Published in final edited form as:

DNA Repair (Amst). 2020 February ; 86: 102764. doi:10.1016/j.dnarep.2019.102764.

Macromolecular crowding induces compaction and DNA binding in the disordered N-terminal domain of hUNG2

Gaddiel Rodriguez[a], **Benjamin Orris**[a], **Ananya Majumdar**[b], **Shridhar Bhat**[a], **James T. Stivers**[a,*]

[a]Department of Pharmacology and Molecular Sciences, The Johns Hopkins University School of Medicine, 725 North Wolfe Street, Baltimore, MD 21205, United States

[b]Biomolecular NMR Center, Johns Hopkins University, Baltimore, MD 21218, United States

Abstract

Many human DNA repair proteins have disordered domains at their N- or C-termini with poorly defined biological functions. We recently reported that the partially structured N-terminal domain (NTD) of human uracil DNA glycosylase 2 (hUNG2), functions to enhance DNA translocation in crowded environments and also targets the enzyme to single-stranded/double-stranded DNA junctions. To understand the structural basis for these effects we now report high-resolution heteronuclear NMR studies of the isolated NTD in the presence and absence of an inert macromolecular crowding agent (PEG8K). Compared to dilute buffer, we find that crowding reduces the degrees of freedom for the structural ensemble, increases the order of a PCNA binding motif and dramatically promotes binding of the NTD for DNA through a conformational selection mechanism. These findings shed new light on the function of this disordered domain in the context of the crowded nuclear environment.

Keywords

hUNG2; N-terminal domain; Intrinsically disordered protein; DNA binding; Conformational selection; Molecular crowding; NMR spectroscopy

1. Introduction

Functional protein domains devoid of higher order structure are abundant in the human proteome [1]. These disordered domains often serve a regulatory role in cellular processes such as RNA-chaperone activity [2,3], targets for post translational modification [4], regulatory elements for multiprotein complex self-assembly [5], and mediators of signal transduction [6]. Furthermore, dysregulation of disordered proteins, through mechanisms involving damaged splicing, aberrant post-translational modification, and proteolytic degradation, is often involved in the pathogenesis of human disease [7].

*Corresponding author. jstivers@jhmi.edu (J.T. Stivers).

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.dnarep.2019.102764.

Mammalian DNA repair glycosylases that catalyze the excision of damaged DNA bases possess disordered protein domains located at either their N or C termini [8–11]. One key example is the 92-amino acid N-terminal domain (NTD) of nuclear uracil DNA glycosylase 2 (hUNG2) that removes uracil bases from U/A and U/G pairs in DNA [12]. Although this domain contains an alpha helix that spans residues 66–88 [13], the rest of the domain (residues 1–65) is predicted to be disordered based on prior NMR chemical shift data and sequence-based structure prediction algorithms [13]. The biological function of the NTD is broad: it carries the nuclear localization signal sequence and 14 sites that can be acetylated, ubiquitinated, or phosphorylated [14–16]. In addition, binding motifs for proliferating cell nuclear antigen (PCNA; res 4–11) and replication protein A (RPA; residues 66–88) [17] are also present and phosphorylation within these binding motifs disrupts the interactions of each protein with hUNG2 [16,18]. Cell cycle-dependent proteolysis of hUNG2 during late S phase is triggered via phosphorylation of Thr60 and Ser64 [19].

We have recently shown that the NTD facilitates several enzymatic functions of hUNG2 and its contributions increase dramatically in the presence of macromolecular crowding [9,10,16]. These facilitated activities specifically attributed to the NTD include enhanced translocation along DNA chains and preferential targeting to uracil sites near ssDNA-dsDNA junctions [9,10]. To understand these effects and the nuclear function of the NTD it is necessary to elucidate its structure and dynamic properties under conditions that mimic the crowded cellular environment. Accordingly, we now report high-resolution heteronuclear NMR experiments of the isolated NTD using solution conditions that contain the inert macromolecular crowder polyethylene glycol 8000 (PEG8K) [20–24].

2. Materials and methods

2.1. DNA oligonucleotides

DNA oligonucleotides were purchased from Integrated DNA Technologies. All DNA substrates were prepared by PAGE purification of the respective ssDNA as described in detail in the Supplemental materials.

2.2. Cloning, expression and purification of isotopically labeled hUNG NTD

The sequence encoding the 92-amino acid N-terminal domain (NTD) of hUNG2 was inserted into the pET21A vector containing an 8x His-tag, a GSGSG linker, with the SUMO protein SMT3 (from S. cerevisiae) fused to its N-terminus and purified by affinity and ion exchange chromatography (Supplemental materials).

2.3. NMR experiments and analysis

Triple resonance heteronuclear NMR experiments were performed at 600 and 800 MHz field strengths using ^{13}C and ^{15}N-labeled NTD in a buffer consisting of 25 mM sodium phosphate pH 7.0, 50 mM NaCl, 90/10 % H_20/D_20 at 22°C. Experiments and assignments were preformed using three conditions: no PEG8K, in the presence of 20 % (w/v) PEG8K, and 20 % (w/v) PEG8K in the presence of saturating ssDNA (5′-GCT CTG TAC ATG AGC AGT G −3′). Concentrations of the isotopically labeled NTD peptide were in the range 166–300 μM. The data were processed and analyzed as described in the Supplemental materials.

The analysis of the chemical shifts in terms of secondary structure and dynamics and are described in the text and Supplemental materials.

