Title

Identification and CRISPR/Cas9 Inactivation of the C1s Protease Responsible for Proteolysis of Recombinant Proteins Produced in CHO Cells.

Journal

Biotechnology and Bioengineering, 116(9)

Authors

Li, Sophia

Yu, Bin

Byrne, Gabriel

et al.

Publication Date

2019-09-01

DOI

10.1002/bit.27016

Peer reviewed

HHS Public Access

Author manuscript

Biotechnol Bioeng. Author manuscript; available in PMC 2020 September 01.

Published in final edited form as:

Biotechnol Bioeng. 2019 September ; 116(9): 2130–2145. doi:10.1002/bit.27016.

Identification and CRISPR/Cas9 Knockout of the Endogenous C1s Protease in CHO Cells Eliminates Aberrant Proteolysis of Recombinantly Expressed Proteins

Sophia Li[1], **Bin Yu**[2], **Gabriel Byrne**[2], **Meredith Wright**[2], **Sara O'Rourke**[2], **Kathryn Mesa**[2], **Phillip W. Berman**[2]

[1]Department of Chemistry, University of California Santa Cruz

[2]Department of Biomolecular Engineering, University of California Santa Cruz

Abstract

Proteolysis associated with recombinant protein expression in Chinese Hamster Ovary (CHO) cells has hindered the development of biologics including HIV vaccines. When expressed in CHO cells, the recombinant HIV envelope protein, gp120, undergoes proteolytic clipping by a serine protease at a key epitope recognized by neutralizing antibodies. The problem is particularly acute for envelope proteins from clade B viruses that represent the major genetic subtype circulating in much of the developed world, including the US and Europe. In this paper, we have identified complement component 1s (C1s), a serine protease from the complement cascade, as the protease responsible for the proteolysis of gp120 in CHO cells. CRISPR/Cas9 knockout of the C1s protease in a CHO cell line was shown to eliminate the proteolytic activity against the recombinantly expressed gp120. Additionally, the $C1s^{-/-}$ $MGAT1^{-}$ CHO cell line, with the C1s protease and the MGAT1 glycosyltransferase knocked out, enabled production of unclipped gp120 from a clade B isolate (BaL-rgp120) and enriched for mannose-5 glycans on gp120 that are required for the binding of multiple broadly neutralizing monoclonal antibodies (bN-mAbs). The availability of this technology will allow for the scale-up and testing of multiple vaccine concepts in regions of the world where clade B viruses are in circulation. Furthermore, the proteolysis issues caused by the C1s protease suggests a broader need for a C1s-deficient CHO cell line to express other recombinant proteins that are susceptible to serine protease activity in CHO cells. Similarly, the workflow described here to identify and knockout C1s in a CHO cell line can be applied to remedy the proteolysis of biologics by other CHO proteases.

Keywords

CHO cells; Protease; C1s; CRISPR/cas9; Gene Editing; HIV; Env Protein; HIV; Vaccines; Cell Engineering; MGAT1; Glycosylation

Corresponding Author: Sophia Li, 1156 High St., University of California Santa Cruz, Santa Cruz, CA, 95064, USA, Tel: (831) 459-3720, sli140@ucsc.edu.

[7]·Author Contributions

S.L. and P.B. conceived the experiments and wrote the manuscript. B.Y. conducted protease inhibition experiments. S.L. and B.Y. conducted mass spectroscopy experiments and analysis of data. S.L. G.B., M.W., S.O. and K.M. created plasmids, developed the CRISPR/Cas9 knockout method, created stable cell lines and expressed protein. S.L. and B.Y. purified protein. S.L. conducted immunoassays and additional proteolysis experiments with purified protein.

2. Introduction

Despite the availability of anti-retrovirals, there is still an urgent need for a vaccine that protects against HIV (Fauci & Marston, 2014). The HIV envelope (Env) glycoprotein, gp120, was a major component of the multivalent vaccine used in the RV144 clinical trial, the only trial to demonstrate protection in humans (Berman et al., 1999; Rerks-Ngarm et al., 2009). Subsequent studies suggested that protection could be attributed to antibodies to gp120 rather than cellular immune responses (Haynes et al., 2012; O'Connell, Kim, Corey, & Michael, 2012). There is interest in improving gp120s and Env-based vaccines such as developing a scalable and cost-effective recovery process suitable for large scale manufacturing, and increasing the potency and breadth of protection. However, problems with proteolysis have prevented current research developments from being applied to vaccine immunogens from clade B viruses, which circulate in North America, Europe, and other regions of the world (Junqueira & Almeida, 2016). Here, we address the problem of proteolysis of clade B vaccine immunogens in Chinese Hamster Ovary cells (CHO), the standard cell line used in the biopharmaceutical production of recombinant proteins.

Although proteolysis of gp120 expressed in CHO cells has been observed for over 30 years, the protease responsible for this has never been identified due to the lack of an annotated CHO genome (Berman et al., 1990; Stephens, Clements, Yarranton, & Moore, 1990; Clements et al., 1991; Werner & Levy, 1993; Scandella et al., 1993; Schulz et al., 1993; Du, Xu, Viswanathan, & Whalen, 2008; Pugach et al., 2015). When clade B gp120 is produced in CHO cells, it is degraded into 70 kDa and 50 kDa fragments by a secreted, endogenous serine protease. Trimers, such as the B41 SOSIP.664 trimer, also undergo proteolysis at the V3 domain (Pugach et al., 2015). Our experiments are with the gp120 monomer, but our findings should apply to both monomeric and trimeric forms of the vaccine. Proteolysis occurs at the Gly-Pro-Gly-Arg (GPGR↓AF) motif at the crown of the V3 domain of gp120. This sequence is present in 71% of clade B envelope proteins and is also present in other clades of HIV (Foley & Korber, 1996). The V3 domain is important as V3 domain-specific antibodies directed to both the crown and the stem of the V3 loop effectively neutralize multiple strains of HIV (LaRosa et al., 1990; Zolla-Pazner et al., 2004; Burton, Stanfield, & Wilson, 2005).

Eliminating this proteolytic activity is necessary for producing unclipped clade B antigens with the correct antigenic structure required to elicit protective antibody responses. Solutions to mitigate proteolysis of clade B envelope proteins are not ideal due to reduced protein recovery, the need for media additives or alteration of the antigen sequence. Purification by a monoclonal antibody that spans the protease cleavage site resulted in low yields of unclipped protein from the proteolyzed, secreted protein (Pitisuttithum et al., 2006; Berman, 2015). Changes to the fermentation process including the lowering of cell culture temperatures; the addition of cell culture additives such as fetal calf serum, EDTA or EGTA (Du et al., 2008; I. K. Srivastava et al., 2002; Chakrabarti, Barrow, Kanwar, Ramana, & Kanwar, 2016); and the use of continuous flow bioreactor systems (Indresh K. Srivastava et al., 2008, 2003) were associated with lower protein expression and recovery. Lastly, substitution of the GPGR cleavage motif for GPGQ, a common sequence in other clades, removes the cleavage site recognized by the protease and makes gp120 no long susceptible to proteolysis, but would

prevent elicitation of 447–52D-like antibodies (Sullivan et al., 2017; Wen et al., 2018). 447–52D effectively neutralize 92% of 32 primary isolates containing the GPGR motif (Stanfield, Gorny, Williams, Zolla-Pazner, & Wilson, 2004; Zolla-Pazner et al., 2004). Thus, a strategy to develop a vaccine effective against clade B strains is to retain the GPGR motif and induce 447–52D-like antibodies.

In this paper, we describe the use of conventional column chromatography, tandem mass spectroscopy and the annotated CHO genome to identify complement component 1s (C1s) as the serine protease that proteolyzes clade B gp120s. We use CRISPR/Cas9 to engineer the $C1s^{-/-}$ $MGAT1^{-}$ CHO cell line to show that proteolysis of gp120 is eliminated with the knockout of the C1s protease. Additionally, we show that knockout of the C1s protease can be combined with knockout of the MGAT1 glycosyltransferase (mannosyl alpha-1,3 -glycoprotein beta-1,2-N-acetylglucosaminyl-transferase) to express gp120 that is unclipped and possesses the oligomannose structures required for the binding of multiple glycan dependent bNAbs. This cell line will enable the development and testing of vaccine concepts based on clade B vaccine immunogens and the production of unclipped, properly glycosylated Env proteins in quantities sufficient for clinical trials.The C1s-deficient CHO cell line is also a proof-of-concept demonstrating that engineered CHO cells lines with knockout of C1s or other CHO proteases can be used to mitigate the proteolysis of biologics.

3. Materials and Methods

3.1 HIV Env Proteins/Antibodies:

A synthetic gene encoding BaL-rgp120 was created using GenBank#AY713409. Recombinant MN, A244 and TZ97008 gp120s were produced in HEK293, CHO-S and MGAT1- CHO-S cell lines as described previously (Byrne et al., 2018; Doran, Tatsuno, et al., 2018). PG9, PGT128, VRC01 were produced based on published sequences. PGT121 and 10–1074 were acquired from the NIH AIDS Reagent Program. The 34.1 mAb to the Herpes Simplex Virus glycoprotein D (gD) was produced in-house (O'Rourke et al., 2019). 11G5 was a gift from Global Solutions for Infectious Diseases (S. San Francisco, CA). The 447–52D antibody was from the NIH AIDS Reagent Program/Dr. Susan Zolla-Pazner.

3.2 Protein Expression and Quantification:

Cell culture supernatants or purified proteins were run on NuPAGE 4–12% Bis-Tris gels (Thermo Fisher Scientific, Invitrogen, Carlsbad, CA) in MES buffer (Thermo Fisher) and stained with SimplyBlue (Thermo Fisher). For immunoblots, PAGE gels were transferred using iBlot2 (Thermo Fisher). Membranes were blocked with 5% milk for 1 hour. The primary antibody was incubated at 1.5 ug/ml in 5% milk. The secondary antibody used was Peroxidase AffiniPure bovine anti-goat IgG (Jackson ImmunoResearch, West Grove, PA) at a 1:5000 dilution in 5% milk.

3.3 Protease inhibition:

5ug of purified MN-rgp120 protein in PBS was mixed with protease inhibitors at a 1:1 w/w ratio. They were further mixed at 1:9 v/v ratio with CHO supernatant and set at 37ºC overnight. Samples were run on SDS-PAGE with and without dithiothreitol (DTT).

Inhibitors: N-(1-Naphthalenylsulfonyl)-Ile-Trp-aldehyde (Enzo Life Sciences, Farmingdale, NY), Z-Phe-Ala-fluoromethylketone (Sigma, St. Louis, MO), protease inhibitor cocktail cOmplete Mini (Roche, Basel, Switzerland), 4-aminobenzamidine (Sigma), chymostatin (Sigma).

3.4 Enrichment of protease through conventional column chromatography:

A 200ml Superdex S200 size exclusion column (GE Healthcare, Sunnyvale, CA) was used to purify 15 ml of supernatant using 0.05 M Tris buffered saline, pH 8. 5ug of MN-rgp120 was incubated with 1 ml of 20x concentrated pooled fractions and incubated at 37C overnight. Fractions with proteolytic activity were pooled and loaded onto a 5ml QHP anion exchange column (GE). Buffer A was 20mM $NaPO_4$ and buffer B was 20mM $NaPO_4$, 1M NaCl. 1.4ug of MN-rgp120 was incubated with 60x concentrated QHP pooled fractions at 37°C overnight.

