Available online at www.sciencedirect.com
ScienceDirect
Vibrational Spectroscopy 48 (2008) 28–36
VIBRATIONAL SPECTROSCOPY
www.elsevier.com/locate/vibspec

Systematic approach to group-specific isotopic labeling of proteins for vibrational spectroscopy

Bettina Warscheid [a], Sven Brucker [b], Angela Kallenbach [b], Helmut E. Meyer [a], Klaus Gerwert [b], Carsten Kötting [b,*]

[a] *Medizinisches Proteom-Center, Ruhr-Universität Bochum, D-44780 Bochum, Germany*
[b] *Lehrstuhl für Biophysik, Ruhr-Universität Bochum, D-44780 Bochum, Germany*

Received 25 October 2007; received in revised form 9 November 2007; accepted 9 November 2007
Available online 22 November 2007

Abstract

Group-specific isotopic labeling of proteins using stable isotopes such as ^{15}N, ^{13}C, or ^{2}H is a widely applied method for the accurate assignment of signals in FTIR spectroscopy. In this work, we report a systematic study on the metabolic incorporation of various stable isotope-coded amino acids into proteins of interest during organism growth. Target proteins were overexpressed in *Escherichia coli* cells that were grown in modified M9 medium supplemented with the respective stable isotope-coded amino acid. Through state of the art mass spectrometric analyses by means of in-gel tryptic digestion, nano-reversed-phase liquid chromatography and electrospray ionization tandem mass spectrometry we were able to quantify the degree of both the incorporation and the spreading of the isotopic label into target proteins in a fast and efficient manner. Our analyses show that specificity and/or efficiency of such metabolic labeling experiments are often limited. The efficient incorporation of stable isotopes can be facilitated by use of *E. coli* strains auxotrophic for the respective isotope-coded amino acid. We tested various auxotrophic strains for their potential to yield specifically labeled proteins suitable for FTIR or Raman measurements and optimized the culturing conditions. As an application example of our method we show the assignment of tyrosine bands in the time-resolved difference spectrum of the GTPase reaction of Ras by use of group-specific isotopic labeling. In conclusion, our work provides a valuable guideline for the production of group-specific isotopically labeled proteins for FTIR spectroscopic analysis.

Keywords: Proteins; Group-specific isotopic labeling; Time-resolved FTIR; Mass spectrometry; Auxotrophic strains; GTPases

1. Introduction

Application of time-resolved Fourier-transform infrared (trFTIR) difference spectroscopy can reveal detailed insight into the reaction mechanisms of proteins with atomic resolution [1]. In order to obtain accurate information about participating amino acids, H-bonding, protonation states, charge distribution and the time dependence of the reaction, bands observed in the FTIR spectra must be correctly assigned to specific groups of the protein. To this extent, group-specific isotopic labeling is an effective tool, enabling the assignment of bands obtained by FTIR spectroscopy [1–4]. In contrast to other methods such as site directed mutagenesis, this approach is non-invasive and all the changes observed in the respective FTIR spectra must have originated from the labeled groups of the protein. Nonetheless, despite the fact that isotopic labeling methods are frequently used in FTIR studies, these studies often lack an accurate, independent measure quantifying the incorporation and spreading of the label.

In order to yield group-specific isotopic labeling of expressed proteins, cells are usually grown in a modified M9 medium containing all naturally occurring amino acids except for one or more amino acids which are replaced by their labeled variants. While it is anticipated that isotope-coded amino acids are fully incorporated into expressed proteins, incomplete labeling of proteins may result from the organism's ability to produce the selected amino acids by metabolizing unlabeled precursors. For example, in *Escherichia coli*, which is one of the most commonly used systems for overexpression of proteins, all amino acids are non-essential. Consequently, all

* Corresponding author. Tel.: +49 234 32 24873; fax: +49 234 32 14238.
E-mail address: carsten.koetting@rub.de (C. Kötting).

0924-2031/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.vibspec.2007.11.003

amino acids may be incompletely incorporated into proteins in the respective labeling experiments [5–7]. Furthermore, the labeled amino acid may be metabolized to other amino acids during organism growth, eventually yielding in a spreading of the label. Both incomplete incorporation of isotope-coded amino acids and spreading of the label will in all likelihood lead to the incorrect assignment of spectroscopic bands in FTIR spectra. In order to prevent such issues, an independent method for determining the degree of the incorporation and/or spreading of the label has to be sought. Since the incorporation of stable isotopes into proteins induces distinct mass shifts, mass spectrometry (MS) represents an adequate technique to monitor those labeling experiments. In previous studies radiotracers like ^{14}C [8,9], MS after complete degeneration of the protein [10] or other spectroscopic techniques [11] were used for the estimation of the incorporation and spreading.

In this paper, we report methods for the group-specific stable isotopic labeling of target proteins using *E. coli* expression systems. We demonstrate the usefulness of modern MS methodologies to obtain detailed information on the degree of both incorporation and spreading of different isotopic labels into the expressed proteins. After in-gel tryptic digestion of labeled and unlabeled target proteins, the resulting peptide mixtures were individually analyzed by nano-reversed-phase liquid chromatography (nano-RPLC) directly coupled to electrospray ionization tandem mass spectrometry (ESI-MS/MS) [12]. Relative quantitative evaluation of the respective mass spectra provides detailed information on the degree of incorporation and/or spreading of the isotopic label in proteins. Additional site-specific information was obtained from the corresponding peptide fragmentation (MS/MS) spectra. We further report the adjustment of culturing conditions and the use of auxotrophic strains to yield stable isotope-labeled proteins suitable for FTIR spectroscopic investigations. Finally, we provide an applied example in which we assign bands in the difference spectra of the GTPase reaction of Ras.

2. Materials and methods

2.1. Expression and purification

Ras, Ran and Rna1p were overexpressed in different *E. coli* strains, namely CK600K [13] (Ras), BL21 DE3 [14] (Ran) and B834 DE3 pLysS [15] (Rna1p), with transformed ptac c-h-ras, pET3d ran and pET11d rna1p (A. Wittinghofer, MPI Dortmund) by using an optimized M9 medium [16] consisting of thiamine–HCl (0.2 mM), thymine (0.5 mM), glucose (50 mM), $MgSO_4$ (2 mM), $CaCl_2$ (0.2 mM), Na_2HPO_4 (48 mM), KH_2PO_4 (22 mM), NaCl (12 mM), NH_4Cl (12.5 mM) and the appropriate antibiotics (pH = 7.6). The media were used as full media, to which 0.834 mM of each amino acid was added. For metabolic labeling experiments the selected amino acid was replaced by its isotopically labeled variant. All the labeled amino acids used in this work were obtained from Cambridge Isotope Laboratories, Inc. (Andover, MA, USA). The purification of the target proteins was performed by ion exchange chromatography and gel filtration chromatography as described previously [13,17,18].

