Biochimica et Biophysica Acta 1834 (2013) 568–582

Contents lists available at SciVerse ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbapap

Influenza A virus protein PB1-F2 from different strains shows distinct structural signatures

Sara M.Ø. Solbak [a,1], Alok Sharma [b,c,1], Karsten Bruns [b,c,1], René Röder [c,d], David Mitzner [c], Friedrich Hahn [c], Rebekka Niebert [a], Anni Vedeler [e], Petra Henklein [d], Peter Henklein [d], Ulrich Schubert [c], Victor Wray [b], Torgils Fossen [a,*]

[a] Department of Chemistry and Centre of Pharmacy, University of Bergen, N-5007 Bergen, Norway
[b] Helmholtz Centre for Infection Research, Department of Structural Biology, Braunschweig, Germany
[c] University of Erlangen-Nürnberg, Institute of Clinical and Molecular Virology, Germany
[d] Humboldt University, Institute of Biochemistry, Berlin, Germany
[e] Department of Biomedicine, University of Bergen, N-5009 Bergen, Norway

ARTICLE INFO

Article history:
Received 15 June 2012
Received in revised form 26 November 2012
Accepted 28 November 2012
Available online 7 December 2012

Keywords:
Influenza
PB1-F2
2009 Pandemic flu
1918 Spanish flu
Bird flu
PR8

ABSTRACT

The proapoptotic influenza A virus PB1-F2 protein contributes to viral pathogenicity and is present in most human and avian influenza isolates. The structures of full-length PB1-F2 of the influenza strains Pandemic flu 2009 H1N1, 1918 Spanish flu H1N1, Bird flu H5N1 and H1N1 PR8, have been characterized by NMR and CD spectroscopy. The study was conducted using chemically synthesized full-length PB1-F2 protein and fragments thereof. The amino acid residues 30–70 of PR8 PB1-F2 were found to be responsible for amyloid formation of the protein, which could be assigned to formation of β-sheet structures, although α-helices were the only structural features detected under conditions that mimic a membranous environment. At membranous conditions, in which the proteins are found in their most structured state, significant differences become apparent between the PB1-F2 variants investigated. In contrast to Pandemic flu 2009 H1N1 and PR8 PB1-F2, which exhibit a continuous extensive C-terminal α-helix, both Spanish flu H1N1 and Bird flu H5N1 PB1-F2 contain a loop region with residues 66–71 that divides the C-terminus into two shorter helices. The observed structural differences are located to the C-terminal ends of the proteins to which most of the known functions of these proteins have been assigned. A C-terminal helix–loop–helix motif might be a structural signature for PB1-F2 of the highly pathogenic influenza viruses as observed for 1918 Spanish flu H1N1 and Bird flu H5N1 PB1-F2. This signature could indicate the pathological nature of viruses emerging in the future and thus aid in the recognition of these viruses.

1. Introduction

The influenza A virus (IAV) H1N1 pandemic (pH1N1) that circulated widely in 2009–10 caused nearly 600 000 deaths worldwide from respiratory or cardiorespiratory causes [1,2]. Although the pH1N1 virus spread quickly to all continents, it was not as virulent as the Spanish flu (sH1N1), which caused more than 50 million deaths worldwide in 1918–20 [3]. Avian IAV strains sharing the lineage of the H5N1, commonly named as "bird flu" viruses crossed the species barrier in 1997 and have since then resulted in 562 infections in humans causing a total of 329 deaths [4], thus demonstrating an unusually severe pathogenic mechanism in humans.

The influenza viral protein PB1-F2, encoded by an alternative +1 reading frame within the coding region of the IAV RNA polymerase subunit PB1, is known as an important virulence factor of IAV [5,6]. However, its precise role within the context of viral pathogenicity remains unknown [7]. To date, PB1-F2 has been found to be immunopathological by enhancing the lung inflammatory responses to IAV infections [8–10], to interact with PB1 [11,12], and to cause cell death in specific cell types [5,13,14] by creating non-selective ion channels through mitochondrial membranes [15]. More recent studies found that some of these functions are strain specific and not a general feature of all PB1-F2 proteins [8–10,12]. PB1-F2 is mainly localized to the inner and outer mitochondrial membranes [13] and several of its known biological functions are suggested to arise through its direct interaction with membranes [15–17].

The PB1-F2 protein is expressed as full-length (87–90 amino acid residues) in nearly all avian IAV strains [18], but becomes truncated to

Abbreviations: ACN, acetonitrile; CD, circular dichroism; COSY, correlation spectroscopy; CSI, chemical shift index; IAV, influenza A virus; MTS, mitochondrial targeting sequence; NMR, nuclear magnetic resonance; NOESY, Nuclear Overhauser effect spectroscopy; RMSD, root mean square deviations; SDS, sodium dodecyl sulfate; TFE, trifluoroethanol; TOCSY, total correlation spectroscopy

* Corresponding author at: Department of Chemistry and Centre of Pharmacy, University of Bergen, N-5007 Bergen, Norway. Tel.: +47 55583463; fax: +47 55589490.
E-mail address: Torgils.Fossen@kj.uib.no (T. Fossen).
[1] These authors contributed equally to this work.

1570-9639/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.bbapap.2012.11.009

various lengths during evolutionary adaption to mammalian hosts [19]. Common for sH1N1, H5N1 and two other pandemic IAV strains from the last century, H2N2 1957 and H3N2 1968, is that PB1-F2 of these viruses has remained as a full-length version [10]. In contrast, the PB1-F2 of swine-origin pH1N1, appears to have a C-terminal truncated form where the PB1-F2 biosynthesis was terminated after 11 or 57 amino acid residues [20,21]. However, a full-length pH1N1 PB1-F2 could arise as a consequence of genetic reassortment, and it was speculated that such a change could increase the pathogenicity of the virus. Nevertheless, a study investigating the influence of a potential full-length pH1N1 PB1-F2, which was produced by removing the three stop codons that prevent synthesis of the full-length protein, on pathogenicity reported only a minor effect on virulence, although up-regulation of genes encoding proinflammatory cytokines was described [20].

Previously, structural characterization of H1N1 PR8 PB1-F2 (PR8 PB1-F2) carried out under membrane mimicking solution conditions by nuclear magnetic resonance (NMR) and circular dichroism (CD) spectroscopy, revealed the presence of an extensive α-helical structure in the C-terminal region, which was connected via a random coil domain to a much weaker N-terminal helix [22]. However, in aqueous solution the structure of the protein including the C-terminus was found to be mainly in a random state, as evidenced by CD spectra [22]. A recent report investigating PB1-F2 structure by CD spectroscopy in acetonitrile (ACN), sodium dodecyl sulfate (SDS) and asolectin described that PB1-F2 adopts a β-sheet structure under these solution conditions, and suggested that PB1-F2 is able to switch from a disordered random coil to α-helical and β-sheet structure in membranes [16]. In addition, it was reported that PB1-F2 forms amyloid fibers in membranes of infected cells, which indicates that the microenvironment surrounding PB1-F2 strongly influences secondary structure throughout its folding pathway [16].

The emergence of the pH1N1 strain in 2009 has increased the interest in understanding the molecular signatures of virulence of influenza viruses. Although PB1-F2 is widely accepted as a major virulence factor, its effect on pathogenicity has been shown to be very strain specific [8–10,12]. The mechanism through which PR8 PB1-F2 causes apoptosis has been assigned to the extensive C-terminal α-helix, which forms non-selective channels through mitochondrial membranes [17]. Additionally, β-sheet aggregation of PB1-F2 under certain solution conditions is associated with amyloid formation of the protein [16]. Taken together, this suggests that the secondary structure of PB1-F2 is closely related to its specific functions. This prompted us to investigate PB1-F2 structural characteristics of the IAV strains that differ in pathogenicity, to reveal a putative molecular (structure–function rationale) basis for explaining the differences in pathogenicity and mortality, which is currently lacking in the existing literature. Consequently, we have developed novel procedures to produce representative synthetic PB1-F2 proteins of sH1N1, H5N1 and PR8 [23,24]. Furthermore, we have examined the structure of the constructed full-length PB1-F2 of pH1N1 (in this paper referred to as PB1-F2 of pH1N1), which was previously found to have a minor influence on virulence of the virus [20]. We have analyzed three equal overlapping fragments of the proteins by combining CD and ^{1}H NMR data to determine the extent and position of structural regions in the PB1-F2 proteins. In addition, we have determined the secondary structure of the full-length PB1-F2 proteins by ^{1}H NMR. Interestingly, the present study reveals that the secondary structure adopted by PB1-F2 under membranous solution conditions correlates with the pathogenicity of the viruses from which the proteins are derived.

2. Materials and methods

2.1. Protein sequences

The PB1-F2 sequences used in this paper are those from the H1N1 IAV isolate A/Puerto Rico/8/34 isolate IAV (PR8) [25], the reconstituted sH1N1 human isolate A/Brevig Mission/1/1918 [26], H5N1 avian isolate A/duck/Guangdong/12/2000 [27], a predicted full-length protein of pH1N1 2009 human isolate A/California/04/2009 [20] and H5N1 avian A/Swan/FR/06299/2006 [16] referred to as PR8 PB1-F2, sH1N1 PB1-F2, H5N1 PB1-F2, pH1N1 PB1-F2 and FR06 PB1-F2, respectively.

2.2. Peptide synthesis and purification

The methodology for synthesis, purification and molecular characterization of the three overlapping fragments (residues 1–40, 30–70 and 50–90/50–87) and full-length proteins of sH1N1, H5N1, FR06 (fragments only) and pH1N1 have been described in detail elsewhere [23,24].

