Conformational Determinants for the Recruitment of ERCC1 by XPA in the Nucleotide Excision Repair (NER) Pathway: Structure and Dynamics of the XPA Binding Motif

Elisa Fadda*
School of Chemistry, National University of Ireland Galway, Galway, Ireland

ABSTRACT XPA is an essential protein in the nucleotide excision repair (NER) pathway, in charge of recruiting the ERCC1-XPF endonuclease complex to the DNA damage site. The only currently available structural insight into the binding of XPA to ERCC1 derives from the solution NMR structure of a complex between the ERCC1 central fragment and a 14-residue peptide, corresponding to the highly conserved binding motif of the XPA N-terminus, XPA_{67-80}. The extensive all-atom molecular-dynamics simulation study of the XPA_{67-80} peptide both bound to the ERCC1 central fragment and free in solution presented here completes the profile of the structural determinants responsible for the ERCC1/XPA_{67-80} complex stability. In addition to the wild-type, this study also looks at specific XPA_{67-80} mutants in complex with the ERCC1 central domain and thus contributes to defining the conformational determinants for binding, as well as all of the essential structural elements necessary for the rational design of an XPA-based, ERCC1-specific inhibitor.

INTRODUCTION

Platinum drugs are currently the most potent chemotherapeutic agents used to treat most types of cancer. The progenitor of all platinum drugs, *cis*-diamminedichloroplatinum(II) (*cis*-DDP or cisplatin), was discovered in 1965 (1) and today is still one of the most effective drugs in chemotherapy. Unfortunately, its efficacy is not the same for all types of cancer: although it is extremely successful in the treatment of testicular cancer, resulting in >90% of complete recoveries, for most other tumors the success rate decreases dramatically. In particular, in ovarian cancer the tumor response to cisplatin is very poor, with a 5-year survival rate of only 48% (GLOBOCAN 2008, http://globocan.iarc.fr). One of the main causes of the failure of cisplatin is intrinsic or acquired resistance to the drug, with the latter causing relapse. The efficacy of cisplatin and all platinum drugs hinges on their ability to bind DNA irreversibly by forming cross-links that bend and unwind the double helix (2). Intrastrand 1,2d[GpG] cross-links are the most abundant platinum-DNA covalent adducts and are probably the most effective for triggering cell death (3). The nucleotide excision repair (NER) pathway is principally responsible for recognizing and excising all bulky DNA lesions. It involves >30 proteins, three of which—the excision-repair cross-complementing group 1 (ERCC1), Xeroderma pigmentosum group F (XPF), and Xeroderma pigmentosum group A (XPA)—are essential for the removal of platinum-DNA adducts (4). Indeed, it is known that ERCC1, XPF, and XPA are overexpressed in most platinum-resistant cancer cells (2,4,5), which suggests an increased DNA repair activity. Enhanced levels of ERCC1 have been directly linked to clinical resistance to platinum chemotherapy (6,7). Conversely, these enzymes are expressed at very low levels in testicular cancer cells, which are very sensitive to cisplatin (2,5). Recent work indicates that downregulation (8) or suppression (9) of the XPF-ERCC1 endonuclease enhances the efficacy of cisplatin. Additionally, mutations in XPA that inhibit its interaction with ERCC1 directly hinder DNA repair (10,11), most likely because XPA is known to recruit the ERCC1-XPF complex to the DNA damage site so that the repair can proceed (12,13).

XPA has a structurally well-characterized zinc-finger central domain (sequence shown in *black* in Fig. 1) comprising residues 98–219 (14). The ERCC1-binding N-terminus includes residues 1–84 and is known to be poorly structured in solution (14–17). Experimental evidence shows that the 14-amino-acid sequence (shown in *red* in Fig. 1) between Lys-67 and Glu-80 comprises all the essential residues necessary for binding the ERCC1-XPF endonuclease (10,15). A 14-residue peptide with the same sequence, XPA_{67-80}, was shown not only to bind ERCC1 but also to inhibit its interaction with XPA (15).

Li et al. (10) investigated the role of two small fragments of the XPA-binding sequence by means of mutagenesis and in vitro binding assays. Their results show that deletion of the Gly-72 to Phe-75 region, also known as the G motif, abolishes binding to ERCC1 completely, whereas deletion of the Glu-78 to Glu-84 region, or E motif, reduces binding by ~70% (10). Additionally, XPA mutants lacking the G motif were found to be unable to confer UV resistance to the cells (10).

Because the ERCC1 binding domain of XPA is poorly structured, there is a lack of structural information that would allow us to identify the specific interactions that are

Submitted November 20, 2012, and accepted for publication April 10, 2013.

*Correspondence: elisa.fadda@nuigalway.ie

Editor: Michael Feig.

0006-3495/13/06/2503/9 $2.00

http://dx.doi.org/10.1016/j.bpj.2013.04.023

>gi|4507937|ref|NP_000371.1|
DNA repair protein complementing
XP-A cells [Homo sapiens]
MAAADGALPEAAALEQPAELPASVRASIERKRQRA
LMLRQARLAARPYSATAAAATGGMANVKAAPKIID
TGGGFILEEEEEEEQKIGKVVHQPGPVMEFDYVIC
EECGKEFMDSYLMNHFDLPTCDNCRDADDKHKLIT
KTEAKQEYLLKDCDLEKREPPLKFIVKKNPHHSQW
GDMKLYLKLQIVKRSLEVWGSQEALEEAKEVRQEN
REKMKQKKFDKKVKELRRAVRSSVWKRETIVHQHE
YGPEENLEDDMYRKTCTMCGHELTYEKM

FIGURE 1 XPA sequence (*Homo sapiens*). The zinc finger domain is highlighted in black. The N- and C-termini are shown in gray above and below the zinc finger domain sequence, respectively. The ERCC1-binding region is shown in red.

critical for binding. In a seminal work, Tsodikov et al. (15) used a combination of NMR-derived distance restrains and x-ray crystallography to determine the structure of the complex between the central domain of ERCC1 and a synthetic $XPA_{67\text{-}80}$ peptide. In this complex, shown in Fig. 2, the $XPA_{67\text{-}80}$ peptide occupies a narrow pocket in the central domain of ERCC1. The NMR data in this work highlight only a limited number of hydrogen bonds and hydrophobic contacts. Additionally, the resolution of the x-ray diffraction data is too low to allow unequivocal identification of the hydrogen-bonding network that holds the $XPA_{67\text{-}80}$ peptide in the ERCC1-binding site (15).

