Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Electronic Supplementary Information (ESI)

Xiakun Chu[1], and Victor Muñoz[1,2,3*]

[1] IMDEA Nanosciences, Faraday 9, Ciudad Universitaria Cantoblanco, Madrid, 28049, Spain
[2] National Biotechnology Center, Consejo Superior de Investigaciones Cientficas, Darwin 3, Campus de Cantoblanco, Madrid, 28049, Spain
[3] Department of Bioengineering, University of California, 95343 Merced, CA, USA
[*]Corresponding Author: vmunoz3@ucmerced.edu

Models and Methods

1. Coarse-grained structure-based model

The coarse-grained structure-based model (CGSBM) includes a description of the folding process of EngHD folding and specific and non-specific binding to DNA. The CGSBM takes into account only native interactions and electrostatics and reduces the molecular representation by coarse-graining. This type of model is based on energy landscape theory[1-4] and has achieved significant successes in describing a variety of biomolecular dynamics processes including protein folding, binding and conformational transitions[5-10]. Our CGSBM is similar to those used in a large number of previous publications[11-22]. Briefly, each amino acid of EngHD is represented by two beads, one corresponding to the backbone (CA bead) and another corresponding to the side-chain (CB). The first one for all the aminoacids and the latter for all but Glycine. Each strand of the DNA molecule is represented with three beads, one corresponding to the sugar, another to the nitrogenous base and a third bead corresponding to the phosphate group. The CGSBM Hamiltonian includes three general terms: 1) a structure based potential for the folding of EngHD; 2) a structure based potential for the binding of EngHD to the specific DNA binding site; 3) a general electrostatic potential defining the interactions between EngHD and any non-specific binding site in the DNA.

$$V = V_{SBM}^{EngHD} + V_{SBM}^{EngHD-DNA} + V_{Ele}^{EngHD-DNA}$$

where V_{SBM}^{EngHD} and $V_{SBM}^{EngHD\text{-}DNA}$ are SBM folding and specific DNA binding potentials that have been derived using the SMOG webserver[23]. $V_{Ele}^{EngHD\text{-}DNA}$ is the non-specific DNA interaction potential, which is described by a simple ionic strength dependent Debye-Hückel model. V_{SBM}^{EngHD} is a typical SBM potential with the following terms:

$$\begin{aligned} V_{SBM}^{EngHD} = & \sum_{bonds} K_r (r - r_0)^2 + \sum_{angles} K_\theta (\theta - \theta_0)^2 + \sum_{chirality} K_\chi (\chi - \chi_0)^2 \\ & + \sum_{dihedral} \varepsilon_\phi K_\phi^{(n)} [1 + \cos(n \times (\phi - \phi_0))] \\ & + \sum_{i<j-3}^{native} \varepsilon_f [5(\frac{\sigma_{ij}}{r_{ij}})^{12} - 6(\frac{\sigma_{ij}}{r_{ij}})^{10}] + \sum_{i<j-3}^{non-native} \varepsilon_{NN} (\frac{\sigma_{NN}}{r_{ij}})^{12} \end{aligned}$$

The SBM potential includes bond stretching, angle bending, chiral, torsional, and nonbonded interactions. The parameters K_r, K_θ, K_χ, K_ϕ, ε_ϕ, ε_f, σ_{NN} represent the strength of each type of interaction. R are the bond lengths and θ the bond angles. ϕ and χ are the proper and improper dihedral angles. In all cases the zero subscript is used to represent the exact values adopted in the native configuration. Nonbonded interactions are subdivided into native interactions and nonnative interactions. For native interactions, $\sigma_{i,j}$ is the distance between beads *i* and *j* in the native 3D structure. For non-native contacts, σ_{NN} represents a excluded volume repulsive term. ε_ϕ determines the strength of the dihedral energy term, which is related to the local native stabilization energy. ε_f determines the strength of the native interactions, which are related to the non-local native stabilization energy. Reduced units were used for all

calculations, so K_r = 100.0, K_θ= 20.0, $K_\phi^{(1)}$= 1.0, $K_\phi^{(3)}$= 0.5, ε_ϕ= 1.0, ε_f = 1.0 and ε_{NN} = 1.0. Further details about our CGSBM can be found here[24].

The 3D structures of EngHD in its free state and of EngHD bound to DNA are very similar (Fig. S10A). We thus assumed that the native state of EngHD is invariant upon DNA binding and used the structure of the free state to generate the folding potential for EngHD. Native contacts within EngHD were computed using the Contacts of Structural Units (CSU) software[25] while for EngHD-DNA, a contact will be formed as any pair of residues (bases) in which at least one heavy atom pair is found at a distance shorter than 5 Å. In the X-ray structure of the EngHD-DNA complex, the DNA molecule is a 21bp-long DNA double duplex including the consensus specific binding sequence TAATTA.

To be able to examine the non-specific DNA search process in depth we used a DNA molecule of 100bp-long including one specific binding site bearing the interactions defined from the atomic coordinates of the complex (Fig. S10B) located in the center of the DNA molecule (Fig. S10C). In addition to the specific binding site, EngHD can bind to any non-specific 10-bp site of the DNA via non-specific electrostatic interactions. Specific DNA-protein interactions only occur when the protein is located in the specific binding site. In the simulations the 100bp DNA segment was placed along Z-axis direction and kept rigid.

Although we used reduced units to describe energy and time scale, it is still useful to compute the "real" times from SBM simulations, which may help to infer the kinetic timescale for the system. We can estimate the connection between the "real" and reduced time scale (τ_{SBM}) by the expression[26, 27]:

$$\tau_{SBM} = (\frac{m_0 a_0^2}{\varepsilon})^2$$

, where m_0 and a_0 are the mass of the beads and bond length, and ε is the energy unit. For our two-bead SBM, the typical values of m_0 and a_0 are $1.5X10^{-22}$ g and $3.8X10^{-8}$ cm and ε = 1.0 kJ/mol, so $\tau_{SBM} \approx 3.6$ ps.

We collected the folding trajectories of EngHD at folding temperature in our structure-based model simulations, and calculated the folding time of EngHD is 62.1 in reduced units, corresponding to 2.2 ns in "real" time scale. In reality, the folding kinetics of EngHD was measured by temperature jump experiments and the folding time at the folding temperature was found to be 20.8 μs[28], which is 4 orders of magnitude larger than the simulation value, in good accordance with the work done by Kouza et al.[26]. It implies that our estimation on "real" times in our SBM simulations is reasonable.

2. Analysis of the DNA-Binding Simulations

One-dimensional sliding and gliding along DNA. Here we use definitions for the different DNA search modes that are similar with those used by Givaty and Levy[13]. 3D diffusion is defined as any snapshot in the trajectory in which EngHD is at least 30 Å away from the closest atoms of the DNA molecule. 1D Sliding is defined as any snapshot that fullfils these three criteria simultaneously: (1) at least 9 of the 14 residues (70%) in the DNA recognition helix of EngHD are in atomic contact with the DNA major groove; (2) the distance of the centroid of the DNA recognition helix of EngHD to the DNA molecule is shorter than 20 Å, which defines a DNA recognition helix of EngHD that is longer by 10 Å than that defined strictly by the interactions

observed in the X-ray complex structure; (3) the orientation angle, which is the angle defined by the centroid of the DNA recognition helix of EngHD, the centroid of the entire EngHD, and the point on the DNA's main (Z) axis perpendicular to the EngHD's centroid is less than 90°. 1D gliding is defined as any snapshot in which EngHD falls within 30 Å distance of the DNA molecule but is not performing 1D sliding (not fullfiling the criteria for sliding).

Quantifying the search speed on DNA. Experimentally, the DNA search efficiency of a protein is measured by the 1D diffusion coefficient (D_1) of the protein along the DNA molecule direction[29, 30]. If the 1D motion of the protein along DNA is Brownian, D_1 can be derived from a single trajectory by calculating the mean square displacement (MSD) with the expression[31]:

$$MSD(n,N) = \sum_{i=1}^{N-n} \frac{(Z_{i+n} - Z_i)^2}{N - n} = 2D_1 n\Delta t \qquad (1)$$

where *N* is the number of frames (observable points) in the trajectory, *n* is the index of the frame, and *Δt* is the time interval between two consecutive frames. Equation (1) is valid if the trajectory obeys Brownian motion (Fig. S11B). If MSD(n,N) is found to be linearly dependent on *t* within a certain time threshold (defined by frame n_c), D_1 can be obtained directly from the slope of equation (1). In practice, choosing the cutoff n_c is difficult because there are no objective criteria currently defined. Alternatively, one can make use of the expression for the variance of MSD(n,N), which has the following expression[30]:

$$\sigma^2_{(n,N)} = \frac{(2D_1 n\Delta t)^2 (2n^2 + 1)}{3n(N - n + 1)}$$

n_c is then chosen to set the uncertainty of MSD(n,N), calculated as $\sigma(n,N)/(2D_1 n\Delta t)$, below 20%. Practically, we analyzed trajectory segments of 100 frames and set n_c to 5.

