

# The transient expression of recombinant proteins in plant cell packs facilitates stable isotope labelling for NMR spectroscopy

Patrick Opdensteinen<sup>1,2</sup> , Laura E. Sperl<sup>3,4</sup>, Mariam Mohamadi<sup>3,4</sup>, Nicole Kündgen-Redding<sup>1</sup>, Franz Hagn<sup>3,4</sup>  and Johannes Felix Buyel<sup>1,2,\*</sup> 

<sup>1</sup>Fraunhofer Institute for Molecular Biology and Applied Ecology IME, Aachen, Germany

<sup>2</sup>Institute for Molecular Biotechnology, RWTH Aachen University, Aachen, Germany

<sup>3</sup>Bavarian NMR Center (BNMRZ) at the Department of Chemistry, Technical University of Munich, Garching, Germany

<sup>4</sup>Institute of Structural Biology, Helmholtz Zentrum München, Neuherberg, Germany

Received 3 November 2021;

revised 5 April 2022;

accepted 3 June 2022.

\*Correspondence (Tel +49 241 6085 13162; fax +49 241 6085 10000; emails johannes.buyel@rwth-aachen.de; johannes.buyel@ime.fraunhofer.de)

## Summary

Nuclear magnetic resonance (NMR) spectroscopy can be used to determine the structure, dynamics and interactions of proteins. However, protein NMR requires stable isotope labelling for signal detection. The cells used for the production of recombinant proteins must therefore be grown in medium containing isotopically labelled substrates. Stable isotope labelling is well established in *Escherichia coli*, but bacteria are only suitable for the production of simple proteins without post-translational modifications. More complex proteins require eukaryotic production hosts, but their growth can be impaired by labelled media, thus reducing product yields and increasing costs. To address this limitation, we used media supplemented with isotope-labelled substrates to cultivate the tobacco-derived cell line BY-2, which was then cast into plant cell packs (PCPs) for the transient expression of a labelled version of the model protein GB1. Mass spectrometry confirmed the feasibility of isotope labelling with <sup>15</sup>N and <sup>2</sup>H using this approach. The resulting NMR spectrum featured a signal dispersion comparable to recombinant GB1 produced in *E. coli*. PCPs therefore offer a rapid and cost-efficient alternative for the production of isotope-labelled proteins for NMR analysis, especially suitable for complex proteins that cannot be produced in microbial systems.

**Keywords:** alternative expression host, defined cultivation media, isotope labelling, plant cell culture, structural analysis.

## Introduction

The analysis of protein structures and interactions often provides functional information that can help to determine disease mechanisms and facilitate the development of therapeutics (Kiss *et al.*, 2018; Phyo *et al.*, 2021; Wand, 2001), as recently shown by the analysis of the SARS-CoV-2 spike protein (Mycroft-West *et al.*, 2020). Methods such as X-ray crystallography and cryo-electron microscopy provide high-resolution models of large, static structures (Matthies *et al.*, 2018; Sun *et al.*, 2021). In contrast, nuclear magnetic resonance (NMR) offers the unique advantage of determining protein structures in solution, studying protein dynamics at multiple time scales (Chao and Byrd, 2020) and investigating weak protein interactions (Chao and Byrd, 2020; Zuiderweg, 2002).

NMR detects the frequency of nuclear spins in (protein-) samples in the presence of high magnetic fields and therefore requires that proteins are labelled with spin-½ nuclei such as <sup>13</sup>C or <sup>15</sup>N. Large protein systems above 30 kDa in size require additional labelling with <sup>2</sup>H (Gardner and Kay, 1998; Grzesiek and Bax, 1992) to minimize signal loss caused by fast spin–spin relaxation. Recombinant proteins can be labelled by growing bacteria in media containing defined labelled substrates such as ammonium chloride (<sup>15</sup>NH<sub>4</sub>Cl). When eukaryotic cell cultures are used instead, complex substrates such as extracts of yeast or algae (originally cultivated with defined labelled substrates) may

be required to meet nutritional demands (Egorova-Zachernyuk *et al.*, 2011; KyungRyun and Ho, 2020). Whereas bacteria such as *Escherichia coli* are widely used to produce isotope-labelled proteins (Goto and Kay, 2000), they cannot produce properly folded complex proteins due to the absence of disulfide bonds and the lack of post-translational modifications such as glycans (Goto and Kay, 2000). This can be addressed to some extent by *in vitro* refolding, but at the cost of lower yields (Eiberle and Jungbauer, 2010). Such issues limit the structural analysis of membrane proteins and other complex targets by solution-state NMR. Animal cell lines can be used as an alternative production system, but the cells may not grow well in media containing a large volume fraction of deuterium oxide (D<sub>2</sub>O, <sup>2</sup>H<sub>2</sub>O, ‘heavy water’ or D; Takahashi and Shimada, 2010; Egorova-Zachernyuk *et al.*, 2011; Dutta *et al.*, 2012; Clark *et al.*, 2018), and many medically relevant proteins interfere with cell viability, thus reducing yields (Favot *et al.*, 2005; Gengenbach *et al.*, 2018; Gengenbach *et al.*, 2019). Another alternative is the production of labelled proteins in plant cells. Plants often tolerate the expression of medically relevant proteins better than animal cells either because there are no targets for the recombinant protein in plant cells or because the protein is directed to accumulate in a subcellular compartment where it is isolated from potential interaction partners (Di Fiore *et al.*, 2002; Streatfield, 2007). So far, plants and plant cells have been used, for example, to produce <sup>15</sup>N labelled proteins (Bindschedler *et al.*, 2008; Ohki

*et al.*, 2008; Yagi *et al.*, 2015). However, to our knowledge, plants or plant cells have not been tested in the context of protein deuteration yet. Additionally, the plants or plant cells that have been tested up to now were stably genetically modified, which requires lead times of several months, expert knowledge for transformation and dedicated laboratory facilities.

Here, we conducted a proof-of-concept study to determine whether protein labelling in culture medium containing D<sub>2</sub>O or/and <sup>15</sup>NH<sub>4</sub><sup>15</sup>NO<sub>3</sub> (N, DN) can be achieved by transient expression in plant cells. We used Bright Yellow 2 (BY-2) cells derived from *Nicotiana tabacum* (tobacco) roots as a model plant cell line and chose plant cell packs (PCPs) as a transient expression system because this allows the simple, rapid and parallelized production of proteins (Gengenbach *et al.*, 2020; Rademacher *et al.*, 2019). We selected the B1 domain of the immunoglobulin binding protein G (GB1) from *Streptococcus sp.* as a well-established model protein for NMR-based structural studies (Bouvignies *et al.*, 2006; Gronenborn *et al.*, 1991; Li *et al.*, 2015). We found that plant cells allow the incorporation of <sup>15</sup>N and <sup>2</sup>H into GB1, offering promising insights into the use of plant cells for the production of isotope-labelled proteins (particularly deuterated proteins) for NMR, which could potentially be extended to more challenging targets such as membrane proteins in the future.

## Results and discussion

### Transient expression of GB1 in PCPs

First, we tested the expression of different GB1 variants (Table 4) in PCPs to identify the subcellular compartment/localization achieving the highest protein accumulation, which is a routine procedure in transient protein production (Gengenbach *et al.*, 2020). We included the carbon source during BY-2 cell cultivation as an additional variable because using glucose instead of the standard sucrose had more than doubled transient expression of proteins in PCPs in the past (Opdensteinen *et al.*, 2021). Similarly, we found that glucose increased GB1 accumulation in PCPs up to twofold compared with sucrose and that overall plastid-targeting in combination with a CHS 5'UTR resulted in the highest product levels of ~100 mg/kg biomass when using glucose as carbon source (Figure S1). We therefore used the plastid-targeted GB1 for all subsequent expression experiments.

### BY-2 cell growth in labelled media

We next investigated how well BY-2 cells can tolerate the replacement of water with deuterium oxide in their cultivation media. We used Evans blue staining to detect cells with compromised cell membrane integrity as an indicator of viability loss. Even though this method is not directly linked to a physiological property such as intracellular ATP levels (Riss *et al.*, 2004), it is well established and in use in plant biotechnology investigations since more than four decades (Ikegawa *et al.*, 2000; Kim *et al.*, 2018; Taylor and West, 1980). Also, whereas cell aggregates can substantially distort the determination of the cell count, the BY-2 cell line we used here exhibited only moderate formation of such structures, for example stretches of cells, so that live and dead cell counts were reliably determined (Figure S2A and B). Specifically, cell counts differed <10% when comparing numbers for partially aggregated cells with those obtained after enzymatic dissolution of aggregates (data not shown).

We did not find any living cells 7 days after replacing the water with deuterium oxide in one step (data not shown). In contrast, the cells were able to tolerate 20% v/v deuterium oxide, with only a ~5% difference in viability compared with a water control after 7 days (Figure 1a). Gradually increasing the proportion of deuterium oxide after each cell passage (after 7–11 days) caused a decline in cell viability and the correlation was linear if the pure water control was excluded and viabilities at the end of each passage were compared (Figure 1b, Pearson's  $r = -0.991$ , adj.  $R^2 = 0.986$ ). Accordingly, we were able to derive a no-effect volume fraction of ~11% for deuterium oxide based on the intersection of the linear regression with a horizontal 100% viability threshold. Importantly, BY-2 cells survived for ~14 days in 50% v/v deuterium oxide with a viability >50%. This indicated that such conditions can be used to cultivate cells in media with a meaningful content of labelled components and at the same time obtain sufficient biomass for subsequent PCP casting and transient protein expression.

