Biochemistry

 Volume 37, Number 38 September 22, 1998

Articles

Functional and Structural Properties of the Mitochondrial Outer Membrane Receptor Tom20[†]

Enrico Schleiff*,[‡] and Joanne L. Turnbull*,[§]

Department of Biochemistry, McIntyre Medical Science Building, McGill University, Montreal, Canada H3G 1Y6, and Department of Chemistry & Biochemistry, Concordia University, Montreal, Canada H3G 1M8

Received April 2, 1998; Revised Manuscript Received July 14, 1998

ABSTRACT: Tom20 is an outer mitochondrial membrane protein that functions as a component of the import receptor complex for cytoplasmically synthesized mitochondrial precursor proteins. The human homologue, hTom20, consists of an N-terminal membrane anchor region predicted between aa5−25 and a soluble cytosolic domain from aa30 to 145. To analyze the properties of hTom20, we have expressed several truncations of the cytosolic domain as fusion proteins with glutathione *S*-transferase. Our studies reveal that the cytosolic region of hTom20 is a monomeric protein in solution containing two domains which are involved in different functions of the receptor. The N-terminal region is involved in membrane binding (aa30−60) and recognition of the cleavable matrix targeting signals (aa50−90). In addition, we have demonstrated that the receptor recognizes the α-helical state of the matrix targeting signal. The dissociation constant for this interaction in the presence of a detergent which induces this secondary structure is 0.6 μM, one-fifth the value in the absence of detergent. In aqueous solution, the region between aa30 and 60 is loosely folded and stabilized against proteolytic cleavage by interaction with detergents or a matrix targeting signal. Our work further shows that the remainder of the cytosolic domain of hTom20, aa60−145, is a compactly folded globular domain containing a region (aa90−145) that is critical for the recognition of proteins bearing internal signal sequences such as the uncoupling protein and porin.

The majority of mitochondrial proteins are nuclear encoded and are synthesized as preproteins in the cytosol prior to translocation into the mitochondria. Import occurs through a translocation apparatus, which is composed of proteins located in both the inner and outer membranes of the organelle (*1*−*3*). The composition of this complex in *Saccharomyces cerevisiae* and *Neurospora crassa* is reasonably well understood, although not all proteins have been identified to date (*4*−*6*). The translocation complex found in the outer membrane can be divided into a receptor complex made up of the heterodimers Tom20/Tom22 and Tom37/Tom70 (translocase on the mitochondrial outer membrane),[1] which is responsible for the specific recognition of prepro-

[†] This work was supported by operating grants from the Medical Research Council and the National Cancer Institute of Canada to Dr. G. C. Shore and from the Natural Sciences and Engineering Research Council to Dr. J. Turnbull. E. Schleiff is a recipient of a fellowship from the HSP/III program from the German Academic Association Service (DAAD).

* Corresponding authors. E-mail: schleiff@med.mcgill.ca and jturn@vax2.concordia.ca.

[‡] McGill University.

[§] Concordia University.

[1] Abbreviations: GST, glutathione *S*-transferase; MTS, matrix targeting signal; NTTS, N-terminal targeting signal; pOCT, preornithine carbamyl transferase; pODHFR, targeting signal of pOCT fused to dihydrofolate reductase; Tom, translocase on the outer mitochondrial membrane; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; aa, amino acid.

S0006-2960(98)00745-4 CCC: $15.00
Published on Web 08/29/1998

teins (*7*,*8*), and the general insertion pore (GIP) comprised of Tom40, Tom7, Tom6, and Tom5 (*9*). Tom72, a homologue of Tom70, is also believed to be involved in the receptor complex as a minor component (*10*, *11*).

In vivo and in vitro studies have indicated that Tom20 recognizes two classes of proteins: one class bearing N-terminal targeting signals (NTTS), such as a cleavable N-terminal matrix targeting signal (MTS) as well as a noncleavable N-terminal signal anchor sequence, and another class comprised of proteins containing internal targeting signal sequences (*12*−*16*). It has been proposed that N-terminal targeting signals bind to the Tom20 via electrostatic (*17*−*19*) or hydrophobic (*16*) interactions. Moreover, it is believed that the positively charged MTSs interact electrostatically with the "acidic bristle" (*17*, *18*), an NTTS binding area formed by Tom22 and the C-terminus of Tom20. The MTS forms an amphipathic α-helix under hydrophobic conditions (*20*, *21*), which has led to the speculation that the receptor proteins interact with the NTTS in its α-helical state (*22*). The Tom37−Tom70 complex functions as a receptor for proteins carrying internal targeting signal sequences (*23*) and interacts with chaperones which are involved in the transport of NTTS-containing preproteins to the membrane-associated receptor complex (*24*). In addition, Tom70 binds to the mature part of the protein (*16*) and also to the N-terminal MTS but in an ATP-dependent fashion (*15*).

There have been very few reports assigning functions to specific regions of primary sequence in Tom20. One such study by Haucke et al. (*25*) indicates that the B part of the tetratricopeptide repeat motif from yTom20 (aa89−101) interacts with yTom70 in a yeast two-hybrid system. Work by Schleiff et al. (*14*) shows that an α-helix spanning aa105−113 of hTom20 is involved in the recognition of precursor proteins containing internal targeting sequences. Iwahashi et al. (*26*) have proposed that the region aa25−90 in rTom20 might form a domain involved in the import of MTS-containing proteins, although in the absence of any binding studies, a distinction could not be made between a role in the arrangement of the receptor complex versus direct precursor recognition. Until now there have been almost no in vitro studies addressing the structure of the individual components of the import machinery.

In the present study we have characterized the full-length cytosolic domain of hTom20 (aa30−145) and selected truncated versions of the receptor. We provide the first experimental evidence that the receptor contains two precursor binding sites, each associated with distinct regions of the polypeptide chain and each possessing distinct structural features. The N-terminal region from aa30−60 is highly flexible and interacts with the membrane surface, followed by the region from about aa50−90 which recognizes proteins, such as preornithine carbamyl transferase (pOCT), containing a positively charged NTTS. Furthermore, we provide the first direct evidence that the receptor interacts with the α-helical state of the MTS. The dissociation constant for the receptor−preprotein complex at about 0.6 μM is one-fifth the value in the absence of a nonionic detergent, implying that the MTS can preferentially adopt an α-helical structure to facilitate its binding to the receptor. The C-terminal region commencing at amino acid 60 forms a globular, more compact domain of which a portion, aa90−145, recognizes proteins such as the uncoupling protein or the voltage-dependent anion channel containing an internal targeting signal. Moreover, the C-terminal five residues (aa141−145) are buried within the protein core and hence are unlikely to participate in the binding of NTTS-containing proteins, as previously speculated (*27*).

