

# Solution NMR and molecular dynamics reveal a persistent alpha helix within the dynamic region of PsbQ from photosystem II of higher plants

Petr Rathner,<sup>1,2</sup> Adriana Rathner,<sup>1</sup> Michaela Horníčáková,<sup>1,3</sup> Christian Wohlschlager,<sup>4</sup> Kousik Chandra,<sup>1</sup> Jaroslava Kohoutová,<sup>2,5</sup> Rüdiger Etrich,<sup>2,5</sup> Reinhard Wimmer,<sup>6</sup> and Norbert Müller<sup>1,2\*</sup>

<sup>1</sup> Institute of Organic Chemistry, Johannes Kepler University Linz, Linz 4040, Austria

<sup>2</sup> Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic

<sup>3</sup> Lohmann Animal Health, Cuxhaven 27472, Germany

<sup>4</sup> Scientific Computing, Johannes Kepler University Linz, Linz 4040, Austria

<sup>5</sup> Center of Nanobiology and Structural Biology, Institute of Microbiology, Academy of Sciences of the Czech Republic, Nové Hradky, Czech Republic

<sup>6</sup> Department of Biotechnology, Chemistry and Environmental Engineering, Aalborg University, Aalborg 9220, Denmark

## ABSTRACT

The extrinsic proteins of photosystem II of higher plants and green algae PsbO, PsbP, PsbQ, and PsbR are essential for stable oxygen production in the oxygen evolving center. In the available X-ray crystallographic structure of higher plant PsbQ residues S14–Y33 are missing. Building on the backbone NMR assignment of PsbQ, which includes this “missing link”, we report the extended resonance assignment including side chain atoms. Based on nuclear Overhauser effect spectra a high resolution solution structure of PsbQ with a backbone RMSD of 0.81 Å was obtained from torsion angle dynamics. Within the N-terminal residues 1–45 the solution structure deviates significantly from the X-ray crystallographic one, while the four-helix bundle core found previously is confirmed. A short  $\alpha$ -helix is observed in the solution structure at the location where a  $\beta$ -strand had been proposed in the earlier crystallographic study. NMR relaxation data and unrestrained molecular dynamics simulations corroborate that the N-terminal region behaves as a flexible tail with a persistent short local helical secondary structure, while no indications of forming a  $\beta$ -strand are found.

Proteins 2015; 83:1677–1686.

© 2015 The Authors. Proteins: Structure, Function, and Bioinformatics Published by Wiley Periodicals, Inc.

**Key words:** dynamic N-terminus; extrinsic photosynthetic protein; hydrogen bond dynamics; intrinsic disorder; solution structure; *Spinacia oleracea*.

## INTRODUCTION

Photosystem II (PSII) accomplishes arguably the most essential process for aerobic life on Earth, which is the water splitting reaction. The mechanism of oxygen generation has remained conserved during the evolution for more than two billion years.<sup>1</sup> In higher plants and cyanobacteria PSII consists of several intrinsic membrane-spanning and a number of extrinsic proteins on the lumenal side of thylakoid. In contrast to the highly conserved intrinsic proteins, the extrinsic photosynthetic proteins forming the oxygen evolving complex (OEC) vary considerably between various photosynthetic organisms. To sustain a high rate of oxygen evolution in

Additional Supporting Information may be found in the online version of this article.

**Abbreviations:** OEC, oxygen evolving complex; PSII, Photosystem II; RMSE, root mean square fluctuations

Grant sponsor: JKU scholarship; Grant sponsor: ÖAD; Grant number: AKTION 69p26; Grant sponsor: Austrian Science Funds FWF project NanoCell; Grant number: W1250; Grant sponsor: Austrian Science Funds FWF; Grant number: M1404 (to K.C.).

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

N.M. and R.W. dedicate this article to Prof. Karl Grubmayr on the occasion of his 65<sup>th</sup> birthday.

\*Correspondence to: Norbert Müller, Institute of Organic Chemistry, Johannes Kepler University Linz, Altenbergerstraße 69, Linz 4040, Austria. E-mail: norbert.mueller@jku.at

Received 27 April 2015; Revised 16 June 2015; Accepted 24 June 2015

Published online 2 July 2015 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/prot.24853

higher plants, the extrinsic proteins PsbO (33 kDa), PsbP (23 kDa), PsbQ (16.5 kDa), and PsbR (10 kDa) are essential.<sup>2,3</sup> In cyanobacteria, the PsbP and PsbQ homologues (denoted as CyanoP and CyanoQ, respectively) are present together with the PsbU and PsbV extrinsic proteins.<sup>4,5</sup> The main functions of these loosely attached extrinsic proteins are perceived to be the protection of the highly reactive manganese cluster from exogenous reductants and the regulation of the ionic environment.<sup>6,7</sup> Among the extrinsic proteins of PSII, PsbO is so far the most extensively studied one, a fact probably owed to its ubiquity in all photosynthetic species. Preliminary NMR investigations of PsbO from *T. elongatus* indicated, that PsbO is a folded protein with numerous disordered regions.<sup>8</sup>

Concerning the PsbQ and PsbP proteins, it is supposed that their interactions play a crucial role in the regulation of calcium and chloride ion concentrations within the PSII complex.<sup>9</sup> Recent studies showed that PsbQ can compensate for a common structural defect of PsbP, that is, truncation of the 15 N-terminal residues which are required for binding to PSII. PsbQ thus contributes to the protection of the catalytic manganese cluster.<sup>10</sup> It has been shown that the PsbP and PsbQ proteins may also directly associate with PSII intrinsic subunits.<sup>11</sup> High resolution crystal structures of PsbP and PsbQ from higher plants<sup>12–15</sup> and cyanobacteria<sup>16</sup> have been reported. In all crystallographic structures of PsbP and PsbQ, including Psb31 from diatom algae, significant parts of the sequences are missing.<sup>12–17</sup> This applies also to the PsbQ protein from *Spinacia oleracea* (spinach). A previous crystallographic study of spinach PsbQ found a well-defined C-terminal four-helix-bundle and a loosely packed N-terminus, from which 20 residues (S14–Y33) are missing.<sup>14</sup> Recent cross-linking experiments of PsbP and PsbQ proteins revealed a close interaction of K176 in PsbP with D28 in PsbQ.<sup>18</sup> Since this cross-linked site of PsbQ is not visible in the electron density map, it was supposed that the N-terminal stretch of free PsbQ is very flexible and probably extended in solution.<sup>14</sup> Previous NMR studies of the secondary structure of PsbQ in solution indicated some residual order in the so far unresolved dynamic N-terminal region<sup>19</sup> which was however inconsistent with the short parallel  $\beta$ -sheet (I5–V7 and F38–L40) reported in the crystallographic structure.<sup>14</sup>

In this article, we present the nearly complete <sup>1</sup>H, <sup>15</sup>N and <sup>13</sup>C backbone and side-chain NMR assignments of recombinant PsbQ from *Spinacia oleracea* together with its high resolution NMR structure in solution. Supporting information obtained from <sup>15</sup>N NMR relaxation, {<sup>1</sup>H}<sup>15</sup>N NOE experiments and unrestrained MD simulations yielded first information on internal backbone dynamics. The N-terminal segment, which apparently is essential for the interactions of PsbQ with other constit-

uents at the luminal side of the thylakoid near PSII, is a main point of interest in this investigation.

