Production of isotope-labeled proteins in insect cells for NMR

Bastian Franke[1,§], Christian Opitz[1,§], Shin Isogai[1], Anne Grahl[1], Leonildo Delgado[1], Alvar D. Gossert[2*], Stephan Grzesiek[1*]

[1]Focal Area Structural Biology and Biophysics, Biozentrum, University of Basel, 4056 Basel, Switzerland

[2]Department of Biology, ETH Zürich, 8093 Zürich, Switzerland

[§]These authors contributed equally to this work.

[*]Address correspondence to:

Stephan Grzesiek, Focal Area Structural Biology and Biophysics, Biozentrum, University of Basel, 4056 Basel, Switzerland. E-mail: stephan.grzesiek@unibas.ch; Phone: +41 61 207 21 00

Alvar D. Gossert, Department of Biology, ETH Zürich, 8093 Zürich, Switzerland, E-mail: alvar.gossert@mol.biol.ethz.ch; Phone: +41 44 633 3455

Keywords

Summary

Baculovirus-infected insect cells have become a powerful tool to express recombinant proteins for structural and functional studies by NMR spectroscopy. This article provides an introduction into the insect cell/baculovirus expression system and its use for the production of recombinant isotope-labeled proteins. We discuss recent advances in inexpensive isotope-labeling methods using labeled algal or yeast extracts as the amino acid source and give examples of advanced NMR applications for proteins, which have become accessible by this eukaryotic expression host.

Introduction

Proteins expressed in *E. coli* dominate structural and biophysical studies due to the ease of genetic manipulation, rapid growth, and high expression yields. This is reflected in the large fraction of entries in the Protein Data Bank (PDB) derived from this system, which presently (12/2017) amounts to 74% (Table 1). The simple, efficient, and inexpensive ^{15}N, ^{13}C, and ^{2}H isotope labeling in this expression host from only ammonium chloride, glucose and water make *E. coli*-expressed proteins particularly accessible to analysis by heteronuclear NMR.

However, for more complex proteins originating from higher organisms, heterologous *E. coli* expression often fails to provide properly folded and functional material suitable for structural analysis. These proteins frequently require the complex machineries for protein folding, post-translational modification, or membrane insertion of higher eukaryotes, which are less developed or absent in *E. coli*. For this reason, the use of eukaryotic expression hosts like yeast, mammalian, or insect cells has increased steadily over the last 25 years (Figure 1) and accounts now for about 8% of all PDB entries. Half of these (4.3%) derive from insect cell expression, whereas much smaller fractions account for yeast and mammalian cells. For proteins with molecular weights higher than 100 kDa or of human origin, the fraction from insect cell expression increases to 7.3% and 10.4%, respectively, whereas for the 'notoriously difficult to make' G protein-coupled receptors (GPCRs), it even dominates with 70.5% (Table 1).

Despite this clear success for structure determination, the insect cell expression system has not seen wide use in NMR. Thus, the PDB currently comprises only two entries (1T50, 1QLO) derived from NMR data using insect cell-expressed, but unlabeled proteins. This is certainly due to the required more elaborate expression protocols as well as the higher costs for the incorporation of isotopes, which need to be provided as isotope-labeled amino acids. Nevertheless, insect cell expression with isotope-labeling has now been used successfully in a number of NMR studies on complex eukaryotic proteins, not accessible from *E. coli*, to provide valuable mechanistic insights into protein kinases[1-3] and GPCRs[4-7]. This is concomitant with the development of robust expression and isotope labeling protocols[1,2,8-17], which are easy to apply in a typical NMR protein production lab.

In this article, we give an introduction into the basic concepts of protein production for NMR using the insect cell system, highlight recent developments of inexpensive ^{15}N, ^{13}C, and ^{2}H-labeling methods and give several examples of advanced NMR applications.

The baculovirus expression vector system: baculovirus and insect cell lines

The baculovirus expressing vector system (BEVS) emerged in the 1980s as a novel technology for heterologous protein production[18]. It is based on insect cell lines in suspension cultures, which are infected by recombinant baculoviruses (BVs) to express the protein of interest.

The baculovirus

The best studied BV is the *Autographa californica* multiple nuclear polyhedrosis virus, which has been isolated from the alfalfa looper *Autographa californica*[18,19]. Once inside a cell, the BV infection cycle spans three major phases – termed the early, late and very late phase. During the early phase, the budded baculovirus enters the cell by absorptive endocytosis. The late phase occurs 6 hours post infection (hpi) and is accompanied by viral DNA replication and gene expression as well as budded BV production. In the very late phase at 24–96 hpi, the cell starts to express the highly abundant polyhedrin, an envelope protein forming the so-called BV occlusion bodies that protect virus particles from the environment once they have been released from the host cell. In most BEVS, the viral polyhedrin gene is replaced by the gene of interest. The expression of the latter is then driven by the strong polyhedrin promoter in the very late phase of baculovirus gene expression in virus-infected cells. Alternatively, there is a second viral promoter with very similar characteristics, the p10 promoter. It is typically used when expressing multiple proteins on a single bacmid (e.g. MultiBac™, Geneva Biotech), as repeated use of the polyhedrin promoter may increase the risk of gene instability due to recombination[20]. Historically, the polyhedrin promoter was favored, because replacing the polyhedrin gene leads to a visible phenotype lacking occlusion bodies, which allows identifying recombinant virus plaques. Replacing the p10 gene, however, has the advantage that the Baculovirus does not lyse the cells towards the end stage of infection, potentially leading to higher expression yields and expression times beyond 72 hours.

Generation of recombinant baculovirus

Today two main BEVS platforms are available: (i) the original baculovirus shuttle vector system[21], commercially available under the name Bac-to-Bac™ (Invitrogen). This system is based on a modified *E. coli* strain (DH10Bac™), which produces recombinant, viral DNA (the bacmid) from a donor plasmid containing the gene of interest. The purified bacmid is then used to transfect insect cells that generate the virus particles. For production of multi-component protein complexes, a modified version of the Bac-to-Bac system (MultiBac™, Geneva Biotech) can be used, which provides expression of multiple proteins on a single bacmid[20]. (ii) The *flash*BAC™ system (Oxford Expression Technologies) allows direct transfection of insect cells simultaneously with a bacmid and a vector containing the target DNA. Recombination then occurs directly in the insect cells[22].

In both approaches, recombinant virus particles are collected from the supernatant of cultures and amplified by repeated rounds of infection. Usually three rounds (V_0, V_1 and V_2) are sufficient to obtain ~200 mL of a high-titer virus stock. The titer is determined by adding diluted virus stock to insect cells grown in an adherent monolayer and counting the plaques of lysed cells caused by the infection. The titer value is the number of plaque forming units (pfu) per volume of virus, which for a high-titer V_2 amounts to typically ~2 x 10^8 pfu/mL.

The virus stocks can be stored at 4°C for several weeks without reduction in titer. However, to achieve stability for several months and even years, the recombinant BV can be stored in form of cryo-preserved baculovirus-infected insect cells (BIICs)[23]. The following protocol works well in our hands. Insect cells are infected at a density of 2.0 x 10^6 cells/mL with V_1 and harvested when the cell diameter increases by 20-30%, i.e. usually after 24-48 h. Cells are then centrifuged at 200 g for 10 min at room temperature and dissolved at a density of 10^7 cells/mL in a mixture of 45% of the used medium, 45% fresh medium, and 10 % DMSO for cryo storage. Expression yields of insect cells infected with either freshly made virus or BIICs are typically comparable within 20%.

Cell lines and their characteristics

At present, three main types of insect cells are used in research and biotechnological applications with BV infection – Sf21, Sf9 and Tn5. (*Drosophila melanogaster* S2 cell lines can also be used for stable transfection with target genes for protein production, but we limit this article to transient transfection.) Sf21 cells were originally isolated from primary explants of the American fall army worm's (*Spodoptera frugiperda*) ovarian pupal tissue and their clonal derived isolates were named Sf9[24]. Tn5 (also called high five™) cells originate from the ovarian tissue of the cabbage looper (*Trichoplusia ni)*[25]. These three strains have different characteristics in terms of growth rate, virus production capacity and protein expression. Tn5 cells have faster doubling times than Sf9 and Sf21 cells (18–24 h vs. 24–30 h, respectively), but produce less BV. Sf9 cell lines are regarded as the cell line with highest virus production. They form excellent monolayers due to their small size, making virus plaque determination easier in the initial steps of virus production. With respect to protein expression yield, the larger Sf21 cells may express more cytosolic protein compared to Sf9, whereas Tn5 cells have been reported to express 5–10-fold higher levels of secreted proteins than the Sf cell lines. Despite these general trends, both Sf and Tn cells should be tested in order to determine the highest expression yields for an individual protein construct.

