Improving production of *Streptomyces griseus* trypsin for its application in processing insulin precursor

YunFeng Zhang (yf.zhang@siat.ac.cn)
Jiangnan University https://orcid.org/0000-0003-2333-142X
Qixing Liang
Jiangnan University
Chuanzhi Zhang
Chia Tai Tianqing Pharmaceutical Group Co Ltd
Juan Zhang
Jiangnan University
Guocheng Du
Jiangnan University
Zhen Kang
Jiangnan University

Research

Keywords: Streptomyces griseus trypsin, autolysis, unfolded protein response (UPR), Pichia pastoris, insulin

Posted Date: December 30th, 2019

DOI: https://doi.org/10.21203/rs.2.19562/v1

Abstract

Background: Trypsin has a plenty application in food and pharmaceutical manufacture. While, the commercial trypsin is usually extracted from pork pancreas, which has the risk of infectious and immunogenicity. Therefore, the microbial *Streptomyces griseus* trypsin (SGT) is a prior alternation because it processes efficient hydrolysis activity without the aforementioned risk. The remarkable hydrolysis efficiency of SGT caused its autolysis, and five autolysis sites R21, R32, K122, R153, and R201 were identified from its' autolysate.

Results: The tbcf (K101A, R201V) mutant was screened by directed selection approach for improved activity in flask culture (60.85 ± 3.42 U·mL -1 , increased 1.5-fold). From the molecular dynamics simulation, the K101A/R201V mutation shortened the distant between catalytical residues D102 and H57 from 6.5 Å vs 7.0 Å, which afforded the improved specific activity 1527.96 ± 62.81 U·mg -1 . Further, the production of trypsin was increased 302.8% (689.47 ± 6.78 U·mL −1) in 3-L bio-reactor, with co-overexpression of chaperones SSO2 and UBC1 in *Pichia pastoris*.

Conclusions: The SGT protein could be an adequate trypsin for insulin production. When working with hydrolysates analysis and direction selection, the tbcf (K101A, R201V) mutant increased 1.5-fold activity. Further, the production of trypsin was improved 3-fold by overexpressing chaperone protein in Pichia pastoris . The future study should be emphasized on the application of SGT in insulin manufacture and pharmaceutical.

Background

Trypsin (EC 3.4.21.4) is an alkaline protease and has stringent cleavage specificity to the carboxyl terminus of arginine (R) and lysine (K). Similar with other serine proteases, trypsin shares the canonical catalytic triad H57, D102, and S195 [1–3]. To date, trypsin has been widely used in leather batting, detergents, food industry, and pharmaceuticals. Especially, in insulin manufacture, the trypsin was applied for activating insulin precursor into insulin ester by digesting mini C-chain peptide [4–7]. In food procession, trypsin was applied for preparing nutritional protein-peptides [8], reducing the allergen and enhancing the digestibility of baby food [9]. Moreover, trypsin was also evaluated for wound healing in the pharmaceutical field [10]. Currently, the commercial trypsin is usually extracted from the porcine and bovine pancreas. However, due to the potential risk of infectious agent contamination, the animal-derived trypsin is under control in pharmaceuticals and food manufacture. Thus, microbial Streptomyces griseus trypsin (SGT) is a potential alternative for sharing similarities with bovine trypsin (BT) in three-dimensional structure and catalytical property. Moreover, SGT showed higher hydrolysis performance than BT in proteomics application, because of producing a amount of matching tryptic peptides [11]. However, the high hydrolysis efficiency of SGT caused autolysis [12], owing to 16 potential autolysis residues in its polypeptide sequence. Thus, high-yield production of trypsin might be hindered by autolysis of active trypsin. The mutant tbcf (K101A), with R145I and R201V mutations, afforded the

higher stability against autolysis [13]. However, the autolysis residues of SGT have not been investigated in detail comparing with BT [14, 15].

Over the past decade, SGT has been investigated in different microbial hosts, such as Escherichia coli [16], Bacillus subtilis [17], S. griseus [18, 19] and Pichia pastoris [20, 21]. Notably, by engineering the N-terminal peptide, the members of Kang group achieved high-yield production of trypsin (227.65 $U \cdot mL^{-1}$) in P. pastoris [13, 22]. However, the expression of trypsin caused the unfolding protein response (UPR), and trigged endoplasmic reticulum-associated degradation (ERAD) in P. pastoris. When induced by pAOX1 promoter, the human trypsinogen protein stacked as inclusion body in the endoplasmic reticulum, which up-regulated the UPR then caused cascade degradation of the intracellular un-folded trypsinogen by Bip protein [23]. On the other hand, the folding of disulfide bonds in trypsin was also the burden for P. pastoris. Because the generated peroxide toxicity further induced ERAD response [24]. In this study, our objective was to develop a strategy for the high-yield production of SGT in P. pastoris, by directed selection and co-overexpression of chaperones. Moreover, the SGT mutant could be applied in manufacturing insulin product from the insulin precursor.

Results And Discussion

Identification of the autolysis sites in tbcf (K101A)

The SGT contains 16 potential autolysis residues (Arg, Lys). In previous study, the stability and production of SGT were improved with tbcf (K101A) mutant [13]. For further investigating the autolysis of tbcf (K101A), its' hydrolysate was analyzed by MALDI-TOF-MS. Based on the autolysis fragments of tbcf (K101A), five autolysis residues were identified as R21, R32, K122, R153, and R201 (Fig. 1A). This result indicated that the SGT slightly prefers to hydrolyze at R than K. Because the hydrogen bond interaction between R and substrate binding D189 was more stable than a water molecule bridged interaction between K and D189 [25]. Štosová et al. significantly reduced the autolysis of SGT by modifying R or K residues with chemical reagents phenylglyoxal and formaldehyde, but the specific activity of SGT dropped to 12% of the parent enzyme [12]. Because the R and K interacted with other residues to form hydrogen bond, salt bridge and π-interaction. These interactions played essential roles in SGT folding, three-dimensional structure, and catalytic activity. In secondary structure of SGT, except R201V (β-Sheet), the other four autolysis residues (R21, R32, K122, and R153) were located in the loop regions (Fig. 1B). Specifically, R21 interacted with Y131 by a hydrogen bond, which was also the case for R32-T55 and R32-Q128 interactions. Whereas K122 formed a salt bridge with D184 and E188, so as for R153 and D60. These results indicated that R21, R32, K122 and R153 residues helped lock the three-dimensional conformation.

Directed selection of SGT

The direct selection method was widely applied to engineer proteins for specific properties [26]. With the help of online bioinformatic tools, bespoke candidate of mutations could be suggested according to the overall protein stability, residue physio-chemical environment, and interaction network of protein [27–29]. The SGT candidates were sorted for higher trypsin activity than parent tbcf (K101A) in flask culture. In the C-terminal residue R201, the tbcf (K101A, R201V) mutant showed the highest activity 60.85 ± 3.42 $U{\cdot}mL^{-1}$, increased 1.5-fold than the parent (Fig. 2A). Among six mutations of R32, the R32A was outstanding with increased trypsin activity 45.15 ± 0.99 $U{\cdot}mL^{-1}$ (Fig. 2B). While, the other mutations of R21, K122 and R153 decreased the trypsin activity (Fig. 2CDE).The R32A mutation was then added into tbcf (K101A, R201V) to afford tbcf (K101A, R201V, R32A) mutant. Even tbcf (K101A, R201V, R32A) showed higher expression level than tbcf (K101A, R201V) (Fig. S1), its activity decreased by 50% (30.67 ± 2.63 $U{\cdot}mL^{-1}$). Further, the specific activity of trypsin mutants was compared, with substrate BAPNA using spectrophotometric assay. The mutant tbcf (K101A, R201V) showed the highest specific activity of 1527.96 ± 62.81 $U{\cdot}mg^{-1}$, increased 13.76% than parent. Eventually, the mutant tbcf (K101A, R201V) was discovered with increased production of trypsin and its specific activity.

