

An improved protocol for amino acid type-selective isotope labeling in insect cells

Yaqun Zhang¹ · Hui Wei² · Dianlin Xie¹ · Deepa Calambur¹ · Andrew Douglas¹ · Mian Gao¹ · Frank Marsilio¹ · William J. Metzler³ · Nicolas Szapiel¹ · Ping Zhang¹ · Mark R. Witmer¹ · Luciano Mueller⁴  · David Hedin⁵

Received: 5 March 2017 / Accepted: 15 May 2017 / Published online: 15 July 2017
© Springer Science+Business Media Dordrecht 2017

Abstract An improved expression protocol is proposed for amino acid type-specific [¹³C], [¹⁵N]-isotope labeling of proteins in baculovirus-infected (BV) insect cell cultures. This new protocol modifies the methods published by Gossert et al. (J Biomol NMR 51(4):449–456, 2011) and provides efficient incorporation of isotopically labeled amino acids, with similar yields per L versus unlabeled expression in rich media. Gossert et al. identified the presence of unlabeled amino acids in the yeastolate of the growth medium as a major limitation in isotope labeling using BV-infected insect cells. By reducing the amount of yeastolate in the growth medium ten-fold, a significant improvement in labeling efficiency was demonstrated, while maintaining good protein expression yield. We report an alternate approach to improve isotope labeling efficiency using BV-infected

insect cells namely by replacing the yeast extracts in the medium with dialyzed yeast extracts to reduce the amount of low molecular weight peptides and amino acids. We report the residual levels of amino acids in various media formulations and the amino acid consumption during fermentation, as determined by NMR. While direct replacement of yeastolate with dialyzed yeastolate delivered moderately lower isotope labeling efficiencies compared to the use of ten-fold diluted undialyzed yeastolate, we show that the use of dialyzed yeastolate combined with a ten-fold dilution delivered enhanced isotope labeling efficiency and at least a comparable level of protein expression yield, all at a scale which economizes use of these costly reagents.

Keywords Baculovirus · Insect cell culture · Labeling · NMR · HC-HSQC · HNCQ-TROSY · Mass spectrometry · Yeastolate

Electronic supplementary material The online version of this article (doi:10.1007/s10858-017-0117-6) contains supplementary material, which is available to authorized users.

✉ Luciano Mueller
luciano.mueller@bms.com

¹ Department of Protein Science, Bristol-Myers Squibb, Route 206 and Province Line Road, P.O. Box 4000, Princeton, NJ 08543-4000, USA

² Mass Spectrometry COEI, Bristol-Myers Squibb, Route 206 and Province Line Road, P.O. Box 4000, Princeton, NJ 08543-4000, USA

³ Lead Discovery and Optimization, Bristol-Myers Squibb, Route 206 and Province Line Road, P.O. Box 4000, Princeton, NJ 08543-4000, USA

⁴ PCO DAS NMR, Bristol-Myers Squibb, Route 206 and Province Line Road, P.O. Box 4000, Princeton, NJ 08543-4000, USA

⁵ Expression Systems, LLC, 2537 Second Street, Davis, CA 95618, USA

Introduction

Expression of proteins for study by structural biology is widely achieved using *Escherichia coli* due to several advantages relative to other expression hosts. Factors include high expression yields, rapid doubling time, ease in handling and manipulating strains, and the ability to use growth media of low complexity. This last advantage is particularly important for proteins targeted for study by NMR methods, since it facilitates incorporation of stable isotopic labels such as [¹⁵N], [¹³C] and [²H] into the expressed proteins. Many proteins produced for structural and mechanistic studies require post-translational modifications to achieve biologically relevant states, which is particularly true for mammalian proteins. Unfortunately, many putative drug discovery targets are difficult to express in functional

or soluble forms in *E. coli*, which lacks the cellular machinery to affect requisite folding and post-translational modifications. Therefore, proteins that are not expressible in *E. coli* or require post-translational modifications, such as phosphorylation or glycosylation, are typically expressed in eukaryotic host systems such as insect cells (e.g. Sf9, Sf21, *T. ni* and Hi5), mammalian cells (e.g. CHO, HEK, COS) or yeast (e.g. *Pichia*, *S. cerevisiae*). Significant advances in protein expression for pharmaceutical research have been made in recent years (Assenberg et al. 2013).

Among the eukaryotic expression systems, BV-infected insect cells is perhaps the most straightforward, least labor-intensive and most commonly employed system. In BV-infected insect cells, heterologous over-expression of proteins is achieved by replacing the naturally highly expressed polyhedrin gene with the gene of interest. This mode of protein expression imparts additional flexibility to expressing isotopically-labeled proteins in insect cell by enabling cell growth and infection in unlabeled optimal growth media prior to transferring infected cells to a type-specifically labeled medium (Kelly et al. 2007). Overall, the relative ease of use has rendered the baculovirus expression system a preferred method for protein expression for X-ray crystallography. The use of the baculovirus expression system for the expression of isotopically enriched proteins for NMR applications has proved more challenging due to the considerable complexity of the required growth media in insect cell cultures (Weisset et al. 1981; Maiorella et al. 1988; Schlaeger 1996; Gossert and Jahnke 2012). The development of serum-free growth media enabled the incorporation of common amino acid substitutes (such as selenomethionine, or Se-Met) or isotopically enriched amino acids for naturally occurring amino acids. Initial uses of baculovirus expression for incorporating isotopically labeled amino-acids were carried out by DeLange et al. who incorporated [²H]-labeled tyrosine into rhodopsin (DeLange et al. 1998), and Creemer et al. who reported the incorporation of [¹⁵N]-labeled lysine into rhodopsin (Creemers et al. 1999). The use of BV-infected insect cell cultures for the production of Se-Met labeled protein for crystallographic phasing was first reported by (Fremont et al. 1998; Bellizzi et al. 1999).

The baculovirus expression system is particularly well suited for type-specific labeling proteins with one or more of its ten essential amino acids (i.e. Arg, His, Ile, Leu, Lys, Met, Phe, Thr, Trp, and Val). Selective isotope labeling of these residue types can be achieved without the risk of isotope scrambling due to the inability of eukaryotic cells to produce these amino acids. [¹⁵N]-label dilution and scrambling occurs only for amino acids that are related biosynthetically to the citric cycle (i.e. Ala, Asn, Asp, Cys, Gln, Glu, Gly, Pro, and Ser). [¹³C]-dilution originates from unlabeled carbohydrates to a varying degree (Strauss et al.