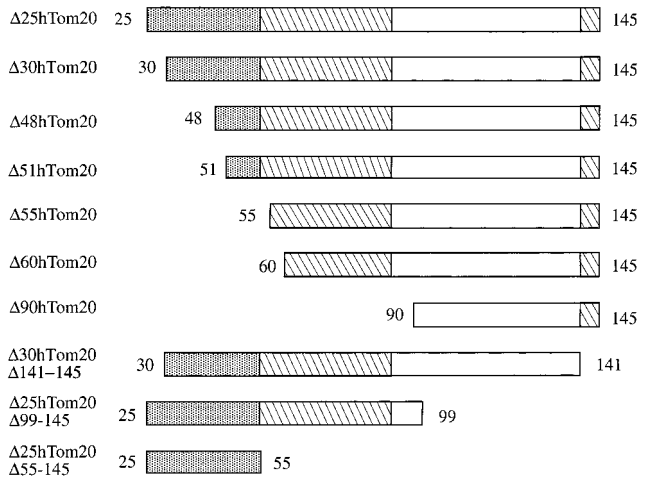

FIGURE 1: Variant forms of hTom20. Regions of hTom20 expressed as fusion proteins with glutathione-*S*-transferase and its thrombin cleavage site include the following: (1) aa25−145hTom20; (2) aa30−145hTom20; (3) aa48−145hTom20; (4) aa51−145hTom20; (5) aa55−145hTom20; (6) aa60−145hTom20; (7) aa90−145hTom20; (8) aa30−141hTom20; (9) aa25−99hTom20; and (10) aa25−55hTom20. The gray box represents a stretch of amino acids that are mostly positively charged while the hatched box indicates those mainly negatively charged.

MATERIALS AND METHODS

Materials and Biochemical Procedures. Chymotrypsin was purchased from Sigma. V8 protease (endoproteinase Glu-C) was from Boehringer/Mannheim. The plasmid pGEX was purchased from Invitrogen. Superdex 12 column and glutathione Sepharose 4B were obtained from Pharmacia. Ultrapure guanidine−HCl was ordered from Schwarz/Mann Biotech, and 1,1,1-trifluoroethanol was obtained from Sigma. Lipid-coated beads were prepared by NIMBUS GmBH, Leipzig, Germany. The pO(1−27) peptide and variants were synthesized as described elsewhere (*28*). Previous articles (*14*, *19*) describe the routine procedures used in this study for in vitro transcription and translation, SDS−PAGE, fluorography, and quantification of radioactive bands.

Plasmid Construction. The plasmid pGST-Δ30hTom20 (*14*) was manipulated by standard polymerase chain reaction techniques to create pGST-Δ48hTom20, pGST-Δ51hTom20, pGST-Δ55hTom20, pGST-Δ60hTom20, pGST-Δ99hTom20, pGST-Δ30hTom20Δ141−145, and pGST-Δ30hTom20Δ91−145. Plasmids pGST-Δ25hTom20 and pGST-Δ25hTom20Δ55−145 were generated using the protocol outlined for the construction of pGST-Δ30hTom20 (*14*). The corresponding polypeptides encoded by these plasmids (after cleavage of the GST domain) are described in Figure 1. Each of the plasmids were finally transformed separately into competent TOPP2 *Escherichia coli* cells (Stratagene). The nucleotide sequences of the hTom20 inserts were verified by sequencing and partly by mass spectroscopy.

Protein Purification. The GST-fusion proteins were purified in buffer A (10 mM Hepes, 32 mM KCl, 0.8 mM

magnesium acetate, 15% glycerol, pH 7.5) for binding studies and buffer B (20 mM sodium phosphate pH 7.4, 150 mM NaCl) for purification of GST-free hTom20 following a modified protocol described in ref *19*. GST-free forms of hTom20 variants were isolated by resuspending the glutathione Sepharose 4B beads saturated with the fusion protein in 1.5 mL of buffer B followed by incubating with 10 U of thrombin (Sigma) overnight. The beads were then pelleted by centrifugation, and the supernatant was incubated for 1 h at room temperature with 0.5 mL of a 50% slurry of glutathione Sepharose 4B. GST-bound Sepharose 4B beads were removed by centrifugation. The purity of the cytosolic domain in the supernatant was monitored by SDS−PAGE and Coomasie blue staining, concentration was determined using Biorad reagent, and results were confirmed by amino acid analysis.

Size Chromatography. One hundred micrograms of the appropriately purified hTom20 variant was loaded onto a Superdex 12 column at 4 °C using a fast protein liquid chromatography system (Pharmacia). Protein was eluted at flow rate of 0.05 mL/min and detected at 214 nm. The hydrodynamic radii and the molecular weights of the purified proteins were calculated using the appropriate protein standards (calibration proteins II, Boehringer/Mannheim; gel filtration calibration kit, Pharmacia). The molecular weight of pODHFR protein was determined using a 14 mL Sephadex G-75 column equilibrated in buffer A with and without 1% Triton X-100. Molecular weights were calculated using the appropriate protein standards.

Determination of Isoelectric Point. The isoelectric point was determined at room temperature by native isoelectric focusing using a gel composed of 1 M urea, 4% polyacrylamide, 5% ampholytes, and 0.1% Triton X-100. The results were verified by monitoring the binding of the protein to Mono S and Mono Q columns (Pharmacia) equilibrated at various pH values.

Limited Proteolysis. The purified cytosolic domain of hTom20 was digested with chymotrypsin or V8 protease using a protocol described by Edwards et al. (*29*). Following digestion, the fragments were purified using C_6 reverse-phase chromatography. The molecular weight of the peptides were determined by mass spectroscopy (Biological Mass Spectrometry Laboratory, University of Watterloo), and the identity of each of the corresponding fragments was deduced by comparing its size to those of proteolytic fragments predicted from the primary sequence of the protein.

