Protein Cell, 2025, **16**, 905–910
https://doi.org/10.1093/procel/pwaf028
Advance access publication 25 May 2025
Letter

Protein & Cell

Letter

Therapeutic implications of synonymous gene recoding: insights into mechanisms controlling protein biogenesis and activity

Brian C. Lin[1,†], Katarzyna I. Jankowska[1,†], Upendra K. Katneni[1,†], Randilu Amarasinghe[1], Nigam Padhiar[1], Nobuko Hamasaki-Katagiri[1], Wells W. Wu[2], Haojie Zhu[3], Hideki Taguchi[3], Arnab Ghosh[4,7], David D. Holcomb[1], Je-Nie Phue[2], Sarah E. Fumagalli[1], Darón I. Freedberg[5], Ofer Kimchi[6], Rong-Fong Shen[2], Anton A. Komar[4], Zuben E. Sauna[1], Chava Kimchi-Sarfaty[1,*]

[1]Division of Hemostasis, Office of Therapeutic Products, Center for Biologics Evaluation and Research, US FDA, Silver Spring, MD, 20993, United States
[2]Facility for Biotechnology Resources, Center for Biologics Evaluation and Research, US FDA, Silver Spring, MD, 20993, United States
[3]Cell Biology Center, Institute of Innovative Research, Tokyo Institute of Technology, Yokohama, 226-8503, Japan
[4]Center for Gene Regulation in Health and Disease and the Department of Biological, Geological and Environmental Sciences, Cleveland State University, Cleveland, OH, 44115, United States
[5]Laboratory of Bacterial Polysaccharides, Division of Bacterial, Parasitic, and Allergenic Products, Center for Biologics Evaluation and Research, US FDA, Silver Spring, MD, 20993, United States
[6]Center for the Physics of Biological Function, Princeton University, Princeton, NJ, 08544, United States
[7]Present address: Integral Molecular, Philadelphia, PA, 19104, United States

[†]These authors contributed equally: Brian C. Lin, Katarzyna I. Jankowska, and Upendra K. Katneni.
[*]Correspondence: Chava.Kimchi-Sarfaty@fda.hhs.gov (C. Kimchi-Sarfaty)

Dear Editor,

Recombinant protein therapeutics and gene therapies can help rescue monogenic disease phenotypes (Ebrahimi and Samanta, 2023). Sufficient/high expression of a therapeutic product is an important consideration in this regard, which is usually achieved by gene redesign, focused on altering synonymous codon usage/codon context, aimed at enhancing translational rates (Alexaki et al., 2019). Different approaches exist to adjust synonymous codon usage and boost protein expression levels (Katneni et al., 2022). As these strategies do not alter the primary sequence of a protein (Liu et al., 2021) they are not expected to alter protein structure and function. However, many recent studies have demonstrated that codon optimization can influence mRNA and protein biogenesis (Hunt and Kimchi-Sarfaty, 2022).

Studies in the field additionally revealed that orientation-specific codon-pairs can alter the translational elongation rate, separately from the effects of modifying individual codon usage biases (Gamble et al., 2016) and optimizing genetic sequences through "codon-pair biases" (codon-pair optimization) became a novel strategy to boost protein expression (Huang et al., 2021).

In this study, we examined the effects of gene recoding through codon (CO) and codon-pair (CPO) optimization strategies on human ADAMTS13 (a disintegrin-like and metalloproteinase with a thrombospondin type 1 motif, member 13) properties and evaluated the cellular mechanisms underlying biogenesis of CO and CPO ADAMTS13 variants in Flp-In single-copy targeted integration system. We found that synonymous gene recoding alters numerous protein and cellular attributes, which were not uncovered for this protein earlier, including protein structure, function, immunogenicity, and cellular bioenergetics.