3. Manipulating the Folding Behavior of EngHD

Variation in Conformational Disorder. We generated versions of the EngHD folding potential in which the protein exhibits different degrees of conformational disorder by changing the strength of ε_f (strength of native folding interactions) from 0.8 to 1.2. We performed protein folding simulations (no DNA) at the folding temperature defined by ε_f=1.0 and plotted the folding energy landscapes along native similarity reaction coordinate $Q_{Folding}$ (Fig. S3) to examine the degree of folding present in each condition. As ε_f increases from 0.8 to 1.2, EngHD changes from being completely unfolded to populate the unfolded and folded ensembles equally (at ε_f=1.0) and them to be permanently folded.

Variation in Folding Scenario. We control the folding scenario of EngHD in our CGSBM by changing the strength of each dihedral term relative to the native contact potential ($R=\varepsilon_\phi/\varepsilon_f$). To guarantee that the folding temperature remains invariant we also modulated the overall strength of the folding potential V_{SBM}^{EngHD}. Varying R between 0.1 and 3 results in a gradual decrease in the folding free energy barrier obtained by projecting the free energy surface onto the parameter $Q_{folding}$. At R=0.1 EngHD exhibits a free energy barrier of 4.50 kT at the folding temperature, which corresponds to an incipient two-state folding scenario. At R=3 EngHD exhibits a free energy barrier of 0.26 kT, which practically approaches the one-state downhill folding scenario (Fig. S7). The original implementation of the model (R=1) results in a free energy

barrier of 1.28 kT, which corresponds to a downhill folding scenario. Varying R also induce changes in the structural properties of the unfolded state. Particularly, the higher the value of R the more residual local structure is present in the unfolded ensemble (increasing the residual helical content of unfolded EngHD). This change shifts the overall folding mechanism from a nucleation-condensation-like for the lowest R to a diffusion-collision one at the highest R (Fig. S1). To better represent the properties of the folding scenario of EngHD we introduce the parameter "downhillness" as a measure of the degree of folding cooperativity in the protein. Downhillness is simply the value of R normalized according to the expression $(R-R_{min})/(R_{max}-R_{min})$, in which R_{min} is 0.1 and R_{max} is 3.

Defining Room Temperature in the Simulations. EngHD is a protein from the fruitfly *Drosophila melanogaster*. We thus define the physiological temperature as room temperature (T=298K). There is no direct relation between the reduced temperature used in SBM and the real temperature. However, relative temperature scales can be obtained using some experimental observables, such as B-factor and folding temperature, as reference. The experimental folding temperature T_f^{Exp} of EngHD obtained by differential scanning calorimetry (DSC) is 325K[28]. From the peak in the heat capacity profile obtained from our simulations we can determine the theoretical folding temperature T_f^{Sim} in the SBM, which is 1.4 (in energy units by multiplying it by Boltzmann constant, k). Assuming a linear temperature dependence of the energy we estimate that room temperature in the simulations corresponds to:

$$T_r^{Sim} = \frac{T_r^{Exp}}{T_f^{Exp}} \times T_f^{Sim} = 0.92 T_f^{Sim} = 1.28$$

Here is worth noting that because of solvent effects, there is in fact a change in heat capacity of denaturation which results in non-linear temperature dependence. However the experimentally determined change in heat capacity of EngHD is relatively small[28], and the temperature range between room temperature and the experimental temperature (27 K) is narrow enough to make the linear approximation reasonably accurate.

References

1. J. D. Bryngelson, J. N. Onuchic, N. D. Socci and P. G. Wolynes, *Proteins: Struct., Funct., Bioinf.*, 1995, 21, 167-195.
2. J. D. Bryngelson and P. G. Wolynes, *Proc. Natl. Acad. Sci. U.S.A*, 1987, 84, 7524-7528.
3. P. E. Leopold, M. Montal and J. N. Onuchic, *Proc. Natl. Acad. Sci. U.S.A*, 1992, 89, 8721-8725.
4. J. N. Onuchic, Z. Luthey-Schulten and P. G. Wolynes, *Annu. Rev. Phys. Chem.*, 1997, 48, 545-600.
5. C. Clementi, H. Nymeyer and J. N. Onuchic, *J. Mol. Biol.*, 2000, 298, 937-953.
6. J. Karanicolas and C. L. Brooks, *Protein Sci.*, 2002, 11, 2351-2361.
7. Y. Levy, S. S. Cho, J. N. Onuchic and P. G. Wolynes, *J. Mol. Biol.*, 2005, 346, 1121-1145.
8. Y. Levy, P. G. Wolynes and J. N. Onuchic, *Proc. Natl. Acad. Sci. U.S.A.*, 2004, 101, 511-516.
9. P. C. Whitford, J. K. Noel, S. Gosavi, A. Schug, K. Y. Sanbonmatsu and J. N. Onuchic, *Proteins: Struct., Funct., Bioinf.*, 2009, 75, 430-441.
10. P. C. Whitford and J. N. Onuchic, *Curr. Opin. Struct. Biol.*, 2015, 30, 57-62.
11. A. Azia and Y. Levy, *J. Mol. Biol.*, 2009, 393, 527-542.
12. X. Chu, F. Liu, B. A. Maxwell, Y. Wang, Z. Suo, H. Wang, W. Han and J. Wang, *PLOS Comput.*

Biol., 2014, 10, e1003804.
13. O. Givaty and Y. Levy, *J. Mol. Biol.*, 2009, 385, 1087-1097.
14. Y. Levy, J. N. Onuchic and P. G. Wolynes, *J. Am. Chem. Soc.*, 2007, 129, 738-739.
15. A. Marcovitz and Y. Levy, *Biophys. J.*, 2009, 96, 4212-4220.
16. A. Toth-Petroczy, I. Simon, M. Fuxreiter and Y. Levy, *J. Am. Chem. Soc.*, 2009, 131, 15084-15085.
17. D. Vuzman, A. Azia and Y. Levy, *J. Mol. Biol.*, 2010, 396, 674-684.
18. D. Vuzman and Y. Levy, *Proc. Natl. Acad. Sci. U.S.A*, 2010, 107, 21004-21009.
19. D. Vuzman and Y. Levy, *Mol. Biosyst.*, 2012, 8, 47-57.
20. D. Vuzman and Y. Levy, *Isr. J. Chem.*, 2014, 54, 1374-1381.
21. D. Vuzman, M. Polonsky and Y. Levy, *Biophys. J.*, 2010, 99, 1202-1211.
22. L. Zandarashvili, A. Esadze, D. Vuzman, C. A. Kemme, Y. Levy and J. Iwahara, *Proc. Natl. Acad. Sci. U.S.A*, 2015, 112, E5142-E5149.
23. J. K. Noel, P. C. Whitford, K. Y. Sanbonmatsu and J. N. Onuchic, *Nucleic Acids Res.*, 2010, 38, W657-W661.
24. X. Chu, Y. Wang, L. Gan, Y. Bai, W. Han, E. Wang and J. Wang, *PLOS Comput. Biol.*, 2012, 8, e1002608.
25. V. Sobolev, A. Sorokine, J. Prilusky, E. E. Abola and M. Edelman, *Bioinformatics*, 1999, 15, 327-332.
26. M. Kouza, M. S. Li, E. P. O'Brien, C.-K. Hu and D. Thirumalai, *J. Phys. Chem. B*, 2006, 110, 671-676.
27. T. Veitshans, D. Klimov and D. Thirumalai, *Folding Des.*, 1997, 2, 1-22.
28. U. Mayor, C. M. Johnson, V. Daggett and A. R. Fersht, *Proc. Natl. Acad. Sci. U.S.A*, 2000, 97, 13518-13522.
29. A. Granéli, C. C. Yeykal, R. B. Robertson and E. C. Greene, *Proc. Natl. Acad. Sci. U.S.A*, 2006, 103, 1221-1226.
30. Y. M. Wang, R. H. Austin and E. C. Cox, *Phys. Rev. Lett.*, 2006, 97, 048302.
31. H. Qian, M. P. Sheetz and E. L. Elson, *Biophys. J.*, 1991, 60, 910.
32. D. De Sancho and R. B. Best, *Mol. Biosyst.*, 2012, 8, 256-267.
33. S. Lifson and A. Roig, *J. Chem. Phys.*, 1961, 34, 1963-1974.
34. G. Hummer, *J. Phys. Chem.*, 2004, 120, 516-523.
35. R. B. Best and G. Hummer, *Proc. Natl. Acad. Sci. U.S.A*, 2005, 102, 6732-6737.
36. N. D. Clarke, C. R. Kissinger, J. Desjarlais, G. L. Gilliland and C. O. Pabo, *Protein Sci.*, 1994, 3, 1779-1787.
37. E. Fraenkel, M. A. Rould, K. A. Chambers and C. O. Pabo, *J. Mol. Biol.*, 1998, 284, 351-361.

Tables

Table S1: Structural and energetic properties for the folded and unfolded ensembles of EngHD when performing sliding and gliding.

