The Carboxy-Terminal Domain of Xeroderma Pigmentosum Complementation Group C Protein, Involved in TFIIH and Centrin Binding, Is Highly Disordered†

Simona Miron,*,‡ Patricia Duchambon,‡ Yves Blouquit,‡ Dominique Durand,§ and Constantin T. Craescu*,‡

Integrative Imaging Unit, INSERM U759/Institut Curie Centre de Recherche, Centre Universitaire Paris-Sud, Bâtiment 112, 91405 Orsay, France, and Institut de Biochimie et Biophysique Moléculaire et Cellulaire, CNRS UMR 8619, Université Paris-Sud, 91405 Orsay Cedex, France

Received September 12, 2007; Revised Manuscript Received November 21, 2007

ABSTRACT: Xeroderma pigmentousum group C protein (XPC) is involved in the first step of nucleotide excision repair, with multiple functional roles including DNA damage recognition and recruitment of the repair machinery. This human protein of 940 residues forms a strong heterotrimeric complex with Rad23B and centrin 2. The structure of XPC is actually not known, and lack of significant sequence homology with proteins from structural data bases precludes any relevant prediction. Here, we present the molecular and structural characterization of a C-terminal fragment of XPC (C-XPC: 126 residues, 815−940), which was shown to be involved in centrin 2 and TFIIH binding. C-XPC may be highly expressed in *E. coli*, but because of its limited solubility it was purified under 6 M urea. Using bioinformatics tools, and a combination of several experimental methods (circular dichroism, fluorescence, nuclear magnetic resonance, and small-angle X-ray scattering), we show that C-XPC has a highly flexible structure under native physiological conditions, with a propensity to form helical secondary structures. Isothermal titration calorimetry experiments show that the C-XPC fragment binds human centrin 2 with high affinity and a 1:1 stoichiometry. NMR analysis indicates that the physical interaction between C-XPC and centrin 2 induces only minor conformational changes into XPC, localized around the 17-mer segment (847−863), showed to be critically involved in the centrin binding.

DNA molecules of living cells are permanently exposed to a variety of endogenous metabolic byproducts and exogenous aggressive agents including free radicals, electromagnetic radiation, and highly reactive chemicals. The damages produced under the action of these agents may seriously threaten the integrity of the genetic material and have important consequences on organism survival. Highly specific mechanisms have evolved to recognize and repair multiple forms of DNA lesions. Among them, nucleotide excision repair (NER[1]) is a versatile and efficient DNA repair system, essentially handling the bulk damages created on a single DNA strand by UV-light or chemical reagents (*1*). Among the lesions recognized by NER, the pyrimidine dimers (such as (6-4)pyrimidine-pyrimidone photoproduct), the chemical adducts, and the intrastrand cross-links (such as those induced by cisplatin) are the most relevant.

There are two distinct pathways of the NER: the global genome repair (GG-NER), which concerns all the genome, and the transcription-coupled repair (TC-NER) operating only on the transcribed strand of active genes. More than 30 different proteins participate in the GG-NER pathway, working in a precise sequential mechanism in the damaged region of the genomic DNA. During the first steps of the repair mechanism, a 125 kDa protein, named XPC (xeroderma pigmentosum group C), plays a critical role in the recognition of the target lesions and recruitment of the DNA incision complex (*2*, *3*). In contrast, this NER factor is not required for the preferential repair of template strands in transcribed genes (*4*). Two other proteins, Rad23B, one of the two human homologues of yeast Rad23, and centrin 2 (HsCen2), an EF-hand protein, were shown to be strongly associated to XPC and to modulate the DNA lesion recognition (*5*, *6*), by mechanisms which are not yet fully understood. Gene mutations in a number of NER factors are responsible for the human autosomal recessive disease xeroderma pigmentosum (XP), characterized by a high photosensitivity, a 1000-fold increase in basal cell carcinomas and squamous cell carcinomas incidence, and in some cases, neurological abnormalities (*7*). These symptoms are due to variable defaults in NER function, caused by abnormalities in one of the seven XP genes (XPA trough XPF). In the case of XPC, several mutations, generally leading to premature termination of the protein sequence, have been identified as the molecular cause of the XP phenotype (*8*).

† This work was supported by the Centre National de la Recherche Scientifique, the Institut National de la Santé et de la Recherche Médicale, and the Institut Curie-Section de Recherche.

* To whom correspondence should be addressed. Tel: 33 1 69863163; fax: 33 1 69075327; e-mail: Simona.Miron@curie.u-psud.fr and Gil.Craescu@curie.u-psud.fr.

‡ INSERM/Institut Curie.

§ Université Paris-Sud.

[1] Abbreviations: HsCen2, human centrin 2; C-HsCen2, the C-terminal half of HsCen2; DN25-HsCen2, HsCen2 deleted of the first 25 residues; CaM, calmodulin; CD, circular dichroism; HSQC, heteronuclear single-quantum correlation spectroscopy; ITC, isothermal titration calorimetry; NER, Nucleotide Excision Repair; P17-XPC, centrin binding peptide from XPC with the sequence N847-R863; TFE, 2,2,2-trifluoroethanol; XPC, xeroderma pigmentosum group C protein.

10.1021/bi701863u CCC: $40.75
Published on Web 01/05/2008

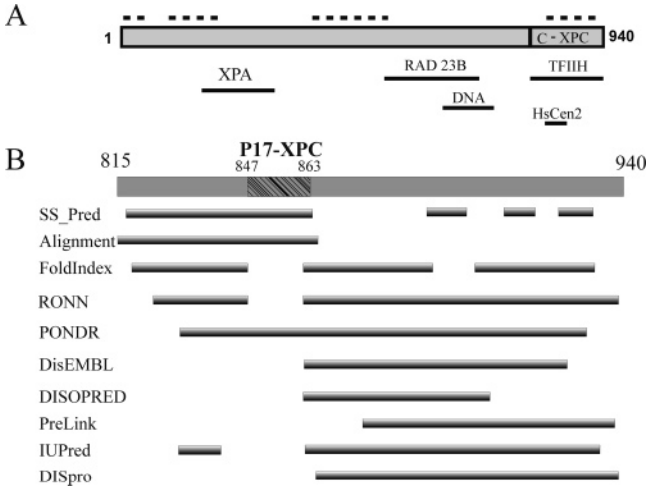

FIGURE 1: (A) Schematic representation of the localization of C-XPC and the binding domains on the XPC sequence. The dashed lines indicate the predicted (PONDR) unstructured fragments. (B) Sequence analysis of C-XPC using bioinformatics tools. The localization of the centrin binding site (P17-XPC) is indicated by the hatched segment. The predicted secondary structure elements were obtained using the Jpred program in the Proteomic tools of ExPASy website. The sequence portion with the best sequence alignment within 48 XPC homologues resulted from a multisequence alignment (CLUSTALW) (*19*) visualized with BioEdit (http://www.mbio.ncsu.edu/BioEdit). The predicted disordered fragments in C-XPC were obtained with various predictors in the DisProt website (http://www.disprot.org/predictors.php).

The three-dimensional structure of the human XPC is not known, but recent X-ray studies of a yeast homologue (Rad4) construct revealed the global structural architecture and the conformational features involved in Rad23 and DNA binding (*9*). The studied Rad4 construct lacks the first 100 residues and the last 122 residues which are highly susceptible to proteolysis, probably to the high flexibility and/or disorder. The structure contains an N-terminal (310 residues) α/β domain that binds to the undamaged DNA segment and three consecutive 50−90 residue α/β domains including a characteristic long β-hairpin motif. The last two β-hairpin domains play a key role in the damage recognition, by inserting into the DNA duplex and expelling the two damaged base pairs out of the double helix. The protein/DNA interactions involve the undamaged unpaired bases and not the damaged bases, thus explaining the large diversity of lesions repaired by the NER pathway.

As suggested by the sequence similarity, the structural features of Rad4 should be highly relevant for XPC as well. However, the human protein has a central insert of about 200 residues, which, together with the N- and C-terminal fragments, may constitute important regulatory domains. The molecular basis of their regulatory function (*9*) remains to be elucidated.

In addition to the Rad23B and HsCen2 binding, the recruitment of other NER factors on damaged DNA sites requires direct physical interactions of XPC with XPA (*10*), UV-DDB (*11*), thymine DNA-glycosylase (*12*), and TFIIH (*13*) (Figure 1A). Binding to XPA was recently shown to involve an N-terminal fragment (154−331) (*10*). DNA and Rad23B binding sites are situated in a central portion of the XPC sequence, while TFIIH binding is localized in the last 125 residue fragment (*14*) for which no structural information is available (*9*). The centrin binding site was the first to be precisely defined (847−863), allowing the high-resolution structural analysis of the XPC/HsCen2 interaction (*15*, *16*). As HsCen2 and TFIIH binding sites are situated within the same relatively small XPC fragment, the question arises whether there is any physical interference between the two binding events that suggests a possible regulatory process.