3.5 Protein Digestion for Mass Spectroscopy:

Protease-enriched samples were concentrated 60x in centrifugal filters with a 10 kDa cutoff (EMD Millipore, Hayward, CA). 15ul of 50 mM NH_4HCO_3, 1.5 ul of 100mM DTT, 10ul of protein sample and 0.5 ul milliQ water were combined and incubated at 55°C for 60 minutes. 3ul 100mM iodoacetamide was added and incubated in the dark at room temperature for 45 minutes. 4ul trypsin at 40 ng/ul was added and incubated overnight at room temperature, followed by incubation for 8 hours with 4 ul chymotrypsin at 40 ng/ul. 4.5 ul of 5% formic acid was added and incubated at room temperature for 10 minutes.

3.6 Tandem Mass Spectroscopy:

Digested peptides were analyzed on an LTQ-Orbitrap Velos Pro-MS (Thermo Fisher) with a 50mm x 2.1 mm Hypersil GOLD VANQUISH C18 UHPLC column (Thermo Fisher). A linear acetonitrile gradient was run from 5 to 10% for 2 minutes, 10% for 3 minutes, 10 to 80% for 50 minutes, 80 to 95% for 2 minutes, at 95% for 3 minutes, from 95 to 5% in 1 minute and at 5% for 4 minutes. Full-scan MS spectra was acquired (300–2000 m/z). Charge states ≥ 1 were fragmented by collision-induced dissociation. Spectra was processed using Proteome Discoverer v1.4.1.14 (Thermo Fisher) and analyzed against the CHO genome (UniProt-CHO). Parameters: no enzyme specified; dynamic carbamidomethyl modification of cysteine, +57.021 Da; dynamic modification of oxidized methionine, +15.995 Da; strict target FDR of 0.01, relaxed target FDR of 0.05.

3.7 Development of a MGAT1- CHO cell line to express BaL-rgp120:

A MGAT1$^-$ CHO-S cell line expressing BaL was developed by a protocol described previously (O'Rourke et al., 2018), utilizing the MGAT1- CHO-S (Byrne et al., 2018). The modified pCDNA3.1 expression vector encoded a codon optimized BaL-rgp120 gene and an N-terminal purification tag from Herpes Simplex Virus glycoprotein D (gD) as described previously (Berman 1998).

3.8 CRISPR-Cas9 knockout of C1s in the stable BaL-expressing MGAT1$^-$ CHO cell line:

The C1s gene was sequenced and verified against the NCBI mRNA RefSeq XM_007646821.2 and protein RefSeq XP_007645011.1. Guide RNAs were cloned into the CRISPR Nuclease Vector (GeneArt, Thermo Fisher) and selected using CRISPy Cas9 target finder (Novo Nordisk, Technical University of Denmark): sgRNA1 GATAATCTCAGGAGGCGTCG, sgRNA2 GCCAGTAAACCGCTCTTCGT, sgRNA3 GGGCTCTTATTGACGAGTAC and sgRNA4 GTTGACAGCCGCTCATGTTG. To sequence indels, 0.5×10^6 cells were spun down and boiled in 10 ul of milliQ water. 5 ul of cell lysate was used for PCR. For single cell cloning, cells were diluted to 0.5 cells/well in 96 well tissue-culture treated microplates (Corning, Corning, NY) and grown for 14–21 days at 37C, 8% CO_2 and 85% humidity. For scale-up protein production, media was supplemented at 10% of culture volume every 3 days with a feed consisting of Proyield Cotton CNE50M-UF (FrieslandCampina, Delhi, NY) and TC Yeastolate Ultra-Filtered (BD Biosciences, San Jose, CA). For blue-white screening, PCR products of the genetic knockout from clones was ligated into the T-Vector pMD20 (Takara Bio, Mountain View, CA). One Shot TOP10 Competent cells were transformed and plated in IPTG and XGal.

3.9 Electroporation:

Using the Maxcyte STX (MaxCyte Inc., Gaithersburg, MD), 80×10^6 cells were transfected with 120 ug of CRISPR Vector DNA ul in the OC-400 cuvette. Cells were incubated at 37°C for 40 minutes. 30 ml of CHO Growth A media (Irvine Scientific, Santa Ana, CA) was added to flasks and transferred to Kuhner shakers at 125 rpm.

3.10 Purification of BaL-rgp120:

Supernatant from the BaL-rgp120 stable CHO cell line was passed through a 0.2 um filter (Millipore) and purified by the gD tag using an immunoaffinity column consisting of the 34.1 mAb conjugated to AminoLink Plus Coupling resin (Thermo Fisher). Buffer A was 50 mM Tris, 0.5 M NaCl, pH 7.4 and buffer B was 0.1 M glycine, pH 3. 1M Tris was added to eluate at 1:10 ratio of eluate:buffer for neutralization. Subsequent size exclusion chromatography was performed on an S200 gel filtration column (GE) with TBS, pH 8. Protein concentration was measured by BCA assay (Thermo Fisher).

3.11 Thrombin proteolysis of BaL-rgp120:

Purified BaL-rgp120 expressed in the MGAT1$^-$ CHO cell line was digested by thrombin for use in the fluorescence immunoassay comparing monoclonal antibody binding to intact and proteolyzed BaL. 60 ug of BaL-rgp120 was digested with 8.45 units of thrombin in a total volume of 929.5 ul for 92 hours.

3.12 Fluorescence immunoassay (FIA):

Fluortrac high binding 96 well plates (Grenier Bio One, Kremsmunster, Austria) were coated for 2 hours with 2 ug/ml of 34.1 mAb. Plates were blocked for 1 hour with 1% BSA and washed 4x with PBS-T. Purified protein was incubated for 2 hours. Antibodies were prepared at 10 ug/ml to 0.003 ug/ml in duplicate. After a 4x wash with PBS-T, secondary antibodies were added at a 1:3000 dilution. Alexa Fluor 488 Goat α-human IgG (Jackson

ImmunoResearch), Donkey α-goat IgG (Invitrogen), or Goat α-mouse (Invitrogen) IgG were used. After a 4x PBS-T wash, wells were filled with 200 ul PBS. Plates were read on an EnVision plate reader (Perkin Elmer Inc, Waltham, MA) at excitation/emission wavelengths of 35¾85 nm. For direct FIAs, protein was incubated with plates for 2 hours at 2 ug/ml. Subsequent steps were similarly to the capture FIA. Assays were performed in triplicate.

4. Results

4.1. Characterization of the CHO protease responsible for gp120 degradation using protease inhibitors

To characterize the type of protease responsible for gp120 degradation, protease inhibitors were added to growth-conditioned cell culture supernatants from normal CHO cells and to inhibit proteolysis of MN-rgp120. MN-rgp120 has a molecular weight of 120 kDa. Proteolyzed 70k and 50kDa fragments were observed by SDS-PAGE in reduced samples. The cathepsin L inhibitor, N-(1-Naphthalenylsulfonyl)-Ile-Trp-aldehyde; the cysteine protease inhibitor, Z-Phe-Ala-fluoromethylketone; and chymostatin failed to inhibit proteolysis (Fig. 1A). However, proteolysis was successfully inhibited by the serine and cysteine protease inhibitor, complete mini EDTA-free protease inhibitor cocktail, and by serine protease inhibitor, 4-Aminobenzamidine dihydrochloride (4-AMB) (Fig. 1A, 1B). The inhibition of proteolysis by serine protease inhibitors is consistent with studies that have shown gp120 is cleaved by a thrombin-like, serine protease (Clements et al., 1991; Schulz et al., 1993). Proteolysis was also inhibited by the calcium-chelating agent, egtazic acid (EGTA) and has been previously observed (Fig. 1C) (Scandella et al., 1993). These inhibition experiments indicated that the protease in CHO cell supernatant responsible for degrading gp120 was a calcium-dependent serine protease.

4.2 Enrichment of gp120-cleaving protease from CHO supernatant by column chromatography.

The gp120-cleaving protease secreted from CHO cells was enriched using column chromatography to remove CHO host cell proteins before mass spectroscopy analysis. CHO supernatant was first fractionated by size exclusion chromatography (Fig. 1D). Fractions were concentrated, incubated with purified MN-rgp120 overnight to determine fractions containing proteolytic activity. Fractions 2 and 3 contained the major proteolytic activity, while lesser proteolytic activity was observed in fraction 10 (Fig. 1E). Fractions 2 and 3 were pooled and further separated over a QHP, strong anion exchange column (Fig. 1F). The fractions were concentrated, incubated with purified MN gp120 overnight and run on immunoblot. Fractions 1 and 2 from the QHP eluate displayed greater proteolytic activity than fractions 3 and 4 (Fig. 1G). Subsequently, fractions 1 and 2 were pooled to be run on mass spectroscopy.

4.3 Identification of the calcium-dependent serine protease, C1s, by mass spectroscopy.

MS/MS spectroscopy was used to analyze the proteins recovered from the two-column chromatography method described above (Supplemental Fig. 1A). The MS/MS spectrum was analyzed with the Uniprot CHO genome (Xu et al., 2011). The proteins identified were

mainly secreted extracellular proteins or those found on the cell surface (Frantz, Stewart, & Weaver, 2010; Park et al., 2017) (Fig. 2A). Only one protease was identified: complement component 1s (C1s). Six unique peptides from C1s were identified (Supplemental Fig. 2), and coverage of the protein was 14.5%, providing high confidence in the identity of this protease. The six unique peptides are labeled by retention time (Supplemental Fig. 1B), and the fragment ion spectrum is shown for peptide SNTLDIVFQTDLTEQR (Supplemental Fig. 1C). Not only was C1s the only protease identified in the sample, it was both a calcium-dependent protease and a serine protease as predicted by the protease inhibition experiments discussed above.

C1s is a serine protease in the complement cascade and circulates within the blood plasma. It cleaves the complement proteins, C4 and C2, into their active forms, C4b and C2a, which are required for the progression of the complement cascade and result in opsonization and cell lysis of foreign pathogens (Noris & Remuzzi, 2013). The cleavage site sequences for Chinese hamster C2 and C4 were aligned with the sequence for BaL and consensus sequences for clade B, C and AE viruses. Arginine (R) is required in the P1 position and small amino acids, glycine (G) or alanine (A), are conserved in the P2 position (Fig. 2B). The lack of R at the P1 positions for clade CRF01-AE and clade C gp120s account for their resistance to proteolysis when expressed in CHO cells. While no structural information exists for Chinese hamster C1s, BLAST alignment of the residues in human and Chinese hamster C1s showed 77% homology. The similar cleavage site sequences of clade B gp120 and the complement proteins, C2 and C4, support the observation that these proteins are cleaved by C1s.

The presence of C1s in CHO cells has been previously observed in CHO proteomic experiments, which identified all proteins in CHO supernatant by mass spectroscopy (Baycin-Hizal et al., 2012; Park et al., 2017). Quantitative analysis of these proteins showed that C1s is a relatively highly expressed protein; in the supernatant of batch run cultures, C1s in was in the top 150 most highly expressed proteins over an 8-day culture. The high expression of C1s in CHO supernatant may correlate with the rapid proteolysis observed of clade B gp120s in CHO cells. More recently, a group used next-generation sequencing to observe differential expression of CHO serine proteases between a CHO-K1 cell line and a CHO-DUXB11 cell line, which produces protein with less proteolytic degradation. C1s was one of the proteases that had significant differences in gene expression (Laux et al., 2018). Thus, our identification of C1s as the gp120-cleaving protease is consistent with other CHO proteomic and protease experiments.