Stages	**1D Sliding**		**1D gliding**	
States	Folded	Unfolded	Folded	Unfolded
Energy[1]	-55.5 ± 6.0	-53.5 ± 8.9	-45.6 ± 6.6	-50.4 ± 9.3
D(all)[2]	12.6 ± 4.3	12.5 ± 3.0	14.1 ± 4.6	12.7 ± 2.8
D(Charged)[2]	10.9 ± 4.5	10.7 ± 2.8	12.6 ± 5.0	10.7 ± 2.0
Salt bridge [3]	5.2 ± 1.6	5.1 ± 2.0	3.8 ± 1.7	4.6 ± 2.0

[1] Energy is the non-native electrostatic energy between EngHD and DNA in reduced units.

[2] D is the average minimum distance from EngHD residues to their closest DNA atom. The label “all” indicates all the residues in the protein. The label “charged” applies to the positively charged residues in EngHD. D is in Å.

[3] Number of salt bridges defined as contacts between positively charged residues in EngHD and phosphate groups in DNA.

Figures

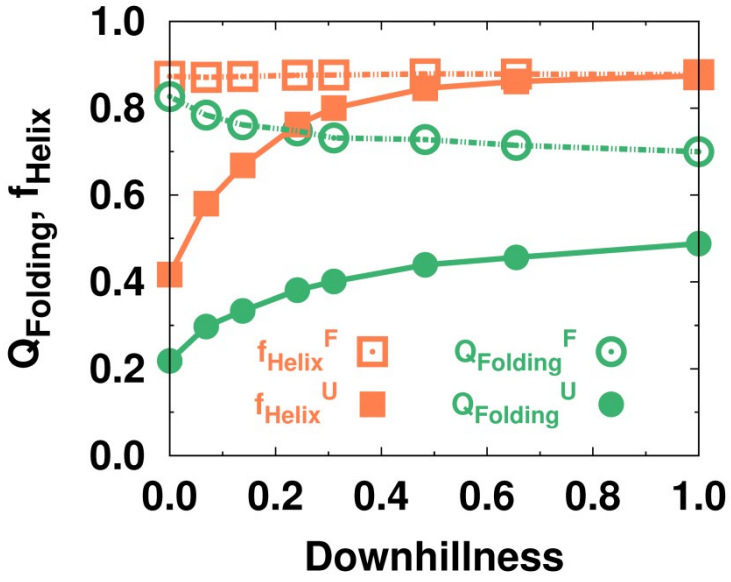

Figure S1: Fraction of native contacts and helical structure as a function of downhillness. f_{Helix} is the number of residues in helical formation divided by 44 (number of native helical residues). An alpha-helix segment is defined as a set of four consecutive CA beads with dihedral angles corresponding to the α-helical region: from -45° to 135° [32, 33].

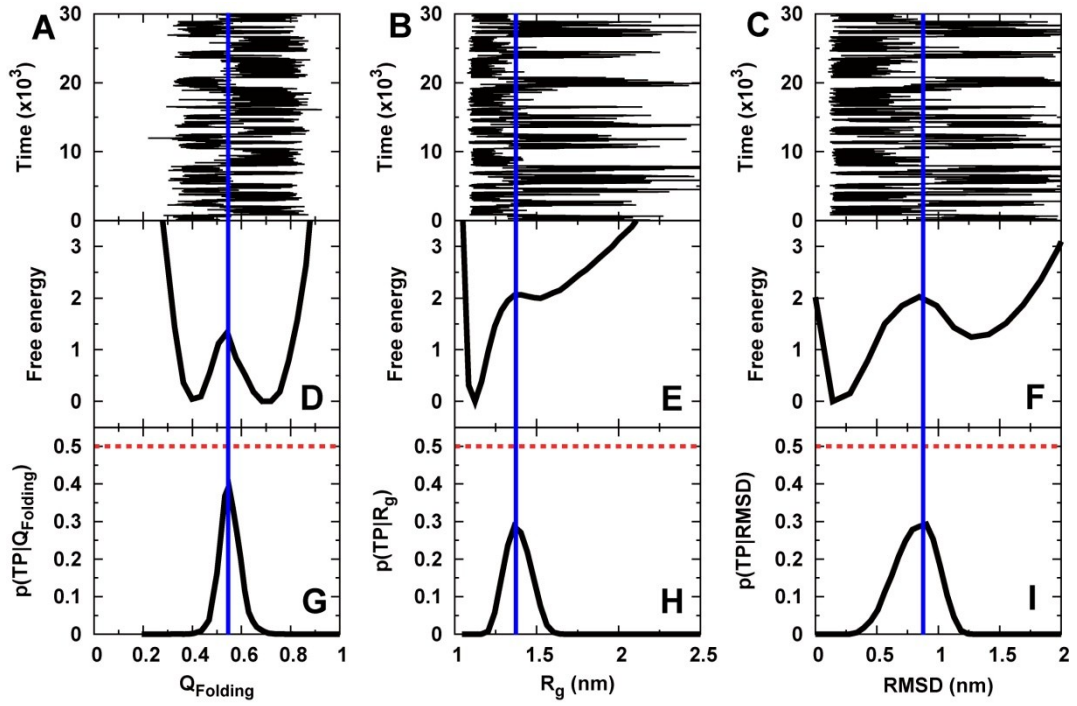

Figure S2. Different reaction coordinates of (left) fraction of native contacts $Q_{Folding}$, (middle) radius of gyration R_g and (right) RMSD for EngHD folding. (A-C) Typical trajectories are shown along different reaction coordinates. Time is in reduced units. (D-F) One-dimensional free energy landscapes at midpoint folding temperature (derived from heat capacity) are projected as a function of each order parameter. Free energy is in kT units. (G-I) Conditional probabilities of being on a transition path $p(\mathrm{TP}|r)$, where r is the reaction coordinate, are shown with different reaction coordinates. Dashed lines signal the theoretical maximum (0.5)[34, 35]. $p(\mathrm{TP}|r)$ has the highest maximal values with $Q_{Folding}$ (0.39) at the transition state according to the free energy surface (barriertop), implying that $Q_{Folding}$ is the best reaction coordinate among these three order parameters. The blue lines indicate the position of transition states, which can be characterized both by free energy landscapes and $p(\mathrm{TP}|r)$.

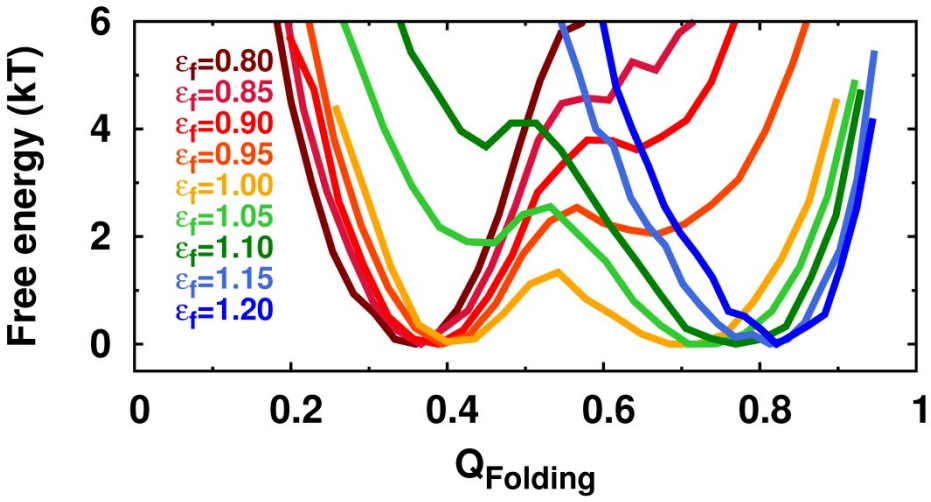

Figure S3: Folding free energy landscape of EngHD as a function of ε_f. $Q_{Folding}$ is the fraction of native folding contacts.

