Ssp1, a Site-specific Parvulin Homolog from *Neurospora crassa* Active in Protein Folding*

(Received for publication, May 20, 1998, and in revised form, August 31, 1998)

Oliver Kops, Christoph Eckerskorn‡, Sandra Hottenrott§, Gunter Fischer§, Huai Feng Mi, and Maximilian Tropschug†

From the Institut für Biochemie und Molekularbiologie der Universität Freiburg, Hermann-Herder-Strasse 7, D-79104 Freiburg, Germany, ‡Max-Planck-Institut für Biochemie, Am Klopfsitz, D-81377 München, Germany, and §Max-Planck-Gesellschaft, Forschungsstelle "Enzymologie der Proteinfaltung," Weinbergweg 52, D-06120 Halle/Saale, Germany

**Peptidyl-prolyl cis-trans-isomerases (PPIases)** are enzymes capable of isomerizing a Xaa-Pro peptide bond. Three families of PPIases are known: cyclophilins, FK-BPs, and parvulins. The physiological functions of the PPIases are only poorly understood. Eucaryotic members of the parvulin family have recently been shown to be essential for regulation of mitosis.

Here we describe the purification and characterization of Ssp1, an abundant parvulin homolog from *Neurospora crassa*, which is unique among the known eucaryotic parvulins in containing a polyglutamine stretch between the N-terminal WW domain and the C-terminal PPIase domain. Ssp1 is a site-specific PPIase with respect to the amino acid N-terminal to the proline residue. Peptides with glutamate, phosphoserine, or phosphothreonine in the −1-position proved to be the best substrates. Ssp1 is not only able to isomerize small peptides but is also active in protein folding, as shown with mouse dihydrofolate reductase. Using the substrate specificity of Ssp1, we could identify Glu81-Pro82 as a PPIase-sensitive site in folding of dihydrofolate reductase.

These results demonstrate that Ssp1 is a potent mediator of protein folding and that parvulins can serve as tools to elucidate rate-limiting steps in protein folding reactions.

* This work was supported by Deutsche Forschungsgemeinschaft Grant SFB 388/B3 and the Fonds der Chemischen Industrie. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank®/EBI Data Bank with accession number(s) AJ006023.

†To whom correspondence should be addressed. Tel.: 49 761 2033244; Fax: 49 761 2035283; E-mail: tropschu@uni-freiburg.de.

‡The abbreviations used are: PPIase, peptidyl-prolyl cis-trans-isomerase; CyP, cyclophilin; DHFR, dihydrofolate reductase; FKB, FK-506-binding protein; PCR, polymerase chain reaction; pNA, p-nitroanilide.

---

Peptidyl-prolyl *cis*-trans-isomerases (PPIases; 1 EC 5.2.1.8.) are enzymes able to catalyze the *cis*-trans-isomerization of Xaa–Pro bonds in oligopeptides and proteins (for reviews, see Refs. 1–5). PPIases are ubiquitous and abundant proteins and have been found in bacteria, fungi, plants, and mammals. They belong to three structural unrelated families, namely the cyclophilins (CyPs), the FK-BPs, and the parvulin-like proteins. CyPs and FK-BPs are characterized by their ability to bind the immunosuppressant drugs cyclosporin A and FK506/rapamycin, respectively.

Due to their enzymatic activity, CyPs and FK-BPs are able to accelerate slow refolding steps in certain proteins in *vitro* (see Ref. 2). Mitochondrial CyPs from *Neurospora crassa* (6, 7) and yeast (8) have been shown to be part of the protein folding machinery of the organelle involving molecular chaperones (6, 7).

The first member of the parvulin family was discovered in *Escherichia coli* (9–11). Eucaryotic parvulins were found in yeast (Ess1/Ptf1; Refs. 12 and 13), *Drosophila* (dodo; Ref. 14), and humans (Pin1; Ref. 15). In contrast to the small *E. coli* parvulin (10 kDa), which consists only of a PPlase domain, the other members of the eucaryotic parvulin family have an additional WW domain at their N terminus. The WW domain is a structural motif containing two invariant tryptophan residues thought to be involved in protein-protein interactions by binding to short proline-rich segments of target proteins (16, 17).