Binding Measurements. All binding studies were performed using GST fusions of hTom20 unless otherwise specified. Binding measurements were obtained as described elsewhere (*14*, *19*). The dissociation constant for the MTS preprotein, pODHFR, from the binary complex with receptor or the ternary complex with detergent and receptor was determined using 10 μM GST-Δ30hTom20 and 11 fixed concentrations (ranging from 0.1 to 20 nM) of purified in vitro translated pODHFR, previously quantified by scintillation counting. After 30 min, the complex was incubated with fresh glutathione Sepharose for 30 min at room temperature. The beads were washed and resuspended in buffer containing 10 mM reduced glutathione. The supernatant was decanted into sample buffer and boiled for 5 min at 95 °C, and then aliquots were subjected to SDS−PAGE. Binding was quantified using phospho-imaging to count the amount of radioactivity. The amount of total GST-Δ30hTom20 was determined by Coomassie blue staining of the beads and supernatant. Dissociation constants (K_d) were determined using the following equation:

$$[\text{complex}]/[\text{GST-}\Delta30\text{hTom20}]_{\text{total}} = [\text{pODHFR}]_{\text{free}}/(K_d + [\text{pODHFR}]_{\text{free}}) \quad (1)$$

Binding of the cytosolic domain of hTom20 to liposomes was monitored using a new method (Schmitz, personal communication). Briefly, phosphatidylcholine (PC)-coated beads (Nimbus GmBH−lipid concentration was determined by standard phosphate analysis) were incubated for 1 h with 2.5 μM hTom20 in buffer B in a volume of 100 μL and then peleted by centrifugation at 12000*g* for 2 min. The beads were resuspended in 20 μL loading buffer and boiled for 10 min, and the soluble fraction was subjected to SDS−PAGE. The protein bands were detected by Coomassie blue staining of the gel, and the amount of protein was quantified by laser densitometry using proteins with known concentrations as standards.

Circular Dichroic Spectroscopy. Circular dichroic (CD) spectra were recorded in a Jasco 710 spectropolarimeter. The instrument was routinely calibrated with (+)-10 camphorsulfonic acid. Protein samples were placed in a thermostated 0.1 or 1 cm path length quartz cuvette for measurements from 184 to 320 nm. A nonthermostated double-chamber 1 cm path length quartz cell (Hellma, Muehlheim, Germany) was used for selected readings in the near-UV (250−320 nm). All measurements were recorded at 23 °C with 1 nm bandwidth, 0.2 nm resolution, and multiple accumulations averaged for each sample. Contributions to the signal by buffer were subtracted. Ellipticity in millidegrees was converted to mean residue ellipticity (deg cm^2 dmol^{-1} residue^{-1}) using SigmaPlot for Windows.

Data Analysis. The data obtained from CD spectroscopy and from binding assays were transferred to the SigmaPlot program (Jandel Scientific, San Rafael, CA) and analyzed using the appropriate functions as described in the text. The secondary structure prediction was obtained interactively through the PredictedProtein PHD e-mail server at EMBL using the program PHDsec (*30*), the e-mail server Protein Sequence Analysis (PSA) (*31*), and the PSSP prediction program (*32*). The accuracy of these prediction programs is reported to be between 70 and 80% (*33*). The secondary structural content of the CD spectra recorded from 190 to 240 nm was calculated using a program written by E.S.

RESULTS

The Cytosolic Region of hTom20 Is Monomeric and Contains a Globular Core Domain and a Loosely Folded N-Terminal Domain. Our report details a comparison of the properties of the full-length cytosolic domain, Δ30hTom20 (aa30−145), with selected constructs of this domain truncated at either the N- or C-terminus (Figure 1). The majority of our work will focus on two truncations, Δ30hTom20Δ141−145 and Δ60hTom20. The first truncated form contains the entire cytosolic domain except for aa141−145. These C-terminal five residues of the cytosolic domain are negatively charged and have been hypothesized to interact with aa40−70 of Tom22 to form a structure called an "acidic

Table 1: Values of the Hydrodynamic Radius, Molecular Weight, and Isoelectric Point of the Cytosolic Domain and Variants

protein	hydrodynamic radius (nm)	M_r^{app} (Da)	$M_r^{predict}$ (Da)	ratio[a]	pI[b] (app)	pI (predict)
GST-Δ30hTom20	3.1	42 500	41 793	1.02		
Δ30hTom20	1.9	16 100	13 251	1.22	9.3 ± 0.3	5.9
Δ30hTom20Δ141−145	1.7	13 200	12 663	1.04	9.6 ± 0.4	9.2
Δ60hTom20	1.4	9200	9679	0.95	4.5 ± 0.4	4.4

[a] Ratio between measured and predicted molecular weight. [b] Values represent the mean ± SD of at least three separate determinations.

bristle" (*27*) which serves as a binding site for N-terminal targeting signals. The Δ60 construct (aa60−145) begins just after a strongly predicted loop at aa55−59 (*30*) which is likely to be a domain boundary. Residues aa30−60 preceding this loop are mainly basic, while the region from aa62−90 following the loop contains many acidic groups. Therefore, characterization of the entire cytosolic domain and the N- and C-terminal deleted forms would allow us to address the function of these regions and their contribution to the overall structure of the cytosolic domain.

In this report, the cytosolic domain of hTom20 was overexpressed as a GST-fusion protein as described previously (*19*). Both the GST-fusion protein and the GST-free protein obtained by thrombin cleavage were used to investigate the properties of hTom20. Far-UV CD spectroscopy, which monitors protein secondary structure, was used to verify that the structure of the cytosolic domain was not altered by its fusion with GST or by reaction with thrombin during the course of purification.[2] The functional integrity of the cytosolic domain before and after cleavage from GST has been demonstrated in earlier work (*19*, *34*).

The hydrodynamic radii of the thrombin-cleaved and GST-associated cytosolic domain of hTom20 were measured using gel exclusion chromatography (Table 1). The molecular weights calculated from these measurements establish that the free cytosolic domain, Δ30hTom20, as well as the GST-fusion protein are monomeric. Surprisingly, however, Δ30hTom20 migrated at a weight of 16.1 kDa, 22% higher than expected based on the weight predicted from the protein's primary sequence and from mass spectrometry data,[2] 13.25 kDa. Our interpretation of these results is that the cytosolic domain is not completely spherical and therefore must contain regions of more loosely folded polypeptide. To determine if Δ30hTom20 contains distinctly folded subdomains in solution, we measured the hydrodynamic radii of two truncated cytosolic domains, Δ60hTom20 and Δ30hTom20Δ141−145, by size exclusion chromatography (Figure 1). Unlike the full cytosolic domain, Δ30hTom20, the ratio of the apparent to predicted weight for the two truncated proteins was 1.00 ± 5%. These results indicate qualitatively that the removal of either aa30−60 or aa141−145 from the cytosolic domain yields a more compact, globular protein. Consistent with the result for Δ60hTom20, limited proteolysis experiments (Table 2) involving incubation of Δ30hTom20 with chymotrypsin for 4 h yield only one stable fragment (aa51−145). Similarly, cleavage of the domain by V8 protease after 4 h produced three fragments—aa46−145, aa48−145, and aa55−145 (Table 2). The polypeptide prior to any one of the N-terminal residues was cleaved completely, implying that this N-terminal region is

[2] Data not shown.

