Published in final edited form as:

Plant Cell Rep. 2015 March ; 34(3): 507–516. doi:10.1007/s00299-014-1730-4.

C-Terminally fused affinity *Strep*-tag II is removed by proteolysis from recombinant human erythropoietin expressed in transgenic tobacco plants

Farooqahmed S. Kittur,

Department of Pharmaceutical Sciences, Biomanufacturing Research Institute & Technology Enterprise, North Carolina Central University, Durham, NC 27707, USA

Mallikarjun Lalgondar,

Center for Agribusiness Excellence, Tarleton State University, Stephenville, TX 76402, USA

Chiu-Yueh Hung,

Department of Pharmaceutical Sciences, Biomanufacturing Research Institute & Technology Enterprise, North Carolina Central University, Durham, NC 27707, USA

David C. Sane, and

Carilion Clinic and Virginia Tech Carilion School of Medicine, Roanoke VA 24014, USA

Jiahua Xie

Department of Pharmaceutical Sciences, Biomanufacturing Research Institute & Technology Enterprise, North Carolina Central University, Durham, NC 27707, USA. 1801 Fayetteville Street, Department of Pharmaceutical Sciences, Biomanufacturing Research Institute & Technology Enterprise, North Carolina Central University, Durham, NC 27707, USA; Phone:+1 919 530 6705; Fax: +1 919 530 6600

Jiahua Xie: jxie@nccu.edu

Abstract

Asialo-erythropoietin (asialo-EPO), a desialylated form of EPO, is a potent tissue-protective agent. Recently, we and others have exploited a low cost plant-based expression system to produce recombinant human asialo-EPO (asialo-rhuEPO^P). To facilitate purification from plant extracts, *Strep*-tag II was engineered at the C-terminus of EPO. Although asialo-rhuEPO^P was efficiently expressed in transgenic tobacco plants, affinity purification based on *Strep*-tag II did not result in the recovery of the protein. In this study, we investigated the stability of *Strep*-tag II tagged asialo-rhuEPO^P expressed in tobacco plants to understand whether this fused tag is cleaved or inaccessible. Sequencing RT-PCR products confirmed that fused DNA sequences encoding *Strep*-tag II were properly transcribed, and three-dimensional protein structure model revealed that

Correspondence to: Jiahua Xie, jxie@nccu.edu.

Author contributions

Conceived and designed the experiments: J.H. Xie, F.S. Kittur and D.C. Sane. Performed the experiments: F.S. Kittur, M. Lalgondar and C.-Y. Hung. Analyzed the data: F.S. Kittur, M. Lalgondar, C.-Y. Hung and J.H. Xie. Wrote the paper: F.S. Kittur, J.H. Xie and D.C. Sane.

Conflict of interest

The authors declare that they have no conflict of interest.

the tag must be fully accessible. However, Western blot analysis of leaf extracts and purified asialo-rhuEPO^P revealed that the *Strep*-tag II was absent on the protein. Additionally, no peptide fragment containing *Strep*-tag II was identified in the LC-MS/MS analysis of purified protein further supporting that the affinity tag was absent on asialo-rhuEPO^P. However, *Strep*-tag II was detected on asialo-rhuEPO^P that was retained in the endoplasmic reticulum, suggesting that the *Strep*-tag II is removed during protein secretion or extraction. These findings together with recent reports that C-terminally fused *Strep*-tag II or IgG Fc domain are also removed from EPO in tobacco plants, suggest that its C-terminus may be highly susceptible to proteolysis in tobacco plants. Therefore, direct fusion of purification tags at the C-terminus of EPO should be avoided while expressing it in tobacco plants.

Keywords

Erythropoietin; Plant-based expression; Affinity tag; Proteolysis

Introduction

Asialo-erythropoietin (asialo-rhuEPO) is an erythropoietin (EPO) derivative obtained by removal of sialic acid residues from mammalian cell-produced recombinant human EPO (rhuEPO^M) (Erbayraktar et al. 2003). It is reported to be non-erythropoietic, but tissue-protective (Erbayraktar et al. 2003). It has been shown to protect brain, heart and kidneys from different injuries (Leist et al. 2004; Okada et al. 2007; Takeyama et al. 2012) without causing any adverse effects associated with its erythropoietic activity. Recent studies have also shown that asialo-rhuEPO can benefit diabetes patients by promoting β -cell growth (Choi et al. 2010). Since the current method for the production of asialo-rhuEPO from rhuEPO^M by enzymatic desialylation is not economically viable, we (Kittur et al. 2012; 2013) and Parson et al. (2012) stably co-expressed human *EPO* and β 1,4-galactosyltransferase (*GalT*) genes in tobacco plants and moss, respectively to produce asialo-rhuEPO^P. These studies have demonstrated that asialo-rhuEPO can accumulate at high levels in transgenic plants (Parson et al. 2012; Kittur et al. 2013), and that plants could be used as a bioreactor for the production of this medically important glycoprotein.

Following advances in optimization of *EPO* gene transcription and translation in plants, the focus is now shifted to downstream processing, which represents up to 80% of overall production costs (Chen 2008; Wilken and Nikolov 2012). Development of efficient extraction and purification processes is therefore, essential for favorable economics. The most effective and commonly used method for protein purification is affinity purification, wherein the protein of interest is appended with a suitable affinity tag, and enriched by virtue of its specific binding properties to an immobilized ligand (Lichty et al. 2005). This strategy was employed to facilitate purification of EPO from transgenic plant leaf tissues. The EPO polypeptide chain was C-terminally fused with a *Strep*-tag II (Conley et al. 2009; Jez et al. 2013) or an IgG Fc domain (Castilho et al. 2011; Castilho et al. 2013). Surprisingly, in all these studies, purification based on these tags did not result in the recovery of the protein from plant extracts. Castilho et al. (2013) observed that a large amount of Fc fragment accumulated in tobacco plants, suggesting that EPO-Fc fusion

protein may be cleaved in plants. The fact that C-terminally His-tagged EPO (Debeljak et al. 2006) or EPO-Fc fusion protein (Lattenmayer et al. 2007) expressed in mammalian cells are stable, suggests that C-terminal region of EPO may be susceptible to proteolysis only in plant cells. Furthermore, it is not known whether these tags are removed by proteolysis during protein secretion or extraction since unintended proteolysis has been recognized recently as a serious problem in utilizing plant-based platforms for the production of pharmaceutical proteins (Benchabane et al. 2008).