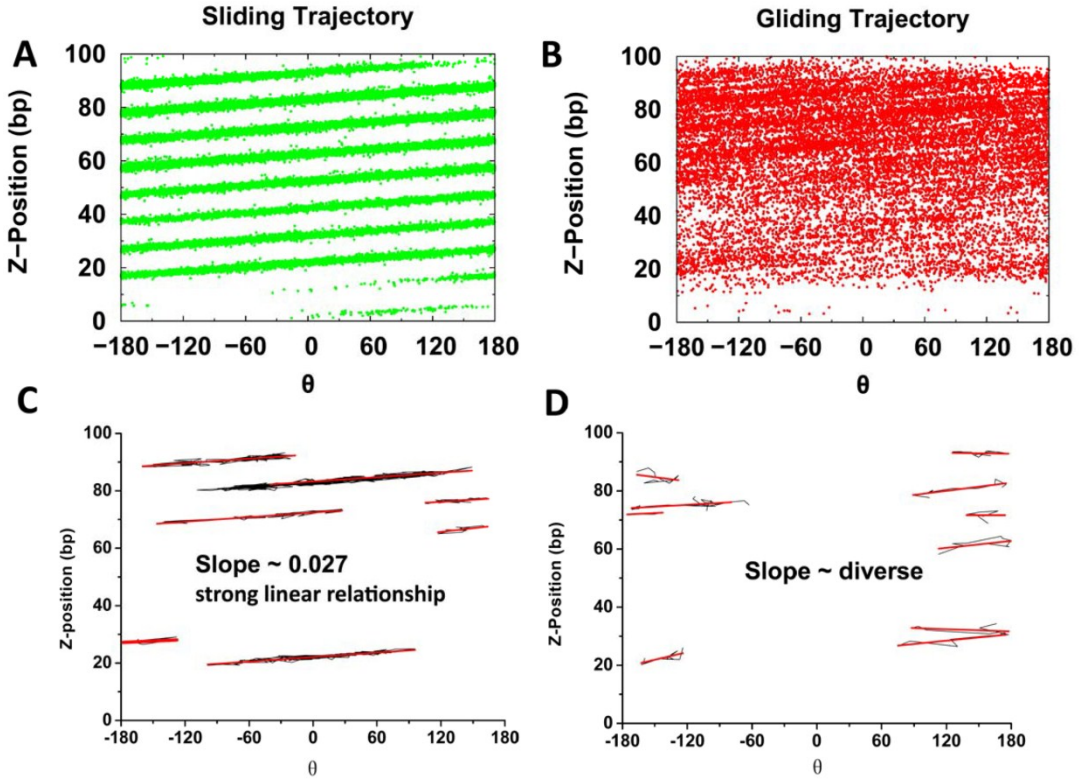

Figure S4: Displacement along the 100-bp DNA length of sliding and gliding trajectories. Top panels: composite of all the sliding (A) and gliding (B) trajectories. θ is the polar coordinate to signal orientation of EngHD around the DNA axis of rotation. Bottom panels: 10 exemplary trajectories of sliding (C) and gliding (D) 1D motion with linear fits to extract the average displacement. Sliding motion has a constant displacement of 0.027 bp/degree that matches the contour of the DNA major groove. Gliding motion fluctuates in the displacement plot indicasting that it does not involve defined rotation about the DNA axis.

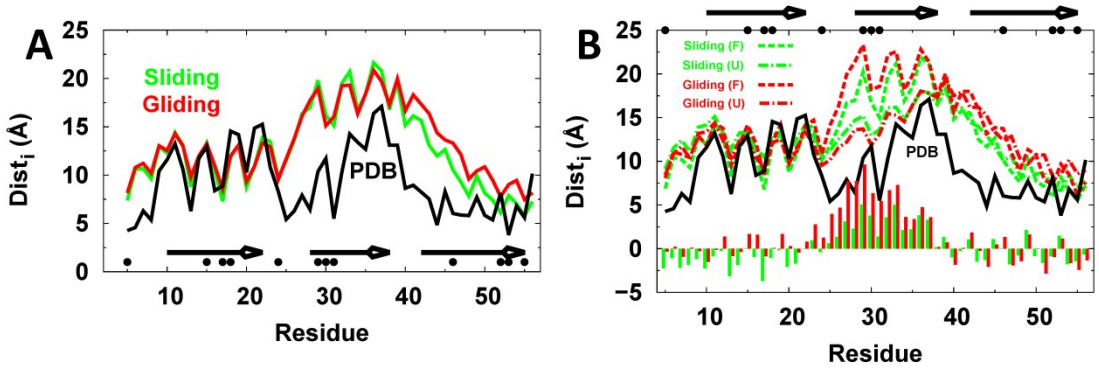

Figure S5: Structural analysis of EngHD during sliding and gliding motions. (A) Minimum distance from each EngHD residue to the closest DNA atom during sliding and gliding. The distances from the X-ray structure of the complex are included as reference. Black circles signal the positions in the protein of the positively charged residues of EngHD, and the three arrows the positions of helices I, II and III, sequentially. The distances between protein residues and DNA during sliding and gliding have lineasr regression coefficients of 0.61 and 0.60 with respect to the X-ray structure, respectively. (B) Same as A but differenciated between folded and unfolded conformations of EngHD. The histograms show the differences in distance to the DNA between the folded and unfolded states calculated as $Dist_i(F) - Dist_i(U)$.

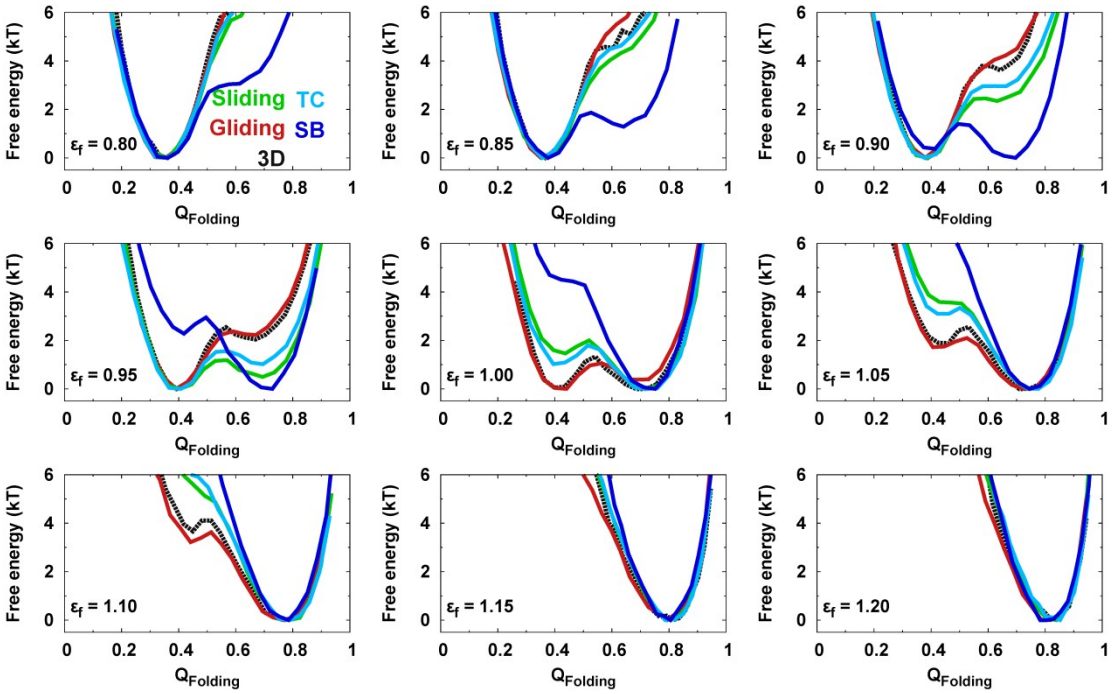

Figure S6: Folding free energy landscapes of EngHD binding to DNA as a function of ε_f at the folding temperature. The colors indicate the four different DNA binding modes: non-specific sliding and gliding, and specific transition complex (TC) and specific binding site (SB). Data for 3D diffusion were obtained from simulations of EngHD in the absence of DNA and are shown as dashed lines.

