LyGo

A Platform for Rapid Screening of Lytic Polysaccharide Monooxygenase Production

Hernández-Rollán, Cristina; Falkenberg, Kristoffer B.; Rennig, Maja; Bertelsen, Andreas B.; Ipsen, Johan; Brander, Søren; Daley, Daniel O.; Johansen, Katja S.; Nørholm, Morten H.H.

Published in:
ACS Synthetic Biology

Link to article, DOI:
[10.1021/acssynbio.1c00034](https://doi.org/10.1021/acssynbio.1c00034)

Publication date:
2021

Document Version
Early version, also known as pre-print

[Link back to DTU Orbit](#)

Citation (APA):

Hernández-Rollán, C., Falkenberg, K. B., Rennig, M., Bertelsen, A. B., Ipsen, J., Brander, S., Daley, D. O., Johansen, K. S., & Nørholm, M. H. H. (2021). LyGo: A Platform for Rapid Screening of Lytic Polysaccharide Monooxygenase Production. *ACS Synthetic Biology*, 10(4), 897-906. <https://doi.org/10.1021/acssynbio.1c00034>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LyGo

A Platform for Rapid Screening of Lytic Polysaccharide Monooxygenase Production

Hernández-rollán, Cristina; Falkenberg, Kristoffer B.; Rennig, Maja; Bertelsen, Andreas B.; Ipsen, Johan Ø.; Brander, Søren; Daley, Daniel O.; Johansen, Katja S.; Nørholm, Morten H. H.

Published in:
A C S Synthetic Biology

DOI:
[10.1021/acssynbio.1c00034](https://doi.org/10.1021/acssynbio.1c00034)

Publication date:
2021

Document version
Early version, also known as pre-print

Citation for published version (APA):
Hernández-rollán, C., Falkenberg, K. B., Rennig, M., Bertelsen, A. B., Ipsen, J. Ø., Brander, S., Daley, D. O., Johansen, K. S., & Nørholm, M. H. H. (2021). LyGo: A Platform for Rapid Screening of Lytic Polysaccharide Monooxygenase Production. *A C S Synthetic Biology*, 10(4), 897-906.
<https://doi.org/10.1021/acssynbio.1c00034>

1 **LyGo: A platform for rapid screening of lytic polysaccharide monooxygenase production**

2
3 Cristina Hernández-Rollán^{1†}, Kristoffer B. Falkenberg^{1†}, Maja Rennig^{1,2}, Andreas B. Bertelsen¹, Johan Ø.
4 Ipsen³, Søren Brander⁵, Daniel O. Daley^{2,4}, Katja S. Johansen⁵, and Morten H. H. Nørholm^{1,2*}.

5
6 ¹The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, 2800
7 Kongens Lyngby, Denmark

8 ²Mycropt ApS, 2800 Kongens Lyngby, Denmark

9 ³Department of Plant and Environmental Sciences, University of Copenhagen, 1871 Frederiksberg,
10 Denmark

11 ⁴Center for Biomembrane Research, Department of Biochemistry and Biophysics, Stockholm University,
12 10691 Stockholm, Sweden

13 ⁵Department of Geosciences and Natural Resource Management, University of Copenhagen, 1958
14 Frederiksberg, Denmark

15
16 *Corresponding author

17
18 Addresses of all authors

19 Cristina Hernández-Rollán: crirol@biosustain.dtu.dk

20 Kristoffer B. Falkenberg: kribaf@biosustain.dtu.dk

21 Maja Rennig: rennig@biosustain.dtu.dk

22 Andreas B. Bertelsen: andbir@biosustain.dtu.dk

23 Johan Ø. Ipsen: Jip@plen.ku.dk

24 Søren Brander: sbd@ign.ku.dk

25 Katja S. Johansen: ksj@ign.ku.dk

26 Daniel O. Daley: ddaley@dbb.su.se

27 Morten H. H. Nørholm: morno@biosustain.dtu.dk

28

29 **Author contributions:**

30 CHR established the LyGo cloning protocol and cloned, established, and performed experiments, and
 31 processed data for proteins produced in the periplasm of *E. coli*. KBF cloned, established, and performed
 32 experiments, and processed data for proteins produced in *B. subtilis* and performed experiments and
 33 processed data for proteins produced in *K. phaffii*. ABB cloned, established, and performed experiments,
 34 and processed data for proteins produced in the cytoplasm of *E. coli*. MR cloned, established, and
 35 performed experiments, and processed data for proteins produced on the surface of *E. coli* and cloned
 36 and established experiments for proteins produced in *K. phaffii*. ABB and JØI performed gel filtration of
 37 the cytoplasmically produced samples and JØI performed the subsequent activity assays and processed
 38 the associated data. All authors read and approved the content of the manuscript.

39

40 [†]These authors contributed equally

41

42 Keywords: Protein production, Expression vector, Cloning, Lytic Polysaccharide Monooxygenase,
43 LPMO.

Abstract

Environmentally friendly sources of energy and chemicals are essential constituents of a sustainable society. An important step towards this goal is the utilization of non-edible biomass as supply of building blocks for future biorefineries. Lytic polysaccharide monooxygenases (LPMOs) are enzymes that play a critical role in breaking the chemical bonds in the most abundant polymers found in recalcitrant biomass, such as cellulose and chitin. Predicting optimal strategies for producing LPMOs is often non-trivial, and methods allowing for screening several strategies simultaneously are therefore needed. Here, we present a standardized platform for cloning LPMOs. The platform allows users to combine gene fragments with different expression vectors in a simple 15-minute reaction, thus enabling rapid exploration of several gene contexts, hosts and expression strategies in parallel. The open-source LyGo platform is accompanied by easy-to-follow online protocols for both cloning and expression. As a demonstration, we utilize the LyGo platform to explore different strategies for expressing several different LPMOs in *Escherichia coli*, *Bacillus subtilis*, and *Komagataella phaffii*.

Introduction

Transforming society from being dependent on fossil resources, to relying on sustainable resources is one of the most important challenges of the 21st century. Overcoming this challenge is highly dependent on the displacement of petrochemicals with renewable biochemicals such as bioethanol, polylactic acid, and biosuccinic acid^{1,2}. Currently, the majority of biorefineries in the EU utilize food crops, which gives rise to the food versus fuel debate and ultimately limits the implementation of these types of biorefineries^{3,4}. This issue is mitigated by 2nd generation biorefineries that rely on non-edible biomass as

a resource, although depolymerization of this type of biomass constitutes a significant technical challenge^{5,6}.

Lytic Polysaccharide Monooxygenases (LPMOs) are carbohydrate active enzymes that oxidize α and β -(1,4) glycosidic bonds in recalcitrant biopolymers, such as chitin and cellulose^{7,8,9}. Due to the chemistry and architecture of the substrate binding surface, LPMOs generally act more readily on crystalline substrates compared to glycoside hydrolases (GHs)^{7,10}. LPMO-induced cleavage weakens the structure of the biomass, and facilitates attack by other enzymes in a cascading mechanism^{11,12}. Consequently, addition of LPMOs has been shown to increase the saccharification efficiency of commercially available cellulose cocktails^{9,13}. Furthermore, LPMOs and LPMO-like proteins are involved in the virulence of plant¹⁴ and animal (including human) pathogens¹⁵. These factors make LPMOs highly interesting from both an industrial and academic perspective.

Heterologous expression of LPMOs is essential for characterization and industrial production of LPMOs. LPMOs have been produced in a range of organisms¹⁶. However, multiple expression strategies for the same enzymes are rarely explored or compared and previous efforts to standardize cloning into expression vectors in the LPMO field have been limited¹⁷. Ideally, a comprehensive strategy for heterologous expression of LPMOs enables simultaneous exploration of several strategies including e.g. different production hosts, gene sequence variants, signal peptides for secretion, localization within the hosts, solubility- and affinity tags, etc. Such a workflow would benefit from a simple standardized DNA editing approach and a diverse and accessible vector collection, which in addition would facilitate collaborations, automation, and data comparison.

Here, we describe an open-source platform, called “LyGo” (Lytic polysaccharide monooxygenase Golden gate cloning). The platform includes functionally validated expression vectors for three widely used protein production host: *Escherichia coli*, *Bacillus subtilis*, and *Komagataella phaffii* (formerly known as *Pichia pastoris*) and is accompanied by easy-to-follow online protocols for cloning and expression in each of the included organisms. It is our hope that the community will embrace, share, and expand the collection in the future. To demonstrate how the LyGo platform can be utilized, we explore a variety of expression strategies in the three production hosts. In all cases, the expression strategies show significant differences in performance, thus underlining why this type of synthetic biology resource is useful for improving production yields and for enabling sustainable biotech.

Results

LyGo cloning enables efficient and scarless assembly of expression vectors

Inspired by Golden Gate cloning¹⁸ and the Electra Vector System® (ATUM, Newark, CA, USA), the LyGo platform utilizes the type IIS restriction endonuclease SapI to generate compatible DNA fragments. Because SapI cuts outside its recognition sequence, this allows for scarless assembly of DNA fragments containing LPMO-coding sequences (called “LyGo fragments”) into compatible expression vectors (called a “LyGo vectors”). SapI-treatment of a LyGo fragment creates three-nucleotide single stranded overhangs corresponding to the codon of the N-terminal histidine (essential for the activity of LPMOs¹⁹) and a stop codon in the 3’ end (Figure 1A). LyGo vectors contain a cloning cassette containing a *ccdB* counter selection marker²⁰ flanked by SapI recognition sites (called a “LyGo cassette”). Upon SapI digestion, the LyGo cassette is released from a LyGo vector, leaving single stranded overhangs

106 complementary to the digested LyGo fragments. This design ensures that any LPMO gene-of-interest,
 107 upon treatment with SapI and T4 ligase, can replace the cloning cassette without leaving cloning scars in
 108 the final vector. When a LyGo vector is combined with a LyGo fragment the SapI recognition sites are
 109 eliminated, and *in vitro* the correct assembly reaction is therefore favored over re-ligation of the LyGo
 110 cassette. Should any background vectors persist, the resulting transformants are selected against *in vivo*
 111 by the presence of the *ccdB* gene (Figure 1B, and Supplementary Figure S1A).

112 Digestion and ligation are performed in a one-pot reaction at room temperature and can be
 113 completed in 15 minutes with high assembly efficiencies (Supplementary Figure S1B). The assembled
 114 DNA can be transformed directly into *E. coli*, following routine protocols²¹. The assembly is not
 115 dependent on DNA amplification or purification, which can be time-consuming, error-prone, and
 116 inefficient. A clear strength of this vector design is that any vector can be “LyGo-fied”, simply by
 117 removing unwanted SapI recognition sites and inserting the LyGo cassette. This allows users to utilize
 118 their favorite vectors and to contribute to the collection.