2.4. Measurements of DNA binding using fluorescence anisotropy

Binding of the hUNG2 NTD peptide to 5′-FAM labeled duplex and ssDNA was performed at 25 °C using a buffer consisting of 25 mM sodium phosphate pH 7.0 and 50 mM NaCl using a SPEX Fluoromax-3 fluorimeter. Settings for the anisotropy measurements were λ^{ex} = 494 nm, λ^{em} = 518 nm using magic angle conditions (excitation polarizer = 0°, emission polarizer = 57.5°), with G factor corrections for the polarizers. All binding reactions were performed using either the single-stranded or double stranded form of the NS19FAM DNA (5′ - FAM - GCT CTG TAC ATG AGC AGT G - 3′). The fraction bound DNA was plotted against total protein concentration and fit to a quadratic binding isotherm as described in detail in the Supplemental materials.

3. Results

3.1. The hUNG2 NTD has a more compact average structure in the presence of macromolecular crowding

A suite of heteronuclear NMR experiments (^{1}H, ^{15}N -TROSY-HSQC, HNCACB and HNCO) were performed to make residue specific HN, $C_{<}$, $C_{®}$ and C’ assignments of the NTD in the presence and absence of the inert crowding agent PEG8K (20 % w/v) (Supplemental Table 1). Resonance assignments in the absence of PEG8K were facilitated by the prior assignments deposited in the Biological Magnetic Resonance Bank [13]. Assignment of residues 2–66 in the disordered region of the sequence were made challenging by the presence of thirteen structure-disrupting proline residues (Fig. 1c; residues 13, 15, 22, 24, 26, 37, 47, 52, 58, 61, 62, 65, and 91). Nevertheless, consecutive backbone assignments for over 96 % of the peptide chain were made for both conditions (93 % of residues have assignments for 3 or more of the 5 sampled NMR nuclei), which allowed for comparison of the backbone chemical shifts, amide proton linewidths, backbone dynamics, and secondary structure propensity in the absence and presence of PEG8K (Fig. 1) (Supplemental Figures S1, S2).

In the presence of PEG8K, significant changes in the backbone amide ^{1}H and ^{15}N chemical shifts of the NTD were observed, indicating that the average structure of the ensemble was altered (Fig. 1a,b). The largest shift changes were seen in the disordered region (residues 2–65) and in the first few residues leading into the α-helix (Ser67, Ala68). The flexible regions of the backbone most likely experience the largest shift changes because their shifts most resemble the random coil values in the absence of crowding. Thus, upon addition of the crowding agent the compaction of the conformational ensemble weights the shifts to values that indicate increased structure. Structural compaction is consistent with expectations from simple entropically driven hard sphere models for crowding, where solutes respond by maximizing the available volume [1,25–27].

Despite the increase in macroscopic solution viscosity with 20 % PEG8K [22], the backbone amide proton linewidths were uniformly reduced in its presence (Fig. 2a). Narrower

linewidths may indicate (i) a decrease in the chemical shift distribution of the conformational ensemble leading to a decrease in the exchange contribution to R2, (ii) a decrease in the correlation time for molecular tumbling due to compaction, or (iii) a decrease in transient aggregation of the peptide in the presence of PEG8K. Transient aggregation in the absence of PEG8K (reduced molecular tumbling) seems unlikely because the linewidth decreases are not uniform across the sequence and the largest changes are associated with regions of known function (discussed below).

The absence of viscosity-induced increases in linewidths is once again consistent with simple polymer theory, where nanometer scale rotational or translational movements can occur within the spaces between the macromolecule crowders [23,28,29]. The reduction in linewidths in the presence of PEG8K provides further evidence that peptide interactions with the polymer are not significant because such interactions would increase the rotational correlation time and lead to broader lines. An especially large decrease in the linewidth of residue Ser64 was observed, despite Ser64 showing only a modest chemical shift change (Fig. 2a). This observation indicates a decrease in the effective correlation time at Ser64 (i.e. increase in rapid internal motions). These dynamic properties of Ser64 may be relevant to its phosphorylation and biological function. Moreover, the dynamic behavior of Ser64 may be related to the backbone motions of the nearby residue Ala68, which is discussed separately below in the context of DNA binding and conformational selection.

3.2. The NTD exists in multiple conformations in slow exchange on the NMR chemical shift time scale

Some of the assigned resonances in the ^{15}N, ^{1}H-TROSY-HSQC spectra in the presence of PEG8K were accompanied by minor peaks, suggesting that slow exchange between peptide conformations was occurring in some sequences. Given the large number of proline residues, we suspected that these peaks might reflect proline cis-trans isomerization [30]. However, the chemical shifts of the shadow peaks were not consistent with the well-characterized effects of Pro isomerization on the amide chemical shifts of adjacent residues (i ± 2; where i = Pro) [30,31]. To confirm our analysis, the proline chemical shift data were input into the Promega server [32], which predicted that all the prolines were in the trans conformation. We conclude that minor peptide conformers are stabilized by crowding, but these conformations do not result from proline cis-trans isomerization.

The chemical shift data (HN, N, Cα, Cβ, and C') for the NTD in the absence and presence of PEG8K were used to calculate the chemical shift-derived backbone order parameters S2(δ) with the TALOS + software suite (Fig. 2a) [32–34]. The S2(δ) values for each residue reflect the averaging of chemical shifts between the residue-specific random coil values and the values observed in folded proteins. The parameter S2(δ) is strongly correlated with NMR relaxation-based dynamic measurements [32–34]. TALOS + S2(δ) values greater or less than 0.5 indicate high or low average order, respectively. The backbone S2(δ) values for residues 2–65 show an interesting oscillation between regions of order and disorder in the presence of PEG8K that are only weakly present in the absence of PEG8K (Fig. 2b). The peaks and valleys of the oscillations occur roughly every 10 residues and could reflect the formation of transient tertiary interactions or an increase in local secondary structure. Of

note, the largest increase in order arising from crowding involved the first six amino terminal residues, which comprises part of the putative PCNA interacting protein (PIP) motif (Fig. 2b).

