

RESEARCH

Open Access



Efficient production of ^2H , ^{13}C , ^{15}N -enriched industrial enzyme *Rhizopus chinensis* lipase with native disulfide bonds

Meng Zhang¹, Xiao-Wei Yu^{1,2*}, G. V. T. Swapna^{3,4,5}, Rong Xiao^{3,4,5}, Haiyan Zheng⁶, Chong Sha¹, Yan Xu^{1,2*} and Gaetano T. Montelione^{3,4,5}

Abstract

Background: In order to use most modern methods of NMR spectroscopy to study protein structure and dynamics, isotope-enriched protein samples are essential. Especially for larger proteins (>20 kDa), perdeuterated and Ile ($\delta 1$), Leu, and Val methyl-protonated protein samples are required for suppressing nuclear relaxation to provide improved spectral quality, allowing key backbone and side chain resonance assignments needed for protein structure and dynamics studies. *Escherichia coli* and *Pichia pastoris* are two of the most popular expression systems for producing isotope-enriched, recombinant protein samples for NMR investigations. The *P. pastoris* system can be used to produce ^{13}C , ^{15}N -enriched and even ^2H , ^{13}C , ^{15}N -enriched protein samples, but efficient methods for producing perdeuterated proteins with Ile ($\delta 1$), Leu and Val methyl-protonated groups in *P. pastoris* are still unavailable. Glycosylation heterogeneity also provides challenges to NMR studies. *E. coli* expression systems are efficient for overexpressing perdeuterated and Ile ($\delta 1$), Leu, Val methyl-protonated protein samples, but are generally not successful for producing secreted eukaryotic proteins with native disulfide bonds.

Results: The 33 kDa protein—*Rhizopus chinensis* lipase (RCL), an important industrial enzyme, was produced using both *P. pastoris* and *E. coli* BL21 *trxB* (DE3) systems. Samples produced from both systems exhibit identical native disulfide bond formation and similar 2D NMR spectra, indicating similar native protein folding. The yield of ^{13}C , ^{15}N -enriched r27RCL produced using *P. pastoris* was 1.7 times higher than that obtained using *E. coli*, while the isotope-labeling efficiency was ~15 % lower. Protein samples produced in *P. pastoris* exhibit O-glycosylation, while the protein samples produced in *E. coli* were not glycosylated. The specific activity of r27RCL from *P. pastoris* was ~1.4 times higher than that produced in *E. coli*.

Conclusions: These data demonstrate efficient production of ^2H , ^{13}C , ^{15}N -enriched, Ile ($\delta 1$), Leu, Val methyl-protonated eukaryotic protein r27RCL with native disulfides using the *E. coli* BL21 *trxB* (DE3) system. For certain NMR studies, particularly efforts for resonance assignments, structural studies, and dynamic studies, *E. coli* provides a cost-effective system for producing isotope-enriched RCL. It should also be potential for producing other ^2H , ^{13}C , ^{15}N -enriched, Ile ($\delta 1$), Leu, Val methyl-protonated eukaryotic proteins with native disulfide bonds.

Keywords: *Rhizopus chinensis* lipase, ^2H , ^{13}C , ^{15}N -enriched protein production, Native disulfide bonds, Nuclear magnetic resonance spectroscopy

*Correspondence: bioyuxw@aliyun.com; yxu@jiangnan.edu.cn

¹The Key Laboratory of Industrial Biotechnology, Ministry of Education, School of Biotechnology, Jiangnan University, 1800 Lihu Avenue, Wuxi 214122, Jiangsu, China

Full list of author information is available at the end of the article

Background

Nuclear magnetic resonance (NMR) spectroscopy is a powerful technique for biophysical studies of proteins, providing atom-specific information about protein structure, dynamics and protein interactions. In addition to providing a method complementary to X-ray crystallography for three-dimensional (3D) structure determination, NMR has evolved as a tool for characterizing protein–ligand and protein–protein interactions, measuring affinities and specificities of interactions, identifying binding epitopes, and for characterizing structural rearrangements and allosteric changes induced by ligand binding [1, 2]. Furthermore, NMR relaxation experiments can uniquely characterize and quantify internal dynamics of proteins on nano- to millisecond timescales [3–7].

For most modern NMR spectroscopy methods, isotope labeled (e.g., ^{15}N , ^{13}C , ^{19}F and/or ^2H) protein samples are required. In the last several years, new methodologies and technologies have been implemented for automated resonance assignments and structure determination [8–10], dynamics studies using nuclear relaxation measurements [11–14], and for studying larger proteins and complexes [15–17]. This creates unique challenges for isotope-enrichment with ^2H , ^{13}C , and/or ^{15}N , which is required for determining resonance assignments and for structural and dynamic studies, particularly for larger (>20 kDa) proteins.

The majority of protein NMR studies have used proteins produced in recombinant *E. coli* expression systems [18]. However, many key protein samples needed for biomolecular NMR studies require eukaryotic expression hosts for protein sample production. For these reasons, technologies to produce ^2H , ^{13}C , ^{15}N -enriched proteins in eukaryotic expression hosts have high potential impact for protein NMR studies [19, 20]. However, using eukaryotic expression systems, such as *Pichia*, baculovirus, or HEK293 present significant challenges, particularly for ^2H enrichment [21–26]. In addition, glycosylation in these eukaryotic host systems can be highly heterogeneous. While enzymatic methods are available to remove N-linked glycans, these methods often have very low yields [27, 28]. Methods for removal of O-linked glycans are even less efficient and routine. The large variety of O-glycans dictates the use of several enzymes for removing intact O-linked sugars of a single sample. Thus, chemical deglycosylation is more applicable for removing O-glycans. However, harsh alkali treatment will also disrupt peptide bonds and degrade the protein sample [29]. For such proteins, it may be preferable to produce perdeuterated and Ile ($\delta 1$), Leu, Val methyl-protonated protein samples in *E. coli* systems, without glycosylation.

In recent years, various fusion tags have been developed for solubility and stability enhancement and protein

folding of partner proteins. Maltose binding protein (MBP) [30] is a highly effective fusion tag for protein sample production in *E. coli* because of its efficient translation initiation, remarkable ability to enhance partner solubility, and mild elution condition [31]. However, soluble MBP-fusion proteins are not always properly folded, especially for those with multiple cysteine residues (and presumably forming disulfide bonds) [32, 33]. According to Nozach et al., for most disulfide-rich proteins they tested, fusions without redox properties (e.g., MBP, GST) could not form correctly folded peptides [34]. Intracellular *E. coli* expression systems for producing ^{13}C , ^{15}N -enriched, perdeuterated, and Ile ($\delta 1$), Leu, Val methyl-protonated enriched protein samples are well established and generally highly efficient [8]. However, secreted eukaryotic proteins also often require native disulfide bond formation, which can be challenging to accomplish using these intracellular *E. coli* expression hosts. In particular, reduction and oxidative refolding of secreted eukaryotic proteins into native conformations, while sometimes successful, can often fail or be very challenging [35, 36]. Thus, for certain NMR studies, it is valuable to develop methods that can provide perdeuterated samples of secreted proteins with native disulfide bonds but lacking native glycosylation. These samples can then be used for resonance assignments and initial NMR studies, which can subsequently be extended to studies of the corresponding glycosylated proteins if required.

Here, we describe the production of ^2H , ^{13}C , ^{15}N -enriched, non-glycosylated *Rhizopus chinensis* lipase with native disulfide bonds using the *E. coli* BL21 *trxB* (DE3) expression system. RCL is a 33 kDa monomeric protein with extensive industrial enzyme applications [37–40]. ^{13}C , ^{15}N -enriched RCL was produced using both *P. pastoris* and BL21 *trxB* (DE3) expression systems. We have developed an *E. coli* expression vector coding for a fusion construct in which this lipase gene *proRCL* is fused to the C-terminus of a maltose binding protein (MBP) expression and solubility enhancing tag, *MBP3-proRCL*. NMR and mass spectrometry data demonstrate heterogeneous glycosylation of RCL samples produced using the *P. pastoris* expression system. However, using *MBP3-proRCL* with the *E. coli* BL21 *trxB* (DE3) system, non-glycosylated, ^{13}C , ^{15}N -enriched MBP-*proRCL* with native disulfide bonds was produced at high yields. Following removal of the MBP expression enhancement tag, NMR, MS, and enzyme activity studies demonstrate a native-like structure for the resulting purified RCL enzyme. The resulting protocol was then used to produce ^2H , ^{13}C , ^{15}N -enriched RCL for NMR studies. These data demonstrate production of a ^2H , ^{13}C , ^{15}N -enriched eukaryotic glycoprotein RCL, in non-glycosylated form, with native disulfides, native tertiary structure, and high

specific activity using the MBP fusion system combined with *E. coli* BL21 *trxB* (DE3) expression hosts. This expression system provides a potential approach for producing isotope-enriched samples of isotope-enriched secreted, glycosylated eukaryotic proteins in glycan-free forms suitable for NMR studies.