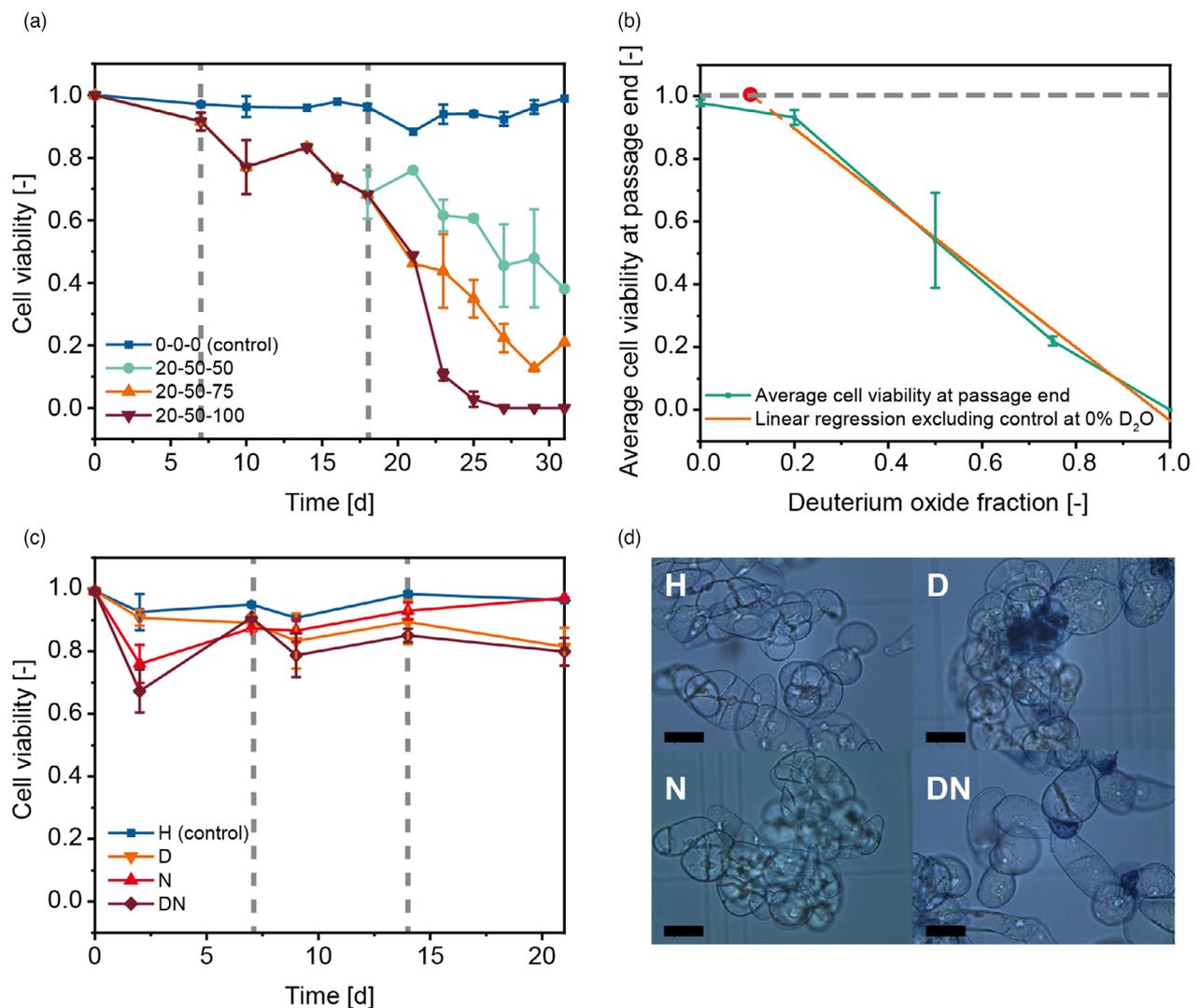
We used the cells grown in 50% v/v deuterium oxide (i.e. the 20–50–50 regime shown in Figure 1a) for further cultivation and found that the average cell viability increased in subsequent passages and stabilized at ~90% compared with the water control (Figure 1c), which was close to a significant difference (two-sided, two-sample Welch test,  $\alpha = 0.05$ ,  $N = 12$ ,  $P = 0.075$ ). Our interpretation was that the gradual decline in cell viability observed when increasing the concentration of deuterium oxide was not an inevitable effect of heavy water accumulation, but instead indicated that the cells can adapt when given sufficient time. Therefore, cultivating BY-2 cells in >50% deuterium oxide can probably be achieved in the future, although additional passages with smaller increments (such as <5% per step) may be necessary.

Replacing ammonium nitrate with a labelled counterpart (<sup>15</sup>NH<sub>4</sub><sup>15</sup>NO<sub>3</sub>) did not significantly affect cell viability (two-sided, two-sample Welch test,  $\alpha = 0.05$ ,  $N = 12$ ,  $P = 0.222$ ), but was close to significance when combined with 50% v/v deuterium oxide (two-sided, two-sample Welch test,  $\alpha = 0.05$ ,  $N = 12$ ,  $P = 0.053$ ; Figure 1c and Table 1). We concluded that statistically significant differences in BY-2 viability can probably be identified between labelled and non-labelled media if the sample size increases (e.g. through long-term parallel cultivation), but these differences will be experimentally irrelevant due to the small absolute difference.

Light microscopy did not reveal substantial morphological differences between cells growing in the four different media (Figure 1d), although automated image analysis could be applied in the future in an attempt to find quantitative differences (Müsken *et al.*, 2010). We also measured the wet and dry cell mass, and found that only the deuterium oxide alone reduced the biomass yield (Table 1), but this difference was only significant compared with the labelled ammonium nitrate medium due to the substantial scattering and limited number of data points. Again, extended cultivation experiments will be necessary to establish the precise quantitative difference between these settings.

### PCP-based expression of GB1 and DsRed using BY-2 cells grown in labelled media

We next used BY-2 cells cultivated in media containing deuterium oxide, labelled ammonium nitrate or both (Table 2) to cast PCPs (Gengenbach *et al.*, 2020), and investigated the transient expression of the model protein GB1, which has been well characterized



**Figure 1** BY-2 cell viability in labelled and control media. (a) A gradually increasing volume fraction of deuterium oxide reduced cell viability over time. Cell passing times are indicated by dashed vertical lines. Numbers in the legend indicate the percentage deuterium oxide in the medium at the start of each passage *v/v*, for example 20–50–50 corresponds to cells cultured in 20% D<sub>2</sub>O in passage 1 and then at 50% D<sub>2</sub>O in passages 2 and 3. (b) The correlation between the deuterium oxide volume fraction and loss of cell viability was linear (Pearson's  $r = -0.991$ , adj.  $R^2 = 0.986$ ) if the water control was excluded, indicating a deuterium oxide no-effect volume fraction of ~11% *v/v* (red dot). (c) BY-2 cells in media containing different labelled components and control medium. Cells in 50% *v/v* deuterium oxide had been passaged with this volume fraction ~27 days before starting the cultivation shown here (see panel a, 20–50–50 regime). Details on the cell properties at the time of harvest are given in Table 1. (d) Microscopic images of BY-2 cells cultivated in labelled and control media after staining with Evans blue. Scale bar = 50 μm. H—control medium, regular water; D—medium prepared using 50% *v/v* deuterium oxide; N—medium prepared using regular water but containing labelled ammonium nitrate (15NH<sub>4</sub>15NO<sub>3</sub>); DN—medium prepared using 50% *v/v* deuterium oxide and labelled ammonium nitrate. Error bars in a–c indicate the standard deviation ( $n \geq 2$ ).

by NMR (Kirsten Frank *et al.*, 2002; Li *et al.*, 2012). The presence of deuterium oxide (50% *v/v*) with or without labelled ammonium did not impair the casting of PCPs. We concluded that any potential isotope-induced changes in BY-2 morphology (Figure 1d) did not interfere with the technical aspects of high-throughput expression in PCPs. BY-2 cell morphology after casting PCPs and infiltration with *A. tumefaciens* showed no visible signs of lost vitality, even at 4 dpi (Figure S2C).

GB1 was successfully expressed in PCPs prepared from BY-2 cells cultivated in all the media we tested, with the highest accumulation of ~55 mg/kg wet biomass in the presence of deuterium oxide (Figure 2 and Table 1), which was in the same

order of magnitude as the expression during screening (Figure S1C) and also in good agreement with expression of labelled GB1 in *E. coli* (Cheng and Patel, 2004). In terms of volumetric yield based on the combined GB1 accumulation and wet biomass yield of the cells (Table 1), the combined labelling approach (DN medium) performed best with a value of ~5.4 mg/L.

Interestingly, GB1 accumulated to significantly higher levels in the cells cultured in any of the isotope-containing media compared with the control medium (Table S1) despite the lower cell viability and growth rate in the presence of isotopes (Figure 1c and Table 1). A similar reduction in growth rate in the presence of isotopes was also reported for *E. coli* (Filiou

**Table 1** The accumulation, extraction and purification of GB1 from BY-2 PCPs with media compositions as listed in Table 2

Component	Medium	H	D	N	DN
	Unit				
Cell count at harvest	[10 <sup>9</sup> L <sup>-1</sup> ]	2.38 ± 0.13	0.95 ± 0.26	2.05	1.12 ± 0.02
Cell viability at harvest <sup>a</sup>	[-]	0.97 ± 0.01	0.82 ± 0.06	0.97 ± 0.01	0.80 ± 0.04
BY-2 cell wet mass at harvest <sup>b</sup>	g/L	244 ± 41	83 ± 11	286 ± 16	227 ± 60
BY-2 cell dry mass at harvest	g/L	53	6	47	45
Average growth rate <sup>c</sup>	d <sup>-1</sup>	0.58 ± 0.08	0.39 ± 0.52	0.41 ± 0.06	0.45 ± 0.41
GB1 accumulation in PCPs <sup>d</sup>	mg/kg (wet cell biomass)	2.7 ± 2.8	56.4 ± 39.7	8.8 ± 6.6	23.8 ± 15.3
Theoretical volumetric GB1 yield <sup>e</sup>	mg/L	0.66 ± 0.68	4.67 ± 3.29	2.51 ± 1.87	5.42 ± 3.48
Volume BY-2 PCP extract used for purification	mL	10	5	15	15
GB1 purity after IMAC (densitometric evaluation) <sup>f</sup>	% peak area of stained bands	69	95	87	80
GB1 recovery after IMAC	% initial	101	81	91	41

<sup>a</sup>Data as shown in Figure 1c.

<sup>b</sup>*n* = 2.

<sup>c</sup>Average of last two cell passages before harvest; the growth rate of the same cells in medium H was 0.54 (Opdensteinen *et al.*, 2021).

<sup>d</sup>Quantification using an authentic his-tagged GB1 standard.