2.3. CD spectroscopy

CD spectra of all peptides dissolved at a concentration of 0.1 mg/ml or 0.4 mg/ml under various solution conditions (TFE concentration and 50 mM phosphate buffer pH 7.2) were recorded at ambient temperature under a nitrogen atmosphere on a Jasco J-810 spectropolarimeter using a Hellma Suprasil® quartz cuvette with a path length of 0.1 cm or 0.5 cm. Six scans were accumulated at a wavelength range of 260 to 180 nm and a step resolution of 0.1, 0.2 or 0.5 nm, respectively. The resulting spectra were corrected for the respective solvents, noise was reduced applying a standard smoothing algorithm, and spectra were finally deconvoluted using the DICROPROT 2000 program [28] in order to determine nature, quantities and positions of secondary structure elements. For some of the spectra the voltage of the high tension channel exceeded 700 V for wavelengths shorter than 185–193 nm. This observation may be rationalized by taking into account the absorbance of the buffer component of the solvent at these short wavelengths. Thus, only the wavelength range 260 to 185 nm has been included in the CD spectra presented in this paper.

2.4. Preparation of samples for NMR spectroscopy

The peptides were dissolved without pH adjustment (pH ~3.0) to final concentrations of 0.7–2 mM in 1:1 mixtures of H_2O and CF_3CD_2OH (50% aqueous TFE-D_2) or 80% ACN-D_3. For NMR studies investigating the influence of pH, peptides were dissolved in 50 mM phosphate buffer pH 7.0 containing 10% D_2O and 100 mM DPC-D_{38} buffer pH 7.4 or 2:8 mixtures of 50 mM acetate buffer pH 5 and ACN. TFE-D_2 (99.5%) and ACN-D_3 (99.8%) were purchased from Aldrich, while DPC-D_{38} was purchased from Cambridge Isotope Laboratories.

2.5. NMR spectroscopy

Two-dimensional (2D) ^{1}H correlation spectroscopy (COSY), total correlation spectroscopy (TOCSY) and Nuclear Overhauser effect spectroscopy (NOESY) NMR spectra of the H5N1 PB1-F2 peptides were recorded without spinning at 300 K on a Bruker Avance DMX 600 MHz instrument using a triple resonance probe head with gradient unit. Measurements were carried out with mixing times of 110 ms for the TOCSY and 250 ms for the NOESY experiment, respectively. A similar set of spectra was recorded under identical conditions for the sH1N1 and pH1N1 related peptides and full-length protein of PR8, sH1N1, H5N1 and pH1N1, (1.1–1.6 mM) on a Bruker Avance 600 MHz instrument equipped with a UltraShield™ Plus magnet and a triple resonance cryo-probe head with gradient unit. Data acquisition, processing and spectral analysis were in all cases performed with standard Bruker software. All spectra were internally referenced to the residual TFE-DH signal at 3.95 ppm. The unambiguous amino acid spin systems and the sequential assignments were established using a standard procedure [29] combining homonuclear 2D ^{1}H TOCSY and 2D ^{1}H NOESY NMR spectral data. Individual spin systems were identified from 2D ^{1}H TOCSY spectra, starting from the backbone amide protons and sequence-specific assignments were determined from cross-peaks in the 2D ^{1}H NOESY spectra

based on short observable distances between ^1H_N, $^1H_\alpha$ and $^1H_\beta$ nuclei of amino acid residue *i* and ^1H_N of residue $i+1$. This procedure provided the unambiguous assignments of all the spin systems for each of the three PB1-F2 fragments of sH1N1, H5N1 and pH1N1 (Supplemental Tables S1–S13) dissolved in 50% TFE-d_2 and allowed a direct comparison with the data for PR8 PB1-F2 reported previously under the same solution conditions [22]. The same procedure was used to determine the secondary structure of FR06 PB1-F2 in 50% aqueous TFE-d_2, and FR06 PB1-F2 and pH1N1 PB1-F2 in 80% aqueous ACN and 80% aqueous ACN-acetate buffer pH 5 (Supplemental Tables S14–S23). A general upfield chemical shift effect was observed for all the NMR signals of PB1-F2 peptides obtained in 80% ACN compared to the peptides in 50% TFE, that was not correlated to any structural changes. Reference values of Hα chemical shifts of random coil peptides are not available in current literature. Therefore, a correction factor of −0.079 ppm, calculated from the average differences in Hα chemical shift between peptides in 80% ACN and in 50% TFE, were applied to all the Hα chemical shifts from peptides in 80% ACN used to perform these CSI plots.

2.6. Structural calculations

The ^1H TOCSY and NOESY spectra of the N-terminal, central, and C-terminal PB1-F2 peptides of pH1N1 were assigned using CCP NMR Analysis version 2.1.5 [30]. Distances were derived from NOE peak heights in the ^1H–^1H NOESY spectrum and structure calculations were performed by using CYANA 2.1 [31]. The structures of the middle and C-terminal peptides of sH1N1 and H5N1 were determined from quantitative NOE data as described in detail elsewhere [32]. Structures were calculated on a Silicon Graphics Octane work station using the program CNS 1.0 with standard CNS parameters for data sets [33]. Distance restraints were used to generate 100 conformations of which 20 conformations, exhibiting no restraint violations greater than 0.2 Å and having the lowest energy values, were used for the final fitting analysis. The heterogeneity within the final set of 20 structures was visualized using the consecutive segment approach which allows reasonable regions for alignments to be defined [34]. The central structure for the defined fitting region with the lowest energy was then determined using the programs LSQMAN [35] and MOLEMAN2 [36] (Uppsala Software Factory). Finally, alignments were performed by comparing the central structure with all other structures and these were visualized with the PYMOL program [37].

2.7. Helix population plot

In order to gain structural information about the N-termini of the other IAV isolates, we performed helix population plots, calculated from ratio of the inter-proton distances *d*NN (i, i + 1) and *d*αN (i, i + 1). The distances *d* which strictly correlate with signal intensities I (I ~ 1/*d*6) were obtained by transferring the intensities of the respective NOE signals into interproton distances using CCP NMR [30]. Only unambiguous signals were used for this analysis. An equation given by Bradley et al. [38] was then used to calculate probabilities for helical or extended conformations.

2.8. Dynamic light scattering

Dynamic light scattering (DynaPro 801TC system, ProteinSolutions) was used to determine the oligomeric high molecular weight species for PB1-F2 proteins and its fragments. The measurements were carried out at 20 °C using a helium-neon laser wavelength of 633 nm and detection angle of 173°. The results were presented as size distribution.

2.9. Chemical cross-linking of synthetic PB1-F2 peptides and fragments thereof

Confirmation of purity and integrity of all synthesized peptides was assessed by SDS-PAGE with subsequent Coomassie Brilliant Blue staining as previously described [24]. Chemical cross-linking of the overlapping fragments as well as the full length peptides were performed similarly to previously established protocols [22]. To avoid potential problems with insufficient antibody reactivity on cross-linked species of PB1-F2, increased amounts of peptide were used to allow detection by Coomassie staining. Peptides were diluted to a final concentration of 500 μg/ml in phosphate-buffered saline and treated with a 2- or 20-fold molar ratio of disuccinimidyl suberate (DSS, Pierce) and incubated for 30 min. The cross-linking reaction was stopped by adjusting the reaction to a final concentration of 100 mM glycine. After addition of equal volumes of SDS sample buffer (0.4% SDS, 10% β-mercaptoethanol, 10% glycerol, 0.02% bromophenol blue, 100 mM Tris/HCl pH 6.8) the samples were separated on 15% SDS/PAA gels. Proteins were visualized by Coomassie staining (0.5% Coomassie Brilliant Blue G-250, 10% acetic acid, 25% isopropanol) with subsequent destaining (25% methanol, 10% acetic acid).

2.10. Electron microscopy

For transmission electron microscopy thin carbon support films were prepared by sublimation of carbon on freshly cleaved mica. Using 200 mesh copper grids, the samples of the N-terminal PB1-F2$^{1-40}$, the central PB1-F2$^{30-70}$ and C-terminal PB1-F2$^{50-87}$ fragments of the influenza strain PR8 at concentration 1 mg/ml in aqueous solution pH ~ 3 were negatively stained with 2% (w/v) aqueous uranylacetate, according to the method of Valentine et al. [39] and examined in a transmission electron microscope (TEM910, Zeiss, Germany) at an acceleration voltage of 80 kV at calibrated magnifications. Images were recorded digitally with a Slow-Scan CCD camera (ProScan 1024 × 1024, Scheuring, Germany) with ITEM software (Olympus Soft Imaging Solutions, Münster, Germany). Images were corrected for brightness and contrast applying Adobe photoshop CS3.

3. Results

3.1. Secondary structure prediction of PB1-F2 from PR8, sH1N1, H5N1 and pH1N1

The use of secondary structure prediction servers is an effortless way to achieve information on secondary structure of a protein, and has been frequently used. Previously, we have shown that secondary structure predictions are limited not only in their inability to simulate the molecules adaptability to environmental conditions but also for prediction of the position of secondary structures [22]. Similar structural predictions have been performed on the PB1-F2 variants included in this paper using the five different structure prediction methods Porter [40], Sable [41,42], Psipred, jnet and sspro [43]. All the predicted structures of the PB1-F2 variants were relatively similar, indicating the presence of three helices; one middle and two C-terminal helices (Figs. S1-S3). The structure predictions derived from Psipred, jnet and sspro (Phyre server) (Figs. S1-S3) also suggest that the N-terminal end contains regions with secondary structure. In accordance with our previous observation [22] all these predicted structures of PB1-F2 variants are of limited usefulness. It is now clear that the conformational properties of such proteins are difficult to predict without due regard to the osmolality, protein concentration, pH, buffer, and temperature conditions of the solutions employed [44]. Hence such techniques are unable to predict reliable structures that could have revealed the expected subtle differences that probably occur between the different isolates. We believe more convincing systematic experimental data can be gained from the use of CD and ^1H NMR spectroscopic data.