2.2. Auxotrophic strains

In addition to the above reported *E. coli* expression strains, auxotrophic strains were also used, namely *E. coli* TH16#4 [19] for glutamine, *E. coli* ER [20,21] for asparagine and *E. coli* DL39 [22] for aspartate. These strains have knocked out the genes for the synthetase(s) of the specific amino acid they are auxotrophic for, namely asnA and asnB for asparagine, glnA for glutamine, aspC and tyrB for aspartate, and gdh and gltB for glutamate [5]. All auxotrophic strains used in this work were checked for their inability to grow in the absence of the amino acid for which they should be auxotrophic for. Since λDE3 prophage was integrated with a λDE3 lysogenization kit (Novagen®), the lysogenized host could be used to express target genes cloned in the pET vectors under the control of the T7 promoter. To ensure adequate cellular growth and protein expression, the M9 media was optimized for the strains TH16#4 and ER by using the twofold amount of isotopically labeled glutamine or asparagine, respectively. Additionally the NH_4Cl concentration was increased to 60 mM in the case of the ER strain. In the case of the strain DL39 the concentrations of the M9 medium were doubled and the fourfold amount glutamine was used.

2.3. Gel electrophoresis and in-gel digestion of proteins

A denatured and discontinuous SDS-PAGE [23] (10% acrylamide for Rna1p, 15% acrylamide for Ras and Ran) was performed. The gels were loaded with 1.5 μg of each labeled and unlabeled target protein. After gel electrophoresis, the proteins were visualized by Coomassie® brilliant blue G250 staining [24] and the respective bands were manually excised from the gel and then transferred into glass mini tubes. Gel pieces were immediately destained by alternately incubating them with 20 μl of 10 mM ammonium hydrogencarbonate (NH_4HCO_3) and 20 μl of 5 mM NH_4HCO_3/50% ACN, each for 10 min intervals. This step was performed thrice. Afterwards, gel pieces were dried *in vacuo* and either stored at −80 °C or directly subjected to proteolytic digestion with trypsin (Promega, Mannheim, Germany) dissolved in 25 mM NH_4HCO_3 (pH 7.8) at a final concentration of 0.03 μg/μl. For in-gel digestion of proteins, gel pieces were incubated overnight at 37 °C with 2 μl of trypsin solution and slight agitation. Subsequently, tryptic peptides were extracted twice with 10 μl of ACN and 5% FA mixed 1:1 (v/v); extracts were combined and ACN was removed *in vacuo*. For mass spectrometric analysis, samples were acidified by addition of 5% FA to a final volume of 20 μl.

2.4. Mass spectrometry

Tryptic peptides from target proteins were analyzed by nano-HPLC/ESI-MS/MS using a Dionex LC Packings system (Dionex LC Packings, Idstein, Germany) coupled to a QSTAR XL instrument (Applied Biosystems, Foster City, CA, USA). Peptide mixtures were separated by online RP capillary HPLC as previously described by Schaefer et al. [25]. The MS

instrument was equipped with a nano-electrospray ion source (SCIEX, Toronto, Canada) and distal-coated SilicaTips (FS360–20–10-D; New Objective, Woburn, MA, USA). For external calibration of the QSTAR XL instrument in the enhanced product ion mode, reserpine (*m/z* 609.280; Agilent Technologies, Santa Clara, CA, USA) was used. The general mass spectrometric parameters were set as follows: ion spray voltage (IS), 1800–2000 V; curtain gas (CUR), 10–14; gas 1.0; declustering potential (DP), 50 V; focusing potential (FP), 220 V; declustering potential 2 (DP2), 15 V. In the Analyst QS software (Applied Biosystems, Foster City, CA, USA), the ''information dependent acquisition'' method was chosen, consisting of a survey MS scan (*m/z* 400–1200) followed by sequential isolation and fragmentation of the three most intense peaks (enhanced product ion scans, *m/z* 100–2000). Only multiple charged peptide ions were isolated and subjected to collision-induced dissociation using nitrogen as collision gas. Collision energies for peptide fragmentation were dynamically adjusted by the software (rolling collision energy) and previously fragmented ions were dynamically excluded for the following 16 s.

2.5. Analysis of mass spectrometric data

Efficient incorporation of stable isotopes into proteins results in a distinct mass shift, which can be directly observed in the respective MS of their proteolytic peptides. In the case that metabolic labeling of proteins by stable isotope-coded amino acids was not fully complete, the degree of isotope incorporation (n) can be quantified by comparing the peak intensities of the corresponding labeled and unlabeled proteolytic peptides in the MS spectrum of the isotopically enriched peptide.

However, if, for example, only a single atom of an amino acid is replaced by its heavy isotope, or that isotope scrambling occurs during metabolic labeling experiments which generate small mass shifts, an overlap of the isotope distributions of unlabeled peptides and their labeled counterparts occurs. In these cases we calculated the degree of isotope incorporation by using the following method:

The peptide is divided into the completely unlabeled amino acids and the partially labeled amino acids. The unlabeled part A would have the natural isotope abundance which is described by the vector $\vec{A}$, which elements A_i are the relative abundance of the masses i in the mass spectrum. If the fraction of labeling is n ($0 \leq n \leq 1$), the partially labeled part can be described by a sum of n times the labeled isotope distribution $\vec{C}$ and $(1 - n)$ times the natural isotope abundance $\vec{B}$. The mass distributions $\vec{A}$, $\vec{B}$ and $\vec{C}$ can be determined by available tool programs like isopro [26]. From this, the MS spectrum of the complete peptide can be calculated by

$$M_i^{\mathrm{s}} = \sum_{l=0}^{i} A_l[(1-n)B_{i-l} + nC_{i-l}] \tag{1}$$

where M_i^{s} is the relative abundance of mass i in the simulation of the measured MS spectrum of the complete peptide.