Figure 1: Schematic overview of LyGo cloning. **(A)** SapI recognition sites (orange), conserved histidine- (green), and stop codons (red) are shown in the LyGo-fragment, -cassette, and -vector before and after SapI digestion. **(B)** In the cloning procedure, LyGo fragment and vector are mixed with SapI restriction enzyme and T4 ligase, digested, and ligated in a 15-minute reaction, and subsequently transformed into a *ccdB* sensitive *E. coli* strain. This way, undigested or religated products are eliminated by *ccdB* counter selection.

Nomenclature for LyGo vectors

The current vector collection consists of 14 vectors in total (Table 1), and we established a simple abstraction and nomenclature, in order to easily recognize plasmids and translate their names into meaningful concepts. LyGo vectors are designated: “pLyGo” followed by two letters indicating the

131 expression host (*Ec* for *E. coli*, *Bs* for *B. subtilis*, and *Kp* for *K. phaffii*), as well as a number that refers to
 132 information about the features of the plasmid. Any modules inserted by LyGo cloning of LyGo fragments
 133 are appended to the name separated by hyphens. For instance, pLyGo-*Ec*-2 is a LyGo vector for
 134 expression in *E. coli* that includes the T7 promoter and the signal peptide MalE^{SP} targeting LPMOs-of-
 135 interest to the periplasm. Cloning of a LyGo fragment encoding the LPMO LsAA9A into pLyGo-*Ec*-2,
 136 would be named pLyGo-*Ec*-2-LsAA9A.

137

138 **Table 1:** LyGo vectors constructed in this study. These vectors are available through plasmid repositories.

139 TEV* denotes that the last codon of the TEV sequence was modified to a histidine codon.

LyGo nomenclature	Classical nomenclature	Description
<i>E. coli</i> vectors		
pLyGo- <i>Ec</i> -1	pET39b-Ub(his10)-TEV*- <i>ccdB</i>	Vector encoding Ubiquitin with an internal His10-tag and the LyGo cassette, Km ^R
pLyGo- <i>Ec</i> -2	pET28a(+)-MalE ^{SP} - <i>ccdB</i>	Vector encoding the MalE signal peptide and the LyGo cassette, Km ^R
pLyGo- <i>Ec</i> -3	pET28a(+)-OmpA ^{SP} - <i>ccdB</i>	Vector encoding the OmpA signal peptide and the LyGo cassette, Km ^R
pLyGo- <i>Ec</i> -4	pET28a(+)-PhoA ^{SP} - <i>ccdB</i>	Vector encoding the PhoA signal peptide and the LyGo cassette, Km ^R
pLyGo- <i>Ec</i> -5	pET28a(+)-DsbA ^{SP} - <i>ccdB</i>	Vector encoding the DsbA signal peptide and the LyGo cassette, Km ^R
pLyGo- <i>Ec</i> -6	pET28a(+)-PelB ^{SP} - <i>ccdB</i>	Vector encoding the PelB signal peptide and the LyGo cassette, Km ^R
pLyGo- <i>Ec</i> -7	pBAD42-PelB ^{SP} - <i>ccdB</i> -TEV-NB-C-IgAP	Vector encoding the C-IgAP surface expression anchor, nanobody, and the LyGo cassette, Km ^R
pLyGo- <i>Ec</i> -8	pBAD42-Lpp ^{SP} -OmpA-TEV*- <i>ccdB</i>	Vector encoding the Lpp ^{SP} -OmpA surface expression anchor and the LyGo cassette, Km ^R
<i>B. subtilis</i> vectors		

pLyGo-Bs-1	pBS293C-amyE-P _{3P} -AmyL ^{SP} - <i>ccdB</i>	Vector encoding the AmyL signal peptide and the LyGo cassette, Km ^R (<i>E. coli</i>), Cm ^R (<i>B. subtilis</i>)
pLyGo-Bs-2	pBS293C-amyE-P _{3P} -AmyQ ^{SP} - <i>ccdB</i>	Vector encoding the AmyQ signal peptide and the LyGo cassette, Km ^R (<i>E. coli</i>), Cm ^R (<i>B. subtilis</i>)
pLyGo-Bs-3	pBS293C-amyE-P _{3P} -AprE ^{SP} - <i>ccdB</i>	Vector encoding the AprE signal peptide and the LyGo cassette, Km ^R (<i>E. coli</i>), Cm ^R (<i>B. subtilis</i>)
pLyGo-Bs-4	pBS293C-amyE-P _{3P} -BatLPMO10 ^{SP} - <i>ccdB</i>	Vector encoding the BatLPMO10 signal peptide and the LyGo cassette, Km ^R (<i>E. coli</i>), Cm ^R (<i>B. subtilis</i>)
<i>K. phaffii</i> vectors		
pLyGo-Kp-1	pPIC9K- α -MF ^{SP} - <i>ccdB</i>	Vector encoding the α -MF pre-sequence and the LyGo cassette, Amp ^R (<i>E. coli</i>), Km ^R , <i>HIS4</i> (<i>K. phaffii</i>)
pLyGo-Kp-2	pPIC9K-Amy ^{SP} - <i>ccdB</i>	Vector encoding the α -amylase signal peptide and the LyGo cassette, Amp ^R (<i>E. coli</i>), Km ^R , <i>HIS4</i> (<i>K. phaffii</i>)

140

141 Exploring protein localization strategies in *E. coli* using LyGo

142 *E. coli* is an important model organism in molecular biology, and a popular expression host for bacterial
143 LPMOs¹⁶. Proteins can be produced in different cellular compartments in *E. coli*, each strategy with its
144 own advantages and disadvantages. In order to explore different approaches, we developed a range of
145 vectors for producing LPMOs in three different cellular compartments of *E. coli*: The cytoplasm, the
146 periplasm, and attached to the surface of the cell. The LyGo vectors were tested with four different
147 LPMOs: *Thermobifida fusca* TfLPMO10A²², *Streptomyces coelicolor* ScLPMO10B²³, BatLPMO10 from *Bacillus*
148 *atrophaeus*²⁴, and LsAA9A from *Lentinus similis*²⁵. TfLPMO10A, ScLPMO10B, and BatLPMO10 were
149 chosen because they previously were produced successfully in *E. coli*^{22,23,24}. LsAA9A was chosen as an
150 example of a fungal enzyme and because it had the added benefit of being compatible with a simple
151 chromogenic assay (based on AZCL-HEC) available²⁶. All sequences were codon optimized for *E. coli*,

except for BatLPMO10 because the native sequence previously was expressed in *E. coli*²⁴ at relatively high titers.

Cytoplasmic expression

We based a pLyGo-*Ec* cytoplasmic vector (pLyGo-*Ec*-1) on the pET39b backbone encoding an N-terminal ubiquitin solubility tag, a poly His-tag, and a TEV cleavage site²⁷, which together has the potential to facilitate soluble expression, purification, and cleavage to release the N-terminal histidine (Supplementary Figure S2A).

The four different LPMO genes were inserted into pLyGo-*Ec*-1, transformed into *E. coli* BL21(DE3), and expression induced in late exponential phase by the addition of IPTG followed by incubation for 20 hours at 18 °C with shaking. Since TfLPMO10A, ScLPMO10B, and LsAA9A all contain disulfide bonds, we expressed them in parallel in the disulfide-bond enhancing strain *E. coli* SHuffle²⁸. Using this setup, we observed that the different LPMOs were produced at titers ranging from approximately 100 to 800 mg/L (Figure 2B and Supplementary Figure S3), and that the disulfide containing LPMOs expressed considerably better in the SHuffle strain. However, the SHuffle strain grew to lower density, which reduced the difference in volumetric titers produced by the two strains.

It was possible to isolate all LPMOs by immobilized metal affinity chromatography (IMAC), but only TfLPMO10A and ScLPMO10B showed the expected fragmentation pattern after treatment with TEV protease (Supplementary Figure S4). The processed TfLPMO10A and ScLPMO10B were separated from the ubiquitin tag and the TEV protease by reverse IMAC purification (Supplementary Figure S4). Afterwards the enzymes were copper-loaded overnight and incubated with phosphoric acid swollen

cellulose (PASC) and ascorbate for 25 hours and analyzed using high performance liquid chromatography with pulsed amperometric detection (HPLC-PAD). The assay showed a clear appearance of peaks corresponding to soluble oligosaccharides as observed previously for active LPMOs²⁹. Furthermore, these peaks were absent when incubating the substrate with CuCl₂ or with the non-catalytic Cu(II) chaperone CopC from *Pseudomonas fluorescens* SBW25 (PfCopC)^{30,31} (Figure 2C). This demonstrates cellulolytic activity from both LPMO-containing samples. Correct processing was confirmed by quantitative amino acid analysis (data not shown). Together, these results show that this expression system constitutes a viable strategy for producing active LPMOs in the cytoplasm of *E. coli*, although the efficiency of the crucial TEV cleavage step varies depending on the specific LPMO.

Periplasmic expression

Different signal peptides are known to influence translocation and folding kinetics of different proteins in an unpredictable manner^{32,33}. Thus, screening of several different signal peptides is often key for optimizing expression and secretion into the periplasm. We based LyGo vectors for periplasmic expression (pLyGo-*Ec*-2 to pLyGo-*Ec*-6) on a pET28a(+) vector³⁴ harboring sequences encoding the signal peptides MalE^{SP}, OmpA^{SP}, PhoA^{SP}, and PelB^{SP} that target proteins to the periplasm post-translationally via the Sec-pathway. In addition, we included the DsbA^{SP} signal peptide as a representative of the co-translational SRP-pathway, as it previously has been shown to increase secretion of some proteins^{35,36}. In an attempt to ensure high production titers, the translation initiation regions (TIRs) embedded in the coding sequence of the signal peptides, were replaced with the optimized versions described by Mirzadeh et al.³⁷ (Supplementary Figure S2B).

The four LPMO constructs were transformed into BL21(DE3) and expression induced in late exponential phase by the addition of 1 mM IPTG and incubated for 20 hours at 18 °C with shaking, except for BatLPMO10 because pilot experiments showed improved expression at 30 °C. All the LPMOs were successfully expressed with titers of at least 40 mg/L culture (Figure 2E, 2F, and Supplementary Figure S5 to S8). The different signal peptides varied in performance for the different LPMOs, although PelB^{SP} showed the best overall performance. The periplasmic fractions collected from the expression of LsAA9A were subjected to AZCL-HEC activity assay (Figure 2F), which showed that all constructs produced functional versions of the enzyme with a good correlation between protein titer and enzymatic activity. This provides a demonstration that a fungal LPMO can be produced in a functional state in a bacterium using the LyGo platform.