To evaluate the effects of gene recoding, we compared wild-type (WT), the previously described single synonymous (P118P) variant (Hunt et al., 2019) used in some experiments and CO and CPO variants of *ADAMTS13*

Accepted 16 March 2025.

designed using different algorithms (detailed in Supplementary Materials; sequences in Supplementary Data File 1) that overall introduced more frequent and faster-translating codons throughout the ADAMTS13 sequence, predicted to enhance protein expression (Figs. 1A,1B, and S1; Table S1). The enrichment of common codons throughout the optimized sequences was reflected in the increase of codon adaptation index (CAI) %MinMax, relative synonymous codon usage, and relative synonymous codon-pair usage values (Figs. 1A, 1B, S1A and S1B). However, while overall patterns in expected translational rates based on %MinMax remained similar between CO and CPO, CPO uniquely retained many trough-like patterns (present in WT), suggesting that CPO may maintain the co-translational folding pathway similar to WT (Table S1). Regions that involve post-translational modifications (PTMs) exhibited more negative %MinMax values for WT than for optimized variants (Fig. 1B). We also found that CO variant has slightly higher mRNA minimum free folding energy levels (–1,903.70 kcal/mol) (being thus less stable) compared to CPO (–2,140.10 kcal/mol) and WT (–2,051.90 kcal/mol) variants (Figs. S1C–E).

Cell-free *in vitro* translation experiments (Figs. 1C, 1D, S2A and S2B) revealed that synonymous mutations had a significant impact on ADAMTS13 translation kinetics: translation rate constants of CPO and P118P (shown previously to have elevated expression levels compared to WT) (Hunt et al., 2019) were twice that of CO (Fig. 1C). Consequently, translation yields were significantly lower in CO and higher in CPO (Fig. 1D). In addition, analysis of ADAMTS13 solubility by ultracentrifugation showed that all variants had relatively high solubility (Fig. S2B), and minimal aggregation-prone misfolding.

We then performed cellular expression experiments using Flp-In HEK293 cell lines. We determined that intracellular mRNA levels were not significantly different between WT, CO, and CPO (Fig. 1E), however, extracellular expression of CO was significantly higher and CPO significantly lower than WT (Fig. 1F), suggesting that intracellular environment affects protein expression and biogenesis in a different way compared to the *in vitro* system. Apparent specific activity measured, using the fluorogenic fluorescence resonance energy transfer-von Willebrand factor (FRETS-VWF73) assay, appeared to be similar for CPO and WT variants (Figs. 1G and S2C). To analyze ADAMTS13 activity in more detail we performed subsequent substrate binding kinetic studies using biolayer interferometry (BLI) and additional enzymatic assays. We further found that WT and CPO ADAMTS13 had similar V_{max} and K_m rates, while CO ADAMTS13 had showed significantly higher V_{max} and K_m values (Figs. 1H, 1I, and S2D) indicating a lower affinity to VWF. We further determined that CO had a smaller dissociation constant (K_d), compared to WT (Figs. 1J, 1K, and S2E), and a significantly higher association rate, k_{on} and non-significantly lower dissociation rate, k_{off} compared to WT, while CPO had a similar binding affinity compared to WT (Fig. 1L and 1M).

As altered enzymatic kinetics and activities have been reported to be associated with altered co-translational folding (Jiang et al., 2022), we thus further used circular dichroism (CD) to evaluate the ADAMTS13 structure. While CD (Figs. 2A, 2B, and S3) followed by structure determination using BestSel software predicted similar secondary structure composition of variants, denaturation and refolding experiments (Greenfield, 2006) nevertheless revealed some differences in folding dynamics (Figs. 2B and S3), suggesting also potential differences in protein stabilities.

We further performed cycloheximide-chase (CHX) assays and tracked ADAMTS13 intracellular stability over a 6 h period. We observed that the CO variant was significantly more stable compared to WT and CPO, which both were barely detectable after 6 h (Fig. 2C and 2D). Conformational differences in recoded ADAMTS13 were also probed through recombinant VWF (rVWF) digestion assays (Fig. S4) and an inhibitory antibody assay (Fig. S5), but both assays demonstrated no statistically significant differences.

The observed differences of recoded ADAMTS13 variant features may result from altered protein folding and secretion, thus altering cellular fitness. We therefore evaluated whether cellular phenotypes were altered depending on the expressed variant. To evaluate bioenergetics, we performed Seahorse respiration assays. We observed that cells expressing the WT variant had similar respiration rates and ATP production as control HEK293 cell lines, while CO and CPO-expressing cells had higher ATP production and maximal oxygen consumption rate compared to WT (Fig. 2E and 2F).

