Glycoconj J (2010) 27:461–476
DOI 10.1007/s10719-010-9291-4

High-yield production, refolding and a molecular modelling of the catalytic module of (1,3)-β-D-glucan (curdlan) synthase from *Agrobacterium* sp.

Maria Hrmova · Bruce A. Stone · Geoffrey B. Fincher

Received: 20 December 2009 / Revised: 14 April 2010 / Accepted: 14 April 2010 / Published online: 16 May 2010

Abstract Biosynthesis of the (1,3)-β-D-glucan (curdlan) in *Agrobacterium* sp., is believed to proceed by the repetitive addition of glucosyl residues from UDP-glucose by a membrane-embedded curdlan synthase (CrdS) [UDP-glucose: (1,3)-β-D-glucan 3-β-D-glucosyltransferase; EC 2.4.1.34]. The catalytic module of CrdS (cm-CrdS) was expressed in good yield from a cDNA encoding cm-CrdS cloned into the pET-32a(+) vector, containing a coding region for thioredoxin, and from the Champion™ pET SUMO system that possesses a coding region of a small ubiquitin-related modifier (SUMO) partner protein. The two DNA fusions, designated pET-32a_cm-CrdS and SUMO_cm-CrdS were expressed as chimeric proteins. High yields of inclusion bodies were produced in *E. coli* and these could be refolded to form soluble proteins, using a range of buffers and non-detergent sulfobetaines. A purification protocol was developed, which afforded a one-step on-column refolding and simultaneous purification of the recombinant 6xHis-tagged SUMO_cm-CrdS protein. The latter protein was digested by a specific protease to yield intact cm-CrdS in high yields. The refolded SUMO_cm-CrdS protein did not exhibit curdlan synthase activity, but showed a circular dischroism spectrum, which had an α/β-type-like conformation. Amino acid sequences of tryptic fragments of the SUMO_cm-CrdS fusion and free cm-CrdS proteins, determined by MALDI/TOF confirmed that the full-length proteins were synthesized by *E. coli*, and that no alterations in amino acid sequences occurred. A three-dimensional model of cm-CrdS predicted the juxtaposition of highly conserved aspartates D156, D208, D210 and D304, and the QRTRW motif, which are likely to play roles in donor and acceptor substrate binding and catalysis.

Keywords *Escherichia coli* · Family GT2 transferases · Heterologous expression · Homology model · Integral membrane proteins · pET-32a(+) and Champion™ pET SUMO vectors

Abbreviations

CAZy	Carbohydrate-active enzymes
CHES	*N*-cyclohexyl-2-aminoethanesulfonic acid
CD	Circular dichroism
cm-CrdS	Catalytic module of CrdS
CrdS	Curdlan synthase
EDTA	Ethylenediaminetetraacetic acid
EPS	Extracellular and capsular polysaccharides
GT	Glycoside transferase(s)
HCA	Hydrophobic cluster analysis
IMAC	Immobilized metal affinity chromatography
IPTG	Isopropyl-β-D-thiogalactopyranoside
MALDI-TOF	Matrix-assisted laser desorption/ionization-time of flight
MS	Mass spectrometry
M_r	Relative molecular mass
NDSB	Non-detergent sulfobetaine

This article is dedicated to the memory of the late Professor Bruce Stone, whose encouragement was a continuing source of inspiration.

Electronic supplementary material The online version of this article (doi:10.1007/s10719-010-9291-4) contains supplementary material, which is available to authorized users.

M. Hrmova (✉) · G. B. Fincher
Australian Centre for Plant Functional Genomics,
School of Agriculture, Food and Wine, University of Adelaide,
Waite Campus,
Glen Osmond, SA 5064, Australia
e-mail: maria.hrmova@adelaide.edu.au

B. A. Stone
School of Biochemistry, La Trobe University,
Bundoora,
Melbourne, Australia

PDB	Protein data bank
PMSF	Phenylmethylsulphonyl fluoride
SDS	Sodium dodecylsulfate
SUMO	Small ubiquitin-related modifier
TCEP	Tris-(2-carboxyethyl)-phosphine
UDP-Glc	Uridine-diphosphate glucose
3D	Three-dimensional

Introduction

Extracellular bacterial (1,3)-β-D-glucans, known as curdlans, are synthesized by a variety of bacteria, such as *Alcaligenes* and *Agrobacterium* that are classified in the Gram-negative α-Proteobacteria. The extracellular curdlans alongside other polysaccharides form pellicles known as extracellular capsular polysaccharides (EPS) that protect the bacteria from a range of adverse environmental conditions, or mediate associations between bacteria and biotic/abiotic surfaces [1, 2]. Some EPS represent major virulence determinants in animal pathogens, while others play roles in symbiotic interactions between bacteria and plants [3].

The bacterial curdlans are neutral and essentially linear, water-soluble or water-insoluble polysaccharides composed of (1-3)-β-linked D-glucosyl residues. The bacterial (1-3)-β-D-glucans stain with a fluorophore (sodium 4,4′-[carbonylbis-(benzene-4,1-diyl)bis(imino)]bisbenzene-sulphonate) from the triphenylmethane dye aniline-blue and form aggregates of triple-stranded helices [3, 4]. The water insoluble curdlans consists of up to 12,000 glucosyl residues [5] and may have some intra- or inter-chain (1-6)-linkages that substitute individual (1-3)-β-D-glucan chains or cross-link them [3, 6]. Some extracellular (1-3)-β-D-glucans have found industrial applications, for example as thermal thickeners and gel-forming agents in the food and building industries [1–3].

The molecular genetics of curdlan production have been studied extensively in *Agrobacterium* sp. [3]. Studies on transposon-insertion mutants identified four genes (*crdA*, *crdS*, *crdC* and *crdR*) that are indispensable for curdlan production [7], and another pss_{AG} gene encoding a phosphatidylserine synthase that controls curdlan production. The *crdsASC* genes cover a region of approximately 5,000 bp on the *Agrobacterium* sp. genome and are organized in an operon-like arrangement [8]. The two other remaining genes *crdR* and pss_{AG} [9] occur at separate loci of the genome and are not linked with the *crdsASC* cluster. Based on the sequence homology with other β-D-glycan synthases, including plant cellulose and (1,3;1,4)-β-D-glucan synthases, the *crdS* gene of an approximate size of 2,000 bp was shown to encode a curdlan synthase [7].

A good deal of information has been published on the biosyntheses of (1-3)-β-D-glucans in yeasts [10], fungi [11] and higher plants [2, 12–14]. The biosynthesis is believed to proceed by the repetitive addition of glucosyl residues from the sugar nucleotide donor UDP-glucose to form polymeric (1-3)-β-D-glucan chains. The synthesis is expected to be mediated by a (1,3)-β-D-glucan synthase, designated UDP-glucose: (1,3)-β-D-glucan 3-β-D-glucosyltransferase (EC 2.4.1.34), which in bacteria is also known as curdlan synthase (CrdS) [1]. Curdlan synthases are confined within cell membranes and based on the sequence homology with other β-glycan synthases, including plant cellulose and (1,3;1,4)-β-D-glucan synthases, the crdS enzyme has been classified in the GT2 family of glycosyl transferases [CAZY; 7,15]. As of April 2010, the GT2 group of glycosyl transferases [2, 15] contains around 16,000 viral, pro- and eukaryotic entries, and represents one of the largest GT groups in the CAZy database. The GT2 enzymes are predicted to adopt a GT-A fold and to use an inverting reaction mechanism that is mediated through a single displacement reaction *via* a glycosyl-enzyme intermediate [15, 16]. The GT2 family is distantly related to the GT12, GT21, GT27 and GT55 families [15]. However, it should be noted that the barley [17] and other plant (1,3)-β-D-glucan synthases [18] were discovered first and are classified within the GT48 group of enzymes [2, 15, 19].

The cDNA of CrdS from *Agrobacterium* sp. [7] translates into a protein of 654 amino acid residues and the topology predictions of CrdS show multiple membrane spanning α-helices [20], while a single large cytoplasmic region embodies a putative catalytic module (cm-CrdS) that contains characteristic D, D, D and QxxRW signatures found in other glucosyltransferase classified in the GT2 family [1]. CrdS, predicted by DISULFIND [21], contains at least one pair of bonded cysteines located in cm-CrdS, although the experimental evidence for its existence is lacking. The known 3D structures of three GT2 enzymes, all elucidated by X-ray crystallography, include a soluble bacterial spore-coat protein [22], a chondroitin polymerase enzyme from *Escherichia coli* [23] involved in synthesis of chondroitin sulfate, and a putative glycosyltransferase from *Bacteroides fragilis* [Protein Data Bank (PDB) accession number 3BCV, Polani, Kumaran, Burley and Swaminathan, unpublished data]. All three enzymes lack membrane spanning helices and essentially represent catalytic modules that fold into mixed α/β architectures.

The objective of the present work was to express in heterologous systems the catalytic module cm-CrdS of the *CrdS* gene from *Agrobacterium* sp. This work was undertaken with the long-term goal of characterizing the 3D structure of the full-length CrdS enzyme. Here, the

coding region of cm-CrdS was sub-cloned into pET-32a (+) and Champion™ pET SUMO expression vectors, and these DNA fusions are designated pET-32a_cm-CrdS and SUMO_cm-CrdS. The former vector was chosen, because it facilitates disulfide bond formation in the reducing environment of *E. coli* cytoplasm [24], alleviates toxic effects of recombinant protein on host cells, prevents target proteins from proteolytic degradation and accelerates expression of proteins in a soluble form [25]. The pET SUMO expression system is known to increase expression levels of recombinant proteins, it enhances the solubility of partially soluble proteins [26, 27], and it has also been used for the production of membrane proteins [28]. The screening of protein expression conditions showed that both DNA constructs produced high yields of inclusion bodies that could successfully be refolded. The refolded, chimeric SUMO_cm-CrdS protein showed a circular dischroism spectrum that suggested it folded into a native-like state. The refolded and purified chimeric SUMO_cm-CrdS protein was digested by a specific SUMO-protease to yield cm-CrdS in high yields.

Materials and methods

Chemicals, reagents and expression plasmids

Oligonucleotide primers were synthesised by GeneWorks (Thebarton, SA, Australia). Restriction and DNA-modifying enzymes were obtained from New England Biolabs (Beverly, MA, USA) and Roche Applied Science (Castle Hill, NSW, Australia), which also supplied the EDTA-free Complete Protease Inhibitor Cocktail tablets, containing trypsin, chymotrypsin, papain, thermolysin and a pancreas-extract. Plasmid extraction kits were obtained from Qiagen (Hilden, Germany) and Macherey Nagel (Duren, Germany), and a mouse anti-polyHistidine-alkaline phosphatase IgG2a isotype antibody and the alkaline phosphotase chromogen BCIP/NBT-purple liquid reagent for membranes were from Sigma Chemical Company (St. Louis, MO, USA). All *E. coli* strains, the pET-32a (+) expression vector and iFOLD® Protein Refolding System 2 were purchased from Novagen-Merck (San Diego, CA, USA), pGEM-T Easy vector system was from Promega (Madison, WI, USA) and the Champion™ pET SUMO expression system, SUMO-protease, *E. coli* competent cells, strain DH5α and One Shot Match1®-T1 were from Invitrogen Life Technologies (Carlsbad, CA, USA). A pure nitrocellulose transfer membrane (0.22 μm) was from Thermo Fisher Scientific Inc (Pittsburgh, PA, USA), the Co^{2+}-charged Talon resin was from Clontech (Mountain View, CA, USA) and NDSB-256 was from Anatrace (Maume, OH, USA).

