Stockholm University

This is a submitted version of a paper published in *The FEBS Journal*.

Citation for the published paper:

Ye, W., Spånnig, E., Unnerståle, S., Gotthold, D., Glaser, E. et al. (2012)
"NMR investigations of the dual targeting peptide of Thr-tRNA synthetase and its
interaction with the mitochondrial Tom20 receptor in *Arabidopsis thaliana*"
The FEBS Journal, 279(19): 3738-3748

Access to the published version may require subscription.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-82129>

<http://su.diva-portal.org>

NMR Investigations of the Dual Targeting Peptide of Thr-tRNA Synthetase and its Interaction with the Mitochondrial Tom20 Receptor in *Arabidopsis thaliana*

**Weihua Ye, Erika Spåanning, Sofia Unnerståle, David Gotthold, Elzbieta
Glaser and Lena Mäler***

Department of Biochemistry and Biophysics, The Arrhenius Laboratory, Stockholm
University, 106 91 Stockholm, Sweden

* To whom correspondence should be addressed. Email: lenam@dbb.su.se; Phone:
+46 8 162448; Fax: +46 8 155597

Running title

NMR investigations of the ThrRS dually targeting peptide

Abbreviations

dTP, dually targeting peptide; mTP, mitochondrial targeting peptide; cTP chloroplastic targeting peptide; aaRS, amino acyl-tRNA synthetase; ThrRS, threonyl tRNA synthetase; TOM, translocase of the outer mitochondrial membrane; TIM, translocase of the inner mitochondrial membrane; TOC, translocase of the outer envelope membrane of chloroplasts; TIC, translocase of the inner envelope membrane of chloroplasts; HSQC, heteronuclear single-quantum coherence; TOCSY, total correlation spectroscopy; NOESY, nuclear Overhauser enhancement spectroscopy, 2D, two-dimensional; 3D, three-dimensional; CLEANEX, clean chemical exchange.

Keywords

NMR, mitochondria, protein import, dual targeting peptide, structure

Database

The backbone chemical shift assignments for ThrRS-dTP(2-60) have been deposited with the Biological Magnetic Resonance Bank (BMRB) under the accession code 18248.

Summary

Most mitochondrial proteins are synthesized in the cytosol as precursor proteins containing an N-terminal targeting peptide and are imported into mitochondria through the import machineries, the Translocase of the Outer Mitochondrial Membrane (TOM) and the Translocase of the Inner Mitochondrial Membrane (TIM). The N-terminal targeting peptide of precursor proteins destined to the mitochondrial matrix is recognized by the Tom20 receptor and plays an important role in the import process. Protein import is usually organelle specific, but several plant proteins are dually targeted into mitochondria and chloroplasts using an ambiguous dual targeting peptide. We present NMR studies of the dual targeting peptide of Thr-tRNA synthetase and its interaction with Tom20 in *Arabidopsis thaliana*. Our findings show that the targeting peptide is mostly unstructured in buffer, with a propensity to form α -helical structure in one region, S6-F27, and a weak tendency for a β -strand structure in another region, Q34-Q38. The α -helical structured region has an amphiphilic character and a $\phi\chi\chi\phi\phi$ motif, both of which have previously been shown to be important for mitochondrial import. Using NMR we have mapped out two regions in the peptide that are important for Tom20 recognition, and one of them, F9-V28, overlaps with the amphiphilic region, and the other comprises residues L30-Q39. Our results show that the targeting peptide may interact with Tom20 in several ways. Furthermore, our results indicate a weak, dynamic interaction. The results provide for the first time molecular details on the interaction of the Tom20 receptor with a dual targeting peptide.

Introduction

Most mitochondrial and chloroplastic proteins are synthesized on cytosolic ribosomes as precursor proteins equipped with N-terminal targeting peptides (TPs) that function as recognition signals for organellar import. The basic properties of TPs, such as amino acid composition, charge and structural propensities, of mitochondrial targeting peptides (mTP) and chloroplastic targeting peptides (cTP) are very similar [1]. Both mTPs and cTPs contain many hydroxylated, hydrophobic and positively charged amino acid residues and only a few negatively charged amino acid residues [1-3]. However, there are also some differences including length, amino acid distribution and structural properties. The mTPs are on average 42 amino acid residues long and contain more arginines especially in the N-terminal portion of the mTPs, while the cTPs are on average 58 amino acid residues long and contain more serines and prolines [1,4]. Structurally, most mTPs are predicted to form an amphiphilic α -helix, which has been shown to play an important role in mitochondrial import [5], while cTPs appear to be unstructured in most studies [6-9]. Nevertheless, some cTPs have been reported to form helical structures in a membrane mimetic environment [9-11].

Protein import is generally very specific, however, and in plants there is a growing number of proteins that are products of a single gene, but can be imported into both mitochondria and chloroplasts, i.e. proteins that are dually targeted to both organelles [3,12]. This is understandable as mitochondria and chloroplasts share several overlapping functions including DNA replication, transcription, translation, and protection against oxidative stress, which implies that several enzymes such as e.g. the aminoacyl-tRNA synthetases (aaRSs) are required in both organelles. There are dual targeting peptides (dTPs) in which alternative translation origin results in

separate organellar domains that control the import into mitochondria or chloroplasts [13]. More common, however, is the use of an ambiguous targeting sequence, which within the same sequence contains information for both mitochondrial and chloroplastic targeting [3,12,14], found in e.g. dTPs of glutathione reductase [15], RNA polymerase RpoT:2 [16] and PresequenceProtease [17] and many aaRSs [11].