We hypothesized that higher ADAMTS13 variant expression levels could incur greater stress on the secretory system. We note that Flp-In cells express higher levels of CO ADAMTS13, while CPO levels are not significantly different, compared to WT (Fig. 2G and 2H). Consequently, we further found that the intracellular levels of endoplasmic reticulum (ER) stress markers, BiP, and phosphorylated-eIF2α are higher in cells expressing the CO variant (Fig. 2G, 2H, and S6). BiP is an ATP-dependent (ER-resident) chaperone, and thus increased chaperone surveillance may be required to support protein folding challenges associated with recoded proteins. CHX administration, which reduces protein translation stress, resulted in greater secretion of CO ADAMTS13 (Figs. 2G, 2H, and S6). In addition, greater ADAMTS13-BiP interaction with CO ADAMTS13 was found via immunoprecipitation assays (Fig. 2I). The ratio of BiP:ADAMTS13 binding was ~3–5-fold greater with CO ADAMTS13 than WT (Fig. 2J). These experiments clearly

A

*	# Nucleotide differences	% Nucleotide change	Total codons changed	% Codon optimized	CAI	CPAI	Average translation rate	Average % MinMax	CpG count	GC%	Predicted NEW splice sites within CDS*
WT	0	0	0	0	0.826	0.598	0.503	49.2	208	0.651	1
CO	853	19.93	739	51.79	0.921	0.634	0.524	83	185	0.625	2
CPO	630	14.72	534	37.42	0.918	0.88	0.513	76.3	79	0.66	0

