Teila Pajazetovic, BSc

Recombinant bacterial milk proteins: The future of synthetic biomaterials?
Cross-linking caseins using active L-DOPA

MASTER'S THESIS

to achieve the university degree of

Diplom-Ingenieurin

Master's degree programme: Biotechnology

submitted to

Graz University of Technology

Supervisor

Univ.-Prof. Dipl.-Ing. Dr.techn. Helmut Schwab

Institute of Molecular Biotechnology
Graz University of Technology

Dipl.-Ing. Dr.techn. Birgit Wiltschi
Austrian Centre of Industrial Biotechnology

Graz, May 2015

AFFIDAVIT

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present master‘s thesis dissertation.

Date

Signature

Table of Contents

Acknowledgments

I would like to thank Univ.-Prof. Dipl.-Ing. Dr.techn. Helmut Schwab for being the main examiner of this work, without whose permission it would not have been possible to conduct my research within the Junior Group Synthetic Biology.

I would also like to extend my gratitude to my supervisor, Dipl.-Ing. Dr.techn. Birgit Wiltschi, for her tremendous guidance and support. Her expertise in non-canonical amino acids was crucial to further my work. I am also thankful for her meticulous nature and constructive scrutiny of my thesis, which helped me prepare my document in a professional manner.

Special and kind thanks go to Patrik Fladischer, MSc, who was with me every step of the way, even with his educational leave to Germany. I am grateful for the countless hours of discussions we held (face-to-face and on Skype) whilst planning experiments and analyzing their results. The dedication he has towards the students he advises is remarkable and inspirational.

I also want to thank the members of the Junior Group of Synthetic Biology for their support in discussing lab issues or troubleshooting. It has been a pleasure working with everyone. In particular, I want to thank Carinna Sommer, whose idea it was to work with casein proteins and Alexandra Weingartner for her ground-laying initial experiments with L-DOPA.

Thank you to the in-house technicians, Karin Reicher and Inge Eiteljörg, who spent many hours teaching and guiding me through all of my bioprocessing and size exclusion experiments.

Furthermore, I would like to thank the Austrian Centre of Industrial Biotechnology and FFG for funding this project.

A special thanks goes to Florian Borsodi, for being there throughout my entire master studies and for his encouraging support, especially on trying days.

Last, but certainly not least, I would like to thank my parents, who made it possible for me to study and reside abroad. Both inspired, comforted, and pushed me to finish my studies, even in times of weakness. They have had to endure the most. Words cannot even describe how grateful I am for their never-ceasing love and support. They are my rock.

Graz, May 2015

Teila Pajazetovic, BSc

1 Abstract

Non-degradable plastic waste is accumulating at an alarming rate. It would, therefore, be of great environmental interest if we could produce biodegradable plastic-like materials without draining our natural resources. Our ambition is to fuse synthetic biotechnology and materials science in order to achieve our goal in generating a novel and degradable biomaterial that is based on Galalith plastics. Galalith was made by treating milk proteins with formaldehyde to cross-link them and form a stable matrix, and was used mostly as an ivory and horn substitute. Its discontinuation in the 1940s was mainly due to economic issues and the carcinogenic effects of formaldehyde. However, the concept of this casein-derived plastic is still interesting on account of its elastic properties. We drew our idea from that of Galalith, but intended to omit the toxic formaldehyde and replace it by synthetic milk protein containing a naturally occurring, reactive amino acid for cross-linking. Therefore the main purpose of this work was to cross-link synthetic milk proteins (caseins) by globally incorporating a non-canonical amino acid, L-DOPA, in pursuance of synthesizing a new plastic substitute. Successfully, we used *E. coli* as a production host to express synthetic L-DOPA-containing caseins, upon which the bidentate ligands of L-DOPA were susceptible to nucleophilic attack by oxidative and coordinating reagents. We further managed to cross-link these synthetic caseins with such agents in the anticipation of gaining a new, biodegradable, Galalith-like, biomaterial in the future without the use of poisonous chemicals such as formaldehyde.

2 Kurzfassung

Die Ansammlung von nicht abbaubaren Kunststoffmüll ist an einem kritischen Punkt angelangt. Deshalb wäre es für die Umwelt von großem Nutzen, biologisch abbaubare Kunststoffe zu produzieren, welche die natürlichen Rohstoffressourcen nicht belasten. Unser Ziel ist es, durch die Integration von synthetischer Biologie in die Materialwissenschaft einen neuartigen, biologisch abbaubaren Kunststoff zu entwickeln, welcher auf Galalithkunststoff basiert. Galalith wurde durch die Behandlung von Milchproteinen (Kaseine) mit Formaldehyd hergestellt und als Elfenbein - und Hornersatz verwendet. Das Formaldehyd wirkt dabei als Quervernetzer, der die Milchproteine in Form einer stabilen Matrix fixiert. In den 40er Jahren wurde die Produktion von Galalith aufgrund der Wirtschaftskrise und kanzerogenen Effekte von Formaldehyd eingestellt. Wegen der elastischen Eigenschaften ist das Konzept eines auf Kaseinen basierenden Kunststoffes weiterhin von Interesse. Wir haben das Konzept eines Galalithkunststoffes weiterentwickelt, indem wir das toxische Formaldehyd durch synthetische Kaseinproteine ersetzen, die eine natürlich vorkommende, reaktive Aminosäure für die Quervernetzung enthalten. Der Fokus dieser Arbeit liegt auf der Vernetzung von Kaseinen durch den Einbau von nicht kanonischen Aminosäuren (L-DOPA), um so einen neuartigen Kunststoff herzustellen. Für die erfolgreiche Expression von Kaseinen mit L-DOPA wurde *E. coli* als Wirt verwendet. Die bidentaten Liganden des L-DOPA ermöglichten den nuklearen Angriff von oxidativen und koordinativen Reagenzien. Es ist uns gelungen, die synthetischen Kaseine mit derartigen Reagenzien miteinander zu vernetzen, um auf diese Weise ein neues, biologisch abbaubares, Galalithähnliches Biomaterial herzustellen, ohne den Einsatz von giftigen Chemikalien wie Formaldehyd.

3 Introduction

Environmentally-sensitive substitutes, such as biodegradable plastics, have become essential innovations in limiting global plastic pollution. Human beings produce ~34 million tons of polyethylene terephthalate (PET) plastic waste per year, of which only 7% is recycled (Pathak, Sneha, and Mathew 2014). The remainder of the plastic debris is distributed across all oceans due to its buoyancy and persistence, and is thus contaminating marine and terrestrial habitats (Eriksen et al. 2014). These synthetic polymers are regarded as hazardous waste by researchers due to their sorption of toxicants, such as polycyclic aromatic hydrocarbons. In turn, these are ingested by more than 260 species, triggering detrimental physical effects (Teuten et al. 2007). As plastic pollution is becoming a global issue, bioplastics may provide relief, if not solve, the plastic pollution problem altogether.

3.1 Bioplastics

Since the 1940s, PET plastics have dominated foremost the packaging industry due to their lightweight, strong, durable, and corrosion-resistant properties. They are valued because of their high industrial versatility and low cost. However, they constitute the largest petroleum use, followed only by energy (Gervet 2007). Considering that fossil fuels are finite, the utilization of sustainable bioplastics and bio-based polymers is gradually progressing to substitute and perhaps even discontinue the use of PET plastics. The bioplastic industry is still in its infancy stages, and with the current growth rate, it is expected that this industry will account for ~1% of biopolymers by the year 2015 (Babu, O'Connor, and Seeram 2013).

The production of bio-based materials relies on the use of renewable, food-based resources (Mekonnen et al. 2013). Nevertheless, there has been a desire to move away from food-based resources in recent years. This is a result of new methodologies that have emerged in the fields of biotechnology and synthetic biology, and many other life sciences that make use of bacterial systems for the production of bio-based polymers (Babu et al. 2013).

Synthetic biology, in particular, is a multidisciplinary approach when it comes to researching and producing biological compounds that are not found in nature. It makes use of biological and chemical engineering, as well as the fields in the fundamental sciences such as chemistry, biology, biophysics, and mathematics (Voigt 2012). Certain tasks of a synthetic biologist include combining multiple genes and/or newly constructed "biological parts," or the use of unnatural molecules to establish new pathways and functions (König et al. 2013).

3.2 Unnatural amino acids and their incorporation into proteins

There are 20 natural, or canonical, amino acid (cAA) building blocks that are encoded by the genetic code in eukaryotic as well as prokaryotic proteins. In view of the limited number of cAAs, there is not much chemical diversity that proteins can exhibit (Voloshchuk and Montclare 2010). Therefore, native proteins do not usually offer the desired chemical properties that can be exploited for various biotechnological applications. In nature, these limitations are overcome by the introduction of post-translational modifications. Nevertheless, they are usually avoided in heterologous protein modification due to difficulties in controlling them (Seo and Lee 2004). However, unnatural amino

acids (or rather non-canonical amino acids, ncAA) are not encoded by the genetic code, but are plenteous in nature and can be implemented to spark different side-chain chemistries within naturally-occurring proteins (Zheng and Kwon 2012). Presently, there are more than 700 known ncAAs which are produced as secondary metabolites, particularly in plants and fungi (Hunt 1985).

The method replacing any amino acid in a protein with an ncAA analog, that can have a vast structural diversity, is interesting to the field of synthetic biology. Incorporation of ncAAs into target proteins can be achieved with the native protein translation machinery, assuming the implemented ncAA is similar in structure to the natural amino acid. The ncAA charges native aminoacyl-tRNA-synthetases prior the translation of the mRNA transcript. The charged tRNA is then brought to the ribosome with the help of elongation factors, upon which protein translation occurs (Zheng and Kwon 2012).

There are established methods for incorporating many ncAAs into target proteins, either site- or residue-specifically. The site-specific method of introducing ncAAs *in vitro* into proteins makes use of chemically misacylated tRNAs and cell-free translation systems, whereas the *in vivo* incorporation involves an *E. coli* strain that has a heterologous orthogonal pair of suppressor tRNA/aminoacyl-tRNA synthetase specific for an ncAA. For the incorporation site, four-base codons and a stop codon (mostly amber stop codon) are used (Zheng and Kwon 2012).

Within this work, we focus on residue-specific incorporation of unnatural analogs. This method involves the use of an auxotrophic bacterial strain that is starved for the cAA and supplemented with its ncAA analog (Wiltschi 2012). Therefore, an ncAA is incorporated by sacrificing one cAA from the pool of amino acid building blocks. Prior to implementing this method, also called the supplementation-based incorporation (SPI) method, it is imperative to perform yeast extract titration experiments to determine how much yeast extract (in g/L) is needed in the cultivation medium. The particular auxotrophic strain ought to go into growth arrest at a predefined cell density due to depletion of the essential amino acid. When cells enter this depletion phase, the medium is then supplemented with the ncAA as its cAA substitute. This protocol of incorporating ncAAs into proteins is beneficial in that the target protein's physical properties are changed, and new chemical modifications are enabled at multiple sites. Cross-linking chemistry via side chains of the incorporated ncAAs is such an example (Link, Mock, and Tirrell 2003).

3.2.1 Covalent cross-linking of the L-DOPA ncAA

L-3,4-dihydroxyphenylalanine (L-DOPA) is a tyrosine analog. It is not prescribed for protein translation by the genetic code and is therefore an ncAA (see Figure 1). With its bidentate enediol structure, L-DOPA is an intriguing candidate for protein cross-linking applications. Proteins containing L-DOPA can be oxidized via their hydroxyl groups by reactive agents such as sodium periodate ($NaIO_4$), thus forming o-quinone intermediates that can be attacked by a neighboring nucleophile. These intermediates can form adducts with lysines, cysteines, or histidines. Generally, these reactions result in covalently cross-linked nonspecific products (Umanah et al. 2009).

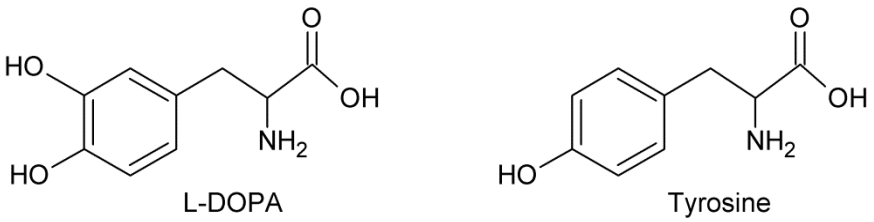

Figure 1. Chemical structures of the ncAA L-DOPA and the cAA tyrosine.

Hence, L-DOPA is an appealing ncAA that we will be using within this work, through which we will attempt to produce cross-linked proteins.

3.2.2 Coordinating L-DOPA to metal ions

Metal complexation is a functional method for assembling and reinforcing biopolymeric materials, as is the case in the mussel proteins. The byssal threads of marine mussels are protected from deformation in turbulent and violent bodies of water by an outer cuticle that possesses a five-fold hardness of the thread core (Harrington et al. 2010). Within these threads, L-DOPA and inorganic Fe^{3+} ions are co-localized, forming a polymeric scaffold of catecholato-iron chelate complexes. Holten-Andersen *et. al.* suggest that iron metals play a key role in mussel foot adhesion (Yang, Cohen Stuart, and Kamperman 2014).

In order to form an iron chelate coordinated complex with L-DOPA, a catechol (L-DOPA) donates a non-bonding electron pair to the iron. Three different degrees of coordinated structures can be formed, including the mono-, bis-, and tris-L-DOPA-Fe^{3+} complexes. Which complex will be formed is dependent on pH via the deprotonation of the hydroxyl groups of L-DOPA. A pH below 5.6 is the source of mono-L-DOPA-Fe^{3+} complexes, whereas the bis-complexes are formed at a pH between 5.6 and 9.1. All pH conditions above 9.1 are favorable for the formation of tris-complexes (Yang et al. 2014).

Nevertheless, the iron concentration is also a crucial factor in coordination-complex formation. Lower concentrations of iron yield the tris-complex, while higher iron concentrations cause mono-complex formation (Zeng et al. 2010).

3.3 Galalith-like bioplastics

A type of bioplastics made from milk by coagulating and precipitating milk proteins (caseins) followed by the immersion thereof in formaldehyde, is called Galalith. This method of making plastics was first developed in 1904 (Siegfeld 1904) and was used until its discontinuation in 1939, due to the scarcity of milk during the Second World War (Krätz 2004). Its production was rather simple. Caseins were extracted from milk by precipitation with an acid, such as sulfuric, hydrochloric, lactic, or sulfurous acid. They were then washed to get rid of traces of fat and were subsequently dried. Upon being pressed and molded, the caseins were immersed in the aforementioned formaldehyde to form a tough, white product. The hardened product was further colored by stirring in dyes or soot (Anonymous 1913). The advantage as well as the disadvantage of Galalith production was the use of milk. Milk is renewable; therefore its use was justified. However, after the Second World War, intercontinental milk import and export ceased, making it a scarcity, and therefore causing Galalith production come to a stand-still.

PET plastics gained popularity in the 1940s due to their high versatility and economical production. However, their production makes use of crude oil as a raw material, a non-renewable resource on the planet, which contributes substantially to thermal pollution. Heat emissions constitute a whopping 55% of total global warming. It has been reported that the net heat generation from plastic production between 1939 and 2000 was 0.41014 kWh. This value constitutes 0.5% of total global warming, inasmuch being almost of the same magnitude as gas flaring. Strikingly, heat generation from plastic production is more than that of coal fires (Gervet 2007).

However, because of the drastic consequences of PET plastics use on the environment, an eco-friendly plastic substitute is of dire need. Hence, it would be beneficial to resurrect Galalith-like bioplastics that are additionally biodegradable, and made employing biotechnological methods instead of fossil fuels. The production of recombinant caseins utilizing bacterial hosts is possible. In this manner food-based resources, such as milk, would be avoided. Further implementation of synthetic biology methods to cross-link caseins via incorporated non-canonical amino acids could yield a bioplastic-like polymer. By doing so, the use of cancerogenic formaldehyde could altogether be bypassed.

3.4 Thesis objectives

With new emerging and constantly developing technologies, we can explore former innovations and apply them differently. We are, therefore, venturing to produce Galalith-like structures from *E. coli* using synthetic biology. Previously, we have investigated *Bos taurus* α-S1-, and β-caseins, by cloning these genes into *E. coli* using pET26b(+) vectors. We were able to express and purify these heterologous proteins, and characterize them using size exclusion chromatography. After these experiments, we postulated that we could further incorporate an active ncAA (L-DOPA) into the caseins, and specifically cross-link them with oxidative and coordinating reagents. In order to obtain high amounts of proteins, we performed high cell density fermentations. We were successful in fermenting *E. coli* cultures expressing synthetic parent caseins (casein[Tyr]) as well as synthetic L-DOPA-containing casein (casein[DOPA]) proteins. Therefore, we were able to implement the supplementation-based incorporation method for incorporating the ncAA L-DOPA into caseins on a bioreactor scale, which proved to be a better method in producing high amounts of protein for further cross-linking studies. We made great strides in our search for a new biomaterial, by being able to observe cross-linking effects of casein[DOPA] variants.

4 Materials and Methods

4.1 Cloning, expression and cultivation of synthetic caseins

The early stages of this thesis began in an earlier laboratory project, where the initial cloning of the α-S1-casein [*E. coli*$_{opt,}$ $N_{truncated}$, C_{term}-H_6] and the β-casein [*E. coli*$_{opt,}$ $N_{truncated}$, C_{term}-H_6] proteins (with genbank accession numbers ACG63494.1 and AAA30431.1, respectively) was achieved in the *E. coli* BL21(DE3)Gold strain (Invitrogen, Carlsbad, CA) (Pajazetovic 2014). The sequences of the genes that we used in this study are listed in Table S1. Casein gBlocks® were successfully subcloned into pET26b(+) vectors via Gibson assembly (Gibson et al. 2009) using primers listed in Table S2. For visual purposes, a vector map of the pET26b(+) plasmid with one of the target genes can be seen in Figure S1. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and size exclusion chromatography (SEC) confirmed successful target protein expression and the purification thereof (see Figure S2).

4.2 SDS-PAGE analysis

Sodium dodecyl sulfate polyacryl gel electrophoresis (SDS-PAGE) was performed as described by Laemmli (Laemmli 1970). All SDS gels, with the exception of NuPAGE® 4-12% Bis-Tris precast gels (NuPAGE®, Life Technologies, Carlsbad, CA) were self-made. The stacking gel and the separating gel comprised of 5% and 14% polyacrylamide, respectively.

To follow protein expression, 0.25 D_{600} samples were harvested from cell cultures at 13,000 rpm, at room temperature for 2 minutes. The cell pellet was resuspended in 30 µL ddH_2O and 10 µL 4 X SDS loading dye (NuPAGE®). Prior to performing an SDS-PAGE, the SDS samples were heated at 98°C for 10 min. 10 µL of the samples were loaded onto the gel. In order to estimate the size of the protein bands, a commercial protein standard was loaded onto each gel prior to electrophoresis (PageRuler prestained protein ladder, Thermo Scientific, Waltham, MA). The gels were run at 100 V for 20 min, followed by 180 V for 40 min. Staining of the gels occurred with Coomassie Brilliant Blue G250 (Roth, Kalsruhe, Germany), after which the gels were scanned for figure editing.

To analyze the protein purification samples, 30 µL of each purification fraction were collected, to which 10 µL of 4 X SDS loading dye (NuPAGE®) were added. Unless otherwise specified, 1 µg of purified protein samples were loaded onto the gels. An SDS-PAGE was performed under the same conditions as stated above.

4.3 Protein concentration determination

The protein concentrations within this work were determined with a Bradford protein assay (BioRad, Vienna, Austria). The manufacturer's instruction manual was followed without any deviations.

4.4 Shake flask cultivation of *E. coli* cells expressing synthetic casein variants

Cultivation of *E. coli* cells expressing synthetic caseins with incorporated L-3,4-dihydroxyphenylalanine (L-DOPA) was initially tested in a previously conducted laboratory project (Weingartner 2014). The tyrosine auxotrophic *E. coli* BL21(DE3)Gold strain, with the B F- *dcm*+ Hte *ompT hsdS(rB- mB-) gal* λ (DE3) *endA* Tet^r Δ*tyrA*::0 genotype, was used for the expression of the

synthetic caseins, with the kind permission of Nikolaus Anderhuber. This strain was then annotated as the BWEC48(TyrΔ) strain.

It is known that L-DOPA oxidizes to a dark color in solutions (Ayyadurai et al. 2011). Therefore, Alexandra Weingartner tested the effect that 10 mM ascorbic acid has on L-DOPA as an antioxidant (Figure S3) in medium and culture (Weingartner 2014). She reported that ascorbic acid inhibits the oxidation of L-DOPA. After also performing yeast extract titration experiments (as explained in section 3.2) with and without ascorbic acid, she reported that the optimal yeast extract concentration for the *E. coli* BWEC48(TyrΔ) strain was 1.75 g/L (see Figure S4) and that 10 mM ascorbic acid does not influence its growth, as can be seen in Figure S5.

Overnight cultures (ONC) of *E. coli* BWEC48(TyrΔ){pET26b(+)-α-casein-H6} and BWEC48(TyrΔ){pET26b(+)-β-casein-H6} were prepared in LB medium, comprised of 10 g/L tryptone, 5 g/L yeast extract, and 5 g/L NaCl (Lennox, Karlsruhe, Germany), with 50 mg/mL kanamycin. Main cultures, comprising of a 500 mL M9-based medium including 1.75 g/L yeast extract (see Table S3 and Table S4 for M9 medium composition) and 50 mg/mL kanamycin, were then inoculated in baffled shake flasks with the corresponding ONCs to a starting D_{600} of 0.1, and cultivated at 37 °C and 120 rpm until the D_{600} reached ~3 (depletion phase). For subsequent expression of the target proteins, the temperature was lowered to 25 °C. The expression of the synthetic caseins was then induced with 0.5 mM IPTG (isopropyl β-D-thiogalactoside, Roth). Incorporation of L-DOPA (ABCR GmbH, Karlsruhe, Germany) into the caseins was then carried out using the SPI (supplementation-based incorporation) method, as described in Wiltschi 2012. Both, α-casein[DOPA] as well as β-casein[DOPA], cultures were supplemented with 1 mM of L-DOPA following their induction and again after four hours of expression. To follow the casein expression, samples for SDS-PAGE were taken before, after four hours, and after overnight induction. The whole-cell SDS samples were normalized to a D_{600} of 0.25 by harvesting the cell cultures at 13,000 rpm and room temperature, for 2 minutes. The supernatant was discarded, and the cell pellet was resuspended in 30 µL H_20 and 10 µL 4 X SDS loading dye (NuPAGE®). Afterward, the cell cultures were harvested by centrifugation (10,000 rpm, 45 min, 4 °C) using an Avanti J-20 XP centrifuge (Beckmann Coulter Inc., Brea, CA) following the overnight incubation for a further Ni-NTA (Qiagen, Hilden, Germany) gravity flow purification (Qiagen 2003). The cell pellet was resuspended in 5 mL of lysis buffer per gram cell wet weight (50 mM $NaH_2PO_4 \cdot H_20$, 300 mM NaCl, 10 mM imidazole, pH 8.0), containing 1 mg/mL lysozyme and 4 µg/mL DNAse, and pre-treated for 1 h on ice. One column volume (CV) of nickel resin was packed into a column and was calibrated with 20 CVs of lysis buffer. The wash buffer used for the purification was composed of 50 mM $NaH_2PO_4 \cdot H_20$, 300 mM NaCl, and 20 mM imidazole, and its pH was adjusted to 8.0. After the crude lysate was passed through the resin, the resin was washed with 50 CVs of the aforementioned wash buffer. The purified protein was then eluted with 4 CVs of elution buffer (same composition as lysis buffer only containing 250 mM imidazole).

The proteins were re-buffered from the imidazole elution buffer to 10 mM phosphate-buffered saline (PBS, 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4) using a HisTrap 26/10 desalting column (GE Healthcare, Buckinghamshire, UK). The PBS buffer also included 3.75 M urea and 1 mM DTT (dithiothreitol).

The purified proteins were stored at 4 °C.

4.5 Specific detection of L-DOPA

As reported in different papers, proteins containing quinones or o-catechols can be detected by redox-cycling staining using nitroblue tetrazolium (NBT) (Ayyadurai et al. 2011, Pazs et al. 1991, Rodgers et al. 2004). To detect if L-DOPA was indeed incorporated into the synthetic caseins, an SDS-PAGE of the purified proteins was performed. The proteins on the gel were then electroblotted onto a polyvinylidene difluoride membrane (PDVF, Immun-Blot® PDVF Membrane; Bio-Rad, Vienna, Austria) following the method described in the work of Haid and Suissa (1983) Briefly, we assembled a blotting sandwich that included fiber pads, a Whatman filter paper, the SDS gel, the PDVF membrane, Whatman filter paper (GE Healthcare) and fiber pads, in that particular order. The fiber pads were previously soaked in transfer buffer (192 mM glycine, 0.1% SDS, 25 mM tris, 20% v/v methanol). The gel was electroblotted at 160 mA for 1 h. The membrane was then washed twice with 1 X TBS buffer (0.05 M Tris, 0.15 M NaCl, pH was adjusted to 8.0 with HCl), followed by a three-hour incubation with the NBT staining solution (2 M sodium glycinate, 0.24 mM NBT, pH 10) in the dark.