A group of 18 aminoacyl-tRNA synthetases (aaRSs) identified in *Arabidopsis thaliana* are dually targeted to both mitochondria and chloroplasts by their dTPs [18]. The N-terminal truncation of prolyl-RS and aspartyl-RS dTPs abolished targeting to either chloroplasts or mitochondria, indicating the presence of a domain structure in these dTPs, i.e. different parts of the sequence is needed for import into the different organelles and do not overlap. However, for most aaRSs, including threonyl-tRNA synthetase dTP (ThrRS-dTP), the N-terminal portions of the dTPs are essential for import into both organelles [11]. Further studies of ThrRS-dTP using C-terminal truncations of different lengths revealed that the shortest sequence with dual targeting capacity was 60 amino acid residues long, ThrRS-dTP(2-60), although the predicted processing site suggested by ChloroP for chloroplasts was after 38 residues and by MitoProt for mitochondria after 35 residues [19]. In accordance with observation for other dTPs, ThrRS-dTP(2-60) has significantly more Ser residues in the N-terminal portion than usually found in mTPs. In addition, it also has three Ser residues at the C-terminus. Previous results by CD spectroscopy revealed that ThrRS-dTP(2-60) is mainly unstructured in aqueous environment [19]. However, the molecular mechanisms of how the ambiguous dTPs mediate import of dually targeted proteins to both organelles have not yet been elucidated.

mTPs and cTPs are recognized by different organellar import machineries. In mitochondria, the Translocase of the Outer Mitochondrial Membrane (TOM) and the

Translocase of the Inner Mitochondrial Membrane (TIM) [20-22] and in chloroplasts, the Translocase of the Outer envelope membrane of Chloroplasts (TOC) and the Translocase of the Inner envelope membrane of Chloroplasts (TIC) [23,24] mediate import of the unfolded precursor proteins.

Tom20 and Tom70 components of the mitochondrial TOM complex function as import receptors [22,25]. Tom20 is the receptor that directly interacts with the mTPs. Although the Tom20 receptors in different organisms (plants, animals, and yeast) are non-homologous, they have equivalent structural features and function [26-28]. The topology of Tom20 is also different depending on species, however in plants the C-terminus is anchored to the outer membrane by a single transmembrane segment, and the hydrophilic N-terminal cytosolic domain recognizes the presequences in an early step of the mitochondrial protein import. The cytosolic domains of Tom20s from both rat and *Arabidopsis thaliana* have an α -helical structure with a groove at the concave surface [26,27]. In *Arabidopsis thaliana*, there are 4 Tom20 isoforms with sequence identities varying from 35% to 60% [29,30]. The interaction between Tom20 and mTPs has been suggested to occur between a hydrophobic patch in the Tom20 groove and hydrophobic residues in mTPs [27]. The presence of amphiphilic helices in mTPs and the extended structure of cTPs raises the question how the properties of the ambiguous dTPs of proteins that are imported into both mitochondria and chloroplasts are recognized by the organellar receptors, such as Tom20.

In this study we wished to elucidate the molecular properties of the dually targeting peptide from threonyl-tRNA synthetase (ThrRS-dTP), and we have for the first time studied the interaction between a dual targeting peptide and Tom20. We have used different nuclear magnetic resonance (NMR) methods for structural characterization of ThrRS-dTP(2-60). Based on backbone assignments, we have

analyzed the structural features of this peptide in aqueous solution. Furthermore, we have investigated the key Tom20 interaction points in ThrRS-dTP(2-60), in the presence of the cytosolic domain of *At*Tom20. To our knowledge, this is the first time that the initial interaction between the mitochondrial import receptor Tom20 and a dual targeting peptide has been studied.

Results

Backbone assignment of ThrRS-dTP(2-60)

In order to perform secondary structure analysis of a dTP and its binding to Tom20, we have purified the shortest peptide, which has dual targeting capacity, ThrRS-dTP(2-60), by overexpressing the peptide as a GST construct in M9 minimal media containing ^{15}N and then isolating the ThrRS-dTP(2-60) as previously described by Berglund et al. [19]. The structural properties of this sequence and its interaction with *At*Tom20 were studied using NMR spectroscopy.

Figure 1 shows the assigned ^{15}N -HSQC spectrum of the purified ThrRS-dTP(2-60) at pH 6. The narrow chemical shift dispersion in the proton dimension indicates that ThrRS-dTP(2-60) is largely unstructured under these conditions, which has previously been demonstrated by CD spectroscopy [19]. ThrRS-dTP(2-60) is composed of 59 amino acid residues and among them five are Pro residues and a very large number are Ser residues (25% of the sequence). The many sequential Ser residues made the assignment more complicated, but nevertheless, using 3D data, ^{15}N -TOCSY-HSQC, ^{15}N -NOESY-HSQC, HNCA and HN(CO)CA, it was possible to assign 47 out of 53 expected backbone cross-peaks (89%), and the corresponding C^α resonances. Although 4 additional peaks in the ^{15}N -HSQC were assigned to Ser

residues, Ser3, Ser4, Ser11, Ser18 and Ser19 remained unassigned, since they were not possible to unambiguously separate. Furthermore, resonances belonging to His5 were not found in the spectra, possibly due to an unfavorable equilibrium between the protonated and non-protonated forms at pH 6.

Analysis of the secondary structure of ThrRS-dTP(2-60)

The chemical shifts of proteins and peptides are routinely used for characterizing secondary structure, since they are very sensitive to local environment. For intrinsically disordered proteins, there is no fully formed α -structure or β -structure, but they possess transient structure and structural propensities to different degrees [31,32]. Marsh et al. have developed a method to calculate the secondary structure propensity (SSP) score at a given position based on different chemical shifts [33]. Here, the SSP scores for individual residues in ThrRS-dTP have been analysed with C^α and H^α chemical shifts as inputs (Figure 2).

