Biomol NMR Assign (2016) 10:13–23
DOI 10.1007/s12104-015-9629-8

ARTICLE

# Solid-state NMR sequential assignments of the N-terminal domain of *Hp*DnaB helicase

**Thomas Wiegand[1] · Carole Gardiennet[2] · Francesco Ravotti[1] · Alexandre Bazin[2] · Britta Kunert[2] · Denis Lacabanne[2] · Riccardo Cadalbert[1] · Peter Güntert[1,3] · Laurent Terradot[2] · Anja Böckmann[2] · Beat H. Meier[1]**

Received: 29 March 2015 / Accepted: 11 August 2015 / Published online: 18 August 2015


**Abstract** We present solid-state NMR assignments of the N-terminal domain of the DnaB helicase from *Helicobacter pylori* (153 residues) in its microcrystalline form. We use a sequential resonance assignment strategy based on three-dimensional NMR experiments. The resonance assignments obtained are compared with automated resonance assignments computed with the ssFLYA algorithm. An analysis of the $^{13}C$ secondary chemical shifts determines the position of the secondary structure elements in this α-helical protein.

**Keywords** *Hp*DnaB · Assignments · Solid-state NMR · Secondary chemical shifts · ssFLYA

T. Wiegand and C. Gardiennet have equally contributed to this work.

✉ Laurent Terradot
laurent.terradot@ibcp.fr

✉ Anja Böckmann
a.bockmann@ibcp.fr

✉ Beat H. Meier
beme@ethz.ch

[1] Laboratorium für Physikalische Chemie, ETH Zürich, Vladimir-Prelog-Weg 2, 8093 Zurich, Switzerland

[2] Institut de Biologie et Chemie des Protéines, Bases Moléculaires et Structurales des Systèmes Infectieux, Labex Ecofect, UMR 5086 CNRS, Université de Lyon, 7 passage du Vercor, 69007 Lyon, France

[3] Institute of Biophysical Chemistry, Center for Biomolecular Magnetic Resonance, Goethe University Frankfurt am Main, Max-von-Laue-Str. 9, 60438 Frankfurt am Main, Germany

## Biological context

DnaB helicases are bacterial ATP-driven enzymes which unwind double-stranded DNA in the presence of ATP during the fork movement in 5′–3′ direction in DNA replication (LeBowitz and McMacken 1986). Structurally, DnaB is a two-domain helicase with an amino-terminal domain and a carboxy-terminal domain separated by a linker region. The full-length protein forms ring-shaped hexameric assemblies which encircle single-stranded DNA (LeBowitz and McMacken 1986). The C-terminal domain supports the ATPase activity and is involved in the ring formation, while the N-terminal domain is forming an α-helical globule which has the function to activate the helicase. A fundamental step in DNA replication is the strand synthesis which is initiated by an interaction between the N-terminal domain of the helicase DnaB and the primase DnaG (Corn and Berger 2006).

This work focuses on the investigation of the N-terminus of DnaB extracted from *Helicobacter pylori* (*H. pylori*), a gram-negative microaerophilic spiral shaped bacterium which is, with a worldwide prevalence of approximately 50 %, responsible for the most common chronic bacterial infections, such as gastric ulcer diseases and gastric adenocarcinoma (Parsonnet et al. 1991; Peterson 1991). The replication system in *H. pylori* exhibits significant differences compared to other microorganisms which were investigated in detail, e.g. *Escherichia coli*. The most relevant differences are the absence of the *recF* gene, the presence of the dnaA gene ~600 kb away from the dnaN-gyrB genes and most importantly the absence of the dnaC gene (Soni et al. 2003). In *E. coli*, DnaC is essential for loading DnaB helicase at *oriC*, which is the origin of the chromosomal DNA replication, whereas in the case of *H. pylori* *Hp*DnaB itself is able to take over the

DnaC function (Soni et al. 2005). The crystal structure of the C-terminal domain (Stelter et al. 2012), as well as of the 121 N-terminal residues of *Hp*DnaB were determined (Kashav et al. 2009). The latter one was found to consist of a dimer and two further degradation peptide fragments in the asymmetric unit (Kashav et al. 2009). This manuscript presents the solid-state NMR spectroscopic investigation of the N-terminus of *Hp*DnaB as a step towards an integrated structural biology approach aiming at a detailed structural description of the full-length protein. In this context, solid-state NMR studies on a sedimented sample of the full-length protein have been described (Gardiennet et al. 2012). Previous biochemical investigations suggest that the N-terminal domain and the linker region play an important role in multimerisation, quaternary state transition and activity of *Hp*DnaB (Kashav et al. 2009; Nitharwal et al. 2007).

## Methods and experiments

### Protein expression and purification, sample preparation

The DNA fragment corresponding to the N-terminal domain (NTD) of the *H. pylori* helicase DnaB (strain 26695) was amplified by PCR (forward 5′-caccatggatcatttaaagcatttgcag-3′ and reverse 5′-gcaccatagaaggctttaggaattag-3′) from genomic DNA and inserted into the plasmid pET151/DTopo (Invitrogen™). The resulting vector was introduced into *E. coli* BL21(DE3) cells (One Shot® BL21 Star™ (DE3) Chemically Competent *E. coli,* Invitrogen™) and protein overexpression was performed in minimal M9 medium (Studier 2005) containing D-[U-$^{13}$C]glucose 2 g $L^{-1}$ (Cambridge Isotope Laboratories, Inc. CLM-1396-PK) and $^{15}NH_4Cl$ 2 g $L^{-1}$ (Sigma-Aldrich® 299251) as the only nitrogen and carbon sources. After cell lysis by a microfluidization process, $^{13}$C-$^{15}$N-*Hp*DnaB-NTD was purified by $Ni^{2+}$-agarose affinity chromatography (Qiagen™). The pseudo-affinity tag was subsequently cleaved with the TEV (Tobacco Etch Virus) protease by dialysis. Six additional residues of the tag remain in the sequence (see Fig. 1).

For crystallization, $^{13}$C-$^{15}$N-*Hp*DnaB-NTD was concentrated to 23.5 mg/ml using a centrifugal concentrator with a 10 kDa cut-off (Vivaspin® 20 VS2001 Sartorius), and the buffer was exchanged during the concentration step with the final buffer (50 mM Tris–HCl pH 6.5, 100 mM NaCl). Crystallization of the protein was performed by mixing an equal volume of protein and crystallization buffer (100 mM HEPES pH 7.0, 0.1 % (m/v) sodium azide, 10 % (v/v) polyethylene glycol 10000) in a nine-well glass plate with 2.3 M NaCl solution in the reservoir.

```
     0
GIDPFT

       10         20         30         40         50         60
MDHLKHLQQL QNIERIVLSG IVLANHKIEE VHSVLEPSDF YYPPNGLFFE IALKLHEEDC

       70         80         90        100        110        120
PIDENFIRQK MPKDKQIKEE DLVAIFAASP IDNIEAYVEE IKNASIKRKL FGLANTIREQ

      130        140        150
ALESAQKSSD ILGAVEREVY ALLNGSTIEG FRN
```

**Fig. 1** Amino-acid sequence of the N-terminus of *Hp*DnaB as extracted from the uniprot database (The UniProt Consortium-Activities at the Universal Protein Resource (UniProt) 2014). At the N-terminal domain a part of the tag (amino acid sequence GIDPFT) is still present after the cleavage of the tag with TEV protease (shown in *grey*). Residues highlighted in *red* are located in α-helices as determined from NMR secondary chemical shifts (see Fig. 9)

Microcrystals were obtained after 1 week at 20 °C and were harvested and centrifuged into the NMR rotor (25,000×*g* during 1 h at 4 °C) using a homemade device (Böckmann et al. 2009).