Table 2: Peptide Fragments Produced by Limited Proteolysis of Δ30hTom20

enzyme	fragment size (Da)	corresponding protein fragment
V8 protease	11 052	Δ46hTom20
	10 810	Δ48hTom20
	10 084	Δ55hTom20
chymotrypsin	10 482	Δ51hTom20

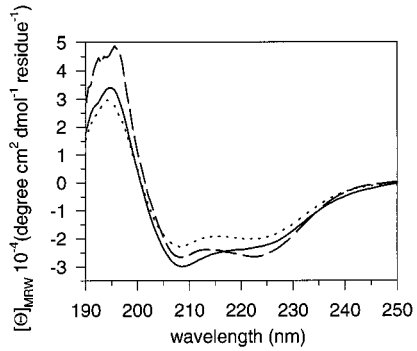

FIGURE 2: CD spectra of the cytosolic domains of hTom20. CD spectra of native proteins △30hTom20 (5 μM, ···), △30hTom20-Δ141−145 (5 μM, —), and △60hTom20 (10 μM, - - -) were recorded in the far-UV (190−250 nm). All measurements were recorded of protein in a buffer of 20 mM sodium phosphate pH 7.4, 150 mM NaCl, and 1 mM DTT. A 0.1 cm cuvette was used. Values were buffer-subtracted and further corrected for differences in protein concentration and primary sequence. Spectral tracings shown represent the average of 15 accumulations.

the most loosely folded and protease-exposed portion of the protein. Secondary structural information of the full-length and truncated cytosolic domains was elucidated by CD spectroscopy. The far-UV CD spectra (195−250 nm) of purified Δ30hTom20 and two truncated versions (Figure 2) showed negative bands at 222 and 207 nm, features that are representative of α-helical structure (*35*, *36*). Estimations of the relative amounts of secondary structure derived from far-UV CD data were comparable to the values obtained from five independent programs predicting secondary structure from primary sequence (Table 3). The full-length cytosolic domain is about 48% α-helical; it contains almost no β-sheet structure, and the remainer of the components include a mixture of random coil, β-turns, and other less common protein secondary structure. While both N- and C-terminal truncations perturb the far-UV CD (Figure 2), only the removal of the larger of the two sections, aa30−60, causes a significant increase (∼15%) in the α-helical content of the cytosolic domain (Table 3).[3] These results support the idea

[3] We have tested the effect of the pH from pH 5 to 9 as well as the effect of ionic strength between 50 and 300 mM NaCl on the secondary structure for these constructs. We did not detect a significant change in the CD spectra under these conditions.

Table 3: Secondary Structural Components of the Full-Length and Truncated Cytosolic Domain of hTom20 and the pO(1−27) Peptide[a]

	n	α-helical (%)	*β*-sheet (%)	other (%)
Δ30hTom20	23	48 ± 1 (40)	3 ± 1 (10)	48 ± 2 (50)
Δ30hTom20Δ141−145	14	50 ± 2	4 ± 1	45 ± 2
Δ60hTom20	11	64 ± 2	0 ± 1	33 ± 2
pO(1−27)/PB	2	16 ± 3	17 ± 2	67 ± 4
pO(1−27)/1% OG	2	49 ± 4	11 ± 3	41 ± 4
pO(1−27)/30% TFE	2	91 ± 2	1 ± 1	9 ± 3

[a] Methods for predicting the secondary structure from the primary sequence of the protein and for calculating experimentally derived secondary structure components are described in Materials and Methods. Values represent the mean ± SD of *n* separate determinations. Predicted values are shown in parentheses. PB, OG, and TFE indicate phosphate buffer, octylglucoside, and 1,1,1-trifluoroethanol, respectively.

that the N-terminal region is more loosely folded and less helical than the region aa60−145.

We obtained information on the accessibility of charged residues in the N- and C-terminal regions of the cytosolic domain by determining the isoelectric points (pI) of Δ30hTom20 and the two truncated variants in solution with gel isoelectric focusing. The entire cytosolic domain is much more positively charged than predicted on the basis of its primary sequence; the apparent pI of Δ30hTom20 is basic (9.3 ± 0.3, Table 1), contrary to the predicted acidic value of 5.9. Truncation of aa141−145 does not significantly alter this value (9.6 ± 0.4, Table 1), implying that the five negative charges of the C-terminus are not exposed at the protein surface in the native conformation of the cytosolic domain. By contrast, the apparent basic pI suggests that the positive charges of the N-terminus *are* surface exposed. The pI is reduced dramatically to 4.5 ± 0.4 by deleting aa30−60, of which 10 are likely positively charged at neutral pH.

Membrane Binding of the Cytosolic Domain Is Mediated by the N-Terminal Region aa30−60 of hTom20. The polypeptide between residues 30 and 60 of hTom20 contains some hydrophobic residues but is strikingly enriched with positively charged amino acids (*34*, *37*, *38*), suggesting that this region might interact with the mitochondrial membrane surface. To address this, we investigated the behavior of Δ30hTom20 during limited proteolysis with chymotrypsin in the absence or presence of detergent or pO(1−27), a 27 amino acid peptide containing the MTS of pOCT. Figure 3 shows that cleavage of the cytosolic domain with chymotrypsin (compare lanes 1 and 2) is markedly reduced or completely inhibited by preincubation of the domain with 1% octylglucoside (lane 3) or pO(1−27) peptide (lane 4), respectively. This ability to stabilize the protein against degradation is a property specific for these two ligands since preincubation with bovine serum albumin does not affect the pattern of cleavage (lane 5). Furthermore, the inhibition of proteolysis by detergent and pO(1−27) peptide is specific for hTom20, since an unrelated protein, Cut repeat 3 + homeodomain protein, is cleaved under the same conditions (compare lane 7 with lanes 8 and 9). Keeping in mind that chymotrypsin's action under conditions of limited proteolysis is directed toward the first 50 amino acids of the cytosolic domain (Table 2), we conclude from the protection experiments that it is the N-terminal of the cytosolic domain of hTom20 that is involved in binding to both the membrane surface and to a MTS peptide.

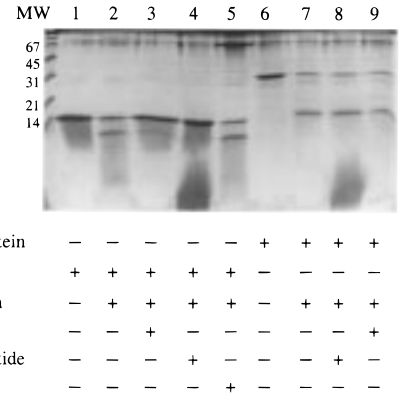

FIGURE 3: SDS−PAGE of △30hTom20 cleaved by limited proteolysis in the absence and presence of detergent and a matrix targeting signal peptide. 30 μM △30hTom20 alone (lane 1) or incubated for 1 h at 37 °C with chymotrypsin at a ratio of 1:1000 (lane 2), and △30hTom20 preincubated for 30 min with 1% octylglucoside (lane 3), or with 100 μM pO(1−27) peptide (lane 4), or with 20 μM BSA (lane 5). As control, 15 μM Cut repeat 3 + homeo domain protein (lane 6) was incubated for 1 h at 37 °C with chymotrypsin at a ratio of 1:1000 (lane 7) or preincubated for 30 min with 100 μM pO(1−29) peptide (lane 8) or 1% octylglucoside (lane 9).

