

Auto-inducing media for uniform isotope labeling of proteins with ^{15}N , ^{13}C and ^2H

Nicolas Guthertz¹ · Julia Klopp² · Aurélie Winterhalter² · César Fernández² · Alvar D. Gossert² 

Received: 25 February 2015 / Accepted: 10 April 2015 / Published online: 19 April 2015
© Springer Science+Business Media Dordrecht 2015

Abstract Auto-inducing media for protein expression offer many advantages like robust reproducibility, high yields of soluble protein and much reduced workload. Here, an auto-inducing medium for uniform isotope labelling of proteins with ^{15}N , ^{13}C and/or ^2H in *E. coli* is presented. So far, auto-inducing media have not found widespread application in the NMR field, because of the prohibitively high cost of labeled lactose, which is an essential ingredient of such media. Here, we propose using lactose that is only selectively labeled on the glucose moiety. It can be synthesized from inexpensive and readily available substrates: labeled glucose and unlabeled activated galactose. With this approach, uniformly isotope labeled proteins were expressed in unattended auto-inducing cultures with incorporation of ^{13}C , ^{15}N of 96.6 % and ^2H , ^{15}N of 98.8 %. With the present protocol, the NMR community could profit from the many advantages that auto-inducing media offer.

Keywords Auto-induction · Protein expression · Isotope labeling · NMR · Deuteration · ^{13}C -labeling · Lactose

Introduction

Contemporary NMR studies of proteins are based on a variety of isotope labeling protocols, which enable to highly enrich proteins with the stable isotopes ^2H , ^{13}C and

^{15}N in uniform manner or with selective patterns (McIntosh and Dahlquist 1989; Muchmore et al. 1989). The most widely used expression host for isotope labeled proteins are *E. coli* bacteria, which are able to grow on minimal media with defined sources of stable isotopes of nitrogen (e.g. $^{15}\text{NH}_4\text{Cl}$), carbon (e.g. ^{13}C -glucose or ^{13}C -glycerol) and hydrogen ($^2\text{H}_2\text{O}$ for ~80 % incorporation of ^2H , or $^2\text{H}_2\text{O}$ in combination with ^2H -glucose or ^2H -glycerol for full incorporation) (McIntosh and Dahlquist 1989; Muchmore et al. 1989; LeMaster 1989; Gardner and Kay 1998).

Auto-induction as an alternative to IPTG

The most popular system for producing heterologous proteins in bacteria is based on induction of protein expression under control of the *lac* operon (Jacob and Monod 1961; Dickson et al. 1975). This induction system takes advantage of elements of the natural lactose catabolism system of *E. coli*. As soon as lactose is present in the bacterial growth medium, it induces translation of several genes involved in the active intake of lactose through the cell membrane (*lacY*) (Kaback et al. 2001; Abramson et al. 2003) and the metabolization (e.g. *lacZ*) of lactose into small high-energy molecules. For the purpose of heterologous protein expression, a copy of the *lac* promoter is added to the 5' end of the target gene. The target protein will therefore be expressed in presence of lactose, which can be added to a culture during the growth phase once a substantial cell density is reached. There are two major refinements to this initial system: (1) replacing lactose as an inducer with a non-metabolizable analog, typically iso-propyl-thiogalactoside (IPTG), leading to stronger induction of protein expression and (2) letting the *lac* promoter act on T7 RNA polymerase instead of controlling expression of the target gene directly. T7 RNA polymerase will transcribe the gene

✉ Alvar D. Gossert
alvar.gossert@novartis.com

¹ Present Address: Division of Structural Biology, Institute of Cancer Research, London, UK

² Novartis Institutes for BioMedical Research, Basel, Switzerland

of interest with a T7 promoter at a much higher rate than the endogenous *E. coli* RNA polymerases, therefore boosting expression rates (Studier and Moffatt 1986). The T7-based system has become the standard system for high-level protein production.

While IPTG has become the standard inducer in laboratories focused on isotope labeling, lactose remains an important alternative due to its low cost and the possibility of formulating auto-inducing media with it. In auto-inducing media, protein expression is initiated at a pre-defined time point during the growth of a culture. The scientist therefore only needs to inoculate the culture and can leave it alone without monitoring, to return the next day to harvest the culture containing expressed protein (Studier 2005).

Auto-inducing media offer several advantages over IPTG induction considering the workload associated when running an expression culture, but also concerning the quality of the results, in terms of reproducibility and levels of soluble protein expression. The workload is alleviated because no constant monitoring of the culture by measuring OD₆₀₀ is needed in order to add IPTG at an appropriate time point. Especially with cultures in minimal media, lag times are variable and growth can be slow, therefore induction often needs to be triggered during nighttime. On the quality side, auto-inducing media offer advantages, since the time point of induction is very reproducible, rendering the whole process of protein expression more robust, while in manual induction the optimal time point is sometimes missed. Moreover, in a high throughput setup it is typically needed to run several cultures in parallel or to perform a high number of parallel small-scale expression tests. Under these circumstances, using auto-inducing media provides numerous advantages in terms of flexibility when planning and performing the experiments and increases efficiency. Regarding protein expression, the levels of soluble protein, especially for difficult proteins, tend to be higher with auto-inducing media. This is probably due to prolonged periods of expression at lower intensity and higher final cell densities in general. The principal carbon source in auto-inducing media is glycerol, which doesn't lead to acidification of the growing culture, as does glucose. Therefore rather high cell densities can be obtained in shake flasks (final OD₆₀₀ > 5). Because of all these advantages, most of the proteins for X-ray studies in high throughput setups, including our labs at Novartis, are expressed using auto-inducing media (Studier 2005; Sreenath et al. 2005; Peti and Page 2007; Assenberg et al. 2013).