In this study, we characterize the molecular and functional properties of the 14 kDa C-terminal XPC domain (residues 815−940). This fragment is largely unfolded in aqueous solution, as observed by CD, NMR, and SAXS methods. In less polar solvents the domain shows a propensity to form α-helical secondary structure. Calorimetric and fluorescence titration experiments reveal a high affinity binding to the C-terminal domain of HsCen2, very similar to the previously studied 17-mer centrin-binding peptide (P17-XPC). The conformational changes in XPC induced by centrin binding were observed only in a limited segment centered on P17-XPC.

MATERIALS AND METHODS

Expression and Purification of C-XPC Protein. Nucleotide sequence coding for residues 815−940 was PCR-amplified using a plasmid containing the full sequence coding for XPC, a 5′ primer (5′ GGAATTCCATATGGAGGAATTCAAAGACGTGCT 3′) and a 3′ primer (5′ GGCGAGCTCAGTGTCGAAGAGTTTACC 3′). The purified PCR product was cloned into the expression vector pET24a+ at the NdeI and XhoI restriction sites. The resulting plasmid was sequenced. The recombinant protein C-XPC E815-L940 was overexpressed in the bacterial strain BL21DE3, as described previously (*17*). For uniformly ^{15}N-labeled proteins, a culture medium containing $^{15}NH_4Cl$ as the sole source of nitrogen was used.

The supernatant of the bacterial lysate was dialyzed against 10 mM Na_2HPO_4 buffer, pH 8.2, containing 1.5 mM β-mercaptoethanol. In these conditions C-XPC precipitated and was recovered by dissolving the pellet in the same buffer including 6 M urea. The obtained solution was chromatographed on a DEAE-TSK column (650S) using a NaCl gradient (0 to 200 mM) for elution. The C-XPC-containing fraction was concentrated on a PM_{10} column (Millipore) and applied on a SP Trisacyl column equilibrated in 50 mM Na_2HPO_4 buffer, pH 6.35, 1.5 mM β-mercaptoethanol, 6 M urea. The column was eluted using a NaCl gradient (0 to 600 mM), and after concentration on the PM_{10} column, the protein was renatured and desalted on a column Trisacryl GF05, decreasing progressively the urea from 6 to 0 M, in the presence of 0.5 M arginine. C-XPC was finally eluted from the column with 1.5% NH_4HCO_3 and lyophilized.

Sequence Analysis and Secondary Structure Prediction. Secondary structure analysis of C-XPC was conducted using Jpred, a consensus method of protein secondary structure prediction developed at University of Dundee (*18*) and accessible *via* ExPASy Proteomic tools server (http://www.expasy.org/tools). The multiple sequence alignment, using 48 XPC homologues, was done with the CLUSTALW software (*19*). The degree of structural organization of the C-XPC domain was analyzed with various disorder predictors using the links in the DisProt server (http://www.disprot.org/predictors.php) (*20*).

Circular Dichroism. A Jasco 715 spectropolarimeter equipped with a Peltier temperature control unit was used

to obtain circular dichroism spectra. Far-UV CD spectra were recorded between 195 and 260 nm, at 20 °C, using 1-mm quartz cells. The spectra were acquired as an average of five scans, with a scan speed of 50 nm/min and a response time of 2 s. The spectrum of the buffer was removed from the sample spectra. The CD samples were at 10 μM, in 50 mM phosphate buffer, pH 7.2 or 6.7, at 0, 100, and 200 mM NaCl. Temperature denaturation curves were recorded between 5 and 95 °C, with a temperature increase rate of 1 °C/min. The secondary structure variation upon addition of different concentrations of TFE (10%, 20%, 30%) was monitored, recording the CD spectra between 195 and 260 nm. The CD spectrum in the presence of 30% TFE displays a characteristic signature of a α-helical reach protein, with two negative bands at 208 and 222 nm. The secondary structure analysis was based on the CD data from 195 to 250 nm and used the CDPro software package (*21*), with CONTIN and CDSSTR tools.

Small-angle X-ray scattering (SAXS) experiments were carried out on the laboratory equipment (Nanostar, Bruker-AXS) installed at the Institut de Biophysique et Biologie Moléculaire et Cellulaire. The instrument includes a small-angle X-ray camera adapted to a rotating anode X-ray source (Cu Kα, $\lambda = 1.54$ Å), all placed in a vacuum chamber, in order to reduce the background. The scattering vector range was 0.012 Å^{-1} < q < 0.35 Å^{-1}, where $q = 4\pi \sin \theta/\lambda$, with 2θ being the scattering angle. The sample to detector distance was 662 mm. Samples were enclosed into 2 mm diameter quartz capillaries mounted in capillary holders directly inserted into vacuum and kept at constant temperature (18 °C). The sample solution was very dilute with a protein concentration of 1.4 mg/mL (about 100 μM). Both buffer (Tris HCl 20 mM, NaCl 200 mM, pH 6.7) and sample scattering were recorded for six 2-h frames. Data were collected using a two-dimensional position-sensitive gas detector (HiSTAR). The SAXS data were azimuthally averaged and corrected for variations in detector efficiency and for spatial distortions. Buffer measurements were used as background and subtracted from the recorded integrated curves.

The scattering curve $I(Q)$ is affected by a slight aggregation in the small angle region. This effect becomes absolutely negligible in the range $q > 0.031$ Å^{-1}. Nevertheless it is difficult to evaluate the radius of gyration R_g using the Guinier approximation (*22*) because, in the case of an unstructured protein, the Guinier approximation holds true in a very restricted angular range corresponding to $R_gQ < 0.8$. The radius of gyration is then evaluated using the indirect transform package GNOM (*23*) which provides the distance distribution function $p(r)$ of the protein that represents the histogram of distances within the particle. Its value is equal to zero when r exceeds D_{max}, the maximum dimension of the protein.

The conformation in solution of C-XPC was determined using two *ab initio* approaches. The program GASBOR represents a protein by an assembly of dummy residues (DR) centered at the Cα positions (*24*). By using a simulated annealing protocol, it builds a locally "chain compatible" dummy residue model inside a sphere with diameter D_{max}. The program BUNCH implements a combination of *ab initio* and rigid body modeling, using the high-resolution structures of protein domains (when available) and describing the missing parts as DR chains (*25*).

Fluorescence Titration. The Trp fluorescence measurements were performed on a Jasco FP777 spectrofluorimeter using a 295 nm excitation, at 20 °C, in a buffer containing 50 mM MOPS, pH 7.2, 100 mM NaCl. Fluorescence emission was monitored with a band-pass of 5 nm between 305 and 400 nm. For titration of C-XPC (1 μM) with HsCen2 or C-HsCen2, fluorescence intensity at 330 nm as a function of protein concentration was fitted to a one-site binding model using Origin 7.0 software (OriginLab Corp.).

NMR Experiments. Because of the limited solubility of C-XPC, the NMR samples were less than 100 μM. The lyophilized protein samples were dissolved in Trisd_{11} buffer (20 mM), pH 6.7, containing 0 or 100 mM NaCl, 0, 100, and 300 μM of $CaCl_2$ and 7% ^2H_2O. The NMR spectra were acquired at 298 K on a 500 MHz Varian Unity spectrometer and on a 700 MHz Bruker NMR spectrometer equipped with a cryoprobe (CEA, Saclay). The NMR data were processed and analyzed using Felix software (Accelrys, San Diego), running on Silicon Graphics Indigo workstations (Silicon Graphics, Mountain View, CA).

The (^{1}H−^{15}N) heteronuclear NMR spectra for centrin/C-XPC complexes were realized with only one of the partners uniformly ^{15}N-labeled: ^{15}N-HsCen2 in the complex with C-XPC and ΔN25-HsCen2 in the complex with the ^{15}N-C-XPC. Utilization of the truncated centrin form (ΔN25-HsCen2) was motivated by the lower aggregation tendency of this construct (*26*). The HSQC spectra are not significantly affected by the presence of the NaCl or the $CaCl_2$. The (^{1}H−^{15}N)HSQC data were apodized with a $\pi/4$ shifted sine-bell function in both dimensions, zero filled to a matrix of 1024 × 1024 real points prior to Fourier transformation. The 3D (^{1}H−^{15}N)NOESY-HSQC spectra were acquired with 8000 Hz (^{1}H) and 1500 Hz (^{15}N) spectral width. 128, 32, and 2048 complex points were collected in the t_1 (^{1}H), t_2 (^{15}N), and t_3 dimension, respectively.

Isothermal Titration Calorimetry. The physical interaction between the C-terminal domain of XPC protein and HsCen2 or C-HsCen2 were analyzed by ITC using a MCS ITC calorimeter (MicroCal Inc., Northampton, MA). Because of the solubility problem, C-XPC (15 μM) was placed in the cell and titrated by HsCen2 or C-HsCen2 (200 μM) using 10 μL automatic injections. Experiments were done at 30 °C, in 50 mM MOPS buffer, pH 7.2, and 100 mM NaCl. The first injection of 2 μL was ignored in the final data analysis. Integration of the peaks corresponding to each injection and correction for the baseline were done using Origin-based software provided by the manufacturer. Fitting of the data to an interaction model results in the stoichiometry (n), equilibrium binding constant (K_a), and enthalpy of complex formation (ΔH). The experimental data allow calculation of the free energy change (ΔG) and of the entropy term ($T\Delta S$) according to the classical thermodynamic formulas:

$$\Delta G = -RT \times \ln K_a$$

$$\Delta G = \Delta H - T\Delta S$$

where R is the universal gas constant and T is the absolute temperature.