In disordered proteins, a SSP score for a given residue reflects the fraction of a certain conformer in the disordered ensemble at that position. Positive values indicate α -helical structure and negative values indicate β -strand structure. In the ThrRS-dTP peptide, residues S6-F27, L30, R33, Q39-F47, V49 and S56 have an α -helical population, whereas residues V28-Y29, T32, Q34-R38, A48, A50-I55 and S57-P60 have a (weak) β -strand population. Among them, only one segment with more than two residues in a row has a population of helical structure greater than 20% and this is S6-F27, with F20-L24 having the highest scores (Figure 2). However, two consecutive residues F35 and W36 in the Q34-R38 segment also have a population above 20% of β -sheet. The result is similar when using a different algorithm, $\delta 2D$ [34]. Predictions by Jpred [35] indicated that helical structure may be formed between

T21-V28. Hence, we conclude that the sequence has a tendency to form helical structure in the N-terminal part, S6–F27, and that weak propensities for alternating β -structure (Q34-R38) and α -helix (Q39-F47) is observed in the remaining part of the sequence.

Since earlier studies have indicated that an amphiphilic helix may be of importance for targeting peptide-Tom20 interaction, we evaluated the amphiphilic helix propensities for ThrRS-dTP(2-60) using HeliQuest [36]. The whole sequence of ThrRS-dTP(2-60) was scanned through a sliding window of 18 residues, and hydrophobic faces were found for L7-P37, especially so for S11-V28 and F20-P37, where there were 6 residues involved in the hydrophobic faces (Figure 3). One of these segments, S11-V28 showed a characteristic amphiphilic character, while the other one did not. This region corresponds to the sequence that had the highest propensity for α -helix in our secondary chemical shift analysis. The second region, F20-P37 does not correspond to a unique structural propensity as judged from chemical shifts. Notably, Y29 was always separated from the hydrophobic face by at least four residues.

To further study the induced structure, CLEANEX experiments with mixing times of 20 ms, 40 ms, 60 ms and 80 ms were acquired to monitor the water-amide proton exchange in ThrRS-dTP(2-60) in aqueous solution. Only the exchange between the H^N-proton of an unassigned Ser residue and water was fast enough to manifest itself already at the shortest mixing time (Figure 4A). However, in the spectrum with a 40 ms mixing time, amide protons from 17 additional amino acid residues had a significant exchange with water (Figure 4B). Notably, the C-terminus (I55, S56, S57 and S58) exchanges relatively fast. The other peaks that appear in the spectrum originate from F9, S10, S12, S15, F20, S22-S23, R25, Q34, F43-S44-T45,

E52 and an unassigned Ser. Increasing the mixing time to 60 ms results in the appearance of cross-peaks for additional 18 amino acid residues that indicate exchange: S6, F13, L14, K16, L24, R26, F27, T32, R33, R38-Q39-R40, V46, A48, A50, T51 and G59 and one unassigned Ser residues (Figure 4C).

By comparing a summary of the measured peak intensities in the CLEANEX spectra as a function of mixing time (not shown) with the secondary structural propensity (Figure 2) it is evident that the residues that appear to display somewhat slower exchange are not entirely correlated with the regions that are identified to have a structural propensity. Hence, the results indicate again that the peptide is largely unstructured and that the tendency to form helical structure is very weak.

Mapping the *At*Tom20 interaction sites in ThrRS-dTP(2-60)

The interaction between *At*Tom20 and ThrRS-dTP(2-60) was studied by titrating a ThrRS-dTP(2-60) sample (100 μ M) with a concentrated solution of overexpressed *At*Tom20 (1 mM) and recording 15 N-HSQC spectra with different *At*Tom20 concentrations ranging from 0 to 1:1 molar ratio. The resulting spectra are shown in Figure 5. Adding *At*Tom20 induced chemical shift changes and intensity loss for some cross-peaks, reflecting the interaction positions in the ThrRS-dTP(2-60) sequence. Generally, the largest effects on the cross-peaks are related to broadening and intensity loss, rather than large chemical shift changes. Some peaks disappear even before the highest *At*Tom20 concentration (1:1), indicating a slower overall tumbling rate or exchange broadening due to an on-off binding event. However, peaks from the C-terminus and L7, L8 and Y29 are still clearly seen at the highest *At*Tom20 concentration (Figure 5C). Hence, the NMR spectral properties of these residues in

ThrRS-dTP(2-60) are not largely affected by *At*Tom20 and most likely do not take part in the interaction.

The intensity changes relative to the sample without *At*Tom20 were calculated for samples with *At*Tom20/ThrRS-dTP(2-60) molar ratios of 0.2:1 and 0.5:1 (Figure 6A). Pronounced changes upon addition of *At*Tom20 were observed mainly in the F9-Q39 region, involving residues F9-K16, S22-V28 and L30-Q39. These regions correspond well with the parts of the sequence seen to have high secondary structure propensity, including both the α -helical structure S6-F27 and the β -strand structure Q34-R38. The signal intensities for these positions decrease to less than half at a molar ratio of 0.2:1 and most of them diminish in intensity even more at the next titration point (0.5:1). These results suggest a loss of mobility in these three regions upon addition of *At*Tom20. We can conclude that these positions are probably where *At*Tom20 binds to ThrRS-dTP(2-60). In addition, other residues, F47 and A48 also show obvious intensity loss, although they do not constitute a specific region. Notably, residues in the C-terminus of ThrRS-dTP(2-60) are not greatly affected by *At*Tom20 (V49-G59), and neither are the N-terminal residues (S6-L8).