3.2. Secondary structures of PB1-F2 from PR8, sH1N1, H5N1 and pH1N1 obtained by CD spectroscopy

The first insights into the structure and folding of the peptides were achieved by CD spectroscopy, which was carried out under various solution conditions (Fig. S3–S6). Initially, none of the peptides showed any significant proportions of secondary structure in phosphate buffered solutions at pH 7.2 (Fig. 1), with one exception. The CD spectrum of pH1N1 PB1-F2$^{30–70}$ showed a weak β-sheet structure propensity with a characteristic negative band at approximately 218 nm and positive band at 195 nm (Figs. 1B, S2B and S5C). In contrast, by increasing the hydrophobicity of the solvent by adding trifluoroethanol (TFE) to 50% (v/v), the helical content of all peptides increased significantly. The CD spectra were similar for analogous peptides dissolved in 50% aqueous TFE at pH 3 and in 50% TFE and phosphate buffer (pH 7.2), indicating that the observed α-helical structures were independent of the pH. The secondary structure of the N-terminal fragments in 50% TFE increased only moderately up to 13–20% (Table 1) and no clear differences were observed between the four peptides (Fig. 2A). A more pronounced increase in helical content was evident for the central fragments PB1-F2$^{30–70}$ from PR8, sH1N1, H5N1 and pH1N1, as observed from the CD spectra that showed characteristic negative ellipticities at ~208 nm and ~222 nm as well as maxima at ~192 nm (Fig. 2B). While ~30% helix was calculated for the PR8 PB1-F2$^{30–70}$ and sH1N1 PB1-F2$^{30–70}$ peptides, significantly higher proportions of helical structure (>46%) were present in the H5N1 PB1-F2$^{30–70}$ peptide and pH1N1 PB1-F2$^{30–70}$. However, the CD spectra of the C-terminal PB1-F2 fragments of PR8 and sH1N1 were almost identical and showed a relatively high helical content (~44%). The relative proportions of helical structure were somewhat lower for the C-terminal PB1-F2$^{50–90}$ fragments of H5N1 and pH1N1 (Fig. 2C, Table 1).

3.3. Secondary structure of full-length PR8 PB1-F2

The most obvious structural changes are likely to occur in the most structured state of the molecule. Although a weak β-sheet structure was observed for the middle fragment of PB1-F2 from one of the IAV isolates (pH1N1) by CD spectroscopy, the optimum conditions for achieving the most structured state for all PB1-F2 fragments were at the hydrophobic conditions using the membrane mimicking aqueous organic solvent TFE (50% v/v). Our previous report on the structure of PR8 PB1-F2 [22] was based on chemical shift index (CSI) plots obtained experimentally for 50% TFE-d_2 solutions of each of three overlapping peptide fragments that were combined to predict the secondary structure of the full-length protein. CSI plots of the α-proton (Hα) chemical shifts relative to those of residues in a random coil have proved to be an appropriate and powerful method for determining the presence of secondary structure in peptides and proteins [29,45]. It has been shown experimentally that Hα-proton chemical shifts greater than 0.1 ppm relative to the random coil values are qualitative indicators of protein secondary structure [45]. A minimum of four adjacent residues with an upfield shift relative to random coil (negative CSI) are indicative of an α-helix, whereas β-sheets require a minimum of three residues with downfield shifts (positive CSI) [45]. The access to full-length PR8 PB1-F2$^{1–87}$ and availability of a sensitivity-enhanced NMR instrument equipped with a cryogenic probe and an ultrashielded magnet, providing high-resolution NMR spectra, allowed the secondary structure determination of the full-length protein (Fig. 3A). The secondary structure of the full-length PR8 PB1-F2 was shown to correspond to the combination of the structures of the N-terminal, middle and C-terminal fragments previously determined.

Table 1
Percentage helical content of the PB1-F2 fragments under various solution conditions. Mean percentage helical content from the average of five algorithms (Chou-Fasman, Chen, Bolotina, Chang, Yang) [65–67] and then combined (averaged) the resulting value with the two other methods: K2D [68] and CfE (Calculted from Ellipticity) [65].

PB1-F2 fragment	TFE:H_2O (1:1)	TFE:Buffer pH 7.2 (1:1)	Buffer pH 7.2
PR8 PB1-F2 (1–40)	13	20	5
pH1N1 PB1-F2 (1–40)	19	15	5
sH1N1 PB1-F2 (1–40)	13	14	6
H5N1 PB1-F2 (1–40)	18	13	5
PR8 PB1-F2 (30–70)	28	39	6
pH1N1 PB1-F2 (3–70)	50	48	15
sH1N1 PB1-F2 (30–70)	33	34	5
H5N1 PB1-F2 (30–70)	46	48	10
PR8 PB1-F2 (50–87)	44	48	2
pH1N1 PB1-F2(50–90)	37	40	11
sH1N1 PB1-F2 (50–90)	48	48	9
H5N1 PB1-F2 (50–90)	35	49	6

3.4. Positions of secondary structure of PB1-F2 from sH1N1, H5N1 and pH1N1 compared to PR8

In order to investigate the structural characteristics of PB1-F2 from sH1N1, H5N1 and pH1N1 comparative to that of PR8 PB1-F2 in more detail with the aim of revealing potential significant structural differences among these proteins that may be related to functional differences, we have recorded a set of one- and two-dimensional (1D, 2D) ^{1}H NMR

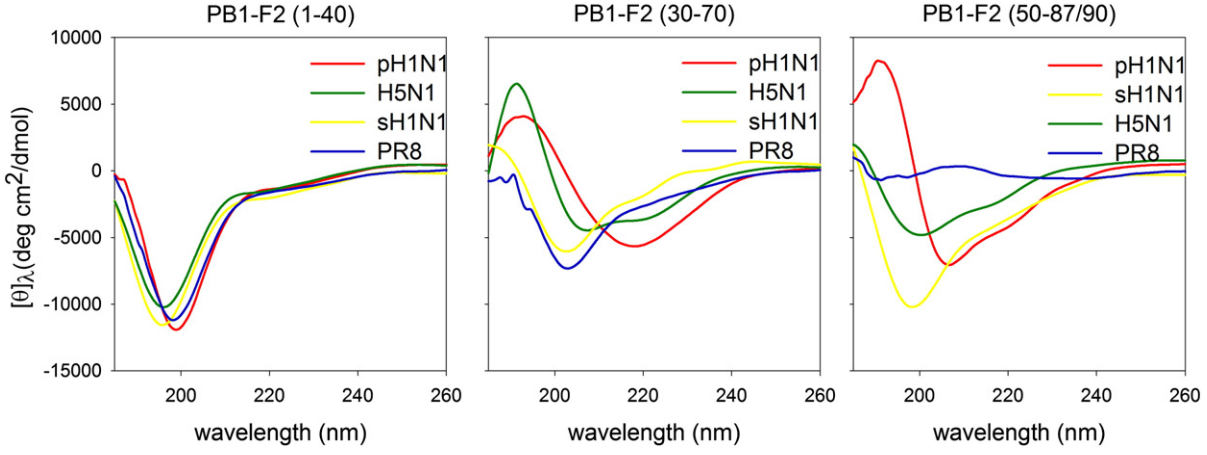

Fig. 1. Secondary structure of PB1-F2 fragments from IAV PR8, sH1N1, H5N1 and pH1N1 at pH 7.2. Comparison of the far-ultraviolet CD spectra of N-terminal PB1-F2$^{1–40}$ fragments of PR8, sH1N1, H5N1 and pH1N1(A), central PB1-F2$^{30–70}$ fragments of PR8, sH1N1, H5N1 and pH1N1(B), and C-terminal PB1-F2$^{50–87/90}$ fragments of PR8, sH1N1, H5N1 and pH1N1(C) recorded in buffer solution pH 7.2 at 20 °C. Spectra were background corrected and are the average of 6 scans.

spectra of each of the peptide fragments and full-length proteins under the same solution and temperature conditions as previously used for peptides of PR8 PB1-F2 [22]. After spectral assignment, CSI plots were performed for both fragments and full-length proteins. However, due to signal broadening caused by oligomerization in the middle region of the full-length H5N1 PB1-F2, some Hα chemical shifts of this protein were not observable in the NMR spectra. Therefore, to approximate the structure of full-length H5N1 PB1-F2 protein, fragment data for residue Q28 (derived from the N-terminal fragment H5N1 PB1-F2$^{1-40}$), S35-R48 and G50-M51 (derived from the middle fragment H5N1 PB1-F2$^{30-70}$) were combined with full-length data of H5N1 PB1-F2 to achieve a complete CSI for this full-length protein. The results are depicted in Fig. 3 and compared with the analogous data for PR8 PB1-F2. CSI from PB1-F2 fragments was in accordance to full-length protein CSI (CSI of fragments is shown in the Supplemental material, Figs. S7–S10). Each of the four proteins contains a substantial content of helical structure, which, in all cases, is located mainly in the middle and C-terminal regions of the molecules.