RESULTS

Sequence Analysis of C-XPC Predicts a Highly Disordered Structure. Analysis of various structural data allows the

prediction that a large fraction of the eukaryotic proteins (up to 33%) contain long fragments (more than 30 residues) lacking any regular secondary and tertiary structure under physiological conditions and in the absence of a binding partner (*27*). During recent years, a lot of computer algorithms were developed to predict the disordered regions in proteins (for a review see ref *28* and references therein). Some of them are based on the experimental information on poorly structured proteins, and on highly flexible loops in crystal structures (PONDR, DisEMBL, Disopred2, RONN, PreLink, DISpro), while others use the physicochemical properties of the amino acid residues of the considered sequence (FoldIndex, IUPred). We first applied these predictors to the whole XPC sequence: several unstructured fragments are detected in the N-terminal, central, and C-terminal regions of the protein (Figure 1A). Therefore, the core of the global regular structure should be composed of two sequence stretches (an N-terminus and a C-terminus), localized on the two sides of the central disordered fragment of about 200 residues (*10*). The binding sites for some of the interacting partners, including XPA, Rad23B, and ssDNA, are localized in the folded region of XPC. The C-XPC fragment (815−940), involved in TFIIH (*14*) and HsCen2 (*29*), is predicted to be largely disordered. We focused the bioinformatics analysis on this part of sequence, combining the disorder predictors with the secondary structure predictions and the multiple sequence alignment.

Secondary structure prediction using a consensus method (*18*) identifies several segments with a significant propensity to form α-helix (Figure 1B): 817−863, 892−903, 911−917, and 924−932. The multiple sequence alignment, including 48 XPC homologues from various species, shows that the N-terminal part (815−865) is significantly conserved (Figure 1B). Some of the sequences stop shortly after this position, while others exhibit long (more than 300 residue) insertions. Overall, the two analysis procedures suggest that the first 47 residues (37% of the sequence) of C-XPC have a propensity to form α-helical structure and are significantly conserved.

Eight different disorder predictors, listed on the DisProt website (http://www.disprot.org/), were used to analyze the propensity of the C-XPC sequence to generate poorly defined, flexible conformations. As shown in Figure 1B, the consensus tendency is that a large C-terminal continuous fragment, representing about 62% of the sequence, is mainly disordered. Half of the prediction procedures also show that the N-terminal sequence, preceding the centrin binding motif, may also contain poorly structured regions. With one exception, the fragment harboring the centrin binding site is predicted to have an ordered structure.

The presence of a highly uncompensated net charge and a low content of hydrophobic amino acids are an important global indication for the absence of compact structure in proteins under physiological conditions (*30*). C-XPC domain has an amino acid composition rich (49.1%) in disorder-promoting residues (Arg, Lys, Glu, Gln, Pro, Ser) and a poor content (20.1%) of order-promoting residues (Trp, Phe, Ile, Tyr, Val, Leu), as compared with the corresponding composition of globular well-folded proteins (33.3% and 31.6%, respectively) (*31*).

C-XPC Exhibits α-Helical Secondary Structure but Denaturates as a Noncooperative Unit. CD spectroscopy is a method of choice for an initial experimental evaluation of the secondary structure content of a soluble protein. Figure 2A shows the far-UV CD spectrum of C-XPC in the absence or the presence of increasing concentrations of TFE. The spectral characteristics of helical secondary structure are two minima (222 and 208 nm) of comparable intensity, while the main feature of the unfolded state is a negative peak at 198 nm. In the absence of TFE, the C-XPC spectrum shows two unequal minima at 204 and 222 nm, suggesting the presence of a low proportion of α-helix, and a larger content of random coil. Similar CD spectra were obtained at different NaCl concentrations (from 0 to 200 mM).

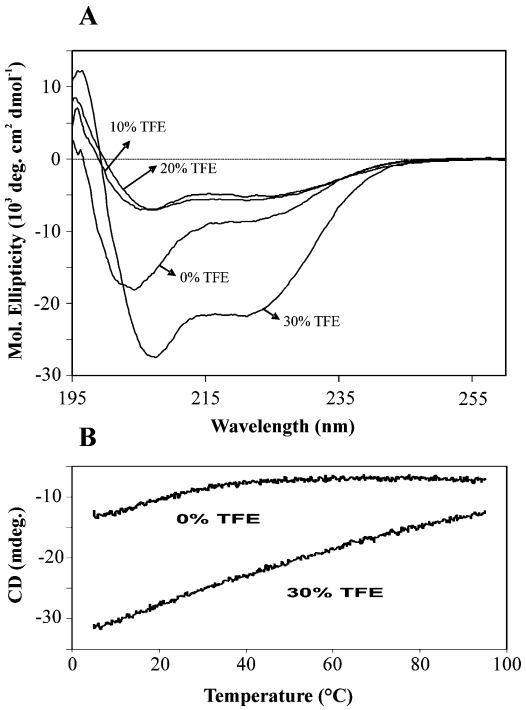

FIGURE 2: CD spectra of C-XPC. (A) Far-UV CD spectra at 20 °C of C-XPC (10 μM) alone or in the presence of various concentrations of TFE. (B) Thermal denaturation curves recorded between 5 °C and 95 °C at 222 nm, in the absence or the presence of TFE.

A quantitative analysis of the far-UV CD spectra of C-XPC domain was made by fitting the experimental data to an appropriate combination of "pure" secondary structure spectra. Utilization of two different fitting procedures, CONTIN (*32*) and CDSSTR (*33*), gave similar values for the secondary structure content of C-XPC: 33% and 37% α-helix, respectively. Thermal denaturation monitored by the ellipticity at 222 nm (Figure 2B) shows that the protein loses progressively (noncooperative unfolding) its helical structure between 5 °C and 40 °C, suggesting a low degree of tertiary structural organization (*34*).

The alcohol 2,2,2-trifluoroethanol (TFE) is often used as a cosolvent stabilizing the secondary structures (α-helix, β-turn, β-hairpin) in peptides and unfolded proteins (*35*, *36*). In a water/TFE mixture, the organic molecules accumulate around the structural elements, removing some of the water/protein hydrogen bonds, thus favoring the intramolecular bonds (*37*). Generally, TFE stabilizes helical conformations only in those parts of proteins which are helical in the native state (*38*). In some cases, at higher concentrations, TFE may destabilize native tertiary structures in intact proteins (*36*,

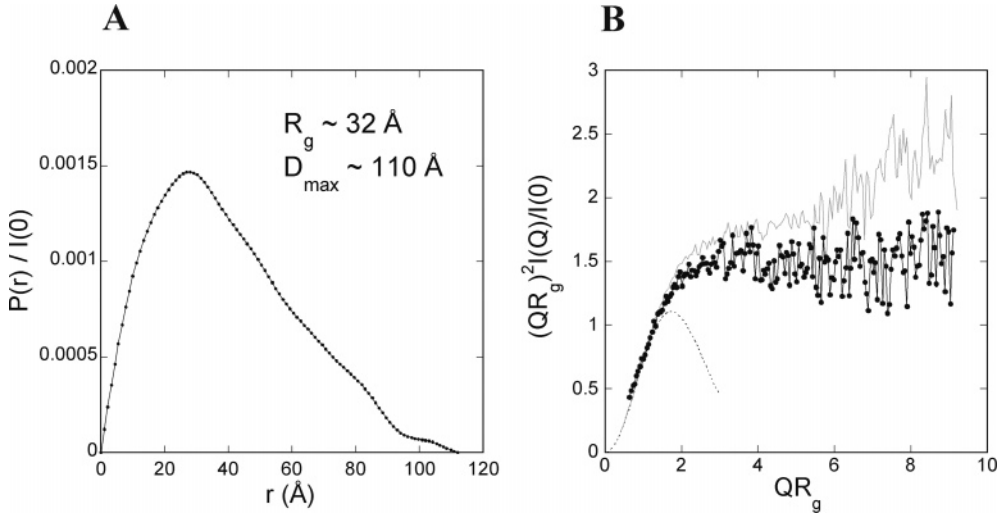

FIGURE 3: (A) Distance distribution function of C-XPC. The radius of gyration R_g and the maximum extension of the molecule D_{max}, estimated from the distance distribution function $P(r)$, are shown in the upper right side of the panel. (B) Dimensionless Kratky plots. Solid line with dots: experimental curve of C-XPC. Dashed line: theoretical curve corresponding to Guinier's law, describing a globular, compact protein. Gray line: experimental curve for an almost completely unstructured small domain (PIR) from the protein Grb14 (*42*).