The inhibition of the interaction between XPA and ERCC1-XPF binding was found to prevent the repair of cisplatin-DNA adducts (11,15). Therefore, the development of high-affinity synthetic ERCC1-XPF substrates can be explored as an avenue to counteract resistance to platinum

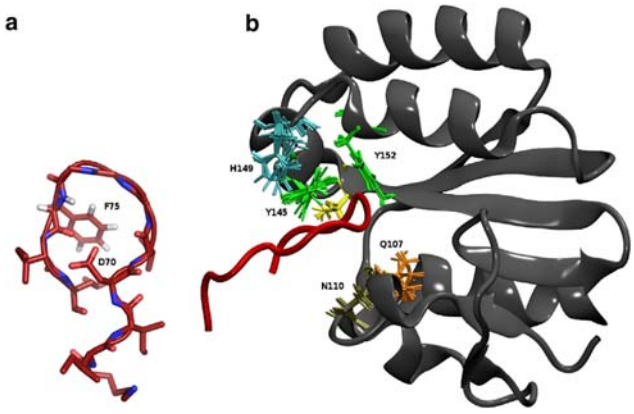

FIGURE 2 Solution NMR structure of the ERCC1-$XPA_{67\text{-}80}$ complex (PDBid 2JNW). The position of the residues represented with sticks is given for all 10 structures deposited in the PDB. SER142 is not labelled for clarity purposes and it is shown in yellow. The $XPA_{67\text{-}80}$ peptide is shown in red on panel a and in complex with the ERCC1 central domain (shown in grey) on panel *b*.

chemotherapy. The rational design of a competitive inhibitor of the XPA/ERCC1-XPF interaction hinges on the identification of the structural and dynamics determinants for binding. The aim of this study was twofold: 1), to highlight the structural and dynamic features that contribute to the stability of the complex between the $XPA_{67\text{-}80}$ section and the ERCC1 central domain; and 2), to determine the degree of disorder versus order in the $XPA_{67\text{-}80}$ peptide free in solution and how the conserved residues might influence the propensity of the peptide secondary structure. Each of these themes was addressed by extended MD simulations and analyzed in view of the available NMR, crystallographic, and mutagenesis data (10,11,15). The insights gained regarding the essential structural and conformational determinants for ERCC1/$XPA_{67\text{-}80}$ binding are discussed in the following sections.

MATERIALS AND METHODS

To characterize both the structure and dynamics of the ERCC1-$XPA_{67\text{-}80}$ complex, I conducted an extended molecular-dynamics (MD) simulation study of the $XPA_{67\text{-}80}$ peptide, both free and in complex with the ERCC1 central domain.

The conformational dynamics of the free $XPA_{67\text{-}80}$ peptide in water was monitored for 10 μs. The $XPA_{67\text{-}80}$ peptide was built with version 4.0.1 of ArgusLab (18) in a fully extended conformation, i.e., $\phi = \psi = 180°$, with sequence KIIDTGGGFILEEE, corresponding to the section K_{67}-E_{80} of the human XPA shown in Fig. 1 (NCBI Reference Sequence: NP_000371.1). The peptide terminal residues were capped with acetyl (ACE) and N-methylamide (NME) groups. The extended peptide was inserted into a cubic simulation box with 70 Å sides. To reach neutrality, three Na^+ counterions were added to the system with the *genion* tool available in version 4.0.7 of the GROMACS simulation package (19). The position of the water molecules and counterions was minimized with 50,000 steps of the steepest-descent algorithm, and then equilibrated for 500 ps in the NVT ensemble, followed by 500 ps in the NPT ensemble with T = 300 K and p = 1 bar as target values. The peptide was released from all position restraints and equilibrated in the NPT ensemble for 1 ns before a 50 ns production run. From this production run, 10 snapshots (one every 5 ns) were collected (see Fig. S1 in the Supporting Material). Each one of these snapshots was inserted into a new, smaller (48 Å sides) cubic simulation box. Each snapshot was first minimized and then equilibrated by following the same protocol used in the initial simulation. The production run for each snapshot was extended to 1 μs, for a total of 10 μs MD simulation.

The role of specific residues in the stability of the ERCC1/$XPA_{67\text{-}80}$ complex was analyzed through a series of 500 ns MD simulations involving wild-type (WT) and two mutant $XPA_{67\text{-}80}$ peptides (F75A and D70A) bound to the ERCC1 central domain. The first one of the 10 structures deposited in the Protein Data Bank (PDB) with PDB ID 2JNW (15) was used as starting structure (or template in the case of the $XPA_{67\text{-}80}$ mutants) for all MD simulations of the complex. The relative root mean-square deviation (RMSD) for all heavy atoms in the $XPA_{67\text{-}80}$ peptide calculated for the 10 structures deposited in the PDB is 0.4 Å, and the backbone RMSD is 0.2 Å. The PDB structure includes the Lys-67 to Leu-77 section of the XPA peptide and the ERCC1 central domain, i.e., from Asn-99 to Ala-214. The residues considered in all of the MD simulations match the PDB 2JNW structure (i.e., to avoid the introduction of conformational biases, residues Glu-78, Glu-79, and Glu-80 were not added). In the PDB structure (15), $His\text{-}149_{ERCC1}$ is protonated at Nδ1. To thoroughly investigate the role of the hydrogen bond between $Asp\text{-}70_{XPA}$ and $His\text{-}149_{ERCC1}$ in the stability of the complex, all MD simulations of the WT and F75A mutant $XPA_{67\text{-}80}$

were also run with His-149$_{ERCC1}$ protonated at Nε2. For the complex with the D70A XPA$_{67-80}$ mutant, only the Nδ1 protonation state was considered, because the mutated Asp-70 is His-149$_{ERCC1}$ hydrogen-bonding partner. The following protocol was used for all five MD simulations: All missing hydrogen atoms were added to the PDB structure. In the case of the mutants, the side-chain atoms were manually removed up to Cβ. The terminal residues of both ERCC1 and XPA were capped with ACE and NME groups. The complex was centered in a cubic box built so that the minimum distance between the protein and the side of the simulation box was 13 Å. Two Cl^- counterions were added to the complexes with WT and the F75A XPA$_{67-80}$ mutant to reach neutrality, and three Cl^- counterions were added to the complex with the D70A XPA$_{67-80}$ mutant. The positions of all water molecules, counterions, and hydrogen atoms were optimized through 100,000 steps of steepest descent and then equilibrated in the NVT ensemble for 500 ps at 300 K. The solvent, hydrogen atoms, and counterions were then equilibrated for another 500 ps in the NPT ensemble at 300 K and 1 bar. The side chains of the ERCC1 central domain were then released from the position restraints and the system was equilibrated for 1 ns. Another equilibration phase of 1 ns followed, during which the side chains of the XPA$_{67-80}$ peptide were also released from all position restraints. During the last equilibration phase, only the Cα atoms of the ERCC1 central domain were kept restrained and the system ran for 5 ns. The conformational dynamics of each one of the five complexes was extended to 500 ns. The temperature was held constant at 300 K by a Langevin thermostat (20) with coupling time constant of 0.1 ps. A Berendsen barostat (21) was used to hold the pressure constant at 1 bar, with a time constant of 0.5 ps. All MD simulations were performed with version 4.0.7 of the GROMACS software package (19). The equations of motion were integrated using a leapfrog stochastic dynamics integrator (22) with a 2 fs timestep. The linear constraint solver (LINCS) was used to constrain all bonds with hydrogen atoms (23). Long-range electrostatics were treated with the particle mesh Ewald (PME) method (24,25). The maximum spacing for the fast Fourier transform (FFT) grid was chosen as 1 Å. Cutoff values for Coulomb were set to 12 Å, and van der Waals interactions were switched off between 10 and 11 Å.