In our recent asialo-rhuEPO expression study (Kittur et al. 2013), EPO protein was appended with a C-terminal TEV protease cleavage site, a *Strep*-tag II, followed by an endoplasmic reticulum (ER) retention signal KDEL. Western blot analysis of transgenic tobacco plant leaf extracts showed high levels of asialo-rhuEPO^P. However, our efforts to recover protein using the *Strep*-Tactin column failed, consistent with the observations made by others (Conley et al. 2009; Castilho et al. 2011, Castilho et al. 2013; Jez et al. 2013). In this study, we set out to determine the stability of *Strep*-tag II tagged asialo-rhuEPO^P expressed in transgenic tobacco plants. The objective was to understand whether this fused tag is inaccessible or removed from the protein. If the *Strep*-tag II is removed, whether the removal of tag occurs *in planta* or *ex planta*.

Materials and Methods

Plant materials

Leaf tissues harvested from greenhouse grown transgenic tobacco line A56-5 with high expression of asialo-rhuEPO^P (Kittur et al. 2013) were used for DNA, RNA and protein purification throughout the studies unless otherwise stated.

PCR and RT-PCR analyses

To confirm the presence of nucleotide sequences encoding TEV protease cleavage site, *Strep*-II tag and KDEL fused to the 3'-end of *EPO* gene sequence in transgenic plants, PCR was performed with primers EPOF: 5'-GCATGTGGATAAAGCCGTCAGT-3' and NosTR: 5'-TATATGATAATCATCGCAAGAC-3'. The genomic DNAs of transgenic tobacco leaves were isolated using DNeasy Plant Mini Kit (Qiagen, Germantown, MD, USA). Plasmid DNA CEJ902 containing *EPO* gene and genomic DNA isolated from transgenic tobacco plant containing *GUS* gene (A1K1) (Musa et al. 2009) were used as positive and negative controls, respectively. Either 100 ng genomic DNA or 10 ng plasmid DNA was used as template. RT-PCR was performed to detect *EPO* transcripts whether containing nucleotide sequences for TEV protease cleavage site, *Strep*-II tag and KDEL. The RNA samples were first isolated from transgenic tobacco leaves using RNeasy Plant Mini Kit (Qiagen) and used for making first strand cDNA by the High-Capacity cDNA Archive kit (Applied Biosystems, Foster City, CA, USA) according to the manufacture instructions with the oligo(dT)18 primer. The subsequent PCR had the same reaction conditions as genomic DNA PCR except that cDNA made from 25 ng of original total RNA was used as template. The RT-PCR product was ligated to TA pCR[®]2.1 vector (Invitrogen, Carlsbad, CA, USA) and sequenced by Eurofins Genomics (Huntsville, AL, USA).

Purification of asialo-rhuEPO^P via *Strep*-tag II

For purification of *Strep*-tag II tagged asialo-rhuEPO^P, 200 g of leaf tissues were grounded into fine powder in liquid nitrogen, and extracted with 0.1 M Tris-HCl buffer (pH 8.0) containing 150 mM NaCl, 1 mM EDTA and plant protease inhibitor cocktail (Sigma-Aldrich, Saint Louis, MO, USA). The extract was then passed through a double layer of miracloth (EMD Millipore, Bellerica, MA, USA), followed by centrifugation at 20,000 g for 15 min to remove insoluble plant debris. The centrifugation step was repeated once in order to remove the remaining insolubles. The clear extract was then applied to a *Strep*-Tactin Sepharose column (5 ml, GE Healthcare BioSciences, Pittsburgh, PA, USA), which had been equilibrated with 10 column volumes (CV) of 0.1 M Tris-HCl buffer (pH 8.0) containing 150 mM NaCl (binding buffer). After washing the column with 6 CV of the binding buffer, bound asialo-rhuEPO^P was eluted with binding buffer containing 2.5 mM desthiobiotin. Asialo-rhuEPO^P elution was monitored using a sandwich ELISA as described previously (Kittur et al. 2012).

Purification of asialo-rhuEPO^P using ion-exchange and immunoaffinity chromatography

To obtain pure asialo-rhuEPO^P for Western blotting and LC-MS-MS analysis, the protein was purified from A56-5 leaf tissues using cation-exchange and immunoaffinity chromatography (IAC) (Kittur et al. 2014). Briefly, 100 g of leaf tissue was extracted with 0.1 M acetate buffer (pH 5) containing 1 mM EDTA, 2% PVPP and plant protease inhibitor cocktail. The extract was clarified by centrifugation and then applied to a SP Sepharose FF column. The asialo-rhuEPO^P peak fractions were collected and further purified using an IAC column. Asialo-rhuEPO^P in each IAC fraction was determined using a sandwich ELISA as described previously (Kittur et al. 2013). Residual BSA in the purified fraction was removed using MelonTM gel BSA removal kit (Thermo Scientific, Waltham, MA, USA).

Microsomes isolation and apoplast protein extraction

Microsomes were isolated from fresh leaves of A56-5 transgenic tobacco plants following the method described by Bakker et al. (2006). Microsomes were reconstituted in PBS before analysis by ELISA and Western blotting. For the apoplast protein extraction, a vacuum infiltration procedure described by Dani et al. (2005) was adopted with some modifications. Briefly, 2 cm² strips of leaf tissues were excised with a razor blade avoiding the major veins, washed with water and submerged in a chilled vacuum infiltration buffer (0.1 M sodium phosphate buffer, pH 6.5) placed in a vacuum flask. Vacuum was applied for 75 s to remove the gas from the apoplastic spaces. The flask was shaken during evacuation to remove air bubbles from the leaf surfaces. The vacuum was released gradually; then excess buffer was removed by drying the leaf strips with Kimwipe tissues. Leaf strips were carefully rolled into a swiss-roll cylinder. Cylinders with leaf tissues were centrifuged on VivaSpin@20 centrifugal devices fitted with wire mesh inserts (made in the lab) at 900 g for 10 min at 4 °C. The vacuum infiltrate (VI) was collected at the bottom of the tubes. Since the protein concentration of VI was low, it was concentrated using a Vivaspin@2 10 kD cutoff centrifugal device. The concentrate was stored at -20 °C until analysis by ELISA.