4.6 High cell density fermentation of caseins

The DASGIP® Parallel Bioreactor System (DASGIP Information and Process Technology GmbH, Jülich, Germany) was used for all fermenting processes within this project.

These bioreactors were equipped with an oxygen and a pH electrode for on-line process monitoring. A pre-set pH of 7.0 was automatically maintained for all fermentations with 10% (v/v) H_3PO_4 and 12.5% (v/v) NH_3. The partial pressure of oxygen (pO_2) was kept to the minimum dissolved oxygen (30%), which was regulated by the impeller velocity. The stirring was done with a flat, six-blade disc turbine for all fermentations.

4.6.1 Fermentation of casein[Tyr] parent proteins

An up-scaled expression of the α-casein[Tyr] as well as β-casein[Tyr] parent proteins was done in 400 mL M9-based medium (medium components are listed in Table S3 and Table S5). Two bioreactors were assigned per casein to total four fermented cultures. Each fermentation medium was inoculated with the corresponding ONC, *E. coli* BWEC48(TyrΔ){pET26b(+)-α-casein-H_6} and BWEC48(TyrΔ){pET26b(+)-β-casein-H_6}, to a starting D600 of 0.1, and cultivated at 37 °C. After the cells reached a D_{600} of 0.8, the bioreactors, and therefore the cultures also, were cooled to 30 °C and the protein expression was induced with 0.5 mM IPTG. Only one casein culture of each type received a glucose substrate feed after induction in order to observe the effect of the feed on the protein expression. The substrate feed comprised of D-glucose monohydrate (55 mM), 2 mM of each of the 20 canonical amino acids (cAA), 50 µg/mL thiamine, 10 mg/mL biotin, and sterile water filled to a total volume of 310 mL. The feed rate was set to 10 mL/h. The expression was performed for 35 hours, with a pO_2 minimum of 30% at pH 7.0.

SDS samples for a subsequent SDS-PAGE were taken before, after four hours, and after overnight induction in order to follow protein expression. These whole-cell samples were normalized to a D_{600} of 0.25 as described above.

The cell cultures were then harvested by centrifugation using an Avanti J-20 XP centrifuge (Beckmann Coulter Inc.) at 10,000 rpm and 4° C, for 45 min after the overnight fermentation. The cell pellets were resuspended in 5 mL lysis buffer (50 mM $NaH_2PO_4 \cdot H_20$, 300 mM NaCl, 10 mM imidazole, pH 8.0), containing 1 mg/mL lysozyme and 4 µg/mL DNAse, per gram CWW and pre-treated for 1 h on ice. Cultured cells were further lysed with a sonifier (Branson, Danbury, CT), while constantly cooled, using the following parameters: duty cycle at 70%, output control micro tip limit at 8. Sonication was performed for 5 minutes, after which the crude lysate cell debris was removed by high speed centrifugation (Avanti J-20 XP centrifuge, 40 min, 20,000 x g, 4 °C). The cleared lysates were used for a subsequent purification with gravity flow Ni-NTA affinity as described in the previous section (4.4).

The proteins were also re-buffered from the imidazole elution buffer to a 10 mM phosphate-buffered saline (PBS) buffer using a PD10 Desalting Column (GE Healthcare). The PBS buffer included 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4, 3.75 M urea, and 1 mM DTT.

The purified proteins were stored at 4 °C.

4.6.2 Fermentation of casein[DOPA] variants with non-canonical amino acid supplementation

A fermentation of the α-casein[DOPA] and the β-casein[DOPA] was partaken in M9-based medium (see medium components in Table S5). The applied conditions for these fermentations were the same as for those of the casein[Tyr] parent proteins. Four bioreactors (two per casein) were used to cultivate the *E. coli* BWEC48(TyrΔ){pET26b(+)-α-casein-H_6} and BWEC48(TyrΔ){pET26b(+)-β-casein-H_6} cells. The fermentation media were inoculated with previously prepared ONCs to a starting D_{600} of 0.1. The cells were grown at 37 °C until they reached their depletion phase at D_{600} of ~21. This value is based on Alexandra Weingartner's yeast titration for the BWEC48 strain (Weingartner 2014). Yeast extract (4.9 g) was added to the M9 medium in order for the cells to start depleting at the desired attenuance. The point of depletion was reached after the D_{600} stagnated at ~21 for 30 minutes. The cultures in the bioreactors were then cooled to 30 °C, after which protein expression was induced with 0.5 mM IPTG. All cultures were supplied with the same substrate feed as the casein[Tyr] parent proteins. However, the feed substrate deviates from the above mentioned one. Instead of using an amino acid stock comprising of all canonical amino acids, tyrosine was omitted. Alternatively, the tyrosine analog, L-DOPA, was supplemented (10 mM). In order to keep the L-DOPA from oxidizing in the feed, as well as in the fermentation medium, 10 mM ascorbic acid were also added. The feed rate of the supplied substrate solution was set to 10 mL/h.

The expression was performed for 35 hours, with a pO_2 minimum of 30% at pH 7.0.

Also here, SDS samples were taken for a successive SDS-PAGE before induction, after 2, 4, 8, 17 and 24 hours of induction. The whole-cell samples were normalized to a D_{600} of 0.25, as described in the above 4.2 section. The cell cultures were then harvested by centrifugation (10,000 rpm, at 4° C, for 45 min) after the overnight fermentation for subsequent purification.

Fresh casein containing cell pellets were resuspended in 5 mL lysis buffer (contents same as stated in section above, only with 6 m urea) per gram CWW and pre-treated for 1 h on ice. The casein[DOPA]-containing cells were lysed with a sonifier (Branson), after which the crude lysate cell debris was

removed by high speed centrifugation (40 min, 20000 x g, 4 °C). The cleared lysates (supernatant) were purified with the ÄKTA Pure system (GE Healthcare) using a 5 mL HiTrap column (GE Healthcare). Only the lysis and elution buffers (7.5 CVs, same contents as lysis buffer only 300 mM imidazole, 3.75 M urea and 1 mM DTT) were used in the ÄKTA Pure system, where during the 7.5 column volumes both buffers were mixed to a steadily increasing gradient (from 1 to 100%) of the elution buffer.

The proteins had to be re-buffered from the imidazole elution buffer to 10 mM phosphate-buffered saline (PBS) using a HisTrap 26/10 desalting column (GE Healthcare). The PBS buffer also included 3.75 M urea and 1 mM DTT.

The purified proteins were stored at 4 °C.

4.7 Comparison of caseins produced in shake flasks to fermented caseins

To compare the casein protein expression between casein[Tyr] parent and casein[DOPA] protein variants, 1 g cell pellets of each protein (produced in shake flasks and bioreactors) were purified in the ÄKTA Pure system (GE Healthcare) after disrupting the cells with a sonifier (Branson). Caseins were purified with purification buffers (as described in the 4.4 section) in the presence and absence of urea and DTT. A subsequent SDS-PAGE and size exclusion chromatography (SEC) was performed to scrutinize the purification fractions of each casein that was purified under different conditions.

4.8 Oxidative bioconjugation of casein[DOPA] variants

4.8.1 Casein[DOPA] variants from shake flasks

A microdialysis technique was applied to cross-link the α-casein[DOPA] obtained from a shake flask cultivation in the presence of sodium periodate ($NaIO_4$). This experiment was performed with the α-casein[DOPA] proteins that were purified in the presence of ascorbic acid and the absence thereof. 100 µL of the purified casein[DOPA] variant were pipetted into an Eppendorf tube, which had a hole perforated in the lid that was covered by a dialysis membrane (SnakeSkin® Dialysis Tubing, Thermo Scientific). The tube was placed head-down in a beaker containing 2 L of 10 mM PBS buffer and different concentrations of $NaIO_4$ (1 mM, or 5 mM, pH ~6). This set-up was incubated at room temperature under stirring for 4 hours. The same was done for an overnight microdialysis sample. The α-casein[Tyr] parent protein, treated under the same conditions, served as a negative control. An SDS-PAGE analysis was performed of these samples, with a subsequent redox-cycling stain.

Cross-linking trials also included adding 1 mM $NaIO_4$ directly to the protein solutions and incubating the mixture at room temperature under vigorous shaking (750 rpm) for 4 hours, as well as overnight. To quench the reaction and inactivate the reactive $NaIO_4$, 100 mM DTT was added. A size exclusion chromatography (SEC, described in section 4.9) was performed in order to determine the molecular weights of the $NaIO_4$-treated casein samples.

4.8.2 Casein[DOPA] variants produced in bioreactors

Fermented casein[DOPA] variants underwent oxidative cross-linking with the direct addition of 1, 2, or 5 mM $NaIO_4$ to the protein solution. A protein concentration of 10 mg/mL was used for these

trials. After treating the caseins with $NaIO_4$, an SDS-PAGE was performed along with SEC. Casein[Tyr] proteins were treated the same and served as negative controls.

4.9 Size exclusion chromatography

Size exclusion chromatography (SEC) was performed on the ÄKTA Pure system (GE Healthcare), using a Superdex 200 FPLC column (Sigma-Aldrich, St. Louis, MO), with all cross-linked casein[DOPA] variants and the casein[Tyr] parent proteins as negative controls. All samples were filtered before loading onto the column with 0.2 µg polyethersulfone membrane filters (Sartorius, Göttingen, Germany). The elution pattern of a gel filtration protein standard (Bio-Rad, Vienna, Austria) was used to plot a calibration curve. The proteins within the gel filtration standard included bovine thyroglobulin (670 kD), bovine γ-globulin (158 kD), chicken ovalbumin (44 kD), horse myoglobin (17 kD), and vitamin B_{12} (1.35 kD). Chromatograms of all samples were evaluated and the molecular weights of the eluted fractions (peaks) were calculated according to the standard calibration curve.

Initial SEC of the of α-casein[Tyr] and of α-casein[DOPA] were executed with the down-flow parameter, meaning the flow of the proteins and all buffers entered at the top of the Superdex column, and exited at the bottom. Due to difficulties with the delta column pressure in down-flow operation, all successive SEC runs were executed with the up-flow setting. With the up-flow parameter, the protein solution and all buffers were pumped through the column while entering at the bottom and exiting at the top of the column. To have a representative calibration curve of the protein standard, the commercial standard was also run in the up-flow mode of operation, and a calibration curve was plotted according to the elution pattern of the protein standard.

4.10 Mass analysis

In-gel samples of β-casein[DOPA] were provided for mass analysis to the ACIB Core Facility Metabolomics (Medical University of Graz, Graz, Austria). The samples were analyzed using LC-MS/MS after a proteolytic digest with trypsin, chymotrypsin, and pepsin.

4.11 Biomimetic coordination of casein[DOPA] with iron

Fermented casein proteins were cross-linked by using ferric chloride, $FeCl_3$. Both, casein[Tyr] parent proteins as well as casein[DOPA] protein variants, were treated under the same conditions, whereas the parent proteins served as a negative control. Different concentrations (10 and 100 µM) of $FeCl_3$ were mixed with 10 mg/mL of casein protein. After the mixture turned an orange color, 420 µM of NaOH was added to the protein solution to elevate the pH to above 9.0. The mixture was manually stirred, after which it turned a darker orange to red color. SDS samples of the iron-mediated, cross-linked proteins were taken for a subsequent SDS-PAGE analysis.

4.12 Instruments and chemicals

All instruments, devices, and chemicals used within the scope of this project are listed in Table S8 and Table S9.

5 Results

To enable intracellular heterologous expression in *E. coli*, the secretion signals at the N-termini of α- and β-casein were removed in the gBlock® genes. Their protein sequences are listed inTable S1. The protein sequences of the caseins used in this study. Table S1. For this study, it was pertinent to know that α-casein contains a total of ten tyrosine amino acid residues, whereas β-casein has only four.

5.1 Shake flask cultivation of synthetic casein[DOPA] variants

Cultivation of synthetic caseins incorporating L-DOPA was initially tested in the previously conducted laboratory project of Alexandra Weingartner. In order to prove its reproducibility, we performed an up-scaled expression (500 mL) of the same proteins in shake flasks. The supplementation-based protocol, from Wiltschi (2012), for globally incorporating non-canonical amino acids was followed with a few deviations in order to incorporate L-DOPA into the synthetic caseins. Shortly, we grew the cells in M9 medium containing yeast extract until they consumed the limiting tyrosine in the medium. This depletion of tyrosine occurred at a D_{600} of ~3, upon which the expression of the caseins was induced. The cells were then supplemented with 10 mM L-DOPA. SDS gels of α- and β-casein[DOPA] whole-cell samples are represented in Figure 2, where an overexpression of the target proteins can clearly be seen in the overnight induced cell cultures, as is indicated by the black arrows.

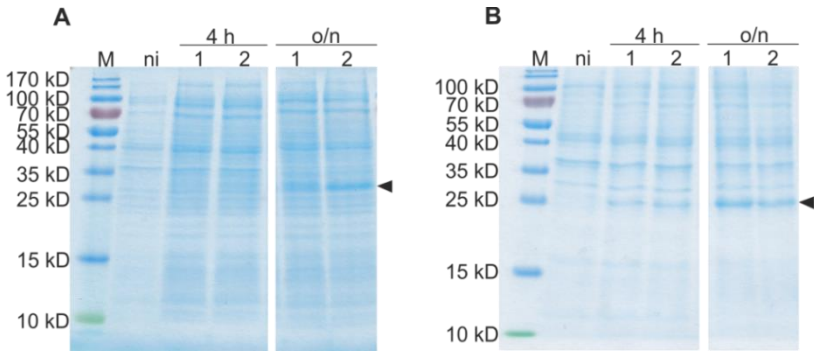

Figure 2. Successfully expressed α- and β-casein[DOPA] in 500 mL shake flask cultures.
14% Coomassie-stained SDS gels picturing whole-cell expression of α- and β-casein[DOPA]. Total protein equaling 0.1 D_{600} was loaded onto each lane. Irrelevant lanes were omitted. (A) α-casein[DOPA]; (B) and β-casein[DOPA]; M, molecular weight marker; ni, non-induced cells; 4h, cells induced for 4 hours; o/n, overnight induction. The arrows indicate the calculated molecular weights of the expected proteins (24.5 kD for A, 25 kD for B).

The 500 mL cultures were divided into 250 mL cultures for easier handling. Then, 250 mL cultures were harvested for each purification, of which we determined the protein concentration. The total amount of α-casein[DOPA] protein produced was 1.59 mg after being purified with Ni-chelate chromatography and re-buffered in PBS. In contrast, the α-casein[DOPA] protein purified in the presence of ascorbic acid yielded only 0.95 mg of total protein. We back-calculated the protein amounts of 500 mL cultures to be 3.18 mg and 1.9 mg for α-casein[DOPA] in PBS and in PBS with ascorbic acid, respectively.

L-DOPA has the tendency to oxidize in medium (Figure S3), and in order to prevent this phenomenon, we added 10 mM ascorbic acid to the medium. Therefore, we also decided to test the purification of α-casein[DOPA] with Ni-NTA affinity column buffers with and without ascorbic acid.

After observing no change in purification (see Figure 3A and B) with buffers in the presence of ascorbic acid and the lack thereof, we decided to omit this antioxidant in further target protein purifications.

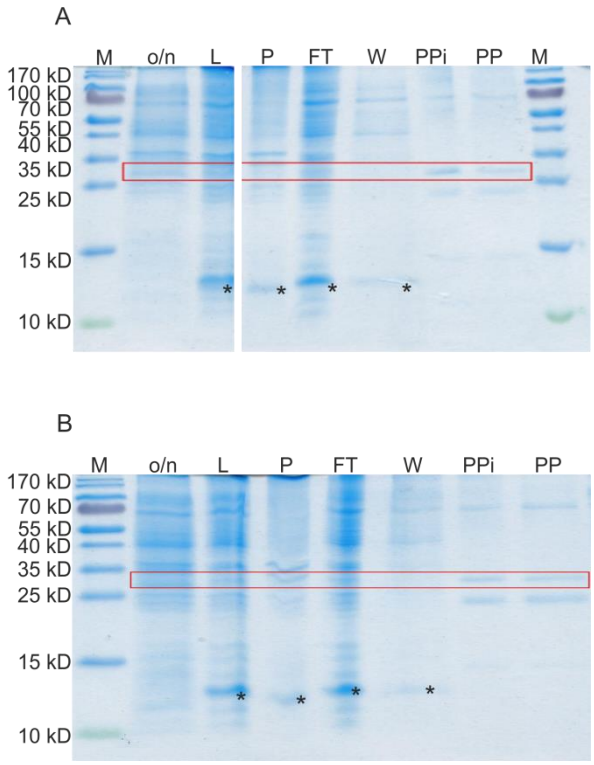

Figure 3. No effect of ascorbic acid in the purification of α-casein[DOPA] from 250 mL shake flask cultures.
14% Coomassie-stained SDS gels. (A) α-casein purified without ascorbic acid, irrelevant lane was omitted; (B) α-casein purified with ascorbic acid; M, molecular weight marker; o/n, overnight induction; L, soluble protein fraction; P, insoluble protein fraction; FT, column flow-through; W, wash fraction; PPi, casein in imidazole elution buffer; PP, casein re-buffered in PBS; asterisk, lysozyme. Total protein loaded for the o/n fraction was 0.1 D_{600}, and total protein loaded for the L, P, FT, and W fractions was 4 µg. The protein amount loaded for the PPi and PP fractions was 1 µg. The red rectangle shows the calculated molecular weight of the expected protein (24.5 kD).

To further detect the global incorporation of L-DOPA in the target proteins, we performed a redox-cycling staining of L-DOPA within these proteins, which were separated by SDS-PAGE and electroblotted onto a PDVF membrane. The casein[DOPA] proteins were then made visible by staining the membrane with NBT (see Figure 4). The arrows next to the blotted dark bands indicate the casein[DOPA] target proteins, which confirms the successful incorporation of L-DOPA.

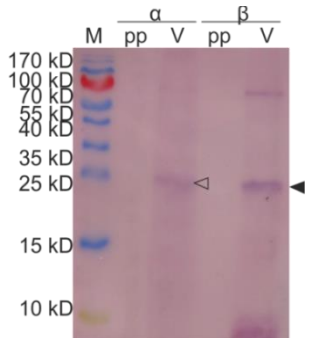

Figure 4. Specific detection of incorporated L-DOPA in α- and β-caseins.
Purified α- and β-caseins (parent protein and DOPA variant) were blotted onto a PVDF membrane and stained by NBT. The total protein loaded onto each lane was 4 µg. M, molecular weight marker; pp, casein[Tyr] parent protein; V, casein[DOPA] variant; open arrow, calculated molecular weight of α-casein[DOPA] (24.5 kD); closed arrow, calculated molecular weight of β-casein[DOPA] (24 kD).

The casein[Tyr] proteins serve as a negative control. These protein bands were not stained, which is further supporting evidence that we produced synthetic proteins incorporating L-DOPA.

An LC-MS/MS analysis was performed on the DOPA-incorporated proteins, but did not yield representative data.

5.2 High cell density fermentation of caseins

We decided upon fermenting the casein target proteins in DASGIP stirred tank reactors in order to produce high amounts of synthetic proteins for further studies.

5.2.1 Fermentation of casein[Tyr] parent proteins

The optimal yeast extract concentration for tyrosine depletion at D_{600} of 3 was determined to be 1.75 g/L, as is demonstrated in Figure S4 (Weingartner 2014).

Four 400 mL cultures were inoculated with *E. coli* BWEC48(TyrΔ) cells harboring the pET26(b)+ plasmids with the casein target genes. The cultures were cultivated to a D_{600} between 0.8 and 1, after which we induced the cellular expression by adding 0.5 mM IPTG (see growth curves in Figure 5). The dashed lines in Figure 5A and B also indicate the time when the substrate feed was initiated for only one of each casein culture in order to observe the effect the substrate feed had on cellular growth and protein expression. The cultures receiving a substrate feed had an end volume of ~650 mL. Cultures without feed had a final volume of ~420 mL before harvesting (see Table 1).

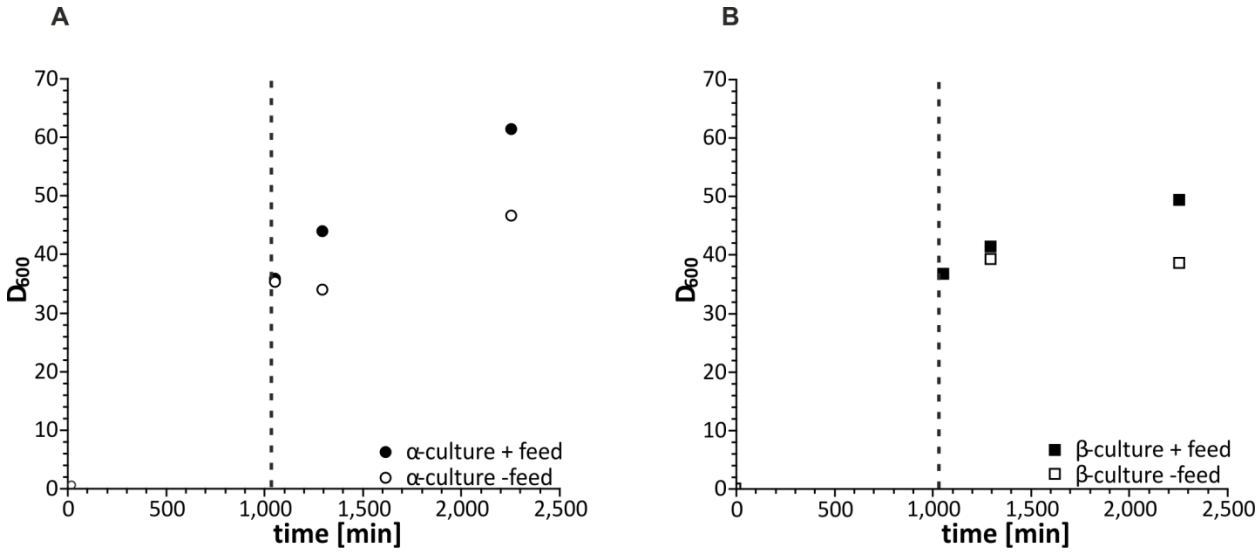

Figure 5. Cell growth of *E. coli* cells expressing parent caseins in DASGIP stirred tank reactors is optimal with a glucose feed containing all 20 cAAs.
Growth of the tyrosine auxotrophic *E. coli* strain BWEC48 (TyrΔ){pET-α-casein-H_6} (A), and BWEC48 (TyrΔ){pET-β-casein-H_6} (B) supplemented with and without a substrate feed containing glucose and all 20 cAAs. All values were determined in triplicates, of which the average is plotted. D_{600}, attenuance of cells at 600 nm; dashed line, time of induction with IPTG and start of glucose feed.

As expected, the cells supplied with a substrate feed (composed of D-glucose monohydrate, each of the 20 canonical amino acids, thiamine, and biotin, see the Materials and Methods 4.6.1 section for details) grew to a higher D_{600} than those that did not receive a substrate feed, which is also an indication of higher cell-density. The cell wet weight obtained from each culture after harvesting is listed in Table 1. Logically, the cell wet weight of those cultures that received a substrate feed was substantially higher than of those that were not fed.

Table 1. Cell wet weight from BWEC48 (TyrΔ){pET-α-casein-H_6} and (TyrΔ){pET-β-casein-H_6} *E. coli* cells.
The obtained cell wet weight is from cell cultures cultivated in bioreactors.

Casein sample	Final volume [mL]	Cell wet weight [g]
α^1	654	27.48
α^2	437	16.28
β^1	667	22.97
β^2	415	17.99

[1] Culture supplemented with a glucose substrate feed (see section 4.6.1, components are listed in Table S5), including all 20 cAA, thiamine and biotin.
[2] No feed added.

SDS samples were taken from each culture before, after four hours and overnight induction with IPTG. It is visible in Figure 6A and B that only the cultures that were fed with substrate were able to express the synthetic proteins (as indicated by the black arrows). Also, the α-casein proteins seem to be more overexpressed than the β-caseins, as is to be seen in the overnight expressed SDS sample in Figure 6A.

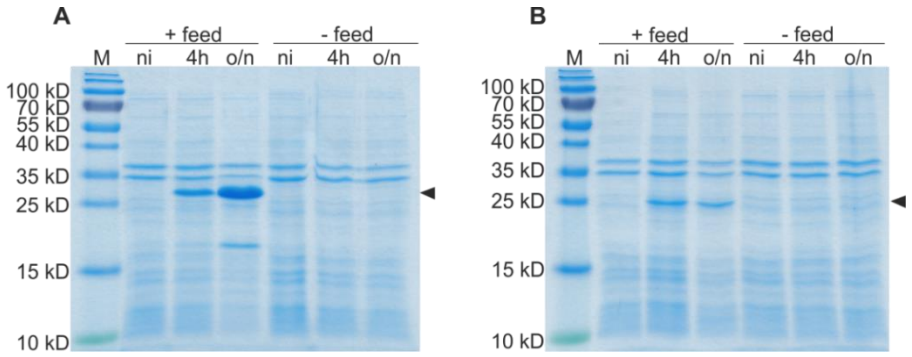

Figure 6. Expression of α- and β-caseins in stirred tank reactors is possible only with a glucose feed containing cAAs.
14% Coomassie-stained SDS gels picturing whole-cell expression of α- and β-caseins in fermented cell cultures. Total protein equaling 0.1 D_{600} was loaded onto each lane. (A) α-casein; (B) β-casein; M, molecular weight marker; ni, non-induced cells; 4h, cells induced for 4 hours; o/n, overnight induction. The arrows indicate the calculated molecular weight of the expected proteins (24.5 kD for A, 25 kD for B).