The N-termini of the four proteins appeared from CSI to be largely unstructured apart from two short regions each about four to six residues in length and centered at positions 11 and 18, respectively, which, to some extent, show propensities for helix formation. By comparing the N-terminal regions of the four IAV strains, a similar extent of weak α-helical structure involving residues 9 to 30 was found. The Hα CSI plots indicate that the structure of the N-terminus of PR8 PB1-F2 is most similar to the N-terminus of sH1N1 PB1-F2 (Figs. 3A,C and 4A,C) whereas the structure of the N-terminus of pH1N1 PB1-F2 strongly resembles to that of H5N1 PB1-F2 (Figs. 3B,D and 4B,D).

The CSI plots of the central and C-terminal regions of the proteins, reveal that PR8 PB1-F2 harbors an extensive α-helix starting directly after Pro-52 and comprising residues Lys-53 to Ser-84 (Fig. 3A) which is in accordance with the calculated NMR structure of the C-terminus of this protein published previously [22]. A second Pro residue located approximately in the center of the helix (Pro-67) does not seem to act as a helix breaker as can readily be seen by the upfield shifts of Hα of residues in its direct proximity. The downfield shift observed for Hα of Asn-66 (0.120 ppm) can be explained by the intrinsic, unusual downfield shifts of 0.28 ppm ± 0.1 ppm that Pro residues generally cause for Hα of the preceding residue [46]. This finding is further supported by the presence of a considerable number of medium range cross peaks in the NOESY spectrum that are indicative of helical conformation in this region of the molecule ($^1H_\alpha$ Arg-65/^1H_N Ile-68, $^1H_\alpha$ Asn-66/^1H_N Leu-69, $^1H_\alpha$ Pro-67/^1H_N Val-70, $^1H_\alpha$ Arg-65/$^1H_\beta$ Ile-68, $^1H_\alpha$ Pro-67/$^1H_\beta$ Val-70, $^1H_\alpha$ Leu-64/^1H_N Ile-68, $^1H_\alpha$ Arg-65/^1H_N Leu-69, and $^1H_\alpha$ Pro-67/^1H_N Phe-71). Therefore, it seems very likely that Asn-66 and Pro-67 are integral elements of a well-defined, continuous α-helix, at most contributing to a limited flexibility at this particular site of the helix.

The pH1N1 PB1-F2 contains an even longer α-helical structure than PR8 PB1-F2, starting directly after Ser-35 and ending at Lys-85 (Fig. 3B). As with PR8, pH1N1 also has a Pro residue located in the center of the helix (Pro-67) giving the same downfield shift effect on Hα of the preceding residue Asn-66 (0.110 ppm). This is, however, as with PR8, unlikely to represent a helix break, which is confirmed by the observed cross peaks in the NOESY spectrum indicative of α-helical conformation in this region.

Compared to PR8 the CSI plot of the sH1N1 PB1-F2 protein (Fig. 3C) shows a similar pattern of upfield shifts. Again a more extensive helix begins directly after Pro-52. However, one striking difference becomes immediately apparent in the 66–71 region of the sequence. In this region, Hα of three successive residues shows downfield shifts relative to random coil values. The most pronounced downfield shift is observed for Ser-66 (0.210 ppm) and particularly Thr-68 (0.360 ppm), which clearly indicate a break of the helix. Although both of these residues precede a Pro residue their pronounced downfield shifts may not solely be explained by the intrinsic downfield shift usually caused by Pro residues on Hα of the preceding residue. The fact that a second Pro residue is introduced at position 69 of the sH1N1 PB1-F2 sequence, only two residues apart from Pro-67 renders it difficult to stabilize secondary structure through H-bond formation and these two Pro residues are likely to act as helix breakers.

For H5N1 a break in the C-terminal helix in the region comprised of residues 66–71 is also observed even though only one Pro residue, Pro-67, is present (Fig. 3D). There is also a notable upfield change in the Hα chemical shifts of residues 36 to 49 compared to sH1N1 indicating a stabilization of helical conformations in this region of the molecule. However, compared to the PR8 sequence the residues I^{68}LVF^{71} have been replaced with T^{68}QGS^{71}. Among these residues, Thr, Gly and Ser are commonly reckoned as helix breakers whereas Leu and Phe are beneficial for α-helix formation [47].

3.5. Structural elements in the N-terminal regions of PB1-F2 from pH1N1, PR8, sH1N1 and H5N1

The Hα CSI plots of full-length PB1-F2 proteins and their respective N-terminal peptides indicate that a relatively random structure predominates in this region, even under solution conditions where formation of α-helical conformation is favored (50% aqueous TFE) (Figs. 3–4). However, a substantial number of medium-range NOE

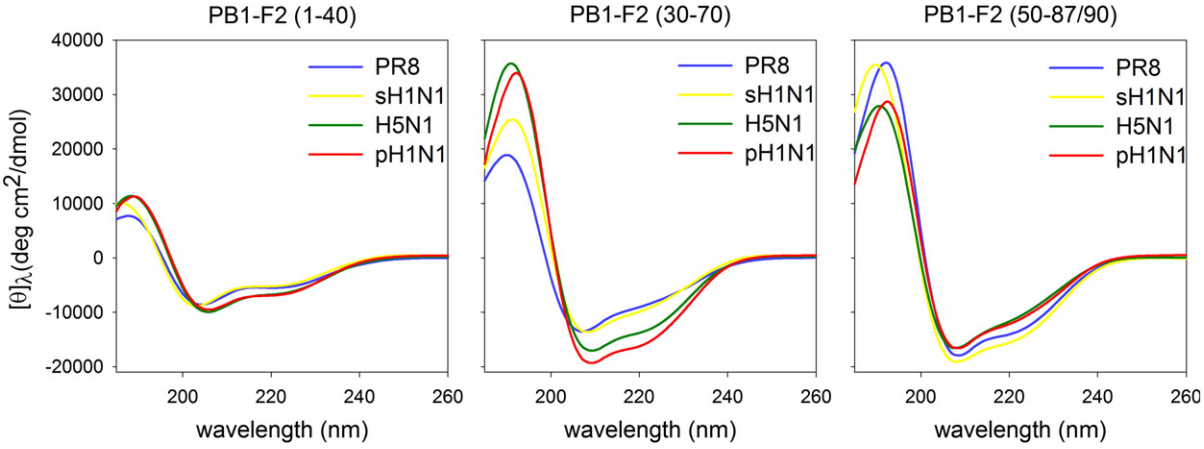

Fig. 2. Secondary structure of PB1-F2 fragments of IAV PR8, sH1N1 and H5N1. Comparison of the far-ultraviolet CD spectra of N-terminal PB1-F2$^{1-40}$ fragments of PR8, sH1N1, H5N1 and pH1N1 (A), central PB1-F2$^{30-70}$ fragments of PR8, sH1N1, H5N1 and pH1N1 (B), and C-terminal PB1-F2$^{50-87/90}$ fragments of PR8, sH1N1, H5N1 and pH1N1 (C) recorded in 50% aqueous TFE. Spectra were background corrected and are the average of 6 scans.

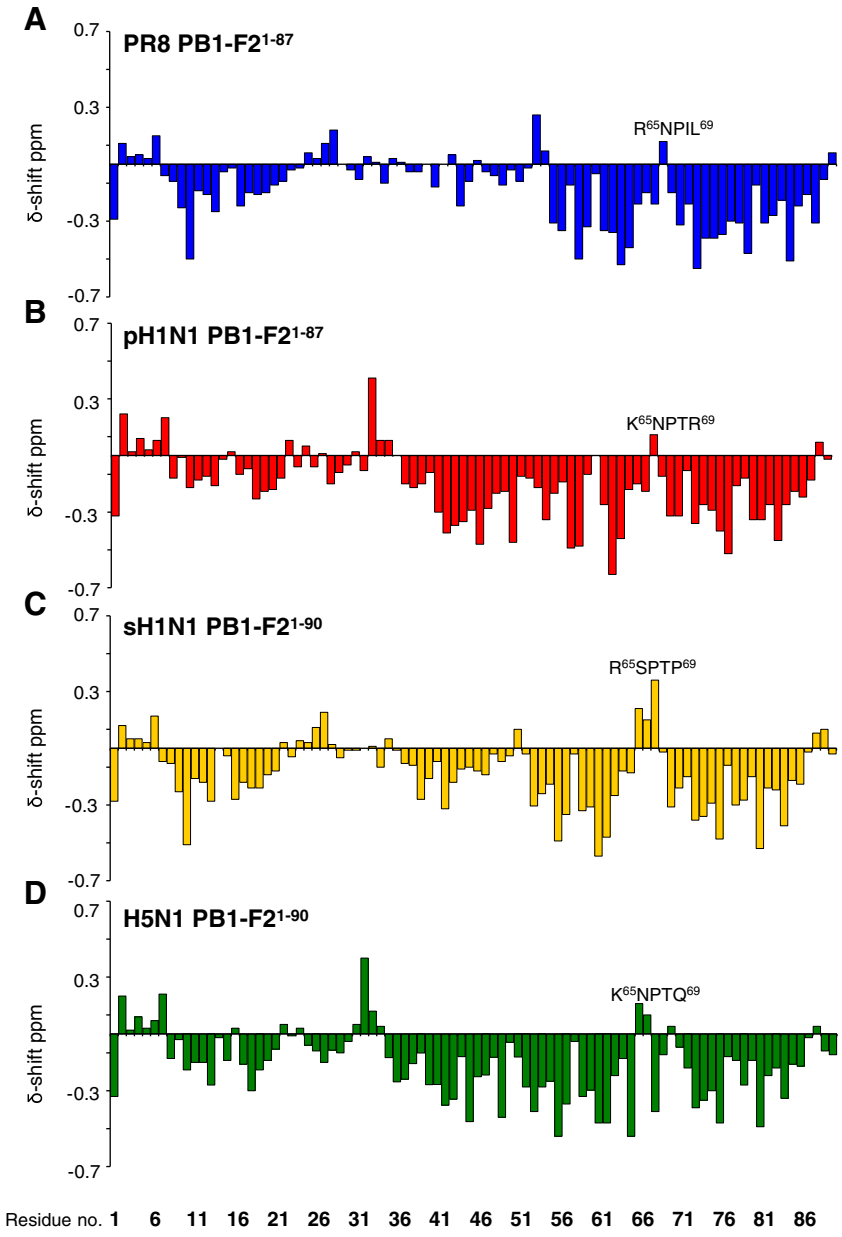

Fig. 3. CSI plots of full-length PB1-F2 of PR8 (A), pH1N1 (B), sH1N1 (C) and H5N1 (D). Due to oligomerization and overlapping signals in the NMR spectra of the middle region of full-length H5N1 PB1-F2 (D), it was not possible to obtain all αH chemical shift. Therefore, fragment data for residues Q28 (N-terminal fragment), S35-R48 and G50-M51 (Central fragment) were combined with full-length data to perform the CSI plot of H5N1 PB1-F2.