2005). In addition to amino acids, insect cell growth media contain multiple carbon sources, including lipids, carbohydrates and yeastolate. Yeastolate, a nutritional supplement generally comprised of a mixture of peptides, amino acids, carbohydrates, and vitamins, is included in standard insect cell media at levels up to 6 g per liter, and therefore constitutes a major source of dilution of isotope labels. Due to the restricted metabolism in insect cells, for [¹⁵N]-labeling, only amino acids and yeastolate need to be considered as sources of isotopic dilution; for [¹³C]-labeling, carbohydrate may be an additional source for the dilution of [¹³C]-enrichment.

Two approaches have been used for preparing growth media for BV-infected protein expression in insect cell culture: preparing media from scratch (Bruggert et al. 2003), and augmenting a commercial medium that has been depleted in a targeted set of amino acids with isotopically enriched amino acid(s) (Creemers et al. 1999; Strauss et al. 2003; Walton et al. 2006; Gossert et al. 2011). The former approach enables high levels of uniform isotope enrichment, but suffers the drawback of being expensive if uniform enrichment is the desired outcome. In addition, preparing media from scratch may adversely impact reproducibility of protein expression (Gossert and Jahnke 2012). In contrast, augmenting a commercially available medium that has been depleted of targeted amino acids affords good reproducibility and high target protein expression yields; hence, it appears most suited for the preparation of type-specifically labeled protein samples. Most recently, Gossert and coworkers (Sitarska et al. 2015) developed protocols for uniform stable isotope labeling of proteins using [¹⁵N]- and [¹³C,¹⁵N]-labeled algal extracts supplemented with the addition of [¹⁵N]-ammonium salt as a nitrogen source and [¹³C]-glucose as a carbon source. This protein expression method represents the most cost effective approach to uniform labeling yet published.

In this study we refine the method of producing type-specifically labeled protein samples utilizing modified versions of the insect cell growth media ESF 921 (Expression Systems, LLC) in which the media were depleted in targeted amino acid types. Similar to Gossert et al. (2011), we focused on improving the growth medium by altering the yeastolate content. We first quantify the level of residual unlabeled amino acids in various media and their consumption during cell growth and protein expression. We then explore the use of yeastolate that had been dialyzed to remove ingredients having molecular weights less than 1 kDa. Our rationale was to eliminate many low molecular weight compounds in yeast extracts that may constitute unlabeled sources for amino acid synthesis during protein biosynthesis. We then show that the use of dialyzed yeastolate combined with a ten-fold dilution delivered enhanced isotope labeling efficiency as well as at least maintaining

similar protein expression yields relative to expression in rich, unlabeled medium.

Results and discussion

In recognition that yeastolate may represent a major source of unlabeled biomaterial limiting the percentage of incorporation of [^{13}C]- and [^{15}N]- isotopes in baculovirus-expressed proteins, we pursued the strategy of using modified yeast extracts that were dialyzed against a filter with a cutoff pore size of 1 kDa. The rationale was to remove amino acids and low molecular weight peptidic fragments from the yeastolate which could potentially be hydrolyzed and converted to amino acids by the insect cells. This step should remove a source of unlabeled amino acids during the biosynthesis of a heterologously expressed protein. Our chosen protein expression protocol resembles the method by Cristiani, Pharmacia (Cristiani 2003). Sf9 or Hi5 cells were grown in commercial ESF 921 growth medium. Once the cell density reached the targeted level, cells were infected with baculovirus. Following an infection period of 16 h the cells were pelleted, washed and transferred to the amino acid-depleted medium. After a 4 h “starvation” period, labeled amino acid stocks were added (see Experimental section for further details). This sequence of events differs from the protocol of Gossert et al. who initiated viral induction after the transfer of cells to the labeled growth medium. In a more recent paper, Gossert et al. (2011) cited improved infection yields in the rich medium by employing the protocol of inducing viral infection 16 h prior to the transfer of cells to the labeled growth medium.

Quantitation of amino acid levels in yeastolate by analysis of growth media

To quantify the residual unlabeled (natural abundance [^{13}C]) amino acid content in growth media, HC-HSQC

spectra were recorded on four media formulations which were derived from the following stock solutions: (a) ESF 921; the “rich” undepleted commercial growth medium containing standard level of 8 g/l (100%) yeastolate, (b) ESF 921- Δ 11, 10%D, ESF 921 medium depleted in 11 amino acids (i.e. Arg, His, Ile, Lys, Leu, Met, Phe, Thr, Tyr, Trp, Val) and containing 10% (0.8 g/l) dialyzed yeastolate, (c) dialyzed yeastolate stock, (d) undialyzed yeastolate stock. From these stock solutions, the following four media test solutions were examined by NMR:

1. ESF 921; the “rich” undepleted ESF 921 medium containing 100% undialyzed yeastolate,
2. ESF 921- Δ 11 10%D stock (containing 0.8 g/l dialyzed yeastolate),
3. ESF 921- Δ 11 10%D stock+100% undialyzed yeastolate stock (containing 0.8 g/l dialyzed yeastolate + 8 g/l undialyzed yeastolate),
4. ESF 921- Δ 11 10%D stock+90% dialyzed yeastolate stock (containing 8 g/l yeastolate).

