Domain Architecture of the p62 Subunit from the Human Transcription/Repair Factor TFIIH Deduced by Limited Proteolysis and Mass Spectrometry Analysis†

Anass Jawhari,‡,§ Stéphanie Boussert,‖ Valérie Lamour,‡,⊥ R. Andrew Atkinson,‡ Bruno Kieffer,‡ Olivier Poch,‡ Noelle Potier,‖ Alain van Dorsselaer,‖ Dino Moras,*,‡ and Arnaud Poterszman‡

Laboratoire de Biologie et Génomique Structurale, Institut de Génétique et de Biologie Moléculaire et Cellulaire, BP 163, 67404 Illkirch Cedex, France, and Laboratoire de Spectrométrie de Masse Bio-organique ECPM, Université Louis Pasteur−25, Rue Becquerel, 67087 Strasbourg Cedex 2, France

Received June 1, 2004; Revised Manuscript Received August 31, 2004

ABSTRACT: TFIIH is a multiprotein complex that plays a central role in both transcription and DNA repair. The subunit p62 is a structural component of the TFIIH core that is known to interact with VP16, p53, Erα, and E2F1 in the context of activated transcription, as well as with the endonuclease XPG in DNA repair. We used limited proteolysis experiments coupled to mass spectrometry to define structural domains within the conserved N-terminal part of the molecule. The first domain identified resulted from spontaneous proteolysis and corresponds to residues 1−108. The second domain encompasses residues 186−240, and biophysical characterization by fluorescence studies and NMR analysis indicated that it is at least partially folded and thus may correspond to a structural entity. This module contains a region of high sequence conservation with an invariant FWxxΦΦ motif (Φ representing either tyrosine or phenylalanine), which was also found in other protein families and could play a key role as a protein−protein recognition module within TFIIH. The approach used in this study is general and can be straightforwardly applied to other multidomain proteins and/or multiprotein assemblies.

Originally identified as a basal transcription factor, human TFIIH was subsequently found to contain XPB and XPD, two helicases involved in nucleotide excision repair (NER)[1] (*1*−*4*), as well as the recently identified TTDA protein (*48*). Mutations in these subunits are responsible for three rare human genetic disorders, xeroderma pigmentosum (XP), Cockayne syndrome, and trichothiodystrophy (*5*, *6*). TFIIH is a multiprotein complex composed of nine large subunits ranging from 32 to 89 kDa. These subunits are assembled into two subcomplexes: the core TFIIH, composed of six subunits (XPB, XPD, p62, p52, p44, and p34) and the cyclin-dependent kinase-activating kinase (CAK), composed of cdk7, cyclin H, and MAT1 (see Figure 2A). Recently, the molecular structures of both human and yeast TFIIHs have been determined by electron microscopy and show similarities in size, shape, and architecture (*7*, *8*).

To understand the functions of TFIIH in transcription, efforts have been directed at systematically characterizing all of its subunits. Several functions have been determined for individual components: XPB and XPD are ATP-dependent helicases indispensable for opening the DNA around a promoter and/or a DNA lesion (*9*−*11*). Cdk7 is a serine/threonine kinase, which is regulated by cyclin H and MAT1 and phosphorylates several substrates including the C-terminal domain of RNA polymerase II (*12*). The N-terminal part of p44, a subunit of core TFIIH, has been shown to positively regulate XPD helicase activity, whereas the C-terminal part is involved in promoter escape (*13*, *14*). XPD helicase regulation is lacking in a majority of XPD patients because mutations in the C-terminal part of XPD abolish the XPD/p44 interaction (*14*). Recently, p52 was shown to regulate the function of XPB by conditioning its anchoring to the TFIIH complex through pairwise interactions (*15*). The two remaining subunits, p62 and p34, contain no enzymatic activities, and their function, as components of the TFIIH structure, remains unclear.

The p62 subunit of TFIIH binds a variety of transcription and DNA repair factors and is involved in the recruitment of TFIIH within large transcription initiation or NER complexes. TFIIH is recruited at a late stage of the formation of the preinitiation complex formation by the general transcription factor TFIIE (for a review, see ref *16*). This functional and physical interaction, characterized in yeast and mammalian cells, requires the direct binding of the p62 subunit of TFIIH to the α subunit of TFIIE (*17*). The N-terminal part of the p62 subunit is important for the recruitment of transcription activators such as the acidic transactivators p53 and VP16 (*18*, *19*). The nuclear hormone receptor ERα is phosphorylated by TFIIH and associates with TFIIH via the p62 subunit (*20*). Furthermore, p62 participates

† This work was supported by Université Louis Pasteur de Strasbourg, le Centre National de la Recherche Scientifique, l'Institut National de la Santé et de la Recherche Médicale, and SPINE (QLG2-CT-2002−00988). J.A. was supported by a grant from the Association pour la Recherche sur le Cancer.

* To whom correspondence should be addressed. E-mail: moras@igbmc.u-strasbg.fr.

‡ Institut de Génétique et de Biologie Moléculaire et Cellulaire.

§ Present address: The Scripps Research Institute, 10550 North Torrey Pines Rd., La Jolla, CA 92037.

‖ Université Louis Pasteur.

⊥ Present address: The Rockefeller University, 1230 York Avenue, New York, NY 10021.

[1] Abbreviations: NER, nucleotide excision repair; XP, xeroderma pigmentosum; ES, electrospray; MALDI−TOF, matrix-assisted laser desorption ionization−time of flight.

10.1021/bi048884c CCC: $27.50
Published on Web 10/20/2004

in the recognition of XPC-HR23B (*21*), an initial damage recognition factor, as well as XPG (*22*, *23*), a structure-specific endonuclease that cleaves DNA at the 3′ end of the lesion.

The identification of domain boundaries in multidomain proteins is of practical interest for the identification of protein regions exhibiting structural independence and folding autonomy (*24*), suitable for structural studies. Domain borders may be estimated by computational approaches based on multiple sequence alignments and/or homology-modeling studies. However, when a limited number of homologous sequences are available and in absence of structural data, the precision of these theoretical approaches is limited and experimental data are required. Limited proteolysis and the subsequent identification of protease-resistant fragments offer an experimental route for the determination of exact boundaries of the domains. The basic premise of this approach is that proteolysis of multidomain proteins occurs more readily in the less structured linker regions between the more densely packed domains than within the domains themselves (*25*).

In the present paper, we have used a systematic approach that combines limited or spontaneous proteolysis and mass spectrometry to identify structural domains within the p62 subunit of TFIIH. This approach allowed us to identify two subdomains. The first corresponds to residues 1−108 and was observed during spontaneous proteolysis of the sample. The second, encompassing residues 186−240, contains a strictly conserved FWxxΦΦ motif (Φ representing either a tyrosine or phenylalanine). Fluorescence and nuclear magnetic resonance (NMR) experiments indicate that this FWxxΦΦ-containing domain corresponds to a structural unit. This work describes a general approach that can be used as a systematic tool for domain structural studies and may be particularly useful for structural genomics of multidomain proteins.

EXPERIMENTAL PROCEDURES

Purification of Recombinant TFIIH. To purify recombinant TFIIH by FLAG strategy, a virus derived from the pSK277 transfer vector expressing Flag-p34, a p34 subunit with a FLAG peptide (MTKDDDDKH) fused at its N terminus was used together with viruses expressing p44, p52, p62, His-XPB, XPD, cdk7, cyclinH, and MAT1. At 48 h after infection, 10^9 cells were collected and dounced in buffer A (20 mM Tris-HCl at pH 7.8, 10% glycerol, 250 mM NaCl, and 2 mM β-mercaptoethanol). DNA and cell membranes were pulled down by centrifugation at 14000*g* during 30 min. The supernatants were incubated for 1 h at 4 °C with $^1/_{40}$ fraction volume of cobalt chelate affinity resin (Talon, CLONTECH). After a 10-resin volume wash with buffer A containing 5 mM imidazole, the TFIIH complex was eluted with the same buffer containing 250 mM imidazole. The fractions from the Talon column, were incubated for 4 h at 4 °C with protein A-Sepharose beads cross-linked with anti-FLAG antibodies (FLAG-M2, Sigma). After three washes with buffer A containing 0.1% Nonidet P40, proteins were eluted in one bed volume of buffer A containing 1 mg/mL epitope peptide for 12 h (*26*).

Construction, Expression, and Purification of p62 Domains. The cDNAs corresponding to different fragments of the p62 subunit ($p62_{1-108}$, $p62_{1-309}$, $p62_{159-389}$, and $p62_{180-253}$) were amplified by PCR using the full-length version of the p62 gene as a matrix and a GATEWAY compatible version of primers (4Gs, 25-bp *att*B1 site-18 to 25-bp gene-specific sequence as a sense primer and 4Gs, 25-bp *att*B2 site-18 to 25-bp gene-specific sequence as an antisense primer). Then, the cDNAs were subcloned using GATEWAY Cloning Technology (Invitrogen). The BP reaction catalyzed by BP clonase led to pENTR entry vectors, and the LR reaction catalyzed by LR clonase led to pDEST 15 expression vectors with glutathione-*S*-transferase (GST) fusion.