Enzyme kinetics and molecular modeling analysis of SGT mutant

The MD simulation was applied to analyze the three-dimensional structure of tbcf (K101A, R201V) compared with tbcf. The root-mean-square deviations (RMSD) of protein backbone atom indicated the stability of protein [30]. After the 7 ns MD simulation, the backbone of tbcf (K101A, R201V) mutant showed the lower deviation of RMSD value (Fig. S2). This result indicated the K101A/R201V mutations could improve the stability of SGT backbone. Then, the internal interaction of catalytical triad (H57, D102 and S195) was analyzed. Interestingly, the tbcf (K101A, R201V) mutant showed a shorter distance between H57 and D102 (6.5 Å vs 7.0 Å) in the catalytic center (Fig. 3AB). Consequently, the tbcf (K101A, R201V) mutant, with a k_{cat}/K_m value of 1.53×10^7 $min^{-1}{\cdot}mM^{-1}$, afforded higher catalytical activity than parent tbcf (K101A) (Supplementary Table 3). Especially, the increased specific activity might attribute to shortened distance between D102 and H57 in catalytical triad, which could consolidate the hydrogen bond between carboxylic oxygen of D102 and δ-nitrogen of H57 [31]. Because the hydrogen bond stabilized the structure of H57 in catalytical transient state, which facilitated H57 to accept the proton from S195 [32]. Moreover, the Km value of tbcf (K101A, R201V) and tbcf were similar, which were 5.39 ± 0.36×10^{-2} mM and 5.86 ± 0.16×10^{-2} mM respectively. This result indicated that K101A/R201V mutations retained the conserved internal interaction at substrate binding domain.

High-yield production of SGT with co-overexpression chaperones in P. pastoris

Protein expression was known to be regulated by UPR [33] or ERAD [34] in P. pastoris. And expression of trypsinogen triggered UPR and ERAD in P. pastoris, because of the unfolded trypsinogen in endoplasmic reticulum (ER) and peroxide toxicity by forming disulfide bond [23, 24, 35]. It was known that protein expression could be improved by upregulating the endogenous proteins [36]. Therefore, twelve proteins were individually overexpressed, involved in transcription regulation, disulfide bond formation and protein secretion. The ER located chaperones processed diverse functions during polypeptides folding into the biologically active protein. And these chaperones included the oxidative reaction in protein folding (Ero1), disulfide bond forming (GLR1, PDI, and GSH2), and degradation of the unfolded protein (UBC1) [34, 35]. Interestingly, the production of trypsin was increased by 17.0% and 31.6% with overexpression of GSH2 and UBC1, respectively (Fig. 4). Moreover, the transport of polypeptides was known to be critical for secretory proteins. The Bip, SLY1 and SEC53 chaperones were responsible for transporting and recognizing of nascent polypeptides in ER and Golgi membrane [37]. And, the SEC1 and SSO2 promoted the extracellular secretory of the folded protein [38, 39]. So, overexpression of SEC1 and SSO2 increased the trypsin activity by 24.1% and 41.5%, respectively. Then, the SEC1, SSO2 and UBC1 were co-overexpressed, because they contributed more than 20% increase of trypsin activity. Finally, the trypsin production of strain GS115-tbcf (K101A, R201V)_SU showed highest production 109.25 ± 4.76 $U{\cdot}mL^{-1}$ (increased by 79.5%), with co-overexpression of SSO2 and UBC1 in flask culture (Fig. 4). This indicated that the bottle-neck for high-yield production of SGT might be hindered by secretory transportation and degradation of unfolded trypsin in P. pastoris.

Application of high-yield trypsin to processing insulin precursor

For scale-up production of the trypsin, the strain GS115-tbcf (K101A, R201V)_SU was cultured in 3-L bio-reactor, according to the published method [13]. After glycerol fed-batch cultivation, a higher density of cells (68.02 ± 1.5 g/L, DCW) was achieved with glycerol feeding and high agitation speed (850 $r{\cdot}min^{-1}$) (Fig. 5). Then, the fermentation entered methanol-feeding cultivation phase, when the glycerol was depleted with the indication of increased DO (over 50%). After induction for 156 h, the trypsin production reached 689.47 ± 6.78 $U{\cdot}mL^{-1}$, which was increased 302.8% than parent GS115-tbcf (K101A).

The mammalian trypsin was generally applied for the preparing insulin from its precursor, because of the canonically tryptic cleavage of lysine for removal of C-chain [4, 40]. While, the traditional method for preparing trypsin is extraction from the mammalian pancreas, which has the risk of the bioactive compound, infectious virus, and heath-harmful proteases [41]. Although the heterologous expression of mammalian trypsin could avoid the aforementioned problems, it still suffered from immunogenicity issues, low expression level, activation of the zymogen, and autolysis [15, 42, 43]. Importantly, the SGT mutant tbcf (K101A, R201V) showed higher hydrolysis performance, with no immunogenicity, high production, autoactivation and stability against autolysis. The tbcf (K101A, R201V) was mixed with insulin precursor rPI to afford insulin precursor with Asp30 deleted B-chain (PI-B^{D30}) (Fig. 6A). And the

tbcf (K101A, R201V) was compared with commercial porcine trypsin at the identical condition. After hydrolysis for 19 h, the rPI was converted to PI-B^{D30} as demonstrated in the HPLC chromatograph. The elution time of rPI was 18.75 min (Fig. 6B). After cleavage by commercial porcine trypsin and tbcf (K101A, R201V), the rPI was converted to PI-B^{D30}, which was eluted out at 21.40 min (Fig. 6CD). So, the engineered SGT mutant tbcf (K101A, R201V) performed the potential application in insulin manufacture, due to the same hydrolysis capacity with commercial porcine trypsin.

Conclusions

In this study, five autolysis residues were identified in of SGT. And the mutant tbcf (K101A, R201V) was identified from a library of 35 mutants with improved hydrolysis performance and specific activity. Furthermore, the production of trypsin was increased 302.8% (689.47 ± 6.78 $U{\cdot}mL^{-1}$) with co-overexpression of chaperone proteins SSO2 and UBC1 in P. pastoris. Consequently, the engineered SGT tbcf (K101A, R201V) showed the same hydrolysis capacity with commercial poricne trypsin. The future study should be emphasized on the application of SGT in insulin manufacture and pharmaceutical [44].

Methods

Chemical and materials

PCR kit was purchased from Takara (Takara Ltd., Otsu, Japan). Restriction enzyme, 12% SDS-PAGE gel, Gel-extraction kit, and high-performance liquid chromatography (HPLC) column (Thermo scientific accucore C4 2.6 μm 50 × 4.6 mm 150A) were purchased from Thermo Scientific (Shanghai, China). DNA ligation kit (CloneExpress™) was purchased from Vazyme Biotech Co., Ltd. (Nanjing, China). P. pastoris GS115 (His^{-}) and E. coli JM109 strain were stocked in the laboratory. Nα-benzoyl-DL-arginine-p-nitroanilide (BAPNA), Nα-benzoyl-L-arginine ethyl ester hydrochloride (BAEE) and porcine trypsin were purchased from Sigma Aldrich (Louis, USA); Oligo synthesis, sequencing and modified Bradford protein assay kit were purchased from Sangong Biotech Co., Ltd. (Shanghai, China). Ultrafiltration 3-kDa cutoff filter was purchased from Millipore (Bedford, MA, USA); Benzamidine column (1 mL Φ1.6 × 2.5 cm), Gel column (HiLoad 16/60 Superdex 200 pg), and ÄKTA™ were purchased from GE Healthcare (Shanghai, China). The ion-exchange chromatography CM Sepharose FF, reversed-phase chromatography (SOURCE 30RPC) and other reagents were purchased from local businesses.