The mechanism of auto-induction

For auto-inducing cultures the same vectors and cell strains as for IPTG-based induction can be used, e.g. BL21 (DE3)

strains and vectors with a T7 promoter, only the medium needs to be specially formulated. Auto-inducing media are based on using a mixture of carbohydrates, namely glycerol, lactose and glucose (Fig. 1). Glycerol acts as a general carbon source and is consumed steadily during the entire time of the cell culture. Lactose is the natural inducer of the *lac* operator and will eventually induce protein expression. Lactose is also consumed, notably in a glucose dependent manner, and should be present in high enough amounts in order to ensure robust induction of protein expression until the end of the culture. Glucose is a natural inhibitor of the *lac* operon and therefore acts as a repressor of protein expression during the initial growth phase of the culture and at the same time represses uptake and consumption of lactose. Once the glucose is used up, *E. coli* switches to the less preferred lactose as the primary carbon source. Thereby the *lac* operon is activated and with it heterologous protein expression in cells with the target protein or T7 RNA polymerase under control of this operon. Since glucose is consumed at a rate at about 0.8 g per OD₆₀₀ unit (g/OD₆₀₀), the point of induction can be precisely controlled from the onset of the culture. For example, adding 0.5 g of glucose to the auto-inducing mixture will result in induction at about OD₆₀₀ = 0.6, using 1 g at OD₆₀₀ = 1.2. F. W. Studier formulated a generally applicable auto-inducing carbohydrate mixture consisting of 5 g/l of glycerol, 0.5 g/l of glucose and 2 g/l of lactose (Studier 2005). Using these carbon sources in an M9 background will lead to induction at OD₆₀₀ = 0.6 (Fig. 1).

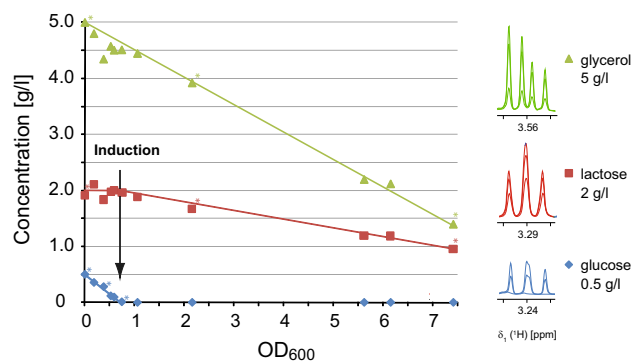


Fig. 1 Consumption of glycerol, lactose and glucose in auto-inducing medium. Glycerol (green) acts as a general carbon source. It is steadily consumed at a constant rate of about 0.6 g per OD₆₀₀ unit (g/OD₆₀₀) in an M9-based auto-induction medium. Lactose (red) serves both as a carbon source and also as the inducer of protein expression (~0.5 g/OD₆₀₀). Glucose (blue) in turn, inhibits the *lac* operon and suppresses protein expression when it is present (~0.8 g/OD₆₀₀). Once glucose is consumed, the *lac* operon is activated by lactose and protein expression is induced (indicated with arrow labeled “Induction”). Concentrations of the individual carbohydrates were monitored by 1D ¹H NMR, as illustrated by selected spectral excerpts on the right. The linear traces were added manually. The initial concentrations of the carbohydrates are indicated on the right

Auto-induction for isotope labeling

Setting up auto-induction protocols for uniform ^{15}N labeling and fractional deuteration is straightforward. The standard auto-inducing carbohydrate mixture supplemented with $^{15}\text{NH}_4\text{Cl}$ as a single nitrogen source can be used to produce ^{15}N labeled protein in auto-inducing media with excellent ^{15}N incorporation (Studier 2005; Tyler et al. 2005). When replacing H_2O with D_2O , fractional deuteration can be achieved with comparable incorporation as with traditional M9 media containing a protonated carbon source.

However, for uniform ^{13}C and ^2H labeling, all three components of the auto-induction mixture need to be labeled (LeMaster 1989). ^{13}C - and/or ^2H -labeled glucose and glycerol can be readily purchased from multiple commercial sources. For lactose, however, custom synthesis is required, which according to the offers we obtained, would bring the price per liter of auto-inducing medium above 10,000 USD for ^{13}C labeling and above 20,000 USD for simultaneous ^2H and ^{13}C labeling. Alternatively, if unlabeled lactose is used, incorporation of isotopes will be maximally 80 %, which is not sufficient for many applications like, for instance, recording of triple resonance experiments or ^{13}C TOCSY-based experiments (Cavanagh et al. 1996).

Here we present a direct way to reduce the cost of labeled lactose and therefore of auto-inducing media. Lactose is a disaccharide consisting of a galactose and a glucose moiety. Labeled glucose is relatively inexpensive, thus galactose is the main cause for the high price of labeled lactose. Interestingly, the most commonly used *E. coli* expression strain, BL21, has a deficiency in the galactose metabolism as indicated by the “gal” phenotype. Therefore, the galactose subunits of lactose are not metabolized and its carbon and hydrogen atoms are not incorporated directly into target proteins. This means that only the glucose subunit would need to be labeled, making it possible to use more economic selectively glucose-labeled lactose.

We show high incorporation (97–99 %) of ^{13}C and ^2H into proteins, using auto inducing media based on lactose with an appropriately labeled glucose moiety. The protocol is robust as protein production is induced with a high amount of lactose and it is generally applicable. Additionally to the high quality in terms of incorporation and robustness, all other advantages of auto-inducing media concerning reproducibility, convenience and expression of high levels of soluble protein are fully conserved.