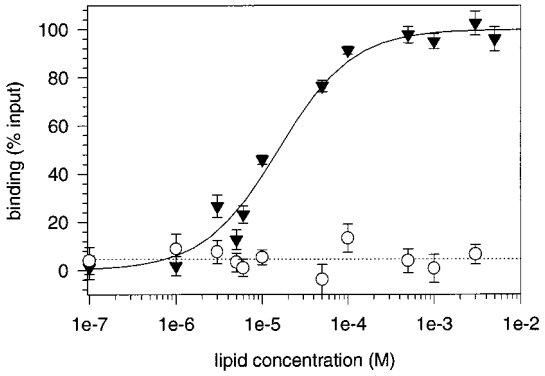

FIGURE 4: Binding of △30hTom20 and △60hTom20 to lipid-coated beads. △30hTom20 or △60hTom20 (2.5 μM) were incubated with increasing amounts of phosphatidylcholine-coated beads and processed as described in Materials and Methods. The data for △30hTom20 (▼) are the averaged values ± SD of five measurements and are fit to the equation for a sigmoidal curve (—) using least-squares analysis. Data for △60hTom20 (○) represent the averaged values ± SD of three measurements and are fit to the equation for a straight line by linear regression.

To establish that the interaction of hTom20 with the membrane surface involves residues between aa30 and 60 of the cytosolic domain, we monitored the binding of Δ30hTom20 and Δ60hTom20 to lipid-coated beads containing similar properties to liposomes (Figure 4, solid line). The data clearly show that these beads bind to Δ30hTom20 with a sigmoidal dependence, yielding a partition coefficient, K_p, of (1.3 ± 0.2) × 10^5 M^{-1}. Such data indicate that the receptor binds the beads with a high affinity. The full-length cytosolic domain also binds to negatively charged phosphatidyl-choline/serine (4:1) vesicles.[2] By contrast, Δ60-hTom20 does not bind to the lipid-coated beads (Figure 4, dotted line). We conclude that the positively charged domain of hTom20 (aa30−60) most likely associates with the mitochondrial surface.

To determine if the binding of aa30−60 to a membrane-like surface induces a structural change in the cytosolic domain, we recorded the CD spectrum of Δ30hTom20 from

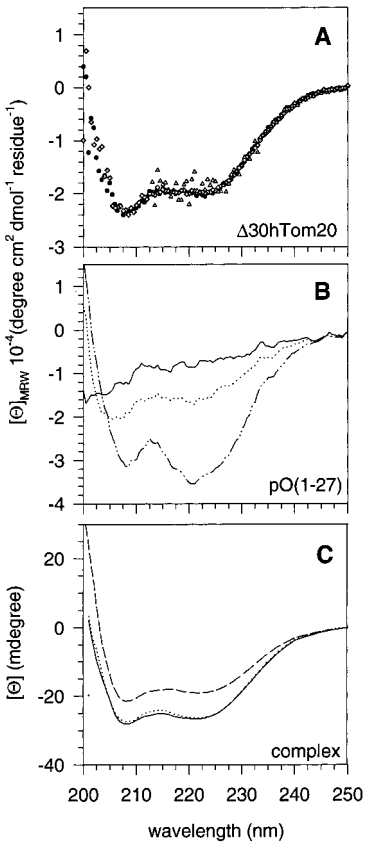

FIGURE 5: The effect of detergent on the CD spectra of △30hTom20, pO(1−27) peptide, and △30hTom20−pO(1−27) peptide complex. The CD spectrum of △30hTom20 (20 μM) was recorded in 20 mM sodium phosphate pH 7.4, 150 mM NaCl, 1 mM DTT (●) and in a buffer containing 0.2% Triton X-100 (△) or 1% octylglucoside (◇) (A). The spectrum of the pO(1−27) peptide (10 μM) was recorded in the same buffer as above (—) and in buffer containing 1% octylglucoside (•••) or 30% trifluoroethanol (TFE) (-•••-) (B). All data were obtained using a 0.1 cm. The summation spectra of △30hTom20 (2.5 μM) and the pO(1−27) peptide (2.5 μM) were recorded from 200 to 250 nm in the absence (- - -) and in the presence of 1% octyglucoside (—) using a 1 cm double-chambered cuvette. The spectrum of the complex (•••) was determined in the presence of 1% octylglucoside. The protein was preincubated with the peptide for 30 min at 25 °C (C). Each curve represents the average of 15 accumulations.

190 to 250 nm in the absence and presence of 1% octylglucoside or 0.2% Triton X-100 (Figure 5A). All three spectra were virtually superimposable, implying that there are no global secondary structural changes in hTom20 induced by its association through aa30−60 with detergents.

Characterization of the Domains of hTom20 Recognizing N-Terminal and Internal Targeting Sequences (NTTS). The cytosolic domain of hTom20 contains two charged regions: a C-terminal acidic region between aa141−145 and an N-terminal region between aa25 and 90 containing acidic and basic clusters (Figure 1). As we expect the positively charged N-terminal targeting signals to interact with negative charges on the receptor, the recognition of these targeting signals likely involves the cytosolic domain between aa25 and 90, since truncation of the C-terminal five amino acids does not affect the binding of this class of preproteins (*14*, *17*). To delineate further the region of interaction of hTom20 with NTTS, we have conducted binding studies using the MTS-containing protein pOCT and several truncations of the cytosolic domain of hTom20 (Figure 6A). While the number