SDS-PAGE and LC-MS/MS analysis

Immunoaffinity purified asialo-rhuEPO^P was separated on a 12.5% SDS-PAGE gel under reducing conditions following method described by Laemmli (1970). After protein separation, the gel was stained with Coomassie brilliant blue. A 30 kD band corresponding to asialo-rhuEPO^P was excised, and processed at Duke University Proteomics Facility to identify EPO peptides using LC-MS/MS following digestion with trypsin.

Western blot analysis

For Western blot analysis, the crude extracts, microsomes, and purified asialo-rhuEPO^P were first separated on a 12.5% SDS-PAGE gel as described above. Following separation, the proteins were transferred onto a PVDF membrane using a CAPS (pH 11) transfer buffer. The blots were then blocked with 5% BSA in 1X PBST for 1 h at room temperature. For detection of *Strep*-tag II, the blot was incubated with a mouse monoclonal anti-*Strep*-tag II antibody (EMD Millipore, Bellerica, MA, USA) at a dilution of 1:1000 in 5% BSA in 1X PBST overnight at 4°C. Five to 15 ng recombinant human protein Jumonji Domain Containing 2E (JMJD2E) with *Strep*-tag II (JMJD2E+*Strep*-tag II) (Cayman Chemicals, Ann Arbor, MI, USA) was used as a positive control, whereas rhuEPO^M without *Strep*-tag II tag was used as a negative control. After incubating the blot with the above antibody, the blot was washed thrice for 10 min each with 1X PBST. Anti-mouse secondary antibody conjugated to horse radish peroxidase (1:2000 diluted in 1X PBST containing 5% BSA) was then added and incubated at room temperature for 1 h. Following washing with 1X PBST, the immunoreactive bands were revealed using a Super Signal West Pico chemiluminescent substrate (Thermo Scientific, Waltham, MA, USA). The blot was then stripped and probed with anti-EPO antibody as described previously (Kittur et al. 2012) to reveal asialo-rhuEPO^P.

Protein structure modeling

To generate a three-dimensional homology model of EPO containing the synthetic peptide, the protein sequence was submitted to The Protein Model Portal (<http://www.proteinmodelportal.org>) (Haas et al. 2013). The portal attempts to obtain homology model from different automated structure prediction tools. Among them the homology model generated by HHpred (Söding et al. 2005) was used for further analysis. The HHpred implements protein sequence comparison of profile hidden markov models (HMMs), which is sensitive in identifying homologs to select template structures and generates homology model utilizing the MODELLER package (Sali et al. 1995). The quality of the structural models was assessed with PROCHECK (Laskowski et al. 1993). Superimposition of structures and rendering figures was performed with the molecular modeling software Chimera (Pettersen et al. 2004).

Results and Discussion

Plants lack sialylating capacity (Wee et al. 1998), but can perform complex *N*-glycosylation (Gomord and Faye 2004). We exploited this unique property of plant-based expression system to express asialo-rhuEPO (Kittur et al. 2012; Kittur et al. 2013). The expression levels of asialo-rhuEPO^P reached 2.4 µg/g of fresh leaf tissue in the highest expressing

transgenic tobacco line (Kittur et al. 2013). To facilitate purification, *Strep*-tag II was fused to the C-terminus of EPO while a TEV protease cleavage site was also engineered to aid in the removal of tag after purification and, a KDEL sequence was appended after *Strep*-tag II to retain the protein in the ER (Kittur et al. 2013). *Strep*-tag II was selected because it is short, biologically inert, and proteolytically stable (Schmidt and Skerra 2007). It exhibits intrinsic affinity towards *Strep*-Tactin, a specifically engineered streptavidin. By exploiting the highly specific interaction, *Strep*-tag II fused proteins can be isolated in one step from crude extracts under gentle conditions, which is especially suited for purification of functional proteins. To purify asialo-rhuEPO^P, leaf extracts of transgenic line A56-5 containing ~66 µg of asialo-rhuEPO^P (as determined by ELISA) were applied to a *Strep*-Tactin Sepharose column. Fig. 1a shows the elution profile of asialo-rhuEPO^P. As is seen, the total calculated asialo-rhuEPO^P collected from fraction 8, 9, 10 and 11 was about 1.6 µg. These results indicate that only 2.4% of asialo-rhuEPO^P was recovered from the *Strep*-Tactin Sepharose column. Since it is known that plants contain high levels of free and conjugated biotin (Baldet et al. 1993), and that biotin can irreversibly inactivate *Strep*-Tactin, we suspected that biotin may be responsible for poor binding of asialo-rhuEPO^P to *Strep*-Tactin Sepharose column. To rule out this possibility, asialo-rhuEPO^P was first precipitated with 70% ammonium sulfate to remove free biotin. The precipitated asialo-rhuEPO^P was then dissolved in *Strep*-Tactin binding buffer and applied to the column. Despite the removal of free biotin, binding of asialo-rhuEPO^P to the *Strep*-Tactin Sepharose column did not improve, only 3% was recovered (data not shown). In another experiment, purified asialo-rhuEPO^P (as described in “Materials and Methods” section) which was devoid of free and conjugated biotin was applied to the column. Only small amount of asialo-rhuEPO^P was bound to the column, whereas the rest was seen in pass through and wash fractions (Fig. 1b). Based on the quantification from ELISA, only about 4% of pure asialo-rhuEPO^P recovered from the column. These results indicate that free or conjugated biotin was not responsible for poor binding of asialo-rhuEPO^P to the *Strep*-Tactin Sepharose column. Previous studies have also shown that EPO-*Strep*-tag II fusion protein expressed in tobacco plants could not be recovered using *Strep*-Tactin (Conley et al. 2009; Jez et al. 2013). Our results together with previous reports suggest that *Strep*-tag II on rhuEPO is either inaccessible or not present on the protein.