Materials and methods

Synthesis of lactose

To synthesize lactose with a ^{13}C -labeled glucose moiety a commercial kit from Sigma was used as a reference for the

galactosyl transferase reaction [$\beta(1 \rightarrow 4)$ Galactosyltransferase Kit Sigma Life Science, Prod Nr. 59,505]. For lactose synthesis at preparative scale, a modified reaction system was established using a carbon-free buffer: stock solutions of the reagents glucose (Sigma-Aldrich), $\beta(1 \rightarrow 4)$ galactosyltransferase (GalT, Sigma-Aldrich) and α -lactalbumin (Sigma-Aldrich) were prepared in reaction buffer (12 mM borate pH 7.4, 30 mM Mn^{2+}). UDP-galactose (Carbosynth) was prepared without Mn^{2+} , because it precipitated in presence of Mn^{2+} . The carbon-free buffer, pH, time, temperature, volume, antioxidant, magnesium and different amounts of each component were optimized in order to improve the yield of the reaction. Optimizations were performed at 100 μl scale in 1.5 ml Eppendorf tubes under agitation (1400 rpm).

To assess the yields of the reaction ^1H NMR spectroscopy was used, where glucose and lactose signals are easily distinguishable. In order to be able to measure spectra of the reaction solution, Mn^{2+} was removed by $\text{H}_2\text{K}_2\text{PO}_4$ precipitation.

The resulting reaction conditions were: 100 mM glucose, 160 mM UDP-galactose, 12 mM borate pH 7.4, 30 mM MnCl_2 , 125 $\mu\text{g/ml}$ $\beta(1 \rightarrow 4)$ GalT, 500 $\mu\text{g/ml}$ α -lactalbumin and alkaline phosphatase. All ingredients, except for $\beta(1 \rightarrow 4)$ GalT, were mixed well. To start the reaction the $\beta(1 \rightarrow 4)$ GalT enzyme was added and the mixture was incubated at 37 °C for 8 h. The reaction was stopped by heating to 90 °C for 2 min.

The efficiency of the reaction was 65 % in molarity of glucose, i.e. with 1 g of glucose, 1.3 g of lactose were produced, and 0.35 g did not react. The limitation can be explained by the creation of pyrophosphate during the reaction, which co-precipitates with manganese, potentially depleting it and stopping the reaction.

Purification of lactose

Since at high pH glucose and lactose were the only two non-charged substances in the resulting reaction mixture, quantitative purification was possible with ion-exchange chromatography. The pH of the samples was adjusted to 10. The mixture was run over a SP-Sepharose column in order to remove Mn^{2+} and subsequently over a Q-Sepharose column to remove uridine, PP_i and unreacted UDP-galactose.

Expression of proteins

Pre-culture: A colony of freshly transformed cells grown on LB agar with the appropriate antibiotics was used to inoculate a 2 ml LB culture. After 2 h incubation at 30 °C, cells were centrifuged and transferred into a 10 ml pre-culture of carbohydrate-free base medium containing 8 g/l

of glucose as a single carbon source. For the labeling experiments glucose enriched with ^{13}C or ^2H was used (98 % Sigma). The carbohydrate-free base medium contained 50 mM KH_2PO_4 , 50 mM Na_2HPO_4 , 5 mM Na_2SO_4 , 2 mM MgSO_4 and 2.5 g/l $^{15}\text{NH}_4\text{Cl}$ in H_2O or $^2\text{H}_2\text{O}$. The cells were grown over night at 30 °C.

Main culture: Main cultures were inoculated with 1 % v/v of the over-night pre-culture. The main culture medium consisted of carbohydrate-free base medium, to which appropriately labeled auto-inducing carbohydrate mixture was added. Cells were grown at 37 °C for 3 h, then the temperature was lowered to 18 °C for over night expression. Cells were harvested by centrifugation at $6000 \times g$ for 10 min at 4 °C.

Purification of proteins

All proteins were produced with an N-terminal His₆-tag. Cells were lysed by sonication. After centrifugation, the soluble fraction was purified by Ni-NTA (1 ml Ni resin cartridge QIAGEN).

Protein analytics

The concentration of the expressed protein in the eluate was determined by HPLC-UV₂₁₅ (Agilent Technologies 1200 series). The molecular weight of the proteins was identified by LC-ESI-MS/LC-TOF (Waters) using a Poros R1H 1*15 mm column equilibrated at 75 °C.

NMR spectroscopy

All spectra were recorded on a 600 MHz Bruker Avance spectrometer equipped with a TCI cryoprobe with magnetic field gradients. All measurements were carried out at 23 °C.

Results

We show here that incorporation of 97–99 % of ^{13}C and ^2H can be achieved in auto-inducing media, in a robust and highly reproducible manner. These results are based on the use of lactose where only the glucose moiety is isotope labeled. On the track to these results, we proved (1) that the galactose moiety of lactose is not consumed by BL21 strains and can therefore remain unlabeled. Additionally (2), we investigated different ratios of the carbohydrates in the auto-induction mixture in order to find the one leading to highest levels of protein production. Finally (3), suitably labeled lactose was produced in an enzymatic reaction and (4) used to express u- ^{13}C and u- ^2H labeled proteins.