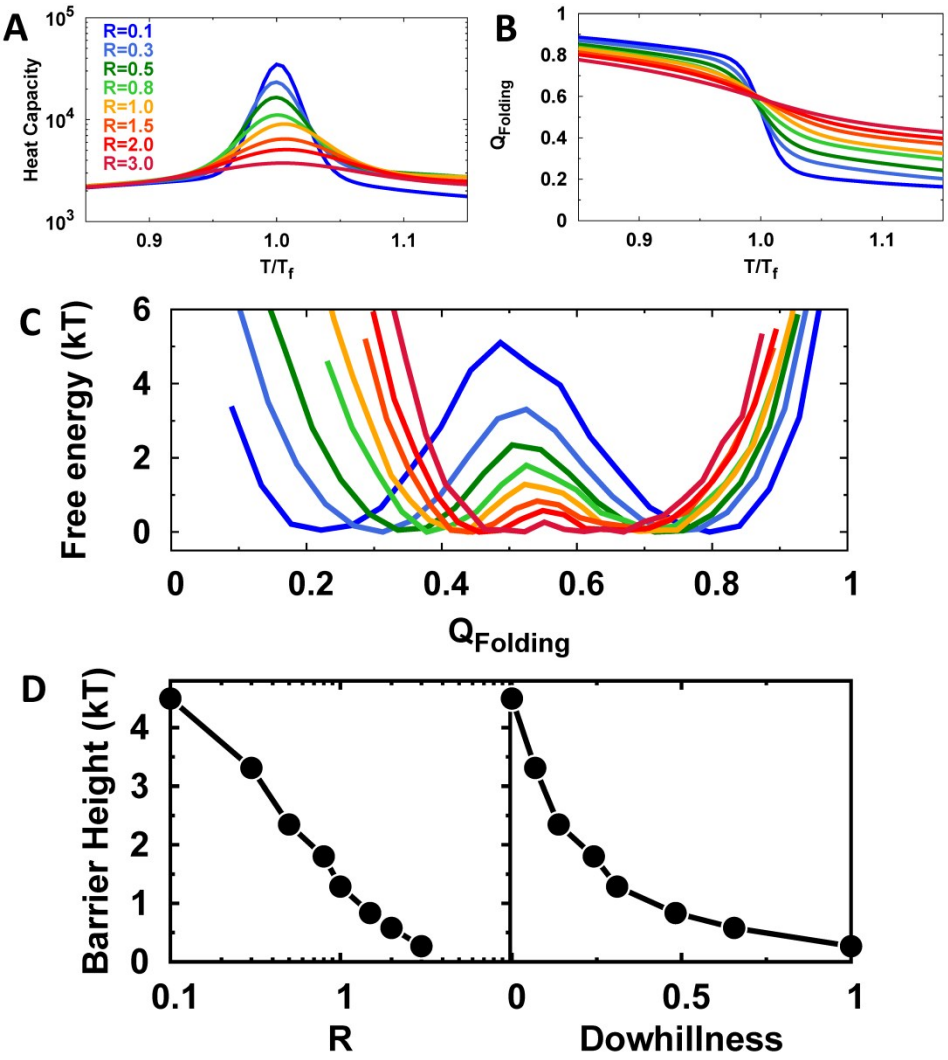

Figure S7: Thermodynamic folding behavior of EngHD as a function of R or downhillness. R is the strength of the dihedral term relative to that of the native contact energy ($\varepsilon_\phi/\varepsilon_f$). (A) Heat capacity as a function of temperature for different values of R. (B) Fraction of native contacts ($Q_{Folding}$) as a function of temperature for different values of R. The width of the heat capacity profile and the broadness of the $Q_{folding}$ unfolding curve as R increases signal the decrease in folding cooperativity. (C) Folding free energy landscapes of EngHD as a function of $Q_{Folding}$ calculated at the folding temperature for different values of R. (D) Height of the folding free energy barrier at the folding temperature as a function of R and downhillness.

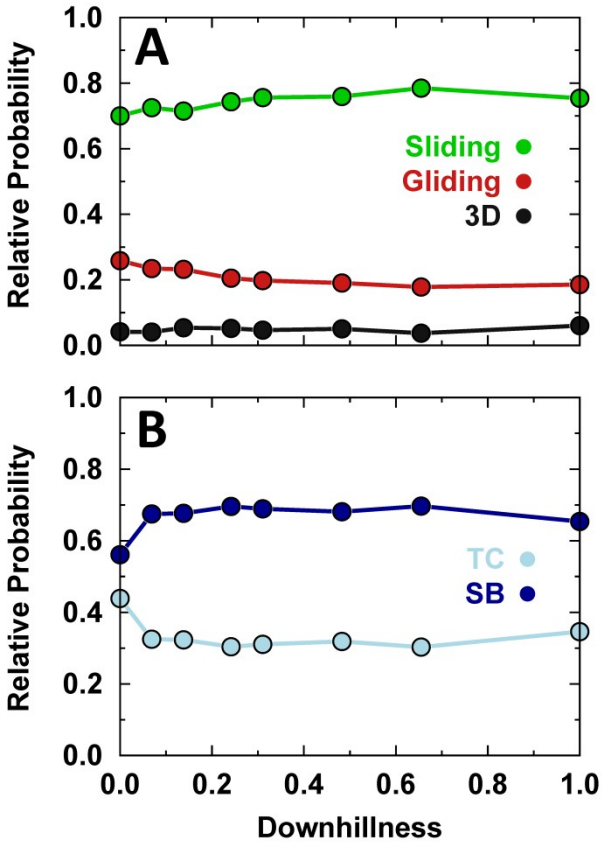

Figure S8: Relative probability of finding the various binding modes at the folding temperature as a function of downhillness. (A) Relative probability for non-specific binding: 3D diffusion, gliding, and sliding. (B) Relative probability for specific binding site: transition complex and specific binding.

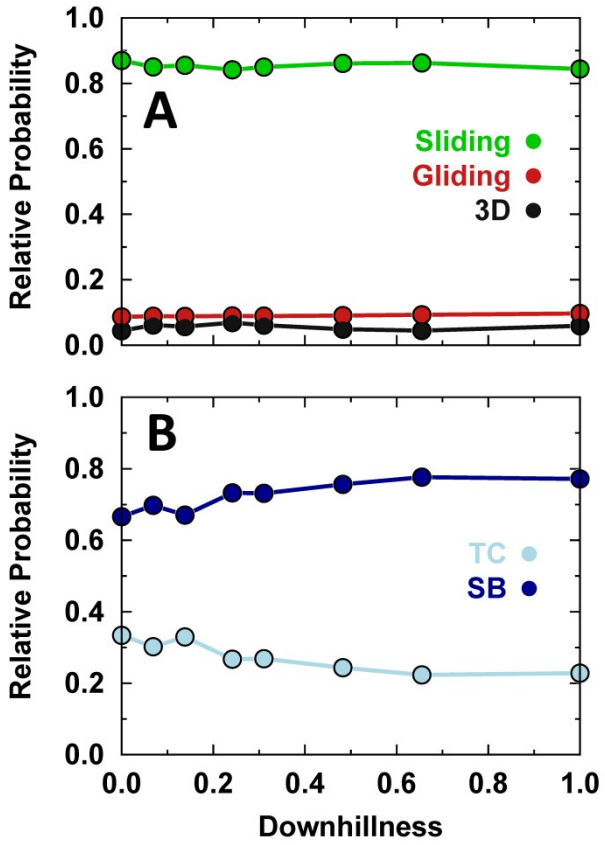

Figure S9: Relative probability of different binding modes at room temperature as a function of folding downhillness. (A) Relative probability for non-specific binding: 3D diffusion, gliding, and sliding. (B) Relative probability for specific binding site: transition complex and specific binding.

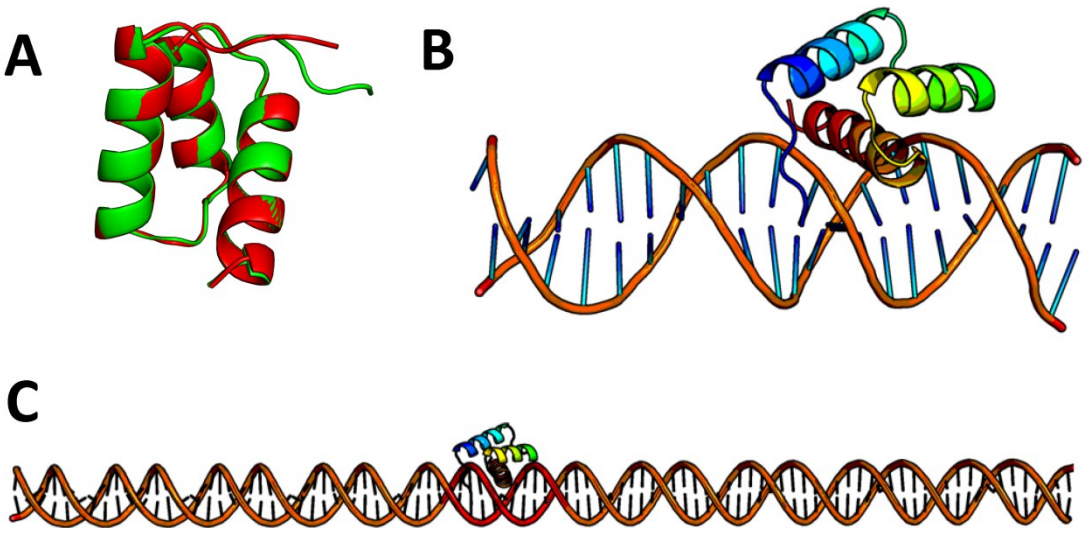

Figure S10: Structural details of folded EngHD and the EngHD-DNA complex. (A) Structural superposition of the EngHD structure in the free state (green) (PDB: 1ENH)[36] and the DNA bound state (red). RMSD between the two structures is 0.7 Å for C_α atoms and 1.7 Å for all heavy atoms. (B) Crystal structure of EngHD-DNA complex (PDB:3HDD)[37]. (C) A 100bp-long DNA duplex was used in DNA binding simulations. The specific binding site is colored in red and non-specific binding sites in orange. EngHD in panels B and C is colored according to the light spectrum from the N-terminus (blue) to C-terminus in (red).