Methods

Production of unlabeled and ^{13}C , ^{15}N -enriched r27RCL in *P. pastoris*

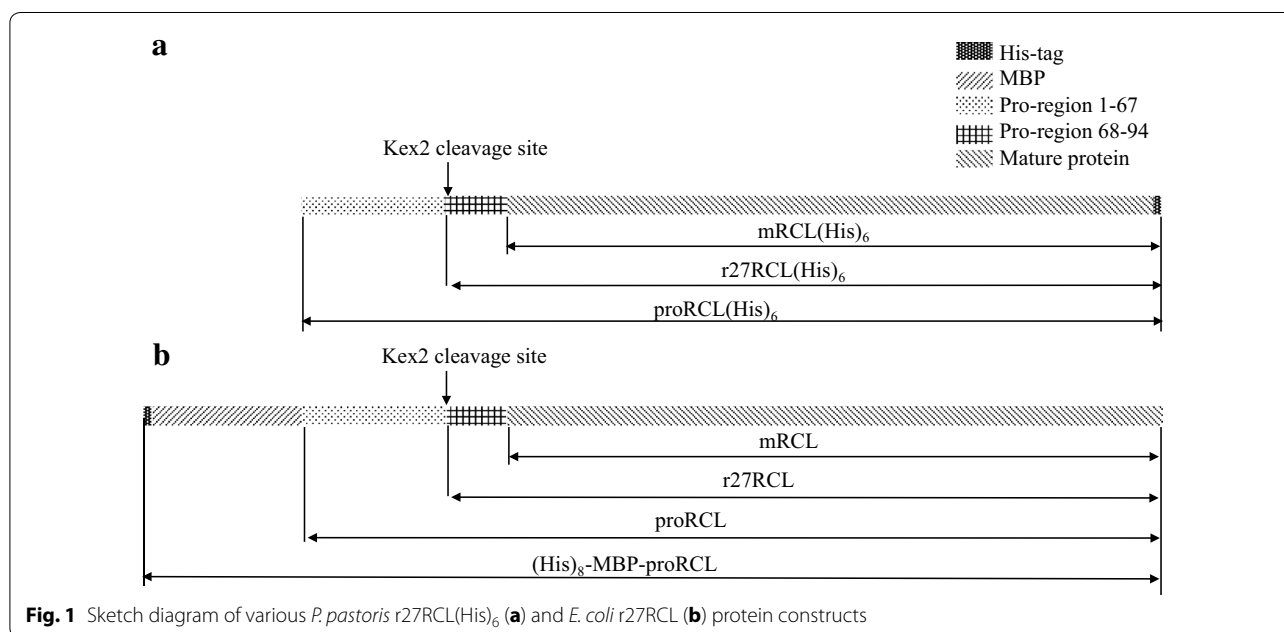
The various forms of RCL produced in *P. pastoris* and *E. coli* are sketched in Fig. 1. Here we adopt the naming conventions of ppRCL and ecRCL for RCL produced in *P. pastoris* and *E. coli*, respectively. The expression and production of ppRCL followed previously described procedures [41]. The PCR-amplified gene products were inserted into pPIC9 K expression vector. The plasmids were electroporated into *P. pastoris* strain GS115, and expressed in BMGY and BMMY media at 30 °C. After 84 h of induction at 28.5 °C, the cell culture was centrifuged at 6000×g for 30 min. The supernatant was filtered and purified on an ÄKTA purifier system (GE Co.) using a Ni-NTA column, equilibrated in binding buffer containing 25 mM Tris-HCl pH 8.0, 150 mM NaCl, 20 mM imidazole. The target protein was eluted with 250 mM imidazole. The fractions containing target protein were then dialysed twice against 25 mM Tris-HCl, 150 mM NaCl buffer (pH 8.0), and concentrated using an Amicon Ultra-15 Centricon (Millipore).

The production of ^{13}C , ^{15}N -enriched ppRCL followed a similar protocol with some modifications [23]. A

single colony containing recombinant plasmid was inoculated from a fresh plate to 25 mL of growth medium 1 [100 mM potassium phosphate buffer solution, pH 6.0, YNB (0.34 %, w/v), ammonium sulphate (1 % w/v), biotin (4×10^{-5} %, w/v), D-glucose (1 % w/v)], cultivated at 30 °C and 200 rpm for 24 h until the OD₆₀₀ reached 2~6. The cell culture was then centrifuged (3000×g, 5 min) and the cells were resuspended in 40 mL of growth Medium 2 [100 mM potassium phosphate buffer solution, pH 6.0, YNB (0.34 %, w/v), $^{15}\text{NH}_4\text{Cl}$ (0.02 % w/v), biotin (4×10^{-5} %, w/v), ^{13}C -D-glucose (0.1 % w/v)]. Cell growth was allowed to continue for 12 h, and stopped by centrifugation (3000×g, 5 min). Cells were then washed with a 0.2 % glycerol solution once followed by centrifugation and transferred into 100 mL induction medium [100 mM potassium phosphate buffer solution, pH 6.0, YNB (0.34 %, w/v), $^{15}\text{NH}_4\text{Cl}$ (1 % w/v), biotin (4×10^{-5} %, w/v)]. Protein expression was induced by the addition of 0.5 % v/v ^{13}C -methanol every 12 h. During the induction phase, the pH was adjusted to 6 by 4.4 M KOH and 7.4 M NaOH. After 84 h of induction at 28.5 °C, culture was centrifuged at 6000×g for 30 min. The supernatant was filtered, and purified on an ÄKTA purifier (GE Co.) using a Ni-NTA column as described above. Protein purity was quantified using AlphaView SA, ProteinSimple [42].

Production of unlabeled and ^{13}C , ^{15}N -enriched r27RCL in *E. coli*

Expression of MBP-proRCL in *E. coli* BL21 *trxB* (DE3) (Novagen) was started with inoculation a single colony



containing recombinant plasmid *MBP3-proRCL* from a fresh selection plate to 500 μ L LB media at 37 °C and 200 rpm for 4–6 h to obtain a pre-culture. 500 μ L of pre-culture was then used to inoculate 50 mL of LB media, and incubated overnight. The full content of this culture was then transferred into a 2 L flask, and incubated at 37 °C and 200 rpm until OD₆₀₀ reached 0.6–0.8 units. The incubation temperature was then decreased to 17 °C, and protein expression was induced by the addition of 1 mM IPTG. After overnight incubation at 17 °C, cells were harvested by centrifugation, the supernatant was discarded, and the pellet was stored at –80 °C. The cell paste was resuspended in 25 mL of binding buffer (50 mM Tris–HCl, 500 mM NaCl, 40 mM imidazole, 1 mM TCEP and 0.02 % NaN₃, pH 7.5) containing a protease inhibitor cocktail (complete protease inhibitor cocktail, Roche), and sonicated in an ice water bath using a Dual Horn 3/4" probe (Qsonica, LLC) for 10 min using a 30 s on/30 s off program. The cell debris was cleared by centrifugation at 15,000 \times g for 45 min, followed by filtration (0.45 μ m). The supernatant was then loaded onto the ÄKTAexpress system, and a one-step Ni-Gel filtration purification protocol was performed. The 8-His-tagged MBP-proRCL was eluted from the HisTrap HP column (5 mL) using five column volumes of elution buffer (50 mM Tris–HCl, 500 mM NaCl, 500 mM imidazole, 1 mM TCEP and 0.02 % NaN₃ at pH 7.5). The ÄKTAexpress system collected major peaks into internal storage loop, which was then automatically injected onto a Superdex 200 26/600 gel filtration column equilibrated in low-salt buffer (10 mM Tris–HCl, 100 mM NaCl, 5 mM DTT and 0.02 % NaN₃ at pH 7.5). The elution was monitored by A₂₈₀ detection, and major peaks were collected in 96-well plate. The elution fractions were analyzed by SDS-PAGE. Fractions containing target protein were pooled together and concentrated using Amicon ultrafiltration concentrators (Millipore). After being concentrated to 8–9 mg/ml, MBP-proRCL was hydrolyzed by Kex2 protease (Pep-rotech Inc.), using a ratio of 1:100 (protease: protein) at room temperature for 3 h. Purified r27RCL was collected in flow through from Ni–NTA column equilibrated in binding buffer 2 (50 mM Tris–HCl, 500 mM NaCl, 1 mM TCEP and 0.02 % NaN₃, pH 7.5).

The expression of ¹³C, ¹⁵N-enriched ecRCL in BL21 *trxB* (DE3) followed a similar, modified protocol. A single colony was picked from a fresh selection plate to 500 μ L LB media and incubated for 4–6 h at 37 °C. An aliquot (50 μ L) of this rich media preculture was then inoculated to 50 mL of ¹⁵N- and ¹³C-enriched MJ9 minimal medium [43] and incubated overnight at 37 °C on a shaker (250 rpm). The entire volume of overnight culture was then used to inoculate a 1 L of ¹⁵N- and ¹³C-enriched MJ9 medium. The culture was incubated at 37 °C (shaking at

250 rpm) until the OD₆₀₀ reaches 0.6–0.8 units and then transferred to 17 °C. Following equilibration at this temperature (~10 min), protein expression was induced with IPTG (1 mM final concentration). After overnight incubation at 17 °C, the cells were harvested by centrifugation and the protein was purified as described above for unlabeled ecRCL purification.