Surface display and shaving

Surface display systems are useful tools for variant screening³⁸ and whole cell catalysis³⁹. Furthermore, controlled release from the membrane, as presented by Ahan et al.⁴⁰, could be used as a method for selectively purifying the protein of interest. These concepts were materialized into two different designs: A C-terminal and an N-terminal fusion construct, which were both tested using LsAA9A. The pLyGo-*Ec* surface display vectors were based on a pBAD expression vector, which had previously been used for surface display⁴¹.

The C-terminal construct (pLyGo-*Ec*-7) uses the C-terminal translocation unit of the *Neisseria gonorrhoeae* autotransporter IgA protease (C-IgAP) and a single domain antibody (nanobody, NB), and should display active LPMOs (i.e. with the N-terminal histidine exposed) on the surface by

215 fusing the NB to the C-terminus of the LPMO⁴¹. The C-terminal tag is obviously not compatible with the
 216 stop codon overhang in the standard LyGo design, so this overhang was changed to the first codon of a
 217 TEV site (glutamate, GAA) (Supplementary Figure S2C). The LsAA9A-encoding DNA sequence was
 218 cloned into this vector and transformed into BL21(DE3). Expression was induced with L-rhamnose in
 219 mid-exponential phase and the LPMO produced for 24 hours followed by functional AZCL-HEC assays
 220 in both liquid culture and agar plate format (Figure 2H). In both cases, activity was observed from the
 221 surface displayed LsAA9A, confirming that this design has the potential to be used for variant screening
 222 and whole-cell catalysis.

223 The N-terminal construct (pLyGo-*Ec*-8) uses a hybrid protein, consisting of the Lpp signal
 224 peptide and residues 66 to 180 of the outer membrane protein OmpA, and displays the LPMO as a fusion
 225 between the LPMO N-terminus and the OmpA C-terminus⁴¹. Furthermore, a sequence encoding a TEV
 226 site was inserted between the LyGo cassette and the Lpp^{SP}-OmpA coding sequence, to allow for the
 227 release of the LPMO and the N-terminal histidine from the fusion protein upon treatment with TEV
 228 protease (Supplementary Figure S2C). LsAA9A was cloned into the construct, expressed (as described
 229 above), and released into the medium by addition of TEV protease. Following a centrifugation step, the
 230 supernatant was subjected to the AZCL-HEC assay. A small but significant increase in LPMO activity
 231 was observed indicating some successful cleavage and release of the surface displayed LsAA9A (Figure
 232 I top). However, due to the low signal observed in the assay, we speculate that TEV cleavage of LsAA9A
 233 is inefficient, as observed for LsAA9A TEV fusions expressed in the cytoplasm. To more convincingly
 234 demonstrate the surface “shaving”, we therefore decided to display and cleave a His-tagged version of
 235 ScLPMO10B with the same approach. Cleavage was induced by different concentrations of purified TEV

236 protease (Figure 2I). Immunoblotting of the cleavage reaction mix with an anti-His tag antibody (α -His)
237 showed that the fraction of anchored ScLPMO10B decreased while the fraction of free ScLPMO10B
238 increased with increasing TEV concentrations. Together, this shows that LyGo vectors can facilitate
239 functional surface display of LPMOs and that these can be released from the surface in a simplistic
240 purification procedure.

241

242 **Figure 2:** LPMO production in different compartments in *E. coli* using LyGo. EV = Empty vector, N.D. =
 243 Not determined. (A) Illustration of LyGo cloning for production of LPMOs in the cytoplasm of *E. coli*. (B)
 244 Results from cytoplasmic expression of four different LPMOs (see main text for further detail) in
 245 BL21(DE3) and the SHuffle strain. Protein production normalized by cell density was analyzed by

246 densitometry of InstantBlue stained protein bands on an SDS-PAGE gel from biological duplicates
 247 (Supplementary Figure S3). A representative SDS-PAGE is shown in the lower panel. **(C)** Activities of
 248 TEV-cleaved LPMOs on PASC using HPLC-PAD, compared with the catalytically inactive control
 249 samples of PfCopC or 0.75 μ M CuCl₂. **(D)** Illustration of LyGo cloning for production of LPMOs in the
 250 periplasm with a range of different signal peptides. **(E)** Results from periplasmic expressions of
 251 BatLPMO10, TflLPMO10A, and ScLPMO10B. Production titers were quantified by densitometry of
 252 InstantBlue stained protein bands on an SDS-PAGE gel from biological duplicates (Supplementary
 253 Figure S5 to S7). A representative SDS-PAGE is shown in the bottom. **(F)** LsAA9A produced in the
 254 periplasm of *E. coli* with different signal peptides quantified by densitometry of InstantBlue stained
 255 protein bands on an SDS-PAGE gel from biological duplicates (Supplementary Figure S8), and activity
 256 of LsAA9A extracts quantified by the AZCL-HEC assay from biological duplicates. A representative SDS-
 257 PAGE from each construct is shown in the bottom. **(G)** Illustration of LyGo cloning for surface display
 258 of LPMOs with two different strategies. **(H)** Activity of the NB-C-IgAP surface displayed LsAA9A in
 259 liquid culture (left) and directly on LB agar plates using AZCL-HEC as a substrate (right). **(I)** LsAA9A
 260 expressed on the surface of *E. coli* using the N-terminal construct (pLyGo-Ec-8). Activity of the cleaved
 261 LsAA9A was analyzed by the AZCL-HEC assay (upper half). Western blot using an anti-His tag antibody
 262 for a His-tagged ScLPMO10B expressed on the surface of *E. coli* (bottom half). TEV protease was titrated
 263 to cleave the surface displayed protein.

264

265 **Exploring the effects of signal peptides and DNA sequence variants in *B. subtilis***

266 Bacteria in the *Bacillus* genus are industrially relevant, due to desirable traits such as high capacity for
 267 secreting proteins, ability to grow on cheap carbon sources, GRAS status, and robustness in industrial
 268 settings^{42,43,44}. *B. subtilis* is widely used in both industry and academia, and the recent advances in
 269 standardizing genetic parts and strain libraries^{45,46,47,48}, makes it an attractive protein production host.

270 LyGo plasmids for *B. subtilis* (pLyGo-Bs-1 to pLyGo-Bs-4) were constructed based on the
 271 integrative plasmid pBS293C-amyE from the *Bacillus* SEVA sibling collection⁴⁵. Two SapI sites were
 272 removed from the backbone, and the LyGo cassette was inserted downstream of a triple promoter P_{amyL}-
 273 P_{amyQ}-P_{cryIIIA}-cryIIIA_{stab}⁴⁹ and a predicted strong ribosome binding site (R0 from Guiziou et al.⁴⁸). Like *E.*
 274 *coli*, signal peptide performance in *B. subtilis* is difficult to predict⁵⁰, so we decided to generate four
 275 different plasmids with different signal peptides: AmyL^{SP} from *Bacillus licheniformis*, AmyQ^{SP} from
 276 *Bacillus amyloliquefaciens*, AprE^{SP} from *Bacillus clausii*, and BatLPMO10^{SP} from *Bacillus atrophaeus*. Initially,
 277 the plasmids were constructed similarly to the *E. coli* counterparts, but while assembling these we
 278 observed an unusual high number of frame-shift mutations in *ccdB*, despite the use of a *ccdB*-tolerant *E.*
 279 *coli* DB3.1. We hypothesized that this could be a result of toxicity caused by increased expression of *ccdB*
 280 from P_{3P} or other upstream sequences in the cloning host. In order to mitigate this, a stop codon was
 281 introduced in-frame with the signal peptide after the upstream SapI site, followed by a terminator from
 282 the BioBrick collection (BBa_B1002)⁵¹ (Supplementary Figure S2D). This updated design allowed for
 283 assembly of constructs with an intact *ccdB* gene in the LyGo cassette.

284 In addition to screening different signal peptides for, the pLyGo-Bs vectors were utilized
 285 to assess the expression of two different sequence variants of each of the sequences encoding BatLPMO10
 286 and LsAA9A (Figure 3A): The native sequence and a sequence codon-optimized for *B. subtilis*. The

287 resulting plasmids were transformed into the extracellular protease- and sporulation deficient *B. subtilis*
288 KO7-S strain, thereby integrating the expression cassettes into the *amyE* locus by homologous
289 recombination. The cells were grown in Cal18-2 media for 72 hours at 20 °C, the cells were harvested by
290 centrifugation, and the supernatants analyzed by SDS-PAGE. BatLPMO10 produced at concentrations
291 ranging from approximately 0.5 to 2 g/L in the supernatant (Figure 3B and Supplementary Figure S9).
292 The signal peptides caused significant variation in expression levels, with the native BatLPMO10 signal
293 peptide producing the highest amount of protein. The presence of LsAA9A in samples was barely
294 detectable on SDS-PAGE gels stained by InstantBlue (data not shown). Thus, a His-tag was added to the
295 sequences allowing for sensitive detection by Western blotting. Additionally, the activity of LsAA9A was
296 assayed using the AZCL-HEC assay. The codon optimized variant was expressed at higher levels than
297 the native sequence, roughly showing double the yield and activity compared to the native sequence
298 (Figure 3C). The different signal peptides only yielded minor differences in activity, except for AprE^{SP}
299 that performed poorly. Together, these results demonstrate that the pLyGo vectors can be utilized to
300 produce and optimize secretion of functional LPMOs from both fungal and bacterial origin in *B. subtilis*
301 and that industrially relevant (g/L) levels can be reached.

302

303 **Figure 3:** Exploration of sequence and signal peptide variants in *B. subtilis* using LyGo. EV = Empty
 304 vector, N.D. = Not determined. **(A)** Schematic overview of the combinatorial cloning strategy: Four
 305 different LyGo fragments encoding two sequence variants of two different LPMOs were cloned into four
 306 LyGo vectors encoding different signal peptides, resulting in a total of 16 expression vectors. **(B)** Analysis
 307 of expression of BatLPMO10 variants estimated by densitometry of InstantBlue stained protein bands on
 308 an SDS-PAGE gel from two biological replicates (Supplementary Figure S9). A representative gel picture

is shown below. (C) Activity of LsAA9A was quantified by AZCL-HEC assay from two biological replicates. A representative Western-blot of each variant is shown below.

Exploring the effects of signal peptides and sequence variants in *K. phaffii*

The methylotrophic yeast *K. phaffii* is the most frequently used yeast species for production of recombinant proteins⁵². This is due to its ability to grow to high cell density, to express recombinant genes in a tightly controlled manner, and to efficiently secrete proteins^{52,53,54,55}. Despite its biotechnological importance and wide use in industry, relatively few genetic tools are readily available. However, *K. phaffii* has proven to be a promising production host for carbohydrate active enzymes like LPMOs^{16,56,57,58,59} and could be a valuable addition to the LyGo platform.