Prediction of the point mutations

The mutations were predicted by web servers (PoPMuSiC 2.1 [28], DUET [27], NeEMO [29]). All of the predictions were used 1SGT as the template. Based on trypsin mutant tbcf (K101A), three web servers (PoPMuSiC 2.1, DUET and NeEMO) were used to predict the positive mutation of five autolysis sites with a cutofff for increased ΔΔG value (PoPMuSiC, DUET) or Kcal/mol value over 1.4 (NeEMO). For NeEMO

prediction, the pH and temperature were set at 6.0 and 30 ºC. And R32A, K122A, and R153A mutation was individually intraduced into tbcf (K101A), for the increased activity in previous study [13].

Plasmid and strain construction

The PCR product was generated from pPIC9K-tbcf (K101A) plasmid with the corresponding primers (Supplementary Table 1). The ligated plasmids were transformed into E. coli strain JM109, then selected by 50 $mg{\cdot}L^{-1}$ kanamycin. Moreover, the constructed plasmids were linearized with restrictive enzyme SalI, then transformed into P. pastoris GS115 (His^{-}) competent cell and screened by histidine autotrophic phenotype. Consequently, the recombinant yeast, with the high copy number of trypsin cassette, was screened by streaking the single colony on YPD plate with gradient geneticin (1 $mg{\cdot}mL^{-1}$, 2 $mg{\cdot}mL^{-1}$, 3 $mg{\cdot}mL^{-1}$ and 4 $mg{\cdot}mL^{-1}$). The q-PCR quantification was applied to quantify the copy number of trypsin cassette in recombinant yeast from 4 $mg{\cdot}mL^{-1}$ geneticin plate [22].

The endogenous proteins were expressed with constitutive promoter pGAP. Firstly, the backbone vector was amplified from the pGAPZB plasmid by FpGAPZB and RpGAPZB primers. Then, the gene fragments were amplified from genomic DNA of P. pastoris GS115 (His^{-}) with the respective primers, and the hac1s (intron split hac1 gene) was synthesized with coding sequence of hac1. Finally, the construction of plasmids was performed by DNA ligation kit. Moreover, the co-expressing plasmids pGAP-SSO2/SEC1 and pGAP-SSO2/UBC1 were constructed. Firstly, the backbone vector was generated from the pGAP-SSO2 plasmid with FpGAP-C and RpGAP-C. Then the expressing cassettes of $pGAP^{m}$-SEC1-tAOX1 and $pGAP^{m}$-UBC1-tAOX1 were separately amplified by primer FpGAPm and RAOX1 from plasmids $pGAP^{m}$-SEC1 and $pGAP^{m}$-UBC1, whose restrictive enzyme site AvrII was mutated by primers FpGAPmAvrII and RpGAPmAvrII. Finally, the backbone plasmid and expression cassettes were assembled by DNA ligation kit. With the same method, the plasmid pGAP-SEC1/UBC1 was constructed with corresponding primers. For the construction of plasmid pGAP-SSO2/SEC1/UBC1, the pGAP-SSO2/SEC1 plasmid was digested by restrictive enzyme BamHI, and the expression cassette $pGAP^{m}$-UBC1-tAOX1 was amplified from the $pGAP^{m}$-UBC1 plasmid with primers FpGAPm-UBC1 and RpGAPm-UBC1. Finally, the digested plasmid and $pGAP^{m}$-UBC1 expression cassette were assembled by DNA ligation kit. For constructing recombination yeast strain, the 2 ug of AvrII-linearized DNA fragment was transformed into GS115-tbcf (K101A, R201V) competent cells by electroporation. Then the transformants were incubated in YPD medium plus 1 M sorbitol (YPDS) for 2 h at 30 °C. Finally, the pellets were spread on YPDS agar plate with 80 $mg{\cdot}L^{-1}$ Zeocin and grew for 3 days at 30 °C. The strains used in this study were shown in supplementary Table 2.

Media and cultivation

The media included Luria–Bertani (LB) medium, Yeast Extract Peptone Dextrose (YPD) medium, Buffered Methanol-complex (BMMY) medium and Basal Salts (BSM) medium [13]. The yeast cells were pre-cultured in YPD medium for 24 h, at 30 °C and 220 rpm. Then the pellets were resuspended in BMMY

medium for 144 h cultivation with the same condition. Also, 1% methanol (v/v) was added into medium every 24 h for inducing the pAOX1 promoter. The scale-up cultivation was carried out in 3-L bio-reactor (INFORS, Switzerland), with 800 mL BSM medium. The cultivation process was divided into three phases. During the glycerol batch cultivation, the yeast cells were cultured under pH 5.5, 30 °C, and the dissolved oxygen (DO) controlled over 30% by constant agitation speed. Glycerol fed-batch cultivation was carried out, when the glycerol was depleted with DO over 50%. And the feeding solution (50% glycerol with 1.2% PTM1 solution) was gradiently added into medium to confer high-density cultivation with higher DO level by increased agitation speed (800 rpm). In the methanol fed-batch phase, the trypsin was expressed by methanol induced promoter pAOX1. In order to avoid the repression effect of pAOX1 by the glycerol, the inducer methanol was added 2 hours later when DO was over 60%. And the methanol was gradually fed, according to the method developed by Wang et al. [45].

Expression of human insulin precursor

The codon-optimized coding sequence of recombinant human insulin precursor (rPI) was ligated into the pPICK-9K plasmid, by the restrictive enzyme site EcoRI and BamHI. The recombinant strain GS115-rPI was screened by 4 $mg{\cdot}mL^{-1}$ geneticin for the high copy number of expression cassettes. And the scale-up cultivation and preparation of insulin precursor were according to the reported method. [4, 46]. The insulin precursor sample was purified by ion-exchange chromatography and reversed-phase chromatography, then lyophilized after isoelectric precipitation.

Purification of trypsin

The culture supernatant was separated for purifying the trypsin, under centrifugation at 5000 g for 10 min. And the filtered (0.22 μm) sample was loaded into 1 mL benzamidine column, then equilibration with buffer A (pH 7.4 50 mM Tris–HCl, 0.5 M NaCl). Finally, the trypsin was eluted with 60% buffer B (pH 2.0, 10 mM HCl, 20 mM NaAc) [13]. Moreover, the purified trypsin was further separated using the HiLoad 16/60 Superdex 200 pg column. Finally, the concentration of purified trypsin was determined by the modified Bradford protein assay kit.

MALDI-TOF-MS analysis

MALDI-TOF-MS was applied to identify the autolysis fragments. The concentrated hydrolysate of tbcf (K101A) [13] was loaded on the MALDI-TOF-MS plate with the control without hydrolysis. Moreover, the autoproteolytic peptides were analyzed with the Swiss-Prot database [47].

Molecular dynamics (MD) analysis

The three-dimensional model of trypsin mutant was simulated by NAMD 2.11 with CHARMM27 force field [48, 49]. For the simulation of trypsin mutants tbcf and tbcf (K101A, R201V), the modeling template SGT (PDB, 1SGT) was download from RCSB Protein Data Bank [50]. The protein was set in a cubic water box of 70 × 70 × 70 Å with the criterion that the more than 7 Å water layer in each dimension and 12 Å cut-off for non-bonded interactions. The default pH was set at 7.0 and the Na^+ was added for charge neutralization. After water equilibration (1 ns) and minimization (1000 steps), the 60 ps of heating was performed from 0 K up to 300 K before each primary molecular simulation [51]. The temperature was kept at 300 K for 150 ps for equilibration and 10 ns simulation for data sampling at constant temperature and pressure. Finally, the three-dimensional structure was analyzed and presented by PyMOL Molecular Graphics system [52].