Expression

Suspension cultures of insect cells can be grown in media containing a complex formulation of inorganic salts, carbohydrates, lipids, vitamins, and amino acids[26]. Commercial media with optimized performance are further supplemented with algal or yeast extracts and possibly other components of plant or animal origin such as fetal calf serum. Typically, the cultures are grown at a temperature of 27°C in conical flasks or centrifugal tubes shaken at ~90 rpm and ~50 mm shaking diameter. Maintenance requires weekly serial passaging with about ten-fold dilution, and cells are kept in exponential growth phase (≤4 x 10^6 cells/ml). Prior to infection, cells are grown to a density of 1.5–2.0 x 10^6 cells/ml. To optimize large-scale expression, functional titers of the virus stock should be determined using freshly made V_2 stocks or BIICs in small-scale cultures in order to determine the optimal virus-to-cell ratio – the multiplicity of infection – which yields the highest amounts of expressed protein. Cell diameter and viability should be checked every 24 hpi with an inverted microscope or a suitable automated cell counter. A functional virus infection leads to an increase in cell diameter from ~16 μm to above 20 μm, whereas a too strong virus infection is detected by a fast decline in viability (e.g. as assayed by trypan blue) to below 90% within less than 48 hpi.

Heterologous protein expression usually starts 24–48 hpi and peaks around 72 hpi. For optimal protein yield, insect cells should be harvested when cell viability drops to 80–90%.

The green fluorescent protein (GFP) is convenient to establish optimal expression conditions in small cell cultures without any purification by quantifying protein levels from the intrinsic fluorescence. In our hands, the truncated, cycle3 variant[27] of GFP provided high expression and at the same time high-quality NMR spectra, which are not affected by the exchange broadening present in other GFP variants.

The implementation of the baculovirus/insect cell system within a standard structural laboratory environment requires only a sterile hood, but no further modifications or expensive additional equipment, e.g. the same type of shaker incubators as for *E. coli* can be used. This is an important advantage over mammalian cell cultures, which are cultivated at 5% CO_2 and 75% humidity, requiring special incubators. Although BV is pathogenic to insect cells, it does not pose any threats to vertebrates allowing handling at standard biosafety level 1.

Isotope labeling

Since insect cells have a very limited capacity to synthesize amino acids from simpler precursors, the production of isotope-labeled proteins from insect cells is usually based on the supplementation of entire labeled amino acids to growth media depleted in amino acids and other proteinaceous material. Complete formulations of media for uniform ^{15}N and ^{13}C labeling are commercially available (BioExpress, CIL), but very costly. For in-house preparation of labeling media, custom-made media without amino acids, carbohydrates, yeast extracts or other sera from biological sources can be purchased from a variety of suppliers (e.g. Bioconcept, Lonza, Expression Systems, Gibco). In the simplest case, commercially available isotope-labeled amino acids are added to these media[1-8,11,12,14,17]. This is very efficient for selective labeling of a few amino acids with specific labeling patterns, but becomes prohibitively expensive for uniform labeling. In particular, deuterated amino acids are very expensive and at the moment some are not even commercially available in ^{2}H,^{13}C,^{15}N-labeled form. As a more economical alternative, the supplementation of labeled algal or yeast extracts as an amino acid source (Figure 2) has been developed in recent years[9,10,13-16]. In this way, deuteration can also be achieved at moderate extra costs.

Selective labeling

Selective labeling of a few amino acid types with specific isotope patterns has the advantage of reducing spectral complexity and simplifies resonance assignments. However, due to the loss of backbone connectivity, application of conventional triple resonance NMR assignment strategies may not be possible. Not all amino acid types lend themselves equally well for selective isotope labeling. High labeling efficiency (>90%) can be achieved for a limited set of essential or semi-essential amino acids of the insect cell metabolism (i.e. isoleucine, leucine, lysine, methionine, phenylalanine, tyrosine, threonine, cysteine, tryptophan, and valine). Some further non-essential amino acids can be labeled with acceptable levels of scrambling, e.g. in a pioneering study on Abelson kinase ^{15}N-glycine was incorporated to 78% and produced 7 additional signals on top of 16 expected ones[1]. Scrambling values for serine, proline, arginine and histidine have not been reported so far. In contrast, amino acids directly related to glycolysis and the citric acid cycle (i.e. alanine, asparagine, aspartate, glutamine, and glutamate) are significantly affected by scrambling, which also causes ^{13}C dilution if carbohydrates such as glucose are not provided in their

labeled form. For these amino acids, selective labeling can be achieved by suppressing unwanted metabolism (see section on amino acid metabolism below).

The selective labeling approach in combination with solution NMR has provided interesting functional information for a number of systems. In one of the first applications, the Abelson kinase domain, a drug target for chronic myelogenous leukemia (CML), was labeled uniformly and selectively in various isotope patterns by the addition of pure isotope-labeled amino acids[1,2,28]. Due to the lack of deuteration, the sensitivity of NMR experiments involving $^{13}C_\beta$ or other side chain nuclei was very low for this 277-amino acid protein, but almost complete backbone assignment was achieved. The hitherto unknown conformation of the kinase activation loop in solution could then be derived from ^{1}H-^{15}N RDC data for various inhibitor complexes (Figure 3a)[3]. In a further study using backbone NMR, 26 ^{1}H-^{15}N resonances could be observed for the G protein-coupled turkey β_1-adrenergic receptor (β_1AR), which was selectively labeled by ^{15}N-valine[7]. Due to the 100-kDa molecular weight of the receptor-detergent micelle and the absence of deuteration, the sensitivity of triple resonance assignment experiments was too low, but 21 ^{1}H-^{15}N resonances could be assigned from point mutations. The response to six different agonist and antagonist ligands as well as the binding of a G protein mimetic nanobody could be followed on these valine resonances, revealing in particular the activating motion induced by the agonists (Figure 3b).

The favorable relaxation properties of methyl groups[29] provide higher sensitivity of NMR detection than backbone resonances albeit at the cost of a less clear correlation between chemical shift and the underlying structural determinants. Selective labeling by (methyl-^{13}C)-methionine has been applied to the β_2-adrenergic receptor (β_2AR) revealing changes in the equilibrium populations of methionine-82, which correlate to modulations of the signal transduction level in several ligand-receptor complexes[4]. More recently, (methyl-^{13}C)-methionine labeling has also been applied to the β_1AR[30], and multiple equilibria of ligand-bound and G protein mimetic nanobody-bound forms as well as respective changes in receptor dynamics could be detected. Extending this approach, labeling by ($\alpha\beta\gamma$-^{2}H, methyl-^{1}H-^{13}C)-methionine has been combined with the use of commercial, highly deuterated algal amino acid mixtures to obtain selectively (methyl-^{1}H-^{13}C)-methionine labeled β_2AR on a deuterated background. The protocol requires the careful supplementation of several, additional deuterated amino acids at well-defined time points during expression and achieved incorporation levels of ^{2}H into side chain and non-exchangeable backbone hydrogen positions as high as 90% for 14 amino acids. The approach enabled studies of the β_2AR dynamics in nanodiscs[6] and has also been applied to the μ-opioid GPCR[31].

Uniform labeling using algal and yeast extracts

To overcome the high costs for uniform labeling that arise from adding pure isotope-labeled amino acids to the insect cell medium, several protocols have been described, which use isotope-labeled algal[15] or yeast[9,10,13,16] extracts as the principal amino acid source (Figure 2).

