**Supplementary Methods for Genome scale modeling of the protein secretory pathway reveals novel targets for improved recombinant protein production in yeast**

Feiran Li1, Yu Chen1, 6, Qi Qi1, 6, Yanyan Wang1, 6, Le Yuan1, 2, Mingtao Huang1,3, Ibrahim E. Elsemman1, 4, Amir Feizi1, \*, Eduard J Kerkhoven1, 2, Jens Nielsen1, 5, \*

1 Department of Biology and Biological Engineering, Chalmers University of Technology, Kemivägen 10, SE412 96 Gothenburg, Sweden

2 Novo Nordisk Foundation Center for Biosustainability, Chalmers University of Technology, Kemivägen 10, SE-412 96, Gothenburg, Sweden

3 Current affiliation: School of Food Science and Engineering, South China University of Technology, Guangzhou, 510641, China

4 Current affiliation: Department of Information Systems, Faculty of Computers and Information, Assiut University, Assiut, Egypt

5 BioInnovation Institute, Ole Måløes Vej 3, DK2200 Copenhagen N, Denmark

6 These authors contributed equally: Yu Chen, Qi Qi and Yanyan Wang.

\* Correspondence to: afeizi@gmail.com, nielsenj@chalmers.se

# Model Generation and Formulation

Table of Contents

[Model Generation and Formulation 2](#_Toc97061394)

[Summary 4](#_Toc97061395)

[Table S1.xlsx 4](#_Toc97061396)

[Protein\_information.xlsx 4](#_Toc97061397)

[Main MATLAB file: 4](#_Toc97061398)

[1. Information collection 5](#_Toc97061399)

[1.1 Protein sequence 5](#_Toc97061400)

[1.2 Protein information 5](#_Toc97061401)

[1.3 Protein stoichiometry 5](#_Toc97061402)

[2. Reformulation of metabolic model 5](#_Toc97061403)

[2.1 Metabolic model origin 5](#_Toc97061404)

[2.2 Adding reactions for production of post-translational modification precursors 5](#_Toc97061405)

[2.3 Splitting isoenzymes and reversible reactions 6](#_Toc97061406)

[3. Protein related process reconstruction 7](#_Toc97061407)

[3.1 Peptide translation 7](#_Toc97061408)

[3.2 Protein processing and complex formation for proteins that are not processed by secretory pathway 11](#_Toc97061409)

[3.3 Protein processing and complex formation for proteins that are processed by secretory pathway 12](#_Toc97061410)

[3.3.1 Protein translocation 12](#_Toc97061411)

[3.3.2 Disulfide bond formation 17](#_Toc97061412)

[3.3.3 GPI formation 17](#_Toc97061413)

[3.3.4 ER *O-glycosylation* 20](#_Toc97061414)

[3.3.5 ER *N*-glycosylation 21](#_Toc97061415)

[3.3.6 Misfolding and ERAD 22](#_Toc97061416)

[3.3.7 Protein and signal peptide degradation 28](#_Toc97061417)

[3.3.8 COPII transport 29](#_Toc97061418)

[3.3.9 Golgi *N*-glycosylation 33](#_Toc97061419)

[3.3.10 Golgi *O*-glycosylation 35](#_Toc97061420)

[3.3.11 Mature 36](#_Toc97061421)

[3.3.12 Sorting 36](#_Toc97061422)

[3.4 Enzyme complex formation 40](#_Toc97061423)

[3.5 Complex dilution 41](#_Toc97061424)

[4 Turnover rates in the model 41](#_Toc97061425)

[4.1 Turnover rates for metabolic complexes 41](#_Toc97061426)

[4.2 Turnover rates for secretory complexes 44](#_Toc97061427)

[4.3 Turnover rates for translation machinery complexes 44](#_Toc97061428)

[5. Constraints 46](#_Toc97061429)

[6. Comparison of pcSecYeast with other models. 48](#_Toc97061430)

# Summary

This manual describes how the data were collected, how the model was reconstructed, and how the optimization problem was generated. In the beginning, we write an index for the main Excel and MATLAB files used to generate and simulate the model. All excel files can be found in the Github repository [https://github.com/SysBioChalmers/pcSecYeast].

## Table S1.xlsx

contains all collected information.

* **Annotation & Annotation\_extra:** gene names and sequence information of all *S. cerevisiae* proteins.
* **Machinery:** proteins which makes ribosome, ribosome assembly complex and proteasome complex
* **Secretory:** proteins which makes secretory machinery complexes
* **Secretory\_ref:** references for how those secretory complexes are summarized
* **kdeg:** kdeg information collected from reference1,2
* **kcat\_info\_metabolic:** *k*cat values used in the metabolic part
* **kcat\_info\_sectory:** *k*cat values used in the secretory part
* **kcat\_info\_machinery:** *k*cat values used in the ribosome, ribosome assembly and proteasome

## Protein\_information.xlsx

Contains the Protein Specific Information Matrix (PSIM) data for all proteins in *S. cerevisiae*, including whether the protein is processed by secretory pathway, the existence of the signal peptide, number of disulfide bond sites, *N*-glycosylation sites, *O*-glycosylation sites, transmembrane domains, GPI sites, protein localization, protein amino acid sequence and signal peptide sequence. The PSIM data were collected from literature3 and UniPort database4.

## Main MATLAB file:

Run the function buildModel to get pcSecYeast model.

# 1. Information collection

## 1.1 Protein sequence

We downloaded protein sequence information from the UniProt database4, which is stored as in the Annotation&Annotation\_extra sheet in Table S1. This information is also stored as Protein\_sequence.mat in the GitHub repository.

## 1.2 Protein information

All protein information was collected from the literature3 and UniProt database4. For proteins that are annotated with multiple compartments, the first annotated compartment was used as its localization. For proteins that were denoted with signal peptide and the signal peptide length were not mentioned, we used the first 21 aa as its signal peptide.

## 1.3 Protein stoichiometry

For each functional protein, we should determine whether its functional unit is a monomer or oligomer. To do so, we collected protein stoichiometry information from the PDBe database (https://www.ebi.ac.uk/pdbe/) as well as the Complex Portal website (<https://www.ebi.ac.uk/complexportal/home>). This information is stored as Protein\_stoichiometry.mat in the GitHub repository.

# 2. Reformulation of metabolic model

## 2.1 Metabolic model origin

The latest GEM Yeast85 for *S. cerevisiae* was used as the basis for the model pcSecYeast. Yeast8.3.5 was downloaded from the GitHub repository: <https://github.com/SysBioChalmers/yeast-GEM>.

## 2.2 Adding reactions for production of post-translational modification precursors

Yeast8.3.5 was firstly curated by adding reactions to ensure the production of all precursors for the secretory pathway. 92 reactions were added into Yeast8.3.5, which contains glycan synthesis, GPI anchor synthesis and transport reactions to shuttle currency metabolites between cytosol and other compartments in the secretory pathway. As for the GPI anchor synthesis, we chose *1-phosphatidyl-1D-myo-inositol* (1-16:0, 2-18:1) as the starting phosphatidyl-inositol according to the literature report6. Other lipids attached to GPI anchor were added according to the reference7. Those reactions can be found in the Yeast8\_Modification.xlsx in the GitHub repository. Function modifyYeast8 was used to add these reactions into Yeast8.

## 2.3 Splitting isoenzymes and reversible reactions

The updated model was then reformulated by splitting reversible reactions into forward and reverse reactions. Besides that, reactions catalyzed by isozymes were also split into multiple identical reactions with various isozymes. This step was performed to facilitate later *k*cat match and enzyme constraining step. Function splitModel was used to perform this change.

# 3. Protein related process reconstruction



Fig. 1 Workflow for adding reactions for enzyme complexes in the pcSecYeast. \*\*\* repetitive process for all subunits in one enzyme complex.

## 3.1 Peptide translation

We formulated the translation process for all proteins in the model. The substrates of translation include charged amino acid tRNAs, while the products include uncharged tRNA and the translated peptide. Besides the charged tRNAs, the translation process requires energy. Three steps in protein translation were considered: translation initiation, translation elongation, and translation termination. The translation initiation process requires one ATP molecule for binding mRNA with initiation factors, one ATP molecule for every step in the scanning process, and two GTP molecules for initiation factors8. The translation elongation process has been studied intensely in yeast8. In summary, for each amino acid, two ATPs and two GTPs are required during the elongation. As for the energy cost for the translation termination, one ATP and one GTP are required for each peptide. Translation initiation, elongation, and termination are lumped into one reaction in the model to improve simulation efficiency. The energy cost is calculated as 2N+3 ATP and 2N+3 GTP for each peptide, where N is the number of amino acids. The energy molecules were included as substrates in the translation reactions. Note that we simplified the model by assuming that all proteins are translated in the cytoplasm. Since the charged and uncharged tRNA are already metabolites in the original Yeast8, we did not update the tRNA charging process. The protein translation is catalyzed by the ribosome. We compiled all ribosome subunits from the reference9. The ribosome is assembled by ribosome assembly factors9, which are also complied as a complex in the model.

Example of translation reaction for protein YAL012W.

|  |
| --- |
| reaction id: r\_YAL012W\_peptide\_translation  reaction equation: 1582 H2O[c] + 791 ATP[c] + 791 GTP[c] + 41 Ala-tRNA(Ala)[c] + 13 Arg-tRNA(Arg)[c] + 22 Asn-tRNA(Asn)[c] + 21 Asp-tRNA(Asp)[c] + 15 Gln-tRNA(Gln)[c] + 21 Glu-tRNA(Glu)[c] + 28 Gly-tRNA(Gly)[c] + 14 His-tRNA(His)[c] + 24 Ile-tRNA(Ile)[c] + 40 Leu-tRNA(Leu)[c] + 22 Lys-tRNA(Lys)[c] + 4 Met-tRNA(Met)[c] + 14 Phe-tRNA(Phe)[c] + 17 Pro-tRNA(Pro)[c] + 29 Ser-tRNA(Ser)[c] + 28 Thr-tRNA(Thr)[c] + 2 Trp-tRNA(Trp)[c] + 10 Tyr-tRNA(Tyr)[c] + 30 Val-tRNA(Val)[c] -> 1582 H+[c] + 1582 phosphate[c] + 791 ADP[c] + 791 GDP[c] + 41 tRNA(Ala)[c] + 13 tRNA(Arg)[c] + 22 tRNA(Asn)[c] + 21 tRNA(Asp)[c] + 15 tRNA(Gln)[c] + 21 tRNA(Glu)[c] + 28 tRNA(Gly)[c] + 14 tRNA(His)[c] + 24 tRNA(Ile)[c] + 40 tRNA(Leu)[c] + 22 tRNA(Lys)[c] + 4 tRNA(Met)[c] + 14 tRNA(Phe)[c] + 17 tRNA(Pro)[c] + 29 tRNA(Ser)[c] + 28 tRNA(Thr)[c] + 2 tRNA(Trp)[c] + 10 tRNA(Tyr)[c] + 30 tRNA(Val)[c] + YAL012W\_peptide[c]  catalyst: Mach\_Ribosome\_complex |

Synthesis of ribosome:

|  |
| --- |
| reaction id: Mach\_Ribosome\_complex\_formation  reaction equation: YGR214W\_folding[c] + YLR048W\_folding[c] + YLR441C\_folding[c] + YML063W\_folding[c] + YGL123W\_folding[c] + YNL178W\_folding[c] + YJR145C\_folding[c] + YHR203C\_folding[c] + YJR123W\_folding[c] + YPL090C\_folding[c] + YBR181C\_folding[c] + YOR096W\_folding[c] + YNL096C\_folding[c] + YBL072C\_folding[c] + YER102W\_folding[c] + YPL081W\_folding[c] + YBR189W\_folding[c] + YOR293W\_folding[c] + YMR230W\_folding[c] + YDR025W\_folding[c] + YBR048W\_folding[c] + YOR369C\_folding[c] + YDR064W\_folding[c] + YCR031C\_folding[c] + YJL191W\_folding[c] + YOL040C\_folding[c] + YMR143W\_folding[c] + YDL083C\_folding[c] + YML024W\_folding[c] + YDR447C\_folding[c] + YDR450W\_folding[c] + YML026C\_folding[c] + YOL121C\_folding[c] + YNL302C\_folding[c] + YHL015W\_folding[c] + YKR057W\_folding[c] + YJL136C\_folding[c] + YJL190C\_folding[c] + YLR367W\_folding[c] + YGR118W\_folding[c] + YPR132W\_folding[c] + YER074W\_folding[c] + YIL069C\_folding[c] + YGR027C\_folding[c] + YLR333C\_folding[c] + YGL189C\_folding[c] + YER131W\_folding[c] + YKL156W\_folding[c] + YHR021C\_folding[c] + YLR264W\_folding[c] + YLR388W\_folding[c] + YDL061C\_folding[c] + YLR287C\_A\_folding[c] + YOR182C\_folding[c] + YLR167W\_folding[n] + YMR116C\_folding[c] + YPL220W\_folding[c] + YGL135W\_folding[c] + YFR031C\_A\_folding[c] + YIL018W\_folding[c] + YOR063W\_folding[c] + YBR031W\_folding[c] + YDR012W\_folding[c] + YPL131W\_folding[c] + YML073C\_folding[c] + YLR448W\_folding[c] + YGL076C\_folding[c] + YPL198W\_folding[c] + YHL033C\_folding[c] + YLL045C\_folding[c] + YGL147C\_folding[c] + YLR075W\_folding[c] + YPR102C\_folding[c] + YGR085C\_folding[c] + YEL054C\_folding[c] + YDR418W\_folding[c] + YDL082W\_folding[c] + YMR142C\_folding[c] + YKL006W\_folding[c] + YHL001W\_folding[c] + YLR029C\_folding[c] + YMR121C\_folding[c] + YIL133C\_folding[c] + YNL069C\_folding[c] + YKL180W\_folding[c] + YJL177W\_folding[c] + YOL120C\_folding[c] + YNL301C\_folding[c] + YBR084C\_A\_folding[c] + YBL027W\_folding[c] + YBR191W\_folding[c] + YPL079W\_folding[c] + YLR061W\_folding[c] + YFL034C\_A\_folding[c] + YBL087C\_folding[c] + YER117W\_folding[c] + YGL031C\_folding[c] + YGR148C\_folding[c] + YOL127W\_folding[c] + YLR344W\_folding[c] + YGR034W\_folding[c] + YHR010W\_folding[c] + YDR471W\_folding[c] + YGL103W\_folding[c] + YFR032C\_A\_folding[c] + YGL030W\_folding[c] + YDL075W\_folding[c] + YLR406C\_folding[c] + YBL092W\_folding[c] + YPL143W\_folding[c] + YOR234C\_folding[c] + YER056C\_A\_folding[c] + YIL052C\_folding[c] + YDL191W\_folding[c] + YDL136W\_folding[c] + YMR194W\_folding[c] + YPL249C\_A\_folding[c] + YLR185W\_folding[c] + YDR500C\_folding[c] + YLR325C\_folding[c] + YJL189W\_folding[c] + YIL148W\_folding[n] + YKR094C\_folding[n] + YDL184C\_folding[c] + YDL133C\_A\_folding[c] + YNL162W\_folding[c] + YHR141C\_folding[c] + YPR043W\_folding[c] + YJR094W\_A\_folding[c] + YLR340W\_folding[c] + YDL081C\_folding[c] + YOL039W\_folding[c] + YDR382W\_folding[c] + YOR267C\_folding[c] + YNL067W\_folding[c] + YMR242C\_folding[c] + YOR312C\_folding[c] + YDL130W\_folding[c] -> Mach\_Ribosome\_complex[c]  catalyst: Mach\_Ribosome\_Assembly\_Factors\_complex |