4.4 Proteolysis of clade B gp120 occurs in CHO cells, but not in HEK293 cells.

The extent of proteolysis of gp120s antigens with the GPGR motif varies considerably between Envs expressed in CHO cells and another commonly used mammalian cell line, Human Embryonic Kidney (HEK293) cells. To confirm this, the clade B isolates, BaL-rgp120 and EN3-rgp120, both containing the GPGR motif, were transiently expressed in CHO and HEK293 cells. The BaL- and EN3-rgp120s remained intact when expressed in HEK293 cells but were proteolyzed in CHO cells (Fig. 3A), with BaL-rgp120 being completely proteolyzed and EN3-rgp120 partially degraded. The differences in proteolytic

activity between CHO and HEK293 cells were consistent with C1s RNA expression data provided by the Human Protein Atlas (Fig. 3B). Although these values are derived from human tissues, the 10-fold higher expression of C1s in ovary tissues over that in kidney tissues may explain why proteolysis of clade B gp120s is observed when expressed in CHO cells, but not in HEK293 cells. Differences in C1s expression have also been observed in mouse tissues where the gene expression of C1s was shown to be 80-fold higher in ovary tissues over kidney tissues (Supplemental. Fig. 3).

4.5 Development of a stable CHO cell line expressing the BaL-rgp120 for use in screening for C1s gene knockout experiments.

To show that C1s was responsible for proteolysis, we created a stable CHO cell line expressing a proteolysis-sensitive gp120 to be used for protease gene knockout experiments. The gp120 gene from the BaL isolate of HIV (Hwang, Boyle, Lyerly, & Cullen, 1991) was transfected into a recently described CHO cell line, MGAT1$^-$ CHO, containing a mutation in the MGAT1 glycosyltransferase gene that limits glycosylation to early high-mannose intermediates in the N-linked glycosylation pathway (Byrne et al., 2018; O'Rourke et al., 2018). BaL-rgp120 was selected based on its binding profile to bN-mAbs. In particular, BaL-rgp120, when expressed in MGAT1- CHO cells, binds the prototypic PG9, PGT128, 10–10-74 and VRC01 bN-mAbs that target epitopes in the V1/V2, V3 stem, and CD4 binding site of gp120. Antigenicity of BaL-rgp120 was improved over MN-rgp120, the original clade B vaccine used in the RV144 trial. MN-rgp120 binds strongly to VRC01, minimally to PG9, and does not bind to PGT121, PGT128 or 1010–74. (The bN-mAb binding data for MN-rgp120 is shown by Doran et al., (Doran, Tatsuno, et al., 2018) and Fig. 7, while the antigenicity of BaL-rgp120 is shown in Figs. 6 and 7). Virion associated Env proteins primarily possess high-mannose glycans (Bonomelli et al., 2011; Doores et al., 2010). Restricting glycosyation improved binding of bN-mAbs dependent on the presence of high-mannose glycans (Sok et al., 2014; Doran, Tatsuno, et al., 2018; Byrne et al., 2018). Using the ClonePix2 colony selection robot, 30,000 rgp120-secreting colonies were imaged an d ranked based on fluorescence intensity and geometric parameters (Fig. 4A). Seven colonies were grown in shake-flask cultures. After eight days, all seven cultures exhibited significant degradation of BaL-rgp120 (Fig. 4B) with productivities of 200–300mg/L for four out of seven cultures (2.4D, 1.5D, 2.3D, 1.3D) (Fig. 4C). In these experiments, clone 1.5d MGAT1$^-$ CHO cell line stably expressing BaL-rgp120 was shown to be sensitive to proteolysis and produced sufficient gp120 for use in gene knockout experiments to assess the role of C1s in the proteolysis of clade B Env proteins.

4.6 CRISPR/cas9 knockout of C1s in a stable MGAT1$^-$ CHO cell line expressing BaL-rgp120 eliminates proteolysis.

Based on the identification of C1s, CRISPR/Cas9 was used to knock out the C1s gene in the BaL-rgp120 1.5d MGAT1$^-$ CHO cell line. Four guide RNA targets (Fig. 4D, 4E) were selected using CRISPy target selection software (Ronda et al., 2014) and targeted exon 3 and exon 11, the serine protease domain of C1s (Fig. 4D). A screen of the four gRNAs was done to determine gene-editing efficiency as not all gRNAs are equally effective (Peng, Lin, & Li, 2016). Knockout efficiency in the population of cells post-transfection was determined by the T7 endonuclease assay and TIDE (tracking of indels by decomposition) analysis (Figure

4F, 4G) (Vouillot et al., 2015; Brinkman, Chen, Amendola, & van Steensel, 2014). Only sgRNA4 exhibited any knockout efficiency as qualitatively shown by the T7 assay (Fig. 4F) and quantitatively by TIDE (Fig. 4G). The indel introduced by sgRNA4 occurs at nucleotide 1627 of the NCBI RefSeq XM_007646821.2. The 11.4% knockout efficiency measured by TIDE analysis suggested that forty clones should be screened to identify clones with the genetic knockout.

Single cell cloning was carried out to isolate clones with Cas9-mediated mutations of the C1s gene. Supernatants from thirty-eight clones transfected with sgRNA4 were recovered between days 18 to 25 and analyzed by immunoblot for the detection of intact and proteolyzed BaL-rgp120 (Fig. 5A). Of thirty-eight clones analyzed, three clones lacked protease activity as indicated by the accumulation of intact gp120 and no visible 50 kDa or 70 kDa fragments. Blue-white lacZ screening was used to individually sequence each of the two C1s alleles present in each clone (Fig. 5C). Sequencing confirmed that clones 1 and 2 that displayed proteolyzed Bal-rgp120 were not gene-edited, while clones 3 and 4 that expressed non-proteolyzed BaL did have Cas9-modified gene edits. Sequencing and TIDE analysis on clone 4 showed two gene-edits: a −3 basepair and −4 basepair deletion (Fig. 5B, 5C). C1s is part of the complement pathway and is not directly invovled in metabolic pathways. We did not observe significant differences in cell doubling times (~25 hours) or protein expression between the normal CHO cell line and the $C1s^{-/-}$ $MGAT1^{-}$ CHO cell line (100–300 mg/L). The clone 4 BaL-rgp120 $C1s^{-/-}$ $MGAT1^{-}$ CHO cell line identified above will be henceforth referred to as BaL-rgp120 $C1s^{-/-}$ $MGAT1^{-}$ CHO cell line and was used for the subsequent antibody-binding experiments.

4.7 Proteolyzed Bal-rgp120 displays diminished binding to V3-directed mAbs and bN-mAbs relative to intact BaL-rgp120.

To compare differences in neutralizing antibody binding between intact and proteolyzed gp120, BaL-rgp120 was purified from the parental 1.5D BaL-rgp120 MGAT- CHO cell line and the BaL-rgp120 $C1s^{-/-}$ $MGAT1^{-}$ CHO cell lines (Fig. 6J). However, BaL-rgp120 produced in the stable $MGAT1^{-}$ 1.5d CHO cell line contained a mixture of proteolyzed and non-proteolyzed proteins. To fully distinguish differences in antibody binding between intact and proteolyzed BaL-rgp120, cleaved BaL-gp120 was produced by additional treatment with thrombin. Treatment of gp120 with exogenous thrombin has been shown to cleave rgp120 into fragments of 50 and 70 kDa at the same GPGR sequence as the C1s protease (Clements et al., 1991; Schulz et al., 1993). The intact BaL-rgp120 expressed in the BaL-rgp120 $C1s^{-/-}$ $MGAT1^{-}$ CHO cell line and the thrombin-cleaved BaL-rgp120 expressed in the $MGAT1^{-}$ 1.5d CHO cell line were used in the antibody binding assays.

Fluorescence immunoassay (FIA) was used to quantitate differences in antibody binding and determine structural changes due to proteolysis between intact BaL-rgp120 and proteolyzed BaL-rgp120 cleaved in the V3 domain. V3-antibodies had reduced binding due to the loss of the V3 epitope, while binding of bN-mAbs to the V1/V2 domain and CD4 binding site were not affected (Fig. 6A–I). PGT128 displayed the most dramatic loss of binding, followed by the 11G5 mAb. PGT128 recognizes a glycan dependent epitope at the base of the V3 domain, while 11G5 recognizes an epitope that spans the cleavage site in the V3 domain

(Pejchal et al., 2011; Nakamura et al., 1992). We did not observe significant differences in binding to the human 447–52D mAb, which recognizes the V3 GPxR crown, or for 10–1074, which recognizes the base of the V3 stem (Zolla-Pazner et al., 2004; Mouquet et al., 2012). Epitopes to the V1/V2 domain and the CD4 binding site were unaffected as shown by the PG9 and CD4 bN-mAbs (McLellan et al., 2011; Zhou et al., 2010). These antibody binding experiments demonstrate that V3-directed antibodies are sensitive to structural changes due to proteolysis at the V3 domain.

4.8 Comparative bN-mAb binding of unclipped MN-rgp120 used in the RV144 trial to glycan restricted clade B Bal-rgp120 and clade C TZ97008-rgp120.

Fluorescence immunoassay was used to compare the antigenic structure of unclipped clade B MN-rgp120, similar to that used in the RV144 trial, with unclipped BaL-rgp120 and clade C TZ97008-rgp120 (Doran, Yu, et al., 2018). Because both the BaL-rgp120 and the TZ97008-rgp120 were expressed in the MGAT1^{-} CHO cell line, N-linked glycosylation was restricted high-mannose glycans. In contrast, MN-rgp120 was expressed in HEK293 cells to prevent proteolysis and possessed a mixture of hybrid and complex sialic acid containing glycans (Yu, Morales, O'Rourke, Tatsuno, & Berman, 2012). BaL-rgp120 exhibited robust binding to the PG9, PGT128, 10–1074 and VRCO1 (Fig. 7A, 7C, 7D and 7E). In contrast, the clade B MN-rgp120 immunogen exhibited robust binding to VRC01, bound only moderately to PG9, but did not bind to PGT128 or 10–1074. The clade C TZ97008-rgp120 expressed in the MGAT1^{-} CHO cell line also bound to PG9, PGT128, 10–1074, VRCO1, and additionally bound to PGT121 (Fig. 7B). While PGT121 has been found to be a strongly neutralizing antibody (Walker et al., 2011), identifying an antigen that binds to PGT121 has been difficult. Although several bN-mAbs require mannose-5 for binding, PGT121 requires a mixture of high mannose and sialic acid glycans for binding (Mouquet et al., 2012). However, 10–1074 is in the same family as PGT121 and binds to Bal-rgp120 at an overlapping epitope. Further screening of clade B antigens, including additional Tier 2 viruses, may result in identification of an antigen that is additionally recognized by PGT121. BaL-rgp120 in combination with TZ97008-rgp120 expressed in the C1s$^{-/-}$ MGAT1^{-} and MGAT1^{-} CHO cell lines, represent candidates for a bivalent vaccine effective against clade B and clade C viruses.

5. Discussion

Proteolysis associated with CHO cells, the preferred cell line for biopharmaceutical production, has hindered the development of therapeutic biologics including HIV vaccines. In these studies, we described the identification of C1s as the protease responsible for the cleavage of HIV vaccine immunogens. This conclusion is supported by the finding that knockout of C1s in CHO cells prevents proteolysis of BaL-gp120, a clade B isolate of HIV. Two CRISPR/Cas9 induced mutations (C1s and MGAT1) were combined in a single CHO cell line to improve the quality of a recombinant protein intended for large scale production. This cell line overcomes two major problems in HIV vaccine production: 1) proteolysis of clade B HIV vaccine immunogens expressed in CHO cells and 2) glycan heterogeneity that has impaired the antigenic structure of recombinant Env proteins. The strategy we have developed provides a practical method for the large-scale production of unclipped, glycan

restricted, Env proteins to enable the testing of clade B HIV vaccine concepts including monomeric and trimeric envelope proteins, as well as guided immunization strategies. Further studies may also elucidate whether a C1s-deficient CHO cell line or other CHO protease-deficient CHO cell line is beneficial for the expression of other recombinant proteins found to be proteolyzed in CHO cells including but not limited to monoclonal antibodies, human Factor VIII or IFN-γ (Bee et al., 2015; Dorai et al., 2011; Clincke, Guedon, Yen, Ogier, & Goergen, 2011; Kaufman, Wasley, & Dorner, 1988).