The possibility that the *Strep*-tag II is inaccessible to *Strep*-Tactin is very unlikely because crystallographic studies (Syed et al. 1998) have shown that the extreme C-terminus (Arg¹⁶²-Arg¹⁶⁶) of EPO protrudes out of the compact protein core and is fully exposed (Fig. 2a). Furthermore, EPO expressed in mammalian cells with a C-terminal His-tag has been reported to readily bind to the Nickel column (Debeljak et al. 2006), indicating that the C-terminus is fully accessible. In some cases, however, fusion of certain affinity tags has been reported to alter protein conformation and activity (Sabaty et al. 2013; Majorek et al. 2014). It is therefore, possible that the synthetic peptide containing *Strep*-tag II, TEV protease cleavage site and an ER retention signal may affect the conformation of EPO such that the *Strep*-tag II becomes less accessible to *Strep*-Tactin. To investigate this possibility, we generated a three-dimensional homology model of EPO containing the synthetic peptide using HHpred. The HMMs profile based sequence search identified human erythropoietin (EPO) with high significance, and the homology model was generated using the crystal

structure of EPO protein (PDB-ID: 1EER, chain A) (Syed et al. 1998). Analysis of the predicted structure showed 94% of amino acids in the allowed region of the Ramachandran plot which is comparable to 96% of amino acids in the allowed region for the template structure. The predicted three-dimensional model indicated that the synthetic peptide, like the native C-terminus of EPO protein (Fig. 2a), remains flexible and fully exposed (Fig. 2b). Furthermore, superimposing it with native EPO structure (Fig. 2a) using molecular modeling software Chimera gave a RMSD value of 0.2 Å, indicating that the main chain conformation of EPO protein is unaffected by the fusion of synthetic peptide (Fig. 2c). The above analysis suggests that fused synthetic peptide does not alter the EPO structure, and that the synthetic peptide containing *Strep*-tag II could be still fully exposed. Therefore, it is very unlikely that the C-terminally fused *Strep*-tag II on asialo-rhuEPO^P is inaccessible to *Strep*-Tactin.

The other possibility that can explain the lack of binding of asialo-rhuEPO^P to *Strep*-Tactin is the absence of *Strep*-tag II on the protein. This could happen because of two reasons: (1) the truncation of *EPO* gene during integration into plant genome or transcription; (2) proteolytic removal of *Strep*-tag II. To rule out the likelihood of gene truncation, a forward primer (EPOF) based on *EPO* sequence and a reverse primer (NosTR) based on *Nos* terminator were designed and used to amplify genomic DNA sequences containing the synthetic peptide sequences (Fig. 3a). A PCR product of ~500 bp was amplified from transgenic tobacco line A56-5 harboring the transgene containing the synthetic peptide sequences (Fig. 3b, lane 2). The size of genomic PCR product was same as that of a PCR product amplified from plasmid DNA containing the sequence encoding EPO-*Strep*-tag II fusion protein (Fig. 3b, lane 3). These results indicate that the *EPO*, including the sequences encoding the synthetic peptide, was incorporated into plant genome. To determine whether *Strep*-tag II was lost because of generation of partial *EPO* transcripts, RT-PCR was performed on total RNA isolated from transgenic tobacco line A56-5 with same set of primers used above. A ~500 bp RT-PCR product was amplified (Fig. 3c) and subsequently sequenced. The results showed that nucleotide sequences encoding TEV protease cleavage site, *Strep*-tag II and ER retention signal were present in the *EPO* transcript (Fig. 3d), indicating that *EPO* with built in nucleotide sequences coding for the synthetic peptide was fully transcribed.

The above results implies that *Strep*-tag II may be absent on asialo-rhuEPO^P. To investigate this possibility, Western blot analysis was performed on crude plant extracts and purified asialo-rhuEPO^P from transgenic line A56-5. JMJD2E+*Strep*-tag II was used as a positive control. Western results showed that as low as 5 ng of JMJD2E+*Strep*-tag II could be detected using mouse monoclonal anti-*Strep*-tag II antibody (Fig. 4, upper panel). In the crude protein extracts prepared from A56-5 transgenic line, however, no signals were detected (Fig. 4, upper panel) even though the samples had 5, 10 and 15 ng of asialo-rhuEPO^P (determined by sandwich ELISA). When the blot was stripped and probed with mouse monoclonal anti-EPO antibody, robust EPO protein signals were detected in crude plant extracts containing 5, 10 and 15 ng of asialo-rhuEPO^P (Fig. 4, lower panel), suggesting that *Strep*-tag II may be absent on asialo-rhuEPO^P. Since it is known that different proteins with same affinity tag are detected with different sensitivities (Debeljak et al. 2006), we used 35 ng of purified asialo-rhuEPO^P for Western blot analysis to rule out the possibility of

lack of detection resulting from insufficient protein loading and/or interference from other components present in the crude protein extracts. Ten nanograms each of JMJD2E+*Strep*-tag II and mammalian cell-made recombinant human EPO (rhuEPO^M) were used as positive and negative controls, respectively. When the blot was probed with anti-*Strep*-tag II antibody, a robust signal was detected only for JMJD2E+*Strep*-tag II (Fig. 5, upper panel, lane 1) whereas no signal was detected for purified asialo-rhuEPO^P (Fig. 5, upper panel, lane 2) and rhuEPO^M (Fig. 5, upper panel, lane 3). Re-probing the blot with anti-EPO antibody however, showed robust signal corresponding to three EPO glycoforms of purified asialo-rhuEPO^P (Fig. 5, lower panel, lane 2) as reported previously (Kittur et al. 2013). These results suggest that *Strep*-tag II is absent on asialo-rhuEPO^P, and could have been removed by proteolysis.