Complex formation of ribosome assembly factors:

|  |
| --- |
| reaction id: Mach\_Ribosome\_Assembly\_Factors\_complex\_formation  reaction equation: YAL005C\_folding[c] + YBR247C\_folding[n] + YCL031C\_folding[n] + YCL059C\_folding[n] + YCR057C\_folding[n] + YEL026W\_folding[n] + YER082C\_folding[n] + YER127W\_folding[n] + YHR148W\_folding[n] + YHR196W\_folding[n] + YIL035C\_folding[c] + YKL143W\_folding[c] + YKR060W\_folding[n] + YDL014W\_folding[n] + YDL148C\_folding[n] + YDL153C\_folding[n] + YDL213C\_folding[n] + YDR324C\_folding[n] + YDR339C\_folding[n] + YDR449C\_folding[n] + YGL120C\_folding[n] + YGR090W\_folding[n] + YGR128C\_folding[n] + YJL010C\_folding[n] + YJL069C\_folding[n] + YJL109C\_folding[n] + YJR002W\_folding[n] + YLL011W\_folding[n] + YLR129W\_folding[n] + YLR186W\_folding[n] + YLR197W\_folding[n] + YLR222C\_folding[n] + YLR409C\_folding[n] + YLR430W\_folding[n] + YML093W\_folding[n] + YMR093W\_folding[n] + YMR128W\_folding[n] + YMR229C\_folding[n] + YMR290C\_folding[n] + YNL075W\_folding[n] + YOL010W\_folding[n] + YOR004W\_folding[m] + YOR078W\_folding[n] + YOR310C\_folding[n] + YPL126W\_folding[n] + YPL217C\_folding[c] + YPR137W\_folding[n] + YPR144C\_folding[n] + YAL025C\_folding[n] + YBR267W\_folding[c] + YCL054W\_folding[n] + YCR072C\_folding[n] + YER002W\_folding[n] + YER006W\_folding[n] + YER126C\_folding[n] + YHR052W\_folding[n] + YHR066W\_folding[n] + YHR088W\_folding[n] + YHR170W\_folding[c] + YIR012W\_folding[c] + YIR026C\_folding[n] + YKL009W\_folding[n] + YKL082C\_folding[n] + YKL172W\_folding[n] + YKL186C\_folding[n] + YKR081C\_folding[n] + YDR087C\_folding[n] + YDR101C\_folding[c] + YDR312W\_folding[n] + YGL029W\_folding[n] + YGL099W\_folding[c] + YGL111W\_folding[n] + YGL173C\_folding[c] + YGR103W\_folding[n] + YGR245C\_folding[n] + YJL050W\_folding[n] + YLR002C\_folding[n] + YLR009W\_folding[c] + YLR074C\_folding[c] + YLR106C\_folding[n] + YLR175W\_folding[n] + YLR387C\_folding[c] + YLR449W\_folding[n] + YML074C\_folding[n] + YMR049C\_folding[n] + YNL002C\_folding[n] + YNL061W\_folding[n] + YNL110C\_folding[c] + YNR053C\_folding[n] + YOL077C\_folding[n] + YOR272W\_folding[n] + YOR294W\_folding[n] + YPL093W\_folding[n] + YPL146C\_folding[n] + YPL169C\_folding[c] + YPL211W\_folding[c] + YPR016C\_folding[c] + YPR143W\_folding[n] + YBL004W\_folding[c] + YEL055C\_folding[n] + YHR065C\_folding[n] + YHR169W\_folding[n] + YIL091C\_folding[n] + YKL078W\_folding[n] + YKL099C\_folding[n] + YDR021W\_folding[n] + YDR365C\_folding[n] + YGL019W\_folding[c] + YGL171W\_folding[n] + YGR081C\_folding[n] + YGR159C\_folding[n] + YGR251W\_folding[n] + YGR280C\_folding[n] + YJL033W\_folding[n] + YLR068W\_folding[n] + YLR336C\_folding[n] + YNL224C\_folding[c] + YNR054C\_folding[n] + YOR039W\_folding[c] + YOR061W\_folding[c] + YOR287C\_folding[n] + YPR112C\_folding[n] + YBR142W\_folding[n] + YHR085W\_folding[n] + YHR197W\_folding[n] + YKL014C\_folding[n] + YKL021C\_folding[n] + YKR024C\_folding[n] + YFL002C\_folding[n] + YFR001W\_folding[n] + YDL031W\_folding[n] + YDR060W\_folding[n] + YDR412W\_folding[n] + YDR496C\_folding[n] + YGL078C\_folding[n] + YGL246C\_folding[n] + YGR276C\_folding[n] + YJL122W\_folding[c] + YJR041C\_folding[n] + YLL008W\_folding[n] + YLL024C\_folding[e] + YLL034C\_folding[n] + YLR022C\_folding[c] + YLR059C\_folding[m] + YLR221C\_folding[n] + YLR276C\_folding[n] + YLR397C\_folding[c] + YMR285C\_folding[c] + YNL112W\_folding[c] + YNL163C\_folding[c] + YNL182C\_folding[n] + YNL227C\_folding[n] + YNR038W\_folding[n] + YOL041C\_folding[n] + YOL144W\_folding[n] + YOR048C\_folding[n] + YOR206W\_folding[n] + YPL012W\_folding[c] + YPL043W\_folding[n] -> Mach\_Ribosome\_Assembly\_Factors\_complex[c]  catalyst: - |

Function addTranslationRxns is used to add translation reactions for proteins.

## 3.2 Protein processing and complex formation for proteins that are not processed by secretory pathway

Translated nascent peptides need to be folded to form functional complexes for performing specific functions. As for proteins that are not processed in the secretory pathway, we added a simplified process for folding, misfolding and complex formation, similar to what has been done in proteome-constrained models for *S. cerevisiae*10 and *Lactococcus lactis*11. The folding process in the model is assumed to be in the cytosol and complex formation is assumed in the localization of the complex. As for complexes with multiple subunits, the subunit stoichiometry information collected before was used in the complex formation reaction.

Example of protein processing and complex formation for YBR153W

|  |
| --- |
| reaction id: YBR153W\_folding\_c  reaction equation: YBR153W\_peptide[c] -> YBR153W\_folding[c]  catalyst: -  reaction id: YBR153W\_degradation\_misfolding\_c  reaction equation: YBR153W\_misfolding[c] -> YBR153W\_subunit[c]  catalyst: -  reaction id: YBR153W\_misfold\_c  reaction equation: YBR153W\_folding[c] -> YBR153W\_misfolding[c]  catalyst: -  reaction id: YBR153W\_dilution\_misfolding\_c  reaction equation: YBR153W\_misfolding[c] ->  catalyst: -  reaction id: r\_0015\_complex\_formation  reaction equation: 2 YBR153W\_folding[c] -> r\_0015\_complex[c]  catalyst: - |

## 3.3 Protein processing and complex formation for proteins that are processed by secretory pathway

As for proteins that are processed by the secretory pathway, folding and further processing are described in detail as follows.

### 3.3.1 Protein translocation

The nascent translated peptide is then translocated into the ER for further modification. There are three different pathways for protein translocation from the cytosol to the ER: co-translational translocation, post-translational translocation, and post-translational translocation-tail- targeting12,13.

Co-translational translocation in *S. cerevisiae* involves the interplay of the nascent peptide, the ribosome, the signal recognition particle (SRP), the signal recognition receptor (SR), and either the Ssh1 or the Sec61 translocon pore. SRP is a complex of six proteins (Srp14, Srp21, Srp54, Srp68, Srp72, and Sec65) and a 7S single RNA (SCR1)14–16. SRP interacts with the signal peptide of nascent peptides to form ribosome–nascent chain (RNC) complex, thus by the interaction of SPR and a signal receptor complex (SR), encoded by *SRP101* and *SRP102,* the RNC complex can attach to ER membrane. Finally, the RNC complex is transferred to the translocon, and then SRP and SR dissociate. Nascent peptides are then translocated into the ER. GTP bind to both SRP (via the Srp54 subunit) and the SR, which is critical for their interaction.

In the model, this process was divided into six template reactions. Only proteins with signal peptide were assumed to adopt this pathway in the model. If the signal peptide sequence has not been annotated in UniProt, we use the first 21 amino acids of the protein as the signal peptide. The degradation of signal peptide is described in further section together with protein degradation.

1. signal peptide recognition
2. ER receptor biding to peptide-SRPC
3. binding of peptide -SRPC-SRC to the translocator (SEC61C)
4. binding of peptide -SRPC-SRC to the translocator (SSH1C)
5. signal peptidase
6. export the signal peptide out of ER for degradation

Example of co-translation reaction of YDR367W

|  |
| --- |
| reaction id: YDR367W\_co\_translation\_TC\_sec\_SRPC\_complex  reaction equation: YDR367W\_peptide[c] -> YDR367W\_translocate\_1[c]  catalyst: sec\_SRPC\_complex  reaction id: YDR367W\_co\_translation\_TC\_sec\_SRC\_complex  reaction equation: YDR367W\_translocate\_1[c] -> YDR367W\_translocate\_2[c]  catalyst: sec\_SRC\_complex  reaction id: YDR367W\_co\_translation\_TC\_sec\_SEC61C\_complex  reaction equation: 2 H2O[c] + 2 GTP[c] + YDR367W\_translocate\_2[c] -> H+[c] + 2 phosphate[c] + 2 GDP[c] + YDR367W\_translocate\_3[c]  catalyst: sec\_SEC61C\_complex  reaction id: YDR367W\_co\_translation\_TC\_sec\_SSH1C\_complex  reaction equation: 2 H2O[c] + 2 GTP[c] + YDR367W\_translocate\_2[c] -> H+[c] + 2 phosphate[c] + 2 GDP[c] + YDR367W\_translocate\_3[c]  catalyst: sec\_SSH1C\_complex  reaction id: YDR367W\_co\_translation\_TC\_sec\_SPC\_complex  reaction equation: H2O[c] + YDR367W\_translocate\_3[c] -> YDR367W[er] + YDR367W\_sp[er]  catalyst: sec\_SPC\_complex  reaction id: YDR367W\_export\_sp\_to\_c  reaction equation: YDR367W\_sp[er] -> YDR367W\_sp[c]  catalyst: - |

Enzyme complex formation for enzymes used in the co-translational translocation process:

|  |
| --- |
| reaction id: sec\_SRPC\_complex\_formation  reaction equation: DL092W\_folding[n] + YKL122C\_folding[n] + YPR088C\_folding[c] + YML105C\_folding[c] + YPL243W\_folding[n] + YPL210C\_folding[er] -> sec\_SRPC\_complex[erm]  catalyst: -  reaction id: sec\_SRC\_complex\_formation  reaction equation: YDR292C\_folding[erm] + YKL154W\_folding[erm] -> sec\_SRC\_complex[erm]  catalyst: -  reaction id: sec\_SEC61C\_complex\_formation  reaction equation: YER087C\_B\_folding[erm] + YLR378C\_folding[erm] + YDR086C\_folding[erm] -> sec\_SEC61C\_complex[erm]  catalyst: -  reaction id: sec\_SSH1C\_complex\_formation  reaction equation: YDR086C\_folding[erm] + YER019C\_A\_folding[erm] + YBR283C\_folding[erm] -> sec\_SSH1C\_complex[erm]  catalyst: -  reaction id: sec\_SPC\_complex\_formation  reaction equation: YIR022W\_folding[erm] + YJR010C\_A\_folding[erm] + YML055W\_folding[erm] + YLR066W\_folding[erm] -> sec\_SPC\_complex[erm]  catalyst: - |

Post-translational translocation is equivalently important with co-translational translocation. Nascent peptides exit the ribosome with the help of RAC chaperones (Ssb1, Ssz1 and Zuo1)17. After that, nascent peptides remain in an unfolded or loosely folded state, bind to the cytosolic chaperones Ssa1 and Ydj1 to avoid aggregation, which are then released before the translocation initiate. The translocation is mediated by the SEC complex, which comprises Sec61, Sbh1, Sss1, Sec62, Sec63, Sec71, and Sec72. The chaperon Kar2 in the ER lumen is suggested to drive the nascent protein into the ER18.

In the model, this process was divided into four template reactions. Coefficient of ATP in step 4 was set as length/40, which is based on the assumption that an ATP molecule that bound to the chaperone Kar2, is hydrolysed to ADP for every 40 amino acids that pass through the translocon pore19.

1. exit the ribosome
2. bind to the cytosolic chaperone
3. translocation
4. pulling of nascent protein

Example of post-translation reaction of YDR453C

|  |
| --- |
| reaction id: YDR453C\_Post\_translation\_PSTA\_sec\_RAC\_complex  reaction equation: YDR453C\_peptide[c] -> YDR453C\_translocate\_1[c]  catalyst: sec\_RAC\_complex  reaction id: YDR453C\_Post\_translation\_PSTA\_sec\_Ssa1\_Ydj1\_Snl1\_complex  reaction equation: H2O[c] + ATP[c] + YDR453C\_translocate\_1[c] -> H+[c] + phosphate[c] + ADP[c] + YDR453C\_translocate\_2[c]  catalyst: sec\_Ssa1\_Ydj1\_Snl1\_complex  reaction id: YDR453C\_Post\_translation\_PSTA\_sec\_SEC61SEC63C\_complex  reaction equation: YDR453C\_translocate\_2[c] -> YDR453C\_translocate\_3[c]  catalyst: sec\_SEC61SEC63C\_complex  reaction id: YDR453C\_Post\_translation\_PSTA\_sec\_BIP\_NEFS\_complex  reaction equation: 5 H2O[c] + 5 ATP[c] + YDR453C\_translocate\_3[c] -> 5 H+[c] + 5 phosphate[c] + 5 ADP[c] + YDR453C[er]  catalyst: sec\_BIP\_NEFS\_complex |

Enzyme complex formation for enzymes used in the post-translational translocation process:

|  |
| --- |
| reaction id: sec\_RAC\_complex\_formation  reaction equation: YDL229W\_folding[c] + YGR285C\_folding[c] + YHR064C\_folding[c] -> sec\_RAC\_complex[c]  catalyst: -  reaction id: sec\_Ssa1\_Ydj1\_Snl1\_complex\_formation  reaction equation: YAL005C\_folding[c] + YNL064C\_folding[c] + YIL016W\_folding[erm] -> sec\_Ssa1\_Ydj1\_Snl1\_complex[erm]  catalyst: -  reaction id: sec\_SEC61SEC63C\_complex\_formation  reaction equation: YER087C\_B\_folding[erm] + YLR378C\_folding[erm] + YDR086C\_folding[erm] + YBR171W\_folding[erm] + YOR254C\_folding[erm] + YLR292C\_folding[c] + YPL094C\_folding[erm] -> sec\_SEC61SEC63C\_complex[erm]  catalyst: -  reaction id: sec\_BIP\_NEFS\_complex\_formation  reaction equation: YJL034W\_folding[er] + YOL031C\_folding[er] + YKL073W\_folding[er] -> sec\_BIP\_NEFS\_complex[er]  catalyst: - |

Post-translational translocation-tail targeting is a unique translocation process, especially for tail-anchored (TA) proteins13. This process is also termed as the GET pathway, which involves Sgt2, Get4, Get5 and Get3 proteins. This process initiates from loading the TA proteins from the ribosome by the complex composed of Sgt2, Get4 and Get5. Then, the complex binds to Get3, a cytosolic transmembrane domains (TMD) recognition complex. This Get3 complex delivers the protein to the ER receptor composted of Get1 and Get2.

This process was formulated into three template reactions in the model, proteins with GPI anchor adopt this pathway for translocation.

1. load the TA proteins
2. bind to Get3
3. bind to ER receptor

Example of post-translation reaction-tail-targeting of YGR216C

|  |
| --- |
| reaction id: YGR216C\_Post\_translation\_PSTB\_sec\_Sgt2\_Get4\_Get5\_complex  reaction equation: YGR216C\_peptide[c] -> YGR216C\_translocate\_1[c]  catalyst: sec\_Sgt2\_Get4\_Get5\_complex  reaction id: YGR216C\_Post\_translation\_PSTB\_sec\_Get3\_complex  reaction equation: H2O[c] + ATP[c] + YGR216C\_translocate\_1[c] -> H+[c] + phosphate[c] + ADP[c] + YGR216C\_translocate\_2[c]  catalyst: sec\_Get3\_complex  reaction id: YGR216C\_Post\_translation\_PSTB\_sec\_Get1\_Get2\_complex  reaction equation: YGR216C\_translocate\_2[c] -> YGR216C[er]  catalyst: sec\_Get1\_Get2\_complex |

Complex formation for enzymes in post-translation reaction-tail-targeting pathway

|  |
| --- |
| reaction id: sec\_Sgt2\_Get4\_Get5\_complex\_formation  reaction equation: YOR007C\_folding[c] + YOR164C\_folding[c] + YOL111C\_folding[c] -> sec\_Sgt2\_Get4\_Get5\_complex[c]  catalyst: -  reaction id: sec\_Get3\_complex\_formation  reaction equation: 2 YDL100C\_folding[c] -> sec\_Get3\_complex[c]  catalyst: -  reaction id: sec\_Get1\_Get2\_complex\_formation  reaction equation: 2 YGL020C\_folding[erm] + 2 YER083C\_folding[erm] -> sec\_Get1\_Get2\_complex[erm]  catalyst: - |

### 3.3.2 Disulfide bond formation

This process is required for proteins annotated with disulfide bonds. The nascent peptide is captured by the chaperone Kar2 to mediate folding. Sulfhydryl groups are then oxidized by protein disulfide isomerases (Pdi1). Reoxidation of Pdi1 is mediated by ER oxidoreductin (Ero1), which in turn transfers electrons to O2, thereby generating reactive oxygen species (ROS)12.