<sup>e</sup>The values are theoretical because BY-2 cells did not express the protein in a bioreactor, but in a subsequent step in PCPs, accumulation based on authentic standard quantitation is used.

<sup>f</sup>GB1 purity after IMAC was estimated based on the densitometric evaluation of stained LDS gels using AIDA software. GB1 recovery was estimated based on the densitometric evaluation of dot blots using standards with a known concentration and the software ImageJ (Schneider *et al.*, 2012). Data values with variance indicate standard deviation.

*et al.*, 2012). Furthermore, differences in GB1 accumulation between the labelled media were significant for each pair of media (Table S1). Specifically, GB1 accumulated to more than double the level in the deuterium-labelled media compared with the media containing labelled ammonium nitrate. The presence of labelled nitrogen has been shown to alter protein expression levels (compared to <sup>14</sup>N) in *E. coli* (Filiou *et al.*, 2012), but not (to our knowledge) in BY-2 cells. Given that the transient expression of proteins in BY-2 cells is substantially affected by the nitrogen supply (Holland *et al.*, 2010), an option to increase the labelling of plant (cell)-derived proteins can be the use of <sup>13</sup>C-labelled glucose in the media because carbon is one of the main components of proteins. This strategy has been successfully applied to bacteria-derived proteins used for high-resolution NMR studies (Sattler *et al.*, 1999). Furthermore, using this carbon source instead of sucrose boosted GB1 expression by a factor of two (Figure S1). Similar results have been reported for the expression of IgG1 and IgG3 antibodies in PCPs derived from BY-2 cells (Opdensteinen *et al.*, 2021).

Interestingly, the accumulation of GB1 inversely correlated with the BY-2 cell wet mass in the cultures at the time of PCP preparation (i.e. 90.5 g/L in medium D compared with 280.0 ± 14.7 g/L in the other media). This phenomenon will need to be investigated in more detail by analysing nutrient uptake during cultivation. We also observed a strong linear correlation between GB1 accumulation in cells grown in labelled media and the accumulation of DsRed expressed in cells of identical origin (Figure 2b). We concluded that other proteins are likely to perform similarly to GB1, and we will validate the generalizability of our results by testing additional proteins in the future.

### Purification of GB1

Due to the highest accumulation in PCPs (Figure S1), we purified plastid-targeted GB1 for NMR analysis. The GB1 we expressed featured a C-terminal His<sub>6</sub> tag for rapid IMAC purification from

PCP extracts. After loading the column, a washing step with 30 mM imidazole was used to remove residual BY-2 HCPs that bound non-specifically to the IMAC resin, followed by single-step elution with 300 mM imidazole. Using this strategy, we recovered labelled GB1 from PCP extracts with a purity of up to 95% based on densitometric evaluation, despite low product levels in the PCPs (≤5 mg/kg) and minimal extract volumes (≤15 mL) used for chromatography (Figure 2c,d and Table 1). The average GB1 recovery was 80 ± 28% based on densitometric evaluation.

We observed two additional protein bands during the analysis of purified GB1 produced in plastids by LDS gel electrophoresis, and both were also detected with the anti-His<sub>6</sub> antibody suggesting they are related to GB1 (Figure 2c, d). We reasoned whether the bands reflected the presence of post-translational modifications (Grabsztunowicz *et al.*, 2017; Lehtimäki *et al.*, 2015), that were not present in cytosolic GB1, where we only detected a single product band (Figure S1). However, neither *in silico* tools (Gupta and Brunak, 2002; Hamby and Hirst, 2008) for post-translational modification site prediction nor, most importantly, mass spectrometric analysis revealed possible modifications or a relevant mass difference, respectively. Specifically, the mass difference between GB1 produced in BY-2 plastids and *E. coli* was <2 Da or <80 Da for the unlabeled and [<sup>2</sup>H, <sup>15</sup>N]-labelled variants, respectively, which was too small for any glycosylation or incomplete signal peptide processing (Shen *et al.*, 2017; Figure 3a, b). We concluded that the apparent size shift was an artefact of the LDS-PAGE analysis.

### Analysis of plant-derived GB1 by ESI-MS and NMR spectroscopy

Next, we used ESI-MS to determine the degree of GB1 isotope labelling that was achieved in the four media, as well as of GB1 obtained from *E. coli* with identical isotope labelling conditions. As anticipated, the molecular mass of GB1 (Figure 3a) increased when the protein was produced in a medium containing heavy

**Table 2** Cultivation media tested for the production of labelled GB1 in BY-2 cells

Component	Medium				
	Unit	H	D	N	DN
2,4-Dichlorophenoxyacetic acid	mg/L ( $\mu\text{M}$ )	0.2 (0.9)	0.2 (0.9)	0.2 (0.9)	0.2 (0.9)
Glucose	g/L (M)	30 (0.167)	30 s(0.167)	30 (0.167)	30 (0.167)
MS medium mix	g/L	4.3	4.3	0	0
MS macro salts <sup>a</sup>	g/L	0	0	1.65	1.65
MS micro salts <sup>b</sup>	g/L	0	0	1.00	1.00
Myo-inositol	g/L (mM)	0.1 (0.555)	0.1 (0.555)	0.1 (0.555)	0.1 (0.555)
Potassium di-hydrogen phosphate	g/L (mM)	0.2 (1.470)	0.2 (1.470)	0.2 (1.470)	0.2 (1.470)
Thiamine	mg/L ( $\mu\text{M}$ )	1.0 (3.0)	1.0 (3.0)	1.0 (3.0)	1.0 (3.0)
Ammonium nitrate (unlabelled)	g/L (mM)	1.65 (20.6)	1.65 (20.6)	0	0
Ammonium nitrate (labelled, $^{15}\text{NH}_4^{15}\text{NO}_3$ )	g/L (mM)	0	0	1.65 (20.6)	1.65 (20.6)
Deuterium oxide	% v/v	0	50	0	50
Fraction of labelled nitrogen	% m/m	0	0	69	69
Medium costs	€ L <sup>-1</sup>	0.78	200.78	233.72	433.72

<sup>a</sup>This component contained 0.97 g/L (9.6 mM) potassium nitrate and thus unlabeled nitrogen.

<sup>b</sup>This component contained 0.93 g/L (9.2 mM) potassium nitrate and thus unlabeled nitrogen.

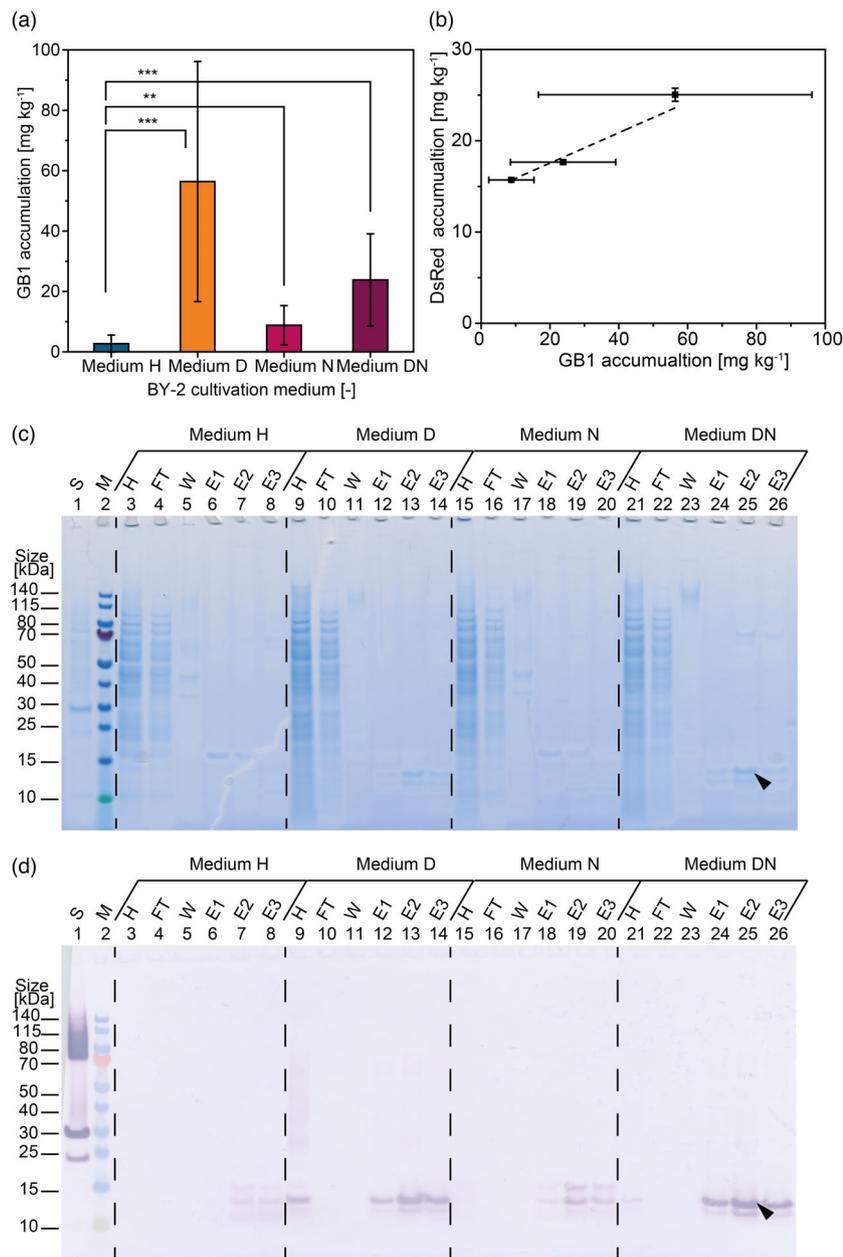
isotopes ( $^2\text{H}$  or  $^{15}\text{N}$ ). The detected increase in molecular mass (Figure 3b and Table 3) was then used to calculate the degree of isotope incorporation.