These results confirm that expression of synthetic caseins is indeed possible in stirred tank reactors. Subsequent purification of the fermented synthetic caseins (from 50 mL cell cultures) that received a substrate feed, was possible with Ni-chelate chromatography. We chose to purify only 50 mL of each culture containing the parent casein proteins in order to avoid overloading the agarose resin during Ni-chelate chromatography. The aim of this purification was not to obtain as much protein as possible, but rather to test the protein purification fractions of the parent proteins obtained from a bioreactor fermentation. The purification fractions of both α- and β-caseins are pictured in Figure 7.

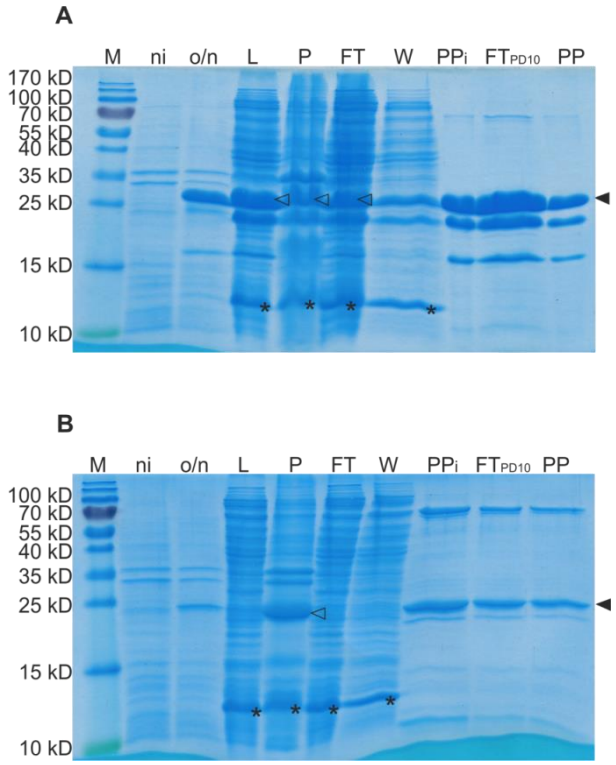

Figure 7. Successful purification of α-, and β-caseins produced in the bioreactor.
14% SDS polyacrylamide gels stained with Coomassie blue, depicting all purification fractions from a Ni-NTA agarose affinity column. (A) α-casein; (B) β-casein; M, molecular weight marker; ni, non-induced cells, o/n, overnight induction; L, soluble fraction; P, insoluble fraction; FT, column flow-through; W, wash fraction; PPi, eluted protein in imidazole buffer; FT_{PD10}, flow-through of the PD10 desalting column; PP, casein re-buffered in PBS buffer; asterisk, lysozyme. Total protein loaded for the ni and o/n fractions was 0.1 D_{600}, and total protein loaded for the L, P, FT, FT_{PD10}, and W fractions was 4 µg. The protein concentration loaded for the PP fraction was 1 µg. Closed arrows, calculated molecular weight of expected protein (24.5 kD for A, 25 kD for B); open arrows, indicate overloaded protein.

The gel lanes are clearly overloaded. The open arrows seen in Figure 7 indicate the loss of protein during the purification with the gravity flow affinity columns. Nevertheless, we were able to purify the parent proteins and calculate the total protein in each fraction (see Table 2). It is notable to state that purified and re-buffered α-casein total protein is much higher than β-casein, with 24 mg and 1.5 mg, respectively. These results are discussed in further detail in the Discussion section 6.2.1.

Table 2. Concentration of total proteins of each purification fraction of fermented α- and β-caseins.
Proteins were purified from 50 mL cell cultures, and the concentrations were determined via a Bradford assay.

Purification fraction	Concentration [mg/mL]	Total protein [mg]
	α-casein	
L[1]	3.55	17.5
FT[2]	2.41	120.5
PPi[3]	0.86	43.0
FT_{PD10}[4]	0.41	20.5
PP[5]	0.48	24.0
	β-casein	
L[1]	2.77	138.5
FT[2]	2.11	105.5
PPi[3]	0.04	2.0
FT_{PD10}[4]	0.02	1.0
PP[5]	0.03	1.5

[1] Soluble protein fraction.
[2] Column flow-through fraction.
[3] Eluted casein in imidazole buffer.
[4] Flow-through of the PD10 desalting column.
[5] Purified casein re-buffered in PBS.

Using the concentrations from Table 2, we were able to back-calculate the amount of protein obtained from roughly 650 mL high cell density cultures, which was 192 mg and 12 mg of α- and β-casein[Tyr], respectively.

Since the fermentation of the parent caseins was possible, we performed a stirred tank fermentation of the casein[DOPA] variants.

5.2.2 Production of casein[DOPA] variants with non-canonical amino acid supplementation

In the present study, the tyrosine auxotrophic *E. coli* BWEC48(TyrΔ) strains harboring the pET-α-casein and pET-β-casein plasmids were fermented in DASGIP stirred tank reactors. Four hundred milliliters of M9-based medium containing 4.9 g yeast extract were inoculated to a D_{600} of 0.1 and grown to a D_{600} of ~21, after which we induced the heterologous protein expression by adding IPTG. The conditions were kept the same as for the casein[Tyr] parent protein production. The yeast extract that was added to the medium was calculated for the depletion phase of the tyrosine auxotrophic cells (D_{600} 21). Following induction of target protein expression, the D_{600} of all fermented cultures stagnated between 25 and 33. After 35 hours of cultivation, the cells were harvested after a decline in attenuance for both α- and β-casein[DOPA] was observed (see Figure 8A and B).

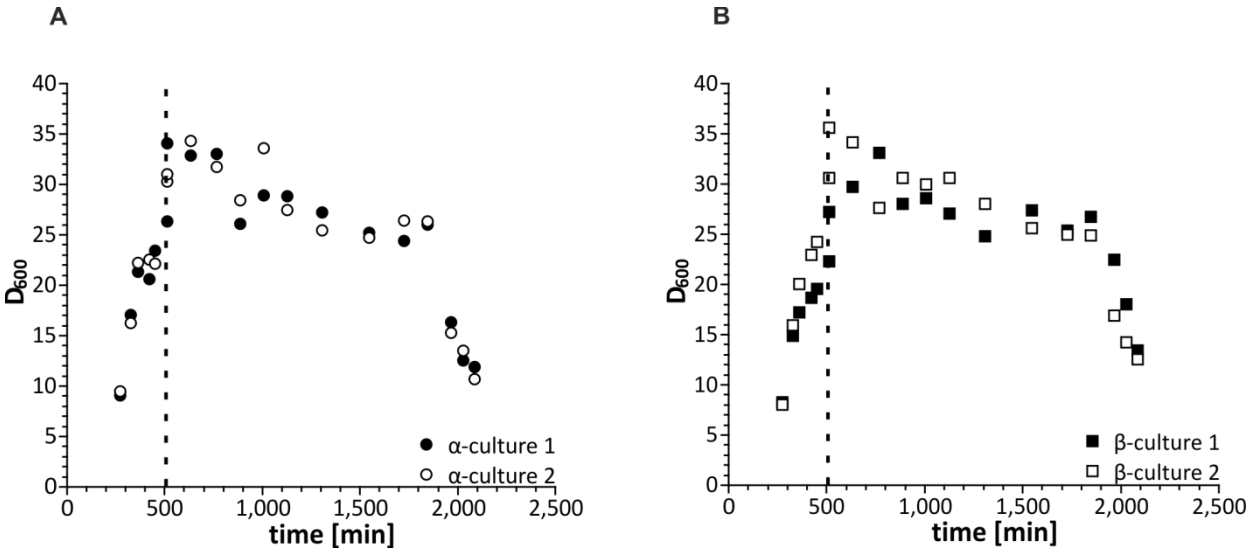

Figure 8. Cell growth curves of casein[DOPA] variants in DASGIP stirred tank reactors depicted as attenuance at 600 nm vs. time do not differ.
Growth of the *E. coli* BWEC48(TyrΔ){pET-α-casein-H_6} (A), and BWEC48(TyΔ){pET-β-casein-H_6} (B) strains in 400 mL cultures, supplemented with a substrate feed containing glucose, all cAAs (except tyrosine), biotin, thiamine, and L-DOPA. Both variants show comparable growth behavior. All values were determined in triplicates, of which the average is depicted in the graphs. D_{600}, attenuance of cells at 600 nm; dashed line, time of induction with IPTG and start of substrate feed.

Process data that was recorded from one of the bioreactors can be viewed in Figure S6. It demonstrates that the pO_2 is lowered (red curve) in bioreactor 1 (that included α-casein[DOPA]) after roughly 4 hours of fermenting, after which the impeller speed was automatically increased. This plummeting pO_2 curve was an indication that the cells were in the log phase and consumed more oxygen than was provided. In order to supply the cells with the oxygen they require for growth, the impeller must stir faster. This is visualized by the abrupt spike of the black curve (stirrer speed) in Figure S6. After ten hours of fermenting, the cells then went into growth arrest after the D_{600} reached ~25 (slightly above the calculated depletion phase), which correlates to the recorded process data where the pO_2 and the stirrer velocity leveled off. All other process data looked very similar to the one shown in Figure S6. Notably, the recorded process data mirrors the D_{600} values of the cultures (Figure 8) very well.

During the course of fermentation of both α- and β-casein[DOPA], SDS samples were taken before and after induction in different time intervals. Both A and B panels of Figure 9 reveal a detectable overexpression of target proteins only after 24 hours of induction.

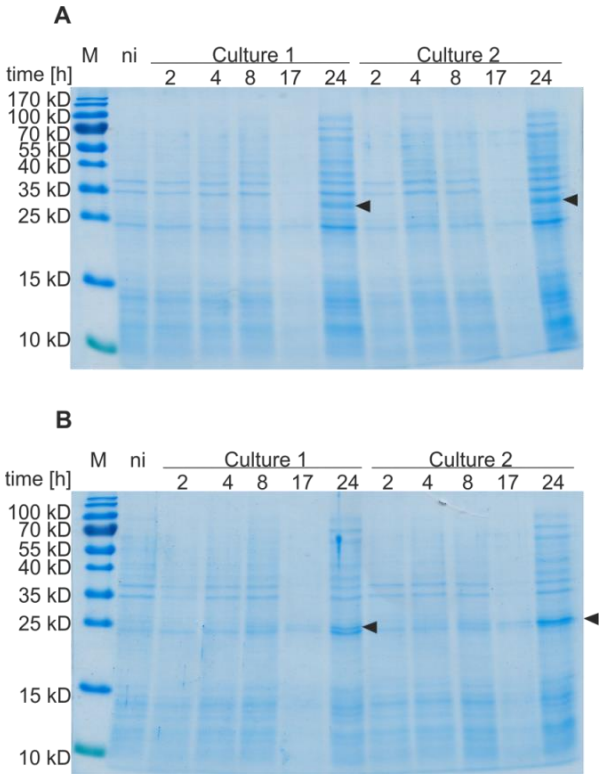

Figure 9. Expression of α- and β-casein[DOPA] in DASGIP stirred tank reactors.
14% Coomassie-stained SDS gels picturing whole-cell expression of α- and β-casein[DOPA] at different times. Total protein equaling 0.1 D_{600} was loaded onto each lane. (A) α-casein[DOPA]; (B) and β-casein[DOPA]; M, molecular weight marker; ni, non-induced cells. The arrow indicates the calculated molecular weight of the expected proteins (24.5 kD for A, 25 kD for B).

After harvesting the cells, the cell wet weight of each culture was determined. The cell wet weight for both casein variants ranges between 11.03 g and 12.35 g (see Table 3).

Table 3. Cell wet weight of fermented BWEC48 (TyrΔ){pET-α-casein-H_6} and (TyrΔ){pET-β-casein-H_6} cells that were supplemented with L-DOPA.

Casein[DOPA] sample	Final volume [mL]	Cell wet weight [g]
α (bioreactor 1)	770	11.03
α (bioreactor 2)	758	11.04
β (bioreactor 3)	762	12.35
β (bioreactor 4)	785	11.24

The fermented caseins were then purified with Ni-chelate crhomatography. The total casein[DOPA] produced in ~770 mL cultures is 34.66 mg and 34.30 mg for α- and β-casein[DOPA], respectively.

To prove the incorporation of L-DOPA into the target casein proteins in bioreactors, a redox-cycling staining of these L-DOPA-containing proteins was done. The proteins were separated by SDS-PAGE, and subsequently electroblotted onto a PDVF membrane. Figure 10A illustrates the SDS gel of the casein[Tyr] parent proteins and their variants stained with Coomassie blue as a negative control. The DOPA detection of the casein[DOPA] variants is visualized after the electroblotted PDVF membrane was stained with NBT (see Figure 10B).

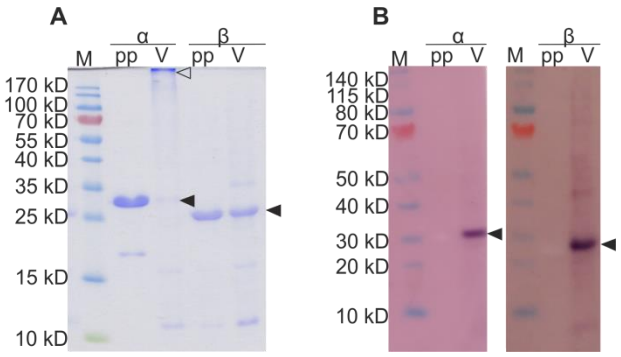

Figure 10. DOPA-stain reveals successful incorporation of L-DOPA into caseins produced in bioreactors.
14% SDS polyacrylamide gel stained with Coomassie blue showing purified α- and β- parent caseins and their DOPA variants (A); (B) blotted PDVF membrane stained with NBT, illustrating only stained DOPA variants. The lanes were loaded with a total protein of 4 µg. M, molecular weight marker; pp, casein[Tyr] parent protein; V, casein[DOPA] variant; black arrows; calculated molecular weight of the expected protein (24.5 kD for α-casein; 25 kD for β-casein).

As an additional confirmation of incorporated L-DOPA in the casein proteins, a photometric analysis of a mixture of α- and β-casein[DOPA] variants was done (see Figure 11). A UV-Vis spectrum between 300 and 700 nm was recorded, since it is known that L-DOPA absorbs at λ_{max} = 395 nm (Yu, Hwang, and Demin 1999). There is a slight peak detected between 340 and 400 nm confirming the presence of L-DOPA.

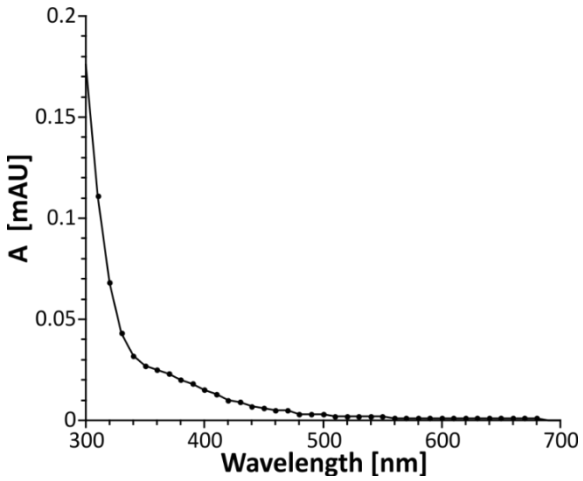

Figure 11. UV-Vis absorbance of a mixture of α- and β-casein[DOPA].
A spectrum of an α- and β-casein[DOPA] mixture yields a slight peak between 340 and 400 nm. This correlates to the known absorbance of L-DOPA.

5.3 Comparison of purification conditions of caseins produced in shake flasks to fermented caseins

5.3.1 Size exclusion chromatography confirms that urea keeps caseins from aggregating

To prevent aggregating behavior of the native proteins, we decided to add urea to all buffers during protein purification.

We decided to perform a size exclusion chromatography (SEC) of the casein[Tyr] parent protein and the casein[DOPA] variant to observe the effect of urea on the molecular weight of the proteins. We

worked with a Superdex 200 FPLC column, onto which we initially loaded a commercial protein standard. To detect the influence of urea on the proteins, we had to include urea also in the buffer of the protein standard. Figure 12 illustrates the chromatogram of the protein standards with (red curve) and without (black curve) urea. With the exception of the first peak (aggregates), the standard has the same elution behavior in the absence and presence of this aggregate-deterring reagent.

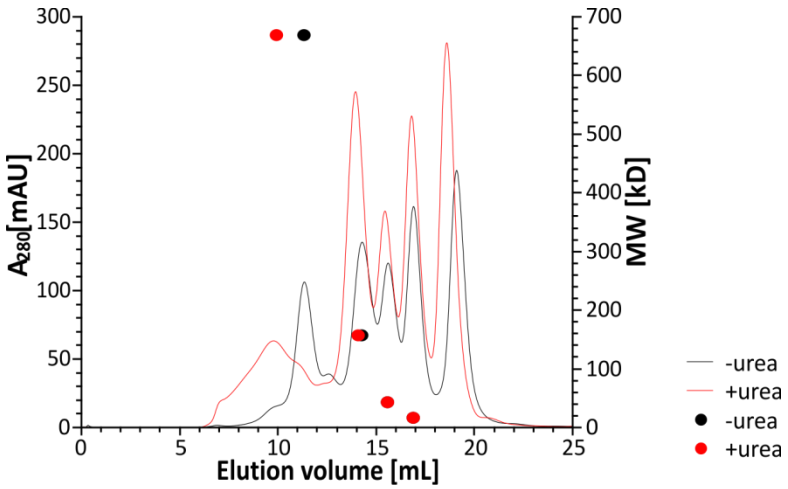

Figure 12. Size exclusion chromatogram of a commercial protein standard with and without urea. Black curve, elution pattern of protein standard without urea in the PBS buffer; black dots, molecular weight of the eluted standard without urea; red curve, elution pattern of protein standard with 3.75 M urea in the PBS buffer; red dots, molecular weight of the eluted standard with urea.

The calibration curves of both standards from Figure 12 can be found in the supporting information (Figure S8A and B).

We then loaded the α-casein[Tyr] parent protein onto the Superdex 200 FPLC column and eluted the sample with 10 mM PBS buffer (Figure 13A, black curve). We ran the same sample, but instead eluted with PBS including 3.75 M urea (Figure 13A, red curve). Casein[Tyr] elutes later in the presence of urea, indicating that the protein is smaller (perhaps even monomeric). The molecular weights of the eluted peaks were calculated using the equation from Figure S8. The molecular weight of α-casein[Tyr] in PBS without urea is ~380 KD (eluted at 12.7 mL), whereas the calculated molecular weight of α-casein[Tyr] in PBS in the presence of urea is ~ 20 kD (eluted at 17.3 mL), which is close to the documented molecular weight of the protein (24.5 kD according to the UniProt database, P02662).

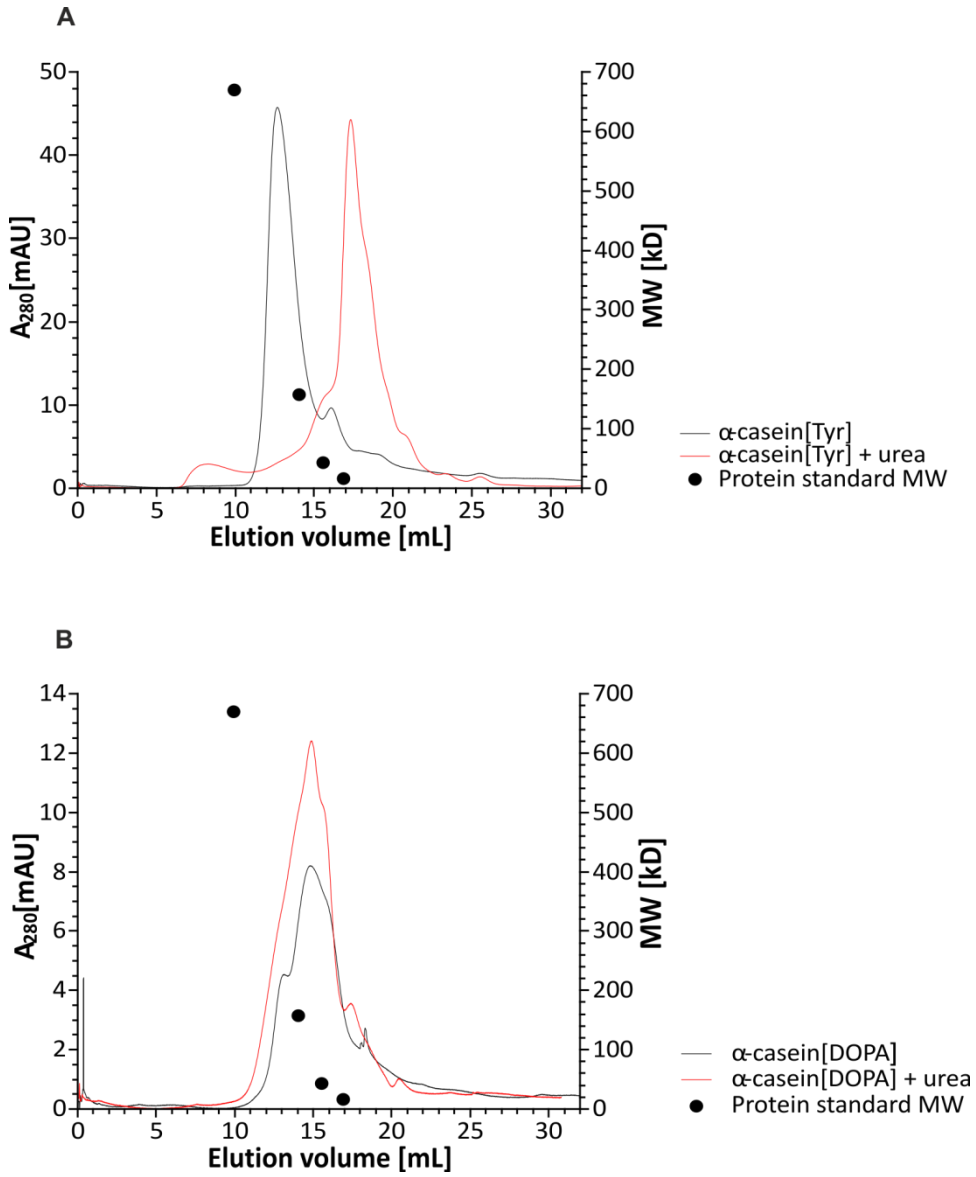

Figure 13. Size exclusion chromatograms of the influence of urea on the elution behavior of α-casein[Tyr] parent protein and α-casein[DOPA] variant.
Both curves of each graph were normalized so that the signal starts at zero and is not a negative value. (A) α-casein[Tyr]; (B) α-casein[DOPA]; black curves, caseins eluted without urea in the elution buffer; red curves, caseins eluted with 3.75 M urea in the elution buffer; black dots, molecular weight of the protein standard.

The same was done for the α-casein[DOPA] protein. It was loaded onto the Superdex 200 FPLC column and eluted with PBS in the absence and presence of urea. It is visible from Figure 13B that urea has little to no influence on the casein[DOPA] variants. Their elution pattern is very similar. We calculated the molecular weight (using the calibration equation from Figure S8) of α-casein[DOPA] that eluted in PBS with urea to be ~60 kD (elution at 14.91 mL), whereas α-casein[DOPA] eluted in PBS only is calculated to be 107 kD (eluted at 13.81 mL).

From this point forward, 3.75 M urea was added to all the buffers in the purification of the casein proteins.

We then switched the flow of the Superdex column in ÄKTA to up-flow due to delta column overpressure in the Superdex column, as was described in section 4.9. Afterward, we had to run the

commercial protein standard again on the column with this changed flow parameter and compare the elution patterns of up-flow to down-flow. As can be seen in Figure 14, the elution pattern of up-flow is very similar to that of down-flow.