LyGo vectors for *K. phaffii* (pLyGo-Kp-1 and pLyGo-Kp-2) were created based on the pPIC9K plasmid hosting the methanol-inducible P_{AOX1} promoter. A multiple cloning site downstream of the α -mating factor pre-sequence (α -MF^{SP}) from *Saccharomyces cerevisiae* was replaced by the LyGo cassette and a SapI site was removed from the backbone by site-directed mutagenesis. A second vector (pLyGo-Kp-2) was designed, where α -MF^{SP} was replaced with the α -amylase signal peptide (Amy^{SP}) from *Aspergillus niger* (PichiaPink™ Secretion Signal Kit, Thermo Fisher) (Supplementary Figure 2E).

Two sequence variants of the gene encoding LsAA9A were cloned into the two pLyGo-Kp vectors: the native sequence and a sequence variant codon optimized for *K. phaffii*. The resulting plasmids were linearized and integrated into *K. phaffii* GS115⁶⁰. This type of integration results in two distinct phenotypes: Mut⁺ and Mut^S, which have shown to impact recombinant protein production⁶¹. In total, this leads to eight different strain variants which were tested for activity on the AZCL-HEC substrate (Figure

4A). Protein production was performed over 4 days at 28 °C with daily addition of methanol, before the cells were harvested and the supernatants were analyzed by SDS-PAGE. LsAA9A was produced at concentrations up to approximately 0.3 g/L (Figure 4B). The expression was largely unaffected by codon variation and strain genotype, and a trend was that the α -MF^{SP} signal peptide produced at higher levels than Amy^{SP}. Protein amounts were found to correlate poorly with the cellulolytic activity measured by the AZCL-HEC assay (Figure 4B): despite the relatively low production titers from the Mut⁺ strain harboring LsAA9A with the Amy^{SP}, the sample showed the highest overall activity. In summary, the *K. phaffii* LyGo platform expands the possibilities of the LyGo platform to include screening of expression strategies in an eukaryotic host.

Figure 4: Production and activity of LsAA9 with sequence and signal peptide variants in *K. phaffii* using LyGo. EV = Empty vector, N.D. = Not determined. **(A)** Schematic overview of the combinatorial cloning strategy: LyGo fragments encoding the native sequence and a codon variant of LsAA9A were cloned into two LyGo vectors encoding different signal peptides and the resulting constructs integrated in *K. phaffii* resulting in Mut⁺ and Mut^s phenotypes. **(B)** Expression was analyzed by densitometry of InstantBlue

stained protein bands on an SDS-PAGE gel, and activity quantified by AZCL-HEC assay from two biological replicates (Supplementary Figure S10). A representative gel picture is shown in the lower panel.

Discussion

Progress in the LPMO field ultimately depends on the availability and quality of the enzyme samples used in different laboratories. Thus, failed or low titer expression is a limiting factor for advancement in the field. Typically, most research groups will work with a limited number of expression strategies, and may sometimes fail in expressing a specific gene or accept low-titer samples. The LyGo platform provides easy access to a range of expression strategies that can be explored in parallel in order to increase the likelihood of high-titer and high-quality expression. Furthermore, standardization of the cloning and expression workflows may improve sample uniformity and comparability of results across different laboratories. In order to facilitate the use of LyGo, the vectors are made available through plasmid repositories such as Addgene, and detailed protocols are available at the online platform protocols.io (Supplementary figure S11).

In this work we demonstrate that LyGo can be used for identifying an optimal expression context for different LPMOs in three different organisms. For *E. coli*, we developed expression vectors for cytoplasmic, periplasmic, and surface expression. Cytoplasmic expression is the most common strategy for the production of heterologous proteins in *E. coli*⁶², and is often advantageous over other expression strategies as it mitigates the need for screening signal peptides and has the potential to result in product yields exceeding 50% of the total cellular protein amounts⁶³. Furthermore, cytoplasmic

expression supports the use of several genetic tools such as co-expression of chaperones⁶⁴ and folding sensors⁶⁵, as demonstrated in this work using the *E. coli* SHuffle strain. However, cytoplasmic expression is not the most obvious strategy for production of LPMOs for two reasons: (1) a functional LPMO requires a free N-terminal histidine residue¹⁹, which is made impossible by the start-codon-encoded methionine in standard cytoplasmic expression strategies. (2) LPMOs frequently contain disulfide bonds⁶⁶, which are not readily formed in the cytoplasm of *E. coli* due to the reducing environment⁶⁷.

We mitigated the first issue by inserting an N-terminal tag, which can be removed by TEV cleavage leaving an N-terminal histidine residue and demonstrated how this results in an untagged and functionally active LPMO. Furthermore, an N-terminal tag has several potential advantages: It can enhance expression and solubility of the fusion protein, and it can allow for purification by IMAC purification for both initial separation prior to TEV cleavage and for subsequent removal of the tag and the TEV protease. Finally, it could quench unwanted oxidations within the production host, which may lead to production of reactive oxygen species⁶⁸. Similar methods have been used to express a range of LPMOs^{69,70,71,72}, although none of these methods take advantage of a solubility tag. Unfortunately, we did not observe any TEV cleavage for BatLPMO10 and LsAA9A, potentially due to steric hindrance at the Ubiquitin/LPMO interface. Thus, this strategy is not a one-size-fits-all strategy for LPMOs. On the other hand, we were able to express and effectively TEV-cleave TflLPMO10A and ScLPMO10A at significantly increased titers compared to the same LPMOs expressed by the periplasmic expression strategy, demonstrating that this is an attractive strategy for preparing high titer samples of LPMOs.

In order to facilitate formation of disulfides, we expressed LPMOs in the *E. coli* SHuffle strain, which enhances the formation of disulfide bonds in the cytoplasm²⁸. This significantly improved

the overall titer of TfLPMO10A, but had little effect on the titers of ScLPMO10B and LsAA9A due to the poorer growth of the Shuffle strain.

Periplasmic expression is a popular strategy for expressing LPMOs in *E. coli*¹⁶, as it naturally provides the means for N-terminal processing and disulfide bond formation. However, secretion has previously been shown to constitute a bottleneck resulting in low yields⁷³. Furthermore, this strategy often requires laborious screening of different signal peptides in order to optimize performance³⁷. In order to simplify such a workflow, we provide vectors for screening several signal peptides in parallel, thereby reducing the time and effort required for this type of optimization. Of the tested signal peptides, we found PelB^{SP} to consistently perform the best. It is unclear whether this is a prevalent property of PelB^{SP}, or an anecdotal observation due to the small number of tested LPMOs.

Production of proteins on the surface generally suffers from similar bottlenecks as periplasmic expression. A previous study has shown that different signal peptides affect the display efficiency of the N-terminal surface display construct⁴¹. In the specific case of LPMOs, while the N-terminal surface display construct ensures proper processing, the C-terminal surface display construct requires cleavage by TEV protease for the LPMO to be active. Unfortunately, in line with our observations from cytoplasmic production, TEV cleavage of LsAA9A was inefficient, providing a cautionary tale that the C-terminal surface display construct might not be suitable for all LPMOs. While TEV protease can be produced at low cost in *E. coli*, simultaneous production of a surface displayed TEV protease might present an attractive extension of the system.

For *B. subtilis* and *K. phaffii* we both explored the effect of signal peptides and sequence variants. As observed many times previously, the results from *B. subtilis* expression show that both signal

peptides and codon optimization efforts influence expression in a protein-dependent manner. This underlines the usefulness of a simple and efficient platform for exploration of expression strategies and importantly our pLyGo-Bs vectors showed capabilities to express functional LPMOs at industrially relevant titers.

The *K. phaffii* LPMO production from pLyGo vectors also showed variations in performance between signal peptides, codon variants, and phenotypes, but with complex interplay between these factors. Furthermore, the production titers and measured activities were found to correlate poorly. One explanation could be the promiscuity in the processing of the α -MF^{SP} signal peptide^{74,75}, leading to inefficient release of the N-terminal histidine, but it is also possible that the high concentration of copper sulphate used in the AZCL-based assay was suboptimal for a proper quantitative readout.

In addition to expressing LPMOs at different titers, different hosts can also give rise to different protein modifications (such as methylation and glycosylation⁷⁶). The choice of organism(s) can therefore be a strategy to optimize protein stability, solubility or activity⁷⁷ - or as a means to study the importance of these modifications. Some fungal LPMOs have been found to be selectively methylated at the N-terminal histidine⁹. This modification has been suggested to protect the enzymes from oxidative inactivation⁷⁶, although this is still actively being investigated. The LyGo platform currently supports the non-glycosylating and non-methylating hosts *E. coli* and *B. subtilis*, and the glycosylating and non-methylating host *K. phaffii*. A methylating host would therefore be a valuable future addition to the platform, as LyGo then would cover the most important modifications for LPMOs.

It is our hope that the community will adopt and expand the LyGo platform with vectors and protocols for additional expression strategies. For this, careful design of the LyGo vectors is

necessary. During construction of the pLyGo vectors, we learned that building constructs with *ccdB* is sensitive to the context of the LyGo cassette – possibly due to increased expression of *ccdB* facilitated by upstream sequences. In one case, this was solved by inserting an in-frame stop codon and a terminator in the pLyGo-*Bs* vectors (Supplementary Figure S2). Future additions to this or similar vector collections should consider this potential issue early in the design phase.

Materials and Methods

Strains

The strains used in this study are listed in Supplementary Table 1. *E. coli* DB3.1 (Invitrogen, Carlsbad, CA, USA) was used to construct and propagate *ccdB*-containing plasmids. *E. coli* NEB5 α (New England Biolabs, Ipswich, MA, USA) was used to construct and propagate of LPMO-containing constructs. Expression experiments were performed with *E. coli* BL21(DE3) (Novagen, Merck KGaA, Darmstadt, Germany), *B. subtilis* KO7-S (The Bacillus Genetic Stock Center, Columbus, OH, USA), and *K. phaffii* GS115 (Thermo Fisher Scientific, Waltham, MA, USA). Bacterial cells were routinely cultivated in lysogeny broth (LB) at 37 °C with 250 RPM of shaking or on LB agar plates at 37 °C. When appropriate, the media was supplemented with kanamycin (50 μ g/mL) for *E. coli* and chloramphenicol (5 μ g/mL) for *B. subtilis*. Competent *E. coli* cells were obtained from (Invitrogen, California, United States) and transformed following standard protocols. Competent *B. subtilis* cells were obtained as described elsewhere⁷⁸, although without adding histidine to the SM1 and SM2 media and with at least 2 hours of recovery. Competent *K. phaffii* cells were obtained using the Pichia EasyComp Kit (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA).