## MATERIALS AND METHODS

### Protein preparation

Cloning, overexpression and purification of uniformly doubly labeled <sup>15</sup>N, <sup>13</sup>C PsbQ from *Spinacia oleracea* in *E. coli* were described previously.<sup>19</sup> The samples for NMR spectroscopy contained 800  $\mu$ M <sup>15</sup>N,<sup>13</sup>C PsbQ, 20 mM Na<sub>2</sub>HPO<sub>4</sub>, 50  $\mu$ M NaN<sub>3</sub>, 1 mM EDTA (pH = 7.0) in H<sub>2</sub>O: <sup>2</sup>H<sub>2</sub>O 90:10.

### NMR spectroscopy

All NMR experiments were performed on 700 MHz Bruker Avance III spectrometer equipped with TCI cryogenically cooled probe (20 K) at 298 K sample temperature. HC(C)H-TOCSY,<sup>20</sup> (H)C(C)(CO)NH,<sup>21</sup> HBHA(CO)NH,<sup>22</sup> HC(C)H-COSY,<sup>23–25</sup> TOCSY-HSQC,<sup>26</sup> (HB)CB(CGCD)HD and (HB)CB(CGCDCE)HE<sup>27</sup> experiments were used to complement the proton and carbon shifts published previously (BMRB entry 17357).<sup>19</sup> The improved and extended set of assignments has been submitted to BMRB (entry 25350). To obtain information about the internal dynamics of PsbQ, <sup>15</sup>N backbone relaxation experiments were investigated. <sup>15</sup>N-T<sub>1</sub> and <sup>15</sup>N-T<sub>2</sub> relaxation rates<sup>28</sup> were computed by exponential fitting of peak heights in sets of different <sup>15</sup>N-HSQC phase sensitive experiments with different relaxation delays. T<sub>1</sub> relaxation delays: 10, 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, and 1200 ms. T<sub>2</sub> relaxations delays: 10, 15, 20, 25, 35, 40, 50, 60, 70, 80, 90, 100, and 120 ms. The backbone {<sup>1</sup>H}<sup>15</sup>N NOEs were calculated from ratios of peak heights from two heteronuclear NOESY <sup>15</sup>N-<sup>1</sup>H sensitivity enhanced HSQC-type experiments.<sup>28</sup>

Subsequently, three-dimensional <sup>15</sup>N-edited (mixing time 60 ms) and <sup>13</sup>C-edited (aliphatic and aromatic, mixing time 100 ms) NOESY-HSQC experiments<sup>26,29,30</sup> were involved to obtain distance restraints. All NMR data were recorded and processed with Bruker Topspin software (v. 3.1 and 3.2). Spectra analysis including resonance assignment was carried out manually with the help of CARA software.<sup>31</sup>

### Structure calculation

NOESY cross-peaks were integrated by the NEASY subroutine<sup>32</sup> of CARA and converted into upper distance constraints using the CALIBA subroutine.<sup>33</sup> Dihedral angle constraints were estimated from backbone chemical shifts using TALOS-N.<sup>34</sup>  $\phi$  and  $\psi$  angles of residues whose multiple-database prediction did not match within the same region of the Ramachandran map, were discarded. Upper-distance- as well as dihedral-angle-constraints were used as input for the structure

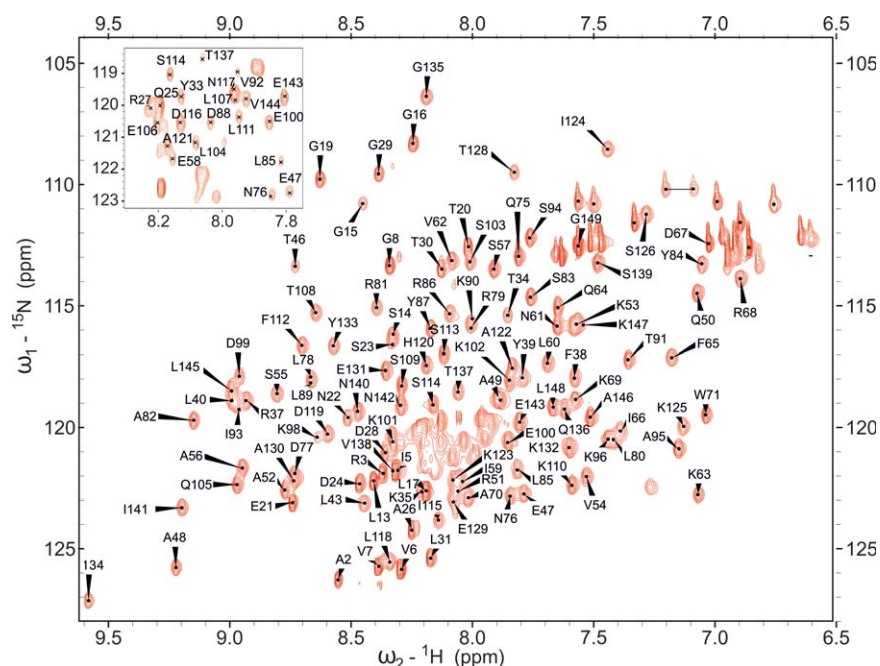
calculations using the torsion angle dynamics program CYANA (v. 3.0).<sup>35</sup> In total 100 structures were calculated in 10,000 annealing steps. The 20 conformers with lowest CYANA target function values were refined in a water box by restrained MD simulations using the YASARA force field<sup>36</sup> implementation in YASARA 12.1.19.<sup>37</sup> The atomic coordinates of the refined bundle of 20 conformers shown in Figures 2 and 3(A) have been deposited in the protein database under PDB ID: 2MWQ.