Galactose is not metabolized by *E. coli* BL21 strains, therefore the lactose only needs to be labeled on the glucose moiety

The central problem in establishing an auto-inducing medium for isotope labeling is the availability of suitably labeled lactose. We propose here to use lactose, which is only labeled on the glucose moiety, because the galactose moiety should not be metabolized by *E. coli* B707 derived strains like BL21, as indicated by the “gal” phenotype (Daegelen et al. 2009). In order to prove this hypothesis, a commercial BL21 (DE3) strain was grown on three different carbon sources: Galactose, lactose and glucose (M9 medium with 1 g/l of the respective carbohydrate). As expected, BL21 bacterial cells did not grow in the galactose based medium, suggesting that galactose was not even partially consumed (Fig. 2). Growth on lactose and glucose led to cell densities of 0.35 and 0.7, further showing that only the glucose moiety of lactose was consumed.

In order to assess whether individual atoms of galactose would not be incorporated into proteins, the protein FKBP12 was expressed in cultures with similar carbohydrate composition as ^{13}C -enriched auto-inducing media and increasing amounts of ^{12}C -galactose (5 g of ^{13}C -glycerol and 1.5 g of glucose and 0, 1 and 3 g of ^{12}C -galactose per liter). The medium with 1 g/l of galactose would therefore mimic the standard auto-inducing medium by replacing 2 g of lactose with 1 g glucose and 1 g of galactose. Cells grew to same final densities, with notable slower growth rate for the culture with 3 g/l of galactose. After purification, the proteins were subjected to MS analysis. ^{13}C -incorporation levels obtained in the different cultures were 98.2, 98.1 and 97.9 %, respectively and were therefore hardly affected by the presence of ^{12}C -galactose (Fig. 2). Also the secondary modification of the protein (~40 % gluconoylation, corresponding to +178 Da for the non- ^{13}C -labeled sample) took place with a labeled sugar moiety, leading to an added mass of 184.1 Da (Geoghegan et al. 1999). We attribute the slight decline of ^{13}C -incorporation to carbohydrate impurities in the galactose employed in these experiments, rather than metabolization of galactose itself. In any case, the resulting incorporation is well within satisfactory limits.

In summary, in an auto-inducing medium only the glucose moiety of the lactose needs to be isotope labeled, as the galactose is not being metabolized.

The auto-inducing carbohydrate mixture leading to highest protein yields consists of 5 g of glycerol, 0.5–1 g of glucose and 2 g of lactose per liter

The relative amounts of the individual carbohydrates in the auto-inducing mixture has been thoroughly established by

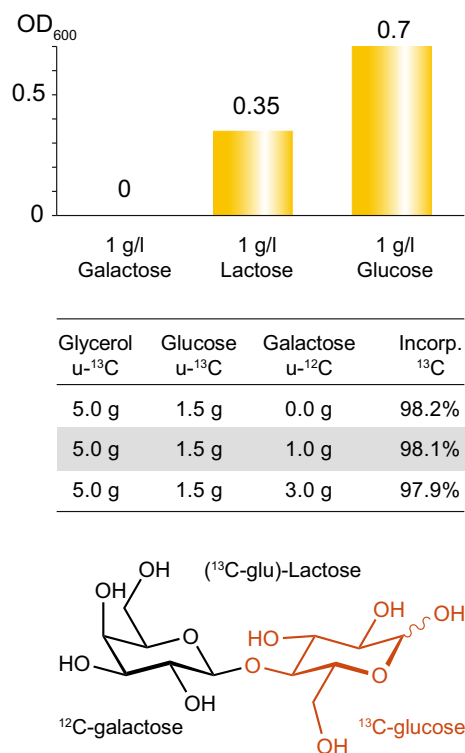


Fig. 2 The galactose moiety of lactose is not consumed by *E. coli* BL21 and not incorporated into proteins. In the *upper panel* final cell densities in OD₆₀₀ units of bacterial cultures grown on different single carbon sources are shown, demonstrating that galactose is not able to support growth. In the *table in the middle*, ¹³C-incorporation into the protein FKBP12 is shown depending on the composition of the growth medium. The content of the carbohydrates glycerol, glucose and galactose in 1 l of medium are given in the *first three columns*, respectively. The *row highlighted with a grey band* should mimic the composition of standard auto-inducing media, i.e. 5.0 g of glycerol, 0.5 g of glucose and 2.0 g of lactose per liter, where lactose is replaced by 1.0 g of galactose and 1.0 g of glucose. ¹³C incorporation into FKBP12 as assessed by MS is given in the *last column* (Incorp. ¹³C). As ¹³C-incorporation is hardly affected by ¹²C-galactose, using (¹³C-glu)-lactose (*lower panel*) instead of u-¹³C-lactose in auto-inducing media should yield identical ¹³C-incorporation into expressed proteins

Studier (2005) in order to obtain the highest protein yields. For the purpose of isotope labeling, the composition of this mixture needs to be optimized in order to balance yields with cost of isotope labeled carbohydrates. Here, the most expensive component is lactose. It can be reduced to 1.0 g/l but expression yields drop practically linearly with decreasing lactose concentration (Table 1). At the lower end, reducing it further to 0.125 g/l leads to low expression yields and results that are difficult to reproduce (Tyler et al. 2005), with values between 0 and 5 % of what was obtained with the standard recipe with 2 g/l of lactose in standard culturing conditions. Since oxygen limitation can lead to induction of protein expression with the *lac*-operator system, experiments were also performed under

oxygen limiting conditions, i.e. 100 ml of culture in a 200 ml shake flask (Studier 2005; Tyler et al. 2005; Li et al. 2011). Protein expression under these conditions ranged from 10 up to 25 % of the original expression levels. Surprisingly, similar expression levels were obtained with lactose-free auto-inducing mixture. Indeed, in media with 0.125 g/l of lactose, lactose is actually depleted at cell densities above one OD₆₀₀ unit, as seen by NMR analysis of the media, and it can therefore not act anymore as an inducer. Therefore, we preferred to use 2 g/l of lactose to ensure robust high level protein production until the end of the culturing period.

