Research Article

Journal of PeptideScience

Received: 7 March 2011 Revised: 20 April 2011 Accepted: 20 April 2011 Published online in Wiley Online Library: 10 June 2011

(wileyonlinelibrary.com) DOI 10.1002/psc.1383

Chemical synthesis of the third WW domain of TCERG 1 by native chemical ligation

Zerrin Fidan,[a] Aylin Younis,[a] Peter Schmieder[b] and Rudolf Volkmer[a]*

The human transcription elongation regulator 1 (TCERG1), with its modular architecture of 3 WW and 6 FF domains, inhibits RNA polymerase II elongation through these WW and FF domains, and interacts with pre-mRNA splicing factors such as SF1, U2 snRNP and U2AF. WW domains are known as the smallest naturally occurring, monomeric, triple stranded, anti-parallel β-sheet structures, generally spanning only about 40 amino acids. The first and second WW domains of TCERG1 have been synthesized and successfully applied for screening cellular targets. In contrast, until now syntheses of the third WW domain yielded oligopeptides with an undefined fold, proving useless for screening cellular targets. This implied that sequence elongation to include the α-helical structure is crucial for proper folding. We describe here the chemical synthesis of such a 53 residue long TCERG1 WW-3 domain sequence that exhibits the typical WW domain fold, and could be useful for discovering cellular targets.

Supporting information may be found in the online version of this article

Keywords: native chemical ligation; WW domain; human CA150 transcription factor; human transcription elongation regulator 1; TCERG 1; CD spectroscopy

Introduction

In general, cellular processes are mediated *via* protein–protein interactions, and a considerable number of protein interactions exploit families of relatively small protein interaction domains (PID), such as SH2, SH3, PDZ or WW domains, that act as receptors to accommodate short linear sequences of partner proteins in their binding pockets. One of these modules, ~30–40 amino acid WW domain, specifically recognizes ligands containing proline-rich motifs [1–3]. WW domains adopt three-stranded anti-parallel β-sheet structures, and in most cases are capable of folding autonomously [4–6]. This β-sheet structure comprises a concave face containing the ligand-binding site, and a convex face contributing to a hydrophobic core involving the termini of the domain. These features have made WW domains popular candidates for studying structure–function relationships. Fortunately, such investigations are simplified by the relatively short sequence length and reliable chemical synthesis protocols for successful solid-phase synthesis of WW domains [7–11].

The human transcription elongation regulator 1 (TCERG1, UniProtKB/Swiss-Prot: accession No. O14776), also called the TATA box-binding protein-associated factor 2S, or human transcription factor CA150, comprises 1098 amino acids and the fold is mainly formed by a modular architecture of 3 WW domains and 6 FF domains. This protein interacts with the pre-mRNA splicing factor SF1, and the phosphorylated *C*-terminal repeat domain (CTD) of RNA polymerase II (RNAP II) through its WW and FF domains [12–15]. The FF domains are essential for co-localization to nuclear speckles, but both the WW and FF domains are required to assemble a transcription/splicing complex [14]. More recently, cellular targets were identified for TCERG1, thus consolidating its role in mRNA processing and coupling transcription and processing [16,17]. Furthermore, TCERG1 is co-deposited with huntingtin protein in Huntington's disease [18], and has been proposed to participate in regulating the onset of neurodegeneration [19].

The WW-1 domain of TCERG1 was found to be a member of the phosphoserine/phosphothreonine (poS/poT) group [11], correlating with the observed direct binding to the phosphorylated CTD of RNAP II [13,20]. In addition, Goldstrohm *et al*. [13] revealed its role in recognizing the pre-mRNA splicing factor SF1.

The TCERG1 WW-1 domain was synthesized by standard Fmoc SPPS [11,21] and the product formed a typical WW domain fold. The sequence of the second WW domain of TCERG1 is identical to the murine FBP-28 WW domain, whose structure has already been solved [4] [Protein Data Bank (PDB) ID code 1E0L]. The domain belongs to the L-group and/or poly P group and recognizes PPPLIPPPP and PPLIPPPP motifs by a novel binding mode [8,10,11]. Similar to the WW-1 domain, the TCERG1 WW-2 domain specifically recognizes pre-mRNA splicing factors possessing proline-rich motifs such as SF1, U2 snRNP and U2AF [13,17]. For a long time

* *Correspondence to: Rudolf Volkmer, Institut für Medizinische Immunologie, Charité – Universitätsmedizin Berlin, Hessische Str. 3-4, 10115 Berlin, Germany. E-mail: rve@charite.de*

a *Institut für Medizinische Immunologie, Charité – Universitätsmedizin Berlin, Hessische Strasse 3-4, 10115 Berlin, Germany*

b *Leibniz-Institut für Molekulare Pharmakologie, Robert-Rössle-Strasse 10, 13125 Berlin, Germany*