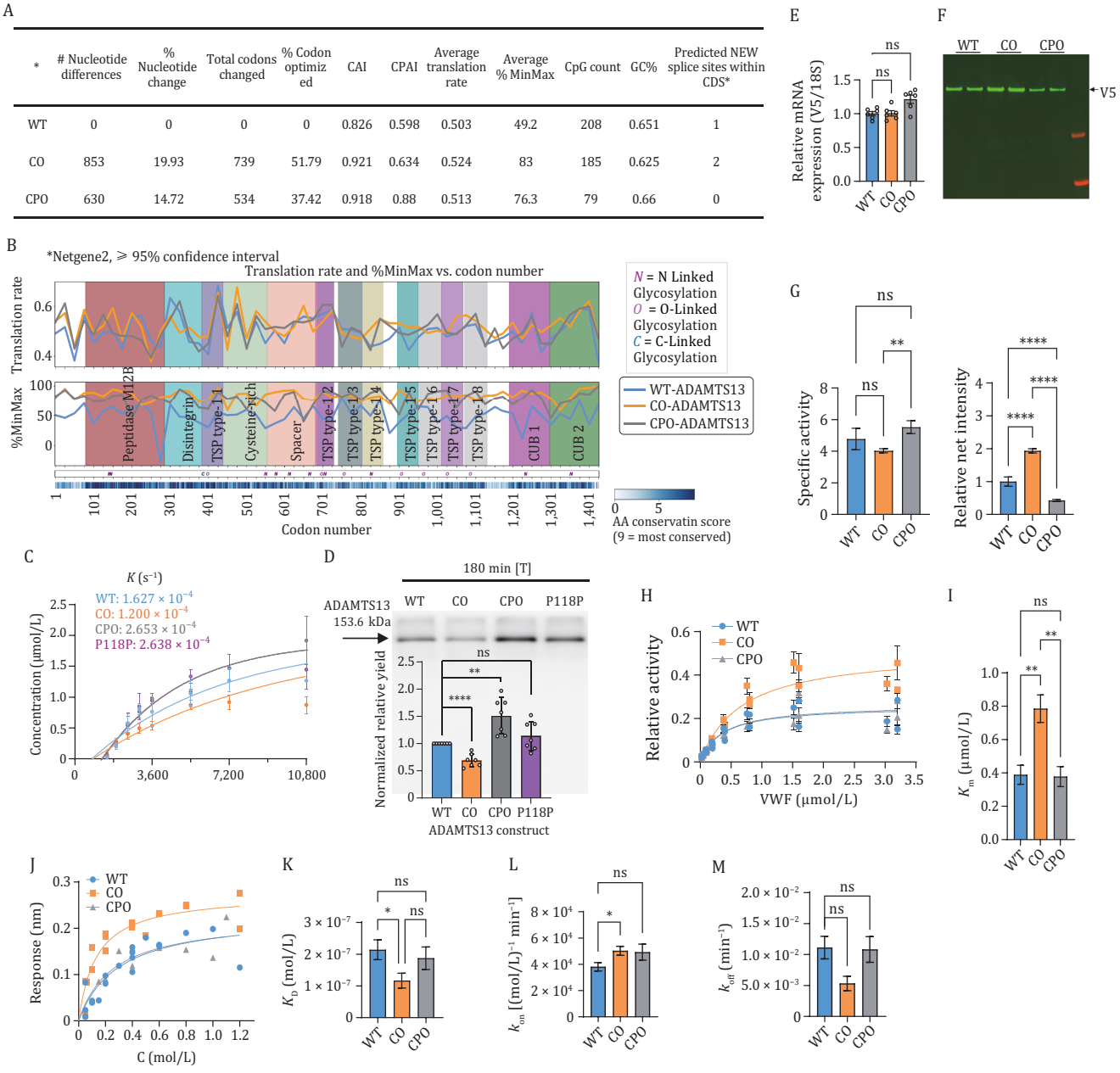

Figure 1. Properties of *ADAMTS13* variants. (A) Nucleotide, codon changes, and additional sequence properties, including codon and codon-pair adaptation indices (CAI, CPAI) of transcripts, calculated using *H. sapiens* codon usage frequencies in codon and codon-pair optimized ADAMTS13 relative to WT. Interestingly, both CO and CPO exhibit similar CAI indices even though CO ADAMTS13 was more extensively optimized (by ~15% more codons). (B) Graphs of translational rates (61) and %MinMax values of CO and CPO ADAMTS13 transcripts plotted against codon number. Calculations of codon usage metrics have been smoothed by averaging ~13 codons. Domains in ADAMTS13. Below the graphs are bars with marked locations of post-translational modification sites and a heatmap of amino acid conservation scores from Consurf with darker colors signifying more conserved sequences. Based on %MinMax, many sequence motifs within conserved regions, predicted to be potentially large areas of translational pauses, such as around codons 280 and 897, were eliminated from the recoded sequences. (C) Kinetic analysis of *ADAMTS13* translation in the rabbit reticulocyte lysate (RRL) system. Dots represent experimental values of ADAMTS13 expression levels, and the curve represents the fitted value. *ADAMTS13* translation rate constants are shown. As a positive control, we included a P118P synonymous variant of ADAMTS13 that was previously demonstrated to result in a 1.2-fold increase in expression (Hunt et al., 2019). (D) Representative blot of *ADAMTS13* translation and the total ADAMTS13 expression levels at 180 min from RRL assay. For each set of experiments, the intensity of the total expression level of WT was set to 1. (E) Relative mRNA expression and (F) Extracellular ADAMTS13 expression relative to WT and normalized to GAPDH (gel image and quantification data). (G) Apparent specific activity, as determined by activity relative to expression levels in Flp-In systems. (H) ADAMTS13 binding affinity to VWF73 substrate and kinetics evaluated by FRETS activity assay with increasing VWF concentration and data was fitted with Michaelis Menten equation to derive kinetic values for K_m in (I). (J) Binding kinetics of ADAMTS13 variants to VWF73 substrate measure by BLI using an Octet RED96 (ForteBio). (K) Dissociation constant (K_d) determined from steady state analysis (nonlinear fitting of response at equilibrium (R_{eq}) at different ADAMTS13 concentrations), from different sets of proteins along with association rate k_{on} in (L), and dissociation rate k_{off} (M) determined from nonlinear fitting, from four different sets of protein batches, measured at various concentrations (Statistics: mean ± SEM, One-way analysis of variance (ANOVA)).