We also used the sequence-based order prediction program Espritz [35], along with the NMR train module, to calculate the order propensity of the NTD (P_{order}, Fig. 2c). This analysis successfully identified the helix containing the RPA binding motif (residues 66–88), the higher order present in the first six residues, and the higher average order in the region beginning around residue thirty-five and continuing until residue fifty (Fig. 2c). However, this prediction did not identify the subtle oscillations determined from the NMR chemical shift data. Similar results to Espritz were obtained with other sequence-based algorithms (SPOT-dis and Disopred 3)(Supplemental Figure S3) [36,37]. As recently reviewed, these programs are among the best at predicting regions of order and disorder and the combined results are largely consistent with the NMR-based S2(δ) method [38]. Finally, the residue specific chemical shifts in PEG8K were also used to predict the NTD secondary structure using TALOS+ (Fig. 2d). This analysis predicted low levels of extended structure across the disordered region of the NTD, which was confirmed in similar chemical shift analyses using the programs SSP and δ2D (Supplemental Fig. S4) [39,40].

3.3. The NTD binds duplex and ssDNA only in the presence of macromolecular crowding

Our previous studies of full-length hUNG2 have established that the NTD confers special DNA recognition properties to the catalytic domain in the presence of crowding that lead to enhanced DNA translocation and targeted engagement of uracils next to ssDNA-duplex junctions [9,10]. To determine if the NTD interacts with DNA on its own and if PEG8K enhances this interaction, we measured its equilibrium binding affinity for duplex and ssDNA using fluorescence anisotropy of the labeled DNA (Fig. 3a). In the absence of PEG8K the NTD had little affinity for either double-stranded or ssDNA. The addition of 20 % PEG8K to the solution dramatically increased binding to both types of DNA with measured KD values of 23 μM (ssDNA) and 10 μM (duplex DNA). These findings indicate that the NTD modulates the DNA binding behavior of the catalytic domain by forming weak non-specific DNA interactions.

To probe the regions of the NTD that might be involved in DNA binding we used the 1H,15 N-HSQC experiment to follow perturbations in the backbone NMR chemical shifts upon addition of ssDNA to a solution of the NTD in the presence of PEG8K (Supplemental Figure S5). These data were then compared to an identical sample in the absence of DNA (Fig. 3b). In general, small changes were observed except for (i) residues S67 and A68, which are located at the junction between the extended coil region and the RPA-binding helix, and (ii) residues 84–89, which are located at the C-terminus of the helix. Shown in Fig. 3c and d, is a comparison of the chemical shift changes for Ala68 that occur upon the addition of PEG8K and ssDNA. Upon addition of PEG8K, the single resonance of A68 (black, Fig. 3c) is split into two major upheld resonances and two minor downfield resonances (red, Fig. 3c). Intriguingly, upon addition of ssDNA the major upheld resonances for A68 merge and the minor downheld resonances become dominant (blue, Fig. 3d). Thus, these data provide evidence for conformational dynamics at the coil-helix junction and a

binding reaction that stabilizes a minor conformation that is already present in the free NTD. It is intriguing that this region, which is most affected by DNA binding, is also a known site for serine phosphorylation (Ser67). Phosphorylation of N-terminal serine residues like Ser67 has been shown to stabilize model helices and could have the same effect in this context. It would also be expected to negatively impact DNA binding.

We also cloned and purihed the NTD region containing the RPA binding helix (residues 66–95) and measured its affinity for duplex DNA in the presence of PEG8K (Supplemental Methods and Fig. S10). This helix showed much reduced DNA binding affinity with an estimated K_D of 850 μM. This result indicates that while RPA $^{66\text{-}95}$ can weakly engage DNA, the full DNA binding affinity of the NTD in the presence of macromolecular crowding also requires NTD residues 2–65. The weak binding affinity of the isolated helix may be partially related to its reduced structural stability in the absence of the NTD residues 2–65 [17]. This notion is consistent with a previous NMR study indicating that the isolated peptide did not have stable helical structure until it bound to RPA [17]. In contrast, our chemical shift data indicate high helical propensity when the helix is in the context of the entire NTD sequence. We suggest that the compaction induced by PEG8K in the unstructured region spanning residues 2–65 produces structural elements that act in concert with the RPA helix to bind both ssDNA and duplex DNA.

We used the docking program HDOCK to model the binding mode of the RPA binding helix within the major groove of B DNA (Fig. 4 and Supplemental methods). The *ab initio* calculated structure suggests that several basic groups are used to make electrostatic interactions with the phosphate backbone (R76, K78, R84, R88). These residues are both conserved and implicated in DNA binding based on the backbone chemical shift perturbations (Supplemental Fig. S11 and Fig. 3b). Although this structural aspect of binding to duplex DNA is intuitive, it is much less clear how the RPA helix and the remainder of the NTD interact with ssDNA. This function is important because preferential excision of uracil bases located within 10 bp of a single strand/duplex DNA junction containing a 5′ single strand overhang is known to require the entire NTD [9].

4. Discussion

This is the first study to examine how intracellular macromolecular crowding could aff ;ect the conformational properties and function of disordered domains found on many DNA repair proteins. Our findings indicate that crowding reduces the conformational dynamics of the NTD backbone, promotes conformations that resemble those populated during DNA binding, and increases the DNA binding activity of the NTD. In the context of the entire protein, the NTD promotes DNA translocation in crowded environments by serving as a flexible tether to the DNA that increases the probability that the enzyme will move along the DNA chain rather than depart to bulk solution [10,41,42]. Since macromolecular crowding appears to trigger collapse of the NTD into a more compact ensemble average structure, the entropic price for forming the DNA binding conformation might be reduced, leading to the observed increase in binding affinity. We speculate that the ebb and flow of NTD ordering observed in the presence of PEG8K (Fig. 2a) could indicate that the overall internal configurational entropy is conserved in the crowded environment. Such a trade-off would

allow the important functional regions to gain ordered structure, while minimizing the internal entropic penalty by increasing the disorder of other regions, all the while gaining the entropic advantage of compaction in the crowded environment. We note that crowding induced compaction has been reported for many but not all intrinsically disordered proteins as recently reviewed by Theillet et al. [1].