450

451 **Vector construction**

452 DNA manipulations were performed using uracil excision cloning as described elsewhere⁷⁹. The
453 plasmids used and constructed in this work are listed in Table 1 and Supplementary Table 2. PCRs were
454 performed using Phusion U Hot Start polymerase (Thermo Fisher Scientific, Waltham, MA, USA),
455 according to manufacturer's instructions. Primers were ordered from Integrated DNA technologies (IDT,
456 Coralville, IA, USA). The primers used in this work are listed in Supplementary Table 3. Plasmids
457 sequences were confirmed by sequencing (Eurofins MWG operon, Germany).

458

459 **LyGo cloning**

460 Inserts were either synthesized as gBlocks (IDT, Coralville, IA, USA) or Strings (Thermo Fisher Scientific,
461 Waltham, MA, USA), or amplified by PCR using Phusion Hot Start II DNA polymerase (Thermo Fisher
462 Scientific, Waltham, MA, USA), according to manufacturer's instructions. The sequence of the codon
463 variants were obtained using IDT's Codon Optimization Tool or Thermo Fisher's GeneOptimizer Tool⁸⁰
464 LyGo cloning was performed by mixing vector and insert in molar ratios between 1:3 to 1:6, together
465 with 0.5 µL FastDigest LguI (SapI) (Thermo Fisher Scientific, Waltham, MA, USA), 1 µL FD buffer
466 (Thermo Fisher Scientific, Waltham, MA, USA), 1 µL T4 DNA Ligase (Thermo Fisher Scientific, Waltham,
467 MA, USA), and 1 µL T4 DNA Ligase buffer (Thermo Fisher Scientific, Waltham, MA, USA). The reaction
468 volume was adjusted to 10 µL using MilliQ water, incubated at room temperature for at least 15 minutes,
469 and subsequently used for transformation.

470

471 **Expression in *E. coli***

472 The expression of LPMOs in the periplasm and cytoplasm of *E. coli* performed similarly as in Hemsworth
 473 et al.,⁸¹ a single colony of the desired strain was inoculated in 50 or 25 mL of LB medium supplemented
 474 with kanamycin and grown O/N at 37 °C with 250 RPM. The O/N culture was back-diluted 1:100 in 50
 475 mL of fresh LB medium with kanamycin and incubated at 37 °C with 250 RPM. When an OD₆₀₀ value
 476 between 0.5 and 0.6 was reached, the cultures were transferred to an 18 °C incubator at 180 RPM (unless
 477 otherwise stated) and grown until an OD₆₀₀ value between 0.8 and 1. At this point, the expression was
 478 induced with 1 mM IPTG and the cultures were grown for 20 hours at 18 °C incubator at 180 RPM (unless
 479 otherwise stated). Cultures were normalized to equal OD₆₀₀ values corresponding to 25 OD units (ODU)
 480 using LB medium, spun down at 8000 g for 20 minutes at 4 °C, and the supernatant was discarded.

481 For surface expression the cultures were treated similarly, although the culture was induced
 482 with 5 mM L-rhamnose once an OD₆₀₀ of 0.5 was reached. The culture was incubated at 30 °C at 250 RPM
 483 for 20 hours, and the cells directly plated on LB agar plates containing the AZCL-HEC substrate or
 484 harvested at 4000 g for 5 minutes, washed twice and resuspended in 10 mM Tris-HCl pH 7.5 (100 µL per
 485 ODU) for activity assay in liquid culture. For cleavage of the surface displayed protein between 1 and 25
 486 µL of purified TEV protease was added and the reaction incubated at room temperature overnight.
 487 Subsequently, cells were pelleted at 5000 g for 10 minutes and the supernatant was collected for analysis
 488 by SDS PAGE, western blot or activity assay.

489

490 **Whole-cell lysis**

491 For the screening of production conditions, extraction of the soluble fraction of the whole-cell lysate was
 492 done by chemical lysis using Cellytic™ B (Sigma-Aldrich, Saint Louis, MO, USA). Every ODU of cells
 493 was lysed using 16 µl of Cellytic™ B in 50 mM tris pH 8, 10mM imidazole, 150 mM NaCl buffer
 494 (supplemented with benzonase and lysozyme according to batch size). Following lysis, the insoluble
 495 fraction was pelleted by centrifugation at 25,000 g for 25 min at 4 °C and the supernatant stored at 4 °C
 496 until use. For the SDS-PAGE gels of whole-cell lysates, 0.046 ODU was loaded in each lane. For
 497 expressions in 500 mL culture volume for production of TflPMO10A and ScLPMO10B for activity
 498 testing, cells were lysed by freeze-thawing. The pelleted cultures were frozen at –80 °C for an hour before
 499 the cells were thawed at RT and resuspended in 50 mL lysis buffer (50 mM tris pH 8, 10mM imidazole,
 500 150 mM NaCl buffer). The lysis was now achieved by two cycles of at least 1 hour at -80 °C followed by
 501 thawing at room temperature. Following lysis, the insoluble fraction was pelleted by centrifugation at
 502 15,000 g for 25 min at 4 °C and the supernatant stored at 4 °C until use.

503

504 **Periplasmic protein extraction**

505 The periplasmic fraction was isolated following a modified version of an existing protocol⁸² by adding
 506 12 µl of TSE buffer (200mM Tris-HCl pH 8, 500mM sucrose, 1mM EDTA) per ODU collected. The pellet
 507 was resuspended by mild pipetting, and the mixture was incubated at room temperature for 10 minutes.
 508 Afterwards, the cultures were subjected to a cold shock by adding 12 µl of ice-cold water per ODU.
 509 Finally, the mixture was subjected to a centrifugation step of 8000 g for 20 minutes at 4 °C and the pellet
 510 was discarded. Supernatants containing the periplasmic fraction were kept at 4 °C for enzymatic assays.
 511 Following this modified protocol, we were able to obtain substantially cleaner and purer periplasmic

fractions in which lysozyme was not added to the buffer mixture (Supplementary Figure S12). 0.16 ODU was loaded in each lane of the gels for SDS-PAGE analysis.

IMAC and reverse IMAC purification

Initial small test purifications were performed using Ni-NTA Spin Column for His-Tagged proteins (Qiagen, Hilden, Germany) following the protocol for 6xHis-Tagged Proteins under Native Conditions from *E. coli* Cell Lysates. Large batch purifications were performed using Ni-NTA Superflow (Qiagen, Hilden, Germany). A volume of lysate equivalent to 10-20 mg of the target protein was mixed with lysis buffer (50 mM Tris-HCl pH 8, 10 mM imidazole, 150 mM NaCl) to a final volume of 16 mL. This mixture was added to 2 mL of equilibrated Ni-NTA Superflow beads and incubated at 4 °C for 25 min with slow stirring. The flow-through was discarded, and the sample was washed twice with 15 mL wash buffer (50 mM Tris-HCl pH 8, 20 mM imidazole, 150 mM NaCl). Finally, the target fusion protein was eluted from the column with 5 mL elution buffer (50 mM Tris-HCl pH 8, 300 mM imidazole, 150 mM NaCl). The total volume of lysate from where the protein was purified was 9, 12, and 48 mL for PfCopC, TfLPMO10A, and ScLPMO10B, respectively. After TEV-cleavage, EDTA was removed by gel filtration on a 220 mL Sephadex-G25 (Cytiva, São Paulo, Brazil) column setup to an ÄKTA pure chromatography system. 30 mL of TEV treated sample was loaded on the column, followed by elution with dialysis buffer (50 mM tris pH 8, 200 mM NaCl). The protein content of the sample was collected between 70 to 110 mL retention volume. Afterwards, the target protein was separated from the His-tagged ubiquitin and TEV protease by running the sample over 5 mL equilibrated Ni-NTA Superflow beads and collecting the flow-through. Effective removal of the Ub(his10)-tag and TEV protease was ensured by SDS-PAGE (Supplementary

Figure S4), and confirmed by amino acid analysis as described by Barkholt et al.⁸³ Subsequently, the samples were concentrated using Amicon ultra spin filters (Sigma-Aldrich, Saint Louis, MO, USA) (10.000 MWC for TFLPMO10A and ScLPMO10B and 5.000 MWC for PfCopC).

Production and purification of TEV protease

pTEVprotease was transformed into *E. coli* BL21(DE3) cells. A fresh colony was inoculated into LB medium supplemented with ampicillin at 37 °C with 250 RPM shaking overnight. The next day, the culture was diluted to an OD₆₀₀ of 0.1 and incubated at 37 °C with 250 RPM shaking until an OD₆₀₀ of 0.9 at which point the cells were induced with 1 mM IPTG and incubated overnight at 37 °C. The next day, cultures were collected by centrifugation and the pellets resuspended in lysis buffer (50 mM Tris-HCl, 200 mM NaCl, 10 mM imidazole, pH 7.8). The resuspended pellets were subjected to sonication in order to obtain the protein fraction (10 intervals of sonication for 10 seconds at the time with 30 seconds pause in between) while maintaining the cell lysates on ice. Afterwards, the sonicated slurry was centrifuged for 1 hour at 15,000 g at 4 °C and the supernatant containing the protein fraction was kept at 4 °C until purification. TEV protease purification was performed by IMAC purification. The resulted in purified TEV protease fractions which were loaded on SDS-PAGE to identify the fractions. The TEV protease was buffer exchanged (50 mM Tris pH 8, 150 mM NaCl) and concentrated in Amicon ultra spin filters (Sigma-Aldrich, Saint Louis, MO, USA) (10.000 MWC) to an ABS₂₈₀ ≥ 3.0, before activity was confirmed and the sample stored as aliquots with 50% glycerol at -80 °C.

Cleavage with TEV protease

554 Immediately before the cleavage, the TEV protease was reduced with 10 mM of 1,4-dithiothreitol (DTT,
555 Sigma-Aldrich, Saint Louis, MO, USA) at room temperature for 30 minutes. For the cytoplasmically
556 produced proteins, TEV protease was mixed with the target fusion protein in a 1:10 ratio based on Abs₂₈₀.
557 Before adding the protease to the reaction mixture, 2 mM EDTA was added to the reaction. The reaction
558 was incubated at 16 °C overnight. For the proteins expressed on the surface 5 µL of TEV protease were
559 mixed with 1 ODU of cell suspension (unless otherwise stated). The reaction was incubated at RT
560 overnight.

561

562 **HPLC-PAD-based cellulose degradation assay**

563 Proteins were preloaded at 4 °C overnight with CuCl₂ at 1:0.75 stoichiometry. The assay was performed
564 in 200 µL samples containing 50 mM citrate phosphate buffer pH 7.25, 0.25 % phosphoric acid swollen
565 cellulose (PASC), 1 mM ascorbic acid and 0.75 µM copper-loaded enzyme. The samples were prepared
566 in microtiter plates in technical triplicates. Samples were incubated for 25 hours in a thermomixer at 50
567 °C with 750 RPM shaking. Samples were filtered and analyzed by HPLC-PAD as described
568 by Westereng et al.⁸⁴.

