Chapter 10
Isotope Labeling in Insect Cells

Alvar D. Gossert and Wolfgang Jahnke

Abstract Target proteins in contemporary NMR are becoming increasingly complex. The interest lies on membrane proteins, multi-domain proteins, secreted proteins with secondary modifications etc., which can often only be expressed in eukaryotic expression hosts. Although *E. coli* is the organism of choice for expression of proteins in isotope labeled form for NMR studies, bacterial cells have limitations concerning their ability of producing soluble and well-folded proteins of human origin. Insect cells are eukaryotic cells and therefore evolutionary closer to human cells than bacterial cells. Therefore a larger share of human proteins can be functionally expressed in them, for example multi-domain proteins, protein complexes, membrane proteins and proteins requiring post-translational modifications. In order to study these proteins by NMR, they are ideally prepared in isotope labeled form. In this chapter different strategies for isotope labeling in insect cells are described – uniform and amino acid specific. A general introduction to expression with baculovirus infected insect cells is given followed by a detailed descriptions of labeling approaches. The chapter is concluded with case studies, describing successful application of isotope labeling in insect cells for NMR studies including solid-state experiments, ligand binding studies and protein dynamics.

10.1 Introduction

Target proteins in contemporary NMR are becoming increasingly complex. The interest lies on membrane proteins, multi-domain proteins, secreted proteins with secondary modifications etc., which can often only be expressed in eukaryotic expression hosts. Although *E. coli* is the organism of choice for expression of proteins in isotope labeled form for NMR studies, bacterial cells have limitations concerning their ability of producing soluble and well-folded proteins of human origin. Insect cells are eukaryotic cells and therefore evolutionary closer to human cells than *i.e.* bacterial cells. Therefore a larger share of human proteins can be functionally expressed in them, especially multi-domain proteins and protein complexes are correctly folded and active. Furthermore, post-translational modifications are introduced into the polypeptides, like phosphorylation. Very similar secretion pathways as in human cells enable expression of membrane proteins and secretory proteins with correct disulfide formation and glycosylation patterns. These post-translational modifications are often essential

A.D. Gossert (✉) • W. Jahnke
Novartis Institutes for BioMedical Research, Novartis AG,
4002 Basel, Switzerland
e-mail: alvar.gossert@novartis.com

H.S. Atreya (ed.), *Isotope labeling in Biomolecular NMR*, Advances in Experimental Medicine and Biology 992,
DOI 10.1007/978-94-007-4954-2_10,

for protein folding, stability and activity. Another advantage of insect cells is that small molecules can pass their delicate plasma membranes. Proteins can therefore be expressed in presence of stabilizing compounds or co-factors. All-in-all, a larger fraction of the human proteome is accessible for expression in insect cells compared to bacteria.

With more complex cells, however, come more complex protocols for heterologuous protein expression. This can be appreciated in a comparison of the general workflow with that of bacterial expression. Culturing of insect cells in shake flasks is practically as simple as growing bacteria, just with the added complication that all manipulations need to be carried out in a sterile environment. However, most other aspects of the expression workflow require higher efforts and more dedication. Bacteria have (i) very fast growth rates and high protein synthesis capacities, (ii) foreign genes can be introduced in matter of minutes and – important for isotope labeling – (iii) they are able to synthesize all molecules important for life from basic building blocks. On the insect cell side, this looks quite different: (i) The slow growth rate of insect cells leads to expression times of typically 2–3 days, up to 7 days, and protein yields can be expected to be in the range of a few mg/l. Continuous maintenance of cell stocks and lengthy building up of expression sized pre-cultures is also laborious and time consuming. (ii) For high levels of expression, the gene of interest needs to be delivered in an infectious particle, the baculovirus. Therefore, a second line of work is required for producing highly infectious recombinant baculovirus, which takes weeks. (iii) Growth media are more complex – and more expensive. They easily contain more than 60 ingredients, as nearly all nutrients like amino acids and vitamins need to be given to insect cells on a silver plate, due to their poor metabolic capacities.

However, despite these disadvantages, insect cells represent a very powerful tool for expression of difficult proteins for structural studies, making accessible a larger fraction of the proteome for NMR studies.

In order to give an overview of the applicability of insect cells for NMR studies, the further text is divided into three sections: A description of the general culturing process (Sect. 10.2), the adaptations needed for isotope labeling (Sect. 10.3) and a number of case studies of successful application of isotope labeling in insect cells (Sect. 10.4).

10.2 Insect Cells as Expression System

In this section a brief overview over the general procedures is given, which are necessary for successful over-expression of proteins in insect cells. Mainly the points are touched, which are important for understanding the special considerations required for isotope labeling. These are described in the next section. For actual comprehensive protocols the reader is referred to the standard works [10–12], where detailed references can be found.

10.2.1 Insect Cells

In the late 1980s to early 1990s, insect cells emerged as an alternative eukaryotic expression system for recombinant proteins. During the last decade insect cells have established their firm place as an expression platform for structural biology. The main insect cell lines for the purpose of expression are derived from two organisms. In 1977, ovary cells of the fall army worm, *spodoptera frugiperda,* were isolated and cultured resulting in the cell lines sf21 and sf9, where sf9 is a monoclonal derivative of sf21 [17]. In 1983 the Hi5 line was established from cells of the cabbage looper, *Trichopulsia ni* [6]. In conjunction with a virus vehicle, the baculovirus, these became the most popular eukaryotic cell lines for protein over-expression for structural biology.

10.2.2 The Baculovirus Life-Cycle and Its Use for Heterologuous Protein Expression

Baculoviruses are a family of arthropode specific viruses. They are rod shaped particles with a genome of 80–200 kbp of double stranded, circular DNA, the so-called bacmid. A prototypical member of this family is the *Autographa californica* multiple nuclear polyhedrosis virus (AcNPV), which is used as the main vehicle for heterologuous protein production in insect cells. The life-cycle of this virus starts with the primary infection of a host cell, where it is replicated. After 10–20 h, new virions emerge from the infected cell and infect other cells in the same organism, where the replication is re-initiated in a secondary infection phase, which ultimately leads to death of the host. The free virion form, however, is not a stable entity outside of the insect organism. Therefore, in the very late stage of infection, *i.e.* after more than 30 h, the virus induces production of copious amounts (up to 50% of total protein) of the proteins polyhedrin and p10 that form a cube-shaped protein matrix, in which virus particles are embedded. These are the so-called occlusion bodies. Occluded viruses are long-term stable outside of the host cells and are, eventually, ingested by another insect larva. In the alkaline conditions of the gut lumen, the polyhedrin matrix dissolves and releases virions that infect the gut cells. With this event, a new cycle in the baculoviral life is initiated.

For the purpose of over-expression of heterologuous proteins, the naturally highly expressed polyhedrin or p10 genes are replaced by the gene of interest. In culture conditions, the polyhedrin/p10 occlusion bodies are not vital for the viral life-cycle, as free virions are stable in cell culture media and conserve their infectivity. Therefore, baculoviruses carrying a heterologuous protein are as infective and proliferative as unaltered ones in cell-culture conditions and can be used to infect insect cells. The result is, that a foreign gene under the control of the strong polyhedrin or p10 promotor is expressed in high amounts by the insect cells 1–3 days post infection.