Production of [²H, ¹³C, ¹⁵N; ¹H-Ile δ 1, Leu- δ , Val- γ]-r27RCL in *E. coli*

²H, ¹³C, ¹⁵N-enriched ecRCL with ¹³C-enriched, protonated Ile (δ 1), Leu, and Val methyl groups ([U-²H, ¹³C, ¹⁵N; ¹H-Ile δ 1, Leu- δ , Val- γ]-r27RCL) was also produced in *E. coli* strain BL21 *trxB* (DE3) using procedures similar to those described in the previous section. Transformed cell cultures were grown at 37 °C in MJ9 minimal media [43] in ²H₂O containing 0.1 % (w/v) (¹⁵NH₄)₂SO₄ and 0.3 % (w/v) U-²H, ¹³C-glucose. When the culture reached an OD of 0.35, the culture was moved to 17 °C, and [U-¹³C₄, 3,3-²H₂]- α -ketobutyrate (50 mg/l), [U-¹³C₅, 3-²H]- α -ketoisovalerate (CIL Inc.) (100 mg/l) added for methyl labeling [44]. Protein expression was induced overnight by addition of IPTG to a final concentration of 1 mM. Cells were harvested by centrifugation at 6000 \times g for 30 min. [²H, ¹³C, ¹⁵N; ¹H-Ile δ 1, Leu- δ , Val- γ]-ecRCL was then purified as described above for ¹³C, ¹⁵N-labeled ecRCL.

Lipase activity and protein concentration measurements

Lipase activity was measured according to Kordel et al. [45]. A 7.9 mM solution of *p*NPP in isopropanol and 50 mM PBS buffer pH 8.0, containing 1.16 g/L sodium deoxycholate and 0.56 g/L arabic gum was mixed prior to use with a rate of 1:9 as substrate. The reaction was carried out by adding 0.1 mL of 3–4 μ M ppRCL or ecRCL solution into 2.4 mL of the substrate mixture. The absorbance at 410 nm against a blank was monitored using a SpectraMax[®] Plus 384 (molecular devices) after reacted at 40 °C for 2 min. One enzyme unit was defined as the amount of enzyme releasing 1 μ mol of *p*-nitrophenol per minute under the assay conditions (pH 8.0, 40 °C). The protein concentration was determined by Pierce BCA protein assay kit (Thermo Scientific) using bovine serum albumin (BSA) as a standard.

Disulfide bond mapping and glycosylation analysis

For disulfide bond mapping, 56 μ g each of ppRCL and ecRCL were incubated with 10 mM iodoacetamide (IAM), 6 M urea in Tris–HCl buffer pH 7.5 at room temperature for 1 h to alkylate cysteines. After alkylation, excess iodoacetamide was removed using 10 kDa cutoff ultrafiltration device. The samples were then hydrolyzed into peptide fragments with chymotrypsin (molarities were

chymotrypsin: protein = 1:25) in ammonium acetate buffer pH 6.0. Two equal fractions were created. The first fraction was reduced with 10 mM TCEP to reduce the disulfide bonds and create free cysteine groups. The second fraction was served as a disulfide-intact control. Peptides were reconstituted in 0.1 % TFA and analyzed by LC-MS/MS on a Dionex RSLC system (ThermoFisher, San Jose, CA) interfaced with a LTQ Orbitrap Velos (ThermoFisher, San Jose, CA) online with a Thermo LTQ Orbitrap Velos mass spectrometry (Thermo Fisher Scientific, San Jose, CA). The peptide mixtures were loaded onto self-packed 100 $\mu\text{m} \times 2\text{ cm}$ trap packed with Magic C18AQ, 5 μm 200 A (Michrom Bioresources Inc, Auburn, CA) and washed with Buffer A (0.2 % formic acid) for 5 min at a flow rate of 10 $\mu\text{l}/\text{min}$. Separation was achieved with an analytical reverse-phase chromatography column (Magic C18AQ, 3 μm 200 A, 75 $\mu\text{m} \times 50\text{ cm}$) and peptides were fractionated at 300 nl/min with a multi-stepped gradient [4–15 % buffer B (0.16 % formic acid 80 % acetonitrile) in 10 min and 15–50 % buffer B in 40 min]. Mass spectrometry data was acquired using a data-dependent acquisition procedure with a cyclic series of a full scan acquired in Orbitrap with resolution of 60,000 followed by MS/MS scans (CID 35 % of collision energy) of 20 most intense ions with a repeat count of two and the dynamic exclusion duration of 60 s. The MS/MS data were searched against a custom fasta database including target protein sequences using X!tandem (SLEDGEHAMMER (2013.09.01), thegpm.org) with carbamidomethylation on cysteine as fixed modification and oxidation of methionine and deamidation on asparagine as variable modifications using a 10 ppm precursor ion tolerance and a 0.4 Da fragment ion tolerance.

For the glycosylation analysis, 10 μg of ppRCL was first separated by SDS-PAGE. Gel bands were cut and divided into two fractions: one fraction was in-gel digested by trypsin, whereas the other fraction was digested with Endo-proteinase Asp-N. For Asp-N digested sample, a portion was digested again with trypsin for 2 h (Asp-N + trypsin sample), while the other sample's portion was also digested with chymotrypsin in solution. For this experiment, 10 μg of protein was denatured by addition of 6 M guanidine HCl, reduced with 20 mM DTT at 60 $^{\circ}\text{C}$ for 30 min, alkylated with 40 mM iodoacetamide, and incubated for 45 min at room temperature in the dark. Afterwards, the samples were desalted using spin filter with MWCO10KD, followed by digestion with chymotrypsin on the membrane at room temperature for 18 h and eluted by passing the filter again. All samples were analyzed by nano-LC-MS/MS. Data was manually interpreted.

NMR sample preparation, data collection and processing

^{13}C , ^{15}N -enriched pp27RCL and ecRCL samples were prepared at 0.5 mM protein concentrations in 0.1 M

sodium phosphate buffer pH 6.0 containing 10 % $^2\text{H}_2\text{O}$ and 50 μM 4,4-dimethyl-4-silapentane-1-sulfonate sodium salt (DSS). The [^2H , ^{13}C , ^{15}N ; ^1H -Ile $\delta 1$, Leu- δ , Val- γ]-ecRCL was prepared at 0.12 mM concentration using the same pH 6.0 sodium phosphate buffer. [^1H , ^{15}N]-TROSY-HSQC and [^1H , ^{13}C]-HSQC 2D spectra were collected on Bruker Avance II 800 MHz spectrometer at 35 $^{\circ}\text{C}$. NMR data were processed and analyzed with NMRPipe, NMRDraw [46] and SPARKY [47]. NMRPipe was used for data processing, and both NMRDraw and SPARKY for spectral display and peak analysis.

Results

Unlabeled and isotope-enriched r27RCL expression in *P. pastoris* and *E. coli*

To compare the levels of heterologous ppRCL and ecRCL expression, the expression levels and enzymatic activities were compared. As described in detail in the “Methods” section, after 84 h of induction by methanol, proRCL was expressed, hydrolyzed by Kex2 protease (which is expressed naturally by *P. pastoris*) at the recognition site Lys-Arg between -29 and -28 of the prosequence, and secreted into culture supernatant as ppRCL (Fig. 1). The supernatant (Fig. 2, lane 1) was purified using a Ni-NTA column with gradient elution. Purity was over 95 % (Fig. 2, lane 2) as analyzed by AlphaView SA, ProteinSimple [42]. When produced in *P. pastoris*, the heterologous protein secreted into the culture was relatively pure and the yield was $124.8 \pm 5.3\text{ mg}/\text{L}$ after one-step Ni-NTA purification.

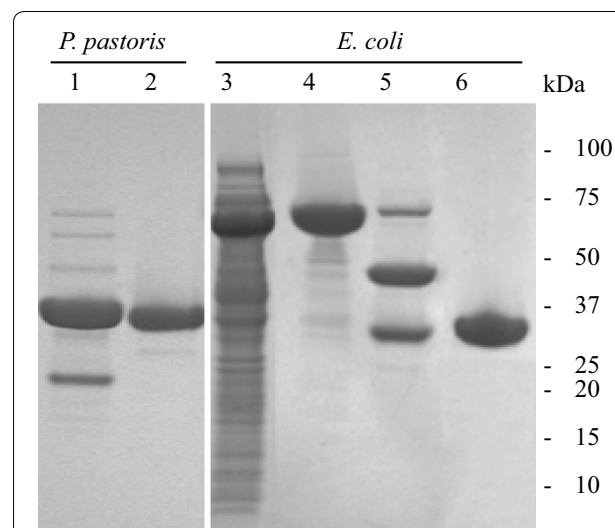
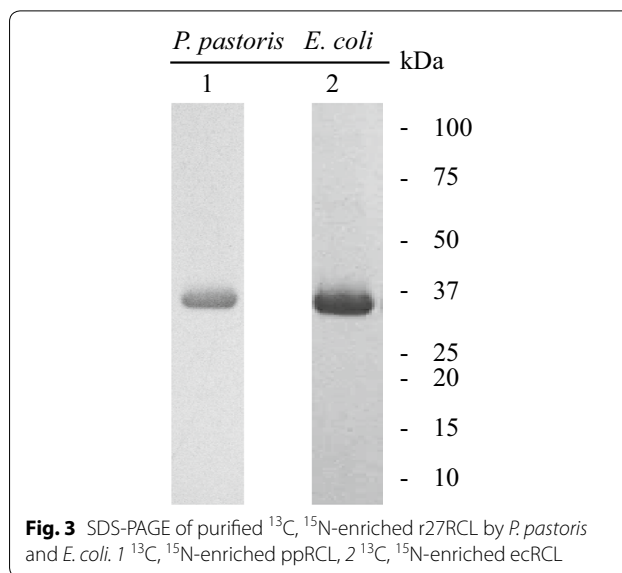