Algal extract

^{13}C-labeling of algae is cost-effective since $^{13}CO_2$ can be used as the sole carbon source. However, algae cultivation requires special bioreactors equipped with a light-source[32] and for ^{13}C-labeling a closed supply of $^{13}CO_2$. Algal extracts are commercially available in all combinations of uniform ^{15}N, ^{13}C, and ^{2}H labeling (e.g. ISOGRO® Sigma-Aldrich, Celtone® Cambridge Isotope Laboratories). Due to the acid hydrolysis during the commercial production, the amino acids cysteine and tryptophan are lost, whereas asparagine and

glutamine are hydrolyzed to the corresponding aspartic and glutamic acids. Supplementing the missing four amino acids in labeled form would be very costly. However, insect cells can produce cysteine from methionine and serine, which both are present in the medium. Furthermore, glutamine and asparagine can be produced upon addition of ammonium (see below) from glutamate and aspartate, respectively, which are abundant in the extract. In contrast, tryptophan is an essential amino acid and needs to be supplemented. A recent approach exploits this idea[15] using commercially available algal extract (10 g/L), tryptophan (20 mg/L), glucose (5 g/L), and ammonium chloride (250 mg/L) with the desired isotope-labeling patterns as supplements to a serum-free medium depleted in amino acids and glucose (Figure 2). Incorporation levels of 80% were achieved for uniform ^{15}N-labeling, 77% for ^{15}N,^{13}C-labeling, and ~73% for ^{2}H,^{15}N and ^{2}H,^{13}C,^{15}N-labeling. Yields ranged between 2–5 mg/L for the tested membrane protein and 12–22 mg/L for various soluble human proteins. Using labeled ammonium as a substrate for producing labeled glutamine and asparagine in insect cells is cost-effective, but has the disadvantage that ammonium hinders virus entry into the cell. Therefore, when using this protocol, cells need to be infected in full medium and are only transferred to the labeling medium at 16 hpi, before the protein is expressed.

Yeast extract

Isotope-labeled yeast extracts as the amino acid source for insect cells can be produced with moderate effort using standard protein expression equipment[9,10,16] or obtained commercially[13] (Cortecnet). A step-by-step procedure to produce such yeast extracts from wild type *Pichia pastoris* has been described recently[16]. The yeast is grown in a fermenter under control of glucose feed, oxygen, pH, and temperature using a minimal medium based on glucose, ammonium, and water as the sole carbon, nitrogen and hydrogen sources. All desired isotope-labeling patterns can be achieved using these basic ingredients in suitable isotope-labeled forms (Figure 2). In particular, yeast growth is possible on 100% ^{2}H$_2$O. After fermentation, yeast cells are autolysed during a high-temperature (50°C, 5 days) incubation step with papain before being micro-filtered with a 10 kDa cut-off and lyophilized to yield yeastolate. This procedure avoids the harsh acid hydrolysis used for commercial algal lysates and largely preserves all amino acids including asparagine and tryptophan (cysteine was not quantified) as determined by quantitative amino acid analysis[16]. In addition to free amino acids, yeastolate also contains incompletely cleaved polypeptides, sugars, and other cellular compounds, which may be metabolized by the insect cells. About 6.6 g yeastolate is obtained per liter of yeast culture from 27.5 g glucose and 4.3 g ammonium chloride. Although glutamine is present in low concentration in the yeast extract, its supplementation considerably boosted expression levels (see below). Thus a supplement of 8 g/L of ^{13}C, ^{15}N-yeastolate and 1 g/L of ^{15}N$_2$-glutamine to the drop-out medium provided high yields (40–80 mg/L for soluble proteins such as the human Abl kinase domain or GFP, 1–2 mg/L for GPCRs) and high (~90%) uniform-^{15}N/^{13}C incorporation[16]. Deuterium incorporation levels of ~60% were achieved when supplementing ^{2}H-labeled yeastolate obtained from yeast grown on 100% ^{2}H$_2$O, but protonated glucose. Figures 3c,d show as examples the ^{1}H-^{15}N TROSY and one-dimensional ^{15}N T$_1$ and T$_2$ relaxation data obtained on the ^{15}N/^{2}H-labeled HIV coreceptor CCR5.

Algal vs yeast extracts

Both algal and yeast extracts have respective advantages and disadvantages when used as isotope-labeled amino acid sources. Commercial extracts provide a good starting point for

initial uniform labeling experiments. At the moment, only ^{15}N-labeled yeast extracts are commercially available (Cortecnet). In contrast, commercial algal extracts can be obtained in all uniform labeling patterns including ^{2}H, ^{13}C and ^{15}N at high isotope incorporation levels (98%). As they are a byproduct of commercial ^{13}C glucose production from [13]CO_2, ^{13}C-labeled algal extracts are relatively inexpensive. However, we observed batch-to-batch variations within the commercial extracts, which occasionally resulted in significant reductions of expression levels. The acid hydrolysis step during the production destroys several amino acids, which need to be supplemented. In principle, the harsh hydrolysis could be avoided by producing the lysate from suitable algal starting material in the home laboratory or by an adaptation of the commercial production process. However, home growth of isotope-labeled algae itself seems very demanding.

In contrast, the production of yeast and yeastolate is less complicated and requires only standard equipment. Once established, it provides full control over all ingredients for the labeling. In particular, the mentioned acid hydrolysis can be avoided and most amino acids preserved. Furthermore, increased capabilities for customization of the labeling scheme become available. For example, yeast extracts with protonated methyl groups on a deuterated background may be produced from yeast cultures grown on [2]H_2O and protonated pyruvate as described for *E. coli*[33].

Due to the commercial production process, algal extracts from *Spirulina* contain a higher total fraction of amino acids (~65% dry weight) than yeast extracts (~33% dry weight). In the published protocols, algal extracts were added at 10 g/L concentration with the aim of highest isotope incorporation. This is close to the maximally tolerated concentration of about 12 g/L, at which the too high osmolarity of the medium severely affects the cell viability. In contrast, the yeast extracts were added at 8–10 g/L, which corresponds to about half of the concentration of amino acids in the final insect cell medium derived from algal extracts. Nevertheless, high ^{15}N and ^{13}C incorporation rates of ~90% are achieved under these conditions from the yeast extracts.

In our experience, expression of soluble proteins such as Abelson kinase and GFP and also of membrane proteins such as GPCRs worked well with either algal or yeast extracts, but we have not carried out a systematic analysis. It appears worth testing both extracts as well as their combination to obtain optimal expression of a certain protein.

Amino acid metabolism

While the glucose metabolism can be ignored for ^{15}N labeling, it must be considered for ^{13}C labeling and ^{2}H labeling. Glucose acts as precursor for alanine, glutamate/glutamine, and aspartate/asparagine through the citric acid cycle and further for serine and glycine through glycolysis. Therefore, ^{13}C labeling is considerably reduced for these amino acids, when glucose is supplemented in ^{13}C-labeled form to the medium, with isotope dilution being most severe for alanine, and decreasingly important for glutamate/glutamine, aspartate/asparagine, serine, and glycine.

For [2]H-labeling, the situation is further complicated, since deuterated glucose considerably reduces viability of insect cells. We have not systematically tried to adapt insect cells to [2]H-glucose, but initial results were discouraging. Thus, the ^{2}H incorporation level cannot easily be increased further by the use of deuterated glucose in the medium. This also represents a challenge for increasing deuteration levels using yeast extracts, since protein expression

levels in insect cells decreased considerably using extracts from yeast grown on deuterated glucose and ^2H_2O as compared to protonated glucose and ^2H_2O.

Of the amino acids that are directly linked to carbohydrate metabolism, alanine is particularly strongly affected by isotope dilution. Incorporation levels of ^{13}C are typically below 30% if unlabeled glucose is used. However, ^{13}C-incorporation levels of alanine can be restored to ~80% for growth on ^{13}C-labeled algal extracts in the presence of unlabeled glucose when cycloserine is added to inhibit the pyruvate-alanine transaminase[15].

Like glucose, glutamine acts as a metabolic precursor for glutamate, aspartate/asparagine as well as alanine and should be added in isotope-labeled form to obtain high labeling efficiencies for these amino acids.

Analysis of isotope incorporation by mass spectrometry

Detailed, quantitative knowledge on the isotope incorporation level of all amino acids is crucial as quality control and as input for strategies to improve labeling efficiencies. Total isotope incorporation of an entire biomolecule is best studied by ESI mass spectrometry, which provides the indiscriminate mass increase due to all ^{15}N, ^{13}C, and ^2H isotopes[1,2,6,16]. In principle, atom-specific isotope incorporation levels may be obtained on intact proteins from a careful analysis of J-coupling transfer or splitting effects in NMR spectra[2,6,16]. In practice, the NMR approach is limited by the spectral resolution and becomes rapidly insensitive for higher molecular weight biomolecules. Acid hydrolysis to single amino acids improves the spectral resolution in NMR[34], but still both sensitivity and resolution remain serious problems for the detection of the various isotopologues and isotopomers.