Abbreviations used: *DCM; DIC; DMF; HOBt; MALDI-TOF ms; NMM; NMR; PDZ domain, (PSD-95, Dlg, ZO-1) domain; PyBOP; RP-HPLC; SH2 domain, Src Homology 2 domain; SH3 domain, SRC Homology 3 domain; WW domain, protein domain with two conserved tryptophans (W); TFA.*

Journal of PeptideScience

synthesis of the WW-2/FBP28 WW domain remained a challenge, but it was mastered by the Bienert lab in 2005 [22].

The third WW domain of TCERG1 was predicted to be a member of the (poS/poT) group [11], but unfortunately SPPS of the 43-mer domain (523–565) yielded an unstructured peptide. Consequently, our efforts to validate WW-3 domain ligand specificity failed [11]. Similarly, other efforts to detect cellular targets of the WW-3 domain of TCERG1 were also unsuccessful, most likely due to the unfolded 41-mer sequence (526–566) used in those studies [17]. Therefore, cellular targets of the TCERG1 WW-3 domain are as yet unknown. Recently, an NMR structure of the TCERG1 WW-3 domain, deposited in the PDB (PDB ID code 2dk7), was derived from an elongated 73-mer sequence of the TCERG1 WW-3 domain. This sequence elongation seems to be crucial for a properly folded TCERG1 WW-3 domain.

We surmised that the *C*-terminal short α-helical region (567–573) stabilizes the TCERG1 WW-3 domain fold. Therefore, in contrast to the 73-mer sequence, we predicted that a shorter TCERG1 sequence truncated just beyond the α-helical region would also form a typical WW domain fold. Here, we not only describe the first successful chemical synthesis of a 53-mer TCERG1 WW-3 domain but using biophysical analyses also demonstrate that this domain is folded correctly.

Materials and Methods

General Remarks

Fmoc amino acids and coupling reagents (PyBOP) were obtained from Novabiochem (Schwalbach, Germany). Pseudoprolines were purchased from Jupiter Bioscience AG (Laufelfingen, Switzerland). All other chemicals were purchased from Sigma (Fluka) or Acros and were used without further purification.

General Procedure for SPPS

Peptides were prepared by SPPS on a 0.05 mmol scale using a peptide synthesizer (MultiPep RS, Intavis AG, Cologne, Germany) and a standard Fmoc/*t*-Bu protocol, with double coupling at every cycle. The synthesis of the peptides was performed on a TentaGel S Ram resin (RappPolymere, Tübingen, Germany) and carried out in plastic syringes at room temperature. Couplings were achieved by reacting four equivalents of Fmoc-aa-OH with four equivalents of PyBOP and eight equivalents of NMM in DMF. The following side-chain protection was used: Boc for the amino acids Lys, Trp; *t*-Bu for the amino acids Thr, Ser; O*t*Bu for the amino acids Glu, Asp; Trt for the amino acids Asn, Gln, His, and Pbf for the amino acid Arg. Coupling of the pseudoproline (oxazolidine) dipeptides Fmoc-Ala-Thr-($\Psi^{Me,Me}$pro)-OH, Fmoc-Leu-Ser-($\Psi^{Me,Me}$pro)-OH and Trp(Boc)-Thr($\Psi^{Me,Me}$pro)-OH was performed as with standard Fmoc amino acids. A solution of 20% piperidine in DMF was used to remove the Fmoc protection group. Peptides were deprotected and cleaved from the resin using a mixture of 10 ml TFA, 0.75 g phenol, 0.5 ml water, 0.5 ml methylphenylsulfide and 0.25 ml ethane-1,2-dithiol. After 3 h at room temperature, the cleavage solution was collected, and the crude peptides precipitated from ice-cold methyl *tert*-butyl ether. Crude peptides were washed five times with dry diethyl ether. Final RP-HPLC (Waters, Eschborn, Germany) purification and analysis were achieved using a linear solvent gradient (eluent A: 0.05% TFA in water; eluent B: 0.05% TFA in acetonitrile; linear gradient from 5 to 60% B over 30 min at a flow rate of 20 ml·min^{-1}, RT) over a Vydac C18 column (Hesperia, CA, USA) and detection at 214 nm. Identities of peptides were validated by MALDI-TOF ms (VoyagerLT, Applied Biosystems, Weiterstadt, Germany).