The AMBER99SB force field (26) was chosen for all of the MD simulations discussed in this work to represent the protein and ion atoms. MD simulations of strong electrolyte (1:1) solutions have shown that the adjusted Åqvist parameters for ions used in the Amber99 force field suffer from several shortcomings and can produce simulation artifacts, such as the formation of ion insoluble aggregates in alkali chlorides below their solubility limit (16,27). No unphysical behavior of the counterions was observed during the simulations discussed in this work. TIP3P (28) was selected as the water model. The superiority of four- and five-site water models relative to TIP3P in terms of the accuracy of water-binding free-energy calculations was discussed in an earlier work (29). TIP3P originally was chosen as a water model in this work because of its complementarity to the chosen protein force field (30–32). A large portion of the MD simulations discussed here were started before the work by Fadda and Woods (29) was published.

Structure clustering was performed with the program *g_cluster* included in version 4.0.7 of the GROMACS simulation package. Clusters were identified by means of the GROMOS algorithm (33). Based on the RMSD analysis of the trajectories of both free and bound peptide, an RMSD cutoff of 1.5 Å was chosen because it allowed the most significant conformational transitions to be captured. Clustering data obtained with cutoff radii ranging from 2.5 to 1.0 Å were collected for the bound peptide and are shown in Table S1. A least-squares fit was performed over all rotational and translational degrees of freedom.

RESULTS

WT XPA$_{67-80}$ bound to the ERCC1 central domain

The stability of specific interactions between the WT XPA$_{67-80}$ and the ERCC1 central domain was assessed during the course of two separate MD simulations: one with His-149$_{ERCC1}$ protonated at Nε2 and one with His-149$_{ERCC1}$ protonated at Nδ1. In the original PDB structure, His-149$_{ERCC1}$ is protonated at Nδ1 (see Fig. 2 *b*). However, the MD simulation results show that the hydrogen bond between Asp-70$_{XPA}$ and His-149$_{ERCC1}$ cannot form when His-149$_{ERCC1}$ is protonated at Nδ1, because the latter is always too far away from the carboxylic acceptor group. The inability to form this hydrogen bond may affect the complex's conformational stability, ultimately biasing the simulation results.

In the complex with His-149$_{ERCC1}$ protonated at Nε2, the XPA$_{67-80}$ peptide fills the binding pocket stably throughout the trajectory (see Fig. 3). The peptide backbone average RMSD relative to the NMR complex calculated throughout the trajectory after aligning just the ERCC1 central domains is 4.7 Å. The conformational dynamics of the peptide involves mostly the peptide section between residues Lys-67$_{XPA}$ and Gly-73$_{XPA}$, which is one of the sections of the peptide that is directly in contact with the solvent. The clustering analysis allows to identify 15 different conformations occurring during the MD simulation, with only clusters 1–3 having a relative population larger than 10% (see Table 1 and Fig. 3 *a*). The backbone RMSDs calculated for each cluster relative to the NMR bound XPA$_{67-80}$ peptide are indicated in parentheses in Table 1.

The MD simulation highlights a network of hydrogen bonds and hydrophobic contacts that contributes to stabilize the complex. The carboxylate group of Asp-70$_{XPA}$ forms hydrogen bonds with both the protonated Nε2 of

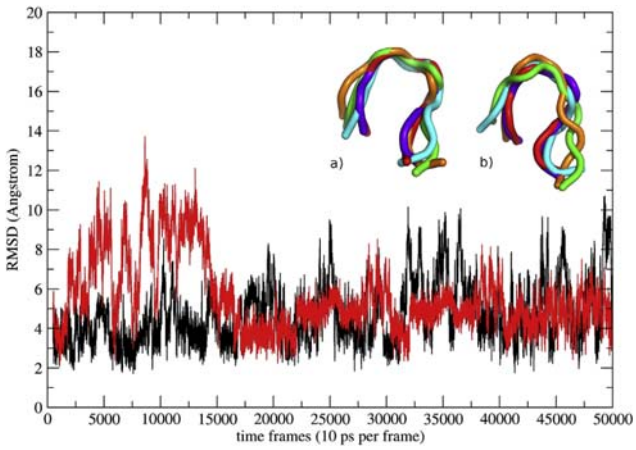

FIGURE 3 WT XPA$_{67-80}$ backbone RMSD relative to the starting structure calculated over the 500 ns MD simulation of the ERCC1-XPA$_{67-80}$ complex. The black line shows the RMSD of the complex where His-145$_{ERCC1}$ is protonated at Nε2 (HIE), and the red line shows the RMSD of the complex where His-145$_{ERCC1}$ is protonated at Nδ1 (HID). The first 5 ns of the trajectory have been discarded as part of the equilibration phase. (*a* and *b*) The first five conformations identified through clustering analysis of the bound XPA$_{67-80}$ peptide in the (*a*) HIE and (*b*) HID complexes. The peptide backbone is shown in red for cluster 1, purple for cluster 2, green for cluster 3, orange for cluster 4, and cyan for cluster 5.

TABLE 1 Clustering analysis results obtained with an RMSD cutoff value of 1.5 Å

Clusters	WT HIE (15)	WT HID (18)	D70A (18)	F75A HIE (14)	F75A HID (26)
Cluster 1	**43.6% (1.5)**	**48.7% (1.5)**	**67.1% (1.3)**	**39.8% (1.1)**	**61.6% (3.9)**
Cluster 2	**18.5% (1.5)**	9.8% (1.7)	6.0% (4.6)	**16.2% (1.6)**	6.0% (3.4)
Cluster 3	**11.1% (2.2)**	9.5% (3.1)	4.0% (1.8)	**10.4% (0.9)**	4.0% (5.0)
Cluster 4	5.1% (2.7)	6.0% (1.8)	3.1% (1.1)	8.0% (2.5)	3.6% (4.0)
Cluster 5	3.6% (1.8)	4.5% (2.7)	2.5% (2.1)	5.6% (1.6)	3.3% (1.5)

The total number of clusters obtained from the analysis of each complex is indicated in parentheses in the first row. HIE indicates that His-149$_{XPA}$ is protonated at Nε2, and HID indicates that His-149$_{XPA}$ is protonated at Nδ1. The backbone RMSD (in Å) calculated for the middle structures of each cluster relative to PDB 2JNW is shown in parentheses beside the cluster populations. Cluster populations above 10.0% are highlighted in bold.