### Solid-state NMR spectroscopy

Solid-state NMR spectra were acquired at 18.8 and 20.0 T static magnetic field strengths using 3.2 mm Bruker Biospin "E-free" probes (Gor'kov et al. 2007). The MAS spinning frequency was set to 17.5 and 17.0 kHz for the two fields, respectively. The 2D and 3D spectra were processed with the software TOPSPIN (version 3.2, Bruker Biospin) with a shifted (2.0–2.8) squared cosine apodization function and automated baseline correction in the direct dimension. The sample temperature was set to 278 K, for more details of the conducted experiments see Table 1. $^{13}$C and $^{15}$N resonance assignments were obtained by using a previously established assignment strategy based on a sequential walk applying 3D NMR spectra (Habenstein et al. 2011; Schuetz et al. 2010) which were analyzed with the software CcpNmr (Fogh et al. 2002; Stevens et al. 2011; Vranken et al. 2005). Many resonances were assigned applying conventional experiments, such as NCACB, NCACX, NCOCX and CANCO. In case of spectral overlap, it was essential to complement the assignment strategy with relayed NMR experiments, such as NcoCACB, CANcoCA and NcaCBCX. Although those experiments are less sensitive due to four polarization transfer steps, they strongly benefit from a larger spectral dispersion (Schuetz et al. 2010), and the signal/noise ratio achieved under the conditions used here (high field, full rotor, moderate-sized protein) is good. A selective C′–Cα polarization transfer in those experiments was achieved by a modified band-selective homonuclear cross-polarization step (Chevelkov et al. 2013). The spectra used for assignment were all recorded on a single sample, whereas reproducibility was carefully checked by 2D measurements

**Table 1** Overview about experimental parameters of the performed solid-state NMR experiments

| Experiment | DARR | NCA | NCACB | NCACX | NCOCX |
|---|---|---|---|---|---|
| *(a)* | | | | | |
| MAS frequency/kHz | 17.0 | 17.0 | 17.5 | 17.0 | 17.0 |
| Field/T | 20.0 | 20.0 | 18.8 | 20.0 | 20.0 |
| Transfer I | HC-CP | HN-CP | HN-CP | HN-CP | HN-CP |
| $^{1}$H field/kHz | 59.4 | 60.8 | 58.5 | 60.8 | 60.8 |
| X field/kHz | 40.9 | 43.1 | 44.0 | 43.1 | 43.1 |
| Shape | Tangent $^{1}$H | Tangent $^{1}$H | Tangent $^{1}$H | Tangent $^{1}$H | Tangent $^{1}$H |
| $^{13}$C carrier/ppm | 95 | – | – | – | – |
| Time/ms | 0.5 | 0.6 | 0.6 | 0.6 | 0.6 |
| Transfer II | DARR | NC-CP | NC-CP | NC-CP | NC-CP |
| $^{1}$H field/kHz | 17.0 | – | – | – | – |
| $^{13}$C field/kHz | – | 6.1 | 10.0 | 6.1 | 6.1 |
| $^{15}$N field/kHz | – | 10.2 | 27.0 | 10.2 | 10.6 |
| Shape | – | Tangent $^{13}$C | Tangent $^{13}$C | Tangent $^{13}$C | Tangent $^{13}$C |
| Carrier/ppm | 95 | 60 | 56 | 60 | 178 |
| Time/ms | 10 | 6 | 4 | 6 | 4 |
| Transfer 3 | – | – | DREAM | DARR | DARR |
| $^{1}$H field/kHz | – | – | – | 25.1 | 25.1 |
| $^{13}$C field/kHz | – | – | 7.9 | – | – |
| $^{15}$N field/kHz | – | – | – | – | – |
| Shape | – | – | Tangent $^{13}$C | – | – |
| Carrier/ppm | – | – | 56 | 178 | 178 |
| Time/ms | – | – | 4 | 60 | 30 |
| t1 increments | 2000 | 2000 | 80 | 108 | 108 |
| Sweep width (t1)/kHz | 100 | 66.7 | 5.7 | 6.0 | 6.0 |
| Acquisition time (t1)/ms | 10 | 15 | 7.1 | 9.0 | 9.0 |
| t2 increments | 2988 | 3072 | 116 | 136 | 116 |
| Sweep width (t2)/kHz | 100 | 100 | 10.9 | 8.6 | 6.4 |
| Acquisition time (t2)/ms | 14.9 | 15.4 | 5.3 | 8.0 | 9.0 |
| t3 increments | – | – | 1988 | 3072 | 3072 |
| Sweep width (t3)/kHz | – | – | 100 | 100 | 100 |
| Acquisition time (t3)/ms | – | – | 9.9 | 15.4 | 15.4 |
| $^{1}$H Spinal64 (Fung et al. 2000) decoupling power/kHz | 89 | 89 | 90 | 89 | 89 |
| Interscan delay/s | 2.1 | 3 | 2.5 | 2.6 | 2.6 |
| Number of scans | 12 | 8 | 24 | 8 | 8 |
| Measurement time/h | 14 | 13 | 157 | 87 | 74 |

| Experiment | CANCO | NcoCACB | CANcoCA | NcaCBCX | CCC |
|---|---|---|---|---|---|
| *(b)* | | | | | |
| MAS frequency/kHZ | 17.0 | 17.0 | 17.0 | 17.0 | 17.5 |
| Field/T | 20.0 | 20.0 | 20.0 | 20.0 | 18.8 |
| Transfer I | HC-CP | HN-CP | HC-CP | HN-CP | HC-CP |
| $^{1}$H field/kHz | 59.4 | 60.8 | 58.4 | 58.4 | 66.0 |
| X field/kHz | 40.9 | 43.1 | 40.8 | 45.2 | 50.0 |
| Shape | Tangent $^{1}$H | Tangent $^{1}$H | Tangent $^{1}$H | Tangent $^{1}$H | Tangent $^{1}$H |
| $^{13}$C carrier/ppm | 95 | – | 103 | – | 58.5 |
| Time/ms | 0.5 | 0.6 | 0.5 | 0.6 | 0.7 |
| Transfer II | CN-CP | NC-CP | CN-CP | NC-CP | DREAM |