We found that the degree of labelling for  $^{15}\text{N}$  was ~70%, which corresponded well to the fraction of labelled nitrogen in the plant cell cultivation medium (Table 2) and implied that the labelling efficacy was close to 100% and thus similar to the >95% reported in *E. coli* grown in 99%  $^{15}\text{N}$  labelled growth medium (Duff *et al.*, 2015). Lower degrees of labelling have been reported for other eukaryotic systems, such as insect cells (~80%  $^{15}\text{N}$ ; Skora *et al.*, 2015), mammalian cells (60%–90%  $^{15}\text{N}$ ; Hansen *et al.*, 1994; Mitri *et al.*, 2018; Sastry *et al.*, 2011) as well as wild-type *Arabidopsis thaliana* (~30–75%; Bindschedler *et al.*, 2008) and transgenic *Nicotiana benthamiana* plants (52%–58%  $^{15}\text{N}$ ; Yagi *et al.*, 2015). Importantly, the cost of media for isotope labelling in animal cells is very high (Dutta *et al.*, 2012; Skora *et al.*, 2015), whereas they were only ~200–450 €/L for the media we tested (Table 2), which was in agreement with the ~15 €/L reported before (Bindschedler *et al.*, 2008), taking into account the absence of  $\text{D}_2\text{O}$  and reduced concentration of non  $^{15}\text{N}$  isotope-labelled ammonium nitrate in the earlier report (KyungRyun and Ho, 2020).

We found that the degree of labelling for  $^2\text{H}$  was only 16%. In part, this was a result of the used maximal 50% v/v  $\text{D}_2\text{O}$  content in the original medium. Furthermore, ammonium nitrate and sucrose were sources of additional protonation, which would reduce  $^2\text{H}$  labelling efficiency. There are no comparable data from plants, but results from *E. coli* indicate that the maximum fraction of deuteration in amino acids is only ~50% when using a defined protonated carbon source and 100%  $\text{D}_2\text{O}$  in the medium (Fogel *et al.*, 2016). In line with these earlier observations, our labelling results in *E. coli* using a M9 minimal growth medium (Miller, 1972) with 50% v/v  $\text{D}_2\text{O}$ ,  $^{15}\text{N}$  ammonium chloride and protonated  $^{12}\text{C}$  glucose, as a sole nitrogen and carbon source, respectively, indicate a degree of deuteration of 35% (Table 3), highlighting the impact of protonated components in the growth medium on the final level of deuteration. Importantly, the hydrogen isotope composition in plants is tissue-specific, with roots (the origin of the BY-2 cells we used) containing 6%–7% less deuterium than other tissues (Sanchez-Bragado *et al.*, 2019). Finally, the cells we used

had been passaged in the labelled medium only twice, indicating that 1%–2% of the cell mass used in PCPs was non-labelled, reducing the maximum efficiency accordingly. Taking these aspects into account, we estimate that the maximum  $^2\text{H}$  labelling efficiency of our current set-up would be ~16%–23%, which is close to the 16% we observed. In the future, we will use mid-term and long-term cultures with more than 50% v/v  $\text{D}_2\text{O}$  to increase the degree of  $^2\text{H}$  labelling. Overall, our data show that plant cells offer a promising alternative to other eukaryotic systems for the high-level incorporation of deuterium in proteins, overcoming the toxicity issues affecting mammalian cells (Kushner *et al.*, 1999). In addition to an estimation of the labelling degree, the mass spectrometry data can be used to assess the presence of post-translational modifications in plants versus *E. coli*. As sample H in plants has a mass of 7257 Da, which is only 2 Da smaller than the expected mass of the GB1 construct, it can be noted that no modification is present in plastid-targeted, plant-derived GB1. This is further confirmed by an almost identical mass of the unlabeled GB1 construct produced in *E. coli*, where post-translational modifications are absent.

Finally, we used 2D-NMR spectroscopy to confirm the structural integrity of plant-derived GB1. Given that we acquired higher yields of  $^2\text{H}$ ,  $^{15}\text{N}$ -labelled GB1 than  $^1\text{H}$ ,  $^{15}\text{N}$ -labelled GB1, we used the former for the acquisition of a 2D- $^{15}\text{N}$ ,  $^1\text{H}$ -HSQC NMR experiment. The resulting high-quality spectrum featured a sharp signal for each backbone amide moiety in the protein (Figure 3c). Furthermore, the large signal dispersion is in agreement with the well-structured  $\beta$ -sheet fold of GB1 (Yao *et al.*, 1997). Next, we compared the NMR spectrum of PCP-derived  $^2\text{H}$ ,  $^{15}\text{N}$ -GB1 to a 2D- $^{15}\text{N}$ ,  $^1\text{H}$ -HSQC spectrum of an identical GB1 construct labelled with  $^{15}\text{N}$  and  $^2\text{H}$  in *E. coli*. GB1 produced in *E. coli* has previously been used for a large number of NMR structure analyses and can be regarded as a model protein for the evaluation of labelling methods (Bouvignies *et al.*, 2006; Gronenborn *et al.*, 1991; Li *et al.*, 2015). When comparing the spectra, we found that the overall signal dispersion and position was almost identical (Figure 3c), suggesting that the plant-derived and bacterial GB1 proteins were folded in an identical manner. The minor differences in the spectra (Figure 3d, e) can be attributed to differences in the sample purity, which was lower in the plant-derived sample (Figure 2).

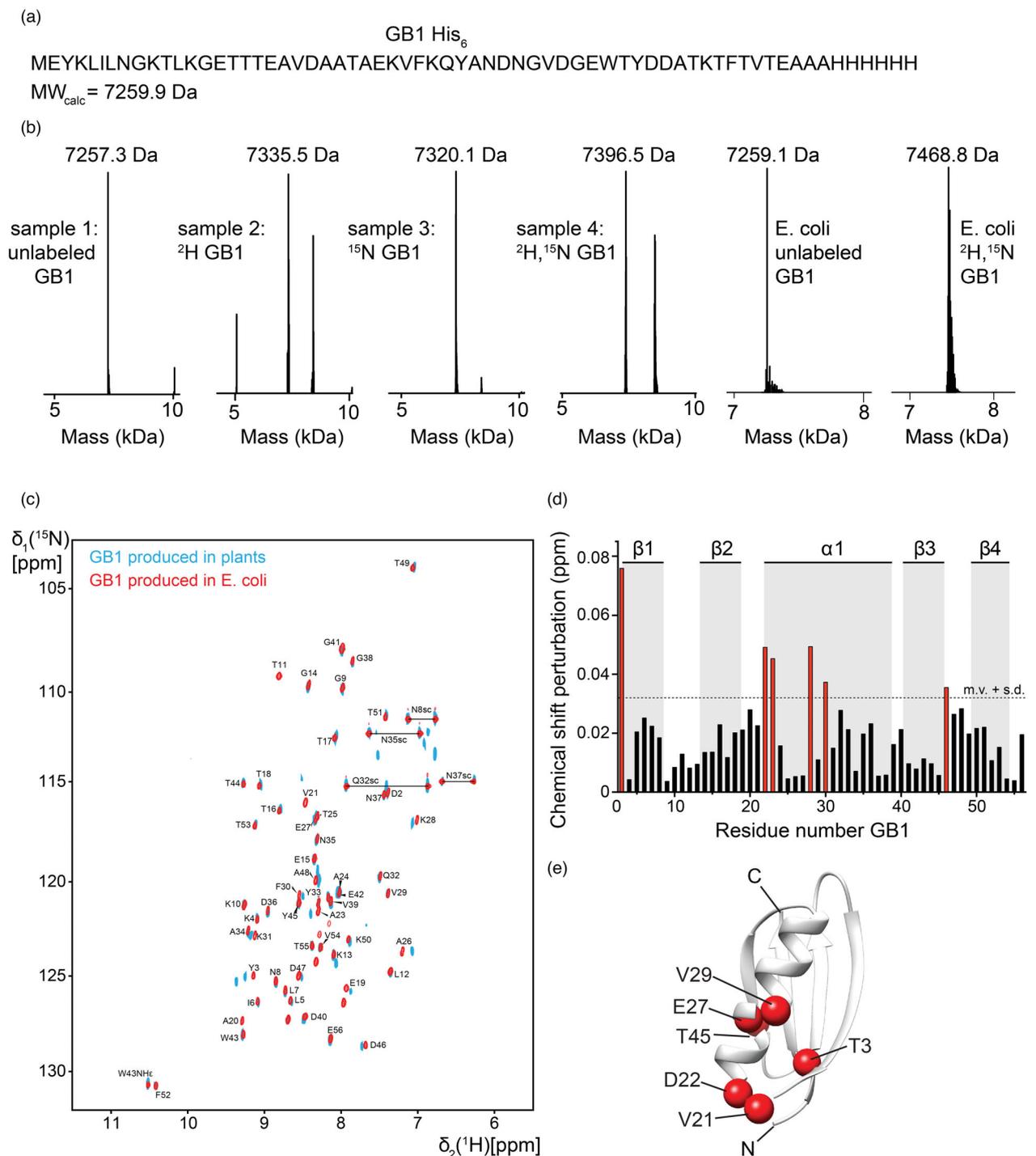


**Figure 2** Expression of plastid-targeted GB1 in PCPs prepared from BY-2 cells grown in labelled and control media, and purification from PCP extracts by immobilized metal affinity chromatography (IMAC). (a) Accumulation of GB1 in BY-2 PCPs cultivated in media with (D, N, DN) and without (H, control) isotopes was estimated from dot blots using a rabbit anti-His<sub>6</sub> primary antibody and an alkaline phosphatase-labelled goat anti-rabbit secondary antibody. GB1 was quantified by densitometry against an authentic GB1 standard in eight dilutions. Error bars show standard deviations from 12–14 PCPs ( $***P < 0.0001$ ;  $**P < 0.01$ ). (b) Accumulation of GB1 and DsRed in plastids of PCPs prepared from BY-2 cells grown in labelled media. DsRed fluorescence was quantified using authentic standards and GB1 was quantified by the densitometric analysis of dot blots against corresponding GB1 standards. The correlation between DsRed and GB1 was linear (Pearson's  $r = 0.980$ ,  $R^2 = 0.961$ , Adj.  $R^2 = 0.921$ ). The error bars for GB1 are considerably larger than those for DsRed because the densitometric quantitation of the former was less precise than fluorescence-based measurements for DsRed. D—medium prepared using 50% v/v deuterium oxide; N—medium prepared using regular water but containing labelled ammonium nitrate ( $^{15}\text{NH}_4^{15}\text{NO}_3$ ); DN—medium prepared using 50% v/v deuterium oxide and labelled ammonium nitrate. (c and d) Samples collected during the purification of plastid-targeted GB1 by IMAC were analysed by LDS gel electrophoresis and subsequent staining with Coomassie Brilliant Blue (c) and Western blotting (d) using the same antibodies as for the dot blot (see panel a). Black arrows indicate the expected size of GB1. S—DsRed standard (10  $\mu\text{g/mL}$ ), M—marker, H—PCP extract, FT—IMAC flow through, WIMAC wash (30 mm imidazole), E1–E3—IMAC elution fractions.