Glycerol is added at 5 g/l to an auto-inducing medium, according to the published protocol. Since only 50–60 % of it is consumed during a culture, it was appealing to reduce its amount in order to make the medium less expensive. However, yields dropped sharply upon reduction of glycerol amounts (Table 1), prompting us to maintain the glycerol concentration at the originally published values (Studier 2005).

Glucose is a minor cost factor as it is employed at only 0.5 g/l. Here, we tested different concentrations, since higher glucose levels could delay protein induction and lead to higher protein amount due to higher cell density of the culture. As in the standard protocol, optimal results were obtained for glucose concentration in the range of 0.5–1.0 g/l.

Synthesis and purification of lactose with an isotope labeled glucose moiety

Lactose was produced in a one-pot enzymatic reaction from ¹³C- or ²H-labeled glucose and non-labeled UDP-galactose (Hudson et al. 1972) (Fig. 3; for details see the materials and methods section). The central enzyme of this reaction is lactose synthase, which is a heterodimer of β(1→4) galactosyltransferase (GalT) and α-lactalbumin. β(1→4) GalT is a relatively unspecific enzyme that transfers activated galactose moieties onto a wide variety of substrates. α-lactalbumin confers specificity towards glucose as a substrate and strongly enhances lactose production. As a secondary enzyme, alkaline phosphatase was used, which pulled the equilibrium reaction of the lactose synthase complex towards the product lactose, by the exergonic hydrolysis of ADP (Fig. 3). After several hours of reaction, molar ratios of unreacted glucose and the product lactose ranging from 1:1 to 1:1.3 were typically obtained. After removal of charged reaction components over an ion exchange resin at pH 10, pure lactose and glucose were isolated. As shown in Table 1, glucose to lactose ratios of 1:1–1:2 yield high level protein expression. Therefore, the glucose and lactose were not further separated and directly employed in protein expression media. If required, the

Table 1 Influence of media composition on expression yields of the protein kRas

	Glucose (g/l)	Lactose (g/l)	Glycerol (g/l)	Cell density (OD ₆₀₀) ^b	Protein yield (mg/l) ^c
Reference ^a	0.5	2.0	5.0	7.2 (0.1)	16.1 (2.7)
Glycerol			3.5	6.2	6.6
			2.25	4.4	2.1
			1.5	3.9	2.7
Lactose		1.0		6.5	9.2
		0.125		4.0 (0.3)	0.4 (0.4)
		0.0		3.5	0.2
O ₂ limitation ^d		0.125		4.0	3.1
		0.0		4.2	3.3
Glucose	1.0			7.1	13.6
IPTG ^e	8.0	0.0	0.0	5.8	18.3

^a Reference auto-inducing medium on M9 background, using 2 g/l of NH₄Cl. For the other media, only the values that were changed compared to this reference medium are shown

^b Final cell density after overnight growth at 18 °C (standard deviations for multiple experiments)

^c Concentration of Ni-affinity purified protein as determined by HPLC-UV₂₁₅ (standard deviation)

^d Culturing under oxygen-limiting conditions using a 200 ml Erlenmeyer flask filled with 100 ml of culture, in order to achieve protein expression with low amounts of lactose. All other experiments were carried out using 100 ml of culture in 500 ml Erlenmeyer flasks

^e M9 medium containing 8 g of glucose for a similar total carbohydrate content induced at OD₆₀₀ of 0.7 with 0.4 mM IPTG (For the protein kRas yields are higher when using 0.4 mM IPTG instead of 1 mM)

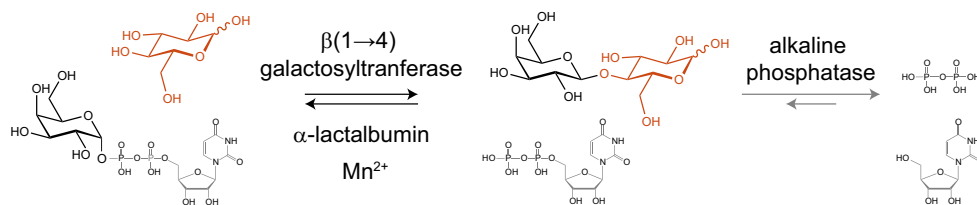


Fig. 3 Enzymatic synthesis of ²H/¹³C-glucose-labeled lactose. In a single reaction, galactosyltransferase in conjunction with α -lactalbumin catalyze the reaction of non-labeled UDP-galactose (black) with ²H/¹³C-labeled glucose (orange) to form (²H/¹³C-glu)-lactose (black and orange for non-labeled and labeled moieties, respectively).

Alkaline phosphatase catalyzes an exergonic reaction, pulling the reaction equilibrium to the right side. The side products (grey) are charged when the pH of the solution is above 10, and can therefore be removed on ion exchange resins, leading to a pure mixture of (²H/¹³C-glu)-lactose together with unreacted ²H/¹³C-glucose

desired glucose to lactose ratio was adjusted by addition of labeled glucose.