A total of 2 ng of pDEST 15 vectors was used to transform 50 μL of BL21 (DE5) competent cells. After growing on Ampicillin LB agar plate (100 μL/mL) at 37 °C overnight, preculture preparation was used to spread 6 L of LB at 37 °C. Cells were induced at an optical density of 0.8, by addition of 0.4 mM IPTG for 4 h at 37 °C.

After harvesting by centrifugation (5000*g* for 20 min), cells were resuspended in 10 mL of buffer A per liter of culture, lysed by sonication, and then centrifuged again at 50000*g* for 1 h. The supernatants were then applied to 1 mL of GST-affinity beads (Clontech) per liter of cells. After washing with buffer A containing 0.1% Nonidet P40, the proteins were eluted by specific thrombin proteolysis on the GST beads. The $p62_{1-108}$, $p62_{1-309}$, and $p62_{159-389}$ proteins were concentrated on Centriprep Amicon 10 and further purified using a gel-filtration S200 16/60 (Amersham) column at a flow rate of 1 mL/min. For the $p62_{180-253}$ fragment, a Centriprep Amicon 3 and a gel-filtration S75 16/60 (Amersham) column were used.

Limited Proteolysis: Proteolysis of Entire TFIIH. A total of 10 μL of purified recombinant TFIIH (0.1 mg/mL) was incubated with 2 μL of chymotrypsin at various concentrations (0, 0.25×10^{-3}, 0.25×10^{-2}, 0.25×10^{-1}, 0.25, and 2.5 mg/mL) during 1 h at room temperature. The reactions were stopped by addition of 5 μL of 60 mM Tris-HCl at pH 6.8, 2% sodium dodecyl sulfate (SDS), 10% glycerol, and 5% β-mercaptoethanol.

Proteolysis of $p62_{159-389}$. A total of 10 μL of purified $p62_{159-389}$ (1.5 mg/mL) was incubated with 2 μL of chymotrypsin, trypsin, pronase, or subtilisin at various concentrations (0.25×10^{-3}, 0.25×10^{-2}, 0.25×10^{-1}, and 0.25 mg/mL) during 1 h at room temperature and stopped as described for proteolysis of the entire TFIIH.

In-Gel Digestion of $p62_{159-389}$. The hydrolyzed fragments were analyzed by polyacrylamide gel electrophoresis (PAGE, 15% homogeneous gel). The separation was performed during 1 h at 10 mA constant amperage. Staining was carried out with Coomassie Blue. Bands were excised from the gel and cut into small pieces.

In-gel digestion with trypsin was adapted from published methods (*27*) to the robotic digestion system (MassPREP, Micromass, Manchester, U.K.). Gel pieces were destained by washing with 50 μL in 25 mM ammonium hydrogenocarbonate followed by washing with 50 μL of acetonitrile. This preparation was repeated twice. Reduction was achieved by treatment for 1 h with 10 mM DTT at 57 °C. Alkylation was performed in 25 mM iodoacetamide for 45 min at room temperature. The gel pieces were then dehydrated with acetonitrile and dried at 60 °C, prior to addition of modified trypsin (Promega, Madison, WI; 10 μL at 12.5 ng/L in 25 mM ammonium hydrogencarbonate). The digestion was performed overnight at 37 °C. The supernatant was removed,

and 10 μL of 60% acetonitrile/35% water/5% formic acid was added for 1 h to each gel piece for peptide extraction.

Matrix-Assisted Laser Desorption Ionization−Time-of-Flight Mass Spectrometry (MALDI−TOF MS) of p62$_{159-389}$. MALDI MS analyses were carried out on a ULTRAFLEX MALDI TOF/TOF mass spectrometer (Bruker−Daltonik GmbH, Bremen, Germany). This instrument was used at a maximum accelerating potential of 20 kV and was operated in reflector positive mode. A saturated solution of α-cyano-4-hydroxycinnamic acid in acetone was used as a matrix. A first layer of fine matrix crystals was obtained by spreading and fast evaporation of 0.5 μL of the matrix solution. On this fine layer of crystals, a droplet of 0.5 μL of an aqueous 5% formic acid solution was deposited. Afterward, 0.5 μL of a sample solution was added, and a second droplet of 0.2 μL of a matrix-saturated (50% H_2O/50% acetonitrile) was added. The preparation was dried under vacuum. The sample was washed by applying 1 μL of aqueous 5% formic acid solution on the target and then flushed after a few seconds. Internal calibration was performed using trypsin autolysis fragments (monoisotopic masses at $m/z = 842.51$, 1045.564, and 2211.105). Monoisotopic peptide masses were assigned using the X-MASS data system.

N-Terminal Microsequencing. Gel patterns were electroblotted to Pro-Blott PVDF membranes using a BioRad Mini Trans-Blot Electrophoretic and stained with Coomassie Blue R, following the recommendations of the manufacturer for the Pro-Blott membrane (Applied Biosystems, Inc). Degradation bands of interest were excised with a razor blade and N-terminal-sequenced by Edman chemistry.

Fluorescence Spectroscopic Measurements. Fluorescence measurements were carried out on a Perkin−Elmer LB50S fluorimeter, using an optical cuvette of 10-mm light path length with a thermostatically controlled cell holder. Tryptophan emission spectra were obtained at a protein concentration of 1.0 μM using an excitation wavelength of 295 nm, with excitation and emission bandwidths of 10 and 2.5 nm, respectively. Tryptophan emission spectra were recorded between 310 and 480 nm at a scan rate of 100 nm/min. Each spectrum is the average of three emission scans, and data were corrected by subtracting a blank from which the enzyme was omitted.

^{1}H NMR Spectroscopy. All spectra were recorded on a 500 MHz (^{1}H frequency) Bruker DRX 500 spectrometer at 15 °C. p62$_{180-253}$ protein was concentrated to 0.8 mM in 50 mM phosphate and 250 mM NaCl buffer. Two-dimensional (2D) NOESY (*28*) and TOCSY (*29*, *30*) spectra were carried out in phase-sensitive mode using the method of States et al. (*31*) and a WATERGATE sequence (*32*) to suppress the water signal. A spectral width of 8012 Hz was set in both dimensions. Mixing times were 100 and 200 ms for NOESY experiments and 40 and 80 ms for TOCSY experiments. A total of 32 transients of 2K data points were acquired per t_1 increment with 256 t_1 increments. Spectra were processed using NMRPipe (*33*) and analyzed using XEASY (*34*). Typically, the acquisition dimension was multiplied by a Gaussian function and the indirect dimension by a 90°-shifted sine-bell function. Each dimension was zero-filled to the next power of two prior to Fourier transformation.

Protein−Protein Interaction Assay. XPD was expressed by infecting Sf9 cells (2.5×10^7) with the corresponding recombinant baculovirus, and cell extracts were prepared in buffer A containing 250 mM NaCl. Purified GST, GST−p62$_{1-108}$, GST−p62$_{158-389}$, or GST−p62$_{180-253}$ produced in *Escherichia coli* (25 μg) were mixed with 1.0 mL of XPG-clarified extract and incubated with 100 μL GST beads during 2 h at 4 °C. After extensive washing with buffer A containing 0.1% NP40 and 250 mM NaCl, proteins were eluted by addition of 200 μL of the same buffer containing 50 mM glutathion. Proteins were analyzed on a 12% polyacrylamide gel and revealed by Western blotting using XPD (Figure 7) and GST-specific monoclonal antibodies (not shown).

RESULTS

Sequence Alignment of p62. A multiple sequence alignment of p62/Tfb1 strict homologues revealed the existence of conserved regions, mainly located within the first half of the sequence (residues 1−280) and within the extreme C terminus (residues 467−548). In particular, the region that extends from residues 180−245 possesses an invariant FWxxΦΦ motif (Figure 1) which is also observed within the most conserved regions of other unrelated protein families (see the Discussion). This prompted us to investigate whether these regions correspond to structural units and to determine the structural organization of the human p62 subunit of TFIIH.

Identification of a Stable p62 N-Terminal Domain. We first produced recombinant TFIIH in insect cells. Sf9 cells were co-infected by recombinant baculovirus expressing a Flag-p34, His-XPX, and the other subunits of TFIIH. The TFIIH complex was purified by double-tag affinity purification, as described in the Materials and Methods. The full set of TFIIH subunits was detected by Western blotting (Figure 2B). To detect protease-resistant domains of p62 in the context of the entire TFIIH complex, we subjected the recombinant TFIIH to chymotrypsin proteolysis and analyzed the degradation products of the reaction by immunoblotting. An antibody directed against the N terminus of p62 indicated that the full-length p62 was completely digested, yielding fragments of 18 and 30 kDa that were further converted into a 12-kDa product corresponding to the most stable N-terminal domain of p62 (Figure 2C).