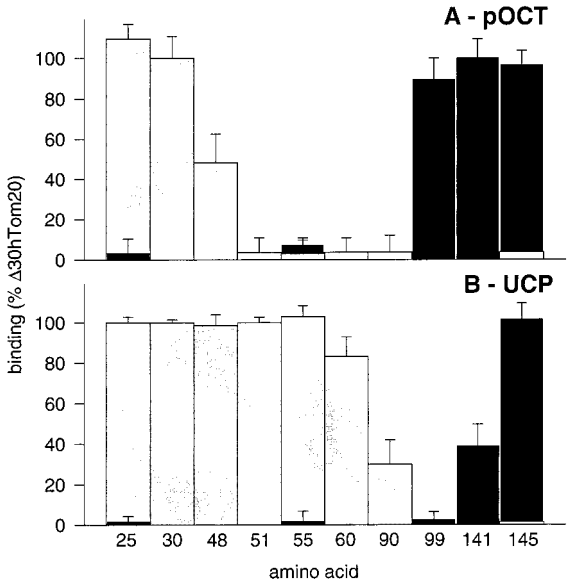

FIGURE 6: Binding of pOCT and the uncoupling protein to full-length and truncated cytosolic domains of GST-hTom20. ^{35}S-radiolabled in vitro translated pOCT (A) or UCP (uncoupling protein) (B) were incubated with glutathione beads saturated with GST fusions of different truncations of the cytosolic domain of hTom20. Binding studies were conducted as described in Materials and Methods. The abscissa indicates the N-terminal (white) or the C-terminal residue (black) of the truncated construct. The order of constructs used, reading the white histograms from left to right, include the following: aa25−145, aa30−145, aa48−145, aa51−145, aa55−145, aa60−145, aa90−145, and GST only. The black histograms, read from right to left, used these constructs: aa30−145, aa30−141, aa25−99, aa25−55, and GST alone (see Figure 1). Each histogram represents the mean ± SD from 3 to 15 separate experiments.

of variants used in these studies is limited, we are able to place aa30 at the N-terminus of the MTS binding domain, while the C-terminal residue was identified between aa55 and 99. Considering the information gained from both the primary sequence and the binding data, we conclude that aa50−90, predominated by a negatively charged cluster, most likely encompass the MTS binding domain. A similar trend was also observed in preliminary binding experiments with another N-terminal targeting sequence, the positively charged signal anchor-containing peptide derived from yTom70.[2]

By using the same experimental template as outlined in Figure 6A, we show in Figure 6B that the region of hTom20 important for recognizing the internal signal sequence of the uncoupling protein lies approximately between aa90 and 145. Similar results were also obtained for another internal signal-containing protein, the voltage-dependent anion channel.[2] Hence, the regions of interaction between hTom20 and either the internal targeting signal (Figure 6A) or MTS (Figure 6B) appear distinct, but we cannot exclude the possibility that they are somewhat overlapping.

Previous studies examining the effect of salt and a nonionic detergent on receptor−ligand interactions (*14*) have clearly established that the binding of hTom20 to internal targeting sequence-containing proteins such as the uncoupling protein and the voltage-dependent anion channel involves hydrophobic interactions. However, similar studies examining the binding of NTTS-containing proteins to Tom20 have not yet led to such definite conclusions. Some results support the involvement of electrostatics (*14*, *15*, *39*), while the study

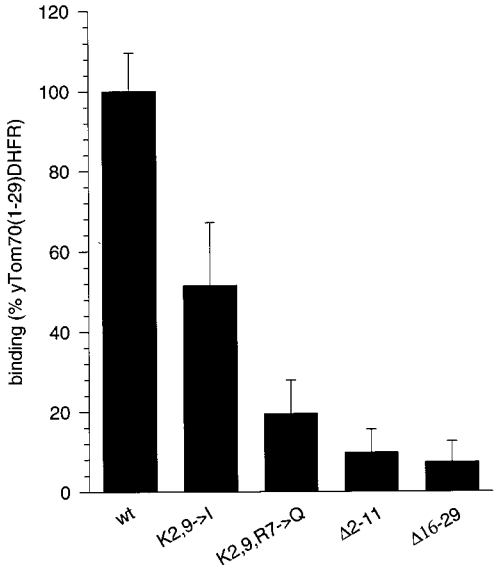

FIGURE 7: Binding of GST-△30hTom20 to various forms of the targeting signal of yTom70. ^{35}S-labeled and in vitro translated yTom70(1−29)DHFR = wt, yTom70(1−29)(K2,9→I)DHFR = K2,9→I, yTom70(1−29)(K2,9,R7→Q)DHFR = K2,9,R7→Q, yTom70(1−29)(△2−11)DHFR = △2−11, yTom70(1−29)(△16−29)DHFR = △16−29 were incubated with glutathione Sepharose previously bound with GST and GST-△30hTom20. Binding is corrected for background and is expressed as wt %. Each histogram represents the mean ± SD from >20 experiments.

by Brix (*16*) disputes the participation of ionic interactions and suggests that hydrophobic interactions are more important. To address this discrepancy, we have compared the binding of hTom20 to yTom70(1−29)DHFR, a hybrid protein bearing the positively charged N-terminal signal anchor sequence of yTom70 (aa1−29) fused to dihydrofolate reductase (DHFR), and mutant forms of the signal protein carrying substitutions and/or deletions in the positively charged and hydrophobic regions of its signal sequence (Figure 7). Modifications within the charged domain (aa1−11) of the signal anchor sequence are as follows: Lys2 and Lys9 were replaced by isoleucines [yTom70(1−29)(K2,9→I)-DHFR]; all charged residues (Lys2, Arg7, Lys9) were replaced by glutamines [yTom70(1−29)(K2,9,R7→Q)-DHFR]; the entire charged domain was deleted (Δ2−11) leaving only the hydrophobic domain [yTom70(12−29)-DHFR]; finally the hydrophobic domain was deleted (Δ16−29) leaving only the charged region [yTom70(1−15)DHFR]. The results suggest that reducing the number of positively charged groups in the targeting signal successively reduces the efficiency of binding to hTom20 (compare lanes 1 with 2, 3, and 4). In addition, hTom20 does not recognize soley the hydrophobic portion of the signal (lane 4). However, the signal bearing only the charged region was not recognized by hTom20 either (lane 5). These results underscore the importance of the whole structure of the yTom70-derived targeting signal in hTom20 recognition, especially the predominant contributions by electrostatic interactions.

Examination of the Interaction between hTom20 and a Matrix Targeting Signal. Our previous studies (*14*) have shown qualitatively that the MTS-containing precursor protein pODHFR binds with a higher affinity to hTom20 in the presence of detergent over the critical micelle concentration (cmc). We have speculated that this increase in efficiency of binding stems from the MTS adopting an amphipathic helix upon binding to detergent micelles (*14*, *20*). Here we further investigate the interaction of MTS with detergent and with its receptor using pO(1−27) peptide and pODHFR as our models for MTS. Binding studies were performed with radiolabeled DHFR-associated targeting signal while specific structural changes in the MTS were monitored by CD spectroscopy using a small signal peptide, such as pO(1−27), to minimize significant CD contributions generated by a large helper protein.

