Targeted expression and purification of fluorine labelled cold shock protein B by using an auxotrophic strategy

Hannah Welte\textsuperscript{a,b}, Michael Kovermann\textsuperscript{a,b,∗}

\textsuperscript{a} Department of Chemistry, University Konstanz, DE-78464 Konstanz, Germany
\textsuperscript{b} Konstanzer Online-Publikations-System (KOPS) URL: http://nbn-resolving.de/urn:nbn:de:bsz:352-2-r8ofamx2vijz9

\textbf{A B S T R A C T}

High resolution NMR spectroscopy is a seminal method in modern structural biology to obtain insights into proteins’ structure, dynamics and function at dilute condition as well as in a cell-like environment or even intracellularly. Usually, $^1$H, $^{15}$N or $^{13}$C nuclei are predominantly used for the characterization of the protein of interest. These measurements are limited due to the wealth of chemical shifts and background signals arising from all molecules present in the NMR test tube. On top of that, the protein under study has to be isotopically enriched in nitrogen and/or carbon nuclei enabling to overcome the inherently low natural abundance of $^{13}$C and $^{15}$N NMR active isotopes. In this way switching to $^{19}$F NMR spectroscopy strongly reduces the total amount of signals seen in an NMR spectrum as it turns off background signals and is for this reason extremely attractive for highly-resolved investigations of proteins performance measured directly in cells or in a cell-like environment.

Here we show the effective expression and purification of cold shock protein B from \textit{Bacillus subtilis} (BcSpB) using fluorine labelled phenylalanine or fluorine labelled tryptophan residues. We reveal that fluorine labelled BcSpB represents the same fold on a secondary as tertiary level as seen for the wild type protein independent of the labelling position illuminating the soft character of fluorine insertion. This experimental setup of targeted fluorine labelling sets a profound ground for a broad range of highly-resolved $^{19}$F NMR applications to be performed in a complex cellular environment.