Fig. 2 Expression and purification of r27RCL by *P. pastoris* and *E. coli*. 1 soluble protein, 2 r27RCL(His)₆ after HisTrap column, 3 soluble protein, 4 MBP-proRCL after one-step HisTrap-size-exclusion chromatography, 5 after Kex2 cleavage, 6 r27RCL after HisTrap column

For the production of ecRCL, soluble fusion protein MBP-proRCL (Fig. 2, lane 3) was purified by sequential Ni-NTA and gel filtration chromatography (Fig. 2, lane 4), followed by tag cleavage by Kex2 protease (Fig. 2, lane 5). The N-terminal His₈-MBP tag together with 67 amino-acid residues of the pro-region of RCL (Fig. 1), and the residual un-hydrolyzed fusion protein, was then removed using a Ni-NTA column. Purification buffers included 1 mM TCEP or 5 mM DTT, which may prevent the formation of dimer or polymer but were not sufficient enough to reduce the disulfide bonds. Pure ecRCL (Fig. 2, lane 6) was pooled together from flow-through fractions. The purity was over 95 % as quantified using AlphaView SA, ProteinSimple [42]. The expression level of MBP-proRCL in *E. coli* was 150.2 ± 3.8 mg/L, but after the sequential Ni-NTA and gel filtration purifications, Kex2 cleavage and subsequent Ni-NTA purification, the final ecRCL yield was 18.7 ± 1.4 mg/L.

Validation of ppRCL and ecRCL were performed by MALDI-TOF mass spectrometry. The ppRCL was partly glycosylated, while the major peak was not glycosylated. Accordingly, the major peak was used for the measurement of molecular weight. The measured molecular weights of ppRCL and ecRCL were 33339.6 ± 25.1 and 32280.8 ± 23.8 Da, in agreement with the expected corresponding theoretical molecular weights of 33325.6 and 32271.5 Da, respectively.

We next compared the production of isotope-enriched r27RCL using the *P. pastoris* and *E. coli* expression systems. Both expression levels and the completeness of isotope incorporation were assessed. As described in detail in the Materials and Methods section, after 12 h of cell adaptation in isotope-enriched growth medium, and 84 h of induction using ¹³C-methanol [48], proRCL was expressed, hydrolyzed by Kex2 protease of *P. pastoris* and secreted into culture supernatant. The supernatant was purified following the same protocol used for the unlabeled protein, providing purity over 95 % (AlphaView SA, ProteinSimple [42]). For the production of ¹⁵N, ¹³C-labeled ecRCL, the fusion protein MBP-proRCL was produced in MJ9 media [43], and purified as described above. The SDS-PAGE gel was similar to the one with unlabeled protein, as shown in Fig. 3, lane 1, 2. For production of ¹³C, ¹⁵N-enriched ppRCL, the yield was 25.6 ± 3.1 mg/L after one-step Ni-NTA purification. The expression level of ¹³C, ¹⁵N-enriched MBP-proRCL in *E. coli* was 121.5 ± 4.1 mg/L, but the final ecRCL yield was only 15.3 ± 2.5 mg/L. The percentages of the labeled species were estimated using the difference between the observed mass and the theoretical average mass. The isotope incorporation completeness was 83.6 ± 3.2 % for ppRCL, while it was 99.2 ± 0.2 % for ecRCL (Table 1).



Validation of native disulfide bond formation

Correct formation of native disulfide bonds is critical for stabilizing protein structures and maintaining protein function. One of the major advantages of *P. pastoris* over *E. coli* is that *P. pastoris* is capable of producing secreted proteins with native disulfide bonds [49]. The *E. coli* trxB strain lacks thioredoxin reductase, which allows formation of disulfide bonds within the *E. coli* cytoplasm [50]. To verify if *E. coli* expressed correctly folded ecRCL with native disulfide bonds, enzymatic digestion followed by mass spectrometry (MS/MS) detection was used for disulfide linkage assignment. The formation of one disulfide bond results in a 2-Da reduction of molecular weight, which can be distinguished by high-resolution mass spectrometry. Native disulfide bonds linking peptide fragments ⁵⁶CRSVVPGTKW⁶⁵ and ²⁸⁸FGINEDSCL²⁹⁶ (Fig. 4b, -TCEP) and forming intra-fragment disulfide bonds in peptides ⁶⁶DCKQ-CLKYVPDGKLIKTF⁸³ (Fig. 4c, -TCEP), and ²⁵²IKED-PADVQICTSNIETKQCSNSIVPF²⁷⁸ (Fig. 4d, -TCEP) were detected in control analyses, which were not treated with the disulfide reducing agent TCEP. The disulfide bonds linking these peptide fragments, and intra-fragment disulfide bonds, were cleaved when treated with TCEP. Thus reduction of disulfide bonds linking peptide fragments resulted in two peptide fragments, and reduction of intra-fragment disulfide bonds resulted in a 2-Da reduction of molecular weight. Since molecular weight shift changes the retention time, those reduced peptides were not detected in the selected molecular weight windows (Fig. 4b, 4c, 4d, +TCEP) with the same retention times as in control groups. As described in the methods

Table 1 Production of isotope-enriched r27RCL in *P. pastoris* and *E. coli*

Expression systems	Expressed proteins	Expression level (mg/l)	Enrichment efficiency
<i>P. pastoris</i>	^{13}C , ^{15}N -enriched ppRCL	25.6 ± 3.1	83.6 ± 3.2 %
<i>E. coli</i>	^{13}C , ^{15}N -enriched ecRCL	15.3 ± 2.5	99.2 ± 0.2 %

section, free cysteines were alkylated during the treatment. Alkylated C204 fragment ($^{201}\text{TVGCPRVGN-NAF}^{212}$) was detected in ppRCL and ecRCL samples (Fig. 4a), indicating that C204 is a free cysteine in both protein samples. These results indicated that native disulfide bonds are formed between cysteines C56–C295, C67–7C0 and C262–C271 in both ppRCL and ecRCL.

$[^1\text{H}, ^{15}\text{N}]$ -TROSY-HSQC and $[^1\text{H}, ^{13}\text{C}]$ -HSQC 2D NMR Spectra

$[^1\text{H}, ^{15}\text{N}]$ -TROSY-HSQC NMR spectra of ^{13}C , ^{15}N -enriched ppRCL and ecRCL showed good resonance dispersion, characteristic of a well-folded, globular proteins (Fig. 5a). Most resonances of proteins produced in either expression system have essentially identical chemical shifts, indicating that the structures of ppRCL and ecRCL are largely identical. However, because of the

lower isotope-enrichment levels, the number of backbone ^1H - ^{15}N correlation peaks identified in the $[^1\text{H}, ^{15}\text{N}]$ -TROSY-HSQC spectra of ppRCL was only 213 out of an expected 276 backbone amides. On the other hand, 271 $[^1\text{H}, ^{15}\text{N}]$ -TROSY-HSQC backbone correlation peaks were identified in the ^{13}C , ^{15}N -enriched ecRCL sample. The ^1H and ^{15}N resonances in spectra of ppRCL are generally broader than those of ecRCL. This may be due to the hexa-His C-terminal tag and/or residual glycosylation (described below) of ppRCL.

As can be observed from the superimposed $[^1\text{H}$ - $^{13}\text{C}]$ -HSQC NMR spectra of ppRCL and ecRCL (Fig. 5b), ppRCL sample also has many extra ^1H - ^{13}C correlation peaks in the region 3.5–4.5 ppm in the ^1H dimension and 65–80 ppm in the ^{13}C dimension. These are assigned to glycans associated with the protein sample,