To further characterize the protease-resistant regions, we cloned residues 1−309 of p62 in a GST fusion expression vector with a thrombin-sensitive linker. Expressed at high levels, the GST−p62$_{1-309}$ fusion protein migrated as a 60-kDa protein. After thrombin digestion at 4 °C during 5 days, we obtain 30-, 18-, and 12-kDa fragments, as observed for the limited proteolysis experiments performed with recombinant entire TFIIH (compare parts C and D of Figure 2 and see Figure 2F). Amino-terminal sequencing of these proteolysis products indicated that all fragment have the same N terminus. The unique sequence detected (GSMATSS...) corresponds to the two last residues (GS) of the GST tag and the N-terminal residues (MATSS) of the human p62 protein. The 12-kDa fragment was by far the most stable proteolysis product observed. The mass of this fragment, determined by MALDI−TOF, was approximatively 12.1 kDa (Figure 2E). This value is close to 12 115.07 Da, the mass calculated for the p62 polypeptide [*GS*MATSS...QQLLPK-FKR], indicating that protease cleavage occurs between R108 and K109, thus defining the carboxy-terminal extremity of the human p62 N-terminal domain.

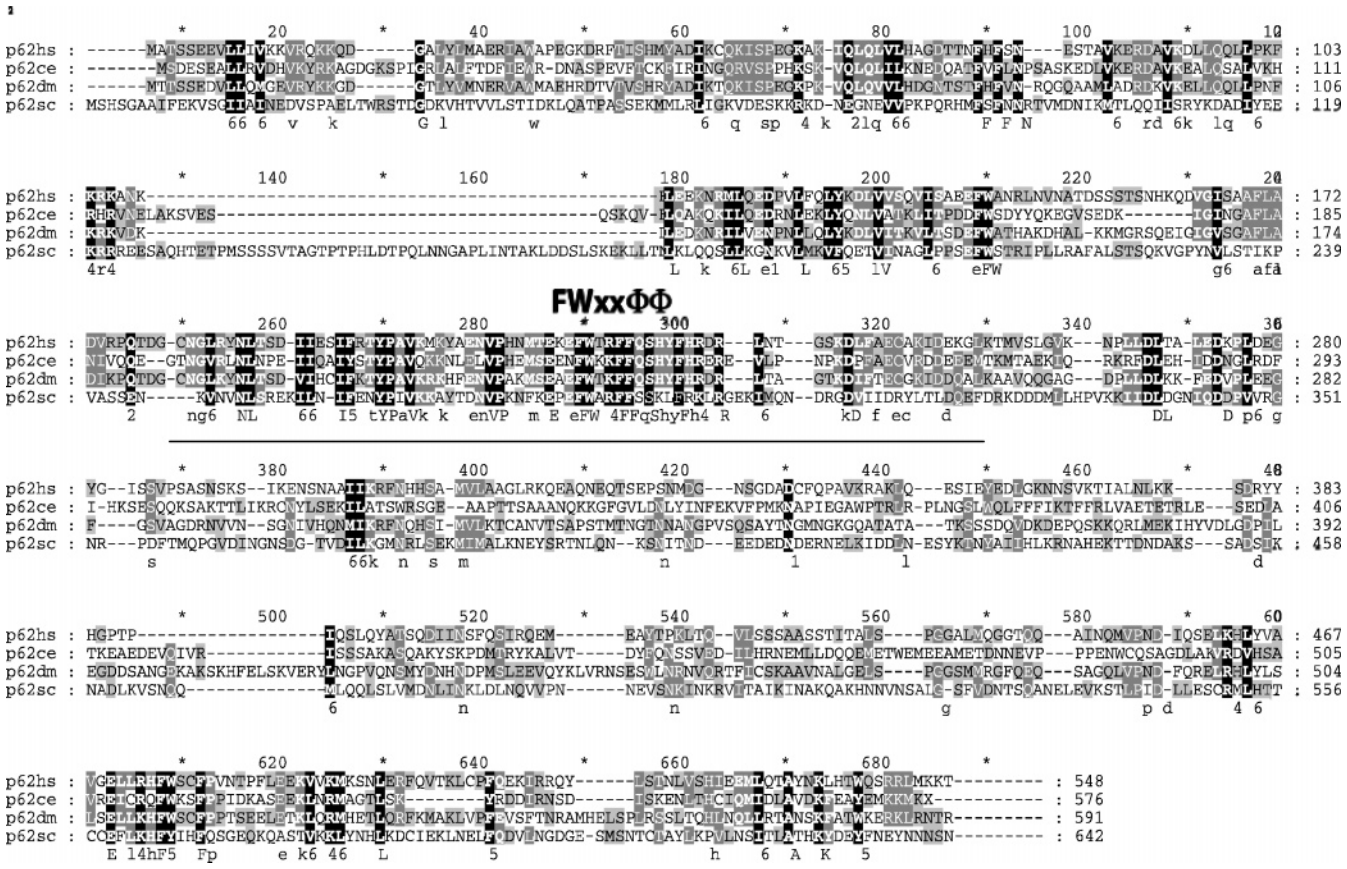

FIGURE 1: Multiple sequence alignment of *Homo sapiens* p62 with its eukaryotic counterparts. The black and gray background indicates conservation levels of 95 and 80%, respectively. The conserved region containing an invariant FWxxΦΦ motif is underlined. The sequences are numbered according to the human protein, and the abbreviations, hs, ce, dm, and sc, denote *Homo sapiens*, *Caenorhabditis elegans*, *Drosophila melanogaster*, and *Saccharomyces cerevisiae* sequences, respectively.

Because $p62_{1-108}$ corresponds to the minimal stable region of the N-terminal part of p62, we examined whether this region corresponded to a structural unit. The domain was cloned, expressed, and purified. Small-angle X-ray-scattering experiments allows us to obtain an estimation of the radius of gyration of $p62_{1-108}$ of 15.1 ± 0.2 Å. This value compares well with the Stokes radius of 17.9 Å calculated for a folded protein of 12 kDa (data not shown), showing that $p62_{1-108}$ behaves as an autonomous structural unit. Its structure was solved by NMR (35).

Limited Proteolysis of $p62_{159-358}$. The median part of the p62 molecule is highly conserved (Figure 1) and contains the FWxxΦΦ motif. To determine whether this region corresponds to a structural unit, cDNA fragments containing the conserved residues from the FWxxΦΦ motif were cloned into an *E. coli* expression vector and the resulting constructs were tested for their ability to produce soluble protein. The construct corresponding to residues 159−389 of the human p62 allowed the production of a soluble protein in fusion with a cleavable amino-terminal His-tag. After metal chelate affinity, thrombin digestion, and gel filtration, the partially purified $p62_{159-389}$ fragment migrated as a 30-kDa protein on SDS−PAGE (lane 1 of Figure 3A). Its molecular weight of 28 064 Da obtained from electrospray ionization−mass spectrometry studies is in agreement with the calculated value of 28 066 Da.

To probe the compactness of $p62_{159-389}$ and identify structured regions, proteolysis experiments were conducted using a set of proteases with different substrate specificities: chymotrypsin, trypsin, pronase, and subtilisin. The patterns of proteolytic digestion obtained after incubation for 1 h with increasing amounts of protease are shown in Figure 3. Digestion with chymotrypsin, trypsin, and subtilisin led to clearly visible bands on the gels, indicating the accumulation of specific protein fragments. Protease-resistant fragments were excised from the gels (bands A−L) and subjected to in-gel trypsination and MALDI−TOF analysis. A list of the peptide mass measurements obtained for representative fragments (bands A−F) are indicated in Table 1.

The protein migrating at 28 kDa (band A of Figure 3) yielded peptides that could be assigned to undigested $p62_{159-389}$ with 11 peptides spread over the entire sequence, yielding a sequence coverage of 60%. The other fragments (bands B−F of Figure 3), with higher apparent electrophoretic mobilites, correspond to chymotrypsin-resistant fragments. All of these fragments produced tryptic peptides derived from residues 186−240 (186−198, 207−218, 224−232, and 233−240). They afforded the same N-terminal peptide ions signals (186−198) but different C termini, indicating progressive degradation of $p62_{159-389}$ from its C terminus (Figure 3). The minimal chymotrypsin-resistant polypeptide (band F of Figure 3) yielded tryptic peptides derived from residues 186−240 only, suggesting that this fragment corresponds to a core fragment. The calculated molecular mass of $p62_{186-240}$ of 6742 Da is in good agreement with the apparent mobility of 7 kDa determined