39), by preferentially binding to hydrophobic side chains and hydrophobic patches. Figure 2A shows the far-UV CD spectra obtained at increasing TFE concentration (10%, 20%, 30%). The TFE effect is biphasic: at lower concentrations (10−20%) the minimum at 204 nm shifts to 208 nm, and the spectrum intensity is globally decreased. In this cosolvent concentration range, the samples were less soluble, and this may partly explain the lower CD intensity. At 30% TFE the sample is no longer aggregated, and the spectrum shows a considerably increased intensity, estimated to represent about 55% α-helix. Thermal denaturation of this state shows a noncooperative unfolding covering the whole temperature range, suggesting that stabilization of the α-helix secondary structure is not associated with the formation of a cooperative folding unit. Therefore, the CD experiments suggest that C-XPC naturally exhibits about 30% of the residues folded in α-helical structure, which may correspond to the N-terminal long helix predicted by Jpred (Figure 1B). Under favorable solvent conditions, additional helical fragments could be stabilized, likely those corresponding to the three short α-helices predicted in the C-terminal half of the domain. The biphasic TFE concentration dependence noted for the far-UV CD spectra was previously observed (*36*) and is consistent with different mechanisms of the structural effect of the alcohol molecules.

C-XPC Is Highly Disordered in the Native State. The SAXS method is particularly well adapted to study flexible or disordered proteins in solution, providing important structural parameters as the radius of gyration (R_g) and the maximum intramolecular distance (D_{max}). The distance distribution function $P(r)$, corresponding to the scattering pattern of a 1.4 mg/mL C-XPC solution in Tris buffer, 200 mM NaCl, is shown in Figure 3A. It exhibits a distinctly asymmetric shape typical of a nonglobular, extended protein conformation. The value of the radius of gyration derived from the $P(r)$ is 32 ± 1 Å and the maximal extension is 110 ± 10 Å. One could argue that the smallest useful q value (q_{min} = 0.031 Å^{-1}) is slightly larger than π/D_{max} (= 0.029 Å^{-1}), a value that is in theory too large for a sound determination of $P(r)$. We therefore checked that $P(r)$ is not significantly altered when calculated using the program Gnom with different cutoff values of q_{min} around 0.031 Å^{-1}. Furthermore, using urate oxidase (MM = 128 kDa) as a standard, an estimate of the molecular mass of C-XPC can be derived from the value of the intensity at the origin $I(0)$. The resulting value of *ca.* 15 kDa is in good agreement with the sequence-derived value (14 kDa), showing that aggregation effects are negligible in the q-range used for data analysis. The R_g value is 2-fold larger than the value expected for a compact, globular protein with the same number (n) of residues which is $R_g \sim 3n^{1/3} = 15$ Å (*40*). It can also be compared with the expected value for a completely unfolded protein. In this case, the protein can be described as a statistical chain with a persistent length, the radius of gyration of which is approximated by:

$$R_g = b\left[\frac{y}{6} - \frac{1}{4} + \frac{1}{4y} - \frac{1}{8y^2}\right]^{1/2}$$

with $y = L/b$, where L is the contour length of the chain and b is twice the persistence length with a typical value of 20 Å (*41*). L is expressed as $L = n \times a \times f$, where n is the number of residues, a is the characteristic dimension of one residue (3.78 Å), and f is a correcting factor taking into account the constraints of a polypeptide chain (f = 0.95). This expression yields a R_g value of 37 Å for a 126 residue random coil polypeptide. The experimental R_g value thus appears to be close but significantly lower than that for a persistent statistical chain.

A dimensionless Kratky plot $(qR_g)^2I(q)/I(0)$ vs qR_g shown in Figure 3B is also very useful to characterize the conformation (globular *versus* extended) of the protein in solution. The dotted line in Figure 3B is the theoretical curve corresponding to Guinier's law $I(q) = I(0)\exp(-q^2R_g^2/3)$, which can describe any compact, globular protein with a well-defined three-dimensional structure. The gray line represents the experimental scattering pattern of a small protein domain named PIR, which is almost completely unstructured, with less than 10% residual secondary structure (*42*). Comparison of all three curves in Figure 3B suggests that C-XPC is very different from a globular protein and behaves essentially as an unfolded chain, although having

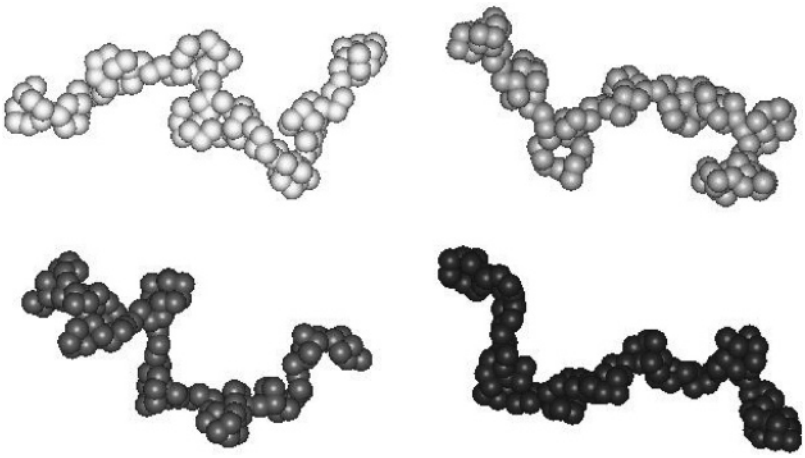

FIGURE 4: Four models of C-XPC obtained by GASBOR. Each dummy residue is represented by a sphere of radius 3.2 Å whose volume is equal to the average volume of an amino acid residue (about 140 Å^3).

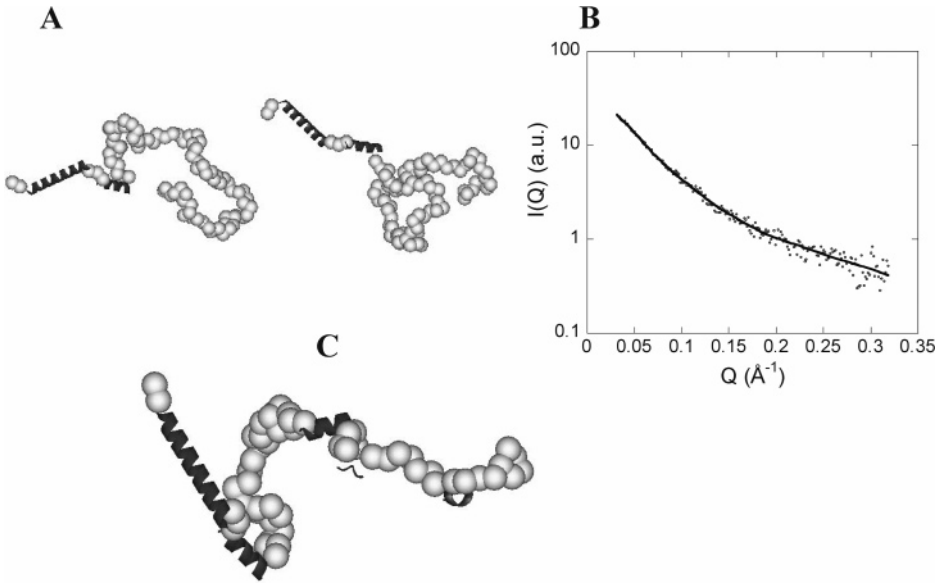

FIGURE 5: Ab initio modeling using the BUNCH program. (A) Two examples of C-XPC conformations obtained using BUNCH, starting from a chain of dummy residues containing 2 α-helices between residues 818 to 843 and 847 to 860. (B) X-ray scattering pattern of C-XPC. Dots: experimental points. Solid line: fit of the experimental data by the BUNCH models shown in panel A. (C) One model of C-XPC obtained by BUNCH, starting from an initial structure including five α-helices in 818−843, 847−860, 891−901, 908−915, and 923−930.

more secondary structure elements than the PIR domain. Therefore, the C-XPC domain is expected to explore an extended conformational space rather than adopt a unique, well-defined three-dimensional conformation in solution.

In order to obtain a structural representation of these results, we used two modeling approaches to optimize the protein conformational distribution under the constraint of the experimental data. The structural models may be considered as representative of the conformations in solution, as their calculated scattering pattern is identical to the average scattering pattern from all conformations in the sample.

The various models obtained using GASBOR program, which describes the protein as a chain of dummy residues, correspond to a mainly unfolded chain with a few amino acid clusters, suggesting the presence of some residual structure (Figure 4). Additional support for this interpretation is brought by the BUNCH modeling using a modified input structure which contains preformed α-helices. We first considered the existence of two α-helices between residues 818 to 843 and 847 to 860, including the centrin binding site, and corresponding to a total α-helical content of 32%, in agreement with circular dichroism data. The two helices were built using the program Pymol (http://pymol.sourceforge.net). The BUNCH program is used to determine the relative position of the two helices while modeling the rest of the molecule as a chain of dummy residues. Two typical models are shown in Figure 5A, and the corresponding calculated scattering pattern is in excellent agreement with the experimental scattering curve ($\chi^2 \sim 0.87$) (Figure 5B). C-XPC adopts multiple dissimilar conformations, with non-interacting α-helices. Finally we tested whether the existence of shorter α-helices within the C-terminal half of the protein modifies the calculated scattering pattern. Using BUNCH, new models were built, containing additional three short α-helices in 891−901, 908−915, and 923−930 (an example is shown in Figure 5C). The agreement of the calculated and experimental scattering pattern is also very good, meaning that the present experimental data could not discriminate between the two conformational models.