Recombinant plasmid design

Based on structural predictions and protein modelling (*vide intra*), truncation sites on the full-length sequence of CrdS were chosen allowing two types of DNA fusion constructs of cm-CrdS to be generated (Fig. 1). The 1st truncation site was positioned behind the end of the 3rd membrane helix for all three constructs (constructs 1 and 2 of pET-32a_cm-CrdS and SUMO_cm-CrdS), or more precisely behind a hydrophobic segment LILL (marked by a left-hand arrow in Fig. 2b). The 2nd truncation site was positioned in front of the membrane helix 4, specifically, behind YLHG and YLHGL motifs for the pET-32_cm-CrdS and SUMO_cm-CrdS constructs, respectively (marked by two right hand arrows in Fig. 2b). The two expression systems pET-32(+) and Champion™ pET SUMO were chosen as the destination vectors containing respective thioredoxin and SUMO folding partners (both 109 residues), other affinity tags (6xHis and S-tag), proteolytic sites (enterokinase, thrombin and SUMO), and cloning overhanging sequences. In summary, the pET-32a_cm-CrdS (constructs 1 and 2) and SUMO_cm-CrdS proteins contained 447 and 403 residues respectively, from which 280 and 281 residues in each expression vector originated from cm-CrdS (Fig. 1).

Recombinant plasmid construction

The pUC19 plasmid with an 8.8 kb *EcoRI Agrobacterium* genomic fragment including the *crdASC* genes, designated pVS1512 plasmid [20] was obtained from Dr Vilma Stanisich (La Trobe University, Bundoora, Australia). DNA constructs encoding cm-CrdS were amplified by PCR, using the forward primer pairs 5′ **GAA TTC** CGC CGA ACC GAC TGG TCA 3′ (forward, construct 1) or 5′ **GAA TTC** CGC CGA ACC GAC TGG 3′ (forward, construct 2) and 5′ **GCG GCC GC**T TAT CCT CAT CCA TGC AGA TAG TG 3′ (reverse, constructs 1 and 2) from the pVS1512 plasmid. Bold characters in oligonucleotide primers indicate restriction sites *EcoRI* (forward, construct 1), *BamHI* (forward, construct 2) and *NotI* (reverse, constructs 1 and 2). The amplified fragments were cloned into the pGEM-T Easy vector system. For DNA amplifications (constructs 1 and 2), *E. coli* strain DH5α (Invitrogen) was transformed as recommended by the manufacturer. The pGEM_cm-CrdS plasmids were digested with *EcoRI* and *NotI* (construct 1) or with *BamHI* and *NotI* (construct 2) and the cm-CrdS DNA fragments were sub-cloned into the pET-32a (+) expression system. The cm-CrdS DNA fragment was ligated with the digested pET-32a (+) vector that was in-frame with the thioredoxin and 6xHis (histidine) tags, a thrombin cleavage site, a S-tag and an enterokinase cleavage site (Fig. 1), and under the transcriptional control of the T7 RNA polymerase, and transformed into *E. coli* strain DH5α using chemical

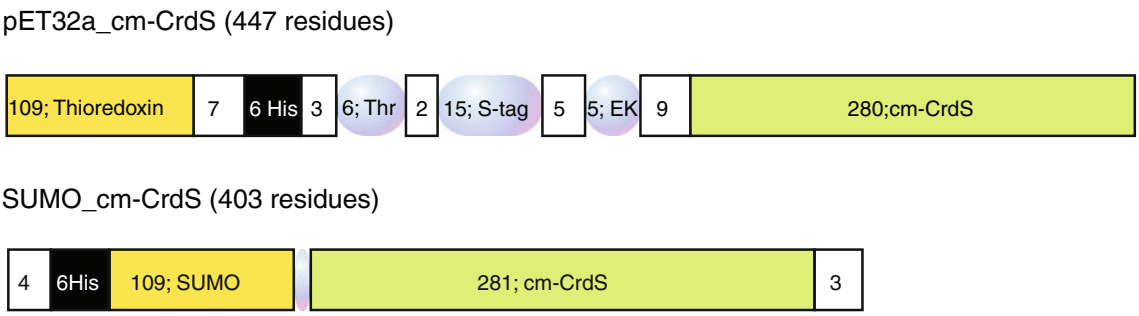


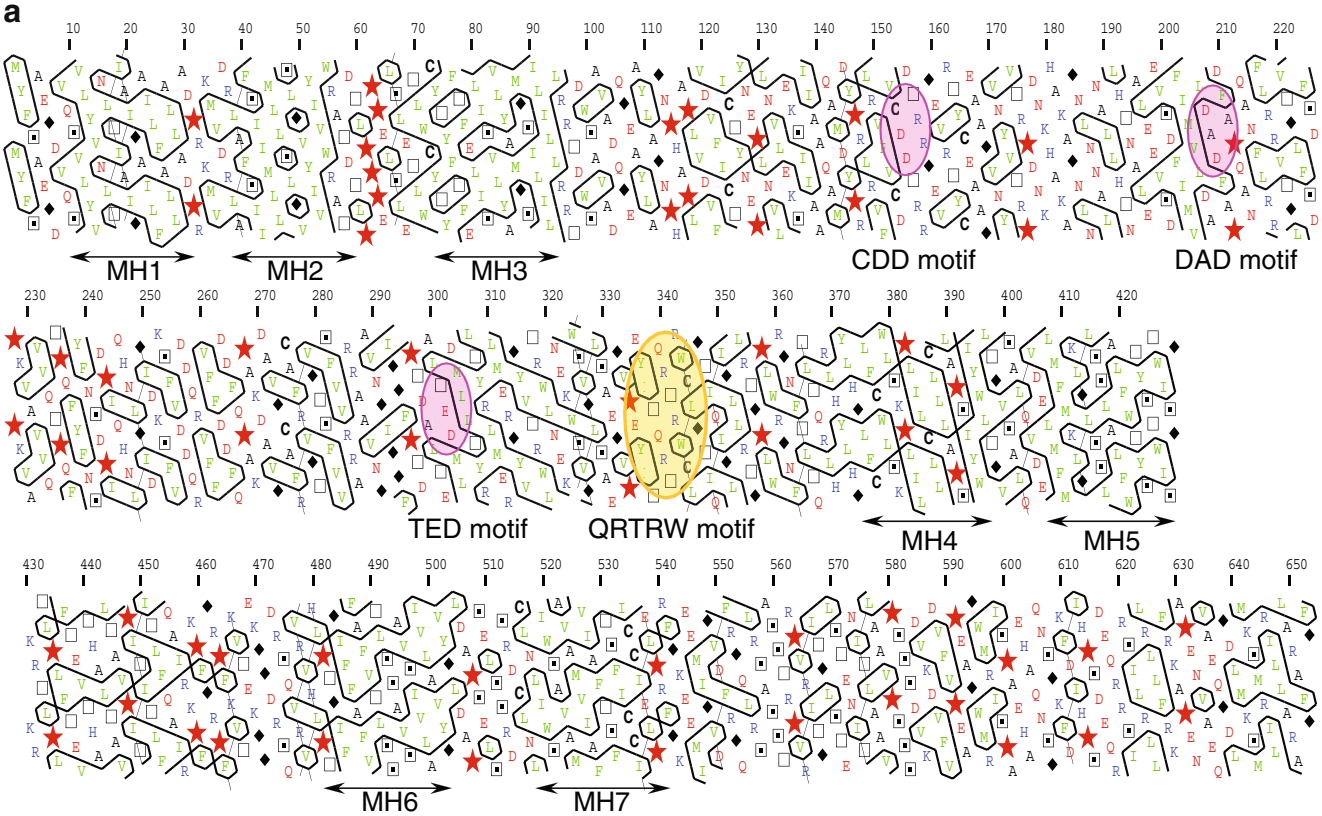
Fig. 1 DNA fusions used for expression of cm-CrdS. The pET-32a_cm-CrdS construct consists of a thioredoxin domain (*yellow*), thrombin site, S-tag and an enterokinase site (all coloured as *blue ellipsoids*), 6His (6x-His-tag, *black*) and cm-CrdS (*green*). The SUMO_cm-CrdS construct consists of 6His (6x-His-tag, *black*), a SUMO domain (*yellow*) and cm-CrdS (*green*). The number of residues in each segment and overhanging residues in both constructs are indicated in respective segments and in *white boxes*, respectively

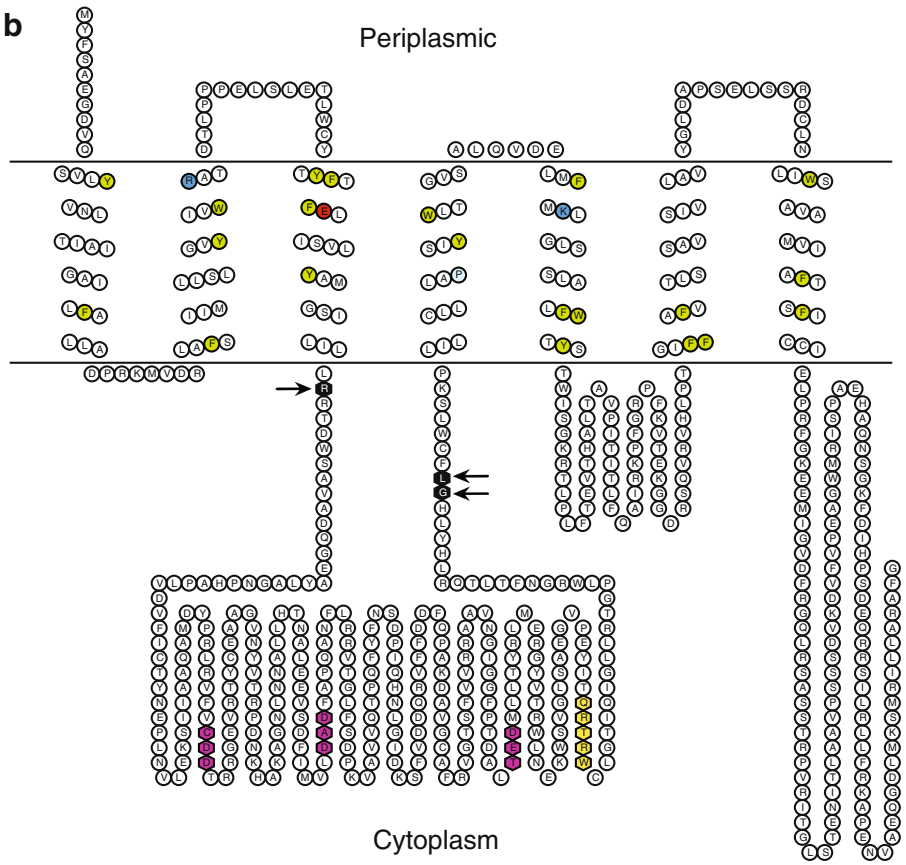
transformation. To generate the SUMO_cm-CrdS fusion, the cDNA of cm-CrdS was amplified by PCR using the primer pairs 5′ CGC CGA ACC GAC TGG TCA 3′ (forward) and 5′ AAG TCC ATG CAG ATA GTG 3′ (reverse). The amplified DNA fragment was ligated with the Champion™ pET SUMO vector that was in-frame with the 6xHis and SUMO tags (Fig. 1) and under the transcriptional control of the T7 RNA polymerase. The ligation mixture was transformed into the competent cells *of E. coli* One Shot Match1®-T1, as recommended by the manufacturer. The successfully transformed cells were selected by colony PCR screening. The DNAs of the pET-32a_cm-CrdS (constructs 1 and 2) and SUMO_cm-CrdS fusions were sequenced in both directions.