cross peaks, characteristic of secondary structure, were present in the NMR spectra of the N-terminal peptides. The NMR structure could be calculated for the N-terminal of pH1N1 in 50% TFE-*d*2 (Fig. 5A). The calculated high resolution secondary structure was better defined than indicated from CSI, with the largest helical propensity from Trp-9 to Ser-30. This was surprising, as this region was previously considered to be unstructured with a very weak propensity for helical structure. The spectral analysis of the secondary structure of the N-terminal PB1-F2 peptides of the other isolates was to some extent complicated by signal overlap in the spectra. This prevented structure calculation and thereby determination of high-resolution structures of N-terminal PB1-F2 fragments of PR8, sH1N1 and H5N1. In order to investigate the N-terminal structure of these IAV isolates we performed helix population plots calculated from ratios of the inter-proton distances *d*NN (i, i + 1) and *d*αN (i, i + 1) (Figs. 5A and 6) as described previously by others [38,48]. These plots revealed similar propensities for helical structure in the N-termini of PB1-F2 of PR8, sH1N1 and H5N1, as determined for pH1N1.

3.6. Structural elements in the central and C-terminal regions of sH1N1, H5N1 and pH1N1

Examination of the 2D NOESY spectra of the central and C-terminal fragments of sH1N1, H5N1 and pH1N1 indicated a substantial number of medium range NOEs that could be resolved or quantified to allow determination of high resolution structures. Consequently, quantitative NOE data from the middle and C-terminal fragments in 50% TFE-*d*2 were used as distance constraints in molecular dynamic/energy minimization calculations to yield high-resolution structures. In brief, a total of between 376 and 547 distance constraints for each fragment were used to generate 100 conformations from which the 20 lowest energy conformers were selected and visualized. The structure of the

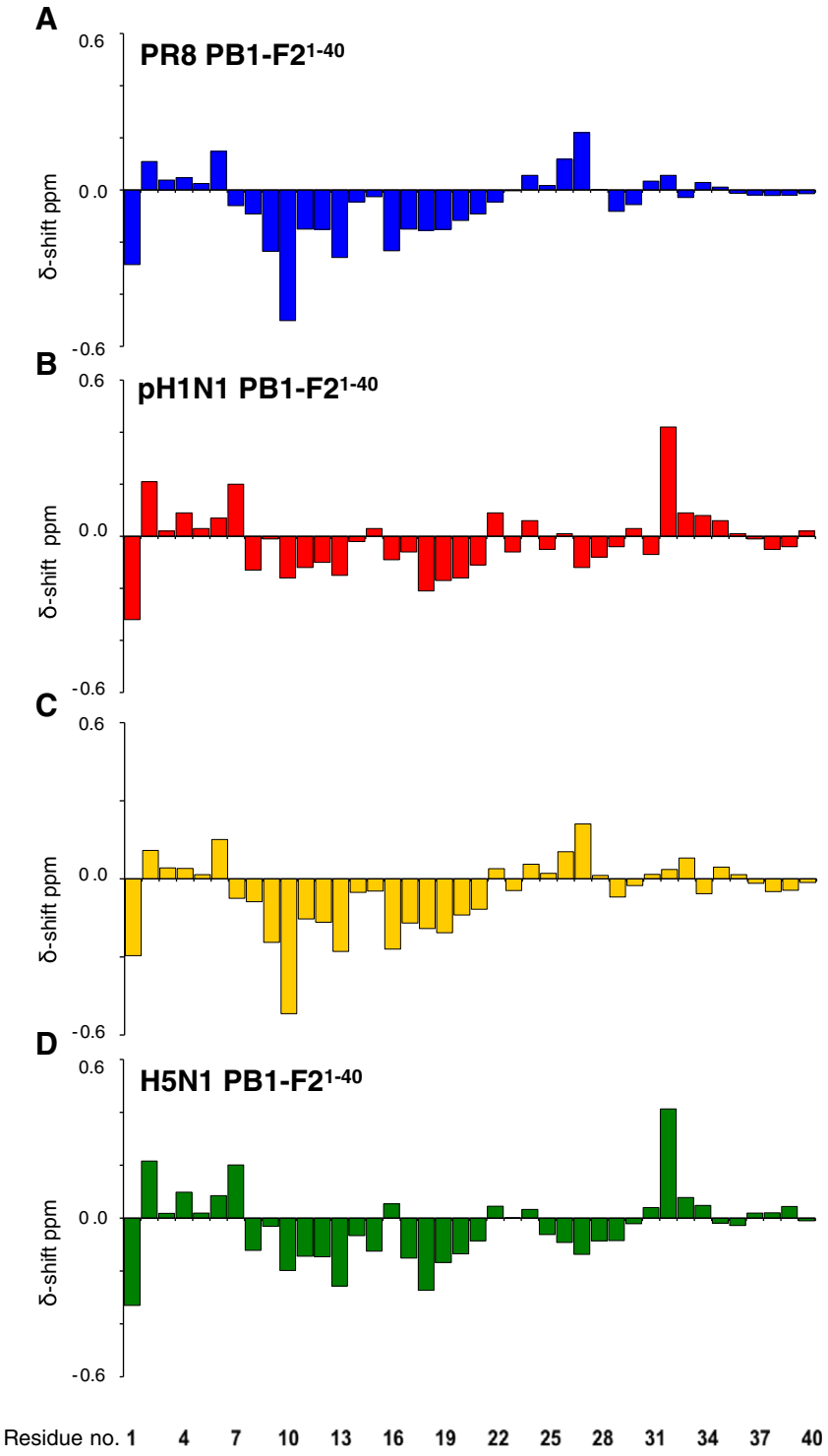

Fig. 4. CSI plots of N-terminal PB1-F2 of PR8 (A), pH1N1 (B), sH1N1 (C) and H5N1 (D).

central and C-terminal PB1-F2 fragments of pH1N1 and corresponding helix population plots are presented in Fig. 5C–F. Root mean square deviations (RMSD) for backbone atoms were calculated for each residue, plotted versus residue number from the 20 final structures and used to furnish the positions of stable structural elements (less than 2 Å) (Supplemental data Figs. S11–S14).

Based on the combination of calculated structures, RMSD plots and CSI plots of the various fragments and full-length PB1-F2 proteins, a complete picture of the secondary structures of PB1F2 of PR8, pH1N1, sH1N1 and H5N1 could be schematically summarized (Fig. 7).

3.7. Variations in PB1-F2 sequences between PR8, pH1N1, sH1N1 and H5N1

In order to reveal the primary structural basis for the pronounced differences between the PB1-F2 sequences of the four influenza strains investigated in this study, sequence alignments were performed (Fig. 8).