10.2.3 Lab Equipment

Insect cells and baculovirus can be cultured and handled in biosafety level 1 facilities. (Caveat: In Europe, Hi5 cells need to be handled in BL2 facilities, as they contain viral DNA potentially harmful to human beings). The basic large equipment needed for culturing of insect cells consists of a sterile bench for carrying out all manipulations, and incubators for culturing the cells in petri dishes and shake flasks of small and large volume. It is important to have enough autoclaving capacity for sterilizing the culture flask and consumables. Further essential equipment comprises an inverse light microscope for assessing cell parameters like viability, cell size and number. A waterbath for pre-warming media to 27°C is also important. Finally a centrifuge is needed for harvesting the cells. A state of the art facility for insect cell expression may have an automatic cell counter that fully replaces the microscope, and more sophisticated devices replacing shake flasks for large scale expression, like wave-bags or fermenters.

In general there are only small differences to a typical *E. coli* lab, except for the sterile bench and the microscope. However, it is strongly advisable not to share equipment with an *E. coli* expression lab, because of the high risk of cross-contamination.

10.2.4 Culturing of Insect Cells

Insect cells (*Spodoptera frugiperda, Trichopulsia ni*) are usually grown in erlenmeyer shake flasks at 27°C and gentle mixing at 70–90 rpm. Basic incubators can be used, as growth takes place in ambient atmosphere. Under these conditions doubling times are about 20–24 h in commercial media.

During culturing, cells are regularly monitored with light microscopy to assess their viability, proliferation and density. Viability is assessed with a dye that penetrates only the non-intact membranes of dead cells, coloring them blue. In this way the fraction of living and dead cells can be counted. The cell diameter, typically 14–18 μm, is a second parameter which allows evaluating viability, the smaller the better. Also, it allows characterizing the efficiency of viral infection, as it is manifested by massively enlarged cells with diameters of up to 24 μm. Proliferation can be monitored by growth rates but also by looking for cell doublets, which indicate cells in the process of doubling and are a sign of a well-growing culture. However, cell lumps of many cells indicate declining viability of the culture and call for measures to improve growth conditions.

The practical process to introduce and establish a new cell line in the lab is the following: Insect cells are obtained in frozen aliquots, *e.g.* from ATCC or Invitrogen, and need to be adapted very gently to growth in suspension in a chosen medium. To this end, the insect cell medium is supplemented with 10% fetal calf serum (FCS). Cells are grown in 50 ml erlenmeyer flasks and every 2–3 days, cells are taken up in fresh medium. At the beginning of an adaptation phase, 50% of the grown culture is mixed with 50% v/v of fresh medium. Progressively larger dilutions up to 10–90% are carried out and after 10–20 such passages a stable stock-culture is established. FCS can slowly be removed over a few passages, and a stock adapted to "serum free" medium is obtained. During this adaptation phase, the cells need to be carefully monitored as described above, and at first signs of cell lumps or reduced viability, adaptation should be slowed down or even reverted one or two steps. Once a stable stock is obtained, it is maintained by regular passaging into fresh medium twice a week. This can be carried on for up to 200 passages, however, generally after 50 passages a backup stock is started.

Most labs keep different stocks, as expression levels can vary largely (up to three fold) depending on combinations of cell strains and media. Adapting strains to different media requires gradual progressive change of the medium, often with the need of re-introducing FCS as a supplement.

For a large scale expression experiment, an expression culture is built up starting from a well-growing individual stock of a density of 4–6 × 10^6 cells/ml. Over the course of days, cells are grown in large volume erlenmeyer flasks, in order to build up the required number cells. Typically, the target density of an expression culture is 1.5 × 10^6 cells/ml. Alternatives to large erlenmeyer flasks are wave-bag devices or fermenters. For expression of a target protein, the culture is infected with baculovirus carrying the gene of interest at a given ratio of virus to cell, termed multiplicity of infection (MOI, typically 1–10). Roughly 20 h after infection, production of the target protein can be detected. Depending on the protein, expression is continued for 24–72 h, or even longer for secreted proteins. Harvesting takes place by gentle centrifugation at 400 g, as insect cells tend to rupture at more than 1,000 g.

10.2.5 Generation of Virus

Introduction of the gene of interest into insect cells is a scientific discipline on itself. There are several ways of delivering DNA to be transcribed to the cells: (1) transient transfection with pure DNA, (2) generation of stable cell lines carrying the gene of interest integrated in the genome and (3) introducing the gene into the host cells by means of a virus, the baculovirus. For structural biology, where high yields are needed (fulfilled by 2 and 3) and a certain flexibility for testing several constructs is often required (fulfilled by 1 and 3), strategy number 3 is the most promising one, and in fact, most often used. Therefore, the rest of the discussion will focus on the baculovirus system, which is used in conjunction with sf9/sf21 and Hi5 cells.

As described in Sect. 10.2.2, the strategy for over-expression of a given target protein consists in inserting the gene of interest in place of the polyhedrin or p10 gene in the baculovirus genome. The molecular biology required for gene insertion into a bacmid is not trivial, as the bacmid is roughly

100 kbp in size, compared to 4–6 kbp for a plasmid used in *E. coli* expression. A plethora of ever refined strategies have been put in place – and have been marketed – for creating recombinant virus. Generally, in a first step the gene of interest is inserted into a bacterial plasmid by conventional techniques. This plasmid contains elements for selection and propagation in *E. coli* and, most importantly, the gene of interest is flanked by sequences that are homologuous to the bacmid sequences flanking the polyhedrin gene. Subsequently the gene on the plasmid needs to be inserted into the bacmid in a sequence specific manner. Homologuous recombination is a process, in which pieces of DNA flanked with identical sequences are exchanged by DNA-repair enzymes, leading in this case to introduction of recombinant DNA into a bacmid. Nowadays, the transfer of the gene of interest from the plasmid into the bacmid is usually accomplished in two major ways: (1) by using *E. coli* as an intermediate host or (2) directly in insect cells. (1) In the first strategy, the plasmid is introduced into special *E. coli* cells, which contain bacmid DNA and a transposition enzyme that will transfer the recombinant gene into the bacmid. The recombinant bacmid can then be isolated and introduced into insect cells (Bac-to-Bac, Invitrogen). (2) The gene of interest can also be introduced in a single experimental step, where insect cells are transfected simultaneously with an "empty" bacmid and the plasmid carrying the gene of interest. In the cell the recombination takes place and elaborate selection mechanisms allow for selecting infectious, recombinant viruses. (*flash*Bac, OET; BacMagic Novagen). A few milliliters containing recombinant virus and infected cells are obtained by these procedures [8, 9, 11].

For large scale expression huge numbers of such viruses are needed, as all cells in the culture should be infected by at least one virus particle. In order to obtain such large numbers of viruses, cells are infected with virus at a certain virus-to-cell ratio (multiplicity of infection, MOI) and after proliferation of the virus the supernatant of the culture containing the new generation of viruses is kept. In order to further increase the concentration of the virus stock solution, it is in turn used to infect a much larger volume of cells. Usually three such amplification steps are required to obtain 50 ml of highly infectious virus stock, containing typically 10^7–10^8 plaque forming units.

The number of viruses is measured in plaque assays, where plaque forming units (pfu) are counted. In a plaque assay, a monolayer of adherent insect cell grown in a petri dish is infected with a given amount of virus solution. After 4–5 days of incubation, plaques become visible. Plaques are circular regions of killed cells in the monolayer, indicating infection with virus. Each plaque therefore represents an original infection event by a single virus. Consequently, the concentration of infectious virus – plaque forming units – can be estimated by plaque assays using appropriate dilutions of virus solution.