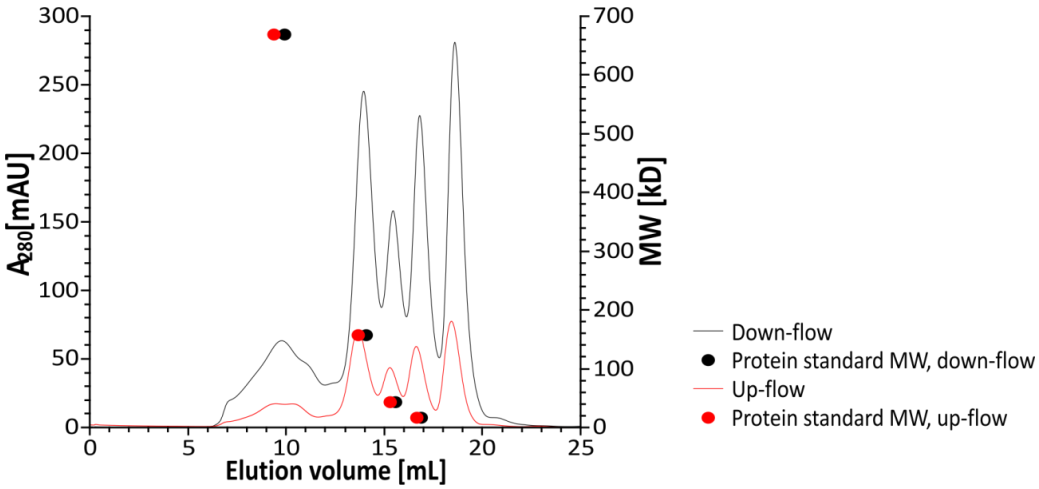

Figure 14. Size exclusion chromatogram of a commercial protein standard eluted with down- and up-flow.
Black curve, elution pattern of protein standard executed in down-flow; black dots, molecular weight of the eluted standard with down-flow; red curve, elution pattern of protein standard executed in up-flow; red dots, molecular weight of the eluted standard with up-flow.

Nevertheless, a new calibration curve was plotted using the up-flow elution peaks from Figure 14. The curve can be seen in Figure S9 and was used from this point on to calculate the molecular weights of all samples that were loaded onto the same column.

5.3.2 SDS-PAGE analysis of caseins from flasks and bioreactors purified with and without urea

Illustrated in Figure 15 are SDS gels of the purification fractions of 1 g fermented cell pellets containing α- and β-casein[Tyr] parent proteins and the casein[DOPA] synthetic variants. Here, we tested the effect that urea and DTT have on the caseins during the purification process. There is a noticeable difference in the insoluble fractions (P) in panels B and C of Figure 15. The largest effect that urea and DTT have is on the casein[Tyr] parent proteins, especially on the β-casein[Tyr] (Figure 15C). According to these gels, urea makes the proteins more soluble.

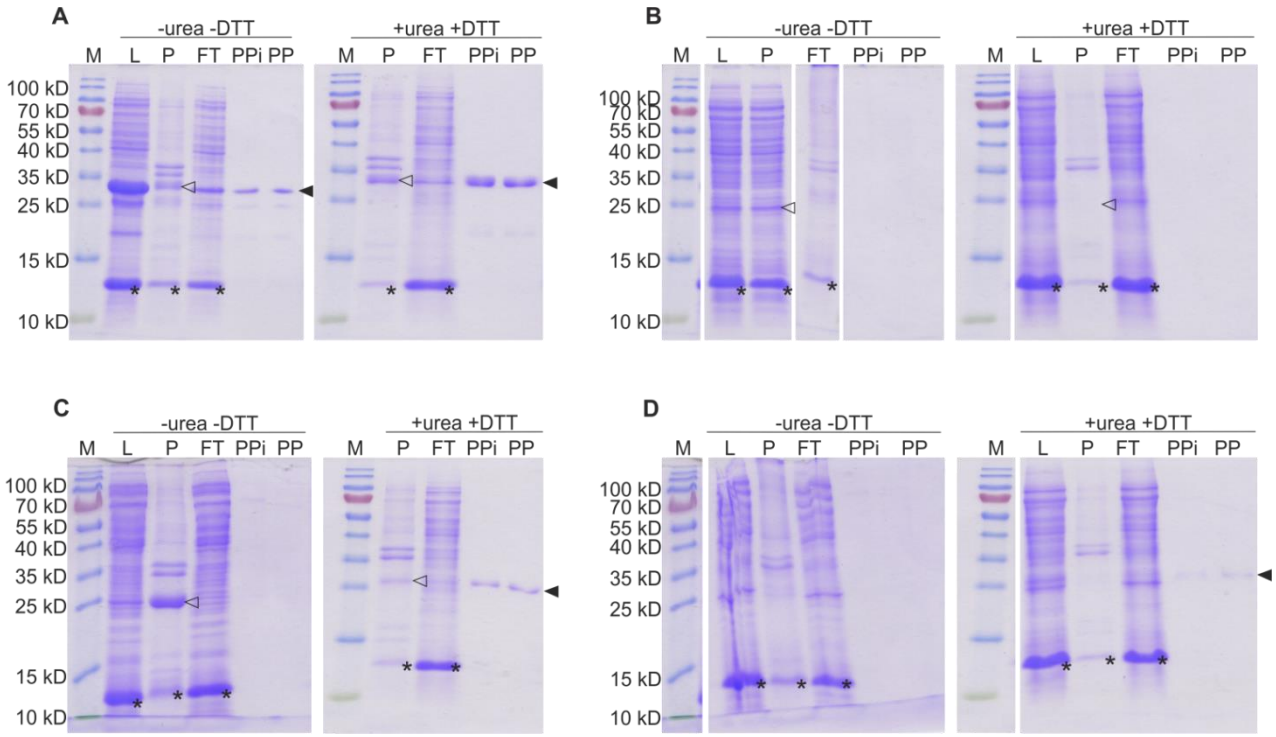

Figure 15. Comparison of purification fractions of fermented α- and β-casein[DOPA] proteins with and without urea and DTT in the purification buffers.
14% Coomassie-stained SDS polyacrylamide gels of purification fractions of 1 g cell pellets of α- and β-casein parent proteins and their DOPA variants. Irrelevant lanes were omitted. (A) α-casein[Tyr]; (B) α-casein[DOPA]; (C) β-casein[Tyr]; (D) β-casein[DOPA]; M, molecular weight marker; L, soluble fraction; P, insoluble fraction; FT, column flow-through; PPi, eluted casein in imidazole buffer; PP, purified casein re-buffered in PBS; asterisk, lysozyme; open arrows, detectable difference in solubility; closed arrows, calculated molecular weight of expected proteins (24.5 kD for α-casein, 25 kD for β-casein). Total protein loaded for the L, P, and FT fractions was 4 µg. The protein concentration loaded for the PPi and PP fractions was 1 µg.

The purified α-casein[DOPA] PPi and PP fractions are not visible (Figure 15B), because the protein concentration was too low. The protein concentrations calculated according to a Bradford assay were faulty. Therefore, an insufficient amount of protein was loaded onto the gel.

We also decided to compare amounts of protein produced in 1 g cell pellet in shake flasks to that produced in stirred tank reactors. The SDS gels in Figure 16 illustrate the expression of the casein proteins.

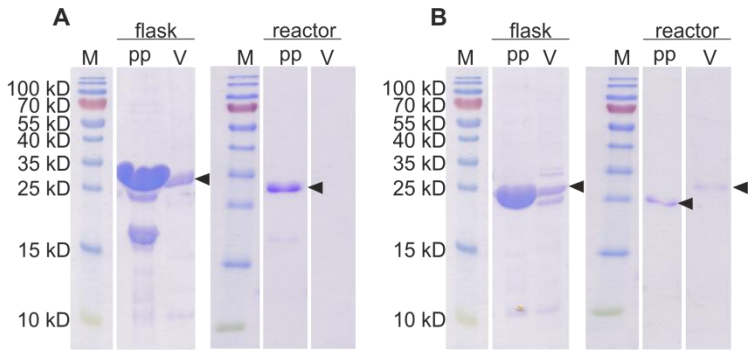

Figure 16. Comparison of α- and β-caseins produced in 1 g cells from shake flasks and bioreactors. 14% SDS polyacrylamide gels stained with Coomassie blue. Depicted is the comparison of expressed caseins in 1 g cell pellet that was cultivated in shake flasks and in bioreactors. The proteins obtained from shake flask cultivations were concentrated prior to loading onto the gels. Irrelevant lanes were omitted. (A) α-casein; (B) β-casein; M, molecular weight marker; pp, casein[Tyr] parent protein; V, casein[DOPA] variant; black arrow, calculated molecular weight of expected proteins (24.5 kD for A, 25 kD for B). Total protein loaded onto each lane was 4 µg according to estimated protein concentrations after a Bradford assay.

Nonetheless, these gels are misleading since the purified proteins obtained in shake flasks were concentrated, and those obtained from the bioreactor were not. All protein concentrations were estimated using a Bradford assay. Even though 4 µg of protein were loaded onto each gel, the protein samples obtained from bioreactors are hardly visible, whereas the shake flask samples seem overloaded. Therefore, the estimated protein concentrations after a Bradford assay must have been faulty. These issues are discussed further in section 6.3.

5.4 Bioconjugation of synthetic casein[DOPA] variants

5.4.1 Oxidative bioconjugation of α-casein[DOPA] obtained from shake flasks

Initial α-casein[DOPA] bioconjugation trials included microdialysis in the presence of ascorbic acid in the elution buffer (see Figure S7). As the microdialysis approach did not lead to apparent protein cross-linking (as seen in Figure S7), we chose to directly add two different $NaIO_4$ concentrations (1 and 5 mM) to the synthetic caseins, and incubate the protein solution under vigorous shaking (750 rpm) at room temperature for four hours and overnight. An SDS-PAGE was performed (see Figure 17A) with a subsequent redox-cycling staining of the electroblotted PDVF membrane (see Figure 17B) to visualize any cross-linked proteins. The SDS gel reveals that the presence of ascorbic acid in PBS hinders the conjugation of the caseins, since there are no bands to be seen in the SDS samples that included the antioxidant (both panels of Figure 17).

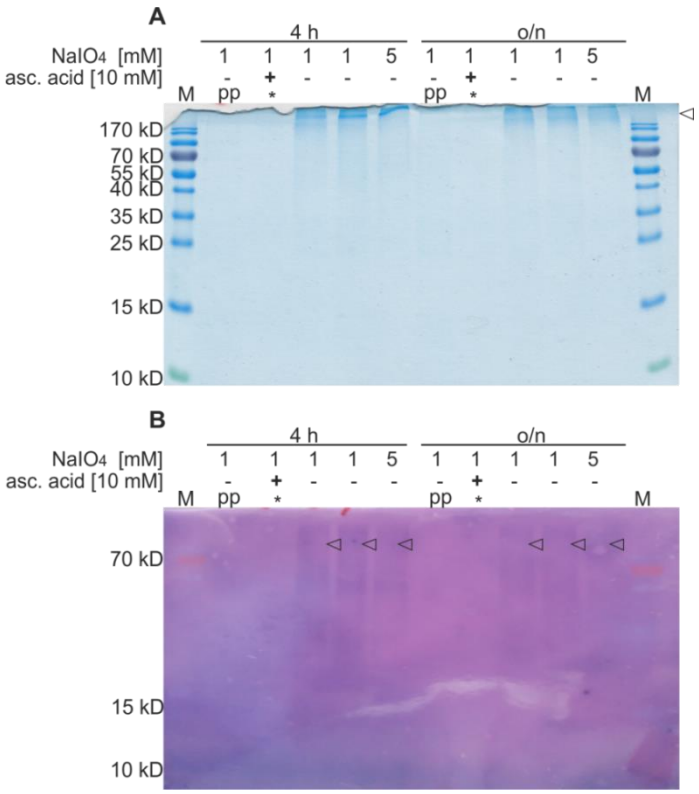

Figure 17. Oxidative cross-linking of α-casein[DOPA] with direct addition of $NaIO_4$.
(A), 14% SDS gel stained with Coomassie blue of α-casein[DOPA] treated with $NaIO_4$ for different times under vigorous shaking; (B), same as (A), only this SDS gel was blotted onto a PDVF membrane and stained with NBT to visualize the L-DOPA-containing proteins. M, molecular weight marker; pp, α-casein parent protein containing tyrosine (negative control); 4h, proteins treated with $NaIO_4$ for four hours; o/n, caseins treated with $NaIO_4$ overnight; open arrows, aggregated L-DOPA-containing proteins; asterisk, an older preparation.

Considering that we did not detect any monomeric protein bands in the casein[Tyr] parent protein, we stipulate that the caseins aggregate in their native state. Literature shows that this is indeed correct, after which we added 3.75 M urea to all the purification and storage buffers, as has already been done by Chaplin in 1986. Urea's aggregation deterring properties helped in keeping the caseins from agglomerating, as is discussed below.

A SEC was also performed with the α-casein[Tyr] and α-casein[DOPA] samples that had 1 mM $NaIO_4$ added to them. The α-casein[Tyr] served as a negative control in this case, on account that it does not have the two protruding adjacent hydroxyl groups needed for cross-linking. Only the $NaIO_4$-treated casein[DOPA] variant should elute much earlier than the untreated casein[DOPA], since it should be cross-linked and therefore have a larger molecular weight. Figure 18 shows that $NaIO_4$ indeed does not have any effects on the α-casein[Tyr] parent protein as expected (panel A). However, Figure 18B does not show a difference in the elution pattern between the α-casein[DOPA] treated with $NaIO_4$ and that of the untreated casein[DOPA]. Nevertheless, the eltuion volumes of the DOPA-containing caseins is different from those of the casein[Tyr] parent proteins. The casein[DOPA] proteins elute much earlier from the Superdex column, indicating that they are larger than their parent species. This elution behavior also implies that the casein[DOPA] proteins tend to aggregate even in the absence of an oxidating agent. The absorption signal of α-casein[DOPA] was not very high. In this case, the protein concentration of the casein[DOPA] variant was too low for

cross-linking studies. Therefore, we turned to use stirred tank reactors to produce more protein for further bioconjugation trials.

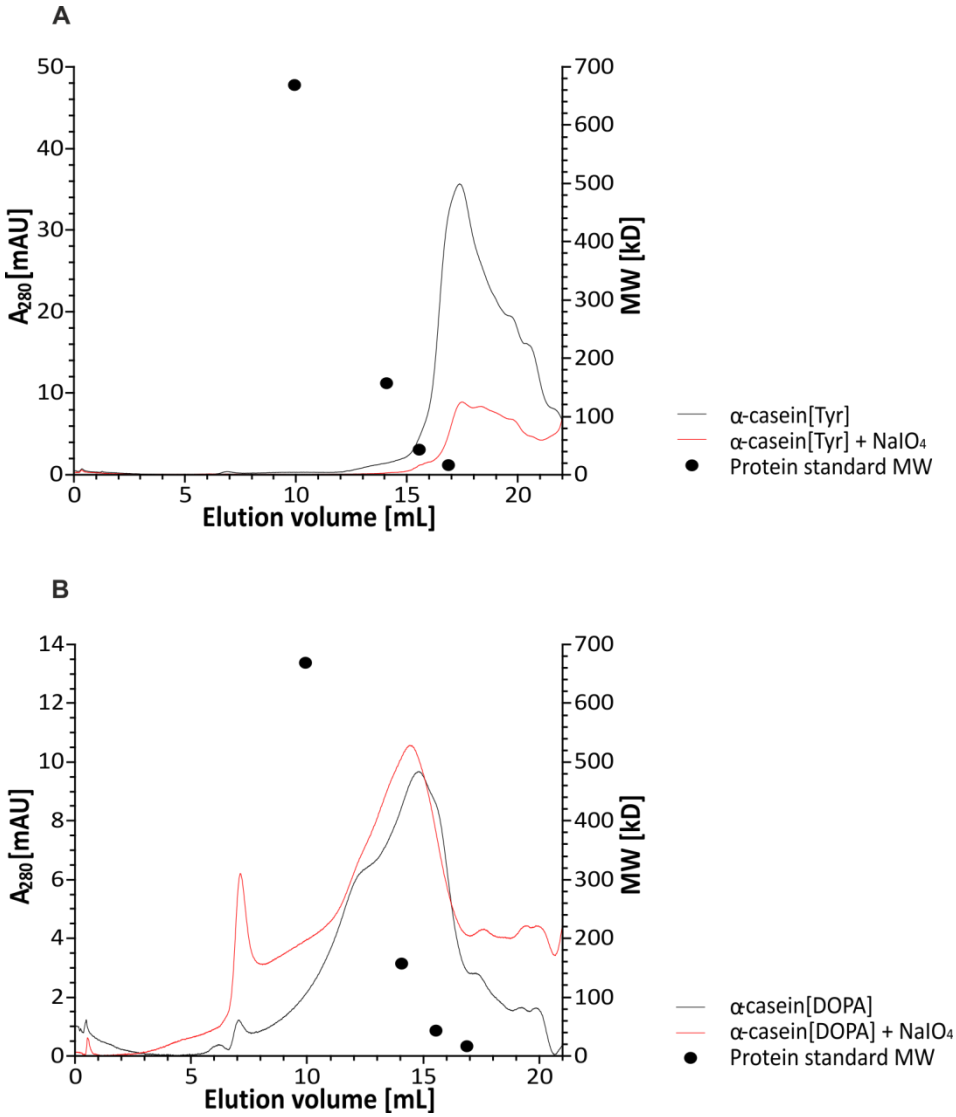

Figure 18. Size exclusion chromatograms of initial oxidative cross-linking trials with $NaIO_4$. Both curves in each graph were normalized so that the signal starts at zero and is not a negative value. Elution behavior of casein[Tyr] parent proteins and DOPA variants from a Superdex 200 FPLC column. (A) α-casein[Tyr]; (B) α-casein[DOPA]; black curves, caseins not treated with $NaIO_4$ as negative controls; red curves, caseins treated with $NaIO_4$; black dots, molecular weight of the protein standard.

5.4.2 Oxidative bioconjugation of β-casein[DOPA] obtained from shake flasks

Only the 1 mM concentration of $NaIO_4$ was chosen for the cross-linking of β-casein[DOPA], since we observed this concentration to have been sufficient for the previous cross-linking trials with the α-casein[DOPA] samples. The size exclusion chromatogram in Figure 19 is the first indication of successful oxidative cross-linking of the casein[DOPA] variants and their two hydroxyl groups conjugated via the aromatic ring system. The red curve in Figure 19 represents the β-casein[DOPA] treated with $NaIO_4$, and the black curve is the untreated protein sample as a negative control.

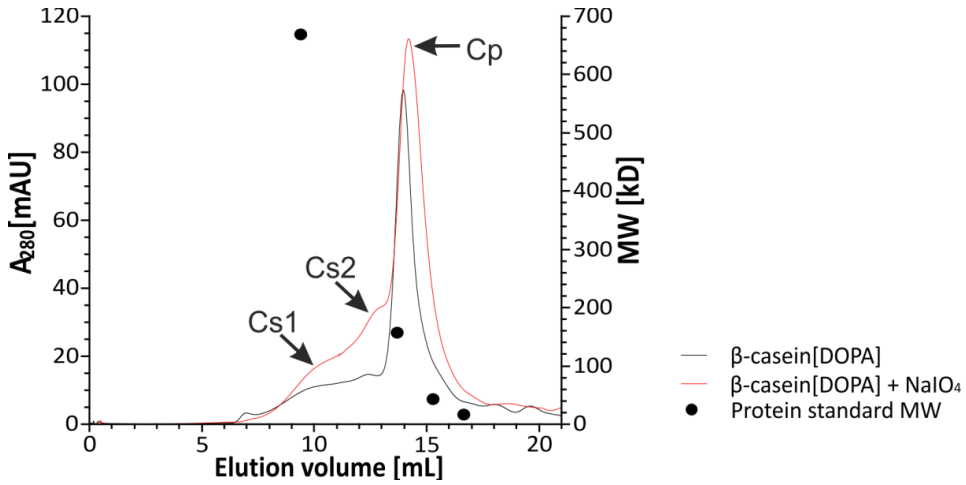

Figure 19. Size exclusion chromatograms of oxidative cross-linking of β-casein[DOPA] with $NaIO_4$. The curves were normalized to each other according to how much protein was loaded onto the Superdex 200 FPLC column, so that the signals obtained are representative. Black curve, casein[DOPA] untreated as a negative control; red curve, casein[DOPA] treated with NaIO4; Cs1, first cross-linked shoulder; Cs2, second cross-linked shoulder; Cp, cross-linked peak; black dots, molecular weight of the protein standard.

The molecular weights of the Cs1 and Cs2 shoulders, as well as that of the Cp peak were calculated using the calibration curve from Figure S9, and are listed in Table 4. In view of the calculated molecular weight of the negative control, which was the β-casein[DOPA] not treated with $NaIO_4$, we also observed an aggregating behavior of the β-casein[DOPA] just as for the α-casein[DOPA]. According to the UniProt database (P02666), the calculated the molecular weight of β-casein is 25 kD. However, after SEC, we calculated the molecular weight to be 128 kD. Nevertheless, referring to the calibration curve from Figure S9, the calculated molecular weights of the $NaIO_4$-treated β-casein[DOPA] samples (with 522 kD and 234 kD of the Cs1 and Cs2 peaks, respectively) are much higher than that of the β-casein negative control (128 kD), as can be seen in Table 4. Therefore, cross-linking is probable.

Table 4. Calculated molecular weights of the eluted β-casein[DOPA] protein from the peaks of the size exclusion chromatogram in Figure 19.
The calculation of molecular weight is based on the calibration curve in Figure S9.

Eluted peaks	Elution volume [mL]	Calculated MW [kD]
Cs1[1]	10.91	522
Cs2[2]	13	234
Cp[3]	14.22	108
Neg. control[4]	13.98	128

[1] First cross-linked shoulder.
[2] Second cross-linked shoulder.
[3] Cross-linked peak.
[4] β-casein[DOPA] not treated with $NaIO_4$.

An SDS-PAGE was performed of the eluted, cross-linked β-casein[DOPA] Cs1 fraction from Figure 19, as well as of the untreated casein[DOPA] sample as the negative control (see Figure 20). The casein monomer is only visible in the untreated negative control (lane U), whereas lane Cs1 shows the protein sitting on top of the gel. The protein hardly migrated, indicating a larger multimeric protein structure.

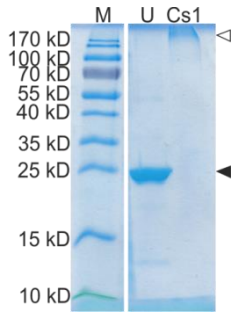

Figure 20. Initial SDS gel of successfully cross-linked β-casein[DOPA] with $NaIO_4$ taken from size exclusion elution fractions.
14% Coomassie-stained SDS polyacrylamide gel. 4 µg of total protein were loaded onto the gel lanes. M, molecular weight marker; U, untreated β-casein[DOPA] peak as a negative control; Cs1, cross-linked sample of the first shoulder; open arrow, protein aggregates; black arrow, calculated molecular weight (25 kD). Both U and Cs1 were taken from size exclusion elution fractions in Figure 19. Irrelevant lanes were omitted.

The untreated, along with all other cross-linked SEC shoulder-peak fractions of β-casein[DOPA] were loaded onto SDS gels for comparison. Aggregating behavior is visible in the $NaIO_4$-treated samples of Figure 21 (panels A and B). Particularly, the lanes of Cs1 do not include any monomeric bands. The asterisks next to the smears on the gels indicate unspecific aggregation. We also used a commercial gel with a gradient in order to get a better separation of the bands. However, as seen in Figure 21B, only the smears persist. Obviously, there are many species of cross-linked proteins.

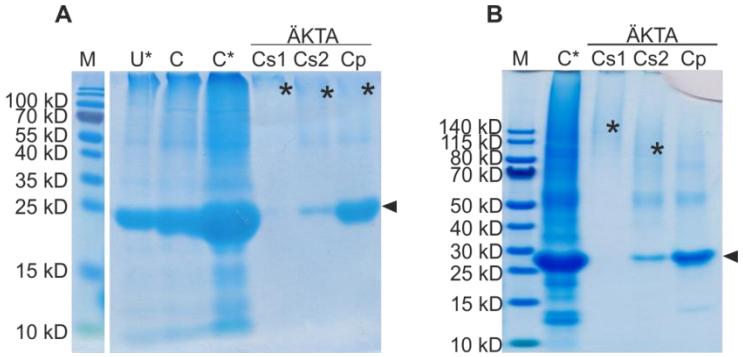

Figure 21. SDS gels showing aggregating behavior of β-casein[DOPA] after treatment with $NaIO_4$.
Coomassie-stained 14% SDS gel (A) and NuPAGE® 4-12% Bis-Tris precast gel (B). Irrelevant lanes were omitted. M, molecular weight marker; U*, concentrated β-casein[DOPA] not treated with NaIO4 as a negative control; C, cross-linked casein; C*, concentrated cross-linked casein; ÄKTA, samples taken from eluted fractions of the ÄKTA Pure system; Cs1, Cs2, Cp, are the cross-linked samples of the first shoulder, the second shoulder, and the peak that are to be seen in the chromatogram in Figure 19, respectively. Asterisk, smears indicating unspecific protein aggregation induced by $NaIO_4$; black arrow, calculated molecular weight of the β-casein[DOPA] protein (25 kD).

We tested different incubation times of the casein[DOPA] variants with 1 mM $NaIO_4$ thereafter, and performed another size exclusion chromatography on a Superdex 200 FPLC column. As can be seen in Figure 22, cross-linking the casein[DOPA] variants for 1 minute has the same effect as cross-linking the samples with the same concentration of $NaIO_4$ for longer periods of time.

