www-wales.ch.cam.ac.uk/software.html.
(40) Nocedal, J. Updating Quasi-Newton Matrices with Limited Storage. *Math. Comput* **1980**, *35*, 773−82.
(41) Liu, D. C.; Nocedal, J. On the Limited Memory BFGS Method for Large-Scale Optimization. *Math. Program.* **1989**, *45*, 503−28.
(42) Trygubenko, S. A.; Wales, D. J. Erratum: A Doubly Nudged Elastic Band Method for Finding Transition States [*J. Chem. Phys.* **120**, 2082 (2004)]. *J. Chem. Phys.* **2004**, *120*, 7820.
(43) Henkelman, G.; Jonsson, H. A Dimer Method for Finding Saddle Points on High Dimensional Potential Surfaces Using Only First Derivatives. *J. Chem. Phys.* **1999**, *111*, 7010−22.
(44) Munro, L. J.; Wales, D. J. Defect Migration in Crystalline Silicon. *Phys. Rev. B: Condens. Matter* **1999**, *59*, 3969−80.
(45) Carr, J. M.; Trygubenko, S. A.; Wales, D. J. Finding Pathways between Distant Local Minima. *J. Chem. Phys.* **2005**, *122*, 234903.
(46) Evans, D. A.; Wales, D. J. Free Energy Landscapes of Model Peptides and Proteins. *J. Chem. Phys.* **2003**, *118*, 3891−7.
(47) Case, D. A.; Darden, T. A.; Cheatham, T. E. I.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Pearlman, D. A.; Crowley, M.; Walker, R. C.; Zhang, W.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Wong, K. F.; Paesani, F.; Wu, X.; Brozell, S.; Tsui, V.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Mathews, D. H.; Schafmeister, C.; Ross, W. S.; Kollman, P. A. *Amber 9*; University of California: San Francisco, CA, USA, 2006.
(48) Onufriev, A.; Bashford, D.; Case, D. A. Exploring Protein Native States and Large-Scale Conformational Changes with a Modified Generalized Born Model. *Proteins* **2004**, *55*, 383−94.
(49) Hess, B.; van der Spoel, D.; Lindhal, E. The Pull Code, *GROMACS USER MANUAL*, version 4.5.3; 2010; ftp://ftp.gromacs.org/pub/manual/manual-4.5.3.pdf, p 140.
(50) Hub, J. S.; de Groot, B. L.; van der Spoel, D. G_Wham—A Free Weighted Histogram Analysis Implementation Including Robust Error

and Autocorrelation Estimates. *J. Chem. Theory Comput.* **2010**, *6*, 3713−20.
(51) Kumar, S.; Rosemberger, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. *J. Comput. Chem.* **1992**, *13*, 1011−21.
(52) Orozco, M.; Luque, F. J. Theoretical Methods for the Description of the Solvent Effect in Biomolecular Systems. *Chem. Rev.* **2000**, *100*, 4187−225.
(53) Hub, J. S.; de Groot, B. L.; Grubmuller, H.; Groenhof, G. Quantifying Artifacts in Ewald Simulations of Inhomogeneous Systems with a Net Charge. *J. Chem. Theory Comput.* **2014**, *10*, 381−90.
(54) Li, Z.; Scheraga, H. A. Monte Carlo-Minimization Approach to the Multiple-Minima Problem in Protein Folding. *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84*, 6611−5.
(55) Wales, D. J.; Doye, J. P. Global Optimization by Basin-Hopping and the Lowest Energy Structures of Lennard-Jones Clusters Containing up to 110 Atoms. *J. Phys. Chem. A* **1997**, *101*, 5111−6.