This step in the model was divided into two template reactions.

1. bind to the chaperone
2. disulfide bond formation

Example of disulfide bond formation of YCL035C

|  |
| --- |
| reaction id: YCL035C\_DSB\_sec\_BIP\_NEFS\_complex  reaction equation: 2.775 H2O[er] + 2.775 ATP[er] + YCL035C[er] -> 2.775 H+[er] + 2.775 phosphate[er] + 2.775 ADP[er] + YCL035C\_Kar2ATPcplx[er]  catalyst: sec\_BIP\_NEFS\_complex  reaction id: YCL035C\_DSB\_PDI\_II\_sec\_PDI1\_ERV2\_Ero1p\_complex  reaction equation: PDI-ox[er] + YCL035C\_Kar2ATPcplx[er] -> PDI[er] + YCL035C\_DSB[er]  catalyst: sec\_PDI1\_ERV2\_Ero1p\_complex |

Complex formation for enzymes used in dilsulfide bond related process

|  |
| --- |
| reaction id: sec\_BIP\_NEFS\_complex\_formation  reaction equation: YJL034W\_folding[er] + YOL031C\_folding[er] + YKL073W\_folding[er] -> sec\_BIP\_NEFS\_complex[er]  catalyst: -  reaction id: sec\_PDI1\_ERV2\_Ero1p\_complex\_formation  reaction equation: YCL043C\_folding[er] + YML130C\_folding[erm] + YPR037C\_folding[erm] -> sec\_PDI1\_ERV2\_Ero1p\_complex[er]  catalyst: sec\_PDI1\_ERV2\_Ero1p\_complex |

### 3.3.3 GPI formation

GPI formation process was formulated into several reactions according to reference6. Metabolic reactions for glycosylphosphatidylinositols (GPIs) synthesis were gap-filled in the Yeast8 modification step. Here, the first step is to transfer synthesized GPIs into proteins, followed by several remodelling of the sugar and lipid moieties in the GPI anchor. The acyl chain from the inositol is firstly removed by the Bst1, then the C18:1 fatty acid of the primary anchor is removed by a phospholipase A2 (Per1)20. Then, a C26:0 fatty acid is attached to the diacylglycerol moiety catalyzed by Gup1. For most GPI anchors, this modified diacylglycerol-based anchor is subsequently transformed into a ceramide-containing anchor catalyzed by Cwh4321. Ted1 further removes a phosphoethanolamine (PEtN) on the second mannose. Even though the GPI anchor is further modified in the Golgi, we do not include the Golgi modification part in the model as the enzyme responsible for this has not been identified. This could be easily added into the model in the future22.

In the model, the GPI formation was divided into six steps:

1. GPI transfer
2. removal of the acyl chain from the inositol
3. removal of the unsaturated acyl chain at the sn-2 position of diacylglycerol to form lyso-GPI
4. transfer C26 saturated acyl chain to the sn-2 position
5. change the lipid moiety to ceramide consisting of PHS with a hydroxy-C26 fatty acid
6. removes a PEtN on the second mannose
7. GPI transfer with recycled GPI anchor from misfolded protein

Example of GPI formation and transfer of YDR437W

|  |
| --- |
| reaction id: YDR437W\_GPIRI\_sec\_GPIR\_complex  reaction equation: 6-O-2-O-((2-aminoethyl)phosphoryl)-alpha-D-mannosyl-(1-2)-{alpha-D-mannosyl-2-O-((2-aminoethyl)phosphoryl)-(1-2)-alpha-D-mannosyl-(1-6)-2-O-((2-aminoethyl)phosphoryl)-alpha-D-mannosyl-(1-4)-alpha-D-glucosaminyl}-O-acyl-1-phosphatidyl-1D-myo-inositol[er] + YDR437W[er] -> H2O[er] + YDR437W\_GPI\_G1[er]  catalyst: sec\_GPIR\_complex  reaction id: YDR437W\_GPIRII\_sec\_Bst1p\_complex  reaction equation: YDR437W\_GPI\_G1[er] -> palmitate[erm] + YDR437W\_GPI\_G2[er]  catalyst: sec\_Bst1p\_complex  reaction id: YDR437W\_GPIRIII\_sec\_Per1p\_complex  reaction equation: H2O[er] + YDR437W\_GPI\_G2[er] -> oleate[er] + YDR437W\_GPI\_G3[er]  catalyst: sec\_Per1p\_complex  reaction id: YDR437W\_GPIRIV\_sec\_Gup1p\_complex  reaction equation: hexacosanoyl-CoA[er] + YDR437W\_GPI\_G3[er] -> H+[er] + coenzyme A[er] + YDR437W\_GPI\_G4[er] catalyst: sec\_Gup1p\_complex  reaction id: YDR437W\_GPIRV\_sec\_Cwh43p\_Gpi7p\_Mcd4p\_complex  reaction equation: ceramide-3 (C26)[er] + YDR437W\_GPI\_G4[er] -> diglyceride (1-26:0, 2-16:0)[er] + YDR437W\_GPI\_G5[er]  catalyst: sec\_Cwh43p\_Gpi7p\_Mcd4p\_complex  reaction id: YDR437W\_GPIRVI\_sec\_Ted1p\_complex  reaction equation: H2O[er] + YDR437W\_GPI\_G5[er] -> O-phosphoethanolamine[er] + YDR437W\_GPI\_G6[er]  catalyst: sec\_Ted1p\_complex  reaction id: YDR437W\_GPIRIB\_sec\_GPIR\_complex  reaction equation: 6-O-2-O-alpha-D-mannosyl-(1-2)-{alpha-D-mannosyl-2-O-((2-aminoethyl)phosphoryl)-(1-2)-alpha-D-mannosyl-(1-6)-2-O-((2-aminoethyl)phosphoryl)-alpha-D-mannosyl-(1-4)-alpha-D-glucosaminyl}-O-inositol-P-ceramide C (C26)[er] + YDR437W[er] -> H2O[er] + YDR437W\_GPI\_G6[er]  catalyst: sec\_GPIR\_complex |

Complex formation of GPI related enzymes

|  |
| --- |
| reaction id: sec\_GPIR\_complex\_formation  reaction equation: 2 YLR088W\_folding[erm] + 2 YDR331W\_folding[erm] + 2 YHR188C\_folding[erm] + 2 YDR434W\_folding[erm] + 2 YLR459W\_folding[erm] -> sec\_GPIR\_complex[erm]  catalyst: -  reaction id: sec\_Bst1p\_complex\_formation  reaction equation: YFL025C\_folding[erm] -> sec\_Bst1p\_complex[erm]  catalyst: -  reaction id: sec\_Per1p\_complex\_formation  reaction equation: YCR044C\_folding[erm] -> sec\_Per1p\_complex[erm]  catalyst: -  reaction id: sec\_Gup1p\_complex\_formation  reaction equation: YGL084C\_folding[erm] -> sec\_Gup1p\_complex[erm]  catalyst: -  reaction id: sec\_Cwh43p\_Gpi7p\_Mcd4p\_complex\_formation  reaction equation: YKL165C\_folding[erm] + YJL062W\_folding[erm] + YCR017C\_folding[erm] -> sec\_Cwh43p\_Gpi7p\_Mcd4p\_complex[erm]  catalyst: -  reaction id: sec\_Ted1p\_complex\_formation  reaction equation: YIL039W\_folding[erm] -> sec\_Ted1p\_complex[erm]  catalyst: -  reaction id: sec\_GPIR\_complex\_formation  reaction equation: 2 YLR088W\_folding[erm] + 2 YDR331W\_folding[erm] + 2 YHR188C\_folding[erm] + 2 YDR434W\_folding[erm] + 2 YLR459W\_folding[erm] -> sec\_GPIR\_complex[erm]  catalyst: - |

### 3.3.4 ER *O-glycosylation*

In yeast, *O*-glycosylation initiates in the ER and extends in Golgi. This section describes how the ER *O*-glycosylation was formulated in the model. The glycan extension in Golgi will be described in the following sections when the peptide is transported into Golgi. More than six proteins in ER are responsible for protein manosylation in yeast. Those proteins make up the complex PMTC, protein O-mannosyltransferase, which transfers mannose residues from dolichyl phosphate D-mannose to protein Ser/Thr residues23.

In the model, this process was formulated into one reaction:

1. ER *O*-glycosylation

Example of ER *O*-glycosylation of YJL137C

|  |
| --- |
| reaction id: YJL137C\_OG\_EROG\_sec\_Pmt2p\_Pmt5p\_Pmt1p\_Pmt6p\_Pmt4p\_Pmt3p\_complex  reaction equation: 3 dolichyl D-mannosyl phosphate[er] + YJL137C[er] -> 3 dolichyl phosphate[er] + YJL137C\_OG\_M1[er]  catalyst: sec\_Pmt2p\_Pmt5p\_Pmt1p\_Pmt6p\_Pmt4p\_Pmt3p\_complex |

Complex formation of ER *O*-glycosylation related enzymes

|  |
| --- |
| reaction id: sec\_Pmt2p\_Pmt5p\_Pmt1p\_Pmt6p\_Pmt4p\_Pmt3p\_complex\_formation  reaction equation: YAL023C\_folding[erm] + YDL095W\_folding[erm] + YDL093W\_folding[erm] + YJR143C\_folding[erm] + YOR321W\_folding[erm] + YGR199W\_folding[erm] -> sec\_Pmt2p\_Pmt5p\_Pmt1p\_Pmt6p\_Pmt4p\_Pmt3p\_complex[erm]  catalyst: - |

### 3.3.5 ER *N*-glycosylation

*N*-glycosylation is one of the most abundant post-translational modifications in yeast. The first step of *N*-glycosylation is to transfer the oligosaccharide precursor Glc3Man9GlcNAc2 to the protein, mediated by the OSTC complex24. After that, three glucose residues in the glycan are trimmed by Cwh4125 and Rot224. Then one of the mannose residues added by Alg9 from N-linked core oligosaccharides is further trimmed by Mns124, which is the last trimming reaction that occurs in the ER before correctly-folded proteins migrate to the Golgi. The further glycan extension in Golgi will be described in the following part when the peptide is transported into Golgi. The oligosaccharide precursor synthesis was added to the model in the previous Yeast8 modification step.

This process was divided into five steps in the model.

1. OSTC\_complex ER *N*-glycan transfer
2. ER Glycan trimming I
3. ER Glycan trimming II
4. ER Glycan trimming III
5. ER demanosylation I

Example: ER *N*-glycosylation for YJL139C

|  |
| --- |
| reaction id: YJL139C\_ERNG\_NG\_sec\_OSTC\_complex  reaction equation: 5 Glucose(3)Mannose(9)GlucoseNAc(2)-PP-dolichol[er] + YJL139C[er] -> 5 dolichyl phosphate[er] + YJL139C\_G3M9[er]  catalyst: sec\_OSTC\_complex  reaction id: YJL139C\_ERNG\_FLI\_NG\_sec\_Cwh41p\_complex  reaction equation: 5 H2O[er] + YJL139C\_G3M9[er] -> 5 D-glucose[er] + YJL139C\_G2M9[er]  catalyst: sec\_Cwh41p\_complex  reaction id: YJL139C\_ERNG\_FLII\_NG\_sec\_Rot2p\_complex  reaction equation: 5 H2O[er] + YJL139C\_G2M9[er] -> 5 D-glucose[er] + YJL139C\_G1M9[er]  catalyst: sec\_Rot2p\_complex  reaction id: YJL139C\_ERNG\_FLIII\_NG\_sec\_Rot2p\_complex  reaction equation: 5 H2O[er] + YJL139C\_G1M9[er] -> 5 D-glucose[er] + YJL139C\_M9[er]  catalyst: sec\_Rot2p\_complex  reaction id: YJL139C\_ERNG\_FLIV\_NG\_sec\_Mns1p\_complex  reaction equation: 5 H2O[er] + YJL139C\_M9[er] -> 5 D-mannose[er] + YJL139C\_M8[er]  catalyst: sec\_Mns1p\_complex |

Complex formation of enzymes used in ER *N*-glycosylation

|  |
| --- |
| reaction id: sec\_OSTC\_complex\_formation  reaction equation: YGL022W\_folding[er] + YJL002C\_folding[erm] + YEL002C\_folding[erm] + YOR085W\_folding[erm] + YML019W\_folding[erm] + YMR149W\_folding[erm] + YOR103C\_folding[er] + YGL226C\_A\_folding[er] + YDL232W\_folding[erm] -> sec\_OSTC\_complex[erm]  catalyst: -  reaction id: sec\_Cwh41p\_complex\_formation  reaction equation: YGL027C\_folding[erm] -> sec\_Cwh41p\_complex[erm]  catalyst: -  reaction id: sec\_Rot2p\_complex\_formation  reaction equation: YBR229C\_folding[er] -> sec\_Rot2p\_complex[er]  catalyst: -  reaction id: sec\_Mns1p\_complex\_formation  reaction equation: YJR131W\_folding[erm] -> sec\_Mns1p\_complex[erm]  catalyst: - |

### 3.3.6 Misfolding and ERAD

Protein misfolding is a common cellular process that can produce intrinsically harmful products. In order to reduce the risk, the cell developed a highly efficient system for protein quality control and endoplasmic reticulum-associated degradation (ERAD) for the degradation of misfolded proteins. The ubiquitin-proteasome system (UPS) is a pathway in the cell responsible for the degradation of proteins. The pathway can be split into two main processes: ubiquitination and degradation. The ubiquitination process requires three enzymes: E1, E2 and E326. The activating enzyme E1 activates a ubiquitin with one ATP molecule. The activated ubiquitin conjugates enzyme E2. The ubiquitin transfers to the ligase E3 from E2 and then to the misfolded protein when E3 binds to the misfolded protein targeted for degradation. The process repeats itself until the misfolded protein acquires a chain of ubiquitin at least four ubiquitin long. Then, the tagged misfolded protein can then be released from E3 and recognized by the proteasome. In yeast, there are mainly three ERAD pathways: ERAD-C, ERAD-L and ERAD-M, they differ in the E3 ligase part. Both ERAD-L and ERAD-M uses Hrd1 ubiquitin ligase complex (Hrd1, Hrd3 and Der1), but luminal factor Yos9 seems dispensable for ERAD-M. ERAD-C uses the Doa10 ubiquitin ligase complex27,28. It is also suggested that Doa10 is also the ubiquitin ligase for ERAD-M29. Thus, as for ERAD-M pathway, we added alternative pathways which can utilize either the Hrd1 or the Doa10 as the ubiquitin ligase in the model. All protein modifications such as disulfide bond, *N*-glycosylation and *O*-glycosylation are reversed in ERAD pathways, which are formulated as the first step in the ERAD pathways in the model. In order to represent the accumulation of misfolded protein, we added two extra reactions to reflect the occupation of misfolded protein with Kar2 and Pdi130.