\section{1. Introduction}

Three homologous cold shock proteins (CSP) are found in \textit{Bacillus subtilis}: CspB, CspC and CspD sharing a sequence identity of $> 72\%$ \cite{1}. Next to \textit{Bacillus subtilis}, CSPs were also found in \textit{Escherichia coli}, \textit{Thermotoga maritima} and \textit{Aquifex aeolicus} \cite{1}. Cold shock proteins are produced as a response to stress induced by a fast reduction of ambient temperature. The abrupt reduction of temperature in a range of 37 °C down to 15 °C or 10 °C is called cold shock \cite{2}. During the acclimatization phase cold shock proteins are intracellularly produced \cite{3}. The cold shock protein B from \textit{Bacillus subtilis} is one of the best characterized cold shock proteins resembling a solid model for the cold shock domain fold \cite{4-7}. BcSpB owns a molecular mass of 7.3 kDa, comprises five beta sheets leading to a prominent beta barrel fold \cite{8}. Notably, seven out of 67 residues are phenylalanines making it to an ideal target for replacing natural phenylalanine with a fluorine labelled variant (Fig. S1) \cite{9}. BcSpB shows a low overall thermodynamic stability of about $\Delta G^\circ = 10 \text{kJ mol}^{-1}$ at ambient temperature \cite{10}, but extremely fast unfolding, $k_u$, and refolding, $k_f$, rate constants of about $k_u = 36 \text{s}^{-1}$ and $k_f = 690 \text{s}^{-1}$, respectively \cite{5}. Both the equilibrium folding-to-unfolding transition and the folding kinetic analysis can be described assuming a two-state folding model comprising the unfolded ensemble and the native state. With this, the relatively low molecular mass and the beneficial thermodynamic and kinetic properties makes the cold shock protein B an ideal candidate to perform protein folding studies not only in vitro but also for conditions found in a living cell. There exist various methods in structural biology like x-ray diffraction, fluorescence or circular dichroism spectroscopy, cryo-electron microscopy or mass spectrometry \cite{1,6} to obtain spatially resolved information about structure, dynamics and function of a protein. However, high resolution nuclear magnetic resonance (NMR) spectroscopy is one of the most powerful techniques getting information on an atomic level \cite{11,12}. This experimental technique enables to obtain precise values for kinetic and thermodynamic parameters like folding and unfolding rate constants, change in enthalpy $\Delta H$ and change in thermal capacity $\Delta C_p$ \cite{13,14} beside the “traditional” acquisition of...
chemical shifts only. Most protein studies using high-resolution NMR spectroscopy focus on \(^1\)H, \(^13\)C or \(^15\)N nuclei. Due to the inherent signal overlap in those NMR spectra, a detailed analysis is limited to proteins being in a range of a molecular mass of up to about 25 kDa or being present in a background free environment lacking other disturbing proton (or carbon and nitrogen) nuclei. Therefore, we focus here on the \(^19\)F nucleus which combines several advantages: (i) no natural occurrence within proteins, allowing investigations of highly selective labelled samples, (ii) a natural abundance of 100%, which allows to perform experiments on a protein present at a relatively low concentration, (iii) a large range of chemical shifts of up to 1200 ppm [15] and (iv) a high sensitivity for potential changes in the local van der Waals environment [13]. To produce a fluorine labelled protein, an artificially labelled aromatic amino acid like phenylalanine, tryptophan or tyrosine will be added to the medium which is used for protein expression. Thereby, the synthesis of the natural aromatic amino acid can be eliminated in the cells by the addition of glyphosate, which inhibits the Shikimate synthetic pathway responsible for the generation of the aromatic amino acids [13,16]. Contrary, another possible method to introduce \(^19\)F labelled aromatic amino acids to a protein is to simply add the labelled amino acid to the common expression medium [16-18]. However, both strategies potentially lead to an incomplete incorporation of the artificial amino acid into the protein finally resulting in a heterogeneous sample possessing labelled and non-labelled molecules. Experimentally, we have not been successful in using glyphosate applied on BsCspB as no band referring to the molecular mass of BsCspB has been detected in the corresponding SDS page analysis (Fig. S2). Consequently, we moved to a third strategy being able to express fluorine labelled BsCspB in a controlled manner. This method relies on cell strains which are auxotrophic regarding natural phenylalanine (Phe) or tryptophan (Trp). This setup offers the high potential of complete and targeted incorporation of \(^19\)F-Phe or \(^19\)F-Trp into BsCspB. Here we present the results of this labelling strategy successfully applied to all three possible positions in Trp and Phe residues (Fig. 1). Additionally, we demonstrate that adding of fluorine into the aromatic ring of the phenylalanine or tryptophan residue does not disturb the fold of BsCspB neither on a secondary nor on a tertiary level. Consequently, such labelled BsCspB proteins permit to perform reliable \(^19\)F NMR spectroscopic studies not only in vitro but also in complex experimental environments.

2. Materials and methods

2.1. Construction of pET24a-CspB

The pET11a-CspB sequence was verified by DNA sequencing (Fig. S3). A polymerase chain reaction (PCR) was performed to obtain the BsCspB sequence. Therefore, the forward and reverse primers 5’-GCCGAAACAAGGGCTCAT-3’ and 3’-TGTGATGAGCCCCGATCTTC-5’ with restriction sites for BamHI-HF and NdeI have been used. PCR using a Taq polymerase (provided by Erika Isono’s laboratory, University of Konstanz) was performed at a temperature of 95 °C for a time of 3 min and then at 95 °C for t = 30 s, at 57 °C for t = 30 s and at 72 °C for t = 20 s for a total of 30 cycles, followed by a T = 72 °C for t = 4 min. The PCR product and the pET24a vector was digested with the restriction endonucleases NdeI and BamHI-HF after purification (GeneJET PCR purification Kit, Thermo scientific) and subsequently purified again. Ligation was done for t = 10 min at room temperature and then over night at T = 4 °C (T4 Ligase, New England BioLabs). The new plasmid pET24a-CspB was transformed into E. coli DH5α and placed on kanamycin (50 μg/ml) containing plates. The transformation procedure is completely described in the Supporting Information. Positive clones were confirmed by colony-PCR using the primers 5’-GCCGAAACAAGGGCTCAT-3’ and 3’-TGTGATGAGCCCCGATCTTC-5’. The PCR product was detected by using 1.5% agarose gel electrophoresis (Fig. 2). Midori Green Direct (Nippon Genetics) was used for staining of the gel. The plasmid isolated from these clones has been analysed by DNA sequencing (Fig. 3). The plasmids pET24a-CspB and pAR1219 were finally transformed into the auxotrophic cell strains CAG 18455 7371 (tryptophan auxotrophy) and DSMZ 12779 (phenylalanine auxotrophy) for further over-expression. The plasmid for the T7 polymerase (pAR1219) was obtained from Sigma Aldrich and analysed by DNA sequencing (Fig. S4).

