Supplementary Material for

Engineering IgGs to aggregation resistance by two independent mechanisms: lessons from comparison of *Pichia pastoris* and mammalian cell expression

Jonas V. Schaefer and Andreas Plückthun^{*}

Department of Biochemistry, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

*Corresponding author. E-mail address: plueckthun@bioc.uzh.ch.

Charge engineering of Kex2 cleavage site

Comparing the results of the maturation of the various Pichia constructs, it became obvious that while the omission of the EAEA tetrapeptide did not cause any problems in the processing of the light chain (as seen in Fig. 3b bottom panel and from the sequencing results in Fig. 3c), it led to severe heterogeneity for the heavy chain H*. As the global fold of the variable domains of the light and the heavy chain including the arrangement and accessibility of the N-terminal stretch is very similar, the cause of this inefficient processing must lie elsewhere. The most obvious parameter is the primary sequence of the POI following the Kex2 cleavage site. Therefore, the sequences of the light and the heavy chain were compared with each other as well as with the "wrong" species derived from the alternative cleavage site (Fig. 3c). Comparing their aminoterminal residues, a clear difference in the charge status between the N-terminal amino-acids of the light (member of the human $V_{\lambda}3$ family) and heavy chains (a human $V_{H}6$ family member) were found (N-terminal sequences: QVQLV vs. DIELT). The natural human $V_{\lambda}3$ sequences generally start with the sequence SYELT, while most human kappa light chains have acidic residues at their N-termini.¹ However, the consensus sequence of the human $V_{\lambda}3$ family had been engineered in the HuCAL library and thus also in our case to start with DIELT for cloning reasons.² Interestingly, acidic amino acids are present in some V_H families as well at these positions, e.g. in V_H3 and $V_H5.^3$

The overview depicted in Supplementary Data Fig. S2a showed that both the natural Kex2 site, the native N-terminus of the light chain as well as the alternative site within the heavy chain's pro-region exhibit acidic residues directly following the cleavage site. This suggests that the Kex2 endopeptidase seems to require negative charged residues downstream of its cutting site for efficient cleavage. This preference for acidic amino acids was also found by Bader and colleagues in their analysis of various Kex2 cleavage sites.⁴ In a three-dimensional model of this enzyme, its S1' and S3' pockets are characterized by positive charges – thus allowing optimal charge-charge interactions with acidic amino acids in position P1' and P3'.

As the charge of the Kex2 substrates seems to be of great importance, the composition of the heavy chain was engineered, changing the non-charged glutamines (Q) in position P1' and P3' of the heavy chain to their charged counterparts glutamic acids (E), either individual or both at the same time. The resulting QE, EQ and EE variants were subsequently compared to the original QQ construct and analyzed for the distribution of the three species. None of these constructs led to a homogenously processed fraction but all contained molecules still having the full pro-region attached (Supplementary Data Fig. S2b). Nonetheless, the distribution of the three species compared to the QQ variant the fraction of the "mature" species for the EE construct was doubled, while the

percentage of the other two species (panels (c) and (d) of Supplementary Data Fig. S2) was consequently reduced.

Taken together, these data indicate that the cleavage efficiency of Kex2 is clearly influenced by the charge distribution at its substrate-binding site. Although engineering of the residues in the P1' and P3' position towards charged amino acids influenced the processing of the resulting IgG construct, it did not lead to homologous, correctly processed heavy chains. Therefore, Kex2 cleavage efficiency not only depends on the charge distribution upstream of its cleavage site, but also on other characteristics of the POI. However, we cannot yet pinpoint what these features are exactly.

- Kawasaki, K., Minoshima, S., Nakato, E., Shibuya, K., Shintani, A., Schmeits, J. L., Wang, J. & Shimizu, N. (1997). One-megabase sequence analysis of the human immunoglobulin lambda gene locus. *Genome Res.* 7, 250-261.
- Knappik, A., Ge, L., Honegger, A., Pack, P., Fischer, M., Wellnhofer, G., Hoess, A., Wölle, J., Plückthun, A. & Virnekäs, B. (2000). Fully synthetic human combinatorial antibody libraries (HuCAL) based on modular consensus frameworks and CDRs randomized with trinucleotides. *J. Mol. Biol.* **296**, 57-86.
- 3. Matsuda, F., Ishii, K., Bourvagnet, P., Kuma, K., Hayashida, H., Miyata, T. & Honjo, T. (1998). The complete nucleotide sequence of the human immunoglobulin heavy chain variable region locus. *J. Exp. Med.* **188**, 2151-2162.
- 4. Bader, O., Krauke, Y. & Hube, B. (2008). Processing of predicted substrates of fungal Kex2 proteinases from Candida albicans, C. glabrata, Saccharomyces cerevisiae and Pichia pastoris. *BMC Microbiol.* **8**, 116.