General Procedure for SPPS of Protected peptides

The side-chain and *N*-terminally Boc protected peptide acids were synthesized on TentaGel S Trt-resins preloaded with the *C*-terminal amino acid tryptophan (Rapp Polymere, Tübingen, Germany). Loading capacity was 0.20 mmol/g. With the exception of the *N*-terminal amino acid (Boc protected amino acid) and the final cleavage procedure, synthesis was performed as described above. The fully protected peptide acids were cleaved from the resin with acetic acid/trifluoroethanol/DCM (6 : 2 : 2) for 2 h. The cleavage mixture was evaporated in the presence of hexane twice, precipitated in *n*-hexane and washed. The resulting residue was lyophilized from dioxane to remove traces of acetic acid.

General Procedure of Peptide Thioester Synthesis

The peptide thioesters were formed by treating the fully protected peptide acids obtained above (0.1–0.2 mmol) with DIC (2 equiv.), *p*-acetamidothiophenol (15 equiv.) and HOBt (1 equiv.) in dry DCM (15 ml). The solution was stirred overnight at RT. After removing the solvent *in vacuo* the peptides were deprotected as described for the standard SPPS. The peptide thioesters were purified by preparative RP-HPLC (eluent A: 0.05% TFA in water; eluent B: 0.05% TFA in acetonitrile; linear gradient from 5 to 60% B over 30 min at a flow rate of 20 ml·min^{-1}, RT, C-18 column). UV detection was performed at 214 nm. For analytical HPLC the same detection and gradient was used. Identity was confirmed via MALDI-TOF ms.

General Procedure for Native Chemical Ligation

Native chemical ligation (NCL) reactions were performed in aqueous solution containing 6 M guanidinium hydrochloride, 200 mM sodium phosphate and 20 mM *Tris*-(2-carboxyethyl)-phosphine (TCEP) at pH 6.8. The peptide concentration was 1–2 mM. The solution was stirred at 25 °C for 1.5 h and the ligated peptides were purified with RP-HPLC (eluent A: 0.05% TFA in water; eluent B: 0.05% TFA in acetonitrile; linear gradient from 5 to 60% B over 30 min at a flow rate of 20 ml·min^{-1}, RT, C-18 column, detection at 214 nm) and further analyzed via MALDI-TOF ms.

CD Spectroscopy

The folding states of peptides **2** and **5** were determined by CD spectroscopy using a Jasco-J720 spectropolarimeter (Jasco Corp., Hachioji, Japan) at a total peptide concentration of 50 μM. The peptides were dissolved in 30 mM phosphate buffer (pH 7.4, ion strength of 154 mM was set with 88 mM NaF). The samples were measured with a 1 mm quartz cell and the spectra were recorded in the range of 260–195 nm at 4 °C. The mean residue ellipticity, θ, is plotted *versus* the wavelength, λ.

Results and Discussion

The sequence of the TCERG1 WW-3 domain between amino acid 513 and 586 encompasses the 73-mer sequence **1** used to solve the structure of the TCERG1 WW-3 domain by NMR (PDB ID code 2dk7) (Fig. 1). The 43 amino acid-long peptide **2** was previously synthesized, but NMR analysis revealed that this domain was not folded correctly [11]. Thus, the *C*-terminal elongation was