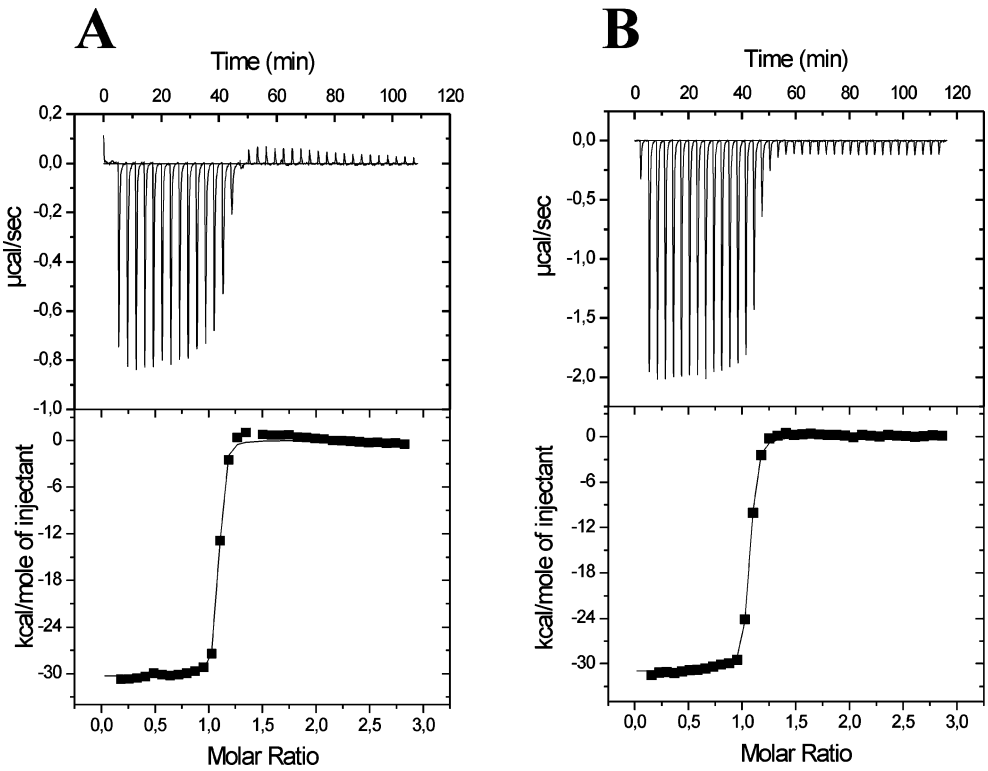

FIGURE 6: C-XPC binding to HsCen2. Row data (upper panel) and the isotherm (lower panel) of the titration of 200 μM HsCen2 (A) or C-HsCen2 (B) into C-XPC (15 μM) at 30 °C. Experimental data were fitted to a single-site binding model to give the stoichiometry (n), the binding constant (K_a) and the reaction enthalpy.

Table 1: Thermodynamic Parameters of the Binding of C-XPC or P17-XPC to HsCen2 and C-HsCen2[a]

protein	Ca^{2+} (mM)	protein (peptide)	K_a (10^6 M^{-1})	ΔG (kcal/mol)	ΔH (kcal/mol)	$T\Delta S$ (kcal/mol)
HsCen2	2	C-XPC	290(100)	−11.9	−30.2(0.2)	−18.3
HsCen2[b]	2	P17-XPC	220(40)	−11.6	−27.2(0.2)	−15.6
C-HsCen2	2	C-XPC	61.5(5)	−10.8	−30.9(0.1)	−20.1
C-HsCen2[b]	2	P17-XPC	120(50)	−11.2	−29.1(0.1)	−17.9

[a] ITC experiments were realized at 30 °C in 50 mM MOPS buffer, pH 7.4, NaCl 100 mM. The stoichiometry of the reactions is close to 1. [b] From ref *29*.

C-XPC Binds Strongly to HsCen2. Calorimetric Titrations. C-XPC construct includes the P17-XPC peptide which represents the centrin binding site (*6*, *29*, *43*). Therefore we should assess whether it preserves the centrin binding properties of the 17-mer peptide. Figure 6A shows the thermogram and the binding isotherm corresponding to the titration of HsCen2 (200 μM) into C-XPC (15 μM). Fitting the isotherm to a single-site binding model gives a stoichiometry of 1:1, a binding constant of 2.9 (1) $\times$ 10^8 M^{-1}, and a reaction enthalpy $\Delta H = -30.2$ (0.2) kcal/mol. Therefore, C-XPC binds HsCen2 with high affinity and thermodynamic parameters similar to that of the 17-mer peptide (Table 1). Titration of C-HsCen2 into the C-XPC solution (Figure 6B) gives a 5-fold lower affinity, indicating that binding of C-XPC is mainly mediated by the centrin C-terminal domain and that the N-terminal domain plays only a marginal role. A difference between P17-XPC binding by HsCen2 and C-HsCen2 was also noted previously (Table 1), but this is barely significant, as the free energy of binding varies by only 0.4 kcal/mol. It is possible that C-XPC, once bound to the centrin C-terminal domain, may establish some additional weak interactions with residues from the N-terminal domain, resulting in an increased affinity relative to P17-XPC.

The calculated entropy contribution ($T\Delta S$) to the free energy change of the C-XPC/HsCen2 interaction is negative (unfavorable), indicating that the interaction is driven by the enthalpy contribution. $T\Delta S$ results from two main components: a hydration effect (which is usually positive and therefore favorable) and a conformational term representing the restriction of the molecular movements of both partners upon complex formation. In our molecular system, the conformational entropy change overcomes the hydration entropy term, resulting in a total negative entropy change upon binding (*16*). The overall entropy change is slightly more negative (by 2−3 kcal/mol) when C-XPC replaces P17-XPC in the titration, suggesting that the larger fragment immobilizes only a few additional residues upon complex formation relative to the 17-mer. If the whole C-XPC polypeptide was folded upon interaction into a stable structure, we would expect a much larger increase of the entropic penalty. Indeed, upon the disorder-to-order transition, each residue loses significant conformational entropy, mainly related to the restriction in the number of accessible main chain dihedral angles and side chain rotamers. For instance, the side chain conformational entropy change upon protein folding was consensually estimated at about 1 kcal/

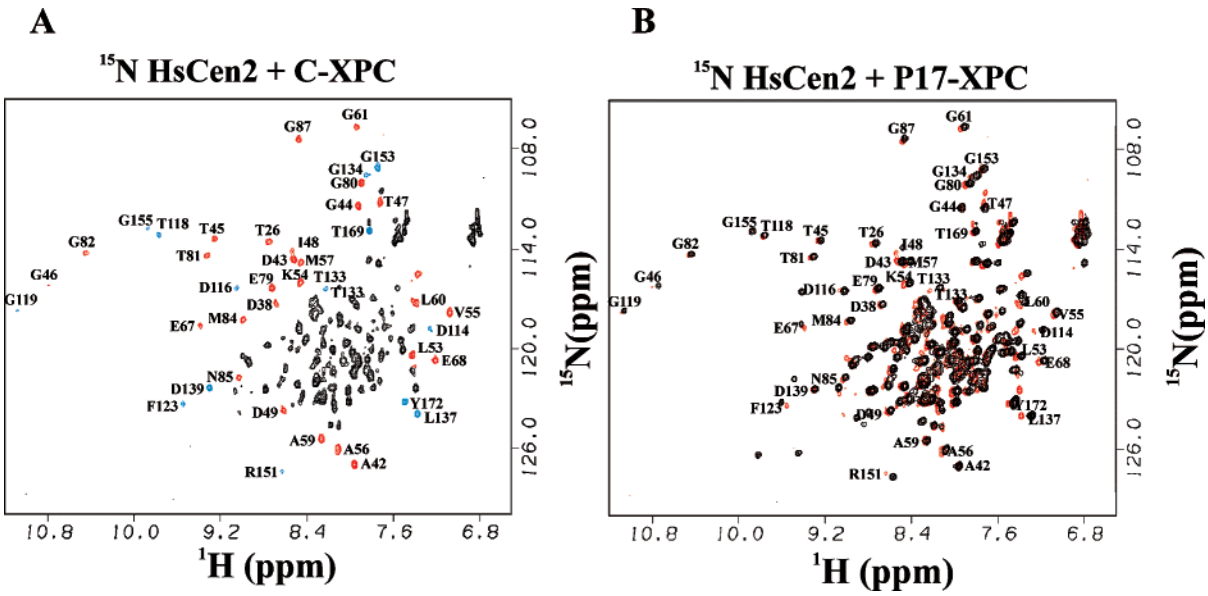

FIGURE 7: (^{1}H,^{15}N)HSQC spectra of HsCen2 complexes with C-XPC and P17-XPC. (A) ^{15}N-uniformly labeled HsCen2 in complex with C-XPC. Samples are in deuterated Tris-d_{11} buffer, pH 6.7, NaCl 100 mM, and the spectra were recorded at 35 °C. The labeled peaks, colored in red or blue, were unambiguously assigned to amide protons from the N- or C-terminal centrin domain, respectively. Because of the low solubility, the complex HsCen2/C-XPC concentration was ~100 μM. (B) Superposition of the HSQC spectra of complexes formed by ^{15}N-uniformly labeled HsCen2 with C-XPC (red) and P17-XPC (black).

mol/residue (*44*). Multiplied by the number of residues (except Gly and Ala), this may result in a substantial contribution to the overall entropy change of the interaction. The moderate change of the $T\Delta S$ term (relative to the 17-mer) observed here suggests that C-XPC interaction is not accompanied by a global folding of the domain.