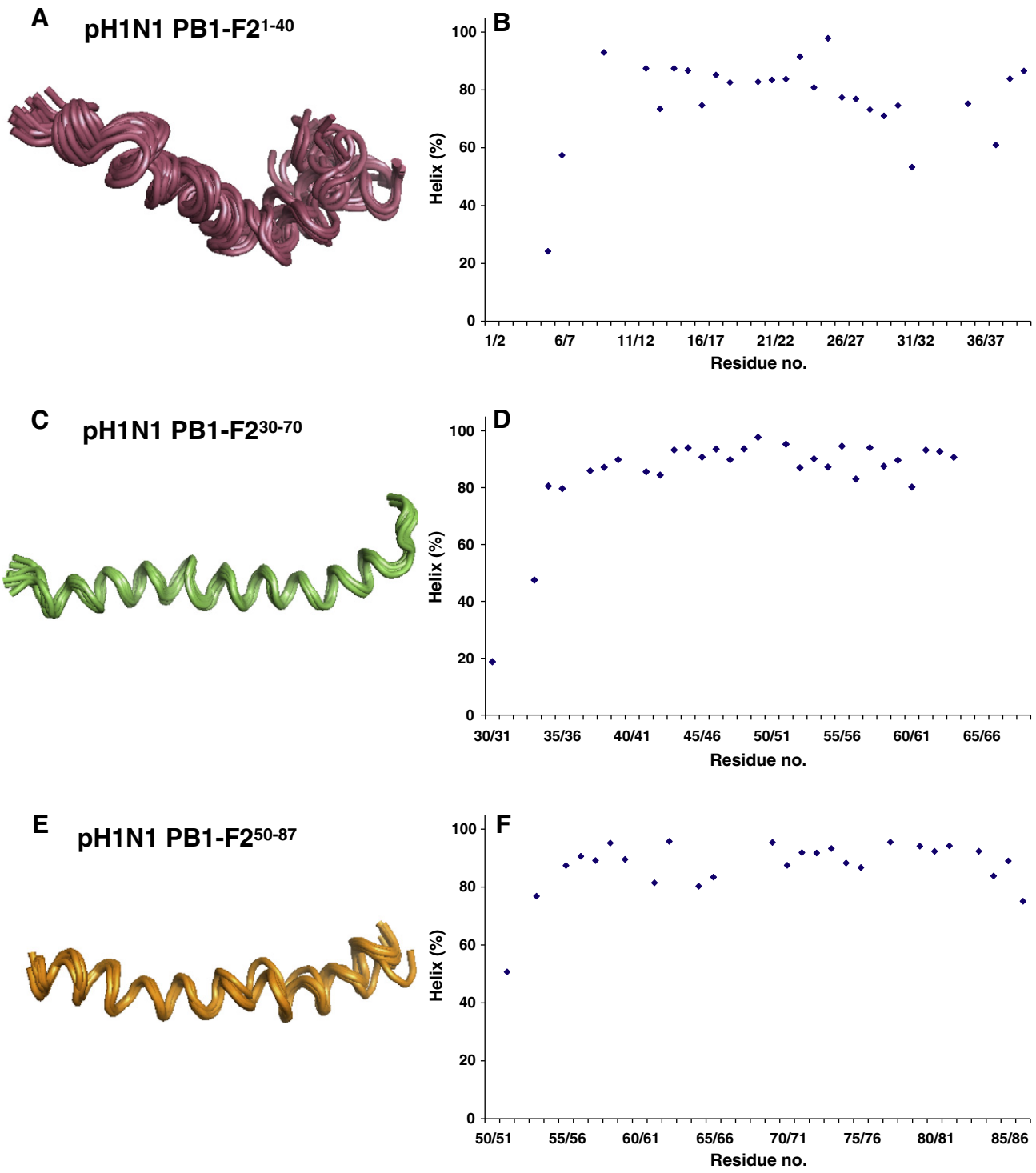

Fig. 5. Calculated structure and helix population estimates of fragments of pH1N1 PB1-F2.

The PB1-F2 sequence of sH1N1 is the one most similar to the sequence of PR8 (Figs. 8 and S19), while the potential full-length PB1-F2 of pH1N1 differ most from the other three IAV strains investigated. A pronounced difference in the region consisting of residues 66 to 71, and a more diffuse region between residues 41 to 53 were observed between the PR8 and H5N1 sequences (Fig. S18). Clearly, both PB1-F2 of sH1N1 and H5N1 show similar types of structural change around residue 70 compared to PR8 (Fig. S18).

3.8. Oligomerization of PB1-F2 proteins

To investigate the self-interaction of PB1-F2 in aqueous solutions, peptide complexes were stabilized by chemical cross-linking and analyzed by SDS-PAGE (Fig. 11). Analyses of the N-terminal fragments revealed that none of the N-termini display oligomerization tendencies. The middle fragments of PR8 and sH1N1 revealed dimerization via disulfide bond formation (Fig. 11B, bands marked with an asterisk).

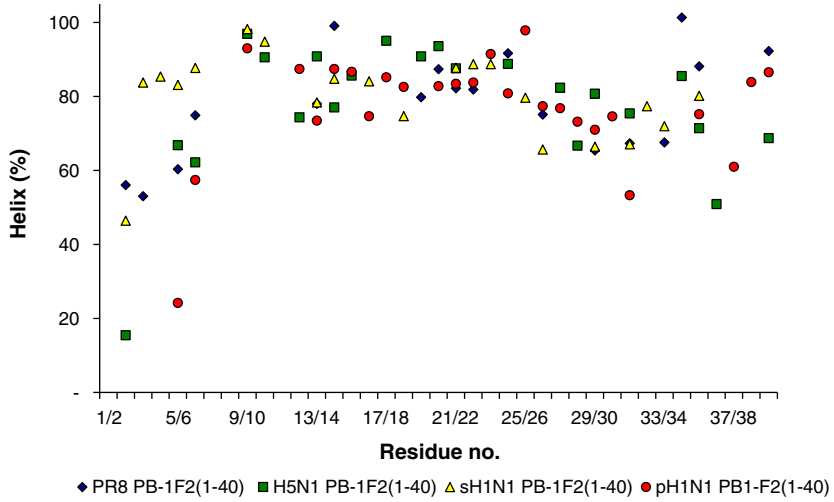

Fig. 6. Helix population estimates of N-terminal PB1-F2 peptides of PR8, H5N1, sH1N1 and pH1N1.

This effect was not observable for the corresponding fragment of H5N1 PB1-F2 as this peptide lacks cysteine residues. Nevertheless, it still displays a pronounced oligomerzation when treated with DSS in contrast to PR8 and sH1N1, independently of the presence of cysteines. The C-terminal fragments of sH1N1 and H5N1 show no oligomerization whereas PR8 reveals multimeric species. Cross-linking of the corresponding full-length peptides reveals that disulfide bond formation in H5N1, PR8, and sH1N1 causes dimerization (Fig. 11C). This finding is unique for full-length H5N1 as it contains a cysteine that is not present in the corresponding middle fragment. This S47C exchange was introduced to allow efficient synthesis of the full-length protein by native chemical ligation [22]. In addition, all full-length peptides show tendencies to oligomerize under influence of DSS independently of cysteines. However, this effect was more pronounced for PR8 and H5N1 where increasing amount of DSS leads to a stronger reduction in the signal intensity of the corresponding monomers as well as an increase of high molecular weight aggregates compared to sH1N1. The differences in the multimerization behavior of the fragments are not found in the corresponding full-length peptides. To provide possible explanations for the observed differences in the oligomerization behavior primary sequence analyses were performed using the TANGO algorithm in order to predict aggregation prone motifs. According to TANGO, a score of $\leq$0.02% indicates no aggregation, 0.02–5.0% indicates moderate aggregation, and $\geq$5.0% indicates high aggregation propensities [49,50]. A weak N-terminal aggregation motif with a score of about 1% was identified, which corresponds to residues 9–13 in the PR8 and sH1N1 sequences (Fig. 11A). In addition, a weak aggregation domain was located more towards the C-terminus in all three protein sequences ranging from residues 54–58 with a score of about 1%. Most interestingly, TANGO identifies a region with high propensity for aggregation at residues 68–72, which is only present in the PR8 sequence (score of 90% per residue). Most intriguingly, this finding correlates with the pronounced oligomerization of the C-terminal PR8 fragment detected by chemical cross-linking.

3.9. Fibril formation of PR8 PB1-F2 peptides in aqueous solution conditions

In order to identify the domains required for amyloid formation, which is normally associated with β-sheet aggregates [51,52], N-terminal PB1-F2$^{1-40}$, middle PB1-F2$^{30-70}$ and C-terminal PB1-F2$^{50-87}$ of the IAV PR8 were studied by electron microscopy in aqueous solution at pH 3. Soluble oligomers from the aqueous solution of PB1-F2$^{30-70}$ were observed shortly after dissolution (Fig. 12). These structures resembled closely the "protofibrils" of the soluble oligomeric intermediates associated with the toxicity of amyloid β and other amyloidogenic proteins [53]. For instance, atomic force microscopic images of amyloid β40 have

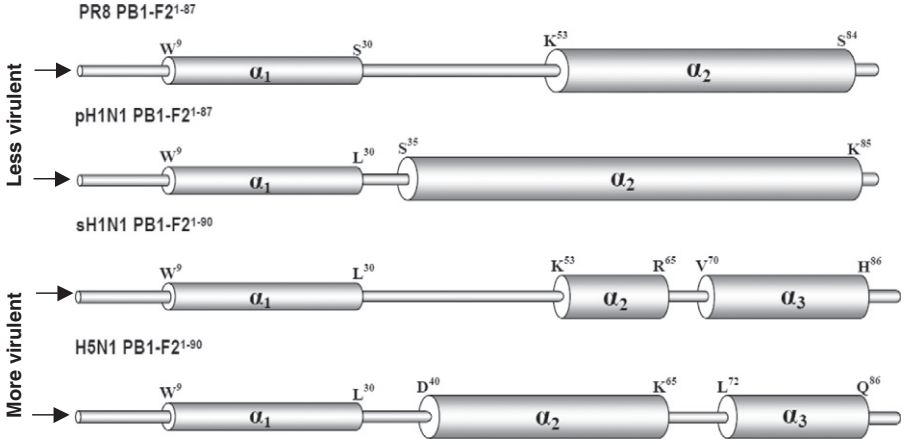

Fig. 7. The overall secondary structure of full-length PB1-F2 of PR8, pH1N1, sH1N1 and H5N1 predicted based on CSI plots, helix population estimates and structural calculations. Regions with the best defined α-helical structure in the C-terminal of the proteins are denoted with broad cylinders, whereas regions with lesser defined α-helical structure in the N-terminals appear as thin cylinders.

Fig. 8. Sequence differences relative to PR8 PB1-F2. Sequence differences of PB1-F2 of pH1N1 (red), sH1N1 (yellow) and H5N1 (green) relative to PR8 (blue font).

shown the same types of curvilinear structures with similar widths and variable lengths. The same morphologies are also evident in electron micrographs of the amyloid β polymer [54]. However, no fibrils were formed in the solutions containing the N- or C-terminal peptides. Interestingly, the behavior of PB1-F2$^{30–70}$ on SDS-PAGE and in dynamic light scattering assays followed that reported for the protofibrils of Aβ42 [54]. Ladders of low molecular weight species ranging from monomer to pentamer were found while light scattering data showed the predominance of polymeric and aggregated proteins.