### Molecular dynamics simulations

As a major part of the N-terminal sequence is unresolved in the X-ray structure of PsbQ, an earlier reported model<sup>38</sup> based on the X-ray structure and loop modeling has been used as representative for the X-ray structure. Of the NMR structures a conformer out of the middle of the conformational ensemble was selected to represent the NMR structure. GROMACS 4.6.3 package<sup>39–41</sup> was used for preparing the system and performing MD simulations using the Amber99 SB force field.<sup>42</sup> Both protein structures were solvated by explicit TIP3P water<sup>43</sup> in a cubic box with a margin of 10 Å between solute and the box walls. Systems were neutralized by addition of sodium counter ions. The particle-mesh Ewald method<sup>44</sup> was applied to calculate long-range electrostatic interactions with a cut-off distance of 10 Å. A Lennard-Jones 6–12 potential was used to evalu-

ate van der Waals interactions within 10 Å cut-off distance. The LINCS algorithm of fourth order expansion was used to constrain bond lengths.<sup>45</sup> After solvation and neutralization steps the system was optimized for 10000 steps using the steepest-descent method to remove steric clashes between atoms. The system was equilibrated for 1 ns with position restraints of 1000 kJ/mol on all heavy atoms. A constant temperature of 300 K or 310 K was maintained using the V-rescale algorithm<sup>46</sup> with a coupling time of 0.1 ps and separate baths for the solute and the solvent. The pressure was kept constant at 1 bar using the Parrinello-Rahman pressure coupling scheme<sup>47</sup> with a time constant of 2 ps. Initial velocities were generated randomly using a Maxwell-Boltzmann distribution corresponding to 300 K and additionally for 310 K in case of the X-ray structural model. Neighbour lists were updated every 10 fs using a group cut-off scheme. Finally the production runs were performed in the isothermal-isobaric (NPT) ensemble without restraints for 225 ns for the X-ray structure at 300 K and 310 K, and for 300 ns for the NMR-structure at 300 K. The grmf tool of the GROMACS package was used to calculate root mean square fluctuations (RMSF) during the last 20 ns of trajectories.

VMD<sup>48</sup> was used to calculate interatomic distances and for manual inspection of trajectories. We employed xmgrace (<http://plasma-gate.weizmann.ac.il/Grace/>) for pre paring the graphs.



**Figure 1**

2D [<sup>1</sup>H-<sup>15</sup>N]-HSQC spectrum of 0.8 mM <sup>15</sup>N, <sup>13</sup>C-PsbQ in 20 mM KH<sub>2</sub>PO<sub>4</sub>, 1 mM EDTA, pH = 7.0 (H<sub>2</sub>O:D<sub>2</sub>O 90:10 v/v), T = 298 K.

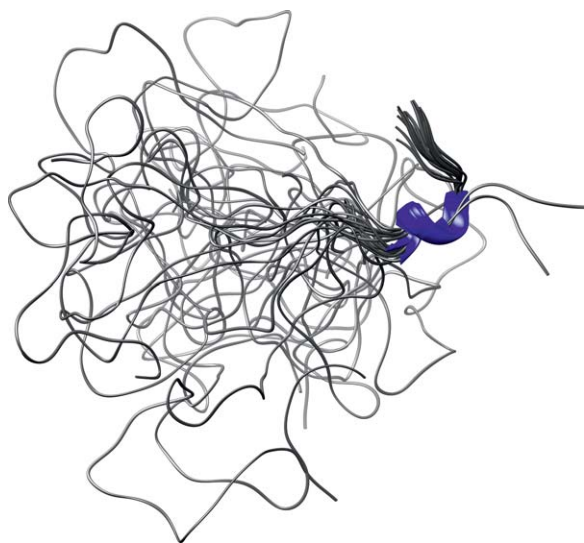
## RESULTS AND DISCUSSION

### NMR assignment

Employing the set of experiments summarized in the section on NMR spectroscopy, the resonance assignments (BMRB entry 17357) of PsbQ from *Spinacia oleracea* were extended and improved. In total 718 of 739 backbone (97.2%) and 839 of 1173 (71.5%) side chain resonance chemical shifts were assigned as compared to 80.7% and 40.1% of backbone and side chain resonances in the earlier assignment,<sup>19</sup> respectively. The backbone amide resonances of five non-proline amide residues (E1, D36, Q41, S73, and L74) could not be identified or assigned unequivocally due to fast NH exchange and/or severe peak overlap in the <sup>15</sup>N-HSQC (Fig. 1). <30% of the side-chain resonances remained unassigned due to signal overlap in HC(C)H-COSY and TOCSY-HSQC spectra.

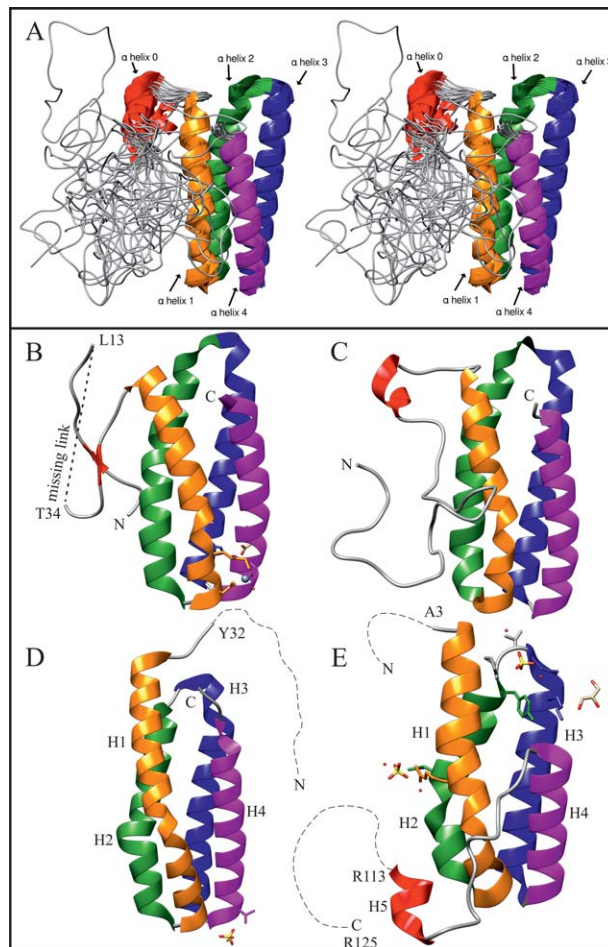
### Restraints, structure calculation, and analysis

Through assignment of the cross-peaks in the <sup>15</sup>N-, <sup>13</sup>C<sub>ali</sub>- and <sup>13</sup>C<sub>aro</sub>-NOESY-HSQC experiments, typical  $\alpha$ -helical NOE patterns [ $d_{NN}(i,i+1)$ ,  $d_{\alpha N}(i,i+1)$ ,  $d_{\beta N}(i,i+1)$ , and  $d_{\alpha N}(i,i+3)$ ] were found in all helices proposed in our previous study<sup>19</sup> largely coinciding with the helix-bundle identified by X-ray crystallography.<sup>14</sup> Only a very limited number of NOE cross-peaks could be found involving the first 35 residues, presumably due to the high mobility of the N-terminus. The majority of NOE constraints was found in the well-defined C-



**Figure 2**

Superimposed backbone structure representations of the N-terminal regions (E1-P45) of 20 water refined conformers aligned at the  $\alpha$  helix 0 (R37-L40).