The misfolding and ERAD is divided into reactions:

1. add Kar2 to the misfolded proteins
2. break the disulfide bond
3. trim one mannose off from the glycan
4. trim the GPI anchor
5. ERAD E3 ligase
6. ERAD ubiquitination
7. trim the glycan
8. misfolding degradation
9. misfolding accumulation with occupation of Pdi1
10. misfolding accumulation with occupation of Kar2

Example of misfolding and ERAD. Since different proteins adopt different reactions based on their protein properties, therefore several proteins are used in this example.

|  |
| --- |
| reaction id: YJL139C\_misfold\_ERAD\_sec\_Kar2p\_complex  reaction equation: 11 H2O[er] + 11 ATP[er] + YJL139C\_M9[er] -> 11 H+[er] + 11 phosphate[er] + 11 ADP[er] + YJL139C\_M9\_misf[er]  catalyst: sec\_Kar2p\_complex  reaction id: YJR104C\_ERAD2A\_sec\_Pdi1p\_complex  reaction equation: 4 glutathione[er] + YJR104C\_DSB\_misf[er] -> 4 H+[er] + 2 glutathione disulfide[er] + YJR104C\_DSB\_misf\_G1[er]  catalyst: sec\_Pdi1p\_complex  reaction id: YJL139C\_ERAD2B  reaction equation: YJL139C\_M9\_misf[er] -> YJL139C\_M9\_misf\_G1[er]  catalyst: -  reaction id: YJL139C\_ERAD3A\_sec\_Mns1p\_complex  reaction equation: 5 H2O[er] + YJL139C\_M9\_misf\_G1[er] -> 5 D-mannose[er] + YJL139C\_M9\_misf\_G2[er]  catalyst: sec\_Mns1p\_complex  reaction id: YJR104C\_ERAD3B  reaction equation: YJR104C\_DSB\_misf\_G1[er] -> YJR104C\_DSB\_misf\_G2[er]  catalyst: -  reaction id: YJL139C\_ERAD4A\_sec\_Mnl1p\_Pdi1p\_complex  reaction equation: 5 H2O[er] + YJL139C\_M9\_misf\_G2[er] -> 5 D-mannose[er] + YJL139C\_M9\_misf\_G3[er]  catalyst: sec\_Mnl1p\_Pdi1p\_complex  reaction id: YJR104C\_ERAD4B  reaction equation: YJR104C\_DSB\_misf\_G2[er] -> YJR104C\_DSB\_misf\_G3[er]  catalyst: -  reaction id: YKL165C\_ERAD5A  reaction equation: H2O[er] + YKL165C\_GPI\_G6\_M9\_misf\_G3[er] -> 6-O-2-O-alpha-D-mannosyl-(1-2)-{alpha-D-mannosyl-2-O-((2-aminoethyl)phosphoryl)-(1-2)-alpha-D-mannosyl-(1-6)-2-O-((2-aminoethyl)phosphoryl)-alpha-D-mannosyl-(1-4)-alpha-D-glucosaminyl}-O-inositol-P-ceramide C (C26)[er] + YKL165C\_GPI\_G6\_M9\_misf\_G4[er]  catalyst: -  reaction id: YJL139C\_ERAD5B  reaction equation: YJL139C\_M9\_misf\_G3[er] -> YJL139C\_M9\_misf\_G4[er]  catalyst: -  reaction id: YJL139C\_ERADL\_sec\_Cue1p\_Ubc6p\_Ubc7p\_Yos9p\_Hrd1p\_Hrd3p\_Der1p\_Usa1p\_complex  reaction equation: YJL139C\_M9\_misf\_G4[er] -> YJL139C\_M9\_misf\_G5[er]  catalyst: sec\_Cue1p\_Ubc6p\_Ubc7p\_Yos9p\_Hrd1p\_Hrd3p\_Der1p\_Usa1p\_complex  reaction id: YJL139C\_ERADL\_sec\_Sbh1p\_Sss1p\_Ssh1p\_Cdc48p\_Ubx2p\_Ufd1p\_Npl4p\_complex  reaction equation: 8 Ubiquitin\_for\_Transfer[c] + YJL139C\_M9\_misf\_G5[er] -> 8 Ubiquitin[c] + YJL139C\_M9\_misf\_G6[c]  catalyst: sec\_Sbh1p\_Sss1p\_Ssh1p\_Cdc48p\_Ubx2p\_Ufd1p\_Npl4p\_complex  reaction id: YKL165C\_ERADM\_sec\_Cue1p\_Ubc6p\_Ubc7p\_Hrd1p\_Hrd3p\_Der1p\_complex  reaction equation: YKL165C\_GPI\_G6\_M9\_misf\_G4[er] -> YKL165C\_GPI\_G6\_M9\_misf\_G5[er]  catalyst: sec\_Cue1p\_Ubc6p\_Ubc7p\_Hrd1p\_Hrd3p\_Der1p\_complex  reaction id: YKL165C\_ERADM2\_sec\_Cue1p\_Ubc6p\_Ubc7p\_Doa10p\_complex  reaction equation: YKL165C\_GPI\_G6\_M9\_misf\_G4[er] -> YKL165C\_GPI\_G6\_M9\_misf\_G5[er]  catalyst: sec\_Cue1p\_Ubc6p\_Ubc7p\_Doa10p\_complex  reaction id: YKL165C\_ERADM\_sec\_Sbh1p\_Sss1p\_Ssh1p\_Cdc48p\_Ubx2p\_Ufd1p\_Npl4p\_complex  reaction equation: 8 Ubiquitin\_for\_Transfer[c] + YKL165C\_GPI\_G6\_M9\_misf\_G5[er] -> 8 Ubiquitin[c] + YKL165C\_GPI\_G6\_M9\_misf\_G6[c]  catalyst: sec\_Sbh1p\_Sss1p\_Ssh1p\_Cdc48p\_Ubx2p\_Ufd1p\_Npl4p\_complex  reaction id: YNL038W\_ERADC\_sec\_Cue1p\_Ubc6p\_Ubc7p\_Doa10p\_complex  reaction equation:  YNL038W\_GPI\_G6\_misf\_G4[er] -> YNL038W\_GPI\_G6\_misf\_G5[er]  catalyst: sec\_Cue1p\_Ubc6p\_Ubc7p\_Doa10p\_complex  reaction id: YNL038W\_ERADC\_sec\_Sbh1p\_Sss1p\_Ssh1p\_Cdc48p\_Ubx2p\_Ufd1p\_Npl4p\_complex  reaction equation:  8 Ubiquitin\_for\_Transfer[c] + YNL038W\_GPI\_G6\_misf\_G5[er] -> 8 Ubiquitin[c] + YNL038W\_GPI\_G6\_misf\_G6[c]  catalyst: sec\_Sbh1p\_Sss1p\_Ssh1p\_Cdc48p\_Ubx2p\_Ufd1p\_Npl4p\_complex  reaction id: YJL139C\_ERAD7A\_sec\_Dsk2p\_Rad23p\_Png1p\_Uba1p\_complex  reaction equation: YJL139C\_M9\_misf\_G6[c] -> 10 N-acetyl-alpha-D-glucosamine 1-phosphate[c] + 35 D-mannose[er] + YJL139C\_misfolding[c]  catalyst: sec\_Dsk2p\_Rad23p\_Png1p\_Uba1p\_complex  reaction id: YKR058W\_ERAD7B\_sec\_Dsk2p\_Rad23p\_Png1p\_Uba1p\_complex  reaction equation: YKR058W\_OG\_M1\_misf\_G6[c] -> 2 D-mannose[er] + YKR058W\_misfolding[c]  catalyst: sec\_Dsk2p\_Rad23p\_Png1p\_Uba1p\_complex  reaction id: YNL038W\_ERAD7C\_sec\_Dsk2p\_Rad23p\_Uba1p\_complex  reaction equation: YNL038W\_GPI\_G6\_misf\_G6[c] -> YNL038W\_misfolding[c]  catalyst: sec\_Dsk2p\_Rad23p\_Uba1p\_complex  reaction id: YJL139C\_degradation\_misfolding\_c  reaction equation: YJL139C\_misfolding[c] -> YJL139C\_subunit[c]  catalyst: -  reaction id: YPL091W\_cycle\_accumulation\_sec\_pdi1p\_ero1p\_complex  reaction equation: 10 oxygen[er] + 20 glutathione[er] + YPL091W\_DSB\_misf[er] -> 10 glutathione disulfide[er] + 10 hydrogen peroxide[er] + YPL091W\_DSB\_misf2[er]  catalyst: sec\_pdi1p\_ero1p\_complex  reaction id: YPL091W\_cycle\_accumulation\_sec\_acc\_Kar2p\_complex  reaction equation: 120 H2O[er] + 120 ATP[er] + YPL091W\_DSB\_misf2[er] -> 120 H+[er] + 120 phosphate[er] + 120 ADP[er] + YPL091W\_DSB\_misfolding\_acc[er]  catalyst: sec\_acc\_Kar2p\_complex |

Complex formation for enzymes in ERAD pathways

|  |
| --- |
| reaction id: sec\_Kar2p\_complex\_formation  reaction equation: YJL034W\_folding[er] -> sec\_Kar2p\_complex[er]  catalyst: -  reaction id: sec\_Pdi1p\_complex\_formation  reaction equation: 2 YCL043C\_folding[er] -> sec\_Pdi1p\_complex[er]  catalyst: -  reaction id: sec\_Mns1p\_complex\_formation  reaction equation: YJR131W\_folding[erm] -> sec\_Mns1p\_complex[erm]  catalyst: -  reaction id: sec\_Mnl1p\_Pdi1p\_complex\_formation  reaction equation: 2 YCL043C\_folding[er] + YHR204W\_folding[er] -> sec\_Mnl1p\_Pdi1p\_complex[er] catalyst: -  reaction id: sec\_Cue1p\_Ubc6p\_Ubc7p\_Yos9p\_Hrd1p\_Hrd3p\_Der1p\_Usa1p\_complex\_formation  reaction equation: YMR264W\_folding[erm] + YMR022W\_folding[erm] + 2 YDR057W\_folding[erm] + YLR207W\_folding[erm] + 2 YOL013C\_folding[erm] + YER100W\_folding[erm] + YBR201W\_folding[erm] + YML029W\_folding[erm] -> sec\_Cue1p\_Ubc6p\_Ubc7p\_Yos9p\_Hrd1p\_Hrd3p\_Der1p\_Usa1p\_complex[erm]  catalyst: -  reaction id: sec\_Sbh1p\_Sss1p\_Ssh1p\_Cdc48p\_Ubx2p\_Ufd1p\_Npl4p\_complex\_formation  reaction equation: YER087C\_B\_folding[erm] + YDR086C\_folding[erm] + YBR283C\_folding[erm] + 6 YDL126C\_folding[er] + YGR048W\_folding[erm] + YBR170C\_folding[er] + YML013W\_folding[erm] -> sec\_Sbh1p\_Sss1p\_Ssh1p\_Cdc48p\_Ubx2p\_Ufd1p\_Npl4p\_complex[erm]  catalyst: -  reaction id: sec\_Cue1p\_Ubc6p\_Ubc7p\_Hrd1p\_Hrd3p\_Der1p\_complex\_formation  reaction equation: YMR264W\_folding[erm] + YMR022W\_folding[erm] + YLR207W\_folding[erm] + 2 YOL013C\_folding[erm] + YER100W\_folding[erm] + YBR201W\_folding[erm] -> sec\_Cue1p\_Ubc6p\_Ubc7p\_Hrd1p\_Hrd3p\_Der1p\_complex[erm]  catalyst: -  reaction id: sec\_Sbh1p\_Sss1p\_Ssh1p\_Cdc48p\_Ubx2p\_Ufd1p\_Npl4p\_complex\_formation  reaction equation: YER087C\_B\_folding[erm] + YDR086C\_folding[erm] + YBR283C\_folding[erm] + 6 YDL126C\_folding[er] + YGR048W\_folding[erm] + YBR170C\_folding[er] + YML013W\_folding[erm] -> sec\_Sbh1p\_Sss1p\_Ssh1p\_Cdc48p\_Ubx2p\_Ufd1p\_Npl4p\_complex[erm]  catalyst: -  reaction id: sec\_Cue1p\_Ubc6p\_Ubc7p\_Doa10p\_complex\_formation  reaction equation: YMR264W\_folding[erm] + YMR022W\_folding[erm] + YER100W\_folding[erm] + YIL030C\_folding[erm] -> sec\_Cue1p\_Ubc6p\_Ubc7p\_Doa10p\_complex[erm]  catalyst: -  reaction id: sec\_Sbh1p\_Sss1p\_Ssh1p\_Cdc48p\_Ubx2p\_Ufd1p\_Npl4p\_complex\_formation  reaction equation: YER087C\_B\_folding[erm] + YDR086C\_folding[erm] + YBR283C\_folding[erm] + 6 YDL126C\_folding[er] + YGR048W\_folding[erm] + YBR170C\_folding[er] + YML013W\_folding[erm] -> sec\_Sbh1p\_Sss1p\_Ssh1p\_Cdc48p\_Ubx2p\_Ufd1p\_Npl4p\_complex[erm]  catalyst: -  reaction id: sec\_Dsk2p\_Rad23p\_Png1p\_Uba1p\_complex\_formation  reaction equation: 4 YMR276W\_folding[n] + YEL037C\_folding[c] + YPL096W\_folding[c] + 2 YKL210W\_folding[c] -> sec\_Dsk2p\_Rad23p\_Png1p\_Uba1p\_complex[c]  catalyst: -  reaction id: sec\_Dsk2p\_Rad23p\_Png1p\_Uba1p\_complex\_formation  reaction equation: 4 YMR276W\_folding[n] + YEL037C\_folding[c] + YPL096W\_folding[c] + 2 YKL210W\_folding[c] -> sec\_Dsk2p\_Rad23p\_Png1p\_Uba1p\_complex[c]  reaction id: YNL038W\_ERAD7C\_sec\_Dsk2p\_Rad23p\_Uba1p\_complex  reaction equation: YNL038W\_GPI\_G6\_misf\_G4[er] -> YNL038W\_GPI\_G6\_misf\_G5[er]  catalyst: -  reaction id: sec\_pdi1p\_ero1p\_complex\_formation  reaction equation: 2 YCL043C\_folding[er] + YML130C\_folding[erm] -> sec\_pdi1p\_ero1p\_complex[er]  catalyst: sec\_pdi1p\_ero1p\_complex  reaction id: sec\_acc\_Kar2p\_complex\_formation  reaction equation: YJL034W\_folding[er] -> sec\_acc\_Kar2p\_complex[er]  catalyst: - |

### 3.3.7 Protein and signal peptide degradation

Translocated misfolded proteins are then degraded into amino acids by the cytosol proteosome. Similar reactions are also added for the cleaved signal peptide. The energy cost for the degradation estimated for eukaryotes is around 0.25-1.3 ATP/aa (the calculation is based on an average protein length of 467 aa)31. We adopted the highest value of 1.3 ATP/aa as the energetic cost for protein degradation and signal peptide. To be noted here, even though that pre and pro sequence in the leader sequence of recombinant protein degraded in ER and Golgi, respectively32. To simplify this process, full leader sequence is cleaved in ER for degradation in the model.