To further investigate the absence of *Strep*-tag II on asialo-rhuEPO^P, LC-MS/MS analysis was performed on purified protein. A 30 kD band corresponding to one of the major bands of asialo-rhuEPO^P (Kittur et al. 2014, in press) was excised from SDS-PAGE gel and subjected to LC-MS/MS analysis after trypsin digestion. The idea was that if EPO had an intact synthetic peptide with *Strep*-tag II as indicated in Fig. 1a, we would expect to see a peptide ENLYFQGWSHPQFEK. If only *Strep*-tag II is removed by proteolysis, we would expect to see peptide ENLYFQ. Therefore, LC-MS/MS analysis after trypsin digestion could be used to differentiate between full-length asialo-rhuEPO^P with *Strep*-tag II and its truncated version lacking *Strep*-tag II. Eight peptides NFYAWK, RMEVGQQAVEVWQGLALLSEAVLR, MEVGQQAVEVWQGLALLSEAVLR, EAISPPDAASAAPLR, TITADTFR, TITADTFRK, VYSNFLR and LYTGEACR corresponding to 46–52, 53–76, 54–76, 117–131, 132–139, 132–140, 144–150 and 155–162 positions in the mature human EPO protein sequences (Supplemental Fig. 1i) could be identified from the MS/MS spectra of tryptic peptides (Supplemental Fig. 1a–h). Neither ENLYFQGWSHPQFEK nor ENLYFQ was identified. The peptide LYTGEACR corresponding to 155–162 position in the mature EPO protein was identified as the C-terminal most peptide. Interestingly, the same C-terminal peptide was also identified in 30 kD and 28 kD asialo-rhuEPO^P bands (glycoforms) in our previous study (Kittur et al. 2013). In addition to synthetic peptide, the Thr¹⁶³-Arg¹⁶⁶ region of EPO was found to be missing in both our present and previous studies (Kittur et al. 2013). These results further confirm that the *Strep*-tag II was absent on asialo-rhuEPO^P, and that it possibly has been removed by proteolytic cleavage in the Thr¹⁶³-Arg¹⁶⁶ region of EPO.

A number of mammalian proteins expressed in tobacco plants have been shown to undergo extensive proteolysis both in *planta* (such as in the ER and apoplast) and during tissue disruption/extraction (Schiermeyer et al. 2005; Benchabane et al. 2009; Hehle et al. 2011). To investigate whether removal of *Strep*-tag II occurred in *planta*, microsomes and apoplastic fluid (AF) from A56-5 were isolated for Western blot analysis. ELISA was performed prior to Western blotting to check the presence of asialo-rhuEPO^P in these fractions. The results showed that asialo-rhuEPO^P was present only in the ER fraction, but not in the AF (data not shown). Absence of asialo-rhuEPO^P in the AF could be due to retention of protein in the ER because of the presence of KDEL sequence. For this reason, only microsomal fraction was used for Western blot analysis to detect *Strep*-tag II on asialo-

rhuEPO^P. The results showed a strong protein signal in microsomal fraction when the blot was probed with anti-*Strep*-tag II antibody (Fig. 6, upper panel, lane 2). An immunoreactive band of same size was also detected when the blot was stripped and probed with anti-EPO antibody (Fig. 6, lower panel, lane 2). These results indicate that the *Strep*-tag II was present on asialo-rhuEPO^P while it was retained in the ER.

To examine whether the compact protein core of EPO, in addition to synthetic peptide, is also susceptible to proteolysis under the conditions of protein extraction, the stability of asialo-rhuEPO^P in the presence and absence of protease inhibitors was investigated. Asialo-rhuEPO^P was extracted in acetate buffer (pH 5.0) with or without plant protease inhibitor cocktail. During indicated intervals of incubation periods (Fig. 7), 50 µl of extracts from each time point were withdrawn and analyzed by Western blot. A total protein extract (TPE) was also prepared by boiling the A56-5 leaf tissue (50 mg) directly with 200 µl of 4X SDS sample buffer. The results of Western blot analysis showed no differences in the electrophoretic mobility of asialo-rhuEPO^P band even after 12 h of incubation at 4 °C irrespective of the presence or absence of protease inhibitors (Fig. 7). These results indicate that compact protein core of asialo-rhuEPO^P is resistance to proteolysis, most likely due to the presence of three *N*-glycosylation sites (Asn²⁴, Asn³⁸ and Asn⁸³) all located in the compact core of the protein (Lai et al. 1986; Yamaguchi et al. 1991) thereby shielding the protein core from proteases. These results also suggest that only synthetic peptide and the exposed *C*-terminus region may be susceptible to proteolysis.

In recent years, unintended proteolysis has been recognized as a serious problem in utilizing plant-based platforms for the production of pharmaceutical proteins (Benchabane et al. 2008). Plant genomes are known to encode hundreds of proteases, which play important roles in regulating embryogenesis, seed coat formation, epidermal cell fate, chloroplast biogenesis, and local and systemic defense responses (van der Hoorn 2008). These however, pose significant challenge to efficient production of recombinant proteins in plants. Proteolysis can result in either complete destruction or only partial trimming of recombinant protein depending on the number of protease “susceptible” cleavage sites within the protein that are accessible to plant proteases. Our results (see Fig. 7) indicate that the compact protein core of EPO is resistant to proteolysis during extraction because of the presence of *N*-glycan chains. The extreme *C*-terminal region (Arg¹⁶²-Arg¹⁶⁶) of EPO and the fused synthetic peptide are fully exposed (Fig. 2a and b), which are readily accessible to plant protease(s) and could be protease “susceptible” cleavage site(s) in asialo-rhuEPO^P. It is possible that the *Strep*-tag II was lost by proteolytic cleavage either in the extreme *C*-terminal region (Arg¹⁶²-Arg¹⁶⁶) of EPO or within the tag itself. Since *Strep*-tag II has been shown to be proteolytically stable in other recombinant proteins expressed in plants (Witte et al. 2004; Noel et al. 2007; Werner et al. 2008), we speculate that the loss of affinity tag might have occurred as a result of proteolytic cleavage in the extreme *C*-terminal region (Arg¹⁶²-Arg¹⁶⁶) of EPO. In a recent study, Castilho et al. (2013) reported that amino acid changes in the hinge region of Fc fragment or mutating Arg¹⁶⁶ of EPO to prevent the breakdown of EPO-Fc fusion proteins in tobacco plants were not successful. Their study suggests that the protease susceptible site must be in the EPO, but away from the Arg¹⁶⁶. The *C*-terminal Arg¹⁶⁶ is reported to be absent (removed proteolytically) in natural urinary

