Protein Expression and Purification 145 (2018) 77–84

Contents lists available at ScienceDirect

Protein Expression and Purification

journal homepage: www.elsevier.com/locate/yprep

In vivo biotinylated calpastatin improves the affinity purification of human m-calpain

Hung Huy Nguyen[a,b], Alexander N. Volkov[a,b,c], Guy Vandenbussche[d], Peter Tompa[a,b,e], Kris Pauwels[a,b,*]

[a] *VIB - VUB Centre for Structural Biology (VIB CSB), Vlaams Instituut voor Biotechnologie (VIB), Brussels, Belgium*
[b] *Structural Biology Brussels (SBB), Department of Biotechnology (DBIT), Vrije Universiteit Brussel (VUB), Brussels, Belgium*
[c] *Jean Jeener NMR Centre, Vrije Universiteit Brussel (VUB), Brussels, Belgium*
[d] *Laboratory for the Structure and Function of Biological Membranes, Centre for Structural Biology and Bioinformatics, Université Libre de Bruxelles (ULB), Brussels, Belgium*
[e] *Institute of Enzymology, Research Centre for Natural Sciences of the Hungarian Academy of Sciences, Budapest, Hungary*

ARTICLE INFO

Keywords:
IDP
BirA
Biotin ligase
BAP-tag
AviTag
Calpain-2
Affinity chromatography
Protein co-expression

ABSTRACT

Recently we established a novel affinity purification method for calpain by exploiting the specific and reversible binding properties of its intrinsically disordered protein inhibitor, calpastatin. The immobilization strategy relied on the strength and specificity of the biotin - streptavidin interaction. Here, we report an improved and optimized method that even enables the general applicability of *in vivo* biotinylated (intrinsically disordered) proteins in any affinity capture strategy. Since *in vitro* chemical biotinylation is only accomplished with reagents that lack exact site specificity, it can not only cause sample heterogeneity but it can also hamper the functionality of the biotinylated molecules. Therefore, we have developed a recombinant expression protocol to produce *in vivo* biotinylated human calpastatin domain 1 (hCSD1) in *Escherichia coli*. We have experimentally verified that the biotinylated polypeptide tag is compatible with the intrinsically disordered state of hCSD1 and that it does not influence the functional properties of this intrinsically disordered protein (IDP). The *in vivo* biotinylated hCSD1 was then used without the need of any prepurification step prior to the affinity capturing of its substrate, human m-calpain. This leads to a simplified purification strategy that allows capturing the calpain efficiently from a complex biological mixture with only a single chromatogaphic step and in a considerably reduced timeframe. Our approach is generally applicable through the *in vivo* biotinylation of any IDP of interest, and its practical implementation will showcase the power to exploit the properties of IDPs in affinity capture strategies.

1. Introduction

The structural and biophysical characterization of proteins primarily relies on substantial amounts of recombinantly expressed protein material that is repeatedly produced to enable various experimental studies. Therefore, efficient protein production and purification is often a bottleneck in the pipeline of ambitious structural biology and drug development projects [1]. Calpains (EC 3.4.22.53) are such a class of difficult to purify proteases [2,3]. Mammalian calpains play an important biotechnological role in the post-mortem tenderization of meat [4], while the human variants are involved in pathogenesis of Alzheimer's disease, cancer and diabetes [1,5]. The activity of these intracellular calcium-dependent proteases is controlled by a specific endogenous inhibitor, called calpastatin [6].

Recently, we have reported a novel affinity purification protocol whereby we exploit the binding specificity and affinity of the intrinsically disordered protein (IDP) human calpastatin domain 1 (hCSD1) to purify its target protein, human m-calpain [7]. Upon further considering the advantage of biotin-based immobilization, we have explored the feasibility of using *in vivo* biotinylation for a simplified and improved high-affinity one-step purification. Particularly, the reproducibility and heterogeneity of the *in vitro* chemical biotinylation can be a limiting factor in the functionality of the capturing agent in the affinity chromatography. To this end, we have developed a BirA protein-biotin ligase (EC 6.3.4.15) co-expression strategy in *Escherichia coli* for the recombinant production of an avi-tagged hCSD1 construct. The AviTag™ (also known as biotin acceptor peptide, BAP-tag) is a 15 residues-long sequence (GLNDIFEAQKIEWHE) that is specifically recognized by the biotin ligase BirA [8]. BirA adds D-biotin covalently to this polypeptide tag via a biotinyl 5′-adenylate intermediate that

* Corresponding author. Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium.
E-mail address: kris.pauwels@vub.be (K. Pauwels).

https://doi.org/10.1016/j.pep.2018.01.002
Received 15 November 2017; Received in revised form 14 December 2017; Accepted 5 January 2018
Available online 13 January 2018

ultimately results in the formation of an amide-linkage between the carboxyl group of biotin and the epsilon-amino group of the modified lysine [9]. Interestingly, we wish to exploit the fact that only a single endogenous protein in *E. coli* has been reported to become biotinylated by BirA [10]. This is the biotin carboxyl carrier protein (BCCP; UniProt P0ABD8), which is a 156 residues-long protein. While recombinantly expressed IDPs can be purified easily upon boiling the resuspended bacterial pellet [7,11], we anticipate that the well-folded BCCP will precipitate upon boiling the bacterial lysate.

Here, we show that the avi-tagged hCSD1 construct can be biotinylated efficiently and specifically in *E. coli*. We also show that the avi-tag is fully compatible with the intrinsically disordered nature of hCSD1 and hence has no negative effect on its interaction with m-calpain. This *in vivo* biotinylated hCSD1 is then used to efficiently purify its natural binding partner in a single chromatographic step.

2. Materials and methods

2.1. Bacterial strains and plasmid transformation

As recombinant expression vector we used the pET22b plasmid (Novagen) whereby the hCSD1 gene (residues 137 - 277 of the human calpastatin (UniProt ID P20810)) is fused with an N-terminal AviTag™ sequence, extended with a stop codon and cloned within the NdeI and EcoRI restriction sites through a commercial gene synthesis service (Genscript). This plasmid (which we designated pAvihCSD1) contains an ampicillin resistant gene and an inducible T7 promoter system. For the co-expression of the BirA protein and the *in vivo* biotinylation experiments, we employed two commercially available bacterial strains: *E. coli* AVB100 and AVB101 (Avidity LLC). AVB100 is an *E. coli* K12 strain [*MC1061 araD139 δ(ara-leu)7696 δ(lac)l74 galU galK hsdR2(rK-mK+) mcrB1 rpsL(StrR)*], whereby the *birA* gene is integrated into the bacterial chromosome. In this strain the BirA expression can be induced with arabinose. The AVB101 is an *E. coli* B strain (*hsdR, lon11, sulA1*) that contains pBirAcm, an IPTG inducible plasmid comprising the *birA* gene and a chloramphenicol-resistance gene.

Since the *E. coli* K12 strain (AVB100) does not contain the T7 RNA polymerase system we resorted to the pT7pol23 and pT7pol26 plasmids (generously provided by Dr Henri De Greve), which both contain a kanamycin-resistance gene, to express T7 polymerase under heat- and IPTG-inducible promoter, respectively [12]. For the co-transformation of the plasmids, the pAvihCSD1 and each of the pT7 plasmids were transformed in two separate steps into chemically competent *E. coli* AVB100 and AVB101 cells. In first instance, 20 ng of the pT7pol23 or pT7pol26 was transformed into the above-mentioned *E. coli* cells by the heat shock method. The successful transformants were selected using Lysogeny broth (LB)-agar plates that were supplemented with 50 μg/mL kanamycin for the AVB100 strain or 10 μg/mL chloramphenicol and 50 μg/mL kanamycin for the AVB101 strain. A single colony of those transformants was used to generate chemically competent cells for the subsequent transformation of the pAvihCSD1 plasmid. The final transformants were selected using the above LB-agar plate supplemented with 100 μg/mL ampicillin.