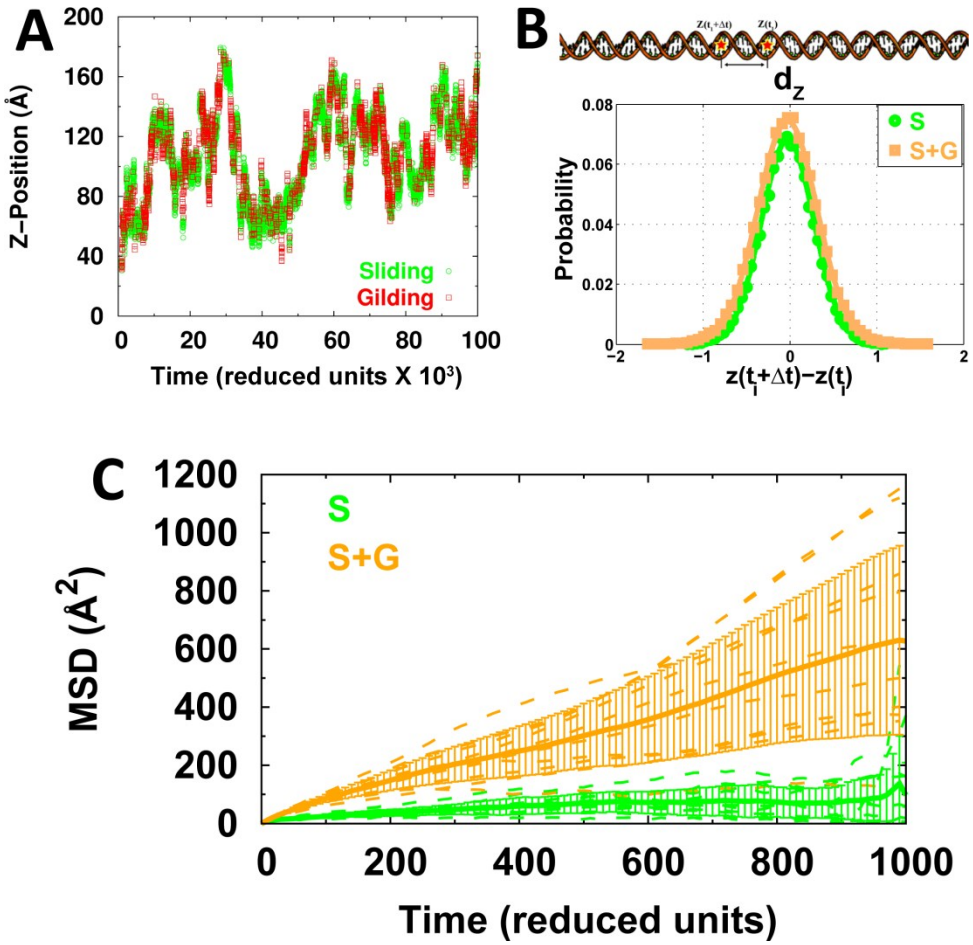

Figure S11: Mean square displacement (MSD) analysis of sliding and gliding motions. (A) Decomposition of a single trajectory into sliding and gliding segments. (B) Calculation of the displacement between two consecutive trajectory frames, where t_1 is the time for the first frame and Δt is the time difference between two consecutive frames. Panel B also shows the distributions of distance displacements along the DNA long axis (Z-direction) for consecutive time frames. "S" and "S+G" stand for sliding dynamics and integrated dynamics of sliding and gliding together. The distributions of displacements are roughly Gaussian, implying the underlying dynamics are Brownian. (C) MSD for sliding (S) and integrated dynamics of sliding and gliding (S+H) as a function of time. The trajectories were saved every 10 reduced time units, producing 100 frames for each trajectory (from 0 to 1000 reduced time units). MSD of 10 individual sliding and sliding-gliding trajectories are shown in dashed lines. The mean MSD and standard deviation for the entire dataset are shown as a thick continuous line and a dashed swath, respectively.

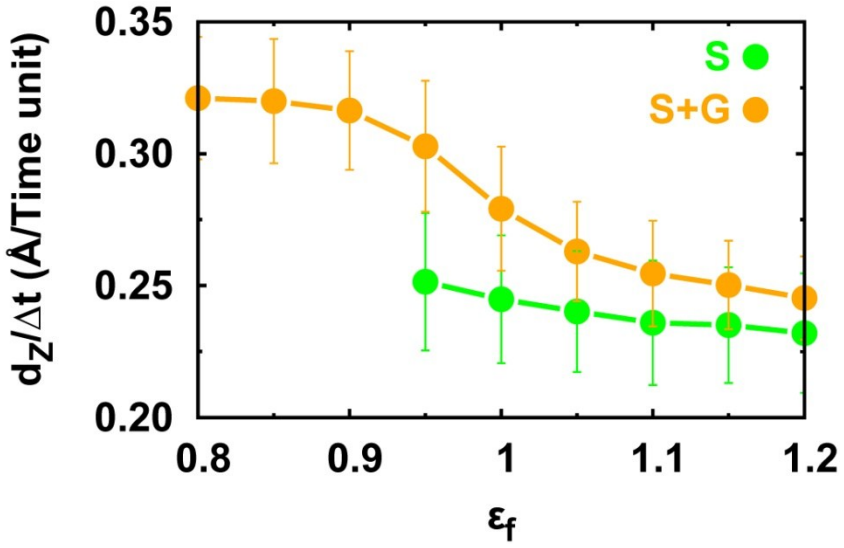

Figure S12: Displacement along the DNA long as a function of ε_f for the sliding and integrated sliding-gliding modes.

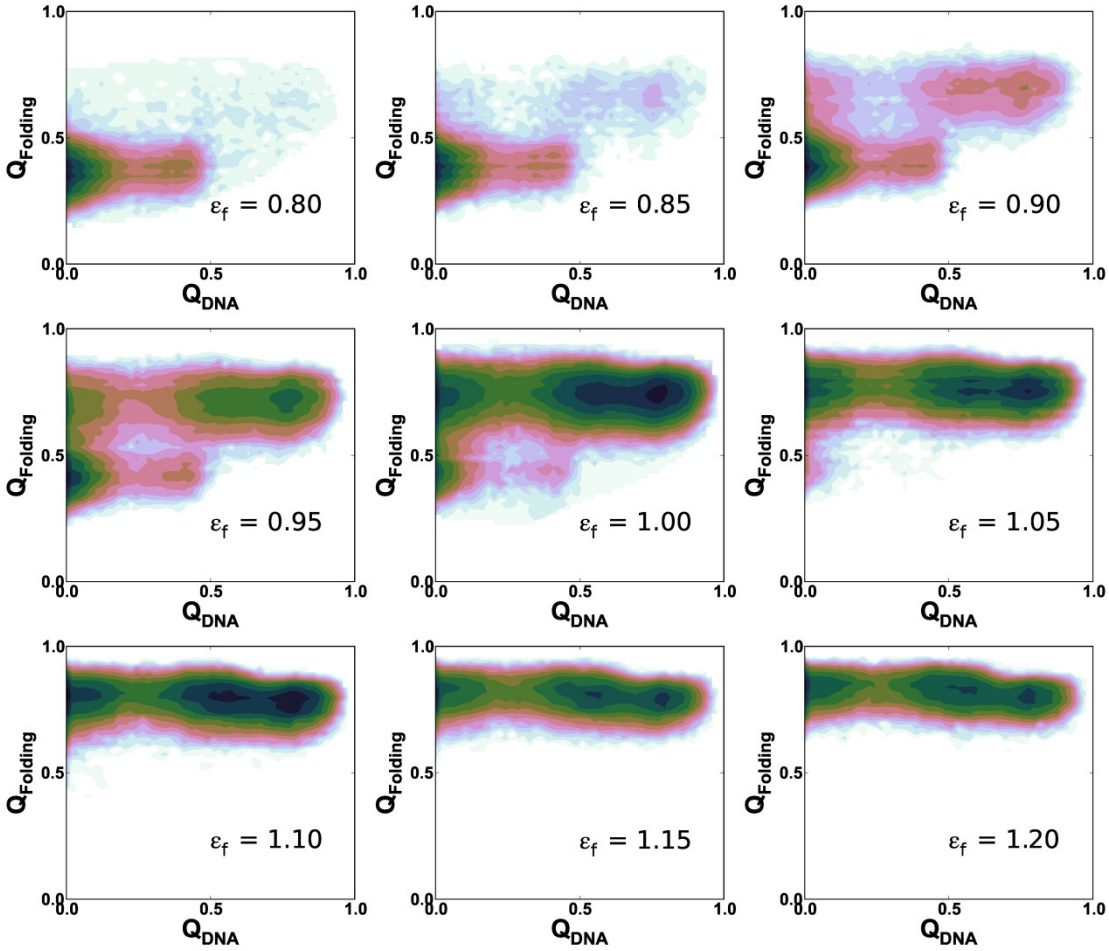

Figure S13: 2D folding and specific binding free energy landscapes for various values of ε_f.