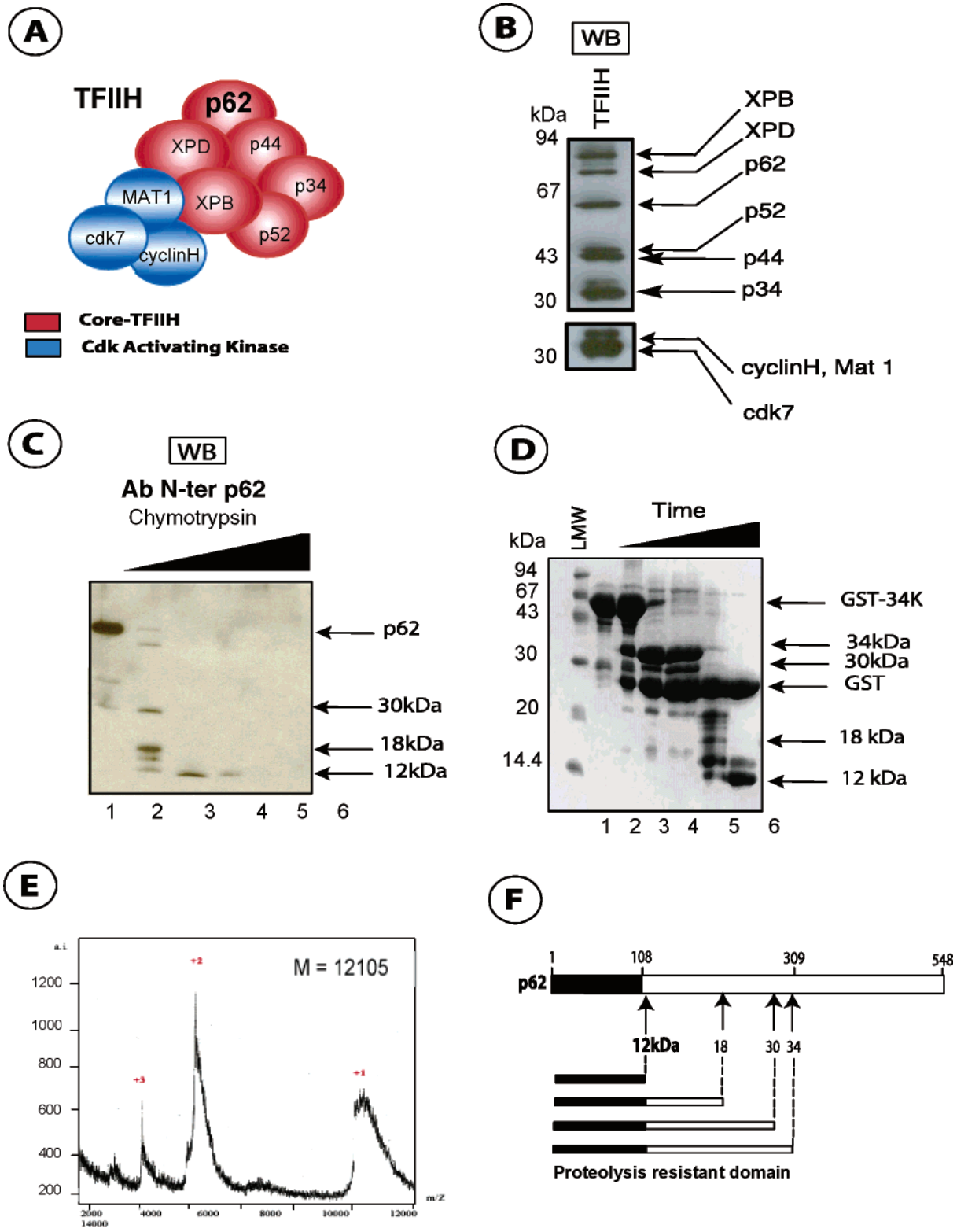

FIGURE 2: Identification of stable p62 N-terminal domains. (A) Schematic drawing of the TFIIH complex showing the core-TFIIH and the CAK subunits. (B) Analysis of purified recombinant TFIIH by Western blotting using a specific antibody for each TFIIH subunit as described in ref *26*. (C) Digestion of recombinant TFIIH with chymotrypsin. TFIIH was submitted to increasing amounts of chymotrypsin (the final concentration of 0, 0.25×10^{-3}, 0.25×10^{-2}, 0.25×10^{-1}, 0.25, and 2.5 mg/mL correspond respectively to lanes 1, 2, 3, 4, 5, and 6), subjected to SDS−PAGE (12.5% acrylamide) and analyzed by Western blotting using monoclonal antibodies directed against the N-terminal end of p62. The 12-, 18-, and 30-kDa protease-resistant fragments that contain the p62 N-terminal epitope accumulate. (D) Identification of a protease-resistant domain of the N-terminal part of p62. $p62_{1-309}$ fused to GST was submitted to thrombin digestion leading to different degradation products (lane 1, $t = 0$; lane 2, $t = 1$ day; lane 3, $t = 2$ days; lane 4, $t = 3$ days; lane 5, $t = 4$ days; and lane 6, $t = 5$ days) and analyzed by Coomassie-stained SDS−PAGE. (E) MALDI−TOF mass spectrum of the 12-kDa degradation product. (F) Schematic representation of proteolysis-resistant regions that contain the 12-kDa degradation product.

by SDS−PAGE for band F. At this point, a more precise definition of the domain borders would have required the identification of peptides, cleaved at one end by chymotrypsin and by trypsin at the other, but despite extensive trials,

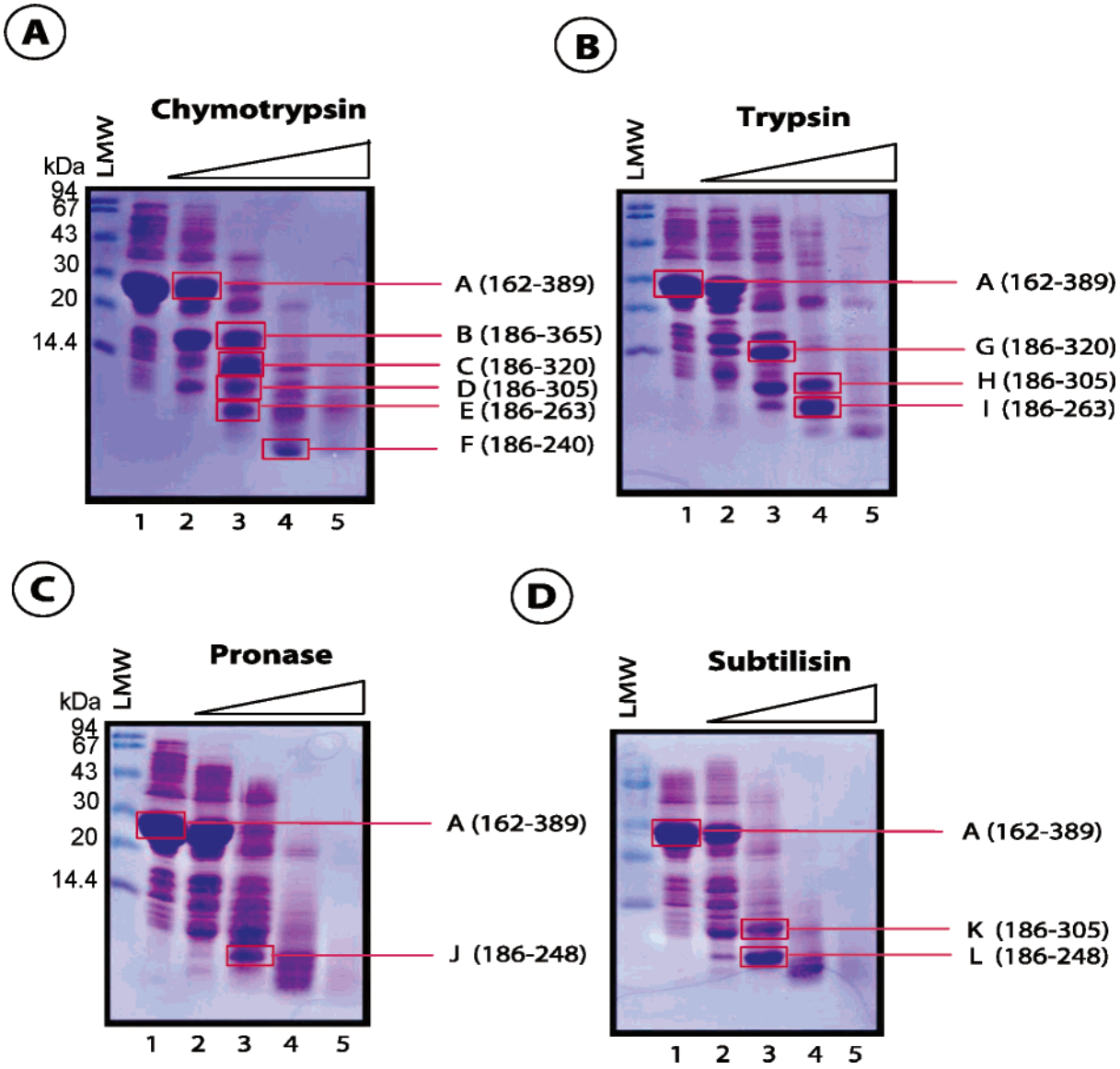

FIGURE 3: Limited proteolysis of the median region of p62. Purified $p62_{159-389}$ protein was digested with increasing amounts of chymotrypsin (A), trypsin (B), pronase (C), or subtilisin (D) (final concentrations of 0, 0.25×10^{-3}, 0.25×10^{-2}, 0.25×10^{-1}, and 0.25 mg/mL for lanes 1, 2, 3, 4, and 5) and subjected to SDS−PAGE (12.5% acrylamide) followed by Coomassie Blue staining. Bands corresponding to digested proteins (A−L) were analyzed by MALDI−TOF MS.

such peptides could not be detected. Limited proteolysis of $p62_{159-389}$ with trypsin and subtilisin yielded results similar to those obtained with chymotrypsin. Highly sensitive protein cleavage sites could be identified within the 160−186 and 240−389 fragments, with the minimal resistant core containing residues 186−263 in the case of trypsin cleavage (band I of Figure 3) and residues 186−248 for both pronase (band J of Figure 3) and subtilisin (band L of Figure 3) proteolysis.