The untagged version of hCSD1 is produced in *E. coli* BL21-AI, while the heterodimeric C105A mutant of human m-calpain was produced in BL21(DE3), as reported previously for both proteins [7].

2.2. Recombinant protein production and in vivo biotinylation of hCSD1

Recombinant protein expression was performed in 2 L baffled flasks containing 1 L LB (Duchefa) while shaking at 180 rpm at 37 °C in a minitron orbital shaker-incubator (InFors HT). Before the inoculation, D-biotin (Sigma) was added to the media to the final concentration of 50 μg/mL. For the pT7pol23 containing cells, the culture temperature was raised to 42 °C for 1 h when the OD_{600} reached 0.4 and subsequently reduced to 37 °C prior to the addition of 1 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG). For the pT7pol26 containing cells, 0.4% L-arabinose was added to the culture at OD_{600} of 0.4, prior to induction with 1 mM IPTG when the OD_{600} reached 0.6. The cell cultures were collected after 5 h of induced protein expression. The collected cell cultures were centrifuged at 5000 g for 15 min, the supernatants were discarded and the cell pellets were stored at -20 °C.

For the ^{15}N-isotope labelling of hCSD1 and avi-hCSD1, the *E. coli* BL21-AI and BL21(DE3) strains that contained the corresponding plasmids were grown in M9 minimal medium (1 L contained 6.0 g Na_2HPO_4, 3.0 g KH_2PO_4, 0.5 g NaCl, 1.0 g $^{15}NH_4Cl$, 1 mM $MgSO_4$, 0.6 mM $CaCl_2$, 1 mg/L D-biotin, 1 mg/L thiamine and supplemented with the required antibiotics) using ^{15}N-isotope NH_4Cl as the sole nitrogen source (Cambridge Isotopes Laboratories). As carbon source we added 0.2% D-glucose for *E. coli* BL21(DE3)) or 0.2% glycerol for *E. coli* BL21-AI to the minimal medium. The cells were grown at 37 °C as described earlier and protein expression was induced with 1 mM IPTG at OD_{600} of 0.6. The cell cultures were collected after 5 h of induction.

2.3. Protein purification procedures

2.3.1. Purification of various hCSD1 constructs

The hCSD1-containing cell pellets (i.e. untagged hCSD1, avi-hCSD1 and b-avi-hCSD1) were resuspended in a 50 mM Tris-HCl pH 7.5 buffer containing 0.1 mM ethylenediaminetetraacetic acid (EDTA), 2 mM dithiothreitol (DTT) and protease inhibitor cocktail (Roche tablets), followed by heating at 90 °C for 15 min in a ThermoScientific water-bath. The lysates were centrifuged at 21119 g for 30 min. After filtering the cleared lysates through a 0.2 μm cut-off polyethersulfone membrane (Sarstedt), the samples were kept on ice and all subsequent protein purification steps were performed at 4 °C. The sample was applied to a 5 mL HiTrap DEAE anion exchange column (GE Healthcare) equilibrated with 50 mM Tris-HCl, 0.1 mM EDTA, 2 mM DTT pH 7.5 and the protein was eluted with a step elution of 150 mM NaCl [7]. Depending on the sample purity an optional size exclusion chromatography was performed using a Superdex 75 10/30 GL (GE Healthcare) in 20 mM 3-(N-morpholino)propanesulfonic acid (MOPS), 1 mM Tris(2-carboxyethyl)phosphine (TCEP) pH 7.5.

2.3.2. Affinity purification of human m-calpain with in vivo biotinylation hCSD1

The lysate containing *in vivo* biotinylated avi-tagged hCSD1 (b-avi-hCSD1) and the purified b-avi-hCSD1 was loaded into a HiTrap Streptavidin HP 1 ml column (GE Healthcare) equilibrated with the binding buffer MOPS pH 7.5, 1 mM TCEP and 150 mM NaCl) supplemented with 5 mM $CaCl_2$. The column was washed with 5 column volumes of the same binding buffer, followed by loading the C105A-calpain lysate supplemented with 5 mM $CaCl_2$ to the column. The calpain was eluted from the b-avi-hCSD1-containing resin by applying 10 mM EDTA (in binding buffer).

2.4. In vitro chemical biotinylation of hCSD1

The chemical biotinylation of hCSD1 was accomplished using EZ-Link™ NHS-PEG4-Biotinylation kit (Pierce Biotechnology) according to the manufacturer's instructions and as described previously [7]. The extent of the biotinylation was determined by a colorimetric quantification kit provided by the same manufacturer (Pierce Biotechnology) utilizing the absorbance-based 2-(4′-hydroxyazobenzene) benzoic acid (HABA) dye.

2.5. Electrophoresis and western blotting

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was run using 12.5% polyacrylamide gel with a PageRuler™ prestained protein ladder (ThermoFisher Scientific) in Tris-glycine buffer at 200 V for 40 min. To directly visualize all the proteins, the gel was

stained using PageBlue™ staining solution (ThermoFisher Scientific).

After performing SDS-PAGE, the gel was transferred to a 0.2 µm polyvinylidene fluoride (PVDF) membrane using Trans-blot Turbo blotting system (Bio-Rad) at 1.3 A for 14 min. The transferred membrane was blocked with phosphate-buffered saline (PBS) containing 0.1% Tween-20 and 5% non-fat dried milk for 30 min, followed by incubating with horseradish peroxidase (HRP)-conjugated streptavidin (ThermoFisher Scientific) at a 1:10000 dilution, on a vertical rotator for 2 h at room temperature. The membrane was further rinsed three times with PBS, and washed two times with PBS supplemented with 0.1% Tween-20 for 20 min each. The membrane was incubated with Pierce ECL Western blotting substrate for 2 min after visualization using Image Lab software with ChemiDoc XRS System (Bio-Rad).

2.6. Mass spectrometry sample preparation and analysis

2.6.1. Protein intact mass analysis

The four proteins that were analyzed (hCSD1, (chemically biotinylated) b-hCSD1, avi-hCSD1 and (*in vivo* biotinylated) b-avi-hCSD1) were solubilized in 0.5% TFA, desalted on ZipTip C18 (Millipore) and eluted in 5 µL 50% acetonitrile/1% formic acid (v/v). The samples were loaded into a nanoflow capillary (Thermo Fisher Scientific). Electron spray ionisation (ESI) mass spectra were acquired on a quadrupole time-of-flight instrument (Q-Tof Ultima – Waters/Micromass) operating in the positive ion mode, equipped with a Z-spray nanoelectrospray source. Data acquisition was performed using a MassLynx 4.1 system. The spectra were recorded in the V mode and represented the combination of 1 s scans. The molecular mass of the proteins was determined after processing of the raw data with the software MaxEnt1.