Further, chemical shift changes between the spectra recorded with and without 0.2:1 molar ratio of *At*Tom20/ThrRS-dTP(2-60) were calculated. These chemical shift differences for the $^1\text{H}^{\text{N}}$ and $^{15}\text{N}^{\text{H}}$ resonances are shown in Figure 6B. Several residues display small but significant shift changes, and the ones displaying the largest shifts are L14, S15, R25, F27-V28 L30, T32, W36, and F47. These residues are mainly located in the first N-terminal part of the peptide that showed a propensity for α -helical structure. W36, on the other hand, is in the short sequence that showed a weak β -sheet propensity. Additional residues also have larger than average shift changes, such as for F47. The chemical shift changes after addition of *At*Tom20 indicate that a

small structural change of ThrRS-dTP(2-60) occurs and it is mainly in the weakly structured regions (F9-V28 and L30-Q39).

A summary of the line-broadening and chemical shift effects of adding *At*Tom20 to ThrRS-dTP(2-60) is shown in Figure 6C, and notably, taking the results together, the most pronounced effects are seen for the region comprising residues F9-V28 and L30-Q39, which agrees very well with the regions identified in the secondary chemical shift analysis as having propensity for forming α -helical or β -strand structure. The first of these regions was also identified by helical wheel calculations as being able to form an amphiphilic structure, although both regions present hydrophobic faces (Figure 3). It is also worth noting that the main effect of adding *At*Tom20 is to increase line-widths rather than inducing chemical shift differences, indicating that the interaction causes dynamic events that increase the apparent T_2 relaxation rates (presumably exchange).

Discussion

Several studies have to date identified and analyzed apparent dTPs in precursor proteins, which are being imported into both mitochondria and chloroplasts [12,18,37] and many more have been predicted to be dually targeted [38]. The determinants for import into mitochondria and chloroplasts appear to be very similar, although some differences in the characteristics of mTPs and cTPs have been found, especially in the N-terminal portion of the targeting peptides [1]. The content of Pro and Ser residues is higher in cTPs, while the prevalence of basic amino acid residues is higher in mTPs. mTPs have been shown to form amphiphilic α -helices in membrane mimetic media [5,7], and it has been demonstrated that they bind with their hydrophobic side into a

hydrophobic groove in the Tom20 receptor [26,27]. An N-terminal α -helix has been observed in the targeting peptides in both alternative oxidase (AOX) [39] and in $F_1\beta$ [5] in SDS micelles, which also appears to be important for targeting. Contrary to this, cTPs appear not to form helical structures to the same degree [6,9] although helical structure has been observed in membrane mimicking media [10,40] and it has been suggested that the binding to the receptor does not require the formation of the an amphiphilic helical structure as observed in mTPs [9].

In the case of a chloroplastic transit peptide (cTP) from the Rubisco activase enzyme it was observed that a rather stable helix in the C-terminal part of the sequence was formed in mixed micelles [9]. However, the cTP of ferredoxin does not adopt any significant amount of helical structure although a small central region is observed to have helical features [40].

The nature of the structure of dTPs is not known. Here we observed that the dual targeting peptide from *At*ThrRS is largely unstructured in aqueous solution, but does show a propensity of forming helical structure in the N-terminal part of the sequence, which is followed by a short region with a weak propensity for β -strand structure. The exchange data also clearly supports an unstructured peptide, with no regions clearly displaying slower NH proton exchange than others. The properties of the N-terminal region of the dTP from *At*ThrRS, which has the capacity to form an amphiphilic structure (Figure 3) corresponds well to regions in mitochondrial targeting sequences that have been implicated in the binding to a hydrophobic patch in the Tom20 receptor.

Here we see that the interactions between ThrRS-dTP(2-60) and Tom20 appear to involve more than one region, however. The largest part that is affected by Tom20 spans residues S11-V28, but also L30-Q39 shows chemical shift changes and line

broadening upon Tom20 interaction. In previous studies of the interaction between the AOX-mTP and Tom20, chemical shift changes in two regions were also observed, and it was concluded that extensive regions in the targeting peptide were involved in binding [39]. Furthermore, there are two discontinuous binding sites at the concave surface of Tom20, which is suggested to explain that the interaction region for plant presequences in general being longer than the animal counterpart.

Here we see that the dually targeted peptide also appears to bind to Tom20 using a large part of the sequence, through two different regions. Hence, the interaction between the dually targeted peptide in this study may be similar in nature as that observed in a mitochondrial targeting peptide. Based on the present findings and other studies, it appears likely that there are at least two different ways that a dual targeting peptide can be recognized by Tom20. From our results we see that out of the two regions that are affected by Tom20 one, S11-V28, has the possibility to adopt an amphiphilic structure, and that mainly residues in the hydrophobic face of the putative helix are affected by Tom20. Notably, the hydrophobic face of the first region is flanked in both ends by two basic amino acid residues, K16 and R25 (Figure 3). It has previously been demonstrated that such an arrangement may serve as an anchor in positioning transmembrane helices in a lipid bilayer [41], and here these residues may be important for locating the sequence in a hydrophobic pocket of Tom20. Both of these basic residues are among the ones that are most affected by Tom20. No such arrangement is observed for the second region of the peptide (L30-Q39).