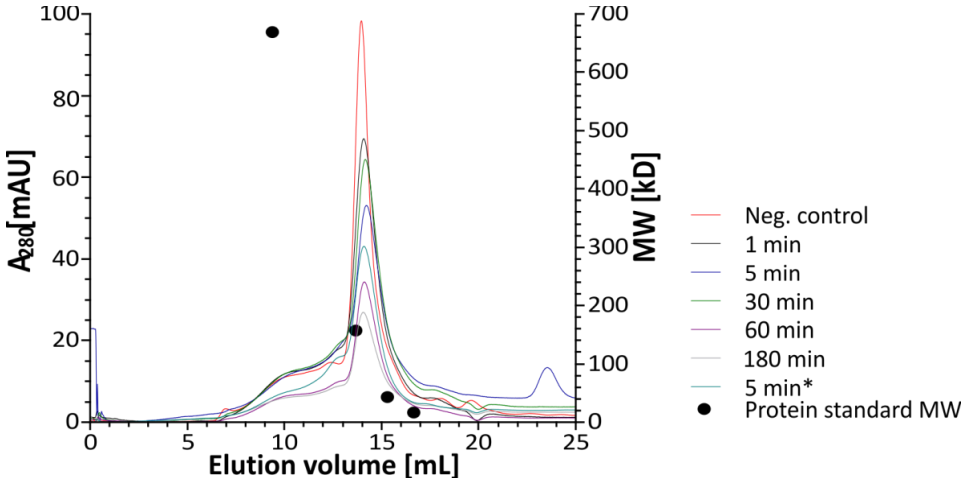

Figure 22. Size exclusion chromatogram of β-casein[DOPA] cross-linked with $NaIO_4$ for various incubation times.
Red curve, negative control (untreated β-casein[DOPA]); all other curves indicate casein[DOPA] variant cross-linked with $NaIO_4$ for different times under vigorous shaking; asterisk; freshly purified casein[DOPA] and cross-linked; black dots, molecular weight of the protein standard.

An SDS-PAGE analysis of the cross-linked β-casein[DOPA] samples from Figure 22 was performed (see Figure 23). All samples illustrate unspecific aggregating behavior on top of the SDS gels, as is indicated by the open arrows. In order to detect if the DTT that was present in the commercial SDS loading dye (NuPAGE®, Life Technologies) had an effect on the protein cross-links, we prepared an in-house SDS loading dye with and without DTT, and tested its influence on the $NaIO_4$-treated β-casein[DOPA] proteins. DTT in the loading dye did not affect the cross-linked samples, since all samples migrated the same on the gels in Figure 23.

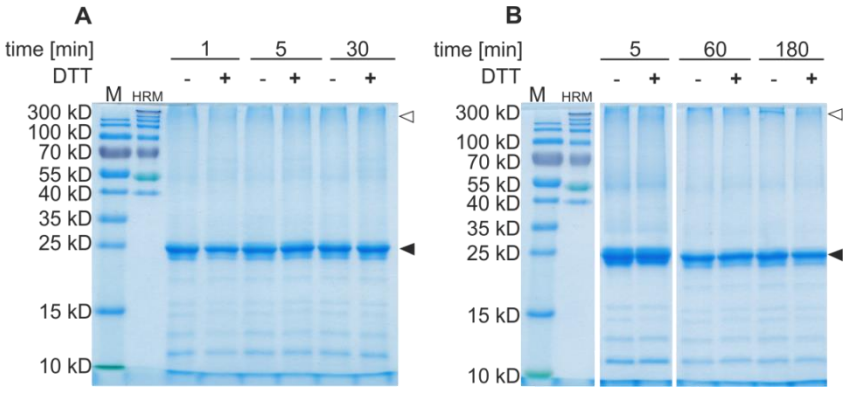

Figure 23. SDS gels of β-casein[DOPA] cross-linked for different incubation times.
14% SDS polyacrylamide gels stained with Coomassie blue. 1 µg of protein samples was loaded with and without DTT onto the gels to detect its influence on cross-links. Irrelevant lanes were omitted. (A) Cross-linked casein[DOPA] for 1, 5, and 30 minutes; (B) cross-linked casein[DOPA] for 5 (freshly purified protein and corss-linked), 60 and 180 minutes; M, molecular weight marker; HRM, high range molecular weight marker; open arrows, protein aggregates due to cross-linking; closed arrow, calculated molecular weight of the target protein (25 kD).

5.4.3 Mass analysis

For protein identification and incorporation verification, β-casein[DOPA] variant samples were submitted to the ACIB Core Facility Proteomics (Medical University of Graz), where they were analyzed by LC-MS/MS after a proteolytic digest, as indicated in the Materials and Methods section 4.10.

We wanted more supporting evidence that we managed to express caseins and incorporated L-DOPA into these target proteins; therefore, we sent excised band samples for in-gel mass analysis. Figure 24 shows the SDS gel bands that were excised and analyzed (indicated by an asterisk).

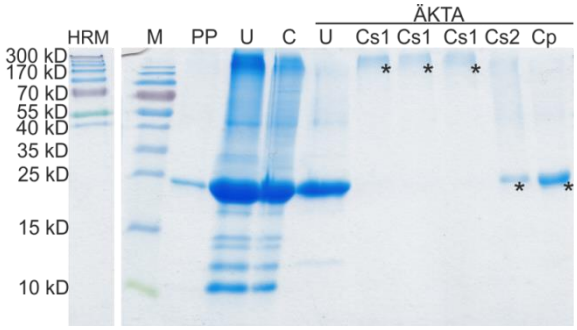

Figure 24. SDS gel of $NaIO_4$-mediated, cross-linked β-casein[DOPA] that was sent for mass analysis. 14% Coomassie-stained SDS polyacrylamide gel of cross-linked peak fractions from the SEC in Figure 19. Irrelevant lanes were omitted. HRM, high range molecular weight marker; M, molecular weight marker; PP, purified and untreated casein[DOPA] as a negative control; U, untreated and concentrated casein[DOPA]; C, cross-linked casein[DOPA] with $NaIO_4$; ÄKTA, samples taken from eluted fractions of the ÄKTA Pure system; Cs1, Cs2, and Cp are the cross-linked samples of the first shoulder, the second shoulder, and the peak that are to be seen in the chromatogram in Figure 19, respectively. Asterisks indicate smears that were excised from the gel and sent for mass analysis. 1 µg of PP were loaded onto the gel. The remaining lanes were loaded with the maximum volume of the pockets.

A positive identification of caseins is made by mass analysis. These results also reveal only one DOPA modification present in the Cp protein band that was analyzed (see Table 5). Results of the Cs1 and the Cs2 protein bands can be seen in Table S6 and Table S7. In these samples, however, only oxidation modifications were found.

Table 5. Mass analysis results of the excised Cp band from Figure 24.
The sequence coverage of Cp was 45.87%. The casein sequences were digested with trypsin, chymotrypsin, and pepsin.

Sequence	Modification	Hits
DmPIQAFLL**Y**QEPVLGPVR[1]	M2(Oxidation); **Y10(DOPA)**	1
DmPIQAFLL**Y**QEPVLGPVR	M2(Oxidation)	112
DmPIQAFLL**Y**QEPVLGPVRGPFPIIVGSHHHHHH	M2(Oxidation)	1
EmPFPK	M2(Oxidation)	2
mRELEELNVPGEIVESLSSSEESITR	M1(Oxidation)	53
HPFAQTQSLVYPFPGPIPNSLPQNIPPLTQTPVVV PPFLQPEVMGVSKVKEAMAPKHK	No fragments detected	0
YPVEPFTESQSLTLTDVENLHLPLPLLQSWMHQP HQPLPPTVMFPPQSVLSLSQSK	No fragments detected	0

[1] Labeling at Y10 is extremely low. This is the only modification found.

5.4.4 Oxidative bioconjugation of casein[DOPA] variants from fermented cultures

The fermented casein[DOPA] variants were used for cross-linking studies with different concentrations of $NaIO_4$, as well. We used 10 mg/mL of each of the α- and β-casein[DOPA] variants and mixed 2 mM and 5 mM $NaIO_4$ with each protein sample. These mixtures were then analyzed after performing an SDS-PAGE (Figure 25). It is visible that smears appear only in the casein[DOPA] variants after the addition of $NaIO_4$. This is another indication that there is covalent cross-linking behavior after the addition of this oxidative reagent.

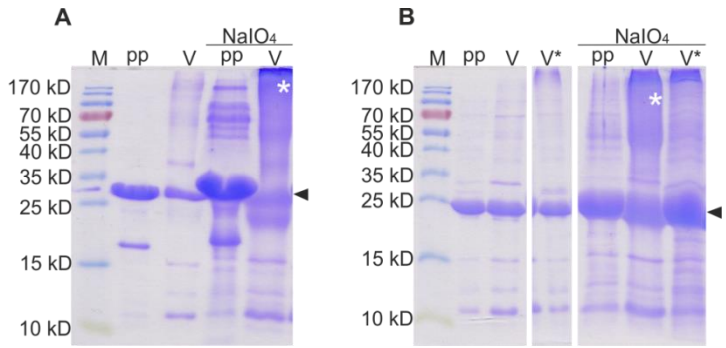

Figure 25. SDS gels of fermented α- and β-casein[DOPA] cross-linked with $NaIO_4$.
(A) α-casein[DOPA]; (B) β-casein[DOPA]. Irrelevant lanes were omitted. M, molecular weight marker; pp, untreated casein[Tyr] parent protein as a negative control; V, casein[DOPA] variant; asterisk, protein aggregates; black arrow, calculated molecular weight of target protein (24.5 kD for A, 25 kD for B). 4 µg of pp and V were loaded onto the gel. About 75 µg of the $NaIO_4$-treated samples were loaded onto each pocket of the gel.

A size exclusion chromatography was performed on the cross-linked α- and β-casein[DOPA] variants, as well as on the untreated parent proteins (from Figure 25A and B). Figure 26 shows the difference between adding 2 and 5 mM $NaIO_4$ to the α-casein[Tyr] parent protein (negative control) and the casein[DOPA] variant. There is no detectable difference between the chromatograms seen in Figure 26A and B. Only in panel A is the negative control (α-casein[Tyr]) half of the signal as that of the negative control in panel B, due to loading difficulties onto the Superdex 200 FPLC column. Some of the protein volume was lost due to handling.

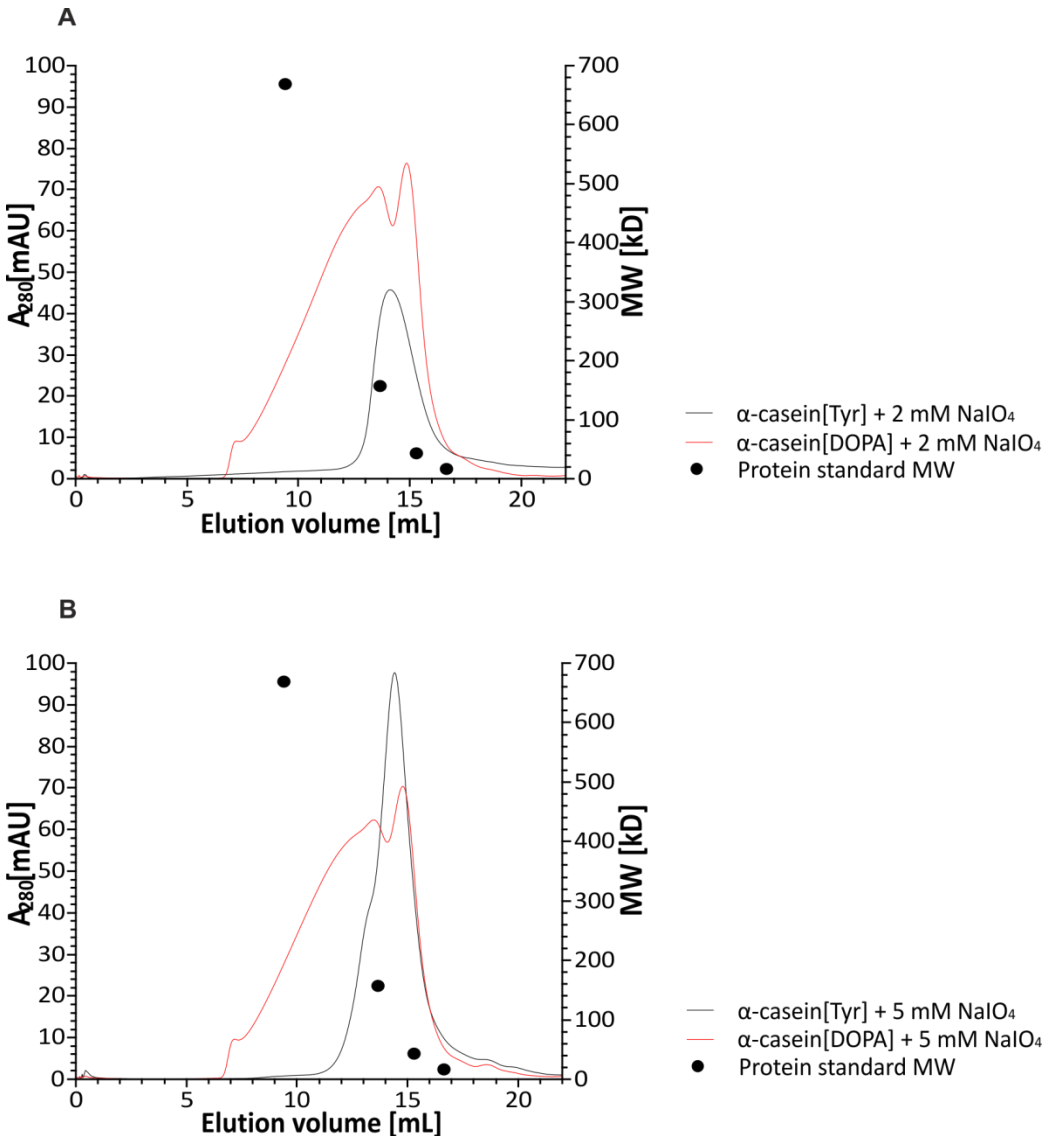

Figure 26. SEC of oxidative cross-linking trials of fermented α-casein with different concentrations of $NaIO_4$.
Elution behavior of casein[Tyr] parent proteins and DOPA variants from a Superdex 200 FPLC column. (A) α-casein[Tyr]; (B) α-casein[DOPA]; black curves, caseins not treated with $NaIO_4$ as negative controls; red curves, caseins treated with $NaIO_4$; black dots, molecular weight of the protein standard.

It is notable that the $NaIO_4$-treated casein[DOPA] samples elute much earlier than the parent casein[Tyr] proteins from the Superdex column, which is an indication of larger species of casein[DOPA] variants.

The elution pattern of the untreated α-casein[Tyr] and casein[DOPA] proteins can be seen in the chromatogram in Figure S10A. The DOPA variant exhibited mild aggregating behavior, even in the absence of the $NaIO_4$ oxidative reagent. However, the elution pattern of the casein parent protein was the same as that of the $NaIO_4$-treated casein[Tyr] samples in Figure 26, indicating once more that the periodate has no effect on the parent casein[Tyr] proteins.

An SDS-PAGE was performed on some of the eluted fractions from Figure 26 and Figure S10A. The fractions taken for an SDS gel are listed in Table 6.

Table 6. Eluted SEC fractions from Figure 26 and Figure S10A of α-casein that were taken for an SDS-PAGE (gel is pictured in Figure 29).

α-casein sample	Fraction	$NaIO_4$ concentration [mM]	Elution volume [mL]
[Tyr]	B8	-	9.01
[Tyr]	C6	-	14.01
[DOPA]	B4	-	7.01
[DOPA]	B8	-	9.01
[DOPA]	B12	-	11.01
[DOPA]	C4	-	13.01
[DOPA]	C7	-	14.51
[Tyr]	C6	2	8.01
[DOPA]	B5	2	7.51
[DOPA]	B9	2	9.51
[DOPA]	B12	2	11.01
[DOPA]	C5	2	13.51
[DOPA]	C7	2	14.01
[Tyr]	C6	5	14.01
[DOPA]	B4	5	7.01
[DOPA]	B8	5	9.01
[DOPA]	B12	5	11.01
[DOPA]	C4	5	13.01
[DOPA]	C7	5	14.51

Further size exclusion chromatography was performed on the cross-linked β-casein[DOPA] variants. Figure 27 shows the difference between adding 2 and 5 mM $NaIO_4$ to the β-casein[Tyr] parent protein (negative control) and its casein[DOPA] variant. As for the $NaIO_4$-treated β-casein[DOPA] variant, there was no detectable difference between the chromatograms seen in Figure 27A and B. Adding 2 mM $NaIO_4$ to the casein[DOPA] variants is sufficient for oxidative cross-linking.

The casein[DOPA] variants eluted much earlier than their parent proteins, which is a strong implication that the casein[DOPA] variants are indeed cross-linked, as previously seen in Figure 25B, and therefore have a larger molecular mass.

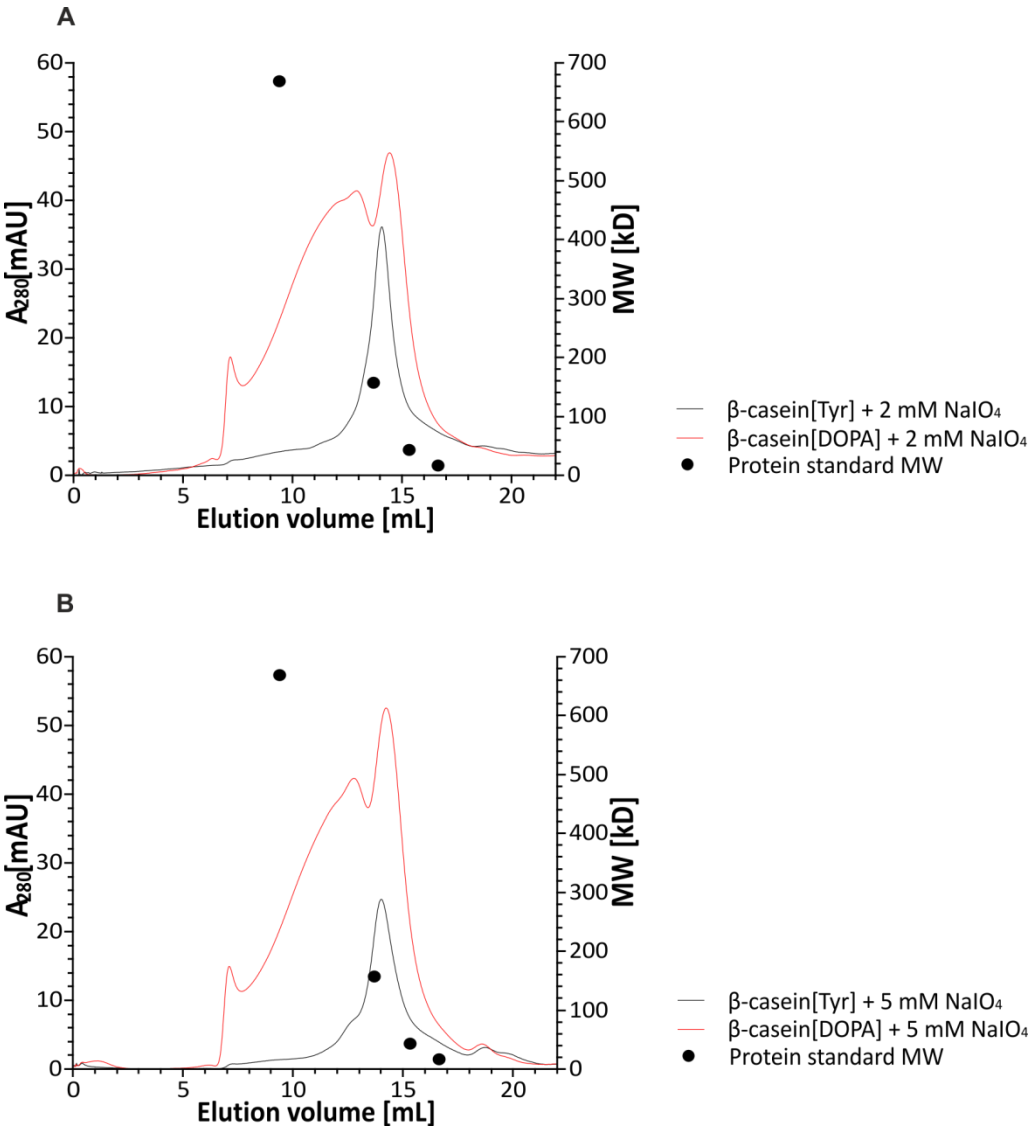

Figure 27. SEC of oxidative cross-linking trials of fermented β-casein with different concentrations of $NaIO_4$.
Elution behavior of casein[Tyr] parent proteins and DOPA variants from a Superdex 200 FPLC colum. (A) β-casein[Tyr]; (B) β-casein[DOPA]; black curves, caseins not treated with $NaIO_4$ as negative controls; red curves, caseins treated with $NaIO_4$; black dots, molecular weight of the protein standard.

The chromatogram in Figure S10B illustrates the elution pattern of the untreated β-casein[Tyr] and casein[DOPA] proteins. The chromatogram curve of the untreated DOPA variant also depicts aggregating behavior, even in the absence of the $NaIO_4$ oxidative reagent. However, the elution pattern of the casein parent protein is the same as that of the $NaIO_4$-treated casein[Tyr] samples in Figure 27, which confirms that the sodium periodate does not have an influence on the parent proteins.

An SDS-PAGE analysis was performed on some of the eluted fractions from Figure 27 and Figure S10B. The fractions taken for an SDS gel are listed in Table 7.

Table 7. Eluted SEC fractions from Figure 27 and Figure S10B of β-casein that were taken for an SDS-PAGE (see Figure 29).

β-casein sample	Fraction	$NaIO_4$ concentration [mM]	Elution volume [mL]
[Tyr]	B8	-	9.01
[Tyr]	C6	-	14.01
[DOPA]	B3	-	7.01
[DOPA]	B12	-	11.51
[DOPA]	C4	-	13.51
[Tyr]	C5	2	14.01
[DOPA]	B4	2	7.01
[DOPA]	B7	2	8.51
[DOPA]	B11	2	10.51
[DOPA]	C3	2	12.51
[DOPA]	C6	2	14.01
[Tyr]	C6	5	14.01
[DOPA]	B4	5	7.01
[DOPA]	B7	5	8.51
[DOPA]	B11	5	10.51
[DOPA]	C3	5	12.51
[DOPA]	C6	5	14.01

The α- and β-casein[DOPA] samples were then mixed together, and another size exclusion chromatography was performed on these mixed, periodate-treated variants from Figure 25B. There is not much difference between the two curves of the mixed samples when treated with 2 or 5 mM $NaIO_4$ (see Figure 28).

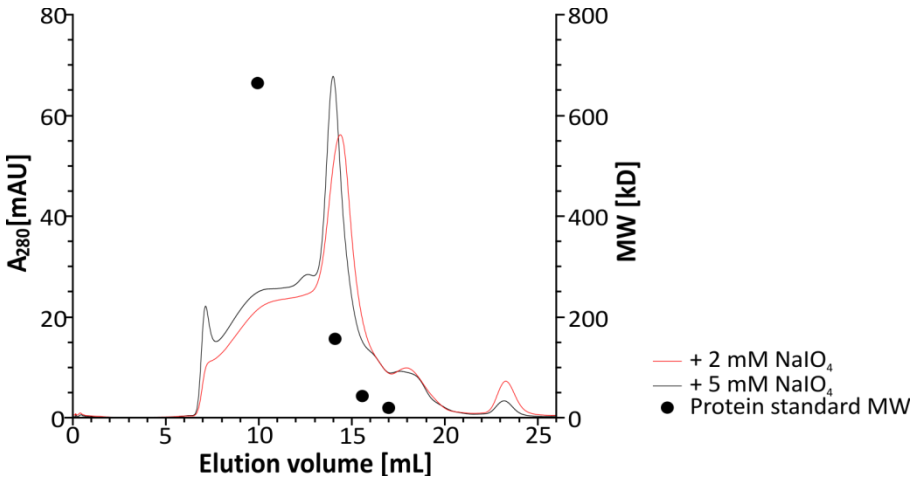

Figure 28. SEC of oxidative cross-linking trials of a fermented α- and β-casein mixture with different concentrations of $NaIO_4$.
Elution behavior of casein[DOPA] variants from a Superdex 200 FPLC column. Red curve, casein[DOPA] mixture treated with 2 mM $NaIO_4$; black curve, casein[DOPA] mixture treated with 5 mM $NaIO_4$; black dots, molecular weight of the protein standard.

To check the SEC peaks from Figure 28, several eluted fractions of the α- and β-casein[DOPA] mixture were taken for an SDS-PAGE analysis (see Table 8).

Table 8. Eluted SEC fractions of the α- and β-casein[DOPA] mixture (see Figure 28) that were taken for an SDS-PAGE (illustrated in Figure 29).