2.2. Expression and purification of \(^19\)F-Trp-labelled cold shock protein B

The tryptophan auxotrophic E. coli cells (CAG 18455 7371) transformed with pAR1219 and pET24a-CspB were grown overnight in LB
medium (shaking speed of 130 rpm) containing concentrations of 50 mg/l kanamycin and 100 mg/l carbenicilin at $T = 37^\circ C$. $V = 100$ ml of this subculture were used for $V = 11$ of the main culture. For the main culture, M9 minimal medium [4] containing 50 mg/l tryptophan, 50 ml/l kanamycin and 100 mg/l carbenicilin was used. The cells were cultured (shaking speed of 130 rpm) at $T = 37^\circ C$ until an OD value of $0.6-0.8$ was reached. Then the medium was changed to M9 minimal medium (see Supporting Information) containing 50 mg/l $^{19}$F labelled tryptophan instead of natural tryptophan. See the Supporting Information for the composition of the M9 medium which has been used in this study. Protein expression was induced by the addition of isopropyl-$\beta$-D-thiogalactopyranoside (IPTG) having a final concentration of $c = 1$ mM after $t = 30$ min. The cells were further incubated (shaking speed of 130 rounds/min) at $T = 25^\circ C$ overnight.

The CAG cells were harvested by centrifugation at 7000 rpm for $t = 20$ min at $T = 4^\circ C$. The pellet was resuspended in $V = 5$ ml lysis buffer (50 mM TRIS, pH 8.0) per $m = 1$ g of pellet and the cells were disrupted by sonication (sonifier W-250, Branson) for 400 s/100 ml of cell-dispersion in cycles of $t = 2$ s work and $t = 3$ s break interval on ice. The cell lysate was centrifuged at 15000 rpm for $t = 30$ min and the supernatant was purified. An anion exchange chromatography (hand-made column with 100 ml Fractogel EMD TMAE (M) (Merck)), a hydrophobic exchange chromatography (HiPrep Butyl FF 16/10, 20 ml, GE Healthcare) and a size exclusion chromatography (Superdex 75 pg HiLoad 16/600, GE Healthcare) have been used for protein purification using an ÄKTA pure system (GE Healthcare).

The purified protein was subsequently analysed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) [19] stained with Coomassie brilliant blue R-250, mass spectrometry, fluorescence and high-resolution NMR spectroscopy.

The concentration of purified fluorine labelled BsCspB was determined applying UV/Vis spectroscopy (Agilent 8453 UV–visible Spectroscopy System, Agilent Technologies) using extinction coefficients of $\varepsilon_{280} = 2705$ M$^{-1}$ cm$^{-1}$ (4-19F-Trp BsCspB), $\varepsilon_{280} = 2887$ M$^{-1}$ cm$^{-1}$ (5-19F-Trp BsCspB), $\varepsilon_{280} = 2575$ M$^{-1}$ cm$^{-1}$ (6-19F-Trp BsCspB) and $\varepsilon_{280} = 5800$ M$^{-1}$ cm$^{-1}$ (2-19F-Phe BsCspB, 3-19F-Phe BsCspB and 4-19F-Phe BsCspB). We have named all six $^{19}$F labelled protein samples (2-19F-Phe, 3-19F-Phe, 4-19F-Phe, 4-19F-Trp, 5-19F-Trp, 6-19F-Trp) according to the position of fluorine present in the aromatic ring of phenylalanine and tryptophan, respectively (Fig. 1).