Using the normalized intensities of the measured spectrum $\vec{M}$ and the calculated abundances of $\vec{A}$, $\vec{B}$ and $\vec{C}$ one can directly calculate the incorporation, n: for an accurate estimation only the peaks of the MS spectrum with a relative intensity >10% between the masses j and k should be used

$$n = \sum_{i=j}^{k} \frac{M_i - \sum_{l=0}^{i} A_i B_{i-l}}{(k-j)\sum_{l=0}^{i} A_i (C_{i-l} - B_{i-l})} \tag{2}$$

2.6. ATR-FTIR

For the measurements of the isolated amino acids, 200 mM solutions of the unlabeled and the labeled amino acid were prepared and measured with a horizontal diamond ATR setup (Resultec) using a Bruker IFS 88 spectrometer. Water at the respective pH was taken as the reference. A linear ATR correction was performed.

2.7. trFTIR

As an example for the application of group-specific isotopic labeling in trFTIR spectroscopy we used the intrinsic GTPase reaction of the GTPase Ras, for which we obtained a photolysis and a hydrolysis spectrum [27]:

$$\mathrm{Ras \cdot caged\text{-}GTP} \overset{\text{photolysis}}{\underset{-\text{cage}}{\longrightarrow}} \mathrm{Ras \cdot GTP} \overset{\text{hydrolysis}}{\longrightarrow} \mathrm{Ras \cdot GDP} + P_i$$

Ras was loaded with (pHP)caged-GTP [28] according to John et al. [29]. The sample composition was 10 mM 1:1 complex of Ras with the caged nucleotide, 25 mM $MgCl_2$, 10 mM DTT and 200 mM Mes (pH 6). Experiments were performed at 283 K using 0.1% NF1 to catalyze the hydrolysis. This results in a much better signal-to-noises (S/N)-ratio for our measurement but does not affect the spectrum. The sample was prepared between two CaF_2 windows as detailed by Cepus et al. [30]. Photolysis of the caged compound was performed with an LPX 240 XeCl excimer laser (308 nm, Lambda Physics, Göttingen, Germany) by 12 flashes within 24 ms. More than 90% of the sample is photolyzed under these conditions. A modified Bruker IFS 66v/s spectrometer in the fast scan mode [31] was used for the measurement.

The data was analyzed between 1800 cm^{-1} and 950 cm^{-1} with a global fit method [32]. In this analysis, the absorbance changes ΔA were analyzed with one exponential ($n_{\mathrm{r}} = 1$) with apparent rate constant k_1 and amplitude a_1:

$$\Delta A(\nu, t) = \sum_{l=1}^{n_{\mathrm{r}}} a_l(\nu) \mathrm{e}^{-k_l t} + a_\infty(\nu)$$

The negative amplitude spectrum ($-a_1\nu$) is the hydrolysis spectrum; the disappearing bands face downwards and the appearing bands face upwards.

An amplitude spectrum $a_0(\nu)$, calculated by

$$a_0(\nu) = a_\infty(\nu) - \sum_{l=1}^{n_{\mathrm{r}}} a_l(\nu)$$

compares the state before the triggering of the reaction by the irradiation with the state after the flashes, extrapolated to time $t = 0$. This is the photolysis spectrum.

3. Results and discussion

In the following section we firstly describe the workflow of our analysis. Subsequent to the experimental outline, we report the results obtained for the group-specific metabolic labeling of target proteins using stable isotope-coded amino acids, which was systematically evaluated by mass spectrometric analysis. Finally, we will illustrate the high applicability of such isotopically labeled proteins for the accurate assignment of bands in FTIR analysis.

Fig. 1 shows a schematic flow chart of our methodological approach to the group-specific labeling of proteins for FTIR analysis. Target proteins were expressed in *E. coli* using a modified M9-medium [16] containing all amino acids one of which was specifically labeled with stable isotopes. After metabolic labeling and protein purification using ion exchange chromatography and gel filtration, both the degree of the incorporation of stable isotopes into target proteins and the potential spreading of the label into other amino acids were determined by nano-RPLC/ESI-MS/MS analyses. We usually considered that stable isotope-labeled proteins were well suited for FTIR measurements if the incorporation of the label was greater 90%, and the spreading was less than 10%. Since in such cases the unspecifically labeled or non-labeled protein variants contribute to less than 10% each, it can be expected that they yield only minor signals in FTIR spectra. In the event that these labeling requirements were not met, media were refined and/or specific auxotrophic strains were used to obtain adequate group-specific labeling of target proteins (Fig. 1).

In Table 1 we summarize our results obtained for the group-specific labeling of target proteins (Ras, Ran and Rna1p) using nine different amino acids employed in various stable isotope-coded forms. Since this approach is largely independent of the expressed protein, providing that the same expression system is used, it should be equally well applicable to the group-specific labeling of other target proteins using the listed amino acids. In addition, we provide information on both the precursors of the different amino acids and their metabolically related amino acids in *E. coli* [7]. Since the latter can be subject of spreading, they were generally checked for potential isotope modification in peptide MS analyses. As a result of our systematic approach we can classify three distinct groups of amino acids that show different characteristics for the group-specific labeling of proteins. The first one includes the amino acids which were incorporated into proteins with a degree of >90% and did not cause any spreading of the label into other amino acids. These were ($\eta^{15}N_2$)-arginine, lysine, (^{15}N)-proline and tyrosine. The second group consists of the amino acids, which showed equally high degrees of incorporation but were also partially converted to other amino acid resulting in the spreading of the label. This was found for glycine and threonine in this work. The third group is the most problematic one and consists of asparagine, aspartate and glutamine. These amino acids generally showed low incorporation rates and in the case of aspartate and glutamine also additional significant spreading of the label was observed. However, by using auxotrophic strains and refining the media conditions as described in Section 2, it was possible to obtain sufficiently high degrees of label incorporation into proteins with high group-specificities. Note that in the case of aspartate even the auxotroph strain could not prevent the spreading into asparagine.