All proteases cleaved in the same regions, yielding minimum fragments of approximately 8 kDa (Figure 4). These minimum protease-resistant fragments contain the FWxxΦΦ motif and correspond to a region of high sequence conservation (Figure 1).

Biophysical Characterization of $p62_{180-253}$. The results presented above pointed to the sequence containing the FWxxΦΦ motif as a probable structural unit. For further characterization, the cDNA coding for residues 180−253, a region slightly larger than the minimum fragment obtained by proteolysis, was inserted into a GST fusion expression vector with a thrombin-sensitive linker. After GST-affinity purification and thrombin digestion, the $p62_{180-253}$ fragment was recovered and further purified by gel-filtration chromatography. The $p62_{180-253}$ fragment is stable (no degradation) and elutes as a single peak (Figure 5A).

The $p62_{180-253}$ sequence possesses a single tryptophan residue that is part of the FWxxΦΦ motif. The intrinsic fluorescence emission spectrum of $p62_{180-253}$ therefore depends mainly on the microenvironment of the tryptophan and can be used to indicate the extent to which this residue is exposed to the solvent. The emission spectrum of $p62_{180-253}$ was recorded in native conditions with an excitation wavelength of 280 nm. The spectrum peaked at 333 nm, with a full width of 64 nm at half-maximum, indicating a partially buried environment (Figure 5B). When, the emission spectrum of $p62_{180-253}$ was measured at various concentrations of guanidinium chloride, a bathochromic fluorescence shift of the peak was observed (not shown). In 8 M guanidinium chloride, the maximum of intrinsic

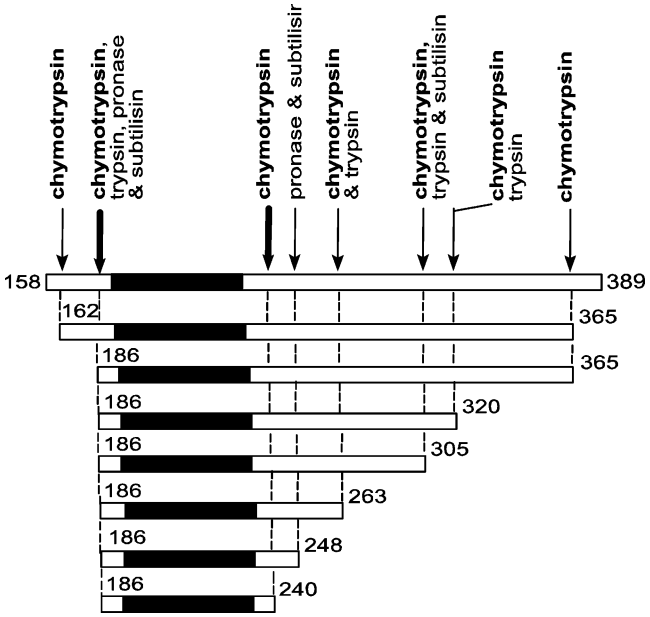

FIGURE 4: Schematic representation of identified sites of proteolysis of the $p62_{158-389}$ region. The most stable domain is indicated in black. Protease cleavage sites are shown by arrows, with chymotrypsin cleavage sites in bold.

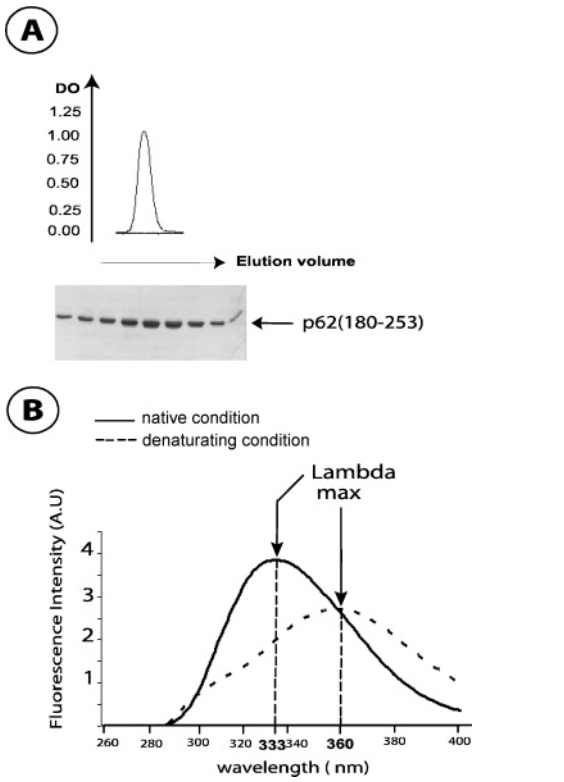

FIGURE 5: Fluorescence characterization of the proteolytically stable $p62_{180-253}$ fragment. (A) Gel-filtration profile (elution volume of 88 mL) and SDS−PAGE analysis of the $p62_{180-253}$ domain. (B) Fluorescence of the tryptophan residue under native conditions (—) and denaturating conditions (8 M guanidinium chloride, - - -). The tryptophan emission maximum was shifted from 333 to 360 nm upon denaturation.

fluorescence was found at 357 nm, a value typically observed for exposed tryptophans (*36*), suggesting that the protein has unfolded. Under native conditions, the indole chromophore from the FWxxΦΦ motif is buried, suggesting the presence of a structured core in $p62_{180-253}$.

Table 1: Fragments Found by MALDI−TOF Analysis of Bands A−F[a]

	gel bands[a]					
fragment	A	B	C	D	E	F
162−185	2560.26					
186−198	1570.87	1570.81	1570.82	1570.84	1570.81	1570.83
207−218	1448.65[b]	1432.67	1432.69	1448.67[b]	1432.65	1448.68[b]
224−232	1268.62	1268.61	1268.63	1268.63	1268.60	1268.64
233−240	890.47	890.48	890.46	890.48	890.47	890.50
241−248	953.45	953.42	953.44	953.46	953.47	
256−263	834.47		834.49	834.5	834.45	
264−293	3130.57	3130.60	3130.66	3130.51		
297−305	959.61	959.54	959.59	959.54		
307−320	1523.77	1523.80	1523.81			
354−365	1423.69	1423.67				

[a] Observed masses of the tryptic fragments from protease-resistant bands A−F (see Figure 3). The numbers in the first column indicate the residue numbers of the fragments. [b] Oxidized peptide. The tryptic fragments 186−198, 207−218, 224−232, and 233−240 are present in all protease-resistant bands and are indicated in bold.

To further characterize the $p62_{180-253}$ module, we carried out NMR experiments. In the ^{1}H NMR spectrum of $p62_{180-253}$, upfield-shifted resonances of methyl groups and a small number of downfield-shifted HN resonances (from amide protons of backbone peptide moieties) indicated the presence of a defined structure and a hydrophobic core. Two-dimensional NOESY spectra revealed a pattern of NOEs involving side-chain atoms of aromatic and aliphatic residues (Figure 6). In particular, the ring protons of the lone tryptophan residue (W221) give NOEs to ring protons of two other aromatic residues and to the methyl protons of an aliphatic side chain. One of the two other aromatic residues gives NOEs to the same aliphatic side chain, while the other gives NOEs to a second such side chain. This set of observations reflects a defined hydrophobic core for $p62_{180-253}$. A number of downfield-shifted Hα resonances suggest a small degree of extended β structure.

The NOESY spectra show clear signs of heterogeneity in line width, with only a small number of HN resonances giving rise to well-defined sets of NOEs. Moreover, the TOCSY spectra were particularly poor (data not shown), hindering attempts at sequential assignment of the spectra of $p62_{180-253}$. The NOEs described above could not, therefore, be interpreted in residue-specific terms, because the aromatic and aliphatic side chains involved could not be unambiguously identified (with the exception of W221). The heterogeneity in line widths and poor quality of TOCSY spectra are uncharacteristic of monomeric, globular proteins of this molecular weight. Oligomerization could be excluded on the basis of NMR diffusion measurements (data not shown), which suggested a monomeric species in solution. It seems probable, therefore, that conformational exchange processes are responsible for degradation of the quality of the spectra.