**Table 1** continued

| Experiment | CANCO | NcoCACB | CANcoCA | NcaCBCX | CCC |
|---|---|---|---|---|---|
| $^{1}$H field/kHz | – | – | – | – | – |
| $^{13}$C field/kHz | 6.0 | 6.1 | 6.0 | 5.9 | 5.4 |
| $^{15}$N field/kHz | 20.0 | 10.6 | 11.4 | 11.4 | – |
| Shape | Tangent $^{13}$C | Tangent $^{13}$C | Tangent $^{13}$C | Tangent $^{13}$C | Tangent $^{13}$C |
| $^{13}$C Carrier/ppm | 60 | 178 | 60 | 60 | 56 |
| Time/ms | 4 | 4 | 5.5 | 5.5 | 4 |
| Transfer 3 | NC-CP | Mod. band-selective CP (Chevelkov et al. 2013) | NC-CP | DREAM | DARR |
| $^{1}$H field/kHz | – | – | – | – | 17.5 |
| $^{13}$C field/kHz | 6.0 | 8.5 | 6.0 | 7.6 | – |
| $^{15}$N field/kHz | 20.0 | – | 11.9 | – | – |
| Shape | Tangent $^{13}$C | Tangent $^{13}$C | Tangent $^{13}$C | Tangent $^{13}$C | – |
| $^{13}$C Carrier/ppm | 178 | 176 | 178 | 51 | 40 |
| Time/ms | 4 | | 3.5 | 2 | 80 |
| Transfer 4 | – | DREAM | Mod. band-selective CP (Chevelkov et al. 2013) | DREAM | – |
| $^{1}$H field/kHz | – | – | – | – | – |
| $^{13}$C field/kHz | – | 7.7 | 8.3 | 7.6 | – |
| $^{15}$N field/kHz | – | – | – | – | – |
| Shape | – | Tangent $^{13}$C | Tangent $^{13}$C | Tangent $^{13}$C | – |
| $^{13}$C Carrier/ppm | – | 51 | 178 | 26 | – |
| Time/ms | – | 2.75 | 5 | 4 | – |
| t1 increments | 130 | 62 | 150 | 68 | 172 |
| Sweep width (t1)/kHz | 8.5 | 3.4 | 8.6 | 4.3 | 15.7 |
| Acquisition time (t1)/ms | 7.6 | 9.0 | 8.8 | 7.9 | 5.5 |
| t2 increments | 98 | 154 | 80 | 156 | 172 |
| Sweep width (t2)/kHz | 6.03 | 8.6 | 4.3 | 12.8 | 15.7 |
| Acquisition time (t2)/ms | 8.1 | 9.0 | 9.3 | 6.1 | 5.5 |
| t3 increments | 3072 | 3072 | 3072 | 3072 | 1864 |
| Sweep width (t3)/kHz | 100 | 100 | 100 | 100 | 93.8 |
| Acquisition time (t3)/ms | 15.3 | 15.4 | 15.4 | 15.4 | 9.9 |
| $^{1}$H Spinal64 (Fung et al. 2000) decoupling power/kHz | 89.0 | 89.0 | 91.1 | 91.1 | 90.0 |
| Interscan delay/s | 2.7 | 3 | 3 | 3 | 2.1 |
| Number of scans | 8 | 16 | 16 | 16 | 8 |
| Measurement time/h | 77 | 128 | 162 | 143 | 145 |

on samples from different preparations which yield comparable spectra in all cases. Spectra were also recorded on samples with the His-tag, and showed differences large enough to discard the further use of this sample.

TALOS+ calculations were performed using version 3.8 (Shen et al. 2009). The DSSP algorithm (Kabsch and Sander 1983) was applied using the corresponding web interface (http://www.cmbi.ru.nl/dssp.html) with the 3D atomic coordinates extracted from the pdb file 3GXV (Kashav et al. 2009).

Solid-state FLYA calculations (Schmidt et al. 2013; Schmidt and Güntert 2012) were performed with CYANA version 3.97. The tolerance value for chemical shift matching was set to 0.55 ppm for $^{13}$C and $^{15}$N. The calculations are based on experimental peak lists as obtained from the manual assignment procedure.

## Assignment and data deposition

The solid-state NMR spectra of the N-terminal domain of *Hp*DnaB (residues 1–153) reveal significant spectral overlap as expected for a protein of 153 amino acids, even though also many isolated, well resolved signals are