As shown in Figure 5B, when the MTS-containing peptide pO(1−27) was incubated with 1% octylglucoside, a concentration above the cmc, there was a decrease in the mean residue ellipticity between 200 and 222 nm over that observed in phosphate buffer alone. A similar spectrum was found using 0.2% Triton X-100.[2] Secondary structure analysis of the CD data estimates that detergent induces a greater than 2-fold increase in the α-helical content of the peptide (Table 3). Although significant, this increase is less than the 4.5-fold achievable by incubating the peptide with the helix-inducing solvent trifluoroethanol (at 30%) in solution (Table 3) or ethanol (at 80%).[2] Nevertheless, the results we obtain with detergent micelles are comparable to the changes in secondary structure of pO(1−27) peptide in liposomes reported earlier by Epand (*20*). The presence of the receptor does not induce any additional conformational changes in the peptide over that caused by the detergent alone. This is illustrated in Figure 5C, where the CD study outlined in Figure 5B was repeated in a double-chambered cuvette (peptide in one compartment and receptor in the other). As noted in the summation spectrum, the addition of 1% octylglucoside separately to hTom20 and the peptide clearly results in an increase in α-helical content of the proteins. This increase is contributed by the pO(1−27) peptide since detergent does not induce a secondary structural change in the cytosolic domain of hTom20 (see Figure 5A). The far-UV CD (Figure 5C) and near-UV CD spectra[2] of the *complex* in 1% octylglucoside, however, are not different from spectra obtained prior to mixing the protein components, indicating that there are no further changes of the secondary structure in the peptide or hTom20 during complex formation.

To verify that the observed secondary structural changes were due to a specific binding interaction between the detergent micelle surface and a MTS, we monitored directly the formation of the complex between radiolabeled pODHFR and 1% Triton X-100 with size exclusion chromatography.[2] The preprotein migrates in the presence of Triton X-100 with a molecular weight that is ∼90 kDa higher than that in the absence of the nonionic detergent—a weight corresponding to that of a Triton X-100 micelle. Furthermore, the micellar complex is specific for the MTS fusion protein since DHFR migrates with the same molecular weight in the presence and in the absence of detergent.[2] Finally we found by direct binding studies that the interaction between pODHFR and GST-Δ30hTom20 is increased 5-fold in the presence of Triton X-100 (K_d for pODHFR = 0.59 ± 0.03 μM) versus in the absence of nonionic detergent (K_d = 2.62 ± 0.05 μM) (data not shown).

DISCUSSION

In the present study we have characterized the entire cytosolic domain of hTom20 (aa30−145) and selected

truncated versions in order to assess their functional and structural significance. It should be noted that in vivo Tom20 is a member of e receptor complex. Hence, the in vivo properties of Tom20 may differ from those presented in this study due to interactions with other components of this complex. Nevertheless, these findings represent a starting point for the biochemical understanding of the receptor complex and the recognition process.

Structural Properties of the Outer Membrane Receptor Tom20. Our findings illustrate that the cytosolic domain of hTom20 has at least two regions with different structural and functional features: an N-terminal region (aa30−60) and a C-terminal domain (aa60−145). In solution, the N-terminal region is more loosely folded than the C-terminal region since the removal of these 30 residues yields a receptor that is more globular (Table 1) and more α-helical (Figure 2, Table 3) than the entire cytosolic domain, as well as resistant to proteolysis (Table 2). About one-third of the amino acids in this N-terminal region are positively charged (*37*) and solvent exposed, contributing to the unusually basic apparent pI of the full-length cytosolic domain (Table 1). Moreover, it is this surface-exposed basic region in the receptor that likely provides a recognition site for the negatively charged phospholipid surface. This idea is supported by the observation that the N-terminal aa30−60 of the receptor, rather than the remaining region, bind to a membrane-like surface (Figure 4). Interaction with detergent or with N-terminal MTSs appears to stabilize the N-terminal aa30−60 region against proteolytic inactivation (Figure 3). However, stabilization by detergent binding does not accompany global changes in secondary structure in the receptor (Figure 5).

The C-terminal region of hTom20 (aa140−145) appears to play an important structural role since the deletion of these five amino acids stabilizes a more globular form of the receptor (Table 1). Moreover, these C-terminal acidic residues must be buried in the protein since the intact cytosolic receptor has a much more basic pI than expected and this value is not altered by deletion of these residues (Table 1). Surprisingly, truncation of aa141−145 did not eliminate or reduce the binding of the N-terminal MTS-containing protein, pOCT, to hTom20 (Figure 6). Such a result implies that the C-terminal portion of Tom20 does not interact with residues on Tom22 to form the "acidic bristle" as previously speculated (*27*).

Functional Domains within the Receptor. (a) NTTS Binding Domain. Our binding studies further reveal that the aa50−90 of the cytosolic domain of hTom20, rich in acidic residues, recognizes positively charged amphipathic NTTS as found within the MTS of pOCT (Figure 6) and the signal anchor region of yTom70.[2] Hence, the NTTS binding region of the receptor may overlap to some extent with the membrane recognition site. The nature of the interaction of hTom20 with the N-terminal targeting signals most certainly involves electrostatic interactions (Figure 7) and hence is consistent our previous findings (*14*) showing that high salt concentrations rather than nonionic detergents dramatically reduce the binding of proteins containing these targeting signals. However, hTom20 does not recognize the targeting signal bearing only the charged region of a signal anchor sequence (Figure 7), the so-called "cryptic targeting signal" [yTom70(1−10)] (*40*). Furthermore, we have shown (Figures 5C, Table 3) that hTom20 recognizes the α-helical structure of the matrix targeting signal peptide of pOCT, and the dissociation constant of the pODHFR-hTom20 complex found in the presence of detergent ($K_d = 0.59 \pm 0.03$ μM) is five times higher than that in absence of detergent ($K_d = 2.62 \pm 0.05$ μM). From these results we conclude hTom20 requires not only the positive charges of an NTTS sequence for preprotein recognition but also the correct tertiary structure associated with the sequence.

We provide experimental evidence that the receptor binds preferentially to the α-helical structure of the signal peptide, a structure only adopted in hydrophobic environment (Figure 5B and refs *20* and *21*). Furthermore, there are no further changes of the overall secondary structure of the targeting signal upon the formation of the ternary complex with the receptor (Figure 5C).