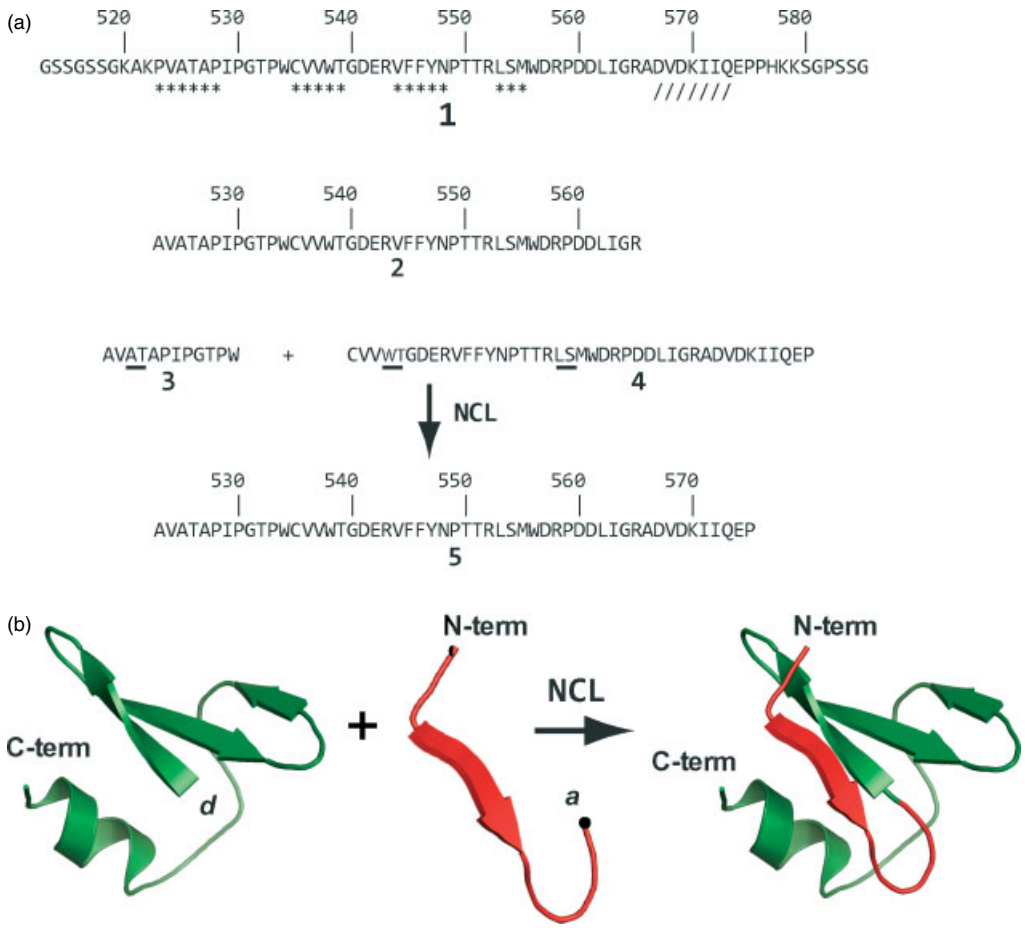

Figure 1. Panel A. Sequences of TCERG1 WW-3 domain variants. (**1**) Sequence used to solve the structure of the TCERG1 WW-3 domain by solution NMR (PDB ID code 2dk7). This 73-mer spans TCERG1 amino acids 513–586 as numbered above. Positions of the β-strand (* * *) and the helix (///) are shown. (**2**) Sequence of the 43-mer spanning residues 523–565 previously used for stepwise SPPS [11]. (**3**) and (**4**) Sequences derived from our chemical synthesis strategies and used here for NCL. Pseudoprolines (underlined residues) were used during SPPS for better synthesis quality. (**5**) Sequence of the NCL product, the 53-mer peptide used in subsequent analyses. Panel B. On the right, design of the new *C*-terminally extended WW domain peptide (**5**) with all the secondary structure elements as resulted from the NCL approach (at the left) (a denotes acceptor-site and d stands for donor-site).

apparently crucial for proper folding of the TCERG1 WW-3 domain (Fig. 1; compare sequences **1** and **2**). We surmised that WW-3 domain folding could be stabilized by the short α-helical region (residues 567–573) or by a mutual stabilization of both secondary structure elements and we decided to truncate the 73-mer sequence **1**. We anticipated that a proper WW domain fold can also be conferred by the α-helix contained within the 53-mer sequence **5** spanning residues 523–575 (Fig. 1).

Although using oxazolidine dipeptide building blocks, a first attempt to synthesize the domain **5** in one piece by SPPS was not successful. HPLC analysis of the raw product shows a highly complex peptide mixture and we did not found the calculated mass of the WW domain **5** by MALDI-TOF ms. Fortunately, sequence **5** contains a Cys residue at position 535 that allows its preparation using a NCL approach [23]. Consequently, we synthesized two fragments for ligation: first the short *N*-terminal peptide thioester **3**, and second the Cys containing fragment **4** (Fig. 1).

Synthesis of the peptide thioester **3** was performed according to the procedure of Eggelkraut-Gottanka *et al*. [24]. Briefly, fragment **3** was completed on a Tentagel Trityl resin *via* Fmoc SPPS and PyBOP activation, whereby the *N*-terminus was Boc protected (see Materials and Methods). The side-chain and *N*-terminal protected peptide was cleaved from the resin with a mixture of acetic acid and trifluoroethanol in DCM. Subsequently, thioester **3** was formed by treating the fully protected peptide acid with *p*-acetamidothiophenol (*p*-Aatp) using DIC activation. After the final side-chain deprotection procedure and HPLC purification, the desired peptide thioester fragment **3** was obtained in 36% yield with excellent purity. The identity of peptide **3** was confirmed by mass spectrometry (Table 1, Fig. 2(A)).