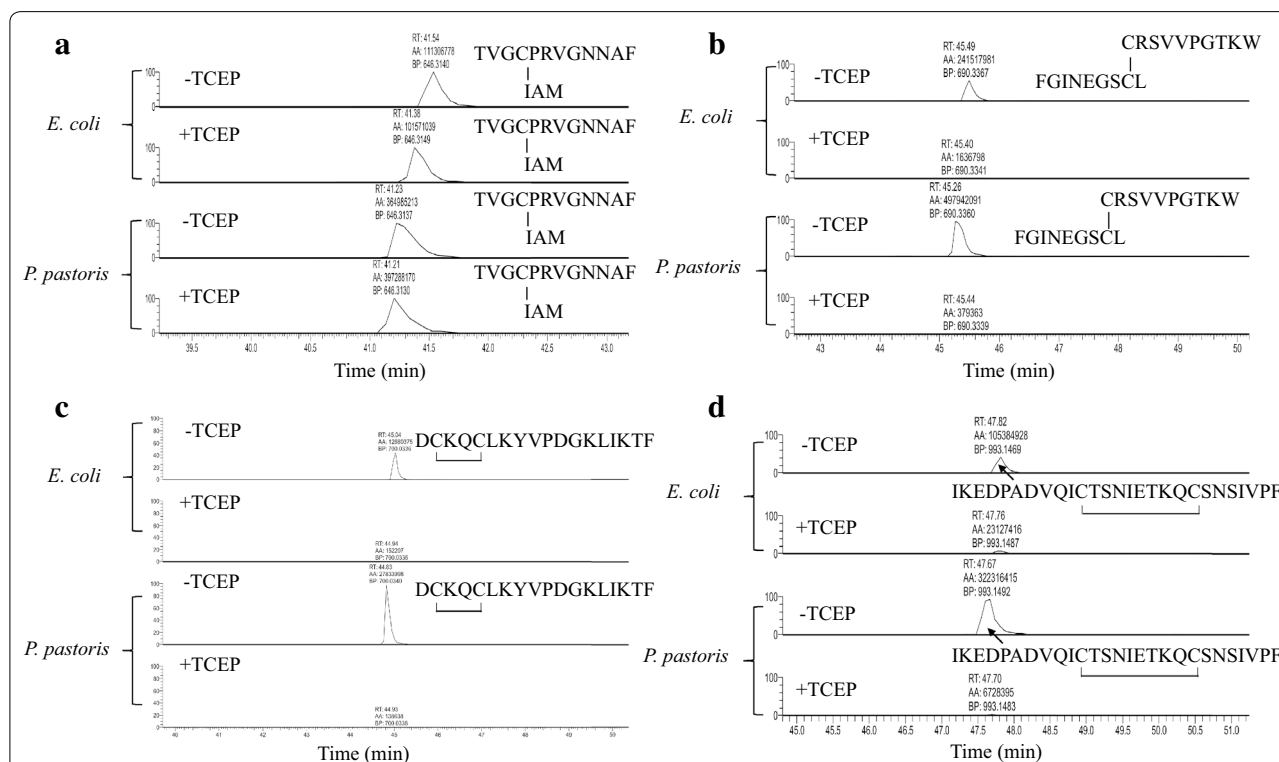
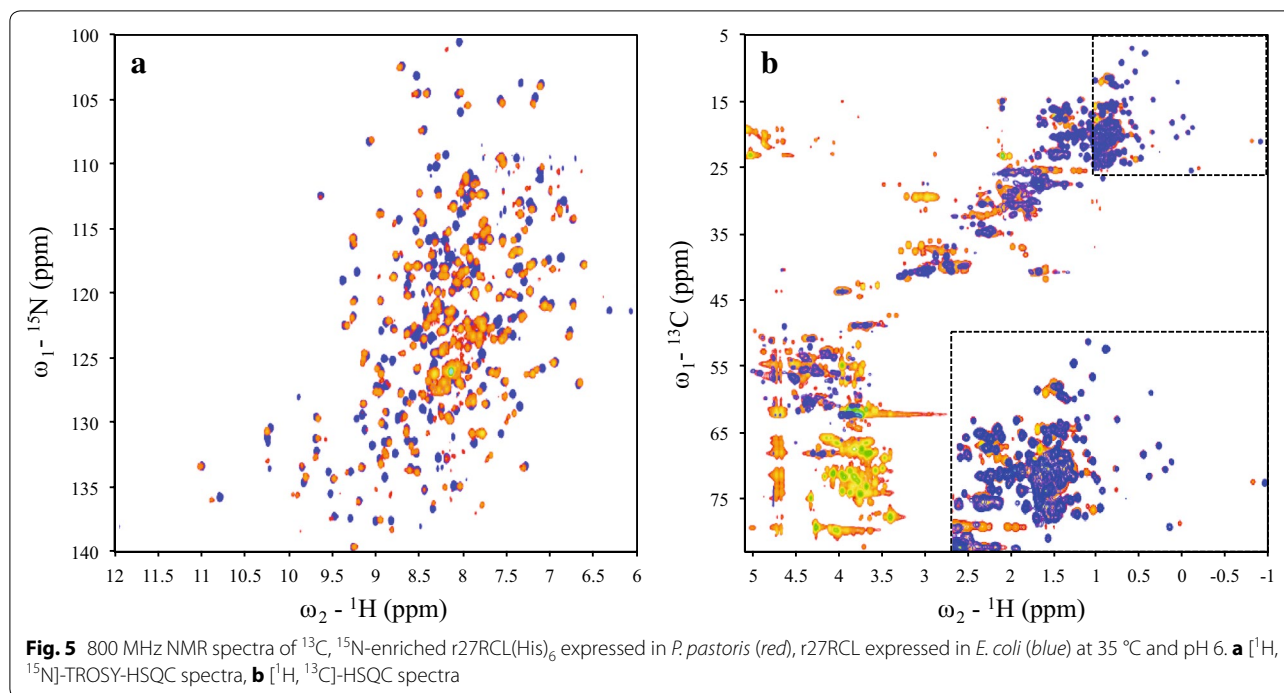


Fig. 4 Disulfide bond mapping of r27RCL(His)₆ from *P. pastoris* and r27RCL from *E. coli*. LC-MS fragments of *P. pastoris* and *E. coli* w/o TCEP treatment **a** C204 [alkylated with iodoacetamide (IAM)]; **b** C56–C295 disulfide bond pair; **c** C67–C70 intra-disulfide bond peptide; **d** intra-disulfide bond peptide indicating C262–C271



demonstrating that ppRCL protein using the methods outlined in this work remains partially glycosylated.

Glycosylation analysis

Protein glycosylation is a common post-translational modification, which can play key roles in the function of enzymes and regulation of their enzymatic activities. 2D [^1H , ^{13}C]-HSQC spectra of ppRCL clearly indicated some glycosylation (Fig. 5b). To further profile the glycan microheterogeneity, purified ppRCL and ecRCL were analyzed by electrospray mass spectrometry (Fig. 6a, b). The spectrum of ecRCL exhibits only one peak of 32264.7 Da (Fig. 6b). However, ppRCL exhibits multiple MALDI-TOF mass spectrometry peaks, with molecular weights of 33318.4 Da (peak 1), 33642.6 Da (peak 2), and 33805.3 Da (peak 3), respectively (Fig. 6a). These mass increments correspond to masses of hexose units (162 Da). Furthermore, after digestion of ppRCL with Endoproteinase Asp-N and analysis by LC-MS/MS, the glycosylation site was revealed to be a non-hydrolyzed O-glycosylation mapped to a Ser or Thr residue within the polypeptide segment $^{-17}\text{DLPENPPPIPATSTAPSS}^1$ (Fig. 6c, d).