One additional aspect of cellular crowding is the possibility of localized regions of high crowding that could facilitate specific interactions [43]. In the case of replication coupled DNA repair involving hUNG2 [18], formation of a large multiprotein DNA replication complex could provide a local macromolecule concentration that specifically triggers compaction of the NTD onto DNA. Such localized crowding could facilitate quick and specific localization of the enzyme to replication foci to facilitate detection of pro-mutagenic uracil bases. Furthermore, transient collapse of the flexible NTD onto DNA without forming strong DNA contacts is compatible with a free energy landscape for the damage search process that involves low energetic barriers that promote rapid inspection of DNA integrity through the sliding and hopping pathways [44].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding

This work was supported by NIH grant GM056834 (to J.T.S). B. O. was supported by NIH Training Grant T32 GM007445.

References

[1]. Theillet F-X, et al., Physicochemical properties of cells and their effects on intrinsically disordered proteins (IDPs), Chem. Rev 114 (2014) 6661–6714. [PubMed: 24901537]

[2]. Dunker AK, et al., Intrinsically disordered protein, J. Mol. Graph. Model 19 (2001) 26–59. [PubMed: 11381529]

[3]. Ward JJ, Sodhi JS, McGuffin LJ, Buxton BF, Jones DT, Prediction and functional analysis of native disorder in proteins from the three kingdoms of life, J. Mol. Biol 337 (2004) 635–645. [PubMed: 15019783]

[4]. Pejaver V, et al., The structural and functional signatures of proteins that undergo multiple events of post-translational modification, Protein Sci. 23 (2014) 1077–1093. [PubMed: 24888500]

[5]. Babu MM, van der Lee R, de Groot NS, Gsponer J, Intrinsically disordered proteins: regulation and disease, Curr. Opin. Struct. Biol 21 (2011) 432–440. [PubMed: 21514144]

[6]. Darling AL, Uversky VN, Intrinsic disorder and posttranslational modifications: the darker side of the biological dark matter, Front. Genet 9 (2018) 158. [PubMed: 29780404]

[7]. Kulkarni P, Uversky VN, Intrinsically disordered proteins: the dark horse of the dark proteome, Proteomics 18 (2018) e1800061. [PubMed: 30218496]

[8]. Hegde ML, et al., The disordered C-terminal domain of human DNA glycosylase NEIL1 contributes to its stability via intramolecular interactions, J. Mol. Biol 425 (2013) 2359–2371. [PubMed: 23542007]

[9]. Weiser BP, Rodriguez G, Cole PA, Stivers JT, N-terminal domain of human uracil DNA glycosylase (hUNG2) promotes targeting to uracil sites adjacent to ssDNA-dsDNA junctions, Nucleic Acids Res. 46 (2018) 7169–7178. [PubMed: 29917162]

[10]. Rodriguez G, et al., Disordered N-Terminal domain of human uracil DNA glycosylase (hUNG2) enhances DNA translocation, ACS Chem. Biol 12 (2017) 2260–2263. [PubMed: 28787572]

[11]. Hegde ML, Hazra TK, Mitra S, Functions of disordered regions in mammalian early base excision repair proteins, Cell. Mol. Life Sci 67 (2010) 3573–3587. [PubMed: 20714778]

[12]. Friedman JI, Stivers JT, Detection of damaged DNA bases by DNA glycosylase enzymes, Biochemistry 49 (2010) 4957–4967. [PubMed: 20469926]

[13]. Buchinger E, et al., Backbone (1)H, (13)C and (15)N chemical shift assignment of full-length human uracil DNA glycosylase UNG2, Biomol. NMR Assign (2017) 0, 0–0.

[14]. Otterlei M, et al., Nuclear and mitochondrial splice forms of human uracil-DNA glycosylase contain a complex nuclear localisation signal and a strong classical mitochondrial localisation signal, respectively, Nucleic Acids Res. 26 (1998) 4611–4617. [PubMed: 9753728]

[15]. Hagen L, et al., Cell cycle-specific UNG2 phosphorylations regulate protein turnover, activity and association with RPA, EMBO J. 27 (2008) 51–61. [PubMed: 18079698]

[16]. Weiser BP, Stivers JT, Cole PA, Investigation of N-Terminal phospho-regulation of uracil DNA glycosylase using protein semisynthesis, Biophys. J 113 (2017) 393–401. [PubMed: 28746850]

[17]. Mer G, et al., Structural basis for the recognition of DNA repair proteins UNG2, XPA, and RAD52 by replication factor RPA, Cell 103 (2000) 449–456. [PubMed: 11081631]

[18]. Otterlei M, et al., Post-replicative base excision repair in replication foci, EMBO J. 18 (1999) 3834–3844. [PubMed: 10393198]

[19]. Fischer JA, Muller-Weeks S, Caradonna S, Proteolytic degradation of the nuclear isoform of uracil-DNA glycosylase occurs during the S phase of the cell cycle, DNA Repair (Amst.) 3 (2004) 505–513. [PubMed: 15084312]

[20]. Esadze A, Stivers JT, Facilitated diffusion mechanisms in DNA base excision repair and transcriptional activation, Chem. Rev 118 (2018) 11298–11323. [PubMed: 30379068]

[21]. Candotti M, Orozco M, The differential response of proteins to macromolecular crowding, PLoS Comput. Biol 12 (2016) el005040.