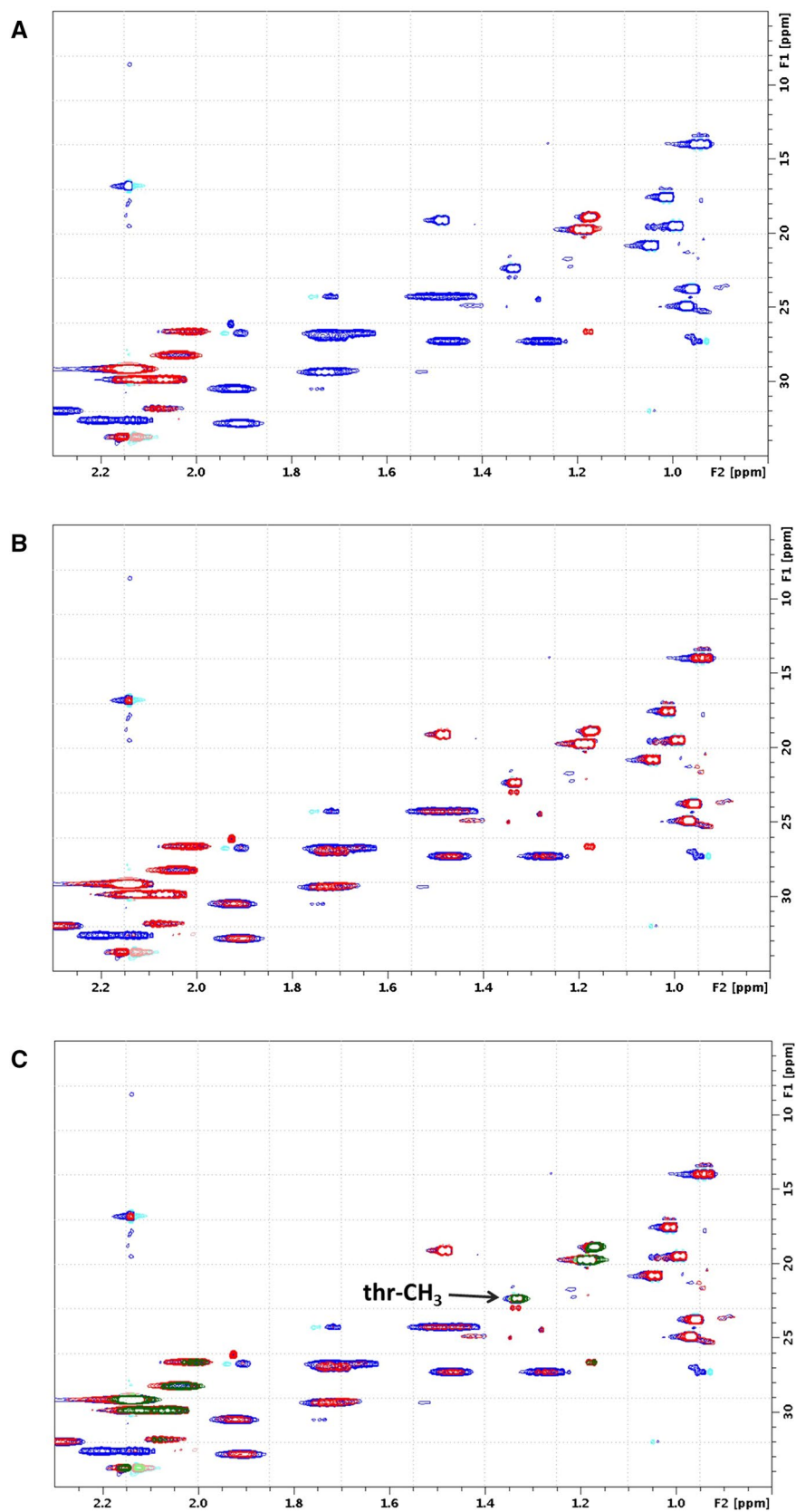
Using the procedure described in the Experimental Details section employing the extraction of cross-peak volumes in HC-HSQC spectra, the percentages of amino acids in the various media formulations were calculated relative to the amino acids content in ESF 921 and listed in columns 3–5 of Table 1. Figure 1 qualitatively illustrates that the undialyzed yeastolate in the ESF 921 growth medium contains a considerable amount of residual amino acids. Figure 1a depicts the super-position of the expanded region of the HC-HSQC spectrum of the rich medium indicated by the blue contours and the ESF 921- Δ 11 medium indicated by the red contours illustrating the disappearance of the methyl peaks of the depleted amino acids. Figure 1c shows how cross-peaks were assigned by spiking ESF 921- Δ 11 media samples with individual amino acids. With the exception of methionine, all tested amino acids are present in sizeable amounts in ESF 921. This helps to explain why

Table 1 Quantitation of unlabeled amino acids in ESF 921- Δ 11 with various formulations of dialyzed or undialyzed yeastolate

aa	ESF 921-11 100%U (%)	ESF 921- Δ 11 10%U (%)	ESF 921- Δ 11 100%D (%)	ESF 921- Δ 11 10%D (%)	Background (%)
Thr	47.9	5.3	6.4	1.15	0.56
Leu	46.2	4.9	5.6	0.81	0.29
Val	32.4	3.2	3.6	0.71	0.39
Lys	32.0	5.4	5.5	2.78	2.48
Arg	26.2	3.1	3.5	0.82	0.53
Ile	23.5	2.4	2.7	0.22	0.00
Phe	23.2	2.7	3.6	0.77	0.45
Met	6.2	0.9	0.9	0.39	0.34

Residual amino acid content was determined from cross-peak volumes in HC-HSQC spectra acquired in media test sample—detailed description of analysis is provided in the Supplemental Materials section

Fig. 1 Expanded region of HC-HSQC spectra in ESF 921. **a** *blue contours*: ESF 921, *red contours*: ESF 921- Δ 11 containing 10% dialyzed yeastolate, **b** *blue contours*: ESF 921, *red contours*: ESF 921- Δ 11 containing 10% dialyzed yeastolate + 100% undialyzed yeastolate, **c** Same as B + *green contours*: spectrum of ESF 921- Δ 11 + 0.25 gr L-threonine



initial attempts to produce labeled proteins in baculovirus expression system yielded relatively poor incorporations of isotope labels (Creemers et al. 1999; DeLange et al. 1998; Strauss et al. 2003). Column 3 in Table 1 shows that the approach by Gossert et al. of diluting the content undialyzed yeastolate ten-fold does indeed significantly reduce the percentage of amino acids originating from the yeastolate. In each case, the level of residual amino acids is reduced close to 10-fold, indicating the yeastolate is the primary source of unlabeled amino acids in the media. Our original approach of using undiluted, dialyzed yeastolate (ESF 921- Δ 11) also led to a significant reduction in all tested amino acids in the 11-amino acid type depleted medium (Table 1, column 4). On average, the use of 100% of dialyzed yeastolate led to a 10–20% *higher* amino acid content than the medium which contained 10% undialyzed yeastolate (compare columns 3 vs. 4). This finding is in qualitative agreement with our observation of moderately lower ^{13}C -enrichment levels than those reported by Gossert et al. (see below) in the growth media containing undiluted, dialyzed yeastolate. The Δ 11 growth medium containing 10% dialyzed yeastolate, Table 1, Column 5, exhibited at least another four-fold reduction in background amino acids, suggesting that an expression medium containing 10% dialyzed yeastolate should produce the highest levels of isotope enrichments.