## Conclusions

We have achieved the plant-based production of isotope-labelled recombinant proteins in as little as 3 days using PCPs prepared

from BY-2 cells cultivated in the presence of deuterium oxide and/or  $^{15}\text{N}$ -labelled ammonium nitrate. The small-scale (~10 mL) shake flask cultivation of BY-2 cells (~2.5 g wet biomass) allowed cost-efficient isotope labelling (Marley *et al.*, 2001) while producing



**Figure 3** ESI-MS and NMR analysis of GB1 samples derived from plant cells. (a) Amino acid sequence of GB1 after the proteolytic removal of the signal peptide. (b) Mass spectra of the intact protein from the four GB1 samples produced in different isotopic media. The molecular mass of the GB1 species is indicated above the spectrum in each case and the corresponding degree of isotope labelling is shown in Table 3. (c) Overlay of the 2D-[<sup>15</sup>N,<sup>1</sup>H]-HSQC spectra of the <sup>2</sup>H,<sup>15</sup>N-labelled GB1 plant-derived sample (medium DN, plastid-targeted, blue) and <sup>2</sup>H,<sup>15</sup>N-labelled GB1 produced in *E. coli* (red). Sequence-specific resonance assignments are labelled (Ikeya *et al.*, 2016). (d) Chemical shift perturbation (CSP) plot indicating only minor spectral changes between the two samples in c. (e) Significant CSPs mapped onto the structure of GB1 (2n9k.pdb (Ikeya *et al.*, 2016)) indicates slight changes at the N-terminal end of the protein.

sufficient quantities of protein for initial screening (~12.5 µg). Depending on the size of the target protein, 1–10 mg is required to prepare an NMR sample with a concentration exceeding 100 µM, which would enable multidimensional NMR experiments.

This amount of labelled protein can be produced by scaling up the PCPs to 100 g (Rademacher *et al.*, 2019) or more and by converting plant cell cultivation into stirred tank bioreactors (which is done routinely; Holland *et al.*, 2010) without losing the

**Table 3** ESI-MS data and calculated  $^2\text{H}$  and  $^{15}\text{N}$  isotope labelling degrees of GB1 samples obtained from PCPs and *E. coli*

Sample	Unit	H: $^1\text{H}, ^{14}\text{N}$	D: $^2\text{H}, ^{14}\text{N}$	N: $^1\text{H}, ^{15}\text{N}$	DN: $^2\text{H}, ^{15}\text{N}$	<i>E. coli</i> $^1\text{H}, ^{14}\text{N}$	<i>E. coli</i> $^2\text{H}, ^{15}\text{N}$
Observed mass	Da	7257.3	7335.5	7320.1	7396.5	7259.1	7468.8
$\Delta m$ $^2\text{H}$	Da	n.a.	D-H: 78.2	n.a.	DN-N: 76.4	n.a.	DN-H <sup>a</sup> : 121.8
Labelling degree $^2\text{H}$	%	n.a.	<b>16.3</b>	n.a.	<b>15.9</b>	n.a.	<b>34.8</b>
$\Delta m$ $^{15}\text{N}$	Da	n.a.	n.a.	N-H: 62.8	DN-D: 61	n.a.	n.a.
Labelling degree $^{15}\text{N}$	%	n.a.	n.a.	<b>71.4</b>	<b>69.3</b>	n.a.	n.a.

<sup>a</sup>Assuming a  $^{15}\text{N}$  labelling degree of >95%, commonly observed in *E. coli* (Hoopes *et al.*, 2015).

benefit of a rapid protein synthesis. Furthermore, mid-term and long-term cell line adaptation may help to increase the biomass and volumetric product yield. We found that GB1 accumulation even increased when cultivating BY-2 cells in the presence of isotopes and that deuterium oxide doubled the target protein accumulation levels compared with  $^{15}\text{N}$ -labelled ammonium nitrate, although the cause of this improvement is unclear. The NMR spectra of the GB1 proteins expressed in *E. coli* and PCPs were almost identical, confirming that the plant-derived GB1 was properly folded and the 70% degree of  $^{15}\text{N}$  labelling (close to 100% efficacy given the fraction of labelled nitrogen) achieved in PCPs was sufficient for an NMR analysis. Importantly, deuteration will be necessary to analyse large proteins (>30 kDa; Grzesiek and Bax, 1992), so the current 16% degree of deuterium labelling must be improved to >70% (Sattler and Fesik, 1996), for example by increasing the volume fraction of deuterium and using a deuterated carbon source. We conclude that BY-2 cells can be used for the  $^{15}\text{N}$ -labelling and deuteration of recombinant proteins for NMR analysis using transient expression. This expands the utility of plants cells beyond the use of transgenic plants or plant cells reported before (Ohki *et al.*, 2008; Yagi *et al.*, 2015). Specifically, the fast high-throughput screening of labelling conditions and protein constructs in PCPs will be very beneficial for the high-level and rapid production of challenging isotope-labelled protein samples for subsequent NMR studies.

## Methods

### Protein production in *E. coli*

Amino acids 229–282 of *Streptococcus sp.* derived from immunoglobulin G-binding protein G, Uniprot ID P06654, additionally comprising the amino acids ME at the N-terminus and AAAHHHHHH at the C-terminus (GB1) was cloned into a modified pET21a plasmid (Merck, Germany) was transfected into *E. coli* BL21 (DE3) cells. Protein expression was induced at an  $\text{OD}_{600\text{nm}}$  of 0.6–0.8 with 1.0 mM IPTG and cells were shaken for 5 h at 37 °C. For the production of  $^2\text{H}$ ,  $^{15}\text{N}$ -labelled GB1, bacteria were grown in M9 medium supplemented with 1.0 g/L [98%  $^{15}\text{N}$ ] ammonium chloride and 2.0 g/L unlabeled glucose in 50% v/v  $\text{D}_2\text{O}$  (Merck). After harvesting by centrifugation, cells were resuspended in buffer A (50 mM Tris pH 8.0 and 200 mM sodium chloride) and GB1 was purified from the soluble fraction with a Ni-NTA gravity flow affinity column, equilibrated with buffer A. The beads were washed with buffer A and buffer A supplemented with 10 mM imidazole, followed by elution with buffer A containing 500 mM imidazole. Eluted GB1 was concentrated with a 3 kDa cut-off Amicon centrifugal device before size exclusion chromatography (SEC) on a HiLoad 16/600 Superdex 75 pg column in NMR buffer (20 mM sodium phosphate pH 7.0, 50 mM sodium chloride and 1.0 mM EDTA).

### Cloning of plant expression constructs

The GB1 coding sequence (amino acids 229–282 of *Streptococcus sp.* derived immunoglobulin binding protein G B1 domain, Uniprot ID P06654, additionally comprising the amino acids ME at the N-terminus and AAAHHHHHH at the C-terminus (i.e. identical amino acid sequence as the *E. coli*-produced GB1)) was codon-optimized for *Nicotiana benthamiana* and synthesized by BioCat (Heidelberg, Germany) with flanking 5' NcoI and 3' NotI sites for subcloning. The sequence was then inserted into pTRAc expression vectors featuring the tobacco mosaic virus omega prime sequence (omega), the tobacco etch virus leader sequence (TL), or the 5' untranslated region (5' UTR) of the *Petroselinum hortense* chalcone synthase gene (CHS) (Gengenbach *et al.*, 2020). GB1 was targeted to the cytosol (no leader peptide), the plastids (N-terminal leader peptide of the ribulose-1,5-bisphosphate carboxylase-oxygenase small subunit, rbcS (Maclean *et al.*, 2007)), the apoplast (leader sequence of antibody mAb24 heavy chain, LPH) or the endoplasmic reticulum (ER) (LPH combined with a C-terminal SEKDEL sequence; Munro and Pelham, 1987; Fischer *et al.*, 1999; Table 4). All constructs were assembled and propagated in *E. coli* strain DH5 $\alpha$  (restriction cloning) and introduced into competent *Agrobacterium tumefaciens* (*Rhizobium radiobacter*) GV3101:pMP90RK cells by electroporation (2400 V, 25  $\mu\text{F}$  and 200  $\Omega$ ) as previously described (Buyel *et al.*, 2013). Restriction enzymes, Q5 high-fidelity DNA polymerase and T4 DNA ligase (all from New England Biolabs, Ipswich, MA) were used according to the manufacturer's recommendations. DsRed constructs featuring the double enhanced 35S promoter, CHS 5' UTR and cytosol, plastid (chloroplast), ER or apoplast targeting elements described above were used as controls for all expression experiments (Gengenbach *et al.*, 2020). Plant-derived GB1 was fused to a C-terminal His<sub>6</sub> tag, whereas an *E. coli*-derived GB1 control contained an 18-residue N-terminal peptide with a His<sub>6</sub> tag (MSYYHHHHHHDY-DIPTTA).

### Plant cell cultivation

BY-2 cells were cultivated at 26 °C and ~50% relative humidity, shaking at 160 rpm in 100-mL shake flasks containing 12 mL of medium (Table 2) for 7–11 days until the next passage, which was carried out at a seeding cell density of  $2 \times 10^8$  cells/L for all cultures and conditions. The volume fraction of deuterium oxide in the medium was increased in one step (0%–100%  $^2\text{H}_2\text{O}$ ) or gradually (steps of 20%, 50%, 75% and 100%) at each passage, with 50% used during BY-2 cultivation for GB1 production and subsequent NMR analysis. Cells were concentrated to a wet mass of 200 g/L for the preparation of PCPs (~60 mg/PCP) by sedimentation as previously described (Gengenbach *et al.*, 2020).