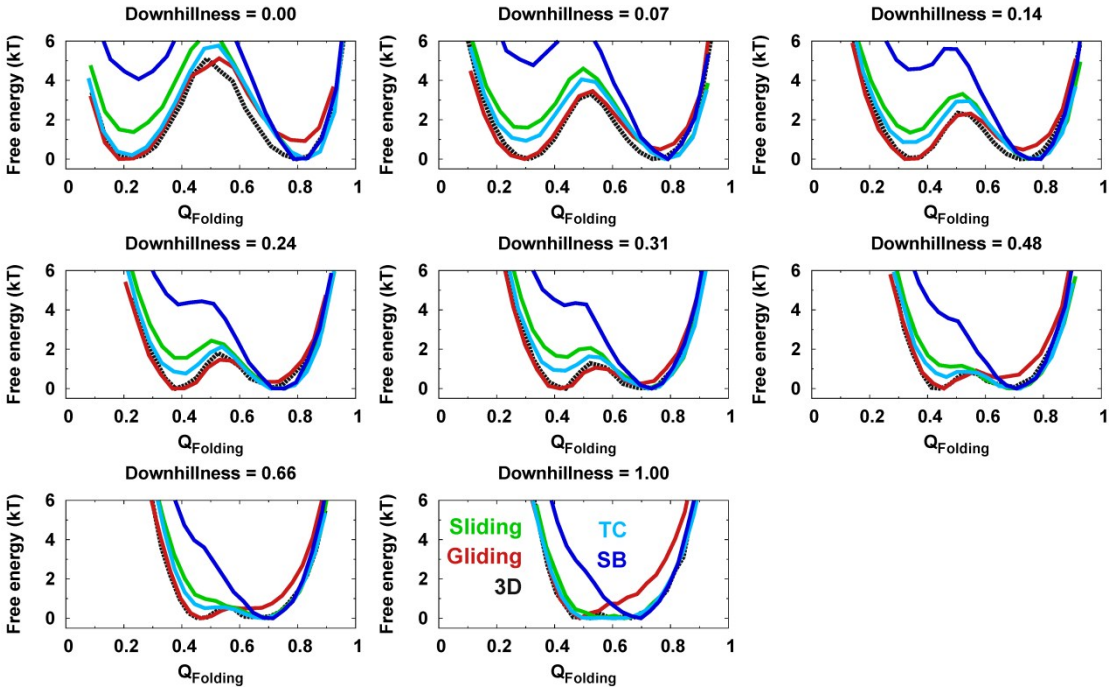

Figure S14: Folding free energy landscapes of EngHD as a function of Qfolding for different values of downhillness. The landscapes for EngHD under different binding modes are shown in different colors: green – sliding, red – gliding, black – 3D diffusion, cyan – transition complex, and blue – specific binding to target.

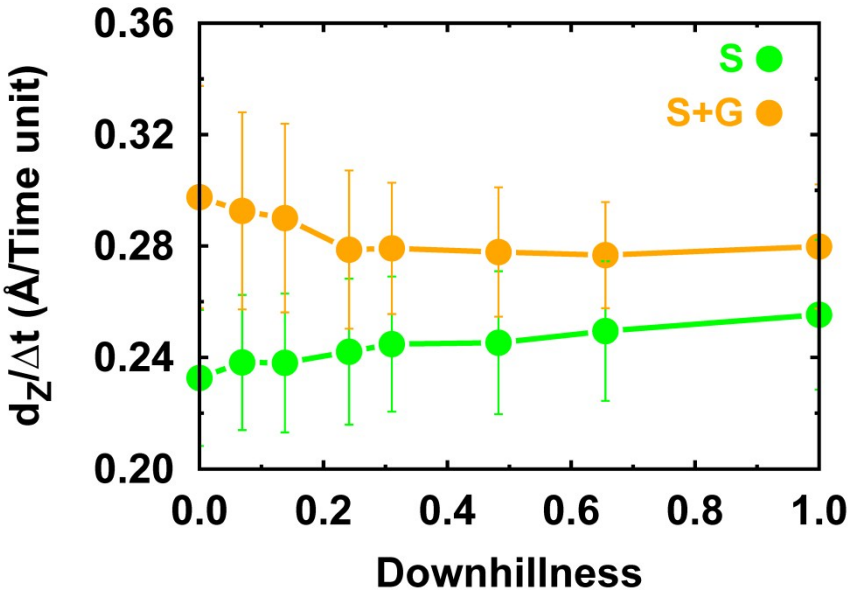

Figure S15: Displacement along the long DNA axis during sliding and gliding as a function of downhillness obtained at the folding temperature.

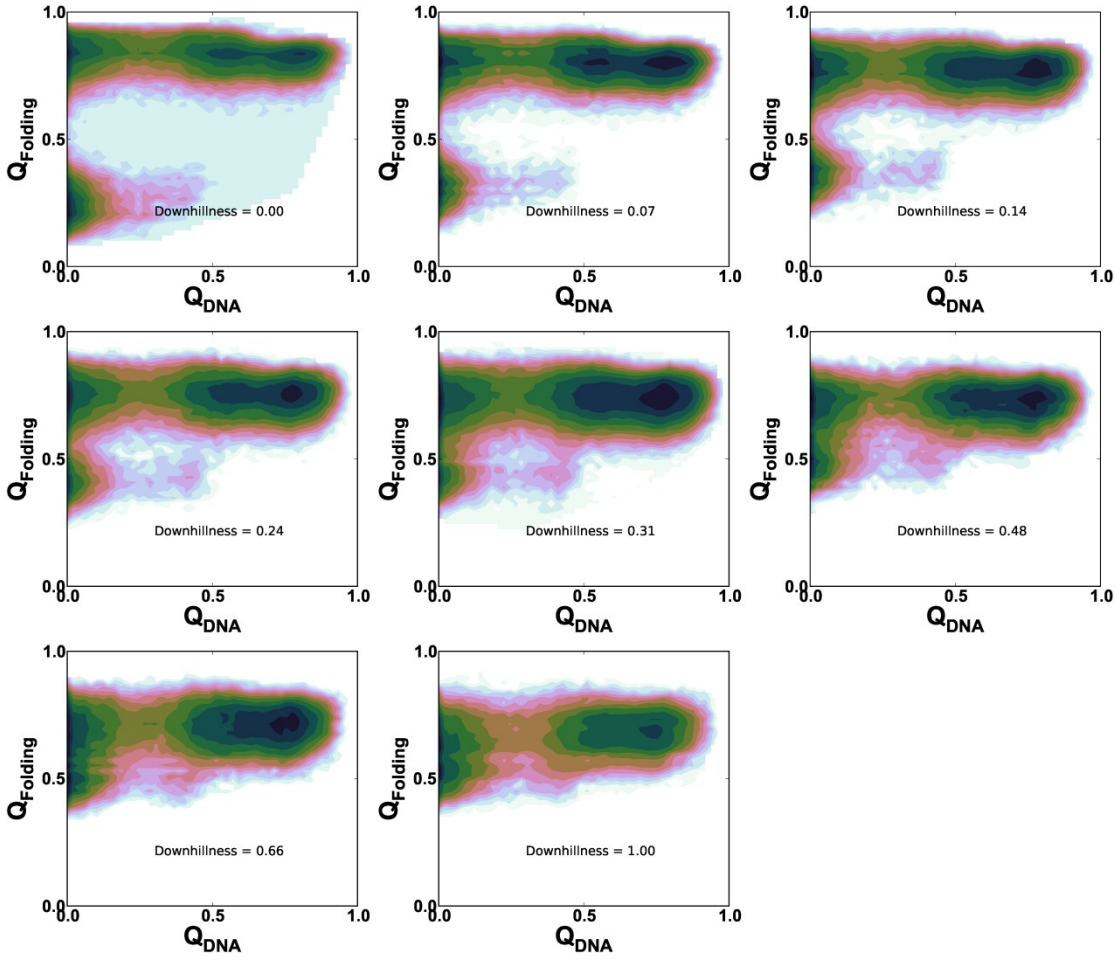

Figure S16: 2D folding and specific binding free energy landscapes for various values of downhillness obtained at the folding temperature.

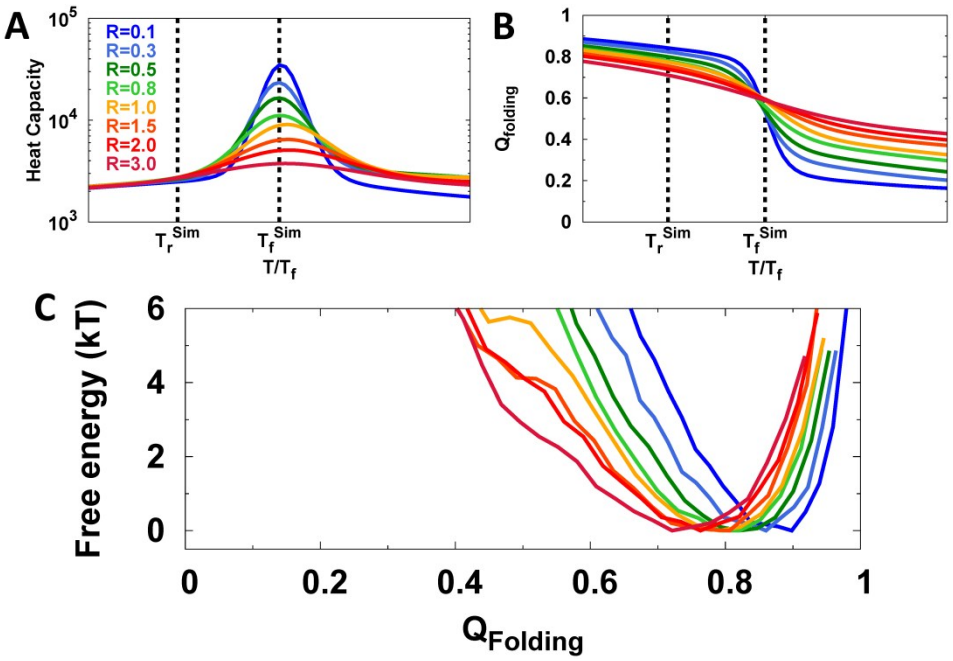

Figure S17: Thermodynamic folding behavior of EngHD at room temperature. R is the strength of the dihedral term relative to that of the native contact energy ($\varepsilon_\phi/\varepsilon_f$). (A) Heat capacity as a function of temperature for different values of R. T_r^{Sim} and T_r^{Sim} signal the room temperature and folding temperature conditions in the simulations (B) Fraction of native contacts ($Q_{Folding}$) as a function of temperature for different values of R, signaling the simulation room temperature and folding temperature conditions (C) Folding free energy landscapes of EngHD as a function of $Q_{Folding}$ calculated at room temperature for different values of R.