Table 1. Characteristics of the synthesized peptides

Peptide	Calculated mass $[M+H]^+$	Measured mass	HPLC t_r (min)[a]	Purity (%Area)	Yield (%)
3	1329.5	1352.2 $[M+Na]^+$	13.07	>95	36
4	4854.4	4853.9 $[M+H]^+$	13.04	>95	22
5	6016.7	6016.2 $[M+H]^+$	14.23	>95	19

Sequences of the compounds are given in Figure 1. Masses were measured by MALDI-TOF mass spectrometry. Purity is calculated by integrating the HPLC peak area.
[a] Measurement errors of un-cooled HPLC column correspond to 0.09 min. Yields were calculated from the first amino acid loading; the yield of ligation product **5** was calculated from the corresponding reacting peptides **3** and **4**.

Journal of PeptideScience

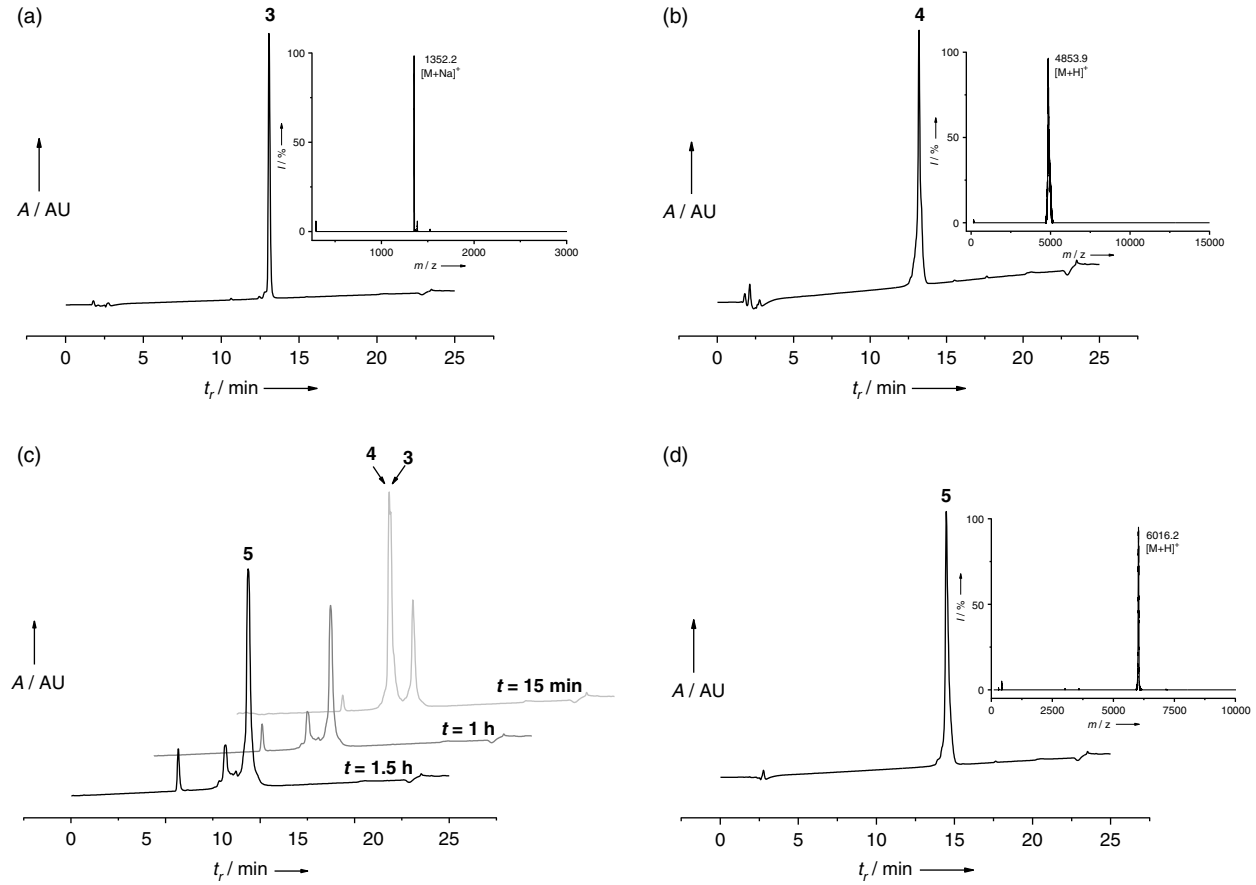

Figure 2. RP-HPLC and mass spectrometry analyses of peptide fragments **3**, **4** and**5**. A: HPLC chromatogram of the peptide thioester **3** with a measured retention time of $t_r = 13.07$ min. (peak area = <95%). Shown as an inset on the right is the MALDI-TOF mass spectra of peptide **3** (mass calculated: $m/z = 1329.5$ $[M + H]^+$, mass found: $m/z = 1352.2$ $[M + Na]^+$). B: HPLC chromatogram of fragment **4** with a measured retention time of $t_r = 13.04$ min. The inset shows the MALDI-TOF mass spectra of **4** (mass calculated: $m/z = 4854.4$ $[M + H]^+$; mass found: $m/z = 4854.9$ $[M + H]^+$). C: HPLC monitoring of the native chemical ligation reaction between thioester fragment **3** and peptide **4**. Fragments **3** and **4** are superimposed but clearly separated from the ligation product **5** at the retention time $t_r = 14.23$ min. D: HPLC and mass spectrometry analyses of the ligation product **5**. HPLC chromatogram of the purified 53-mer TCERG1 WW-3 domain **5** with a measured retention time of $t_r = 14.23$ min (peak area = >95%). The inset shows the MALDI-TOF mass spectrum of the ligation product **5** (mass calculated: $m/z = 6016.75$ $[M + H]^+$; mass found: $m/z = 6016.20$ $[M + H]^+$).

Fragment **4** was synthesized by Fmoc SPPS on a TentaGel SRam resin by PyBOP activation. To improve synthesis quality, oxazolidine building blocks were incorporated at certain sequence positions (Fig. 1; underlined residues). The peptide fragment **4** was obtained in good yield (22%) and purity after HPLC purification, and its identity was validated by mass spectrometry (Table 1, Fig. 2(B)).