### PCP preparation and infiltration with *A. tumefaciens*

Pre-cultures of *A. tumefaciens* (500 µL/well in 96-deepwell plates) were inoculated from glycerol stocks to an OD<sub>600nm</sub> of 0.04 in PAM4 medium (Houdelet *et al.*, 2017) containing 50 mg/L carbenicillin, 25 mg/L kanamycin and 25 mg/L rifampicin, and were incubated at 28 °C for 24 h, shaking at 1000 rpm. Main cultures were inoculated from the pre-cultures to an OD<sub>600nm</sub> of 0.10 and were cultivated under the same conditions. We applied an automated protocol to generate PCPs from BY-2 cells followed by infiltration with *A. tumefaciens* as previously described (Gengenbach *et al.*, 2020) using 100 µL infiltration solution per PCP (0.5 g/L Murashige-Skoog medium M0221 (Duchefa Biochemie, Haarlem, Netherlands), 50.0 g/L (146 mM) sucrose, 2.0 g/L (10 mM) glucose monohydrate, 0.0392 g/L (0.2 mM) acetosyringone, 2.928 g/L (15 mM) 2-(*N*-morpholino)ethanesulfonic acid (MES), pH 5.6) and an infiltration OD<sub>600nm</sub> of 0.4. The infiltration suspension was incubated for 1 h at 22°C before use and remained in contact with the PCPs for 1 h before removal by centrifugation. Infiltrated PCPs were incubated at 26 °C and 80% relative humidity for 96 h over a water reservoir as previously described (Gengenbach *et al.*, 2020).

### Protein extraction from BY-2 cells

Extraction was carried out 3 days (GB1 expression in different compartments) or 4 days (expression of labelled GB1) post-infiltration. Each 60-mg PCP was extracted with a 3 v/m ratio of extraction buffer (7.8 g/L (50 mM) sodium phosphate buffer, 1.9 g/L (10 mM) sodium bisulfite, 29.2 g/L (500 mM) sodium chloride, pH 8.0) in 1.2-mL collection microtube strips (Qiagen, Hilden, Germany) containing a single 3-mm steel bead per microtube, using an MM300 bead mill (Retsch, Han, Germany) at 28 Hz for 2 × 3 min (Gengenbach *et al.*, 2020) Extracts were clarified by centrifugation (5100 g for 8 min at 4 °C) and supernatants were stored at –20 °C. For preparative purposes, 24–72 PCPs were pooled for extraction and subsequent purification by chromatography. PCP extracts were passed through a 0.2-µm filter before the purification of GB1.

### Purification of GB1 by immobilized metal affinity chromatography (IMAC)

His<sub>6</sub>-tagged GB1 protein was purified on pre-packed Ni<sup>2+</sup>-charged 1-mL Chelating Sepharose FF columns (Cytiva, Uppsala, Sweden) using an ÄKTApure L system (Cytiva). The running buffer (20 mM sodium phosphate, 500 mM sodium chloride, pH 7.5) was supplemented with 30 mM imidazole for washing and 300 mM imidazole for elution. The flow rate was 0.5 mL/min (0.78 m/h) at all times (2 min contact time). An extract volume of 5 mL (BY-2 cells cultivated on medium D), 10 mL (BY-2 cells cultivated on medium H) or 15 mL (BY-2 cells cultivated on medium N or DN) was passed through a 0.2-µm filter and GB1 was purified by chromatography (Table 2).

### Sample analysis

The fraction of living BY-2 cells was determined every 7 days by Evans blue staining (Taylor and West, 1980) using a trinocular microscope (BMS, Breukhoven, Netherlands) equipped with a 40× objective and a 10× eyepiece. Briefly, 100-µL cell culture samples were mixed with 850 µL water and 50 µL 0.5% m/v Evans blue. Stained (dead) and total cells were immediately counted in two counting areas in a Fuchs Rosenthal chamber and then averaged. Of the 16 large squares of the chamber, we used

**Table 4** Overview of cloned constructs for *Agrobacterium*-mediated transient expression of GB1 in plant cells. All constructs contain a C-terminal His<sub>6</sub> tag including three alanine residues as a spacer

Construct ID	Vector	5' UTR	5' signal sequence	3' signal sequence	Targeted compartment
000190	pTRAc	CHS	None	None	Cytosol
000191	pTRAc	CHS	LPH	None	Apoplast
000192	pTRAc	CHS	TP	None	Plastids
000193	pTRAc	CHS	LPH	KDEL	ER
000194	pTRAc	Omega	None	None	Cytosol
000195	pTRAc	Omega	TP	None	Plastids
000196	pTRAc	TL	None	None	Cytosol
000197	pTRAc	TL	TP	None	Plastids

Signal peptide	Amino acid sequence	Mass [kDa]	Removal of signal peptide during processing
LPH	MEWSWIFLFLSGTAGVHS	2.1	Yes
TP	MASSVISSAAVATRTNVTQ AGSMIAPFTGLKSAATF PVSRRKQNLDTISIASNGGRVRA	5.8	Yes
KDEL	SEKDEL	0.7	No

CHS, *Petroselinum hortense* chalcone synthase gene 5' UTR, ER, endoplasmic reticulum, LPH, leader peptide of the antibody mAb24 heavy chain, omega, omega prime sequence from tobacco mosaic virus, SEKDEL, signal for target protein retention in the ER, TL, tobacco etch virus leader sequence, TP, transit peptide from the ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit of *Solanum tuberosum*.

the four squares forming the top left to bottom right diagonal plus the bottom left square for cell counting which was equivalent to a total counting volume of 1 µL. Live and dead cell concentrations were then calculated by multiplying live and dead cell counts by a factor of 10 to correct for the initial dilution and by a factor of 10<sup>6</sup> to convert the counts to a per litre basis. The growth  $\mu$  rate was calculated based on these cell concentrations  $c_{t,i}$  using Equation 1.

$$\mu = \frac{c_{t,j} - c_{t,i}}{t_j - t_i} \times \frac{1}{c_{t,i}} \quad (1)$$

where  $t_j$  and  $t_i$  are the time points at which the cell concentrations were determined with  $j > i$ .

Samples from extraction and purification steps were separated on 4%–12% Bis-Tris polyacrylamide gels before staining with SimplyBlue SafeStain (Thermo Fisher Scientific, Waltham, MA) as previously described (Arfi *et al.*, 2016) or analysis by Western blotting (Kastilan *et al.*, 2017). GB1 was detected using a rabbit-anti-His<sub>6</sub> primary antibody (0.1 mg/L, GenScript, Piscataway, NJ) and an alkaline phosphatase-labelled goat-anti-rabbit detection antibody (0.06 mg/L, Jackson ImmunoResearch, West Grove, PA). GB1 purity was determined by densitometric analysis using AIDA Image Analyser (Raytest, Straubenhardt, Germany) as previously described (Menzel *et al.*, 2016). The peak area for bands corresponding to GB1 and the total peak area of all bands (i.e. stained host cell proteins (HCPs) and GB1) in a sample were used to estimate GB1 purity. Samples were also analysed by dot blot (Opendenstein *et al.*, 2021) using ImageJ (Schneider *et al.*, 2012) for quantitative comparison to His<sub>6</sub>-tagged GB1 standards at

concentrations of 0.0, 1.5, 3.5, 7.5, 10.0, 20.0, 40 and 70.0 mg/L. DsRed fluorescence in the expression controls was quantified as previously described using 12 standards in the 0–225 mg/L range (Knödler *et al.*, 2019).

### Electrospray ionization mass spectrometry (ESI-MS)

GB1 samples were analysed using an LCQ-Fleet system (Thermo Fisher Scientific) equipped with a 3D ion trap and ESI device. The MS instrument was connected to an UltiMate 3000 high-performance liquid chromatography (HPLC) system (Dionex, Thermo Fisher Scientific) with a 10 × 2.1 mm HyperSep Retain PEP column (MZ Analysetechnik, Mainz, Germany). The labelling efficiency achieved for GB1 in heavy isotope-containing media was calculated using Equation 2.

$$E_L = \frac{m_L - m_R}{\Delta m_i \times n_i} \times 100\% \quad (2)$$

where  $E_L$  is the labelling efficiency,  $m_L$  is the mass of the protein obtained from cells grown in labelled cultivation medium [Da],  $m_R$  is the mass of the protein obtained from a reference medium [Da],  $\Delta m_i$  is the mass difference between the common and heavy isotope [Da] and  $n_i$  is the number of atoms of the element in the protein, that is 88 for N and 362 for H (only non-exchangeable protons (Hoopes *et al.*, 2015) are considered). For  $^{15}\text{N}$  labelling,  $\Delta m_i$  was 1 Da and the two pairs of media were H (reference)/N (labelled) and D (reference)/DN (labelled), for  $^2\text{H}$  labelling,  $\Delta m_i$  was also 1 Da and the two pairs of media were H (reference)/D (labelled) and N (reference)/DN (labelled).

### NMR spectroscopy

For NMR experiments, plastid-targeted GB1 extracted from cells grown in medium DN was exchanged into NMR buffer (20 mM sodium phosphate pH 7.0, 50 mM sodium chloride, 1 mM EDTA) using a 3-kDa cut-off Amicon centrifugal device (MerckMillipore, Burlington, MA) and the  $^2\text{H}$ ,  $^{15}\text{N}$ -labelled GB1 was concentrated to 1  $\mu\text{M}$  in the same device. For NMR experiments, we used a Bruker Avance III HD instrument (Bruker, Billerica, MA) equipped with a cryogenic probe operating at 900 MHz proton frequency. A 2D- $^{15}\text{N}$ ,  $^1\text{H}$ -heteronuclear single quantum coherence (HSQC) experiment was recorded at 313 K with 2800 transients and 74 increments in the indirect  $^{15}\text{N}$  dimension. Furthermore, a reference HSQC spectrum with [50%]- $^2\text{H}$ , [98%]- $^{15}\text{N}$ -labelled GB1 at a concentration of 200  $\mu\text{M}$  at 313 K with 12 transients and 256  $^{15}\text{N}$ -increments was measured on a Bruker Avance III HD NMR spectrometer operating at 800 MHz proton frequency (Bruker, Billerica, MA) equipped with a cryogenic probe.