The largest interaction domain of ThrRS-dTP(2-60) (S11-V28) contains two Arg and one Lys residue, and no negatively charged residues, and the second short region (L30-Q39) has two Arg residues, implying that these positively charged residues may

be of importance in addition to hydrophobic interactions. The interaction between mTPs and Tom20 from both yeast and plants is reported to be hydrophobic in nature, and it has even been pointed out that the hydrophobic residues of the amphiphilic structure and not the positively charged residues that are recognized by Tom20 [27]. The plant F₁β mTP, the AOX mTP and the rat mTP of aldehyde dehydrogenase all contain positively charged residues [5,26,39], and although it has been observed that the peptides bind with a hydrophobic surface, it has also been pointed out that the Arg residues may form contacts with polar residues in Tom20 [26]. Muto et al. (2001) identified five-residue binding regions in different mTPs that all contain Arg residues, although the pattern did not strictly require a basic amino acid residue [42,43]. The pattern of this sequence is described as $\phi\chi\chi\phi\phi$, where ϕ is a hydrophobic/aromatic amino acid residue and χ is often a polar residue. In our case, the sequence L₂₄RRFV₂₈ is the only one in the ThrRS-dTP sequence that fulfils this pattern, and notably, these residues are among the ones that were most affected by the addition of Tom20 (Figure 6).

The main effect of Tom20 on the targeting peptide is to cause selective line-broadening for certain residues. This effect is much more pronounced than the chemical shift perturbations even at low concentrations of Tom20, and several peaks diminish greatly at a molar concentration ratio of ThrRS-dTP(2-60) to Tom20 of 1:0.2 (Figure 6). This clearly indicates that the complex between the two is dynamic and causes exchange broadening to interfere with the NMR spectrum. Even at a high Tom20 concentration (1:1), the spectrum is broadened beyond detection for most residues, which is consistent with a dynamic equilibrium that cannot be stabilized by excess Tom20. This result is reasonable, since it is not expected that the targeting peptide should bind tightly to the receptor, as this would have impact on the import

efficiency. The lack of large shift changes may also indicate that there is no need for a large structural conversion of the targeting peptide, in contrast to rat Tom20 – presequence interactions, where a well-defined, stable helical presequence peptide has been observed [26]. Increasing amounts of evidence that intrinsically disordered proteins do not have to adopt a well-defined structure upon interaction with other proteins are emerging [44-47]. Sometimes local and transient structural preferences in an ensemble of peptides dictate function, and even protein-protein interactions may not always lead to well-formed secondary or tertiary structures, but indicate a novel mode of action of these intrinsically disordered proteins [32,45,46].

In summary, we have for the first time characterized the interaction between a dual targeting peptide and the mitochondrial import receptor Tom20. The dTP from *At*ThrRS was investigated by solution NMR, and we found that the peptide is mainly unstructured, but with tendencies to form an amphipathic α -helical structure in the N-terminal part of the sequence and a weak tendency for a β -strand structure in the midpart of the sequence, a trait characteristic for intrinsically disordered proteins (IDPs) [31,32,47]. The dual targeting peptide does not appear to bind rigidly to Tom20, but is in a kinetic equilibrium, leading to line-broadening effects in NMR spectra. The two regions identified as having structural propensities are the ones mainly responsible for the binding. This implies that the interaction with Tom20 may share features with the mTP – Tom20 interaction, but that also other parts of the targeting sequence take part in the interaction.

Materials and methods

Sample preparation

Cloning of the constructs of ThrRS-dTP and purification of ThrRS-dTP were done as described previously by Berglund et al. [19]. In short, the peptide was produced by overexpressing a GST-ThrRS-dTP fusion construct. The inclusion bodies containing the insoluble protein were purified and ThrRS-dTP(2-60) was obtained by CNBr cleavage after the first Met in ThrRS-dTP as described earlier [5,19]. ^{15}N labeled ThrRS-dTP was overexpressed in ^{15}N -enriched minimal culture medium. Dialyzed peptide was lyophilized and dissolved in a solution containing 50 mM sodium phosphate buffer (pH 6.0) and 10% (v/v) D_2O . The final concentration of the peptide was 80 - 100 μM .

AfTom20 $\Delta(159-187)$ (isoform Tom20-4) was overexpressed as a GST fusion protein in Luria-Bertani medium and grown at 37°C to a density of A600 between 0.5 and 0.6. Expression of the fusion protein was achieved by induction with IPTG at a final concentration of 1 mM for 5 h, at 37°C. *AfTom20* $\Delta(159-187)$ was purified by GSTrap (GE Healthcare) followed by ion exchange chromatography on a Resource QTM column (GE Healthcare). The purified protein was dialyzed against 20mM sodium phosphate buffer (pH 6.0), lyophilized and dissolved in H_2O to the concentration of 1mM in 100mM sodium phosphate buffer (pH 6.0).

NMR spectroscopy

NMR experiments for chemical shift assignments for ThrRS-dTP(2-60) as well as water-amide proton exchange experiments were collected on a Bruker Avance spectrometer operating at a ^1H frequency of 600 MHz, equipped with a 5 mm triple resonance, pulsed-field z-gradient probe head, using samples containing either ^{15}N - or ^{15}N , ^{13}C -labeled peptide (80 – 100 μM) in 50 mM phosphate buffer. The assignments were based on a combination of 2D ^{15}N HSQC, a 3D ^{15}N edited NOESY-HSQC [48-

50], two ^{15}N edited TOCSY-HSQC [48], a HNCA [51-53] and a HN(CO)CA [51,54]. The NOESY-HSQC spectrum was recorded with a mixing time of 200 ms and the TOCSY-HSQC spectra were recorded with mixing times of 40 ms and 80 ms, respectively and collected as 2048 x 56 x 256 data point matrices using 8 scans. HNCA and HN(CO)CA were collected as 2048 x 56 x 128 data point matrices using 16 scans. For water-amide proton exchange studies a Fast-HSQC with a CLEANEX-PM filter was used [55].