3.1. Successful incorporation

As an example of the complete incorporation of isotope-coded amino acids into proteins, we report the metabolic labeling of bacterially expressed Ras protein using ring-$^2H_4$-tyrosine. As a control, Ras protein was expressed under the same conditions using *E. coli* grown in medium containing tyrosine in its natural form. Following protein purification and sample preparation as described in Section 2, tryptic digests of both unlabeled and ring-$^2H_4$-tyrosine-labeled Ras were subjected to nano-RPLC/ESI-MS/MS analysis in order to assess the efficiency as well as the specificity of the labeling. Fig. 2A shows the doubly charged tryptic peptide SFEDIHQYR from Ras protein observed at a mass-to-charge (m/z) ratio of 597.8 (monoisotopic peak) in the respective MS spectrum. The natural isotopic peaks (^{13}C) of the peptide are observed at higher m/z values, each showing a mass shift of $m/z = 0.5$. Due to isotopic labeling, a mass shift of 4 Da was introduced into the corresponding tryptic peptide from ring-$^2H_4$-tyrosine-labeled Ras protein. Accordingly, the doubly charged peptide

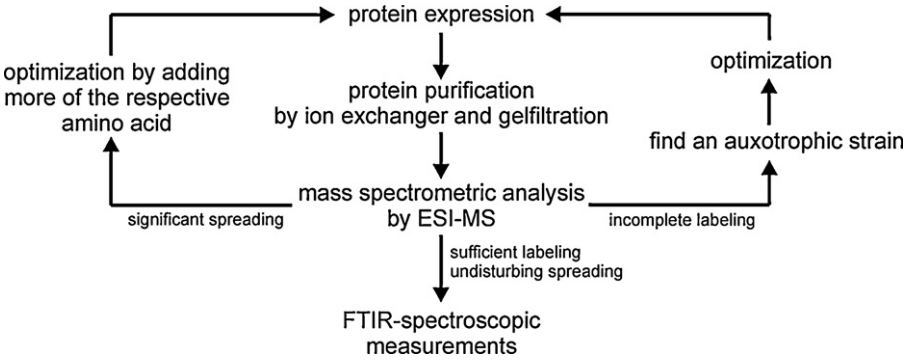

Fig. 1. Flowchart of our methodological approach to control the group-specific metabolic labeling of proteins for FTIR spectroscopic analysis.

Table 1
Summary of the mass spectrometric investigations to assess the group-specific labeling of target proteins expressed in *Escherichia coli* using media supplemented with different stable isotope-coded amino acids

Labeled aa	Kind of labeling	*E. coli* strain	Incorporation (%)	Synthesis	Theoretical spreading	Observed spreading
Arginine	$\eta(^{15}N_2)$	CK600K	>95	Glu	Pro	–
	$\eta(^{15}N_2)$	BL21 DE3	>95			–
Asparagine	U-$^{13}C_4{}^{15}N_2$	B834	<20	Asp	Thr, Lys, Met	–
		ER*	>95			–
Aspartate	U-$^{13}C_4{}^{15}N$	CK600K	None	Oxalacetate	Asn, Thr, Lys, Met	–
	U-$^{13}C_4{}^{15}N$	DL39*	50			Asn (40%)
Glutamine	U-$^{13}C_5{}^{15}N_2$	CK600K	<20	α-Ketoglutarate	Glu, Asp	Glu (∼50%) Asp (∼20%)
	U-$^{13}C_5{}^{15}N_2$	TH16#4*	85–95			Glu (40%) Asp (10%)
	$\varepsilon(^{15}N)$	TH16#4*	>95			–
Glycine	U-$^{13}C_2{}^{15}N$	CK600K	80	Ser, Thr	Cys, Ser	Ser (20%)
Lysine	U-$^{13}C_6{}^{15}N_2$	CK600K	>95	Asp	–	–
	U-$^{13}C_6{}^{15}N_2$	BL21 DE3	>95			–
	$\xi(^{15}N)$	CK600K	>95			–
Proline	U-^{15}N	CK600K	>95	Glu	–	–
Threonine	U-$^{13}C_4{}^{15}N$	CK600K	>95	Asp	Ile, Gly, Ser	Gly (5%) Ser (30%)
	U-$^{13}C_4{}^{15}N$	BL21 DE3	>95			Gly (5%) Ser (30%)
Tyrosine	U-$^{13}C_9{}^{15}N$	CK600K	>95	Chorismate	–	–
	U-$^{13}C_9{}^{15}N$	BL21 DE3	>95			–
	ring-$^2H_4$	CK600K	>95			–

Information about amino acids (aa), stable isotope labels, *E. coli* strains, amino acid precursors and metabolically related amino acids are provided
* Auxotrophic *E. coli* strain.

SFEDIHQYR labeled with $^2H_4$-tyrosine was detected at *m/z* 599.8, while virtually no signal was observed at *m/z* 597.8 (unlabeled peptide form), indicating the almost complete incorporation of the label (Fig. 2B). The specificity of the metabolic incorporation of ring-$^2H_4$-tyrosine into Ras protein could be verified by partial sequencing of the tryptic peptide SFEDIHQYR in its unlabeled and labeled form using MS/MS (Supplementary Fig. 1).

3.2. Spreading of the label

We further investigated the capability of using amino acids, such as $^{13}C^{15}N$-threonine, which can be metabolically converted into other amino acids for the group-specific labeling of Ran protein. Since it is well known that threonine can readily be transformed to serine in bacterial metabolism, one has to consider the potentiality of spreading of the label into serine during protein synthesis and turnover. We therefore examined MS spectra of tryptic peptides from unlabeled Ran and $^{13}C_4{}^{15}N$-threonine-labeled Ran, which contained either threonine or serine in their amino acid sequences. Evaluation of the respective mass spectra of tryptic peptides containing threonine but no serine indicated almost complete incorporation (>95%) of the label into Ran protein (data not shown). We next inspected the MS spectra of the doubly protonated tryptic peptide NLQYYDISK ($[M + 2H]^{2+}$ = 607.8 Da) originating from unlabeled and $^{13}C_4{}^{15}N$-threonine-labeled Ran protein. In comparison with the control the isotopic pattern of the peptide from $^{13}C_4{}^{15}N$-threonine-labeled Ran protein was clearly shifted to higher masses, reflecting a modified isotope distribution (Fig. 3A and B). This finding indicated the spreading of the label by conversion of $^{13}C_4{}^{15}N$-threonine into serine, which was further confirmed by evaluation of the corresponding peptide MS/MS spectra (Supplementary Fig. 2). Using Eq. (2), the degree of metabolic conversion of threonine to serine was estimated to be 30% under the conditions applied in this work. In order to reduce the degree of spreading, we increased the amount of serine added to the M9 medium by 10-fold when using $^{13}C_4{}^{15}N$-labeled threonine for the group-specific labeling of Ran protein. As a result, spreading of the label into serine could be sufficiently reduced (<10%) facilitating the production of $^{13}C_4{}^{15}N$-threonine-labeled Ran protein suitable for FTIR analysis. Using the same approach, the spreading of the label into glycine was quantified to be 5% and no spreading into isoleucine was found.