### Statistical data analysis

GB1 accumulation in BY-2 PCPs cultivated in different media was compared using two-sided two-sample  $t$ -tests. Shapiro–Wilk and  $F$ -tests were used to assess normality and homoscedasticity, respectively. Welch correction was applied to samples with unequal variances. An  $\alpha$ -level of 0.05 was selected as the significance threshold.

### Acknowledgements

We thank Dr. Henrik Nausch for cell culture support and Dr. Richard M Twyman for editorial assistance. NMR experiments were conducted at the Bavarian NMR Center (BNMRZ) supported by the Technical University of Munich and Helmholtz Munich. We

thank Drs. Gerd Gemmecker and Sam Asami (TUM) for general NMR support. Open Access funding enabled and organized by Projekt DEAL.

### Funding

This work was funded in part by the Fraunhofer-Gesellschaft Internal Programs under grant no. Attract 125–600164, and the state of North-Rhine-Westphalia under the Leistungszentrum grant no. 423 ‘Networked, adaptive production’. The work was also supported by the Deutsche Forschungsgemeinschaft (DFG) in the framework of the Research Training Group ‘Tumor-targeted Drug Delivery’ grant 331065168. L.E.S. and M.M. are supported by the DFG Collaborative Research Center 1035 (CRC1035, project number 201302640, project B13). F.H. acknowledges support from Helmholtz Munich and the Helmholtz Society (grant no. VG-NG-1039) and the Center for Integrated Protein Science Munich (CIPS<sup>M</sup>).

### Conflict of interest

The authors have no conflict of interest to declare.

### Author contributions

PO, LS, MM, FH and JB designed the experiments. PO, LS and MM conducted the experiments and analysed the data. FH and JB supported data analysis. PO, LS and MM wrote the manuscript. FH and JB revised the manuscript and secured the funding.

### Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

### References

- Arfi, Z.A., Hellwig, S., Drossard, J., Fischer, R. and Buyel, J.F. (2016) Polyclonal antibodies for specific detection of tobacco host cell proteins can be efficiently generated following RuBisCO depletion and the removal of endotoxins. *Biotechnol. J.* **11**(4), 507–518.
- Bindschedler, L.V., Palmblad, M. and Cramer, R. (2008) Hydroponic isotope labelling of entire plants (HILEP) for quantitative plant proteomics; an oxidative stress case study. *Phytochemistry*, **69**(10), 1962–1972.
- Bouvignies, G., Markwick, P., Brüschweiler, R. and Blackledge, M. (2006) Simultaneous determination of protein backbone structure and dynamics from residual dipolar couplings. *J. Am. Chem. Soc.* **128**(47), 15100–15101.
- Buyel, J.F., Kaefer, T., Buyel, J.J. and Fischer, R. (2013) Predictive models for the accumulation of a fluorescent marker protein in tobacco leaves according to the promoter/5'UTR combination. *Biotechnol. Bioeng.* **110**(2), 471–482.
- Chao, F.-A. and Byrd, R.A. (2020) Protein dynamics revealed by NMR relaxation methods. *Emerging Top. Life Sci.* **2**(1), 93–105.
- Cheng, Y. and Patel, D.J. (2004) An efficient system for small protein expression and refolding. *Biochem. Biophys. Res. Commun.* **317**(2), 401–405.
- Clark, L., Dikiy, I., Rosenbaum, D.M. and Gardner, K.H. (2018) On the use of *Pichia pastoris* for isotopic labelling of human GPCRs for NMR studies. *J. Biomol. NMR*, **71**(4), 203–211.
- Di Fiore, S., Li, Q., Leech, M.J., Schuster, F., Emans, N., Fischer, R. *et al.* (2002) Targeting tryptophan decarboxylase to selected subcellular compartments of tobacco plants affects enzyme stability and in vivo function and leads to a lesion-mimic phenotype. *Plant Physiol.* **129**(3), 1160–1169.
- Duff, A.P., Wilde, K.L., Rekas, A., Lake, V. and Holden, P.J. (2015) Chapter one - robust high-yield methodologies for  $^2\text{H}$  and  $^2\text{H}/^{15}\text{N}/^{13}\text{C}$  labelling of proteins for structural investigations using neutron scattering and NMR. In

- Methods in Enzymology: Isotope Labelling of Biomolecules – Labelling Methods* (Kelman, Z., ed), pp. 3–25. Cambridge, MA: Academic Press.
- Dutta, A., Saxena, K., Schwalbe, H. and Klein-Seetharaman, J. (2012) Isotope labelling in mammalian cells. In *Protein NMR techniques* (Shekhtman, A. and Burz, D.S., eds), pp. 55–69. Totowa, NJ: Humana Press.
- Egorova-Zachernyuk, T.A., Bosman, G.J. and DeGrip, W.J. (2011) Uniform stable-isotope labelling in mammalian cells: formulation of a cost-effective culture medium. *Appl. Microbiol. Biotechnol.* **89**, 397–406.
- Eiberle, M.K. and Jungbauer, A. (2010) Technical refolding of proteins: do we have freedom to operate? *Biotechnol. J.* **5**(6), 547–559.
- Favot, L., Gillingwater, M., Scott, C. and Kemp, P.R. (2005) Overexpression of a family of RPEL proteins modifies cell shape. *FEBS Lett.* **579**(1), 100–104.
- Filiou, M.D., Varadarajulu, J., Teplytska, L., Reckow, S., Maccarrone, G. and Turck, C.W. (2012) The <sup>15</sup>N isotope effect in *Escherichia coli*: a neutron can make the difference. *Proteomics*, **12**(21), 3121–3128.
- Fischer, R., Schumann, D., Zimmermann, S., Drossard, J., Sack, M. and Schillberg, S. (1999) Expression and characterization of bispecific single-chain Fv fragments produced in transgenic plants. *Eur. J. Biochem.* **262**(3), 810–816.
- Fogel, M.L., Griffin, P.L. and Newsome, S.D. (2016) Hydrogen isotopes in individual amino acids reflect differentiated pools of hydrogen from food and water in *Escherichia coli*. *Proc. Natl Acad. Sci.* **113**(32), E4648–E4653.
- Gardner, K.H. and Kay, L.E. (1998) The use of <sup>2</sup>H, <sup>13</sup>C, <sup>15</sup>N multidimensional NMR to study the structure and dynamics of proteins. *Annu. Rev. Biophys. Biomol. Struct.* **27**, 357–406.
- Gengenbach, B.B., Keil, L.L., Opendenstein, P., Müschen, C.R., Melmer, G., Lentzen, H., Bührmann, J. et al. (2019) Comparison of microbial and transient expression (tobacco plants and plant-cell packs) for the production and purification of the anticancer mistletoe lectin viscumin. *Biotechnol. Bioeng.* **116**(9), 2236–2249.
- Gengenbach, B.B., Müschen, C.R. and Buyel, J.F. (2018) Expression and purification of human phosphatase and Actin regulator 1 (PHACTR1) in plant-based systems. *Protein Expr. Purif.* **151**, 46–55.
- Gengenbach, B.B., Opendenstein, P. and Buyel, J.F. (2020) Robot cookies – plant cell packs as an automated high-throughput screening platform based on transient expression. *Front. Bioeng. Biotechnol.* **8**, 393.
- Goto, N.K. and Kay, L.E. (2000) New developments in isotope labelling strategies for protein solution NMR spectroscopy. *Curr. Opin. Struct. Biol.* **10** (5), 585–592.
- Grabsztunowicz, M., Koskela, M.M. and Mulo, P. (2017) Post-translational modifications in regulation of chloroplast function: recent advances. *Front. Plant Sci.* **8**, 240.
- Gronenborn A.M., Filpula D.R., Essig N.Z., Aniruddha A., Marc W. & Wingfield Paul T. (1991) A novel, highly stable fold of the immunoglobulin binding domain of streptococcal protein G. *Science*, **253**(5020), 657–661.
- Grzesiek, S. and Bax, A.D. (1992) Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. *J. Magn. Reson.* **96**, 432–440.
- Gupta, R. and Brunak, S. (2002) Prediction of glycosylation across the human proteome and the correlation to protein function. *Pac. Symp. Biocomput.* **7**, 310–322.
- Hamby, S.E. and Hirst, J.D. (2008) Prediction of glycosylation sites using random forests. *BMC Bioinform.* **9**(1), 500.
- Hansen, A.P., Petros, A.M., Meadows, R.P., Nettesheim, D.G., Mazar, A.P., Olejniczak, E.T., Xu, R.X. et al. (1994) Solution structure of the amino-terminal fragment of urokinase-type plasminogen activator. *Biochemistry*, **33**, 4847–4864.
- Holland, T., Sack, M., Rademacher, T., Schmale, K., Altmann, F., Stadlmann, J. et al. (2010) Optimal nitrogen supply as a key to increased and sustained production of a monoclonal full-size antibody in BY-2 suspension culture. *Biotechnol. Bioeng.* **107**(2), 278–289.
- Hoopes, J.T., Elberson, M.A., Preston, R.J., Reddy, P.T. and Kelman, Z. (2015) Chapter two – protein labelling in *Escherichia coli* with <sup>2</sup>H, <sup>13</sup>C, and <sup>15</sup>N. In *Methods in Enzymology: Isotope Labelling of Biomolecules - Labelling Methods* (Kelman, Z., ed), pp. 27–44. Cambridge, MA: Academic Press.
- Houdelet, M., Galinski, A., Holland, T., Wenzel, K., Schillberg, S. and Buyel, J.F. (2017) Animal component-free agrobacterium tumefaciens cultivation media for better GMP-compliance increases biomass yield and pharmaceutical protein expression in *Nicotiana benthamiana*. *Biotechnol. J.* **12**(4), 1600721.
- Ikegawa, H., Yamamoto, Y. and Matsumoto, H. (2000) Responses to aluminium of suspension-cultured tobacco cells in a simple calcium solution. *Soil Sci. Plant Nutr.* **46**(2), 503–514.
- Ikeya, T., Hanashima, T., Hosoya, S., Shimazaki, M., Ikeda, S., Mishima, M., Güntert, P. et al. (2016) Improved in-cell structure determination of proteins at near-physiological concentration. *Sci. Rep.* **6**(1), 38312.
- Kastilan, R., Boes, A., Spiegel, H., Voepel, N., Chudobová, I., Hellwig, S., Buyel, J.F. et al. (2017) Improvement of a fermentation process for the production of two PfAMA1-DiCo-based malaria vaccine candidates in *Pichia pastoris*. *Sci. Rep.* **7**(1), 11991.
- Kim, B.M., Lotter-Stark, H.C.T., Rybicki, E.P., Chikwamba, R.K. and Palmer, K.E. (2018) Characterization of the hypersensitive response-like cell death phenomenon induced by targeting antiviral lectin griffithsin to the secretory pathway. *Plant Biotechnol. J.* **16**(10), 1811–1821.
- Kirsten Frank, M., Dyda, F., Dobrodumov, A. and Gronenborn, A.M. (2002) Core mutations switch monomeric protein GB1 into an intertwined tetramer. *Nat. Struct. Biol.* **9**(11), 877–885.
- Kiss, R., Fizil, Á. and Szántay, C. (2018) What NMR can do in the biopharmaceutical industry. *J. Pharm. Biomed. Anal.* **147**, 367–377.
- Knödler, M., Rühl, C., Emonts, J. and Buyel, J.F. (2019) Seasonal weather changes affect the yield and quality of recombinant proteins produced in transgenic tobacco plants in a greenhouse setting. *Front. Plant Sci.* **10**, 1245.
- Kushner, D.J., Baker, A. and Dunstall, T.G. (1999) Pharmacological uses and perspectives of heavy water and deuterated compounds. *Can. J. Physiol. Pharmacol.* **77**(2), 79–88.
- KyungRyun, L. and Ho, L.J. (2020) Stable isotope labelling of proteins in mammalian cells. *J. Korean Magnetic Resonance Soc.* **24**(3), 77–85.
- Lehtimäki, N., Koskela, M.M. and Mulo, P. (2015) Posttranslational modifications of chloroplast proteins: an emerging field. *Plant Physiol.* **168** (3), 768–775.
- Li, F., Grishaev, A., Ying, J. and Bax, A. (2015) Side chain conformational distributions of a small protein derived from model-free analysis of a large set of residual dipolar couplings. *J. Am. Chem. Soc.* **137**(46), 14798–14811.
- Li, Y.D., Lamour, G., Gsponer, J., Zheng, P. and Li, H. (2012) The molecular mechanism underlying mechanical anisotropy of the protein GB1. *Biophys. J.* **103**(11), 2361–2368.
- Maclean, J., Koekemoer, M., Olivier, A.J., Stewart, D., Hitzeroth, I.I., Rademacher, T., Fischer, R. et al. (2007) Optimization of human papillomavirus type 16 (HPV-16) L1 expression in plants: comparison of the suitability of different HPV-16 L1 gene variants and different cell-compartment localization. *J. Gen. Virol.* **88**, 1460–1469.
- Marley, J., Lu, M. and Bracken, C. (2001) A method for efficient isotopic labelling of recombinant proteins. *J. Biomol. NMR*, **20**(1), 71–75.
- Matthies, D., Bae, C., Toombes, G.E., Fox, T., Bartesaghi, A., Subramaniam, S. and Swartz, K.J. (2018) Single-particle cryo-EM structure of a voltage-activated potassium channel in lipid nanodiscs. *Elife*, **7**, e37558.
- Menzel, S., Holland, T., Boes, A., Spiegel, H., Bolzenius, J., Fischer, R. and Buyel, J.F. (2016) Optimized blanching reduces the host cell protein content and substantially enhances the recovery and stability of two plant-derived malaria vaccine candidates. *Front. Plant Sci.* **7**, 159.
- Miller, J.H. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
- Mitri, E., Barbieri, L., Vaccari, L. and Luchinat, E. (2018) <sup>15</sup>N isotopic labelling for in-cell protein studies by NMR spectroscopy and single-cell IR synchrotron radiation FTIR microscopy: a correlative study. *Anal. J. Royal Soc. Chem.* **143**, 1171–1181.
- Munro, S. and Pelham, H.R.B. (1987) A C-terminal signal prevents secretion of luminal ER proteins. *Cell*, **48**(5), 899–907.
- Müsken, M., Di Fiore, S., Römling, U. and Häussler, S. (2010) A 96-well-plate-based optical method for the quantitative and qualitative evaluation of *Pseudomonas aeruginosa* biofilm formation and its application to susceptibility testing. *Nat. Protoc.* **5**(8), 1460–1469.
- Mycroft-West, C., Su, D., Elli, S., Li, Y., Guimond, S. & Miller, G. et al. (2020) The 2019 coronavirus (SARS-CoV-2) surface protein (spike) S1 receptor binding domain undergoes conformational change upon heparin binding. *BioRxiv*, 2020.02.29.971093. Available from: <https://doi.org/10.1101/2020.02.29.971093>.