[22]. Kozer N, Kuttner YY, Haran G, Schreiber G, Protein-protein association in polymer solutions: from dilute to semidilute to concentrated, Biophys. J 92 (2007) 2139–2149. [PubMed: 17189316]

[23]. Cravens SL, et al., Molecular crowding enhances facilitated diffusion of two human DNA glycosylases, Nucleic Acids Res. 43 (2015) 4087–4097. [PubMed: 25845592]

[24]. Cravens SL, Stivers JT, Comparative effects of ions, molecular crowding, and bulk DNA on the damage search mechanisms of hOGGl and hUNG, Biochemistry 55 (2016) 5230–5242. [PubMed: 27571472]

[25]. Kang H, Pincus PA, Hyeon C, Thirumalai D, Effects of macromolecular crowding on the collapse of biopolymers, Phys. Rev. Lett 114 (2015) 068303. [PubMed: 25723249]

[26]. Soranno A, et al., Single-molecule spectroscopy reveals polymer effects of disordered proteins in crowded environments, Proc. Natl. Acad. Sci. U. S. A 111 (2014) 4874–4879. [PubMed: 24639500]

[27]. Szasz CS, et al., Protein disorder prevails under crowded conditions, Biochemistry 50 (2011) 5834–5844. [PubMed: 21634433]

[28]. Kuttner YY, Kozer N, Segal E, Schreiber G, Haran G, Separating the contribution of translational and rotational diffusion to protein association, J. Am. Chem. Soc 127 (2005) 15138–15144. [PubMed: 16248654]

[29]. Schreiber G, Haran G, Zhou H-X, Fundamental aspects of protein-protein association kinetics, Chem. Rev 109 (2009) 839–860. [PubMed: 19196002]

[30]. Alderson TR, Lee JH, Charlier C, Ying J, Bax A, Propensity for cis-proline formation in unfolded proteins, Chembiochem 19 (2018) 37–42. [PubMed: 29064600]

[31]. Shen Y, Bax A, Prediction of Xaa-Pro peptide bond conformation from sequence and chemical shifts, J. Biomol. NMR 46 (2010) 199–204. [PubMed: 20041279]

[32]. Shen Y, Bax A, SPARTA +: a modest improvement in empirical NMR chemical shift prediction by means of an artificial neural network, J. Biomol. NMR 48 (2010) 13–22. [PubMed: 20628786]

[33]. Berjanskii MV, Wishart DS, The RCI server: rapid and accurate calculation of protein flexibility using chemical shifts, Nucleic Acids Res. 35 (2007) W531–7. [PubMed: 17485469]

[34]. Berjanskii MV, Wishart DS, Application of the random coil index to studying protein flexibility, J. Biomol. NMR 40 (2008) 31–48. [PubMed: 17985196]

[35]. Walsh I, Martin AJM, Di Domenico T, Tosatto SCE, ESpritz: accurate and fast prediction of protein disorder, Bioinformatics 28 (2012) 503–509. [PubMed: 22190692]

[36]. Hanson J, Yang Y, Paliwal K, Zhou Y, Improving protein disorder prediction by deep bidirectional long short-term memory recurrent neural networks, Bioinformatics (2016), 10.1093/bioinformatics/btw678btw678.

[37]. Jones DT, Cozzetto D, DISOPRED3: precise disordered region predictions with annotated protein-binding activity, Bioinformatics 31 (2015) 857–863. [PubMed: 25391399]

[38]. Nielsen JT, Mulder FAA, Quality and bias of protein disorder predictors, Sci. Rep 9 (2019) 5137. [PubMed: 30914747]

[39]. Camilloni C, De Simone A, Vranken WF, Vendruscolo M, Determination of secondary structure populations in disordered states of proteins using nuclear magnetic resonance chemical shifts, Biochemistry 51 (2012) 2224–2231. [PubMed: 22360139]

[40]. Marsh JA, Singh VK, Jia Z, Forman-Kay JD, Sensitivity of secondary structure propensities to sequence differences between α- and γ-synuclein: implications for fibrillation, Protein Sci. 15 (2006) 2795–2804. [PubMed: 17088319]

[41]. Krepel D, Gomez D, Klumpp S, Levy Y, Mechanism of facilitated diffusion during a DNA search in crowded environments, J. Phys. Chem. B 120 (2016) 11113–11122. [PubMed: 27723976]

[42]. Levy Y, Onuchic JN, Wolynes PG, Fly-casting in protein- DNA binding: frustration between protein folding and electrostatics facilitates target recognition, J. Am. Chem. Soc 129 (2007) 738–739. [PubMed: 17243791]

[43]. Du M, Chen ZJ, DNA-induced liquid phase condensation of cGAS activates innate immune signaling, Science 361 (2018) 704–709. [PubMed: 29976794]

[44]. Friedman JI, Majumdar A, Stivers JT, Nontarget DNA binding shapes the dynamic landscape for enzymatic recognition of DNA damage, Nucleic Acids Res. 37 (2009) 3493–3500. [PubMed: 19339520]