For Tom20 interaction studies, *At*Tom20 was titrated into a ThrRS-dTP(2-60) sample. 2D ^{15}N - ^1H HSQC spectra were recorded on a 700 MHz Bruker Avance spectrometer, equipped with a 5 mm triple resonance, pulsed-field z-gradient, cryogenically cooled probe. The titration was started by recording a 2D ^{15}N HSQC spectrum of 400 μl 100 μM ^{15}N -labeled ThrRS-dTP sample in the absence of *At*Tom20. Later, 8 μl , 12 μl and 20 μl of 1 mM, unlabeled *At*Tom20 with the same buffer condition were added to the ThrRS-dTP sample in steps, giving molar ratios of *At*Tom20:ThrRS-dTP of 0:1, 0.2:1, 0.5:1 and 1:1. The overall volume change using this procedure was less than 10%. 2D ^{15}N HSQC spectra were recorded for all the steps. All NMR experiments were carried out at 298 K. The data were processed with TopSpin 2.1 and analyzed with CCPN [56] and Sparky 3 [57].

Web tools

SSP program [33] and $\delta 2\text{D}$ program [34] were applied to calculate the secondary structure propensities of the IDP ThrRS-dTP with $\text{C}\alpha$ and $\text{H}\alpha$ chemical shifts. HeliQuest [36] was used to analyze the helical properties of the sequence, and to investigate the potential amphiphilicity of these. Jpred [35] was used to predict the secondary structure propensity for the ThrRS-dTP(2-60) sequence.

Acknowledgements

This work was supported by grants from the Swedish Research Council (to LM and EG) and by a grant from the Carl Trygger Foundation.

References

- 1 Bhushan S, Kuhn C, Berglund A, Roth C & Glaser E (2006) The role of the N-terminal domain of chloroplast targeting peptides in organellar protein import and miss-sorting. *FEBS Lett* **580**, 3966-3972.
- 2 Glaser E & Soll J (2004) Targeting signals and import machinery of plastids and plant mitochondria. In *Molecular Biology and Biotechnology of Plant Organelles Chloroplasts and Mitochondria* (Daniell, H. and Chase, C., eds.) pp. 385-417. Springer, Netherlands.
- 3 Peeters N & Small I (2001) Dual targeting to mitochondria and chloroplasts. *Biochim Biophys Acta* **1541**, 54-63.
- 4 Zhang XP & Glaser E (2002) Interaction of plant mitochondrial and chloroplast signal peptides with the Hsp70 molecular chaperone. *Trends Plant Sci* **7**, 14-21.
- 5 Moberg P, Nilsson S, Stahl A, Eriksson A, Glaser E & Mäler L (2004) NMR Solution Structure of the Mitochondrial F1beta Presequence from *Nicotiana plumbaginifolia* . *J Mol Biol* **336**, 1129-1140.
- 6 von Heijne, G & Nishikawa, K (1991) Chloroplast transit peptides. The perfect random coil? *FEBS Lett* **278**, 1-3.
- 7 Lancelin J, Gans P, Bouchayer E, Bally I, Arlaud GJ & Jacquot J (1996) NMR structures of a mitochondrial transit peptide from the green alga *Chlamydomonas reinhardtii* . *FEBS Lett* **391**, 203-208.

- 8 Roise D, Theiler F, Horvath SJ, Tomich JM, Richards JH, Allison DS & Schatz G (1988) Amphiphilicity is essential for mitochondrial presequence function. *EMBO J* **7**, 649-653.
- 9 Krimm I, Gans P, Hernandez J, Arlaud GJ & Lancelin, J (1999) A coil–helix instead of a helix–coil motif can be induced in a chloroplast transit peptide from *Chlamydomonas reinhardtii*. *Eur J Biochem* **265**, 171-180.
- 10 Bruce, BD (2000) Chloroplast transit peptides: structure, function and evolution. *Trends Cell Biol* **10**, 440-447.
- 11 Berglund, A, Pujol C, Duchene AM & Glaser E (2009) Defining the determinants for dual targeting of amino acyl-tRNA synthetases to mitochondria and chloroplasts. *J Mol Biol* **393**, 803-814.
- 12 Silva-Filho MC (2003) One ticket for multiple destinations: dual targeting of proteins to distinct subcellular locations. *Curr Opin Plant Biol* **6**, 589-595.
- 13 Chabregas SM, Luche DD, Van Sluys M, Menck CFM & Silva-Filho MC (2003) Differential usage of two in-frame translational start codons regulates subcellular localization of *Arabidopsis thaliana* THI1. *J Cell Sci* **116**, 285-291.
- 14 Karniely S & Pines O (2005) Single translation—dual destination: mechanisms of dual protein targeting in eukaryotes. *EMBO Rep* **6**, 420-425.
- 15 Creissen G, Reynolds H, Xue Y & Mullineaux P (1995) Simultaneous targeting of pea glutathione reductase and of a bacterial fusion protein to chloroplasts and mitochondria in transgenic tobacco. *Plant J* **8**, 167-175.