His-149$_{ERCC1}$ and the hydroxyl group of Tyr-145$_{ERCC1}$. These two hydrogen-bond interactions occur at different times (Fig. S1) but are also seen to form concurrently. The hydrogen bond between the Gly-74$_{XPA}$ backbone carboxylic group and the Ser-142$_{ERCC1}$ backbone amide nitrogen, also identified by NMR (15), remains stable throughout the trajectory, together with the hydrogen bond between the Gly-73$_{XPA}$ backbone carbonyl group and the Gln-107$_{ERCC1}$ side-chain amide group, and the hydrogen bond between the Gly-72$_{XPA}$ backbone carbonyl and the Arg-156$_{ERCC1}$ side chain. The hydrophobic stacking between the Phe-75$_{XPA}$ and Asn-110$_{ERCC1}$ side chains is conserved throughout the trajectory. A summary of all stable direct interactions identified during the MD simulation is shown in Fig. 4. Also in agreement with the analysis of the NMR data (15), the MD simulation also shows that the Ile-68$_{XPA}$ side chain fills a hydrophobic pocket on the side of the ERCC1-binding site, with an average RMSD of 2.1 Å relative to the starting structure. Another interesting hydrophobic interaction is the one involving His-149$_{ERCC1}$, Tyr-145$_{ERCC1}$, and Tyr-142$_{ERCC1}$ (Fig. 2). These side chains are often found stacked, and such an interaction orients the side chains of both Tyr-145$_{ERCC1}$ and His-149$_{ERCC1}$ in an ideal position to form hydrogen contacts with the Asp-70$_{XPA}$ carboxylate.

The MD simulation of the WT complex with the His-149$_{ERCC1}$ protonated at Nδ1 shows that the hydrogen bond between His-149$_{ERCC1}$ and Asp-70$_{XPA}$ cannot form due to the longer distance between the donor and the acceptor. This results in the peptide being partially unbound and conformationally disordered for the first 150 ns (Fig. 3). As Asp-70$_{XPA}$ forms a stable hydrogen bond with the Tyr-145$_{ERCC1}$ hydroxyl group, the peptide folds back into the binding site, where it remains stable until the end of the simulation. The clustering algorithm identifies 18 clusters, with cluster 1 being the most populated conformation at 49% (see Table 1). Clusters 1 and 2 are populated only after the formation of the stable hydrogen bond between Tyr-145$_{ERCC1}$ and Asp-70$_{XPA}$.

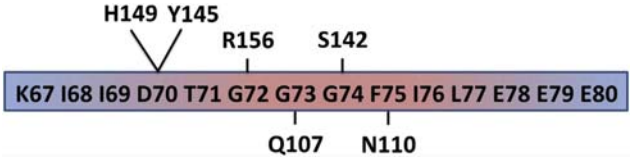

FIGURE 4 Contact map showing the most stable direct interactions identified throughout the 500 ns MD of the WT complex with H149$_{ERCC1}$ protonated at Nε2. The XPA$_{67-80}$ peptide residue IDs are shown in a box in which the background color indicates the part of the peptide that is in contact with the solvent (blue) and the part of the peptide that is buried in the binding pocket (red).

D70A XPA$_{67-80}$ bound to the ERCC1 central domain

The simulation of the D70A XPA$_{67-80}$ mutant is aimed at defining the role of the hydrogen bond between Asp-70$_{XPA}$ and His-149$_{ERCC1}$ and/or Tyr-145$_{ERCC1}$ in the stability of the complex. The results are qualitatively similar to those obtained for the WT complex with His-149$_{ERCC1}$ protonated at Nδ1. Indeed, as shown in Fig. 5, during the first 150 ns of the simulation, the XPA$_{67-80}$ D70A mutant is conformationally disordered and partially unbound. Again, the disorder is largely localized on the peptide section between Lys-67$_{XPA}$ and Gly-73$_{XPA}$ in contact with the solvent (see Fig. 4). After this time frame, the conformational stability is regained and then maintained

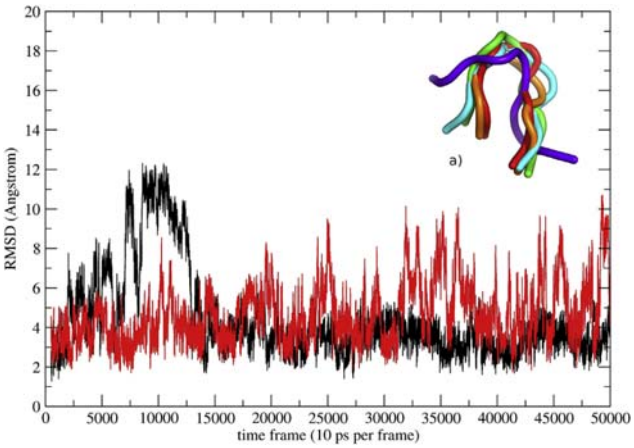

FIGURE 5 The backbone RMSD value of the D70A mutant of XPA$_{67-80}$ bound to the ERCC1 central domain is shown in black. The RMSD of the complex with the WT XPA$_{67-80}$ where His-149$_{ERCC1}$ is protonated at Nε2 is shown in red for comparison. All RMSD values are relative to the starting structure (PDB ID: 2JNW). The middle structures of the first five most-populated clusters are shown in the upper-right corner.

until the end of the trajectory. The clustering algorithm identifies 18 different conformations visited during the 500 ns MD (the first five are shown in Fig. 5 *a*). Cluster 1 is largely the most populated and represents the only conformational ensemble visited in the time frame between 150 and 500 ns. This stable conformation is characterized by internal hydrogen bonds between the peptide backbone atoms, i.e., Phe-75$_{XPA}$ and Thr-71$_{XPA}$, and Leu-77$_{XPA}$ and Ile-69$_{XPA}$. Other stable hydrogen bonds are formed between the His-149$_{ERCC1}$ Nε2 and the hydroxyl group of Thr-71$_{XPA}$, and between the Tyr-145$_{ERCC1}$ and the Thr-71$_{XPA}$ backbone carbonyl. All other stable interactions identified previously for the WT complex are also maintained here, i.e., the stacking between Phe-75$_{XPA}$ and Asn-110$_{ERCC1}$, the hydrogen bond between the Gly-74$_{XPA}$ backbone carbonyl and the Ser-142$_{ERCC1}$ backbone amide nitrogen, and the hydrogen bond between the Gly-73$_{XPA}$ backbone carbonyl and the Gln-107$_{ERCC1}$ side-chain amide group (see Fig. 4). Additionally, the nonspecific hydrophobic contact between Ile-76$_{XPA}$ and the hydrophobic pocket at the mouth of the ERCC1 binding site is also conserved.