Determination of trypsin activity and hydrolysis of insulin precursor

The culture supernatant or the purified enzyme were prepared to measure the trypsin activity, according to the published method [22]. And, the purified recombinant SGT and commercial porcine trypsin were prepared to hydrolyze the insulin precursor. The trypsin, with 3000 U BAEE activity, was added into 10 mL insulin precursor solution (60 $mg{\cdot}mL^1$ in 50 mM pH 8.0 Tris-HCl solution with EDTA·2Na). The hydrolysis was performed at 25 ºC for 19 h, then ceased by adjusting pH value to 3.0 with HCl. At the same time, the hydrolysis solution without trypsin was set as control. Consequently, the hydrolysis product was analyzed by HPLC with the modified method developed by Richard et al. [53].

All experiments were carried out with triplicate and data was shown as means ± standard deviation.

Abbreviations

SGT
Streptomyces griseus trypsin
UPR
unfolded protein response
R
arginine
K
lysine
BT
bovine trypsin
MD
Molecular dynamics
ERAD
reticulum-associated degradation

RMSD
root-mean-square deviations
ER
endoplasmic reticulum

DCW
dried cell weight
DO
Dissolved oxygen

Declarations

Acknowledgements

We thank Dr. Zhongyi Chen (School of Biotechnology, Jiangnan University) for the provision and instruction of the molecular dynamics analysis and also Dr. Xiaozhou Luo (Department of Chemical and Biomolecular Engineering, University of California, Berkeley) for study instruction. We thank the Xinhui Yao and Hao Huang (School of Biotechnology, Jiangnan University) for experiment support.

Authors' contributions

K. and Y. Z. conceived and designed the experiment. Y. Z. and Q. L. acquired the data. C. Z. instructed the experiment. Y. Z. and Z. K. analyzed the data and wrote the manuscript. J. Z. and G. D. contributed to revise the manuscript. All authors read and approved the final manuscript.

Funding

This work was financially supported by the grant from the Fundamental Research Funds for the Central Universities (JUSRP51707A), the Program for Changjiang Scholars and Innovative Research Team in University (grant No. IRT_15R26), the National Key Research and Development Program of China (2017YFB0308401), the Priority Academic Program Development of Jiangsu Higher Education Institutions, and Jiangnan University Graduate Scholarship for Overseas Study.

Availability of data and materials

All data generated or analyzed during this study are included in this published article and its additional file.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors claim no competing interests.

Author details

[1]Key Laboratory of Industrial Biotechnology, Ministry of Education, Jiangnan University, 1800 Lihu Road, Wuxi, Jiangsu 214122, China;

[2]The Key Laboratory of Carbohydrate Chemistry and Biotechnology, Ministry of Education, School of Biotechnology, Jiangnan University, Wuxi 214122, China

[3]Center for Synthetic Biochemistry, Institute of Synthetic Biology, Shenzhen Institutes of Advanced Technologies, Shenzhen, China.

[4]Bio-Pharmaceutical Research Institute Lian Yun Gang, Chia Tai Tianqing Pharmaceutical Group Co., Ltd., Lianyungang, Jiangsu, China.

References

1. Carter P, Wells JA: Dissecting the catalytic triad of a serine protease. *Nature* 1988, 332:564-568.
2. Kraut J: Serine proteases-structure and mechanism of catalysis. *Annu Rev Biochem* 1977, 46:331-358.
3. Ash EL, Sudmeier JL, De Fabo EC, Bachovchin WW: A low-barrier hydrogen bond in the catalytic triad of serine proteases? Theory versus experiment. *Science* 1997, 278:1128-1132.
4. Liu H, Zhou X, Tian S, Hao X, You J, Zhang Y: Two-step transpeptidation of the insulin precursor expressed in *Pichia pastoris* to insulin ester via trypsin-catalyzed cleavage and coupling. *Biotechnol Appl Biochem* 2014, 61:408-417.
5. Liu H, Zhou X, Xie F, You J, Zhang Y: An efficient trypsin digestion strategy for improving desB30 productivity from recombinant human insulin precursor fusion protein. *Process Biochem* 2013, 48:965-971.
6. Zhang Z, Chen H, Tang YH, Feng YM: A simple, economical method of converting gene expression products of insulin into recombinant insulin and its application. *Progress in Natural Science* 2003, 13:596-600.
7. Jakubke HD, Kuhl P, Konnecke A: Basic principles of protease-catalyzed peptide-bond formation. *Angew Chem Int Ed Engl* 1985, 24:85-93.
8. Ketnawa S, Benjakul S, Martinez-Alvarez O, Rawdkuen S: Fish skin gelatin hydrolysates produced by visceral peptidase and bovine trypsin: bioactivity and stability. *Food Chem* 2017, 215:383-390.
9. Mao YH, Krischke M, Hengst C, Kulozik U: Comparison of the influence of pH on the selectivity of free and immobilized trypsin for beta-lactoglobulin hydrolysis. *Food Chem* 2018, 253:194-202.

10. Gudmundsdottir A, Hilmarsson H, Stefansson B: Potential use of Atlantic cod trypsin in biomedicine. *Biomed Res Int* 2013, 2013:1-11.
11. Walmsley SJ, Rudnick PA, Liang Y, Dong Q, Stein SE, Nesvizhskii AI: Comprehensive analysis of protein digestion using six trypsins reveals the origin of trypsin as a significant source of variability in proteomics. *J Proteome Res* 2013, 12:5666-5680.
12. Stosová Tá, Sebela M, Rehulka P, Sedo O, Havlis J, Zdrahal Z: Evaluation of the possible proteomic application of trypsin from *Streptomyces griseus*. *Anal Biochem* 2008, 376:94-102.
13. Zhang Y, Huang H, Yao X, Du G, Chen J, Kang Z: High-yield secretory production of stable, active trypsin through engineering of the N-terminal peptide and self-degradation sites in *Pichia pastoris*. *Bioresour Technol* 2017, 247:81-87.
14. Vestling MM, Murphy CM, Fenselau C: Recognition of trypsin autolysis products by high-performance liquid chromatography and mass spectrometry. *Anal Chem* 1990, 62:2391-2394.
15. Harris WA, Janecki DJ, Reilly JP: Use of matrix clusters and trypsin autolysis fragments as mass calibrants in matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom* 2002, 16:1714-1722.
16. Nohara D, Sugiura H, Sakakibara H, Matsubara M, Kojima S, Miura K, Sakai T: High performance in refolding of *Streptomyces griseus* trypsin by the aid of a mutant of *Streptomyces* subtilisin inhibitor designed as trypsin inhibitor. *J Biochem* 1999, 125:343-347.
17. Page MJ, Wong S-L, Hewitt J, Strynadka NC, MacGillivray RT: Engineering the primary substrate specificity of *Streptomyces griseus* trypsin. *Biochemistry* 2003, 42:9060-9066.
18. Oh EA, Kim MS, Chi WJ, Kim JH, Hong SK: Characterization of the *sgtR1* and *sgtR2* genes and their role in regulating expression of the *sprT* gene encoding *Streptomyces griseus* trypsin. *FEMS Microbiol Lett* 2007, 276:75-82.
19. Chi WJ, Song JH, Oh EA, Park SW, Chang YK, Kim ES, Hong SK: Medium optimization and application of affinity column chromatography for trypsin production from recombinant *Streptomyces griseus*. *J Microbiol Biotechnol* 2009, 19:1191-1196.
20. Ling Z, Ma T, Li J, Du G, Kang Z, Chen J: Functional expression of trypsin from *Streptomyces griseus* by *Pichia pastoris*. *J Ind Microbiol Biotechnol* 2012, 39:1651-1662.
21. Ling Z, Liu Y, Teng S, Kang Z, Zhang J, Chen J, Du G: Rational design of a novel propeptide for improving active production of *Streptomyces griseus* trypsin in *Pichia pastoris*. *Appl Environ Microbiol* 2013, 79:3851-3855.
22. Zhang Y, Ling Z, Du G, Chen J, Kang Z: Improved production of active *Streptomyces griseus* trypsin with a novel auto-catalyzed strategy. *Sci Rep* 2016, 6:23158.
23. Hohenblum H, Gasser B, Maurer M, Borth N, Mattanovich D: Effects of gene dosage, promoters, and substrates on unfolded protein stress of recombinant *Pichia pastoris*. *Biotechnol Bioeng* 2004, 85:367-375.
24. Tu BP, Ho-Schleyer SC, Travers KJ, Weissman JS: Biochemical basis of oxidative protein folding in the endoplasmic reticulum. *Science* 2000, 290:1571-1574.