Quantitation of isotopic label incorporation

To overcome the reported dilution in isotope labeling efficiency in proteins by the required presence of yeastolate in insect cell cultures, we tested the replacement of yeastolate with dialyzed yeastolate. As shown above, undialyzed yeastolate contains a significant level of amino acids as illustrated in Fig. 1b that can be incorporated into an expressed protein during bio-synthesis. An initial customized medium from Expression Systems, LLC was depleted in seven essential amino acids (i.e. Ile, Leu, Val, Met, Thr, Phe, Tyr) as well as yeastolate (ESF 921- Δ 7). Subsequent formulations of the growth medium included the depletion of all ten essential amino acids plus tyrosine (i.e. Arg, His, Ile, Lys, Leu, Met, Phe, Thr, Tyr, Trp, Val) as well as yeastolate (ESF 921- Δ 11). Prior to use, ESF 921- Δ 7 and ESF 921- Δ 11 were replenished with the unlabeled missing amino acids plus dialyzed yeastolate except for those amino acids which were to be added in labeled form. Initial tests proved encouraging. ^{13}C -incorporation efficiency in ubiquitin type-specifically labeled in threonine and in valine was compared by performing protein expressions with ESF 921- Δ 7 supplemented with either dialyzed or undialyzed yeastolate. As expected the ^{13}C -incorporation efficiency was significantly improved when dialyzed yeastolate was used. For example, in threonine labeled ubiquitin expressed while supplying the medium with 0.2 gram

per liter of labeled amino acid, the ^{13}C -incorporation was 62.4% (Table 2A) when using undialyzed yeastolate, whereas the ^{13}C -incorporation level jumped to 89.6% when using dialyzed yeastolate. Having determined the residual levels of unlabeled amino acid of the depleted residue types in ESF 921- Δ 11 media (Table 1), we also calculated the maximum achievable ^{13}C -incorporation levels when a set amount of labeled amino acid was added to the depleted medium. For example, the addition of 0.2 grams per liter of labeled threonine is expected to produce a maximum ^{13}C -level of 62.5% in the medium containing undialyzed yeastolate (Table 2 row 2, column 5). This closely matches the achieved incorporation of 62.4% (column 4). In expression runs in which ubiquitin was type-specifically labeled with valine, the observed ^{13}C -incorporation levels were 53.3 and 70.5% in media supplied with undialyzed yeastolate and 0.2 and 0.4 grams of labeled valine per liter of media respectively. The estimated upper limits of ^{13}C -incorporation based on the residual levels of unlabeled valine in the medium were 50.7 and 67.3%. These results suggest that measuring the residual level of unlabeled amino acids, including random coil peptides containing amino acids, provides a good measure of the maximum level of stable isotope enrichment which may be achievable.

Additional expressions of single amino acid type-specific labeled ubiquitin using the ESF 921- Δ 7 or ESF 921- Δ 11 were performed with dialyzed yeastolate at the concentration present in the commercial ESF 921 growth medium to assess ^{13}C -incorporation levels (Table 2B). The replacement of undialyzed yeastolate with dialyzed yeastolate significantly improved the level of ^{13}C -enrichment in baculovirus expressed type-specific labeled ubiquitin samples. For most expression runs, the achieved ^{13}C -incorporation levels were within two percentage points of the calculated upper limits of ^{13}C -incorporation levels. In expression runs of isoleucine labeled ubiquitin the difference between observed and expected ^{13}C -incorporation levels was about 5%. In the valine expression runs, a large gap of ~10% between expected and observed ^{13}C -incorporation levels in the expression run in which valine was supplied at 0.2 g per liter. Conversely, in the expression run where labeled valine was supplied at 0.4 g per liter there again was a close match between expected and observed ^{13}C -incorporation level (92.65% observed with 94.9% expected). The highest level of ^{13}C -incorporation level was achieved for tyrosine; the ^{13}C -incorporation level in tyrosine labeled ubiquitin was 92% and 94% in expressions where the growth medium supplied with 0.2 and 0.4 grams per liter of labeled tyrosine respectively. Interestingly, ^{13}C -labeled tyrosine was already present in ESF 921- Δ 11 upon addition of cell-paste while the other ^{13}C -labeled amino acid types were added following a four hour “starvation period” (see [Materials and methods](#) section for details).

Table 2 ^{13}C and ^{15}N incorporation efficiency

(A)							
Amino acid	^{13}C -aa added [g/l]	ESF 921- Δ 7 Undialyzed residual aa [g/l]	ESF 921- Δ 7 Undialyzed ^{13}C labeling (%)	$\%^{13}\text{C}$ labeling upper limit (%)	ESF 921- Δ 7 Dialyzed residual aa [g/l]	ESF 921- Δ 7 Dialyzed ^{13}C labeling (%)	$\%^{13}\text{C}$ labeling upper limit (%)
^{13}C , ^{15}N -Thr	0.2	0.12	62.4	62.5	0.016	89.60	92.6
^{13}C , ^{15}N -Thr	0.4	0.12	77.2	77.0	0.016	94.20	96.2
^{13}C , ^{15}N -Val	0.2	0.19	53.3	50.7	0.022	82.60	90.2
^{13}C , ^{15}N -Val	0.4	0.19	70.5	67.3	0.022	92.60	94.9
(B)							
Amino acid	^{13}C -aa added [g/l]	Residual aa [g/l]	^{13}C labeling (%)	^{13}C upper limit (%)	Dialyzed yeastolate (%)		
^{13}C , ^{15}N -Phe	0.2	0.029	85.90	87.4	100		
^{13}C , ^{15}N -Phe	0.4	0.029	92.00	93.3	100		
^{13}C , ^{15}N -Tyr	0.2		94.00		100		
^{13}C , ^{15}N -Tyr	0.4		96.80		100		
^{13}C , ^{15}N -Ile	0.2	0.022	85.00	90.3	100		
^{13}C , ^{15}N -Ile	0.4	0.011	92.60	97.4	100		
^{13}C , ^{15}N -Thr	0.2	0.016	89.60	92.6	100		
^{13}C , ^{15}N -Thr	0.4	0.016	94.20	96.2	100		
^{13}C , ^{15}N -Val	0.2	0.022	82.60	90.2	100		
^{13}C , ^{15}N -Val	0.4	0.022	92.60	94.9	100		
^{13}C , ^{15}N -Leu	0.2	0.045	79.90	81.7	100		
^{13}C , ^{15}N -Leu	0.4	0.045	86.10	89.9	100		
^{13}C , ^{15}N -Leu	0.2	0.006	90.50	96.9	10		
^{13}C , ^{15}N -Leu	0.4	0.006	95.30	98.4	10		