3.10. Potential formation of β-sheet structure in PB1-F2

In order to investigate the potential existence of β-sheet secondary structure in further detail at atomic resolution, we recorded NMR spectra of the N-terminal, middle and C-terminal fragments of PB1-F2 from the isolates of pH1N1 and FR06 in 80% aqueous ACN-d_3 solution, which was regarded as the optimum condition for the formation of β-sheet structure of PB1-F2 [16].

After completion of the spectral assignment, Hα CSI was plotted for these peptides and compared with CSI of the same peptides in 50% TFE (Figs. 9 and 10). The nature and extent of the secondary structures of these peptides were similar in both solvents (Figs. 9 and 10). β-sheet structure could not be observed in 80% aqueous ACN, but rather a pronounced α-helical structure similar to that seen in 50% TFE was observed. The α-helical structure in 80% ACN was also confirmed by the presence of NOE cross peaks characteristic of α-helical conformations.

Chevalier et al. [16] reported that the most prominent β-sheet structure was observed for PB1-F2 FR06 in 80% aqueous ACN, where the aqueous component was 50 mM acetate buffer at pH 5. In order to determine if β-sheet structure is the predominant folded structure of this isolate of PB1-F2 at this particular pH, NMR spectra of the middle fragment of the isolate FR06 were recorded in 80% aqueous ACN-acetate buffer at pH 5. The middle fragment was chosen since this region is responsible for amyloid formation of PB1-F2 (Fig. 12). However, contrary to what was reported by Chevalier et al., the Hα CSI revealed that the predominant secondary structure of the middle fragment of FR06 PB1-F2 was α-helical (Fig. S20). The observation that the middle fragment of PR8 PB1-F2 forms amyloids indicates that some proportions of PB1-F2 is present as β-sheet aggregates. However, the NMR-spectra recorded of various variants of the proteins demonstrate that the predominant secondary structure of the protein is α-helical (Figs. 3, 5, 9 and 10) under the solution conditions investigated.

4. Discussion

Some functions of PB1-F2 are strain specific and not a general feature of the PB1-F2 protein. The ability to cause cell death through mediated mitochondrial interaction is found to be specific to PR8 [10], while enhanced inflammation has only been reported for PR8, H5N1 and the sH1N1 and some other strains with the PB1 gene segment directly derived from avian reservoirs [8–10]. Interaction between PB1 and PB1-F2 is only seen for PR8 and sH1N1 and not for other IAVs [12]. Of particular importance is the discovery of McAuley et al. (2007) [9] who reported that an influenza virus engineered to express a PB1-F2 with coding changes matching the 1918 pandemic strain induced more pulmonary immunopathology, led to more severe secondary bacterial pneumonia and was more virulent than WT in mice.

The fact that some functions of PB1-F2 are found to be strain specific and that PB1-F2 is proposed to be an important virulence factor encouraged us to investigate the possibility of a structural signature in PB1-F2 common for the more virulent IAV strains. Consequently, we have investigated the structure of PB1-F2 proteins of sH1N1, H5N1, PR8 and pH1N1 by NMR and CD spectroscopy. Our experimental data have confirmed that there are substantial differences in the secondary structure of the C-terminal regions of PB1-F2 in 50% TFE solution, with a characteristic helix break seen for sH1N1 and H5N1, but not for PR8 and pH1N1. Instead, PB1-F2 proteins of PR8 and pH1N1 contain one extensive continuous α-helix under membranous solution conditions. When comparing the N-terminal regions of the four IAV strains, we found that they exhibit a similar extent of weak α-helical structure. The Hα CSI plots indicate that the structure of the N-terminus of PR8 PB1-F2 is most similar to sH1N1 PB1-F2 (Figs. 3–4), which is a direct precursor of the PR8 virus. More surprisingly, the structure of the N-terminus of pH1N1 PB1-F2 is most similar to that of H5N1 PB1-F2 (Figs. 3–4). The structural calculations and helix population plots of the N-terminal regions showed that this region of PR8 is more structured than previously recognized (Figs. 5–7). However, the α-helical structure of the N-terminus is not as well defined as the α-helical structure of the middle and C-terminal regions (Figs. 5–6).

Most of the known functions of PB1-F2 are associated with the C-terminal region of the protein. The functional mitochondrial targeting sequence (MTS), consisting of residues 46–75 of PB1-F2, known to be involved in causing cell death through mediated mitochondrial interaction [13,14,56], and the protein's ability to enhance the inflammatory responses are both associated with the C-terminal region [9,55,57]. Furthermore, the C-terminal peptide PR8 PB1-F2$^{61–87}$ acted synergistically with bacterial pneumonia co-infection in mice [9,55]. A C-terminal Ser residue at position 66 in the sequence of sH1N1 PB1-F2 has been associated with increased virulence by suppressing the host innate immune response pathways early in viral infection [8,58,59]. Similar immune suppression is seen when the mutation N66S is introduced in PB1-F2 for PR8 and H5N1 [8,58]. In this context it is therefore not surprising that we find significant structural differences in PB1-F2 in the C-terminal region among different IAVs. Residue Ser-66 is situated in the center of the region where the C-terminal helix break is observed for sH1N1 (Fig. 3). When comparing the primary sequences of PB1-F2 from sH1N1 with that of PR8 it becomes evident that the region showing the greatest differences consists of residues 66–71, i.e. the residues that form the interhelical loop region of sH1N1 PB1-F2 (Fig. 8). The residue mainly responsible for the observed helix break in sH1N1 is the additional Pro residue present at position 69. Thus, the functionally important region containing Ser-66 comprises an interhelical loop region of sH1N1 PB1-F2. Previous reports have suggested that introduction of a Ser residue at position 66 of PB1-F2 pH1N1 may lead to a more virulent virus. However, Ser-66 is not the only difference between PR8 and sH1N1 PB1-F2, but the effect of substituting other residues in the region 66–71 has hitherto not been reported. The main structural difference between PR8 PB1-F2 and sH1N1 PB1-F2 is located within these C-terminal residues, where four out of six residues are different (Fig. 8). The N-terminal regions of these proteins are relatively similar. The primary 40-residue N-terminal sequence only harbors 4 mutations (at positions 22, 29, 33 and 40), which according to our structural data do not have any significant effect on the secondary structure of the N-terminal region of PB1-F2 (Figs. 4–6).

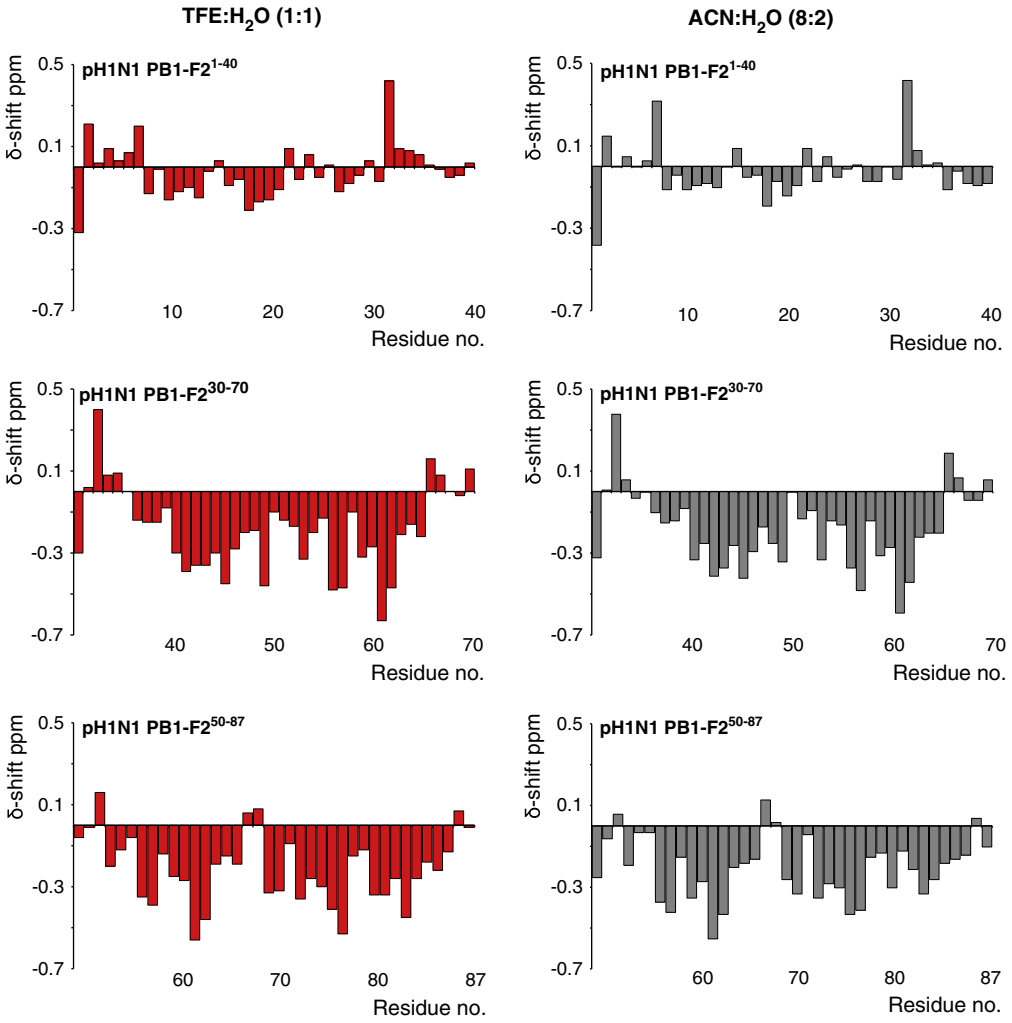

Fig. 9. CSI plots of pH1N1 PB1-F2 fragments determined from NMR spectra of peptides dissolved in TFE:H_2O (1:1) (red) and ACN:H_2O (8:2) (gray). No differences were observed in the secondary structure at the two different solution conditions.