569

570 **AZCL-HEC-based cellulose degradation assay**

571 Activity of LsAA9A samples was measured by mixing a reaction solution consisting of 1 mg/mL AZCL-
572 HEC substrate (Megazyme, County Wicklow, Bray, Ireland), 2 mM Ascorbic acid (Sigma-Aldrich, Saint
573 Louis, MO, USA), 200 µM Copper sulphate (Sigma-Aldrich, Saint Louis, MO, USA), and adjusted to the
574 desired volume with 100 mM Sodium acetate (pH 5) (Sigma-Aldrich, Saint Louis, MO, USA). 100 µL

sample was mixed with 400 μ L reaction mix and incubated at 50 °C with 1500 RPM shaking until a blue color had developed in the samples with LsAA9A. The reactions were centrifuged at 17000 g for 5 min and the absorbance of the supernatant was measured at 590 nm. Controls with the buffer or media the enzymes were suspended in were included, and the absorbance values of these samples were subtracted from the other measurements. For plate experiments, a 2x LB agar solution was mixed in a 1:1 ratio with a 2x reaction buffer (final concentration as described above) and poured very thinly to assure that the AZCL-HEC substrate was accessible from the surface.

Expression in *B. subtilis*

B. subtilis transformants were streaked on LB agar plates supplemented with 1% starch, incubated at 37 °C for 24 hours, and a strain which had not produced a visible halo (indicating a correct double crossover event) was selected. A single colony was inoculated in 3 mL of LB medium supplemented with chloramphenicol and grown O/N at 37 °C with 250 RPM shaking in an Axygen 24-deepwell plate (Corning Life Science, Corning, New York, USA). The O/N culture was back-diluted 1:300 in 3 mL fresh Cal18-2 medium⁸⁵ (Glucidex 12 was exchanged for Maltodextrin DE 13-17 (Sigma-Aldrich, Saint Louis, MO, USA) supplemented with chloramphenicol also in an Axygen 24-deep well plate. The culture was grown for 72 hours at 20 °C with 250 RPM shaking. The cells were harvested by centrifugation at 6000 g for 5 minutes at 4 °C, and the pellet was discarded. The samples were stored at 4 °C before use. 10 μ L were loaded on gels for SDS page and western blotting.

Expression in *K. phaffii*

596 A single colony of the desired strain was inoculated in 25 mL pH 6 buffered Glycerol-complex (BMGY)
 597 medium (*Pichia* Expression Kit, Life Technologies, Carlsbad, CA, USA) and grown at 28 °C with 250 RPM
 598 shaking until saturation. The preculture was then diluted 1:40 into 100 mL of pH 6 buffered BMGY media
 599 and grown O/N at 28 °C with 250 RPM shaking. Cells were pelleted, and resuspended in Buffered
 600 Methanol-complex (BMMY) medium (*Pichia* Expression Kit, Life Technologies, Carlsbad, CA, USA) to a
 601 final OD₆₀₀ of 50. This suspension was back-diluted 1:50 in a 500 mL baffled shake flask containing 100
 602 mL BMMY medium. Expression was performed for 4 days at 28 °C with 250 RPM shaking. Induction
 603 was maintained by addition of 1% methanol to the cultures once a day. The cells were harvested by
 604 centrifugation at 4000 g for 20 minutes at 4 °C and the pellets were discarded. The samples were stored
 605 at 4 °C before use. 10 µL were loaded on gels for SDS page.

606

607 **SDS-PAGE analysis**

608 Enzyme samples were mixed 1:1 with sample buffer (8 M urea, 0,0105 % (w/v) bromophenol blue, 5 mM
 609 EDTA, 100mM Tris-HCl pH 6.8, 4 % (w/v) SDS, and 25% (v/v) glycerol), and heated to 98 °C for 10
 610 minutes. Specific volumes of samples (described previously) were loaded on a 4–20 % Mini-PROTEAN-
 611 TGX gel (BioRad, Hercules, CA, USA) and run at 180 V for 45 min. Gels were stained for at least 4 hours
 612 using InstantBlue Protein Stain (Expedeon Inc., San Diego, CA, USA), and destained O/N using
 613 demineralized water. Protein amounts were estimated by densitometry using the Fiji software⁸⁶ using a
 614 dilution series of the *Thermoascus aurantiacus* LPMO TaAA9A with a known concentrations (unpublished
 615 results). For *E. coli* expression, the protein titers were normalized by the culture volumes.

616

617 **Western blot analysis**

618 Enzyme samples were handled as described in “SDS-PAGE analysis”, although instead of staining the
 619 gel the proteins were transferred to a nitrocellulose membrane using an iBlot Dry Blotting System
 620 (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA) at 20V for 7 minutes. The membrane was
 621 blocked with 2 % skim milk at 4 °C for at least 24 hours, washed three times for 5 minutes each in TBS-T
 622 (20 mM Tris-HCl pH 7.6, 150 mM NaCl, 0.1 % (v/v) Tween-20), and the membrane was incubated with
 623 primary antibody (anti-his tag, 1:1000, Merck Millipore, Merck KGaA, Darmstadt, Germany) diluted in
 624 2 % skim milk for 1 hour. The washing steps were repeated, and the membrane was incubated with the
 625 secondary antibody (anti-Mouse-HRP IgG, 1:10,000, Sigma-Aldrich, Saint Louis, MO, USA) diluted in
 626 TBS-T for 1 hour. The washing steps were repeated again, and the protein:antibody complexes were
 627 visualized using Amersham ECL Prime Western Blotting Detection Reagent (GE Healthcare, Chicago,
 628 IL, USA).

630 **Acknowledgements**

631 We thank Leila Lo Leggio for her inputs on the production of LPMOs in *E. coli*, Kenneth Jensen for his
 632 inputs on the production of LPMOs in *B. subtilis*, and Radina Tokin for inputs on the AZCL-HEC assay.
 633 This work was supported by the Novo Nordisk Foundation (Grant no. NNF17SA0027704).

635 **Conflicts of interest**

636 The TIRs used in the vectors pLyGo-*Ec*-2 through pLyGo-*Ec*-6 are patent pending (2030038-0, 2030039-8
637 and 2030040-6). These patents are shared by CloneOpt AB and Xbrane Biosciences. DOD and MHHN are
638 shareholders in CloneOpt.

639

640 **References**

- 641 (1) Spekrijse, J., Lammens, T., Parisi, C., Ronzon, T., & Vis, M. (2019). Insights into the European Market of
642 Bio-Based Chemicals. *Publications Office of the European Union, Luxembourg*.
- 643 (2) Masson-Delmotte, V., Zhai, P., Pörtner, H. O., Roberts, D., Skea, J., Shukla, P. R., ... Waterfield, T. (2018).
644 Summary for Policymakers. In: Global Warming of 1.5°C. An IPCC Special Report on the impacts of
645 global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission
646 pathways, in the context of strengthening the global response to. In *IPCC - SR15*. Retrieved from
647 https://report.ipcc.ch/sr15/pdf/sr15_spm_final.pdf<http://www.ipcc.ch/report/sr15/>
- 648 (3) Hassan, S. S., Williams, G. A., & Jaiswal, A. K. (2019). Lignocellulosic Biorefineries in Europe: Current State
649 and Prospects. *Trends in Biotechnology*, 37(3), 231–234.
- 650 (4) Thompson, P. B. (2012). The agricultural ethics of biofuels: The food vs. fuel debate. *Agriculture*, 2(4), 339–
651 358.
- 652 (5) Ramos, J., & Valdivia, M. (2016). Benefits and perspectives on the use of biofuels. *Microbial Biotechnology*,
653 9(4), 436–440.
- 654 (6) Ahorsu, R., Medina, F., & Constantí, M. (2018). Significance and Challenges of Biomass as a Suitable
655 Feedstock for Bioenergy and Biochemical Production: A Review. *Energies*, 11(12), 3366.
- 656 (7) Vaaje-Kolstad, G., Westereng, B., Horn, S. J., Liu, Z., Zhai, H., Sørli, M., & Eijsink, V. G. H. (2010). An
657 Oxidative Enzyme Boosting the Enzymatic Conversion of Recalcitrant Polysaccharides. *Science*,

- 658 330(6001), 219–222.
- 659 (8) Harris, P. V., Welner, D., McFarland, K. C., Re, E., Navarro Poulsen, J. C., Brown, K., ... Lo Leggio, L. (2010).
660 Stimulation of lignocellulosic biomass hydrolysis by proteins of glycoside hydrolase family 61: Structure
661 and function of a large, enigmatic family. *Biochemistry*, 49(15), 3305–3316.
- 662 (9) Quinlan, R. J., Sweeney, M. D., Lo Leggio, L., Otten, H., Poulsen, J.-C. N., Johansen, K. S., ... Walton, P. H.
663 (2011). Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits
664 biomass components. *Proceedings of the National Academy of Sciences*, 108(37), 15079–15084.
- 665 (10) Karkehabadi, S., Hansson, H., Kim, S., Piens, K., Mitchinson, C., & Sandgren, M. (2008). The First Structure
666 of a Glycoside Hydrolase Family 61 Member, Cel61B from *Hypocrea jecorina*, at 1.6 Å Resolution. *J. Mol.*
667 *Biol.*, 6(383), 144–154.
- 668 (11) Villares, A., Moreau, C., Bennati-Granier, C., Garajova, S., Foucat, L., Falourd, X., ... Cathala, B. (2017).
669 Lytic polysaccharide monooxygenases disrupt the cellulose fibers structure. *Scientific Reports*, 7, 40262.
- 670 (12) Eibinger, M., Ganner, T., Bubner, P., Rošker, S., Kracher, D., Haltrich, D., ... Nidetzky, B. (2014). Cellulose
671 surface degradation by a lytic polysaccharide monooxygenase and its effect on cellulase hydrolytic
672 efficiency. *Journal of Biological Chemistry*, 289(52), 35929–35938.
- 673 (13) Müller, G., Várnai, A., Johansen, K. S., Eijsink, V. G. H., & Horn, S. J. (2015). Harnessing the potential of
674 LPMO-containing cellulase cocktails poses new demands on processing conditions. *Biotechnology for*
675 *Biofuels*, 8, 187.
- 676 (14) Morales-cruz, A., Amrine, K. C. H., Blanco-ulate, B., Lawrence, D. P., Travadon, R., Rolshausen, P. E., ...
677 Cantu, D. (2015). Distinctive expansion of gene families associated with plant cell wall degradation,
678 secondary metabolism, and nutrient uptake in the genomes of grapevine trunk pathogens. *BMC*
679 *Genomics*, 16, 469.