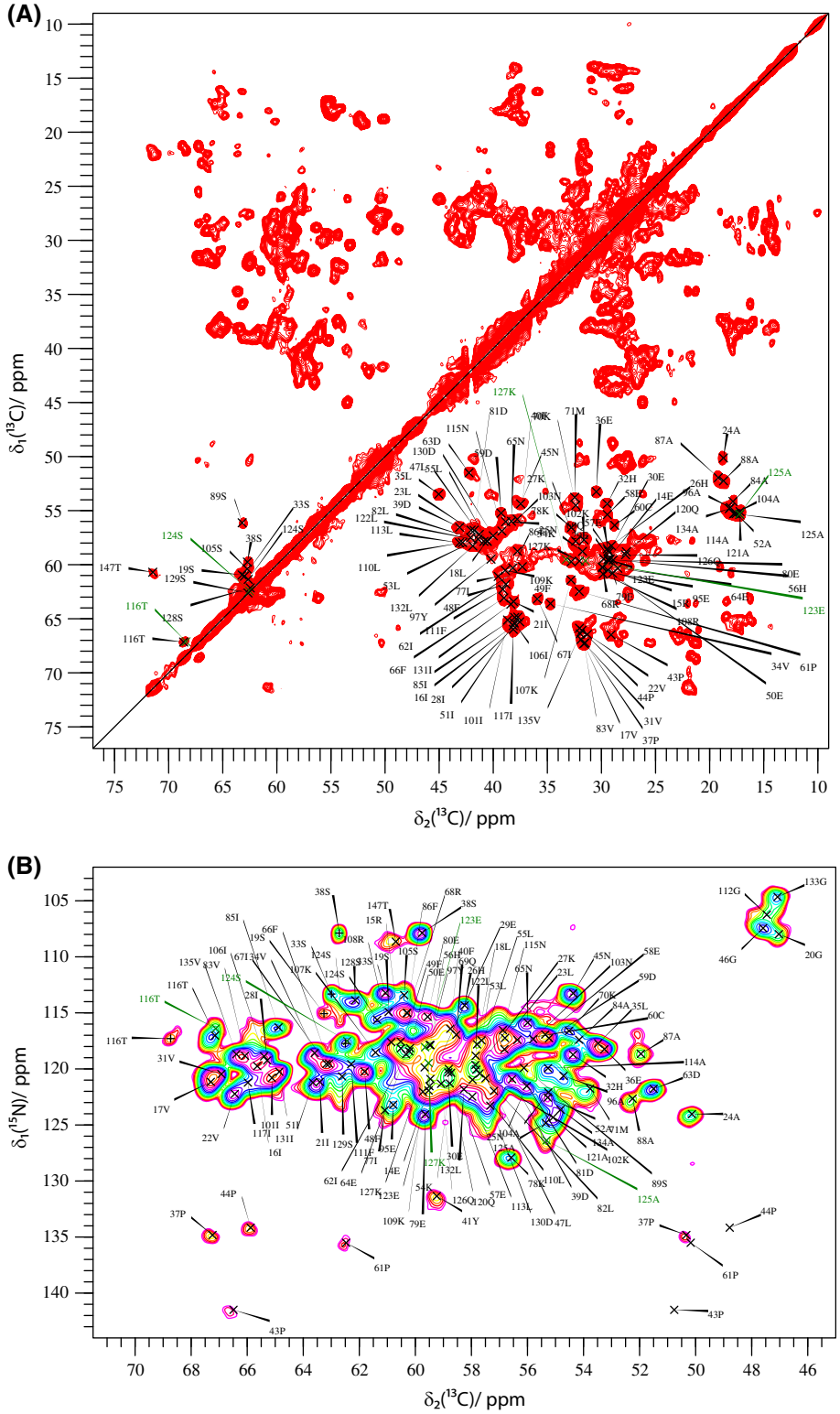


**Fig. 2** **a** 2D $^{13}C$,$^{13}C$ DARR spectrum of the N-terminus of *Hp*DnaB measured at 20.0 T with a spinning frequency of 17.0 kHz and 10 ms DARR mixing. The spectrum includes the labels for the Cα–Cβ peaks as predicted from the manually created shift list using the CcpNmr software (*black* chain A, *green* chain B). In the Cα/Cβ region seven isolated peaks could not be assigned, most probably because the corresponding residues are located in flexible parts of the protein. **b** 2D NCA spectrum of the N-terminus of *Hp*DnaB acquired at 20.0 T with a spinning frequency of 17.0 kHz. The spectrum includes the labels for the peaks as predicted from the manually created shift list using the CcpNmr software (*black* chain A, *green* chain B) assuming that only intraresidual peaks with a through-space limit corresponding to one bond are visible (x), peaks labeled with *plus* indicate N–Cβ resonances

detected as can be seen in the 2D dipolar correlation NMR spectra shown in Fig. 2. Those $^{13}C$ resonances are quite narrow with a line width at half height in the order of 0.6 ppm. Figure 2a shows the 2D $^{13}C$,$^{13}C$ DARR spectrum and Fig. 2b the 2D $^{15}N$,$^{13}C$ NCA spectrum, both with good signal-to-noise ratio which allows also to acquire 3D assignment NMR spectra even with four polarization-transfer steps. The assignment was mainly achieved by

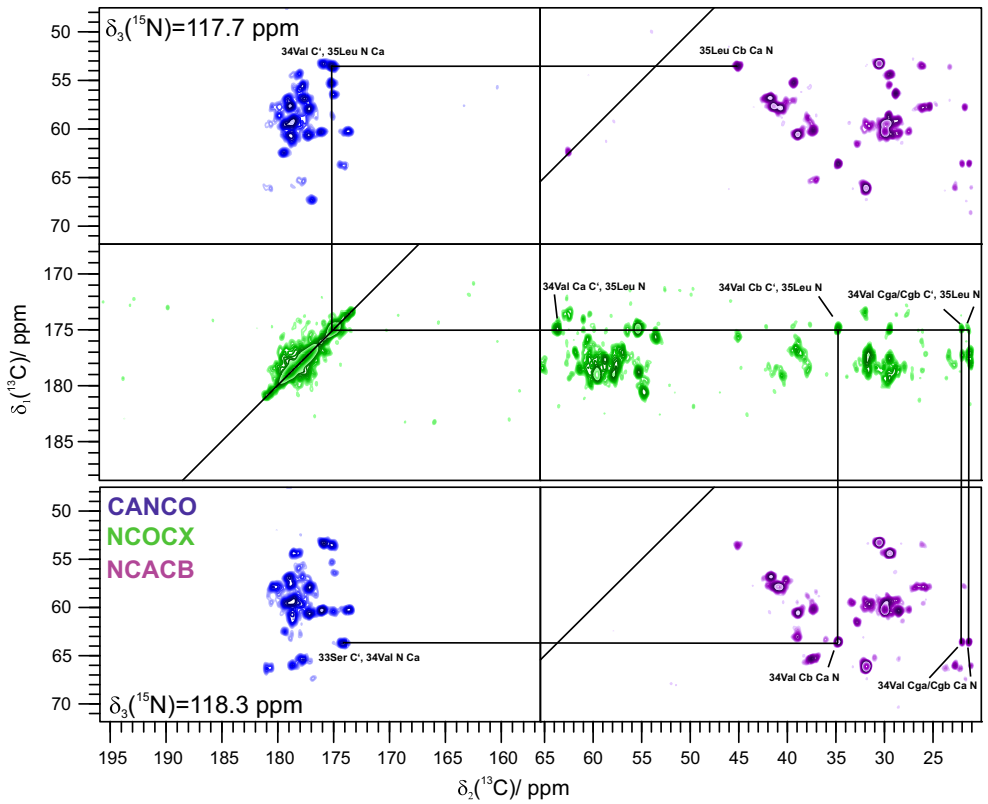


**Fig. 3** Example for the sequential walk along the protein backbone (from the C- to the N-terminus) by using NCACB (negative peaks are shown in *magenta*), CANCO (positive peaks are illustrated in *blue*) and NCOCX (positive peaks are shown in *green*) spectra

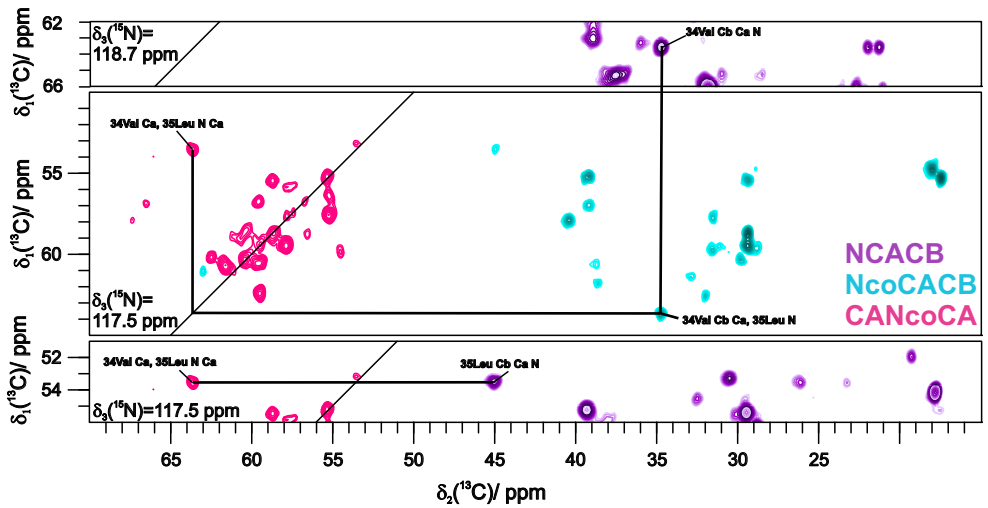


**Fig. 4** Example for the sequential walk along the protein backbone (from the N- to the C-terminus) of the N-terminus of DnaB by using NCACB (negative peaks represented in *magenta*), NcoCACB (negative peaks shown in *turquoise*) and CANcoCA (positive peaks represented in *pink*) spectra