Screening of protein production

The pET-32a_cm-CrdS and SUMO_cm-CrdS DNA fusions were transformed into four competent *E. coli* strains using a chemical transformation procedure. Here, the Origami (DE3) and Origami B (DE3) strains were used for pET-32a_cm-CrdS and the BL21*(DE3) and Rosetta-Gami strains were used for the SUMO_cm-CrdS plasmid. From the matrix of 3×2 possibilities for the pET-32a_cm-CrdS (constructs 1 and 2 in two strains) and SUMO_cm-CrdS (1 construct in two strains) fusions, protein expression screening tests were carried out at three isopropyl β-D-1-thiogalactopyranoside (IPTG) concentrations (0, 0.2 and 0.8 mM) and three different temperatures (4, 14 and 28°C) during 8 h to 120 h induction time intervals as follows. Freshly transformed *E. coli* cells of the Origami (DE3) and Origami B (DE3) strains were grown in Luria-Bertani (LB) medium supplemented with ampicillin (50 μg/ml), kanamycin (15 μg/ml) and tetracycline (12.5 μg/ml) at 37°C on an orbital shaker for 24 h. Protein expression was induced with 0, 0.2 and 0.8 mM IPTG at 14 and 28°C and the cells were grown for up to 4 days. Screening of protein production with the BL21*(DE3) and Rosetta-Gami strains was carried out using the freshly transformed *E. coli* cells that were grown in the LB medium supplemented with either kanamycin (50 μg/ml), or kanamycin (50 μg/ml), tetracycline (12.5 μg/ml) and chloramphenicol (34 μg/ml), respectively, at 37°C on an orbital shaker for 24 h. Protein expression was induced as described above for the Origami cells, except that various temperatures (4, 14 and 28°C) and induction intervals (up to 5 days) were tested. Production of chimeric cm-CrdS proteins was assayed as follows. Induced 160 μl cell aliquots were sampled into the 96-well Multi-screen®HTS-HV plates (Millipore, Billerica, USA) and lysed under intense reciprocal shaking (1,250 rpm/min; 10°C; 20 min) in the presence of 40 μl of the PopCulture reagent (Novagen-Merck, Darmstadt, Germany) containing lysozyme (10 μg/ml), benzonase (25 U/ml), 1 mM PMSF, 2 mM TCEP and 0.1 mM UDP-Glc. The lysed cells were filtered under reduced pressure through the HT Millipore manifold (Millipore) and filtered solutions were collected in microtubes, and stored at 4°C. The cell pellets were resuspended in 100 μl 4% (w/v) SDS and intensely

Fig. 2 Secondary structure analyses of CrdS from *Agrobacterium* sp. **a** A two-dimensional HCA plot. Positions of seven membrane helices (MH1-MH7) are marked by *arrowed lines*. In the HCA plots, proline residues are shown as *red stars*, glycine residues as *black diamonds*, serine residues are *empty squares* and threonine residues are shown as *squares* containing a *black dot* in the centre. Negatively charged residues are coloured in *red* and positively charged residues are in *blue*. Other residues are shown by their single amino acid letter codes. The amino acid numbers are read from the *top* to the *bottom* of the plots (in duplicate) in a *left* to *right* direction. The putative catalytic (CDD, DAD and TED) and the QRTRW motifs are *circled* in *magenta* and *yellow*, respectively. **b** A topology model predicted by PRED-TMR algorithm (Pasquier *et al*., 1999). Truncation sites of CrdS to generate cm-CrdS are marked at sites indicated by *black arrows* and *black hexagons* to yield pET-32_cm-CrdS (residues 96–375) and SUMO_cm-CrdS (residues 96–376). The NH_2-terminal residues in pET-32a_cmCrdS and SUMO_CrdS-cm are arginines (R), while the COOH-terminal residues in pET-32a_cmCrdS and SUMO_CrdS-cm are glycine (G) and leucine (L), respectively. The putative catalytic and the QRTRW motifs are coloured in *magenta* and *yellow*, respectively. The acidic, basic, aromatic and proline residues in membrane helices are shown in *red*, *blue*, *green* and *cyan*, respectively. The topology map was drawn with TOPO (http://www.sacs.ucsf.edu/TOPO/topo.html)


a
MH1
MH2
MH3
CDD motif
DAD motif
TED motif
QRTRW motif
MH4
MH5
MH6
MH7


b
Periplasmic
Cytoplasm

(200 rpm/min) shaken (1,250 rpm/min; 10°C; 20 min). The soluble supernatant and SDS-solubilised cell fractions were evaluated by the immuno-blot (4 µl of fractions were applied), SDS-PAGE (20 µl) and Western immuno-blot (5 µl) analyses.

Large scale production of inclusion bodies

The Origami B (DE3) (pET-32a_cm-CrdS fusion, constructs 1 and 2) and BL21*(DE3) (SUMO_cm-CrdS fusion) cell lines that produced the highest yields of chimeric cm-CrdS proteins were used for large scale production of inclusion bodies. Cells from the glycerol stocks were inoculated into 50 ml of LB, which also contained the required antibiotics for selection of the cells specified above, and were grown with orbital shaking (160 rpm, 37°C, 12 h), when the cells typically entered a stationary growth phase. Approximately 6 ml cells were added into 400 ml LB media and the cell growth continued under shaking (160 rpm, 37°C, 2.5 h) until the cells reached OD_{600nm}=0.6. The cells in the exponential growth phase were treated by 20 mM benzyl alcohol to induce internal chaperone protein production and incubation continued in the presence of benzyl alcohol (160 rpm, 28°C, 30 min). The cells were subsequently induced by 0.4 mM IPTG for 8 h (SUMO_cm-CrdS construct) and 12 h (pET-32a_cm-CrdS constructs) under orbital shaking (160 rpm, 28°C). The cells were then washed by centrifugation (3,500 g, 4°C, 20 min), resuspended in fresh LB media and allowed to grow for another 2 h (160 rpm, 28°C) before harvesting by centrifugation (3,500 g, 4°C, 20 min). The cells were stored at −20°C in the form of dried pellets.

Purification of inclusion bodies

Pellets from 400 ml LB media [Origami B (DE3) for the pET-32a_cm-CrdS constructs 1 and 2, and BL21*(DE3) or the SUMO_cm-CrdS construct] of 1 g wet weight, prepared as described in the previous paragraph, were gently thawed, suspended in 50 ml sonication buffer (50 mM Tris-HCl pH 8, containing 150 mM NaCl, 2 mM NaF, 1 mM EDTA, 0.1% v/v Triton X-100, 5% v/v glycerol, 1 mM TCEP, 2 mM PMSF, 0.1 mM UDP-Glc and one EDTA-free Complete Protease Inhibitor Cocktail Tablet/50 ml of sonication buffer) and applied to an EmulsiFlex-C5 homogenizer (Avestin Inc., Ottawa, Ontario, Canada). The cells were homogenized three times with intermittent cooling and care was taken to prevent frothing. The cell homogenate was spun (15,000 g, 4°C; 30 min), washed once by centrifugation (15,000 g, 4°C; 30 min) in 100 ml sonication buffer, excluding Triton X-100 and including 125 mM NDSB-201 (MP Biomedicals, Solon, OH, USA), and twice in 100 ml sonication buffer lacking Triton X-100 and NDSB-201. The purified inclusion bodies were stored at −20°C in the form of dried pellets.

Refolding of inclusion bodies and refolding optimization

Inclusion bodies, from 100 ml cultures of *E. coli* Origami B (DE3) and BL21*(DE3) cells, prepared as described above, were suspended in 5 ml sonication buffer, excluding Triton X-100 but containing freshly prepared 8 M urea (solubilisation buffer). The protein solution was gently mixed on an orbital shaker (100 rpm) for 2 h at 4°C, spun (15,000 g, 15°C, 30 min) and filtered through the Millex-GV filters (Millipore) with a pore size of 0.22 µm to remove non-solubilised protein aggregates. Solubilised inclusion bodies were adjusted to 5 mg/ml concentration and 10 µl aliquots were distributed in the wells of the flat-bottom 96-well tissue Sarstedt plates (Sarstedt AG & Co., Nümbrecht, Germany) that contained 165 µl of individual unique refolding solutions of iFOLD® Protein Refolding System 2, where the condition 91 (water) is referred to as a negative control. The configuration of the refolding system shown in Fig. S1. After the inclusion bodies were distributed in 96-well tissue plates (3×96 aliquots were typically distributed in 20 min), the plates were sealed and incubated under orbital shaking (160 rpm, 16°C, 5 days). The progress of refolding of inclusion bodies was followed by recording A_{340}, as a measure of protein precipitation (PolarStar Optima Plate Reader, BMG LabTech, Offenburg, Germany). All A_{340} measurements were done in duplicate after 20 min, 8 h and every 24 h up to 5 days, after the refolding trial was set up.

On-column refolding and purification of the SUMO_cm-CrdS protein

About 10 mg of inclusion bodies of the SUMO_cm-CrdS protein solubilised in 3 ml of the sonication buffer, containing 8 M urea and 5 mM imidazole, were added to 3 ml of the Talon resin, equilibrated in 25 mM CHES (*N*-cyclohexyl-2-aminoethanesulfonic acid) buffer, pH 8 containing 8 M urea and 5 mM imidazole. The mixture was incubated for 3 h at 4°C and packed into the 10×1 cm disposable column (BioRad Laboratories, Hercules, CA, USA). The column was washed in 25 mM CHES buffer, pH 8 containing 8 M urea and 5 mM imidazole, until stable A_{260} was recorded, indicating that no unbound protein was eluted. The SUMO_cm-CrdS protein bound to the Talon resin was refolded on-column using a descending gradient of urea (0 and 8 M urea, 100 ml each) in 25 mM CHES, pH 8, containing 0.032 M NDSB-256 and 5 mM imidazole. The elution proceeded for 16 h at 4°C at a flow rate of 0.2 ml/min. After the gradient elution was completed, the Talon resin was washed by 25 mM CHES,

pH 8, lacking urea, but containing 5 mM imidazole, until stable A_{260}. Subsequently, 5–500 mM imidazole gradient (20 ml each) was applied in 25 mM CHES, pH 8 at a flow rate of 0.7 ml/min. The 1.5 ml fractions were collected and evaluated by the immuno-blot (4 µl of fractions), SDS-PAGE (20 µl) and Western immuno-blot (5 µl) analyses. The fractions containing a near homogenous refolded SUMO_cm-CrdS protein were concentrated, washed in 50 mM Tris-HCl, pH 8 and stored at 4°C.

Proteolytic digestion of the SUMO_cm-CrdS protein by SUMO-protease

The refolded and purified SUMO_cm-CrdS protein (10 µg) was suspended in 50 µl 50 mM Tris-HCl, pH 8 containing 0.2% (w/v) NP-40 (Invitrogen), 1 mM DTT and 0.15 M NaCl, and after mixing, 3 U of SUMO-protease were added. The mixture was incubated under shaking (100 rpm, 14°C, 48 h). Aliquots were taken at 8, 24 and 48 h, analysed by Western immuno-blots (5 µl), or mixed with the equal volumes of the SDS-PAGE loading buffer [29] and analysed by SDS-PAGE (20 µl).

Steady-state circular dichroism (CD) spectroscopy of SUMO_cm-CrdS

Far-UV CD spectra (180–260 nm) with a step size of 0.5 nm (1 nm slit and exit parameters) with refolded and purified SUMO_cm-CrdS were recorded at ambient temperature (approximately 21°C) in a Pi-180 spectropolarimeter (Applied PhotoPysics, Surrey, UK) using an adaptive sampling method and a quartz cell of a path length of 1 mm (Hellma GmbH & Co. KG, Müllheim, Germany) under constant nitrogen flush of five liters per min. The protein was concentrated to 0.24 mg/ml (12 µM SUMO_cm-CrdS) in 10 mM Tris-HCl, pH 8 containing 1 mM mercaptoethanol and the spectra of refolded SUM_cm-CrdS were measured without and with 3.1 M urea. Background spectra containing the same buffer of the same composition but lacking protein were subtracted to obtain the baseline-corrected spectra. The data are plotted as values of delta epsilon ($\Delta\varepsilon$) against wavelength (λ).