**Figure 3**

(A) Stereoscopic ribbon representation of 20 spatially aligned (at the region of amino acid residues 45–149) water refined conformers. (B) Crystal structure of spinach PsbQ.<sup>14</sup> (C) NMR solution structure of spinach PsbQ with the lowest restraint violation energy. (D) Crystal structure of PsbQ from *Synechocystis* sp. PCC 6803.<sup>16</sup> (E) Crystal structure of diatom Psb31.<sup>17</sup> The dashed lines represent unresolved regions. The figure was produced using UCSF Chimera.<sup>51</sup>

terminal region of the molecule, mainly from the <sup>15</sup>N-NOESY-HSQC spectrum. 177 long range NOE cross-peaks ( $|i-j| \geq 5$ ) were assigned. Note that PsbQ contains 13 proline residues, which lack amide protons. In total, 872 experimentally derived upper distance limits and 187 angular constraints were used as input information for the CYANA structure calculations. Subsequent rMD simulations using the YASARA force field<sup>36</sup> yielded a final set of 20 conformers without any experimental distance violation larger than 0.1 Å. The ensuing structural statistics are summarized in Table I.

The backbone ( $C^\alpha$ ,  $C'$ , N) RMSD for the region of the four  $\alpha$ -helices (residues 46–149) is  $0.81 \pm 0.15$  Å, while for the N-terminal part (residues 1–36) we find only an RMSD of  $7.74 \pm 1.50$  Å. However, a short  $\alpha$ -helix, which we call 0 remains persistent among the 20 low energy

**Table 1**

Structural Statistics for the Energy-Minimized NMR Solution Structure of PsbQ

Experimental constraints	
NOE based distance constraints	
Total	872
Intra-residue [ $i = j$ ]	262
Sequential [ $ i - j  = 1$ ]	315
Medium range: [ $1 <  i - j  < 5$ ]	118
Long range: [ $ i - j  \geq 5$ ]	177
NOE constraints per restrained residue	6.3
Dihedral-angle constraints <sup>a</sup> :	
Total number of restricting constraints	1059
Total number of restricting constraints per restrained residue	7.6
Violations	
Max. distance violation (Å)	0.10
RMS of distance violation/constraint (Å)	0.01
Maximal torsion angle constraint violation (°)	5.00
RMS of dihedral angle violation/structure (°)	0.19
Structure quality	
CYANA target function value (Å <sup>2</sup> ) <sup>b</sup>	2.24 ± 0.31
Backbone [C <sup>α</sup> , C', N] RMSD from average (Å) <sup>c</sup>	0.81 ± 0.15
Heavy atoms RMSD from average (Å) <sup>c</sup>	1.63 ± 0.20
Total energy (kJ·mol <sup>-1</sup> )	-54429.0 ± 521.2
Ramachandran plot summary for selected residues <sup>d</sup> from Procheck <sup>52</sup>	
Most favored regions	95.9%
Additionally allowed regions	4.0%
Generously allowed regions	0.0%
Disallowed regions	0.0%

<sup>a</sup>Derived from backbone chemical shifts by Talos-N.<sup>34</sup><sup>b</sup>Before water refinement.<sup>c</sup>Calculated for non-mobileresidues (45–149) after water refinement.<sup>d</sup>Residues 9–11, 37–95, 99–148.

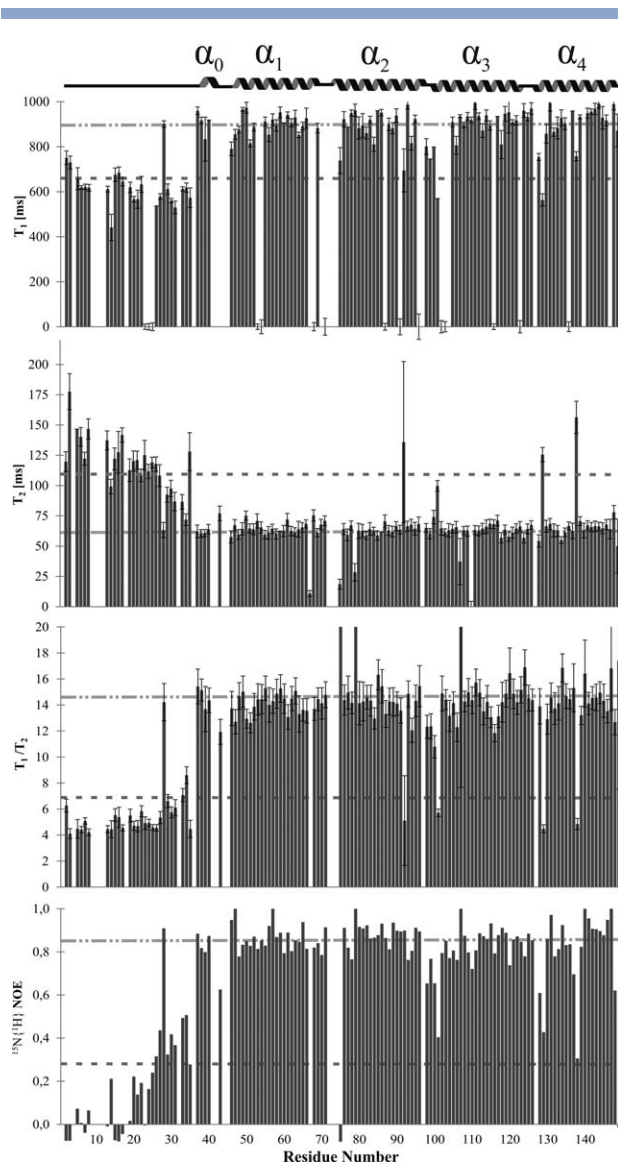
structures (Fig. 2). The backbone (C<sup>α</sup>, C', N) RMSD for these four residues is  $0.24 \pm 0.13$  Å.

In summary, the solution structure of PsbQ from *Spirinacia oleracea* can be described as belonging to an all- $\alpha$  protein class with a bundle of four up-down-up-down  $\alpha$ -helices and a flexible N-terminus [Fig. 3(A)]. In comparison with the most up-to-date available crystallographic structure (PDB: 1VYK), the larger C-terminal part (residues 46–149) corresponding to the four-helix bundle remains largely conserved in the mean solution structure with a backbone (C<sup>α</sup>, C', N) RMSD of 1.27 Å with respect to the X-ray structure, while the N-terminus emerges as predominantly unfolded, with one important exception. Based on the experimental NOE data, residues 37–40 located in the dynamic N-terminus of the molecule exhibit high  $\alpha$ -helical character. Remarkably, this is exactly at the location where one of the two short  $\beta$ -strands had been proposed in the crystal structure [Fig. 3(B)]<sup>14</sup> and by Raman spectroscopy.<sup>38</sup> No indication of any well-defined secondary structure could be found for residues 5–7 where the complementary  $\beta$ -strand was reported in the X-ray structure. In the crystallographic structure of a related cyanobacterial protein (PsbQ from *Synechocystis sp.* PCC 6803)<sup>16</sup> [Fig. 3(D)] helix H1, corresponding to  $\alpha$ -helix 1 in our structure, is extended toward the N-terminus but ends short of the location of the short helix 0 in the solution structure of spinach PsbQ. In the Supporting Information the experimental

NOE-patterns in this range are compared to predictions of NOE patterns corresponding to the X-ray structure, clearly corroborating the existence of this  $\alpha$ -helix. Interestingly, the NMR solution structure appears to be very similar to the crystallographic structure of Psb31 from the diatom *Chaetoceros gracilis*, although the sequence similarity between Psb31 and spinach PsbQ, compared with other PsbQ-like proteins, is very low (25.5%).<sup>17</sup> However, while PsbQ has a flexible N-terminal stretch with an embedded short  $\alpha$ -helix, Psb31 exhibits a similar feature on the C-terminus [Fig. 3(E)].