Another possibility to impede spreading of the label is the use of only partly isotope-coded amino acids. For example, since arginine is a precursor for proline synthesis, we employed arginine that is labeled only at the side chain by ^{15}N, which is then cleaved off in the corresponding metabolic pathway. Accordingly, no spreading of the label from $\eta(^{15}N_2)$-arginine into proline within proteins was observed as indicated in Table 1. Similarly, we used $\varepsilon(^{15}N)$-glutamine instead of uniformly labeled U-$^{13}C_5{}^{15}N_2$-glutamine to prevent spreading of the label into glutamate and/or aspartate residues within expressed proteins in *E. coli* TH16#4 (Table 1).

3.3. Auxotrophic strains

As an example of the incomplete metabolic incorporation of stable isotopes into proteins, we report the labeling of

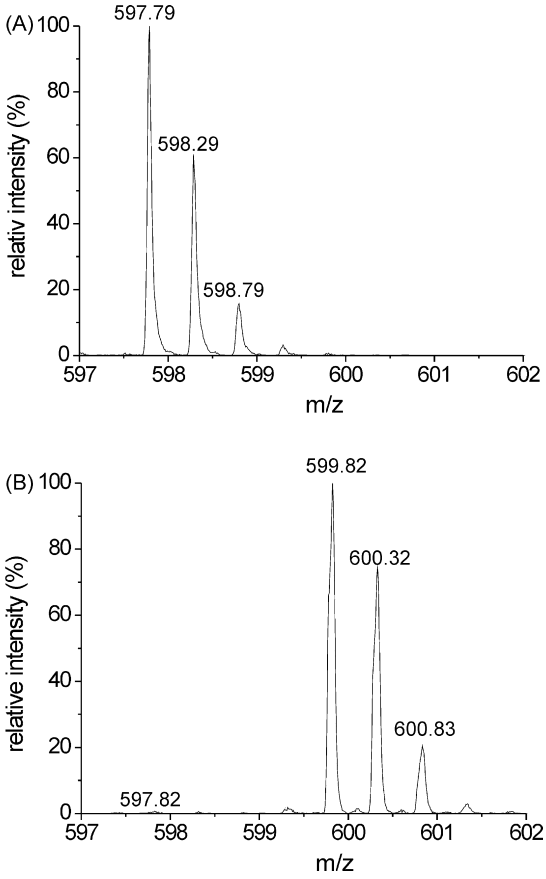

Fig. 2. Mass spectra of the doubly charged tryptic peptide SFEDIHQYR from (A) unlabeled Ras and (B) ring-$^{2}H_4$-tyrosine-labeled Ras with monoisotopic masses detected at *m/z* 597.79 and 599.82, respectively. The observed mass shift of 4 Da between the native and labeled tryptic peptide corresponds with the incorporation of ring-$^{2}H_4$-tyrosine due to metabolic labeling of Ras protein. Since virtually no signal is observed at *m/z* 597.79 in the MS spectrum of the labeled peptide (B), almost complete (>95%) ring-$^{2}H_4$-tyrosine-specific labeling of Ras protein is indicated.

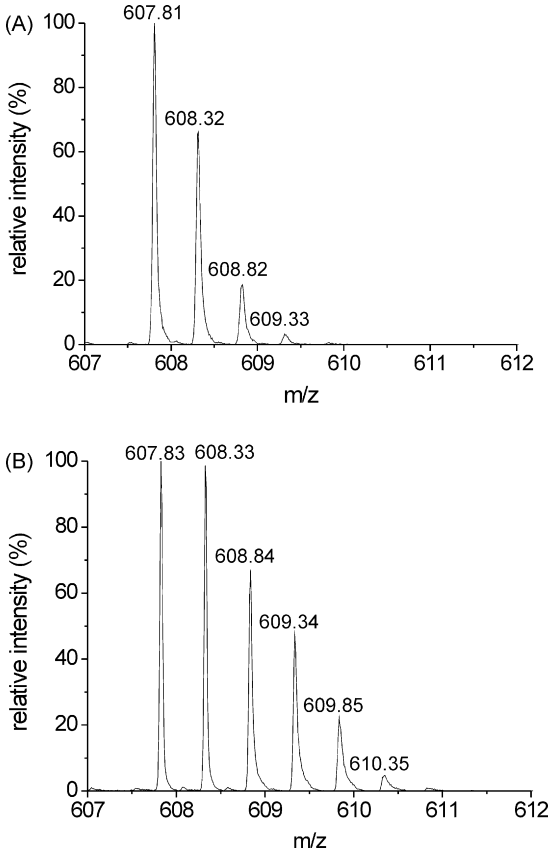

Fig. 3. Mass spectra of the doubly charged tryptic peptide NLQYYDISK ($[M + 2H]^{2+}$ = 607.8 Da) from (A) unlabeled Ran and (B) U-$^{13}C_4{}^{15}N$-threonine-labeled Ran. Since there is no threonine residue in the peptide, no mass shift or change in the isotopic distribution is expected. However, both shifted and broadened isotopic distribution of the peptide from U-$^{13}C_4{}^{15}N$-threonine-labeled Ran (B) was observed, indicating the spreading of the label into serine.

bacterially expressed Rna1p protein using $^{13}C_4{}^{15}N_2$-asparagine. In such cases the degree of label incorporation should be significantly increased by using the respective auxotrophic strains of the bacteria. Fig. 4 shows the MS spectra of the tryptic peptide LENIGLQTLR ($[M + 2H]^{2+}$ = 578.84 Da) of Rna1p in its native form (A) as well as originating from $^{13}C,^{15}N$-asparagine-labeled Rna1p either expressed in an *E. coli* non-auxotrophic strain (B) or a mutant strain auxotroph for asparagine (C). Metabolic labeling using $^{13}C_4{}^{15}N_2$-asparagine should result in a mass shift of 6 Da, resulting in a monoisotopic mass of 581.84 amu of the doubly charged tryptic peptide LENIGLQTLR. However, expression and metabolic labeling of Rna1p protein in the *E. coli* non-auxotrophic strain B834 DE3 pLysS resulted in only about 20% incorporation of isotopically labeled asparagine (Fig. 4B). In order to significantly increase the level of incorporation of $^{13}C_4{}^{15}N_2$-asparagine, we used the auxotrophic *E. coli* strain ER [20,21] in which the genes asnA and asnB are knocked out. Since both genes are essential for the expression of the asparagine synthetase, this strain requires the external supply of asparagine and, thus, expressed proteins should only comprise the $^{13}C_4{}^{15}N_2$-asparagine added into the M9 media. In accordance with this assumption, we could observe virtually complete incorporation (>95%) of the isotopically labeled asparagine as indicated by the intensities of the monoisotopic peaks of the unlabeled and labeled peptide LENIGLQTLR at *m/z* 578.86 and 581.85, respectively (Fig. 4C). The peak observed at *m/z* 581.56 in the spectrum results from ^{14}N and ^{12}C impurities in the supplied $^{13}C_4{}^{15}N_2$-asparagine (98% $^{13}C,^{15}N$).