Example of degradation of YPL091W.

|  |
| --- |
| reaction id: r\_YPL091W\_subunit\_degradation  reaction equation: 627 H2O[c] + 627 ATP[c] + YPL091W\_subunit[c] -> 627 H+[c] + 627 phosphate[c] + 35 L-glutamate[c] + 10 L-methionine[c] + 33 L-alanine[c] + 11 L-glutamine[c] + 627 ADP[c] + 24 L-aspartate[c] + 43 L-glycine[c] + 18 L-arginine[c] + 28 L-asparagine[c] + 29 L-serine[c] + 5 L-cysteine[c] + 15 L-histidine[c] + 32 L-isoleucine[c] + 14 L-proline[c] + 26 L-threonine[c] + 4 L-tryptophan[c] + 19 L-tyrosine[c] + 34 L-leucine[c] + 42 L-lysine[c] + 17 L-phenylalanine[c] + 45 L-valine[c]  catalyst: Mach\_proteasome\_complex |

Example of degradation of signal peptide for YAL053W.

|  |
| --- |
| reaction id: r\_YAL053W\_SP\_degradation  reaction equation: 28 H2O[c] + 28 ATP[c] + YAL053W\_sp[c] -> 28 H+[c] + 28 phosphate[c] + L-methionine[c] + 2 L-alanine[c] + 28 ADP[c] + L-glycine[c] + L-arginine[c] + L-asparagine[c] + L-serine[c] + 3 L-cysteine[c] + L-isoleucine[c] + 3 L-threonine[c] + 4 L-leucine[c] + 3 L-phenylalanine[c] + L-valine[c]  catalyst: Mach\_proteasome\_complex |

Complex formation of proteosome complex.

|  |
| --- |
| reaction id: Mach\_proteasome\_complex\_formation  reaction equation: YMR276W\_folding[n] + YEL037C\_folding[c] + YJL001W\_folding[c] + YOR157C\_folding[c] + YER094C\_folding[c] + YER012W\_folding[c] + YPR103W\_folding[c] + YBL041W\_folding[c] + YFR050C\_folding[c] + YGL011C\_folding[c] + YML092C\_folding[c] + YGR135W\_folding[c] + YOL038W\_folding[c] + YGR253C\_folding[c] + YMR314W\_folding[c] + YOR362C\_folding[c] + YDL097C\_folding[c] + YDL147W\_folding[c] + YDR363W\_A\_folding[c] + YDR427W\_folding[c] + YER021W\_folding[c] + YFR004W\_folding[c] + YFR052W\_folding[c] + YOR261C\_folding[c] + YPR108W\_folding[c] + YKL145W\_folding[c] + YDL007W\_folding[c] + YDR394W\_folding[c] + YOR259C\_folding[c] + YOR117W\_folding[c] + YGL048C\_folding[c] + YHR027C\_folding[c] + YIL075C\_folding[c] + YHR200W\_folding[c] + YLR421C\_folding[c] + YER143W\_folding[c] + YFR010W\_folding[c] + YBR082C\_folding[c] + YFL007W\_folding[c] + YGL141W\_folding[c] + YKL010C\_folding[n] + YHL030W\_folding[c] -> Mach\_proteasome\_complex[c]  catalyst: - |

### 3.3.8 COPII transport

The ER is the first membrane compartment of the secretory pathway and is where secretory and membrane proteins traffic in the “vesicular” mode first occurs. The exit of correctly folded proteins from ER requires a distinct set of coat proteins and accessory factors. The first step is per-budding complex formation. Sar1, Sec23 and Sec24 are common proteins among different cargos and coat formation mechanisms. Selective export of soluble luminal cargo requires specific cargo receptors. Yeast Erv29 is required for efficiently packaging of the glycosylated alpha-factor pheromone precursor (gpaf) into COPII vesicles and for efficiently secretion of carboxypeptidase Y (CPY)33. GPI-anchored proteins require adaptors for cargo selection34. A yeast p24 family protein Emp24, which forms heterotrimeric complex with Erp2, Emp24, and Erv25, has been shown to work as an adaptor for efficiently transport of GPI-anchored cargo35. Then the budding complex recruits Sec13-Sec31 heterotetramer, providing the outer layer of the coat36. Although Sar1, Sec23, Sec24, Sec13 and Sec31 are necessary and sufficient for vesicle formation, additional factors such as Sec16 and Sed4 are also involved in this process37. Through interactions with other COPII proteins, Sec16 is thought to facilitate the assembly of the vesicle coat by stabilizing the pre-budding complex, while Sed4 may regulate the vesicle budding process by inhibiting the GTPase-activating protein (GAP) activity of Sec23.

After budding, vesicles have to move toward the target membrane proceeding through defined and progressive steps: tethering, docking, and fusion. Tethering is mediated by tethering factors such as Uso138 and Bug139 or multisubunit complexes such as TRAPPI complexes. COPII complex would tether to the cis-Golgi membrane through binding to Sec23 mediated by the TRAPPI complex, including Bet3, Bet5, Trs31, Trs23, Trs33, Trs20, and Trs8540. The process of vesicle docking is mediated by another class of proteins named Rabs, which belong to the superfamily of small Ras-like GTPases. Rabs regulate membrane trafficking through interaction with defined effectors. Ypt1 is a Rab protein required ER to Golgi transport and TRAPPI complex activates Ypt1. Fusion is mediated by SNAREs complex such as Bet1 and Bos141.

Here in the model, we divided this process into three steps:

1. COPII pre budding (1A lumen cargo 1B transmembrane cargo 1C GPI anchored cargo)
2. COPII formation
3. COPII tethering docking and fusion

And we formulated three different COPII routes for different types of proteins in the model:

|  |  |
| --- | --- |
| Protein information | Route |
| Transmembrane proteins | coat\_trans\_membrane |
| GPI anchored proteins | coat\_GPI |
| Other proteins | coat\_other |

Since COPII process has alternatives, we use different proteins to show the example.

Example of COPII\_GPI route for protein YBR004C.

|  |
| --- |
| reaction id: YBR004C\_COPII\_GPI\_ERGL1C\_sec\_Sec12p\_Sar1p\_Sec23p\_Sec24p\_Emp24p\_Erp1p\_Erp2p\_Erv25p\_Bos1p\_Bet1p\_complex  reaction equation: H2O[er] + GTP[er] + YBR004C\_GPI\_G6\_M8[er] -> H+[er] + phosphate[er] + GDP[er] + YBR004C\_GPI\_G6\_M8\_GPI\_G1\_COP[er] catalyst: sec\_Sec12p\_Sar1p\_Sec23p\_Sec24p\_Emp24p\_Erp1p\_Erp2p\_Erv25p\_Bos1p\_Bet1p\_complex  reaction id: YBR004C\_COPII\_ERGL\_sec\_Sec13p\_Sec31p\_Sec16p\_Sed4p\_Sec5p\_Sec17p\_complex  reaction equation: YBR004C\_GPI\_G6\_M8\_GPI\_G1\_COP[er] -> YBR004C\_GPI\_G6\_M8\_GPI\_G2\_COP[c]  catalyst: sec\_Sec13p\_Sec31p\_Sec16p\_Sed4p\_Sec5p\_Sec17p\_complex  reaction id: YBR004C\_COPII\_ERGL\_sec\_Ypt1p\_Uso1p\_bug1p\_Bet3p\_Bet5p\_Trs20p\_Trs23p\_Trs31p\_Trs33p\_complex  reaction equation: H2O[c] + GTP[c] + YBR004C\_GPI\_G6\_M8\_GPI\_G2\_COP[c] -> H+[c] + phosphate[c] + GDP[c] + YBR004C\_GPI\_G6\_M8[g]  catalyst: sec\_Ypt1p\_Uso1p\_bug1p\_Bet3p\_Bet5p\_Trs20p\_Trs23p\_Trs31p\_Trs33p\_complex |

Complex formation for enzymes involved in the COPII\_GPI

|  |
| --- |
| reaction id: sec\_Sec12p\_Sar1p\_Sec23p\_Sec24p\_Emp24p\_Erp1p\_Erp2p\_Erv25p\_Bos1p\_Bet1p\_complex\_formation  reaction equation: YNR026C\_folding[erm] + 3 YPL218W\_folding[erm] + 3 YPR181C\_folding[c] + 3 YIL109C\_folding[c] + YIL004C\_folding[erm] + YGL200C\_folding[erm] + YAR002C\_A\_folding[erm] + YAL007C\_folding[erm] + YML012W\_folding[erm] + YLR078C\_folding[erm] -> sec\_Sec12p\_Sar1p\_Sec23p\_Sec24p\_Emp24p\_Erp1p\_Erp2p\_Erv25p\_Bos1p\_Bet1p\_complex[erm]  catalyst: -  reaction id: sec\_Sec13p\_Sec31p\_Sec16p\_Sed4p\_Sec5p\_Sec17p\_complex\_formation  reaction equation: 2 YLR208W\_folding[c] + 2 YDL195W\_folding[c] + 2 YPL085W\_folding[erm] + YCR067C\_folding[erm] + YDR166C\_folding[c] + YBL050W\_folding[v] -> sec\_Sec13p\_Sec31p\_Sec16p\_Sed4p\_Sec5p\_Sec17p\_complex[erm]  catalyst: -  reaction id: sec\_Ypt1p\_Uso1p\_bug1p\_Bet3p\_Bet5p\_Trs20p\_Trs23p\_Trs31p\_Trs33p\_complex\_formation  reaction equation: YFL038C\_folding[gm] + 2 YKR068C\_folding[g] + YML077W\_folding[g] + YDR246W\_folding[g] + YDR472W\_folding[g] + YDL058W\_folding[c] + YDL099W\_folding[c] + YBR254C\_folding[g] + YOR115C\_folding[g] -> sec\_Ypt1p\_Uso1p\_bug1p\_Bet3p\_Bet5p\_Trs20p\_Trs23p\_Trs31p\_Trs33p\_complex[gm]  catalyst: - |

Example of coat\_trans\_membrane route for protein YBR021W.

|  |
| --- |
| reaction id: YBR021W\_COPII\_TransM\_ERGL1B\_sec\_Sec12p\_Sar1p\_Sec23p\_Sec24p\_Bet1p\_Bos1p\_complex  reaction equation: H2O[er] + GTP[er] + YBR021W[er] -> H+[er] + phosphate[er] + GDP[er] + YBR021W\_COP\_coated[er]  catalyst: sec\_Sec12p\_Sar1p\_Sec23p\_Sec24p\_Bet1p\_Bos1p\_complex  reaction id: YBR021W\_COPII\_ERGL\_sec\_Sec13p\_Sec31p\_Sec16p\_Sed4p\_Sec5p\_Sec17p\_complex  reaction equation: YBR021W\_COP\_coated[er] -> YBR021W\_COP\_coated[c]  catalyst: sec\_Sec13p\_Sec31p\_Sec16p\_Sed4p\_Sec5p\_Sec17p\_complex  reaction id: YBR021W\_COPII\_ERGL\_sec\_Ypt1p\_Uso1p\_bug1p\_Bet3p\_Bet5p\_Trs20p\_Trs23p\_Trs31p\_Trs33p\_complex  reaction equation: H2O[c] + GTP[c] + YBR021W\_COP\_coated[c] -> H+[c] + phosphate[c] + GDP[c] + YBR021W[g]  catalyst: sec\_Ypt1p\_Uso1p\_bug1p\_Bet3p\_Bet5p\_Trs20p\_Trs23p\_Trs31p\_Trs33p\_complex |

Complex formation for enzymes involved in the COPII\_TransM

|  |
| --- |
| reaction id: sec\_Sec12p\_Sar1p\_Sec23p\_Sec24p\_Bet1p\_Bos1p\_complex\_formation  reaction equation: YNR026C\_folding[erm] + YPL218W\_folding[erm] + YPR181C\_folding[c] + YIL109C\_folding[c] + YIL004C\_folding[erm] + YLR078C\_folding[erm] -> sec\_Sec12p\_Sar1p\_Sec23p\_Sec24p\_Bet1p\_Bos1p\_complex[erm]  catalyst: -  reaction id: sec\_Sec13p\_Sec31p\_Sec16p\_Sed4p\_Sec5p\_Sec17p\_complex\_formation  reaction equation: 2 YLR208W\_folding[c] + 2 YDL195W\_folding[c] + 2 YPL085W\_folding[erm] + YCR067C\_folding[erm] + YDR166C\_folding[c] + YBL050W\_folding[v] -> sec\_Sec13p\_Sec31p\_Sec16p\_Sed4p\_Sec5p\_Sec17p\_complex[erm]  catalyst: -  reaction id: sec\_Ypt1p\_Uso1p\_bug1p\_Bet3p\_Bet5p\_Trs20p\_Trs23p\_Trs31p\_Trs33p\_complex\_formation  reaction equation: YFL038C\_folding[gm] + 2 YKR068C\_folding[g] + YML077W\_folding[g] + YDR246W\_folding[g] + YDR472W\_folding[g] + YDL058W\_folding[c] + YDL099W\_folding[c] + YBR254C\_folding[g] + YOR115C\_folding[g] -> sec\_Ypt1p\_Uso1p\_bug1p\_Bet3p\_Bet5p\_Trs20p\_Trs23p\_Trs31p\_Trs33p\_complex[gm]  catalyst: - |

Example for the COPII\_other route for protein YBR092C.

|  |
| --- |
| reaction id: YBR092C\_COPII\_normal\_ERGL1A\_sec\_Sec12p\_Sar1p\_Sec23p\_Sec24p\_Erv29p\_Bet1p\_Bos1p\_complex  reaction equation: H2O[er] + GTP[er] + YBR092C\_M8[er] -> H+[er] + phosphate[er] + GDP[er] + YBR092C\_M8\_COP\_coated[er]  catalyst: sec\_Sec12p\_Sar1p\_Sec23p\_Sec24p\_Erv29p\_Bet1p\_Bos1p\_complex    reaction id: YBR092C\_COPII\_ERGL\_sec\_Sec13p\_Sec31p\_Sec16p\_Sed4p\_Sec5p\_Sec17p\_complex  reaction equation: YBR092C\_M8\_COP\_coated[er] -> YBR092C\_M8\_COP\_coated[c]  catalyst: sec\_Sec13p\_Sec31p\_Sec16p\_Sed4p\_Sec5p\_Sec17p\_complex    reaction id: YBR092C\_COPII\_ERGL\_sec\_Ypt1p\_Uso1p\_bug1p\_Bet3p\_Bet5p\_Trs20p\_Trs23p\_Trs31p\_Trs33p\_complex  reaction equation: H2O[c] + GTP[c] + YBR092C\_M8\_COP\_coated[c] -> H+[c] + phosphate[c] + GDP[c] + YBR092C\_M8[g]  catalyst: sec\_Ypt1p\_Uso1p\_bug1p\_Bet3p\_Bet5p\_Trs20p\_Trs23p\_Trs31p\_Trs33p\_complex |

Complex formation for enzymes involved in the COPII\_other

|  |
| --- |
| reaction id: sec\_Sec12p\_Sar1p\_Sec23p\_Sec24p\_Erv29p\_Bet1p\_Bos1p\_complex\_formation  reaction equation: YNR026C\_folding[erm] + 3 YPL218W\_folding[erm] + 3 YPR181C\_folding[c] + 3 YIL109C\_folding[c] + YIL004C\_folding[erm] + YLR078C\_folding[erm] + YGR284C\_folding[erm] -> sec\_Sec12p\_Sar1p\_Sec23p\_Sec24p\_Erv29p\_Bet1p\_Bos1p\_complex[erm]  catalyst: -  reaction id: sec\_Sec13p\_Sec31p\_Sec16p\_Sed4p\_Sec5p\_Sec17p\_complex\_formation  reaction equation: 2 YLR208W\_folding[c] + 2 YDL195W\_folding[c] + 2 YPL085W\_folding[erm] + YCR067C\_folding[erm] + YDR166C\_folding[c] + YBL050W\_folding[v] -> sec\_Sec13p\_Sec31p\_Sec16p\_Sed4p\_Sec5p\_Sec17p\_complex[erm]  catalyst: -  reaction id: sec\_Ypt1p\_Uso1p\_bug1p\_Bet3p\_Bet5p\_Trs20p\_Trs23p\_Trs31p\_Trs33p\_complex\_formation  reaction equation: YFL038C\_folding[gm] + 2 YKR068C\_folding[g] + YML077W\_folding[g] + YDR246W\_folding[g] + YDR472W\_folding[g] + YDL058W\_folding[c] + YDL099W\_folding[c] + YBR254C\_folding[g] + YOR115C\_folding[g] -> sec\_Ypt1p\_Uso1p\_bug1p\_Bet3p\_Bet5p\_Trs20p\_Trs23p\_Trs31p\_Trs33p\_complex[gm]  catalyst: - |

### 3.3.9 Golgi *N*-glycosylation

*N*-glycosylated proteins are further modified in Golgi. Upon their arrival, Och1 adds a mannose moiety to the core N-linked oligosaccharide42. After that, the Golgi *N*-glycosylation modification diverges either to form small core-type oligosaccharides or hyper mannan type43,44. Since this kind of data is not available for each protein, we adopted the hyper mannan type in the model, which can be easily modified in the future with available data. As for the hyper mannan type *N*-glycosylation pathway, a heterodimeric complex M-Pol I, consisting of one copy of Van1 and one copy of Mnn9, is the first enzyme that contributes to the polymerization of mannose in the Golgi45. M-Pol II contains five subunits Mnn9, Anp1, Mnn10, Mnn11, and Hoc1, which further elongates the polysaccharide mannan chain45. The α-1,6-mannose backbone is further modified with the addition of α-1,2-mannoses by Mnn2 and Mnn5, mannosylphosphate residues by Mnn4p and Mnn6p, and α-1,3-mannoses by Mnn146.