Expression of u-¹³C and u-²H labeled proteins with auto-inducing medium containing glucose-labeled lactose

As a proof of concept, two test proteins were selected for uniform ¹³C and ²H isotope labeling, respectively, using the newly designed auto-inducing media. The aim was to produce ¹³C,¹⁵N-FKBP12 (Michnick et al. 1991) and ²H,¹⁵N-MBP (Kapust and Waugh 1999) in order to assess isotope incorporation levels with this protocol. ¹⁵N- and ²H,¹⁵N-enriched carbon-free base media were prepared, and suitably labeled auto-inducing mixtures were added:

i.e. 0.8 g/l of ¹³C-glucose, 5.0 g/l of ¹³C-glycerol and 2.0 g/l of (¹³C-glu)-lactose for ¹³C, ¹⁵N-labeling and 0.8 g/l of ²H-glucose, 5.0 g/l of ²H-glycerol and 2.0 g/l of (²H-glu)-lactose for ²H,¹⁵N-labeling. The purified proteins were subjected to mass spectrometry (MS) analyses, where incorporation of 96.6 and 98.8 % was shown for ¹³C,¹⁵N-FKBP12 and ²H,¹⁵N-MBP, respectively. For ¹³C incorporation, a value of 98 % was expected from the experiments with glucose and galactose as pure carbon sources. We attributed the slightly lower value to trace contaminants of unlabeled products from the lactose synthesis. Overall, these isotope incorporation levels are comparable to the ones obtained with conventional labeling techniques, therefore enabling NMR studies that require high ²H and ¹³C incorporation levels (Fig. 4; Table 2).

Discussion

A robust and inexpensive protocol for isotope labeling with auto-inducing media

Here, an auto-inducing protocol for uniform ^2H , ^{13}C and ^{15}N labeling at high incorporation (>95 %) and reproducible yields is presented. It is based on selectively labeled lactose which can be produced from inexpensive substrates: unlabeled activated galactose and labeled glucose.

Reduction of cost of auto-inducing media for isotope labeling

Whereas auto-inducing media have found wide application in high-throughput expression setups in e.g. X-ray based structural biology laboratories, this approach has not found

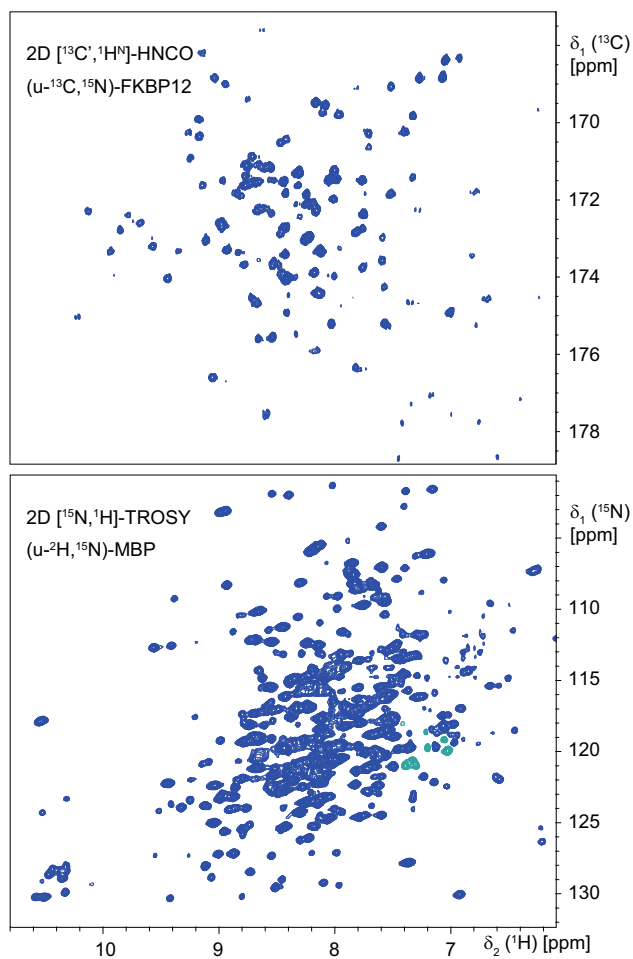
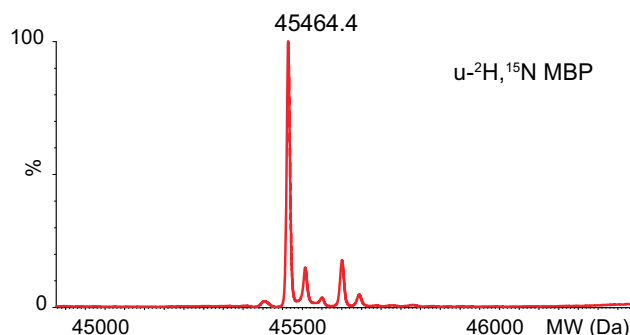
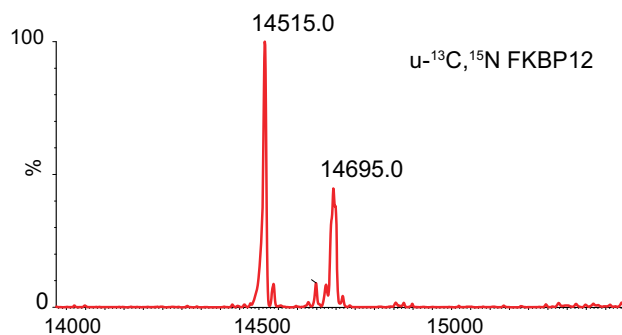


Fig. 4 NMR spectra of $u\text{-}^{13}\text{C}, ^{15}\text{N}$ -FKBP12 and $u\text{-}^2\text{H}, ^{15}\text{N}$ -MBP. The first 2D [^{13}C , ^1H]-plane of an HNCO spectrum of FKBP12 is shown in the *top panel*. In the *lower panel*, a 2D [^{15}N , ^1H]-TROSY of MBP is shown