25. Evnin LB, Vásquez JR, Craik CS: Substrate specificity of trypsin investigated by using a genetic selection. *Proc Natl Acad Sci U S A* 1990, 87:6659-6663.
26. Arnold FH: Directed evolution: bringing new chemistry to life. *Angew Chem Int Ed Engl* 2018, 57:4143-4148.
27. Pires DE, Ascher DB, Blundell TL: DUET: a server for predicting effects of mutations on protein stability using an integrated computational approach. *Nucleic Acids Res* 2014, 42:314-319.
28. Dehouck Y, Kwasigroch JM, Gilis D, Rooman M: PoPMuSiC 2.1: a web server for the estimation of protein stability changes upon mutation and sequence optimality. *BMC Bioinformatics* 2011, 12:151-162.
29. Giollo M, Martin AJM, Walsh I, Ferrari C, Tosatto SCE: NeEMO: a method using residue interaction networks to improve prediction of protein stability upon mutation. *Bmc Genomics* 2014, 15 (Suppl 4):57-67.
30. Xia Y, Cui W, Cheng Z, Peplowski L, Liu Z, Kobayashi M, Zhou Z: Improving the thermostability and catalytic efficiency of the subunit-fused nitrile hydratase by semi-rational engineering. *Chemcatchem* 2018, 10:1370-1375.
31. Craik CS, Roczniak S, Largman C, Rutter WJ: The catalytic role of the active site aspartic acid in serine proteases. *Science* 1987, 237:909-913.
32. Hedstrom L: Serine protease mechanism and specificity. *Chem Re* 2002, 102:4501-4523.
33. Yang Z, Zhang Z: Engineering strategies for enhanced production of protein and bio-products in *Pichia pastoris*: A review. *Biotechnol Adv* 2018, 36:182-195.
34. Vanz AL, Nimtz M, Rinas U: Decrease of UPR- and ERAD-related proteins in *Pichia pastoris* during methanol-induced secretory insulin precursor production in controlled fed-batch cultures. *Microb Cell Fact* 2014, 13:23.
35. Delic M, Rebnegger C, Wanka F, Puxbaum V, Haberhauer-Troyer C, Hann S, Kollensperger G, Mattanovich D, Gasser B: Oxidative protein folding and unfolded protein response elicit differing redox regulation in endoplasmic reticulum and cytosol of yeast. *Free Radic Biol Med* 2012, 52:2000-2012.
36. Gu L, Zhang J, Du G, Chen J: Multivariate modular engineering of the protein secretory pathway for production of heterologous glucose oxidase in *Pichia pastoris*. *Enzyme Microb Technol* 2015, 68:33-42.
37. Ron D, Walter P: Signal integration in the endoplasmic reticulum unfolded protein response. *Nat Rev Mol Cell Biol* 2007, 8:519-529.
38. Aalto MK, Jantti J, Ostling J, Keranen S, Ronne H: Mso1p: A yeast protein that functions in secretion and interacts physically and genetically with Sec1p. *Proc Natl Acad Sci U S A* 1997, 94:7331-7336.
39. Yamamoto T, Yamamoto D, Rokugawa K, Yoshimura K, Imura Y, Yoshimura E, Suzuki M: Decreased aluminium tolerance in the growth of *Saccharomyces cerevisiae* with SSO2 gene disruption. *BioMetals* 2018, 31:1-13.

40. Ding HC, Li DF, Wei XY, Huang YW, Cui S, Xie HJ, Zhou T: Protein-peptide nutritional material prepared from surimi wash-water using immobilized chymotrypsin-trypsin. *J Sci Food Agric* 2016, 97:1746-1752.
41. Chunsheng Leng HC, Shunan Li Application of recombinant enzymes in the production of human insulin. *Journal of Chinese Pharmaceutical Sciences* 2014, 23:335-337.
42. Buettner K, Kreisig T, Strater N, Zuchner T: Protein surface charge of trypsinogen changes its activation pattern. *BMC Biotechnol* 2014, 14:109.
43. Paulova L, Hyka P, Branska B, Melzoch K, Kovar K: Use of a mixture of glucose and methanol as substrates for the production of recombinant trypsinogen in continuous cultures with *Pichia pastoris* Mut^+. *J Biotechnol* 2012, 157:180-188.
44. Deng Y, Gruppen H, Wierenga PA: Comparison of protein hydrolysis catalyzed by bovine, porcine, and human trypsins. *J Agric Food Chem* 2018, 66:4219-4232.
45. Wang Z, Wang Y, Zhang D, Li J, Hua Z, Du G, Chen J: Enhancement of cell viability and alkaline polygalacturonate lyase production by sorbitol co-feeding with methanol in *Pichia pastoris* fermentation. *Bioresour Technol* 2010, 101:1318-1323.
46. Xie T, Liu Q, Xie F, Liu H, Zhang Y: Secretory expression of insulin precursor in *Pichia pastoris* and simple procedure for producing recombinant human insulin. *Prep Biochem Biotechnol* 2008, 38:308-317.
47. Bairoch A, Apweiler R: The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Res* 2000, 28:45-48.
48. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K: Scalable molecular dynamics with NAMD. *J Comput Chem* 2005, 26:1781-1802.
49. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, et al: All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* 1998, 102:3586-3616.
50. Read RJ, James MNG: Refined crystal structure of *Streptomyces griseus* trypsin at 1.7 Å resolution. *J Mol Biol* 1988, 200:523-551.
51. Cheng ZY, Cui WJ, Xia YY, Peplowski L, Kobayashi M, Zhou ZM: Modulation of nitrile hydratase regioselectivity towards dinitriles by tailoring the substrate binding pocket residues. *Chemcatchem* 2018, 10:449-458.
52. DeLano WL: The PyMOL molecular graphics system. DeLano Scientific: San Carlos, CA, USA. 2002.
53. Hayes R, Myers P, Edge T, Zhang H: Monodisperse sphere-on-sphere silica particles for fast HPLC separation of peptides and proteins. *Analyst* 2014, 139:5674-5677.