F75A XPA$_{67\text{-}80}$ bound to the ERCC1 central domain

The role played by the stacking interaction between Phe-75$_{XPA}$ and Asn-110$_{ERCC1}$ in the conformational stability of the complex was assessed by means of two separate MD simulations: one with His-149$_{ERCC1}$ protonated at Nε2 and one with His-149$_{ERCC1}$ protonated at Nδ1. The average backbone RMSDs relative to the NMR structure calculated throughout both simulations are shown in Fig. 6. The results clearly indicate that the XPA$_{67\text{-}80}$ peptide

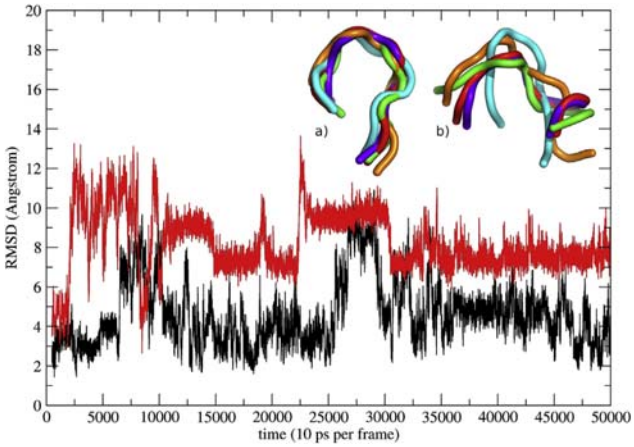

FIGURE 6 Backbone RMSD of the F75A mutant of XPA$_{67\text{-}80}$ bound to the ERCC1 central domain. The RMSD of the complex with His-149$_{ERCC1}$ protonated at Nε2 is shown in black, and the RMSD of the complex with His-149$_{ERCC1}$ protonated at Nδ1 is shown in red. All RMSDs are calculated relative to the corresponding starting structure. (*a* and *b*) Snapshots were collected every 2.5 ns of the F75A XPA$_{67\text{-}80}$ mutant in complex with ERCC1, with His-149$_{ERCC1}$ protonated at (*a*) Nε2 or (*b*) Nδ1.

is less conformationally disordered when bound to ERCC1 with His-149$_{ERCC1}$ protonated at Nε2. The clustering analysis for this complex (see Table 1) identifies only three clusters populated more than 10%. During the first 100 ns of the simulation, the most populated conformational ensemble is cluster 3. Such conformation is stabilized by internal hydrogen bonds between the peptide backbone atom, i.e., between Gly-72$_{XPA}$ and Ala-75$_{XPA}$, and Ile-69$_{XPA}$ and Leu-77$_{XPA}$, but among the interactions highlighted in Fig. 4, it lacks other than the stacking between Phe-75$_{XPA}$ and Asn-110$_{ERCC1}$ due to the F75A mutation, the interaction between Arg-156$_{ERCC1}$ and Gly-72$_{XPA}$, and any interaction involving Asp-70$_{XPA}$. Indeed, Tyr-145$_{XPA}$ is involved in an interaction with the amide backbone atom of Gly-74$_{XPA}$, and His-149$_{ERCC1}$ is not involved in any hydrogen bonding throughout the entire trajectory but remains stacked to the Tyr-145$_{XPA}$ side chain. Clusters 1 and 2 become populated after the initial 70 ns and are characterized by the formation of stable hydrogen bonds between Tyr-145$_{XPA}$ and Asp-70$_{XPA}$, and Arg-156$_{ERCC1}$ and Gly-72$_{XPA}$, which last throughout the remainder of the trajectory.

As shown in Fig. 6, the peptide in complex with ERCC1 with His-149$_{ERCC1}$ protonated at Nδ1 shows the highest degree of disorder among all the complexes studied in this work. The clustering analysis summarized in Table 1 identifies 26 different conformations visited during the simulation, with only cluster 1 significantly populated at 62%. The middle structures for the first five clusters are shown in Fig. 6 *b*. All of these conformations are characterized by a large distance between the two sides of the hairpin, in contrast to the peptide in the other complexes. Indeed, the XPA$_{67\text{-}80}$ peptide is largely unbound throughout the simulations. Among the stable interactions identified and shown in Fig. 4, only two hold throughout the trajectory: the one between Ser-142$_{ERCC1}$ and Gly-74$_{XPA}$, and the one between Gly-73$_{XPA}$ and Gln-107$_{ERCC1}$. In contrast to all of the previous cases, neither Tyr-145$_{ERCC1}$ nor His-149$_{ERCC1}$ is involved in any stable interaction with the peptide residues.

XPA$_{67\text{-}80}$ free peptide in solution

The conformational dynamics of the XPA$_{67\text{-}80}$ peptide unbound in solution was also studied. From an initial trajectory of 50 ns started from a fully extended peptide, 10 uncorrelated snapshots (one each every 5 ns) were selected. These 10 structures are shown in Fig. S2. The conformational dynamics of each snapshot was monitored for 1 μs, for a total of 10 μs MD simulation. The different snapshots are seen to explore specific regions of the XPA$_{67\text{-}80}$ conformational space. Based on the results of the clustering (see Table 2) and conformational analyses (the latter in terms of radius of gyration and RMSD), the XPA$_{67\text{-}80}$ peptide in solution does not show any propensity to form a stable secondary structure motif for ~52% of the total simulation time. The radius of gyration calculated for each snapshot (shown in

TABLE 2 Clustering analysis results obtained with an RMSD cutoff value of 1.5 Å

Cluster sizes per snapshot	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
s1 (172)	**15.7% (3.4)**	**13.1% (2.3)**	7.4% (3.2)	4.1% (2.8)	3.4% (3.4)
s2 (238)	**24.3% (2.4)**	9.6% (3.7)	5.3% (2.8)	2.7% (3.6)	2.5% (3.3)
s3 (36)	**47.7% (3.7)**	**24.3% (2.6)**	8.2% (1.9)	5.6% (3.9)	3.9% (2.6)
s4 (113)	**45.3% (2.8)**	4.1% (4.4)	3.2% (4.1)	3.2% (5.1)	3.0% (4.4)
s5 (186)	**17.8% (2.4)**	8.6% (4.1)	7.4% (3.2)	3.4% (2.4)	3.0% (2.3)
s6 (3)	**93.6% (2.4)**	5.9% (2.4)	0.6% (3.1)		
s7 (8)	**90.0% (2.4)**	4.0% (2.8)	3.7% (2.4)	1.6% (2.1)	0.3% (3.4)
s8 (13)	**92.7% (2.5)**	3.0% (2.7)	0.9% (3.9)	0.8% (3.1)	0.7% (2.2)
s9 (230)	**18.3% (3.9)**	**15.7% (3.4)**	7.9% (2.4)	5.2% (2.4)	3.7% (3.3)
s10 (146)	**23.1% (2.6)**	**10.8% (4.1)**	5.8% (3.1)	5.5% (4.2)	2.7% (5.5)