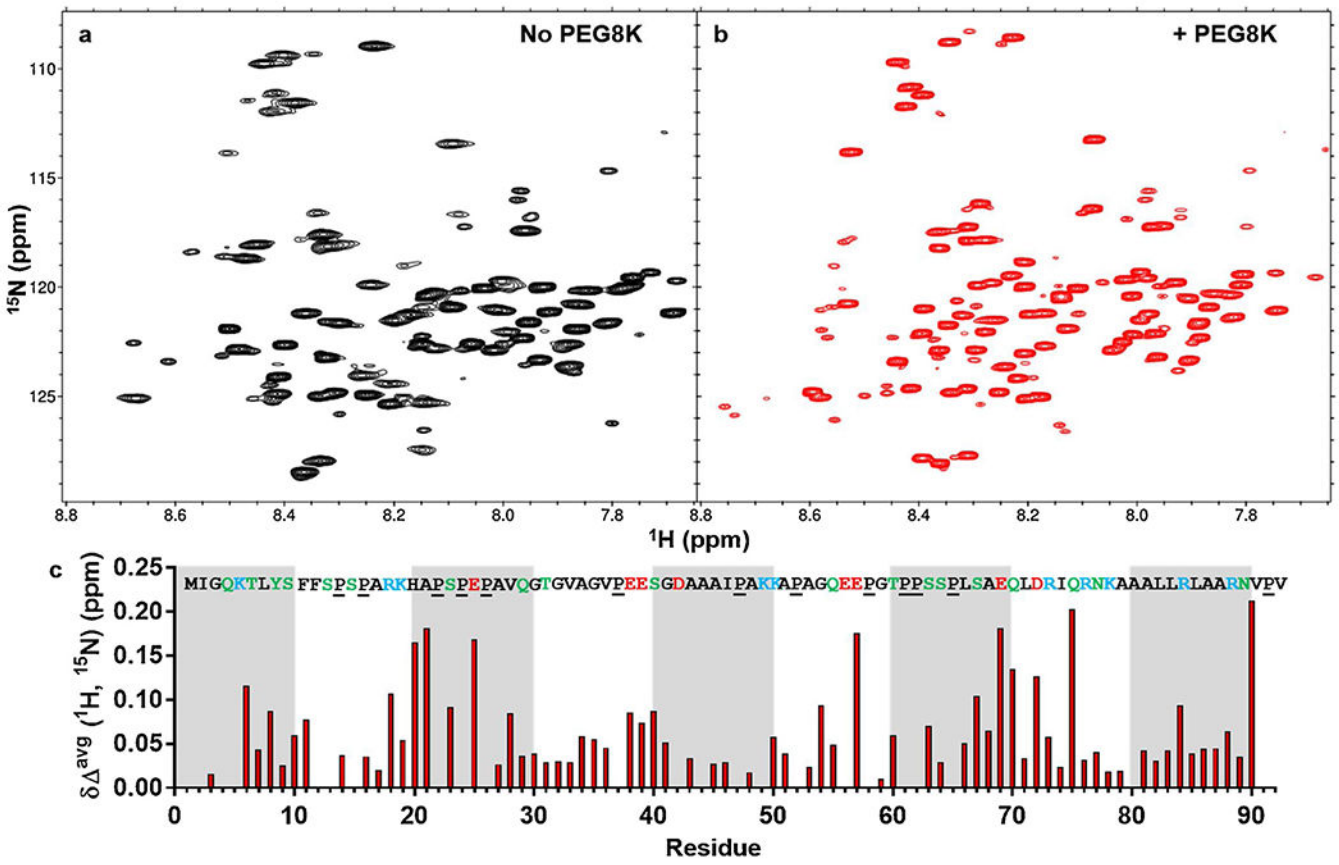

Fig. 1.

NTD backbone amide chemical shift dispersion in the absence and presence of 20 % PEG8K. (**a**) ^{1}H, ^{15}N-TROSY HSQC spectrum of 0.35 mM in NMR buffer A obtained at 800 MHz proton frequency, (**b**) Same spectrum as panel (a) except that 20 % PEG8K was added to the buffer, (**c**) The residue specific weighted average chemical shift changes are defined as $\delta\Delta^{avg}$ = NTD^{+PEG8K} - NTD^{-PEG8K}. The NTD amino acid sequence is shown with negatively charged (red), positively charged (blue), polar residues (green), and prolines indicated (underlined).