Figures

A

R21 R32
FVEFVVGGTRAAQGEFPFMVRLS MG CGGALYAQ DIVLT AA HCVSGSGNNT

SI TATGG VVDLQSS SAVKVR STKV L QA PGY NGTGKDWALI KLAQPI NQ PT

K122 R153
LKI ATTTAYNQGTFTVAGWG ANIEGGSQQRYLLK ANV PFV SDAACRSAYGNE

R201
LVA NEEICAGYPDTGGVDTCQGDSGGPMFRKDNADEWIQVGIVSWGYGCA R

PGYPGVYTE VS T FAS AIASAARTL

B

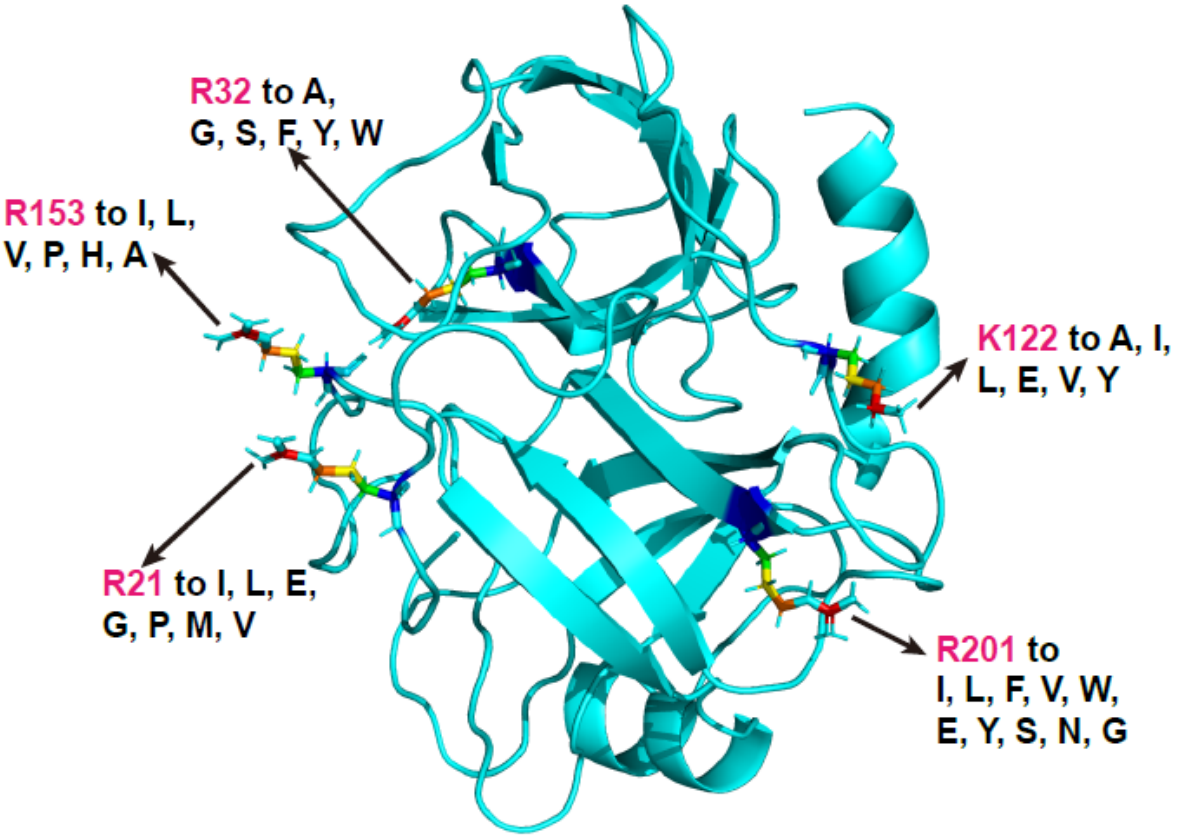

Figure 2

Identification of the autolysis residues and predicted mutations. (A) The peptide fragments detected from hydrolysate of tbcf (K101A). The peptide fragments were presented in colored sequences. (B) The predicted mutations of autolysis residues in tbcf (K101A).

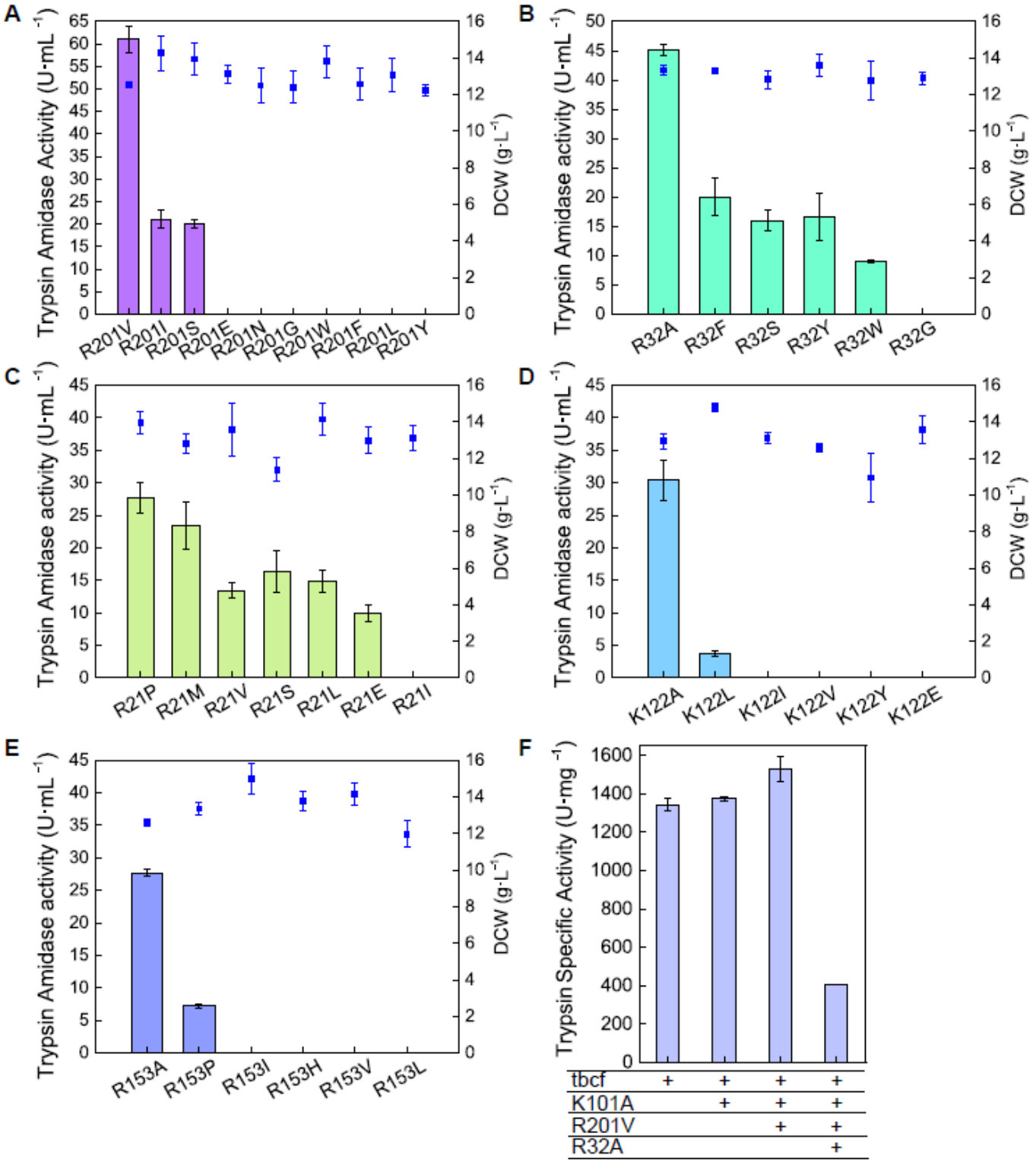

Figure 4

Amidase activity of trypsin variant. Trypsin amidase activity ( ), DCW (□). (A) Variant of R201 mutation, (B) variant of R32 mutation, (C) variant of R21 mutation, (D) variant of K122 mutation, (E) variant of R153 mutation, and (F) the specific activity of trypsin mutant.