Deletions of *ESS1* in yeast are lethal (12). This phenotype can be rescued by expressing either Pin1 or dodo (14, 15). Furthermore, overexpression of Pin1 in HeLa cells inhibits G2/M transition, whereas depletion of Pin1 induces mitotic arrest (15). These data point to an important function of the eucaryotic parvulins in mitotic control. Ranganathan et al. (18) suggested that substrate recognition of Pin1 is phosphorylation-dependent. Yaffe et al. (19) and Shen et al. (20) indeed showed that Pin1 binds mitotic phosphoproteins containing phosphothreonine-proline or phosphoserine-proline bonds.

Here we describe the purification, cloning, and biochemical analysis of a homolog of Pin1 from the fungus *N. crassa*, called Ssp1. Ssp1 is unique in containing a polyglutamine stretch between the WW and the PPIase domains. Like Pin1, Ssp1 is a site-specific PPlase, highly preferring acidic residues (glutamate, phosphoserine, or phosphothreonine) amino-terminal to a proline. We show for the first time that an eucaryotic parvulin (Ssp1) is not only active in peptide PPIase assays but also in protein folding. As a substrate, we chose dihydrofolate reductase (DHFR), which is widely used as a model protein in the analysis of intracellular protein folding and trafficking (see e.g., Refs. 21–27). Yet, all steps of the complex folding pathway of DHFR are not fully resolved. Using Ssp1 with its site-specific prolyl isomerase activity, we could identify the isomerization of Glu81-Pro82 in mouse DHFR as being rate-limiting in the folding reaction.

**EXPERIMENTAL PROCEDURES**

*Preparation of *N. crassa* Lysate—*Cultures of* *N. crassa* were grown as described (28). Hyphae from 10-liter cultures were frozen at −80 °C for 1 h and then homogenized in a Waring blender for 10 min in the cold room with 2 volumes of lysate buffer (50 mM HEPES/NaOH, pH 7.8; 5 mM dithiothreitol; 1 mM EDTA; Boehringer complete protease inhibitors). Cell debris were pelleted by centrifugation (30 min, 15,000 × *g*), and the cloudy supernatant was spun again at 100,000 × *g* for 90 min. The clear cell lysate was withdrawn, frozen in liquid nitrogen, and stored at −80 °C until use. A typical protein yield was 4–5 mg of protein/ml.*
**Ssp1 Purification**—200 ml of *N. crassa* lysate, corresponding to about 1 g of total protein, were loaded onto a 50-ml DEAE-Sepharose Fast-Flow column (Amersham Pharmacia Biotech) at 1 ml/min. The flowthrough was loaded onto a 10-ml Affi-Gel Blue (Bio-Rad) column at 0.5 ml/min. Ssp1 was eluted at salt concentrations between 250 and 500 mM NaCl (in buffer I; 50 mM HEPES/NaOH, pH 7.8). Fractions containing Ssp1 were dialyzed overnight against buffer II (75 mM Tris acetate, pH 9.3) and loaded onto a Mono-P column (Amersham Pharmacia Biotech). Proteins were eluted using Polybuffer 96 (Amersham Pharmacia Biotech). Ssp1 eluted at a pH of about 8. The corresponding fractions were concentrated using a 10-kDa exclusion membrane (Millipore Corp.). Finally, Ssp1 was purified to homogeneity by gel filtration using a Superdex 75 16/60 column equilibrated with 35 mM HEPES/NaOH, pH 7.8, 150 mM NaCl (buffer III).

Throughout the procedure, the presence of Ssp1 was monitored by Western blot analysis using a cross-reaction polyclonal antiserum raised against yeast Ptf1 (13).