The NCL between the peptide thioester **3** and fragment **4** was performed in aqueous sodium phosphate buffer containing 6 M guanidinium hydrochloride (GnHCl), 200 mM sodium phosphate and 20 mM TCEP at pH 6.8 at room temperature. The reaction was monitored by analytical HPLC, and ligation was almost complete after 1 h (Fig. 2(C)). Unfortunately, the thioester fragment **3** and peptide **4** peaks are superimposed in the HPLC chromatogram ($t_r = 13.04$ and 13.07), but distinct separation between product **5** and reactants **3** and **4** enabled monitoring of the ligation reaction. The peak at $t_r = 10.7$ min belongs to *p*-acetamidothiophenol. After HPLC purification, the ligation product **5** was obtained in 19% yield and excellent purity. In addition, mass spectrometry analysis confirmed the identity of the ligation product **5** (Table 1, Fig. 2(D)).

Peptide **2** (short 43-mer TCERG1 WW-3 domain spanning amino acids 523–565) was synthesized according to the procedure of Otte *et al.* [11] and used for comparison in subsequent experiments (data not shown).

To check whether the presence of the α-helical sequence region affects folding of the 53-mer TCERG1 WW-3 sequence, we characterized the folding features and conformational stability of the ligation product **5** obtained by biophysical methods. The CD spectrum of **5** recorded at 4 °C and pH 7.4 (Fig. 3) is characterized by significant absorption with a positive ellipticity at around 230 nm, typical for anti-parallel β-sheet folding of a WW domain [4], and two negative minima around 206 and 220 nm typical for an alpha helical fold. In comparison, CD spectra of the short WW-3 domain **2** reveal no definable secondary structure at all (Fig. 3).

This result was supported by 1D-^{1}H NMR analyses of **5**, which confirmed a folded WW domain (Fig. S1 in Supporting Information). Thus, our presented data clearly demonstrate that the chemically synthesized 53-mer TCERG1 WW-3 domain **5** depicts well defined structure typically for a WW domain. Interestingly, the structure of the WW-3 domain **5** was not significantly influenced by increasing its concentration (Fig. S2).

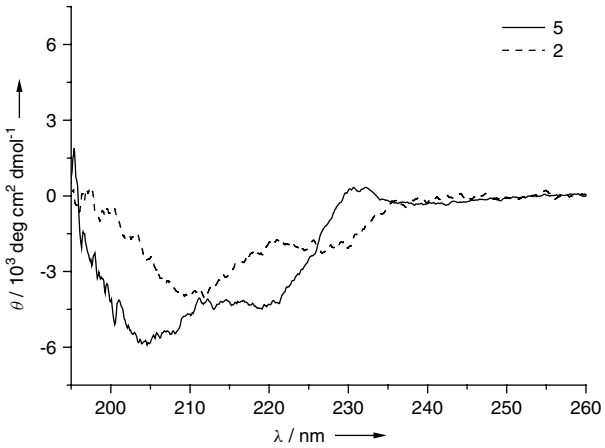

Figure 3. CD spectra of the two TCERG1 WW-3 domain variants **5** and **2**. The solid line (–) plots the CD spectra of the 53-mer WW-3 domain peptide **5**. The dashed line (---) traces the CD spectra of the 43-mer WW-3 domain peptide **2**. Both domains were measured at pH = 7.4 and at a concentration of 50 μM. In contrast to domain **2,** the profile for domain **5** reveals the typical WW domain fold at 230 nm and a helical structure at 205 and 220 nm.