Note, that we tried several auxotrophic strains for each case of incomplete group-specific labeling of target proteins. As detailed in Section 2, we found that the *E. coli* strain ER for asparagine, DL39 for aspartate and TH16#4 for glutamine provided the best protein yields and the best degree of labeling within our modified M9 medium, respectively.

After employing MS(/MS) to ensure the correct group-specific labeling of the target protein, the isotopically labeled protein can be used for FTIR measurements. This can lead to band assignments in the spectra, which allow conclusions on the reaction mechanism as exemplified in the following section.

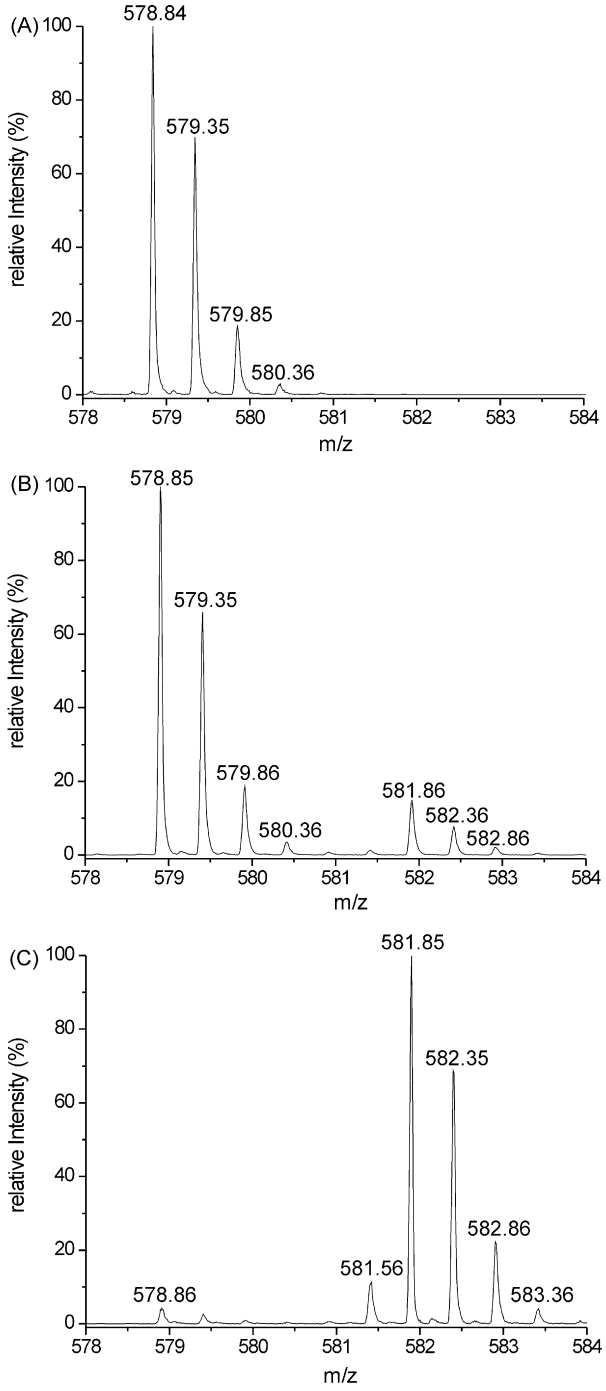

Fig. 4. Mass spectra of the tryptic peptide LENIGLQTLR ($[M + 2H]^{2+}$ = 578.85 Da) from native Rna1p (A) as well as from Rna1p expressed in a non-auxotrophic *Escherichia coli* strain (B) and an asparagine-auxotrophic *E. coli* strain (C) using an U-$^{13}C_4{}^{15}N_2$-asparagine-labeled medium. On the basis of the peak intensities of the U-$^{13}C_4{}^{15}N_2$-asparagine-labeled peptides ($[M + 2H]^{2+}$ = 581.85 Da), the degree of group-specific metabolic labeling was determined to be approximately 20% for Rna1p expressed in the non-auxothropic *E. coli* strain, but over 95% using the respective auxotrophic *E. coli* strain.

3.4. Band assignment in FTIR difference spectroscopy

For a pure stretching vibration of two atoms with the masses m_1 and m_2 without coupling the position of the band of the labeled protein ν_{iso} relative to the position of the band of the unlabeled protein ν can be approximated by

$$\nu_{iso} = \nu\sqrt{\frac{\mu}{\mu_{iso}}} \quad \text{with } \mu = \frac{m_1 m_2}{m_1 + m_2}$$

This gives a good estimate, e.g. for an uncoupled C=O vibration at 1700 cm^{-1} ($\mu \approx 6{,}86$) which shifts upon ^{13}C labeling ($\mu_{iso} \approx 7{,}17$) by about 40 cm^{-1}.