- Ohki, S., Dohi, K., Tamai, A., Takeuchi, M. and Mori, M. (2008) Stable-isotope labelling using an inducible viral infection system in suspension-cultured plant cells. *J. Biomol. NMR*, **42**, 271–277.
- Opdensteinen, P., Dietz, S.J., Gengenbach, B.B. and Buyel, J.F. (2021) Expression of biofilm-degrading enzymes in plants and automated high-throughput activity screening using experimental *Bacillus subtilis* biofilms. *Front. Bioeng. Biotechnol.* **9**, 714.
- Opdensteinen, P., Meyer, S. and Buyel, J.F. (2021) *Nicotiana* spp. for the expression and purification of functional IgG3 antibodies directed against the *Staphylococcus aureus* alpha toxin. *Front. Chem. Eng.* **3**, 50.
- Phyo, P., Zhao, X., Templeton, A.C., Xu, W., Cheung, J.K. and Su, Y. (2021) Understanding molecular mechanisms of biologics drug delivery and stability from NMR spectroscopy. *Adv. Drug Deliv. Rev.* **174**, 1–29.
- Rademacher, T., Sack, M., Blessing, D., Fischer, R., Holland, T. and Buyel, J. (2019) Plant cell packs: a scalable platform for recombinant protein production and metabolic engineering. *Plant Biotechnol. J.* **17**(8), 1560–1566.
- Riss, T.L., Moravec, R.A., Niles, A.L., Duellman, S., Benink, H.A., Worzella, T.J. et al. (2004) Cell viability assays. In *Assay Guidance Manual* (Markossian, S., Grossman, A., Brimacombe, K., Arkin, M., Auld, D., Austin, C.P. et al., eds). Bethesda, MD: Eli Lilly & Company and the National Center for Advancing Translational Sciences.
- Sanchez-Bragado, R., Serret, M.D., Marimon, R.M., Bort, J. and Araus, J.L. (2019) The hydrogen isotope composition  $\delta^2\text{H}$  reflects plant performance. *Plant Physiol.* **180**(2), 793–812.
- Sastry, M., Xu, L., Georgiev, I.S., Bewley, C.A., Nabel, G.J. and Kwong, P.D. (2011) Mammalian production of an isotopically enriched outer domain of the HIV-1 gp120 glycoprotein for NMR spectroscopy. *J. Biomol. NMR*, **50**, 197–207.
- Sattler, M. and Fesik, S.W. (1996) Use of deuterium labelling in NMR: overcoming a sizeable problem. *Structure*, **4**(11), 1245–1249.
- Sattler, M., Schleucher, J. and Griesinger, C. (1999) Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. Nucl. Magn. Reson. Spectrosc.* **34**(2), 93–158.
- Schneider, C.A., Rasband, W.S. and Eliceiri, K.W. (2012) NIH image to ImageJ: 25 years of image analysis. *Nat. Methods*, **9**(7), 671–675.
- Shen, B.-R., Zhu, C.-H., Yao, Z., Cui, L.-L., Zhang, J.-J., Yang, C.-W., He, Z.H. et al. (2017) An optimized transit peptide for effective targeting of diverse foreign proteins into chloroplasts in rice. *Sci. Rep.* **7**(1), 46231.
- Skora, L., Shrestha, B. and Gossert, A.D. (2015) Chapter eleven – isotope labelling of proteins in insect cells. In *Methods in Enzymology: Isotope Labelling of Biomolecules - Labelling Methods* (Kelman, Z., ed), pp. 245–288. Cambridge, MA: Academic Press.
- Streatfield, S.J. (2007) Approaches to achieve high-level heterologous protein production in plants. *Plant Biotechnol. J.* **5**(1), 2–15.
- Sun, B., Feng, D., Chu, M.L.-H., Fish, I., Lovera, S., Sands, Z.A. et al. (2021) Crystal structure of dopamine D1 receptor in complex with G protein and a non-catechol agonist. *Nat. Commun.* **12**(1), 3305.
- Takahashi, H. and Shimada, I. (2010) Production of isotopically labeled heterologous proteins in non-*E. coli* prokaryotic and eukaryotic cells. *J. Biomol. NMR*, **46**, 3–10.
- Taylor, J.A. and West, D.W. (1980) The use of Evan's blue stain to test the survival of plant cells after exposure to high salt and high osmotic pressure. *J. Exp. Bot.* **31**, 571–576.
- Wand, A.J. (2001) Dynamic activation of protein function: a view emerging from NMR spectroscopy. *Nat. Struct. Biol.* **8**(11), 926–931.
- Yagi, H., Fukuzawa, N., Tasaka, Y., Matsuo, K., Zhang, Y., Yamaguchi, T., Kondo, S. et al. (2015) NMR-based structural validation of therapeutic antibody produced in *Nicotiana benthamiana*. *Plant Cell Rep.* **34**(6), 959–968.
- Yao, J., Dyson, H.J. and Wright, P.E. (1997) Chemical shift dispersion and secondary structure prediction in unfolded and partly folded proteins. *FEBS Lett.* **419**(2–3), 285–289.
- Zuiderweg, E.R.P. (2002) Mapping protein–protein interactions in solution by NMR spectroscopy. *Biochemistry*, **41**(1), 1–7.

## Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Figure S1** GB1 expression in different plant cell compartments in dependence of the carbon source used during BY-2 cell cultivation.

**Figure S2** Cell viability in media and PCPs.

**Table S1** Comparison of GB1 accumulation levels in PCPs prepared from BY-2 cells grown in different cultivation media.