### Backbone atom relaxation studies

Since the current and previous experimental studies have indicated high internal mobility of the N-terminal segment of PsbQ, we used dynamic NMR methods <sup>15</sup>N longitudinal and transverse relaxation times  $T_1$ ,  $T_2$ , respectively, and the heteronuclear {<sup>1</sup>H}<sup>15</sup>N NOEs for backbone dynamics analysis summarized in Figure 4. The backbone <sup>15</sup>N  $T_1$ , and  $T_2$  relaxation times as well as the {<sup>1</sup>H}<sup>15</sup>N NOEs data corroborate the mobility information obtained in the experimental <sup>15</sup>N and <sup>13</sup>C filtered <sup>1</sup>H-<sup>1</sup>H NOE, apart from a few outliers. For several residues including the 13 prolines (P4, P9, P10, P11, P12, P18, P32, P42, P44, P45, P72, P97, P127) and the 6 unassigned residues' (E1, D36, Q41, D67, S73, L74) {<sup>1</sup>H}<sup>15</sup>N NOEs and relaxation data could not be



**Figure 4**

Heteronuclear  $\{^1\text{H}\}^{15}\text{N}$  NOEs enhancements,  $^{15}\text{N}$   $T_1$  and  $^{15}\text{N}$   $T_2$  relaxation rates of PsbQ. The NOE enhancements show small values for residues 2–35 indicating high backbone flexibility in this region. Notably, the values for residues 37–41 are higher, indicating conformational stability of the short  $\alpha$  helix 0 [Fig. 3(A)] found by the homonuclear NOE experiments. Correspondingly, the  $T_1$  and  $T_2$  values indicate the reduced mobility of residues 37–41 and the dynamic nature of the conformation within the first 35 residues. The heteronuclear NOEs and relaxation rates confirm rigidity of the four-helix bundle in Figure 3(A).

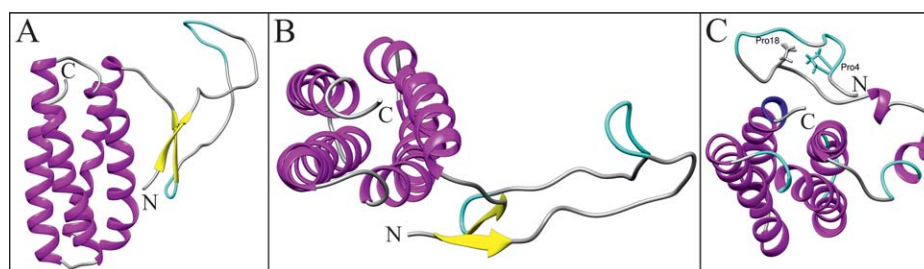
determined. The averaged  $T_1/T_2$  ratio over all measured residues was  $12.6 \pm 5.1$ , which is in the expected range for a near globular protein of this size. By contrast, the average  $T_1/T_2$  ratio for the residues 2–35 was found to be  $6.92 \pm 3.6$ . The  $T_1/T_2$  ratios rapidly increase in the well-defined C-terminal helical bundle region including the residues of the short  $\alpha$ -helix 0 (R37-Q41). The overall rotational correlation time  $\tau_c$  for the four-helix bun-

dle (residues 46–149) of PsbQ was calculated<sup>49</sup> as  $9.91 \pm 1.14$  ns (298 K). Complementary  $\{^1\text{H}\}^{15}\text{N}$  NOE data also prove the high mobility of the first 35 residues. By comparison to the average of the first 45 residues, the numbers of NOE cross-peaks found for all helical parts are significantly higher (Fig. 4).

The heteronuclear NOEs and relaxation rates corroborate the first 35 residues of the N-terminal region and the residues between helix 0 and helix 1 to be highly flexible and intrinsically disordered. The presence of a polyproline II helix (residues 9–12) suggested in the crystallographic structure<sup>14</sup> could not be confirmed by NMR in solution. The heteronuclear  $\{^1\text{H}\}^{15}\text{N}$  NOEs,  $^{15}\text{N}$   $T_1$  and  $^{15}\text{N}$   $T_2$  relaxation rates of the immediately adjacent residues (V6-G8, L13-G15) contradict the notion of any ordered conformation in this region.

### Molecular dynamics simulations

Earlier work based on the X-ray crystal structure<sup>38</sup> suggested that the two-stranded  $\beta$ -sheet would anchor the large N-terminal loop between Leu13 and Thr34 and weak interactions with the rest of the protein would give this loop a random coil structure with a distinct stable fold. The herein reported NMR-results are directly contradicting this hypothesis and therefore new molecular dynamics simulations were set up for both structures, the X-ray-based model structure used in the earlier study, and the newly reported NMR structure. As the above mentioned conclusions were drawn from 20 ns long simulations, that were state-of-the-art then, longer simulations, which are possible nowadays, might show that the conformation the loop adapted could have been only a local minimum and on longer time scales might sample a much larger conformational space. For the X-ray-based model a slightly elevated temperature of 310 K was used additionally to allow for possible jumping over an energy barrier, if the protein would be trapped in a local minimum. Equilibration of the simulations was estimated by root mean square deviation (not shown) and sufficiently long trajectories were calculated to have at least the last 30 ns in equilibrium. Thus, simulations were performed for 225 ns or 300 ns, respectively. Figure 5 shows the X-ray-based crystal structure after 225 ns and the NMR structure after 300 ns of molecular dynamics simulation at 300K in water, and surprisingly the two stranded  $\beta$ -sheet in the X-ray-based model kept its secondary structure. The two hydrogen bonds between Phe38-H Arg3-O and Phe38-O Ile5-H that keep the two  $\beta$ -strands together are persistent throughout the simulations and show only slightly larger fluctuations at elevated temperature of 310K (Fig. 6, left and middle panels). However, Figure 6, left panel, also demonstrates that the loop structure has significantly changed from the initial structure and that there are no weak contacts between the loop