$p62_{180-253}$ Module Interacts with XPD in Vitro. Having shown that residues 180−240 from p62 form a protease-resistant fragment, which buried tryptophan, we wondered wether this identified entity is capable of interactions directly with other TFIIH subunits. To address this point, GST-pull-down experiments were performed using the GST-tagged p62 fragments and showed that the p62 median region interacts

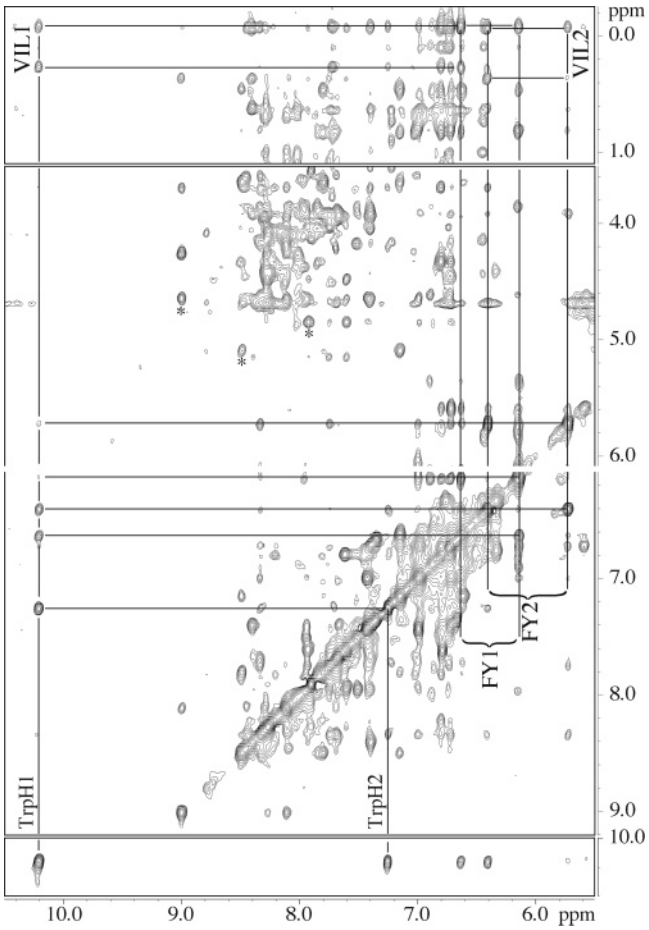

FIGURE 6: Two-dimensional NOESY spectrum of the median stable domain of p62. Downfield portion of the NOESY spectrum of 0.8 mM $p62_{180-253}$ at 15 °C and pH 6.8, recorded with a mixing time of 200 ms at 500 MHz (^{1}H frequency). NOE connectivities described in the text are highlighted. The tryptophan residue contacts aromatic (FY1 and FY2) and aliphatic (VIL1) side chains and additional contacts between this set of residues and additional side chains (e.g, VIL2) define a hydrophobic core. A small number of downfield-shifted Hα resonances suggest some degree of extended structure. For these resonances, NOEs observed at the positions of TOCSY cross peak are indicated by an asterisk.

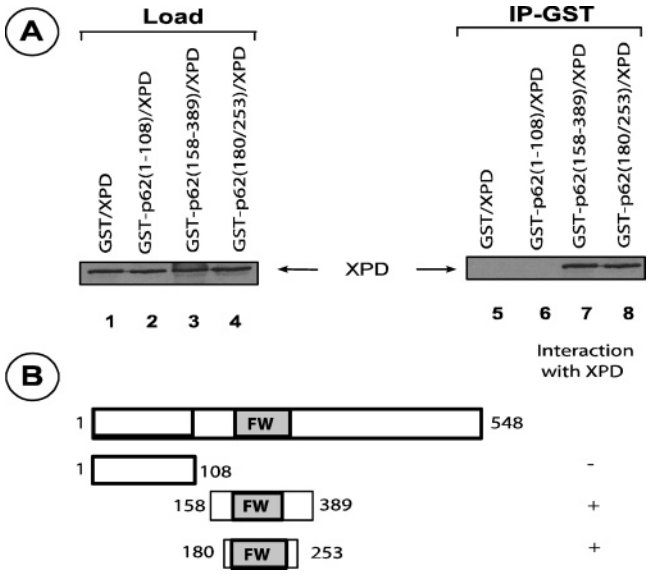

FIGURE 7: Interaction between $p62_{180-253}$ and XPD. (A) Clarified Sf9 cell lysates containing the recombinant XPD TFIIH subunit were mixed with GST, GST−$p62_{1-108}$, GST−$p62_{158-389}$, or GST−$p62_{180-253}$ and subjected to a GST-pull-down assay. The proteins before (lanes 1−4) and after (lanes 5−8) GST-affinity chromatography were resolved by SDS−PAGE (12.5% acrylamide) and revealed by Western blotting using an XPD-specific monoclonal antibody. (B) Schematic drawing of p62. The protease-resistant module (residues 186 and 240), which contains the FWxxΦΦ motif, is represented in gray, and the constructs that bind XPD are indicated.

with XPD. Crude extracts from Sf9 insect cells infected with baculoviruses expressing the XPD protein were mixed with GST, GST−$p62_{1-108}$, GST−$p62_{159-389}$, and GST−$p62_{180-253}$ and further incubated with Glutathione-Sepharose beads. After extensive washing with a buffer containing 250 mM NaCl, coprecipitation of the TFIIH subunits with the Glutathione-Sepharose beads was assessed by immunoblotting with specific antibodies. This set of experiments showed that XPD can be retained on the GST beads in the presence of $p62_{159-359}$ or $p62_{180-240}$ fused to GST (lanes 7 and 8 of Figure 7A). No unspecific retention of XPD on the beads was observed in the control experiment performed with GST alone or with $p62_{1-108}$ (lanes 5 and 6 of Figure 7A). This result clearly demonstrates that the isolated $p62_{180-253}$ module can physically interact with the other TFIIH subunit XPD (Figure 7B).

DISCUSSION

Limited proteolysis has been widely used to delineate the structural organization of multidomain proteins. Recently, approaches combining limited proteolysis with electrospray (ES) (*37*) or MALDI−TOF (*38*) mass spectrometry have shown a high potential for the topological analysis of monomeric proteins as well as multimeric protein complexes. Two strategies are usually considered. The first is based on the identification of peptides released from limited proteolysis by LC−MS, assuming that the peptides formed reflect either sites of accessibility on the molecular surface or unstructured domains. This approach is generally applied to probe the conformation of protein−protein, protein−DNA, or protein−ligand complexes (*39*). The second strategy aims at identifying the proteolysis-resistant core by ES or MALDI, assuming that the undigested protein corresponds to a compact-folded domain. This latter approach was used for the definition of structural domains within the conserved N-terminal part of the p62 subunit from the transcription/DNA repair factor TFIIH.

Here, we have identified two structural domains within the conserved N-terminal half of the molecule (Figure 8A). The first encompasses residues 1−108 and corresponds to a folded compact domain. Its solution structure revealed that it adopts a PH-domain fold, and functional studies showed that this domain of the p62 subunit of TFIIH directly binds XPG, a critical interaction for NER (*35*). The second domain, encompassing residues 180−253, is more sensitive to degradation than the N-terminal PH domain and contains an invariant FWxxΦΦ motif.

Interestingly, we and others (*40*, *41*) have identified within the p62 sequence encompassing residues 180−253 an invariant FWxxΦΦ motif (Φ represents either a tyrosine or phenylalanine residue). A search for related sequences identified six proteins containing the consensus motif: four of these were hypothetical proteins and two were docu-