We divided this process into four steps in the model:

1. Golgi *N*-glycosylation I with Och1
2. Golgi *N*-glycosylation II with MPOI
3. Golgi *N*-glycosylation II with MPOII
4. Golgi *N*-glycosylation II with Mnn1p Mnn5p and Mnn2p

Example of Golgi *N*-glycosylation for protein YJL139C

|  |
| --- |
| reaction id: YJL139C\_GLNG\_Golgi\_N\_linked\_glycosylation\_I\_sec\_Och1p\_complex  reaction equation: 5 GDP-alpha-D-mannose[g] + YJL139C\_M8[g] -> 5 GDP[g] + YJL139C\_M8\_GNG\_G1[g]  catalyst: sec\_Och1p\_complex  reaction id: YJL139C\_GLNG\_Golgi\_N\_linked\_glycosylation\_II\_sec\_MPOLI\_complex  reaction equation: 45 GDP-alpha-D-mannose[g] + YJL139C\_M8\_GNG\_G1[g] -> 45 GDP[g] + YJL139C\_M8\_GNG\_G2[g]  catalyst: sec\_MPOLI\_complex  reaction id: YJL139C\_GLNG\_Golgi\_N\_linked\_glycosylation\_III\_sec\_MPoLII\_complex  reaction equation: 150 GDP-alpha-D-mannose[g] + YJL139C\_M8\_GNG\_G2[g] -> 150 GDP[g] + YJL139C\_M8\_GNG\_G3[g]  catalyst: sec\_MPoLII\_complex  reaction id: YJL139C\_GLNG\_Golgi\_N\_linked\_glycosylation\_II\_sec\_Mnn1p\_Mnn2p\_Mnn5p\_complex  reaction equation: 5 GDP-alpha-D-mannose[g] + YJL139C\_M8\_GNG\_G3[g] -> 5 GDP[g] + YJL139C\_M8\_GNG\_G4[g]  catalyst: sec\_Mnn1p\_Mnn2p\_Mnn5p\_complex |

Complex formation of Golgi *N*-glycosylation involved enzymes.

|  |
| --- |
| reaction id: sec\_Och1p\_complex\_formation  reaction equation: YGL038C\_folding[erm] -> sec\_Och1p\_complex[g]  catalyst: -  reaction id: sec\_MPOLI\_complex\_formation  reaction equation: YPL050C\_folding[erm] + YML115C\_folding[erm] -> sec\_MPOLI\_complex[gm]  catalyst: -  reaction id: sec\_MPoLII\_complex\_formation  reaction equation: YPL050C\_folding[erm] + YEL036C\_folding[erm] + YDR245W\_folding[erm] + YJL183W\_folding[g] + YJR075W\_folding[g] -> sec\_MPoLII\_complex[gm]  catalyst: -  reaction id: sec\_Mnn1p\_Mnn2p\_Mnn5p\_complex\_formation  reaction equation: YER001W\_folding[gm] + YBR015C\_folding[gm] + YJL186W\_folding[g] -> sec\_Mnn1p\_Mnn2p\_Mnn5p\_complex[gm]  catalyst: - |

### 3.3.10 Golgi *O*-glycosylation

In contrast to Golgi *N*-glycosylation, O-glycans are synthesized by the stepwise addition of monosaccharides to the first mannose residue added in ER. α-1, 2-mannosyltransferases, Ktr1, Ktr3 and Kre2/Mnt1, participate in the addition of the second mannose residue onto *O*-linked chains in the Golgi47. Kre2 has been known to be the primary enzyme responsible for adding the third mannose onto O-glycans47. Ktr1 and Ktr3 are also able to add mannose, although to a lesser extent than Kre2. The α-1,3- annosyltransferase Mnn1 attaches the fourth mannose residue in the linear chain of up to five mannose residues.

The process was divided into two steps in the model.

1. *O*-glycosylation mannose extension kre2\_ktr1\_ktr3
2. *O*-glycosylation mannose extension Mnn1

Example for Golgi *O*-glycosylation of YJL137C

|  |
| --- |
| reaction id: YJL137C\_GLOG\_Golgi\_O\_linked\_manosylation\_I\_sec\_Kre2p\_ktr1p\_ktr3p\_complex  reaction equation: 9 GDP-alpha-D-mannose[g] + YJL137C\_OG\_M1[g] -> 9 GDP[g] + YJL137C\_OG\_M1\_GOG\_G1[g]  catalyst: sec\_Kre2p\_ktr1p\_ktr3p\_complex  reaction id: YJL137C\_GLOG\_Golgi\_O\_linked\_manosylation\_II\_sec\_Mnn1p\_complex  reaction equation: 6 GDP-alpha-D-mannose[g] + YJL137C\_OG\_M1\_GOG\_G1[g] -> 6 GDP[g] + YJL137C\_OG\_M1\_GOG\_G2[g]  catalyst: sec\_Mnn1p\_complex |

Complex formation of Golgi *O*-glycosylation involved enzymes.

|  |
| --- |
| reaction id: sec\_Kre2p\_ktr1p\_ktr3p\_complex\_formation  reaction equation: YBR205W\_folding[c] + YDR483W\_folding[gm] + YOR099W\_folding[gm] -> sec\_Kre2p\_ktr1p\_ktr3p\_complex[gm]  catalyst: -  reaction id: sec\_Mnn1p\_complex\_formation  reaction equation: YER001W\_folding[gm] ->sec\_Mnn1p\_complex[gm]  catalyst: - |

### 3.3.11 Mature

We added a mature reaction in the model to indicate the end of the modification.

Example of mature reaction for YJL137C

|  |
| --- |
| reaction id: YJL137C\_Mature  reaction equation: YJL137C\_OG\_M1\_GOG\_G2[g] -> YJL137C\_OG\_M1\_GOG\_G2\_mature[g]  catalyst: - |

### 3.3.12 Sorting

Mature proteins are transported to their destinations by different vesicle transport.

As for ER or ER membrane proteins, those proteins are transported back to ER via COPI48. The COPI complex comprises an ADP-ribosylation factor, Arf1 and the coatomer (Cop1, Sec26, Sec27, Sec21, Ret2, Sec28, and Ret3). The COPI assembly is initiated by the interaction of Arf1 with the Golgi membrane. Arf1 activity is controlled by guanine nucleotide exchange factors (GEF) such as Gea1, Gea2 and Sec7. Once the COPI arrives ER, Arf1 is inactivated to release the cargo proteins. The uncoating and fusion are two separate steps in ER-Golgi vesicle transport, which were lumped in one reaction in the model for simplicity. By hydrolyzing the ARF1-GTP, the uncoating starts, and then the uncoated vesicle binds to the Golgi membrane by the t-SNAREs12.

ALP pathway is one of the known trafficking routes from Golgi to Vacuole. Many proteins involved as detecting (Vps1, Swa2), tethering (ClathrinC, Arf1) and docking (t-SNAREC) of the vesicles ALP pathway have been characterized, including Apm3, Apl6, Aps3 and Apl549.

The CPY pathway is the default route to the Vacuole from Golgi. A two-step process using AP complexes. The pathway is named after it was suggested to traffic carboxypeptidase Y to the vacuole. AP-1 complex vesicles can transfer proteins from the trans-Golgi to the early or late endosome. After this, the AP-3 complex vesicle moves proteins from the Golgi/endosome to Vacuole50.

As for proteins located in the cell membrane or extracellular matrix, there are two types of exocytotic vesicles from the trans-Golgi called light density secretory vesicles (LDSV) and heavy density secretory vesicles (HDSV) (upon density-based separation experiments). LDSV is known to carry constitutively expressed cell membrane proteins such as Bgl2, Pma1 and Gas1; and is believed to emerge from the trans-Golgi and transit directly to the cell membrane. HDSV package soluble, secreted proteins, such as acid phosphates (Pho11, Pho12, Pho5) and invertase (Suc2). Those proteins are usually under transcriptional regulation and induced under certain conditions51,52.

|  |  |
| --- | --- |
| Destination | Pathway |
| ER membrane | COPI |
| ER | COPI |
| Vacuole membrane | ALP |
| Vacuole | CPY |
| Cell membrane | LDSV |
| extracellular | HDSV |
| Other | General |

Example of the COPI for protein YJL196C

|  |
| --- |
| reaction id: YJL196C\_GLER\_COPI\_formation\_sec\_Arf1p\_Gea1p\_Gea2p\_Rer1p\_Erd2p\_Cop1p\_Sec26p\_Sec27p\_Sec21p\_Ret2p\_Sec28p\_Ret3p\_complex  reaction equation: 2 H2O[c] + 2 GTP[c] + YJL196C\_mature[g] -> 2 H+[c] + 2 phosphate[c] + 2 GDP[c] + YJL196C\_mature\_COPI\_G1[c]  catalyst: sec\_Arf1p\_Gea1p\_Gea2p\_Rer1p\_Erd2p\_Cop1p\_Sec26p\_Sec27p\_Sec21p\_Ret2p\_Sec28p\_Ret3p\_complex  reaction id: YJL196C\_GLER\_COPI\_uncoating\_and\_fission\_sec\_Rer1p\_Ret2p\_Cop1p\_Sec27p\_Sec21p\_Bet1p\_complex  reaction equation: YJL196C\_mature\_COPI\_G1[c] -> YJL196C\_mature[er]  catalyst: sec\_Rer1p\_Ret2p\_Cop1p\_Sec27p\_Sec21p\_Bet1p\_complex |

Complex formation of COPI involved enzymes.

|  |
| --- |
| reaction id: sec\_Arf1p\_Gea1p\_Gea2p\_Rer1p\_Erd2p\_Cop1p\_Sec26p\_Sec27p\_Sec21p\_Ret2p\_Sec28p\_Ret3p\_complex\_formation  reaction equation: YDL192W\_folding[g] + YJR031C\_folding[c] + YEL022W\_folding[c] + YDL145C\_folding[gm] + YGL137W\_folding[g] + YFR051C\_folding[g] + YIL076W\_folding[g] + YCL001W\_folding[gm] + YBL040C\_folding[erm] + YDR238C\_folding[g] + YNL287W\_folding[gm] + YPL010W\_folding[g] -> sec\_Arf1p\_Gea1p\_Gea2p\_Rer1p\_Erd2p\_Cop1p\_Sec26p\_Sec27p\_Sec21p\_Ret2p\_Sec28p\_Ret3p\_complex[gm]  catalyst: -  reaction id: sec\_Rer1p\_Ret2p\_Cop1p\_Sec27p\_Sec21p\_Bet1p\_complex\_formation  reaction equation: YIL004C\_folding[erm] + YDL145C\_folding[gm] + YGL137W\_folding[g] + 4 YFR051C\_folding[g] + YCL001W\_folding[gm] + YNL287W\_folding[gm] -> sec\_Rer1p\_Ret2p\_Cop1p\_Sec27p\_Sec21p\_Bet1p\_complex[gm]  catalyst: - |

Example of the ALP pathway for vacuole membrane protein YJR001W

|  |
| --- |
| reaction id: YJR001W\_ALPtransport\_sec\_Apl6p\_Aps3p\_Apm3p\_Apl5p\_Vam3p\_Clc1p\_Chc1p\_Arf1p\_Swa2p\_Vps1p\_complex  reaction equation: 4 H2O[c] + 4 GTP[c] + YJR001W\_mature[g] -> 4 H+[c] + 4 phosphate[c] + 4 GDP[c] + YJR001W\_folding[vm]  catalyst: sec\_Apl6p\_Aps3p\_Apm3p\_Apl5p\_Vam3p\_Clc1p\_Chc1p\_Arf1p\_Swa2p\_Vps1p\_complex |

Complex formation of ALP involved enzymes.

|  |
| --- |
| reaction id: sec\_Apl6p\_Aps3p\_Apm3p\_Apl5p\_Vam3p\_Clc1p\_Chc1p\_Arf1p\_Swa2p\_Vps1p\_complex\_formation  reaction equation: YDL192W\_folding[g] + YOR106W\_folding[vm] + YDR320C\_folding[c] + YGR261C\_folding[c] + YJL024C\_folding[c] + YBR288C\_folding[c] + YPL195W\_folding[c] + YGR167W\_folding[c] + YGL206C\_folding[c] + YKR001C\_folding[c] -> sec\_Apl6p\_Aps3p\_Apm3p\_Apl5p\_Vam3p\_Clc1p\_Chc1p\_Arf1p\_Swa2p\_Vps1p\_complex[c]  catalyst: - |

Example of the CPY pathway for protein YKL103C

|  |
| --- |
| reaction id: YKL103C\_CPYI\_sec\_Gga1p\_Gga2p\_Arf1p\_Apl4p\_Apl2p\_Apm1p\_Aps1p\_Chc1p\_Clc1p\_Pep12p\_Vps45p\_Vps5p\_Swa2p\_complex  reaction equation: 4 H2O[c] + 4 GTP[c] + YKL103C\_M8\_GNG\_G4\_mature[g] -> 4 H+[c] + 4 phosphate[c] + 4 GDP[c] + YKL103C\_M8\_GNG\_G4\_mature\_CPY\_G1[v]  catalyst: sec\_Gga1p\_Gga2p\_Arf1p\_Apl4p\_Apl2p\_Apm1p\_Aps1p\_Chc1p\_Clc1p\_Pep12p\_Vps45p\_Vps5p\_Swa2p\_complex  reaction id: YKL103C\_CPYII\_sec\_Vps4p\_Vps27p\_Apl6p\_Aps3p\_Apm3p\_Apl5p\_Vam3p\_complex  reaction equation: H2O[c] + ATP[c] + YKL103C\_M8\_GNG\_G4\_mature\_CPY\_G1[v] -> H+[c] + phosphate[c] + ADP[c] + YKL103C\_folding[v]  catalyst: sec\_Vps4p\_Vps27p\_Apl6p\_Aps3p\_Apm3p\_Apl5p\_Vam3p\_complex |

Complex formation of CPY involved enzymes.

|  |
| --- |
| reaction id: sec\_Gga1p\_Gga2p\_Arf1p\_Apl4p\_Apl2p\_Apm1p\_Aps1p\_Chc1p\_Clc1p\_Pep12p\_Vps45p\_Vps5p\_Swa2p\_complex\_formation  reaction equation: YDL192W\_folding[g] + YDR320C\_folding[c] + YGR167W\_folding[c] + YGL206C\_folding[c] + YDR358W\_folding[g] + YHR108W\_folding[g] + YPR029C\_folding[g] + YKL135C\_folding[ce] + YPL259C\_folding[c] + YLR170C\_folding[c] + YOR036W\_folding[v] + YGL095C\_folding[c] + YOR069W\_folding[gm] -> sec\_Gga1p\_Gga2p\_Arf1p\_Apl4p\_Apl2p\_Apm1p\_Aps1p\_Chc1p\_Clc1p\_Pep12p\_Vps45p\_Vps5p\_Swa2p\_complex[c]  catalyst: -  reaction id: sec\_Vps4p\_Vps27p\_Apl6p\_Aps3p\_Apm3p\_Apl5p\_Vam3p\_complex\_formation  reaction equation: YOR106W\_folding[vm] + YGR261C\_folding[c] + YJL024C\_folding[c] + YBR288C\_folding[c] + YPL195W\_folding[c] + YPR173C\_folding[erm] + YNR006W\_folding[erm] -> sec\_Vps4p\_Vps27p\_Apl6p\_Aps3p\_Apm3p\_Apl5p\_Vam3p\_complex[c]  catalyst: - |

Example of the LDSV for protein YKL217W

|  |
| --- |
| reaction id: YKL217W\_LDSV\_sec\_Arf1p\_Sec3p\_Sec5p\_Sec6p\_Sec8p\_Sec10p\_Sec15p\_Exo70p\_Exo84p\_Sec4p\_Chc1p\_Clc1p\_complex  reaction equation: H2O[c] + GTP[c] + YKL217W\_mature[g] -> H+[c] + phosphate[c] + GDP[c] + YKL217W\_folding[ce]  catalyst: sec\_Arf1p\_Sec3p\_Sec5p\_Sec6p\_Sec8p\_Sec10p\_Sec15p\_Exo70p\_Exo84p\_Sec4p\_Chc1p\_Clc1p\_complex |