using a combination of two strategies. The first one is based on 3D NMR spectra such as NCACB, NCACX, NCOCX and CANCO (Schuetz et al. 2010). A representative example for the backbone walk using this "classical" strategy is given in Fig. 3. The spectral overlap observed in the NMR spectra requires the largest possible spectral dispersion in all dimensions, which is in terms of 3D NMR spectra given by circumventing the detection of the CO-dimension which possesses the smallest chemical shift dispersion. For fulfilling this objective, an assignment

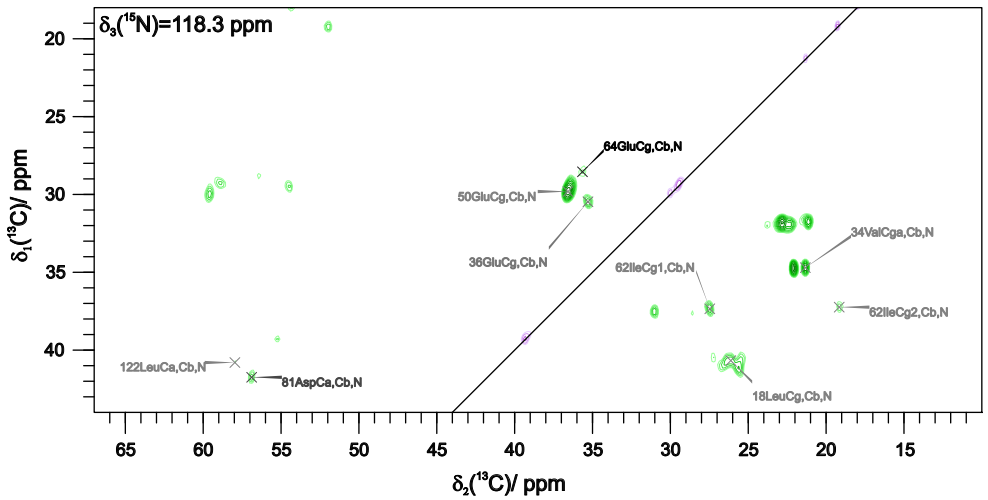


**Fig. 5** Example of a representative plane extracted from the 3D NcaCBCX spectrum ($\delta(^{15}N) = 118.3$ ppm) used for side-chain assignment

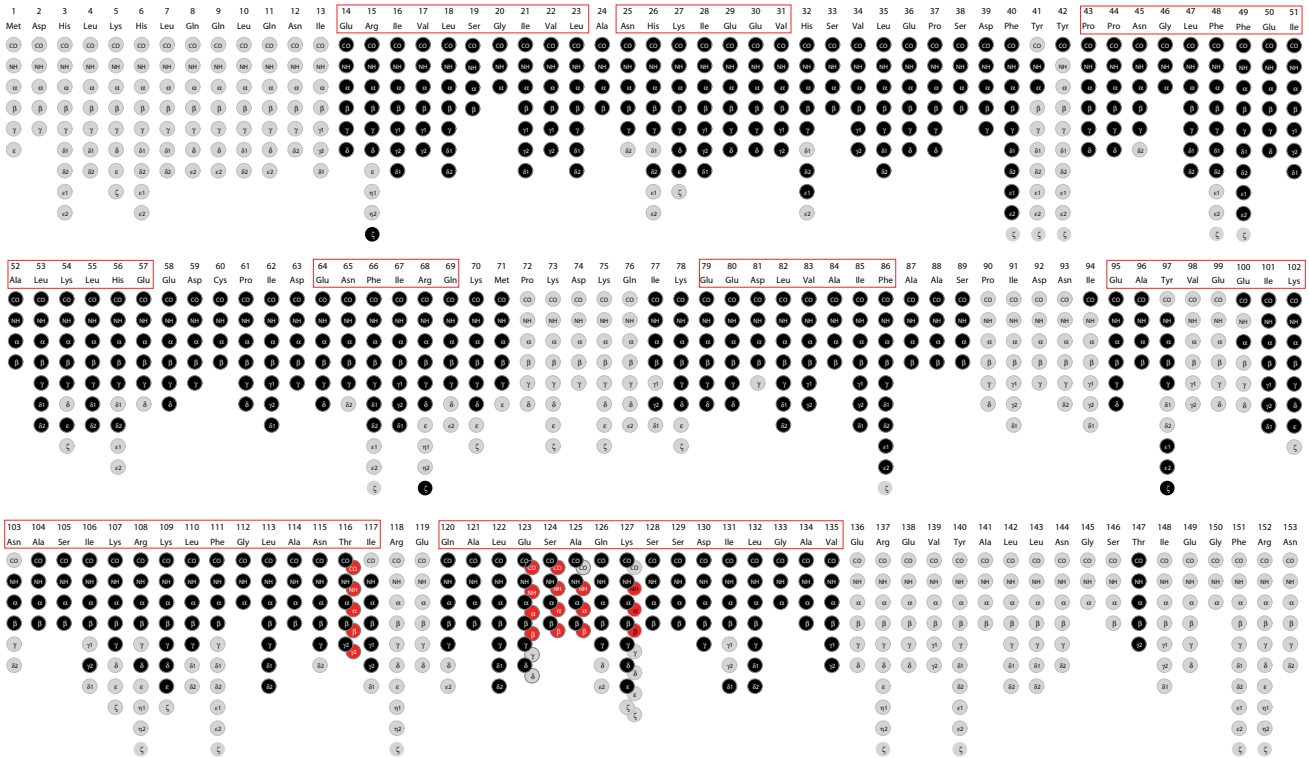

**Fig. 6** Assignment graph of the N-terminus of *Hp*DnaB created using the CcpNmr software. Residues marked in *red show* significant peak doubling. *Black* or *red dots* indicate assigned spins, *grey dots* unassigned spins. Residues highlighted by *red rectangles* are located in α-helices as determined from NMR secondary chemical shifts (see Fig. 9)

strategy based on NCACB, NcoCACB and CANcoCA experiments (Schuetz et al. 2010) was applied, a representative example of a sequential walk is given in Fig. 4. This assignment procedure could clarify many assignments which remained ambiguous after the firstly described assignment strategy. The selective C′–Cα polarization transfer step in the relayed experiments (Chevelkov et al. 2013) yields a good signal-to-noise ratio also for these types of experiments (see Fig. 4). Nevertheless, it has to be noted that a small number of resonances could not be assigned, mainly due to the spectral overlap as especially visible in the leucine, glutamic acid and glutamine region. The sidechains are mainly assigned by analyzing CCC (DREAM and DARR transfer) and NcaCBCX (two