**Amino Acid Analysis and Cloning of Ssp1**—Purified Ssp1 was digested in a gel after electrophoresis using protease Lys-C, and resulting peptides were separated according to Ref. 29. Two peptides (see Fig. 2A) were used to design degenerated oligonucleotides for PCRs using the codon usage of *N. crassa* (30): SSP-A (5′-TGG GAG CCI CCI (A/T)(G/C)I GGI AC-3′) and SSP-B (5′-TTTCTC TGCC ATG TCI CCI (T/G)I CCG AAG TA-3′). PCR reactions using a *N. crassa* cDNA library made by using Marathon cDNA kit (CLONTECH) resulted in a Ssp1-specific band, as judged by sequencing. Further PCR reactions using Ssp1-specific primers SSP-C (5′-AGT CGG ACT GCT CCT CGT-3′) and SSP-E (5′-TAA TCT CGG ACT CGC GCC A-3′) and cDNA adaptor primers AP1 and AP2 (Marathon cDNA kit, CLONTECH) yielded a full-length cDNA for Ssp1. For expression of Ssp1, oligonucleotides SSP forward (5′-GAA TTC ATG TCC AAC ACC ATC GAG ATT AGC TCC TCT CCT CCT ACC GAC ACC-3′) and SSP reverse (5′-GGG CTA CTC GAG CCG TTC AAT CAA ATG AGC TCC TCG TGG TGC TTT TGG GAT TTC TCT ACT GAG AA-3′) were used. All mutations were verified by DNA sequencing of the entire reading frames.

**Miscellaneous**—CyP20 was purified from *N. crassa* as described (7).

**RESULTS**

**Purification, Cloning, and Expression of Ssp1**—In initial experiments, we detected an abundant cytosolic protein in *N. crassa* extracts that cross-reacted with a polyclonal antiserum against yeast Ptf1 (13). We set out to purify this protein (Ssp1) from *N. crassa* hyphae using DEAE-Sepharose chromatography, Affi-Gel Blue chromatography, chromatofocusing, and a final gel filtration step as described under “Experimental Procedures.”

A typical yield of 4–5 μg of Ssp1 starting with 1 g of total protein was obtained. Pure Ssp1 is a protein with an apparent molecular mass on SDS gels of 23 kDa (Fig. 1, lane 5). We prepared antibodies against Ssp1 in rabbits and by using analytical Western blotting, we measured a Ssp1 content of 0.05–0.1% (using total extracts of *N. crassa* hyphae).

In order to get sequence information for cloning the cDNA, purified Ssp1 was excised from a Coomassie-stained gel and digested with protease LysC, and the resulting peptides were purified according to Ref. 29. Two internal peptides were sequenced. Peptides 1 and 2 yielded the sequences KTS(Q)/W-PPS(S/T)D(V/I)V/K and KGGDLGGFGROGD(M/K)Q, respectively (see *underline* in Fig. 2A; amino acids in parentheses represent the most likely amino acid from high performance liquid chromatography).

Based on the sequences of the peptides and by using the codon usage of *N. crassa* (30), two degenerated oligonucleotides (SSP-A and SSP-B) were synthesized and used in PCRs with *N. crassa* cDNA libraries (33–35). Sequencing of a resulting 300-bp PCR product showed that the deduced amino acid had high homology to the three eucaryotic parvulins known to date (data not shown).

Further PCRs were used to identify and clone a cDNA that contained the entire open reading frame coding for Ssp1. The deduced amino acid sequence of Ssp1 is shown in Fig. 2A. Ssp1 is a hydrophilic protein with a molecular mass of 20,674 Da (182 amino acids). The calculated isoelectric point is 7.4.

Ssp1 exhibits a high degree of sequence identity to the three other eucaryotic parvulins known to date (Fig. 3). Sequence identities are 50% (Pin1), 45% (Ess1/Ptf1), and 44% (dodo), respectively.
In contrast to the other three members of the eucaryotic parvulin family, Ssp1 is unique in containing a polyglutamine stretch consisting of 11 glutamine residues and a proline residue between the conserved WW and PPIase domains, respectively (Fig. 2B). Polyglutamine regions are characteristic regulatory modules being present in various proteins involved in replication, transcription, and cellular regulation (see e.g. Refs. 36–40).