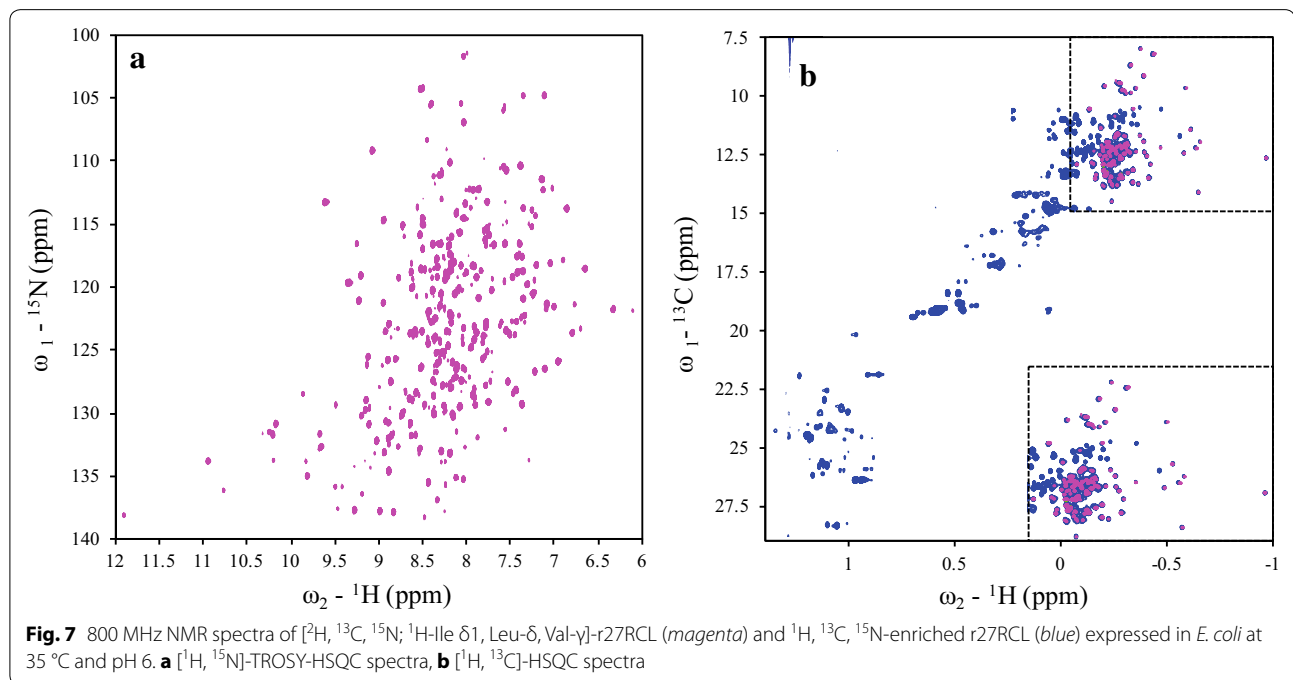
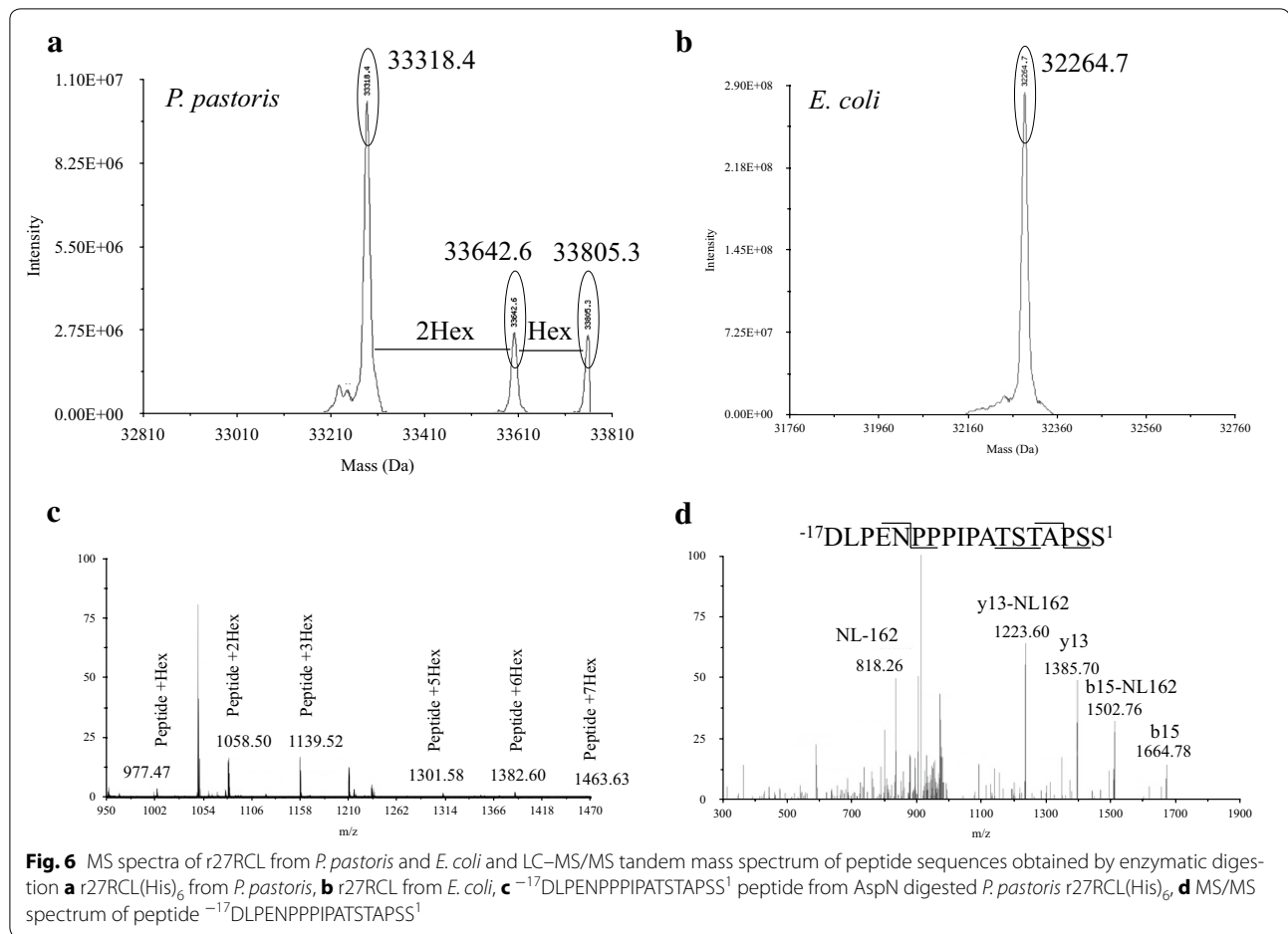
Expression level and activity comparisons

As an enzyme, specific activity is an essential characteristic. To further compare the specific activity of r27RCL from *E. coli* and *P. pastoris*, unlabeled proteins were both buffer exchanged to 25 mM Tris-HCl, 150 mM NaCl, pH

8.0 buffer, which is the optimum pH for r27RCL. Lipase activity was measured at 40 °C for 2 min. The specific activities were 283.52 ± 4.28 and 118.48 ± 0.45 U/mg for ppRCL and ecRCL, respectively.

Expression and NMR studies of [^2H , ^{13}C , ^{15}N ; ^1H -Ile δ 1, Leu- δ , Val- γ]-r27RCL

Samples of perdeuterated [^2H , ^{13}C , ^{15}N ; ^1H -Ile δ 1, Leu- δ , Val- γ]-enriched r27RCL were also prepared in the *E. coli* BL21 *trxB* (DE3) expression system. The growth rate of adapted cells in ^2H , ^{13}C , ^{15}N -enriched MJ9 medium in $^2\text{H}_2\text{O}$ was relatively slow compared with those in ^{13}C , ^{15}N -enriched MJ9 medium in H_2O . It took 8 h before OD₆₀₀ reached 0.35 units, at which point the ^{13}C - ^1H methyl labeling precursors were added. The yield of purified fusion protein MBP-proRCL and r27RCL were 51.2 and 4.3 mg/L, respectively. Based on the [^1H , ^{13}C]-HSQC NMR spectrum (Fig. 7b, magenta), it appears that this protein sample is fully perdeuterated except for the targeted methyl groups, as there is no indication of ^1H - ^{13}C correlation peaks except those of the Ile δ 1 methyl and the Leu- δ , Val- γ isopropyl methyl groups. 96 out of 106 Ile δ 1 methyl and Leu- δ /Val- γ isopropyl methyl groups were observed. Compared with the 2D [^1H , ^{13}C]-HSQC spectrum of ^{13}C , ^{15}N sample (Fig. 7b, blue), the spectrum of [^2H , ^{13}C , ^{15}N ; ^1H -Ile δ 1, Leu- δ , Val- γ]-r27RCL sample also showed significant improvement of resolution and reductions of peak linewidths in both ^1H and ^{13}C dimensions. Furthermore, comparing the one-dimensional



slices of [^1H , ^{15}N]-TROSY-HSQC spectra of ^{13}C , ^{15}N -enriched (Fig. 5a, blue) and ^2H , ^{13}C , ^{15}N -enriched ecRCL (Fig. 7a), the NMR resonances were significantly sharper in perdeuterated sample spectrum. For example, the proton linewidths for the amide ^{15}N - ^1H correlation resonance of Gly7 are ~ 26.4 Hz for the ^{13}C , ^{15}N -enriched sample, and ~ 20.6 Hz for ^2H , ^{13}C , ^{15}N -enriched ecRCL. Hence, [^2H , ^{13}C , ^{15}N ; ^1H -Ile $\delta 1$, Leu- δ , Val- γ]-enriched ecRCL with native disulfide bonds can be produced in sufficient quantities for NMR studies using the *E. coli* BL21 *trxB* (DE3) expression system.

Discussion

Although bacterial host systems have been developed that can sometimes be used successfully to produce protein samples with native disulfide bonds [36, 51, 52], production of isotope-enriched proteins with native disulfide bonds can present a significant challenge which needs to be overcome for NMR studies. In particular, while the *E. coli* BL21 *trxB* (DE3) strain has been used extensively for producing proteins, including unlabeled proteins for NMR studies [53] and ^{15}N , ^{13}C -enriched proteins using specialized Celvone enriched media for triple-resonance NMR studies [54], our study demonstrates the use of BL21 *trxB* (DE3) for production of ^2H , ^{13}C , ^{15}N -enriched RCL with native disulfide bonds in defined MJ9 minimal media. This is the first example of using the *E. coli* BL21 *trxB* (DE3) expression system to produce a disulfide-containing, perdeuterated [^2H , ^{13}C , ^{15}N ; ^1H -Ile $\delta 1$, Leu- δ , Val- γ]-enriched eukaryotic protein. The technology is enabled by using the MBP fusion system. This approach is especially valuable for larger proteins (>20 kDa), since perdeuteration is required to suppress nuclear relaxation properties of the remaining ^1H , as well as ^{13}C and ^{15}N nuclei, providing slower relaxation rates and sharper line widths [55]. Protonated methyl groups often serve as key sensitive probes of protein structure and dynamics and are of special value in NMR studies [8]. Such, perdeuterated [^2H , ^{13}C , ^{15}N ; ^1H -Ile $\delta 1$, Leu- δ , Val- γ]-enriched protein sample provide not only significant improvement of resolution and reductions of peak linewidths, but also allow analysis of backbone and some sidechain methyl resonance assignments that can be used for structural and dynamic studies of important enzymes like RCL.

Although the structures and dynamics of glycosylated proteins are both important and feasible, as a first step towards such studies it is useful to complete resonance assignments and structural analysis of the non-glycosylated protein. As shown in the [^1H , ^{15}N]-TROSY-HSQC and [^1H , ^{13}C]-HSQC spectra in Fig. 5, glycosylation increased protein rotational correlation times (τ_c) and conformational heterogeneity, and can cause extensive line-broadening in the NMR spectra. In many cases,

the sugar moiety is not essential for structural integrity or functional activity of the protein and can be trimmed [56]. Indeed, most crystallographic and/or NMR studies of glycoproteins have been completed using non-glycosylated versions of these proteins. For example, although *P. pastoris* has been used to prepare protein samples for more than 90 NMR structures deposited in the PDB database, in some of these studies glycosylation was suppressed by mutation of N-linked Asn glycosylation sites to Gln [19, 57, 58], while in other cases glycosylation was not characterized [59]. In the solution NMR structure of recombinant Ber e1, a major allergen from Brazil nut containing four disulfide bonds, the sample produced using *P. pastoris* included three O-linked glycans which were retained for the NMR studies. This heterogeneous glycosylation, which is not observed in the native protein purified from Brazil nut, did not prevent the NMR structure analysis, but contributed unnecessary complications and line broadening [60].