The identification of C1s as the protease found in CHO cell culture medium provides an explanation for the proteolysis that has limited the development of recombinant therapeutic proteins. Previously, groups have partially characterized CHO proteases involved in proteolytic degradation and have stated that the CHO protease of interest is a metalloprotease or serine protease based on screening assays with protease inhibitors and limited amino acid sequence from mass spectroscopy experiments. At the time, they were not able to identify the protease due to the lack of an annotated CHO genome (Du et al., 2008; Sandberg et al., 2006; Clincke et al., 2011; Dorai et al., 2011). With the publication of the CHO genome in 2011 (Xu et al., 2011) and common MS/MS techniques, we have determined the identity of the protease and created an engineered knockout cell line to eliminate proteolysis of recombinant proteins. Similar to our work, other groups in recent years have also utilized the CHO genome to identify CHO proteases involved in the degradation of biologics (Cathepsin D, Cathepsin L, matriptase, etc…) or to do large-scale proteomics studies (Bee et al., 2015; Laux et al., 2018; Luo et al., 2019; Park et al., 2017; Valente, Lenhoff, & Lee, 2015). Identifying the myriad of proteases or host cell proteins in CHO cells is necessary due to the continuing use of the CHO cell line for expression of therapeutic proteins. The workflow described here (ie enrichment by column chromatography followed by tandem mass spectroscopy) can be used to identify other CHO proteases involved in the degradation of recombinant proteins. Knowledge of these proteases and host cell proteins is also important for designing purification processes that will remove impurities from the final purified product, as many host cell proteins will co-elute with the product and can be difficult to remove (Valente et al., 2015). We demonstrate separation of C1s by anion exchange chromatography during the enrichment of the protease; adding an anion exchange chromatography step could be used as a method to remove C1s in a protein purification process. Additional studies will determine whether recombinant proteins other than HIV Env such as Factor VIII or IFN-γ are also susceptible to proteolysis by C1s in CHO cells.

Conservation of the GPGR motif recognized by C1s in clade B viruses suggests that this cleavage site provides an advantage in fitness. In previous studies, we observed that cleavage sites recognized by the antigen-processing enzymes, cathepsins S, L and D, were conserved in Envs and occurred at epitopes recognized by bNAbs (Yu, Fonseca, O'Rourke, & Berman, 2010). It was suggested that these protease cleavage sites represent an immune escape mechanism whereby key epitopes are degraded during the antigen presentation process, diminishing an effective immune response to these sites. Similarly, deletion of an overlapping cathepsin S cleavage site in MN-rgp120 facilitated cross presentation and increased binding affinity for MHC class 1 molecules (Frey et al., 2018). Preservation of the C1s cleavage site may also facilitate cellular immune responses as viruses have been known

to use host cell proteases and complement to promote infectivity and immune escape strategies (Ploegh, 1998; Zipfel, Würzner, & Skerka, 2007). While the GPGR motif is present only in clade B viruses, the protease cleavage sites of cathepsins S, L and D are present in Envs of other clades. Proteolysis of clade B Envs at the GPGR supports the broader hypothesis that protease cleavage sites located proximal to epitopes recognized by bNAbs on viral proteins may promote immune escape or enhance the viruses' evolutionary fitness.

Along with the identification of the C1s protease, we have also shown highly antigenic Env proteins suitable for an HIV vaccine. We recently described glycan-optimized gp120s from the A244 strain of HIV (clade CRF01-AE) and the TZ97008 strain of HIV (clade C) (O'Rourke et al., 2018; Doran, Yu, et al., 2018). Here, we describe a novel C1s$^{-/-}$ MGAT1^{-} CHO cell line expressing unclipped, glycan optimized gp120 from the BaL isolate of HIV. The two gp120s described previously along with the BaL-rgp120 described in this paper represent three components of a trivalent vaccine effective against the three major subtypes (clades B, C, and CRF01_AE) circulating in the developed and developing worlds. Further immunogenicity studies will be required to define the full immunogenic potential of these new candidate vaccines. Preliminary studies on the immunogenicity of mannose-5 glycosylated antigens from the MGAT1- CHO cell line show comparable antibody titers in animal sera (unpublished) and we are currently analyzing the antibody repertoire in sera elicited by these antigens.

In summary, the data presented in this report represents a continuation of efforts to improve the quality and manufacturability of recombinant HIV immunogens. This technology can be used to produce other candidate vaccines that have not been tested because they could not be manufactured cost effectively or expeditiously. The C1s-deficient CHO cell line represents a permanent solution to mitigate proteolysis of recombinant HIV immunogens and possibly other recombinant therapeutic proteins susceptible to serine protease activity.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

6. Acknowledegements

We wish to thank Global Solutions for Infectious Diseases (S. San Francisco) and the NIH AIDS Research Reagent Program for providing the 11G5 monoclonal antibody used in this study. This work was supported by a grant to P.B. from the National Institute of Drug Abuse (#RO1DA036335). S.L. and P.B. have filed a patent on C1s-deficient cells for production of vaccines and biopharmaceutical proteins.

8. References

References

Baycin-Hizal D, Tabb DL, Chaerkady R, Chen L, Lewis NE, Nagarajan H, … Betenbaugh M (2012). Proteomic Analysis of Chinese Hamster Ovary Cells. Journal of Proteome Research, 11(11), 5265–5276. 10.1021/pr300476w [PubMed: 22971049]

Bee JS, Tie L, Johnson D, Dimitrova MN, Jusino KC, & Afdahl CD (2015). Trace levels of the CHO host cell protease cathepsin D caused particle formation in a monoclonal antibody product. Biotechnology Progress, 31(5), 1360–1369. 10.1002/btpr.2150 [PubMed: 26259961]

Berman PW (2015, 6). Strategies on Improving the AIDSVAX Vaccine. Presented at the NIH Env Manufacturing Workshop, NIAID, Bethesda, MD Retrieved from https://www.vaccineenterprise.org/sites/default/files/Agenda%20-%20HIV%20Env%20Manufacturing%20workshop%20-%202015.pdf

Berman PW, Gregory TJ, Riddle L, Nakamura GR, Champe MA, Porter JP, … Eichberg JW (1990). Protection of chimpanzees from infection by HIV-1 after vaccination with recombinant glycoprotein gp120 but not gp160. Nature, 345(6276), 622–625. 10.1038/345622a0 [PubMed: 2190095]

Berman PW, Huang W, Riddle L, Gray AM, Wrin T, Vennari J, … Gregory TJ (1999). Development of Bivalent (B/E) Vaccines Able to Neutralize CCR5-Dependent Viruses from the United States and Thailand. Virology, 265(1), 1–9. 10.1006/viro.1999.0031 [PubMed: 10603312]

Bonomelli C, Doores KJ, Dunlop DC, Thaney V, Dwek RA, Burton DR, … Scanlan CN (2011). The Glycan Shield of HIV Is Predominantly Oligomannose Independently of Production System or Viral Clade. PLoS ONE, 6(8). 10.1371/journal.pone.0023521

Brinkman EK, Chen T, Amendola M, & van Steensel B. (2014). Easy quantitative assessment of genome editing by sequence trace decomposition. Nucleic Acids Research, 42(22), e168–e168. 10.1093/nar/gku936 [PubMed: 25300484]

Burton DR, Stanfield RL, & Wilson IA (2005). Antibody vs. HIV in a clash of evolutionary titans. Proceedings of the National Academy of Sciences, 102(42), 14943–14948. 10.1073/pnas.0505126102

Byrne G, O'Rourke SM, Alexander DL, Yu B, Doran RC, Wright M, … Berman PW (2018). CRISPR/Cas9 gene editing for the creation of an MGAT1-deficient CHO cell line to control HIV-1 vaccine glycosylation. PLOS Biology, 16(8), e2005817. 10.1371/journal.pbio.2005817

Chakrabarti S, Barrow CJ, Kanwar RK, Ramana V, & Kanwar JR (2016). Studies to Prevent Degradation of Recombinant Fc-Fusion Protein Expressed in Mammalian Cell Line and Protein Characterization. International Journal of Molecular Sciences, 17(6). 10.3390/ijms17060913

Clements GJ, Price-Jones MJ, Stephens PE, Sutton C, Schulz TF, Clapham PR, … Marsh M (1991). The V3 loops of the HIV-1 and HIV-2 surface glycoproteins contain proteolytic cleavage sites: a possible function in viral fusion? AIDS Research and Human Retroviruses, 7(1), 3–16. 10.1089/aid.1991.7.3 [PubMed: 2015114]

Clincke M-F, Guedon E, Yen FT, Ogier V, & Goergen J-L (2011). Characterization of metalloprotease and serine protease activities in batch CHO cell cultures: control of human recombinant IFN-γ proteolysis by addition of iron citrate. BMC Proceedings, 5(Suppl 8), P115 10.1186/1753-6561-5-S8-P115 [PubMed: 22373384]

Doores KJ, Bonomelli C, Harvey DJ, Vasiljevic S, Dwek RA, Burton DR, … Scanlan CN (2010). Envelope glycans of immunodeficiency virions are almost entirely oligomannose antigens. Proceedings of the National Academy of Sciences, 107(31), 13800–13805. 10.1073/pnas.1006498107

Dorai H, Santiago A, Campbell M, Tang QM, Lewis MJ, Wang Y, … Hancock W (2011). Characterization of the proteases involved in the N-terminal clipping of glucagon-like-peptide-1-antibody fusion proteins. Biotechnology Progress, 27(1), 220–231. 10.1002/btpr.537 [PubMed: 21312369]

Doran RC, Tatsuno GP, O'Rourke SM, Yu B, Alexander DL, Mesa KA, & Berman PW (2018). Glycan modifications to the gp120 immunogens used in the RV144 vaccine trial improve binding to broadly neutralizing antibodies. PLOS ONE, 13(4), e0196370. 10.1371/journal.pone.0196370

Doran RC, Yu B, Wright M, O'Rourke SM, Yin L, Richardson JM, … Berman PW (2018). Development of a Stable MGAT1- CHO Cell Line to Produce Clade C gp120 with Improved Binding to Broadly Neutralizing Antibodies. Frontiers in Immunology, 9 10.3389/fimmu.2018.02313

Du SX, Xu L, Viswanathan S, & Whalen RG (2008). Inhibition of V3-specific cleavage of recombinant HIV-1 gp120 produced in Chinese hamster ovary cells. Protein Expression and Purification, 59(2), 223–231. 10.1016/j.pep.2008.02.002 [PubMed: 18406166]

Fauci AS, & Marston HD (2014). Ending AIDS — Is an HIV Vaccine Necessary? New England Journal of Medicine, 370(6), 495–498. 10.1056/NEJMp1313771 [PubMed: 24499210]

Foley B, & Korber B (1996). Global variation in the HIV-1 V3 region. MS K710, Los Alamos National Laboratory, III: 77–139.