2.3. Expression and purification of $^{19}$F-Phe-labelled cold shock protein B

The expression and purification of $^{19}$F-Phe-labelled cold shock protein B was done as for the $^{19}$F-Trp-labelled variant (see above). However, New Minimal Medium [20] (NMM) was used instead of M9 minimal medium. NMM contains a concentration of $c = 50$ mg/l of all natural abundant amino acids (see Supporting Information for the composition of NMM which has been used in this study). The phenylalanine auxotrophic cell strain DSMZ 12779 was used for the expression of $^{19}$F-Phe-labelled BsCspB.

2.4. Mass spectrometry

Mass spectrometry was performed using a protein concentration of $c = 4 \mu$M loaded on a Thermo LTQ Orbitrap Discovery.

2.5. NMR spectroscopy

All NMR spectra were measured in 20 mM sodiumoctylate buffer containing 90% H2O and 10% D2O at pH = 7. The 5-19F-Trp BsCspB variant was additionally labelled with $^{15}$N using $^{15}$NH4Cl, as a single $^{15}$N source. One-dimensional $^1$H/$^{19}$F and two-dimensional $^1$H–$^{15}$N HSQC spectra have been acquired by using a protein concentration of $c = 650 \mu$M. One-dimensional $^1$H and $^{19}$F spectra were measured for all six possible variants of Trp- or Phe-labelled BsCspB. All one-dimensional $^1$H as well as two-dimensional $^1$H–$^{15}$N HSQC spectra were collected on a 600 MHz Bruker Avance III NMR spectrometer equipped with a TCI cryogenically cooled probe at $T = 298$ K sample temperature. All one-dimensional $^{19}$F spectra were collected on a 400 MHz Bruker Avance III HD NMR spectrometer equipped with a TBO probe operated at $T = 300$ K. All $^1$H spectra have been directly referenced by using the proton resonance frequency of Trimethylsilylproanoic acid (TMS). This value has been used for indirect referencing of $^{19}$F spectra. Chemical shift perturbations, $\Delta\omega$, obtained for 5-19F-Trp BsCspB (5F) compared to wild type BsCspB (wt) were calculated as follows: $\Delta\omega = [(\omega_{288}^{5F}(wt) - \omega_{288}^{5F}(SF))^{2} + 1/25 (\omega_{288}^{5F}(wt) - \omega_{288}^{5F}(SF))^{2}]^{1/2}$. NMR data processing was performed on TOPSPIN 4.0.3 software (Bruker Biospin, Germany).

3. Results and discussion

3.1. Construction of pET24a-CspB

Two plasmids possessing different antibiotic resistance are a prerequisite to obtain $^{19}$F labelled BsCspB using auxotrophic strains. One out of these two is used for the expression of the T7-polymerase which is needed for the reading of the BsCspB sequence. The second one carries the sequence of BsCspB. The sequence for BsCspB has been cut from the pET11a-CspB plasmid and cloned on a pET24a vector for the newly designed BsCspB plasmid. Note, that this strategy is a prerequisite to generate a difference in antibiotic resistance.

The BsCspB sequence has been cut by using PCR methodology. The primers used here contained cutting regions for the restriction endonucleases Ndel and BamHI-HF. Therefore, both the PCR product and the vector pET24a were digested with these restriction enzymes and subsequently ligated to the pET24a-CspB plasmid. After transforming into E. coli DH5a cells and selection on LB medium plates, six colonies were picked and verified by colony PCR methodology. The results obtained from PCR were detected by agarose gel electrophoresis (Fig. 2). It can be seen that all six clones which have been selected were positive and correspond to a theoretical size of 533 base pairs (bp) as expected for the PCR product (Fig. 2). In comparison, the length of the pET24a vector obtained via PCR has been determined to 360 bp.