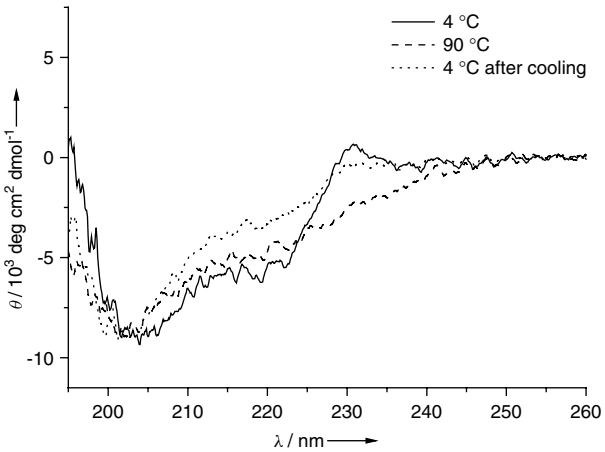

Figure 4. CD spectra and thermal unfolding of the TCERG1 WW-3 domain **5**. CD spectra of **5** were measured at 4, 90 and 4 °C after cooling.

This is in contrast to the TCERG1 WW-2 domain, which is reported to show enhanced self-assembly at increased concentrations [9].

Furthermore, we recorded far-UV CD spectra of peptide **5** during thermal unfolding from 4 to 90 °C at 230 nm and after cooling from 90 to 4 °C (Fig. 4). The revealed ellipticity changes at 230 nm, demonstrating the unfolding and refolding behavior of the WW domain **5,** and suggesting this is a reversible process.

In summary, we successfully synthesised the *C*-terminally truncated WW-3 domain of TCERG1 using NCL. The chosen synthesis strategy was efficient and resulted in good yields and excellent purity of the 53 amino acid-long peptide. CD and NMR studies revealed formation of both α-helical and β-sheet secondary structures, substantiating our prediction that the short α-helical sequence stabilizes the TCERG1 WW-3 domain fold. These data should enable future screening for cellular targets of the TCERG1 WW-3 domain.

Acknowledgements

Financial support from Deutsche Forschungsgemeinschaft (VO 885/3-1) is gratefully acknowledged.

Supporting information

Supporting information may be found in the online version of this article.