We found that separation of single amino acids by liquid chromatography coupled to mass spectrometry (LC/MS) is the most efficient and sensitive way to determine their isotope composition[16]. For this, the free amine groups of amino acid acids are derivatized by phenylisothiocyanate (PITC, Edman's reagent) to phenylthiocarbamyl-amino acids and separated via an HPLC column (Figure 4a) before being injected into an ESI mass spectrometer. The HPLC step separates single amino acids according to their type and also from higher peptides and other compounds. Thus the method can be used to detect single amino acids in crude cellular extracts. Besides cysteine, all amino acids are detectable by this method. The high resolution of modern Fourier transform mass spectrometers (Orbitrap Elite™, Thermofisher) separates the amino acids into individual isotopologues by the distinct mass differences of hydrogen, carbon, and nitrogen isotopes ($\Delta m(^2H-^1H) = 1.00628$, $\Delta m(^{13}C-^{12}C) = 1.00335$, $\Delta m(^{15}N-^{14}N) = 0.99704$). It is thus possible to simultaneously analyze ^2H, ^{13}C and ^{15}N incorporation in a single experiment. Figure 4b shows as an example the separation of phenylthiocarbamyl-histidine into $^2H_{n1}{}^1H_{n2}{}^{15}N_{n3}{}^{14}N_{n4}$ isotopologues. The entire analysis of the $^2H,^{15}N$ incorporation of the free amino acids present in yeastolate grown on 98% ^{15}N ammonium chloride, 98% ^2H_2O, and protonated glucose is depicted in Figure 4c. For almost all amino acids, the ^{15}N incorporation is around 95%. However, due to the protonated glucose the average ^2H incorporation is only about 70% and varies strongly for the different amino acid types.

The identical analysis can be carried out on acid-hydrolyzed proteins. During the acid hydrolysis asparagine and glutamine are converted to aspartate and glutamate, respectively, whereas tryptophan is destroyed. Typical results are shown in Figure 4d for hydrolyzed GFP expressed in Sf9 insect cells grown on $^2H,^{15}N$-yeastolate and $^{15}N_2$-glutamine. Uniformly high (>85%) ^{15}N incorporation is achieved for all detected amino acids. Conversely, ^2H

incorporation is lower (~60% average) and strongly amino acid-dependent. Consistent with their metabolic reactions within the citric acid cycle, nearly complete protonation is observed for alanine, asparagine/aspartate and glutamine/glutamate. Glycine and methionine are also strongly protonated to about 60–70%. The remaining amino acids show only moderate (~50%, serine, threonine, phenylalanine, tyrosine, proline) to low (<40%, isoleucine, leucine, valine, arginine, lysine, histidine) protonation levels.

Such an analysis has not yet been carried out for expression based on algal extracts, where the starting deuteration level is 98% and higher overall incorporation of ~74% is achieved. The relatively low ^{2}H incorporation of about half of the amino acids in the yeast extract and the subsequently lower incorporation into the target protein is the result of the protonated glucose used for yeast growth. As indicated before, yeast can be grown on 100% ^{2}H$_2$O and deuterated glucose, but this caused a severe reduction in insect cell growth. It remains to be tested whether a suitable further fractionation of the fully deuterated yeast extract can yield higher deuteration levels. Furthermore, the citric acid cycle clearly causes additional loss of deuteration in insect cells growing in media containing protonated glucose. Suitable inhibitors may reduce this loss to some extent. Current efforts in our groups are directed to clarify these questions.

Costs

Uniform labeling can be achieved using commercial or homemade media for insect cells. Commercial media (BioExpress, Cambridge Isotope Laboratories) remain prohibitively expensive and so far, no deuterated media are available. The commercial labeling media have a very similar composition as standard full media. This allows passaging of insect cells for several times, thereby potentially increasing the isotope incorporation. However, such an approach is presently not affordable. In contrast, the discussed homemade media are cost-optimized by using bulk extracts as the amino acid source instead of purified amino acids. These roughly 10-fold less expensive media are designed for protein expression but not for continuous culture. Therefore they only sustain one or two passages.

Homemade labeling media are based on commercial unlabeled dropout media depleted in individual amino acids or devoid of all amino acid sources and carbohydrates. Such custom media can be presently obtained for about 50–100 € per L, depending on the vendor. For amino acid-type specific backbone or side chain labeling, the total costs comprise this base medium and the labeled target amino acid. In many cases, the costs for the labeled amino acid can be moderate, e.g. ~200 € per L for ^{15}N-valine, ^{15}N leucine, ^{15}N, phenylalanine or [*methyl*-^{13}C]-methionine, of which ~60–200 mg/L are supplemented. The total costs of homemade media for uniform labeling are dominated by the algal or yeast extracts (Table 2). While the prices for individual components are subject to change, a few general observations can be made.

Obviously, ^{15}N labeling is the cheapest since inexpensive ^{15}NH$_3$/^{15}NH$_4$Cl can be used as the nitrogen source for algae/yeast. Glutamine is required for strong expression in insect cells and presents a source of isotope scrambling through its metabolism. Commercial prices for ^{15}N$_2$-labeled glutamine are currently about ~700 €/g, which presents a considerable cost factor if high ^{15}N labeling of all amino acids connected to the glutamine metabolism is desired. The cost for ^{15}N-labeled glutamine may be reduced to negligible levels (≤30 €/g) by enzymatic synthesis via ^{15}N-glutamate, which can be produced by glutamate dehydrogenase from ^{15}N-labeled ammonium and 2-oxoglutaric acid[35] and subsequently be converted to ^{15}N-glutamine

by glutamine synthetase under the addition of ^{15}N-labeled ammonium[36]. With this procedure, ^{15}N labeling with homemade yeast extract and ^{15}N-glutamine supplementation is considerably cheaper (~100 €/L) than with commercial algal extract (~500 €/L, Table 2).

As mentioned, insect cells can synthesize glutamine from glutamate and added ammonium. This pathway was exploited by Meola et al.[13] who supplemented the amino acid depleted medium with 10 g/L of commercial labeled yeast extract and 5 mM [15]NH_4Cl or 2 mM glutamine, respectively. No difference was reported for protein expression in stably transfected S2 cells. Sitarska et al.[15] confirmed that supplementation of 5 mM [15]NH_4Cl instead of glutamine together with 10 g/L of algal extract did not affect cell growth and viability in Sf cell lines. However, NH_4Cl had an inhibitory effect on virus infection. For optimal expression, insect cells were therefore infected in full medium and transferred only 16 h after infection into labeling medium containing NH_4Cl. In both cases, expression yields were below 25 mg/L for all tested proteins. Using 1 g/L of [15]N_2-glutamine as a supplement to yeast extract, Opitz et al.[16] reported expression yields ≥40 mg/L for several soluble proteins, but the yield dropped below 10% when substituting up to 10 mM NH_4Cl for glutamine. In the latter case, NH_4Cl had been added together with virus and likely inhibited virus entry. Recent ongoing experiments, however, confirm that even when supplying NH_4Cl only 16 h after infection, expression yields are more than two-fold lower compared to the glutamine supplementation.

Compared to ^{15}N-labeling, the additional costs for producing partially deuterated ^{15}N-labeled proteins are relatively small (~500 €/L) for both algal and yeast extracts. Since the additional costs for the yeast extract arise solely from D_2O, they may be reduced substantially by recycling D_2O via distillation. These cost estimates for deuterated yeast extracts are based on the use of protonated glucose, resulting in the mentioned lower overall deuteration of the yeast amino acids and leading to final deuteration levels in insect cell-expressed proteins of ~60% (Figure 4c,d). In contrast, commercial algal extracts have 98% enrichment for all isotopes, including [2]H. However, a significant part of this high deuteration is lost due to the insect cell metabolism, leading to final deuteration levels in insect cell-expressed proteins of ≤75%, making its impact more limited, unless the unwanted dilution can be suppressed by further improvements of labeling protocols.