The plaque assay can also be used for isolating individual, monoclonal baculovirus in a process called plaque cloning. This will be discussed in the following section.

10.2.6 Optimizing Yields of Expressed Protein

As for labeling experiments in *E. coli*, due to cost reasons it is essential to ensure highest yields of expressed protein per liter of – expensive – medium. Two main factors in the culturing process affect the yields in a labeling experiment (1) the characteristics of the expression medium and pre-culture media and (2) the productivity of the baculovirus.

10.2.6.1 Expression Medium Change Leads to Reduced Yields

Insect cell media should promote highest viability of cells in pre-cultures on the one hand, but also ensure efficient infection by the baculovirus in the expression culture. Generally, stocks are maintained in commercial and unlabeled media. For labeling therefore a medium change from pre-culture to labeled expression culture is needed. In this step, attention needs to be paid to the fact, that insect

cells don't tolerate large changes in the medium very well. Changing from a commercial medium to a medium prepared in house usually results in lowered protein yields [5]. Therefore, pre-cultures should be maintained in a very similar medium as the main culture. This can be ensured by preparing all media following the same recipe or by relying on commercial media, as done in the case of BioExpress2000 (Cambridge Isotope Labs), where no change in protein production could be observed for labeled expression and expression in the corresponding unlabeled medium.

Most insect cell laboratories maintain a variety of stocks in different media of different brands that are tested for highest expression. That's because different formulations of media can lead to differences in expression levels of factors of 2–3. Some vendors offer amino acid free media upon request, which can be supplemented by labeled amino acids and yeast extract. This strategy ensures highest yields, as it gives the flexibility of using different media with their proprietary additive formulations, which are crucial for highly efficient infection by baculovirus and concomitant high level protein expression. As mentioned above, with this strategy the cells also don't experience a significant change of medium when changed form pre-culture to main culture.

10.2.6.2 Plaque Cloning of Virus

In our laboratory, for expensive labeling cultures, special attention is given to the baculovirus preparation. In a process called plaque cloning, single virus clones are selected and individual expression levels are determined. It is always astonishing, how different the expression levels of individual clones – all carrying the gene of interest – are. Plaque cloning takes an additional 1–2 weeks, but since expression levels can be more than doubled by this procedure it is often more economical to go through the process. For a labeling experiment it is therefore recommended to use plaque-cloned virus for highest expression levels.

10.3 Strategies for Isotope Labeling in Insect Cells

10.3.1 General Considerations for Isotope Labeling

The much reduced amino acid metabolism of insect cells represents the major difference important for isotope labeling compared to *E. coli*. Ten amino acids are essential and therefore need to be present in the cell culture medium. This offers an advantage for amino acid specific isotope labeling, as label dilution and metabolic scrambling to other amino acids is much reduced for these amino acids. Uniform isotope labeling, on the other hand, suffers from this fact. Insect cells are – by far – not able to synthesize a protein from basic carbon and nitrogen sources. Therefore most amino acids need to be included in the medium in labeled form, rendering uniform labeling a very costly endeavor.

10.3.2 Culture Media for Insect Cells

The typical composition of insect cell media is given in Table 10.1 [11, 20]. The ingredients can be grouped in different classes: Minerals, carbohydrates, amino acids, lipids and yeast extract. Additionally, additives like FCS or antibiotics are added. In order to produce isotope labeled proteins, obviously the carbon, nitrogen and – if deuterium labeling is needed – hydrogen sources need to be replaced by labeled ones. The most important sources for these elements that are incorporated into proteins are carbohydrates, amino acids and yeast extract.

Table 10.1 Composition of IPL41 insect cell medium

Group	Ingredient	Amount [mg/l]
	KCl	1,200
Salts	$MgSO_4$	920
	$MgCl_2 \times 6\ H_2O$	1,000
	NaCl	1,000
	$NaHCO_3$	350
	$NaH_2PO_4 \times H_2O$	1,160
	$CaCl_2$	500
	$FeSO_4 \times 7\ H_2O$	0.55
	$NH_4Mo_6O_{24}$	0.042
	$CoCl_2 \times 6\ H_2O$	0.05
	$CuSO_4$	0.4
	$MnCl_2$	0.02
	$ZnCl_2$	0.04
Carbohydrates	Glucose	2,500–10,000
	Maltose	1,000
	Sucrose	1,650
Amino acids (pure)	L-Alanine (–)	300
	L-Arginine (–)	800
	L-Asparagine (–)	1,300
	L-Aspartic acid (–)	1,300
	L-Glutamic acid (–)	1,500
	L-Glutamine (–)	2,000
	Glycine (–)	200
	L-Histidine (–)	200
	L-Isoleucine (+)	750
	L-Leucine (+)	250
	L-Lysin HCl (+)	700
	L-Methionine (+)	200–1,000
	L-Phenylalanine (+)	1,000
	L-Proline (–)	500
	L-Serine (–)	500
	L-Threonine (+)	200
	L-Tryptophan (+)	200
	L-Tyrosine 2 Na (+ Phe)	360
	L-Valine (+)	120
	L-Cystine 2 HCl (+ Met)	200
	L-Hydroxyproline	800
Vitamins	Inositol	0.4
	Nicotinic acid	0.16
	Pyridoxine HCl	0.4
	Thiamine HCl	0.08
	Ca Panthothenic acid	0.008
	p-Aminobenzenic acid	0.32
	Vitamin B12	0.24
	Biotine	0.16
	Choline chloride	20
	Folic acid	0.08
	Riboflavine	0.08
Organic acids	Succinic acid	4.8
	L-Malic acid	53.6
	Fumaric acid	4.4
	Ketoglutaric acid	30

(continued)

Table 10.1 (continued)

Group	Ingredient	Amount [mg/l]
	Cholesterol	2.81
Lipids	Tween 80	15.62
	Tocopherolacetate	1.25
	Cod liver oil*	6.25
	Pluronic	1,000
Yeast or algal extract	Yeast extract/Yeastolate*	4,000–6,000

The ingredients are divided into major groups (left)
The pH of insect cell media is adjusted to 6.2, and the osmolarity is typically 320 mOsm/kg
(+) and (−) indicate whether an amino acid is essential or not, respectively. Tyrosine is not essential if Phenylalanine is present, Cysteine if Methionine is present
*asterisks indicate undefined mixtures of animal origin

Firstly, as apparent from Table 10.1, labeling in insect cells is therefore very expensive. Secondly, another obstacle is that insect cell media contain components like cod liver oil, FCS, pluronic and yeast extract or yeastolate, which are poorly defined mixtures and are not easily available in isotope labeled form. To our knowledge, it is not possible to formulate a "minimal medium" with pure nitrogen and carbon sources for insect cells as it is done for *E. coli*. While cells can grow well without FCS and cod liver oil, yeast extract and pluronic are essential. Therefore, these two ingredients need to be discussed in greater detail:

Yeast extract or the ultra-filtrated aqueous solution of it, yeastolate, is present in concentrations of up to 6 g/l in standard insect cell media. It represents a major obstacle for isotope labeling because of its high amino acid content. It has been replaced by pure components like nucleic acids, amino acids, lipids etc. in numerous – unsuccessful – attempts. The viability of cells in yeast extract-free media is inevitably reduced, and it is not possible to maintain cells for several passages in such media. However, cells can survive for a few days without yeast extract, which is enough time for an expression experiment. Still, reproducibility of experiments is not satisfactory and infection with baculovirus seems to be less effective without yeast extract as witnessed by three to five fold lowered expression amounts.