Tryptic mapping by MALDI MS, MS/MS and MALDI-TOF/TOF spectrometry

About 5–10 µg cm-CrdS with and without a SUMO folding component, and the pET-32a_cm-CrdS were S-amidomethylated, digested with 100 ng sequencing grade trypsin (Promega) in 5 mM ammonium bicarbonate and concentrated to 5 µl. The digest (0.5 µl) was applied to a 600 µm AnchorChip (Bruker Daltonik GmbH, Bremen, Germany) according to the α-cyano-4-hydroxycinnamic acid (HCCA, Bruker Daltonik) thin-layer method. MALDI TOF mass spectra were acquired using a Bruker Ultraflex III MALDI TOF/TOF mass spectrometer (Bruker Daltonik GmbH) operating in reflectron mode under the control of FlexControl, version 3.0 (Bruker Daltonik GmbH). External calibration was performed using peptide standards (Bruker Daltonik GmbH) that were analysed under the same conditions. Spectra were obtained at random locations over the surface of the matrix spot at an intensity determined by the operator. During data acquisition for MS/MS, four of the most highly abundant non-trypsin and non-keratin sample ions were selected as precursors for MS/MS analysis. MALDI-TOF/TOF was performed in the LIFT mode using the same spot on the target. Data processing for all spectrometry data proceeded with MS and MS/MS spectra that were subjected to smoothing, background subtraction and peak detection using FlexAnalysis, version 3.1 (Bruker Daltonik GmbH), and the spectra and mass lists were exported to BioTools (Bruker Daltonik GmbH). The MS and corresponding MS/MS spectra were combined and submitted to the Mascot database-search (http://www.matrixscience.com). The specifications: Taxonomy-viridiplantae, Database-NCBI non-redundant 20071013, enzyme-trypsin, fixed, modifications-carbamidomethyl (C), variable modifications-oxidation (M), mass tolerance MS-50 ppm, MS/MS tolerance-0.5 Da, missed cleavages-1.

Protein determination, SDS-PAGE and western immuno-blot analyses

Protein determination, SDS-PAGE and protein detection on SDS-PAGE gels with colloidal Coomassie Brilliant Blue G-250 were performed following standard procedures [30]. Western immuno-blot analyses were performed with 0.22 µm nitrocellulose blotting membranes (Millipore), a mouse monoclonal anti-polyHistidine-alkaline phosphatase IgG2a isotype antibody and the BCIP/NBT-purple liquid reagent for membranes, as suggested by the manufacturer (Sigma).

Structural predictions of CrdS

The topology of the CrdS sequence was investigated by the HMMTOPv2.0 [31], TOPPRED [32], DAS [33], TMHMMv2.0 [34], MEMSAT [35], PHD [36], Kyte-Doolittle Hydrophobicity Tool [37] and PRED-TMR [38] programs. The positions of secondary structural elements and hydrophobic clusters were manually examined with hydrophobic cluster analysis (HCA) [39]. The distributions of secondary structural elements in CrdS were predicted through the Structure Prediction Meta-Server [40], SeqAlert (Bioinformatics and Biological Computing, Weizmann

Institute of Science, Israel) and 3D-PSSM (Imperial College of Science, Technology & Medicine, London, UK).

Construction of a 3D model of cm-CrdS

The crystal structure of the A2 domain of chondroitin polymerase from *E. coli* was used as a structural template (PDB accession number 2z86:A) [23]. The template structure 2z86:A2 was aligned with cm-CrdS using PROMALS3D [41] and the alignment was checked manually to maintain the integrity of secondary structural elements, using Robetta [42], 3D-Jury [40] and HCA [39]. The structurally aligned sequences were used as input parameters to generate 3D models of cm-CrdS in complex UDP-Glc and Mn, using Modeller 9v5 [43] running Ubuntu 8.10 on a Linux station. The optimal model of cm-CrdS was selected from 40 models. The model with the lowest value of the Modeller 9v5 objective function was assessed by a DOPE module of Modeller 9v5, and that with the most favourable parameters was chosen for loop optimisation, using Modeller 9v5. The final cm-CrdS model with optimised loop geometry was chosen from 60 models and evaluated. The overall G-factors (estimates of stereochemical parameters) evaluated by PROCHECK [44] were 0.13 and −0.38 for 2z86:A2 and cm-CrdS model, respectively. The Z-score values [45] reflecting combined statistical potential energy, for 2z86:A2 and cm-CrdS were −7.34 and −6.07, respectively. The 'Iterative Magic Fit' algorithm, using combined PAM matrix and iterative structural alignment, in the DeepView browser [46] was used to determine the rmsd values in Cα positions between 2z86:A2 (273 residues) and cm-CrdS (281 residues); the rmsd value was 0.74 Å over 248 residues. Electrostatic potentials were calculated by Adaptive Poisson-Boltzmann Solver (the dielectric constants of solvent and protein were 80 and 2, respectively) [47] implemented in PyMol (http://www.pymol.org) as a plugin, and mapped on protein molecular surfaces generated with a probe radius of 1.4 Å. Molecular graphics was generated with PyMol (http://www.pymol.org).

Results

Structural predictions of CrdS

A range of predictive tools suggested that CrdS, which has 654 amino acid residues [20], was an integral membrane protein with 7 to 9 helices and a single catalytic module (cm-CrdS) composed of 280 to 285 residues (Fig. 2). More specifically, HMMTOPv2.0 [31], TOPPRED [32], DAS [33], TMHMMv2.0 [34], MEMSAT [35], PHD [36], Kyte-Doolittle Hydrophobicity Tool [37] and PRED-TMR [38] programs indicated that CrdS contained 9, 8, 8, 7, 7, 7, 7 and 7 helices, respectively. Their positions were confirmed by HCA, a two-dimensional plot analysis that clusters hydrophobic residues and assigns regular secondary structures [39] (Fig. 2a). Here, only seven *bona fide* membrane-spanning helices flanking cm-CrdS were unambiguously identified in CrdS (Fig. 2b). The predictive tools further indicated that the NH_2-terminus of CrdS was oriented to the periplasmic space, while its COOH-terminus was envisaged to project to the intracellular space (Fig. 2b). The CrdS protein had three 20-residue helices spanning the membrane at the NH_2-terminal part of the sequence, followed by the cm-CrdS region located more-or-less centrally, while four 18–19 residue helices were positioned in the COOH-terminal region of cm-CrdS. The latter four helices symmetrically enclosed a 58-residue intracellular segment, while another 117-residue intracellular fragment formed the COOH-terminal part of CrdS (Fig. 2b). The positions of four highly conserved motifs in cm-CrdS comprised the putative catalytic and divalent cation-binding residues (CDD, DAD, TED), and a consensus QRTRW motif (Fig. 2b); the precise functionalities of these motifs are currently unknown [20].

Molecular model of cm-CrdS

Searches through the Structure Prediction Meta-Server [40], SeqAlert (Bioinformatics and Biological Computing, Weizmann Institute of Science, Israel) and 3D-PSSM (Imperial College of Science, Technology & Medicine, London, UK) benchmarks indicated that the crystal structure of the A2 domain of chondroitin polymerase from *E. coli* (PDB ID 2z86:A2) [23] was the best template for molecular modelling. This enzyme shared 21% and 62% positional sequence identity and similarity with cm-CrdS, respectively [48]. The chondroitin polymerase structural template 2z86:A2 also contained UDP-Glc and a divalent cation Mn. Evaluations of molecular models, based on stereochemical and energy parameters indicated that the constructed cm-CrdS model was of satisfactory quality. The model revealed α/β architecture, which included housed five well-formed helices and 11 β-sheets (Fig. 3a). The model of cm-CrdS allowed the disposition of interacting amino acid residues to be defined with respect to the positions of bound UDP-Glc and Mn. Here, both ligands were accommodated in a central cavity with an overall electro-negative charge, dictated by the presence of a quadruplet of aspartic residues that were suggested to participate in substrate binding and metal coordination (Fig. 3a). In the cavity, the positions of three highly conserved CDD, DAD and TED motifs indicated that they could be involved in binding of UDP-Glc and mediate catalysis. These residues had almost identical orientations compared to the corresponding CDD, DSD and AVD motifs of 2z86:A2 (Fig. 3b).

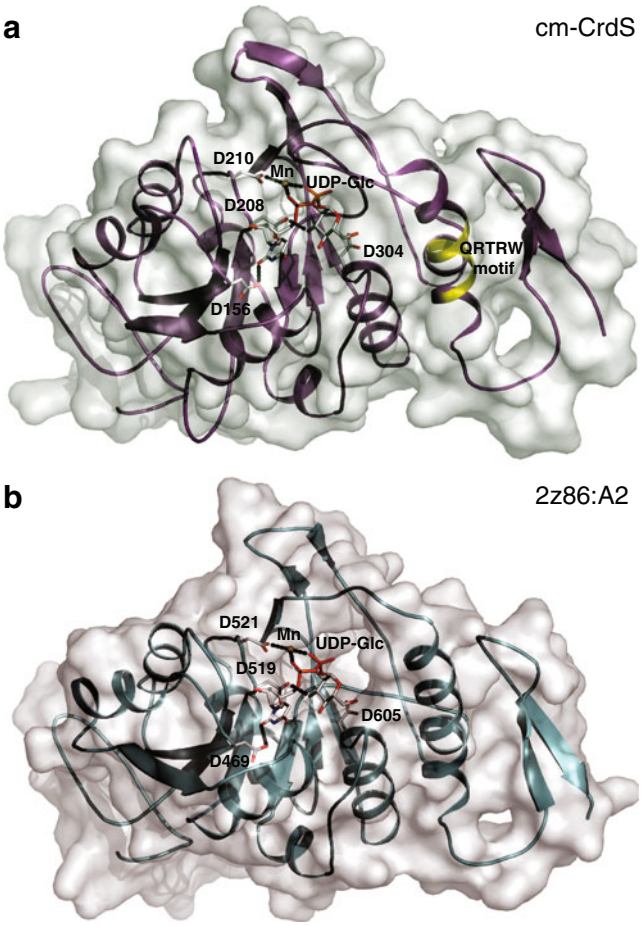


Fig. 3 Molecular model of cm-CrdS. **a** A molecular model of cm-CrdS (*magenta*), and **b** a crystal structure of the A2 domain of chondroitin polymerase 2z86:A2 (cyan), showing dispositions of secondary structure elements and surfaces. Cartoons show α/β folds and accessibilities of UDP-Glc (sticks in atomic colours) and Mn (*wheat spheres*) in central cavities. A root-mean-square-deviation value for 248 structurally aligned residues in cm-CrdS and 2z86:A2 is 0.74 Å over the Cα backbone positions. Four conserved aspartates (atomic colours) that coordinate Mn and UDP-Glc are shown. The separations, marked by *dotted lines*, between Mn and the surrounding residues are within 2.2–2.9 Å (cm-CrdS) and 1.7–2.9 Å (2z86:A2). The position of the QRTRW motif in cm-CrdS is highlighted in *yellow* (*cf*. Fig. 2b)

Superposition of residues of 2z86:A2 and cm-CrdS indicated that the positions of the Mn-coordination sites were highly preserved (Fig. 3). Here, the two conserved aspartates D208 and D210 did not generate steric clashes, and formed close interactions with Mn at separations of 2.2–2.9 Å in cm-CrdS and of 1.7–2.9 Å in 2z86:A2. Other aspartates D156 and D304 from the CDD and TED motifs coordinated the glucosyl residue from UDP-Glc *via* a strong (2.6 Å) and a weaker (3.5 Å) hydrogen bond (Fig. 3). The cm-CrdS model could further predict disposition of a consensus QRTRW motif (Figs. 2 and 3a, highlighted in yellow), which formed a part of the helix surrounding the edge of a central cavity, where the four conserved catalytic aspartates were located.

