Complete $^1$H, $^{13}$C and $^{15}$N NMR assignments and secondary structure of the 269-residue serine protease PB92 from *Bacillus alcalophilus*

Rasmus H. Fogh, Dick Schipper, Rolf Boelens and Robert Kaptein

*Bijvoet Center for Biomolecular Research, Utrecht University, Padualaan 8, NL-3584 CH Utrecht, The Netherlands

Received 16 September 1994
Accepted 29 September 1994

**Keywords:** Triple-resonance spectroscopy; Multidimensional NMR; $^1$H, $^{15}$N and $^{13}$C NMR assignment; Serine protease

**Summary**

The $^1$H, $^{13}$C and $^{15}$N NMR resonances of serine protease PB92 have been assigned using 3D triple-resonance NMR techniques. With a molecular weight of 27 kDa (269 residues) this protein is one of the largest monomeric proteins assigned so far. The side-chain assignments were based mainly on 3D H(C)CH and 3D (H)CCH COSY and TOCSY experiments. The set of assignments encompasses all backbone carbonyl and CH$_n$ carbons, all amide (NH and NH$_2$) nitrogens and 99.2% of the amide and CH$_n$ protons. The secondary structure and general topology appear to be identical to those found in the crystal structure of serine protease PB92 [Van der Laan et al. (1992) Protein Eng., 5, 405-411], as judged by chemical shift deviations from random coil values, NH exchange data and analysis of NOEs between backbone NH groups.

**Introduction**

High alkaline serine protease PB92 (Maxacal®) is a subtilisin-like serine protease belonging to the subtilase family of proteases (Siezen et al., 1991), which are of interest both as well-studied examples of enzyme catalysis (Kraut, 1977) and as molecules of considerable industrial importance, being protein-degrading components of washing powders (Shaw, 1987). Subtilases are currently being used as model systems for protein engineering studies (see Teplyakov et al. (1992) and references cited therein) to improve stability and performance, and mutant enzymes are applied in modern detergent powders. For this purpose, knowing the solution structure and dynamics would be of immediate practical interest. The crystal structure of serine protease PB92 has been determined at 1.75 Å resolution (Van der Laan et al., 1992) and appears to have high similarity with various other members of the subtilisin family, such as subtilisin BPN' (Bott et al., 1988), subtilisin Carlsberg (Bode et al., 1987), thermitase (Gros et al., 1989) and proteinase K (Betzel et al., 1988). However, both information regarding dynamics and a solution structure, as can be obtained from NMR studies, are lacking.

Serine proteases have received much attention as a subject for NMR studies (see Consonni et al. (1992) for references); however, due to the high autolytic capacity of these proteases in the concentration range normally used for NMR studies, some of these studies have been unintentionally performed on partially degraded enzymes. To prevent degradation during the (many) measurements necessary for the assignments of the resonances and the determination of the 3D structure of proteases in general, use has to be made of highly purified protein and a tightly bound inhibitor for successful NMR studies. Diisopropylfluorophosphosphate (used in this study) is such an inhib-
itor that is stable for many weeks after the process of ageing (Van der Drift et al., 1985).

With a molecular weight of 27 kDa (269 residues), serine protease PB92 is one of the largest protein monomers studied in detail by NMR (Wagner, 1993). The constant improvement in NMR techniques (for a review, see Bax and Grzesiek (1993)) has permitted the assignment of the NMR spectra of such large proteins (e.g., Grzesiek et al., 1992; Powers et al., 1993; Xu et al., 1993; Yamazaki et al., 1993). The use of \(^{13}\text{C}^{15}\text{N}\)-labelled proteins in combination with heteronuclear 3D and 4D experiments alleviates the problem of spectral overlap, ultimately allowing the protons involved in an NOE to be distinguished on the basis of the frequencies of both the proton and the attached carbon/nitrogen. Assignment of protein backbone resonances can be done on the basis of triple-resonance \(^1\text{H}^{15}\text{N}^{13}\text{C}\) experiments without recourse to the – inherently more ambiguous – NOE connectivities. For the assignment of amino acid side chains, the larger C-C J-couplings permit bond-bond connections to be observed, even for quite large proteins with correspondingly large line widths.

The assignment of the backbone resonances of serine protease PB92 has been described previously (Fogh et al., 1994). Similar results have been obtained for a closely related serine protease, subtilisin 309 from Bacillus lentinus, by Remerowski et al. (1994). Here we present nearly complete \(^1\text{H},^{13}\text{C}\) and \(^{15}\text{N}\) assignments of serine protease PB92 using 3D triple-resonance NMR techniques. The secondary chemical shifts, NH exchange data and NOE assignments of serine protease PB92 in the crystal. Data from the crystal structure and the general structural topology of serine protease PB92 in the crystal. Data from the crystal structure, however, have not been used at any stage of the assignment procedure.