2.6.2. Peptide microsequencing

The protein bands were excised from the SDS-PAGE gel and crushed in small pieces of about 1 mm^3. The gel pieces were washed with 25 mM NH_4HCO_3 and 50% CH_3CN/25 mM NH_4HCO_3 and then dried in a vacuum centrifuge. The gel pieces were swollen in an ice-cold bath in 10 µL of a digestion buffer containing 25 mM NH_4HCO_3, and 10 ng/µL of sequencing grade modified trypsin (Promega) or Endoproteinase Glu-C (Promega). The digestion was carried out overnight at 37 °C. The supernatant was collected and the peptides remaining in the gel pieces were extracted sequentially by 25 mM NH_4HCO_3, 50% CH_3CN/25 mM NH_4HCO_3 and 50% CH_3CN/5% HCOOH. The supernatant and the wash were pooled and dried in a vacuum centrifuge. The sample hCSD1 was also digested in solution overnight at 37 °C with sequencing grade Endoproteinase Glu-C (protein/enzyme mass ratio 25:1).

Before analysis by mass spectrometry, the sample was desalted on ZipTip C18 (Millipore) and eluted in acetonitrile 50%/formic acid 1% (v/v). The samples were loaded into a nanoflow capillary (Proxeon) and ESI mass spectra were acquired on a quadrupole time-of-flight instrument (Q-Tof Ultima - Waters/Micromass) operating in the positive ion mode, equipped with a Z-spray nanoelectrospray source. The tryptic peptides were analyzed by tandem mass spectrometry (MS/MS). After processing of the MS/MS data by the maximum entropy data enhancement program MaxEnt 3, the amino acid sequences were semi-automatically deduced using the peptide sequencing program PepSeq (Waters, Milford, USA). Based on the peptide sequences, the proteins were identified using the MASCOT Sequence Query.

2.7. Biophysical characterization

2.7.1. Nuclear magnetic resonance analysis

The ^{1}H-^{15}N-transverse relaxation-optimized spectroscopy (TROSY) spectra of 0.2 mM hCSD1 or 0.44 mM avi-hCSD1 constructs were recorded at 25 °C in 20 mM MOPS pH 7.5, 1 mM TCEP, and 10% D_2O (Sigma) for the lock. The experiments were acquired on an 800 MHz Bruker NMR spectrometer equipped with TCI cryoprobe for enhanced sensitivity and processed in TopSpin v3.2.

2.7.2. Bio-layer interferometry

The b-hCSD1 and b-avi-hCSD1 at a concentration of 2.5 and 50 µg/mL, respectively, were immobilized with a similar BLI intensity onto SA biosensors (ForteBio) that contain covalently attached streptavidin. The buffer consisted of 20 mM MOPS pH 7.5, 5 mM $CaCl_2$, 0.1% Tween-20 and 0.1% bovine serum albumin (BSA), which was used to minimize the non-specific interaction of calpain with the biosensors. A calpain concentration range from 60 nM to 500 nM was used during the association phase. Each of those concentrations had its corresponding non-functionalized biosensor as a parallel reference sensor. The binding data were recorded at 25 °C with agitation at 1000 rpm using the ForteBio data acquisition software v9.0. The fitting of individual binding curves was done using ForteBio data analysis software v9.0. The steady state analysis was analyzed in Graphpad Prism 7 using non-linear regression with the equation $Y = B_{max}*X/(K_D + X)$, where B_{max} is the maximum specific binding, and K_D represents the equilibrium dissociation constant.

2.7.3. Homology modelling

The human homology model was generated with the SwissModel server (https://swissmodel.expasy.org/) using 3DF0.pdb as structural template and UNIPROT codes P17655, P04632 and P20810 as sequences for CAPN2, CAPNS1 and calpastatin region 137-277 (i.e. hCSD1), respectively [13]. The figures were generated with Pymol (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC).

3. Results and discussion

3.1. In vivo biotinylation of avi-hCSD1 and purification of different hCSD1 variants

Since the hCSD1 sequence has no natural BAP-sequence [14], we generated an avi-tagged hCSD1 construct (avi-hCSD1) so that *in vivo* biotinylation would occur only specifically at the N-terminal avi-tag lysine site. Therefore, we tested different expression strategies to establish an *in vivo* biotinylation protocol for avi-hCSD1. Since the recombinant overexpression of hCSD1 was reported to be successful with inducible T7 promoter plasmids, we opted for a pET22 plasmid to produce the avi-hCSD1 in *E. coli* [7,15]. Because the two commercially available *E. coli* strains (AVB100 and AVB101) that enable *in vivo* biotinylation do not contain the T7 RNA polymerase, we tested 2 plasmids (pT7pol23 and pT7pol26) that carry an inducible T7 RNA polymerase gene for its applicability in our production strategy [12]. The results of the SDS-PAGE and Western blotting analysis of these expression tests are depicted in Fig. 1.

Even though, the expression of the *in vivo* biotinylated avi-hCSD1 (represented as b-avi-hCSD1) was not clearly observable through the appearance of one additional prominent protein band via SDS-PAGE (Fig. 1A), the Western blotting analysis of the same gel showed that the *in vivo* biotinylated form of the protein expressed well in all combinations of plasmids and *E. coli* strains (Fig. 1B). Upon comparing the pT7pol23 and pT7pol26 plasmids, we noticed that the cells containing pT7pol23 displayed a more pronounced leaky expression of b-avi-hCSD1 prior to induction with IPTG as compared to the pT7pol26 construct. This observation can be rationalized by the fact that pT7pol23 contains a thermally inducible promoter that is not tightly controlled at 30 °C. Thus, this protocol can probably be optimized by reducing the growth temperature from 37 °C to 28 °C (or lower) prior to shifting it to 42 °C to induce expression [16].

From the Western blotting it also becomes clear that the anticipated endogenously biotinylated BCCP protein was present in the induced sample (albeit at variable quantities), as its band runs at the expected molecular mass of 16.7 kDa (Fig. 1B). In the AVB101 strain, another (unexpected) biotinylated contaminant was observed around 35 kDa. Based on those observations, we have selected the combination of the *E. coli* AVB100 strain and the pT7pol26 plasmid for further purification as

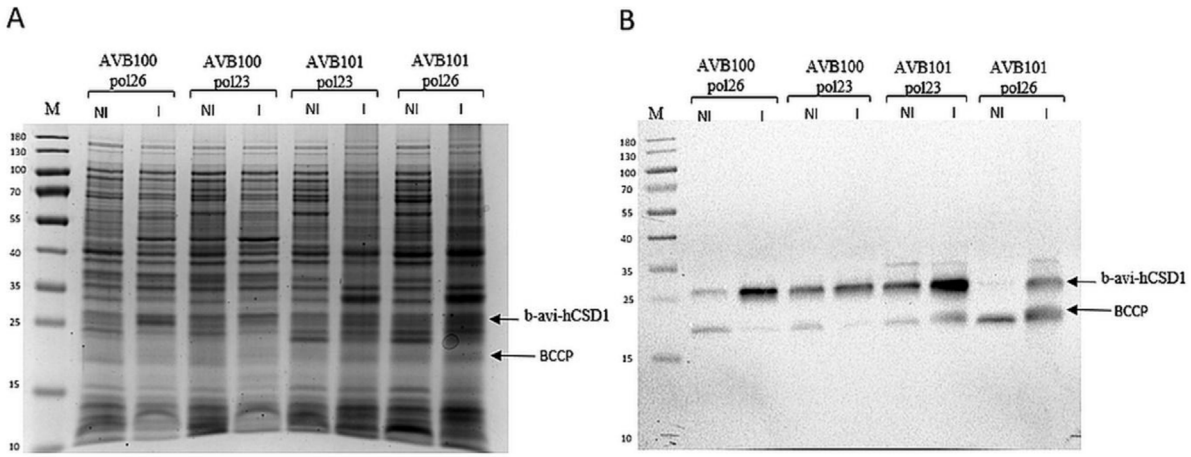

Fig. 1. Recombinant expression analysis of *in vivo* biotinylated avi-tagged hCSD1 (b-avi-hCSD1) based on (A) SDS-PAGE with Coomassie staining and (B) Western blotting with HRP-conjugated streptavidin. The different combinations of *E. coli* strains (the *E. coli* K12-derived AVB100 and the *E. coli* B-based AVB101) that contain the pAvihCSD1 plasmid and either pT7pol23 or pT7pol26, are indicated at the top of the image. NI represents the non-induced cells, while I represents the sample of the cells at 4 h after the induction of protein expression. The proteins of interest are indicated with an arrow and the molecular weight (in kDa) of the marker proteins is indicated on the left hand side of the images.