Lipids and the lipid mixture pluronic are also vital components of insect cell media and cannot be replaced. They are crucial as a protectant against cell rupture in shaking suspension cultures and play an important role in ensuring efficient infection by the baculovirus.

Therefore, for reasons of cost and/or feasibility, only parts of the ingredients are exchanged for isotope labeled ones and some unlabeled components are always present in the medium. It is therefore important to assess, which ingredients are metabolized into amino acids, as these will need to be labeled. Fortunately, lipids don't seem to be significantly metabolized into amino acids. The challenges represented by the yeast extract are much greater.

10.3.3 Isotope Incorporation Ratios

With the published protocols, isotope incorporation of maximally 90–94% can be obtained [5, 13]. Unlabeled components, which are present in the labeling medium as described above, are not the main reason for lowered incorporation. There are three more important sources of unlabeled amino acids and carbohydrates, which need to be considered: (1) The endogenous pool of unlabeled amino acids and carbohydrates of insect cells from carry over from the pre-culture and (2) from the suspension of BV, which is unlabeled. Additionally, (3) metabolic scrambling of the isotope labeled atoms in the labeled compounds, e.g. amino acids or sugars, can occur.

10.3.3.1 Pre-culture

Stocks and pre-cultures are usually grown in unlabeled media or commercial media due to cost reasons. Unlike *E. coli*, insect cell pre-cultures cannot be diluted strongly, e.g. 1:1,000, for inoculating the main expression culture. A maximum of 1:4 is usually attainable. Therefore, cells are gently centrifuged (3–10 min at 400 g) and only the cell pellet is re-suspended in the labeling medium. Optionally, cell pellets can be washed in PBS to further dilute the unlabeled pre-culture medium. By this method the carry over of medium is minimized. However, the endogenous pool of unlabeled amino acids remains. Fortunately, production of the target protein starts typically only 8 h post-infection. During this time the endogenous unlabeled amino acids are strongly diluted.

If cost permit, the pre-culture may be run in labeled medium. However, self made labeling media are often not as proliferative as commercial ones. This may lead to reduced viability in the main culture and lower expression yields.

10.3.3.2 BV Suspension

The second source of unlabeled components is the suspension of baculovirus. Baculovirus is produced in unlabeled media, usually supplemented with 10% FCS. Depending on the multiplicity of infection needed, an important volume of this suspension needs to be added to the expression culture, leading to significant label dilution. It is therefore essential, to produce a highly concentrated virus solution with more than 10^8 pfu. This allows limiting the virus solution to 1–2% of the final culture volume. Such high virus titers can be achieved by extra rounds of virus amplification, or simply by gently concentrating the virus solution. The first of the two methods is the preferred one, as virus may loose infectivity by the concentration process.

Again, here an alternative would be to carry out the last amplification round in labeled medium.

10.3.3.3 Metabolic Scrambling of Isotopes

^{15}N label dilution and scrambling only appears for the few amino acids biosynthetically directly related to the citric acid cycle, *i.e.* Asn, Asp, Gln, Glu, and Ala as well as Gly. For ^{13}C, label dilution from carbohydrates in the medium (foremost glucose) occurs in varying degrees. In an experiment with only carbohydrate labeled BioExpress2000 medium, 77% of carbons of Ala were labeled. This occured also for Gln/Glu (28%) and Asn/Asp (16%) [14]. For ^{2}H we found no examples in literature, but we expect that the same rules apply for it as for carbon labeling.

In general, as mentioned above, less scrambling is happening in insect cells than in bacterial cells. In Fig. 10.1, spectra from a ligand binding experiment based on ^{15}N-Phe labeling in insect cells are shown, the advantage of reduced cross-labeling in insect cells compared to *E. coli* was evident. While in *E. coli* pronounced cross-labeling of Phe to Tyr, and to a lesser degree to Asp and Glu, is usually observed, in insect cells only the expected number of signals arising from Phe was detected.

10.3.4 Amino Acid Specific Labeling

Insect cells lend themselves very well for amino acid specific labeling, as their amino acid metabolism is much reduced compared to *E. coli*. Ten amino acids are essential and most others are not readily metabolized (Table 10.1). Therefore, for most amino acids no special attention needs to be paid to scrambling or label dilution (Fig. 10.1). In practice, in the published protocols the amino acid of

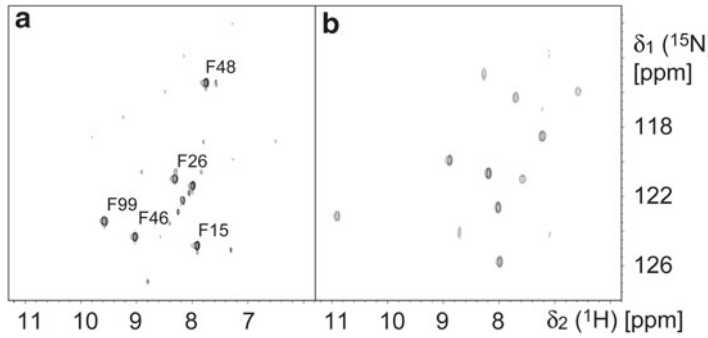

Fig. 10.1 Reduced metabolic scrambling of isotopes in insect cells: example of ^{15}N-Phe labeling. (**a**) [^{15}N,^{1}H]-HSQC spectrum of FKBP labeled with ^{15}N-Phe in *E. coli*, using a optimized protocol based on Muchmore et al. No specific enzymatic inhibitors or auxotroph strains are used. The assignments of the five expected signals are shown. Scrambling of the ^{15}N label to Tyr, and to a lesser degree to Asp and and Glu, leads to additional signals. In [^{15}N,^{1}H]-HSQC spectra of a kinase produced in insect cells (**b**) only the expected number of 11 signals for ^{15}N-Phe are visible

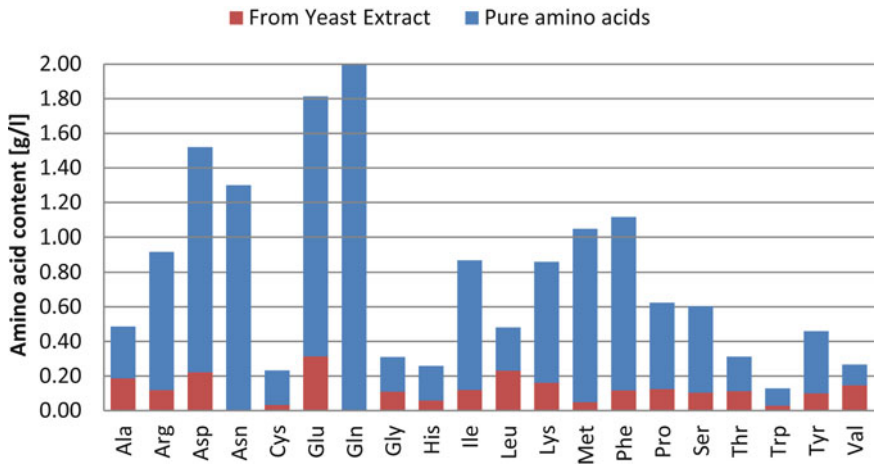

Fig. 10.2 Relative amounts of amino acids in insect cell media originating from pure sources (*blue*) or from yeast extract (*red*). The amino acid composition of a typical insect cell medium is given, containing 6 g/l of yeast extract

interest is just replaced by the isotope labeled version, when preparing the expression medium. Two approaches have been followed: Preparing the media from scratch [1] or ordering amino acid-free versions of commercial media, *i.e.* without amino acids and yeast extract [2, 5, 13, 18].