Design and recombinant plasmid construction

Based on structural predictions and protein modelling (*vide supra*), truncation sites on the full-length sequence of CrdS were chosen allowing two types of DNA fusion constructs of cm-CrdS to be generated (Figs. 1 and 2b). The 1st truncation site was positioned behind the end of the 3rd membrane helix for all three constructs (constructs 1 and 2 of pET-32a_cm-CrdS and SUMO_cm-CrdS), or more precisely behind a hydrophobic segment LILL (marked by a left-hand arrow in Fig. 2b). The 2nd truncation site was positioned in front of the membrane helix 4, specifically, behind YLHG and YLHGL motifs for the pET-32_cm-CrdS and SUMO_cm-CrdS constructs, respectively (marked by two right hand arrows in Fig. 2b). The two expression systems pET-32(+) and Champion™ pET SUMO were chosen as the destination vectors containing respective thioredoxin and SUMO folding partners (both 109 residues), other affinity tags (6xHis and S-tag), proteolytic sites (enterokinase, thrombin and SUMO), and cloning overhanging sequences. In summary, the pET-32a_cm-CrdS (constructs 1 and 2) and SUMO_cm-CrdS proteins contained 447 and 403 residues respectively, from which 280 and 281 residues in each expression vector originated from cm-CrdS (Fig. 1).

Screening of protein expression with the pET-32a_cm-CrdS and SUMO_cm-CrdS vectors

The highest protein yields with the pET-32a_cm-CrdS constructs were recorded with the Origami (DE3) and Origami B (DE3) cell lines, whereas with the SUMO_cm-CrdS construct, the BL21 (DE3) and Rosetta-Gami cell lines produced the highest protein yields (Fig. 4). In all instances, the chimeric proteins formed inclusion bodies or protein aggregates and only negligible amounts of protein were observed in soluble fractions of the lysed *E. coli* cells (data not shown). However, it is well documented that protein solubility can be improved by varying expression conditions other than temperature. With a view forming the chimeric cm-CrdS in soluble forms, various IPTG concentrations and induction intervals were tested (Fig. 4). Extensive screening did not lead to significant increases of the soluble cm-CrdS proteins (data not shown). The experimental approach was therefore shifted to produce inclusion bodies in high yields that could subsequently be refolded (*vide infra*). It was observed that the expression of the pET-32a_cm-CrdS constructs in the Origami (DE3) cells was leaky insofar as protein expression was detected without the addition of IPTG, while in the Origami B (DE3) cells the expression of the pET-32a_cm-CrdS proteins was induced only by IPTG (Fig. 4). The same tight induction control by IPTG was also observed in the BL21 (DE3) and Rosetta-Gami cells using the SUMO_cm-

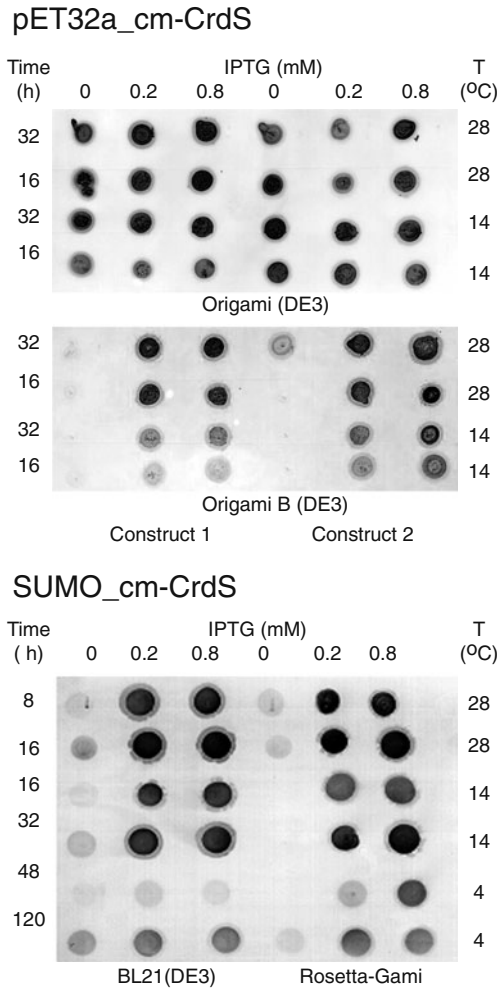


Fig. 4 Immuno-blot analyses of cm-CrdS produced by the pET-32a_cm-CrdS (constructs 1 and 2) and SUMO_cm-CrdS expression vectors. Proteins solubilised in 4% (w/v) SDS were evaluated by immuno-blot analyses (4 µl) using an anti-polyHistidine-alkaline phosphatase IgG2a isotype antibody and a BCIP/NBT-purple liquid substrate system, as described in " Materials and methods". Protein production was induced for 16 h and 32 h (pET-32a_cm-CrdS) and during 8 h to120 h (SUMO_cm-CrdS) at 4, 14 and 28°C in the specified *E. coli* cell lines

CrdS construct (Fig. 4). In both instances, decreases in temperatures to 14 or 4°C led to significant declines in protein production (Fig. 4). The highest yield of inclusion bodies was observed at 28°C with the Origami (DE3) and BL21(DE3) cell lines for the pET-32a_cm-CrdS and SUMO_cm-CrdS DNA constructs, respectively, using 0.8 mM IPTG (Fig. 4).

Refolding of the pET-32a_cm-CrdS and SUMO_cm-CrdS proteins and its optimisation

The relatively pure inclusion bodies as detected by SDS-PAGE (Fig. 5, lanes 4–6) were subjected to refolding screening, using iFOLD® Protein Refolding System 2 (Fig. S1) as follows. The inclusion bodies were distributed in the 96-well tissue plates and incubated at 4°C for up to 5 days. The progress of refolding was followed by quantifying A_{340} that is a measure of protein precipitation [49]. The progress of refolding was also followed by quantifying the amounts of soluble proteins by Bradford assays and by immuno-blots (Fig. S2). Here it was obvious that the inclusion bodies produced by the three DNA fusions followed very similar refolding patterns, where formulations 34, 59, 66, 72 and 81 yielded the highest amounts of refolded proteins (green circles in Fig. S2). As expected, no protein refolded in water (formulation 91; a red circle in Fig. S2). The amounts of solubilised proteins in the most favourable conditions for each DNA fusion were confirmed by SDS-PAGE, shown underneath immuno-blots (Fig. S2). Evaluations of the yields of the refolded proteins showed that formulations 59, 66, 72 and 81 were highly effective. In particular formulation 59 (50 mM CHES buffer, pH 9 with 0.5 M NDSB-256) produced high yields of soluble proteins (Fig. S2).

After the optimal refolding conditions were selected, the conditions were optimized further for all three constructs (Fig. S3). The condition 59 (50 mM CHES buffer, pH 9.0, 0.5 M NDSB-256) was modified in terms of the optimal pH value (pH 7, 8 and 9), NDSB-256 concentration (0.5, 0.25, 0.125, 0.064, 0.032 and 0 M) and the addition of 0.1 mM UDP-Glc that serves as a substrate for CrdS. During optimization it was observed that the original CHES concentration and its pH value (50 mM, pH 9) could be lowered (25 mM and pH

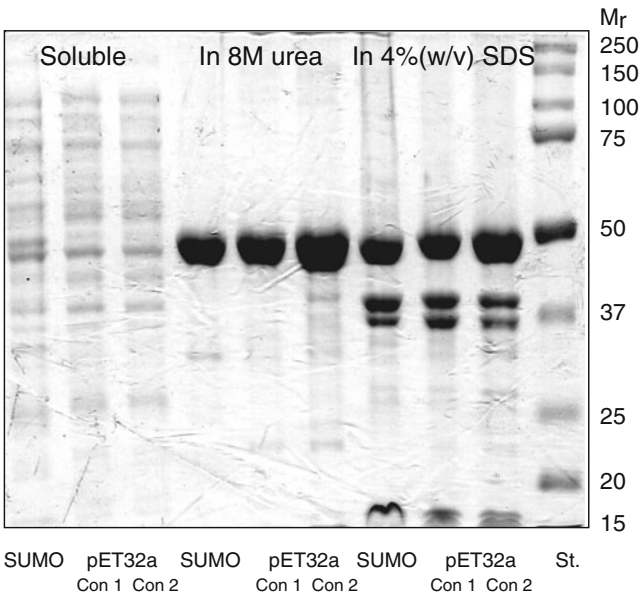


Fig. 5 SDS-PAGE of the purified SUMO_cm-CrdS and pET-32a_cm-CrdS inclusion bodies solubilised by 8 M urea. The preparations in lanes 1–9 represent fractions isolated during purification of inclusion bodies. Lanes 1–3, fractions soluble in the lysis buffer; lanes 4–6 fractions soluble in 8 M urea; lanes 7–9, fractions soluble in 4% (w/v) SDS. Approximately 5 µg (lanes 1–3) or 25 µg (lanes 6–9) of samples were applied. Molecular mass standards (St.) with M_r are indicated in lane 10

8) as well as the concentrations of NDSB-256 (from 0.5 M to 0.064–0.032 M), while the yields of refolded proteins remained unchanged. The addition of UDP-Glc further improved refolding by 5% (data not shown). Thus, the refolding best occurred after 8 h at 4°C using 0.064–0.032 M NDSB-256 in 25 mM CHES, pH 8 (Fig. S3).

On-column refolding and purification of the SUMO_cm-CrdS protein, and digestion by SUMO-protease

The chimeric His_6-tagged SUMO_cm-CrdS fusion protein was solubilised in 8 M urea, bound to the Talon resin with high efficacy, and refolded on-column, using a descending gradient of urea in 25 mM CHES buffer, pH 8 and containing 0.032 M NDSB-256 and 5 mM imidazole, over 16 h at 4°C at a low flow rate (0.2 ml/min). After the urea concentration was decreased to zero, the affinity resin was washed with 25 mM CHES buffer, lacking urea and NDSB-256, and subsequently the SUMO_cm-CrdS protein was eluted by IMAC using approximately 150 mM imidazole (Fig. 6a). Essentially homogenous SUMO_cm-CrdS protein was eluted by a one-step affinity purification procedure, as detected by the SDS-PAGE and Western immuno-blot analyses (Fig. 6b).

The near homogenous SUMO_cm-CrdS protein (molecular mass 45.4 kDa) was digested by a SUMO-protease during 8 h incubation at 14°C to produce two expected protein species cm-CrdS (32.0 kDA) and the SUMO folding partner protein (13.4 kDa) (Fig. 6c). The efficiency of proteolysis was estimated to be approximately 85–90% (Fig. 6c). The overall yield of the soluble SUMO_cm-CrdS protein *via* on-column refolding and IMAC purification was about 15%, that is 10 mg of inclusion bodies (from 400 ml of *E. coli* culture) solubilised in urea, produced about 1.5 mg of refolded SUMO_cm-CrdS.