it had a good level of b-avi-hCSD1 expression and less contaminating biotinylated proteins. Since it was also reported that BirA would be less soluble at elevated temperatures, we preferentially opted for the pT7pol26-based expression system whereby the fermentation temperature is kept constant [17].

With *E. coli* AVB100 that contains the pAvihCSD1 and pT7pol26 plasmids we scaled up the production of b-avi-hCSD1 (*in vivo* biotinylated avi-hCSD1). After the heat treatment of the resuspended cells, b-avi-hCSD1 was isolated from the cleared lysate by an anion exchange chromatography and a polishing size exclusion chromatography (Fig. 2). During this procedure, we observed the presence of a heat-resistant contaminating protein, which we further identified by MS/MS peptide microsequencing as flagellin. We found that flagellin can be separated from b-avi-hCSD1 by size exclusion chromatography using a Superdex 75 10/30 GL column (Fig. 2B and C). Although the presence of flagellin would not affect our downstream calpain purification strategy (see section 3.5) as it does not bind to the streptavidin column and is located in the flow-through (Fig. 7 - lane 2), we have also tried to apply the commercial detergent-based BugBuster® protein extraction reagent (Novagen) as an alternative lysis method in an attempt to remove flagellin, without success.

The purification of the untagged version of hCSD1 and its *in vitro* chemical biotinylation (resulting in b-hCSD1) allowed for further functional and comparative studies with the b-avi-hCSD1. In this context, we can also report that the production of ^{15}N-isotope labelled hCSD1 and avi-hCSD1 through bacterial expression in minimal medium and its subsequent purification did not reveal any additional experimental obstacles e.g. degradation or presence of contaminants …).

3.2. Mass spectrometry analysis of hCSD1 and avi-hCSD1 biotinylation

We confirmed by ESI MS analysis that the experimental molecular mass of purified b-avi-hCSD1 (17047 Da) was increased with a shift of mass of 226 Da as compared to avi-hCSD1 (experimental molecular mass of 16821 Da). This corresponds to the addition of a single D-biotin molecule to the protein by the BirA biotin-protein ligase. By MS/MS peptide microsequencing and the direct comparison of avi-hCSD1 and b-avi-hCSD1 peptides, we further confirmed that the only site of biotinylation is indeed the first lysine in the sequence, which is located within the avi-tag. This experimental result also reveals that the heat treatment that is typically employed for the purification of IDPs does not alter the integrity of the protein and doesn't lead to any further chemical modifications.

With regard to the chemical (*in vitro*) biotinylation, by combining

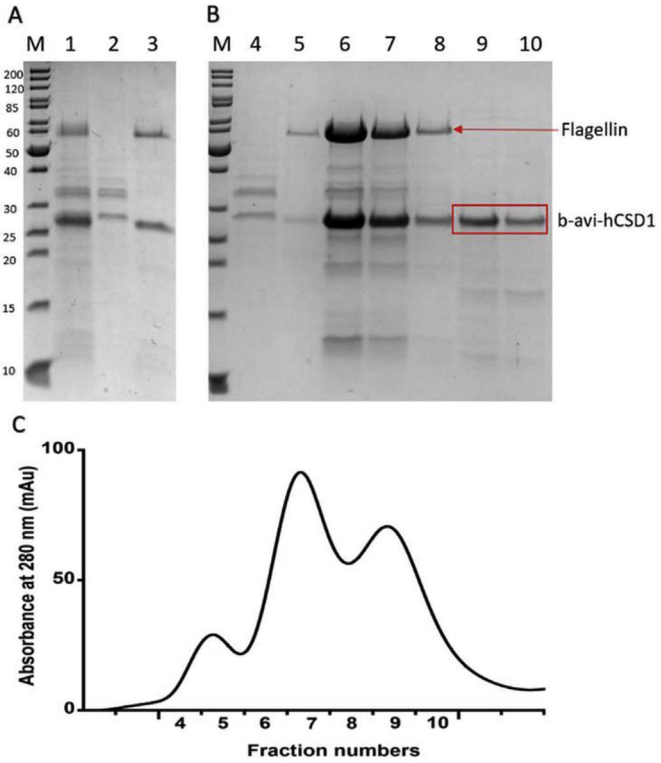

Fig. 2. SDS-PAGE analysis with Coomassie staining of (A) the anion exchange chromatography and (B) the gelfiltration elution fractions of b-avi-hCSD1, which are corresponding to (C) the gelfiltration elution profile. 1: lysate of *E. coli* AVB100 containing pT7pol26 and pAvihCSD1, after heat treatment and clearance by centrifugation; 2: flow-through; 3: elution fraction at 150 mM NaCl; Panel B represents the gelfiltration elution fractions for avi-hCSD1 whereby the numbering corresponds to the elution fractions of the chromatogram that is depicted in Panel C.

peptide mass fingerprinting and MS/MS peptide microsequencing we could determine that most of the lysine residues in b-hCSD1 were biotinylated. For each detected digested peptide that contains one or two lysine residues, we observed a 473 Da or 946 Da shift of its experimental molecular mass, respectively. This is in perfect agreement with the expected mass of 473.22 Da that is added to the target lysine by the reaction with NHS-PEG4-biotin, which is a PEGylated and amine-reactive variant of D-biotin. We could detect the appearance of biotinylated peptides that cover 10 out of 13 lysine residues in the

hCSD1 sequence. Also the concomitant disappearance of the corresponding unmodified peptides was monitored. Unfortunately, in this way, it was not possible to detect the N-terminal peptide comprising residues 1-38 from hCSD1, which contains 3 lysines. The undetectability of this peptide (in both hCSD1 and b-hCSD1) is likely due to its large size (theoretical molecular mass of 3959.32 Da for the non-biotinylated peptide). However, we could compare the experimental molecular mass of the intact unmodified hCSD1 of 14764 Da (in perfect accord with the expected theoretical mass of 14764.28 Da) to the total experimental mass of b-hCSD1 that was 21018 Da. Under the assumption that the 13 lysines in the polypeptide are biotinylated, we expected a molecular mass of 20913 Da (i.e. 14764 + (13 × 473) = 20913 Da). Thus, from the mass spectrometry analysis it is plausible to conclude that the 13 lysines are biotinylated, while an extra mass of 105 Da is present. The nature of this additional 105 Da, likely a posttranslational modification, remains enigmatic. Yet, we can rule out that this (post-translational) modification occurs during the boiling of the cell resuspension since both avi-hCSD1 and hCSD1 had the expected and correct experimental mass. Hence, some additional modification should have occurred only during (or after) the chemical biotinylation reaction. Nevertheless, since we detected the correct masses for all digested peptides except the N-terminal 38-residues long peptide, the modification can only be present in this segment.