Materials and Methods

Serine protease PB92 was isolated from the Bacillus alcalophilus strain PB92 (Van Eekelen et al., 1989; Van der Laan et al., 1992). The protein was produced as described by Teplyakov et al. (1992), except that the yeast extract was replaced either by \(^{15}\text{N}\)-labelled yeast (cultured with \(^{15}\text{NH}_2\text{OH}\)) or by a doubly \(^{13}\text{C}^{15}\text{N}\)-labelled carbohydrate/amino acid medium, obtained from H. Orschkinat (EMBL, Heidelberg). Purification involved only cation exchange chromatography and ultrafiltration with a YM10 membrane (Amicon). The protein was inhibited and aged as described by Van der Drift et al. (1985) and again submitted to ultrafiltration. NMR samples contained 2 mM protein in 25 mM deuterated acetate buffer at pH 5.0 in H\(_2\)O:D\(_2\)O (95:5) or pure D\(_2\)O. Except where otherwise indicated a sample temperature of 315 K was used, and water suppression was done by presaturation.

All NMR experiments were performed on a Bruker 600 MHz AMX spectrometer, equipped with a three-channel NMR interface and a triple-resonance \(^1\text{H}^{15}\text{N}^{13}\text{C}\) probe with an additional gradient coil. 3D HNCA and HN(CA)CO (Clubb et al., 1992a; Grzesiek and Bax, 1992), 3D HNCO and HN(CA)CA (Grzesiek and Bax, 1992), 3D HN(CA)HA (Ikura et al., 1990a; Clubb et al., 1992b), and 3D TOCSY-\(^{15}\text{N},^1\text{H}\)-HSQC (Marion et al., 1989a) experiments were acquired with some modifications relative to the original pulse sequences (Vis et al., 1994). 3D HCA(CO)(CO)NH experiments were derived from the 3D HBBH(CBCA)(CO)NH experiment of Grzesiek and Bax (1993) by removing the pulses effecting magnetisation transfer between the \(\text{C}^\alpha\) and \(\text{C}^\beta\) spins. 3D HCH-COSY and HCCCH-TOCSY experiments (DIPSI-3 spin-lock sequence, duration 15.5 ms), using pulsed field gradients for removal of artifacts and partial water suppression, were carried out according to the description by Bax et al. (1990a,b) with some modifications (Vis et al., 1994). In the 3D HA(CA)(CO)NH experiment, resonances normally suppressed by water presaturation were recovered using the SCUBA method (Brown et al., 1988). Two 3D CB-CA(CO)NH experiments were acquired according to Grzesiek and Bax (1993), with delays optimised for the observation of NH and NH\(_2\) groups, respectively. 3D NOESY-\(^{15}\text{N},^1\text{H}\)-HSQC and NOESY-\(^{13}\text{C},^1\text{H}\)-HSQC experiments were acquired using the general scheme of Marion et al. (1989b) and Fesik and Zuiderweg (1988), and 3D \(^{15}\text{N}\) and \(^{13}\text{C}\) HMQC-NOESY-HMQC experiments were carried out according to Ikura et al. (1990b) and Frenkel et al. (1990). NOESY mixing times were 145 ms for the 3D \(^{15}\text{N}\) HMQC-NOESY-HMQC and 150 ms for other 3D NOESY spectra. 3D spectra were acquired with at least 16 scans per transient. Quadrature detection was obtained with either the TPPI (Marion and Wüthrich, 1983) or the States-TPPI (Marion et al., 1989c) method. In all cases, the numbers of real points acquired were 58–80 for nitrogen and aliphatic carbon, 80–96 for CO and typically 200–256 for indirectly detected protons. In carbon and nitrogen dimensions, the measured points were extended 50–100% using forward linear prediction with unmodified poles (Press et al., 1992), and Fourier transformed using cosine-square windows (typically shifted by 36°), zero-filling to 128 final points. In HCH experiments the \(\omega_c\) carbon dimension was folded once, using a spectral width of 45 ppm. The carrier frequency was mostly positioned in the centre of the aliphatic carbon range (40.0 ppm), but the centre frequency of the final processed spectrum was moved to 22.3 ppm by applying a linear phase correction to the free induction decays before processing. Proton dimensions were typically processed to 256 final points in indirectly detected and 1024 points in directly detected dimensions, using Lorentzian–