Far-UV (spectrum of the refolded and purified SUMO_cm-CrdS protein

Far-UV CD spectroscopy of refolded and purified SUMO_cm-CrdS yielded a spectrum with well defined signal intensities in the 190–205 nm and 210–230 nm regions at the 0.24 mg/ml (12 μM) concentration in 10 mM Tris-HCl buffer, pH 8 (Fig. 7). The spectrum of the refolded

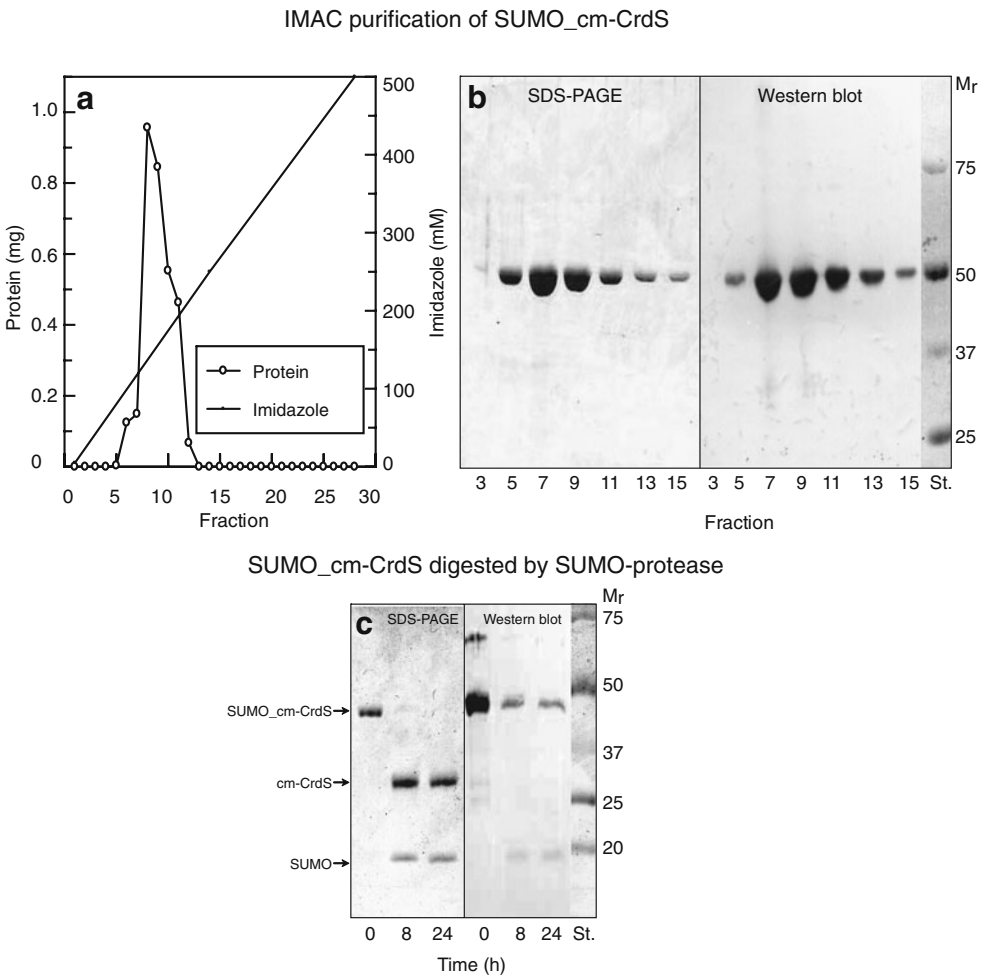


Fig. 6 On-column refolding and purification of the SUMO_cm-CrdS protein. **a** Elution profile of SUMO_cm-CrdS during purification with IMAC. The SUMO_cm-CrdS protein elution and imidazole gradient profiles are shown. **b** SDS-PAGE and Western immuno-blot analyses of the refolded and purified SUMO_cm-CrdS protein. Protein contents (fractions 3–15) eluted by the 0–500 mM imidazole gradient, are between 2–30 μg per lane. **c** SDS-PAGE and Western immuno-blot analyses of the SUMO_cm-CrdS protein (5 μg in lane 1) cleaved by SUMO-protease during the 0 h to 24 h time interval. Molecular mass standards (St.) and M_r are indicated in lanes 15 (**b**) and 7 (**c**)

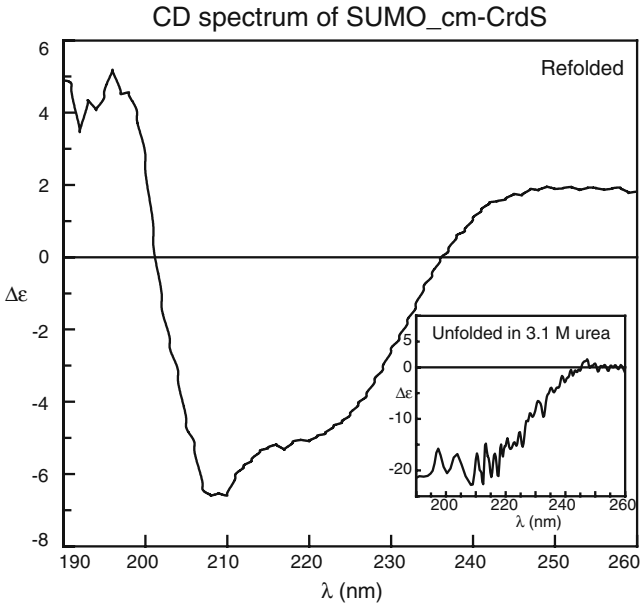


Fig. 7 Far-UV CD spectrum of the refolded SUMO_cm-CrdS protein. The spectrum shows well defined signal intensities in the 190–205 nm and 210–230 nm regions. The inset shows the CD spectrum of refolded SUMO_cm-CrdS in the presence of 3.1 M urea. CD spectrum represents the average of 5 replica scans. The data are plotted as delta epsilon (Δε) against wavelength (λ)

SUMO_cm-CrdS protein was also recorded after the addition of 3.1 M urea, where a noisy profile with undefined intensities in the 190–205 and 210–230 nm regions were observed (inset in Fig. 7).

MALDI-TOF/TOF spectrometry of the pET-32a_cm-CrdS, SUMO_cm-CrdS and cm-CrdS proteins

The purified and refolded recombinant pET-32a_cm-CrdS and SUMO_cm-CrdS proteins (Fig. 6a and b) and cm-CrdS, obtained after digestion by the SUMO-specific protease (Fig. 6c) were digested with trypsin. The molecular masses of individual tryptic fragments were subjected to MALDI-TOF/TOF spectrometry analyses and provided respective coverage of 43%, 31% and 59% of ions with expected *m/z* values (Table 1). Further, eleven, eight and eight peptide sequences (Table 1) deduced from the nucleotide sequences of the pET-32a_cm-CrdS, SUMO_cm-CrdS and SUMO-digested SUMO_cm-CrdS DNA fusions, respectively, matched perfectly with ions of expected *m/z* values during MS analyses.

Discussion

There are few structural studies of enzymes that mediate cell wall polysaccharide biosynthesis, yet additional structural data are needed for complete molecular descriptions of substrate donor and/or acceptor binding events during the catalysis of polysaccharide synthesis. This is particularly true for synthases of the GT2 family, which includes enzymes required for the synthesis of cellulose, but also of (1-3)-β-D-glucans and (1-3;1-4)-β-D-glucans [50]. Because of their membrane location, their relative lowly concentration in many tissues and the tendency of the enzymes to lose activity following tissue disruption, it has proved difficult to purify the synthases from biological material [2]. Protein expression from DNA fragments in heterologous systems offers an alternative route to the preparation of sufficient amounts of common polysaccharide synthases for physico-chemical, kinetic, enzymic and three-dimensional (3D) structural analyses. Nevertheless, successful studies of polysaccharide synthases are rare [51] and progress on the structural descriptions of β-D-glucan synthases has been slow. For these reasons, we have investigated the heterologous expression of full-length CrdS and cm-CrdS DNA fragments, with a view to obtaining sufficient protein for 3D structural and other analyses. Our previous attempts to express the full-length CrdS in a variety of vectors (*e.g.* pET15b, pBAD24, Champion™ pET SUMO) failed, because truncated forms of CrdS were generally produced *via* abortive translation (data not shown). Hence, we concentrated our efforts on the expression of cm-CrdS in a soluble form. Similar experimental approaches have suggested that truncated catalytic domains could successfully be expressed in native conformations, and some of these proteins have been crystallized [52–57].

Here, structural prediction analyses indicated that appropriate truncation sites in full-length CrdS could be found for the generation of a 'cytoplasmic' component or catalytic module (cm-CrdS) of the enzyme (Fig. 2). To this end, cm-CrdS, in contrast to full-length membrane-bound CrdS, might be expressed in a soluble form. Thus, two cm-CrdS DNA fusions were engineered in the vectors pET-32a(+) and Champion™ pET SUMO using ligation-based cloning (Fig. 1). The pET-32a and SUMO expression vectors contain a variety of affinity tags to assist with protein purification from crude extracts of *E. coli* cells and proteolytic sites to remove folding partner proteins. As for the differences between the constructs 1 and 2 of the pET-32a_cm-CrdS fusion generated here, constructs 1 and 2 contained Eco*RI* or Bam*HI* restriction sites, respectively, at their 5′ ends. These restriction sites preceded the coding regions of cm-CrdS (Fig. 1) and were translated into the EF and GS overhanging sequences at the amino acid level. We could therefore explore the effects of small variations on expression and refolding patterns of pET-32a fusions. Both expression vectors have been used previously for expression of recombinant proteins in folded forms [26–28].

Extensive screening of protein expression of cm-CrdS with at least ten different *E. coli* cell lines at various induction

Table 1 Amino acid sequences of tryptic fragments from the pET32a_cm-CrdS and SUMO_cm-CrdS proteins identified by MALDI-TOF/TOF. The sequences in cm-CrdS (32.0 kDa) were identified after SUMO_cm-CrdS (45.3 kDa) was digested by SUMO protease (*cf.* Fig. 6c)

pET32a_cm-CrdS	SUMO_cm-CrdS	cm-CrdS
GIPTLLLFK	GSASMSDSEVNQEAKPEVKPEVK	SIIAAQAMDYPR
HHSSGLVPR	QGKEMDSLRFLYDGIR	TYCEAAGVNYVTRPDNK
SIIAAQAMDYPRLRVFVCDDTR	IQADQTPEDLDMEDNDIIEAHREQ	RVTGLFSDPK
RVTGLFSDPK	IGGRR	VAVVQTPQFYFNSDPIQHNLGIDK
VAVVQTPQFYFNSDPIQHNLGIDK	SIIAAQAMDYPRLR	SFVDDQRVFFDDFQPAK
SFVDDQRVFFDDFQPAK	VAVVQTPQFYFNSDPIQHNLGIDK	RAAVNGIGGFPTDALTEDMLLTYR
DAVGCAFRVGTSFVVR	DAVGCAFRVGTSFVVR	WLNEKWSVGLSAEGVPEYITQR
AAVNGIGGFPTDALTEDMLLTYR	TRWCLGTIQIGLLRTGPLWR	GNFTLTQRLHYLHGLR
WLNEKWSVGLSAEGVPEYITQR		
WCLGTIQIGLLR		
GNFTLTQR		

temperatures and concentrations of IPTG, using the pET-32a and SUMO expression systems, showed that insoluble inclusion bodies were produced and only negligible amounts of protein were observed in soluble fractions. Prior to the IPTG induction, we also treated *E. coli* cells with non-lethal doses of benzyl alcohol to induce internal chaperone production by membrane fluidization [58, 59], and subsequently we allowed refolding of induced proteins in fresh media [59], but these treatments did not lead to the formation of soluble protein. Nevertheless, the insoluble inclusion bodies produced by both expression systems could be refolded efficiently using a range of buffers and non-detergent sulfobetaines (Fig. S2). The original refolding formulation 59 of iFOLD® Protein Refolding System 2 (50 mM CHES buffer, pH 9.0; 0.5 M NDSB-256) could further be optimized (25 mM CHES, pH 8.0, and 0.032–0.064 M NDSB-256) (Fig. S2). It was not surprising that the CHES (*N*-cyclohexyl-2-aminoethanesulfonic acid) buffer was found to be an optimal refolding constituent, because CHES is a known suppressor of protein aggregation, and it has been reported to refold aggregated or misfolded proteins [60]. However, in our hands a low concentration of a non-detergent sulfobetaine NDSB-256, in addition to the CHES buffer, was also needed to obtain soluble and stable chimeric proteins in high yields (Fig. S2). Similarly, other authors observed that pH and the type of detergent impacted on the efficiency of protein refolding [61, 62]. The SUMO_cm-CrdS protein could be separated from its SUMO folding partner protein using a specific SUMO-protease, and thereafter it remained in a soluble form (Fig. 6c). We also observed that the pET-32a_cm-CrdS and SUMO_cm-CrdS inclusion bodies could be refolded in the presence of 0.5 M arginine, methyl β-cyclodextrin and other cyclodextrins, which are also known aggregation suppressors [61], using either a dilution or dialysis refolding techniques. However, the yields of refolded proteins in these instances were lower (data not shown).