With this approach combining molecular mass determination, peptide mass fingerprinting and MS/MS microsequencing, we suggested that all lysine positions in b-hCSD1 have been biotinylated (even identified at a residue-level). This was somewhat surprising based on the result of the HABA-based colorimetric quantification method where on average only 8 biotin molecules were attached per hCSD1 molecule. In our previous study, we reported based on this colorimetric assay that the chemical biotinylation of hCSD1 (i.e. b-hCSD1) yielded 2 mol biotin per mol of hCSD1 [7]. While this could point to the rather non-specific nature of the chemical biotinylation and the possible batch-to-batch dependence, this discrepancy might also be explained through the inaccuracy of determining the concentrations of IDPs [18].

Nonetheless, this analysis confirms that the *in vivo* biotinylation strategy allowed us to strictly control the extent and position in the sequence of the biotin addition, in contrast to the *in vitro* biotinylation that yields a multi-biotinylated protein.

3.3. Effect of the avi-tag on the structural properties of hCSD1 using NMR

We also examined if the avi-tag would interfere with the IDP nature of hCSD1. Therefore, we collected a ^{1}H-^{15}N transverse relaxation-optimized spectroscopy (TROSY) of the avi-tagged hCSD1 and compared it to the spectrum of its untagged counterpart (Fig. 3). The TROSY spectrum displays the low dispersion in the proton dimension that is typical for an intrinsically disordered protein [19,20]. We could observe a number of additional signals in the avi-hCSD1 spectrum, which agrees with the presence of 18 additional residues in the avi-tagged construct as compared to the untagged hCSD1. This also includes the appearance of a Trp indole resonance at 10 ppm δH that is uniquely present in the avi-tag (Fig. 3). Since the chemical shift is very sensitive to the chemical environment, for instance due to structural changes, we conclude from this dataset that the mere presence of the N-terminal avi-tag does not interfere with the IDP nature of hCSD1.

3.4. Effect of the biotinylation of hSCD1 on its binding properties

Since our mass spectrometry results revealed that chemical biotinylation lead to the modification of multiple lysines in hCSD1, we expected that the subsequent streptavidin-based immobilization of b-hCSD1 could display a reduced functionality in binding to its calpain substrate. With the b-avi-hCSD1, we anticipated that the full functionality of hCSD1 would be retained. Therefore, the effect of the biotinylation method on the binding properties of hCSD1 to calpain was

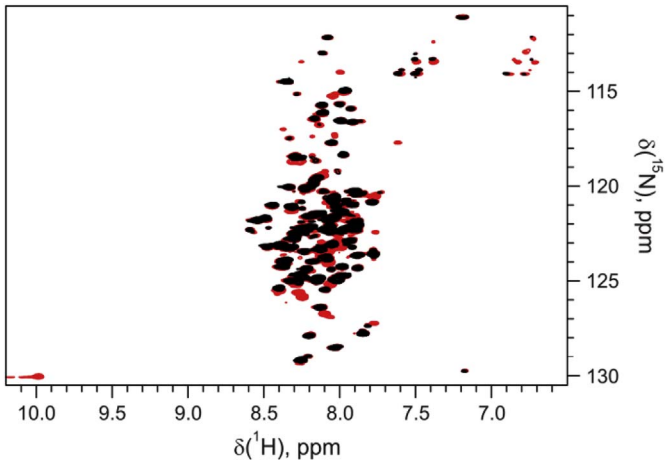

Fig. 3. Overlay of the ^{1}H-^{15}N transverse relaxation-optimized spectroscopy of hCSD1 (black) and avi-hCSD1 (Red). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

analyzed using bio-layer interferometry (BLI). This BLI analysis relied on the immobilization of the biotinylated hCSD1 variants to the biosensor surface with subsequent exposure to various concentrations of m-calpain, whereby the association and dissociation phases were recorded. With this strategy, we could detect the concentration-dependent binding of m-calpain to both hCSD1 variants in real time (Fig. 4). The binding curves of b-avi-hCSD1 suggested that a simple 1:1 binding occurred, as the experimental data could be fitted with this model (Fig. 4A). For the immobilized b-hCSD1, the simple 1:1 model could not be applied to fit the data, whereas the heterogeneous binding model gave the best fitting results (Fig. 4B). This confirmed, while b-avi-hCSD1 displayed a homogeneous binding behavior to m-calpain, that the chemical biotinylation in the case of b-hCSD1 has led to an intrinsic sample heterogeneity. The affinity constant of the interaction of b-avi-hCSD1 to calpain (73.1 nM) derived from the steady state analysis was two times lower that of b-hCSD1 (134.4 nM) (Fig. 4C and D and Table 1). In addition to this apparent higher affinity, the association rate of calpain to b-avi-hCSD1 was also almost four times faster compared to b-hCSD1 (Table 1). These are likely explained as the result of the simultaneous biotinylation of multiple lysine residues along the hCSD1 sequence. This *in vitro* biotinylation might not only interfere with its flexibility when being immobilized onto the streptavidin biosensors, it could also reduce its overall binding capacity to calpain (as reflected in the reduced signal-to-noise for b-hCSD1 curves as compared to b-avi-hCSD1). The exact *modus operandi* of hCSD1 binding to the calcium-activated m-calpain surface was revealed when the crystal structures of the bound rat proteins were solved (Fig. 5A) [21,22]. The intrinsically disordered hCSD1 and its rat counterpart encompassed three segments that became structured upon binding: regions A and C form an alpha-helical element, while the extended region B actually covered the active site cleft of m-calpain (Fig. 6). Based on the x-ray structure (3DF0.pdb) we generated a homology model for the human complex of hCSD1 and m-calpain to visualize the position of the lysine residues in the visible structure (Fig. 5B) (supplementary material). These lysines are distributed along the sequence and 5 lysines (depicted in the spherical representation in orange in Fig. 5B) are actually modelled in the visible regions of hCSD1. It is therefore conceivable that upon biotinylating these lysines, the proper binding of b-hCSD1 to the calpain surface gets affected.

These results provide a strong evidence that biotinylation of hCSD1 specifically and perhaps IDPs in general, through an avi-tag is not only more convenient but also better since it preserves their (un)structured and fully functional state.