Frantz C, Stewart KM, & Weaver VM (2010). The extracellular matrix at a glance. Journal of Cell Science, 123(24), 4195–4200. 10.1242/jcs.023820 [PubMed: 21123617]

Frey BF, Jiang J, Sui Y, Boyd LF, Yu B, Tatsuno G, … Berzofsky JA (2018). Effects of Cross-Presentation, Antigen Processing, and Peptide Binding in HIV Evasion of T Cell Immunity. The Journal of Immunology, 200(5), 1853–1864. 10.4049/jimmunol.1701523 [PubMed: 29374075]

Haynes BF, Gilbert PB, McElrath MJ, Zolla-Pazner S, Tomaras GD, Alam SM, … Kim JH (2012). Immune-Correlates Analysis of an HIV-1 Vaccine Efficacy Trial. New England Journal of Medicine, 366(14), 1275–1286. 10.1056/NEJMoa1113425 [PubMed: 22475592]

Hwang SS, Boyle TJ, Lyerly HK, & Cullen BR (1991). Identification of the envelope V3 loop as the primary determinant of cell tropism in HIV-1. Science, 253(5015), 71–74. 10.1126/science.1905842 [PubMed: 1905842]

Junqueira DM, & Almeida SE de M. (2016). HIV-1 subtype B: Traces of a pandemic. Virology, 495, 173–184. 10.1016/j.virol.2016.05.003 [PubMed: 27228177]

Kaufman RJ, Wasley LC, & Dorner AJ (1988). Synthesis, Processing, and Secretion of Recombinant Human Factor VI11 Expressed in Mammalian Cells. Journal of Biological Chemistry, 263(13), 6352–6392. [PubMed: 3129422]

LaRosa GJ, Davide JP, Weinhold K, Waterbury JA, Profy AT, Lewis JA, … Et A (1990). Conserved sequence and structural elements in the HIV-1 principal neutralizing determinant. Science, 249(4971), 932–935. 10.1126/science.2392685 [PubMed: 2392685]

Laux H, Romand S, Nuciforo S, Farady CJ, Tapparel J, Buechmann-Moeller S, … Bodendorf U (2018). Degradation of recombinant proteins by Chinese hamster ovary host cell proteases is prevented by matriptase-1 knockout. Biotechnology and Bioengineering, 0(0). 10.1002/bit.26731

Luo H, Tie L, Cao M, Hunter AK, Pabst TM, Du J, … Wang WK (2019). Cathepsin L Causes Proteolytic Cleavage of Chinese-Hamster-Ovary Cell Expressed Proteins During Processing and Storage: Identification, Characterization, and Mitigation. Biotechnology Progress, 35(1), e2732 10.1002/btpr.2732 [PubMed: 30320962]

McLellan JS, Pancera M, Carrico C, Gorman J, Julien J-P, Khayat R, … Kwong PD (2011). Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9. Nature, 480(7377), 336–343. 10.1038/nature10696 [PubMed: 22113616]

Mouquet H, Scharf L, Euler Z, Liu Y, Eden C, Scheid JF, … Bjorkman PJ (2012). Complex-type N-glycan recognition by potent broadly neutralizing HIV antibodies. Proceedings of the National Academy of Sciences of the United States of America, 109(47), E3268–E3277. 10.1073/pnas.1217207109 [PubMed: 23115339]

Nakamura GR, Byrn R, Rosenthal K, Porter JP, Hobbs MR, Riddle L, … Berman PW (1992). Monoclonal Antibodies to the Extracellular Domain of HIV-1IIIB gp160 that Neutralize Infectivity, Block Binding to CD4, and React with Diverse Isolates. AIDS Research and Human Retroviruses, 8(11), 1875–1885. 10.1089/aid.1992.8.1875 [PubMed: 1283308]

Noris M, & Remuzzi G (2013). Overview of Complement Activation and Regulation. Seminars in Nephrology, 33(6), 479–492. 10.1016/j.semnephrol.2013.08.001 [PubMed: 24161035]

O'Connell RJ, Kim JH, Corey L, & Michael NL (2012). Human Immunodeficiency Virus Vaccine Trials. Cold Spring Harbor Perspectives in Medicine, 2(12), a007351. 10.1101/cshperspect.a007351

O'Rourke SM, Byrne G, Tatsuno G, Wright M, Yu B, Mesa KA, … Berman PW (2018). Robotic selection for the rapid development of stable CHO cell lines for HIV vaccine production. PLOS ONE, 13(8), e0197656. 10.1371/journal.pone.0197656

O'Rourke SM, Yu B, Morales JF, Didinger CM, Alexander DL, Vollmers C, & Berman PW (2019). Production of a recombinant monoclonal antibody to Herpes Simplex Virus glycoprotein D for immunoaffinity purification of tagged proteins. Journal of Immunological Methods, 465, 31–38. 10.1016/j.jim.2018.11.015 [PubMed: 30502324]

Park JH, Jin JH, Lim MS, An HJ, Kim JW, & Lee GM (2017). Proteomic Analysis of Host Cell Protein Dynamics in the Culture Supernatants of Antibody-Producing CHO Cells. Scientific Reports, 7, 44246 10.1038/srep44246 [PubMed: 28281648]

Pejchal R, Doores KJ, Walker LM, Khayat R, Huang P-S, Wang S-K, … Wilson IA (2011). A potent and broad neutralizing antibody recognizes and penetrates the HIV glycan shield. Science (New York, N.Y.), 334(6059), 1097–1103. 10.1126/science.1213256

Peng R, Lin G, & Li J (2016). Potential pitfalls of CRISPR/Cas9-mediated genome editing. The FEBS Journal, 283(7), 1218–1231. 10.1111/febs.13586 [PubMed: 26535798]

Pitisuttithum P, Gilbert P, Gurwith M, Heyward W, Martin M, van Griensven F, … Tappero JW (2006). Randomized, Double-Blind, Placebo-Controlled Efficacy Trial of a Bivalent Recombinant Glycoprotein 120 HIV-1 Vaccine among Injection Drug Users in Bangkok, Thailand. The Journal of Infectious Diseases, 194(12), 1661–1671. 10.1086/508748 [PubMed: 17109337]

Ploegh HL (1998). Viral Strategies of Immune Evasion. Science, 280(5361), 248–253. 10.1126/science.280.5361.248 [PubMed: 9535648]

Pugach P, Ozorowski G, Cupo A, Ringe R, Yasmeen A, de Val N, … Moore JP (2015). A native-like SOSIP.664 trimer based on an HIV-1 subtype B env gene. Journal of Virology, 89(6), 3380–3395. 10.1128/JVI.03473-14 [PubMed: 25589637]

Rerks-Ngarm S, Pitisuttithum P, Nitayaphan S, Kaewkungwal J, Chiu J, Paris R, … MOPH-TAVEG Investigators. (2009). Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. The New England Journal of Medicine, 361(23), 2209–2220. 10.1056/NEJMoa0908492 [PubMed: 19843557]

Roche. (2004). The Complete Guide for Protease Inhibition. Retrieved from http://wolfson.huji.ac.il/purification/PDF/Protease_Inhibitors/ProteaseInhibitorRoche.pdf

Ronda C, Pedersen LE, Hansen HG, Kallehauge TB, Betenbaugh MJ, Nielsen AT, & Kildegaard HF (2014). Accelerating Genome Editing in CHO Cells Using CRISPR Cas9 and CRISPy, a Web-Based Target Finding Tool. Biotechnology and Bioengineering, 111(8), 1604–1616. 10.1002/bit.25233 [PubMed: 24827782]

Sandberg H, Lütkemeyer D, Kuprin S, Wrangel M, Almstedt A, Persson P, … Mikaelsson M (2006). Mapping and partial characterization of proteases expressed by a CHO production cell line. Biotechnology and Bioengineering, 95(5), 961–971. 10.1002/bit.21057 [PubMed: 16897737]

Scandella CJ, Kilpatrick J, Lidster W, Parker C, Moore JP, Moore GK, … Haigwood NL (1993). Nonaffinity Purification of Recombinant gp120 for Use in AIDS Vaccine Development. AIDS Research and Human Retroviruses, 9(12), 1233–1244. 10.1089/aid.1993.9.1233 [PubMed: 8142140]

Schulz TF, Reeves JD, Hoad JG, Tailor C, Stephens P, Clements G, … Weiss RA (1993). Effect of Mutations in the V3 Loop of HIV-1 gp120 on Infectivity and Susceptibility to Proteolytic Cleavage. AIDS Research and Human Retroviruses, 9(2), 159–166. 10.1089/aid.1993.9.159 [PubMed: 8457383]

Sok D, Doores KJ, Briney B, Le KM, Saye-Francisco KL, Ramos A, … Burton DR (2014). Promiscuous Glycan Site Recognition by Antibodies to the High-Mannose Patch of gp120 Broadens Neutralization of HIV. Science Translational Medicine, 6(236), 236ra63–236ra63. 10.1126/scitranslmed.3008104

Srivastava IK, Stamatatos L, Legg H, Kan E, Fong A, Coates SR, … Barnett SW (2002). Purification and Characterization of Oligomeric Envelope Glycoprotein from a Primary R5 Subtype B Human Immunodeficiency Virus. Journal of Virology, 76(6), 2835–2847. 10.1128/JVI.76.6.2835-2847.2002 [PubMed: 11861851]

Srivastava Indresh K., Kan E, Sun Y, Sharma VA, Cisto J, Burke B, … Barnett SW. (2008). Comparative evaluation of trimeric envelope glycoproteins derived from subtype C and B HIV-1 R5 isolates. Virology, 372(2), 273–290. 10.1016/j.virol.2007.10.022 [PubMed: 18061231]

Srivastava Indresh K., Stamatatos L, Kan E, Vajdy M, Lian Y, Hilt S, … Barnett SW. (2003). Purification, Characterization, and Immunogenicity of a Soluble Trimeric Envelope Protein Containing a Partial Deletion of the V2 Loop Derived from SF162, an R5-Tropic Human Immunodeficiency Virus Type 1 Isolate. Journal of Virology, 77(20), 11244–11259. 10.1128/JVI.77.20.11244-11259.2003 [PubMed: 14512572]

Stanfield RL, Gorny MK, Williams C, Zolla-Pazner S, & Wilson IA (2004). Structural Rationale for the Broad Neutralization of HIV-1 by Human Monoclonal Antibody 447–52D. Structure, 12(2), 193–204. 10.1016/j.str.2004.01.003 [PubMed: 14962380]

Stephens PE, Clements G, Yarranton GT, & Moore J (1990). A chink in HIV's armour? Nature, 343(6255), 219 10.1038/343219b0 [PubMed: 2300166]

Sullivan JT, Sulli C, Nilo A, Yasmeen A, Ozorowski G, Sanders RW, … Doranz BJ (2017). High-Throughput Protein Engineering Improves the Antigenicity and Stability of Soluble HIV-1 Envelope Glycoprotein SOSIP Trimers. Journal of Virology, 91(22). 10.1128/JVI.00862-17

Valente KN, Lenhoff AM, & Lee KH (2015). Expression of difficult-to-remove host cell protein impurities during extended Chinese hamster ovary cell culture and their impact on continuous bioprocessing. Biotechnology and Bioengineering, 112(6), 1232–1242. 10.1002/bit.25515 [PubMed: 25502542]

Vouillot L, Thélie A, & Pollet N (2015). Comparison of T7E1 and Surveyor Mismatch Cleavage Assays to Detect Mutations Triggered by Engineered Nucleases. G3: Genes|Genomes|Genetics, 5(3), 407–415. 10.1534/g3.114.015834 [PubMed: 25566793]