For many other vibrations it is more difficult to estimate the shifts. For example the most intense bands of tyrosine are coupled vibrations with CC-stretching, CH-bending and CO-stretching character and thus the shift upon ^{13}C- or ^{2}H-labeling is lower due to participation of unlabeled atoms. Due to decoupling even shifts to higher wave numbers are possible. Thus we measured the unlabeled and the labeled form of the isolated amino acids dissolved in water. In Fig. 5A, the spectra of unlabeled, $^{13}C_9{}^{15}N$-labeled and ring-$^{2}H_4$-labeled tyrosine at

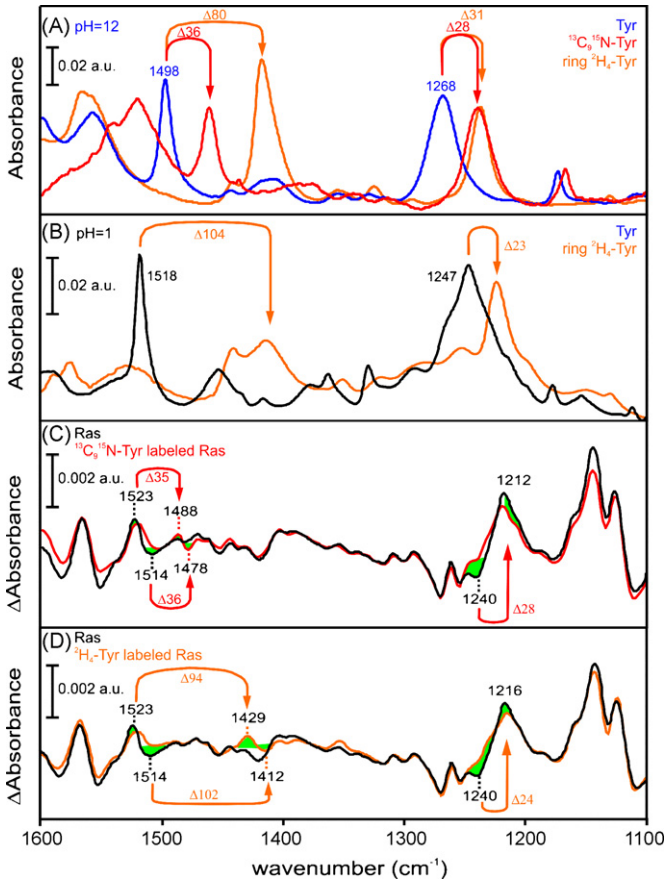

Fig. 5. (A) IR-spectra of tyrosine (blue), U$^{13}C^{15}N$-tyrosine (red) and ring-$^{2}H_4$-tyrosine (orange) at pH 12. (B) IR-spectra of tyrosine (black) and ring-$^{2}H_4$-tyrosine (orange) at pH 1. (C) IR-difference spectra of the photolysis of unlabeled (black) and U-$^{13}C_9{}^{15}N$-tyrosine labeled (red) Ras·cagedGTP. Positive bands are due to the Ras·GTP state, negative bands are due to the Ras·cagedGTP state. Due to an environmental change the two most intense bands of tyrosine appear in the difference spectrum. The band positions are similar to the isolated tyrosine with the protonated side chain as shown in (B). The shifts upon labeling are in agreement with the shifts found for the isolated amino acid. (D) The same IR-difference spectra as in (C) but with unlabeled (black) and ring-$^{2}H_4$-tyrosine labeled (orange) Ras·cagedGTP. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

pH 12 are shown. The two most intense bands of the deprotonated tyrosine side chain are observed at 1498 cm^{-1} (ν(CC)ring + δ(CH)) and 1268 cm^{-1} (ν(CO) + ν(CC)) [33]. We found that by $^{13}C_9{}^{15}N$ labeling of tyrosine these bands shift by 36 cm^{-1} and 28 cm^{-1}, respectively. In the case of the ring-$^2H_4$ labeling of tyrosine the corresponding shifts are 80 cm^{-1} and 31 cm^{-1}. Since costs for the latter labeled amino acid are much lower and it induces even larger shifts, we recommend ring-$^2H_4$-tyrosine for use in FTIR difference spectroscopy.

In Fig. 5B the spectra of native tyrosine and ring-$^2H_4$-labeled tyrosine at pH 1 are shown. Due to the protonation of the side chain, the ν(CC)ring + δ(CH) band shifted towards 1518 cm^{-1} and the ν(CO) + ν(CC) band shifted towards 1247 cm^{-1}. Interestingly, the band shifts due to the ring-$^2H_4$-label are also affected by the protonation. The protonation state of tyrosine can thus be easily determined by both band position and isotopic shift using FTIR spectroscopy.

We further performed trFTIR measurements of Ras as described earlier [27,30,34] using both unlabeled Ras and Ras that was labeled with either of the two different stable isotope-labeled tyrosine variants. The corresponding photolysis spectra and the comparison with unlabeled Ras are shown in Fig. 5C and D. Indeed, we found changes just in the region of the two tyrosine bands. A double band in the unlabeled spectrum at 1523 cm^{-1} and 1514 cm^{-1} is shifted towards 1488 cm^{-1} and 1478 cm^{-1} by the $^{13}C_9{}^{15}N$-tyrosine labeling (Fig. 5C) and towards 1429 cm^{-1} and 1412 cm^{-1} by the ring-$^2H_4$-tyrosine labeling (Fig. 5D). Furthermore, a negative band at 1240 cm^{-1} is shifted to 1212 cm^{-1} by the $^{13}C_9{}^{15}N$-tyrosine labeling and to 1216 cm^{-1} by the ring-$^2H_4$-tyrosine labeling.

3.5. Implications for the reaction mechanism

In our example we assigned the tyrosine vibration of Ras (Fig. 5A–D) and consequently information on the reaction mechanism of the GTPase reaction can now be gained. We firstly examine the negative bands, which are from the Ras·cagedGTP state. Here, the tyrosine bands for unlabeled Ras are at 1514 cm^{-1} and 1240 cm^{-1}. Thus, the band positions are very similar to the isolated amino acid at pH 1 as shown in Fig. 5B but clearly deviate from the bands at pH 12. One can therefore conclude that the observed tyrosine is protonated and the environment of this group is hydrophilic. This assignment is also supported by the observed isotopic shifts.

The positive bands in Fig. 5C and D are bands from the Ras·GTP state after photolysis. Here the 1514 cm^{-1} band shifted towards 1523 cm^{-1} indicating that the tyrosine is still protonated but with less hydrogen bonding in a more buried environment. In the hydrolysis difference spectrum (data not shown) we observed just the opposite band pattern, the 1523 cm^{-1} band disappeared and the original 1514 cm^{-1} was formed again. Thus this band behaves just like the threonine band which we recently assigned as a marker band for the switch I movement of Ras [34] and we can therefore associate the 1523 cm^{-1} absorption with the ‘‘on’’ state and the 1514 cm^{-1} absorption with the ‘‘off’’ state of Ras.

One drawback of the employed metabolic labeling approach is that it is not site-specific. Using difference spectroscopy this is often not a problem, because only the bands from groups involved in the observed reaction are seen. Since this is often only the case for one specific residue of the given type, the assignment is in many cases unambiguous. If this is not the case, a combination of site-directed mutagenesis and isotopic labeling is necessary.