Commercial ^{13}C-labeled algal extracts are by-products of ^{13}C-glucose production from algae grown on [13]CO_2 as the sole carbon source. Thus, they are intrinsically cheaper as a source of ^{13}C-labeled amino acids than yeast grown on ^{13}C-glucose. This is reflected in the costs for ^{13}C,[15]N-labeled growth media derived from the two extracts (Table 2), which amount to ~1800 €/L and ~3200 €/L for algal and yeast extracts, respectively. The additional costs for producing partially deuterated ^{13}C,[15]N-labeled proteins are again only small for the homemade yeast extract (~450 €/L), but considerably larger for the commercial algal extract (~1400 €/L).

Further improvements may reduce the discussed costs. E.g. roughly double the concentration of amino acids was used for the algal extract media as compared to the yeast extract media. Thus it may be possible to reduce the added amount of either the algal extracts alone or in combination with yeast extract. Furthermore, ^{13}C-glucose dominates the cost of ^{13}C-labeled yeast extracts, which may be reduced to some extent via alternative carbon sources. Finally, bacterial production of glutamine[37,38] may make this important supplement available in all desired isotope-labeled forms at very low cost.

Conclusion

Eukaryotic expression systems in combination with simple, inexpensive labeling protocols have a great potential to make biologically interesting, but difficult-to-express proteins accessible for NMR analysis. They are relatively easy to apply, since insect cell expression platforms are often already present in neighboring X-ray crystallography or electron microscopy laboratories, which may readily incorporate the described protocols for integrated structural biology projects with NMR. The current labeling procedures may be further improved and enhanced. In particular, it seems possible to (i) further increase deuteration levels via a proper understanding of the isotope dilution and scrambling, (ii) expand the described protocols to mammalian cell cultures, given the similarity between the expression systems. It is also hoped that inexpensive, commercial media for uniform labeling based on algal and yeast extracts will become available in the near future, which may further reduce the effort of isotope labeling in eukaryotic expression hosts.

Acknowledgments

Figure Legends

Figure 1: Timeline of the number PDB entries expressed in eukaryotic hosts. The number of these entries structures shows a strong increase over the last 25 years. Half of the structures derive from insect cell expression, with *Spodoptera frugiperda* (Sf9, Sf21) being the most used organism.

Figure 2: Overview of isotope-labeled protein expression in insect cells based on isotope-labeled commercial algal (top) or homemade yeast (bottom) extracts. The principal nitrogen, carbon, and hydrogen sources for the amino acids and their pathways are indicated.

Figure 3: NMR analysis of proteins expressed in isotope-labeled form in insect cells. (a) Comparison of $^1H^N$-^{15}N residual dipolar couplings (RDCs) obtained experimentally (red) or derived from crystallographic structures (black) of Abelson kinase domain expressed in Sf9 cells and selectively labeled by ^{15}N-FGMY amino acids. Top: complex with imatinib inhibitor, bottom complex with dasatinib inhibitor. The good agreement between experimental solution NMR RDCs and their prediction from the crystal structures indicates that the activation loop adopts a similar conformation in solution as in the crystal for both imatinib (inactive conformation) and dasatinib (active conformation). Data replotted from Vajpai et al.[3] **(b)** ^1H-^{15}N-TROSY spectra of detergent-solubilized β_1-adrenergic receptor expressed in Sf9 cells and selectively labeled by ^{15}N-valine. Differences in the spectra of the complexes with the agonist isoprenaline (red) and the antagonist atenolol (blue) indicate allosteric signal transmission from the ligand binding site (V172) to the G protein binding site (V226). Data replotted from Isogai et al.[7] **(c)** ^1H,^{15}N-TROSY spectrum (600 MHz, 308 K) of ^2H,^{15}N-labeled CCR5 solubilized in n-dodecyl-β-D-maltopyranoside. ^2H,^{15}N-labeled CCR5 was expressed in Sf9 cells grown on ^2H,^{15}N-yeastolate and $^{14}N_2$-glutamine. **(d)** Determination of average ^{15}N T_1 and T_2 relaxation times of ^2H,^{15}N-labeled CCR5 (900 MHz, other conditions as panel **(c)**). The relaxation times are determined from the intensity ratios of one-dimensional, ^1H projections of standard ^1H-^{15}N-edited ^{15}N relaxation experiments carried out at two different relaxation delays Δ.

Figure 4: LC/MS analysis of isotope content of individual amino acids in yeast extracts and proteins. (a) HPLC separation of amino acids obtained by PITC derivatization of a ^2H,^{15}N-labeled yeast extract. Peaks are annotated by the corresponding amino acid. The derivatized amino acids were separated by HPLC (Agilent) using a Phenomenex Kinetex XB-C18 column (75 mm x 4.6 mm, 2.6 μm, 100 Å) with a buffer gradient from 0.05% formic acid (FA)/water to 0.05% FA/100% acetonitrile at 50°C and detected by UV absorbance at 254 nm. **(b)** LC/MS spectrum of histidine obtained by acid hydrolysis of ^2H,^{15}N-GFP. The different $^2H_{n1}^1H_{n2}^{15}N_{n3}^{14}N_{n4}$ isotopologues are indicated. The protein hydrolysate was derivatized by PITC and then separated on an EASY-nLC™ nano-HPLC system

(Thermofisher) coupled to an Orbitrap Elite™ (Thermofisher) mass spectrometer. HPLC separation was performed at room temperature using a silica (ReproSil-Pur 120 C18-AQ) self-packed column (150 mm x 75 μm, 1.9 μm, 120 Å) with a buffer gradient from 0.05% FA/water to 0.05% FA/80% acetonitrile. **(c)** ^{2}H,^{15}N incorporation of individual amino acids of autolysed ^{2}H,^{15}N-labeled yeast obtained by the LC/MS protocol described in **(b)**. The yeast had been grown on ^{15}N ammonium, D$_2$O, and protonated glucose. Due to the mild autolysis conditions, all amino remain intact, but cysteine is lost during PITC derivatization. **(d)** ^{2}H,^{15}N incorporation of individual amino acids obtained from acid-hydrolyzed GFP expressed in Sf9 cells using ^{2}H,^{15}N-labeled yeast extract supplemented by ^{15}N$_2$-glutamine. The same LC/MS protocol as in **(b)** was used. During the acid hydrolysis of the protein, asparagine and glutamine are converted to aspartate and glutamate, respectively, and tryptophan is degraded to undetected compounds.

Tables

Table 1: Protein structures obtained from samples produced in different expression hosts queried from the Protein Data Bank (PDB) (12/2017). Structures have been filtered for size (>100 kDa), human origin as well as G protein-coupled receptors (GPCRs).

	All structures	**>100 kDa**	**Human**	**GPCRs**
Total	135359 (100.0[a])	29791 (100.0)	37859 (100.0)	220 (100.0)
Bacterial[b]	99765 (73.7)	21639 (72.6)	27169 (71.8)	18 (8.2)
Yeast[c]	2353 (1.7)	582 (2.0)	1052 (2.8)	3 (1.4)
Insect[d]	5811 (4.3)	2174 (7.3)	3945 (10.4)	155 (70.5)
Mammalian[e]	3093 (2.3)	1190 (4.0)	2062 (5.4)	19 (8.6)

[a] in percent
[b] *E. coli*
[c] *P. pastoris*, *S. pombe* and *S. cerevisiae*
[d] *S. frugiperda*, *T. ni* and *D. melanogaster*
[e] *H. sapiens*, *M. musculus* and *C. griseus*

Table 2: Costs (in €) for uniform isotope labeling in insect cells per liter of growth medium.

		Amount/L	^{15}N	^{2}H,^{15}N	^{13}C,^{15}N	^{2}H,^{13}C,^{15}N
Algal extract	Commercial AE[a]	10 g	465	925	1150	2575
	^{15}NH$_4$Cl[b]	250 mg	4	4	4	4
	Tryptophan[c]	20 mg	42	100	250	250
	^{13}C$_6$-Glucose[d,e]	5 g			400	400
	Total AE		511	1029	1804	3229
Yeast extract	^{13}C$_6$-Glucose for YE[e,f]	33.3 g			2667	2667
	^{15}NH$_4$Cl for YE[b]	5.2 g	73	73	73	73
	D$_2$O for YE[g]	1 L		448		448
	^{15}N$_2$-Glutamine[h]	1 g	30	30	30	30
	^{13}C$_6$-Glucose[d,e]	5 g			400	400
	Total YE		103	551	3170	3618

[a] Prices for ISOGRO algal extract (AE, Sigma-Aldrich) with the respective isotope labeling pattern at 98% incorporation. A 50% quantity discount on list prices is used, which was obtained when ordering material for 5 L of insect cell medium in 2017.