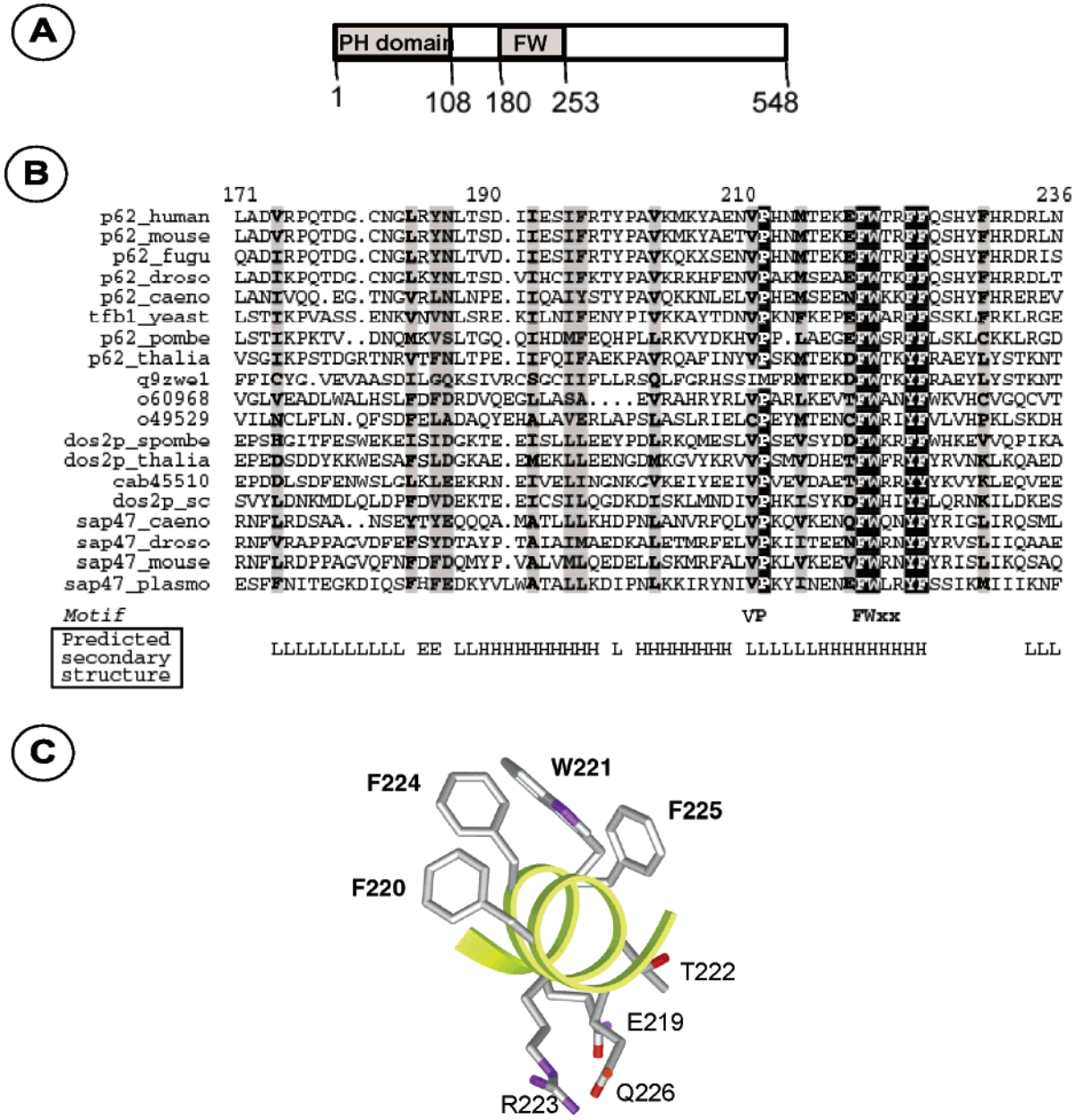

FIGURE 8: Conserved domains of p62. (A) Schematic representation of identified p62 domains. PH denotes the Pleckstrin homology domain. FW denotes the domain containing the FWxxΦΦ motif. (B) Alignment of FWxxΦΦ containing regions of human and eukaryotic counterparts of p62 with other FWxxΦΦ-containing proteins. Residues conserved in more than 90% of the sequences are shown in white on a black background. Residues conserved to a lesser extent are shown in black on a gray background. Secondary structure was predicted using PHD (http://dodo.cpmc.columbia.edu/predictprotein/), where H indicates α helices, E indicates *β* sheets, and L indicates loops. (C) Model of the predicted helix containing the FWxxΦΦ motif constructed using a canonical α helix as a template. Residues from human p62 were constructed with the O package (*46*). The figure was produced with SETOR (*47*). The helix representation shows the amphipathic distribution of the conserved residues.

mented, sap47 and dos2p (Figure 8B). The sap47 protein localizes to synaptic terminals in *Drosophila* and is related to the neuronal transmitter release apparatus. The dos2p protein, encoded by the *Saccharomyces cerevisiae* gene DOS2, is involved in single-copy DNA replication and ubiquitination (*42*). The conservation of the FWxxΦΦ motif in three distincts protein families, sap47, dos2p, and p62/Tfb1, suggests that they might share a common structural unit. Mean secondary structure prediction on these three protein families are in agreement with the sequence data and suggest that the conserved region contains three α helices. The last α helix contains the invariant FWxxΦΦ motif and is preceded by a loop containing the almost invariant VP dipeptide motif (Figure 8B). A model representing a canonical α helix with the FWxxΦΦ motif of human p62 shows that the polar residues (E219, T222, R223, and Q226) lie on one face of the helix, while the aromatic residues (F220, W221, F224, and F225) are clustered on the opposite face (Figure 8C). Such amphipathic helices with aromatic clusters have been found, for example, in all cyclins (*43*), stabilizing their fold, and at the protein−protein interface in the complex between the transactivation domain of p53 and the oncoprotein MDM2 (*44*).

Using a combination of limited proteolysis, fluorescence, and NMR analysis, we have shown that $p62_{180-253}$, which contains the FWxxΦΦ motif, corresponds to a structural unit

with a defined core. However, the high protease sensitivity and the characteristics of the NMR spectra suggest that the isolated $p62_{180-253}$ module is conformationally mobile and might be stabilized upon interaction with a partner molecule. The finding that $p62_{180-253}$ is capable of binding XPD, another TFIIH subunit, supports this hypothesis and suggests that this module corresponds to a protein−protein interaction unit. More importantly, this observation provides initial data regarding the function of this module and its associated FWxxΦΦ motif. Other TFIIH subunits, p44 and MAT1, physically bind the XPD helicase, and the p44/XPD interaction stimulates the XPD activity (*14*, *45*). The p62 FWxxΦΦ containing module could play a similar role in the architecture of TFIIH and/or in the regulation of the XPD activity. Additional data are needed to determine the importance of this module for the stabilization of the complex and its influence on XPD unwinding activity and on the TFIIH transcription/DNA repair functions.

The p62/Tfb1 subunit is essential for both the transcription and DNA repair activities of TFIIH, and its N-terminal part was shown to be important for the recruitment of partners such as the transcription activators VP16 or p53 (*19*) or the nuclear hormone receptor ERα (*20*) within large transcription initiation and/or NER complexes. The experimental definition of two structural domains within the conserved N-terminal half of the p62/Tfb1 subunit presented here provides a framework for the investigation of the functional role of this essential subunit and in particular of its FWxxΦΦ-containing module.

ACKNOWLEDGMENT

We are grateful to J. M. Egly and J. C. Thierry for support and discussions. We thank the IGBMC common facilities for DNA sequencing and oligonucleotide synthesis. We acknowledge I. Kolb and J. L. Weickert for protein production in insect cells, G. Travé for help concerning the use of the spectrofluorimeter, and G. Fagart for advice regarding molecular modeling. We thank Holger Greschik for critical reading of the manuscript.