α- and β-casein[DOPA] fraction	$NaIO_4$ concentration [mM]	Elution volume [mL]
B4	2	7.02
B11	2	10.52
B10	5	10.51
B4	5	7.51

The SDS gels of the eluted fractions from SEC of α- and β-casein parent and DOPA variant proteins, as well as the α- and β-casein[DOPA] mixtures, are pictured in Figure 29. As expected, there were ladders of smeared aggregates present for all $NaIO_4$-treated samples. Not only did the periodate-treated casein[DOPA] variants aggregate, there were also aggregates present in the $NaIO_4$-untreated α-casein[DOPA] protein sample. This is to be seen in the B8, B12, and the C4 lanes of panel A in Figure 29.

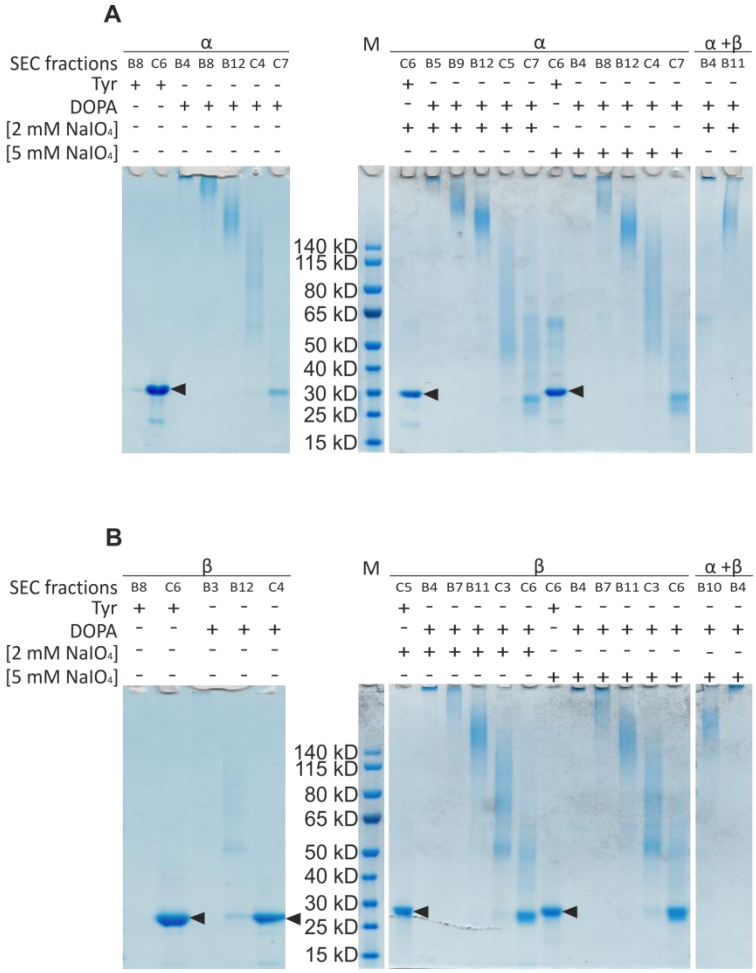

Figure 29. SDS gels of cross-linked α- and β-casein[DOPA] SEC fractions illustrating smears of unspecific aggregates.
Coomassie-stained NuPAGE® 4-12% Bis-Tris gel. The panels are of (A) α-casein[DOPA] and (B) β-casein[DOPA], with the last two lanes being the α- and β-casein[DOPA] mixtures. Irrelevant lanes were omitted. M; molecular weight marker; black arrows, calculated molecular weight of target proteins (24.5 kD for A, and 25 kD for B).

5.4.5 Biomimetic coordination of casein[DOPA] with iron

Another task we wanted to accomplish was to cross-link the casein[DOPA] variants with $FeCl_3$, and herein mimic the coordinating chemistry of the mussel foot proteins. We took 10 mg/mL of each protein and mixed them with 10 and 100 µM $FeCl_3$ and then added 420 µM NaOH to reduce the pH of the buffer solution to 8.0-9.1. After mixing these chemicals with the target proteins on a petri dish, we noticed that the casein[DOPA] variants turned more viscous than their parent proteins. Thereafter inversion of the petri plate demonstrated the viscous behavior of the DOPA variants. These droplets adhered to the plate longer than the parent proteins, which slid rapidly down the dish. A photo series of this can be viewed in Figure S11 and Figure S12.

SDS samples of these coordinated α- and β-casein[DOPA] proteins were also taken for a subsequent SDS-PAGE analysis. These gels can be seen in Figure 30. Both concentrations (10 and 100 µM) of $FeCl_3$ yielded cross-linked casein[DOPA] variants, as can be seen by the white asterisks.

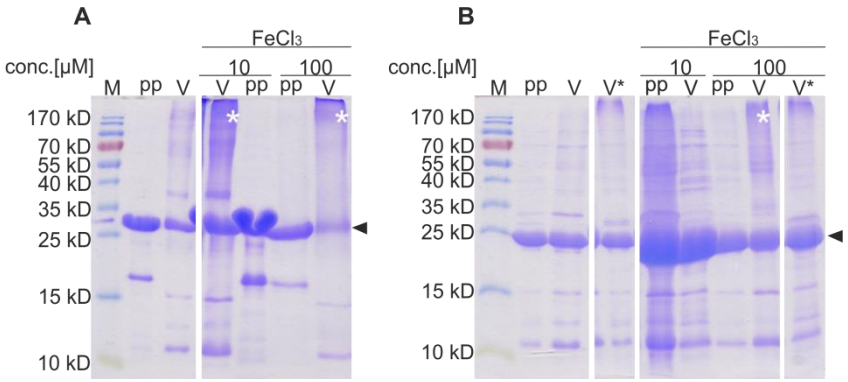

Figure 30. SDS gels of $FeCl_3$-coordinated, cross-linked α- and β-casein[DOPA].
Coordination of casein[DOPA] with different concentrations of $FeCl_3$ depicted on 14% Coomassie-stained SDS gels. Irrelevant lanes were omitted. (A) α-casein[DOPA]; (B) β-casein[DOPA]; M, molecular weight marker; pp, casein[Tyr] parent protein; V, casein[DOPA] variant; V*, α- and β-casein[DOPA] mixture; asterisk, aggregates; black arrows, calculated molecular weight of target proteins (24.5 kD for A, and 25 kD for B). 4 µg of the pp and V samples were loaded onto the gel, whereas about 75 µg of the $FeCl_3$-treated samples were loaded onto each pocket.

6 Discussion

Our idea to use synthetically obtained casein proteins in cross-linking studies to produce a new bioplastic-like material was drawn from Galalith (Siegfeld 1904), in which milk caseins were conjugated upon treatment with formaldehyde. We hypothesized that we could produce bovine α- and β-casein proteins incorporating the tyrosine analog, L-DOPA, and facilitate cross-linking events thereof by implementing oxidative and coordinating chemical reagents to produce a similar material. Cross-linking of caseins via the hydroxyl groups on the aromatic ring of the L-DOPA is crucial in further applying these synthetic proteins to generate a new bioplastic. Proteins that undergo cross-linking treatments have modified mechanical properties (Bastioli 2005) deeming them suitable candidates for biopolymer generation in the material sciences.

Within this study, we demonstrated the successful expression of α- and β-casein[DOPA] proteins. After mixing cross-linking reagents with the target proteins, we were able to detect protein aggregation, which was the main task of this thesis.

6.1 Shake flask cultivation of synthetic casein[DOPA] variants

Shake flask cultivations of *E. coli* BWEC48(TyrΔ){pET26b(+)-α-casein-H_6} and BWEC48(TyrΔ) {pET26b(+)-β-casein-H_6} cells were done in M9-based media containing yeast extract and ascorbic acid. In these experiments, a defined amount of yeast extract served as a tyrosine supplement for this tyrosine auxotrophic strain. After all the tyrosine is used up by the growing cells, the cells enter growth arrest due to the depletion of this cAA. It is at this point that the cultures were supplemented with the tyrosine analog, L-DOPA. Simultaneously, protein expression was induced with the addition of IPTG. Initial SDS-PAGE analyses revealed successful expression of target proteins, as is to be seen in Figure 2. There is no protein expression in the non-induced whole-cell samples, whereas the overnight induced whole-cell SDS samples show distinct bands at the expected molecular weight.

After harvesting and rupturing the cells, the casein[DOPA] target proteins were purified by Ni-chelate chromatography in the presence and absence of ascorbic acid, and re-buffered in PBS. Re-buffering the proteins to PBS was done in order to keep the proteins stable. We chose PBS for its isotonic and non-toxic properties (Green and Sambrook 2012). Since ascorbic acid was present in the cultivation medium as an antioxidant to protect the oxidation of L-DOPA (Weingartner 2014), we wanted to test if its addition to the purification buffers has any effects on the purification of the caseins. Figure 3 indicates no difference in purification of the caseins in the presence of ascorbic acid.

In addition, after determining the concentration of the proteins following their purification under the aforementioned conditions, we obtained less protein when they were purified in the presence of ascorbic acid. The total amount of α-casein[DOPA] protein that was produced was 3.18 mg during Ni-chelate chromatography from 500 mL cultures. Conversely, the α-casein[DOPA] variants purified in the presence of ascorbic acid yielded only 1.9 mg of total protein. Furthermore, we conclude that ascorbic acid has a slight negative effect on the proteins in the purification process. We did not further analyze this effect but concluded that it is sufficient only to add ascorbic acid to the cultivation medium to keep the L-DOPA from oxidizing.

To detect if L-DOPA had been incorporated in the synthetic caseins, we performed a redox-cycling staining of these proteins, which were electroblotted onto PDVF membranes. This simple and easy method is used to specifically detect quinones such as L-DOPA (see section 4.5). To be certain that only DOPA-containing proteins would be visualized, we included the α- and β-casein[Tyr] parent proteins as negative controls. As expected, only the casein[DOPA] variants turned purple, and therefore tested positive (see Figure 4).

6.1.1 Mass analysis

Mass analysis of the $NaIO_4$-treated, DOPA-containing, caseins produced in shake flasks was not successful in proving that we specifically cross-linked these caseins, nor that we incorporated L-DOPA. However, the results did indicate that the proteins we cloned and expressed were indeed caseins. Nevertheless, cross-links could not be detected with this method. LC-MS/MS also has size limitations. It is feasible that we produced aggregates that were too large and therefore could not be detected.

The incorporated L-DOPA could only be detected once, as is listed in Table 5. However, this is not conclusive evidence of its incorporation with this method since less than half of the protein sequence (45.87%) was covered by the analysis. It is probable that after treating the casein[DOPA] variant proteins with $NaIO_4$, cross-linking is indeed facilitated, and therefore the DOPA is not registered by mass analysis.

There is another possibility as to why mass analysis did not yield conclusive results. One of the standard procedures in sample preparation for mass spectrometry is the alkylation of the protein with iodoacetamide. This reaction is common in proteomics for sulfhydryl group determination in proteins. However, the α- and β-caseins do not contain any sulfhydryl groups. Therefore, this alkylating agent could have reacted with the N-terminal amines and the C-terminal carboxylic acid groups, as well as protic functional groups of peripheral amino acids, ultimately blocking the L-DOPA residue for detection (Yang and Attygalle 2007).

Additionally, since the proteins provided for LC-MS/MS analysis were in-gel samples, it is feasible that gel-induced modifications could have occurred. There have been reports that proteins can be modified under electrophoretic conditions, in which the acrylamide in the gel matrix reacts covalently with hydroxyl-compounds, such as serines. Due to these protein-acrylamide adducts, the actual molecular weight is in turn difficult to determine (Jeannot, Zheng, and Li 1999). Therefore, we speculate that all, or at least the majority, of the tyrosine positions are indeed occupied by L-DOPA molecules and that these residues interact with the polyacrylamide gel matrix. This is a plausible explanation for undetected L-DOPA.

We conclude that standard in-gel mass analysis is not the optimal approach to detecting cross-linked casein[DOPA] proteins. Alternatively, the mass analysis of the intact casein variants might be more successful. We tried to analyze the proteins produced in shake flasks in this way but to no success, most probably because of insufficient protein amounts. The analysis will be repeated with higher concentrated protein samples. Nevertheless, a simple redox-cycling staining test proved, however, that we did manage to incorporate L-DOPA into the synthetic caseins. Raman spectroscopy may be an even more suitable method for quantitative aggregate detection (Choi, Huh, and Erickson 2011).

6.2 High cell density fermentation of caseins

6.2.1 Fermentation of casein[Tyr] parent proteins

We decided upon culturing the casein-producing *E. coli* cells in stirred tanks in order to produce higher cell densities than in shake flasks. Fermenting the *E. coli* BL21(DE3)Gold cells harboring the pET26b(+)-α-, and β-casein[Tyr] proteins was successful as described in section 4.6.1.

Proteins were purified from 50 mL of the fermented cultures, from which the protein concentration was determined to be 24 mg and 1.5 mg for α-, and β-casein[Tyr], respectively. If we use these concentrations to back-calculate the protein amount produced in the ~660 mL cell cultures from the bioreactors, we obtain 192 mg and 12 mg of α- and β-casein[Tyr], respectively. Nevertheless, we did not purify the entire fermented cultures of the parent proteins. With the purification of 50 mL cell cultures, rather than the entire culture volume, we wanted to assess if the protein expression of the synthetic parent casein[Tyr] proteins was indeed possible in stirred tanks. Initial expression studies in shake flasks of the parent caseins yielded adequate amounts of protein after back calculating to 50 mL, with 2.86 mg and 1.22 mg of α- and β-casein[Tyr], respectively (Pajazetovic 2014). Thus, we expected even more proteins to be produced under the optimal conditions within the DASGIP bioreactors, and therefore we purified less culture volume to avoid overloading the nickel resin during Ni-chelate chromatography. Nevertheless, the immense difference between these protein amounts is most probably due to overloading the column-resins with crude lysate during the gravity flow purification, even though the manufacturer states that a protein yield of up to 100 mg per mL resin is possible. It is also possible that we did not use enough resin during the purification of the casein[Tyr] proteins. Moreover, it cannot be excluded that the purification has different effects on each protein. We cannot assume, however, that this is indeed true after comparing the lysate fractions of both α- and β-casein parent proteins in Figure 7, where β-casein is evidently expressed to a much lesser degree than α-casein. Inasmuch, we still deemed the expression of the parent proteins to be successful in bioreactors, as well as the purification thereof. This was an indication for us that the expression of caseins incorporating L-DOPA is feasible in bioreactors.

6.2.2 Production of casein[DOPA] variants with L-DOPA supplementation

After cross-linking trials of casein[DOPA] variants (see sections 5.4.1 and 5.4.2), we realized that we need much more of the synthetic proteins in order to conduct further experiments. Thus, we turned to DASGIP bioreactors in hopes of up-scaling the protein production of the DOPA variants.

As described in the above 4.6.2 section, the incorporation of L-DOPA into the synthetic caseins was done using the SPI method to cultivate *E. coli* BWEC48(TyrΔ){pET26b(+)-α-casein-H_6} and BWEC48(TyrΔ){pET26b(+)-β-casein-H_6} cells in the presence of L-DOPA. As expected, the cell wet weight of the harvested casein[DOPA]-containing cells was approximately half of the fermented casein[Tyr] parent proteins.

After rupturing the cells and purifying the casein[DOPA] variants from the crude lysate, we calculated the total casein[DOPA] produced in ~770 mL cultures to be 34.66 mg and 34.30 mg for α- and β-casein[DOPA], respectively. With these protein amounts we proceeded to compare the expression of target proteins under different conditions from shake flasks and bior eactors, as well as conduct cross-linking trials.

6.3 Shake flask production compared to bioreactor production of synthetic casein

Caseins have naturally adhering properties making them aggregate. This is a rather problematic issue that we had to overcome, considering our purpose was to specifically agglomerate the casein[DOPA] variant proteins via the two L-DOPA hydroxyl groups upon the addition of oxidative or coordinating agents. There are publications in which the authors describe that urea can be used as a solution additive to deter protein aggregation (Chen, Sagle, and Cremer 2007). This mechanism is not yet fully understood, although there are two major proposed models: the direct and indirect mechanism. Urea binds to proteins through electrostatic interactions or hydrogen bonding, according to the direct mechanism. In contrast, the indirect mechanism is based on the hypothesis that urea weakens hydrophobic interactions by disturbing water structure. In this manner, the hydrophobic protein residues become better solvated (Chen et al. 2007). There is literature that describes adding particularly 3.75 M urea to buffers when purifying casein proteins (Chaplin 1986). We decided to follow suit, and use the same concentration of urea in purifying the synthetic caseins.

Size exclusion chromatography of the α-casein parent and α-casein[DOPA] variant proteins, eluted in the presence and absence of urea and DTT, confirms that these agents help deter protein aggregation, as can be seen in the insoluble protein purification fractions (panels A and B) of Figure 15. These findings are in accordance with previous studies (Ayyadurai et al. 2011, Umanah et al. 2009, Burdine et al. 2004). The elution pattern of casein[Tyr] parent proteins is fore most affected. Casein[Tyr] elutes much earlier from the Superdex 200 FPLC column when urea is not present in the PBS elution buffer. This is an indication of larger proteins; hence, aggregates. By contrast, when the PBS elution buffer contains urea, the protein elutes later, indicating a smaller molecular weight. Notably, the α-casein[DOPA] synthetic variant has the same elution pattern when eluted with and without urea, meaning that urea does not affect the elution behavior of the synthetic DOPA-containing casein. These observations lead us to conclude that the casein[DOPA] variants are most likely inert to urea.

To further scrutinize the effect of urea on caseins, one gram of fermented synthetic casein cell pellets was taken to juxtapose the proteins purified in the presence of urea and DTT in the purification buffers to those purified without these agents. The SDS gels in Figure 15 reveal that urea and DTT have an impact mostly on the casein[Tyr] parent proteins. Urea solubilizes caseins, which is demonstrated by less target proteins in the pellet (insoluble) fractions on the gels. The purified proteins depicted in panels B, C, and D in Figure 15 are not visible, because these protein samples were not concentrated before the SDS samples were taken. Therefore, the protein amounts loaded onto SDS gels were too dilute. Nevertheless, we were able to confirm that having urea in the buffers during purification is beneficial for the proteins.

We also compared the purified proteins obtained from shake flasks to those produced in bioreactors on SDS gels. It seems that more caseins are produced in shake flasks than in bioreactors after an SDS-PAGE was performed (see Figure 16). However, this cannot be the case. The lack of visible bands on the aforementioned gels could be due to a problematic Coomassie blue staining since the same amount of protein was loaded onto the gels. The protein detection may have failed due to the incorporation of L-DOPA into the caseins. The Coomassie dye binds to basic and aromatic amino acids (Congdon, Muth, and Splittgerber 1993), and if many L-DOPA residues have been incorporated,

it might be that they are interfering in the detection of the casein[DOPA] proteins. Moreover, these SDS gels are conflicting with the calculated total protein amount obtained from the fermented cells. Previously, we ascertained that the *E. coli* cells cultivated in bioreactors produced tenfold more casein[DOPA] proteins than those cultivated in shake flasks. For instance, shake flask cultivation of *E. coli* α-casein[DOPA]-containing cells yielded 3.18 mg of total protein (see section 5.1), whereas cells cultivated in a bioreactor produced a total of 34.66 mg of the same casein variant (see section 5.2.2). Therefore, the SDS gels depicted in Figure 16 most probably do not reflect the real situation.

6.4 Bioconjugation trials of casein[DOPA] variants from shake flasks

Oxidative cross-linking of casein[DOPA] proteins can be done with sodium periodate, $NaIO_4$. Only the hydroxyl groups of L-DOPA can selectively be oxidized by this reagent since most amino acids within the caseins are inert to it. An intermediate, ortho-quinone, would be expected within this reaction, which forms stable cross-links after a nucleophilic attack by sodium periodate (Burdine et al. 2004). Therefore, we added different concentrations of $NaIO_4$ to the casein[DOPA] variants to facilitate cross-linked casein proteins via the hydroxyl groups on the incorporated L-DOPA ncAA. We already determined that the addition of ascorbic acid hinders cross-linking of the caseins, and that microdialysis is not the best method to achieve aggregates, most likely due to insufficient amounts of proteins that were purified from a shake flask cultivation (see Figure S7). However, microdialysis experiments were only performed with the proteins obtained from shake flasks. We cannot definitely conclude that microdialysis is a poor method for conjugating proteins, since we did not repeat this experiment with proteins obtained from bioreactors.

Initial cross-linking attempts of α-casein[DOPA] were not definitive. The first size exclusion chromatogram of $NaIO_4$-treated casein[DOPA] does not show an elution curve indicative of cross-linked caseins (see Figure 18). This is contradicting data to that of Figure 17B, where we see slight smeared bands that migrated in higher molecular weight regions on the stained, electroblotted PDVF membrane containing $NaIO_4$-treated α-casein[DOPA] proteins. We assume that the protein amounts produced in shake flasks were inadequate. If there are few proteins, establishing cross-linking thereof is improbable due to the physical distance between the proteins in solution.

The same was done for the β-casein[DOPA] protein produced in shake flasks. However, treating this protein solution with $NaIO_4$ resulted in detectable molecular weight variation, according to SEC and subsequent SDS-PAGE analyses (see Figure 19, Figure 20, and Figure 21). These data were the first to support positive cross-linking behavior of casein[DOPA] variants. In-gel samples of the SEC shoulders (Figure 19) eluting earlier than the main peak, which corresponds to the casein monomer, were sent for mass analysis at the ACIB Core Facility Metabolomics. It was verified that these samples indeed contained casein proteins, upon which one DOPA modification was found (see Table 5). However, the sequence coverage was too low to determine the degree of L-DOPA incorporation, and inasmuch neither could any crosslinks be detected. The mass analysis of intact L-DOPA variant proteins most probably could yield more conclusive results.

6.5 Oxidative cross-linking of casein[DOPA] protein variants from bioreactors

Oxidative cross-linking of fermented casein[DOPA] variants was carried out in the same manner as the bioconjugation of the DOPA variants obtained from shake flask cultivations. However, as we had

more protein available for analysis, we tested different concentrations of $NaIO_4$ (2 and 5 mM). An SDS-PAGE analysis also revealed smears in the $NaIO_4$-treated samples, whereas the casein[Tyr] parent proteins remained monomeric after the same treatment. Size exclusion chromatography of all $NaIO_4$-treated (see Figure 26, Figure 27, and Figure 28) and untreated samples (see Figure S10) are in accordance with what SDS gels in Figure 29 show. After SEC, we determined that both periodate concentrations had the same effect on the synthetic casein variants. The elution profiles were very similar for all samples treated with both concentrations of $NaIO_4$. We, therefore, conclude that lower concentrations of $NaIO_4$ are sufficient for cross-linking casein[DOPA] variants. It is crucial to have sufficient protein available for cross-linking events to occur. Within the SEC profiles, we also observed higher signals at A_{280}, which is indicative that we loaded much more casein[DOPA] proteins onto the Superdex column. This was only possible after producing the casein[DOPA] protein variants in bioreactor cultures.

Another SDS-PAGE analysis was performed taking particular SEC elution fractions (listed in Table 6, Table 7, and Table 8). The NuPAGE® gels (see Figure 29) reveal what we expected. The periodate-untreated α- and β-casein[Tyr] proteins are monomeric and elute at ~14 mL, implying that the proteins have indeed the calculated molecular weights, 24.5 kD and 25 kD for α- and β-casein[Tyr], respectively.

However, the $NaIO_4$-untreated α- and β-casein[DOPA] variants seem to aggregate even without adding the oxidative reagent. The initial SDS-PAGE analysis of concentrated and untreated α-casein[DOPA] also revealed a high ranging protein band (see Figure 10A, lane V). We stipulate that since these are concentrated samples, the incorporated L-DOPA molecules undergo auto-oxidation. It has been reported that catechols are prone to spontaneous oxidation in the presence of air. The loss of two electrons, and therefore two hydrogens, leads to the formation of o-quinone (Yang et al. 2014). This may be a reason as to why there is a broad peak eluting much earlier than the monomeric peak for the α-casein[DOPA] sample in Figure S10A, and high ranging smears of protein on the NuPAGE® gel in Figure 29A. The SEC and migration behavior of the caseins on SDS gels correlate with the initial SDS-PAGE analysis.

6.6 Coordination of DOPA variant proteins with iron

An emerging versatile cross-linking mechanism in biological systems is metal complexation. Therefore, one of our tasks entailed specifically introducing coordination-based cross-links of casein[DOPA] proteins with iron, mimicking the adhesiveness that mussel foot proteins exhibit. The adhesiveness principle of mussel proteins relies on coordination chemistry between the L-DOPA residues and iron that occurs in marine environments (Silverman and Roberto 2007). Thus, we tried to generate bis- and tris-catechol-Fe^{3+} complexes by adding $FeCl_3$ to the casein[DOPA] proteins to cross-link the synthetic casein[DOPA] proteins, and therefore mimic the same process that mussels undergo naturally.