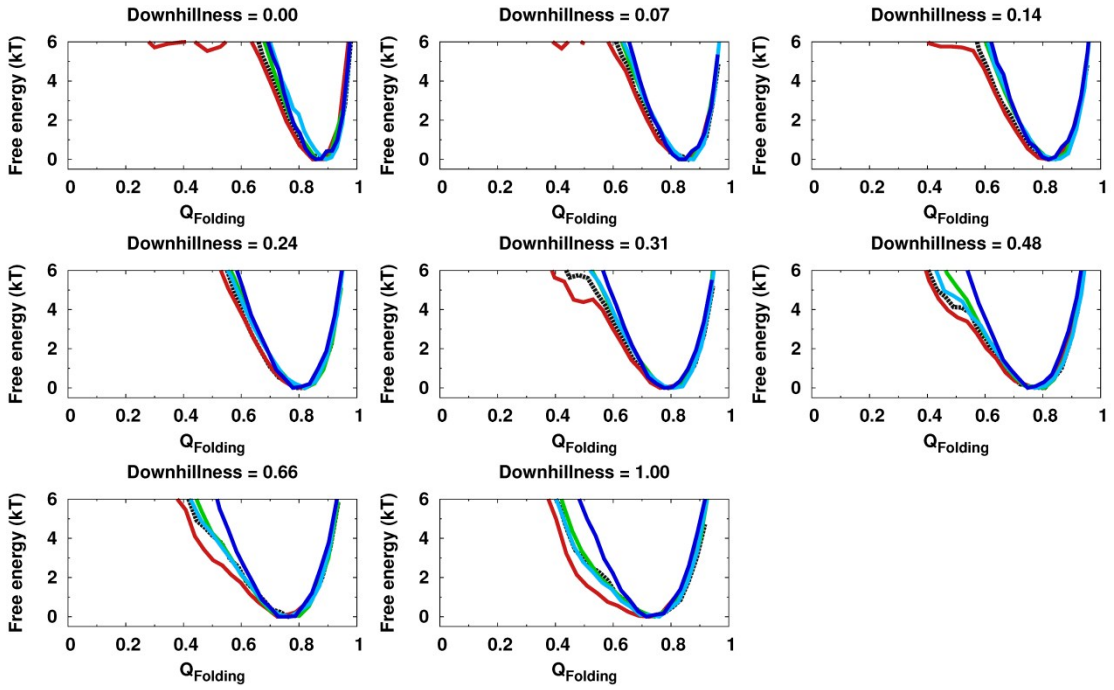

Figure S18: Folding free energy landscapes of EngHD as a function of Qfolding for different values of downhillness at room temperature. The landscapes for EngHD under different binding modes are shown in different colors: green – sliding, red – gliding, black – 3D diffusion, cyan – transition complex, and blue – specific binding to target. Data for 3D diffusion were extracted from simulations of isolated EngH and are shown as dashed black lines.

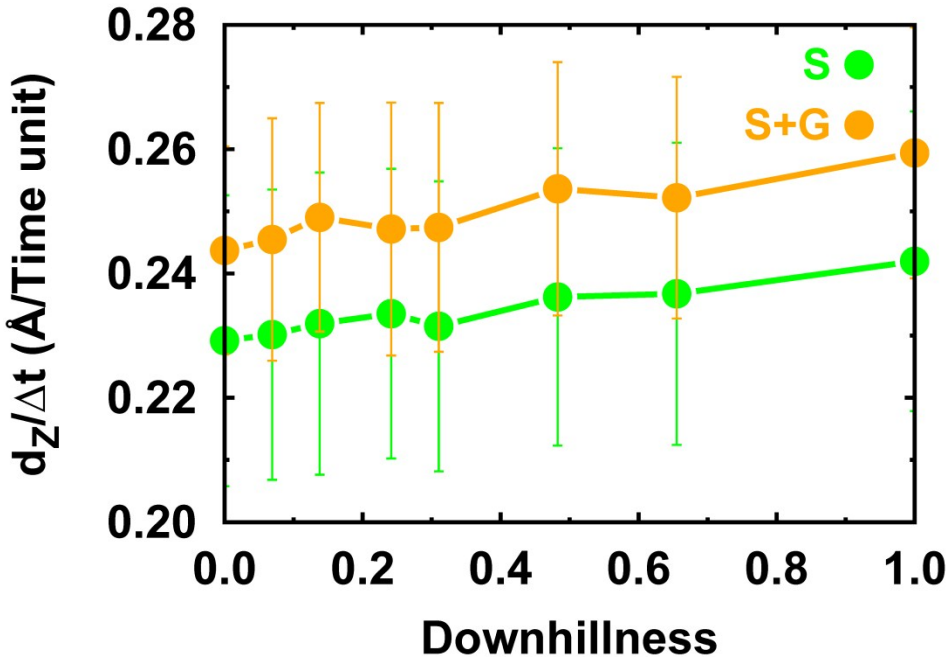

Figure S19: Displacement along the long DNA axis during sliding and gliding as a function of downhillness obtained at room temperature.

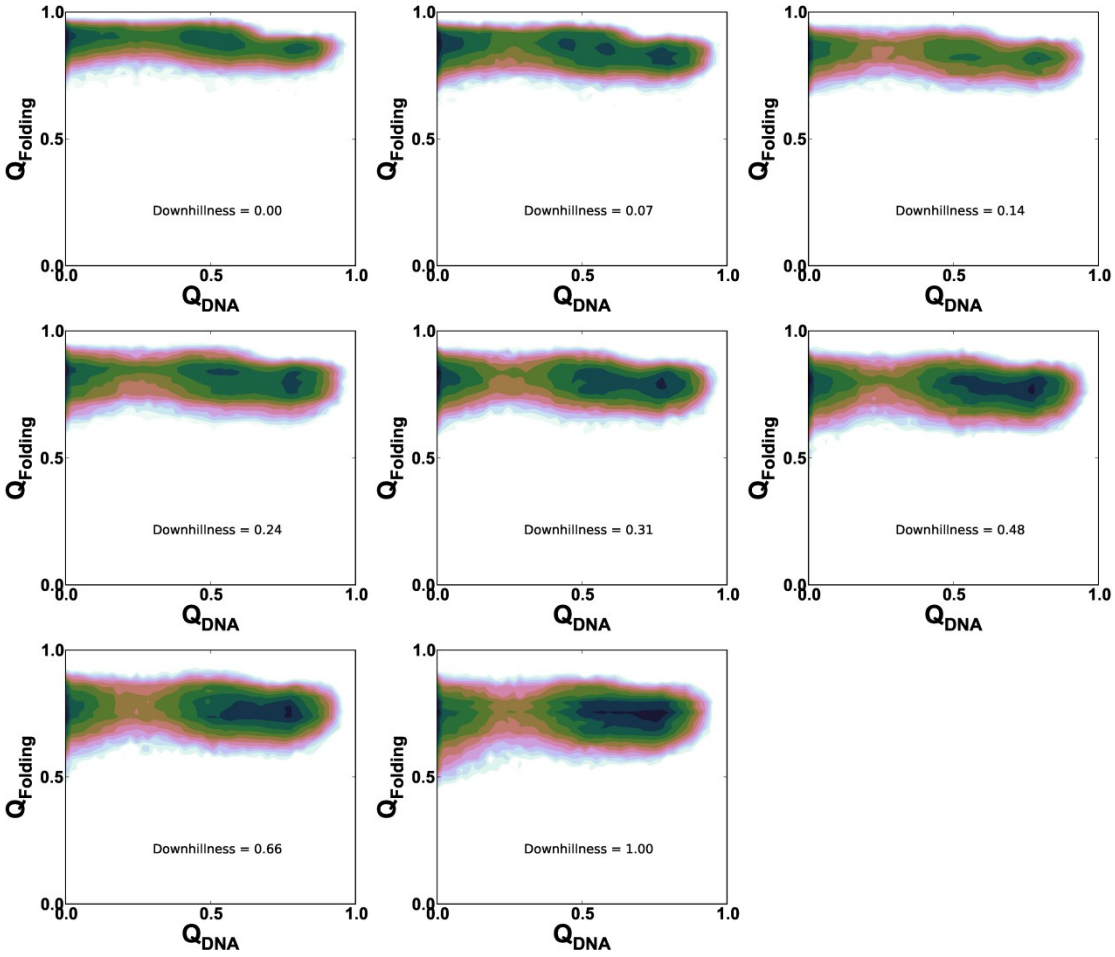

Figure S20: 2D folding and specific binding free energy landscapes for various values of downhillness obtained at room temperature.