Conformational Features of the HsCen2/C-XPC Interaction. Fluorescence and NMR spectroscopy were used to obtain structural information on the complex formation. C-XPC contains three Trp residues, one of them situated in the P17-XPC fragment, while HsCen2 has only one Tyr residue. Titration of 1 μM C-XPC by HsCen2 or C-HsCen2 shows that the maximum emission of the Trp fluorescence shifts progressively from 349 to 330 nm and increases in intensity (Supporting Information, Figure S1), indicating that at least one of the Trp residues of XPC is embedded in an apolar environment. Fitting the titration points (fluorescence emission intensity at 330 nm) to a single-site binding model gives affinity values very close to those obtained by ITC.

NMR structural study of the HsCen2/C-XPC complex is rendered difficult by the relatively large molecular mass (about 34 kDa), the flexibility of HsCen2, and the limited solubility (~100 μM) of the heterodimer. Nevertheless, we were able to obtain qualitative information from heteronuclear spectra of heterodimers where only one of the partners is uniformly ^{15}N-labeled. Figure 7 shows the HSQC spectra of the complexes formed by ^{15}N-HsCen2 with unlabeled C-XPC or the peptide P17-XPC. Despite the lower quality of the spectrum acquired with the larger complex, it appears that the two complexes have a highly similar fingerprint (Figure 7B), strongly suggesting that HsCen2 assumes the same backbone conformation in the two situations. On the basis of the previous NMR analysis of the complex HsCen2/P17-XPC (*43*), we assigned a significant number of peaks of ^{15}N-HsCen2 bound to C-XPC (Figure 7). These peaks constitute useful site-dependent conformational probes for both N-terminal and C-terminal centrin domains. The spectral analysis enabled us to conclude that HsCen2 assumes a very close global structure in the two compared complexes. Therefore, the presence of a large number of additional residues in C-XPC (relative to P17-XPC) does not significantly perturb the centrin conformation.

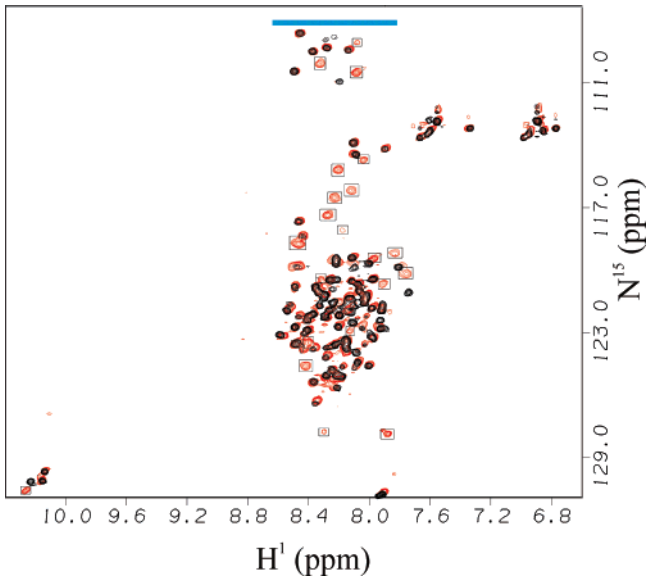

FIGURE 8: **Superposition of the HSQC spectra of** ^{15}N-C-XPC in isolation (black) or in complex (red) with unlabeled ΔN25-HsCen2. Samples are in Tris-d_{11} (20 mM) buffer, 100 mM NaCl, pH 6.7, 298 K. The C-XPC peaks which are significantly perturbed by the complex formation are boxed. The blue bar in the upper side of the figure indicates the range of amide chemical shift in P17-XPC peptide bound to C-HsCen2.

Figure 8 (black peaks) shows the (^{1}H−^{15}N)HSQC spectrum of uniformly ^{15}N-labeled C-XPC domain in the absence or the presence of ΔN25-HsCen2, a truncated form of human centrin 2, in which the first 25 N-terminal residues (which are not structured) have been deleted. This centrin construct avoids the aggregation problems, usually observed with HsCen2. The thermodynamic parameters obtained for HsCen2

or ΔN25-HsCen2 interaction with XPC target peptides are similar (*16*), indicating that the N-terminal fragment has no contribution to the C-XPC binding. The spectrum of C-XPC alone exhibits a very low spectral dispersion of the amide proton resonances (between 7.7 and 8.6 ppm), in a range typically observed for unfolded proteins (*45*). Moreover, in the (^{15}N)NOESY-HSQC spectrum no medium or long-range NOE cross-peaks were observed, suggesting a lack of persistent secondary or tertiary structure, in agreement with the SAXS results. The limited solubility of the sample and the low dispersion of the spectrum render difficult the C-XPC sequential resonance assignment.

In the presence of a slight molar excess of unlabeled ΔN25-HsCen2, C-XPC exhibits moderate spectral changes (Figure 8, red peaks). The chemical shift perturbations, induced by the complex formation, appear to be limited to a low proportion of peaks, representing about 16% of the residues (boxed peaks). The chemical shifts of the perturbed residues remain in the range observed for the isolated protein, but this is not in contradiction to the binding because the amide protons in the bound P17-XPC resonate in the same spectral range (the blue bar) (*43*). This suggests that, despite its binding to HsCen2, C-XPC domain maintains to a large extent its conformational flexibility and that only a small fragment of about 20 residues undergoes a structural rearrangement.

DISCUSSION

An increasing number of proteins are experimentally found to be intrinsically disordered or to contain fragments of 30 residues or longer lacking a persistent spatial structure (*31*, *46*, *47*). Many of the functionally annotated disordered proteins or domains play important biological roles in signaling, molecular assembly, or connection of active modules (*48*).

The ensemble of predictions and the experimental results obtained in this work allow us to conclude that the C-terminus of human XPC is naturally unfolded and functions as a regulatory module, by strongly binding to HsCen2. Existence of secondary structure elements (shown by CD experiments) but absence of tertiary persistent interactions (suggested by the absence of chemical shift dispersion and of long-range NOE interactions in NMR studies, as well as by the noncooperative thermal denaturation) could suggest that C-XPC is a molten (or pre-molten) globule (*27*). However, the large values for the radius of gyration and the maximum diameter, estimated from SAXS experiments, are characteristic for an extended, coil-like conformation. *Ab initio* modeling shows that C-XPC is mostly extended, with transitory helical regions covering at least the P17-XPC sequence and separated by unstructured segments (see Figure 5A).

Most intrinsically unstructured proteins or domains undergo a disorder-to-order transition upon binding to specific targets (*49*). The present experiments show that the strong binding of C-XPC to HsCen2 is not accompanied by a global folding of C-XPC. We may consider that the conformational changes are limited to a short XPC fragment which includes the high affinity P17-XPC peptide. This fragment is probably the single portion of the isolated C-XPC to assume a regular structure, predicted as a α-helix (Figure 1B). Binding to HsCen2 is coupled to a stabilization of the pre-existing helix.

Biological Implications. The involvement of intrinsically disordered proteins or domains in regulation and signaling processes are thought to have several advantages over rigid proteins (*31*, *49*). Disordered regulatory domains may have a larger binding diversity, recognizing different partners by adapting their structure to each particular intermolecular interface. In the absence of pre-existing, rigid binding geometry, the binding affinity is usually lower, but the interaction may be easily reversed. Lack of compact structure also enables faster association rates, because most of the approach geometries lead to molecular recognition and complex formation. In addition, an extended, flexible chain could easily fit to a larger binding surface, thus increasing the binding enthalpy.

C-XPC has the capacity to interact both with HsCen2 (*50*) and one or more subunits of the TFIIH factor (*14*). Its flexibility should be a favorable factor in adapting to the two different protein binding sites. Existence of a flexible C-terminal domain in XPC is evocative of other NER proteins having highly variable and flexible termini like XPA (*51*) or XPB (*52*). While the core of these proteins, which is the structural support for the main biological function, is highly conserved from prokaryotes to eukaryotes, variable sequence extensions on both sides appeared in higher eukaryotes during evolution (*52*). These fragments are usually involved in multiple interactions with other proteins within a functional network having regulatory or recruitment roles.