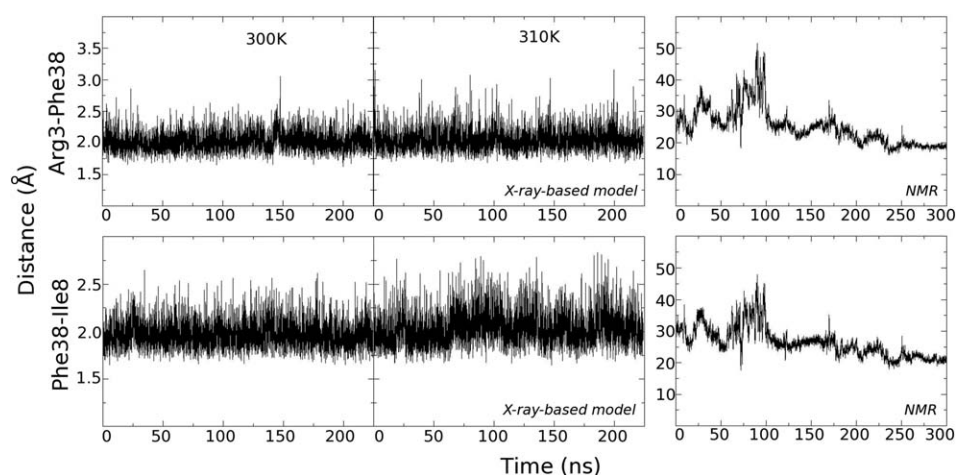


**Figure 5**

Simulated PsbQ structures after unrestrained molecular dynamics in solvent. (A) Side view of the X-ray-based model after 225 ns at 300 K. Helices are coloured in magenta,  $\beta$ -strands in yellow and  $\beta$ -turn structures in cyan.  $3_{10}$  helices are drawn in blue. N and C terminals are labelled accordingly. (B) Top view of the structure in panel A. (C) Top view of the NMR structure after 300 ns at 300 K.

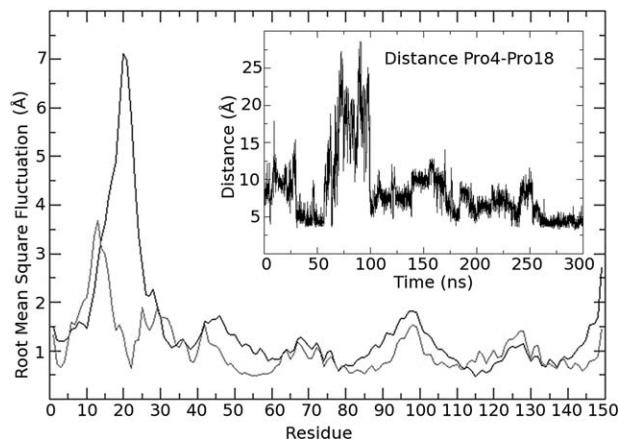
and the helix-bundle of the protein left over, but the loop is rather oriented toward the solvent. In the simulation of the NMR structure, the respective residues get never closer than 20 Å and thus there is no possibility to form hydrogen bonds or the  $\beta$ -sheet (Fig. 6, right panels). Contrary to this, the NMR solution structure shows weak interactions between Pro4 and Pro18, manifest by the distance between the gamma carbons of both proline residues in the insert of Figure 7. The RMSF during the last 20 ns of the simulations, in which both simulations are well equilibrated, show extremely high values of up to 7 Å for the N-terminal loop (Fig. 7, black line) in the X-ray-based model, while the NMR structure shows fluctuations only slightly higher than for the loops connecting the helices in the bundle (Fig. 7, gray line). Although the  $\beta$ -sheet is still present in the X-ray-based model after the simulation, this clearly contradicts the original hypothesis that the  $\beta$ -sheet would anchor the loop in a specific conformation. The

loop in the simulations does not keep any weak interactions with the rest of the protein and does not show a distinct stable fold, but is extremely flexible, giving the impression of unsuccessfully sampling the conformational space for a local minimum. In contrast to that, the loop in the simulations with the NMR structure rearranges in the course of the simulation and finally keeps a metastable distinct fold with comparably low fluctuation. There is no indication in these simulations that the N-terminal would preferentially form a  $\beta$ -strand. On the contrary, the helical structure has been shown to be stable throughout the simulation. This would be consistent with an understanding that the N-terminal, if not complexed by PSII, is rather unstructured and behaves as a tail, sampling a large conformational space with various local minima. This is also supported by the fact that in the recent crystal structure of CyanoQ<sup>50</sup> the first 34 amino acids are not resolved as well, in spite of a generally high resolution of 1.6 Å.



**Figure 6**

Dynamics of hydrogen-bonding between the two N-terminal  $\beta$ -strands. Left and middle panels: Distances between Phe38-H Arg3-O (upper panels) and Phe38-O Ile5-H (lower panels) in the X-ray based model structure in the course of the MD simulations at 300 K and 310 K. Right panels: Distances between Phe38-H Arg3-O (upper panel) and Phe38-O Ile5-H (lower panel) in the MD simulation of the NMR structure.



**Figure 7**

RMSF of PsbQ during the last 20 ns of MD simulation at 300 K. Black curve: Starting from the X-ray-based model structure, grey curve: Starting from the NMR structure. The insert shows the distance between the gamma carbons of Pro4 and Pro18 in the course of the MD simulation of the NMR structure.

## CONCLUSION

In this article we presented the solution structure of the accessory photosynthetic protein PsbQ from *Spinacia oleracea* based on the near complete resonance assignment. Using 3D NMR experiments at 700 MHz on recombinant, uniformly  $^{13}\text{C}$ ,  $^{15}\text{N}$  labeled PsbQ, the resonance assignment levels of 97.2% of the backbone and 71.5% of the side chain residues were achieved. Nuclear Overhauser effect based distance constraints and chemical shift derived torsion angle constraints led to a structure ensemble with backbone ( $\text{C}^\alpha$ ,  $\text{C}'$ ,  $\text{N}$ ) RMSD of  $0.81 \pm 0.15$  Å for the well-defined four-helix-bundle region. A highly flexible N-terminal region is attached to this core. The bundle of four up-down-up-down  $\alpha$ -helices corresponds closely to the previous crystallographic structure<sup>14</sup> with a backbone ( $\text{C}^\alpha$ ,  $\text{C}'$ ,  $\text{N}$ ) RMSD of 1.27 Å with respect to the mean NMR structure. However, a short  $\alpha$ -helix within residues R37-L40 is found in a region, where one half of the two-stranded  $\beta$ -sheet had been derived previously in the crystal structure.<sup>14</sup> It is noteworthy that D28, which has been found to interact with PsbP in a cross-linking study,<sup>18</sup> is part of the flexible tail preceding the short helix. It is thus able to access the large range of orientations, which may be a requirement for the repair functions of PsbQ for the OEC (for example, under redox stress conditions) suggested in a previous study.<sup>10</sup> The unrestrained molecular dynamics simulations for up to 300 ns in water support the understanding that in solution and in absence of other interaction partners the N-terminal region behaves as a flexible tail with an embedded conserved short  $\alpha$ -helix rather than being anchored by a two strand  $\beta$ -sheet. This short helix might play a role in the assembly process of the

OEC luminal protein complex. In the fully assembled PSII complex, the N-terminal part could not only form either of the two proposed structures but assume a different conformation induced by intermolecular interactions. It is noteworthy that in the recently determined crystal structure of the Psb31 protein from *Chaetoceros gracilis*, which shows very low sequence homology to PsbQ but seem to be a functional equivalent, a similar global structure, a four-helix-bundle with a flexible tail with an embedded short  $\alpha$ -helix which, by contrast, is attached at the C-terminal side has been found.<sup>17</sup> This may be an indication of both intrinsically disordered regions performing similar maintenance functions.