**Table 2** Statistics of the manually performed peak assignments

| Category | Assigned/% |
|---|---|
| C | 69.3 |
| CA | 70.6 |
| CB | 69.4 |
| CG | 61.7 |
| CD | 54.5 |
| CE | 34.5 |
| CZ | 17.6 |
| N | 69.9 |
| Residue Ala | 91.7 |
| Residue Arg | 50.0 |
| Residue Asn | 55.6 |
| Residue Asp | 62.5 |
| Residue Cys | 100.0 |
| Residue Gln | 42.9 |
| Residue Glu | 72.2 |
| Residue Gly | 66.7 |
| Residue His | 60.0 |
| Residue Ile | 81.3 |
| Residue Leu | 68.8 |
| Residue Lys | 72.7 |
| Residue Met | 50.0 |
| Residue Phe | 85.7 |
| Residue Pro | 66.7 |
| Residue Ser | 88.9 |
| Residue Thr | 100.0 |
| Residue Tyr | 75.0 |
| Residue Val | 75.0 |

Data created with the CcpNmr software

DREAM transfers) spectra (see Fig. 5). On basis of the acquired spectra and the applied assignment strategies, approximately 70 % of the backbone carbon and nitrogen atoms could be assigned (see Fig. 6; Table 2). 77 % of the residues for which $^{15}$N, $^{13}$Cα and $^{13}$Cβ chemical shifts were assigned are located in α-helices as indicated by the NMR results (vide infra). The resonances of most of the unassigned residues could not be detected in the 3D NMR spectra, most probably because they are located in flexible parts of the protein (those resonances were also not detected in $^{13}$C,$^{1}$H INEPT and $^{15}$N,$^{1}$H HSQC spectra indicating intermediate dynamics in this protein). Notably, the resonances of Thr147 are clearly detected in the 2D DARR spectrum (see Fig. 2a), but appear only very weakly in the 3D NCACB and CCC spectra and are even absent in the other 3D NMR spectra. Since the second threonine residue (Thr116) present in the N-terminus of *Hp*DnaB could be assigned on the basis of the before described assignment strategies, the remaining Thr signal can be assigned to Thr147, and weak correlations visible in the 2D DARR spectrum at long mixing times support this conclusion. The Cα and Cβ chemical shifts of these two threonine residues already reveal that Thr116 is located in an α-helix, whereas Thr147 is most probably located in a loop in agreement with the observed flexible character. The chemical shifts have been deposited in the BMRB database under the accession number 26548.

As indicated in Fig. 6, for some resonances in the amino-acid region Thr116-Lys127 a peak doubling is observed which is exemplarily illustrated by analyzing the unique serine-alanine amino acid pair (residues Ser124 and Ala125). The sequential backbone walk for these residues using NCACB, NcoCACB and CANcoCA spectra is illustrated in Fig. 7 and clearly demonstrates the presence of two sets of resonances. Pronounced spectroscopic differences are mainly observed for the $^{15}$N frequencies. The observed peak doubling might indicate crystallographically distinct molecules in the asymmetric unit, which would agree with the previously published crystal structure for the residues 1–121 consisting of a dimer in the asymmetric unit (Kashav et al. 2009).

The results of the manual assignment procedure are validated by automated peak assignments as implemented in the solid-state FLYA algorithm (Schmidt et al. 2013). Figure 8 illustrates the good agreement between the manually assigned residues and the assignments obtained by FLYA calculations based on the peak lists from the manual assignment procedure. Only a few significant differences (e.g. for the two prolines Pro43 and Pro44, as well as for Thr147, and the C′ resonances of Ser89 and Val135) were observed and in those cases the manually assignment was carefully rechecked and its resonances were taken for the final assignment as they unambiguously result from the performed sequential walk.

## Secondary structure

Secondary chemical shifts were obtained by subtracting the random-coil shifts (Wang and Jardetzky 2002) from the observed solid-state NMR chemical shifts (Wishart et al. 1992) and are visualized in Fig. 9. These data clearly illustrate the dominant α-helical character of the N-terminus of *Hp*DnaB. Seven α-helices are identified in total, the longest one is reaching from residue 95 to 117. Our results indicate an interruption after the sixth α-helix and a succeeding seventh α-helix with positive secondary chemical shifts observed for residues 120–135. It is reasonable to assume that the C-terminal part of the protein is mainly flexible, since most resonances are not detected in the performed NMR experiments. The visible residue Thr147 shows Cα/Cβ chemical shifts more typical for a loop or extended conformation than an α-helical arrangement