The total number of clusters obtained from the analysis of each snapshot is indicated in parentheses in the first column. The relative population and the backbone RMSD (in Å) of the middle structure relative to the NMR 2JNW structure are shown for the five highest-populated clusters. Clusters with populations larger than 10.0% are highlighted in bold.

Figs. S3 and S4) indicates that during this time frame the peptide extends and recoils, forming very short-lived structures. These unstable conformations are usually held together by one or two hydrogen bonds, typically involving terminal residues. For the remaining 48% of the simulation time, the peptide forms different stable hairpin-like structures (see Fig. 7). The relative populations of these structures and their backbone RMSDs relative to the bound peptide from NMR are shown in Table 2. The highest-populated clusters obtained from analysis of the simulations of snapshots 2, 5, and 6–8 correspond to the same hairpin conformation. The alignment of the clusters' middle structures, with an average backbone RMSD of 0.4 Å, is shown in Fig. S5. As shown in Table 2, this hairpin has an average backbone RMSD of 2.4 Å relative to the bound NMR conformation, and based on the clustering analysis is stable for >3.1 μs.

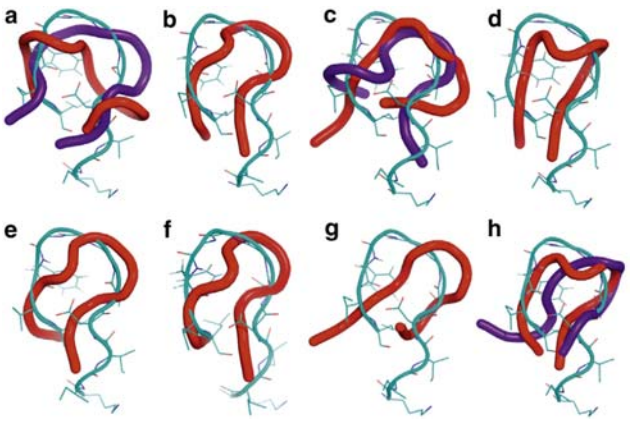

FIGURE 7 Stable hairpin motifs identified through clustering analysis during the 10 μs simulation of the $XPA_{67\text{-}80}$ peptide free in solution. The structure of the bound peptide from the NMR structure 2JNW is shown as a reference in cyan. Each panel shows the middle structure from the highest-populated clusters identified during the simulations (see Table 2). The highest-populated clusters are shown in red and the second highest-populated clusters are shown in purple. Only clusters with a relative population larger than 10.0% are shown. (*a–h*) Clusters from (*a*) snapshot 1, (*b*) snapshot 2, (*c*) snapshot 3, (*d*) snapshot 4, (*e*) snapshot 5, (*f*) snapshots 6–8, (*g*) snapshot 9, and (*h*) snapshot 10.

This stability depends on the formation of three hydrogen bonds between the backbone atoms of Ile-68 and Leu-77, Asp-70 and Gly-74, and Asp-70 and Phe-75. Snapshot 3 shows two stable hairpin conformations with relative populations of 48% and 24%, respectively (see Fig. 7 *c*). As shown in the RMSD plots in Fig. S6 *c*, cluster 1 forms during the first half of the simulation and cluster 2 forms during the second half. Snapshot 4 also shows a stable hairpin (Fig. 7 *d*). This structure remains stable for 48% of the simulation time (Table 2) and is held together by three hydrogen bonds: one between the backbone atoms of Asp-70 and Phe-75, and two between Ile-68 and Leu-7. The same motif is also the highest-populated one during the simulation of snapshot 10 (Fig. 7 *h*). The relative backbone RMSD between the highest-populated clusters from snapshots 4 and 10 is 0.8 Å. This hairpin is held by three hydrogen bonds between Ile-68 and Leu-77, Asp-70 and Phe-75, and Asp-70 and Gly-73 (Fig. 7 *h*). Based on the clustering analysis, this structure is stable for 684 ns. The backbone RMSD obtained from alignment of the middle structures representative of the most stable and second-most stable hairpin formed by the $XPA_{67\text{-}80}$ peptide free in solution is 1.8 Å.

DISCUSSION

Sequence alignment shows that the XPA sequence between residues 67 and 80 ($XPA_{67\text{-}80}$) is highly conserved, with very few exceptions (10,11,34). Highly conserved residues can form essential interactions in the binding site and/or control the conformational propensity of the peptide free in solution. The latter determines the presentation of the peptide to its receptor upon binding.

The simulations of the ERCC1 in complex with WT and mutant $XPA_{67\text{-}80}$ peptides presented in this work allow for the identification of all nonbonded interactions that are critical for the stability of the complex. These interactions are summarized in Fig. 4 and include both hydrogen bonds and hydrophobic contacts. The stability of all of the hydrogen-bonding interactions highlighted from the analysis of the NMR data (15), i.e., the hydrogen bond between the

backbone atoms of Gly-74$_{XPA}$ and Ser-142$_{ERCC1}$, and between the side chains of Asp-70$_{XPA}$ and His-149$_{ERCC1}$, is confirmed by the MD simulations. A critical but previously unidentified hydrogen bond was ubiquitous in all of the simulations. This bond forms between the Gly-73$_{XPA}$ backbone carbonyl and the Gln-107$_{ERCC1}$ amide side chain, and stabilizes the position of the XPA$_{67-80}$ hairpin central loop.