EPO and rhuEPO expressed in mammalian cells (Recny et al. 1987). No such modification however, has been reported yet for rhuEPO expressed in plants. It is very unlikely that the cleavage of EPO-Fc fusion protein occurred as a result of peptide bond hydrolysis at Arg¹⁶⁶ because manipulations at this residue were not able to prevent breakdown of EPO-Fc fusion protein in tobacco plants (Castilho et al. 2013). However, it is possible that proteolytic cleavage further down the Arg¹⁶⁶, such as in the Arg¹⁶²-Arg¹⁶⁶ region in EPO. To predict the potential protease cleavage sites within Arg¹⁶²-Arg¹⁶⁶ region in asialo-rhuEPO^P, an analysis using peptideCutter (http://web.expasy.org/peptide_cutter/) was performed. The results showed cutting sites for Arg-C proteinase and Asp-N endopeptidase. As mentioned earlier, plant genomes are known to encode hundreds of proteases, including arginine/lysine-specific proteases (Vercammen et al. 2004; Watanabe and Lam 2005). Arginine/lysine-specific proteases have been shown to be directly involved in the degradation of Sea Anemone protease inhibitor equistatin expressed in potato plants (Outchkourov et al. 2003). We therefore, speculate that the Arg¹⁶²-Arg¹⁶⁶ region of EPO may be target of plant arginine/lysine-specific proteases. A detail study however, is warranted to identify the exact cleavage site in EPO, and the protease(s) responsible for proteolysis.

Proteolytic degradation of recombinant proteins in plants remains a major challenge despite numerous strategies involving targeting the transgene expression in specific tissue or cellular organelle, stabilizing recombinant protein by fusing with a protein stabilizing domain, co-expressing with protease inhibitor or rational engineering of protein variant with improved resistance to plant proteases are being considered (Benchabane et al. 2008). For the rational design of protein variants with improved resistance to plant proteases, knowledge of protease susceptible sites within the protein is essential. Our results suggest that the extreme C-terminal region (Arg¹⁶²-Arg¹⁶⁶) of EPO may be susceptible to proteolysis in tobacco plants. It has been reported that the Thr¹⁶³-Arg¹⁶⁶ region of EPO is not important for its hematopoietic activity (Wen et al. 1994). We have also observed that asialo-rhuEPO^P lacking this region display even better cytoprotective effects than rhuEPO^M (Kittur et al. 2013). Therefore, altering this region by site directed mutagenesis or deleting this region and then appending the purification tags immediately after Arg¹⁶² could be a good strategy to prevent removal of purification tags by proteolysis when EPO is expressed in plants.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The work was supported by National Institute of General Medical Sciences grant (SC3GM088084) and North Carolina Biotechnology Center Grant (2013-BRG-1207) to J.H. Xie.

References

- Bakker H, Rouwendal GJ, Karnoup AS, et al. An antibody produced in tobacco expressing a hybrid β 1,4-galactosyltransferase is essentially devoid of plant carbohydrate epitopes. *Proc Natl Acad Sci USA*. 2006; 103:7577–7582. [PubMed: 16675551]
- Baldet P, Alban C, Axiotis S, Douce R. Localization of free and bound biotin in cells from green pea leaves. *Arch Biochem Biophys*. 1993; 303:67–73. [PubMed: 8489267]