Only full-length PB1-F2 with an intact C-terminus was found to co-localize with PB1 [11,12]. This interaction has been suggested to correlate with increased influenza polymerase activity and replication, and thus to increase the virulence [11]. However, more recently, McAuley et al. did not find any correlation between the presence of PB1-F2 and polymerase activity and replication [12], and thus the importance of this interaction remains unknown. Intriguingly, Kosik et al. reported that the N-terminal region of PB1-F2 is responsible for an increase in PB1 protein expression [60], and thereby is the first study to connect the N-terminal part of PB1-F2 to any of its biological functions. The structures of the N-terminal regions reported in this paper have similarly weak α-helical structure for all four IAV strains investigated. This suggests that any function associated with this region of PB1-F2, may be conserved in several IAV strains, including those with a C-terminally truncated PB1-F2 protein.

A study investigating the potential influence of a full-length PB1-F2 on pathogenicity of pH1N1 reported only a minor influence on the virulence of the virus [20]. This indicates that a full-length PB1-F2 is not in itself a signature for virulence. However, PB1-F2 of pH1N1 differs from sH1N1 and H5N1 by having only one continuous C-terminal α-helix. This was also found for PB1-F2 of PR8, although it has a shorter C-terminal helix. The ability to cause cell death through mitochondrial interaction appears to be specific to PR8 [10] and pore formation mediated by the C-terminal α-helix is one of the suggested mechanisms [17]. This may indicate that other PB1-F2 variants with a continuous extensive C-terminal α-helix like pH1N1 PB1-F2 could also induce apoptosis by the same mechanism. The fact that sH1N1 PB1-F2 and H5N1 PB1-F2 did not cause cell death through mediated mitochondrial interaction [10] is in accordance with our observations that these proteins do not contain a continuous C-terminal α-helix that could be assigned to this function (Fig. 7).

Recently, Chevalier et al. reported that PB1-F2 of several isolates are able to adopt β-sheet structure as an alternative folding state of the protein in 80% aqueous ACN, 0.01% aqueous SDS and asolectin solutions in addition to amyloid formation observed by electron microscopy [16]. Amyloid formation of a protein is usually associated with the existence of significant proportions of β-sheet aggregates [51,52]. Identification of β-sheet structural elements in PB1-F2 was proposed solely from CD spectroscopy data [16]. CD spectroscopy is considered as a rather indicative method for determining secondary structure of proteins, as this technique does not provide data at atomic resolution [61]. In order to investigate this in further detail by NMR, we prepared pH1N1 and FR06 fragments in 80% ACN, which was reported as optimum conditions for formation of β-sheet structures [16]. Chevalier et al. [16] reported that PB1-F2 from the FR06 strain showed the most pronounced β-sheet structure in 80% ACN. Therefore, this variant was also included in our study. However, when comparing the CSI plots of these fragments dissolved in 50% TFE and 80% ACN, respectively (Figs. 9 and

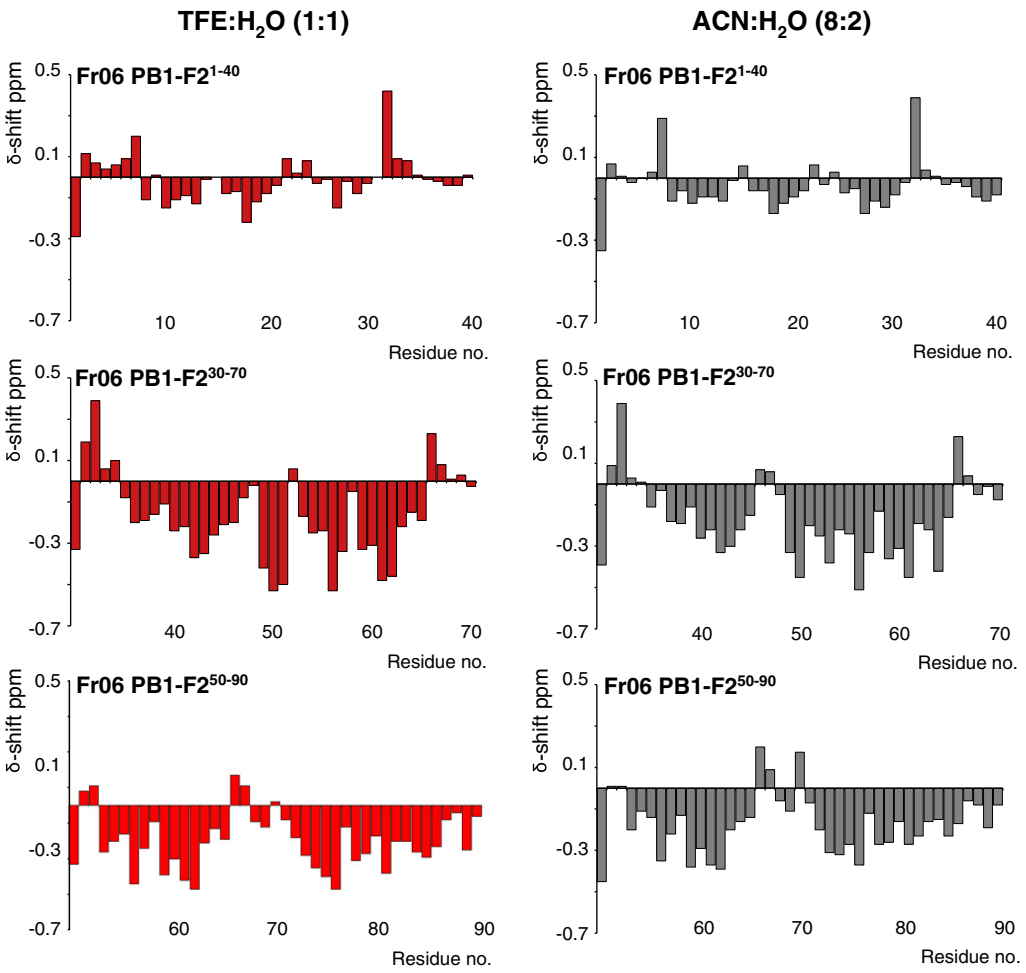

Fig. 10. CSI plots of FR06 PB1-F2 fragments determined from NMR spectra of peptides dissolved in TFE:H_2O (1:1) (red) and ACN:H_2O (8:2) (gray). No differences were observed in the secondary structure of the middle and C-terminal regions at the two different solution conditions.

10), no significant differences between the secondary structures were observed and the predominant structure was α-helical in both media.

Previously, we determined that the secondary structural elements of PB1-F2 are α-helices at membranous solution conditions [22]. Using similar experimental conditions, Chevalier et al. confirmed that PB1-F2 adopts an α-helical structure in membrane-mimicking 50% aqueous TFE solution [16]. However, in 80% aqueous ACN solution at pH 5, CD spectra of several variants PB1-F2 were indicative of the presence of β-sheet conformation [16]. Additionally, CD spectra of PB1-F2 FR06 and MA65 recorded in the presence of a detergent (0.01% SDS) were also indicative of some proportions of β-sheet conformation [16]. Furthermore, Chevalier et al. suggested that the variants of PB1-F2 also form varying proportions of β-sheet conformation in an alternative membrane model solution (0.1 mg/ml asolectin liposomes). However, this suggestion is not supported by the CD spectra presented in the publication (Fig. 2E in Chevalier et al. [16]), which do not show any of the characteristics of CD spectra of proteins with significant proportions of β-sheet conformation (i.e., a maximum around 195 nm and a minimum around 218 nm), but rather resemble that of weak α-helical structures with negative bands at approximately 222 and 208 nm and a positive band around 193 nm [61–63]. Moreover, CD spectra of PB1-F2 from PR8 and sH1N1 in SDS (Fig. 2D in Chevalier et al. [16]) exhibited curves typical of peptides with an α-helical structure, although the authors suggested that this could be related to a β-sheet structure. The CD spectrum of PB1-F2 FR06 in 50% aqueous ACN is typical of a peptide with an α-helical structure (Fig. 2B in Chevalier et al. [16]).

Amyloidal fibers play a role in multiple diseases, most commonly in neurodegenerative diseases but recently these fibers were also found to enhance human immunodeficiency virus infection [64]. Chevalier et al., reported formation of amyloid-like fibrils of the IAV protein PB1-F2 in infected cells in the presence of 0.01% SDS [16]. In this paper we have established that the residues 30–70 of PR8 PB1-F2 are responsible for amyloid formation (Fig. 12). Fibril formation has been coupled to alteration of β-sheet rich conformation that facilitates conversion into a cross-β amyloid structure [51,52]. However, the substantial structural data at atomic resolution for several variants of PB1-F2 presented in this paper show that β-sheets are not the main secondary structure of these proteins in solution. At membranous solution conditions α-helices are the only secondary structural features of PB1-F2 observable by NMR spectroscopy. In general, the structure of PB1-F2 is found to be random at aqueous solution conditions. However, the fact that a CD spectrum typical of the presence of a weak β-sheet structure was recorded for the middle fragment of sH1N1 PB1-F2 in aqueous phosphate buffer at pH 7 and the observations of amyloid formation of the middle fragment of PR8 PB1-F2 in aqueous solution at pH 3 indicate that PB1-F2 may be present in several conformational forms including β-sheet aggregates, α-helical structures and random coil depending upon the environment.