- 16 Hedtke B, Börner T & Weihe A (2000) One RNA polymerase serving two genomes. *EMBO Rep* **1**, 435-440.
- 17 Bhushan S, Lefebvre B, Ståhl A, Wright SJ, Bruce BD, Boutry M & Glaser E (2003) Dual targeting and function of a protease in mitochondria and chloroplasts. *EMBO Rep* **4**, 1073-1077.
- 18 Duchêne A, Giritch A, Hoffmann B, Cognat V, Lancelin D, Peeters NM, Zaepfel M, Maréchal-Drouard L & Small ID (2005) Dual targeting is the rule for organellar aminoacyl-tRNA synthetases in *Arabidopsis thaliana*. *Proc Natl Acad Sci USA* **102**, 16484-16489.
- 19 Berglund A, Spånnig E, Biverståhl H, Maddalo G, Tellgren-Roth C, Mäler, L & Glaser E (2009) Dual targeting to mitochondria and chloroplasts: characterization of Thr-tRNA synthetase targeting peptide. *Mol Plant* **2**, 1298-1309.
- 20 Bolender N, Sickmann A, Wagner R, Meisinger C & Pfanner N (2008) Multiple pathways for sorting mitochondrial precursor proteins. *EMBO Rep* **9**, 42-49.
- 21 Lister R, Hulett JM, Lithgow T & Whelan J (2005) Protein import into mitochondria: origins and functions today. *Mol Membr Biol* **22**, 87-100.
- 22 Neupert W & Herrmann JM (2007) Translocation of proteins into mitochondria. *Annu Rev Biochem* **76**, 723-749.
- 23 Benz JP, Soll J & Bölter B (2009) Protein transport in organelles: the composition, function and regulation of the Tic complex in chloroplast protein import. *FEBS J* **276**, 1166-1176.

- 24 Jarvis P (2008) Targeting of nucleus - encoded proteins to chloroplasts in plants. *New Phytol* **179**, 257-285.
- 25 Lithgow T, Glick BS & Schatz G (1995) The protein import receptor of mitochondria. *Trends Biochem Sci* **20**, 98-101.
- 26 Abe Y, Shodai T, Muto T, Mihara K, Torii H, Nishikawa S, Endo T & Kohda D (2000) Structural basis of presequence recognition by the mitochondrial protein import receptor Tom20. *Cell* **100**, 551-560.
- 27 Perry AJ, Hulett JM, Likic VA, Lithgow T & Gooley PR (2006) Convergent evolution of receptors for protein import into mitochondria. *Curr Biol* **16**, 221-229.
- 28 Brix J, Dietmeier K & Pfanner N (1997) Differential recognition of preproteins by the purified cytosolic domains of the mitochondrial import receptors Tom20, Tom22, and Tom70. *J Biol Chem* **272**, 20730-20735.
- 29 Werhahn W, Niemeyer A, Jansch L, Kruft V, Schmitz UK & Braun H (2001) Purification and characterization of the preprotein translocase of the outer mitochondrial membrane from *Arabidopsis*. Identification of multiple forms of TOM20. *Plant Physiol* **125**, 943-954.
- 30 Werhahn W, Jansch L & Braun H (2003) Identification of novel subunits of the TOM complex from *Arabidopsis thaliana*. *Plant Physiol Biochem* **41**, 407-416.
- 31 Dyson HJ & Wright PE (2005) Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Bio* **6**, 197-208.

- 32 Uversky VN (2011) Intrinsically disordered proteins from A to Z. *Int J Biochem Cell Biol* **43**, 1090-1103.
- 33 Marsh JA, Singh VK, Jia Z & Forman-Kay JD (2006) Sensitivity of secondary structure propensities to sequence differences between α - and γ -synuclein: Implications for fibrillation. *Protein sci* **15**, 2795-2804.
- 34 Camilloni C, De Simone A, Vranken WF & Vendruscolo M (2012) Determination of Secondary Structure Populations in Disordered States of Proteins Using Nuclear Magnetic Resonance Chemical Shifts. *Biochemistry* **51**, 2224-2231.
- 35 Cuff JA, Clamp ME, Siddiqui AS, Finlay M & Barton GJ (1998) JPred: a consensus secondary structure prediction server. *Bioinformatics* **14**, 892-893.
- 36 Gautier R, Douguet D, Antonny B & Drin G (2008) HELIQUEST: a web server to screen sequences with specific α -helical properties. *Bioinformatics* **24**, 2101-2102.
- 37 Carrie C, Giraud E & Whelan J (2009) Protein transport in organelles: dual targeting of proteins to mitochondria and chloroplasts. *FEBS J* **276**, 1187-1195.
- 38 Mitschke J, Fuss J, Blum T, Höglund A, Reski R, Kohlbacher O & Rensing SA (2009) Prediction of dual protein targeting to plant organelles. *New Phytol* **183**, 224-236.
- 39 Rimmer KA, Foo JH, Ng A, Petrie EJ, Shilling PJ, Perry AJ, Mertens HDT, Lithgow T, Mulhern TD & Gooley PR (2011) Recognition of mitochondrial

- targeting sequences by the import receptors Tom20 and Tom22. *J Mol Biol* **405**, 804-818.
- 40 Wienk HLJ, Wechselberger RW, Czisch M & de Kruijff B (2000) Structure, dynamics, and insertion of a chloroplast targeting peptide in mixed micelles. *Biochemistry* **39**, 8219-8227.
- 41 Unnerstle S, Mler L & Draheim RR (2011) Structural characterization of AS1-membrane interactions from a subset of HAMP domains. *Biochim Biophys Acta* **1808**, 2403-2412
- 42 Muto T, Obita T, Abe Y, Shodai T, Endo T & Kohda D (2001) NMR identification of the Tom20 binding segment in mitochondrial presequences. *J Mol Biol* **306**, 137-143.
- 43 Endo T & Kohda, D (2002) Functions of outer membrane receptors in mitochondrial protein import. *Biochim Biophys Acta* **1592**, 3-14.
- 44 Tompa P (2011) Unstructural biology coming of age. *Curr Opin Struc Biol* **21**, 419-425.
- 45 Mittag T, Marsh J., Grishaev A, Orlicky S, Lin H, Sicheri F, Tyers M, & Forman-Kay JD (2010) Structure/function implications in a dynamic complex of the intrinsically disordered Sic1 with the Cdc4 subunit of an SCF ubiquitin ligase. *Structure* **18**, 494-506.
- 46 Sigalov AB (2011) Uncoupled binding and folding of immune signaling-related intrinsically disordered proteins. *Prog Biophys Mol Biol* **106**, 525-536.