References

1 Zarrinpar A, Bhattacharyya RP, Lim WA. The structure and function of proline recognition domains. *Sci. STKE* 2003; (179): RE8.
2 Macias MJ, Wiesner S, Sudol M. WW and SH3 domains, two different scaffolds to recognize proline-rich ligands. *FEBS Lett.* 2002; **513**: 30–37.
3 Sudol M. The WW domain. In *Modular Protein Domains*. Cesareni G, Gimona M, Sudol M, Yaffe M (eds). Wiley-VCH: Weinheim, 2005; 59–72.
4 Macias MJ, Hyvönen M, Baraldi E, Schultz J, Sudol M, Saraste M, Oschkinat H. Structure of the WW domain of a kinase-associated protein complexed with a proline-rich peptide. *Nature* 1996; **382**: 646–649.
5 Koepf EK, Petrassi HM, Sudol M, Kelly JW. WW: an isolated three-stranded antiparallel beta-sheet domain that unfolds and refolds reversibly; evidence for a structured hydrophobic cluster in urea and GdnHCl and a disordered thermal unfolded state. *Protein Sci.* 1999; **8**: 841–853.
6 Jager M, Deechongkit S, Koepf EK, Nguyen H, Gao J, Powers ET, Gruebele M, Kelly JW. Understanding the mechanism of beta-sheet folding from a chemical and biological perspective. *Biopolymers* 2008; **90**: 751–758.
7 Deechongkit S, Nguyen H, Jager M, Powers ET, Gruebele M, Kelly JW. Beta-sheet folding mechanisms from perturbation energetics. *Curr. Opin. Struct. Biol.* 2006; **16**: 94–101.
8 Ramirez-Espain X, Ruiz L, Martin-Malpartida P, Oschkinat H, Macias MJ. Structural characterization of a new binding motif and a novel binding mode in group 2 WW domains. *J. Mol. Biol.* 2007; **373**: 1255– 1268.
9 Ferguson N, Becker J, Tidow H, Tremmel S, Sharpe TD, Krause G, Flinders J, Petrovich M, Berriman J, Oschkinat H, Fersht AR. General structural motifs of amyloid protofilaments. *Proc. Natl. Acad. Sci. U.S.A.* 2006; **103**: 16248–16253.
10 Przezdziak J, Tremmel S, Kretzschmar I, Beyermann M, Bienert M, Volkmer-Engert R. Probing the ligand-binding specificity and analyzing the folding state of SPOT-synthesized FBP28 WW domain variants. *ChemBioChem* 2006; **7**: 780–788.
11 Otte L, Wiedemann U, Schlegel B, Pires JR, Beyermann M, Schmieder P, Krause G, Volkmer-Engert R, Schneider-Mergener J, Oschkinat H. WW domain sequence activity relationships identified using ligand recognition propensities of 42 WW domains. *Protein Sci.* 2003; **12**: 491–500.
12 Carty SM, Goldstrohm AC, Sune C, Garcia-Blanco MA, Greenleaf AL. Protein-interaction modules that organize nuclear function: FF domains of CA150 bind the phosphoCTD of RNA polymerase II. *Proc. Natl. Acad. Sci. U.S.A.* 2000; **97**: 9015–9020.
13 Goldstrohm AC, Albrecht TR, Sune C, Bedford MT, Garcia-Blanco MA. The transcription elongation factor CA150 interacts with RNA polymerase II and the pre-mRNA splicing factor SF1. *Mol. Cell. Biol.* 2001; **21**: 7617–7628.
14 Smith MJ, Kulkarni S, Pawson T. FF domains of CA150 bind transcription and splicing factors through multiple weak interactions. *Mol. Cell. Biol.* 2004; **24**: 9274–9285.
15 Sánchez-Alvarez M, Goldstrohm AC, Garcia-Blanco MA, Suñé C. Human transcription elongation factor CA150 localizes to splicing factor-rich nuclear speckles and assembles transcription and splicing components into complexes through its amino and carboxyl regions. *Mol. Cell. Biol.* 2006; **26**: 4998–5014.
16 Pearson JL, Robinson TJ, Muñoz MJ, Kornblihtt AR, Garcia-Blanco MA. Identification of the cellular targets of the transcription factor TCERG1 reveals a prevalent role in mRNA processing. *J. Biol. Chem.* 2008; **283**: 7949–7961.

17 Lin KT, Lu Rm, Tarn WY. The WW domain-containing proteins interact with the early spliceosome and participate in pre-mRNA splicing in vivo. *Mol. Cell. Biol.* 2004; **24**: 9176–9185.
18 Holbert S, Denghien I, Kiechle T, Rosenblatt A, Wellington C, Hayden MR, Margolis RL, Ross CA, Dausset J, Ferrante RJ, Neri C. The Gln-Ala repeat transcriptional activator CA150 interacts with huntingtin: neuropathologic and genetic evidence for a role in Huntington's disease pathogenesis. *Proc. Natl. Acad. Sci. U.S.A.* 2001; **98**: 1811–1816.
19 Arango M, Holbert S, Zala D, Brouillet E, Pearson J, Regulier E, Thakur AK, Aebischer P, Wetzel R, Deglon N, Neri C. CA150 expression delays striatal cell death in overexpression and knock-in conditions for mutant huntingtin neurotoxicity. *J. Neurosci.* 2006; **26**: 4649–4659.
20 Goldstrohm AC, Greenleaf AL, Garcia-Blanco MA. Co-transcriptional splicing of pre-messenger RNAs: considerations for the mechanism of alternative splicing. *Gene* 2001; **277**: 31–47.
21 Wellings DA, Atherton E. Standard Fmoc protocols. *Methods Enzymol.* 1997; **289**: 44–67.
22 Tremmel S, Beyermann M, Oschkinat H, Bienert M, Naumann D, Fabian H. 13C-labeled tyrosine residues as local IR probes for monitoring conformational changes in peptides and proteins. *Angew. Chem. Int. Ed. Engl.* 2005; **44**: 4631–4635.
23 Dawson PE, Muir TW, Clark-Lewis I, Kent SB. Synthesis of proteins by native chemical ligation. *Science* 1994; **266**: 776–779.
24 Eggelkraut-Gottanka R, Klose A, Beck-Sickinger AG, Beyermann M. Peptide (alpha) thioester formation using standard Fmoc-chemistry. *Tetrahedron Lett.* 2003; **44**: 3551–3554.