Complex formation of LDSV involved enzymes.

|  |
| --- |
| reaction id: sec\_Arf1p\_Sec3p\_Sec5p\_Sec6p\_Sec8p\_Sec10p\_Sec15p\_Exo70p\_Exo84p\_Sec4p\_Chc1p\_Clc1p\_complex\_formation  reaction equation: YDR166C\_folding[c] + YDL192W\_folding[g] + YGR167W\_folding[c] + YGL206C\_folding[c] + YER008C\_folding[c] + YIL068C\_folding[c] + YPR055W\_folding[c] + YLR166C\_folding[c] + YGL233W\_folding[c] + YJL085W\_folding[c] + YBR102C\_folding[c] + YFL005W\_folding[c] -> sec\_Arf1p\_Sec3p\_Sec5p\_Sec6p\_Sec8p\_Sec10p\_Sec15p\_Exo70p\_Exo84p\_Sec4p\_Chc1p\_Clc1p\_complex[c]  catalyst: - |

Example of HDSV for protein YLR155C

|  |
| --- |
| reaction id: YLR155C\_HDSVI\_sec\_Arf1p\_Pep12p\_Swa2p\_Chc1p\_Clc1p\_Apl4p\_Apl2p\_Apm1p\_Aps1p\_complex  reaction equation: H2O[c] + GTP[c] + YLR155C\_M8\_GNG\_G4\_mature[g] -> H+[c] + phosphate[c] + GDP[c] + YLR155C\_M8\_GNG\_G4\_mature[ce]  catalyst: sec\_Arf1p\_Pep12p\_Swa2p\_Chc1p\_Clc1p\_Apl4p\_Apl2p\_Apm1p\_Aps1p\_complex  reaction id: YLR155C\_HDSVII\_sec\_Vps1p\_Chc1p\_Clc1p\_complex  reaction equation: H2O[c] + GTP[c] + YLR155C\_M8\_GNG\_G4\_mature[ce] -> H+[c] + phosphate[c] + GDP[c] + YLR155C\_folding[e]  catalyst: sec\_Vps1p\_Chc1p\_Clc1p\_complex |

Complex formation of HDSV involved enzymes.

|  |
| --- |
| reaction id: sec\_Arf1p\_Pep12p\_Swa2p\_Chc1p\_Clc1p\_Apl4p\_Apl2p\_Apm1p\_Aps1p\_complex\_formation  reaction equation: YDL192W\_folding[g] + YDR320C\_folding[c] + YGR167W\_folding[c] + YGL206C\_folding[c] + YPR029C\_folding[g] + YKL135C\_folding[ce] + YPL259C\_folding[c] + YLR170C\_folding[c] + YOR036W\_folding[v] -> sec\_Arf1p\_Pep12p\_Swa2p\_Chc1p\_Clc1p\_Apl4p\_Apl2p\_Apm1p\_Aps1p\_complex[c]  catalyst: -  reaction id: sec\_Vps1p\_Chc1p\_Clc1p\_complex\_formation  reaction equation: YGR167W\_folding[c] + YGL206C\_folding[c] + YKR001C\_folding[c] -> sec\_Vps1p\_Chc1p\_Clc1p\_complex[c]  catalyst: - |

Example of general pathway for protein YLR240W

|  |
| --- |
| reaction id: YLR240W\_transportFromGolgiToOthercompartment  reaction equation: YLR240W\_mature\_transportFromGolgiToOthercompartment YLR240W\_mature[g] -> YLR240W\_folding[g]  catalyst: - |

## 3.4 Enzyme complex formation

In the model, we formulated reactions to form enzyme complexes, which serve as catalysts for reactions in metabolism and protein modification processes. If this enzyme contains multiple subunits, then the stichometry for the subunits were included in the equation. Stoichiometries for the subunits are from the PDB database as stated in the previous data collection part53.

Example for enzyme complex formation for metabolic reaction r\_2141, the complex contains six copies per subunit, thus the stichometry for each subunit is six.

|  |
| --- |
| reaction id: r\_2141\_complex\_formation  reaction equation: 6 YKL182W\_folding[c] + 6 YPL231W\_folding[c] -> r\_2141\_complex[c]  catalyst: - |

## 3.5 Complex dilution

The complex dilution reactions in the model are to represent cell division process. Diluted complexes can represent partial protein content in the biomass. Dilution reactions were added for all complexes rather for separate subunits.

Example for complex dilution of r\_2141\_complex

|  |
| --- |
| reaction id: r\_2141\_complex\_dilution  reaction equation: r\_2141\_complex[c] ->  catalyst: - |

# 4 Turnover rates in the model

## 4.1 Turnover rates for metabolic complexes

The *k*cat values for metabolic reactions were acquired from the BRENDA database by matching the EC numbers. The *k*cat extraction process used the criteria as follows53. The *k*cat values for all organisms were downloaded from BRENDA, with only wildtype enzymes, and only the maximal values for the multiple measurements are kept. The assignment of *k*cat relies on the matching EC number of the subunit and substrate to the dataset. The following criteria was used to determine the *k*cat:

a) *k*cat values for both substrate and organism matched in the dataset were prioritized.

b) if there is not such fully matched data, then the median of all *k*cat values within the organism from the same EC number was used.

c) if only substrate was matched but not organism, then the median of all *k*cat values with matched substrate from the same EC number was used.

d) if neither organism nor substrate was matched, then the median of all available values within the EC number was assigned.

e) if no *k*cat value was available for the EC number, then the *k*cat value was set with the median of all assigned values.

Note that the *k*cat value should be adjusted based on the protein stoichiometry information, e.g., the *k*cat value should multiply 2 for a dimer enzyme, and the median *k*cat value among subunits was selected for a complex when its subunits had various *k*cat values. We also manually collect several *k*cat values, which are available in the GitHub repository: <https://github.com/SysBioChalmers/pcSecYeast/tree/main/ComplementaryData/manual_update.xlsx>. Besides all these steps, for enzymes with available *in vivo* *k*cat values (*k*max), we updated the *k*cat values in the model to *in vivo* *k*max since it was demonstrated that utilization of *in vivo* *k*max could improve the model prediction54. Functions: collectkcats, updatekcats and matchkappToKcat were used to collect this information and perform changes. Check the corresponding lines in the main function buildModel for detailed information. Fig. 2 describes the kinetic parameter search pipeline.



Fig. 2 Process of *k*cat match for metabolic complexes. \* For reactions does not have enzyme associated, we keep the original reaction in the model and do not add kinetic constraints for those reactions. \*\* Complex formation related process can be found in Fig. 1.

## 4.2 Turnover rates for secretory complexes

To get *k*cat parameters for secretory machinery enzymes, proteome abundance data in PaxDb database55 combining collected PSIM were used to deduct apparent kinetic parameters. The calculation method is based on the machinery abundance and total abundance of proteins processed by the machinery (Fig. 3). Based on the equation, we can calculate the apparent *k*cat values for secretory machineries. Protein stoichiometry information is also considered in the calculation. Function: SimulateSecParam was used to get the *k*cat parameters for the secretory machinery complexes.



Fig. 3 Calculation of secretory machinery *k*cat parameters.

## 4.3 Turnover rates for translation machinery complexes

Ribosome synthesis rate was collected from the literature, which is 2000/min56. Given that the ATP turnover for the proteasome is 110/min (bionumber: 109936)57, and the ATP cost for eukaryote protein degradation is around 100-200 ATP( bionumber: 112155)31, and we therefore assumed the proteasome catalytic rate to be 0.5-1 protein/min.

The calculation of ribosomal catalytic rate follows the same method done for *Escherichia coli*58 and *Lactococcus lactis*11. We assumed the equation of ribosomal catalytic rate follow the Michaelis-Menten-type. We collected the mRNA content, protein content and specific growth rates (Fig. 4)59–61.



Fig. 4 Ribosome/protein ratio correlates with growth rates.

This study62 showed that there was a linear correlation between the specific growth rate and RNA/protein ratio in *S. cerevisiae.* From the Fig. 4, we also found clear correlation with the slope being 0.2523 while the intercept 0.1117. The RNA-to-Protein ratio follows the equation.

in which 𝑅 is total cellular RNA mass (g/gCDW), 𝑃 is total cellular protein mass (g/gCDW).

Accordingly, we can estimate = 1/0.2523 = 3.964, while = 0.1117 based on the equation. Ribosomal catalytic rate (aa/ribosome/s) can be formulated as

in which is protein synthesis rate (aa/s), is number of ribosomes, is the molecular weight of average amino acid (g/mol), is the mass of rRNA per ribosome (g/mol ribosome),   
 is the fraction of rRNA in total RNA.

Using the R/P equation, then

in which:

(4)

(5)

Given that = 109g/mol, = 1.90E6, = 0.85 (bionumber: 105192, 100258)63,64, we can calculate that = 22.6, and = 0.443. The ribosomal catalytic rate(aa/ribosome/s) is hence:

(6)

# 5. Constraints

Constraints are required to perform simulations. All constraints were formatted into a linear programming (LP) file for solving as required by the solver SoPlex. Besides the basic flux balanced analysis (FBA) simulation constraint such as:

which exists in the basic GEM to represent the steady state and the flux for each reaction should be between the lower bound and upper bound.



Fig. 5 Flowchart of the protein related process in the pcSecYeast.

Protein related constraints are added to couple metabolic reactions and the corresponding enzyme enzymes (Fig. 5), which is represented as the rate of a metabolic reaction is constrained by the concentration of the enzyme that catalyzes it:

(9)

where

In which, represents the formation rate for one enzyme complex, while represents the dilution rate. Dilution rate of one enzyme complex is coupled to the growth rate, thus we can calculate the enzyme abundance from the synthesis rate and the growth rate (Eq. 12). To be noted here in Eq. 12, enzyme complexes are only diluted without degradation in the pcSecYeast, since the degradation happened during the folding process (Fig. 5). Combining (Eq. 9-12), we can couple the reaction flux for metabolic reaction with its corresponding enzyme complex formation as:

(13)

This type of inequality constraint has been applied to other reactions in protein biosynthesis process, including coupling translation rate and ribosome synthesis rate (Fig. 5):

(14)

coupling ribosome synthesis rate and ribosome assembly rate (Fig. 5), which presents the ribosome synthesis rate is constrained by abundance of ribosome assembly factors:

(15)

coupling protein degradation rate and proteasome synthesis rate (Fig. 5):

(16)

coupling post-translational modification rate with corresponding enzyme synthesis rate (Fig. 5):

(17)

coupling the misfolded protein degradation with the protein translation (Fig. 5), is collected from the reference. For proteins with no measurement, mean ratio of 30% is used:

(18)

Total proteome is constrained as Eq. 19, in which 0.46g is the protein content of 1g biomass.

(19)

Besides that, we also included extra constraints in the parameter sensitivity analysis part for CPY accumulation simulation (Eq.20-24). Since the protein volume can be roughly considered as the linear correlation with the protein mass65. Thus, we can transfer the volume into simple abundance constraint. Therefore, we calculated the maximum value (0.0786g/gCDW) for total ER proteins from multiple available proteome data for *S. cerevisiae* under diverse conditions1,66–70. Then, we used the value to constrain the ER protein abundance, which is represented as the sum of each ER protein abundance and then converted to the protein synthesis rate according to the Eq. 12.

ER volume constraint: (20)

The similar constraint was added for ER membrane proteins, ERAD pathway proteins, secretory machinery proteins and retro-translocation enzymes.

0.008 (21)

0.0125 (22)

0.0244 (23)

8.08e-5 (24)

# 6. Comparison of pcSecYeast with other models.

pcSecYeast adopts the similar fine-grained proteome constrained concept as in the proteome constrained model pcYeast10 and ME model yETFL71, which is to couple the protein synthesis with the metabolic reactions through the enzyme kinetic capacity. The main difference of pcSecYeast with those models is that pcSecYeast has well-constructed processes for protein synthesis, folding, misfolding and degradation, which are rather lumped into one reaction or not covered in ME models or pc models. Therefore, pcSecYeast is more likely an expanded version of ME model/pc models in terms of the protein folding, misfolding, degradation process. Besides all applications of ME models or pc models, pcSecYeast allows the precise analysis of cellular behaviors for different physiological conditions, especially for the simulation of recombinant protein production.

The main difference of pcSecYeast with the whole cell model of yeast (WM\_S288C72) is the process coverage. WM\_S288C decomposes cell functionality into 26 cellular processes, while pcSecYeast covers six of those processes including metabolism, protein translation, protein folding, protein decay, protein modification and ribosome assembly. However, the protein folding and protein decay in pcSecYeast is more complex and protein-specific compared with that in WM-S288c and includes the secretion pathway in detail. As for the protein modification processes, glycosylation, disulfide bond and GPI considered in pcSecYeast are not considered in WM-S288C. Besides that, proteome-constrained models (including pcYeast73 and yETFL71 and pcSecYeast) are constraint-based optimization frameworks with the steady state assumption, while the whole cell model is a dynamic model which uses ordinary differential equations. Compared with the vast environment and condition dependent parameters requirement as in the WM-S288C72, pcSecYeast requires less parameters, which enables a more efficient simulation of cell behavior.

Compared with other secretory models74,75 which adopts basic GEM concept, pcSecYeast adopts the proteome constrained approach, which links the protein synthesis with the metabolism. This kind of expansion of protein synthesis can improve model prediction in terms of lower variability of simulated fluxes, improved capability of complex phenotype prediction. Besides that, this kind of model can serve as the platform to integrate the transcriptome and proteome data for accurate flux simulation. PcSecYeast also covers more processes compared with previous published secretory model, which is the current most comprehensive secretory model.

Table 1. Comparison of pcSecYeast with other models.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Models** | | **ihGlycopastoris**75 | **Mammalian secretory model**74 | **pcYeast**10 | **yETFL**71 | **pcSecYeast** | **WM\_S288C**72 |
| Model type | | Basic GEM | Basic GEM | Fine-grained proteome-constrained GEM | Fine-grained proteome-constrained GEM | Fine-grained proteome-constrained GEM | whole-cell model |
| Organism | | *P. pastoris* | Mammalian cells | *S. cerevisiae* | *S. cerevisiae* | *S. cerevisiae* | *S. cerevisiae* |
| Model assumption | | Steady state | Steady state | Steady state | Steady state | Steady state | Dynamic |
| Constraint | | Mass balance | Mass balance | Mass balance, kinetic constraint | Mass balance, kinetic constraint, thermodynamic constraints | Mass balance, kinetic constraint | - |
| Processes | Metabolism | Yes | Yes | Yes | Yes | Yes | Yes |
| Transcription | No | No | No | Yes, lumped | No | Yes |
| RNA cleavage | No | No | No | No | No | Yes |
| tRNA modification | No | No | Yes | No | No | Yes |
| tRNA charging | No | No | Yes | Yes | Yes | Yes |
| Translation | For the recombinant protein, NOT for native proteins | For the recombinant protein, NOT for native proteins | Yes | Yes | Yes | Yes |
| Folding | For the recombinant protein, NOT for native proteins | For the recombinant protein, NOT for native proteins | Lumped with protein-specific chaperones | Lumped | Comprehensive, protein-specific | Lumped |
| Misfolding | No | For the recombinant protein, NOT for native proteins | Lumped | Lumped | Comprehensive, protein-specific | Lumped |
| Protein translational modification | *N*-glycosylation | Glycosylation and disulfide bond formation | No | No | Glycosylation, GPI anchor and disulfide bond formation | Phosphorylation, acetylation and ubiquitination |
| Degradation | For the recombinant protein, NOT for native proteins | For the recombinant protein, NOT for native proteins | Lumped | Lumped | Comprehensive, protein-specific | Lumped |
| Sorting | For the recombinant protein, NOT for native proteins | For the recombinant protein, NOT for native proteins | Lumped | No | Comprehensive, protein-specific | Lumped |
| Ribosome assembly | No | No | Yes | Yes | Yes | Yes |
| Protein complex formation | No | No | Yes | Yes | Yes | Yes |
| Other cell processes | No | No | No | No | No | Yes |
| Model application | Simulate proteome changes | No | No | Yes | Yes | Yes | Yes |
| Integrate proteome data | No | No | Yes | Yes | Yes | No |
| Simulate protein misfolding | No | No | No | No | Yes | Yes |
| Native protein competition with recombinant protein | No | No | N/A | N/A | Yes | N/A |
| Simulate engineering targets for improving recombinant targets | Only targets in metabolic pathway | Only targets in metabolic pathway | N/A | N/A | Targets both in secretory and metabolic pathways | N/A |

N/A means that the description is not applicable for the specific model.