In addition to these interactions, the simulations suggest that Asp-70$_{XPA}$ is often involved in a stable hydrogen bond with the side chain of Tyr-145$_{ERCC1}$. The key role of Tyr-145$_{ERCC1}$ in the stability of the complex is shown by the analysis of the MD simulations of all complexes. Indeed, both simulations of the complex with the WT XPA$_{67-80}$, independently of the protonation state of His-149$_{ERCC1}$, indicate that the formation of a hydrogen bond between Asp-70$_{XPA}$ and Tyr-145$_{ERCC1}$ restores the conformational stability of the bound peptide (Fig. 3). Additionally, in the simulations of the complexes with the D70A and F75A XPA$_{67-80}$ mutants, the hydrogen-bonding interactions involving Tyr-145$_{ERCC1}$ are crucial for the return of the peptide to a bound conformation (Figs. 5 and 6). As shown in the simulation of the complex with the F75A XPA$_{67-80}$ mutants and with the His-149$_{ERCC1}$ protonated at Nδ1, the inability of His-149$_{ERCC1}$ and Tyr-145$_{ERCC1}$ to interact with any hydrogen-bond acceptors on the solvent-exposed side of XPA$_{67-80}$ causes the peptide to remain partially unbound throughout the simulation. This crucial structural role played by Tyr-145$_{ERCC1}$ in the stability of the complex is in agreement with experimental evidence (11) that shows a significantly diminished NER activity of the Y145A ERCC1 mutant, most likely due to its inability to bind XPA.

The MD results also show that the His-149$_{ERCC1}$, Tyr-145$_{ERCC1}$, and Tyr-152$_{ERCC1}$ side chains are often found stacked. This stacking interaction orients the residue side chains so that both His-149$_{ERCC1}$ and Tyr-145$_{ERCC1}$ can interact with the section of the peptide in contact with the solvent. The possibility of a relative reorientation of the His-149$_{ERCC1}$, Tyr-145$_{ERCC1}$, and Tyr-152$_{ERCC1}$ side chains is also supported by the structural data (15) and, as shown in Fig. 2, such an orientation can be observed directly from the alignment of the 10 structures deposited in the PDB.

The main hydrophobic interactions that XPA$_{67-80}$ forms in the ERCC1 binding site are the stacking between Phe-75$_{XPA}$ and Asn-110$_{ERCC1}$ and the nonspecific contact between the Ile-67$_{XPA}$ side chain and the side of the binding pocket (Fig. 2). The mutation of Phe-75$_{XPA}$ to Ala was chosen for comparison with functional assays showing that the F75A mutant completely abolished binding and thus NER activity (15). The simulations show that the F75A mutation does not significantly affect the orientation of the Asn-110$_{ERCC1}$ side chain or the position of the section of peptide that is in contact within the interior of the binding pocket. This may be due to the tight packing of the peptide in the binding pocket, as well as the hydrogen bond and hydrophobic interactions being sufficient to hold that section in place during the time frame of the MD analyzed. Moreover, the structural complementarity is most likely only one of the contributing factors behind the critical role of Phe-75$_{XPA}$ in binding. Nevertheless, as shown by the simulation of the complex with the F75A XPA$_{67-80}$ mutant with His-149$_{ERCC1}$ protonated at Nδ1, the lack of a stacking interaction between Phe-75$_{XPA}$ and Asn-110$_{ERCC1}$, together with the absence of the hydrogen bonds involving either His-149$_{ERCC1}$ or Tyr-145$_{ERCC1}$, results in the most conformationally unstable complex. This result is in agreement with previous experimental data (11) showing the inability of both the single N110A and double A110A/Y145A ERCC1 mutants to bind XPA.

In agreement with heteronuclear single quantum correlation (HSQC) NMR data (15), the conformational dynamics of the XPA$_{67-80}$ free in solution shows a high degree of disorder. This work demonstrates the ability of the peptide to form stable secondary structure motifs. As shown in Fig. 7, these motifs are prevalently β-strand turns and hairpins stabilized by a minimum of three hydrogen bonds involving backbone atoms. The most populated hairpin conformation identified is shown in Fig. 7, *b*, *e*, and *f*, and is stable for slightly more than 3.1 μs, or ~1/3 of the total simulation time. The intrinsic propensity to form narrow hairpins, which are quite similar to the bound structure (see Fig. 7), could play a role in the presentation and recognition of the peptide by ERCC1. Although the relative stability of secondary structure motifs can be affected by the choice of force field, the diversity and reoccurrence of the same structural motifs seen in these extensive MD simulations suggest that the highly conserved XPA$_{67-80}$ sequence may determine the preference for a particular fold. More specifically, in the complex with ERCC1, three consecutive glutamates (Glu-78$_{XPA}$, Glu-79$_{XPA}$, and Glu-80$_{XPA}$) and on the opposite end of the peptide a lysine (Lys-67$_{XPA}$) are located outside of the binding pocket. The available structural data (15) indicate that the sections of XPA outside of the binding pocket are quite disordered. Nevertheless, all three glutamates are highly conserved and their deletion has been shown to reduce binding significantly (10). The analysis of the MD simulations of the free peptide in solution suggests that the favorable electrostatic interaction between these highly conserved residues with opposite charge located at both ends of the peptide can contribute to the propensity to form hairpins. Moreover, small flexible residues, such as the three glycines located in the center of the XPA$_{67-80}$ peptide, would also favor a hairpin conformation. Indeed, deletion of the G motif, which includes two of these glycines (Gly-73$_{XPA}$ and Gly-74$_{XPA}$) and Phe-75$_{XPA}$, has been shown to eliminate binding completely (10).

CONCLUSIONS

NMR and crystallographic data (15) highlight a number of specific interactions that the XPA$_{67-80}$ peptide forms

with the ERCC1 central domain. These involve almost exclusively the section of the XPA_{67-80} peptide in contact with the interior of the ERCC1-binding pocket. Based on this evidence, the peptide section from Lys-67_{XPA} to Thr-71_{XPA} that is exposed to the solvent does not appear to form specific interactions with the ERCC1 central domain. This study builds on these structural data (15) and highlights all of the structural determinants that are essential for the conformational stability of the complex between the XPA_{67-80} peptide and the ERCC1 central domain (Fig. 4). The key structural role of specific residues, such as Tyr-145_{ERCC1} and His-149_{ERCC1}, identified in this study is also supported by the results of functional assays. Indeed, the highly diminished NER activity of the single Y145A and double N110A/Y145A mutants reported by Orelli et al. (11) is supported by the critical role played by Tyr-145_{ERCC1} in the conformational stability of the complex. Based on the data presented and discussed in this work within the context of experimental evidence, the design of a pharmacophore profile aimed at identifying an ERCC1/XPA inhibitor is currently in progress.

Furthermore, the clustering analysis of the extensive conformational dynamics of the XPA_{67-80} peptide free in solution points toward a definite propensity to form hairpin motifs similar to the bound conformation (Fig. 7). This intrinsic conformational preference may affect recognition and ultimately binding by being a determinant factor in the presentation of the peptide to the receptor. A conformational study focused on determining the propensity of specific XPA_{67-80} peptide mutants to form hairpin motifs is also currently under way.

SUPPORTING MATERIAL

One table and seven figures are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(13)00456-6.

The Computational Glycoscience Laboratory and Prof. R.J. Woods are gratefully acknowledged for providing access to computational resources.