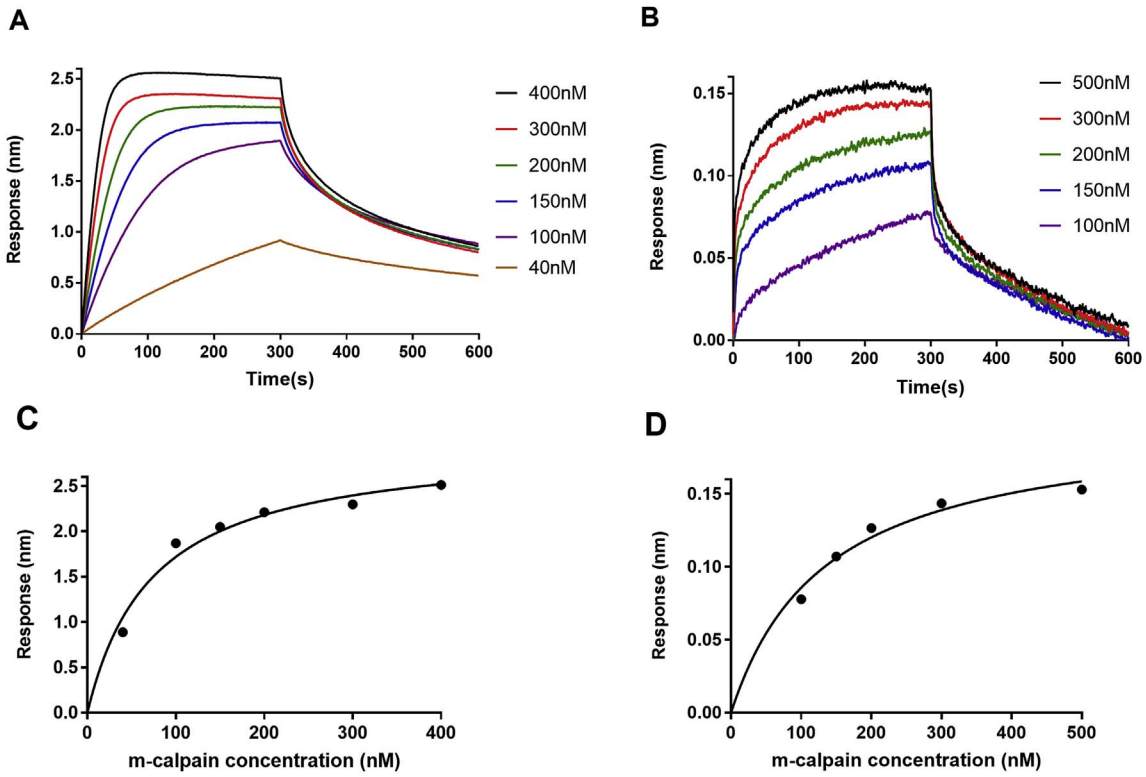

Fig. 4. Biolayer interferometry was used to monitor the binding kinetics (A and B) and steady-state analysis (C and D) of the interaction between immobilized b-avi-hCSD1 (A and C) or b-hCSD1 (B and D) and calcium-bound m-calpain in solution at various concentrations.

3.5. Affinity purification of human m-calpain based on in vivo biotinylated hCSD1

Since we successfully applied an *in vivo* biotinylation strategy and verified the functionality of the avi-hCSD1 in its binding to m-calpain, we wanted to implement the b-avi-hCSD1 in an affinity chromatography to purify its calpain substrate. Our previously published affinity purification method for calpain was based on b-hCSD1 and required the separate purification of hCSD1 and its subsequent chemical biotinylation (Fig. 6) [7]. With the *in vivo* biotinylated b-avi-hCSD1 we envisaged that we could omit the requirement for calpastatin isolation and *in vitro* biotinylation and directly load the b-avi-hCSD1 containing lysate onto the straptavidin-based resin.

By applying a b-avi-hCSD1-containing lysate directly to the streptavidin-containing resin, followed by the cleared lysate that contains the calcium-bound inactive C105A m-calpain and finally the elution of calpain by applying EDTA to the resin, the yield of the calpain purification was increased by two times as compared to applying chemically biotinylated hCSD1 (b-hCSD1). From 1 L LB in which *E. coli* BL21(DE3) expresses the heterodimeric and inactive C105A mutant m-calpain we obtained 2 mg upon this new purification strategy, while the purity is retained at the same level as described before (Fig. 7) [7]. This is probably due to the fact that the *in vivo* biotinylation occurs specifically (and solely) at the N-terminus, thus allowing a more efficient and functional immobilization of hCSD1. The multi-biotinylated b-hCSD1 (13 biotin molecules along the 140-residue-long polypeptide according to our mass spectrometry data) can probably occupy multiple streptavidin sites, thereby (1) causing heterogeneity in its binding behavior (less available binding motifs) or (2) reducing the total number of b-hCSD1 that can be immobilized to the column. The single-biotinylated b-avi-hCSD1 is likely to retain its full capacity to capture m-calpain from biological mixture (see section 3.4) and likely also obtains a more uniform and higher loading density on the resin. In addition to the increased yield, this strategy also entails less hands-on time, less materials and consumables and less energy since there is no need for prior purification of hCSD1 that requires chemical biotinylation. Despite having observed the heat-resistant flagellin and the biotinylated BCCP during the development of this strategy, both contaminants where not observed in the final sample (Fig. 7). Therefore, we can conclude that the combination of *in vivo* biotinylation of hCSD1 with the single-chromatographic-step purification of m-calpain constitutes a real improvement of the original purification strategy (Fig. 6).

4. Conclusion and future perspectives

We presented a novel and improved method to isolate m-calpain based on calpastatin, whereby we have optimized our protein production workflow at the cloning (with a new tag), expression (co-transformation with pT7pol26 and co-expression of BirA) and purification level (single-resin affinity purification). The introduction of an avi-tag at the N-terminus of hCSD1 allowed for the specific *in vivo* biotinylation in an *E. coli* expression system. We characterized the avi-hCSD1 at the

Table 1
Binding kinetics analysis parameters of the calpain-calpastatin interaction.

b-avi-hCSD1			b-hCSD1		
K_D (nM)	k_{on} ($M^{-1}.s^{-1}.10^4$)	k_{off} ($s^{-1}.10^{-3}$)	K_D (nM)	k_{on} ($M^{-1}.s^{-1}.10^4$)	k_{off} ($s^{-1}.10^{-3}$)
73.1 ± 0.3	11.6 ± 0.1	8.5 ± 0.1	134.4 ± 0.8	3.2 ± 0.1	4.4 ± 0.2

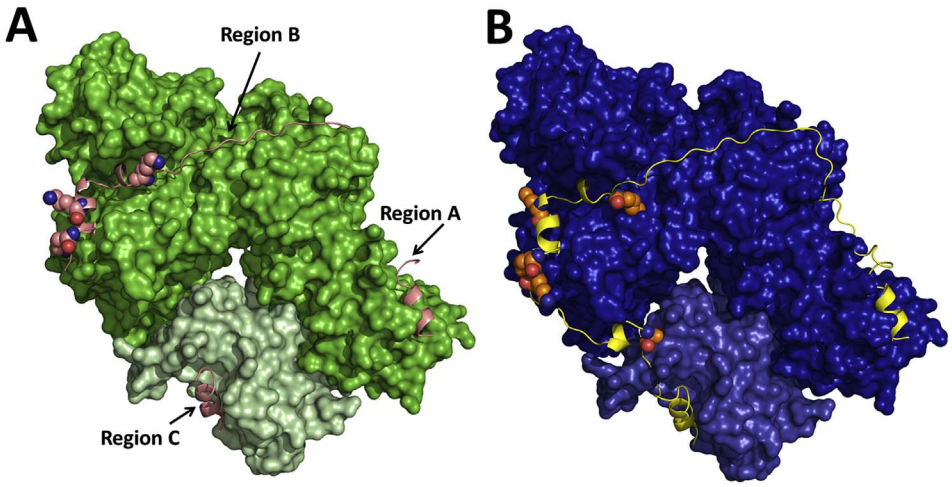

Fig. 5. Representation of the crystal structure (A) of the rat m-calpain structure in complex with the rat calpastatin domain (3DF0.pdb)) and the homology model (B) of the human m-calpain bound to hCSD1 (in cartoon representation). The heterodimeric m-calpain is depicted with the surface representation (rat in green, human in blue), the calpastatin domains are shown in the cartoon representation (pink or yellow) and the lysine residues in the visible calpastatin domains are highlighted in the sphere representation. The large subunits (corresponding to CAPN2) are represented in the dark color; the small subunits (corresponding to CAPNS1) are represented in the lighter color. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