The above discussed examples demonstrate the capabilities of this method which we also successfully applied to the proteins NF1 [35], Rap and RapGAP [36]. Publications on the important catalytic residues of GTPase activating proteins, the arginine-finger of Ras and the asparagine-thumb of Rap are in preparation.

4. Conclusion

Group-specific labeling of proteins is a powerful method for assigning bands in FTIR spectroscopy. The latter is a prerequisite to resolve reaction mechanisms in proteins at the atomic level by vibrational spectroscopy. Since the metabolic incorporation of isotope labels into target proteins can often be low in both efficiency and specificity we advocate the use of ESI-MS of proteolytic peptides to independently assess the correct and efficient labeling of proteins using stable isotopes as a fast and accurate method which is superior to conventional methods, e.g. the determination by radiotracers. The use of adequately labeled proteins then facilitates the accurate assignment of bands in FTIR spectra.

Acknowledgments

We thank the Genetic Stock Center (Yale University) and the NIG collection (NBRP (NIG, Japan): *E. coli*) for the auxotroph *E. coli* strains, Alfred Wittinghofer for the plasmids for Ras, Ran and Rna1p, Romano Hebeler and Heike Piechura for expert technical assistance, and Jon Barbour for critically reading the manuscript. This work was supported by the Ruhr-University School funded by the DFG in the framework of the excellence initiative as well as by the Deutsche Forschungsgemeinschaft within the SFB 642 and the Bundesministerium für Bildung und Forschung.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.vibspec.2007.11.003.

References

[1] C. Kötting, K. Gerwert, Chem. Phys. Chem. 6 (2005) 881.
[2] J. Heberle, Biochim. Biophys. Acta 1458 (2000) 135.
[3] W. Mäntele, Trends Biochem. Sci. 18 (1993) 197.
[4] R. Vogel, F. Siebert, Curr. Opin. Chem. Biol. 4 (2000) 518.
[5] D.C. Muchmore, L.P. McIntosh, C.B. Russell, D.E. Anderson, F.W. Dahlquist, Methods Enzymol. 177 (1989) 44.
[6] L.-Y. Lian, D.A. Middleton, Progr. Nucl. Magn. Reson. Spectrosc. 39 (2001) 171.

[7] L. Reitzer, 2nd ed., Encyclopedia of Microbiology, vol. 1, 2000, p. 134.
[8] H. Kandori, N. Kinoshita, Y. Yamazaki, A. Maeda, Y. Shichida, R. Needleman, J.K. Lanyi, M. Bizounok, J. Herzfeld, J. Raap, J. Lugtenburg, Proc. Natl. Acad. Sci. U.S.A. 97 (2000) 4643.
[9] Y. Xiao, M.S. Hutson, M. Belenky, J. Herzfeld, M.S. Braiman, Biochemistry 43 (2004) 12809.
[10] F. DeLange, C.H.W. Klaassen, S.E. Wallace-Williams, P.H.M. Bovee-Geurts, X.-M. Liu, W.J. DeGrip, K.J. Rothschild, J. Biol. Chem. 273 (1998) 23735.
[11] F. MacMillan, A. Kannt, J. Behr, T. Prisner, H. Michel, Biochemistry 38 (1999) 9179.
[12] S. Sechi, B.T. Chait, Anal. Chem. 70 (1998) 5150.
[13] J. Tucker, G. Sczakiel, J. Feuerstein, J. John, R.S. Goody, A. Wittinghofer, EMBO J. 5 (1986) 1351.
[14] F.W. Studier, B.A. Moffatt, J. Mol. Biol. 189 (1986) 113.
[15] W.B. Wood, J. Mol. Biol. 16 (1966) 118.
[16] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, New York, 1989.
[17] J. Kuhlmann, I. Macara, A. Wittinghofer, Biochemistry 36 (1997) 12027.
[18] R.C. Hillig, L. Renault, I.R. Vetter, T.I.V. Drell, A. Wittinghofer, J. Becker, Mol. Cell 3 (1999) 781.
[19] S. Ueno-Nishio, K.C. Backman, B. Magasanik, J. Bacteriol. 153 (1983) 1247.
[20] H. Cedar, J.H. Schwartz, J. Biol. Chem. 244 (1969) 4112.
[21] K. Von Meyenburg, F.G. Hansen, L.D. Nielsen, E. Riise, Mol. Gen. Genet. 160 (1978) 287.
[22] D.M. LeMaster, F.M. Richards, Biochemistry 27 (1988) 142.
[23] U.K. Laemmli, Nature (London, United Kingdom) 227 (1970) 680.
[24] M.M. Bradford, Anal. Biochem. 72 (1976) 248.
[25] H. Schaefer, J.P. Chervet, C. Bunse, C. Joppich, H.E. Meyer, K. Marcus, Proteomics 4 (2004) 2541.
[26] M.W. Senko, Isopro 2.1, National High Magnetic Field Laboratory, Tallahassee, FL, USA.
[27] C. Allin, K. Gerwert, Biochemistry 40 (2001) 3037.
[28] C.-H. Park, R.S. Givens, J. Am. Chem. Soc. 119 (1997) 2453.
[29] J. John, R. Sohmen, J. Feuerstein, R. Linke, A. Wittinghofer, R.S. Goody, Biochemistry 29 (1990) 6058.
[30] V. Cepus, A.J. Scheidig, R.S. Goody, K. Gerwert, Biochemistry 37 (1998) 10263.
[31] K. Gerwert, G. Souvignier, B. Hess, Proc. Natl. Acad. Sci. U.S.A. 87 (1990) 9774.
[32] B. Hessling, G. Souvignier, K. Gerwert, Biophys. J. 65 (1993) 1929.
[33] A. Barth, C. Zscherp, Quart. Rev. Biophys. 35 (2002) 369.
[34] C. Kötting, A. Kallenbach, Y. Suveyzdis, C. Eichholz, K. Gerwert, ChemBioChem 8 (2007) 781.
[35] M.R. Ahmadian, U. Hoffmann, R.S. Goody, A. Wittinghofer, Biochemistry 36 (1997) 4535.
[36] P.P. Chakrabarti, O. Daumke, Y. Suveyzdis, C. Kötting, K. Gerwert, A. Wittinghofer, J. Mol. Biol. 367 (2007) 983.