In this way ^{2}H, ^{15}N and ^{13}C nuclei can be introduced fairly easily in an amino acid specific manner into the protein. With this simple procedure, already in 1998 and 1999, $^{15}N_2$-Lysine and ring-$^{2}H_4$-Tyrosine labeled samples of rhodopsin were produced with isotope incorporation ratios of 60 and 70%, respectively [2, 4]. In 2003, Brüggert et al. brought forward a very similar protocol, where they presented a complete medium formulation for preparing media from scratch. Incorporation ratios for this medium were probably similar as above, but exact numbers were not stated in the publication [1].

In amino acid specific labeling, isotope incorporation ratios represent the main challenge. Due to the presence of significant amounts of amino acids from yeast extract in the medium, incorporation ratios can be low (Fig. 10.2). For certain applications, notably dual amino acid labeling [19], higher incorporation ratios are needed and strategies have been devised to increase them.

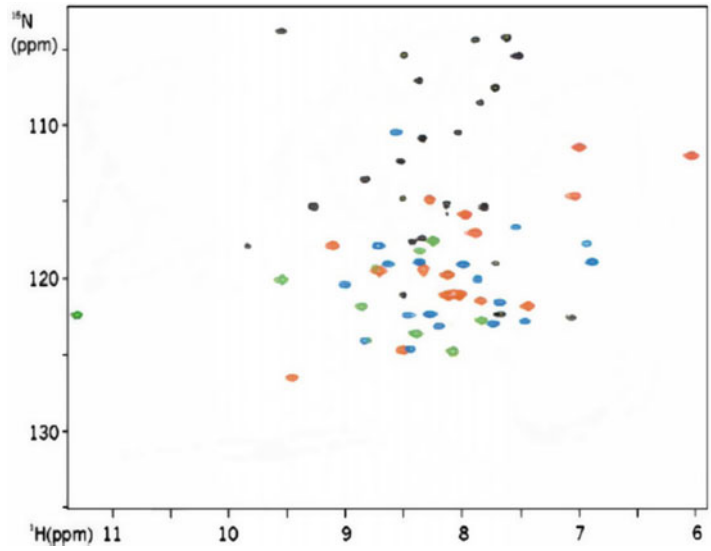

Fig. 10.3 Amino acid type selective labeling of Abl-kinase in BV-infected insect cells. 2D [^{15}N,^{1}H]-HSQC spectra of Abl kinase labeled with ^{15}N-Phe (*green*), ^{15}N-Tyr (*red*), ^{15}N-Val (*blue*) or with ^{15}N-Gly (*black*). Note secondary signals in the Gly spectrum, which arise from cross-labeling to Ser. Gly is one of the few amino acids, where such scrambling is observed in insect cells. (From Strauss et al. [13])

Strauss et al. opted for completely leaving away the yeast extract from the medium. Incorporation ratios of 91–94% are obtainable by this approach for non-metabolized amino acids, and yielded high quality spectra (Fig. 10.3, [13]). This, however, comes at the cost of three to five fold lowered protein yields and less reproducibility. Probably this is due to factors in the yeast extract that facilitate virus entry into the cell. The cells generally have a 1–3 μm smaller diameter at the end of an expression experiment in media without yeast extract than in full media, which may indicate less efficient infection. Nevertheless, this seems to be the most often and most broadly applied strategy for amino acid selective isotope labeling. A number of successful studies are based on this protocol.

In a recent publication based on the above protocol, we have shown that adding 10% of the usual amount of yeast extract can rescue yields and reproducibility to similar levels as in full media. At the same time incorporation ratios are only reduced by 2% compared to the yeast extract-free approach. Knowledge of the amino acid composition of yeast extract enabled adjusting its amount in such a way that no amino acid was present in the medium at more than 5% in unlabeled form [5].

The knowledge of the amino acid composition of yeast extract also makes an approach possible without much work required and high protein yields. The total amount of a given amino acid present in a commercial medium can be estimated. Labeled amino acid is added in such an excess to yield a targeted final incorporation ratio. We showed an incorporation ratio of 80% for ^{15}N-Leu labeling of Abl Kinase with this approach [5].

Feasibility of the approach depends on the solubility of the amino acid of interest, its toxicity to insect cells at a high concentration, and the kind of money available for the experiment. The advantage is that no special medium needs to be mixed from many components, the amino acid of interest is just added to a commercial medium. Expression levels are usually highest with this approach, as commercial media are most productive and the cells don't need to adapt to a different medium for expression.

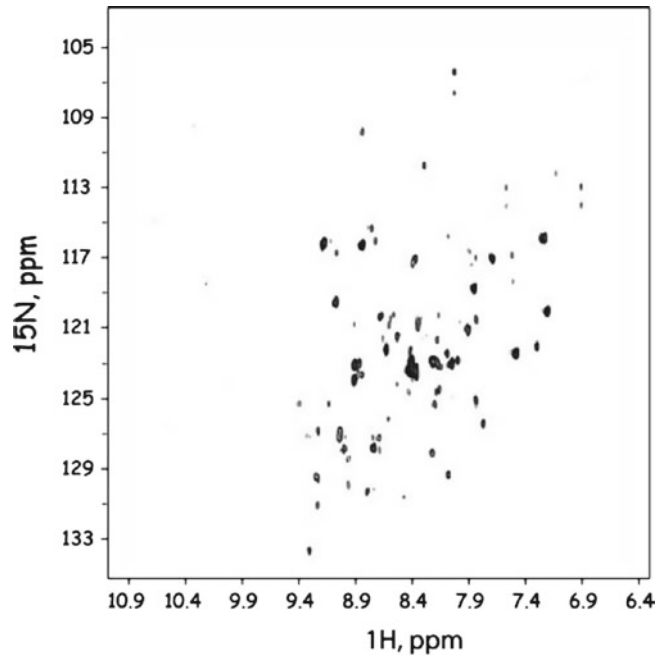

Fig. 10.4 [^{15}N,^{1}H]-HSQC spectrum of uniformly ^{15}N-labeled glycoprotein Thy-1. Additionally, the glycosyl moieties of the protein were ^{13}C-labeled by using ^{13}C-glucose in the medium. (From Walton et al. [18])

10.3.5 Uniform Labeling

Very similar considerations are applicable to uniform labeling. Here however, not only the amino acids present in pure form in the medium need to be replaced by labeled ones – which in itself is already a very costly procedure. On top of this, the yeast extract as well as the carbohydrates need to be isotope labeled. There are two records of uniform labeling in insect cells, one using self-made medium [18] and one relying on a commercial medium [14]. In the first approach, a ^{15}N labeled algal extract was used as the main amino acid source (1.5 g/l). ^{15}N-Cys and Trp were added, as they were not present in the algal extract. Additional ^{15}N sources were included in the medium, $^{15}N_2$-Gln, ^{15}N-Glu and ^{15}N-NH_4Cl, yielding a relatively inexpensive recipe for uniform ^{15}N labeling (Fig. 10.4). Unfortunately, $^{15}NH_4Cl$ seemed to have a negative impact in reducing protein yields by a factor of four. At the same time, yeastolate was added to the medium, thereby deliberately introducing unlabeled amino acids. The levels of incorporation of ^{15}N was not stated in the publication. However, the main focus of the work lied on labeling carbohydrates, which was achieved.