In the strong XPC/HsCen2 complex, the intermolecular interface is constituted by the hydrophobic side of a three-turn α-helix from XPC (mainly W848, L851, and L855) and the apolar surface exposed by the C-terminal domain of centrin (*43*). The deep embedding of the XPC helix and the complementarity of the two molecular surfaces account for the major part of the interacting free energy (*16*). If the centrin binding helix in the isolated XPC is structured and involved in stable intramolecular interactions with the protein core, it is likely that this helix will be arranged such as to expose its hydrophilic face to the solvent and to hide the hydrophobic face. Such a conformation would be hardly compatible with a rapid molecular recognition and a high binding affinity. It is therefore more probable that the C-XPC fragment is structurally disconnected from the XPC core and highly flexible in solution, as predicted by the present work.

The structural similarity between centrin and CaM has suggested that the two proteins may recognize the same target motifs. The hypothesis was experimentally demonstrated by showing that HsCen2 and CaM share a common target motif with the sequence pattern 1-4-8, in which a Trp is in position 1 and hydrophobic side chains are in positions 4 and 8 (*29*). Because of its structural plasticity, CaM may actually bind at least five distinct target motifs and a large variety of sequences (*53*). A comprehensive analysis of structures and sequences of CaM-regulated proteins (*54*) concluded that most of the CaM-binding domains are disordered or are flanked by disordered regions. This is in agreement with the binding mode which requires that the target motif (generally an α-helix) is sterically isolated in order to be engulfed within the hydrophobic pocket of the partner. A common mechanism for many CaM-activated protein kinases or phosphatases was described as an autoinhibitory domain displacement (*55*), in which binding of CaM displaces the pseudo-substrate inhibitory domain and allows for full

enzyme activity. We can imagine that the centrin role in the NER process is performed through a similar mechanism: binding to the C-terminal domain of XPC may remove some sterical constraints from the active site of the protein, thus modulating the affinity and the specificity of damaged DNA recognition.

The centrin binding site is predicted to be the most stable helical element of C-XPC in isolation (Figure 1B). The XPC/centrin interaction is strong ($\sim 10^8$ M^{-1}) and shows a weak dependence on Ca^{2+}, suggesting that the complex is permanent in normal cell conditions. Indeed, XPC is always purified as heterotrimer including HsCen2 and Rad23B (*50*). This may explain why the C-terminal end of XPC is not rapidly degraded by the proteases (there are 28 potential cleavage sites for trypsin, for instance). Using cellular imaging we have recently shown that, in the absence of UV-radiation, overexpression of XPC translocates additional amounts of centrin to the nucleus (*16*) through the physical interaction between the two proteins. Recruitment of the TFIIH complex after damage detection is critically dependent on the interactions between C-XPC and the transcription factor (XPB or p62 subunit) (*13*). This raises the question of the relationships between TFIIH and HsCen2 binding to neighboring (or partially overlapping) XPC sites.

The NER was shown to function in cell-free reconstituted systems including the purified molecular factors (*56*, *57*). The presence of centrin 2 in these systems is not essential for the DNA repair process, suggesting that HsCen2 is dispensable for NER, at least *in vitro* (*6*). In contrast, *in vivo* studies have shown that HsCen2 enhances significantly the efficiency of 6-4PP DNA damage repair (*6*), likely through the interaction with XPC. On the other hand, experimental observations, based on gel shift assays with purified protein factors (*6*), seem to reject the hypothesis of a direct structural correlation between HsCen2 and TFIIH binding to XPC. Understanding of the regulatory role of HsCen2 during the initial stages of NER would require more precise estimation of thermodynamic parameters characterizing the multicomponent interactions including the NER factors and the damaged DNA.

ACKNOWLEDGMENT

We thank Ao Yang for help with ITC data collection, Joël Couprie for help with NMR spectroscopy, and Thierry Magnaldo for the XPC cDNA.

SUPPORTING INFORMATION AVAILABLE

The fluorescence titration of C-XPC fragment by human centrin 2. This material is available free of charge *via* the Internet at http://pubs.acs.org

REFERENCES

1. de Laat, W. L., Jaspers, N. G. J., and Hoeijmakers, J. H. J. (1999) Molecular mechanism of nucleotide excision repair, *Gen. Dev. 13*, 768−785.
2. Sugasawa, K., Ng, J. M. Y., Masutani, C., Iwai, S., van der Spek, P. J., Eker, A. P. M., Hanaoka, F., Bootsma, D., and Hoeijmakers, J. H. J. (1998) Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair, *Mol. Cell 2*, 223−232.
3. Maillard, O., Solyom, S., and Naegeli, H. (2007) An aromatic sensor with aversion to damaged strands confers versatility to DNA repair, *PLoS Biol. 5*, e79.
4. Venema, J., van Hoffen, A., Karcagi, V., Natarajan, A. T., van Zeeland, A. A., and Mullenders, L. H. F. (1991) Xeroderma pigmentosum complementation group C cells remove pyrimidine dimers selectively from the transcribed strand of active genes, *Mol. Cell. Biol. 11*, 4128−4134.
5. Masutani, C., Araki, M., Sugasawa, K., van der Spek, P. J., Yamada, A., Uchida, A., Maekawa, T., Bootsma, D., Hoeijmakers, J. H. J., and Hanaoka, F. (1997) Identification and characterization of XPC-binding domain of hHR23B, *Mol. Cell Biol. 17*, 6915−6923.
6. Nishi, R., Okuda, Y., Watanabe, E., Mori, T., Iwai, S., Masutani, C., Sugasawa, K., and Hanaoka, F. (2005) Centrin 2 stimulates nucleotide excision repair by interacting with xeroderma pigmentosum group C protein, *Mol. Cell Biol. 25*, 5664−5674.
7. Bootsma, D., Kraemer, K. H., E, C. J., and Hoeijmakers, J. H. J. (2001) in *The Metabolic and Molecluar Bases of Inherited Disease* (Beaudet, A. L., Sly, W. S., and Valle, D., Eds.) pp 677−703, McGraw-Hill Book Co., New York.
8. Chavanne, F., Broughton, B. C., Pietra, D., Nardo, T., Browitt, A., lehmann, A. R., and Stefanini, M. (2000) Mutations in the XPC gene in families with Xeroderma Pigmentosum and consequences at the cell, protein, and transcription levels, *Cancer Res. 60*, 1974−1982.
9. Min, J.-H., and Pavletich, N. P. (2007) Recognition of DNA damage by Rad4 nucleotide excision repair protein, *Nature 449*, 570−575.
10. Bunick, C. G., Miller, M. R., Fuller, B. E., Fanning, E., and Chazin, W. J. (2006) Biochemical and structural domain analysis of xeroderma pigmentosum complementation group C protein, *Biochemistry 45*, 14965−14979.
11. Sugasawa, K., Okuda, Y., Saijo, M., Nishi, R., Matsuda, N., Chu, G., Mori, T., Iwai, S., Tanaka, K., Tanaka, K., and Hanaoka, F. (2005) UV-induced ubiquitylation of XPC protein mediated by UB-DDB-ubiquitin ligase complex, *Cell 121*, 387−400.
12. Shimizu, Y., Iwai, S., Hanaoka, F., and Sugasawa, K. (2002) Xeroderma pigmentosum group C protein interacts physically and functionally with thymine DNA glycosylase, *EMBO J. 22*, 164−173.
13. Yokoi, M., Masutani, C., Maekawa, T., Sugasawa, K., Ohkuma, Y., and Hanaoka, F. (2000) The xeroderma pigmentosum group C protein complex XPC-HR23B plays an important role in the recruitment of transcription factor IIH to dammaged DNA, *J. Biol. Chem. 275*, 9870−9875.
14. Uchida, A., Sugasawa, K., Masutani, C., Dohmae, N., Araki, M., Yokoi, M., Ohkuma, Y., and Hanaoka, F. (2002) The carboxy-terminal domain of XPC protein plays a crucial role in nucleotide excision repair trough interactions with transcription factor IIH, *DNA Repair 1*, 449−461.
15. Thompson, J. R., Ryan, Z. C., Salisbury, J. L., and Kumar, R. (2006) The structure of the human centrin 2-xeroderma pigmentosum group C protein complex, *J. Biol. Chem. 281*, 18746−18752.
16. Charbonnier, J.-B., Renaud, E., Miron, S., Le Du, M. H., Blouquit, Y., Duchambon, P., Christova, P., Shosheva, A., Rose, T., Angulo, J. F., and Craescu, C. T. (2007) Structural, thermodynamic and cellular characterization of human centrin 2 interaction with xeroderma pigmentosum group C protein, *J. Mol. Biol. 373*, 1032−1046.
17. Blouquit, Y., Duchambon, P., Brun, E., Marco, S., Rusconi, F., and Sicard-Roselli, C. (2007) High sensitivity of human centrin 2 toward radiolytical oxidation: C-terminal tyrosinyl residue as the main target. *Free Radic. Biol. Med. 43*, 216−228.
18. Cuff, J. A., Clamp, M. E., Siddiqui, A. S., Finlay, M., and Barton, G. J. (1988) Jpred: A Consensus Secondary Structure Prediction Server, *Bioinformatics 14*, 892−893.
19. Higgins, D., Thompson, J., Gibson, T., Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressivemultiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, *Nucleic Acids Res. 22*, 4673−4680.
20. Vucetic, S., Obradovic, Z., Vacic, V., Radivojac, P., Peng, K., Iakoucheva, L. M., Cortese, M. S., Lawson, J. D., Brown, C. J., Sikes, J. G., Newton, C. D., and Dunker, A. K. (2005) DisProt: A Database of Protein Disorder, *Bioinformatics 21*, 137−140.
21. Sreerama, N., and Woody, R. W. (2000) Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with expanded reference set, *Anal. Biochem. 287*, 252−260.