[b] Assuming a price of 14 €/g

[c] Current catalogue prices

[d] Labeled glucose is added for obtaining higher incorporation of ^{13}C for amino acids linked to carbohydrate metabolism (see text).

[e] Assuming a price of 80 €/g

[f] Cost of isotope labeled substrates for producing the amount of yeast extract (YE, 8 g) required for 1 L of insect cell medium. No costs for work or equipment are included. For reasons described in the text non-deuterated glucose is used for ^{2}H labeling. Using deuterated glucose would result in considerably higher costs for ^{2}H,^{15}N and ^{2}H,^{13}C,^{15}N labeling.

[g] Assuming a price of 370 €/L. Costs can be reduced by recycling D$_2$O.

[h] Costs of homemade glutamine are assumed as 30 €/g, using the enzymatic reactions described in the main text. For all labeling patterns, only ^{15}N$_2$-labeled glutamine was used. Commercial prices are considerably higher, e.g. ~700 €/g for ^{15}N$_2$-labeled glutamine.

References

1.   Strauss, A., Bitsch, F., Cutting, B., Fendrich, G., Graff, P., Liebetanz, J., Zurini, M. & Jahnke, W. Amino-acid-type selective isotope labeling of proteins expressed in Baculovirus-infected insect cells useful for NMR studies. *Journal of Biomolecular NMR* **26**, 367-372 (2003).

2.   Strauss, A., Bitsch, F., Fendrich, G., Graff, P., Knecht, R., Meyhack, B. & Jahnke, W. Efficient uniform isotope labeling of Abl kinase expressed in Baculovirus-infected insect cells. *Journal of Biomolecular NMR* **31**, 343-349 (2005).

3.   Vajpai, N., Strauss, A., Fendrich, G., Cowan-Jacob, S.W., Manley, P.W., Grzesiek, S. & Jahnke, W. Solution conformations and dynamics of ABL kinase-inhibitor complexes determined by NMR substantiate the different binding modes of imatinib/nilotinib and dasatinib. *Journal of Biological Chemistry* **283**, 18292-18302 (2008).

4.   Kofuku, Y., Ueda, T., Okude, J., Shiraishi, Y., Kondo, K., Maeda, M., Tsujishita, H. & Shimada, I. Efficacy of the β2-adrenergic receptor is determined by conformational equilibrium in the transmembrane region. *Nature Communications* **3**, 1045 (2012).

5.   Nygaard, R., Zou, Y., Dror, R.O., Mildorf, T.J., Arlow, D.H., Manglik, A., Pan, A.C., Liu, C.W., Fung, J.J., Bokoch, M.P., Thian, F.S., Kobilka, T.S., Shaw, D.E., Mueller, L., Prosser, R.S. & Kobilka, B.K. The dynamic process of β2-adrenergic receptor activation. *Cell* **152**, 532-542 (2013).

6.   Kofuku, Y., Ueda, T., Okude, J., Shiraishi, Y., Kondo, K., Mizumura, T., Suzuki, S. & Shimada, I. Functional dynamics of deuterated β2-adrenergic receptor in lipid bilayers revealed by NMR spectroscopy. *Angewandte Chemie International Edition in English* **53**, 13376-13379 (2014).

7.   Isogai, S., Deupi, X., Opitz, C., Heydenreich, F.M., Tsai, C.-J., Brueckner, F., Schertler, G.F.X., Veprintsev, D.B. & Grzesiek, S. Backbone NMR reveals allosteric signal transduction networks in the β1-adrenergic receptor. *Nature* **530**, 237-241 (2016).

8.   Brüggert, M., Rehm, T., Shanker, S., Georgescu, J. & Holak, T.A. A novel medium for expression of proteins selectively labeled with 15N-amino acids in *Spodoptera frugiperda* (Sf9) insect cells. *Journal of Biomolecular NMR* **25**, 335-348 (2003).

9.   Egorova-Zachernyuk, T.A., Bosman, G.J.C.G.M., Pistorius, A.M.A. & DeGrip, W.J. Production of yeastolates for uniform stable isotope labelling in eukaryotic cell culture. *Applied Microbiology and Biotechnology* **84**, 575-581 (2009).

10.   Egorova-Zachernyuk, T.A., Bosman, G.J.C.G.M., DeGrip, W.J. & Shvets, V.I. Stable isotope labelling of human histamine receptor H1R: prospects for structure-based drug design. *Doklady Biochemistry and Biophysics* **433**, 164-167 (2010).

11.   Gossert, A.D., Hinniger, A., Gutmann, S., Jahnke, W., Strauss, A. & Fernández, C. A simple protocol for amino acid type selective isotope labeling in insect cells with improved yields and high reproducibility. *Journal of Biomolecular NMR* **51**, 449-456 (2011).

12.   Saxena, K., Dutta, A., Klein-Seetharaman, J. & Schwalbe, H. Isotope labeling in insect cells. *Methods in Molecular Biology* **831**, 37-54 (2012).

13.   Meola, A., Deville, C., Jeffers, S.A., Guardado-Calvo, P., Vasiliauskaite, I., Sizun, C., Girard-Blanc, C., Malosse, C., van Heijenoort, C., Chamot-Rooke, J., Krey, T., Guittet, E., Pêtres, S., Rey, F.A. & Bontems, F. Robust and low cost uniform 15 N-labeling of proteins expressed in *Drosophila* S2 cells and *Spodoptera frugiperda* Sf9 cells for NMR applications. *Journal of Structural Biology* **188**, 71-78 (2014).

14.   Skora, L., Shrestha, B. & Gossert, A.D. Chapter Eleven - Isotope Labeling of Proteins

in Insect Cells. in *Methods in Enzymology*, Vol. 565 (ed. Kelman, Z.) 245-288 (Academic Press, 2015).

15. Sitarska, A., Skora, L., Klopp, J., Roest, S., Fernández, C., Shrestha, B. & Gossert, A.D. Affordable uniform isotope labeling with 2H, 13C and 15N in insect cells. *Journal of Biomolecular NMR* **62**, 191-197 (2015).

16. Opitz, C., Isogai, S. & Grzesiek, S. An economic approach to efficient isotope labeling in insect cells using homemade 15N-, 13C- and 2H-labeled yeast extracts. *Journal of Biomolecular NMR* **62**, 373-385 (2015).

17. Zhang, Y., Wei, H., Xie, D., Calambur, D., Douglas, A., Gao, M., Marsilio, F., Metzler, W.J., Szapiel, N., Zhang, P., Witmer, M.R., Mueller, L. & Hedin, D. An improved protocol for amino acid type-selective isotope labeling in insect cells. *Journal of Biomolecular NMR* **68**, 237-247 (2017).

18. Smith, G.E., Summers, M.D. & Fraser, M.J. Production of human beta interferon in insect cells infected with a baculovirus expression vector. *Molecular and Cellular Biology* **3**, 2156-2165 (1983).

19. Kelly, D.C. Baculovirus Replication. *Journal of General Virology* **63**, 1-13 (1982).

20. Berger, I., Fitzgerald, D.J. & Richmond, T.J. Baculovirus expression system for heterologous multiprotein complexes. *Nature Biotechnology* **22**, 1583-1587 (2004).

21. Luckow, V.A., Lee, S.C., Barry, G.F. & Olins, P.O. Efficient generation of infectious recombinant baculoviruses by site-specific transposon-mediated insertion of foreign genes into a baculovirus genome propagated in *Escherichia coli*. *Journal of Virology* **67**, 4566-4579 (1993).

22. Possee, R.D., Hitchman, R.B., Richards, K.S., Mann, S.G., Siaterli, E., Nixon, C.P., Irving, H., Assenberg, R., Alderton, D., Owens, R.J. & King, L.A. Generation of baculovirus vectors for the high-throughput production of proteins in insect cells. *Biotechnology and Bioengineering* **101**, 1115-1122 (2008).

23. Wasilko, D.J., Edward Lee, S., Stutzman-Engwall, K.J., Reitz, B.A., Emmons, T.L., Mathis, K.J., Bienkowski, M.J., Tomasselli, A.G. & Fischer, D.H. The titerless infected-cells preservation and scale-up (TIPS) method for large-scale production of NO-sensitive human soluble guanylate cyclase (sGC) from insect cells infected with recombinant baculovirus. *Protein Expression and Purification* **65**, 122-132 (2009).