In the other study, BioExpress2000 medium for ^{15}N and ^{15}N/^{13}C labeling was used. Isotope incorporation ratios were above 90% for all amino acids except for Gly (81%), Arg (81%) and His (73%) (Figs. 10.3 and 10.8). Judging from published spectra, this method is superior to the one of Walton et al., but it is also more expensive.

We can only speculate about the formulation of this commercial medium. Probably, a yeast or algal extract is used as the main amino acid source and a few selected amino acids are added, which would be underrepresented or not present at all in the medium if only the extract was used. Additionally, glucose as the main carbohydrate component should be present in labeled form, as this is important for ^{13}C labeling of Ala, Gln/Glu and Asn/Asp, and it is not expensive [14]. Other carbohydrates are

probably not labeled, since they are not readily available and are not metabolized into amino acids by insect cells. In this way, theoretically all combinations of uniform ^{2}H, ^{13}C and ^{15}N labeling patterns can be achieved.

10.4 Case Studies

Isotope labeling of proteins in insect cell media have enabled a number of diverse applications with proteins otherwise not amenable to NMR studies. Examples range from simple detection of ligand binding to studies of different conformational states and dynamics of proteins. Even complete backbone assignments were achieved for the 31 kDa protein Abl-kinase based on uniform labeling in insect cells.

10.4.1 First Applications with Rhodopsin

Interestingly, isotope labeling of proteins in insect cells was not used for NMR experiments in the first place, but for fourier transform infrared spectroscopy (FTIR). In a publication of 1998, ring-$^{2}H_{4}$-Tyrosine was introduced into bovine rhodopsin and the involvement of Tyrosine in photo-activation was studied [4]. However, it did not take long until NMR spectroscopists exploited the labeling method developed for FTIR. One year later, 1D solid state spectra of $^{15}N_{2}$-Lysine labeled rhodopsin were published (Fig. 10.5, [2]). In the spectra obtained under magic angle spinning conditions, isotropic chemical shifts of Lysine residues in rhodopsin could be determined. A strong signal of side chain ε-nitrogens was visible as well as a broader signal of backbone amides from Lysines. Probably, natural abundance ^{15}N from other amino acids also contributed to that signal. Finally, a well-separated resonance was observed, which was attributed to the nitrogen involved in the Schiff-base with retinal. The characteristic chemical shift of that nitrogen confirmed a protonated state of the Schiff-base. Further, calculations were carried out that suggested stabilization of the Schiff-base by a counter-ion in close proximity.

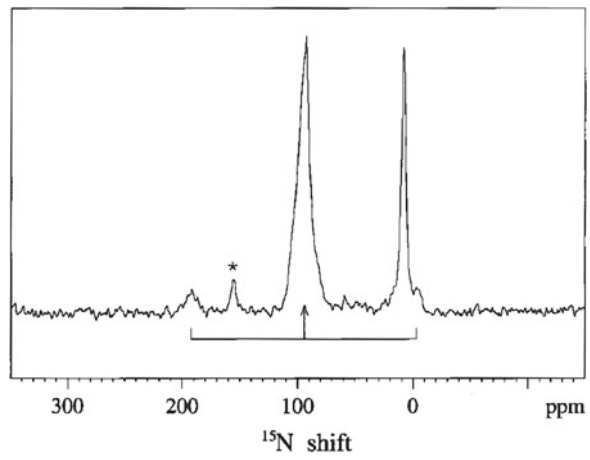

Fig. 10.5 1D ^{15}N solid state spectrum of $^{15}N_{2}$-Lysine labeled rhodopsin. The *asterisk marks* the charachteristic signal of the $^{15}N^{\varepsilon}$ of the lysine involved in the Schiff-base. The *sharp signal* at 10 ppm arised from free $^{15}N^{\varepsilon}$-groups of other Lysines. The *broad central signal* indicated by the *arrow* represents backbone amides, most likely not only from Lysines but also form other amino acids which where labeled by metabolic scrambling of the ^{15}N. Finally, two spinning side bands at $\Delta\delta$ 100 ppm from the major signal are indicated (From Creemers et al. [2]. Figure provided by Dr. HJM de Groot, Leiden University, The Netherlands)

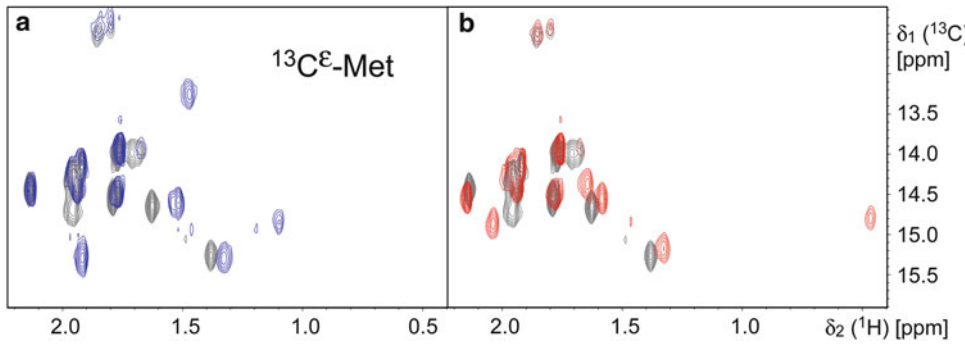

Fig. 10.6 Identification of ligands using a $^{13}C^{\varepsilon}$-Met labeled kinase (33 kDa). 2D [^{13}C,^{1}H]-HMQC spectra of the protein with two different ligands ((**a**) *blue*, (**b**) *red*) are shown, individually superimposed on the spectrum of apo protein (*grey*). Specific binding can be proven by chemical shift perturbations. Spectra were recorded with 50 μM protein samples with 200 μM compound in 30 min (From Gossert et al. [5])

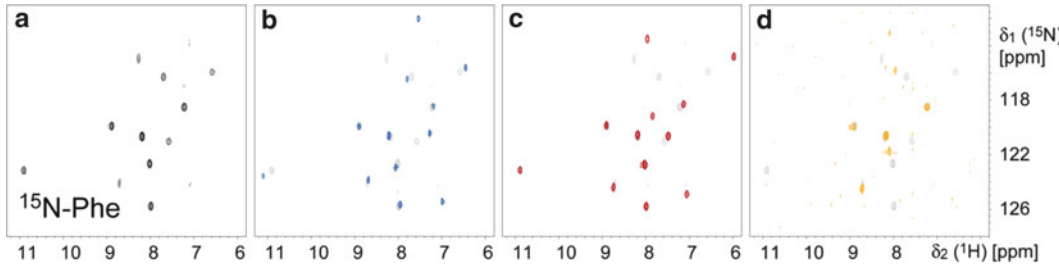

Fig. 10.7 Identification of specific ligand binding using a ^{15}N-Phe labeled kinase. The 2D [^{15}N,^{1}H]-HSQC spectrum of the apo form of the protein is shown in *panel* **a**. The exact number of expected signals is visible in the spectra of the inactive form of the kinase (**a**). Specific binding of different compounds is evidenced by chemical shift perturbations of the resonances, compared to the apo spectrum in *grey* (**b**, **c**). The compound Dasatinib binds to the active state of this kinase [3], leading to a qualitatively different spectrum due to internal dynamics (**d**). Spectra were recorded with 50 μM protein samples with 200 μM compound in 4 h (From Gossert et al. [5])

10.4.2 Amino Acid Selective Labeling for Binding Determination of Ligands

There are a number of NMR methods for detection of ligand binding to a protein without the need to isotope labeled protein. However, these approaches suffer from the risk of false positives. The gold standard in ligand binding detection are protein-observation spectra relying on chemical shift perturbation. In crowded 1D ^{1}H spectra, it is very often difficult to assess such chemical shift perturbations induced by true binding of a ligand. Here, simplified 2D HSQC spectra of amino acid selective labeled proteins represent an ideal alternative.