- 680 (15) Loose, J. S. M., Forsberg, Z., Fraaije, M. W., Eijsink, V. G. H., & Vaaje-Kolstad, G. (2014). A rapid
681 quantitative activity assay shows that the *Vibrio cholerae* colonization factor GbpA is an active lytic
682 polysaccharide monooxygenase. *FEBS Letters*, 588(18), 3435–3440.
- 683 (16) Gaber, Y., Rashad, B., Hussein, R., Dishisha, T., Várnai, A., Abdelgawad, M., ... Várnai, A. (2020).
684 Heterologous expression of lytic polysaccharide monooxygenases (LPMOs). *Biotechnology Advances*, 43,
685 107583.
- 686 (17) Courtade, G., Le, S. B., Sætrom, G. I., Brautaset, T., & Aachmann, F. L. (2017). A novel expression system
687 for lytic polysaccharide monooxygenases. *Carbohydrate Research*, 448, 212–219.
- 688 (18) Engler, C., Kandzia, R., & Marillonnet, S. (2008). A one pot, one step, precision cloning method with high
689 throughput capability. *PLoS ONE*, 3(11), e3647.
- 690 (19) Vaaje-Kolstad, G., Forsberg, Z., Loose, J. S., Bissaro, B., & Eijsink, V. G. (2017). Structural diversity of lytic
691 polysaccharide monooxygenases. *Current Opinion in Structural Biology*, 44, 67–76.
- 692 (20) Bernard, P. (1996). Positive selection of recombinant DNA by CcdB. *Biotechniques*, 21(2), 320–323.
- 693 (21) Froger, A., & Hall, J. E. (2007). Transformation of plasmid DNA into *E. coli* using the heat shock method.
694 *Journal of Visualized Experiments : JoVE*, (6), 253.
- 695 (22) Zerhusen, N. K., Alahuhta, M., Lunin, V. V., Himmel, M. E., Bomble, Y. J., Wilson, D. B., ... Wilson, D. B.
696 (2017). Structure of a *Thermobifida fusca* lytic polysaccharide monooxygenase and mutagenesis of key
697 residues. *Biotechnology for Biofuels*, 10, 243.
- 698 (23) Forsberg, Z., Mackenzie, A. K., Sorlie, M., Rohr, A. K., Helland, R., Arvai, A. S., ... Eijsink, V. G. H. (2014).
699 Structural and functional characterization of a conserved pair of bacterial cellulose-oxidizing lytic
700 polysaccharide monooxygenases. *Proceedings of the National Academy of Sciences*, 111(23), 8446–8451.
- 701 (24) Yu, M. J., Yoon, S. H., & Kim, Y. W. (2016). Overproduction and characterization of a lytic polysaccharide

monooxygenase in *Bacillus subtilis* using an assay based on ascorbate consumption. *Enzyme and Microbial Technology*, 93–94, 150–156.

(25) Frandsen, K. E. H., Simmons, T. J., Dupree, P., Poulsen, J. C. N., Hemsworth, G. R., Ciano, L., ... Walton, P. H. (2016). The molecular basis of polysaccharide cleavage by lytic polysaccharide monooxygenases. *Nature Chemical Biology*, 12(4), 298–303.

(26) Frandsen, K. E. H., Tovborg, M., Jørgensen, C. I., Spødsberg, N., Rosso, M. N., Hemsworth, G. R., ... Leggio, L. Lo. (2019). Insights into an unusual Auxiliary Activity 9 family member lacking the histidine brace motif of lytic polysaccharide monooxygenases. *Journal of Biological Chemistry*, 294(45), 17117–17130.

(27) Rogov, V. V., Rozenknop, A., Rogova, N. Y., Löhr, F., Tikole, S., Jaravine, V., ... Dötsch, V. (2012). A Universal Expression Tag for Structural and Functional Studies of Proteins. *ChemBioChem*, 13(7), 959–963.

(28) Lobstein, J., Emrich, C. A., Jeans, C., Faulkner, M., Riggs, P., & Berkmen, M. (2012). SHuffle, a novel *Escherichia coli* protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm. *Microbial Cell Factories*, 11, 56.

(29) Forsberg, Zarah, Bissaro, B., Gullesen, J., Dalhus, B., Vaaje-Kolstad, G., & Eijsink, V. G. H. (2018). Structural determinants of bacterial lytic polysaccharide monooxygenase functionality. *Journal of Biological Chemistry*, 293(4), 1397–1412.

(30) Wijekoon, C. J. K., Young, T. R., Wedd, A. G., & Xiao, Z. (2015). CopC protein from *Pseudomonas fluorescens* SBW25 features a conserved novel high-affinity Cu(II) binding site. *Inorganic Chemistry*, 54(6), 2950–2959.

(31) Udagedara, S. R., Wijekoon, C. J. K., Xiao, Z., Wedd, A. G., & Maher, M. J. (2019). The crystal structure of the CopC protein from *Pseudomonas fluorescens* reveals amended classifications for the CopC protein family. *Journal of Inorganic Biochemistry*, 195, 194–200.

- 724 (32) Freudl, R. (2018). Signal peptides for recombinant protein secretion in bacterial expression systems.
725 *Microbial Cell Factories*, 17, 52.
- 726 (33) Yang, Y., Li, J., Liu, X., Pan, X., Hou, J., Ran, C., & Zhou, Z. (2017). Improving extracellular production of
727 *Serratia marcescens* lytic polysaccharide monooxygenase CBP21 and *Aeromonas veronii* B565 chitinase
728 Chi92 in *Escherichia coli* and their synergism. *AMB Express*, 7, 170.
- 729 (34) Dubendorf, J. W., & Studier, F. W. (1991). Controlling basal expression in an inducible T7 expression
730 system by blocking the target T7 promoter with *lac* repressor. *Journal of Molecular Biology*, 219, 45–59.
- 731 (35) Zhang, W., Lu, J., Zhang, S., Liu, L., Pang, X., & Lv, J. (2018). Development an effective system to
732 expression recombinant protein in *E. coli* via comparison and optimization of signal peptides: Expression
733 of *Pseudomonas fluorescens* BJ-10 thermostable lipase as case study. *Microbial Cell Factories*, 17, 50.
- 734 (36) Soares, C. R. J., Gomide, F. I. C., Ueda, E. K. M., & Bartolini, P. (2003). Periplasmic expression of human
735 growth hormone via plasmid vectors containing the λ PL promoter: Use of HPLC for product
736 quantification. *Protein Engineering*, 16(12), 1131–1138.
- 737 (37) Mirzadeh, K., Shilling, P. J., Elfageih, R., Cumming, A. J., Cui, H. L., Rennig, M., ... Daley, D. O. (2020).
738 Increased production of periplasmic proteins in *Escherichia coli* by directed evolution of the translation
739 initiation region. *Microbial Cell Factories*, 19, 85.
- 740 (38) Becker, S., Theile, S., Heppeler, N., Michalczyk, A., Wentzel, A., Wilhelm, S., ... Kolmar, H. (2005). A
741 generic system for the *Escherichia coli* cell-surface display of lipolytic enzymes. *FEBS Letters*, 579(5), 1177–
742 1182.
- 743 (39) Rangra, S., Kabra, M., Gupta, V., & Srivastava, P. (2018). Improved conversion of Dibenzothiophene into
744 sulfone by surface display of Dibenzothiophene monooxygenase (DszC) in recombinant *Escherichia coli*.
745 *Journal of Biotechnology*, 287, 59–67.

- 746 (40) Ahan, R. E., Klrpat, B. M., Saltepe, B., & Şeker, U. Ö. Ş. (2019). A Self-Actuated Cellular Protein Delivery
747 Machine. *ACS Synthetic Biology*, 8(4), 686–696.
- 748 (41) Wendel, S., Fischer, E. C., Martínez, V., Seppälä, S., & Nørholm, M. H. H. (2016). A nanobody: GFP
749 bacterial platform that enables functional enzyme display and easy quantification of display capacity.
750 *Microbial Cell Factories*, 15, 71.
- 751 (42) van Dijn, J. M., & Hecker, M. (2013). *Bacillus subtilis*: from soil bacterium to super-secreting cell factory.
752 *Microbial Cell Factories*, 12, 3.
- 753 (43) US Food and Drug Administration. *Carbohydrase and Protease Enzyme Preparations Derived From Bacillus*
754 *subtilis or Bacillus amyloliquefaciens; Affirmation of GRAS Status as Direct Food Ingredients*. , Pub. L. No. 21
755 CFR Part 184, 64 19887 (1999).
- 756 (44) Cui, W., Han, L., Suo, F., Liu, Z., Zhou, L., & Zhou, Z. (2018). Exploitation of *Bacillus subtilis* as a robust
757 workhorse for production of heterologous proteins and beyond. *World Journal of Microbiology and*
758 *Biotechnology*, 34, 145.
- 759 (45) Radeck, J., Meyer, D., Lautenschläger, N., & Mascher, T. (2017). *Bacillus* SEVA siblings: A Golden Gate-
760 based toolbox to create personalized integrative vectors for *Bacillus subtilis*. *Scientific Reports*, 7, 14134.
- 761 (46) Radeck, J., Kraft, K., Bartels, J., Cikovic, T., Dürr, F., Emenegger, J., ... Gebhard, S. (2013). The *Bacillus*
762 BioBrick Box: generation and evaluation of essential genetic building blocks for standardized work with
763 *Bacillus subtilis*. *Journal of Biological Engineering*, 7, 29.
- 764 (47) Popp, P. F., Dotzler, M., Radeck, J., Bartels, J., & Mascher, T. (2017). The *Bacillus* BioBrick Box 2.0:
765 Expanding the genetic toolbox for the standardized work with *Bacillus subtilis*. *Scientific Reports*, 7(1), 1–
766 13.
- 767 (48) Guiziou, S., Sauveplane, V., Chang, H.-J., Clerté, C., Declerck, N., Jules, M., & Bonnet, J. (2016). A part

768 toolbox to tune genetic expression in *Bacillus subtilis*. *Nucleic Acids Research*, 44(15), 7495–7508.

769 (49) Jensen, K., Østergaard, P. R., Wilting, R., & Lassen, S. F. (2010). Identification and characterization of a
770 bacterial glutamic peptidase. *BMC Biochemistry*, 11, 47.

771 (50) Brockmeier, U., Caspers, M., Freudl, R., Jockwer, A., Noll, T., & Eggert, T. (2006). Systematic Screening of
772 All Signal Peptides from *Bacillus subtilis*: A Powerful Strategy in Optimizing Heterologous Protein
773 Secretion in Gram-positive Bacteria. *Journal of Molecular Biology*, 362(3), 393–402.