- Benchabane M, Goulet C, Rivard D, et al. Preventing unintended proteolysis in plant protein factories. *Plant Biotechnol J*. 2008; 6:633–648. [PubMed: 18452504]
- Benchabane M, Saint-Jore-Dupas C, Bardor M, et al. Targeting and post-translational processing of human α_1 -antichymotrypsin in BY-2 tobacco cultured cells. *Plant Biotechnol J*. 2009; 7:146–160. [PubMed: 19055607]
- Castilho A, Gattinger P, Grass J, et al. *N*-glycosylation engineering of plants for the biosynthesis of glycoproteins with bisected and branched complex *N*-glycans. *Glycobiology*. 2011; 21:813–823. [PubMed: 21317243]
- Castilho A, Neumann L, Gattinger P, et al. Generation of biologically active multi-sialylated recombinant human EPO-Fc in plants. *PLoS One*. 2013; 8:e54836. [PubMed: 23372778]
- Chen Q. Expression and purification of pharmaceutical proteins in plants. *Biol Eng*. 2008; 1:291–321.
- Choi D, Schroer SA, Lu SH, et al. Erythropoietin protects against diabetes through direct effects on pancreatic beta cells. *J Exp Med*. 2010; 207:2831–2842. [PubMed: 21149549]
- Conley AJ, Mohib K, Jevnikar AM, Brandle JE. Plant recombinant erythropoietin attenuates inflammatory kidney cell injury. *Plant Biotechnol J*. 2009; 7:183–199. [PubMed: 19055608]
- Dani V, Simon WJ, Duranti M, Croy RRD. Changes in the tobacco leaf apoplast proteome in response to salt stress. *Proteomics*. 2005; 5:737–745. [PubMed: 15682462]
- Debeljak N, Davis KL, Kowel R, Sytkowsky AJ. Variability in the immunodetection of His-tagged recombinant proteins. *Anal Biochem*. 2006; 359:216–223. [PubMed: 17081490]
- Erbayraktar S, Grasso G, Sfacteria A, et al. Asialoerythropoietin is a nonerythropoietic cytokine with broad neuroprotective activity in vivo. *Proc Natl Acad Sci USA*. 2003; 100:6741–6746. [PubMed: 12746497]
- Gomord V, Faye L. Posttranslational modification of therapeutic proteins in plants. *Curr Opin Plant Biol*. 2004; 7:171–181. [PubMed: 15003218]
- Haas J, Roth S, Arnold K, et al. The Protein Model Portal - a comprehensive resource for protein structure and model information. *Database* 2013. 2013; bat031.10.1093/database/bat031
- Hehle VK, Paul MJ, Drake PM, Ma JKC, Dollerweed CJV. Antibody degradation in tobacco plants: predominantly apoplastic process. *BMC Biotechnol*. 2011; 11:128. [PubMed: 22208820]
- Jez J, Castilho A, Grass J, et al. Expression of functionally active sialylated human erythropoietin in plants. *Biotechnol J*. 2013; 8:371–382. [PubMed: 23325672]
- Kittur FS, Hung CY, Darlington DE, Sane DC, Xie J. *N*-glycosylation engineering of tobacco plants to produce asialoerythropoietin. *Plant Cell Rep*. 2012; 31:1233–1243. [PubMed: 22371257]
- Kittur FS, Bah M, Archer-Hartmann S, et al. Cytoprotective effect of recombinant human erythropoietin produced in transgenic tobacco plants. *PLoS One*. 2013; 8(10):e76468. [PubMed: 24124563]
- Kittur FS, Arthur A, Nguyen M, et al. Two-step purification procedure for recombinant human asialoerythropoietin expressed in transgenic plants. *Intl J Biol Macromol*. 2014; 72:1111–1116.
- Laemmli UK. Cleavage of structural proteins during the assembly of the head of the bacteriophage T4. *Nature*. 1970; 227:680–685. [PubMed: 5432063]
- Lai PH, Everette R, Wang FF, et al. Structural characteristics of human erythropoietin. *J Biol Chem*. 1986; 261:3116–3121. [PubMed: 3949763]
- Laskowski RA, MacArthur MW, Moss DS, Thornton JM. Procheck - a program to check the stereochemical quality of protein structures. *J Appl Crystallogr*. 1993; 26:283–291.
- Lattenmayer C, Loeschel M, Schriebl K, et al. Protein-free transfection of CHO host cells with an IgG-fusion protein: selection and characterization of stable high producers and comparison to conventionally transfected clones. *Biotechnol Bioeng*. 2007; 96:1118–1126. [PubMed: 17004273]
- Leist M, Ghezzi P, Grasso G, et al. Derivatives of erythropoietin are tissue protective but not erythropoietic. *Science*. 2004; 305:239–242. [PubMed: 15247477]
- Lichty JJ, Malecki JL, Agnew HD, Michelson-Horowitz DJ, Song T. Comparison of affinity tags for protein purification. *Protein Expr Purif*. 2005; 41:98–105. [PubMed: 15802226]
- Majorek KA, Kuhn ML, Chruszcz M, et al. Double trouble buffer selection and His-tag presence may be responsible for non-reproducibility of biomedical experiments. *Protein Sci*. 2014; 23:1359–1368. [PubMed: 25044180]

- Musa TA, Hung CY, Darlington DE, Sane DC, Xie J. Overexpression of human erythropoietin in tobacco plants does not affect plant fertility or morphology. *Plant Biotechnol Rep.* 2009; 3:157–165.
- Noel LD, Cagna G, Stuttman J, et al. Interactions between SGT1 and cytosolic/nuclear HSC70 chaperones regulates Arabidopsis immune responses. *Plant Cell.* 2007; 19:4061–4076. [PubMed: 18065690]
- Okada T, Sawada K, Kubota K. Asialoerythropoietin has a strong renoprotective effects against ischemia-reperfusion in a murine model. *Transplantation.* 2007; 84:504–510. [PubMed: 17713435]
- Outchkourov NS, Rogelj B, Strukelj B, Jongsma MA. Expression of Sea Anemone equistatin in Potato. Effects of plant proteases on heterologous protein production. *Plant Physiol.* 2003; 133:379–390. [PubMed: 12970503]
- Parson J, Altmann F, Arrenber CK, et al. Moss-based production of asialoerythropoietin devoid of Lewis A and other plant-typical carbohydrate determinants. *Plant Biotechnol J.* 2012; 10:851–861. [PubMed: 22621344]
- Pettersen EF, Goddard TD, Huang CC, et al. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem.* 2004; 25:1605–1612. [PubMed: 15264254]
- Recny MA, Scoble HA, Kim Y. Structural characterization of natural human urinary and recombinant DNA-derived erythropoietin. Identification of des-arginine 166 erythropoietin. *J Biol Chem.* 1987; 262:17156–17163. [PubMed: 3680293]
- Sabaty M, Grosse S, Adiyanczyk G, et al. Detrimental effect of 6 His C-terminal tag on YedY enzymatic activity and influence of the TAT signal sequence on YedY synthesis. *BMC Biochem.* 2013; 14:28. [PubMed: 24180491]
- Sali A, Potterton L, Yuan F, van Vlijmen H, Karplus M. Evaluation of comparative protein modeling by MODELLER. *Proteins.* 1995; 23:318–326. [PubMed: 8710825]
- Schiermeyer A, Schinkel H, Apel S, Fischer R, Schillberg S. Production of Desmodus salivary plasminogen activator alpha1 (DSPAalpha1) in tobacco is hampered by proteolysis. *Biotechnol Bioeng.* 2005; 89:848–858. [PubMed: 15685597]
- Schmidt TG, Skerra A. The *Strep-tag* system for one-step purification and high affinity detection or purification of proteins. *Nat Protoc.* 2007; 2:1528–1535.
- Söding J, Biegert A, Lupas AN. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* 2005; 33:W244–W248. Web Server issue. [PubMed: 15980461]
- Syed RS, Reid SW, Li C, et al. Crystal structure of human erythropoietin complexed to its receptor at 1.9 Angstrom. *Nature.* 1998; 395:511–516. [PubMed: 9774108]
- Takeyama T, Takemura G, Kanamori H, et al. Asialoerythropoietin, a nonerythropoietic derivative of erythropoietin displays broad anti-heart failure activity. *Circ Heart Fail.* 2012; 5:274–285. [PubMed: 22343033]
- Van der Hoorn RA. Plant proteases: from phenotypes to molecular mechanisms. *Annu Rev Plant Biol.* 2008; 59:191–223. [PubMed: 18257708]
- Vercammon D, Cotte BVD, Jaeger GD, et al. Type II metacaspases Atmc4 and Atmca of Arabidopsis thaliana cleaves substrates after arginine and lysine. *J Biol Chem.* 2004; 279:45329–45336. [PubMed: 15326173]
- Watanabe N, Lam E. Two Arabidopsis metacaspases AtMCP1b and AtMCP2b are arginine/lysine-specific cysteine proteases and activate apoptosis like cell death in yeast. *J Biol Chem.* 2005; 280:14691–14699. [PubMed: 15691845]
- Wee EGT, Sherrier DJ, Prime TA, Dupree P. Targeting of active sialyltransferase to the plant Golgi apparatus. *Plant Cell.* 1998; 10:1759–1768. [PubMed: 9761801]
- Wen D, Boissel JP, Showers M, Ruch BC, Bunn F. Erythropoietin structure-function relationship. *J Biol Chem.* 1994; 269:22839–22846. [PubMed: 8077236]
- Werner AK, Sparkes IA, Romeis T, Witte CP. Identification, biochemical characterization and subcellular localization of allantoate amidohydrolases for Arabidopsis and Soybean. *Plant Physiol.* 2008; 146:418–430. [PubMed: 18065556]
- Wilken LR, Nikolov ZL. Recovery and purification of plant-made recombinant proteins. *Biotech Adv.* 2012; 30:419–433.