Fig. S1. Influence of glycosylation on IgG stability. (a) Comparison of the DSF signal of HEK-produced native IgG with that of its unglycosylated counterpart (T299A glycan knock-out). The transition of the C_{H2} unfolding is the dominant signal detected. **(b)** DSC signals recorded for the same constructs as in (a). **(c)** DSF analysis of *Pichia*-produced IgG constructs. **(d)** Comparison of DSF signals recorded for HEK- and *Pichia*-produced IgGs in their glycosylated and un-glycosylated state. **(e)** Thermal denaturation curves of the constructs seen in (d) measured by intrinsic tryptophan fluorescence. The curves were derived from the intensity ratio of the emission spectrum at 330 nm and 350 nm upon excitation at 295 nm, normalized to the plateaus of fully folded and fully denatured proteins. This value is plotted as a function of temperature.





(d)







Fig. S1

Fig. S2. Charge engineering of *Pichia*-produced IgG H* / L construct. (a) Comparison of the mature N-terminal amino acid sequences of different yeast heavy chain variants analyzed. The uppermost sequence is the one primarily found for the H* construct, derived from an alternative cleavage within the pro-region and carrying two negatively charged amino acids at its N-terminus (red color). To test charge influence on cleavage efficiency the residues at the first and third position of the mature heavy chain N-terminus were mutated from glutamine (Q) to glutamic acid (E) in individual combinations or combined in the EE variant. (b) Western blot analysis of IgG H* / L charge variants after incubation in the absence or presence of PNGase F. Shown are the blots detected with antibodies specific to the heavy chain (α -H) and to the α -factor pre-pro-region (α - α MFpp). (c) Comparative reducing SDS-PAGE analysis of yeast IgG H* / L QQ and EE variants after incubation of IgG H* / L forms for the QQ and EE variants, based on the relative signal intensities found in Edman sequencing.











Fig. S2

Fig. S3. Binding affinity of IgG obtained from mammalian expression. Direct interaction studies between purified IgGs and their fluorescently labeled cognate antigen myoglobin (monomer) analyzed using microscale thermophoresis technology. Normalized fluorescence [‰] is plotted against the IgG concentration titrated from 0.03 nM to 1,000 nM. The fitted K_D values lie within the experimental error range and indicate that the binding is fully retained for (a) IgG H / L, (b) IgG H-E / L and (c) IgG H-E / L-E.





Fig. S4. Biophysical characterization of *Pichia*-produced IgG constructs. (a) Thermal denaturation curves measured by intrinsic tryptophan fluorescence. The curves were derived from the intensity ratio of the emission spectrum at 330 nm and 350 nm upon excitation at 295 nm, normalized to the plateaus of fully folded and fully denatured proteins. This value is plotted as a function of temperature, comparing different *Pichia* IgG variants. (b) DSC of the *Pichia*-produced H / L, H-E / L and H-E / L-E constructs. Depicted as "H / L" construct is the H-pro-del variant. (c) Guanidine hydrochloride (GdnHCI)-induced denaturation of IgGs expressed in *Pichia pastoris*. The denaturation was followed by plotting the F_{330}/F_{350} ratio normalized to the plateaus of fully folded and fully denatured proteins.





(b)

Fig. S5. Aggregation behavior of different IgGs obtained from mammalian expression. (a)

Thermal denaturation curves. The denaturation was followed by CD, plotting the signals at 208 nm as a function of temperature. A comparison between different mammalian IgG variants is shown and the values are reported as MRE. (b) High tension (HT) voltage signal corresponding to the curves shown in (a) recorded by the photomultiplier connected to the CD spectrometer. The values were normalized by setting the initial and the highest values of the curves to 0 and 1, respectively. (c) Aggregation of IgG constructs measured by light scattering at 500 nm. Emission was recorded at the same wavelength as excitation. (d) Zoom at the aggregation temperatures of selected constructs. In addition to some of the constructs already shown in (c), also the signals derived from IgGs carrying the control peptide AGIQ on the heavy chain (H-A format) are presented.



Fig. S6. Aggregation behavior of different *Pichia*-produced IgGs. (a) Thermal denaturation curves measured by CD. The signals at 208 nm are plotted as a function of temperature. A comparison of different *Pichia*-produced IgG variants is shown and the values are reported as MRE. Depicted as "H / L" construct is the H-pro-del variant. (b) Photomultiplier high tension (HT) current signal corresponding to the curves shown in (a) recorded by the photomultiplier connected to the CD spectrometer. The values were normalized by setting the initial and the highest values of the curves as 0 and 1, respectively. (c) Aggregation of IgG constructs measured by light scattering at 500 nm.





(b)

Fig. S7. Isoelectric Focusing of HEK and *Pichia* **produced IgGs** Non-reduced IgGs were separated based on their isoelectric point (pl) after incubation in the absence or presence of PNGase F. A pl standard is shown in the left lane, with the corresponding pl values denoted next to it. Based on the primary sequences, the pl's were calculated to be 6.1 for the H-E / L-E and 6.7 for the H / L constructs, respectively. Upon deglycosylation, the asparagine to which the glycan was attached is converted to aspartic acid, resulting in an additional negative charge to the protein, which consequently focuses at a slightly lower pl value.