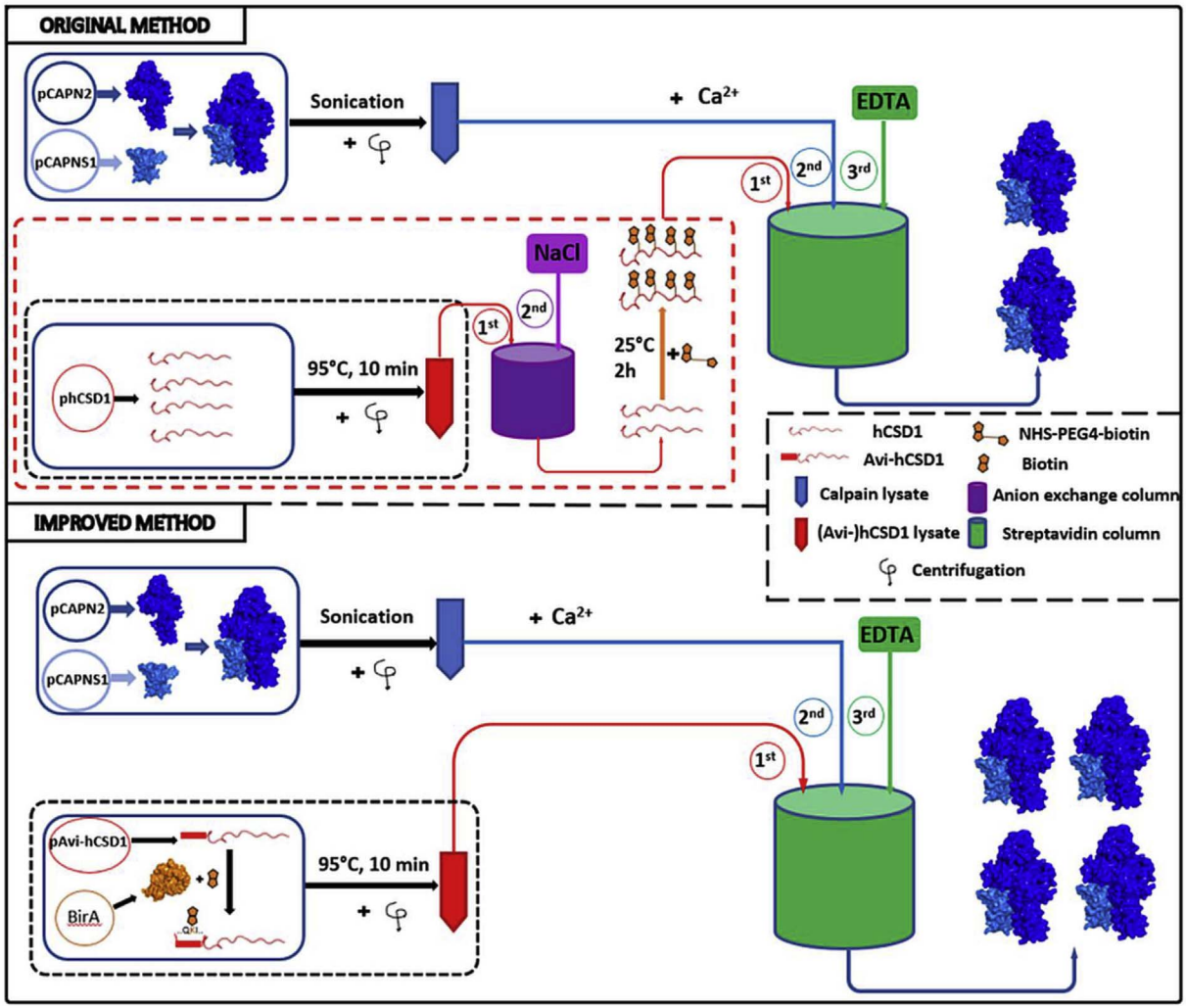

Fig. 6. Schematic comparison of the various steps and procedures needed for the affinity purification of human m-calpain based on biotinylated versions of hCSD1. Our original affinity purification methodology was based on the recombinant expression of hCSD1, followed by a heating step (95 °C for 10 min) and clearance of insoluble material via centrifugation [7]. This sample was then purified by anion exchange chromatography (eluted with NaCl) and subjected to an *in vitro* chemical biotinylation protocol using a commercially available kit. The resulting b-hCSD1 was then immobilized onto a streptavidin-based resin and used to capture recombinantly produced human m-calpain from the bacterial lysate (sonicated, centrifuged and supplemented with calcium). The complex between calpain and b-hCSD1 was dissociated by applying EDTA as a third step onto the column, leading to the elution of calpain. Our novel and improved method is based on the co-expression of avi-tagged hCSD1 and the biotin-ligase BirA that covalently attaches biotin to avi-hCSD1. The b-avi-hCSD1 is compatible with heat treatment for IDP purification and can be immobilized directly onto the streptavidin-based resin without the need for any additional manipulation. The subsequent workflow for calpain affinity purification is the same as the original method, but leads to an increased protein yield.

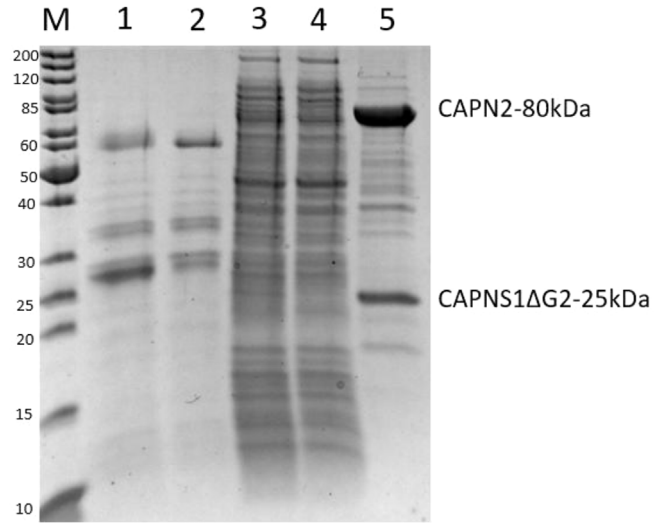

Fig. 7. Affinity purification with streptavidin-loaded resin of the C105A mutant of human m-calpain using b-avi-hCSD1-containing lysate after heat treatment. M: Marker proteins with their corresponding molecular weight (in kDa) indicated on the left; Lane 1: b-avi-hCSD1 lysate after heat treatment and centrifugation; Lane 2: flow-through when b-avi-hCSD1 lysate was loaded onto the streptavidin-loaded resin; Lane 3: bacterial lysate that contains human m-calpain; Lane 4: flow-through when calcium-activated calpain-containing lysate was loaded; Lane 5: when the column was eluted with EDTA, the elution fraction contained predominantly the large (CAPN2) and small (CAPNS1ΔG2) subunits of m-calpain.

residue level by combining NMR, BLI and MS analysis and found that (1) only the avi-tag is modified with biotin, (2) the presence of the avi-tag does not interfere with the intrinsically disordered nature of hCSD1, and (3) its binding potential to the calcium-activated calpain surface is improved. Hence, this approach enabled the direct purification/immobilization of b-avi-hCSD1 on streptavidin-based resin without the need of any additional prepurification step. With this new workflow to ultimately purify m-calpain based on its reversible binding of calcium and calpastatin, not only have we reduced considerably the purification time, but we also reduced the required energy, materials and consumables (and concomitantly the production cost per purified sample), which is important in these times of growing ecological awareness and sustainability (Fig. 6).