^{13}C and ^{15}N incorporation efficiency in single-amino acid type labeled ubiquitin in ESF 921- Δ 7 and ESF 921- Δ 11 containing various formulations of yeastolate, (A) undialyzed yeastolate at 100% level of commercial ESF 921 medium, (B) dialyzed at either 100% or 10% level of ESF 921

Column 2 gram per liter of labeled amino acid added to depleted growth medium

Column 3 residual unlabeled form of amino acid labeling type (calculated from the residual concentrations of amino acids in the ESF 921- Δ 7 or ESF 921- Δ 11 (Table 1)

Column 6 percentage of dialyzed yeastolate added to the depleted growth medium

Column 4 measured ^{13}C -enrichment levels of labeled residue type in expressed ubiquitin, column 5 maximum level of ^{13}C -enrichments based on concentrations of labeled and unlabeled amino acid (col. 2 & 3)

While our work was still ongoing, a paper appeared in print by Gossert et al. (2011) describing a new protocol for enhancing the isotope labeling efficiency. Previously, Gossert et al. showed that a complete removal of yeastolate produced a high isotope incorporation efficiency in excess of 90% at the expense of significant loss in protein expression yields (Strauss et al. 2003). Using a new protocol whereby only 10% of yeast extract was added to the growth medium, an almost complete restoration in protein expression yield resulted compared to expression in a rich medium.

These results by Jahnke and co-workers inspired us to test whether it might be possible to combine the two approaches whereby the amino acid depleted medium was

supplemented with 10% of dialyzed yeastolate. Analysis of ESF 921- Δ 11 supplemented with a ten-fold reduced amount of dialyzed yeastolate relative to the standard addition (ESF 921- Δ 11 10%D) produced on average close to four-fold further reduction in the content of residual “unlabeled”-amino acid compared to using dialyzed yeastolate at 100% of the concentration of the commercial ESF 921 growth medium. Expression of [^{13}C , ^{15}N]-leucine labeled ubiquitin showed a significant increase of [^{13}C]-content from 80% to over 90% using 0.2 g of labeled leucine in ESF 921- Δ 11 10%D (Table 2B, rows 11 & 13).

This [^{13}C]-labeling efficiency in [^{13}C , ^{15}N]-LEU-ubiquitin is comparable to the results of Gossert who expressed

[^{13}C , ^{15}N]-LEU-Abl kinase in a growth medium supplied with 0.25 g of labeled leucine and 10% undialyzed yeastolate (Gossert et al. 2011). This raises the question as to possible causes limiting labeling efficiencies of proteins expressed using $\Delta 11$ media supplied with ten-fold reduced amounts of dialyzed yeastolate.

One factor reducing isotope enrichment level is the incomplete enrichment levels in commercially available [^{13}C , ^{15}N]-amino acids which typically have enrichment levels of about ~99%, giving rise to a ~1% reduction in labeling efficiency. A second factor is the presence of unlabeled amino acids in the cell paste that is transferred to the labeled medium. This should give rise to a further reduction in labeling efficiency of an additional 2–5%. A third potential source of reduction in incorporation levels of isotope labels is the expression of target protein between the time of viral infection of insect cell and the addition of labeled amino acid(s) to the 11-medium. In the presented setup this time lag is 16 h for tyrosine and 20 h for other amino acid types. Only low levels of expression of target proteins is expected during the first 16–20 h post baculovirus-infection since the baculovirus polyhedrin promoter is activated only in the late stages of the viral life cycle (Parsarelli and Guarino 2007), (Jarvis 2009).

Furthermore, a significant drop in protein expression levels was reported when viral incubation period was reduced to less than 16 h (Meola et al. 2014). Possibly, a moderate reduction of the baculovirus incubation period may produce an incremental increase in protein labeling efficiency. Inspection of Table 2B shows the highest [^{13}C]-incorporation yield was observed in tyrosine. Using 0.4 g of [^{13}C]-tyrosine produced a [^{13}C]-incorporation yield of 96.8%. Assuming that the 0.4 g of tyrosine, which is the prescribed amount in the rich medium, produced close to 100% [^{13}C]-labeling efficiency suggests that 1–2% of unlabeled tyrosine may be due to protein expression during the viral incubation period. The lower than 96% [^{13}C]-incorporation yield in [^{15}N , ^{13}C]-leucine labeled ubiquitin which was expressed in the medium containing 10% dialyzed yeastolate may be due to residual expression of ubiquitin during the 4 h “starvation” period prior to the addition of fresh medium containing [^{13}C]-labeled leucine. Another potentially significant source of unlabeled amino acids may be cell paste which are transferred from the rich medium and which may comprise of 2–5% of the labeled media volume.

Depletion of amino acids in the growth medium during baculovirus protein expression

To reduce cost of expression with labeled amino acids, we assessed the level of amino acid consumption during two expression runs. We quantified the depletion

of amino acids by comparing their levels in aliquots of the growth media taken at the beginning and end of fermentation runs. Levels of amino acids were determined both for labeled amino acids and for unlabeled amino acids containing a methyl group by integration of cross peak volumes in the HC-HSQC spectra. Very little consumption of amino acids was observed when adding the amino acids at levels corresponding to the amounts in rich media. Table 3A lists the amount of each amino acid remaining at the end of the expression of [^{13}C , ^{15}N]-Thr, [^{15}N]-Val-CD4 as a percentage of its level at the beginning of the run. In every case, over 85% of the initial amino acid remained (Table 3A). Given the large amounts of amino acids remaining, we lowered the levels of amino acids down to the 0.12–0.20 g per liter range.