- 47 Tompa P (2002) Intrinsically unstructured proteins. *Trends Biochem Sci* **27**, 527-533.
- 48 Marion D, Driscoll PC, Kay LE, Wingfield PT, Bax A, Gronenborn AM & Clore GM (1989) Overcoming the overlap problem in the assignment of proton NMR spectra of larger proteins by use of three-dimensional heteronuclear proton-nitrogen-15 Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: application to interleukin 1. beta. *Biochemistry* **28**, 6150-6156.
- 49 Marion D, Kay LE, Sparks SW, Torchia DA & Bax A (1989) Three-dimensional heteronuclear NMR of nitrogen-15 labeled proteins. *J Am Chem Soc* **111**, 1515-1517.
- 50 Zuiderweg ERP & Fesik SW (1989) Heteronuclear three-dimensional NMR spectroscopy of the inflammatory protein C5a. *Biochemistry* **28**, 2387-2391.
- 51 Grzesiek S & Bax A (1992) Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. *J Magn Reson* **96**, 432-440.
- 52 Kay LE, Ikura M, Tschudin R, & Bax A (1990) Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *J Magn Reson* **89**, 496-514.
- 53 Grzesiek S & Bax A (1992) An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins. *J Magn Reson* **99**, 201-207.

- 54 Bax A & Ikura M (1991) An efficient 3D NMR technique for correlating the proton and ^{15}N backbone amide resonances with the α -carbon of the preceding residue in uniformly $^{15}\text{N}/^{13}\text{C}$ enriched proteins. *J Biomol NMR* **1**, 99-104.
- 55 Hwang TL, van Zijl, PCM & Mori S (1998) Accurate quantitation of water-amide proton exchange rates using the Phase-Modulated CLEAN chemical EXchange (CLEANEX-PM) approach with a Fast-HSQC (FHSQC) detection scheme. *J Biomol NMR* **11**, 221-226.
- 56 Fogh R, Ionides J, Ulrich E, Boucher W, Vranken W, Linge JP, Habeck M, Rieping W, Bhat TN & Westbrook J (2002) The CCPN project: an interim report on a data model for the NMR community. *Nat Struct Biol* **9**, 416-418.
- 57 Goddard T & Kneller D (2004) SPARKY 3. *University of California, San Francisco* .

Figure legends

Figure 1. Assigned ^{15}N HSQC spectrum recorded at 700 MHz ^1H frequency of 100 μM ThrRS-dTP(2-60) in 50 mM sodium phosphate buffer at pH 6, 298 K.

Figure 2. Secondary structural propensity scores for ThrRS-dTP(2-60) calculated using C^α and H^α chemical shifts. Positive values represent the propensities for α -helical structure and negative values represent the propensities for β -structure.

Figure 3. Amphipathic analysis from an α -helical wheel representation of the sequences S11-V28 (A) and F20-P37 (B). The hydrophobic residues are colored orange, whereas the hydrophilic residues are colored green and the positively charged residues are colored blue.

Figure 4. H^N – water exchange monitored by CLEANEX experiments recorded at 600 MHz ^1H frequency, pH 6, 298 K with mixing times of A) a 20 ms, B) 40 ms, C) 60 ms, and D) a 80 ms, respectively.

Figure 5. ^{15}N HSQC spectra from the titration of *At*Tom20 onto 100 μM ^{15}N -ThrRS-dTP(2-60) recorded at 700 MHz ^1H frequency, pH 6, 298 K. A) An overlay of a spectrum recorded for ThrRS-dTP(2-60) without *At*Tom20 (red) and a spectrum recorded for a 0.2:1 mixture of *At*Tom20 and ThrRS-dTP (green). B) shows a spectrum recorded for a 0.5:1 mixture and C) for a 1:1 mixture of *At*Tom20 and ThrRS-dTP. All spectra were recorded under identical conditions and are shown with the same display parameters.

Figure 6. Effects on the ThrRS-dTP(2-60) ^{15}N -HSQC spectrum upon addition of *At*Tom20. A) The relative change in signal intensity of cross-peaks in ^{15}N HSQC spectra upon titration. The bars indicate the relative change in signal intensity compared to intensities in a spectrum recorded for a 100 μM ThrRS-dTP(2-60) sample without *At*Tom20. Grey indicates a 1:0.2 molar ratio and black indicates a 1:0.5 molar ratio of ThrRS-dTP(2-60) and *At*Tom20. B) H^{N} and N^{H} chemical shift differences in the spectrum of 100 μM ThrRS-dTP(2-60) without *At*Tom20 and with 20 μM *At*Tom20. The shifts are reported as $\delta_{\text{with}} - \delta_{\text{without}}$. The cut-off limits for significant chemical shift differences, chosen as $\Delta\delta\text{H}^{\text{N}} > 0.01$ and $\Delta\delta\text{N}^{\text{H}} > 0.2$ are shown. Residues with shift perturbations larger than this cut-off are indicated by their assignment. C) Summary of the effect of adding *At*Tom20 to ThrRS-dTP(2-60) when the molar ratio is 1:0.2. The positions with a significant change in chemical shifts by *At*Tom20 addition have letters in red. The $\phi\chi\chi\phi\phi$ motif, previously reported as a Tom20 recognition element, is indicated by the red frame. The positions that show a decrease in the signal intensity ratio of more than 0.5 are highlighted yellow. Letters in grey are prolines or unassigned residues.