24. Vaughn, J.L., Goodwin, R.H., Tompkins, G.J. & McCawley, P. The establishment of two cell lines from the insect *Spodoptera frugiperda* (Lepidoptera; Noctuidae). *In Vitro* **13**, 213-217 (1977).

25. Hink, W.F. Established insect cell line from the cabbage looper, *Trichoplusia ni*. *Nature* **226**, 466-467 (1970).

26. O'Reilly, D.R., Miller, L.K. & Luckow, V.A. *Baculovirus expression vectors: a laboratory manual*, (Oxford University Press, 1994).

27. Huang, J.-R., Craggs, T.D., Christodoulou, J. & Jackson, S.E. Stable intermediate states and high energy barriers in the unfolding of GFP. *Journal of Molecular Biology* **370**, 356-371 (2007).

28. Vajpai, N., Strauss, A., Fendrich, G., Cowan-Jacob, S.W., Manley, P.W., Jahnke, W. & Grzesiek, S. Backbone NMR resonance assignment of the Abelson kinase domain in complex with imatinib. *Biomolecular NMR Assignments* **2**, 41-42 (2008).

29. Sprangers, R., Velyvis, A. & Kay, L. Solution NMR of supramolecular complexes: providing new insights into function. *Nature Methods* **4**, 697-703 (2007).

30. Solt, A.S., Bostock, M.J., Shrestha, B., Kumar, P., Warne, T., Tate, C.G. & Nietlispach, D. Insight into partial agonism by observing multiple equilibria for ligand-bound and Gs-mimetic nanobody-bound β1-adrenergic receptor. *Nature Communications* **8**, 1795 (2017).

31. Okude, J., Ueda, T., Kofuku, Y., Sato, M., Nobuyama, N., Kondo, K., Shiraishi, Y., Mizumura, T., Onishi, K., Natsume, M., Maeda, M., Tsujishita, H., Kuranaga, T., Inoue, M. & Shimada, I. Identification of a Conformational Equilibrium That Determines the Efficacy and Functional Selectivity of the μ-Opioid Receptor. *Angewandte Chemie International Edition in English* **54**, 15771-15776 (2015).

32. Ugwu, C.U., Aoyagi, H. & Uchiyama, H. Photobioreactors for mass cultivation of algae. *Bioresource Technology* **99**, 4021-4028 (2008).

33. Rosen, M.K., Gardner, K.H., Willis, R.C., Parris, W.E., Pawson, T. & Kay, L.E. Selective methyl group protonation of perdeuterated proteins. *Journal of Molecular Biology* **263**, 627-636 (1996).

34. Hochuli, M., Szyperski, T. & Wuthrich, K. Deuterium isotope effects on the central carbon metabolism of *Escherichia coli* cells grown on a D2O-containing minimal medium. *Journal of Biomolecular NMR* **17**, 33-42 (2000).

35. Kragl, U., Gödde, A., Wandrey, C., Kinzy, W., Cappon, J. & Lugtenburg, J. Repetitive batch as an efficient method for preparative-scale enzymatic-synthesis of 5-azido-neuraminic acid and N15-L-glutamic acid. *Tetrahedron: Asymmetry* **4**, 1193-1202 (1993).

36. Hansen, A.P., Petros, A.M., Mazar, A.P., Pederson, T.M., Rueter, A. & Fesik, S.W. A practical method for uniform isotopic labeling of recombinant proteins in mammalian cells. *Biochemistry* **31**, 12713-12718 (1992).