- Witte CP, Noel LD, Gilbert J, Parker JE, Romeis T. Rapid one-step protein purification from plant material using the eight-amino acid long *Strep II* epitope. *Plant Mol Biol.* 2004; 55:135–147. [PubMed: 15604670]
- Yamaguchi K, Akai k, Kawanishi G, et al. Effects of site-directed removal of *N*-glycosylation sites in human erythropoietin on its production and biological properties. *J Biol Chem.* 1991; 266:20434–20439. [PubMed: 1657925]

Key message

C-terminally fused *Strep*-tag II is removed from rhuEPO expressed in tobacco plants. The finding suggests that direct fusion of purification tags at the *C*-terminus of rhuEPO should be avoided.

Fig. 1.

Elution profile of asialo-rhuEPO^P from *Strep*-Tactin column. **a** Crude extract made from transgenic line A56-5. Bound protein was eluted from the column with desthiobiotin (fractions 7–14). **b** Purified asialo-rhuEPO^P (500 ng) in 0.1 M Tris-HCl buffer (pH 8) containing 150 mM NaCl. Bound protein was eluted as described above. Arrows indicate the beginning of elution.

Fig. 2.

Three-dimensional homology model of human EPO containing synthetic peptide. **a** Crystal structure of native EPO (PDB-ID: 1EER, Chain A). **b** Homology model of EPO with a synthetic peptide containing *Strep*-tag II, TEV protease cleavage site and an ER retention signal. **c** Superimposed structures of native EPO and EPO with synthetic peptide. Letters N and C indicate the N- and C-termini, respectively. Region colored in yellow represent synthetic peptide while the one in red is suspected protease susceptible region, which may be cleaved by endogenous plant protease resulting in the loss of purification tag.

Fig. 3.

Schematic representation of cDNA region encoding *Strep*-tag II (WSHPQFEK), TEV protease cleavage site (ENLYFQG) and ER retention signal peptide KDEL fused to EPO cDNA and its detection. **a** Amino acid sequences of synthetic peptide containing *Strep*-tag II, TEV protease cleavage site, and ER retention signal, and the corresponding nucleotide sequences. Primers used to amplify the EPO-*Strep*-tag II fusion cDNA from genomic DNA and RT-PCR are indicated as EPOF and NosTR. The red arrows indicate the trypsin cleavage sites. **b** Amplification of DNA encoding fusion protein from genomic DNA of transgenic line A56-5. Tobacco plant transformed with a plasmid containing GUS gene (A1K1) and plasmid DNA CEJ902 were used as negative and positive controls, respectively. **c** RT-PCR to amplify EPO cDNA from transgenic plant A56-5. **d** Sequencing result of A56-5 RT-PCR product.

Fig. 4.

Western blot analysis to detect the presence of *Strep*-tag II on asialo-rhuEPO^P. Crude protein extracts containing 25–75 μg of total soluble protein and 5–15 ng of JMJD2E+*Strep*-tag II (positive control) were used. The blot was first probed with anti-*Strep*-tag II antibody (upper panel), followed by stripping and re-probing with anti-EPO antibody (lower panel).

Fig. 5.

Western blot analysis to detect the presence of *Strep*-tag II on purified asialo-rhuEPO^P.

Lanes 1–3: 10 ng of JMJD2E+*Strep*-tag II, 35 ng of purified asialo-rhuEPO^P and 10 ng of rhuEPO^M. The blot was first probed with anti-*Strep*-tag II antibody (upper panel), followed by stripping and re-probing with anti-EPO antibody (lower panel).

Fig. 6. Western blot analysis to detect the presence of *Strep*-tag II on asialo-rhuEOP retained in the ER. Microsomal fraction made from transgenic line A56-5 (lane 2, 20 μ g protein) and 10 ng of JMJD2E+*Strep*-tag II (lane 1, positive control) were used. The blot was first probed with anti-*Strep*-tag II antibody (upper panel), followed by stripping and re-probing with anti-EPO antibody (lower panel).

Fig. 7.

Western blot analysis to detect stability of asialo-rhuEPO^P in crude protein extracts made with acetate buffer (pH 5). Crude protein extracts of A56-5 transgenic line (containing 35 µg proteins) with and without protease inhibitors (PI) were incubated at 4°C for indicated interval of time. The samples were then boiled with 4X SDS-PAGE sample buffer. These together with total protein extract (TPE) and rhuEPO^M (positive control, 10 ng) were applied onto a 12.5% SDS-PAGE gel. The blot was probed with mouse anti-EPO antibody.