Walker LM, Huber M, Doores KJ, Falkowska E, Pejchal R, Julien J-P, … Poignard P (2011). Broad neutralization coverage of HIV by multiple highly potent antibodies. Nature, 477(7365), 466–470. 10.1038/nature10373 [PubMed: 21849977]

Wen Y, Trinh HV, Linton CE, Tani C, Norais N, Martinez-Guzman D, … Barnett SW (2018). Generation and characterization of a bivalent protein boost for future clinical trials: HIV-1 subtypes CR01_AE and B gp120 antigens with a potent adjuvant. PLoS ONE, 13(4). 10.1371/journal.pone.0194266

Werner A, & Levy JA (1993). Human immunodeficiency virus type 1 envelope gp120 is cleaved after incubation with recombinant soluble CD4. Journal of Virology, 67(5), 2566–2574. [PubMed: 8474162]

Xu X, Nagarajan H, Lewis NE, Pan S, Cai Z, Liu X, … Wang J (2011). The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. Nature Biotechnology, 29(8), 735–741. 10.1038/nbt.1932

Yu B, Fonseca DPAJ, O'Rourke SM, & Berman PW (2010). Protease Cleavage Sites in HIV-1 gp120 Recognized by Antigen Processing Enzymes Are Conserved and Located at Receptor Binding Sites. Journal of Virology, 84(3), 1513–1526. 10.1128/JVI.01765-09 [PubMed: 19939935]

Yu B, Morales JF, O'Rourke SM, Tatsuno GP, & Berman PW (2012). Glycoform and Net Charge Heterogeneity in gp120 Immunogens Used in HIV Vaccine Trials. PLOS ONE, 7(8), e43903. 10.1371/journal.pone.0043903

Zhou T, Georgiev I, Wu X, Yang Z-Y, Dai K, Finzi A, … Kwong PD (2010). Structural Basis for Broad and Potent Neutralization of HIV-1 by Antibody VRC01. Science, 329(5993), 811–817. 10.1126/science.1192819 [PubMed: 20616231]

Zipfel PF, Würzner R, & Skerka C (2007). Complement evasion of pathogens: Common strategies are shared by diverse organisms. Molecular Immunology, 44(16), 3850–3857. 10.1016/j.molimm.2007.06.149 [PubMed: 17768102]

Zolla-Pazner S, Zhong P, Revesz K, Volsky B, Williams C, Nyambi P, & Gorny MK (2004). The Cross-Clade Neutralizing Activity of a Human Monoclonal Antibody Is Determined by the GPGR V3 Motif of HIV Type 1. AIDS Research and Human Retroviruses, 20(11), 1254–1258. [PubMed: 15588347]

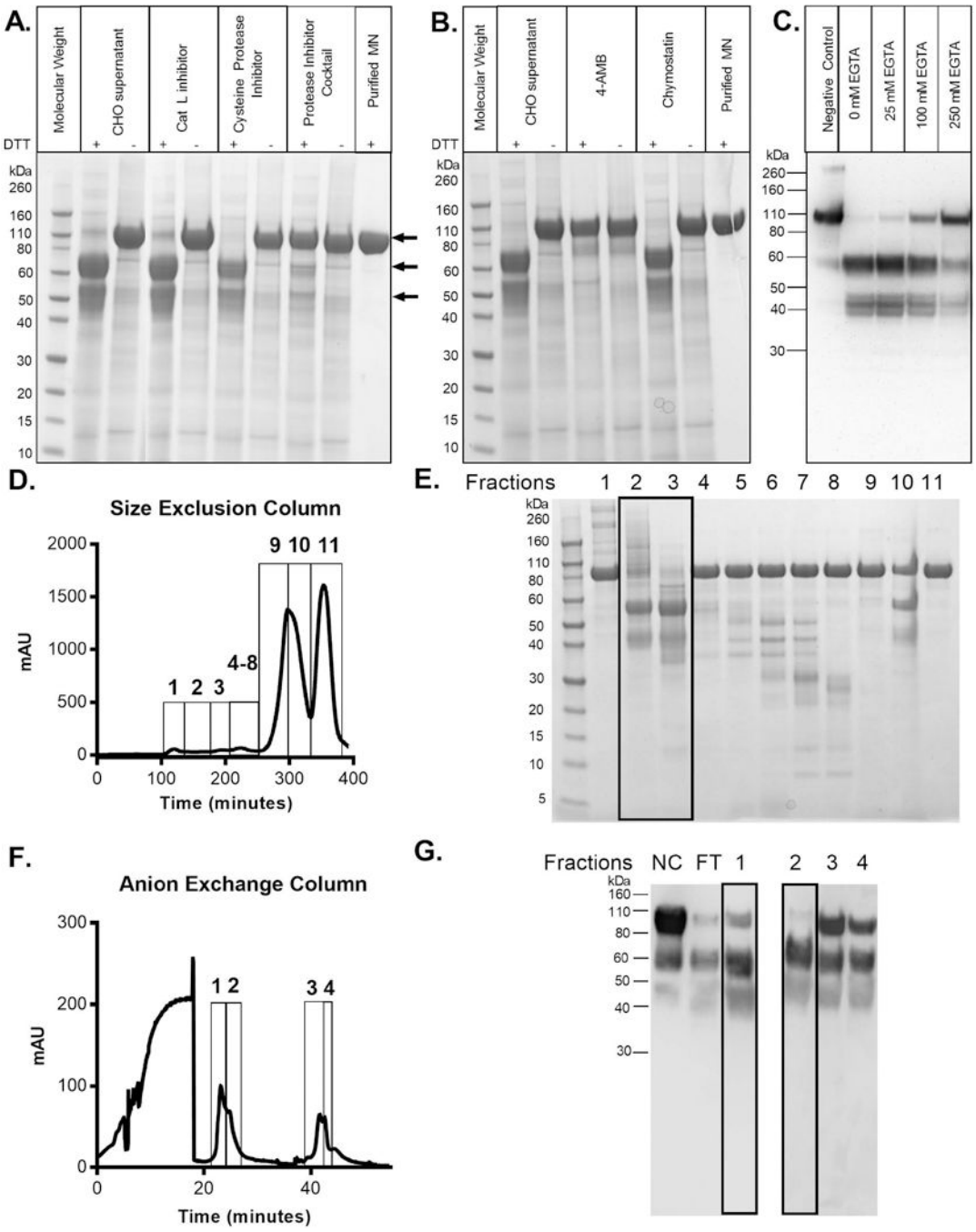


Figure 1. Characterization and enrichment of the CHO cell protease responsible for the degradation of clade B, gp120 immunogens.

(A) Purified, unclipped MN-rgp120 produced in HEK293 cells was incubated at 37°C overnight with growth-conditioned CHO cell culture medium containing the endogenous gp120-cleaving CHO protease. gp120 proteolysis was visualized by the appearance of a 50 kDa and 70 kDa fragment and seen only under reduced conditions by SDS-PAGE. Shown are the results obtained with the cathepsin L inhibitor, N-(1-Naphthalenylsulfonyl)-Ile-Trp-aldehyde; the cysteine protease inhibitor, Z-Phe-Ala-fluoromethylketone; and protease inhibitor cocktail, cOmplete Mini. (B) Panel B shows results obtained with 4-

aminobenzamidine, a serine protease inhibitor, and chymostatin, a serine and cysteine protease inhibitor. (C) Panel C shows results obtained with increasing amounts of EGTA, a calcium-chelating agent. In these studies, MN-rgp120 was added to 20X concentrated CHO supernatant and analyzed by immunoblot using a Protein G-purified goat polyclonal antibody raised against gp120. (D) Elution profile of CHO host cell proteins fractionated by size exclusion chromatography using an S200 gel filtration column. (E) Immunoblots showing the mobilities of gp120 protein fragments after overnight treatment at 37°C of purified MN-rgp120 with fractions from the SEC eluate. (F) Elution profile following anion exchange chromatography with a QHP sepharose column of proteins from fractions 2 and 3 in the SEC. (G) Immunoblots with α-gp120 polyclonal antibody showing the mobilities of gp120 protein fragments after overnight treatment at 37°C of purified MN-rgp120 with fractions from the QHP column. Fractions 1 and 2 were pooled and 60x concentrated for analysis by tandem MS.

A.

Uniprot Identifier	Description	Score	Coverage (%)	# Unique Peptides	# Peptides
G3H0E4	Chondroitin sulfate proteoglycan 4	124.16	18.30	14	34
G3HWE4	Nidogen-1	73.91	25.17	10	20
G3GUR0	**Calcium-dependent serine proteinase**	**45.54**	**14.53**	**6**	**9**
G3HN02	Talin-1	13.23	2.60	1	6
G3GXS2	EMILIN-1	5.57	4.62	1	5
G3H354	Heat shock protein HSP 90-alpha	20.54	16.57	3	5
G3H1K9	Alpha-actinin-1	17.42	6.42	4	5
G3I278	Laminin subunit beta-1	14.71	3.90	1	5
G3H8F4	Dystroglycan	9.90	4.37	1	4
G3IFA4	Multiple epidermal growth factor-like domains 6	2.12	4.23	1	4
G3HWE7	Dickkopf-related protein 3	6.44	4.02	1	2
G3I1Y9	Sulfated glycoprotein 1	5.56	8.43	1	2
G3HHR3	Vimentin	13.44	4.94	1	2
G3HC84	Heat shock protein HSP 90-beta	8.39	6.07	1	2
G3H171	G protein-coupled receptor kinase	5.32	2.90	1	2
G3HSX8	Biglycan	3.38	2.98	1	1
G3GYZ1	CD44 antigen	3.72	2.43	1	1
G3ILK7	Calsyntenin-1	4.04	2.16	1	1

B.

Substrate	P_4	P_3	P_2	P_1	P_1'	P_2'	P_3'	P_4'	Cleaved by C1s
BaL gp120	Gly	Pro	Gly	Arg	Ala	Phe	Tyr	Thr	+
Consensus B gp120	Gly	Pro	Gly	Arg	Ala	Phe	Tyr	Thr	+
Consensus A/E gp120	Gly	Pro	Gly	Gln	Val	Phe	Tyr	Arg	-
Consensus C gp120	Gly	Pro	Gly	Gln	Thr	Phe	Tyr	Ala	-
complement component C2	Asn	Leu	Gly	Arg	Arg	Ile	Gln	Ile	+
complement component C4	Gly	Leu	Ala	Arg	Ala	Gln	Glu	Val	+

Figure 2. CHO cell proteins identified in column fractions enriched for CHO protease by tandem MS and protein sequences sensitive to degradation by the C1s protease.

(A) The proteins identified in fractionated CHO cell supernatants by tandem MS were tabulated and listed along with the UniProt identifiers. The score provided by SEQUEST is the sum of the ion scores of all the peptide identified for the protein. Percent coverage indicates the extent of the protein sequence identified from the peptides. The number of unique peptides and total number of peptides are also provided. (B) Alignment of protein sequences from HIV gp120 and CHO cell complement proteins that are sensitive or resistant to degradation by the C1s protease. Standard P1/P1* nomenclature was used to identify

conserved residues on either side of the scissle bond. Sequences are listed for the V3 crown sequences of BaL-rgp120 and the consensus sequences for clade B, clade CRF01-AE, and Clade C gp120. Also included are alignments of the normal physiologically-relevant substrates of C1s, complement components C2 and C4 (UniProt identifier G3HZE1 and the UniProt identifier UPI0004542963).