REFERENCES

1. Feaver, W. J., Svejstrup, J. Q., Bardwell, L., Bardwell, A. J., Buratowski, S., Gulyas, K. D., Donahue, T. F., Friedberg, E. C., and Kornberg, R. D. (1993) Dual roles of a multiprotein complex from *S. cerevisiae* in transcription and DNA repair, *Cell 75*, 1379−1387.
2. Schaeffer, L., Roy, R., Humbert, S., Moncollin, V., Vermeulen, W., Hoeijmakers, J. H., Chambon, P., and Egly, J. M. (1993) DNA repair helicase: A component of BTF2 (TFIIH) basic transcription factor, *Science 260*, 58−63.
3. Guzder, S. N., Sung, P., Bailly, V., Prakash, L., and Prakash, S. (1994) RAD25 is a DNA helicase required for DNA repair and RNA polymerase II transcription, *Nature 369*, 578−581.
4. Drapkin, R., Reardon, J. T., Ansari, A., Huang, J. C., Zawel, L., Ahn, K., Sancar, A., and Reinberg, D. (1994) Dual role of TFIIH in DNA excision repair and in transcription by RNA polymerase II, *Nature 368*, 769−772.
5. Berneburg, M., and Lehmann, A. R. (2001) Xeroderma pigmentosum and related disorders: Defects in DNA repair and transcription, *Adv. Genet. 43*, 71−102.
6. Vermeulen, W., Rademakers, S., Jaspers, N. G., Appeldoorn, E., Raams, A., Klein, B., Kleijer, W. J., Hansen, L. K., and Hoeijmakers, J. H. (2001) A temperature-sensitive disorder in basal transcription and DNA repair in humans, *Nat. Genet. 27*, 299−303.
7. Schultz, P., Fribourg, S., Poterszman, A., Mallouh, V., Moras, D., and Egly, J. M. (2000) Molecular structure of human TFIIH, *Cell 102*, 599−607.
8. Chang, W. H., and Kornberg, R. D. (2000) Electron crystal structure of the transcription factor and DNA repair complex, core TFIIH, *Cell 102*, 609−613.
9. Tirode, F., Busso, D., Coin, F., and Egly, J. M. (1999) Reconstitution of the transcription factor TFIIH: Assignment of functions for the three enzymatic subunits, XPB, XPD, and cdk7, *Mol. Cell 3*, 87−95.
10. Kim, T. K., Ebright, R. H., and Reinberg, D. (2000) Mechanism of ATP-dependent promoter melting by transcription factor IIH, *Science 288*, 1418−1422.
11. Dvir, A., Conaway, J. W., and Conaway, R. C. (2001) Mechanism of transcription initiation and promoter escape by RNA polymerase II, *Curr. Opin. Genet. Dev. 11*, 209−214.
12. Larochelle, S., Chen, J., Knights, R., Pandur, J., Morcillo, P., Erdjument-Bromage, H., Tempst, P., Suter, B., and Fisher, R. P. (2001) T-loop phosphorylation stabilizes the CDK7−cyclin H-MAT1 complex in vivo and regulates its CTD kinase activity, *EMBO J. 20*, 3749−3759.
13. Tremeau-Bravard, A., Perez, C., and Egly, J. M. (2001) A role of the C-terminal part of p44 in the promoter escape activity of transcription factor IIH, *J. Biol. Chem. 276*, 27693−27697.
14. Coin, F., Marinoni, J. C., Rodolfo, C., Fribourg, S., Pedrini, A. M., and Egly, J. M. (1998) Mutations in the XPD helicase gene result in XP and TTD phenotypes, preventing interaction between XPD and the p44 subunit of TFIIH, *Nat. Genet. 20*, 184−188.
15. Jawhari, A., Laine, J. P., Dubaele, S., Lamour, V., Poterszman, A., Coin, F., Moras, D., and Egly, J. M. (2002) p52 mediates XPB function within the transcription/repair factor TFIIH, *J. Biol. Chem. 277*, 31761−31767.
16. Svejstrup, J. Q., Vichi, P., and Egly, J. M. (1996) The multiple roles of transcription/repair factor TFIIH, *Trends Biochem. Sci. 21*, 346−350.
17. Uetz, P., Giot, L., Cagney, G., Mansfield, T. A., Judson, R. S., Knight, J. R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P., Qureshi-Emili, A., Li, Y., Godwin, B., Conover, D., Kalbfleisch, T., Vijayadamodar, G., Yang, M., Johnston, M., Fields, S., and Rothberg, J. M. (2000) A comprehensive analysis of protein−protein interactions in *Saccharomyces cerevisiae*, *Nature 403*, 623−627.
18. Leveillard, T., Andera, L., Bissonnette, N., Schaeffer, L., Bracco, L., Egly, J. M., and Wasylyk, B. (1996) Functional interactions between p53 and the TFIIH complex are affected by tumour-associated mutations, *EMBO J. 15*, 1615−1624.
19. Xiao, H., Pearson, A., Coulombe, B., Truant, R., Zhang, S., Regier, J. L., Triezenberg, S. J., Reinberg, D., Flores, O., Ingles, C. J., et al. (1994) Binding of basal transcription factor TFIIH to the acidic activation domains of VP16 and p53, *Mol. Cell. Biol. 14*, 7013−7024.
20. Chen, D., Riedl, T., Washbrook, E., Pace, P. E., Coombes, R. C., Egly, J. M., and Ali, S. (2000) Activation of estrogen receptor α by S118 phosphorylation involves a ligand-dependent interaction with TFIIH and participation of CDK7, *Mol. Cell 6*, 127−137.
21. 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 damaged DNA, *J. Biol. Chem. 275*, 9870−9875.
22. Iyer, N., Reagan, M. S., Wu, K. J., Canagarajah, B., and Friedberg, E. C. (1996) Interactions involving the human RNA polymerase II transcription/nucleotide excision repair complex TFIIH, the nucleotide excision repair protein XPG, and Cockayne syndrome group B (CSB) protein, *Biochemistry 35*, 2157−2167.
23. Winkler, G. S., Sugasawa, K., Eker, A. P., de Laat, W. L., and Hoeijmakers, J. H. (2001) Novel functional interactions between nucleotide excision DNA repair proteins influencing the enzymatic activities of TFIIH, XPG, and ERCC1-XPF, *Biochemistry 40*, 160−165.
24. Hubbard, S. J. (1998) The structural aspects of limited proteolysis of native proteins, *Biochim. Biophys. Acta 1382*, 191−206.
25. Carrey, E. A. (1989) *Protein Structure: A Practical Approach* (Creighton, T. E., Ed.) pp 117−144, IRL Press, New York.
26. Jawhari, A., Uhring, M., Crucifix, C., Fribourg, S., Schultz, P., Poterszman, A., Egly, J. M., and Moras, D. (2002) Expression of FLAG fusion proteins in insect cells: Application to the multisubunit transcription/DNA repair factor TFIIH, *Protein Expression Purif. 24*, 513−523.

27. Jeno, P., Mini, T., Moes, S., Hintermann, E., and Horst, M. (1995) Internal sequences from proteins digested in polyacrylamide gels, *Anal. Biochem. 224*, 75−82.
28. Jeener, J. M. B. H., Bachmann, P., and Ernst, R. R. (1979) Investigation of exchange processes by two-dimensional NMR spectroscopy, *J. Chem. Phys. 71*, 4546−4553.
29. Bax, A., and Davis, D. G. (1985) MLEV-17 based two-dimensional homonuclear magnetization transfer spectroscopy, *J. Magn. Reson. 65*, 355−360.
30. Braunschweiler, L., and Ernst, R. R. (1983) Coherence transfer by isotopic mixing: Application to proton coorelation spectroscopy, *J. Magn. Reson. 53*, 521−528.
31. States, D. J., Haberkorn, R. A., and Ruben, D. (1982) A two-dimensional nuclear Overhauser experiment with pure absorption phase in four quadrants, *J. Magn. Reson. 48*, 288−292.
32. Piotto, M., Saudek, V., and Sklenar, V. (1992) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions, *J. Biomol. NMR 2*, 661−665.
33. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR 6*, 277−293.
34. Bartels, C., Xia, T. H., Billeter, M., Güntert, P., and Wüthrich, K. (1995) The program XEASY for computer supported NMR spectral analysis of biological macromolecules, *J. Biomol. NMR 5*, 1−10.
35. Gervais, V., Lamour, V., Jawhari, A., Frindel, F., Wasielewski, E., Dubaele, S., Egly, J. M., Thierry, J. C., Kieffer, B., and Poterszman, A. (2004) TFIIH contains a PH domain involved in DNA nucleotide excision repair, *Nat. Struct. Mol. Biol. 11*, 616−622.
36. Creighton, T. E. (1993) *Proteins: Structures and Molecular Properties*, New York.
37. Bantscheff, M., Weiss, V., and Glocker, M. O. (1999) Identification of linker regions and domain borders of the transcription activator protein NtrC from *Escherichia coli* by limited proteolysis, in-gel digestion, and mass spectrometry, *Biochemistry 38*, 11012−11020.
38. D'Ambrosio, C., Talamo, F., Vitale, R. M., Amodeo, P., Tell, G., Ferrara, L., and Scaloni, A. (2003) Probing the dimeric structure of porcine aminoacylase 1 by mass spectrometric and modeling procedures, *Biochemistry 42*, 4430−4443.
39. Spolaore, B., Bermejo, R., Zambonin, M., and Fontana, A. (2001) Protein interactions leading to conformational changes monitored by limited proteolysis: Apo form and fragments of horse cytochrome *c*, *Biochemistry 40*, 9460−9468.
40. Lamour, V. (1999) Ph.D. Thesis, Université Louis Pasteur.
41. Doerks, T., Huber, S., Buchner, E., and Bork, P. (2002) BSD: A novel domain in transcription factors and synapse-associated proteins, *Trends Biochem. Sci. 27*, 168−170.
42. Singer, J. D., Manning, B. M., and Formosa, T. (1996) Coordinating DNA replication to produce one copy of the genome requires genes that act in ubiquitin metabolism, *Mol. Cell. Biol. 16*, 1356−1366.
43. Noble, M. E., Endicott, J. A., Brown, N. R., and Johnson, L. N. (1997) The cyclin box fold: Protein recognition in cell-cycle and transcription control, *Trends Biochem. Sci. 22*, 482−487.
44. Kussie, P. H., Gorina, S., Marechal, V., Elenbaas, B., Moreau, J., Levine, A. J., and Pavletich, N. P. (1996) Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain, *Science 274*, 948−953.
45. Busso, D., Keriel, A., Sandrock, B., Poterszman, A., Gileadi, O., and Egly, J. M. (2000) Distinct regions of MAT1 regulate cdk7 kinase and TFIIH transcription activities, *J. Biol. Chem. 275*, 22815−22823.
46. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models, *Acta Crystallogr., Sect A 47* (Part 2), 110−119.
47. Evans, S. V. (1993) SETOR: Hardware-lighted three-dimensional solid model representations of macromolecules, *J. Mol. Graphics 11*, 134−138.
48. Giglia-Mari, G., Coin, F., Ranish, J. A., Hoogstraten, D., Theil, A., Wijgers, N., Jaspers, N. G., Raams, A., Argentini, M., van der Spek, P. J., Botta, E., Stefanini, M., Egly, J. M., Aebersold, R., Hoeijmakers, J. H., and Vermeulen, W. (2004) A new, tenth subunit of TFIIH is responsible for the DNA repair syndrome trichothiodystrophy group A, *Nat. Genet. 36*, 714−719.

BI048884C