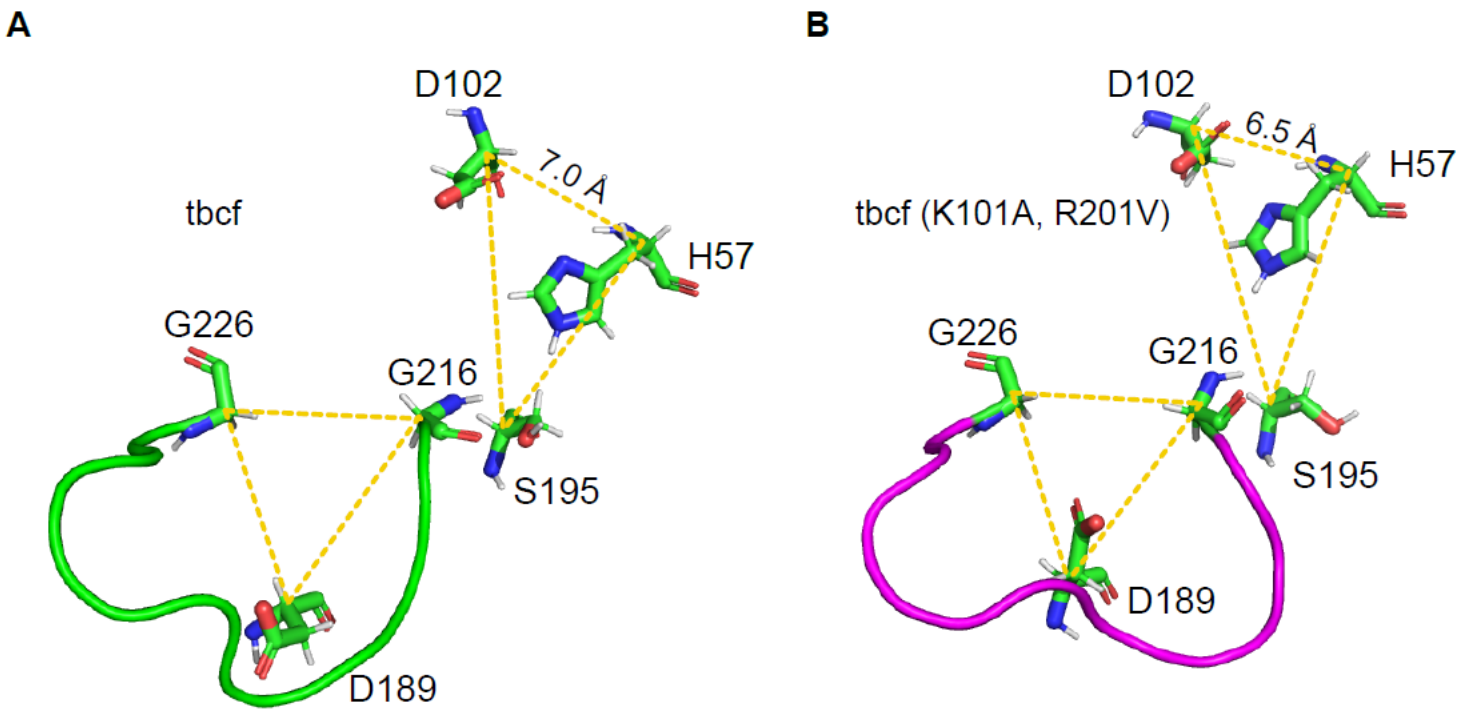

Figure 6

Comparison of the three-dimensional structure of the trypsin mutant. (A) and (B) presented the three-dimensional structure of tbcf and tbcf (K101A, R201V), respectively.

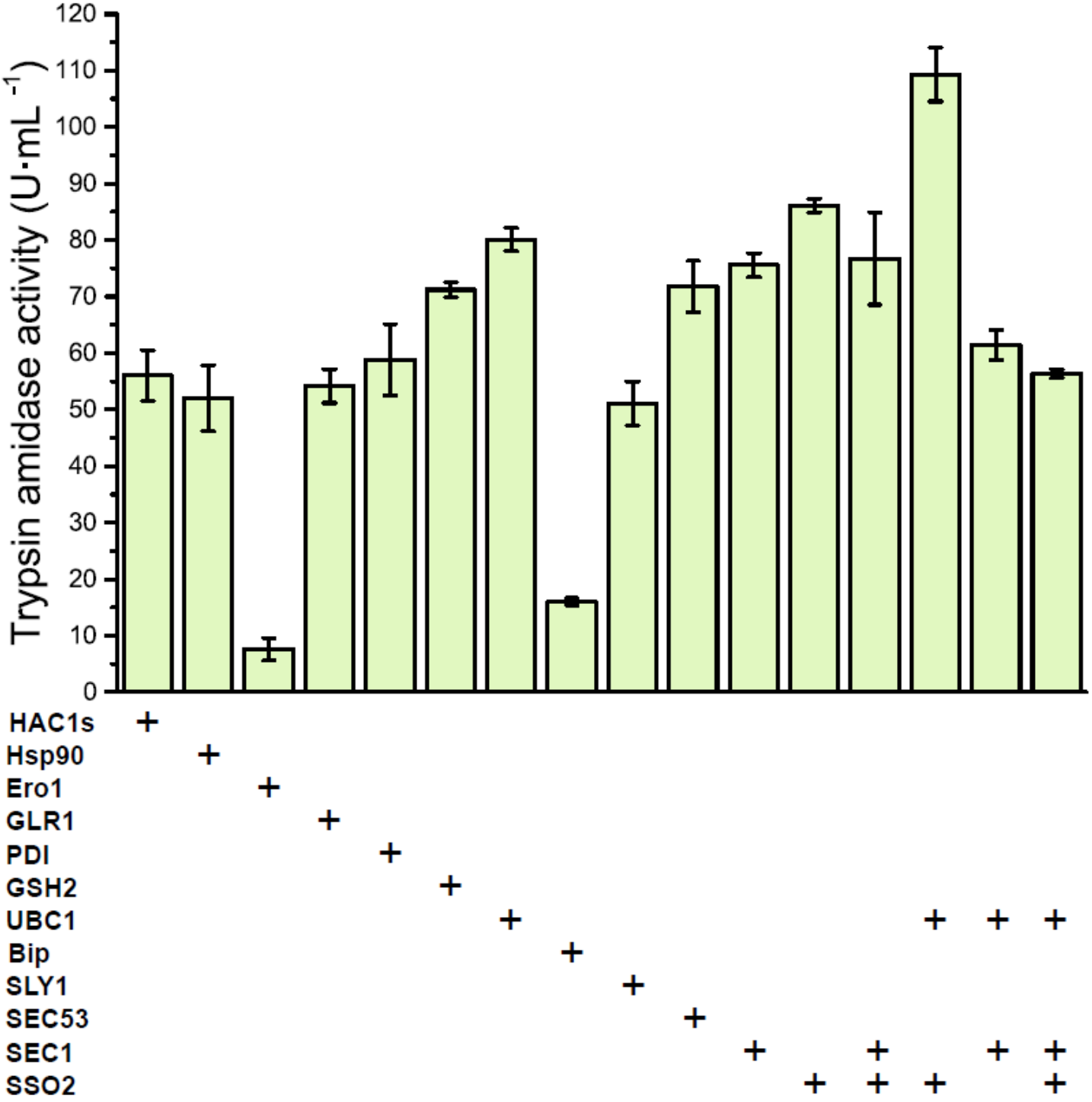

Figure 8

The trypsin amidase activity of yeast strain with overexpression of endogenous proteins.