Table 2 Isotope incorporation as determined from mass spectrometry

Protein name Labelling pattern	FKBP12 $u\text{-}^{13}\text{C}, ^{15}\text{N}$	MBP $u\text{-}^2\text{H}, ^{15}\text{N}$
MW unlabeled ^a (Da)	13757.4	42776.6
MW labeled ^b (Da)	14541.8	45497.1
MW found ^c (Da)	14515.0	45464.4
Incorporation ^d (%)	96.6	98.8



^a Calculated molecular weight (MW) from amino acid sequence with natural abundance isotope distribution

^b Calculated MW assuming 100 % isotope incorporation and complete back-exchange of labile protons in the case of $u\text{-}^2\text{H}, ^{15}\text{N}$ MBP

^c Measured molecular weight of produced proteins, derived from the upper and lower mass spectra shown below the table, for $u\text{-}^{13}\text{C}, ^{15}\text{N}$ FKBP12 and $u\text{-}^2\text{H}, ^{15}\text{N}$ MBP, respectively. The second mass obtained for FKBP12 arises from the gluconoylated form

^d Incorporation level of isotopes

widespread application in NMR due to extremely high prices for labeled lactose. Simply reducing the amount of lactose in media for labeling can lead to lower yields and lower reproducibility of protein expression. Yields are lower because lactose is not present in high enough concentrations in later stages of the culture in order to sustain high-level protein expression. At the same time, reproducibility is lower because, at low lactose concentrations, protein expression is linked to oxygenation levels, which are difficult to control. We propose here to keep the lactose concentration at 2 g/l and to reduce the cost of lactose by

producing it in only partially labeled form, as only the glucose moiety needs to be labeled. The activated galactose moiety, which is the major price driver for synthesis of uniformly labeled lactose, does not need to be labeled. In this work it is shown that lactose for isotope labeling in auto-inducing media can be produced from inexpensive substances in a one-pot reaction. The reaction can probably be much simplified and scaled up economically in an industrial setup specialized in such chemistry. In this work, reaction and purification approaches were restricted to tools available in a small biochemistry research laboratory.

A secondary cost factor for auto-inducing media is the need for 5 g/l of glycerol. For ^2H labeling, glycerol is more cost-effective than glucose, but for ^{13}C and simultaneous ^2H , ^{13}C labeling it is more expensive. Also here, simply reducing the amount of glycerol is not advisable, as the yields drop more steeply than the cost by reducing the glycerol concentration in the media (Table 1). This is probably due to the fact that cells consume lactose at a higher rate if glycerol concentrations are reduced. Thereby, lactose may fall below the concentration needed for robust induction of expression during the culture, resulting in lower amounts of expressed protein.

Limitation of this protocol to *E. coli* B strains

The present protocol is only applicable when using *E. coli* B strains (Daegelen et al. 2009). This does not represent a serious limitation: for bacterially expressed proteins, this sub-strain was reported for 97 % of the proteins deposited in the RCSB's protein data bank (PDB; <http://www.rcsb.org/pdb/>). Within the BL21 strains, cells with a mutation in the lactose transporter (LacY^- , lactose permease), e.g. TunerTM cells (Novagen), are generally not amenable to auto-induction, because lactose does not freely diffuse through bacterial cell membranes, as opposed to IPTG. However, the main reason for using such cell strains is to improve the ratio of soluble over insoluble protein by reducing the strength of induction, which is exactly what is achieved by switching from IPTG to lactose as inducer of protein expression.

Summary and outlook

In summary, we show here a way of producing auto-inducing media for stable isotope labeling in a cost-effective way, by replacing uniformly labeled lactose by glucose-labeled lactose, which can be produced from inexpensive substrates. Using this protocol, uniform isotope labeling with high isotope incorporation levels can be achieved, fully exploiting the advantages of auto-inducing media. This starts in expression scouting in small volumes, where measuring cell densities and inducing several different cultures at their individual optimal cell density is

impractical and work intense. With auto-inducing media, cultures can be all started in parallel and do not need to be monitored until harvested, with the additional benefit that all cultures will be induced at a pre-defined cell density (i.e. $\text{OD}_{600} = 0.7$). Once a suitable construct is identified in small scale, the culturing conditions can be easily scaled up, as the medium composition does not vary and results are independent of oxygenation levels. In large scale, bacterial cell cultures in minimal media with different isotope enrichment display strongly differing growth rates, frequently with an additional lag phase of varying duration after inoculation. Monitoring such cultures is work intense and error prone, since the ideal time point of induction may eventually be missed, especially when several cultures need to be run in parallel, and often requires presence during night hours. With auto-inducing media, these issues are minimized, and the time point of induction is highly reproducible among multiple cultures. Generally, auto-inducing cultures yield higher amounts of soluble protein, because higher cell densities are typically obtained. Additionally, time can be freed up for work on tasks bearing higher scientific impact. Ideally, this work would enable vendors of stable isotope products to offer glucose-labeled lactose or a formulated auto-inducing medium at competitive costs for uniform labeling of ^{13}C and ^2H at high isotope incorporation levels, rendering all advantages of auto-inducing media reachable to the NMR community.

Acknowledgments We thank Patrick Graff for excellent support with mass spectrometry and Jean-Michel Rondeau for helpful discussions.

Conflict of interest The authors declare that they have no conflict of interest.

Ethical standard This article does not contain any studies with human participants or animals performed by any of the authors.