While flexible terminal regions are quite common in protein structures it appears from biological evidence that they have very determined roles within the Psb proteins.<sup>10</sup> The experimental NMR and computational MD results reported thus corroborate that the flexible, intrinsically disordered, regions of the Psb proteins of higher plant PSII and in particular of PsbQ play a key role in protein-protein interactions. Their solution structures and dynamics appear to be essential for understanding the assembly of the OEC and the functions of the extrinsic PSII proteins. The roles of the unstructured parts including the long flexible tails in the Psb protein interactions may be multiple, as is typical for intrinsically disordered regions. They may very likely adopt different conformations and exhibit different dynamics during the assembly process of the OEC, regulation of ion concentration while anchored to the thylakoid, and when performing the aforementioned repair function,<sup>10</sup> respectively. Based on the current results and the recent NMR assignment of PsbP,<sup>53</sup> which both are probably most relevant for the assembly stage, our future investigations will include all extrinsic Psb proteins of the OEC found in higher plants. Exploring their mutual interactions as well as the influence of the ionic and membrane environment on their structural and dynamic features will reveal their functions in greater detail.

## ACKNOWLEDGMENTS

The authors are indebted to Peter Güntert (Frankfurt) for a pre-release version of CYANA 3.0. All NMR experiments were performed at the NMR laboratory of the Upper Austrian—South Bohemian Research Infrastructure Center in Linz, co-financed by the European Union in the context of the project “RERI-uasb”, EFRE RU2-EU-124/100-2010 (ETC Austria-Czech Republic 2007-2013, project M00146).

## REFERENCES

1. Blankenship R. Origin and early evolution of photosynthesis. *Photosynth Res* 1992;33:91–111.



2. Bricker TM, Roose JL, Fagerlund RD, Frankel LK, Eaton-Rye JJ. The extrinsic proteins of Photosystem II. *Biochim Biophys Acta* 2012; 1817:121–142.
3. Roose JL, Wegener KM, Pakrasi HB. The extrinsic proteins of Photosystem II. *Photosynth Res* 2007;92:369–387.
4. Thornton LE, Ohkawa H, Roose JL, Kashino Y, Keren N, Pakrasi HB. Homologs of plant PsbP and PsbQ proteins are necessary for regulation of photosystem II activity in the cyanobacterium *Synechocystis* 6803. *Plant Cell* 2004;16:2164–2175.
5. Kashino Y, Lauber WM, Carroll JA, Wang Q, Whitmarsh J, Satoh K, Pakrasi HB. Proteomic analysis of a highly active photosystem II preparation from the cyanobacterium *Synechocystis* sp. PCC 6803 reveals the presence of novel polypeptides. *Biochemistry* 2002;41: 8004–8012.
6. Ghanotakis DF, Topper JN, Youcum CF. Structural organization of the oxidizing side of photosystem II. Exogenous reductants reduce and destroy the Mn-complex in photosystems II membranes depleted of the 17 and 23 kDa polypeptides. *Biochim Biophys Acta* 1984;767:524–531.
7. Miyao M, Murata N. Role of the 33-kDa polypeptide in preserving Mn in the photosynthetic oxygen-evolution system and its replacement by chloride ions. *FEBS Lett* 1984;170:350–354.
8. Nowaczyk M, Berghaus C, Stoll R, Rogner M. Preliminary structural characterisation of the 33 kDa protein (PsbO) in solution studied by site-directed mutagenesis and NMR spectroscopy. *Phys Chem Chem Phys* 2004;6:4878–4881.
9. Bricker TM. Oxygen evolution in the absence of the 33-kilodalton manganese-stabilizing protein. *Biochemistry* 1992;31:4623–4628.
10. Kakiuchi S, Uno C, Ido K, Nishimura T, Noguchi T, Ifuku K, Sato F. The PsbQ protein stabilizes the functional binding of the PsbP protein to photosystem II in higher plants. *Biochim Biophys Acta* 2012;8:25.
11. Ido K, Nield J, Fukao Y, Nishimura T, Sato F, Ifuku K. Cross-linking evidence for multiple interactions of the PsbP and PsbQ proteins in a higher plant photosystem II supercomplex. *J Biol Chem* 2014;289: 20150–20157.
12. Kopecky V, Jr., Kohoutova J, Lapkouski M, Hofbauerova K, Sovova Z, Ettrichova O, Gonzalez-Perez S, Dulebo A, Kaftan D, Smatanova IK, Revuelta JL, Arellano JB, Carey J, Ettrich R. Raman spectroscopy adds complementary detail to the high-resolution X-ray crystal structure of photosynthetic PsbP from *Spinacia oleracea*. *PLoS One* 2012;7:e46694.
13. Ifuku K, Nakatsu T, Kato H, Sato F. Crystal structure of the PsbP protein of photosystem II from *Nicotiana tabacum*. *EMBO Rep* 2004;5:362–367.
14. Balsera M, Arellano JB, Revuelta JL, de las Rivas J, Hermoso JA. The 1.49 Å resolution crystal structure of PsbQ from photosystem II of *Spinacia oleracea* reveals a PPII structure in the N-terminal region. *J Mol Biol* 2005;350:1051–1060.
15. Cao P, Xie Y, Li M, Pan X, Zhang H, Zhao X, Su X, Cheng T, Chang W. Crystal structure analysis of extrinsic PsbP protein of photosystem II reveals a manganese-induced conformational change. *Mol Plant* 2015;8:664–666.
16. Jackson SA, Fagerlund RD, Wilbanks SM, Eaton-Rye JJ. Crystal structure of PsbQ from *synechocystis* sp. PCC 6803 at 1.8 Å: implications for binding and function in cyanobacterial photosystem II. *Biochemistry* 2010;49:2765–2767.
17. Nagao R, Suga M, Niikura A, Okumura A, Koua FHM, Suzuki T, Tomo T, Enami I, Shen J-R. Crystal structure of Psb31, a novel extrinsic protein of photosystem II from a marine centric diatom and implications for its binding and function. *Biochemistry* 2013; 52:6646–6652.
18. Nagao R, Suzuki T, Okumura A, Niikura A, Iwai M, Dohmae N, Tomo T, Shen J-R, Ikeuchi M, Enami I. Topological analysis of the extrinsic PsbO, PsbP and PsbQ proteins in a green algal PSII complex by cross-linking with a water-soluble carbodiimide. *Plant Cell Physiol* 2010;51:718–727.
19. Hornicakova M, Kohoutova J, Schlagnitweit J, Wohlschlager C, Ettrich R, Fiala R, Schoeberger W, Müller N. Backbone assignment and secondary structure of the PsbQ protein from photosystem II. *Biomol NMR Assign* 2011;5:169–175.
20. Bax A, Clore GM, Gronenborn AM.  $^1\text{H}$ - $^1\text{H}$  correlation via isotropic mixing of  $^{13}\text{C}$  magnetization, a new three-dimensional approach for assigning  $^1\text{H}$  and  $^{13}\text{C}$  spectra of  $^{13}\text{C}$ -enriched proteins. *J Magn Reson* (1969) 1990;88:425–431.
21. Grzesiek S, Anglister J, Bax A. Correlation of backbone amide and aliphatic side-chain resonances in  $^{13}\text{C}/^{15}\text{N}$ -enriched proteins by isotropic mixing of  $^{13}\text{C}$  magnetization. *J Magn Reson Series B* 1993; 101:114–119.
22. Grzesiek S, Bax A. Amino acid type determination in the sequential assignment procedure of uniformly  $^{13}\text{C}/^{15}\text{N}$ -enriched proteins. *J Biomol NMR* 1993;3:185–204.
23. Bax A, Clore GM, Driscoll PC, Gronenborn AM, Ikura M, Kay LE. Practical aspects of proton-carbon-carbon-proton three-dimensional correlation spectroscopy of  $^{13}\text{C}$ -labeled proteins. *J Magn Reson* (1969) 1990;87:620–627.
24. Kay LE, Ikura M, Bax A. Proton-proton correlation via carbon-carbon couplings: a three-dimensional NMR approach for the assignment of aliphatic resonances in proteins labeled with carbon-13. *J Am Chem Soc* 1990;112:888–889.
25. Ikura M, Kay L, Bax A. Improved three-dimensional  $^1\text{H}$ - $^{13}\text{C}$ - $^1\text{H}$  correlation spectroscopy of a  $^{13}\text{C}$ -labeled protein using constant-time evolution. *J Biomol NMR* 1991;1:299–304.
26. Marion D, Driscoll PC, Kay LE, Wingfield PT, Bax A, Gronenborn AM, Clore GM. Overcoming the overlap problem in the assignment of proton NMR spectra of larger proteins by use of three-dimensional heteronuclear proton-nitrogen-15 Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: application to interleukin 1 $\beta$ . *Biochemistry* 1989;28:6150–6156.
27. Yamazaki T, Forman-Kay JD, Kay LE. Two-dimensional NMR experiments for correlating carbon-13  $\beta$  and proton  $\delta/\epsilon$  chemical shifts of aromatic residues in  $^{13}\text{C}$ -labeled proteins via scalar couplings. *J Am Chem Soc* 1993;115:11054–11055.
28. Palmer AG. 3rd. NMR probes of molecular dynamics: overview and comparison with other techniques. *Annu Rev Biophys Biomol Struct* 2001;30:129–155.
29. Marion D, Kay LE, Sparks SW, Torchia DA, Bax A. Three-dimensional heteronuclear NMR of nitrogen-15 labeled proteins. *J Am Chem Soc* 1989;111:1515–1517.
30. Zuiderweg ERP, Fesik SW. Heteronuclear three-dimensional NMR spectroscopy of the inflammatory protein C5a. *Biochemistry* 1989; 28:2387–2391.
31. Keller RJ. Optimizing the process of nuclear magnetic resonance spectrum analysis and computer aided resonance assignment. Diss. ETH No. 15947, 2004.
32. Bartels C, Xia T, Billeter M, Güntert P, Wüthrich K. The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. *J Biomol NMR* 1995;5:1–10.
33. Güntert P, Braun W, Wüthrich K. Efficient computation of three-dimensional protein structures in solution from nuclear magnetic resonance data using the program DIANA and the supporting programs CALIBA, HABAS and GLOMSA. *J Mol Biol* 1991;217:517–530.
34. Shen Y, Bax A. Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J Biomol NMR* 2013; 56:227–241.
35. Güntert P, Mumenthaler C, Wüthrich K. Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J Mol Biol* 1997;273:283–298.
36. Krieger E, Joo K, Lee J, Raman S, Thompson J, Tyka M, Baker D, Karplus K. Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: four approaches that performed well in CASP8. *Proteins* 2009;114–122.