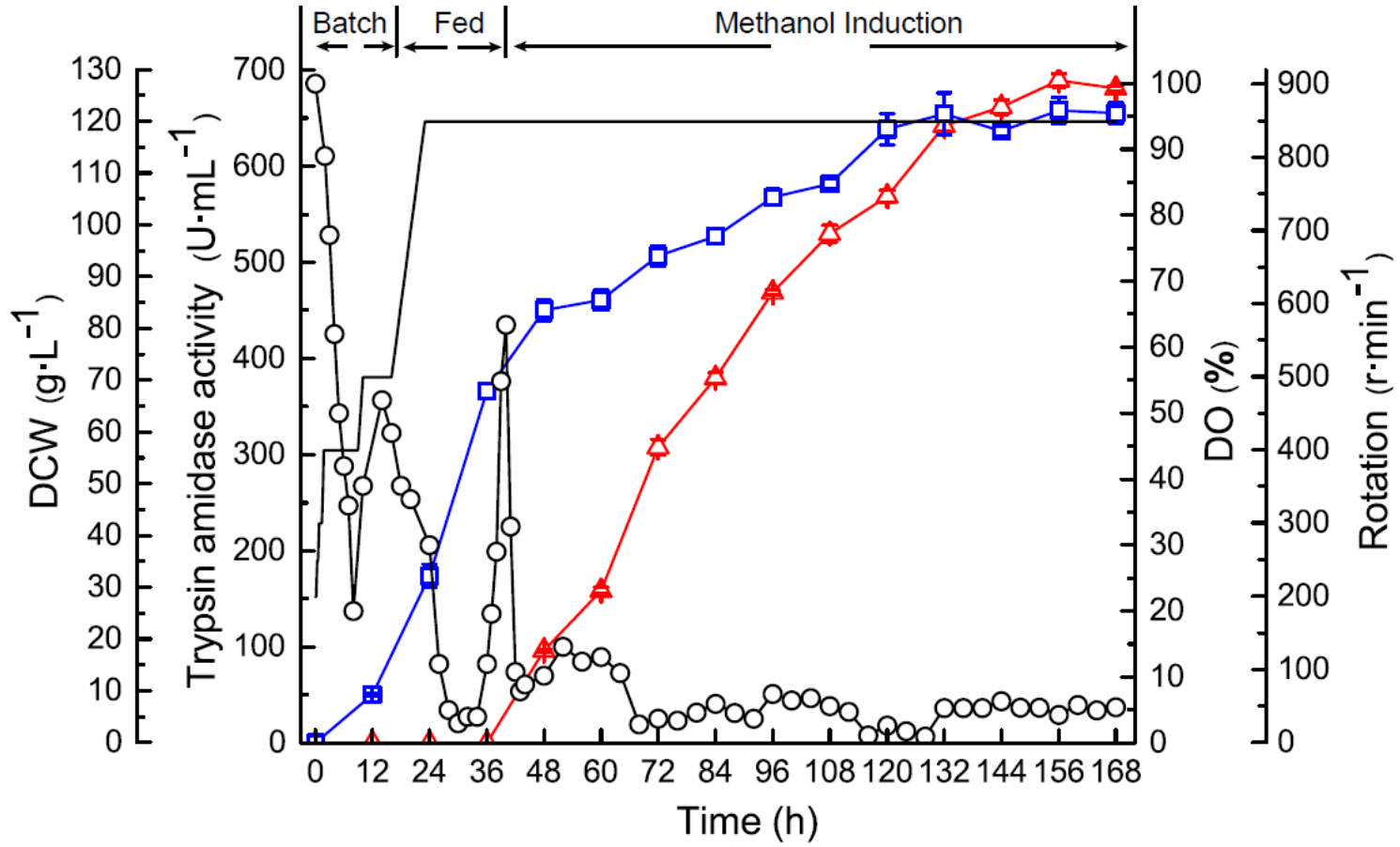

Figure 9

Fed-batch fermentation of yeast strain in 3-L bio-reactor. Trypsin amidase activity (△), dried cell weight (DCW, ▯), dissolved oxygen (DO, ●), and rotation speed

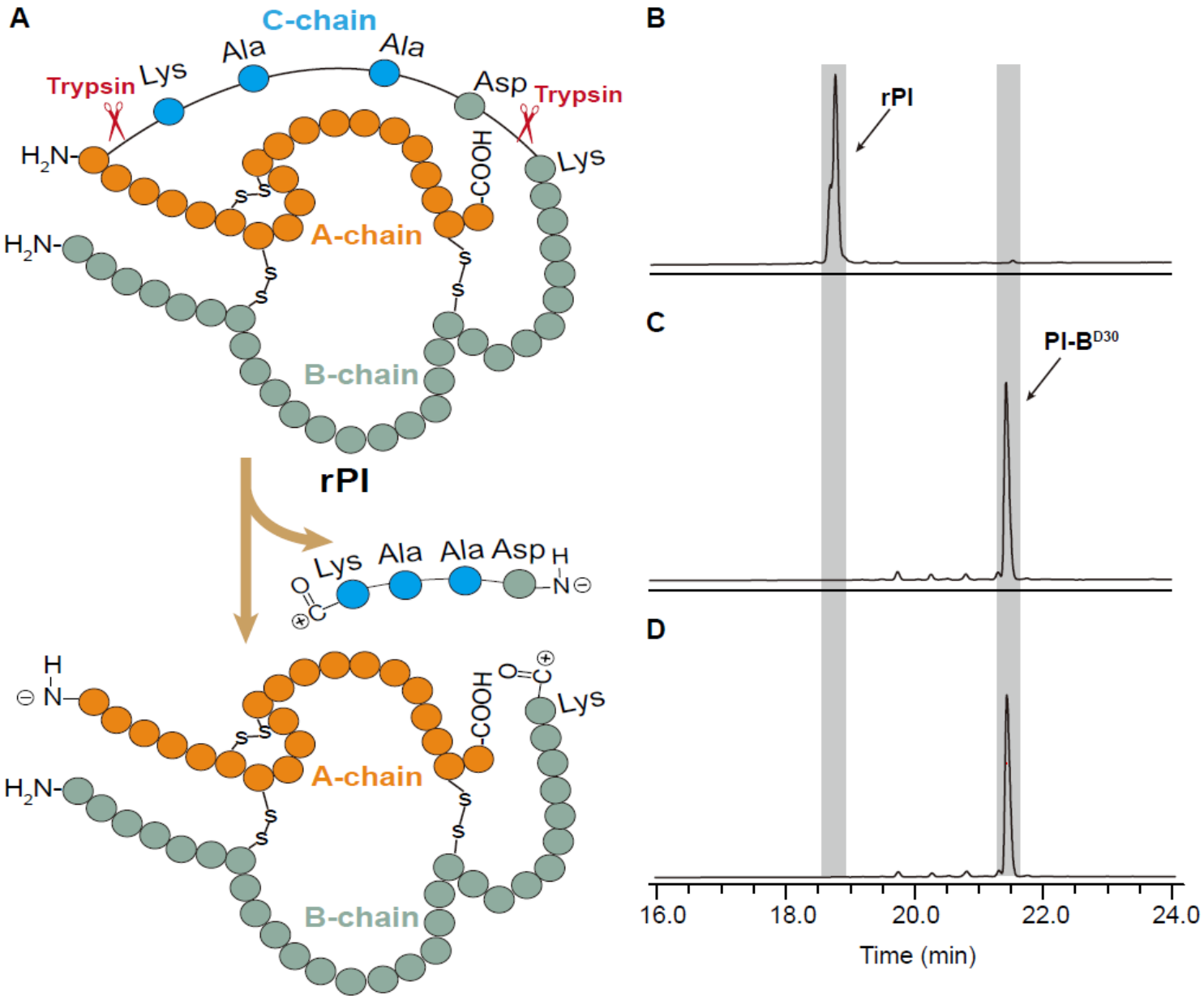

Figure 11

The HPLC chromatography of manufacturing insulin with trypsin. (A) Schematic presentation of hydrolysis procession. rPI, recombination insulin precursor, PI-BD30, insulin precursor with Asp30 deleted B-chain. (B)-(D) The chromatography of hydrolysis product of rPI after incubation with trypsin at 25 ºC for 19 h. (B) rPI, (C) rPI and commercial porcine trypsin, and (D) rPI and tbcf (K101A, R201V).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementrayMaterial.docx
- SupplementrayMaterial.docx

- Graphicabstract.pdf
- Graphicabstract.pdf