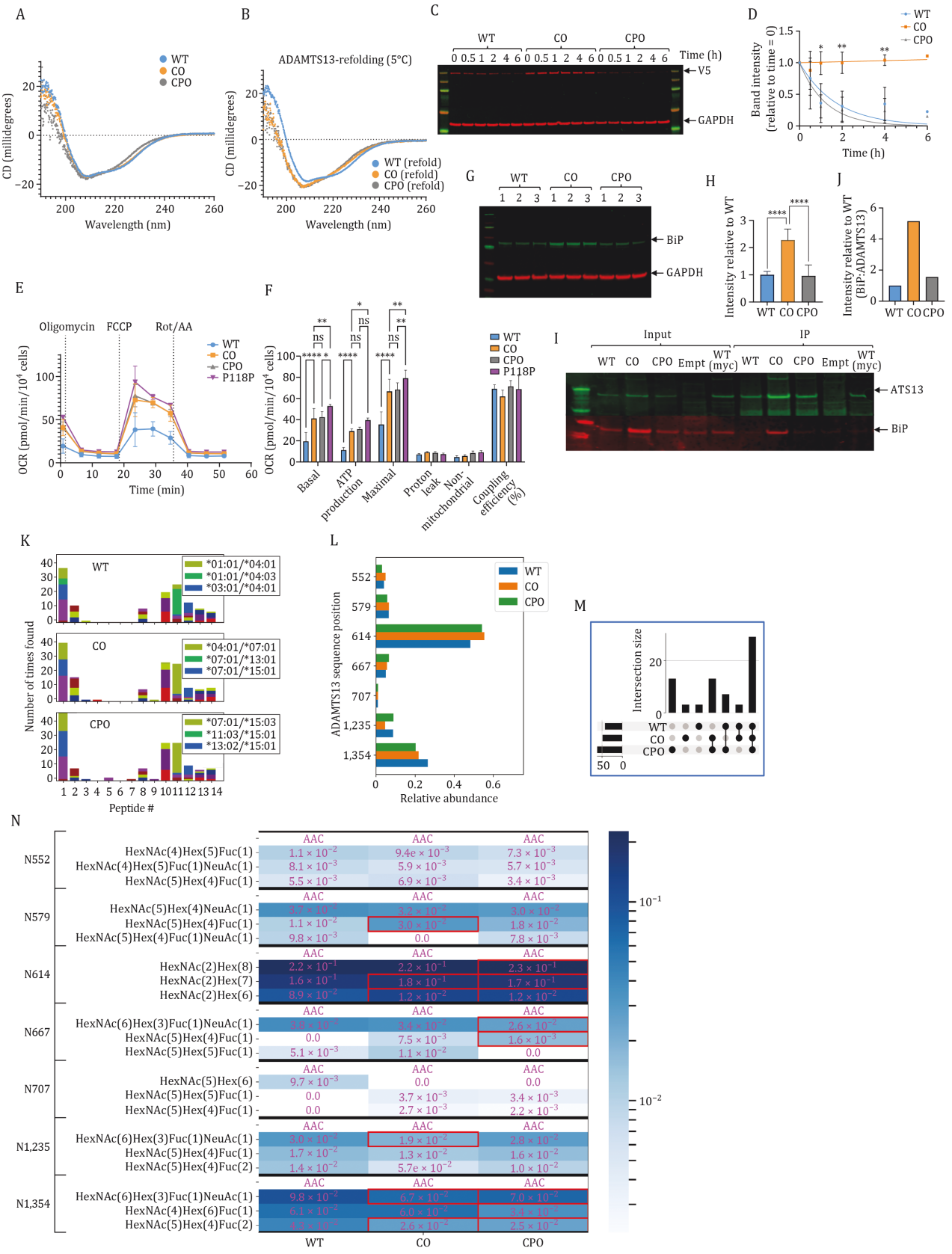

Figure 2. Synonymous gene recoding affects protein structure, stability, cellular bioenergetics, immunogenicity, and PMTs of ADAMTS13. (A) CD structures of WT, CO, and CPO purified ADAMTS13, measured at 5°C. (B) CD structures of ADAMTS13 at 5°C after proteins were subjected to an unfolding-refolding protocol. (C) Cycloheximide (CHX) protein stability assay of Flp-In cells expressing the 3 ADAMTS13 variants after incubation with CHX. Intracellular degradation of ADAMTS13 proteins were measured with immunoblots over a 6-h time course. (D) Quantification of CHX protein stability assay. (E) Seahorse assay of cells expressing WT, CO, and CPO ADAMTS13. (F) Quantification of bioenergetic metrics from (E). (G) Immunoblot of ADAMTS13 Flp-In derived lysates,

suggest altered cellular fitness during the expression of recoded proteins.