37. Krieger E, Koraimann G, Vriend G. Increasing the precision of comparative models with YASARA NOVA—a self-parameterizing force field. *Proteins* 2002;47:393–402.
38. Ristvejova J, Kopecky V, Jr, Sovova Z, Balsera M, Arellano JB, Green M, Ettrich R. Structure and dynamics of the N-terminal loop of PsbQ from photosystem II of *Spinacia oleracea*. *Biochem Biophys Res Commun* 2006;345:287–291.
39. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ. GROMACS: fast, flexible, and free. *J Comput Chem* 2005;26:1701–1718.
40. Berendsen HJC, van der Spoel D, van Drunen R. GROMACS: a message-passing parallel molecular dynamics implementation. *Comp Phys Commun* 1995;91:43–56.
41. Pronk S, Pall S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, Shirts MR, Smith JC, Kasson PM, van der Spoel D, Hess B, Lindahl E. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* 2013;29:845–854.
42. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* 2006;65:712–725.
43. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 1983;79:926–935.
44. Darden T, York D, Pedersen L. Particle mesh Ewald: an  $N \log(N)$  method for Ewald sums in large systems. *J Chem Phys* 1993;98:10089–10092.
45. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. LINCS: a linear constraint solver for molecular simulations. *J Comput Chem* 1997;18:1463–1472.
46. Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. *J Chem Phys* 2007;126:014101
47. Parrinello M, Rahman A. Polymorphic transitions in single crystals: a new molecular dynamics method. *J Appl Phys* 1981;52:7182–7190.
48. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Mol Graph* 1996;14:33–38.
49. Kay LE, Torchia DA, Bax A. Backbone dynamics of proteins as studied by  $^{15}\text{N}$  inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry* 1989;28:8972–8979.
50. Michoux F, Boehm M, Bialek W, Takasaka K, Maghlaoui K, Barber J, Murray JW, Nixon PJ. Crystal structure of CyanoQ from the thermophilic cyanobacterium *Thermosynechococcus elongatus* and detection in isolated photosystem II complexes. *Photosynth Res* 2014;122:57–67.
51. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* 2004;25:1605–1612.
52. Laskowski RA, Macarthur MW, Moss DS, Thornton JM. {PROCHECK}: a program to check the stereochemical quality of protein structures. *J Appl Cryst* 1993;26:283–291.
53. Rathner A, Chandra K, Rathner P, Hornicakova M, Schlagnitweit J, Kohoutova J, Ettrich R, Müller N. Resonance assignment of PsbP; an extrinsic protein from photosystem II of *Spinacia oleracea*. *Biomol NMR Assign*, doi:10.1007/s12104-015-9606-2.