22. Guinier, A., and Fournet, G. (1955) *Small angle scattering of X-rays*, Wiley, New York.
23. Svergun, D. I. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria, *J. Appl. Crystallogr. 25*, 495−503.
24. Svergun, D. I., Petoukhov, M. V., and Koch, M. H. (2001) Determination of domain structure of proteins from X-ray solution scattering, *Biophys. J. 80*, 2946−2953.
25. Petoukhov, M. V., and Svergun, D. I. (2005) Global rigid body modeling of macromolecular complexes against small-angle scattering data, *Biophys. J. 89*, 1237−1250.
26. Tourbez, M., Firanescu, C., Yang, A., Unipan, L., Duchambon, P., Blouquit, Y., and Craescu, C. T. (2004) Calcium-dependent self-assembly of human centrin 2, *J. Biol. Chem. 279*, 47672−47680.
27. Uversky, V. N. (2002) Natively unfolded proteins: a point where biology waits for physics, *Protein Sci. 11*, 735−756.
28. Ferron, F., Longhi, S., Canard, B., and Karlin, D. (2006) A practical overview of protein disorder prediction methods, *Proteins 65*, 1−14.
29. Popescu, A., Miron, S., Blouquit, Y., Duchambon, P., and Craescu, C. T. (2003) Xeroderma pigmentosum group C protein possesses a high affinity binding site for human centrin 2 and calmodulin, *J. Biol. Chem. 278*, 40252−40261.
30. Gast, K., Damaschun, H., Eckert, K., Schulze-Forster, K., Maurer, H. R., Müller-Frohne, M., Zirwer, D., Czarnecki, J., and Damaschun, G. (1995) Protymosin a: a biologically active protein with random coil conformation, *Biochemistry 34*, 13211−13218.
31. Radivojac, P., Iakoucheva, L. M., Oldfield, C. J., Obradovic, Z., Uversky, V. N., and Dunker, A. K. (2007) Intrinsic disorder and functional proteomics, *Biophys. J. 92*, 1439−1456.
32. Provencher, S. W., and Glöckner, J. (1981) Estimation of protein secondary structure from circular dichroism, *Biochemistry 20*, 33−37.
33. Johnson, W. C. (1999) Analysing protein circular dichroism spectra for accurate secondary structures, *Proteins: Struct. Funct. Genet. 35*, 307−312.
34. Uversky, V., N., Gillespie, J., R., and Fink, A., L. (2000) Why are "natively unfolded" proteins unstructured under physiological conditions?, *Proteins Struct. Funct. Genet. 41*, 415−427.
35. Buck, M. (1998) Trifluoroethanol and colleagues:cosolvents come of age. Recent studies with peptides and proteins, *Q. Rev. Biophys. 31*, 297−355.
36. Povey, J. F., Smales, C. M., Hassard, S. J., and Howard, M. J. (2007) Comparison of the effects of 2,2,2-trifluoroethanol on peptide and protein structure and function, *J. Struct. Biol. 157*, 329−338.
37. Thomas, P. D., and Dill, K. A. (1993) Local and nonlocal interactions in globular proteins and mechanisms of alcohol denaturation, *Protein Sci. 2*, 2050−2065.
38. Dyson, H. J., Merutka, G., Waltho, J. P., Lerner, R. A., and Wright, P. E. (1992) Folding of peptide fragments comprising the complete sequence of proteins. Models for initiation of protein folding. I. Myohemerythrin, *J. Mol. Biol. 226*, 795−817.
39. Gast, K., Zirwer, D., Müller-Frohne, M., and Damaschun, G. (1999) Trifluoroethanol-induced conformational transitions of proteins: insights gained from the differences between α-lactalbumin and ribonuclease A, *Prot. Sci. 8*, 625−634.
40. Petoukhov, M. V., and Svergun, D. I. (2005) Global rigid body modeling of macromolecular complexes against small-angle scattering data, *Biophys. J.*, 1237−1250.
41. Rowe, G., and Lopez, A. (1990) Influence of the solvent on the conformational-dependent properties of random-coil polypeptides. I. The mean-square of the end-to-end distance and of the dipole moment, *Biophys. Chem. 36*, 57−64.
42. Moncoq, K., Broutin, I., Craescu, C. T., Vachette, P., Ducruix, A., and Durand, D. (2004) SAXS study of the PIR domain from the Grb14 molecular adaptor: a natively unfolded protein with a transient structure primer? *Biophys. J. 87*, 4056−4064.
43. Yang, A., Miron, S., Mouawad, L., Duchambon, P., Blouquit, Y., and Craescu, C. T. (2006) Flexibility and Plasticity of Human Centrin 2 Binding to the Protein XPC from Nuclear Excision Repair, *Biochemistry 45*, 3653−3663.
44. Doig, A. J., and Sternberg, M. J. (1995) Side-chain conformational entropy in protein folding, *Prot. Sci. 4*, 2247−2251.
45. Dyson, H. J., and Wright, P. E. (2004) Unfolded proteins and protein folding studied by NMR, *Chem. Rev. 104*, 3607−3622.
46. Dunker, A. K., Lawson, J. D., Brown, C. J., Williams, R. M., Romero, P., Oh, J. S., Oldfield, C. J., Campen, A. M., Ratliff, C. M., Hipps, K. W., Ausio, J., Nissen, M. S., Reeves, R., Kang, C., Kissinger, C. R., Bailey, R. W., Griswold, M. D., Chiu, W., Garner, E. C., and Obradovic, Z. (2001) Intrinsically disordered protein, *J. Mol. Graph. Model. 19*, 26−59.
47. Fink, A. L. (2005) Natively unfolded proteins, *Curr. Opin. Struct. Biol. 15*, 35−41.
48. Dunker, A. K., Brown, C. J., and Obradovic, Z. (2002) Identification and functions of usefully disordered proteins, *Adv. Prot. Chem. 62*, 25−49.
49. Dyson, H. J., and Wright, P. E. (2002) Coupling of folding and binding for unstructured proteins, *Curr. Opin. Struct. Biol. 12*, 54−60.
50. Araki, M., Masutani, C., Takemura, M., Uchida, A., Sugasawa, K., Kondoh, J., Ohkuma, Y., and Hanaoka, F. (2001) Centrosome protein centrin 2/caltractin 1 is part of the Xeroderma Pigmentosum group C complex that initiates global genome nucleotide excision repair, *J. Biol. Chem. 276*, 18665−18672.
51. Iakoucheva, L. M., Kimzey, A. L., Masselon, C. D., Bruce, J. E., Garner, E. C., Brown, C. J., Dunker, A. K., Smith, R. D., and Ackerman, E. J. (2001) Identification of intrinsic order and disorder in the DNA repair protein XPA, *Protein Sci. 10*, 560−571.
52. Fan, L., Arvai, A. S., Cooper, P. K., Iwai, S., Hanaoka, F., and Tainer, J. A. (2006) Conserved XPB core structure and motifs for DNA un winding: implications for pathway selection of transcription or excision repair, *Mol. Cell 22*, 27−37.
53. Yap, K. L., Kim, J., Truong, K., Sherman, M., Youan, T., and Ikura, M. (2001) Calmodulin target database, *J. Struct. Funct. Genomics 1*, 8−14.
54. Radivojac, P., Vucetic, S., O'Connor, T. R., Uversky, V. N., Obradovic, Z., and Dunker, A. K. (2006) Calmodulin signaling: analysis and prediction of a disorder-dependent molecular recognition, *Proteins 63*, 398−410.
55. Hoeflich, K. P., and Ikura, M. (2002) Calmodulin in action: diversity in target recognition and activation mechanisms, *Cell 108*, 739−742.
56. Araujo, S. J., Nigg, E. A., and Woody, R. W. (2001) Strong functional interactions of TFIIH with XPC and XPG in human DNA nucleotide excision repair, without a preassembled repairosome, *Mol. Cell Biol. 21*, 2281−2291.
57. Masutani, C., Sugasawa, K., Yanagisawa, J., Sonoyama, T., Ui, M., Enomoto, T., Takio, K., Tanaka, K., van der Spek, P. J., Bootsma, D., Hoeijmakers, J. H. J., and Hanaoka, F. (1994) Purification and cloning of a nucleotide excision repair complex involving the xeroderma pigmentosum group C protein and a human homologue of yeast RAD23, *EMBO J. 13*, 1831−1843.

BI701863U