As recoded ADAMTS13 variants revealed differences in protein conformation, we hypothesized that the proteolytic processing of the variants within antigen-presenting cells could also differ. Major histocompatibility complex (MHC)-associated peptide proteomics (MAPPs) assay, using monocyte-derived dendritic cells (MoDCs) from 12 donors revealed distinct sets of MHC-II peptides presented for all three ADAMTS13 variants (Fig. S8A). Although most donors showed similar peptide presentations, subtle differences were observed in certain peptides (Fig. 2K). For example, peptides 4 and 9 were presented by MoDCs in response to CO, but not to WT. However, small total numbers of peptides precluded any exhaustive statistical analyses. As a result, to further explore immunogenicity differences, we investigated whether recoded ADAMTS13 could elicit $CD4^+$ T-cell proliferation, and statistically significant differences in T-cell stimulation were found between WT and CO and CPO proteins (Fig. S8B and S8C; Table S2).

We also evaluated glycosylation profiles of the ADAMTS13 variants (Figs. 2L–N and S9; Tables S3 and S4) since glycosylation is known to alter secretory protein expression levels (Ricketts et al., 2007; Sorvillo et al., 2014). Out of 10 reported N-linked glycopeptides (Verbij et al., 2016), we identified seven. The three missing N-linked glycopeptides (N142, N146, and N828) were outside the detection limits by mass spectrometry. As shown in Fig. 2L, all ADAMTS13 proteins were glycosylated at similar levels. However, a slight increase in glycosylation abundance were observed in CO and CPO at the N614 site, while a decrease in glycosylation abundance were seen at N1235 for CO and at N1354 for CO and CPO. Further quantitative determination of site-specific N-glycopeptides revealed distinct glycosylation profiles for each ADAMTS13 variant. Comparing all glycans identified across all N-linked sites, we found that WT was less glycosylated than CO and CPO (Fig. 2M and 2N).

Changes to PTM may be caused by protein structural differences, which alter the degree of accessibility of potential PTM sites (Esmail and Manolson, 2021). Interestingly, we found (Figs. 2N and S9A; Table S5) that CO and CPO variants had more complex glycans, from which the most abundant ones were the fucosylated species. The N614 position was the most highly glycosylated site across all variants (Fig. 2L; Table S3). Glycan structures and abundance levels were clearly distinct (Figs. 2N and S9A; Tables S3 and S5) and O-fucosylation differences were also observed between the CO and CPO variants, having different modifications at the S863 and S1027 sites (Fig. S9B; Table S4). Moreover, fucosylation of additional O-sites (S886, T889, and S1288) were uniquely detected in CO and CPO. They were shown to be required for ADAMTS13 secretion (Ricketts et al., 2007). ADAMTS13 proteins were also modified by C-mannosylation. (Fig. S9B; Table S4), representing impactful determinants of protein folding. Changes to mannosylation sites, caused e.g., by W387A mutation were shown to affect ADAMTS13 secretion and activity (Ling et al., 2013). The observed differences in PTMs could thus explain differences in expression/secretion (Fig. S10), and ADAMTS13 properties.

In this study, we carried out a systematic evaluation of the biochemical and biophysical consequences of gene recoding using WT, CO, and CPO ADAMTS13 variants. We revealed multiple changes in protein expression, stability, conformation, PTMs, and immunogenicity of recoded variants.

Our studies indicate that careful evaluation of the recoded transgene within the context of the host cell characteristics is required to inform the optimal design of protein therapeutics. We demonstrate that recoding can result in undesirable protein properties that could affect drug safety. Understanding the molecular and cellular mechanisms that underlie changes in the biophysical and biochemical properties of proteins following recoding is important to circumvent the limitations of gene recoding.

Supplementary data

Supplementary data is available at *Protein & Cell* online https://doi.org/10.1093/procel/pwaf028.

Footnotes

This study was in part supported by grant R01HL151392 (to A.A.K) and AMED-CREST with Grant Number JP21gm1410008 (to H.T.). This work was supported by funds from the Center for Biologics Evaluation and Research, US Food and Drug Administration.