Table 3B lists the percentage of amino acids remaining after the expression of FAK1. Again, a large percentage of the individual amino acids remained at the end of the expression, demonstrating the expression is not limited by the depicted levels of amino acids being added to the medium. Not surprisingly, the depletion of an amino acid during an expression run is correlated with the concentration of the amino acid in the medium. For example, a reduction in the concentration of isoleucine (Ile) from 0.80 to 0.12 g per liter reduced the residual concentration at the end of the expression from 96.5 to 75.6%—see rows 2 and 3 in Table 3A, B respectively. The most pronounced depletion was observed in arginine (Arg), whose concentration declined almost two-fold during the protein expression when its concentration was reduced from 0.80 grams (rich medium) to 0.12 grams. Even so, greater than 50% of the original Arg in the sample remained at the end of the expression when only 0.12 g were added to the media at the start of expression. Interestingly, reducing the concentration of valine (Val) from 0.6 to 0.2 g per liter produced no noticeable reduction in Val content during a fermentation run.

Inspection of HC-HSQC spectra of growth media also revealed that apart from the amino acids, the concentrations of most components changed very little during the expression runs with a few notable exceptions (highlighted in Figure S1). Figure S1a shows a cross peak which indicates the reduction in concentration of one component by almost half. The [^1H] and [^{13}C] chemical shift of this cross peak were subsequently assigned to phosphocholine (PhCh). Figure S2B points to a second component that changes considerably during protein expression. A five-fold increase in intensity was seen for the alanine (Ala) methyl peak due to active synthesis of Ala during the expression run. This observation closely resembles the result by Sitarska et al. (2015) who reported a four-fold increase in Ala concentration during an expression run.

Table 3 Consumption of amino acids during growth and protein expression

(A)			
Type	% aa remaining c_{end}/c_{start} (%)	Amount used	Amount in rich medium
Ala	592.9	0.11	
Ile	96.5	0.80	0.80
Leu	86.5	0.50	0.50
Met	94.9	1.00	1.00
Thr	86.8	0.20	0.25
Val	98.4	0.20	0.60
PhCh	37.2	<i>0.033</i>	
(B)			
Type	% aa remaining c_{end}/c_{start} (%)	Amount aa added (g)	Amount in rich medium
Ala	593.0	0.07	
Arg	54.8	0.12	0.80
Ile	75.6	0.12	0.80
Leu	81.6	0.25	0.50
Met	91.3	0.12	1.00
Phe	76.4	0.12	0.80
Thr	85.6	0.25	0.25
Val	100.3	0.60	0.60
PhCh	48.4	<i>0.025</i>	

(A) Expression of [^{13}C , ^{15}N]-Thr, [^{15}N]-Val-CD4, 11-residue-depleted medium of 200 mg/L of labeled amino acids and the remaining nine depleted amino acids in the respective amounts in the rich medium

(B) Expression of [^{13}C , ^{15}N]-Arg, Ile, Phe, [^{15}N]-Leu, Met, Tyr, Val-FAK1, 11-residue-depleted medium of 100 mg/L of labeled amino acids and the remaining nine depleted amino acids in the respective amounts in the rich medium

% amino acid remaining was calculated from cross-peaks in HC-HSQC spectra on media samples taken at the start and end of a fermentation run. % remaining was measured in labeled amino acids and in methyl group containing amino acids. The listed concentrations of alanine in rows 1 highlighted by italics were measured

To make reasonable estimates in the accuracy of amino acid depletion Table 3 only depicts depletion of either [^{13}C]-labeled amino acids or methyl group containing amino acids. These results suggest that prior to performing protein expression runs using alternate labeling schemes, it may be prudent to run a trial expression to estimate amino acid depletion and assess minimal concentrations of nutrients in the growth medium prior to setting up a large-scale fermentation run.

Protein expression yields

In spite of the fact that insect cells do not grow well in the ESF 921- Δ 11 medium, which was replenished with dialyzed yeastolate and amino acids, we observed good protein expression yields. Expression of labeled KATII, FAK1 and CD4 all produced yields which exceeded what

was observed in the expression runs of unlabeled protein samples. However, all expressions of these proteins in unlabeled form were done at larger scale (i.e. 12–40 L), using wave bag fermentation. These differences in growth and expression conditions hampers accurate and meaningful comparisons in expression yields between labeled and unlabeled protein sample for these proteins.

For ubiquitin, however, expression yields of labeled and unlabeled were both based on shake flask expressions of identical volumes. We observed expression of type-specifically labeled ubiquitin was moderately higher than unlabeled protein.

In summary, it appears that the presented protocol for expressing type-specifically labeled protein appears qualitatively comparable to expression of unlabeled protein in the rich medium (ESF 921).

Application of baculovirus expression system to probe the interaction of an engineered adnectin-based inhibitor to CD4

In order to probe the site of interaction of the extracellular domains of CD4 with an engineered adnectin (Lipovsek 2010), the extracellular domain of CD4 was expressed in insect cell cultures with pairs of amino acid types isotopically enriched such that amino acid type 1 was [^{13}C , ^{15}N]-labeled and type 2 was [^{15}N]-labeled. Residue type selection was guided by choosing sequential $^{13}\text{C}(i-1)$, $^{15}\text{N}(i)$ -residue pairs which occur at a low abundance in CD4 and for which at least one of the sequentially [^{13}C] and [^{15}N]-labeled amino acid residue pairs are expected to reside near the putative binding site of the adnectin. In order to highlight HN-resonance pairs in ^{15}N -labeled amino acid residues which are preceded by a ^{13}C -labeled residue, CO-edited HN-TROSY spectra were recorded on a Bruker Avance III 700 MHz spectrometer equipped with a TCI-cryogenic probe. CO-edited HN-TROSY spectra were recorded on type-specifically [^{13}C , ^{15}N]-Thr, [^{15}N]-Val labeled CD4 in the absence and presence of unlabeled adnectin. At room temperature only poor-sensitivity HN-TROSY spectra were obtained. Raising the sample temperature to 40 °C produced high quality spectra (Fig. 2), which depicts the expanded spectral region of the superposition of CO-edited HN-TROSY spectra in samples of [^{13}C , ^{15}N]-Thr, [^{15}N]-Val-CD4 in the absence (blue contours) and presence (red contours) of unlabeled adnectin which was present at a 1.2-fold molar excess. The HN-resonances of all three expected Thr-Val-pairs were observed with one exhibiting a large shift perturbation highlighting the spatial proximity of respective Val residue to the adnectin binding interface. The quality of the depicted spectra is commensurate with the reported high efficiency of [^{13}C]- and [^{15}N]-labeling.