37. Kusumoto, I. Industrial production of L-glutamine. *Journal of Nutrition* **131**, 2552S-5S (2001).

38. Li, J., Ma, C., Ma, Y., Li, Y., Zhou, W. & Xu, P. Medium optimization by combination of response surface methodology and desirability function: an application in glutamine production. *Applied Microbiology and Biotechnology* **74**, 563-571 (2007).
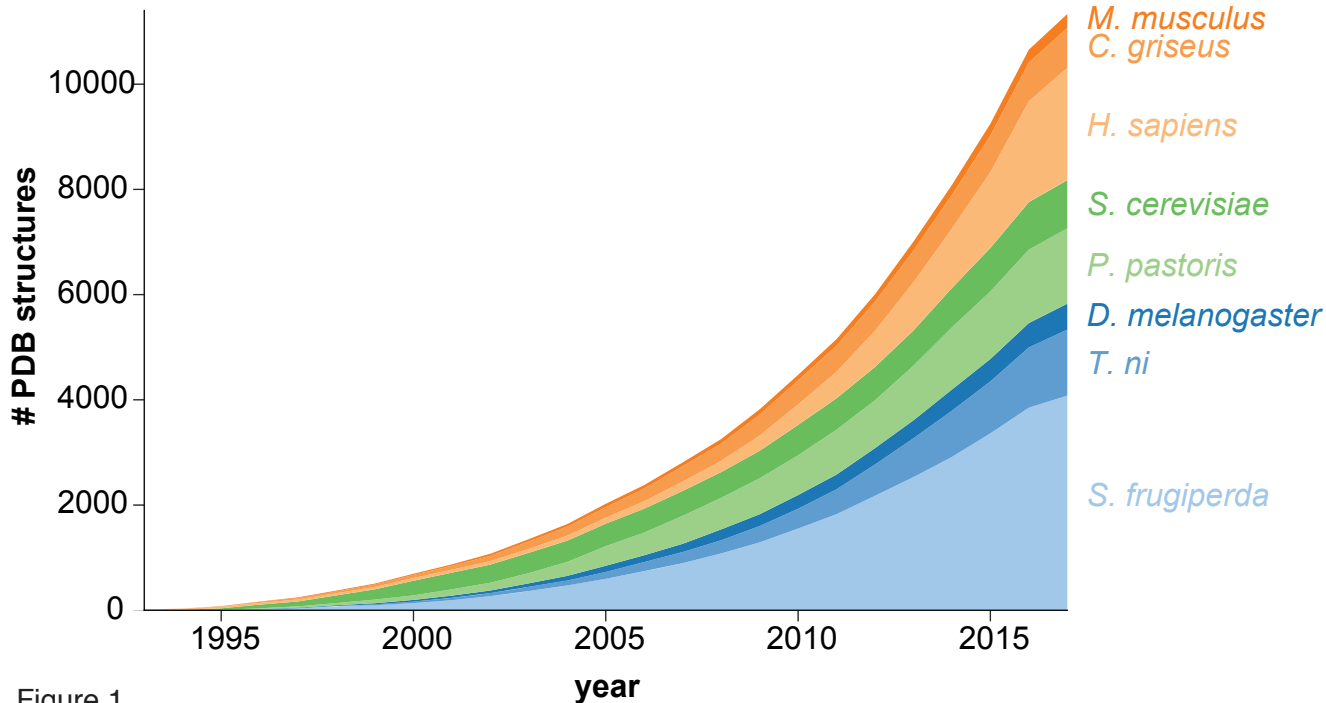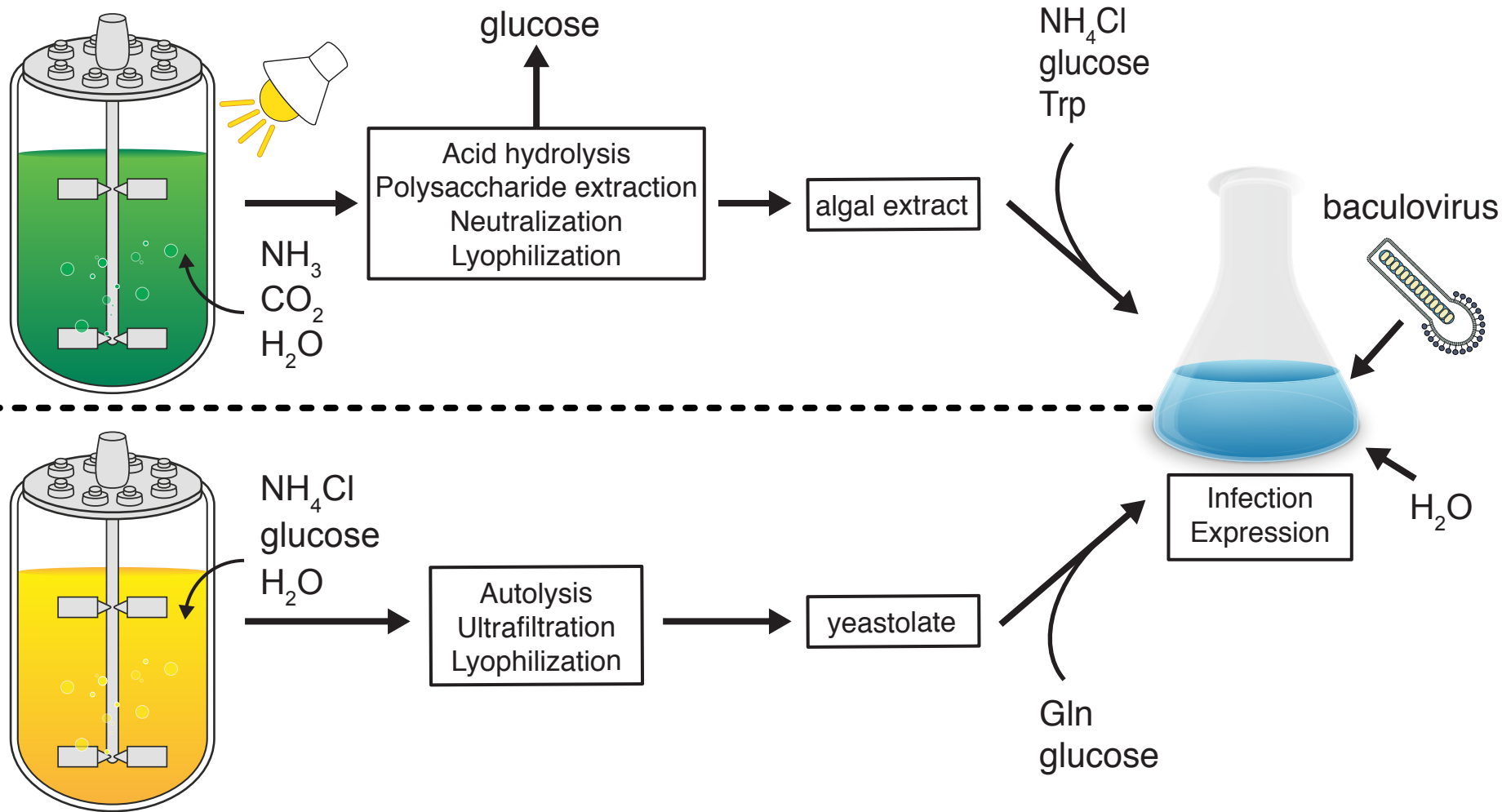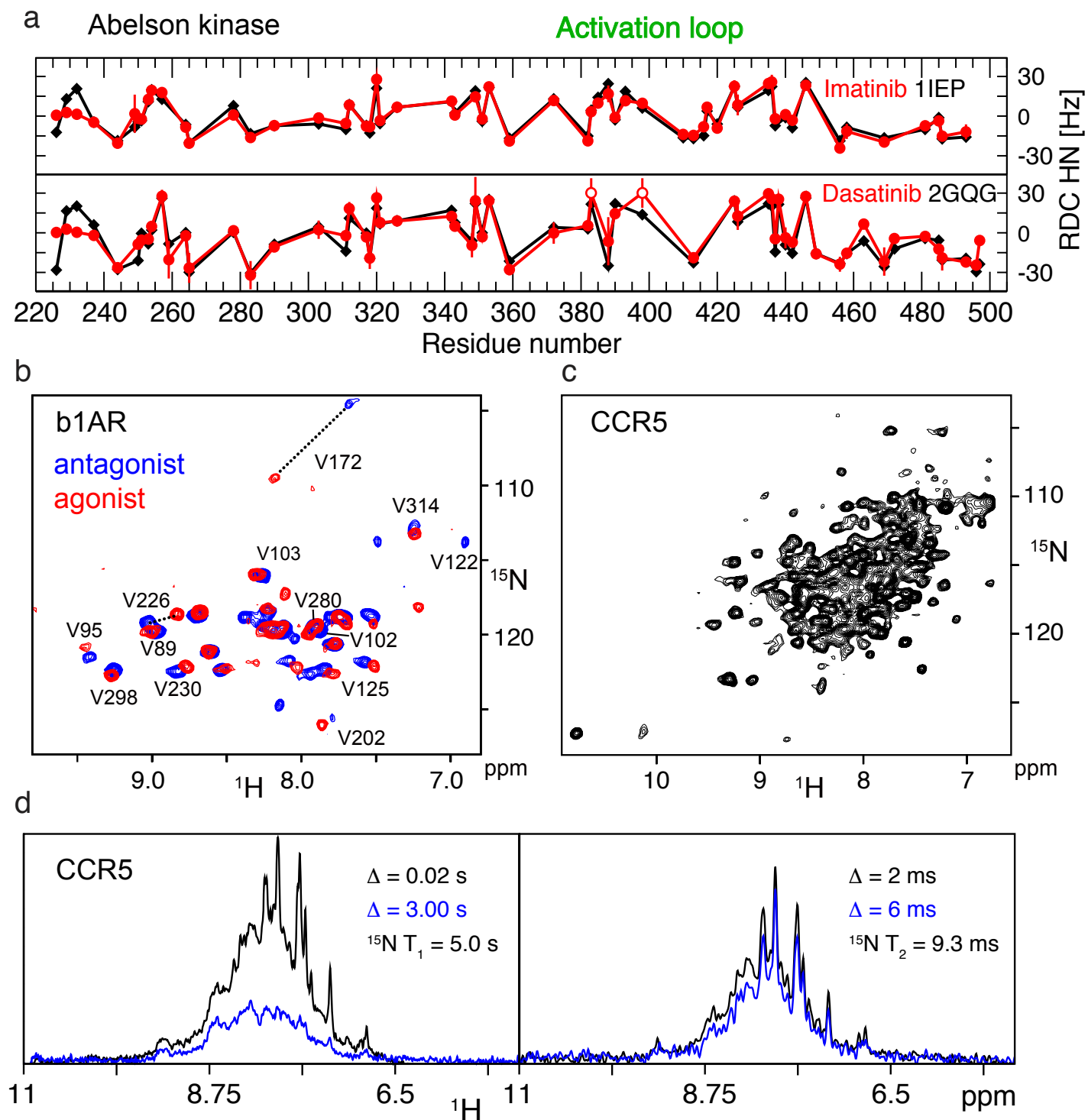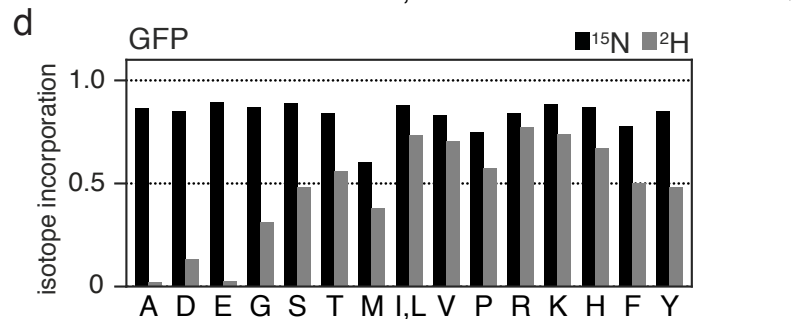
Figure 1

glucose

NH$_4$Cl
glucose
Trp

Acid hydrolysis
Polysaccharide extraction
Neutralization
Lyophilization

algal extract

baculovirus

NH$_3$
CO$_2$
H$_2$O

NH$_4$Cl
glucose
H$_2$O

Autolysis
Ultrafiltration
Lyophilization

yeastolate

Infection
Expression

H$_2$O

Gln
glucose

Figure 2

a Abelson kinase   Activation loop

Imatinib 1IEP

Dasatinib 2GQG

RDC HN [Hz]

Residue number

b b1AR
antagonist
agonist

V172
V314
V103
V122
V226
V280
V95
V89
V102
V298
V230
V125
V202

c CCR5

15N

1H   ppm

d CCR5

Δ = 0.02 s
Δ = 3.00 s
15N T₁ = 5.0 s

Δ = 2 ms
Δ = 6 ms
15N T₂ = 9.3 ms

1H   ppm

Figure 3

Figure 4