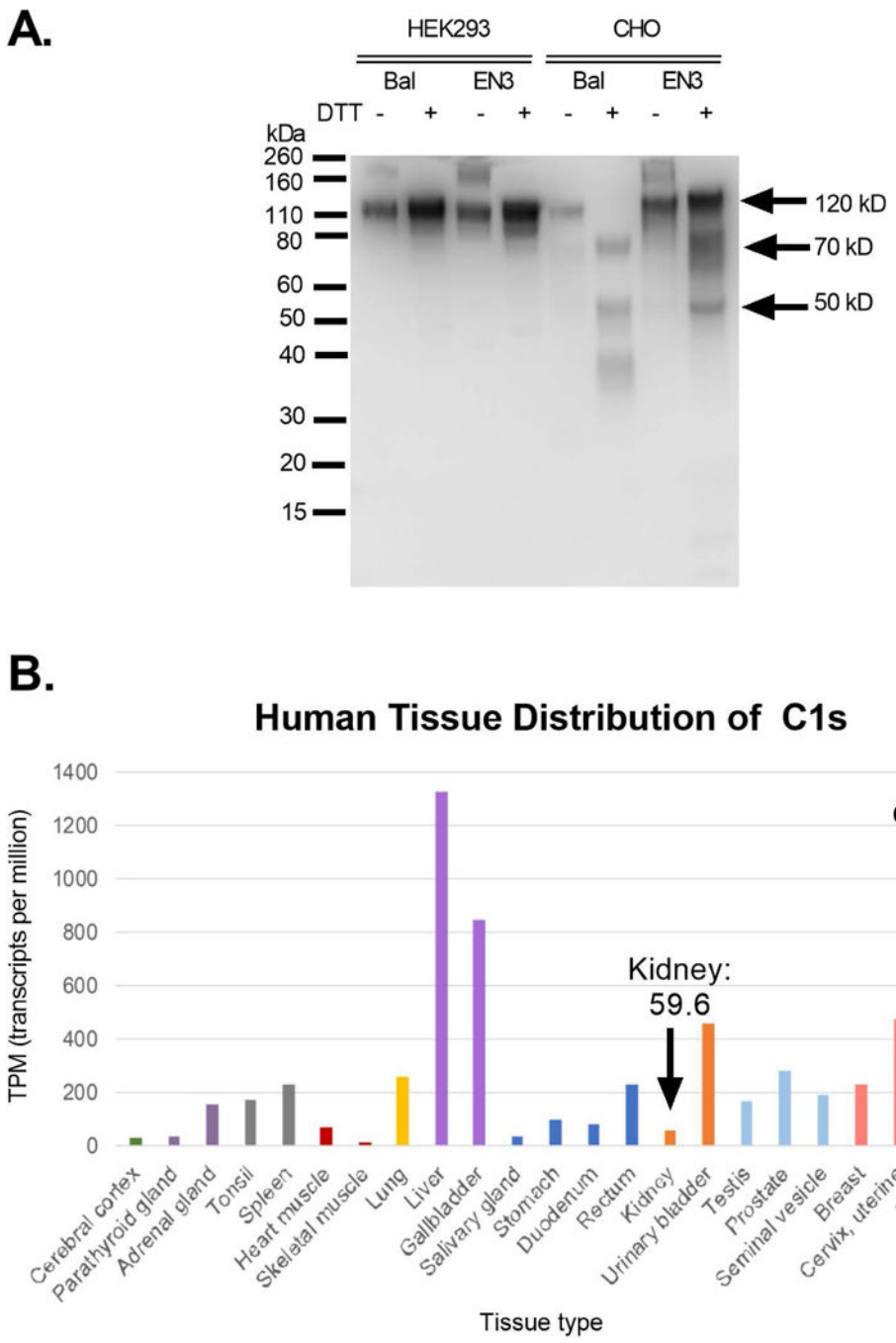


Figure 3. Differential expression of C1s in different cell types.

(A) BaL-rgp120 and clade B, EN3-rgp120 were transiently expressed in HEK293 and CHO cells. Supernatants was analyzed by immunoblot using an α-gp120 polyclonal antibody. Proteolytic degradation was indicated by the appearance of bands at 50 kDa and 70 kDa. (B) RNA seq expression data for C1s in human expression tissues was plotted based on data from the Human Protein Atlas. RNA expression values in transcripts per million for kidney and ovary tissues are labeled.

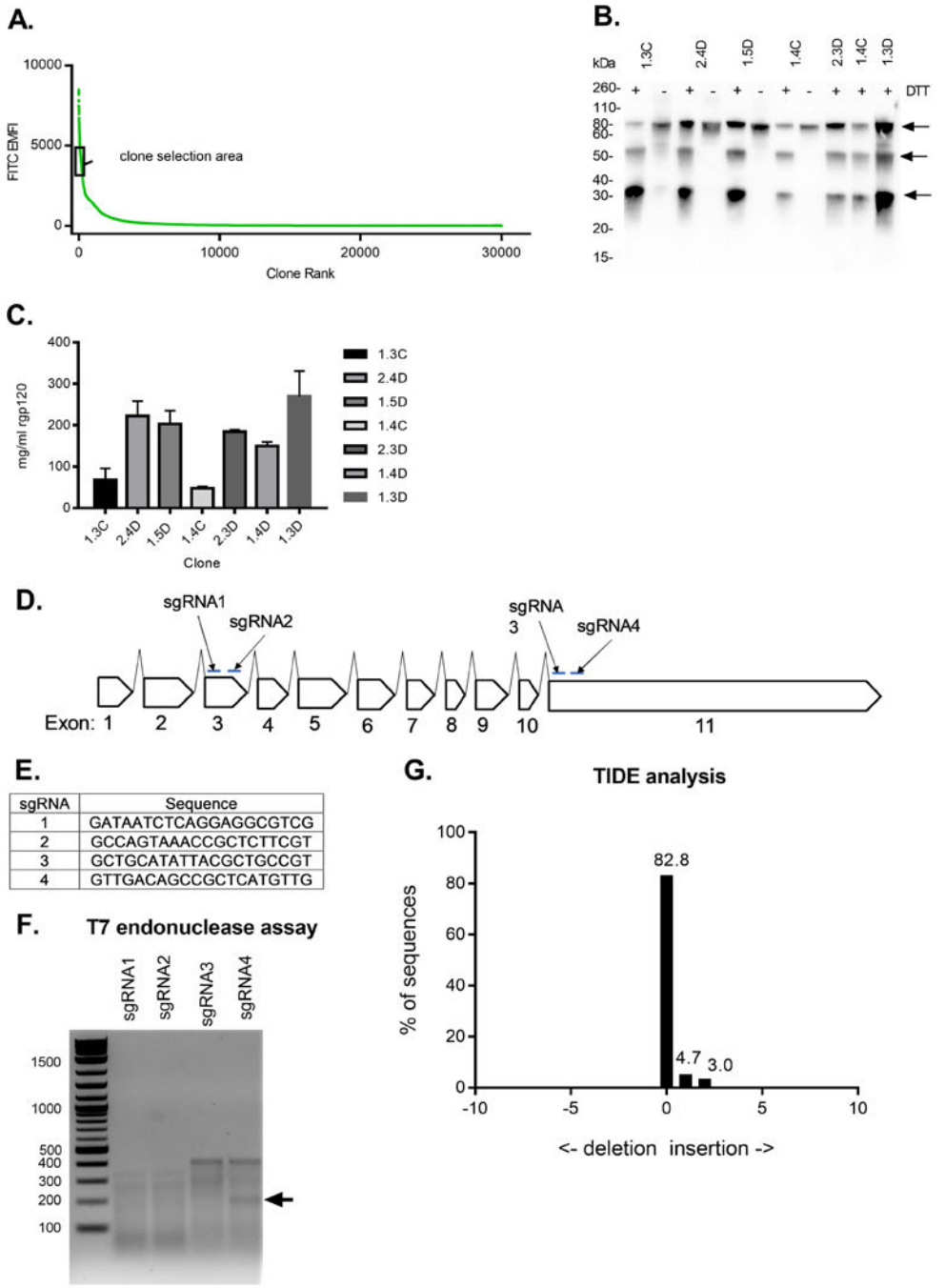
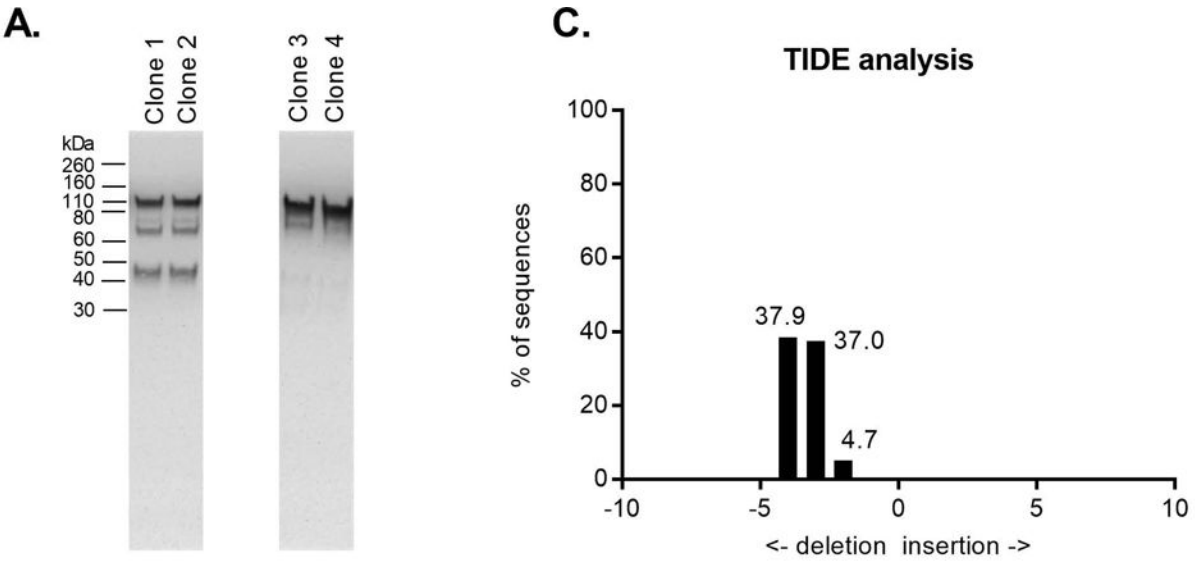


Figure 4. Production of a stable cell line expressing BaL-rgp120 and screening colonies for CRISPR/Cas9 knockout of C1s.

(A) 30,000 colonies of MGAT1$^-$ CHO cells were transfected with a plasmid expressing both BaL-rgp120 and G418 (geneticin) resistance. These were analyzed for rgp120 expression using the Clonepix2 robot. Colonies were ranked by exterior mean fluorescence intensity (EMFI), which was plotted as a function of clone rank. Approximately 0.01% of colonies, or 42 colonies, were selected and expanded. (B) Following selection and expansion, supernatants from seven cell lines were assayed by Western blot for the presence of Bal-rgp120 protease cleavage products (indicated by arrows). (C) Quantitative yield of rgp120

from the eight-day batch-fed cultures as determined by ELISA. (D) Diagram of the C1s gene showing intron and exon structure with the exon length to scale and the location of single guide RNAs (sgRNA) as indicated with arrows. (E) Sequences of sgRNAs that were tested for the ability to knockout the C1s gene. (F) Selection of single guide RNA for CRISPR/Cas9 knockout of C1s. Knockout efficiency was qualitatively measured in a population of cells post-transfection by the T7 endonuclease assay for sgRNA1, sgRNA2, sgRNA3 and sgRN4. A 400 bp PCR product was generated around the Cas9 target site using PCR primers from upstream and downstream of the sgRNA sites. The arrow indicates the 200 bp PCR product generated by the T7 endonuclease at the secondary structures formed by indel-containing DNA. (G) The quantitative TIDE analysis for sgRNA4 in a population of cells post-transfection indicated a Cas9-knockout efficiency of 11.4%.