In preparing glycoproteins for structural studies, glycans are often removed by glycosidase treatment. This glycosylation is relatively routine for N-linked glycans. Unfortunately, there are no general glycosidases that completely remove all O-linked glycans. In this case a mix of exoglycosidases can be followed by a broad specificity O-glycosidase (Protein Deglycosylation Mix, NEB Inc.; Enzymatic Protein Deglycosylation Kit, Sigma-Aldrich Inc.). However, this process is not very efficient. Following glycosidase treatment, the final products may still contain residual glycan, which complicates NMR spectral data.

Although yields are lower per liter, is also less about 50 % less expensive to produce ^{13}C , ^{15}N -enriched r27RCL using the *E. coli* BL21 *trxB* (DE3) system than when using *P. pastoris*. In our experience, ^2H , ^{13}C , ^{15}N -enriched samples could also be produced in *P. pastoris*, as has been reported by other groups [21, 24, 59]. However, we could not biosynthetically incorporate ^1H , ^{13}C methyl groups into otherwise perdeuterated RCL lipase produced in *P. pastoris*, which is also inefficient as reported by Clark et al. [61].

Conclusion

In this work we have shown that [^2H , ^{13}C , ^{15}N ; ^1H -Ile $\delta 1$, Leu- δ , Val- γ]-enriched r27RCL lipase with native disulfide bonds, native tertiary structure, and high specific activity can be produced using the *E. coli* BL21 *trxB* (DE3) system. These samples provide the basis for NMR studies of the structure, dynamics, and function of r27RCL. These results also suggest a potential strategy for NMR studies of secreted, glycosylated proteins like RCL lipase, in which samples are first produced in *E. coli* BL21 *trxB* (DE3) system for NMR resonance assignments

and initial structural and dynamic studies. These data then provide the basis for studies of glycosylated protein samples produced in *P. pastoris*.

Abbreviations

RCL: *Rhizopus chinensis* lipase; mRCL: *Rhizopus chinensis* mature lipase; r27RCL(His)₆: *Rhizopus chinensis* mature lipase with 6 his-tag at C-terminus and 27 amino acids of the C-terminal of the prosequence; r27RCL: *Rhizopus chinensis* mature lipase with 27 amino acids of the C-terminal of the prosequence; proRCL: *Rhizopus chinensis* lipase with prosequence; MBP-proRCL: *Rhizopus chinensis* lipase with prosequence and MBP-tag at the N-terminus; ppRCL: RCL expressed in *P. pastoris*; ecRCL: RCL expressed in *E. coli*; pNPP: *p*-nitrophenyl palmitate; BMGY: buffered glycerol complex medium; BMMY: buffered methanol complex medium; YNB: yeast nitrogen base.

Authors' contributions

MZ carried out the experimental assay and data analysis. MZ, X-WY, YX, RX, and GTM designed the study. MZ, X-WY, YX and GTM wrote the manuscript. GVTS participated in the collection and analysis of the NMR data. H-YZ carried out LC-MS/MS experiment and results analysis. CS constructed MBP-proRCL strain. All authors have read and approved the final manuscript.

Author details

¹ The Key Laboratory of Industrial Biotechnology, Ministry of Education, School of Biotechnology, Jiangnan University, 1800 Lihu Avenue, Wuxi 214122, Jiangsu, China. ² State Key Laboratory of Food Science and Technology, Jiangnan University, 1800 Lihu Avenue, Wuxi 214122, Jiangsu, China. ³ Center for Advanced Biotechnology and Medicine, Department of Molecular Biology and Biochemistry, Rutgers, The State University of New Jersey, Piscataway, NJ, USA. ⁴ Department of Biochemistry and Molecular Biology, Robert Wood Johnson Medical School, Rutgers, The State University of New Jersey, Piscataway, NJ, USA. ⁵ Northeast Structural Genomics Consortium, Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA. ⁶ Biological Mass Spectrometry Facility, Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA.

Acknowledgements

We thank Profs. T. Szyperki and J. Hunt for extensive and useful discussions.

Competing interests

G.T.M. and R.X. are consultants to Nexomics Biosciences, Inc. The remaining authors declared that they have no competing interests.

Availability of data and material

The datasets supporting the conclusions of this article are included within the article.

Funding

Financial support from the National High Technology Research and Development Program of China (863 Program) (2015AA021004), the High-end Foreign Experts Recruitment Program (GDW20123200113), Six Talent Peaks Project in Jiangsu Province (NY-010), 333 Project in Jiangsu Province (BRA2015316), NSFC (21336009), and the 111 Project (111-2-06) are greatly appreciated.

Received: 21 April 2016 Accepted: 3 July 2016

Published online: 13 July 2016

References

- Chen W-N, Kuppam KV, Lee MD, Jaudzems K, Huber T, Otting G. O-tert-butyltyrosine, an NMR tag for high-molecular-weight systems and measurements of submicromolar ligand binding affinities. *J Am Chem Soc.* 2015;137:4581–6.
- Arai M, Ferreon JC, Wright PE. Quantitative analysis of multisite protein–ligand interactions by NMR: binding of intrinsically disordered p53 transactivation subdomains with the TAZ2 domain of CBP. *J Am Chem Soc.* 2012;134:3792–803.
- Eisenmesser EZ, Millet O, Labeikovsky W, Korzhnev DM, Wolf-Watz M, Bosco DA, Skalicky JJ, Kay LE, Kern D. Intrinsic dynamics of an enzyme underlies catalysis. *Nature.* 2005;438:117–21.
- Mulder FA, Mittermaier A, Hon B, Dahlquist FW, Kay LE. Studying excited states of proteins by NMR spectroscopy. *Nat Struct Biol.* 2001;8:932–5.
- Moschen T, Wunderlich CH, Spitzer R, Levic J, Micura R, Tollinger M, Kreutz C. Ligand-detected relaxation dispersion NMR spectroscopy: dynamics of pre-Q1-RNA binding. *Angew Chem Int Ed Engl.* 2015;54:560–3.
- Khirich G, Loria JP. Complexity of protein energy landscapes studied by solution NMR relaxation dispersion experiments. *J Phys Chem B.* 2015;119:3743–54.
- Sanchez-Medina C, Sekhar A, Vallurupalli P, Cerminara M, Muñoz V, Kay LE. Probing the free energy landscape of the fast-folding gpw protein by relaxation dispersion NMR. *J Am Chem Soc.* 2014;136:7444–51.
- Tugarinov V, Kanelis V, Kay LE. Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy. *Nat Protoc.* 2006;1:749–54.
- Qing GL, Ma LC, Khorchid A, Swapna GVT, Mal TK, Takayama MM, Xia B, Phadtare S, Ke HP, Acton T, et al. Cold-shock induced high-yield protein production in *Escherichia coli*. *Nat Biotechnol.* 2004;22:877–82.
- Xiao R, Anderson S, Aramini J, Belote R, Buchwald WA, Ciccocanti C, Conover K, Everett JK, Hamilton K, Huang YJ. The high-throughput protein sample production platform of the northeast structural genomics consortium. *J Struct Biol.* 2010;172:21–33.
- Li Y, Altorelli NL, Bahna F, Honig B, Shapiro L, Palmer AG 3rd. Mechanism of E-cadherin dimerization probed by NMR relaxation dispersion. *Proc Natl Acad Sci USA.* 2013;110:16462–7.
- Rosenzweig R, Farber P, Velyvis A, Rennella E, Latham MP, Kay LE. ClpB N-terminal domain plays a regulatory role in protein disaggregation. *Proc Natl Acad Sci USA.* 2015;112:E6872–81.
- Sanchez-Medina C, Sekhar A, Vallurupalli P, Cerminara M, Muñoz V, Kay LE. Probing the free energy landscape of the fast-folding gpW protein by relaxation dispersion NMR. *J Am Chem Soc.* 2014;136:7444–51.
- Lorieau JL, Louis JM, Bax A. Whole-body rocking motion of a fusion peptide in lipid bilayers from size-dispersed 15 N NMR relaxation. *J Am Chem Soc.* 2011;133:14184–7.
- Duss O, Michel E, Yulikov M, Schubert M, Jeschke G, Allain FH. Structural basis of the non-coding RNA RsmZ acting as a protein sponge. *Nature.* 2014;509:588–92.
- Saio T, Guan X, Rossi P, Economou A, Kalodimos CG. Structural basis for protein antiaggregation activity of the trigger factor chaperone. *Science.* 2014;344:1250494.
- Deshmukh L, Schwieters CD, Grishaev A, Ghirlando R, Baber JL, Clore GM. Structure and dynamics of full-length HIV-1 capsid protein in solution. *J Am Chem Soc.* 2013;135:16133–47.
- Mondal S, Shet D, Prasanna C, Atreya HS. High yield expression of proteins in *E. coli* for NMR studies. *Adv Biosci Biotechnol.* 2013;4:751–767.
- Olson LJ, Peterson FC, Castonguay A, Bohnsack RN, Kudo M, Gotschall RR, Canfield WM, Volkman BF, Dahms NM. Structural basis for recognition of phosphodiester-containing lysosomal enzymes by the cation-independent mannose 6-phosphate receptor. *Proc Natl Acad Sci USA.* 2010;107:12493–8.
- Nyarko A, Singarapu KK, Figueroa M, Manning VA, Pandelova I, Wolpert TJ, Ciuffetti LM, Barbar E. Solution NMR structures of *Pyrenophora tritici-repentis* ToxB and its inactive homolog reveal potential determinants of toxin activity. *J Biol Chem.* 2014;289:25946–56.
- Tomida M, Kimura M, Kuwata K, Hayashi T, Okano Y, Era S. Development of a high-level expression system for deuterium-labeled human serum albumin. *Jpn J Physiol.* 2003;53:65–9.
- Morgan W, Kragt A, Feeney J. Expression of deuterium-isotope-labelled protein in the yeast *Pichia pastoris* for NMR studies. *J Biomol NMR.* 2000;17:337–47.
- Pickford AR, O'Leary JM. Isotopic labeling of recombinant proteins from the methylotrophic yeast *Pichia pastoris*. *Methods Mol Biol.* 2004;278:17–33.
- Massou S, Puech V, Talmont F, Demange P, Lindley ND, Tropis M, Milon A. Heterologous expression of a deuterated membrane-integrated receptor and partial deuteration in methylotrophic yeasts. *J Biomol NMR.* 1999;14:231–9.
- Dutta A, Saxena K, Schwalbe H, Klein-Seetharaman J. Isotope labeling in mammalian cells. *Methods Mol Biol.* 2012;831:55–69.