Further sequencing analysis has proved that the gene coding for BsCspB has been correctly inserted into the vector pET24a (Fig. 3).

3.2. Expression and purification of $^{19}$F labelled BsCspB

BsCspB molecules which have been fluorine labelled at all six different positions tryptophan and phenylalanine inherently offer (Fig. 1) were successfully expressed by growing the auxotrophic bacteria in M9 minimal media or NMM as described above. The addition of $c = 1$ mM IPTG induces the expression of BsCspB. Variations in expression time and temperature showed that incubation overnight at $T = 25^\circ C$ is the most promising condition. After performing cell lysis and centrifugation, the target protein was found in the supernatant. Anion exchange chromatography has been used as the first step in the purification procedure (Fig. 4A, D). The fractions 12–33 for 5-19F-Trp BsCspB have been collected and this merged sample has been used for the second purification step of hydrophobic exchange chromatography which showed significant improvement of the purity of the desired sample (Fig. 4B, E). The fractions 6–17 for 5-19F-Trp BsCspB have been collected to perform a final size exclusion chromatography as third purification step leading to high purity of this fluorine labelled BsCspB variant (Fig. 4C, F).

The elution profiles and SDS page analyses for the remaining five $^{19}$F labelled variants of BsCspB are shown in Figs. S5–S9. Note that yields for all $^{19}$F labelled variants of BsCspB obtained here are in the same order of magnitude compared to the yield obtained for wild type BsCspB (Table S1).

88

H. Wele and M. Kovermann

Protein Expression and Purification 157 (2019) 86–91
3.3. Mass spectrometry of $^{19}$F labelled BsCspB

Mass spectrometry has been applied to confirm $^{19}$F labelling in 5-$^{19}$F-Trp BsCspB quantitatively. The corresponding ESI MS spectrum shows a peak at 7382.3 Da which fits excellently to the theoretical mass of 7382 Da as expected for singly $^{19}$F labelled BsCspB (Fig. 4). The deconvolution of peaks results in a total amount of 5-$^{19}$F-Trp BsCspB of about 80% whereas the amount of wild type BsCspB accounts with 3% corresponding to a molar mass of 7365 Da (Table S2). The remaining part can be attributed to fragmentation (about 5%) and agglomeration processes (about 7%) as to oxidation (about 8%).

3.4. NMR spectroscopy of $^{19}$F labelled BsCspB

One-dimensional $^1$H as well as two-dimensional heteronuclear $^1$H–$^{15}$N HSQC spectra obtained for the $^{15}$N labelled 5-$^{19}$F-Trp BsCspB variant have been acquired to evaluate the overall protein fold of fluorine labelled BsCspB and being able to compare with wild type properties. These comparisons between one-dimensional proton as well as two-dimensional $^1$H–$^{15}$N HSQC spectra indicate no significant structural difference between wild type and the 5-$^{19}$F-Trp variant of BsCspB (Fig. 6). Note the broad dispersion seen for amide protons comprising 5-$^{19}$F-Trp BsCspB which is fully conserved compared to wild type BsCspB. Additionally, the most high-field shifted signals representing methyl groups in 5-$^{19}$F-Trp BsCspB are comparable to wild type BsCspB for both the chemical shift value and the relative signal intensity, respectively. Calculating chemical shift perturbations, $\Delta\omega$, for changes in proton and nitrogen dimensions due to 5-$^{19}$F-Trp labeling in wild type BsCspB results in a mean value of $\Delta\omega = (0.035 \pm 0.032)$ ppm (Fig. S10A). It can be seen that fluorine labelling of the single tryptophan changes the chemical environment of the protein only locally (Fig. S10B). Note the missing $^1$H–$^{15}$N resonance signal for 5-$^{19}$F-Trp BsCspB representing the indole group of wild type BsCspB in the two-dimensional HSQC spectrum (Fig. 6). This missing signal refers to the expression route used here for the generation of the fluorine labelled protein. As artificially fluorine labelled tryptophan possesses $^{15}$N nuclei with natural abundance only, no signal for $^1$H–$^{15}$N correlation can be seen. Qualitatively, the two-dimensional $^1$H–$^{15}$N HSQC spectrum acquired for 5-$^{19}$F-Trp BsCspB indicates a solid purity of the protein sample by representing a single set of resonance signals.