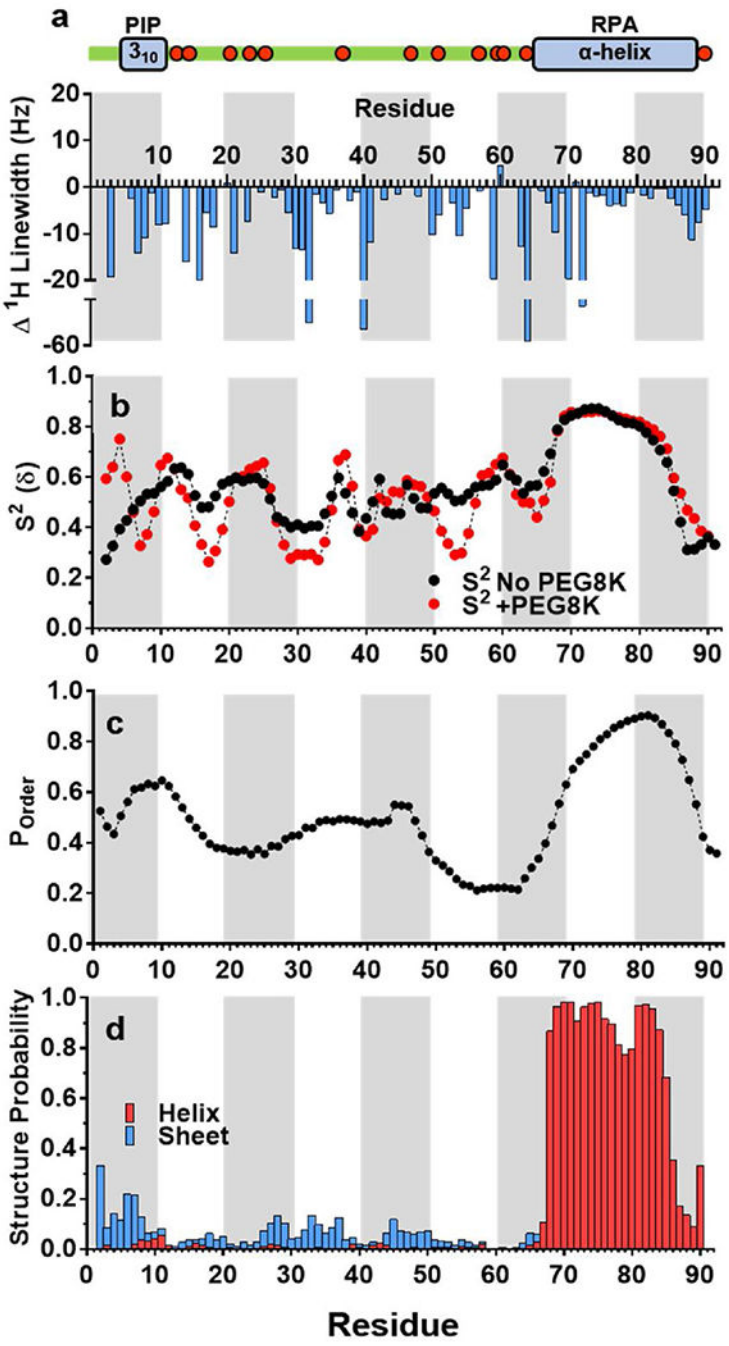

Fig. 2.

Chemical shift linewidths, shift-derived order parameters, sequence-based order probabilities, and calculated secondary structure propensity for the NTD. The structural elements of the NTD are shown graphically, including the location of the twelve structure-disrupting proline residues (red circles) and a putative short 3_{10} helix in the PCNA binding motif (PIP-motif)9. (**a**) The presence of 20 % PEG8K narrows the ^{1}H resonance linewidths as compared to the absence of PEG8K Lw (Hz) = LW^{+PEG8K} - LW^{-PEG8K}. (**b**) The TALOS + program was used to calculate the chemical shift-derived order parameters $S^{2}(\delta)$ using the

residue specific HN, N, Cα, Cβ, and C' chemical shifts in the absence (black) and presence (red) of PEG8K. (**c**) The sequence-based order probability was calculated using the Espritz program, employing the optional NMR train basis set that uses NMR structural models deposited in the PDB. Three regions in the NTD are predicted to have increased order (residues 4–12, 35–50 and 68–88) and these regions correlate well with ordered regions based on the $S^2(\delta)$ analysis (see text). (**d**) The residue specific chemical shifts were used to predict the NTD secondary structure in PEG8K using TALOS + (see text).

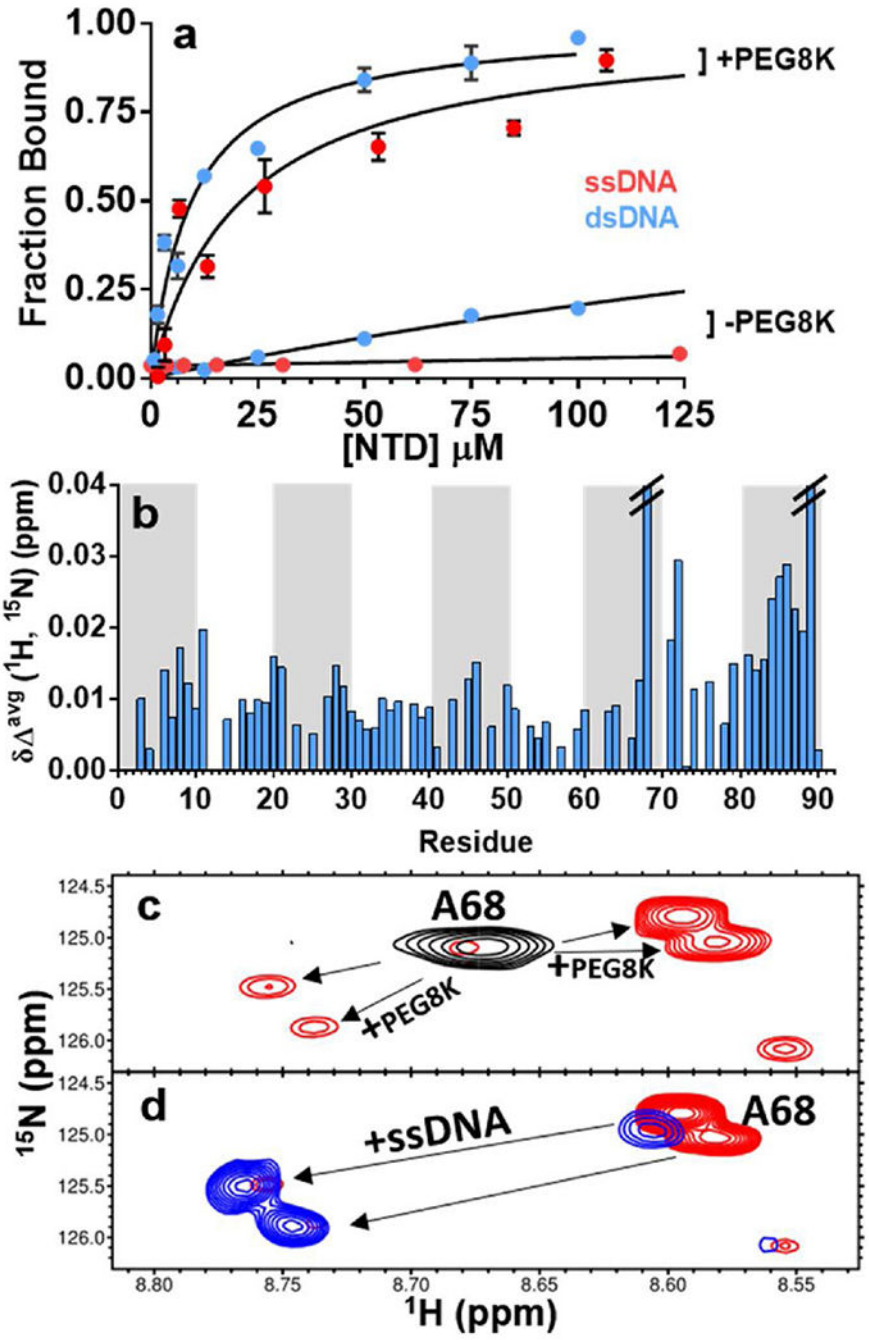

Fig. 3.