In a recent study, we used a selectively $^{13}C^{\varepsilon}$-Met labeled kinase produced in insect cells to study ligand binding. Ligand binding or non-binding was clearly demonstrated in the simplified [^{13}C ^{1}H]-HSQC spectra by presence or absence of chemical shift distortions upon addition of ligand. The same can be very well appreciated in [^{15}N ^{1}H]-HSQC spectra. The first approach, however, has the advantage of much reduced measurement times. (Figs. 10.6 and 10.7, [5])

10.4.3 Abl-Kinase

As the following examples all deal with kinases and their conformational states, a brief introduction is given here. Kinases are proteins with considerable inherent flexibility. Foremost two polypeptide regions, the phosphate binding loop (P-loop) and the activation loop, can adopt different conformations and dynamics, which are associated with activation or inactivation of the kinase. These non-rigid regions are important for the fine-tuning of kinase activity in the living cell.

Many forms of cancer are linked to mutations of kinases. In the case of Abl-kinase, constitutively activated protein leads to life-threatening chronic myelogenous leukemia. Therefore, attempts were made to find inhibitors that bind Abl-kinase and keep it in an inactive state.

10.4.4 Uniform, Dual and Single Amino Acid Labeling for Complete Assignments of the 31 kDa Protein Abl-Kinase

In a tour-de-force Vajpai and co-workers managed to obtain 96% of the assignments of backbone $^{1}H^{N}$, ^{15}N, $^{13}C^{\alpha}$ and $^{13}C'$resonances of the 277 residue protein Abl-kinase (Fig. 10.8). Protein samples with different labeling patterns were produced in insect cells [14]. Notably, uniformly ^{13}C,^{15}N labeled protein enabled backbone assignments based on HNCO, HNCA, HN(CO)CA spectra and an ^{15}N-edited ^{1}H,^{1}H-NOESY [16]. Due to the lack of deuteration, experiments encompassing $^{13}C^{\beta}$ were not sensitive enough. Therefore, the assignment was aided and verified by spectra of a total of 15 selectively labeled samples that had ^{15}N, $^{13}C^{\alpha}$ and $^{13}C'$ nuclei introduced specifically for certain amino acid types or combinations thereof. These assignments enabled functional studies of Abl-kinase.

10.4.5 Quadruple Amino Acid Labeling for Understanding Activation States

With NMR studies based on chemical shift perturbations, residual dipolar couplings, and ^{15}N relaxation, different ATP-site binders of Abl-kinase were characterized [15]. As uniform labeling in insect cells is very expensive, the experiments for functionally characterizing Abl-kinase were focussed on the P-loop and the activation loop. By labeling only four amino acid types, FGMY, most residues of the two loops were labeled. As a consequence of metabolic scrambling from Glycine also Serine residues were labeled. Besides the amino acid-type information, this selective labeling scheme had the advantage that all key resonances were completely free of overlap in the simplified spectra, making possible precise measurements of residual dipolar couplings and ^{15}N-relaxation parameters. These data were in agreement with the binding sites of the ligands that were previously defined in crystal structures. More importantly, the solution NMR data showed that the inactive and active states of the kinase induced by the ligands were present in solution and were not just consequences of crystal packing or preferred conformations in crystals.

10.4.6 A Conformational Assay

A different type of ligand binding assay was brought forward by [7]. In a drug discovery effort, allosteric ligands were found, which induced the activated or the inactivated state of the kinase, depending

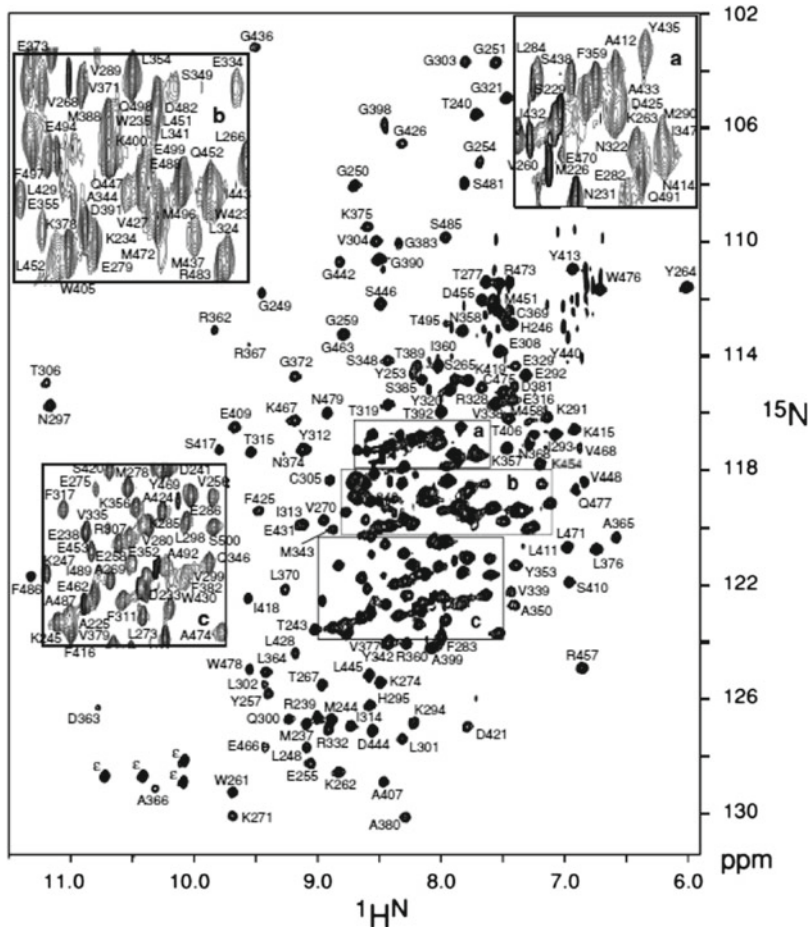

Fig. 10.8 Backbone assignments of Abl-kinase obtained from uniformly and amino acid specifically labeled samples produced in insect cells. A 2D [^{15}N, ^{1}H]-TROSY correlation spectrum is shown, with the assignments indicated (From Vajpai et al. [16])

on the ligand. Only the ligands leading to the inactive state of the kinase were true inhibitors and had a potential as anti-cancer drugs.

Here, not the question of binding and non-binding was central, but the conformations of Abl-kinase that the different ligands induced. When going from the inactive conformation to the active one, the C-terminal alpha helix of Abl kinase would unfold. In the [^{15}N, ^{1}H]-HSQC spectrum of ^{15}N-Val labeled Abl kinase, a single signal would at the same time become much more intense. This signal arised from the Valine in the C-terminal helix, and upon unfolding of the helix the signal became much sharper, due to the fast dynamics of the unfolded state. In summary, a fast assay was implemented based on simplified [^{15}N, ^{1}H]-HSQC spectra, which allowed quickly characterizing the conformation of Abl kinase in presence of different ligands (Fig. 10.9). This was a much faster approach than crystallization of each complex and as measurements were carried out in solution, crystal packing could not influence conformations. Ultimately, this study was enabled by expression of selectively ^{15}N-Valine labeled Abl-Kinase in insect cells by the method of [13].