The one-dimensional $^1$H spectra for the remaining five fluorine labelled variants of BsCspB are shown in Figures S11A, S12A, S13A, S14A and S15A illuminating again wild type like structural properties for these $^{19}$F labelled BsCspB variants independent of the labelling site.

Beside proton and nitrogen dimensions, one-dimensional $^{19}$F NMR spectra have been recorded. The spectrum of 5-$^{19}$F-Trp BsCspB shows one resonance signal at a chemical shift, $\omega$, of $\omega = -122.38$ ppm for the fluorine present at position 5 in the tryptophan ring (Fig. 7). Additionally, the one-dimensional $^{19}$F NMR spectrum of isolated 5-$^{19}$F-Trp has been acquired as well. The resonance signal present in this spectrum is shifted by about $\Delta\omega = -2.3$ ppm to $\omega = -124.68$ ppm indicating the pronounced different chemical environment of fluorine by comparing the free state of 5-$^{19}$F-Trp to the incorporation of 5-$^{19}$F-Trp into BsCspB. To go along these lines, the increase in linewidth FWHM...
Comparing free 5-^{19}F-Trp with 5-^{19}F-Trp_{BsCspB} supports the successful incorporation of the fluorine labelled tryptophan into BsCspB even further. For comparison, natively folded fluorine labelled human transthyretin owning a molecular mass twice as much as BsCspB possesses a linewidth of the tryptophan resonance of about 100 Hz in the {^{19}F} dimension acquired at \( B_0 = 14.1 \) T [17].

The one-dimensional {^{19}F} spectra for the remaining five fluorine labelled variants of BsCspB are shown in Figures S11B, S12B, S13B, S14B and S15B.

4. Conclusion

In summary, we present a straightforward method for the reliable expression and purification procedure used for fluorine labelling of a protein using auxotrophic strains. This method has been successfully applied on BsCspB here which has been fluorine labelled at all six possible positions tryptophan and phenylalanine residues inherently offer. We show that fluorine labelling at these six positions lead to a soft and local change according to the NMR chemical shift scale but preserve the overall protein fold as seen for wild type BsCspB. This finding
is supported by the acquisition of two-dimensional NMR data on a doubly labelled $^{15}$N/$^{19}$F BtCspB sample. Although $^{19}$F labelling of protein samples is limited to specific sites, fluorine labelled protein variants offer the high potential to study protein folding, dynamics and functions directly in a cell or in a cell-like environment by applying $^{19}$F NMR spectroscopy. The main advantage of the method for the preparation of fluorine labelled proteins as presented here lies in using of auxotrophic bacterial strains. This procedure leads to an targeted incorporation of the fluorine label in a controlled manner compared to methods relying either on adding amounts of glyphosate or on adding an excess of the labelled amino acid to the conventional expression medium [13,16–18].

Conflicts of interest

The authors have no competing interests to declare.

Acknowledgement

We thank Dietmar Funck, Caroline Haupt and Iljas Müller for support at the initial state of this project. Financial support from the Juniorprofessurenprogramm handled by the Baden-Württemberg Stiftung (grant ID: 647/16) is gratefully acknowledged.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.pep.2019.02.006.

References

[16] P.B. Crowley, C. Kyne, W.B. Monteith, Simple and inexpensive incorporation of fluorine labelled proteins as presented here lies in using of auxotrophic bacterial strains. This procedure leads to an targeted incorporation of the fluorine label in a controlled manner compared to methods relying either on adding amounts of glyphosate or on adding an excess of the labelled amino acid to the conventional expression medium [13,16–18].