NTD binds 19mer duplex and single-stranded DNA with enhanced affinity in the presence of PEG8K and binding involves conformational selection. (**a**) DNA binding was followed by measuring the fluorescence anisotropy increase of fluorescein end-labeled duplex and ssDNA (25 nM) in the absence and presence of 20 % PEG8K using buffer A. The K_D values in the absence of PEG8K were greater than 100 M, and were reduced to 10 M and 23 M for duplex DNA and ssDNA, respectively. (**b**) Weighted-average chemical shift changes ($\delta\Delta^{avg}$) in the amide resonances of the NTD upon binding ssDNA (see full spectra in Supplemental

Figure S5). Spectra were obtained using 50 M NTD and 250 M 19mer ssDNA at 600 MHz. The large chemical shift changes for residues A68 and N89 extends far beyond the y axis limit. (**c**) Upon addition of PEG8K (red) the amide resonance of residue A68 enters a slow conformational exchange regime. (**d**) Upon addition of ssDNA to the NTD in the presence of PEG8K (blue), the conformational equilibrium is shifted to favor the minor state observed in the absence of ssDNA (compare panels c and d).

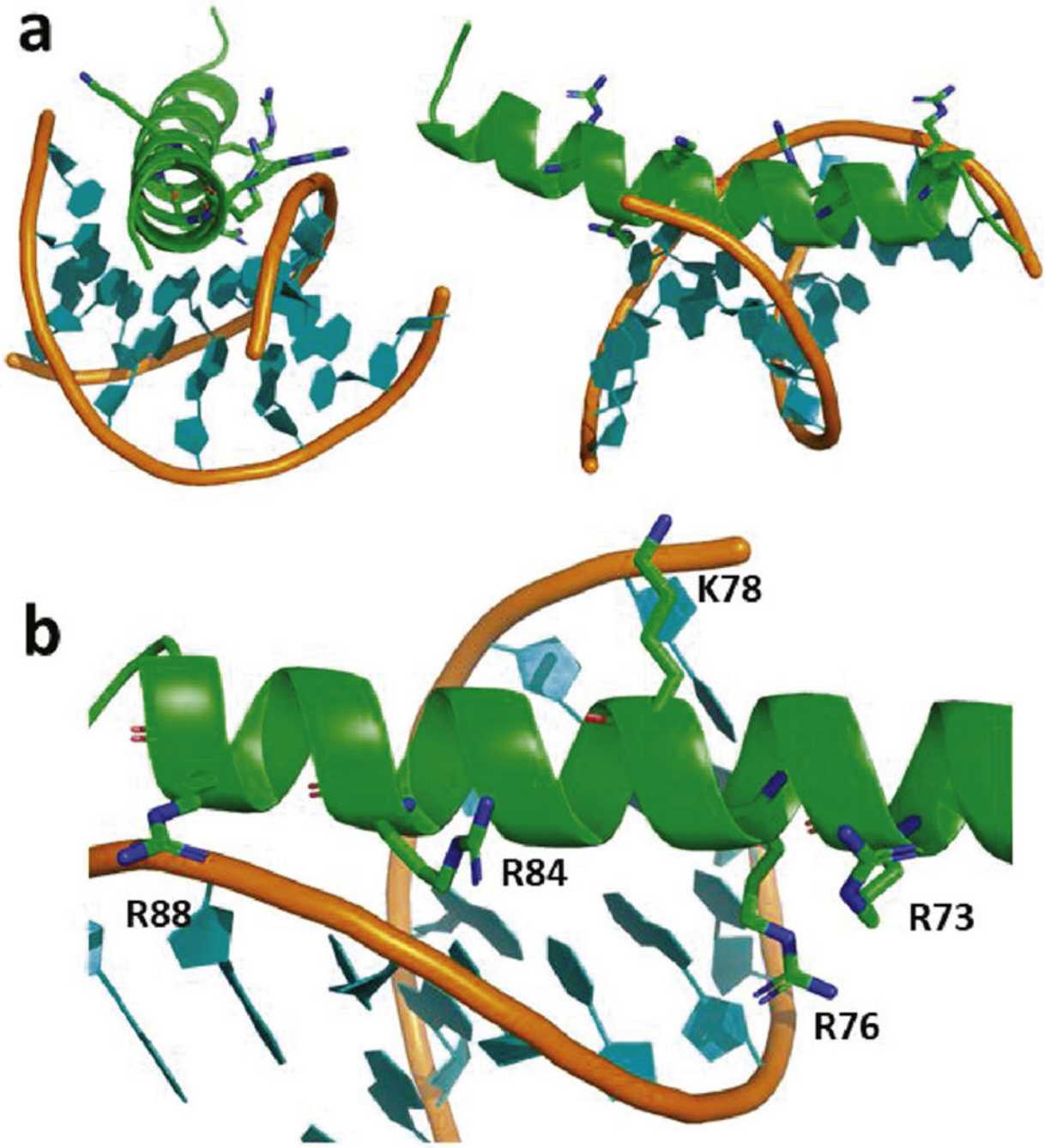

Fig. 4.

Computational model of the helix residues 66–95 bound to a dodecameric B-DNA ligand (PDBID: 1BNA) using HDOCK. (**a**) Orthogonal views of the RPA-helix interacting with the major groove of B-DNA. (**b**) The model indicates that dsDNA binding by the RPA binding helix is mediated by electrostatic interactions between four basic residues with the sugar-phosphate backbone (R76, K78, R84, R88).