Authors declare that they have no competing interests. The authors declare their agreement to publish.

probed for ER stress chaperone BiP. (H) Quantification of immunoblot in (G). (I) Immunoprecipitation experiment for assessing binding interaction between BiP–ADAMTS13. (J) Quantification of ratio of BiP:ADAMTS13 in (I). (K) MAPPS assay with peptides identified based on HLA alleles. (L) Post-translational modification analysis of purified Flp-In derived ADAMTS13 proteins shown by grouped Bar plot depicting relative abundance of glycans at each site, with most abundant glycan annotated for each bar. (M) Upset plot depicting shared glycan species between each group. (N) Heatmap depicting the abundance of the top three N-glycosylations at each site, as calculated by Byologic software. The codon identity at the site is labeled at the top of each block of the heatmapin. The color bar is log-scaled. Red boxes highlight differences > 0.01 compared to WT (Statistics: mean ± SEM, One-way ANOVA).

All data needed to evaluate the conclusions in the paper are presented in the paper. Additional data is available from authors upon request.

Commercially available software was used to do the majority of the analysis, including GraphPad Prism for line charts and bar graphs (with significance testing) and Protein Metrics Byologic software for glycosylation analysis. Publicly available python visualization libraries (matplotlib, seaborn) were used for a small portion of the visualizations. Additional information can be provided upon request.

C.K-S. conceptualized the study. C.K-S., B.C.L., K.I.J., U.K.K., R.A., N.H-K., W.W.W., Z.H., H.T., A.G., J-N.P., D.I.F., R-F.S., A.A.K. and Z.E.S. designed and/or performed experiments. N.P., D.D.H. and S.E.F. performed in silico analysis. O.K. designed codon-pair optimization algorithm. H.T., D.I.F., R-F.S, A.A.K. Z.E.S. and C.K-S supervised the work. All authors reviewed and edited the manuscript. H.T., A.A.K. and C.K-S acquired funding for the project.

References

Alexaki A, Kames J, Holcomb DD *et al*. Codon and Codon-Pair Usage Tables (CoCoPUTs): facilitating genetic variation analyses and recombinant gene design. *J Mol Biol* 2019;**431**:2434–2441.

Ebrahimi SB, Samanta D. Engineering protein-based therapeutics through structural and chemical design. *Nat Commun* 2023;**14**:2411.

Esmail S, Manolson MF. Advances in understanding N-glycosylation structure, function, and regulation in health and disease. *Eur J Cell Biol* 2021;**100**:151186.

Gamble CE, Brule CE, Dean KM *et al*. Adjacent codons act in concert to modulate translation efficiency in yeast. *Cell* 2016;**166**:679–690.

Greenfield NJ. Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. *Nat Protoc* 2006;**1**:2527–2535.

Huang Y, Lin T, Lu L *et al*. Codon pair optimization (CPO): a software tool for synthetic gene design based on codon pair bias to improve the expression of recombinant proteins in Pichia pastoris. *Microb Cell Fact* 2021;**20**:209.

Hunt RC, Kimchi-Sarfaty C. When silence disrupts. *N Engl J Med* 2022;**387**:753–756.

Hunt R, Hettiarachchi G, Katneni U *et al*. A single synonymous variant (c.354G>A [p.P118P]) in *ADAMTS13* confers enhanced specific activity. *Int J Mol Sci* 2019;**20**:5734.

Jiang Y, Neti SS, Sitarik I *et al*. How synonymous mutations alter enzyme structure and function over long timescales. *Nat Chem* 2022;**15**:308–318.

Katneni UK, Alexaki A, Hunt RC *et al*. Structural, functional, and immunogenicity implications of F9 gene recoding. *Blood Adv* 2022;**6**:3932–3944.

Ling J, Su J, Ma Z *et al*. The WXXW motif in the TSR1 of ADAMTS13 is important for its secretion and proteolytic activity. *Thromb Res* 2013;**131**:529–534.

Liu Y, Yang Q, Zhao F. Synonymous but not silent: the codon usage code for gene expression and protein folding. *Annu Rev Biochem* 2021;**90**:375–401.

Ricketts LM, Dlugosz M, Luther KB *et al*. O-fucosylation is required for ADAMTS13 secretion. *J Biol Chem* 2007;**282**:17014–17023.

Sorvillo N, Kaijen PH, Matsumoto M *et al*. Identification of N-linked glycosylation and putative O-fucosylation, C-mannosylation sites in plasma derived ADAMTS13. *J Thromb Haemost* 2014;**12**:670–679.

Verbij FC, Stokhuijzen E, Kaijen PHP *et al*. Identification of glycans on plasma-derived ADAMTS13. *Blood* 2016;**128**:e51–e58.