774 (51) Shetty, R. P., Endy, D., & Knight, T. F. (2008). Engineering BioBrick vectors from BioBrick parts. *Journal of*
775 *Biological Engineering*, 2, 5.

776 (52) Cregg, J. M., Cereghino, J. L., Shi, J., & Higgins, D. R. (2000). Recombinant Protein Expression in *Pichia*
777 *pastoris*. *Molecular Biotechnology*, 16, 23–52.

778 (53) Mattanovich, D., Graf, A., Stadlmann, J., Dragosits, M., Redl, A., Maurer, M., ... Gasser, B. (2009). Genome,
779 secretome and glucose transport highlight unique features of the protein production host *Pichia pastoris*.
780 *Microbial Cell Factories*, 8, 29.

781 (54) Mattanovich, D., Branduardi, P., Dato, L., Gasser, B., Sauer, M., & Porro, D. (2012). Recombinant Protein
782 Production in Yeasts. In *Recombinant gene expression* (pp. 329–358).

783 (55) Heisteringer, L., Gasser, B., & Mattanovich, D. (2020). Microbe Profile: *Komagataella phaffii*: a methanol
784 devouring biotech yeast formerly known as *Pichia pastoris*. *Microbiology*, 166(7), 614–616.

785 (56) Kittl, R., Kracher, D., Burgstaller, D., Haltrich, D., & Ludwig, R. (2012). Production of four *Neurospora*
786 *crassa* lytic polysaccharide monooxygenases in *Pichia pastoris* monitored by a fluorimetric assay.
787 *Biotechnology for Biofuels*, 5, 79.

788 (57) Dell, W. B. O., Swartz, P. D., Weiss, L., & Meilleur, F. (2017). Crystallization of a fungal lytic polysaccharide
789 monooxygenase expressed from glycoengineered *Pichia pastoris* for X-ray and neutron diffraction. *Acta*

- 790 *Crystallographica Section F: Structural Biology Communications*, 73(2), 70–78.
- 791 (58) Breslmayr, E., Daly, S., Požgajčić, A., Chang, H., Rezić, T., Oostenbrink, C., & Ludwig, R. (2019). Improved
792 spectrophotometric assay for lytic polysaccharide monooxygenase. *Biotechnology for Biofuels*, 12, 283.
- 793 (59) Couturier, M., Ladevèze, S., Sulzenbacher, G., Ciano, L., Fanuel, M., Moreau, C., ... Berrin, J.-G. (2018).
794 Lytic xylan oxidases from wood-decay fungi unlock biomass degradation. *Nature Chemical Biology*, 14(3),
795 306.
- 796 (60) Cregg, J. M., Barringer, K. J., Hessler, A. Y., & Madden, K. R. (1985). *Pichia pastoris* as a Host System for
797 Transformations. *Molecular and Cellular Biology*, 5(12), 3376–3385.
- 798 (61) Theron, C. W., Berrios, J., Steels, S., Telek, S., Lecler, R., Rodriguez, C., & Fickers, P. (2019). Expression of
799 recombinant enhanced green fluorescent protein provides insight into foreign gene-expression
800 differences between Mut⁺ and MutS strains of *Pichia pastoris*. *Yeast*, 36(5), 285–296.
- 801 (62) Khosrowabadi, E., Takaloo, Z., Sajedi, R. H., & Khajeh, K. (2018). Improving the soluble expression of
802 aequorin in *Escherichia coli* using the chaperone-based approach by co-expression with artemin.
803 *Preparative Biochemistry and Biotechnology*, 48(6), 483–489.
- 804 (63) Baneyx, F., & Mujacic, M. (2004). Recombinant protein folding and misfolding in *Escherichia coli*. *Nature*
805 *Biotechnology*, 22(11), 1399–1407.
- 806 (64) Rosano, G. L., Morales, E. S., & Ceccarelli, E. A. (2019). New tools for recombinant protein production in
807 *Escherichia coli*: A 5-year update. *Protein Science*, 28(8), 1412–1422.
- 808 (65) Teilum, K., Nielsen, L. H., Gavrilov, Y., Olsen, J., Roche, J., Nielsen, A. T., & Lindorff-Larsen, K. (2020).
809 Selection of Stabilized Variants of Chymotrypsin Inhibitor 2. *The FASEB Journal*, 34(S1), 1.
- 810 (66) Frommhagen, M., Westphal, A. H., & Berkel, W. J. H. Van. (2018). Distinct Substrate Specificities and
811 Electron-Donating Systems of Fungal Lytic Polysaccharide Monooxygenases. 9, 1080.

- 812 (67) Stewart, E. J., Åslund, F., & Beckwith, J. (1998). Disulfide bond formation in the *Escherichia coli* cytoplasm:
813 An in vivo role reversal for the thioredoxins. *EMBO Journal*, 17(19), 5543–5550.
- 814 (68) Tandrup, T., Frandsen, K. E. H., Johansen, K. S., Berrin, J.-G., & Lo Leggio, L. (2018). Recent insights into
815 lytic polysaccharide monooxygenases (LPMOs). *Biochemical Society Transactions*, 46(6), 1431–1447.
- 816 (69) Forsberg, Zarah, Vaaje-Kolstad, G., Westereng, B., Bunæs, A. C., Stenstrøm, Y., MacKenzie, A., ... Eijsink,
817 V. G. H. (2011). Cleavage of cellulose by a CBM33 protein. *Protein Science: A Publication of the Protein*
818 *Society*, 20(9), 1479–1483.
- 819 (70) Ghatge, S. S., Telke, A. A., Waghmode, T. R., Lee, Y., Lee, K. W., Oh, D. B., ... Kim, S. W. (2015).
820 Multifunctional cellulolytic auxiliary activity protein HcAA10-2 from *Hahella chejuensis* enhances
821 enzymatic hydrolysis of crystalline cellulose. *Applied Microbiology and Biotechnology*, 99(7), 3041–3055.
- 822 (71) Paspaliari, D. K., Loose, J. S. M., Larsen, M. H., & Vaaje-Kolstad, G. (2015). *Listeria monocytogenes* has a
823 functional chitinolytic system and an active lytic polysaccharide monooxygenase. *FEBS Journal*, 282(5),
824 921–936.
- 825 (72) Gregory, R. C., Hemsworth, G. R., Turkenburg, J. P., Hart, S. J., Walton, P. H., & Davies, G. J. (2016).
826 Activity, stability and 3-D structure of the Cu(II) form of a chitin-active lytic polysaccharide
827 monooxygenase from: *Bacillus amyloliquefaciens*. *Dalton Transactions*, 45(42), 16904–16912.
- 828 (73) Schlegel, S., Rujas, E., Ytterberg, A. J., Zubarev, R. A., Luirink, J., & de Gier, J. W. (2013). Optimizing
829 heterologous protein production in the periplasm of *E. coli* by regulating gene expression levels. *Microbial*
830 *Cell Factories*, 12, 24.
- 831 (74) Lin-cereghino, G. P., Stark, C. M., Kim, D., Chang, J., Shaheen, N., Poerwanto, H., ... Lin-cereghino, J.
832 (2013). The effect of α -mating factor secretion signal mutations on recombinant protein expression in
833 *Pichia pastoris*. *Gene*, 519(2), 311–317.

- 834 (75) Brake, A. J., Merryweather, J. P., Coit, D. G., Heberlein, U. A., Masiarz, F. R., Mullenbach, G. T., ... Barr,
835 P. J. (1984). α -Factor-directed synthesis and secretion of mature foreign proteins in *Saccharomyces*
836 *cerevisiae*. *Proc Natl Acad Sci U S A*, 81, 4642–4646.
- 837 (76) Petrovi, D. M., Bissaro, B., Chylenski, P., Skaugen, M., Sørli, M., Jensen, M. S., ... Eijsink, V. G. H. (2018).
838 Methylation of the N-terminal histidine protects a lytic polysaccharide monooxygenase from auto-
839 oxidative inactivation. *Protein Science*, 22(6), 1–50.
- 840 (77) Kattla, J. J., Struwe, W. B., Doherty, M., Adamczyk, B., Saldova, R., Rudd, P. M., & Campbell, M. P. (2011).
841 Protein Glycosylation. In *Comprehensive Biotechnology, Second Edition* (Second Edi, Vol. 3).
- 842 (78) Vojcic, L., Despotovic, D., Martinez, R., Maurer, K., & Schwaneberg, U. (2012). An efficient transformation
843 method for *Bacillus subtilis* DB104. *Applied Microbiology and Biotechnology*, 94(2), 487–493.
- 844 (79) Cavaleiro, A. M., Kim, S. H., Seppälä, S., Nielsen, M. T., & Nørholm, M. H. H. (2015). Accurate DNA
845 Assembly and Genome Engineering with Optimized Uracil Excision Cloning. *ACS Synthetic Biology*, 4(9),
846 1042–1046.
- 847 (80) Raab, D., Graf, M., Notka, F., Schödl, T., & Wagner, R. (2010). The GeneOptimizer Algorithm: Using a
848 sliding window approach to cope with the vast sequence space in multiparameter DNA sequence
849 optimization. *Systems and Synthetic Biology*, 4(3), 215–225.
- 850 (81) Hemsworth, G. R., Henrissat, B., Davies, G. J., & Walton, P. H. (2014). Discovery and characterization of a
851 new family of lytic polysaccharide monooxygenases. *Nature Chemical Biology*, 10(2), 122–126.
- 852 (82) Quan, S., Hiniker, A., Collet, J.-F., & Bardwell, J. C. A. (2013). Isolation of Bacteria Envelope Proteins. In
853 *Bacterial Cell Surfaces: Methods and Protocols* (pp. 359–366).
- 854 (83) Barkholt, V., & Jensen, A. L. (1989). Amino acid analysis: Determination of cysteine plus half-cystine in
855 proteins after hydrochloric acid hydrolysis with a disulfide compound as additive. *Analytical*

856 *Biochemistry*, 177(2), 318–322.

857 (84) Westereng, B., Agger, J. W., Horn, S. J., Vaaje-Kolstad, G., Aachmann, F. L., Stenstrøm, Y. H., & Eijsink, V.

858 G. H. (2013). Efficient separation of oxidized cello-oligosaccharides generated by cellulose degrading

859 lytic polysaccharide monooxygenases. *Journal of Chromatography A*, 1271, 144–152.

860 (85) Rasmussen, M. D., Bjoernvad, M. E., & Diers, I. (2000). *Patent No. WO 00/75344*. World Intellectual Property

861 Organization.

862 (86) Schindelin, J., Arganda-carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., ... Cardona, A. (2019).

863 Fiji: an open-source platform for biological-image analysis. *Nature Methods*, 9(7).

864

865