Clone 1	5'-GTTGACAGCCGCTCATG TTGTGG-3'	WT	
Clone 2	5'-GTTGACAGCCGCTCATG TTGTGG-3'	WT	
Clone 3: edit 1	5'-GTTGACAGCCGCTCATG ------TGG-3'	Δ -3bp	Deletion of TTG
Clone 3: edit 2	5'-GTTGACAGCCGCTCATGTTTGTGG-3'	Δ+1bp	Insertion of T
Clone 4: edit 1	5'-GTTGACAGCCGCTCATG -------GG-3'	Δ -4bp	Deletion of TTGT
Clone 4: edit 2	5'-GTTGACAGCCGCT-------TTTGTGG-3' OR	Δ -3bp	Deletion of CATG, Insertion of T

Figure 5. Screening clones for a stable C1s$^{-/-}$ MGAT^{-} CHO cell line expressing BaL-rgp120.

(A) Out of thirty-eight clones, three were identified with Cas9-mediated genetic knockouts of the C1s gene. For simplicity, data from only two are shown. Immunoblot analysis of BaL gp120 supernatants of clones 1 through 4 is shown. Clones 1 and 2 are representative of CHO cell clones where the C1s gene was unmodified, while clones 3 and 4 are examples of clones where the C1s gene was inactivated. (B) Sanger sequencing was used to characterize the genetic edits created by Cas9-editing in cells from clone 3 and clone 4. (C) TIDE was used to confirm the presence of two mutations in the C1s gene of clone 4.

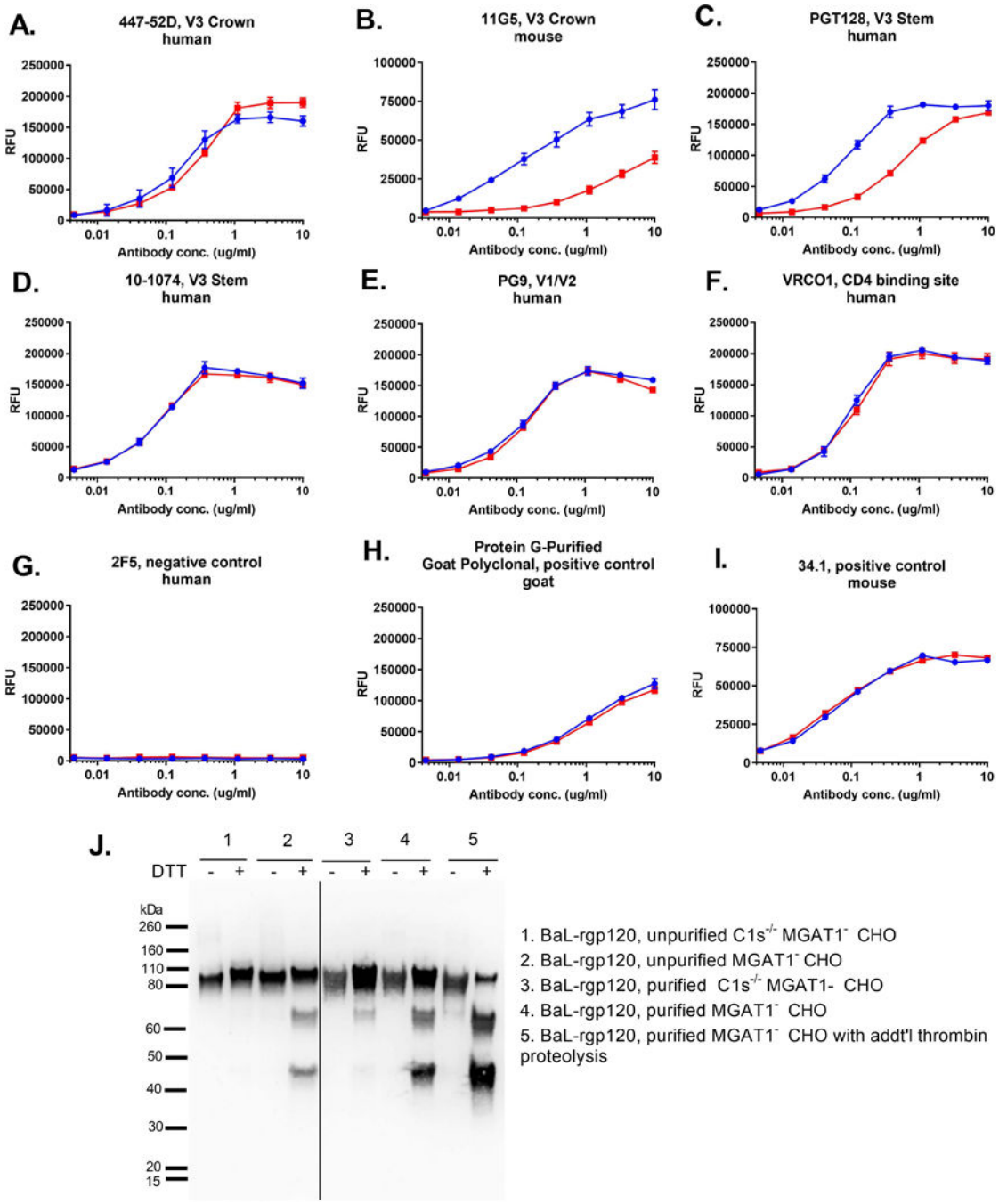


Figure 6. bN-mAb and V3-antibody binding to intact and proteolyzed BaL-rgp120.

A direct coat FIA was performed to measure bNAb binding to Bal-rgp120 produced in C1s $^{-/-}$ MGAT1^{-} CHO cells and normal MGAT1^{-} CHO cells. Gp120s from MGAT1- CHO cells were further cleaved by treatment with thrombin as described in the Results section. The binding of mAbs to proteolyzed BaL-rgp120 (red) and non-proteolyzed BaL-rgp120 is indicated (blue). (A-F) Included were antibodies to the V3 domain (447–52D, PGT128, 10–1074, and 11G5), the V1/V2 domain (PG9) and the CD4 binding site (VRCO1). (G) The human bN-mAb, 2F5, directed to gp41 was included as a negative control. (H, I) Polyclonal

goat antibodies to gp120 and the 34.1 mouse monoclonal antibody to the N-terminal gD purification tag served as positive controls. (J) Immunoblot analysis of the proteins used in these antibody binding assays are shown. The figure shows both protein from supernatant as well as purified protein.

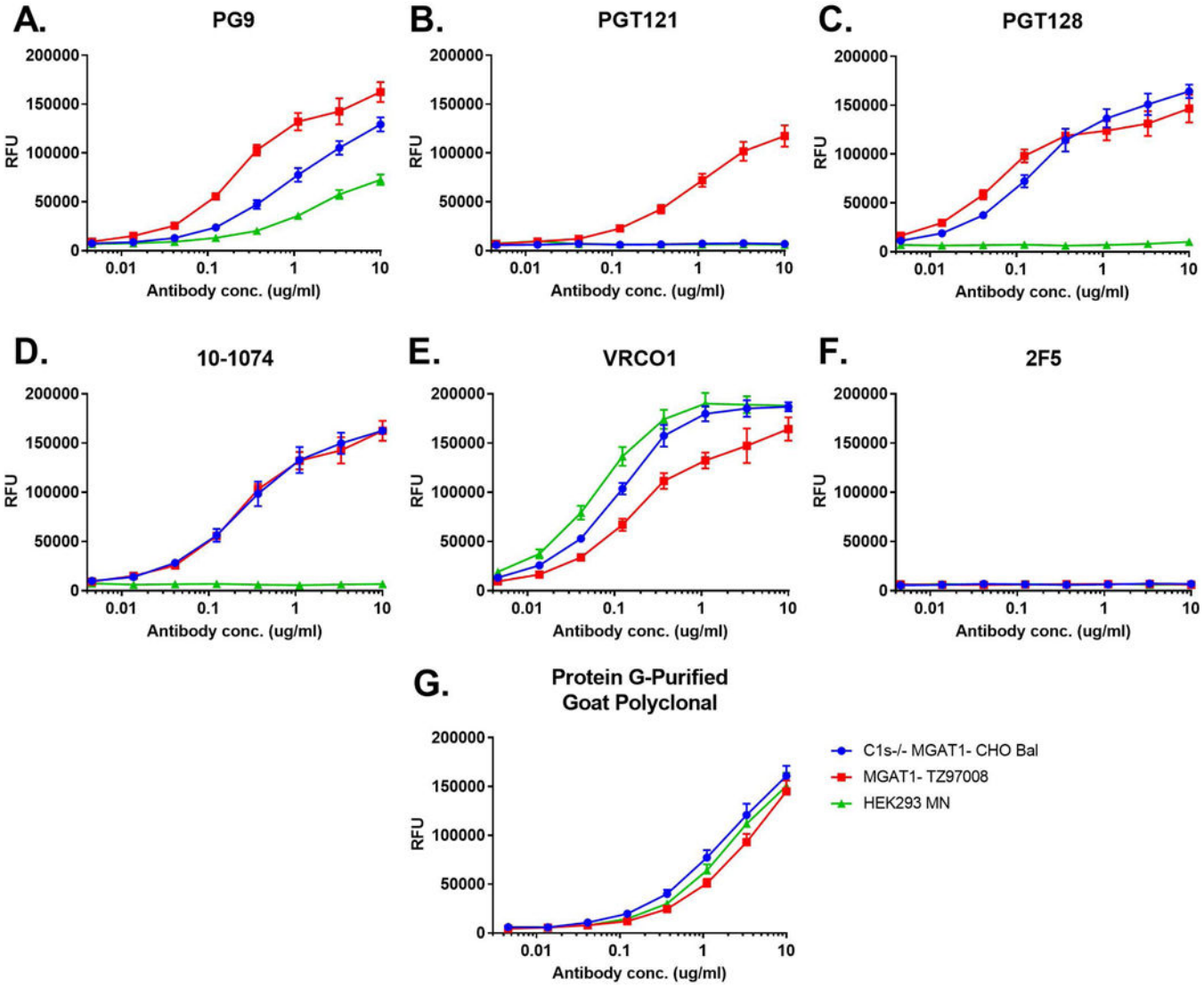


Figure 7. Comparison of bNAb binding to MN-rgp120 used in the RV144 trial with Bal-rgp120 and TZ97008-rgp120 produced in CRISPR-engineered cell lines.

(A-F) Unclipped gp120s possessing normal N-linked glycosylation (MN-rgp120) or N-linked glycosylation restricted to mannose-5 glycans (Bal-rgp120 and TZ97008-rgp120) were compared for the binding of several prototypic bN-mAbs by FIA. Unclipped MN-rgp120 (green) was prepared by expression in HEK293 cells. TZ97008-rgp120 (red) was expressed in MGAT1- CHO cells. Bal-rgp120 (blue) was expressed in $C1s^{-/-}$ $MGAT1^{-}$ CHO cells. (F) The bN-mAb 2F5 was used as a negative control. Bal-rgp120 (blue) was expressed in $C1s^{-/-}$ $MGAT1^{-}$ CHO cells. (G). Polyclonal antibodies to gp120 were used as a positive control.