We had to overcome some challenges within this experiment. Iron chloride is soluble only at a low pH levels. Nevertheless, to facilitate coordinated cross-linking, the pH of the protein solution has to be above 9.1 to obtain tris-species (Holten-Andersen et al. 2011). It has been proposed that mussels cope with this problem by pre-binding Fe^{3+} in mono-L-DOPA-Fe^{3+} complexes by the mussel foot protein (mfp-1) at pH 5 in secretory granules. After secretion of the mono-species into seawater,

spontaneous physical cross-linking occurs via bis- and tris-L-DOPA-Fe^{3+} complex formation due to the pH elevation (Yang et al. 2014). Although Fe^{3+} precipitates at alkaline pH levels, we tested different concentrations of sodium hydroxide that had to be added in order to link the casein[DOPA] proteins in a basic environment without precipitating the Fe^{3+}. Adding 420 µM of NaOH to the protein mixture already including either 10 or 100 µM $FeCl_3$ is sufficient for linking the DOPA-containing proteins. We then observed more viscous behavior of the casein[DOPA] proteins than of the casein[Tyr] parent proteins, which were treated under the same conditions (see Figure S11 and Figure S12).

SDS gels in Figure 30 also illustrate positive cross-linking in the presence of $FeCl_3$. Only the DOPA-containing variants have a smear in the high molecular weight range on the gel, indicating different species of linked proteins. Contrary to what Zeng *et al.* (2010) reported, we observed a better cross-linking effect of the DOPA-proteins with 100 µM $FeCl_3$ compared to our cross-linking trials using 10 µM $FeCl_3$. Our observations are, however, in accordance to what Holten-Andersen *et al*. reported about the cross-linking mechanism being pH-controlled (Holten-Andersen et al. 2011). After elevating the pH to > 9.1, we found coordination reactions to be occurring between *in vivo* generated L-DOPA-containing caseins and Fe^{3+}. Under these pH conditions, we also observed mild precipitation of the $FeCl_3$, presumably due to the alkaline conditions. Nevertheless, we only base our conclusions on the SDS-PAGE analysis that we performed. Further investigation is still needed to quantitatively confirm these results.

7 Conclusions

The expression of synthetic caseins and the subsequent incorporation of L-DOPA into these proteins was validated within this work. Oxidative bioconjugation of the L-DOPA-containing caseins, via sodium periodate, was confirmed, as well as the cross-linking thereof by iron metal complexation. We found that it is crucial, however, to have high amounts of proteins, which in our instance could only be obtained from bioreactor fermentations. Nevertheless, more experimental data is needed to ascertain if the iron-mediated coordination of casein[DOPA] proteins is a more efficient method to link these proteins than the covalent oxidative cross-linking via $NaIO_4$. Nonetheless, considering the progress within this work, we have come at least one step closer to producing the bioplastics for which we strive.

8 References

Anonymous. 1913. "The Casein Industry." *The Journal of Industrial and Engineering Chemistry* 79–80.

Ayyadurai, Niraikulam et al. 2011. "Bioconjugation of L-3,4-Dihydroxyphenylalanine Containing Protein with a Polysaccharide." *Bioconjugate chemistry* 22:551–55.

Babu, Ramesh P., Kevin O'Connor, and Ramakrishna Seeram. 2013. "Current Progress on Bio-Based Polymers and Their Future Trends." *Progress in Biomaterials* 2(1):8.

Bastioli, Catia, ed. 2005. *Handbook of Biodegradable Polymers*. Rapra Technology Limited.

Burdine, Lyle, Thomas G. Gillette, Hai-Jun Lin, and Thomas Kodadek. 2004. "Periodate-Triggered Cross-Linking of DOPA-Containing Peptide-Protein Complexes." *Journal of the American Chemical Society* 126(37):11442–43.

Chaplin, L. C. 1986. "Hydrophobic Interaction Fast Protein Liquid Chromatography of Milk Proteins." *Journal of chromatography* 363(2):329–35.

Chen, Xin, Laura B. Sagle, and Paul S. Cremer. 2007. "Urea Orientation at Protein Surfaces." *Journal of the American Chemical Society* 129(49):15104–5.

Choi, Inhee, Yun Suk Huh, and David Erickson. 2011. "Size-Selective Concentration and Label-Free Characterization of Protein Aggregates Using a Raman Active Nanofluidic Device." *Lab on a chip* 11(4):632–38.

Congdon, R. W., G. W. Muth, and A. G. Splittgerber. 1993. "The Binding Interaction of Coomassie Blue with Proteins." *Analytical biochemistry* 213(2):407–13.

Eriksen, Marcus et al. 2014. "Plastic Pollution in the World's Oceans: More than 5 Trillion Plastic Pieces Weighing over 250,000 Tons Afloat at Sea." *PLoS ONE* 1–15.

Gervet, Bruno. 2007. "The Use of Crude Oil in Plastic Making Contributes To Global Warming." *Renewable Energy* (May).

Gibson, Daniel G. et al. 2009. "Enzymatic Assembly of DNA Molecules up to Several Hundred Kilobases." 6(5).

Green, Michael R., and Joseph Sambrook. 2012. *Molecular Cloning: A Laboratory Manual*. 4th ed. Cold Spring Harbor Laboratory Press.

Haid, A., and M. Suissa. 1983. "Immunochemical Identification of Membrane Proteins after Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis." *Methods in enzymology* 96:192–205.

Harrington, Matthew J., Admir Masic, Niels Holten-Andersen, J. Herbert Waite, and Peter Fratzl. 2010. "Iron-Clad Fibers: A Metal-Based Biological Strategy for Hard Flexible Coatings." *Science (New York, N.Y.)* 328(5975):216–20.

Holten-Andersen, Niels et al. 2011. "pH-Induced Metal-Ligand Cross-Links Inspired by Mussel Yield Self-Healing Polymer Networks with near-Covalent Elastic Moduli." *Proceedings of the National Academy of Sciences of the United States of America* 108(7):2651–55.

Hunt, S. 1985. "The Non-Protein Amino Acids." Pp. 55–137 in *Chemistry and Biochemistry of the amino acids*, edited by G. C. Barrett. Dordrecht: Springer Netherlands.

Jeannot, Michael a., Jing Zheng, and Liang Li. 1999. "Observation of Sodium Gel-Induced Protein Modifications in Dodecylsulfate Polyacrylamide Gel Electrophoresis and Its Implications for Accurate Molecular Weight Determination of Gel-Separated Proteins by Matrix-Assisted Laser Desorption Ionization Time-of-." *Journal of the American Society for Mass Spectrometry* 10(6):512–20.

König, Harald, Daniel Frank, Reinhard Heil, and Christopher Coenen. 2013. "Synthetic Genomics and Synthetic Biology Applications between Hopes and Concerns." *Current genomics* 14(1):11–24.

Krätz. 2004. "S T E I N AU S M I LC H." *Chem. Unserer Zeit* 2(38):133–37.

Laemmli, U. K. 1970. "Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4." *Nature* 227(5259):680–85.

Link, A. James, Marissa L. Mock, and David A. Tirrell. 2003. "Non-Canonical Amino Acids in Protein Engineering." *Current Opinion in Biotechnology* 14(6):603–9.

Mekonnen, Tizazu, Paolo Mussone, Hamdy Khalil, and David Bressler. 2013. "Progress in Bio-Based Plastics and Plasticizing Modifications." *Journal of Materials Chemistry A* 1(43):13379.

Pajazetovic, Teila. 2014. *A New Approach on Bioplastics : Cloning and Expression of Bovine Caseins in E . Coli*.

Pathak, Swati, C. L. R. Sneha, and Blessy Baby Mathew. 2014. "Bioplastics : Its Timeline Based Scenario & Challenges." 2(4):84–90.

Pazs, Mercedes A., Rudolf Fliickigers, Andra Boakll, Herbert M. Kaganll, and Paul M. Gallop. 1991. "Quinoproteins by Redox-Cycling." (19):0–3.

Qiagen. 2003. *The QIA Expressionist*. 5th ed. Retrieved (http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:The+QIA+expressionist+?#0\nhttp://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:The+QIA+expressionist#0).

Rodgers, Kenneth J., Peter M. Hume, Rachael a. Dunlop, and Roger T. Dean. 2004. "Biosynthesis and Turnover of DOPA-Containing Proteins by Human Cells." *Free Radical Biology and Medicine* 37(11):1756–64.

Seo, Jawon, and Kong-Joo Lee. 2004. "Post-Translational Modifications and Their Biological Functions: Proteomic Analysis and Systematic Approaches." *Journal of biochemistry and molecular biology* 37(1):35–44.

Siegfeld, M. 1904. "Galalith." *Zeitschrift für Angewandte Chemie* (48):1816–18.

Silverman, Heather G., and Francisco F. Roberto. 2007. "Understanding Marine Mussel Adhesion." *Marine Biotechnology* 9(6):661–81.

Teuten, Emma L., Steven J. Rowland, Tamara S. Galloway, and Richard C. Thompson. 2007. "Potential for Plastics to Transport Hydrophobic Contaminants." *Environmental Science and Technology* 41(22):7759–64.

Umanah, George K. E., Cagdas Son, FaXiang Ding, Fred Naider, and Jeffrey M. Becker. 2009. "Cross-Linking of a DOPA-Containing Peptide Ligand into Its G Protein-Coupled Receptor." *Biochemistry* 48(9):2033–44.

Voigt, Christopher a. 2012. "Synthetic Biology." *ACS Synthetic Biology* 1(1):1–2.

Voloshchuk, Natalya, and Jin Kim Montclare. 2010. "Incorporation of Unnatural Amino Acids for Synthetic Biology." *Molecular bioSystems* 6(1):65–80.

Weingartner, Alexandra. 2014. *New Approach on Bioplastics: Incorporation of L-DOPA into Bovine Casein*.

Wiltschi, Birgit. 2012. "Expressed Protein Modifications: Making Synthetic Proteins." Pp. 211–25 in *Synthetic Gene Networks*, vol. 813, *Methods in Molecular Biology*, edited by Wilfried Weber and Martin Fussenegger. Totowa, NJ: Humana Press.

Yang, Juan, Martien a. Cohen Stuart, and Marleen Kamperman. 2014. "Jack of All Trades: Versatile Catechol Crosslinking Mechanisms." *Chem. Soc. Rev.* 43(24):8271–98.

Yang, Zhihua, and Athula B. Attygalle. 2007. "LC/MS Characterization of Undesired Products Formed during Iodoacetamide Derivatization of Sulfhydryl Groups of Peptides." *Journal of mass spectrometry* 42:233–43.

Yu, Miaoer, Jungyeon Hwang, and Timothy J. Demin. 1999. "Role of L-3,4-Dihydroxyphenylalanine in Mussel Adhesive Proteins." *J. Am. Chem. Soc.* 121:8825–5826.

Zeng, Hongbo, Dong Soo Hwang, Jacob N. Israelachvili, and J. Herbert Waite. 2010. "Strong Reversible Fe3+-Mediated Bridging between Dopa-Containing Protein Films in Water." *Proceedings of the National Academy of Sciences of the United States of America* 107(29):12850–53.

Zheng, Shun, and Inchan Kwon. 2012. "Manipulation of Enzyme Properties by Noncanonical Amino Acid Incorporation." *Biotechnology Journal* 7(1):47–60.

9 Supporting Material

9.1 Cloning of caseins

The following data were taken from a previous laboratory project report (Pajazetovic 2014).

Table S1. The protein sequences of the caseins used in this study.
Hexa-histidine tags are in bold and the tyrosines are in red. α-casein contains 10 tyrosine amino acids, whereas β-casein contains 4.

Casein	Amino acid sequence
α-S1 casein [*E. coli*$_{opt}$, $N_{truncated}$, C_{term}-H_6]	MRPKHPIKHQGLPQEVLNENLLRFFVAPFPEVFGKEKVNELSKDIGSESTEDQAMEDIKQMEAESISSSEE IVPNSVEQKHIQKEDVPSERYLGYLEQLLRLKKYKVPQLEIVPNSAEERLHSMKEGIHAQQKEPMIGVNQE LAYFYPELFRQFYQLDAYPSGAWYYVPWGTQYTDAPSFSDIPNPIGSENSEKTTMPLWGS**HHHHHH**
β-casein [*E. coli*$_{opt}$, $N_{truncated}$, C_{term}-H_6]	MRELEELNVPGEIVESLSSSEESITRINKKIEKFQSEEQQQTEDELQDKIHPFAQTQSLVYPFPGPIPNSLPQ NIPPLTQTPVVVPPFLQPEVMGVSKVKEAMAPKQKEMPFPKYPVEPFTESQSLTLTDVENLHLPLPLLQS WMHQPHQPLPPTVMFPPQSVLSLSQSKVLPVPQKAVPYPQRDMPIQAFLLYQEPVLGPVRGPFPIIVGS **HHHHHH**

Table S2.Primers used for cloning caseins into a pET26b(+) vector.
Homology regions are underlined, restriction sites are italicized and the coding sequence for the hexa-histidine-tag is in bold.

Primer	Sequence (5' → 3')	Gene	Primer description
pBP760	AGAAA*TAATTT*TGT*TTAA*CT*TTAA*GAAGGA*GATATACATATG*CG*AC CTAAACATCC*TATC	α-casein	Forward
pBP761	AGAAAT*AATTT*TGT*TTAA*CT*TTAAGAA*GGAGATATA*CATATG*AGAG AGCT*GGAAGAACTG	β-casein	Forward
pBP763	TGGTGGTG*GTGCTCGAGTGCGGCCGCAAGCTT*CACCACAGTG*GCA TAGTAGTCTTTTCAC	α-casein	Reverse
pBP764	TGGTGGTGGT*GCTCGAGTGCGGCCGCAAGCTT*CAGACAATAATAG GGAA*GGGTCCC*	β-casein	Reverse
pBP766	**TGGTGGTGGTGGTG*GTG****CTCGAGTGCGGCCGCAAGCTT*CATTAGT GATGGTGATGG	Casein	Casein-histidine-tag reverse

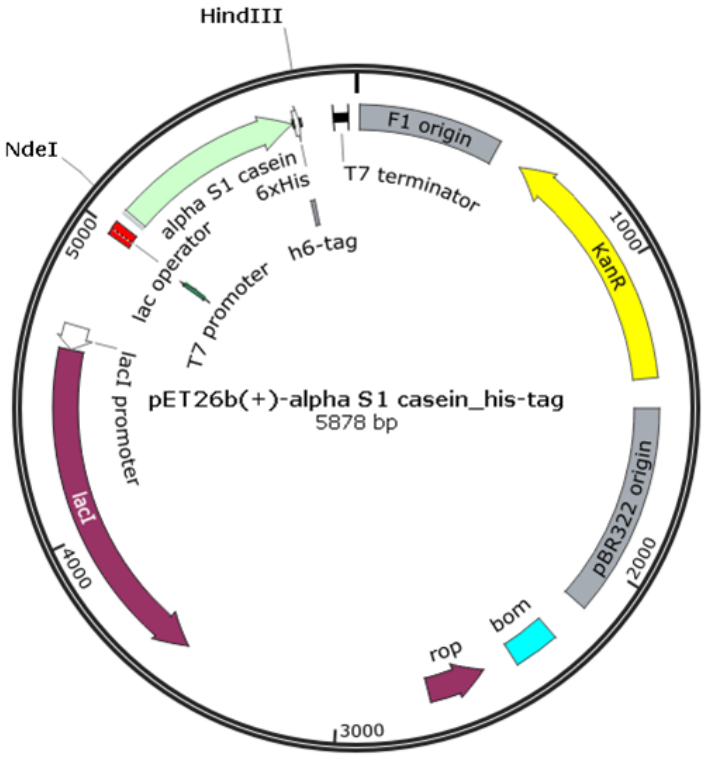

Figure S1. Vector map of pET26b(+) containing the α-casein-H_6 gene of interest.
The important features of this vector are the T7 promoter, the kanamycin resistance marker, and the *lac*I that makes the expression of the casein inducible with IPTG.

9.2 Expression and purification fractions of caseins

This image was taken from a previous laboratory project report (Pajazetovic 2014).

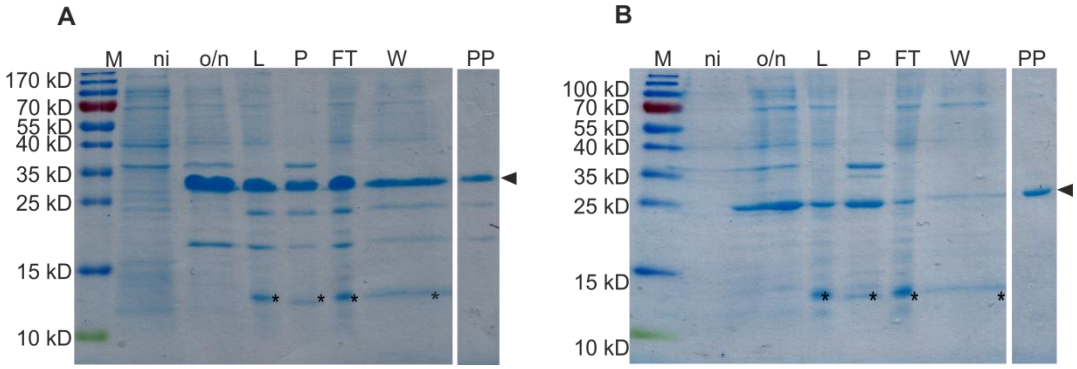

Figure S2. Initial expression and purification of α- and β-caseins.
14% Coomassie-stained SDS gels. (A) α-casein, irrelevant lanes were omitted; (B) β-casein, irrelevant lanes were omitted; M, molecular weight marker; ni, non-induced cells; o/n, overnight induction; L, soluble protein fraction; P, insoluble protein fraction; FT, column flow-through; W, wash fraction; PP, purified casein protein re-buffered in PBS; asterisk, lysozyme; black arrows, calculated molecular weight of target proteins (24.5 kD for A, and 25 kD for B).

9.3 Initial L-DOPA experiments

The following data were taken from a previous laboratory project report (Weingartner 2014).

9.3.1 L-DOPA oxidation

Figure S3. Oxidation of L-DOPA in medium with an *E. coli* culture.
Flasks containing M9 medium and L-DOPA without ascorbic acid (cultures 1-3), and with ascorbic acid (cultures 4-6). Images were taken shortly after inoculation (t=0h, flasks 1 and 4), after 8 hours (t=8h, flasks 2 and 5) and overnight (t=o/n, flasks 3 and 6). Oxidation of L-DOPA is visible in culture 3. The orange color of cultures 5 and 6 is most likely caused by oxidation of ascorbic acid itself. Thomas Stanzer is credited for taking these pictures.

9.3.2 Yeast titration

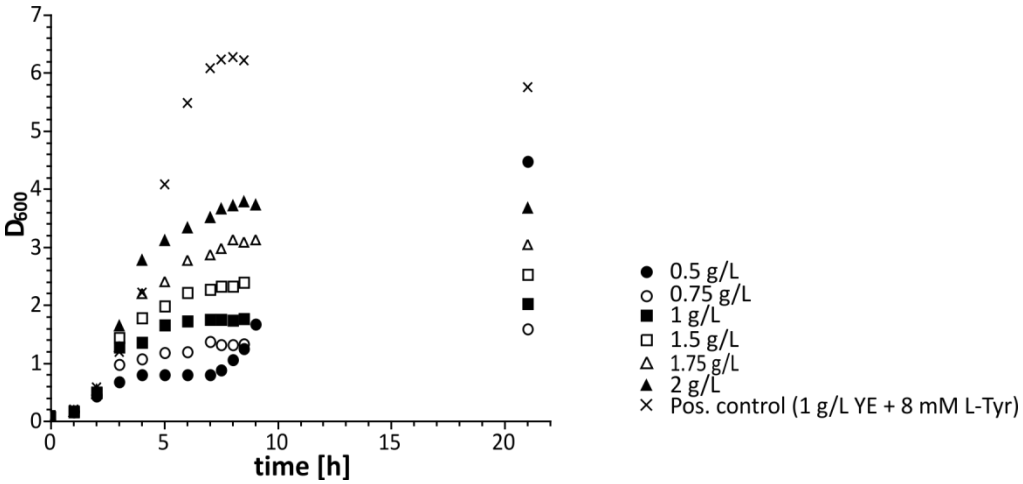

Figure S4. Yeast extract titration of the *E. coli* BWEC48(TyrΔ) strain.
Different cultures of the tyrosine auxothrophic BWEC48(TyrΔ) strain, harboring an inducible pET26b(+)-β-casein expression plasmid, were supplemented with 0.5 g/L (filled circles), 0.75 g/L (open circles), 1 g/L (filled squares), 1.5 g/L (open squares), 1.75 g/L (open triangles), and 2 g/L (filled triangles) yeast extract. As a positive control, a culture of the same strain supplemented with 1 g/L yeast extract and 8 mM Tyr was used (crosses). 8 mM of L-Tyr was added to the culture supplemented with 0.5 g/L yeast extract after growth arrest (t=~7 h).

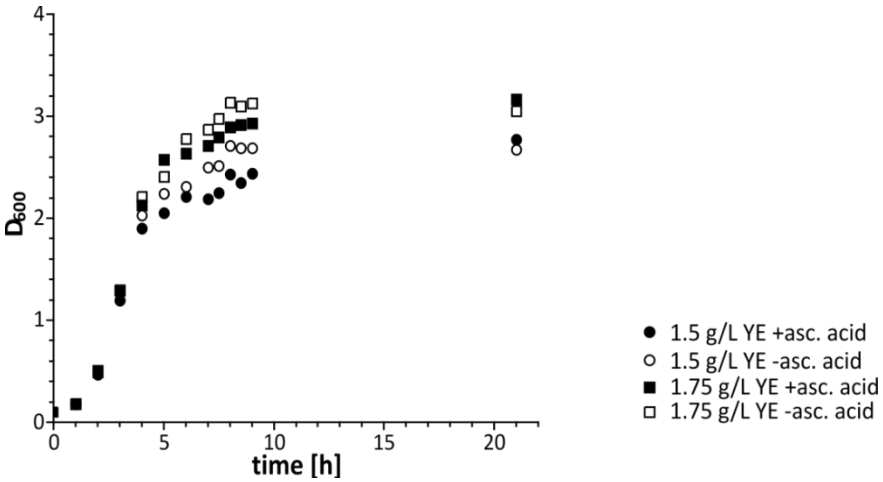

Figure S5. Influence of 10 mM ascorbic acid on the growth of the *E. coli* BWEC48(TyrΔ) strain.
The influence of ascorbic acid on the growth of the *E. coli* strain BWEC48(TyrΔ) containing the expression plasmid pET26b(+)-β-casein was tested. Different yeast extract concentrations were used: 1.5 g/L with (filled circles) and without ascorbic acid (open circles) and 1.75 g/L with (filled squares) and without ascorbic acid (open squares).

9.4 Media used

Table S3. Composition of the M9 salt stock.

Component	Amount [g/L]
Na_2HPO_4	33.9
KH_2PO_4	15
NaCl	2.5
NH_4Cl	5

Table S4. M9 medium components for shake flask culture cultivation.

Component	Stock concentration	Final concentration	$V_{stock\ [mL]}$/500 mL
M9 salts	5 X	1 X	100
$MgSO_4$	1 M	5 mM	0.5
Ascorbic acid	1 M	10 mM	5
$CaCl_2$	1 mg/mL	1 µg/mL	0.5
Glucose	1 M	20 mM	60
Trace elements	~18.7 X	5 X	0.18
Kanamycin	1000 X	1 X	0.5
Yeast extract	-	5.25 g	5.25 g
ddH_2O[1]	-		Up to 500 mL

Table S5. Composition of M9 fermentation medium for the bioreactor.

Component	Stock concentration	Final concentration	$V_{stock\ [mL]}$/400 mL
M9 salts[2]	5 X	1 X	80
$MgSO_4$[2]	1 M	5 mM	2
Ascorbic acid[4]	1 M	10 mM	4
Ca^{2+}/citrate[2]	1 mg/mL/100 mg/mL	5 µg/mL/500 mg/mL	2
Glucose[2]	1 M	125 mM	50
Trace elements[2]	~18.7 X	5 X	0.12
Kanamycin[2]	1000 X	1 X	0.4
Yeast extract[1,5,6]	-	7 g	7 g
Antifoam[2,3]	-	1+9	1
ddH_2O[1]	-	-	Up to 400 mL

[1] 7 g yeast extract were dissolved in ddH_2O and autoclaved (for the fermentation of casein[Tyr]-containing *E. coli* cells).
[2] All other ingredients were supplemented afterwards as sterile stock solutions.
[3] Ca^{2+} and antifoam were added under stirring.
[4] Ascorbic acid was added only in the L-DOPA supplemented casein variant fermentation
[5] For the L-DOPA supplemented casein variant fermentation, 4.9 g of yeast extract were used.
[6] Yeast extract amounts supplied to the M9 medium differed for the fermentation of *E. coli* cells containing casein[Tyr] parent caseins and casein[DOPA] variant proteins, because these amounts were calculated for the desired final D_{600} of each culture.