E.F. received financial support from the School of Chemistry, National University of Ireland Galway. Funding from the Science Foundation of Ireland (08/IN. 1/B2070) and the European Research Development Fund.

REFERENCES

1. Rosenberg, B., L. Vancamp, and T. Krigas. 1965. Inhibition of cell division in Escherichia coli by electrolysis products from a platinum electrode. *Nature.* 205:698–699.
2. Kelland, L. 2007. The resurgence of platinum-based cancer chemotherapy. *Nat. Rev. Cancer.* 7:573–584.
3. Zamble, D. B., Y. Mikata, . , S. J. Lippard. 2002. Testis-specific HMG-domain protein alters the responses of cells to cisplatin. *J. Inorg. Biochem.* 91:451–462.
4. Reed, E. 1998. Platinum-DNA adduct, nucleotide excision repair and platinum based anti-cancer chemotherapy. *Cancer Treat. Rev.* 24:331–344.
5. Siddik, Z. H. 2003. Cisplatin: mode of cytotoxic action and molecular basis of resistance. *Oncogene.* 22:7265–7279.
6. Olaussen, K. A., A. Dunant, . , J. C. Soria; IALT Bio Investigators. 2006. DNA repair by ERCC1 in non-small-cell lung cancer and cisplatin-based adjuvant chemotherapy. *N. Engl. J. Med.* 355:983–991.
7. Yan, Q. W., E. Reed, . , J. J. Yu. 2006. MZF1 possesses a repressively regulatory function in ERCC1 expression. *Biochem. Pharmacol.* 71:761–771.
8. Arora, S., A. Kothandapani, . , S. M. Patrick. 2010. Downregulation of XPF-ERCC1 enhances cisplatin efficacy in cancer cells. *DNA Repair (Amst.).* 9:745–753.
9. Chang, I. Y., M. H. Kim, . , H. J. You. 2005. Small interfering RNA-induced suppression of ERCC1 enhances sensitivity of human cancer cells to cisplatin. *Biochem. Biophys. Res. Commun.* 327:225–233.
10. Li, L., C. A. Peterson, . , R. J. Legerski. 1995. Mutations in XPA that prevent association with ERCC1 are defective in nucleotide excision repair. *Mol. Cell. Biol.* 15:1993–1998.
11. Orelli, B., T. B. McClendon, . , O. D. Schärer. 2010. The XPA-binding domain of ERCC1 is required for nucleotide excision repair but not other DNA repair pathways. *J. Biol. Chem.* 285:3705–3712.
12. Riedl, T., F. Hanaoka, and J. M. Egly. 2003. The comings and goings of nucleotide excision repair factors on damaged DNA. *EMBO J.* 22:5293–5303.
13. Barakat, K., M. Gajewski, and J. A. Tuszynski. 2012. DNA repair inhibitors: the next major step to improve cancer therapy. *Curr. Top. Med. Chem.* 12:1376–1390.
14. McNeil, E. M., and D. W. Melton. 2012. DNA repair endonuclease ERCC1-XPF as a novel therapeutic target to overcome chemoresistance in cancer therapy. *Nucleic Acids Res.* 40:9990–10004.
15. Tsodikov, O. V., D. Ivanov, . , T. Ellenberger. 2007. Structural basis for the recruitment of ERCC1-XPF to nucleotide excision repair complexes by XPA. *EMBO J.* 26:4768–4776.
16. Chen, A. A., and R. V. Pappu. 2007. Parameters of monovalent ions in the AMBER-99 forcefield: assessment of inaccuracies and proposed improvements. *J. Phys. Chem. B.* 111:11884–11887.
17. Wu, Y., N. J. Zacal, . , X. D. Zhu. 2007. XPF with mutations in its conserved nuclease domain is defective in DNA repair but functions in TRF2-mediated telomere shortening. *DNA Repair (Amst.).* 6:157–166.
18. Thompson, M. A. 2004. Molecular docking using ArgusLab, an efficient shape-based search algorithm and the AScore scoring function. ACS Meeting, Philadelphia, PA, 172, CINF 42.
19. Hess, B., C. Kutzner, . , E. Lindahl. 2008. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* 4:435–447.
20. Grest, G. S., and K. Kremer. 1986. Molecular dynamics simulation for polymers in the presence of a heat bath. *Phys. Rev. A.* 33:3628–3631.
21. Allen, M. P., and D. J. Tildesley. 1987. Computer Simulation of Liquids. Oxford University Press, Oxford/New York.
22. Hockney, R. W. 1970. The potential calculation and some applications. *Methods Comput. Phys.* 9:136–211.
23. Blomberg, M. R. A., and P. E. M. Siegbahn. 2012. The mechanism for proton pumping in cytochrome c oxidase from an electrostatic and quantum chemical perspective. *Biochim. Biophys. Acta.* 1817:495–505.
24. York, D. M., T. A. Darden, and L. G. Pedersen. 1993. The effect of long-range electrostatic interactions in simulations of macromolecular crystals. A comparision of the Ewald and truncated list methods. *J. Chem. Phys.* 99:8345–8348.
25. Essmann, U., L. Perera, . , L. G. Pedersen. 1995. A smooth particle mesh Ewald method. *J. Chem. Phys.* 103:8577–8593.
26. Simmerling, C., B. Strockbine, and A. E. Roitberg. 2002. All-atom structure prediction and folding simulations of a stable protein. *J. Am. Chem. Soc.* 124:11258–11259.

27. Joung, I. S., and T. E. Cheatham, 3rd. 2008. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J. Phys. Chem. B.* 112:9020–9041.

28. Jorgensen, W. L., J. Chandrasekhar, . , M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.

29. Fadda, E., and R. J. Woods. 2011. On the role of water models in quantifying the binding free energy of highly conserved water molecules in proteins: the case of concanavalin A. *J. Chem. Theory Comput.* 7:3391–3398.

30. Hornak, V., R. Abel, . , C. Simmerling. 2006. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins.* 65:712–725.

31. Lange, O. F., D. van der Spoel, and B. L. de Groot. 2010. Scrutinizing molecular mechanics force fields on the submicrosecond timescale with NMR data. *Biophys. J.* 99:647–655.

32. Wickstrom, L., A. Okur, and C. Simmerling. 2009. Evaluating the performance of the ff99SB force field based on NMR scalar coupling data. *Biophys. J.* 97:853–856.

33. Daura, X., I. Antes, . , A. E. Mark. 1999. The effect of motional averaging on the calculation of NMR-derived structural properties. *Proteins.* 36:542–555.

34. Basu, A., and S. Krishnamurthy. 2010. Cellular responses to cisplatin-induced DNA damage. *J. Nucleic Acids.* pii:201367.