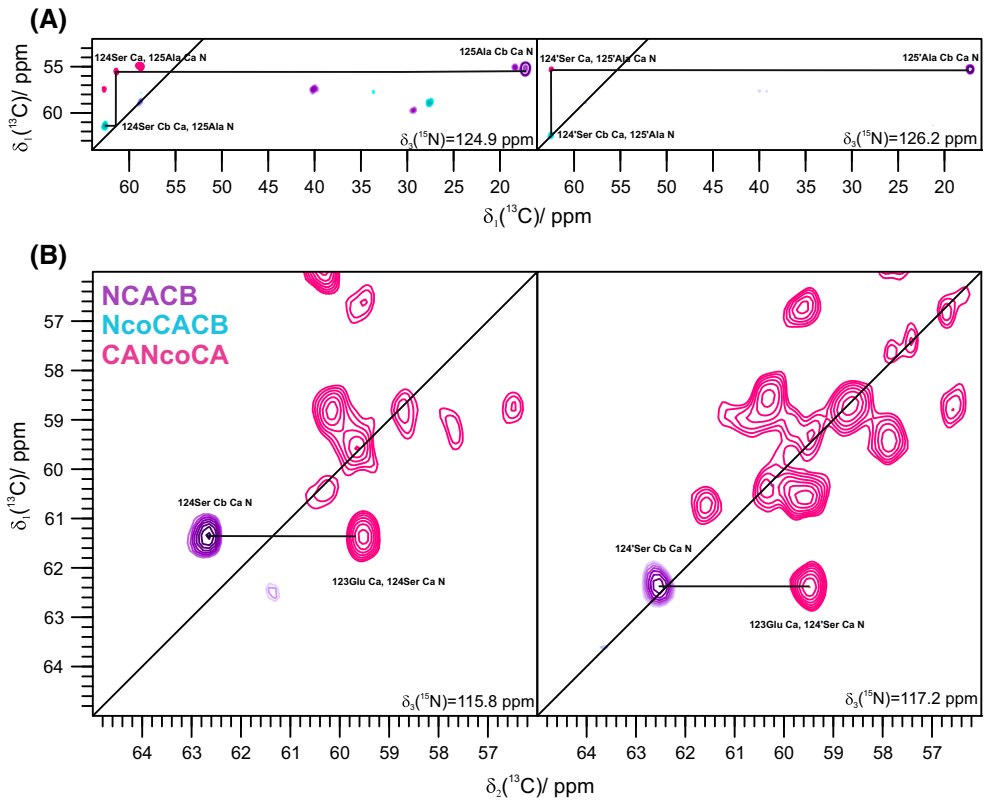


**Fig. 7** Illustration of peak doubling for a unique Ser-Ala amino acid pair (124Ser-125Ala). **a** The sequential walk along the protein backbone based on NcoCACB (negative peaks are illustrated in turquoise), CANcoCA (positive peaks are shown in *pink*) and NCACB (negative peaks are shown in *magenta*) spectra is shown for the two sets of resonances. **b** Correlation between the residues Ser124 and Glu123 in the NCACB and CANcoCA spectra, for the color code see (**a**)

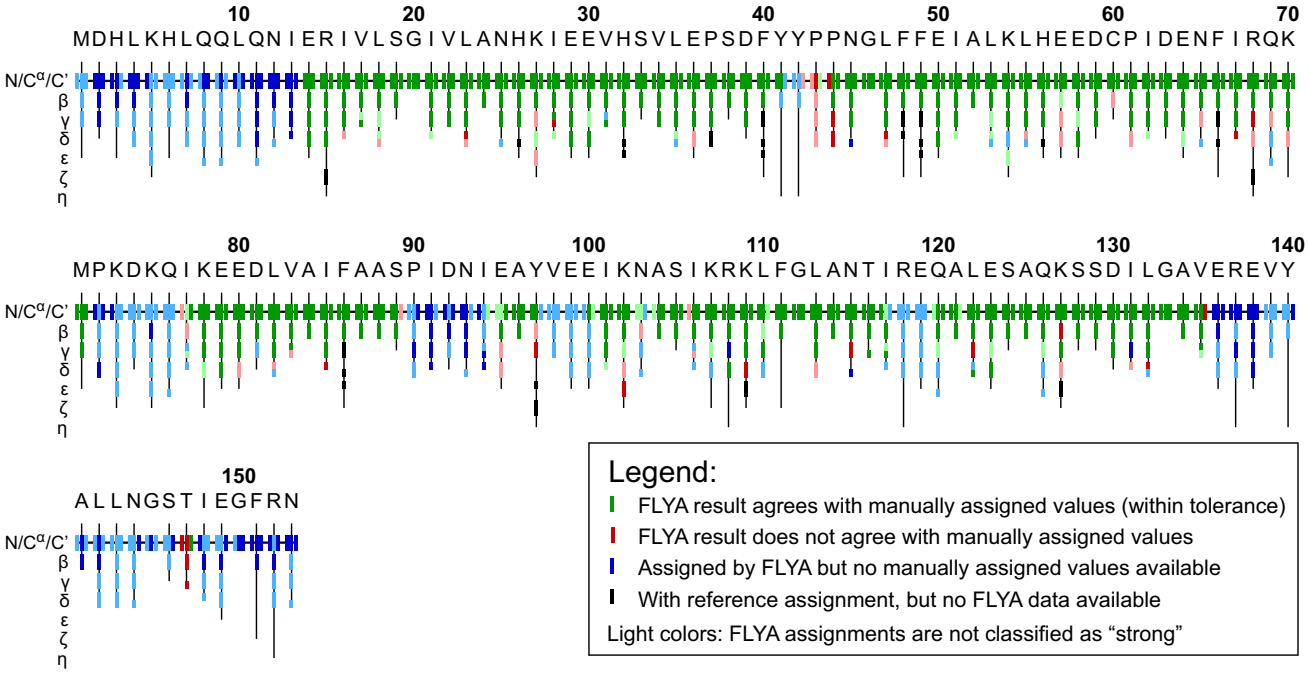


**Fig. 8** Result of an automated peak assignment using ssFLYA calculations on the basis of the manually created peak lists (which were used for the sequential assignment) extracted from NCACB, CANCO, NCOCX, NCACO, CANcoCA, NcoCACB and NcaCBCX spectra

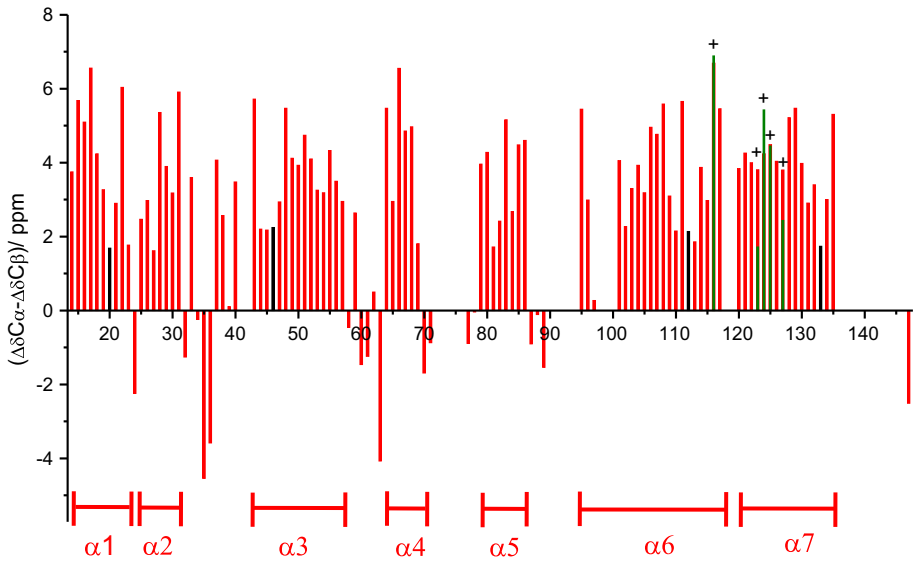


**Fig. 9** Secondary $^{13}C$ chemical shifts (obtained by subtracting the random-coil shifts (Wang and Jardetzky 2002) from the observed chemical shifts). Glycines are marked in *black* (and ΔδCα is plotted) and residues for which significant peak doubling was observed are highlighted with a+ and the corresponding secondary chemical shifts are shown in *green*

**Table 3** α-Helices determined from the secondary chemical shifts and from the 3D atomic coordinates deposited in the pdb file 3GXV using the DSSP algorithm (Kabsch and Sander 1983) via the corresponding web interface

| α-Helix | NMR | TALOS+[a] | DSSP[b] |
|---|---|---|---|
| 1 | Glu14-Leu23 | Arg15-Leu23 | His3-Leu23 |
| 2 | Asn25-Val31[c] | His26-His32 | Ile28-His32[d] |
| 3 | Pro43-Glu57 | Pro44-Glu57 | Pro43-Glu57 |
| 4 | Glu64-Gln69 | Glu64-Gln69 | Glu64-Gln69 |
| 5 | Glu79-Phe86 | Glu79-Phe86 | Glu79-Phe86 |
| 6 | Glu95-Ile117 | Ile101-Ile117 | Glu95-Gln120 |
| 7 | 120Gln–135Val | 120Gln-135Val | – |

[a] Calculated from the solid-state NMR chemical shifts described in this work using the TALOS+ software (Shen et al. 2009)

[b] Data for chain A are given

[c] Data do not allow a distinction between 3/10 helix and α-helix

[d] A 3/10 helix is predicted for residues Asn25-Lys27

(Wang and Jardetzky 2002). The secondary chemical shifts for the doubled resonances have the same sign and approximate magnitude for both partners, indicating no significant differences in the secondary structure elements (see Fig. 9).