9.5 Process data of synthetic casein fermentation

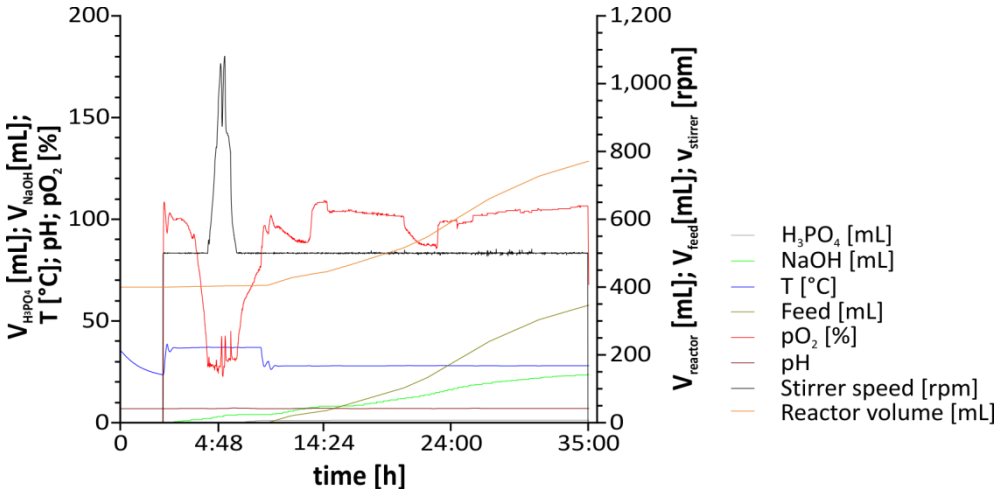

Figure S6. DASGIP stirred tank reactor process data of α-casein[DOPA].
The parameters followed during the fermentation (reactor 1) were the temperature (blue curve), pH (brown curve), stirrer speed (black curve), oxygen partial pressure (red curve), reactor volume (orange curve), and the feed volume (dark green curve). The addition of acid (H_3PO_4, gray curve) and base (NaOH, green curve) were also recorded.

9.6 Oxidative bioconjugation trials

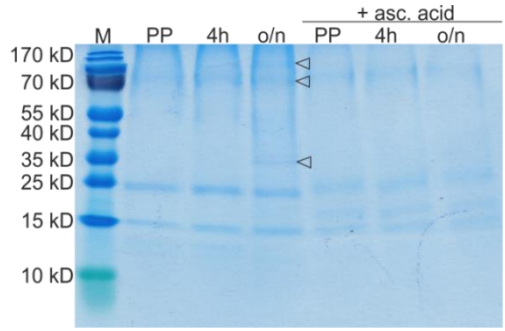

Figure S7. Oxidative cross-linking of α-casein[DOPA] with $NaIO_4$ is not effective with ascorbic acid.
14% Coomassie stained SDS gels picturing α-casein[DOPA] treated with 1 mM $NaIO_4$ during microdialysis. M, molecular weight marker; PP, untreated protein as a negative control; 4h, protein treated with $NaIO_4$ and dialyzed for 4 hours; o/n, overnight microdialysis in the presence of $NaIO_4$.;the open arrows indicate unspecific protein aggregation.

9.7 Size exclusion calibration curves

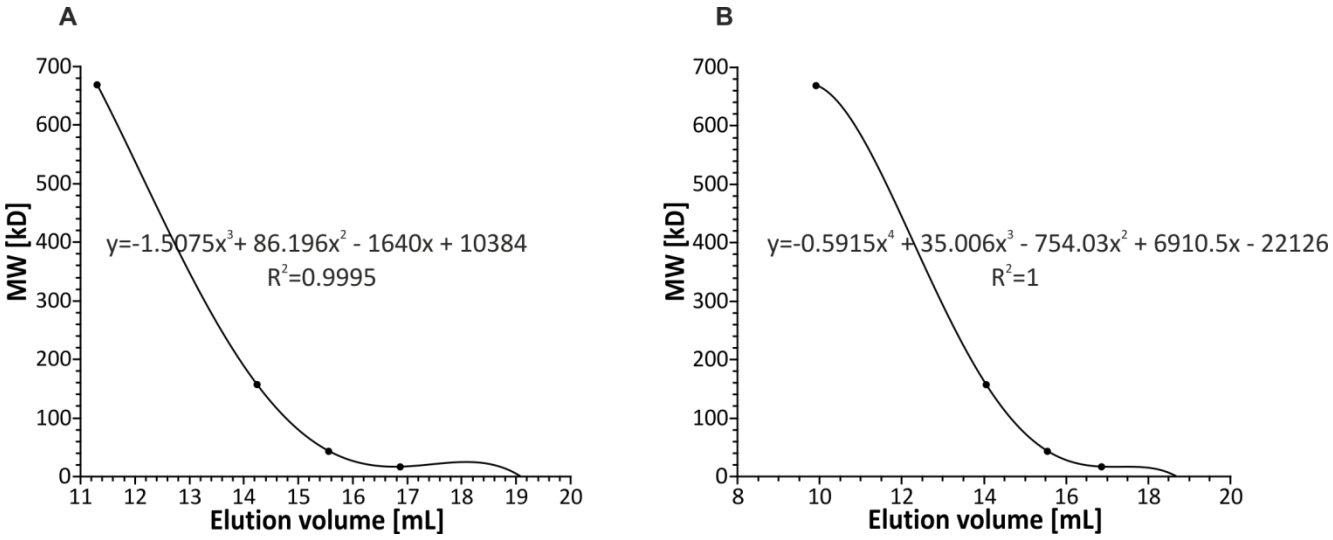

Figure S8. Calibration curves of the protein standard from the size exclusion chromatogram in Figure 12.
(A) Calibration curve without urea in the elution buffer; (B) calibration curve with 3.75 M urea in the elution buffer.

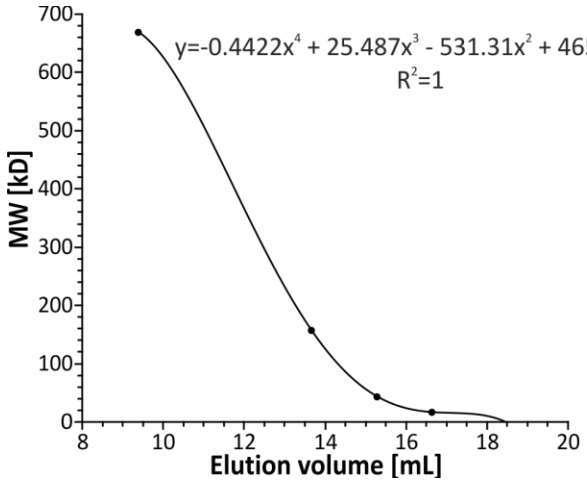

Figure S9. Calibration curve of the protein standard executed in up-flow from the size exclusion chromatogram in Figure 14.

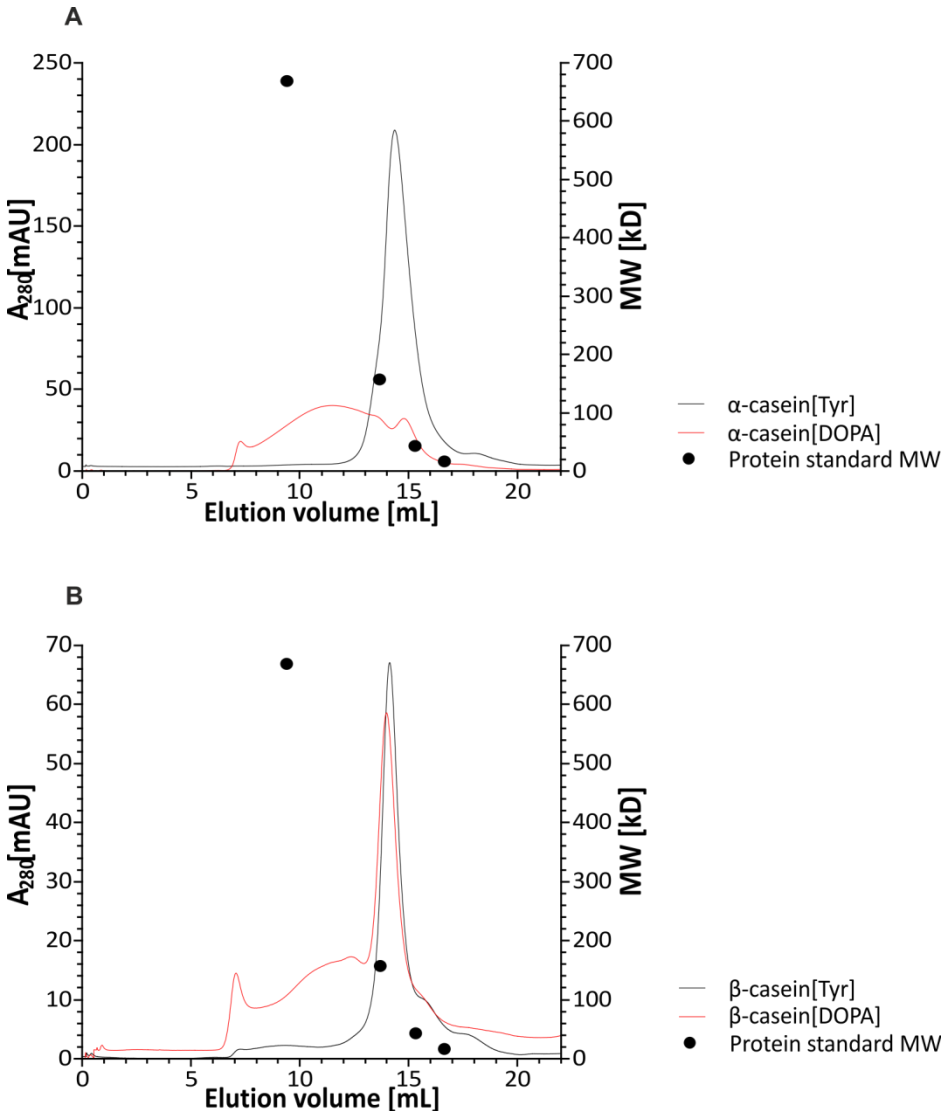

Figure S10. SEC of fermented α- and β-casein parent and DOPA variant proteins illustrate aggregating behavior of sodium periodate-untreated casein[DOPA] variants.
Elution behavior of caseins from a Superdex 200 FPLC column. (A) α-casein; (B) β-casein. Black curve, elution pattern of casein[Tyr] parent proteins; red curves, elution pattern of casein[DOPA] variants; black dots, molecular weight of the protein standard.

9.8 Mass analysis of caseins

Table S6. Mass analysis results of excised Cs1 gel bands band from Figure 24.
Sequence coverage of Cs1 was 91.74%. The casein sequences were digested with trypsin, chymotrypsin, and pepsin.

Sequence	Modification	Hits
DMPIQAFLLYQEPVLGPVR	M2 (Oxidation)	20
EMPFPK	M2 (Oxidation)	1
MHQPHQPLPPTVMFPPQSVL	M13 (Oxidation)	6
MHQPHQPLPPTVMFPPQSVL	M1 (Oxidation); M13 (Oxidation)	27
MRELEELNVPGEIVESL	M1 (Oxidation)	1
MRELEELNVPGEIVE	M1 (Oxidation)	26
MRELEELNVPGEIVESLSSSEESITR	M1 (Oxidation)	5
PVVVPPFLQPEVM	M13 (Oxidation)	22
PVVVPPFLQPEVMG	M13 (Oxidation)	6
PVVVPPFLQPEVMGVS	M13 (Oxidation)	3
PVVVPPFLQPEVMGVSK	M13 (Oxidation)	2
PVVVPPFLQPEVMGVSKVKE	M13 (Oxidation)	1
VVVPPFLQPEVM	M12 (Oxidation)	1
WMHQPHQPLPPTVMFPPQSVL	M14 (Oxidation)	2
WMHQPHQPLPPTVMFPPQSVL	M2 (Oxidation); M14 (Oxidation)	1
PQNIPPLTQ	No fragments detected	0
AMAPKHK	No fragments detected	0

Table S7 Mass analysis results of the excised Cs2 band from Figure 24.
The sequence coverage of Cs2 was 45.41%. The casein sequences were digested with trypsin, chymotrypsin, and pepsin.

Sequence	Modification	Hits
DmPIQAFLLYQEPVLGPVR	M2(Oxidation)	91
EmPFPK	M2(Oxidation)	1
mRELEELNVPGEIVESLSSSEESITR	M1(Oxidation)	10
INKK	No fragments detected	0
HPFAQTQSLVYPFPGPIPNSLPQNIPPLTQTPVVVP PFLQPEVMGVSKVKEAMAPKHK	No fragments detected	0
YPVEPFTESQSLTLTDVENLHLPLPLLQSWMHQPH QPLPPTVMFPPQSVLSLSQSK	No fragments detected	0

9.9 Biomimetic cross-linking with iron

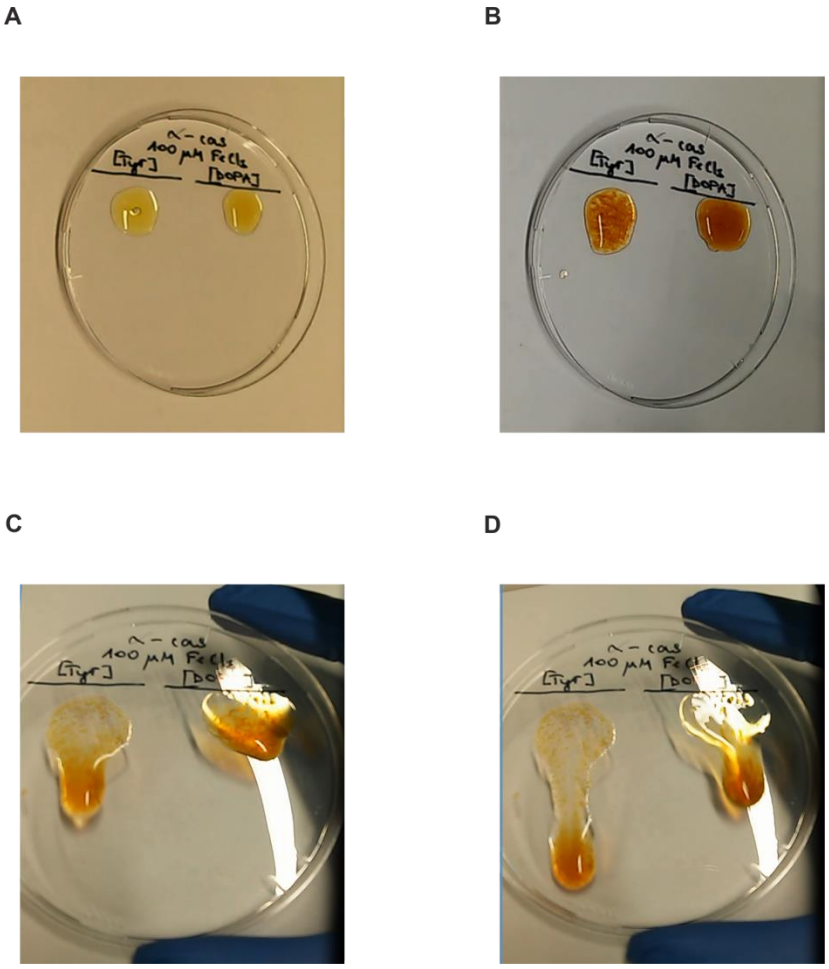

Figure S11. α-casein[DOPA] becomes viscous after cross-linking takes place in the presence of $FeCl_3$. The concentration of protein used was 10 mg/mL. (A) addition of 100 μM $FeCl_3$ to the casein[Tyr] parent and casein[DOPA] variant proteins; (B) addition of 420 μM NaOH; (C) after mixing and slowly lifting the petri plate, the casein[Tyr] drops down rapidly; (D) a few seconds later, the casein[DOPA] starts descending. The orange particles seen in the casein[Tyr] sample are precipitated iron.

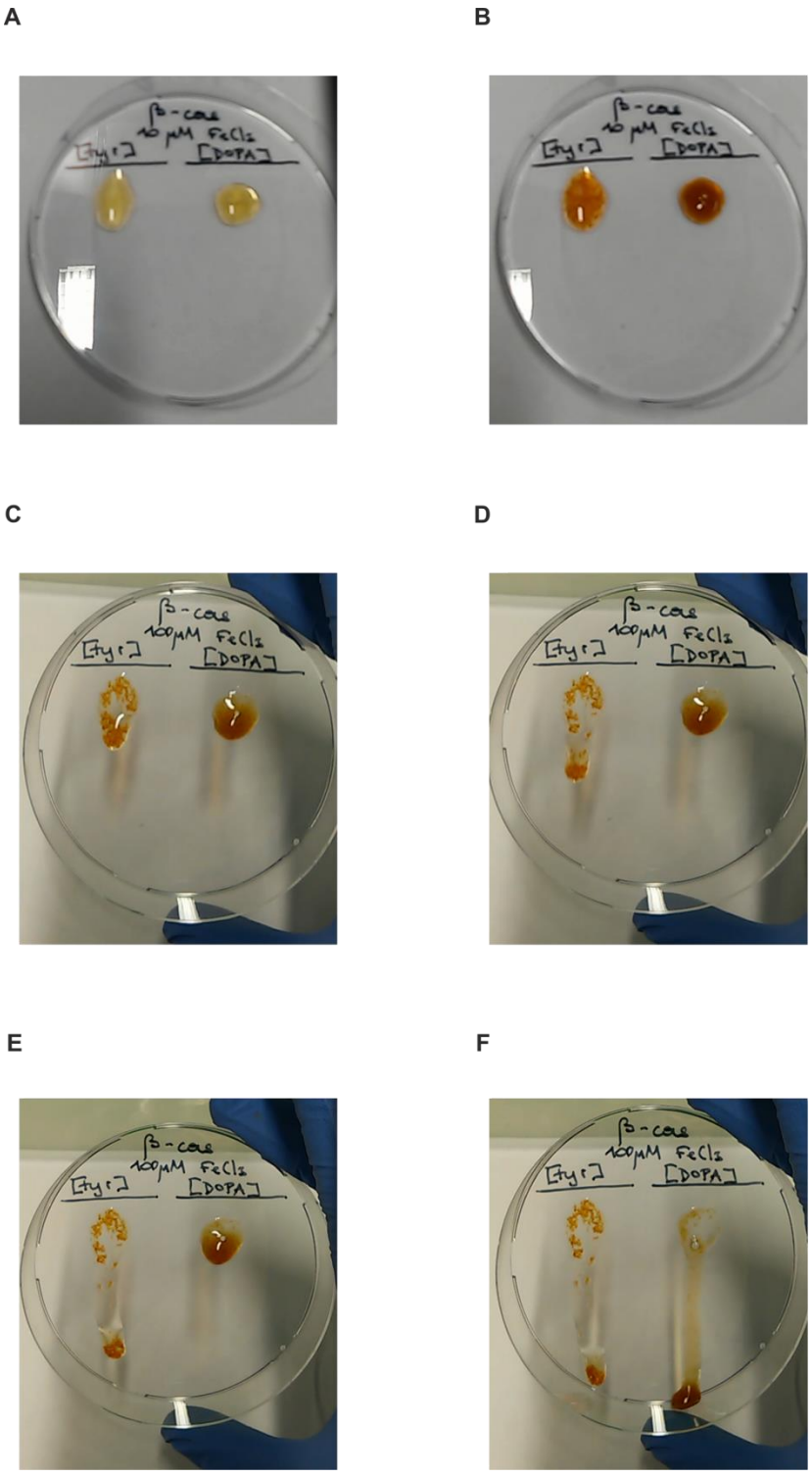

Figure S12. β-casein[DOPA] becomes viscous after cross-linking takes place in the presence of $FeCl_3$. The concentration of protein used was 10 mg/mL. (A) addition of 100 μM $FeCl_3$ to the casein[Tyr] parent and casein[DOPA] variant proteins; (B) addition of 420 μM NaOH; (C) after mixing and slowly lifting the petri plate, the casein[Tyr] starts migrating, the DOPA variant sticks to the plate; (D) a few seconds later, the casein[Tyr] sample descends further; (E) casein[DOPA] still has not yet moved; (F) the DOPA sample finally drops to the bottom of the petri plate. The orange particles seen in the casein[Tyr] sample are precipitated iron.

9.10 Instruments and chemicals used

Table S8. Instruments and devices used for all lab work.

Instrument	Supplier
Analytical scale	Sartorius; Göttingen, Germany
Avanti J-20 XP centrifuge	Beckmann Coulter Inc.; Brea, CA
Bradford reagent (Cat. Nr. 500-0006)	BioRad; Vienna, Austria
Centrifuge tubes, 50 mL and 500 mL	Thermo Scientific Inc.; Waltham, MA
Centrifuges	Centrifuge 5415R: Eppendorf; Hamburg, Germany Centrifuge 5424: Eppendorf;
Column (5 mL HiTrap)	GE Healthcare; Buckinghamshire, UK (Cat. Nr. 17-5247-01)
Desalting column (HisTrap 26/10)	GE Healthcare (cat. Nr. 17-5087-01)
Bioreactor	DASGIP technology; Jülich, Germany
Column (5 mL HiTrap)	GE Healthcare (Cat. Nr. 17-5247-01)
Desalting column (HisTrap 26/10)	GE Healthcare (Cat. Nr. 17-5087-01)
Eppendorf tubes	Sarstedt, Nümbrecht, Germany
Flasks	2000 mL (Schott Duran); Bartelt; Graz, Austria 1000 mL (SIMAX), Bartelt 250 mL (Schott Duran); Bartelt
FPLC	ÄKTA Purifier; GE Healthcare ÄKTA Prime; GE Healthcare
Gravity flow affinity purification equipment	Ni-NTA resin; Quiagen; Hilden, Germany Disposable 10 mL Polypropylene columns; Thermo Scientific
ddH_20 device (arium®basic)	Sartorius
Incubator HT MultitronII	InforsAG, Bottmingen, Switzerland
Lab scale	Binder GmbH, Tuttlingen, Germany
Laminar air flow hood	AirClean, Woerden, The Netherlands
PD10 Desalting columns	GE Healthcare
Petri dishes	Greiner Bio-one International AG; Kremsmünster, Austria
Photometer	Beckman Clouter Inc. BioPhotometer; Eppendorf
Pipette tips	Greiner Bio-one International AG
Pipettes	1000 µL, 200 µL, 20 µL; Denville; South Plainfield, NJ 10 µL (Biohit); Sartorius
Plate reader(SPECTRAmax Plus384)	Molecular Devices, Sunnyvale, CA
Scanner	Tevion USB Scanner; Mühlheim, Germany
Sterile syringe filters	Scientific Strategies; Yukon, OK
Sterile filters	Sartorius
Sonifier	Branson; Danbury, CT
UV cuvettes	Greiner Bio-one International AG
Vortex	IKA®-Werke GembH & Co. K; Staufen, Germany

Table S9. List of enzymes and reagents used in this project.

Reagent	Cat. Nr.	Supplier
acetic acid	6755.2	Roth; Kalsruhe, Germany
agarose LE	840004	Biozyme; Hessisch-Oldendorf, Germany
ammonia	5460.1	Roth
antifoam	LM1207072	Bussetti & Co GmbH; Vienna, Austria
ammonium persulfate (APS)	13375.01	Serva; Heidelberg, Germany
Acrylamide/Bis solution, 37.5:1 (30 % w/v), 2.6 % C	10688.01	Serva
boric acid	1001625000	Merck; Billerica, MA
calcium chloride dihydrate	CN93.1	Roth
Comassie Blue-250R	3862.2	Roth
DNAse I	DN25-100mg	Sigma-Aldrich
1,4-dithiothreitol (DTT)	6908.1	Roth
ethanol	20821.330	VWR International, Radnor, PA
α-D-glucose monohydrate	6780.2	Roth
Gel filtration standard	151-1901	BioRad
glycerol	3908.3	Roth
hydrochloric acid, fuming	4625.1	Roth
imidazole	1047161000	Merck
iron(III)chloride	1039650500	Merck
isopropyl β-D-1-thiogalactopyranoside (IPTG)	CN03.3	Roth
kanamycin sulfate	T832.2	Roth
LB-Agar Lennox	X65.3	Roth
LB-medium Lennox	X964.2	Roth
leucine	1699.1	Roth
lysozyme	8259.2	Roth
magnesium chloride	8.14733.0500	Merck
magnesium sulfate heptahydrate	A537.4	Roth
β-mercaptoethanol	4227.1	Roth
methionine	1702.1	Roth
PageRuler prestained protein ladder	SM0671	Thermo Scientific
di-potassium hydrogen phosphate	T875.2	Roth
potassium dihydrogen phosphate	P018.2	Carl Roth, Germany
SOB medium	AE27.1	Roth
sodium chloride	9265.1	Roth
di-sodium hydrogen phosphate	T876.2	Roth
sodium dihydrogen phosphate monohydrate	T879.2	Roth
sodium dodecyl sulfate (SDS)	2326.1	Roth
sodium hydroxide	P031.2	Roth
sodium hydrogen carbonate	0965.1	Roth
Sodium periodate	7790-28-5	Merck
sterile water	-	Fresenius Cabi, Graz, Austria
tetramethylethylendiamin (TEMED)	35930.01	Serva
tris base	4855.3	Roth
Triton-X100	3051.3	Roth
valin	4879.1	Roth
yeast extract (Charge: 269107355)	2363.4	Roth