Similar refolding or renaturation studies have been performed with a range of soluble or membrane proteins that were notoriously difficult to express in native forms. For example, human antibody fragments [63], morphogenetic proteins [60], G-protein coupled receptors (GPCRs) [64] and glutamate receptors [65] have been refolded using high-throughput refolding trials and fractional folding screens [66]. Often, efficient on-column refolding techniques have allowed simultaneous renaturation and purification of proteins, where misfolded proteins refolded after their aggregation was prevented by immobilization [67, 68]. Refolding strategies for proteins in inclusion bodies have also been comprehensively explored by structural consortia using various high-throughput approaches [69]. These procedures allow the progress of refolding to be monitored, using various systems, in which coding regions of proteins are expressed as chimeric proteins or are co-expressed with chaperones in pro- and eukaryotic hosts [69, 70].

Selected biochemical and biophysical characteristics were investigated with the refolded SUMO_cm-CrdS protein. It was observed that its CD spectrum was significantly different from that of unfolded SUMO_cm-CrdS in the presence of 3.1 M urea, where in particular the signal in the region of 190–205 nm lost its spectral characteristics (Fig. 7). It has previously been observed that α-helices of proteins show CD maxima at 192 nm, minima at 208 nm and a shoulder at 222 nm, while β-sheets typically show a broad maximum at 218 nm [71]. According to these criteria, the spectrum of the refolded SUMO_cm-CrdS protein had significant secondary structure, and thus would contain both α-helices and β-sheets; this observation is in accordance with protein modelling. The identities of recombinant proteins were further confirmed by MALDI/TOF spectrometry. The mass spectra of the trypsin-digested pET-32a_cm-CrdS, SUMO_cm-CrdS proteins and the cm-CrdS module, released from its SUMO_cm-CrdS chimeric state, provided peptide

coverage equivalent to 43%, 31% and 59% of the total protein (Table 1). Hence, the experimentally determined amino acid sequence coverage confirmed that the expected full-length proteins were synthesized by the *E. coli* cells, and that no amino acid changes were detected in tryptic fragments.

Refolded and purified pET-32a_cm-CrdS and SUMO_cm-CrdS, and cm-CrdS were examined for catalytic activities. We were unable to detect any activity with either of the truncated refolded proteins, using traditional assays for measurement of the (1,3)-β-D-glucan synthase activity in the presence or absence of acceptor substrates [10, 14]. A range of reasons could be considered here, in particular the absence of anchoring membrane helices that might contribute to improper packing of secondary structure elements during catalytic cycles of these enzymes, the lack of an appropriate electro-chemical membrane potential that can influence directly or indirectly the activities of polysaccharide synthases [72] or vectorial transport of the (1-3)-β-D-glucan polymer [9, 73]. Thus, if the proper environment for membrane catalysts is not provided, the membrane-associated enzymes may not be catalytically competent. Other possibilities are that certain polysaccharide synthases may require the presence of associated helper proteins for catalysis [73], or that the absence of other genes (*crdA*, *crdC, crdR, pss$_{AG}$*) leads to catalytic incompetence in CrdS or cm-CrdS. Finally, it has been reported on a number of occasions that activation of (1,3)-β-D-glucan synthases in higher plants might require post-translational modification of the nascent protein [2, 18, 74].

To investigate the functional significance of highly conserved aspartic residues and of the QRTRW motif in cm-CrdS, a molecular modelling study of cm-CrdS was initiated (Fig. 3). The model of cm-CrdS in complex with UDP-Glc and cationic Mn revealed that the mixed α/β fold of cm-CrdS could accommodate five well-formed helices and 11 β-strands (Fig. 3a). The model of cm-CrdS allowed dispositions of interacting amino acid residues to be defined with respect to bound UDP-Glc and Mn. Both ligands were located in a central electro-negatively charged cavity, where D156, D208 and D210 in the two conserved CDD and DAD motifs contributed to the binding of UDP-Glc and formed a coordination complex Mn (Fig. 3). On the other hand, the 4th aspartate D304 from the TED motif was predicted to fulfill the role of a general base catalyst in the inverting catalytic mechanism. The latter prediction was in agreement with Osawa *et al*. [23], who suggested that in the crystal structure of the chondroitin polymerase from *E. coli* (PDB 2z86:A2), D605 from the AVD motif could play the role of a general base catalyst in the inverting catalytic mechanism adopted by the enzyme. The corresponding residue to D605 in CrdS was D304 in the TED motif. In addition to the roles of catalytic aspartates, a putative function of the consensus QRTRW motif in cm-CrdS, conserved in the glycosyltransferases of the GT2 family, was predicted (Fig. 3), contrary to previous studies [23]. Based on our structural predictions we suggest that the QRTRW motif could easily be exposed to incoming acceptor substrates and in particular the RT motif could directly, or *via* supporting water molecules, be involved in acceptor binding events.

In conclusion, suitable procedures have been developed that lead to high yields of inclusion bodies of cm-CrdS in free and chimeric forms in pET32a (+) and Champion™ pET SUMO vectors destined for expression in *E. coli*. These inclusion bodies could be refolded in a range of buffers and non-detergent sulfobetaines and free cm-CrdS could subsequently be released by proteolytic digestion. The MALDI/TOF analyses and CD spectroscopy revealed that expected full-length proteins were synthesized in native-like conformations. We are currently attempting to crystallize these proteins for future structural studies. We expect that these rapid and efficient procedures developed here could be used for generation of recombinant proteins, particularly, where large quantities of proteins are required for downstream applications.

Acknowledgements This work was supported by grants from the Australian Research Council (to GBF, BAS and MH). We are grateful to Dr Vilma Stanisich from La Trobe University (Australia) for providing pVS1512 plasmid and to Dr Christopher J. Bagley from the Adelaide Proteomics Centre (Australia) for mass spectrometry analyses.

References

1. Stone, B.A., Clarke, A.E.: The chemistry and biochemistry of (1, 3)-β-glucans. La Trobe University Press Melbourne, Australia (1992)
2. Stone, B.A., Jacobs, A.K., Hrmova, M., Burton, R.A., Fincher, G. B.: The biosynthesis of plant cell wall and related polysaccharides by enzymes of the GT2 and GT48 families. In: Ulvskov, P. (ed.) Plant Polysaccharides Series: Annual Plant Reviews. Blackwell, Danvers (2010). in the press
3. Mcintosh, M., Stone, B.A., Stanisich, V.A.: Curdlan and other bacterial (1, 3)-β-d-glucans. Appl. Microbiol. Biotechnol. **68**, 163–173 (2005)
4. Harada, T., Harada, A.: Curdlan and succinoglycan. In: Dumitriu, S. (ed.) Polysaccharides in Medical Applications, pp. 21–57. Dekker, New York (1996)
5. Nakanishi, I., Kimura, K., Kusui, S., Yamazaki, E.: Complex formation of gel-forming bacterial (1, 3)-β-d-glucans (curdlan-type polysaccharides) with dyes in aqueous solution. Carbohydr. Res. **32**, 47–52 (1974)
6. Saito, H., Misaki, A., Harada, T.: A comparison of the structure of curdlan and pachyman. Agric. Biol. Chem. **32**, 1261–1269 (1968)
7. Stasinopoulos, S.J., Fisher, P.R., Stone, B.A., Stanisich, V.A.: Detection of two loci involved in (1→3)-β-glucan (curdlan) biosynthesis by *Agrobacterium* sp. ATCC31749, and comparative sequence analysis of the putative curdlan synthase gene. Glycobiology **9**, 31–41 (1999)