In order to get high amounts of Ssp1 for biochemical analysis, we expressed a glutathione S-transferase fusion protein in E. coli. Purification of the recombinant protein, cleavage using factor Xa, and purification of recombinant Ssp1 resulted in a pure protein with an apparent molecular mass of 23 kDa comigrating with purified Ssp1 on SDS gels (not shown).

**PPIase Activity and Substrate Specificity of N. crassa Ssp1**—In order to examine the PPIase activity and the substrate specificity of Ssp1, we carried out in vitro PPIase assays (31) using 10 different peptide substrates varying in the -1-position with regard to the proline (see Table 1).

As shown in Table 1, Ssp1 is a site-specific PPIase with regard to the -1-position. Using peptides with His, Leu, Ala, Phe, Gln, and Glu in the -1-position, respectively, Ssp1 catalyzes only the *cis*-trans-isomerization of a peptide containing Glu in front of a proline. Exchanging the Glu N-terminal to the proline with basic or neutral amino acids greatly reduces the PPIase activity.

Interestingly, the highest $k_{cat}/K_m$ value was found with the peptide substrate Ac-Ala-Ala-Ser-Pro-Arg-pNA, where the Ser in the -1-position is phosphorylated (Table 1). In contrast, the nonphosphorylated peptide is a bad substrate for the enzyme. The same is true for a peptide containing Thr in the -1-position. However a peptide containing a phosphorylated Thr is a good substrate for Ssp1 (Table 1).

We conclude that Ssp1 is a site-specific PPIase, highly preferring acidic residues in the -1-position relative to the proline and perhaps also in the -2-position (see below).

**Protein Folding Activity of Ssp1**—Is Ssp1 able to accelerate the refolding of proteins? Scholz et al. (41) and Dolinski et al. (50) have demonstrated that there may be a difference when measuring PPIase activity with either peptide assays or protein folding activity, respectively. Protein folding activity has been shown to date for members of the cyclophilin and the FKBP families, respectively (6–8, 35, 43). The small *E. coli* parvulin, which does not contain a WW domain, is also active in protein folding (42) in *vitro*; for eucaryotic parvulins, no protein folding activity has been demonstrated to date.

We used mouse DHFR as a model protein for folding. This enzyme is widely used as a model protein in studies of cellular protein maturation and trafficking (e.g. Refs. 21–27). The folding pathway of DHFR is complex and not fully understood (44–46). Prolyl isomerization\(^2\) has been demonstrated to be important for the *E. coli* enzyme (47).

As shown in Fig. 4, Ssp1 is indeed able to accelerate the refolding of mouse DHFR. It exhibits about the same catalytic efficiency in the refolding assay as an abundant cyclophilin from *N. crassa*, namely CyP20 (Fig. 4; see also Refs. 6 and 7). This is the first demonstration that a eucaryotic parvulin (Ssp1) is able to accelerate protein folding *in vitro*.

**Use of the Site Specificity of Ssp1 to Identify a PPIase-sensitive Glu-Pro Site in Mouse DHFR**—Taking into account that Ssp1 is very specific with respect to the amino acid in the -1-position relative to proline (see above), we concluded that Glu-Pro (and perhaps Asp-Pro) bonds should be critical PPIase-sensitive sites during folding of mouse DHFR. Mouse DHFR has 13 Xaa-Pro sites (see Fig. 6), which all might be possible candidates for the action of a PPIase. But there is only one acidic (glutamate) residue amino-terminal to a proline in the DHFR sequence (Glu\(^{81}\)-Pro\(^{82}\); see Fig. 5).

Fig. 5 shows a homology alignment of the amino acid sequences of mouse DHFR (32) and *E. coli* DHFR (48). As can be seen from this figure, Glu\(^{81}\)-Pro\(^{82}\) of mouse DHFR aligns with Gln\(^{65}\)-Pro\(^{66}\) of *E. coli* DHFR. Strikingly, *cis*-trans-isomerization of the Gln\(^{65}\)-Pro\(^{66}\) bond in *E. coli* DHFR was suggested to be important for fast folding of the protein (47). This observation and the site specificity of Ssp1 in the peptide assay (see above) prompted us to test whether the homologous Glu\(^{81}\)-Pro\(^{82}\) bond in mouse DHFR is also a critical PPIase-sensitive site for the folding process.