To the best of our knowledge this is the first time that an *in vivo* biotinylation protocol for an IDP (and its purification based on heat-treatment of resuspended cells) is reported. Since it is a cost-effective and a considerably simplified protocol that can be routinely used to purify any protein of interest through reversible binding onto the column, it seems particularly convenient to use with IDPs. Biotinylation might also be carried out *in vitro* with purified avi-tagged IDPs and purified BirA as this is reported to be easily produced at large scale [17]. Apart from the applicability of biotinylation for immobilization, it can be used as a detection method for various biological assays employing avidin/streptavidin-conjugated antibody or biotin-conjugated dye.

In conclusion, we described a successful and easily-applicable protein expression and affinity-capture strategy by combining commercially accessible plasmids and commercial *E. coli* strains. This new strategy not only facilitated the scale-up of calpain purification in particular, but also advocated for the general applicability of *in vivo* biotinylated IDPs in affinity purification strategies.

Acknowledgements

We thank Dr. Henri De Greve (VIB-VUB) for generously providing the plasmids pT7pol23 and pT7pol26. Kris Pauwels is the recipient of a FWO long-term postdoctoral fellowship (#1218713) and this work was partly supported by a pilot grant of Stichting Alzheimer Onderzoek (P#12017). Peter Tompa was supported by the Odysseus grant G.0029.12 from Research Foundation Flanders (FWO-Vlaanderen). The authors thank Ivo Vanoverstraeten for excellent technical assistance, Dr. Jesper S. Oeemig for critically reading the manuscript and Dr. Kenneth Verstraete (VIB-UGent) for useful discussions.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.pep.2018.01.002.

References

[1] Y. Ono, T.C. Saido, H. Sorimachi, Calpain research for drug discovery: challenges and potential, Nat. Rev. Drug Discov. 15 (2016) 854–876, http://dx.doi.org/10.1038/nrd.2016.212.
[2] S. Hata, M. Ueno, F. Kitamura, H. Sorimachi, Efficient expression and purification of recombinant human m-calpain using an Escherichia coli expression system at low temperature, J. Biochem 151 (2012) 417–422, http://dx.doi.org/10.1093/jb/mvs002.
[3] J.S. Elce, Expression of m-calpain in Escherichia coli, Methods Mol. Biol. 144 (2000) 47–54, http://dx.doi.org/10.1385/1-59259-050-0:47.
[4] T. Lian, L. Wang, Y. Liu, A new insight into the role of calpains in post-mortem meat tenderization in domestic animals: a review, Asian-Australasian, J. Anim. Sci. 26 (2013) 443–454, http://dx.doi.org/10.5713/ajas.2012.12365.
[5] S. Chakraborti, M.N. Alam, D. Paik, S. Shaikh, T. Chakraborti, Implication of calpain in health and diseases, Indian J. Biochem. Biophys. 49 (2012) 316–328.
[6] D.E. Goll, V.F. Thompson, H. Li, W.E.I. Wei, J. Cong, The calpain system, Physiol. Rev. 83 (2003) 731–801, http://dx.doi.org/10.1152/physrev.00029.2002.
[7] H.H. Nguyen, M. Varadi, P. Tompa, K. Pauwels, Affinity purification of human m-calpain through an intrinsically disordered inhibitor, calpastatin, PLoS One 12 (2017) e0174125, , http://dx.doi.org/10.1371/journal.pone.0174125.
[8] D. Beckett, E. Kovaleva, P.J. Schatz, A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation, Protein Sci. 8 (2008) 921–929, http://dx.doi.org/10.1110/ps.8.4.921.
[9] P.J. Schatz, Use of peptide libraries to map the substrate specificity of a peptide-modifying enzyme: a 13 residue consensus peptide specifies biotinylation in Escherichia coli, Nat. Biotechnol. 11 (1993) 1138–1143, http://dx.doi.org/10.1038/nbt1093-1138.
[10] J.E. Cronan Jr., Biotination of proteins in vivo, J. Biol. Chem. 265 (1990) 10327–10333.
[11] A.M. Livernois, D.J. Hnatchuk, E.E. Findlater, S.P. Graether, Obtaining highly purified intrinsically disordered protein by boiling lysis and single step ion exchange, Anal. Biochem. 392 (2009) 70–76, http://dx.doi.org/10.1016/j.ab.2009.05.023.
[12] N. Mertens, E. Remaut, W. Fiers, Versatile, multi-featured plasmids for high-level expression of heterologous genes in Escherichia coli: overproduction of human and murine cytokines, Gene 164 (1995) 9–15, http://dx.doi.org/10.1016/0378-1119(95)00505-Z.
[13] M. Biasini, S. Bienert, A. Waterhouse, K. Arnold, G. Studer, T. Schmidt, F. Kiefer, T.G. Cassarino, M. Bertoni, L. Bordoli, T. Schwede, SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information, Nucleic Acids Res. 42 (2014) 252–258, http://dx.doi.org/10.1093/nar/gku340.
[14] E. de Boer, P. Rodriguez, E. Bonte, J. Krijgsveld, E. Katsantoni, A. Heck, F. Grosveld, J. Strouboulis, Efficient biotinylation and single-step purification of tagged transcription factors in mammalian cells and transgenic mice, Proc. Natl. Acad. Sci. U. S. A 100 (2003) 7480–7485, http://dx.doi.org/10.1073/pnas.1332608100.
[15] M. Maki, K. Hitomi, Purification of recombinant calpastatin expressed in Escherichia coli, Methods Mol. Biol. 144 (2000) 85–94, http://dx.doi.org/10.1385/1-59259-050-0:85.
[16] M. El Khattabi, P. Van Gelder, W. Bitter, J. Tommassen, Role of the lipase-specific foldase of Burkholderia glumae as a steric chaperone, J. Biol. Chem. 275 (2000) 26885–26891, http://dx.doi.org/10.1074/jbc.M003258200.
[17] Y. Li, R. Sousa, Expression and purification of E. coli BirA biotin ligase for in vitro biotinylation, Protein Expr, Purif 82 (2012) 162–167, http://dx.doi.org/10.1016/j.pep.2011.12.008.
[18] E. Szollosi, E. Házy, C. Szász, P. Tompa, Large systematic errors compromise quantitation of intrinsically unstructured proteins, Anal. Biochem. 360 (2007) 321–323, http://dx.doi.org/10.1016/j.ab.2006.10.027.
[19] M. Guharoy, K. Pauwels, P. Tompa, SnapShot: intrinsic structural disorder, Cell 161 (2015) e1, , http://dx.doi.org/10.1016/j.cell.2015.05.024 1230–1230.
[20] I.C. Felli, R. Pierattelli, Intrinsically sidordered proteins studied by NMR spectroscopy, Adv. Exp. Med. Biol. 870 (2015) 361–362, http://dx.doi.org/10.1017/CBO9781107415324.004.
[21] R.a Hanna, R.L. Campbell, P.L. Davies, Calcium-bound structure of calpain and its mechanism of inhibition by calpastatin, Nature 456 (2008) 409–412, http://dx.doi.org/10.1038/nature07451.
[22] T. Moldoveanu, K. Gehring, D.R. Green, Concerted multi-pronged attack by calpastatin to occlude the catalytic cleft of heterodimeric calpains, Nature 456 (2008) 404–408, http://dx.doi.org/10.1038/nature07353.