8. Leigh, J.A., Coplin, D.L.: Exopolysaccharides in plant-microbe interactions. Annu. Rev. Microbiol. **46**, 307–346 (1992)
9. Karnezis, V., Fischer, H.C., Neumann, G.M., Stone, B.A., Stanisich, V.A.: Cloning and characterization of the phosphatidylserine synthase gene of *Agrobacterium* sp. strain ATCC 31749 and effect of its inactivation on production of high-molecular-mass (1→3)-β-D-glucan (curdlan). J. Bacteriol. **184**, 4114–4123 (2002)
10. Cabib, E., Bowers, B., Roberts, R.L.: Vectorial synthesis of a polysaccharide by isolated plasma membranes. Proc. Natl Acad. Sci. USA **80**, 3318–3321 (1983)
11. Hrmova, M., Taft, C.S., Selitrennikoff, C.P.: (1, 3)-β-d-Glucan synthase of Neurospora crassa: Partial purification and characterization of solubilized enzyme activity. Exp. Mycol. **13**, 129–139 (1989)
12. Kuribayashi, I., Morita, T., Mitsui, T., Igaue, I.: Purification of β-glucan synthase II from suspension-cultured rice cells. Biosci. Biotech. Biochem. **57**, 682–684 (1993)
13. Kudlicka, K., Brown, R.M.: Cellulose and callose biosynthesis in higher plants. 1. Solubilization and separation of (1,3)- and (1,4)-β-glucan synthase activities from mung bean. Plant Physiol. **115**, 643–656 (1997)
14. Pelosi, L., Imai, T., Chanzy, H., Heux, L., Buhler, E., Bulone, V.: Structural and morphological diversity of (1-3)-β-D-glucans synthesized in vitro by enzymes from *Saprolegnia monoica*. Comparison with a corresponding *in vitro* product from blackberry (*Rubus fruticosus*). Biochemistry **42**, 6264–6274 (2003)
15. Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V., Henrissat, B.: The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. Nucleic Acids Res. **37**, D233–D238 (2009)
16. Coutinho, P.M., Deleury, E., Davies, G.J., Henrissat, B.: An evolving hierarchical family classification for glycosyltransferases. J. Mol. Biol. **328**, 307–317 (2003)
17. Li, J., Burton, R.A., Harvey, A.J., Hrmova, M., Wardak, A.Z., Stone, B.A., Fincher, G.B.: Biochemical evidence linking a putative callose synthase gene with (1, 3)-β-d-glucan biosynthesis in barley. Plant Mol. Biol. **53**, 213–225 (2003)
18. Brownfield, L., Wilson, S., Newbigin, E., Bacic, T., Reid, S.: Molecular control of the glucan synthase-like protein NaGSL1 and callose synthesis during growth of *Nicotiana alata* pollen tubes. Biochem. J. **414**, 43–52 (2008)
19. Brownfield, L., Ford, K., Doblin, M.S., Newbigin, E., Reid, S., Bacic, T.: Proteomic and biochemical evidence links the callose synthase in Nicotiana alata pollen tubes to the product of the NaGSL1 gene. Plant J. **52**, 147–156
20. Karnezis, T., Epa, V.C., Stone, B.A., Stanisich, V.A.: Topological characterization of an inner membrane (1→3)-ß-D-glucan (curdlan) synthase from *Agrobacterium* sp. strain ATCC31749. Glycobiology **13**, 693–706 (2003)
21. Ceroni, A., Passerini, A., Vullo, A., Frasconi, P.: DISULFIND: a disulfide bonding state and cysteine connectivity prediction server. Nucleic Acids Res. **34**, W177–W181 (2006)
22. Charnock, S.J., Davies, G.J.: Structure of the nucleotide-diphospho-sugar transferase, SpsA from *Bacillus subtilis*, in native and nucleotide-complexed forms. Biochemistry **18**, 6380–6385 (1999)
23. Osawa, T., Sugiura, N., Shimada, H., Hirooka, R., Tsuji, A., Shirakawa, T., Fukuyama, K., Kimura, M., Kimata, K., Kakuta, Y.: Crystal structure of chondroitin polymerase from *Escherichia coli* K4. Biochem. Biophys. Res. Commun. **378**, 10–14 (2009)
24. Stewart, E.J., Aslund, F., Beckwith, J.: Disulfide bond formation in the *Escherichia coli* cytoplasm: an *in vivo* role reversal for the thioredoxins. EMBO J. **17**, 5543–5550 (1998)
25. Tsunoda, Y., Sakai, N., Kikuchi, K.: Improving expression and solubility of rice proteins produced as fusion proteins in *Escherichia coli*. Protein Expr. Purif. **42**, 268–277 (2005)
26. Malakhov, M.P., Mattern, M.R., Malakhov, O.A., Drinker, M., Weeks, S.D., Butt, T.R.: SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. J. Struct. Funct. Genomics **5**, 75–86 (2004)
27. Butt, T.R., Edavettal, S.C., Hall, J.P., Mattern, M.R.: SUMO fusion technology for difficult-to-express proteins. Protein Exp. Purif. **43**, 1–9 (2005)
28. Zuo, X., Li, S., Hall, J., Mattern, M.R., Tran, H., Shoo, J., Tan, R., Weiss, S.R., Butt, T.R.: Enhanced expression and purification of membrane proteins by SUMO fusion in *Escherichia coli*. J. Struct. Funct. Genomics **6**, 103–111 (2005)
29. Laemmli, U.K.: Cleavage of structural proteins during assembly of the head of bacteriophage T4. Nature **227**, 680–685 (1970)
30. Hrmova, M., Farkas, V., Harvey, A.J., Lahnstein, J., Wischmann, B., Kaewthai, N., Ezcurra, I., Teeri, T.T., Fincher, G.B.: Substrate specificity and catalytic mechanism of a xyloglucan xyloglucosyl transferase HvXET6 from barley (*Hordeum vulgare* L.). FEBS J. **276**, 437–456 (2009)
31. Tusnady, G.E., Simon, I.: The HMMTOP transmembrane topology prediction server. Bioinformatics **17**, 849–850 (2001)
32. von Heijne, G.: Membrane protein structure prediction: hydrophobicity analysis and the 'positive inside' rule. J. Mol. Biol. **225**, 487–494 (1992)
33. Cserzo, M., Wallin, E., Simon, I., von Heijne, G., Elofsson, A.: Prediction of transmembrane alpha-helices in procariotic membrane proteins: the dense alignment surface method. Protein Eng. **6**, 673–676 (1997)
34. Krogh, A., Larsson, B., von Heijne, G., Sonnhammer, E.L.L.: Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. J. Mol. Biol. **305**, 567–580 (2001)
35. Jones, D.T., Taylor, W.R., Thornton, J.M.: A model recognition approach to the prediction of all-helical membrane protein structure and topology. Biochemistry **33**, 3038–3049 (1994)
36. Rost, B.: PHD: predicting 1D protein structure by profile based neural networks. Meth. Enzymol. **266**, 525–539 (1996)
37. Kyte, J., Doolittle, R.: A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157**, 105–132 (1982)
38. Pasquier, C., Promponas, V.J., Palaios, G.A., Hamodrakas, J.S.: A novel method for predicting transmembrane segments in proteins based on a statistical analysis of the SwissProt database: the PRED-TMR algorithm. Protein Eng. **12**, 381–385 (1999)
39. Callebaut, I., Labesse, G., Durand, P., Poupon, A., Canard, L., Chomilier, J., Henrissat, B., Mornon, J.P.: Deciphering protein sequence information through hydrophobic cluster analysis (HCA): current status and perspectives. Cell. Mol. Life Sci. **53**, 621–645 (1997)
40. Ginalski, K., Elofsson, A., Fischer, D., Rychlewski, L.: 3D-Jury: a simple approach to improve protein structure predictions. Bioinformatics **19**, 1015–1018 (2003)
41. Pei, J., Grishin, N.V.: PROMALS: towards accurate multiple sequence alignments of distantly related proteins. Bioinformatics **23**, 802–808 (2007)
42. Kim, D.E., Chivian, D., Baker, D.: Protein structure prediction and analysis using the Robetta server. Nucleic Acids Res. **32** Suppl. 2, W256-231 (2004)
43. Sali, A., Blundell, T.L.: Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. **234**, 779–815 (1993)
44. Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M.: PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. **26**, 283–291 (1993)

45. Sippl, M.J.: Recognition of errors in three-dimensional structures of proteins. Proteins **17**, 355–362 (1993)
46. Guex, N., Peitsch, M.C.: Swiss-model and the Swiss-PdbViewer: An environment for comparative protein modeling. Electrophoresis **18**, 2714–2723 (1997)
47. Baker, N.A., Sept, D., Joseph, S., Holst, M.J., McCammon, J.A.: Electrostatics of nanosystems: application to microtubules and the ribosome. Proc. Natl Acad. Sci. USA **98**, 10037–10041 (2001)
48. Smith, T.F., Waterman, M.S.: Identification of common molecular subsequences. J. Mol. Biol. **147**, 195–197 (1981)
49. Vincentelli, R., Canaan, S., Campanacci, V., Valencia, C., Maurin, D., Frassinetti, F., Scappucini-Calvo, L., Bourne, Y., Cambillau, C., Bignon, C.: High-throughput automated refolding screening of inclusion bodies. Protein Sci. **13**, 2782–2792 (2004)
50. Hrmova, M., Fincher, G.B.: Plant and microbial enzymes involved in the depolymerization of (1,3)-β-D-glucans and related polysaccharides. In: Bacic, T., Fincher, G.B., Stone, B.A. (eds.) Chemistry, biochemistry and biology of (1,3)-β-D-glucans and related polysaccharides. Elsevier, Philadelphia (2009)
51. Ihara, Y., Takeda, T., Sakai, F., Hayashi, Y.: Transferase activity of GhCesA2 (putative cotton cellulose 4-β-glucosyltransferase) expressed in *Pichia pastoris*. J. Wood Sci. **48**, 425–428 (2002)
52. Lamani, E., McPherson, D.T., Hollingshead, S.K., Jedrzejas, M.J.: Production, characterization, and crystallization of truncated forms of pneumococcal surface Protein A from *Escherichia coli*. Protein Expr. Purif. **20**, 372–378 (2000)
53. Yanagisawa, T., Ishii, R., Fukunaga, R., Nureki, O., Yokoyama, S.: Crystallization and preliminary X-ray crystallographic analysis of the catalytic domain of pyrrolysyl-tRNA synthetase from the methanogenic archaeon *Methanosarcina mazei*. Acta Crystallogr. **F62**, 1031–1033 (2006)
54. Zhao, L., Tang, H.: Crystallographic characterization of a multidomain histidine protein kinase from an essential two-component regulatory system. Acta Crystallogr. **F65**, 346–349 (2009)
55. Jin, Y., Hattori, M., Nisimasu, H., Ishitani, R., Nureki, O.: Crystallization and preliminary X-ray diffraction analysis of the truncated cytosolic domain of the iron transporter FeoB. Acta Crystallogr. **F65**, 784–787 (2009)
56. Guilfoyle, A., Maher, M.J., Rapp, M., Clarke, R., Harrop, S., Jormakka, M.: Structural basis of GDP release and gating in G protein coupled Fe2+ transport. EMBO J. **28**, 2677–2685 (2009)
57. Tenno, T., Goda, N., Tateishi, Y., Tochio, H., Mishima, M., Hayashi, H., Shirakawa, M., Hiroaki, H.: High-throughput construction method for expression vector of peptides for NMR study suited for isotopic labeling. Protein Eng. Des. Sel. **17**, 305–314 (2004)
58. Shigapova, N., Török, Z., Balogh, G., Goloubinoff, P., Vígh, L., Horváth, I.: Membrane fluidization triggers membrane remodeling which affects the thermotolerance in *Escherichia coli*. Biochem. Biophys. Res. Commun. **328**, 1216–1223 (2005)
59. de Marco, A., Vigh, A., Diamant, S., Goloubinoff, P.: Native folding of aggregation-prone recombinant proteins in *Escherichia coli* by osmolytes, plasmid- or benzyl alcohol-overexpressed molecular chaperones. Cell Stress Chaperones **10**, 329–339 (2005)
60. Vallejo, L.F., Rinas, U.: Optimized procedure for renaturation of recombinant human bone morphogenetic protein-2 at high protein concentration. Biotechnol. Bioeng. **85**, 601–609 (2004)
61. Armstrong, N., de Lencastre, A., Gouaux, E.: A new protein folding screen: Application to the ligand binding domains of a glutamate and kainate receptor and to lysozyme and carbonic anhydrase. Protein Sci. **8**, 1475–1483 (1999)
62. Benoit, I., Coutard, C., Oubelaid, R., Asther, M., Bignon, C.: Expression in *Escherichia coli*, refolding and crystallization of Aspergillus niger feruloyl esterase A using a serial factorial approach. Protein Expr. Purif. **55**, 166–174 (2007)
63. Liu, M., Wang, X., Yin, C., Zhang, Z., Lin, Q., Zhen, Y., Huang, H.: One-step on-column purification and refolding of a single-chain variable fragment (scFv) antibody against tumour necrosis factor α. Biotechnol. Appl. Biochem. **43**, 137–145 (2006)
64. Bane, S.E., Velasqueza, J.E., Robinson, A.S.: Expression and purification of milligram levels of inactive G-protein coupled receptors in *E. coli*. Protein Expr. Purif. **52**, 348–355 (2007)
65. Chen, G.Q., Gouaux, E.: Overexpression of a glutamate receptor (GluR2) ligand binding domain in *Escherichia coli*: application of a novel protein folding screen. Proc. Natl Acad. Sci. USA **94**, 13431–13436 (1997)
66. Willis, M.S., Hogan, J.K., Prabhakar, P., Liu, X., Tsai, K., Wei, Y., Fox, T.: Investigation of protein refolding using a fractional factorial screen: a study of reagent effects and interactions. Protein Sci. **14**, 1818–1826 (2005)
67. Wang, C., Wang, L., Geng, X.: Renaturation with simultaneous purification of rhG-CSF from *Escherichia coli* by ion exchange chromatography. Biomed. Chromatogr. **21**, 1291–1296 (2007)
68. Stempfer, G., Hőll-Neugebauer, B., Rudolph, R.: Improved refolding of an immobilized fusion protein. Nat. Biotechnol. **14**, 329–334 (1996)
69. Trésaugues, L., Colline, B., Minard, P., Henckes, G., Aufrère, R., Blondeau, K., Liger, D., Zhou, C.Z., Janin, J., van Tilbeurgh, H., Quevillon-Cheruel, S.: Refolding strategies from inclusion bodies in a structural genomics project. J. Struct. Funct. Genomics **5**, 195–204 (2004)
70. Lundstrom, K., Wagner, R., Reinhart, C., Desmyter, A., Cherouati, N., Magnin, T., Zeder-Lutz, G., Courtot, M., Prual, C., André, N., Hassaine, G., Michel, H., Cambillau, C., Pattus, F.: Structural genomics on membrane proteins: comparison of more than 100 GPCRs in 3 expression systems. J. Struct. Funct. Genomics **7**, 77–91 (2006)
71. Mo, Y., Lee, B.-K., Ankner, J.F., Becker, J.M., Heller, W.T.: Detergent-associated solution conformation of helical and β-barrel membrane proteins. J. Phys. Chem. B **112**, 13349–13354 (2008)
72. Bacic, A., Delmer, D.: Stimulation of membrane-associated polysaccharide synthetases by a membrane potential in developing cotton fibers. Planta **152**, 346–351 (1981)
73. Saier Jr., M.H.: Vectorial metabolism and the evolution of transport systems. J. Bacteriol. **182**, 5029–5035 (2000)
74. Taylor, N.G.: Cellulose biosynthesis and deposition in higher plants. New Phytol. **178**, 239–252 (2008)