References

- Abramson J, Smirnova I, Kasho V et al (2003) Structure and mechanism of the lactose permease of *Escherichia coli*. *Science* 301:610–615. doi:10.1126/science.1088196
- Assenberg R, Wan PT, Geisse S, Mayr LM (2013) Advances in recombinant protein expression for use in pharmaceutical research. *Curr Opin Struct Biol* 23:393–402. doi:10.1016/j.sbi.2013.03.008
- Cavanagh J, Fairbrother WJ, Palmer AG III, Skelton NJ (1996) Protein NMR spectroscopy, principles and practice. Academic Press, Waltham
- Daegelen P, Studier FW, Lenski RE et al (2009) Tracing ancestors and relatives of *Escherichia coli* B, and the derivation of B strains REL606 and BL21 (DE3). *J Mol Biol* 394:634–643. doi:10.1016/j.jmb.2009.09.022
- Dickson RC, Abelson J, Barnes WM, Reznikoff WS (1975) Genetic regulation: the Lac control region. *Science* 187:27–35. doi:10.1126/science.1088926

- Gardner KH, Kay LE (1998) The use of ^2H , ^{13}C , ^{15}N multidimensional NMR to study the structure and dynamics of proteins. *Annu Rev Biophys Biomol Struct* 27:357–406. doi:[10.1146/annurev.biophys.27.1.357](https://doi.org/10.1146/annurev.biophys.27.1.357)
- Geoghegan KF, Dixon HBF, Rosner PJ et al (1999) Spontaneous alpha-N-6-phosphogluconoylation of a “His Tag” in *Escherichia coli*: the cause of extra mass of 258 or 178 Da in fusion proteins. *Anal Biochem* 267:169–184. doi:[10.1006/abio.1998.2990](https://doi.org/10.1006/abio.1998.2990)
- Hudson BG, McKenzie L, Ebner KE (1972) Enzymic synthesis of lactose-14C (galactosyl-14C) and lactose-3H (glucosyl-3H)1. *Anal Biochem* 48:524–528. doi:[10.1016/0003-2697\(72\)90107-8](https://doi.org/10.1016/0003-2697(72)90107-8)
- Jacob F, Monod J (1961) Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol* 3:318–356. doi:[10.1016/S0022-2836\(61\)80072-7](https://doi.org/10.1016/S0022-2836(61)80072-7)
- Kaback HR, Sahin-Toth M, Weinglass AB (2001) The kamikaze approach to membrane transport. *Nat Rev Mol Cell Biol* 2:610–620. doi:[10.1038/35085077](https://doi.org/10.1038/35085077)
- Kapust RB, Waugh DS (1999) *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci* 8:1668–1674. doi:[10.1110/ps.8.8.1668](https://doi.org/10.1110/ps.8.8.1668)
- LeMaster DM (1989) Deuteration in protein proton magnetic resonance. *Methods Enzymol* 177:23–43. doi:[10.1016/0076-6879\(89\)77004-X](https://doi.org/10.1016/0076-6879(89)77004-X)
- Li Z, Kessler W, van den Heuvel J, Rinas U (2011) Simple defined autoinduction medium for high-level recombinant protein production using T7-based *Escherichia coli* expression systems. *Appl Microbiol Biotechnol* 91:1203–1213. doi:[10.1007/s00253-011-3407-z](https://doi.org/10.1007/s00253-011-3407-z)
- McIntosh LP, Dahlquist FW (1989) Biosynthetic incorporation of ^{15}N and ^{13}C . *Q Rev Biophys* 23:1–38. doi:[10.1017/S0033583500005400](https://doi.org/10.1017/S0033583500005400)
- Michnick SW, Rosen MK, Wandless TJ et al (1991) Solution structure of FKBP, a rotamase enzyme and receptor for FK506 and rapamycin. *Science* 252:836–839. doi:[10.1126/science.1709301](https://doi.org/10.1126/science.1709301)
- Muchmore DC, McIntosh LP, Russel CB et al (1989) Expression and nitrogen-15 labeling of proteins for proton and nitrogen-15 nuclear magnetic resonance. *Methods Enzymol* 177:44–73. doi:[10.1016/0076-6879\(89\)77005-1](https://doi.org/10.1016/0076-6879(89)77005-1)
- Peti W, Page R (2007) Strategies to maximize heterologous protein expression in *Escherichia coli* with minimal cost. *Protein Expr Purif* 51:1–10. doi:[10.1016/j.pep.2006.06.024](https://doi.org/10.1016/j.pep.2006.06.024)
- Sreenath HK, Bingman CA, Buchan BW et al (2005) Protocols for production of selenomethionine-labeled proteins in 2-L polyethylene terephthalate bottles using auto-induction medium. *Protein Expr Purif* 40:256–267. doi:[10.1016/j.pep.2004.12.022](https://doi.org/10.1016/j.pep.2004.12.022)
- Studier FW (2005) Protein production by auto-induction in high-density shaking cultures. *Protein Expr Purif* 41:207–234. doi:[10.1016/j.pep.2005.01.016](https://doi.org/10.1016/j.pep.2005.01.016)
- Studier FW, Moffatt BA (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* 189:113–130. doi:[10.1016/0022-2836\(86\)90385-2](https://doi.org/10.1016/0022-2836(86)90385-2)
- Tyler RC, Sreenath HK, Singh S et al (2005) Auto-induction medium for the production of [^{15}N]- and [^{13}C , ^{15}N]-labeled proteins for NMR screening and structure determination. *Protein Expr Purif* 40:268–278. doi:[10.1016/j.pep.2004.12.024](https://doi.org/10.1016/j.pep.2004.12.024)