**References**

1. Lahtvee, P.-J. *et al.* Absolute quantification of protein and mRNA abundances demonstrate variability in gene-specific translation efficiency in yeast. *Cell Syst.* **4**, 495–504 (2017).

2. Christiano, R. *et al.* A Systematic Protein Turnover Map for Decoding Protein Degradation. *Cell Rep.* **33**, 108378 (2020).

3. Feizi, A., Österlund, T., Petranovic, D., Bordel, S. & Nielsen, J. Genome-Scale Modeling of the Protein Secretory Machinery in Yeast. *PLoS One* **8**, e63284 (2013).

4. The UniProt Consortium. UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* **45**, D158–D169 (2017).

5. Lu, H. *et al.* A consensus *S. cerevisiae* metabolic model Yeast8 and its ecosystem for comprehensively probing cellular metabolism. *Nat. Commun.* **10**, 1–13 (2019).

6. Bosson, R., Guillas, I., Vionnet, C., Roubaty, C. & Conzelmann, A. Incorporation of ceramides into *Saccharomyces cerevisiae* glycosylphosphatidylinositol-anchored proteins can be monitored in vitro. *Eukaryot. Cell* **8**, 306–314 (2009).

7. Lopez, S., Rodriguez-Gallardo, S., Sabido-Bozo, S. & Muñiz, M. Endoplasmic Reticulum Export of GPI-Anchored Proteins. *Int. J. Mol. Sci.* **20**, (2019).

8. Dever, T. E. & Green, R. The elongation, termination, and recycling phases of translation in eukaryotes. *Cold Spring Harb. Perspect. Biol.* **4**, 1–16 (2012).

9. de la Cruz, J., Karbstein, K. & Woolford, J. L. Functions of ribosomal proteins in assembly of eukaryotic ribosomes in vivo. *Annu. Rev. Biochem.* **84**, 93–129 (2015).

10. Elsemman, I. E. *et al.* Whole-cell modeling in yeast predicts compartment-specific proteome constraints that drive metabolic strategies. *Nat. Commun.* **13**, 801 (2022).

11. Chen, Y. *et al.* Proteome constraints reveal targets for improving microbial fitness in nutrient-rich environments. *Mol. Syst. Biol.* **17**, e10093 (2021).

12. Delic, M. *et al.* The secretory pathway: exploring yeast diversity. *FEMS Microbiol. Rev.* **37**, 872–914 (2013).

13. Johnson, N., Powis, K. & High, S. Post-translational translocation into the endoplasmic reticulum. *Biochim. Biophys. Acta (BBA)-Molecular Cell Res.* **1833**, 2403–2409 (2013).

14. Brown, J. D. *et al.* Subunits of the *Saccharomyces cerevisiae* signal recognition particle required for its functional expression. *EMBO J.* **13**, 4390–4400 (1994).

15. Ogg, S. C., Poritz, M. A. & Walter, P. Signal recognition particle receptor is important for cell growth and protein secretion in *Saccharomyces cerevisiae*. *Mol. Biol. Cell* **3**, 895–911 (1992).

16. Siegel, V. & Walter, P. Each of the activities of signal recognition particle (SRP) is contained within a distinct domain: analysis of biochemical mutants of SRP. *Cell* **52**, 39–49 (1988).

17. Gautschi, M. *et al.* RAC, a stable ribosome-associated complex in yeast formed by the DnaK-DnaJ homologs Ssz1p and zuotin. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 3762–3767 (2001).

18. Young, B. P., Craven, R. A., Reid, P. J., Willer, M. & Stirling, C. J. Sec63p and Kar2p are required for the translocation of SRP-dependent precursors into the yeast endoplasmic reticulum in vivo. *EMBO J.* **20**, 262–271 (2001).

19. Araki, K. & Nagata, K. Protein folding and quality control in the ER. *Cold Spring Harb. Perspect. Biol.* **3**, a007526 (2011).

20. Fujita, M., Umemura, M., Yoko-o, T. & Jigami, Y. PER1 is required for GPI-phospholipase A2 activity and involved in lipid remodeling of GPI-anchored proteins. *Mol. Biol. Cell* **17**, 5253–5264 (2006).

21. Umemura, M., Fujita, M., Yoko-O, T., Fukamizu, A. & Jigami, Y. *Saccharomyces cerevisiae* CWH43 is involved in the remodeling of the lipid moiety of GPI anchors to ceramides. *Mol. Biol. Cell* **18**, 4304–4316 (2007).

22. Kinoshita, T. & Fujita, M. Biosynthesis of GPI-anchored proteins: special emphasis on GPI lipid remodeling. *J. Lipid Res.* **57**, 6–24 (2016).

23. Girrbach, V. & Strahl, S. Members of the evolutionarily conserved PMT family of protein O-mannosyltransferases form distinct protein complexes among themselves. *J. Biol. Chem.* **278**, 12554–12562 (2003).

24. Herscovics, A. Processing glycosidases of *Saccharomyces cerevisiae*. *Biochim. Biophys. Acta* **1426**, 275–285 (1999).

25. Jiang, B. *et al.* CWH41 encodes a novel endoplasmic reticulum membrane N-glycoprotein involved in beta 1,6-glucan assembly. *J. Bacteriol.* **178**, 1162–1171 (1996).

26. Lopata, A., Kniss, A., Löhr, F., Rogov, V. V & Dötsch, V. Ubiquitination in the ERAD Process. *Int. J. Mol. Sci.* **21**, (2020).

27. Paul, P. *et al.* The protein translocation systems in plants - composition and variability on the example of Solanum lycopersicum. *BMC Genomics* **14**, 189 (2013).

28. Thibault, G. & Ng, D. T. W. The endoplasmic reticulum-associated degradation pathways of budding yeast. *Cold Spring Harb. Perspect. Biol.* **4**, (2012).

29. Schmidt, C. C., Vasic, V. & Stein, A. Doa10 is a membrane protein retrotranslocase in ER-associated protein degradation. *Elife* **9**, (2020).

30. Jakob, C. A., Burda, P., Roth, J. & Aebi, M. Degradation of misfolded endoplasmic reticulum glycoproteins in *Saccharomyces cerevisiae* is determined by a specific oligosaccharide structure. *J. Cell Biol.* **142**, 1223–1233 (1998).

31. Peth, A., Nathan, J. A. & Goldberg, A. L. The ATP costs and time required to degrade ubiquitinated proteins by the 26 S proteasome. *J. Biol. Chem.* **288**, 29215–29222 (2013).

32. Julius, D., Blair, L., Brake, A., Sprague, G. & Thorner, J. Yeast alpha factor is processed from a larger precursor polypeptide: the essential role of a membrane-bound dipeptidyl aminopeptidase. *Cell* **32**, 839–852 (1983).

33. Caldwell, S. R., Hill, K. J. & Cooper, A. A. Degradation of endoplasmic reticulum (ER) quality control substrates requires transport between the ER and Golgi. *J. Biol. Chem.* **276**, 23296–23303 (2001).

34. Gomez-Navarro, N. & Miller, E. Protein sorting at the ER-Golgi interface. *J. Cell Biol.* **215**, 769–778 (2016).

35. Strating, J. R. P. M. & Martens, G. J. M. The p24 family and selective transport processes at the ER-Golgi interface. *Biol. cell* **101**, 495–509 (2009).

36. Stagg, S. M. *et al.* Structure of the Sec13/31 COPII coat cage. *Nature* **439**, 234–238 (2006).

37. Gimeno, R. E., Espenshade, P. & Kaiser, C. A. SED4 encodes a yeast endoplasmic reticulum protein that binds Sec16p and participates in vesicle formation. *J. Cell Biol.* **131**, 325–338 (1995).

38. Noda, Y., Yamagishi, T. & Yoda, K. Specific membrane recruitment of Uso1 protein, the essential endoplasmic reticulum-to-Golgi tethering factor in yeast vesicular transport. *J. Cell. Biochem.* **101**, 686–694 (2007).

39. Behnia, R., Barr, F. A., Flanagan, J. J., Barlowe, C. & Munro, S. The yeast orthologue of GRASP65 forms a complex with a coiled-coil protein that contributes to ER to Golgi traffic. *J. Cell Biol.* **176**, 255–261 (2007).

40. Cai, H. *et al.* TRAPPI tethers COPII vesicles by binding the coat subunit Sec23. *Nature* **445**, 941–944 (2007).

41. Lian, J. P. & Ferro-Novick, S. Bos1p, an integral membrane protein of the endoplasmic reticulum to Golgi transport vesicles, is required for their fusion competence. *Cell* **73**, 735–745 (1993).

42. Nakayama, K., Nagasu, T., Shimma, Y., Kuromitsu, J. & Jigami, Y. OCH1 encodes a novel membrane bound mannosyltransferase: outer chain elongation of asparagine-linked oligosaccharides. *EMBO J.* **11**, 2511–2519 (1992).

43. Munro, S. What can yeast tell us about N-linked glycosylation in the Golgi apparatus? *FEBS Lett.* **498**, 223–227 (2001).

44. Conde, R., Cueva, R., Pablo, G., Polaina, J. & Larriba, G. A search for hyperglycosylation signals in yeast glycoproteins. *J. Biol. Chem.* **279**, 43789–43798 (2004).

45. Stolz, J. & Munro, S. The components of the *Saccharomyces cerevisiae* mannosyltransferase complex M-Pol I have distinct functions in mannan synthesis. *J. Biol. Chem.* **277**, 44801–44808 (2002).

46. De Pourcq, K., De Schutter, K. & Callewaert, N. Engineering of glycosylation in yeast and other fungi: current state and perspectives. *Appl. Microbiol. Biotechnol.* **87**, 1617–1631 (2010).

47. Lussier, M., Sdicu, A. M., Bussereau, F., Jacquet, M. & Bussey, H. The Ktr1p, Ktr3p, and Kre2p/Mnt1p mannosyltransferases participate in the elaboration of yeast O- and N-linked carbohydrate chains. *J. Biol. Chem.* **272**, 15527–15531 (1997).

48. Spang, A. Retrograde traffic from the Golgi to the endoplasmic reticulum. *Cold Spring Harb. Perspect. Biol.* **5**, (2013).

49. Cowles, C. R., Odorizzi, G., Payne, G. S. & Emr, S. D. The AP-3 adaptor complex is essential for cargo-selective transport to the yeast vacuole. *Cell* **91**, 109–118 (1997).

50. Bowers, K. & Stevens, T. H. Protein transport from the late Golgi to the vacuole in the yeast *Saccharomyces cerevisiae*. *Biochim. Biophys. Acta* **1744**, 438–454 (2005).

51. Gurunathan, S., David, D. & JE, G. Dynamin and clathrin are required for the biogenesis of a distinct class of secretory vesicles in yeast. *EMBO J.* **21**, 602–614 (2002).

52. Finger, F. P. & Novick, P. Spatial regulation of exocytosis: lessons from yeast. *J. Cell Biol.* **142**, 609–612 (1998).

53. Chen, Y., Li, F., Mao, J., Chen, Y. & Nielsen, J. Yeast optimizes metal utilization based on metabolic network and enzyme kinetics. *Proc. Natl. Acad. Sci.* **118**, (2021).

54. Chen, Y. & Nielsen, J. In vitro turnover numbers do not reflect in vivo activities of yeast enzymes. *Proc. Natl. Acad. Sci. U. S. A.* **118**, (2021).

55. Wang, M., Herrmann, C. J., Simonovic, M., Szklarczyk, D. & von Mering, C. Version 4.0 of PaxDb: Protein abundance data, integrated across model organisms, tissues, and cell-lines. *Proteomics* **15**, 3163–3168 (2015).

56. Huber, A. *et al.* Sch9 regulates ribosome biogenesis via Stb3, Dot6 and Tod6 and the histone deacetylase complex RPD3L. *EMBO J.* **30**, 3052–3064 (2011).

57. Henderson, A., Erales, J., Hoyt, M. A. & Coffino, P. Dependence of proteasome processing rate on substrate unfolding. *J. Biol. Chem.* **286**, 17495–17502 (2011).

58. O’Brien, E. J., Lerman, J. A., Chang, R. L., Hyduke, D. R. & Palsson, B. Ø. Genome-scale models of metabolism and gene expression extend and refine growth phenotype prediction. *Mol. Syst. Biol.* **9**, 693 (2013).

59. Gombert, A. K., Moreira dos Santos, M., Christensen, B. & Nielsen, J. Network identification and flux quantification in the central metabolism of *Saccharomyces cerevisiae* under different conditions of glucose repression. *J. Bacteriol.* **183**, 1441–1451 (2001).

60. Ertugay, N. & Hamamci, H. Continuous cultivation of bakers’ yeast: change in cell composition at different dilution rates and effect of heat stress on trehalose level. *Folia Microbiol. (Praha).* **42**, 463–467 (1997).

61. Fonseca, G. G., Gombert, A. K., Heinzle, E. & Wittmann, C. Physiology of the yeast Kluyveromyces marxianus during batch and chemostat cultures with glucose as the sole carbon source. *FEMS Yeast Res.* **7**, 422–435 (2007).

62. Karpinets, T. V, Greenwood, D. J., Sams, C. E. & Ammons, J. T. RNA:protein ratio of the unicellular organism as a characteristic of phosphorous and nitrogen stoichiometry and of the cellular requirement of ribosomes for protein synthesis. *BMC Biol.* **4**, 30 (2006).

63. von der Haar, T. A quantitative estimation of the global translational activity in logarithmically growing yeast cells. *BMC Syst. Biol.* **2**, 87 (2008).

64. Warner, J. R. The economics of ribosome biosynthesis in yeast. *Trends Biochem. Sci.* **24**, 437–440 (1999).

65. Erickson, H. P. Size and shape of protein molecules at the nanometer level determined by sedimentation, gel filtration, and electron microscopy. *Biol. Proced. Online* **11**, 32–51 (2009).

66. Yu, R., Vorontsov, E., Sihlbom, C. & Nielsen, J. Quantifying absolute gene expression profiles reveals distinct regulation of central carbon metabolism genes in yeast. *Elife* **10**, (2021).

67. Qi, Q. *et al.* Different Routes of Protein Folding Contribute to Improved Protein Production in *Saccharomyces cerevisiae*. *MBio* **11**, (2020).

68. Björkeroth, J. *et al.* Proteome reallocation from amino acid biosynthesis to ribosomes enables yeast to grow faster in rich media. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 21804–21812 (2020).

69. Yu, R. *et al.* Nitrogen limitation reveals large reserves in metabolic and translational capacities of yeast. *Nat. Commun.* **11**, 1881 (2020).

70. Di Bartolomeo, F. *et al.* Absolute yeast mitochondrial proteome quantification reveals trade-off between biosynthesis and energy generation during diauxic shift. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 7524–7535 (2020).

71. Oftadeh, O. *et al.* A genome-scale metabolic model of *Saccharomyces cerevisiae* that integrates expression constraints and reaction thermodynamics. *Nat. Commun.* **12**, 4790 (2021).

72. Ye, C. *et al.* Comprehensive understanding of *Saccharomyces cerevisiae* phenotypes with whole-cell model WM\_S288C. *Biotechnol. Bioeng.* **117**, 1562–1574 (2020).

73. Elsemman, I. E. *et al.* Whole-cell modeling in yeast predicts compartment-specific proteome constraints that drive metabolic strategies. *bioRxiv* 2021.06.11.448029 (2021). doi:10.1101/2021.06.11.448029

74. Gutierrez, J. M. *et al.* Genome-scale reconstructions of the mammalian secretory pathway predict metabolic costs and limitations of protein secretion. *Nat. Commun.* **11**, 68 (2020).

75. Irani, Z. A., Kerkhoven, E. J., Shojaosadati, S. A. & Nielsen, J. Genome-scale metabolic model of *Pichia pastoris* with native and humanized glycosylation of recombinant proteins. *Biotechnol. Bioeng.* **113**, 961–969 (2016).