(b) Internal Targeting Signal Binding Region. In contrast to the N-terminal region, the aa60−145 region does not interact with the lipid surface (Figure 4). Instead, this region comprises two highly conserved motifs—the B domain of the tetratricopeptide repeat motif (aa89−101) which interacts with Tom70 (*25*), and the Q motif, a nine-residue amphiphilic helix following Pro103 (*37*) which plays a major role in the recognition of proteins such as the uncoupling protein and the voltage-dependent anion channel bearing internal targeting signals (*14*). The Q motif comprises aa105−113 in hTom20, aa126−134 in *N. crassa*, aa129−137 in yeast, and aa120−129 in potato. Our binding studies (Figure 6) show that the recognition site for internal targeting signals likely encompasses aa90−145, a region much larger than that defined by the Q motif, but may possibly be overlapping with the NTTS binding site. Clearly, the binding site for proteins containing an internal targeting signal is distinct from the membrane binding site. Such a mechanism could prevent the hydrophobic preprotein from entering the outer mitochondrial membrane, ensure that the inner membrane protein would not be mistargeted into the outer membrane, or prevent outer membrane proteins from entering the membrane in an incorrect orientation.

ACKNOWLEDGMENT

We are indebted to Dr. A. Edwards and Mr. S. Orlicky for teaching the limited proteolysis technique, Dr. A. Vrielink for the Cut repeat 3+ homeo domain protein, Dr. E. Meighen for the Superdex 12 column, and in particular to Dr. G. C. Shore for helpful discussions and support.

REFERENCES

1. Pfanner, N., and Meijer, M. (1997) *Current Biology 7*, R100−R103.
2. Haucke, V., and Lithgow, T. (1997) *J. Bioenerg. Biomembr. 29*, 11−17.
3. Neupert, W. (1997) *Annu. Rev. Biochem. 66*, 863−917.
4. Tokatlidis, K., Junne, T., Moes, S., Schatz, G., Glick, B. S., and Kronidou, N. (1996) *Nature 384*, 585−588.
5. Sirrenberg, C., Bauer, M. F., Guiard, B., Neupert, W., and Brunner, M. (1996) *Nature 384*, 582−585.
6. Kubrich, M., Dietmeier, K., and Pfanner, N. (1995) *Curr. Genet. 27*, 393−403.
7. Schatz, G. (1996) *J. Biol. Chem. 271*, 31763−31766.
8. Lill, R., Nargang, F. E., and Neupert, W. (1996) *Curr. Opin. Cell Biol. 8*, 505−512.

9. Pfanner, N., Douglas, M. G., Endo, T., Hoogenraad, N. J., Jensen, R. E., Meijer, M., Neupert, W., Schatz, G., Schmitz, U. K., and Shore, G. C. (1996) *Trends Biochem. Sci. 21*, 51−52.
10. Schlossmann, J., Lill, R., Neupert, W., and Court, D. A. (1996) *J. Biol. Chem. 271*, 17890−17895.
11. Bomer, U., Pfanner, N., and Dietmeier, K. (1996) *FEBS Lett. 382*, 153−8.
12. Sollner, T., Griffiths, G., Pfaller, R., Pfanner, N., and Neupert, W. (1989) *Cell 59*, 1061−1070.
13. Millar, D. G., and Shore, G. C. (1996) *J. Biol. Chem. 271*, 25823−25829.
14. Schleiff, E., Shore, G. C., and Goping, I. S. (1997) *J. Biol. Chem. 272*, 17784−17789.
15. Komiya, T., Rospert, S., Schatz, G., and Mihara, K. (1997) *EMBO J. 16*, 4267−4275.
16. Brix, J., Dietmeier, K., and Pfanner, N. (1997) *J. Biol. Chem. 272*, 20730−20735.
17. Bolliger, L., Junne, T., Schatz, G., and Lithgow, T. (1995) *EMBO J. 14*, 6318−6326.
18. Mayer, A., Nargang, F. E., Neupert, W., and Lill, R. (1995) *EMBO J. 14*, 4204−4211.
19. Schleiff, E., Shore, G. C., and Goping, I. S. (1997) *FEBS Lett. 404*, 314−318.
20. Epand, R. M., Hui, S. W., Argan, C., Gillespie, L. L., and Shore, G. C. (1986) *J. Biol. Chem. 261*, 10017−10020.
21. Roise, D., Horvath, S. J., Tomich, J. M., Richards, J. H., and Schatz, G. (1986) *EMBO J. 5*, 1327−1334.
22. Hammen, P. K., Gorenstein, D. G., and Weiner, H. (1994) *Biochemistry 33*, 8610−7.
23. Sollner, T., Pfaller, R., Griffiths, G., Pfanner, N., and Neupert, W. (1990) *Cell 62*, 107−115.
24. Mihara, K., and Omura, T. (1996) *Trends Cell Biol. 6*, 104−108.
25. Haucke, V., Horst, M., Schatz, G., and Lithgow, T. (1996) *EMBO J. 15*, 1231−1237.
26. Iwahashi, J., Yamazaki, S., Komiya, T., Nomura, N., Nishikawa, S., Endo, T., and Mihara, K. (1997) *J. Biol. Chem. 272*, 18467−18472.
27. Lithgow, T., Junne, T., Suda, K., Gratzer, S., and Schatz, G. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 11973−11977.
28. Gillespie, L. L., Argan, C., Taneja, A. T., Hodges, R. S., Freeman, K. B., and Shore, G. C. (1985) *J. Biol. Chem. 260*, 16045−16048.
29. Morin, P. E., Awrey, D. E., Edwards, A. M., and Arrowsmith, C. H. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 10604−10608.
30. Rost, B., and Sander, C. (1993) *J. Mol. Biol. 232*, 584−599.
31. Stultz, C. M., White, J. V., and Smith, T. F. (1993) *Protein Sci. 2*, 305−314.
32. Solovyev, V. V., and Salamov, A. A. (1994) *Comput. Appl. Biosci. 10*, 661−669.
33. Sternberg, M. J. E. (1996) *Protein Structure Prediction*, Vol. 170, IRL Press, Oxford.
34. Hanson, B., Nuttal, S., and Hoogenraad, N. (1996) *Eur. J. Biochem. 235*, 750−753.
35. Greenfield, N., and Fasman, G. D. (1969) *Biochemistry 8*, 4108−4116.
36. Saxena, V. P., and Wetlaufer, D. B. (1971) *Proc. Natl. Acad. Sci. U.S.A. 68*, 969−972.
37. Goping, I. S., Millar, D. G., and Shore, G. C. (1995) *FEBS Lett. 373*, 45−50.
38. Seki, N., Moczko, M., Nagase, T., Zufall, N., Ehmann, B., Dietmeier, K., Schafer, E., Nomura, N., and Pfanner, N. (1995) *FEBS Lett. 375*, 307−310.
39. Haucke, V., Lithgow, T., Rospert, S., Hahne, K., and Schatz, G. (1995) *J. Biol. Chem. 270*, 5565−5570.
40. McBride, H. M., Goping, I. S., and Shore, G. C. (1996) *J. Cell Biol. 134*, 307−313.

BI9807456