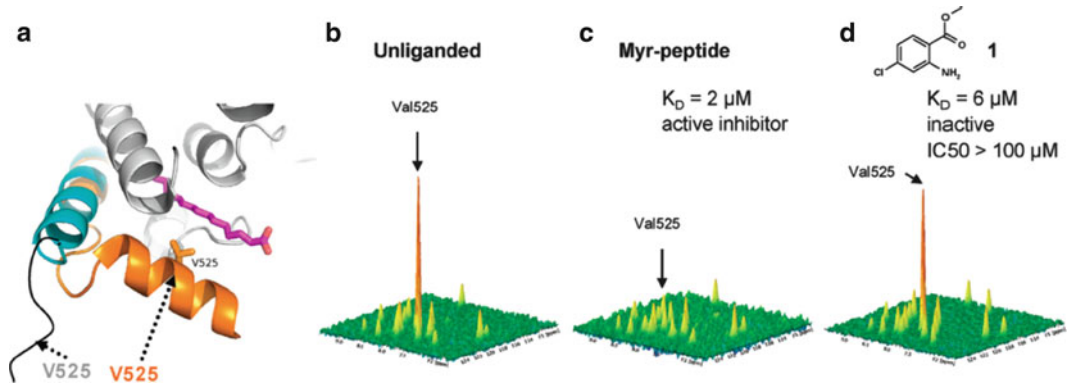

Fig. 10.9 A conformational assay for testing the activation state of Abl kinase upon binding of different compounds. In (**a**) the conformations of the C-terminal helix of Abl kinase domain in the inactive state (*orange*, helix formed) and the active state (*black,* flexibly disordered) are shown in a ribbon model. The position of V525 is indicated. In (**b–d**), [^{15}N, ^{1}H]-HSQC spectra of ^{15}N-Valine labeled Abl-kinase are shown, with the resonance of Val525 indicated. The spectra were recorded without a ligand present (**b**), with a myristoylated peptide (**c**) and with a compound (**d**). The intensitiy of the signal of Val525 indicates the flexibly disordered state of the C-terminus (**b** and **d**) or the well structured helical state (**c**), reflecting the active state and inactive state of the kinase. Therefore, the compound shown in (**d**) doesn't yield an inactive conformation of Abl kinase and is therefore not an inhibitor, although it binds to the myr-pocket (From Jahnke et al. [7])

10.4.7 Conclusion

The above examples show that with expression in insect cells difficult-to-express proteins can be studied by NMR. While uniform labeling is possible, selective labeling is in most cases sufficient to play out all the strengths of NMR – characterizing interactions, conformations and dynamics of proteins in solution.

Acknowledgements We would like to thank André Strauss for his work in establishing isotope labeling in insect cells in our lab and for sharing his knowledge. Further we thank Sébastien Rieffel and Binesh Shrestha for discussions on general insect cell expression techniques.

References

1. Brüggert M, Rehm T, Shanker S, Georgescu J, Holak TA (2003) A novel medium for expression of proteins selectively labeled with N-15-amino acids in Spodoptera frugiperda (Sf9) insect cells. J Biomol NMR 25:335–348
2. Creemers AF, Klaassen CH, Bovee-Geurts PH, Kelle R, Kragl U, Raap J, de Grip WJ, Lugtenburg J, de Groot HJ (1999) Solid state 15N NMR evidence for a complex Schiff base counterion in the visual G-protein-coupled receptor rhodopsin. Biochemistry 38:7195–7199
3. Das J, Chen P, Norris D, Padmanabha R, Lin J, Moquin RV, Shen Z, Cook LS, Doweyko AM, Pitt S, Pang S, Shen DR, Fang Q, de Fex HF, McIntyre KW, Shuster DJ, Gillooly KM, Behnia K, Schieven GL, Wityak J, Barrish JC (2006) 2-aminothiazole as a novel kinase inhibitor template. Structure-activity relationship studies toward the discovery of N-(2-chloro-6-methylphenyl)-2-[[6-[4-(2-hydroxyethyl)-1- piperazinyl)]-2-methyl-4-pyrimidinyl]amino)]-1,3-thiazole-5-carboxamide (dasatinib, BMS-354825) as a potent pan-Src kinase inhibitor. J Med Chem 49:6819–6832
4. 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 273:23735–23739

5. 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 51:449–456
6. Granados RR, Derksen AC, Dwyer KG (1986) Replication of the Trichoplusia ni granulosis and nuclear polyhedrosis viruses in cell cultures. Virology 152:472–476
7. Jahnke W, Grotzfeld RM, Pelle X, Strauss A, Fendrich G, Cowan-Jacob SW, Cotesta S, Fabbro D, Furet P, Mestan J, Marzinzik AL (2010) Binding or bending: distinction of allosteric Abl kinase agonists from antagonists by an NMR-based conformational assay. J Am Chem Soc 132:7043–7048
8. Kitts PA, Possee RD (1993) A method for producing recombinant baculovirus expression vectors at high frequency. Biotechniques 14:810–817
9. Luckow VA, Lee SC, Barry GF, Olins PO (1993) Efficient generation of infectious recombinant baculoviruses by site-specific transposon-mediated insertion of foreign genes into a baculovirus genome propagated in *Escherichia coli*. J Virol 67:4566–4579
10. Murhammer D (2007) Baculovirus and insect cell expression protocols, vol 388, Methods in molecular biology. Humana Press, Totowa
11. O'Reilly DR, Miller LK, Luckow VA (1994) Baculovirus expression vectors: a laboratory manual. Oxford University Press, Oxford
12. Richardson C (1995) Baculovirus expression protocols, vol 39, Methods in molecular biology. Humana Press, Totowa
13. 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 26:367–372
14. 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 31:343–349
15. Vajpai N, Strauss A, Fendrich G, Cowan-Jacob SW, Manley PW, Grzesiek S, Jahnke W (2008) Solution conformations and dynamics of ABL kinase-inhibitor complexes determined by NMR substantiate the different binding modes of imatinib/nilotinib and dasatinib. J Biol Chem 283:18292–18302
16. Vajpai N, Strauss A, Fendrich G, Cowan-Jacob SW, Manley PW, Jahnke W, Grzesiek S (2008) Backbone NMR resonance assignment of the Abelson kinase domain in complex with imatinib. Biomol NMR Assign 2:41–42
17. Vaughn JL, Goodwin RH, Tompkins GJ, McCawley P (1977) The establishment of two cell lines from the insect Spodoptera frugiperda (Lepidoptera; Noctuidae). In Vitro 13:213–217
18. 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 36:225–233
19. Weigelt J, van Dongen M, Uppenberg J, Schultz J, Wikström M (2002) Site-selective screening by NMR spectroscopy with labeled amino acid pairs. J Am Chem Soc 124:2446–2447
20. Weiss SA, Smith GC, Kalter SS, Vaughn JL (1981) Improved method for the production of insect cell-cultures in large volume. In Vitro J Tissue Cult Assoc 17:495–502