26. Saxena K, Dutta A, Klein-Seetharaman J, Schwalbe H. Isotope labeling in insect cells. *Methods Mol Biol.* 2012;831:37–54.
27. O'Neill RA. Enzymatic release of oligosaccharides from glycoproteins for chromatographic and electrophoretic analysis. *J Chromatogr A.* 1996;720:201–15.
28. Trimble RB, Tarentino AL. Identification of distinct endoglycosidase (endo) activities in flavobacterium meningosepticum: endo F1, endo F2, and endo F3. Endo F1 and endo H hydrolyze only high mannose and hybrid glycans. *J Biol Chem.* 1991;266:1646–51.
29. Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, Hart GW, Etzler ME. *Essentials of Glycobiology.* Cold Spring Harbor: Cold Spring Harbor Laboratory Press; 2009.
30. Maina CV, Riggs PD, Grandea AG 3rd, Slatko BE, Moran LS, Tagliamonte JA, McReynolds LA, Guan CD. An *Escherichia coli* vector to express and purify foreign proteins by fusion to and separation from maltose-binding protein. *Gene.* 1988;74:365–73.
31. Kapust RB, Waugh DS. *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci.* 1999;8:1668–74.
32. Sorensen HP, Mortensen KK. Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli*. *Microb Cell Fact.* 2005;4:1.
33. Raran-Kurussi S, Waugh DS. The ability to enhance the solubility of its fusion partners is an intrinsic property of maltose-binding protein but their folding is either spontaneous or chaperone-mediated. *PLoS One.* 2012;7:e49589.
34. Nozach H, Fruchart-Gaillard C, Fenaille F, Beau F, Ramos OHP, Douzi B, Saez NJ, Moutiez M, Servent D, Gondry M, et al. High throughput screening identifies disulfide isomerase DsbC as a very efficient partner for recombinant expression of small disulfide-rich proteins in *E. coli*. *Microb Cell Fact.* 2013;12:1–16.
35. Baneyx F, Mujacic M. Recombinant protein folding and misfolding in *Escherichia coli*. *Nat Biotech.* 2004;22:1399–408.
36. de Marco A. Strategies for successful recombinant expression of disulfide bond-dependent proteins in *Escherichia coli*. *Microb Cell Fact.* 2009;8:26.
37. Sun SY, Xu Y, Wang D. Purification and biochemical characterization of an intracellular lipase by *Rhizopus chinensis* under solid-state fermentation and its potential application in the production of eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA). *J Chem Technol Biotechnol.* 2009;84:435–41.
38. Xu Y, Wang D, Mu XQ, Ni YQ. Efficient esterification of sorbitan oleate by lipase in a solvent-free system. *J Am Oil Chem Soc.* 2003;80:647–51.
39. Sun SY, Xu Y, Wang D. Novel minor lipase from *Rhizopus chinensis* during solid-state fermentation: biochemical characterization and its esterification potential for ester synthesis. *Bioresour Technol.* 2009;100:2607–12.
40. Qin H, Yan X, Yun T, Dong W. Biodiesel production catalyzed by whole-cell lipase from *Rhizopus chinensis*. *Chin J Catal.* 2008;29:41–6.
41. Yu X, Xu Y, Wang R. Enhancement of activity of *Rhizopus chinensis* lipase by directed evolution. *J Biotechnol.* 2010;150:344.
42. AlphaView software. http://www.proteinsimple.com/software_alpha-view.html. Accessed Dec 15, 2015.
43. Jansson M, Li YC, Jendeberg L, Anderson S, Montelione GT, Nilsson B. High-level production of uniformly ¹⁵N- and ¹³C-enriched fusion proteins in *Escherichia coli*. *J Biomol NMR.* 1996;7:131–41.
44. Goto NK, Gardner KH, Mueller GA, Willis RC, Kay LE. A robust and cost-effective method for the production of Val, Leu, Ile (delta 1) methyl-protonated ¹⁵N-, ¹³C-, 2H-labeled proteins. *J Biomol NMR.* 1999;13:369–74.
45. Kordel M, Hofmann B, Schomburg D, Schmid RD. Extracellular lipase of *Pseudomonas* sp. strain ATCC 21808: purification, characterization, crystallization, and preliminary X-ray diffraction data. *J Bacteriol.* 1991;173:4836–41.
46. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax AD. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR.* 1995;6:277–93.
47. Sparky NMR Assignment and Integration Software. <http://www.cgl.ucsf.edu/home/sparky>. Accessed July 10, 2015.
48. Pickford AR, O'Leary JM. Isotopic labeling of recombinant proteins from the methylotrophic yeast *Pichia pastoris*. *Methods Mol Biol.* 2004;278:17–33.
49. Demain AL, Vaishnav P. Production of recombinant proteins by microbes and higher organisms. *Biotechnol Adv.* 2009;27:297–306.
50. Derman AI, Prinz WA, Belin D, Beckwith J. Mutations that allow disulfide bond formation in the cytoplasm of *Escherichia coli*. *Science.* 1993;262:1744–7.
51. Salinas G, Pellizza L, Margenat M, Flo M, Fernandez C. Toned *Escherichia coli* as a host for the expression of disulfide-rich proteins. *Biotechnol J.* 2011;6:686–99.
52. Stewart EJ, Aslund F, Beckwith J. Disulfide bond formation in the *Escherichia coli* cytoplasm: an in vivo role reversal for the thioredoxins. *EMBO J.* 1998;17:5543–50.
53. Jones NG, Nietlispach D, Sharma R, Burke DF, Eyres I, Mues M, Mott HR, Carrington M. Structure of a glycosylphosphatidylinositol-anchored domain from a trypanosome variant surface glycoprotein. *J Biol Chem.* 2008;283:3584–93.
54. Chattopadhyay A, Jones NG, Nietlispach D, Nielsen PR, Voorheis HP, Mott HR, Carrington M. Structure of the C-terminal domain from *Trypanosoma brucei* variant surface glycoprotein MITat1.2. *J Biol Chem.* 2005;280:7228–35.
55. Kontaxis G, Clore GM, Bax A. Evaluation of cross-correlation effects and measurement of one-bond couplings in proteins with short transverse relaxation times. *J Magn Reson.* 2000;143:184–96.
56. Downing AK. *Protein nuclear magnetic resonance techniques.* Heidelberg: Springer Science & Business Media; 2004.
57. Peng H, Hu Y, Zhou A, Jin C, Pan W. Solution structure of a *Plasmodium falciparum* AMA-1/MSP 1 chimeric protein vaccine candidate (PfCP-2.9) for malaria. *Malaria J.* 2010;9:76.
58. Hong Z, Nowakowski M, Spronk C, Petersen SV, Andreassen PA, Kozminski W, Mulder FA, Jensen JK. The solution structure of the MANEC-type domain from hepatocyte growth factor activator inhibitor-1 reveals an unexpected PAN/apple domain-type fold. *Biochem J.* 2015;466:299–309.
59. Zhang L, DeRider M, McCornack MA, Jao SC, Isern N, Ness T, Moyer R, LiWang PJ. Solution structure of the complex between poxvirus-encoded CC chemokine inhibitor vCCI and human MIP-1β. *Proc National Acad Sci.* 2006;103:13985–90.
60. Rundqvist L, Tengel T, Zdunek J, Bjorn E, Schleucher J, Alcocer MJ, Larsson G. Solution structure, copper binding and backbone dynamics of recombinant Ber e 1-the major allergen from Brazil nut. *PLoS One.* 2012;7:e46435.
61. Clark L, Zahm JA, Ali R, Kukula M, Bian L, Patrie SM, Gardner KH, Rosen MK, Rosenbaum DM. Methyl labeling and TROSY NMR spectroscopy of proteins expressed in the eukaryote *Pichia pastoris*. *J Biomol NMR.* 2015;62:239–45.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