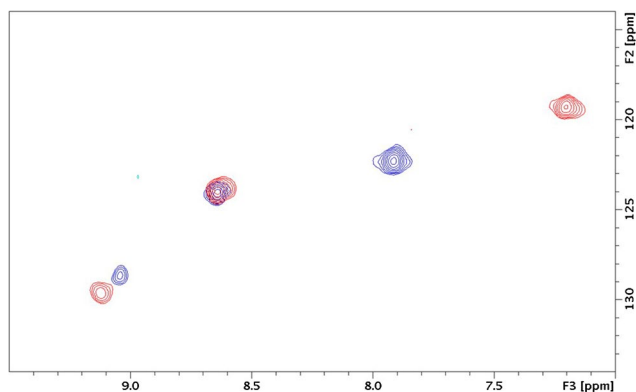


Fig. 2 Superposition of CO-edited HN-TROSY spectra in [^{13}C , ^{15}N]-Thr, [^{15}N]-Val-CD4 in *blue contours*: apo form & *red contours*: in complex with unlabeled adnectin

Conclusions

We have presented an improved protocol for producing type-specifically labeled protein samples in insect cell cultures. Enhancements in isotope labeling efficiency have been demonstrated by replacing yeast extracts in the growth medium by a ten-fold reduced amount of dialyzed yeast extract. In addition, we have shown that the ten-fold reduction in yeastolate reported by Gossert et al. in growth media still produces high level protein expression when dialyzed yeastolate is used. Hence, the use of ten-fold reduced amounts of dialyzed yeastolate offers the potential for enhanced levels of levels of incorporation in type-specifically labeled proteins.

Materials and methods

Protein expression vector generation

Sf9 cells (Invitrogen, Carlsbad, CA) were used to generate recombinant baculovirus and for virus amplification, according to the manufacturer's instructions. Sf9 cells (Expression Systems LLC, Davis, CA) for protein expression and production were routinely maintained in ESF 921 medium (Expression Systems, LLC: ESF 921-96-001) at 27 °C in shake flasks. Further details a provided in the supplemental materials section.

Preparation of type-selectively labeled growth media

Preparation of amino acid depleted stocks are described in detail in the supplemental materials section.

Protein expression

Sf9 or Hi5 cells were grown to the targeted density of 2×10^6 cells/ml in the rich medium. Cell growth was monitored and analyzed under a microscope, and as well by Vi-CELL XR Cell Viability Analyzer (Beckman Coulter, Indianapolis IN). After reaching the targeted density, the cells were inoculated with a high-titer secondary-passage baculovirus. After 16-hour post infection, cells were transferred and spun down in sterile centrifugation bottles (Corning, Tewksbury MA) by mild centrifugation. The cell pellets were washed, followed by a resuspension into media stock 2. Following a 4 h "starvation"-period, the labeled amino acids of concentrated stock solutions were added into the cell culture at the targeted concentration. Hence, labeled amino acids, with the exception of labeled tyrosine were added to the growth media 20 h after viral infection. The fermentation process continues for another 40–48 h followed by harvesting cells by centrifugation. The cell paste

was stored at -80°C for protein purification. For the production of CD4, the conditioned medium from cell culture was processed immediately following harvesting.

Protein purification

Labeled Ubiquitin, FAK1 and KATII samples expressed in Sf9 or Hi5 cells as cleavable N-terminal Hexa-His tagged constructs. CD4 was expressed in Sf9 cells containing a non-cleavable C-terminal Hexa-His tag. The adnectin, ATI-1637, was expressed in *E. coli* with a non-cleavable C-terminal Hexa-His tag. Further details on the respective purification protocols are provided in the supplemental materials section.

NMR

Analysis of residual amino acid contents in the $\Delta 11$ amino acid growth medium

To determine the residual level of depleted unlabeled amino acids in ESF 921-11, test samples were prepared with ESF 921-11 supplemented with various amounts of undialyzed and/or dialyzed yeastolate. Table S3 in the supplemental Material section depicts the compositions of the four media test samples.

HC-HSQC spectra in the growth media test samples were recorded on an AvanceIII HD 600 MHz spectrometer which was equipped with 5 mm TCI cryo probe. Spectra processed with cosine² apodization in both spectral dimensions, twofold zero-filling in both dimension, followed by baseline correction along F_2 using the automatic polynomial baseline correction tool in topspin. Table 1 depicts the residual percentages of unlabeled (natural abundance) of the depleted eleven amino acids were generated as follows: HC-HSQC spectra were uploaded into nmrView version 9.0.-b114 (OneMoon Scientific, Inc.). Cross-peaks were automatically picked using elliptic integration footprints followed by manual editing. Cross-peak integrals were exported into MS-Excel tables for further analysis of amino acid contents in growth media. The translation of cross-peak integrals of hc-hsqc spectra of growth media test samples (depicted in Table S3) into residual amino acid contents is depicted in Table S4.

NMR data on [^{13}C , ^{15}N]-Thr, [^{15}N]-val-CD4 n type-specifically labeled CD4

Sample preparation

Doubly type-specific [^{13}C , ^{15}N]-Thr and [^{15}N]-Val CD4 and unlabeled adnectin were exchanged into a buffer 25 mM HEPES, 300 mM NaCl, pH 7.0, 90% H_2O +10% D_2O ,

NMR sample 1:0.35 mM apo [^{13}C , ^{15}N]-Thr and [^{15}N]-Val CD4 in a 4 mm Shigemimico sample tube, Sample 2: 0.5 mM [^{13}C , ^{15}N]-Thr and [^{15}N]-Val CD4+0.6 mM (unlabeled) adnectin.

NMR data acquisition

All spectra were acquired on a Bruker AvanceIII 700 MHz spectrometer operating at 700.13 MHz equipped with a TCI 5 mm triple resonance cryo probe. Two-dimensional HNCO-TROSY (M. Salzman et al. 1998) were acquired, ^{15}N -evolution in the indirect detection domain and proton detection at 40°C were acquired on an Bruker Avance III 700 MHz spectrometer employing the following NMR data acquisition parameters: number of t_1 -increments=112, number of t_2 -data points=1024, ns=352, relaxation delay=1.6 s, sw3=15 ppm, sw1=35 ppm, total acquisition time=19 h, processed in topspin 3.2 with cosine square apodization in both time domains with two-fold zero filling. Spectra are displayed with geometric spacing of contour-levels with a coefficient of 1.3.