We therefore generated mouse DHFR mutants, namely DHFR E81A and DHFR E81L. The mutant proteins folded efficiently after synthesis in reticulocyte lysates; folding was tested as resistance against proteinase K (6, 7).

We then tested the influence of Ssp1 on the folding of DHFR E81A and DHFR E81L. Despite the fact that Ssp1 is able to accelerate folding of the wild type protein (see above), it is not able to accelerate folding of the respective mutant proteins (Fig. 6, A and B).

Cyclophilin 20 from *N. crassa* (CyP20) was used as a control to see if the mutations had made the resulting protein insensitive for the action of a PPIase. Cyclophilins are known to have a low substrate specificity regarding the -1-position (49). Fig. 6, A and B, shows that CyP20 was able to accelerate folding of the mutant proteins with almost the same refolding efficiency as the wild type protein.

In addition, we constructed mutant DHFR P82A. CyP20 and also Ssp1, surprisingly, are efficiently able to accelerate refolding of this mutant (see Fig. 6C). Uncatalyzed refolding of DHFR P82A is in the same range as wild type DHFR (compare Figs. 4B and 6B).

How can the removal of a proline still lead to effective refolding activity of both PPIases? We suggest that the proline

---

\(^2\) The term “prolyl isomerization” is used throughout this paper for the *cis*-trans-isomerization of the peptide bond preceding proline in an amino acid sequence. Similarly, the term “prolyl bond” is synonymous with the peptide bond preceding proline. -1-position denotes an amino acid amino-terminal to a proline in an amino acid sequence.
following immediately after Pro82 (Pro83; see Fig. 5) might also play a role in the folding pathway of mouse DHFR and in catalysis by CyP20 and Ssp1.

**DISCUSSION**

This paper describes the first biochemical isolation of a eucaryotic member of the parvulin family. In addition, we have characterized this novel PPIase from *N. crassa*. In comparison with the other eucaryotic members of the parvulin family known to date, Ssp1 is unique in two respects. (a) It is an abundant enzyme (0.05–0.1% of total cellular protein), in contrast to Pin1, whose total cellular concentration in HeLa cells...
was estimated to be 0.5 \mu m (20). Ssp1 concentration therefore is in the same range as for other abundant PPIases like CyP20 (33) and FKBP13 (35) in \textit{N. crassa}. In addition, Ssp1 seems not only to be located in the nucleus (like human Pin1; Ref. 15), since we can detect it in and purify it from cytosolic fractions of \textit{N. crassa}.\textsuperscript{3} (b) In contrast to the known eucaryotic members of the parvulin family from \textit{Saccharomyces cerevisiae} (12, 13), \textit{Drosophila melanogaster} (14), or human cells (15), \textit{N. crassa} Ssp1 contains a polyglutamine stretch in the interconnecting region between the conserved WW and the PPIase domains, respectively. Polyglutamine regions are characteristic regulatory modules present in various proteins involved in replication, transcription, and cellular regulation (see e.g. Refs. 36–40). The functional role of the polyglutamine stretch in Ssp1 remains to be clarified.

Taken both points together, we speculate that there might exist more than one parvulin-like protein in higher eucaryotic cells than yeast and that we might have isolated the first member of this new type. In this respect, it is interesting to note that the only parvulin member in yeast (Ess1/Ptf1) is essential (12) but \textit{dodo} in \textit{D. melanogaster} is not (14). Whether Pin1 is essential in human cells is a matter of debate (15, 51, 52), so it might be the case that in higher eucaryotic cells parvulin isoforms exist, as has long been known for CyPs and FKBPs.