A complete comparison of the secondary structure elements obtained from solid-state NMR with those found in the published crystal structure is hampered by the fact that the latter one consists only of residues 1–121 (Kashav et al. 2009) and might crystallize differently than the full-length N-terminal domain (no diffracting crystals could be obtained for the full-length N-terminal domain). Nevertheless, the structure of the globular N-terminal domain seems to be roughly similar between the 1–121 and 1–153 residues samples, since in general a good agreement between NMR data [secondary chemicals shifts and TALOS+ backbone torsion angle calculations (Shen et al. 2009)] and the X-Ray structure is observed (see Table 3).

## Conclusions

We describe the sequential resonance assignment of the N-terminal domain of *Hp*DnaB (residues 1–153) based on 3D solid-state NMR experiments leading to a site-specific chemical shift assignment of approximately 70 % of the backbone resonances. Those manually obtained peak assignments are validated by solid-state FLYA calculations. Seven α-helices where identified by a secondary chemical shift analysis which mostly agrees with the single crystal structure published for the residues 1–121 (PDB 3GXV). The observed peak doubling might point to crystallographically distinct molecules in the asymmetric unit.

**Acknowledgments** This work was supported by the Agence Nationale de la Recherche (ANR-11-BSV8-021-01, ANR-12-BS08-0013-01), the ETH Zurich, the Swiss National Science Foundation (Grant 200020_159707 and 200020_146757). PG is supported by the Lichtenberg Program of the Volkswagen Foundation, AB and LT by the CIBLE program 2011 from the Région Rhône-Alpes.

## References

Böckmann A et al (2009) Characterization of different water pools in solid-state NMR protein samples. J Biomol NMR 45:319–327

Chevelkov V, Giller K, Becker S, Lange A (2013) Efficient CO–CA transfer in highly deuterated proteins by band-selective homonuclear cross-polarization. J Magn Reson 230:205–211

Corn JE, Berger JM (2006) Regulation of bacterial priming and daughter strand synthesis through helicase–primase interactions. Nucleic Acids Res 34:4082–4088

Fogh R et al (2002) The CCPN project: an interim report on a data model for the NMR community. Nat Struct Mol Biol 9:416–418

Fung BM, Khitrin AK, Ermolaev K (2000) An improved broadband decoupling sequence for liquid crystals and solids. J Magn Reson 142:97–101

Gardiennet C, Schütz AK, Hunkeler A, Kunert B, Terradot L, Böckmann A, Meier BH (2012) A sedimented sample of a 59 kDa dodecameric helicase yields high-resolution solid-state NMR spectra. Angew Chem Int Ed 51:7855–7858

Gor'kov PL, Witter R, Chekmenev EY, Nozirov F, Fu R, Brey WW (2007) Low-E probe for 19F–1H NMR of dilute biological solids. J Magn Reson 189:182–189

Habenstein B et al (2011) Extensive de novo solid-state NMR assignments of the 33 kDa C-terminal domain of the Ure2 prion. J Biomol NMR 51:235–243

Kabsch W, Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 22:2577–2637

Kashav T, Nitharwal R, Abdulrehman SA, Gabdoulkhakov A, Saenger W, Dhar SK, Gourinath S (2009) Three-dimensional structure of N-terminal domain of DnaB helicase and helicase–primase interactions in *Helicobacter pylori*. PLoS One 4:e7515

LeBowitz JH, McMacken R (1986) The *Escherichia coli* dnaB replication protein is a DNA helicase. J Biol Chem 261: 4738–4748

Nitharwal RG et al (2007) The domain structure of *Helicobacter pylori* DnaB helicase: the N-terminal domain can be dispensable for helicase activity whereas the extreme C-terminal region is essential for its function. Nucleic Acids Res 35:2861–2874

Parsonnet J, Friedman GD, Vandersteen DP, Chang Y, Vogelman JH, Orentreich N, Sibley RK (1991) *Helicobacter pylori* infection and the risk of gastric carcinoma. N Engl J Med 325:1127–1131

Peterson WL (1991) *Helicobacter pylori* and peptic ulcer disease. N Engl J Med 324:1043–1048

Schmidt E, Güntert P (2012) A new algorithm for reliable and general NMR resonance assignment. J Am Chem Soc 134:12817–12829

Schmidt E et al (2013) Automated solid-state NMR resonance assignment of protein microcrystals and amyloids. J Biomol NMR 56:243–254

Schuetz A et al (2010) Protocols for the sequential solid-state NMR spectroscopic assignment of a uniformly labeled 25 kDa protein: HET-s(1–227). ChemBioChem 11:1543–1551

Shen Y, Delaglio F, Cornilescu G, Bax A (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. J Biomol NMR 44:213–223

Soni RK, Mehra P, Choudhury NR, Mukhopadhyay G, Dhar SK (2003) Functional characterization of *Helicobacter pylori* DnaB helicase. Nucleic Acids Res 31:6828–6840

Soni RK, Mehra P, Mukhopadhyay G, Dhar SK (2005) *Helicobacter pylori* DnaB helicase can bypass *Escherichia coli* DnaC function in vivo. Biochem J 389:541–548

Stelter M et al (2012) Architecture of a dodecameric bacterial replicative helicase. Structure 20:554

Stevens T et al (2011) A software framework for analysing solid-state MAS NMR data. J Biomol NMR 51:437–447

Studier FW (2005) Protein production by auto-induction in high-density shaking cultures. Protein Expr Purif 41:207–234

The UniProt Consortium (2014) Activities at the Universal Protein Resource (UniProt). Nucleic Acids Res 42:D191–D198

Vranken WF et al (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. Proteins Struct Funct Bioinform 59:687–696

Wang Y, Jardetzky O (2002) Probability-based protein secondary structure identification using combined NMR chemical-shift data. Protein Sci 11:852–861

Wishart DS, Sykes BD, Richards FM (1992) The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. Biochemistry 31:1647–1651