Mass spectrometric analysis of labeled protein samples

Protein digestion

Each sample was diluted to ~ 1 mg/mL with 100 mM NH_4HCO_3 (pH 7.5). 25 μg of sample was diluted :1 (v/v) with 0.1% RapiGest (Waters, Milford MA), and heated at 80°C for 15 min. Then each sample was digested with trypsin (Promega, Madison WI) at a ratio of protein: trypsin of 50:1. The digestion was kept at 37°C overnight, and inactivated by 6 μL 5% trifluoroacetic acid. Each digested solution was centrifuged at 14,000 RPM for 5 min and supernatant was analyzed by LC/MS.

LC/MS instrumentation and data analysis

Digested samples were separated using an ultra high pressure LC (ACQUITY UPLC, Waters) equipped with BEH C18 column, 2.1×150 mm, 130 \AA with 1.7 μm particle. The mobile phase A was 0.1% formic acid in water, and mobile phase B was 0.1% formic acid in acetonitrile. The injection volume for each digested sample was 1 μL . Tryptic peptides were separated in a gradient of 1–40% B in 23 min at 45°C column temperature.

Peptides were detected using SYNAPT G2 (Waters) with positive ion electrospray. Mass acquisition range was 50–2000 m/z, resolution was 20,000. Peptide identification was performed with accurate mass at both precursor ion level and fragment ion level. The relative labeling efficiency for each labeling site was calculated from the ratio of the peak intensity of labeled peptide containing the

labeling site divided by the sum intensities of the labeled and unlabeled peptides.

References

- Assenberg R, Wan PT, Geisse S, Mayr LM (2013) Advances in recombinant protein expression for use in pharmaceutical research. *Curr Opin Struct Biol* 23(3):393–402
- Bellizzi, J. J., III, J. Widom, C. W. Kemp, J. Clardy (1999). Producing selenomethionine-labeled proteins with a baculovirus expression vector system. *Structure* 7(11): R263–R267
- Bruggert M, Rehm T, Shanker S, Georgescu J, Holak TA (2003) A novel medium for expression of proteins selectively labeled with 15 N-amino acids in *Spodoptera frugiperda* (Sf9) insect cells. *J Biomol NMR* 25(4):335–348
- Creemers AF, Klaassen CH, Bovee-Geurts PH, Kelle R, Kragl U, Raap J, de Grip WJ, Lugtenburg J, de Groot HJ (1999) Solid state 15 N NMR evidence for a complex Schiff base counterion in the visual G-protein-coupled receptor rhodopsin. *Biochemistry* 38(22):7195–7199
- Cristiani, C (2003) Tailoring protein kinases for structural studies. Poster, 6th Annual Meeting Baculovirus and Insect Cell Culture.
- DeLange F, Klaassen CH, Wallace-Williams SE, Bovee-Geurts PH, Liu XM, DeGrip WJ, Rothschild KJ (1998) Tyrosine structural changes detected during the photoactivation of rhodopsin. *J Biol Chem* 273(37):23735–23739
- Fremont DH, Crawford F, Marrack P, Hendrickson WA, Kappler J (1998) Crystal structure of mouse H2-M. *Immunity* 9(3):385–393
- Gossert, A. D. and W. Jahnke (2012) Isotope labeling in insect cells. *Adv Exp Med Biol* 992:179–196
- Gossert AD, Hinniger A, Gutmann S, Jahnke W, Strauss A, Fernandez C (2011) A simple protocol for amino acid type selective isotope labeling in insect cells with improved yields and high reproducibility. *J Biomol NMR* 51(4):449–456
- Jarvis, D. L. (2009). Baculovirus-insect cell expression systems. *Methods Enzymol* 463:191–222
- Kelly, B. J., L. A. King, R. D. Possee (2007). Introduction to baculovirus molecular biology.” *Methods Mol Biol* 388:25–53
- Lipovsek, D. (2010). Adnectins: engineered target-binding protein therapeutics. *Protein Eng Des Sel* 24(1–2):3–9
- Maiorella B, Inlow D, Shauger A, Harano D (1988) Large-scale insect cell-culture for recombinant protein production. *Bio/Technology* 6(12):1406–1410
- Meola A, Deville C, Jeffers SA, Guardado-Calvo P, Vasiliauskaite I, Sizun C, Girard-Blanc C, Malosse C, van Heijenoort C, Chamot-Rooke J, Krey T, Guittet E, Petres S, Rey FA, Bontems F (2014) Robust and low cost uniform 15 N-labeling of proteins expressed in *Drosophila* S2 cells and *Spodoptera frugiperda* Sf9 cells for NMR applications. *J Struct Biol* 188(1):71–78
- Passarelli, A. L. and L. A. Guarino (2007) Baculovirus late and very late gene regulation. *Curr. Drug Targets* 8(10):1103–1115
- Schlaeger EJ (1996) Medium design for insect cell culture. *CytoTechnology* 20(1–3):57–70
- Sitarska A, Skora L, Klopp J, Roest S, Fernandez C, Shrestha B, Gossert AD (2015) Affordable uniform isotope labeling with 2 H, 13 C and 15 N in insect cells. *J Biomol NMR* 62(2):191–197
- Strauss A, Bitsch F, Cutting B, Fendrich G, Graff P, Liebetanz J, Zurini M, Jahnke W (2003) Amino-acid-type selective isotope labeling of proteins expressed in baculovirus-infected insect cells useful for NMR studies. *J Biomol NMR* 26(4):367–372
- Strauss A, Bitsch F, Fendrich G, Graff P, Knecht R, Meyhack B, Jahnke W (2005) Efficient uniform isotope labeling of Abl kinase expressed in baculovirus-infected insect cells. *J Biomol NMR* 31(4):343–349
- Walton WJ, Kasprzak AJ, Hare JT, Logan TM (2006) An economic approach to isotopic enrichment of glycoproteins expressed from Sf9 insect cells. *J Biomol NMR* 36(4):225–233
- Weiss, S. A., G. C. Smith, S. S. Kalter, J. L. Vaughn (1981). Improved method for the production of insect cell cultures in large volume. *In Vitro* 17(6):495–502