In contrast to CyPs and FKBPs, Ssp1 shows a very high substrate specificity in standard PPIase assays using short peptides (31). Our results reveal that peptides containing acidic residues like glutamate, phosphoserine, or phosphothreonine N-terminal to a proline are highly preferred substrates for the PPIase activity of Ssp1. Therefore, the substrate specificity of Ssp1 and the \(k_{\text{cat}}/K_m\) values are comparable with that of the human homolog Pin1 (18, 19).

This theory implies that Pin1 is not only able to act as a PPIase on peptide substrates but also exhibits the same substrate specificity regarding proteins. Recent publications show, however, that results obtained in the peptide PPIase assays might not hold true for protein substrates (41, 50). Both papers compare the PPIase activity of active site mutant PPIases, as deduced from the peptide PPIase assay, with their ability to refold denatured protein substrates. Mutant cyclophilins and FKBPs, inactive in the peptide assay, however, were active in a protein folding assay. Therefore, one has to be cautious in applying results from the protease-coupled peptide assay to the protein folding activity of a PPIase.

With this in mind, we tested whether Ssp1 is also able as a PPIase to accelerate folding of a protein.

\footnote{O. Kops and M. Tropschug, unpublished results.}
biochemical mechanism of action of Ssp1 during DHFR folding in vivo. The high substrate specificity of Ssp1 for acidic residues N-terminal to a proline, we speculated that Glu81-Pro82 is a crucial site for peptide-prolyl cis-trans-isomerization. The generation of two mutant proteins DHFR E81A and DHFR E81L strengthens our speculation. Ssp1 is inactive in refolding the mutant proteins, whereas refolding is efficiently catalyzed by CyP20, reflecting the low substrate specificity of CyPs.

Surprisingly, refolding of a mutant where Pro82 is exchanged for Ala (DHFR P82A) can also be accelerated by both CyP20 and Ssp1 (Fig. 6C). We speculate that in this mutant Ala82-Pro82 becomes the PPIase-sensitive site; Glu 81 seems to make this region accessible for Ssp1. Further experiments are necessary to unravel this novel finding, that a Glu in the 2-position relative to a proline is able to generate a Ssp1-sensitive site.

Exchanging both Pro82 and Pro83 for Ala resulted in an unstable protein that did not fold correctly (as tested by protease resistance) and therefore was not accessible for refolding assays (data not shown). Obviously, at least one proline is necessary for correct folding in this region of DHFR.

We conclude that the cis-trans-isomerization of the Glu81-Pro82 bond in mouse DHFR represents a rate-limiting step in the folding pathway. This notion is supported by former studies (47) identifying a Gln–Pro bond as a critical site for the fast refolding of the E. coli protein. Furthermore, Ssp1, a member of the eucaryotic parvulin family, is indeed able to accelerate rate-limiting steps in protein folding due to its PPIase activity. Also, the substrate specificity of Ssp1, as determined by the peptide assay, holds true for protein substrates, like DHFR, with the interesting finding that acidic residues N-terminal (and perhaps also in the 2-position relative to a proline) are highly preferred, if not mandatory, for the action of this specific PPIase. In view of these results, we suggest that Ssp1 and its homologs from other organisms might act as site-specific PPIases in vivo. However, whether (a) Glu(Xaa)-Pro (and perhaps Asp(Xaa)-Pro) or (b) Ser(P)-Xaa-Pro and Thr(P)-Xaa-Pro, are the real substrates for Ssp1 and are used with the same efficiency in vivo remains to be clarified.

Further experiments will be necessary to elucidate the exact biochemical mechanism of action of Ssp1 during DHFR folding and to study the role of the two adjacent prolines Pro82-Pro83. Also, the role of acidic residues (Glu and Asp) in a 2-position relative to a proline will be clarified. In any case, the use of the site-specific PPIases of the parvulin family can aid in a better understanding of the refolding pathways of certain proteins that depend on prolyl cis-trans-isomerization.

Acknowledgments—We thank Dr. Horst Domdey for a gift of anti-ase serum against yeast Ptf1 and Dr. von Ahsen for purifying CyP20. We thank Dr. Joachim Rassow for stimulating discussions and critical reading of the manuscript.