^{1}H, ^{15}N and ^{13}C assignments and secondary structure of the EGF-like module pair 3–4 from vitamin K-dependent protein S

Andreas Muranyi[a,*], Johan Evenäs[a,1], Yvonne Stenberg[b,2], Johan Stenflo[b], Torbjörn Drakenberg[a]

[a] *Physical Chemistry 2, Lund University, P.O. Box 124, S-221 00 Lund, Sweden*
[b] *Clinical Chemistry, Lund University, University Hospital Malmö, S-205 02 Malmö, Sweden*

Received 5 May 2000

Edited by Thomas L. James

Abstract Vitamin K-dependent protein S, which is a cofactor for activated protein C and thus important for down-regulation of the coagulation cascade, contains several Ca^{2+}-binding sites with unusually high affinity. The 89 amino acid fragment constituting the third and fourth epidermal growth factor-like (EGF) modules of protein S is the smallest fragment that retains high-affinity Ca^{2+} binding and is therefore useful for investigating the structural basis of this property. Heteronuclear multidimensional nuclear magnetic resonance experiments were used to obtain extensive assignments of the ^{1}H, ^{15}N and ^{13}C resonances of the module pair with one Ca^{2+} bound in EGF 4. In addition, nearly complete assignments of the ^{1}H resonances of the isolated Ca^{2+}-free EGF 3 module were obtained. The assignment process was complicated by broadening of several resonances, spectral heterogeneity caused by *cis-trans* isomerisation of the peptide bond preceding Pro-168, and dimerisation. Analysis of weighted average secondary chemical shifts, $^{3}J_{HNH\alpha}$ coupling constants, and NOE connectivities suggest that both EGF modules in this fragment adhere to the classical secondary structure of EGF modules, consisting of one major and one minor anti-parallel β-sheet.

Key words: Assignment; Blood coagulation; Calcium-binding epidermal growth factor-like module; Protein S

1. Introduction

Blood coagulation is activated as a response to vascular injury and leads to the activation of platelets and coagulation proteins which ultimately lead to the formation of a fibrin plug [1]. This process is regulated by an anticoagulant counterpart, the so-called protein C anticoagulant system [2]. Activated protein C (APC) catalyses the proteolytic degradation of activated factors V and VIII, which have a crucial role in the maintenance of the coagulation cascade. Vitamin K-dependent protein S acts as a cofactor for APC. Protein S is a mosaic protein composed of an N-terminal Gla module, a region sensitive to cleavage by thrombin and factor Xa, four epidermal growth factor-like (EGF) modules, and a module homologous with a steroid hormone-binding protein in plasma [3–5]. The three C-terminal EGF modules have a consensus sequence for Ca^{2+} binding [6–9]. It has also been shown that they bind Ca^{2+} with relatively high affinity and that the affinity is dependent on neighbouring modules. The strongest binding site in recombinant EGF 1–4 has a dissociation constant $K_d = 20 \times 10^{-9}$ M, and this site is probably located in EGF 4 [10]. However, the isolated EGF 4 module has only a K_d of 8.6×10^{-3} M, i.e. a 10^6 times weaker binding, indicating that neighbouring modules are necessary for high-affinity Ca^{2+} binding [11]. The smallest fragment that retains relatively high Ca^{2+} affinity is the EGF 3–4 module pair (pS EGF 3–4), which has a K_d for Ca^{2+} binding to the site in EGF 4 of 1×10^{-6} M [12]. By studying the effects of Ca^{2+} binding and determining of the structure and the dynamics of this module pair we hope to be able to explain the structural basis of the unusually high Ca^{2+} affinity of certain EGF modules in protein S and the role of metal binding and inter-module interactions in this context. In this paper we report the assignments of ^{1}H, ^{15}N, and ^{13}C resonances of recombinant pS EGF 3–4 with one bound Ca^{2+}, as well as the ^{1}H resonances of the Ca^{2+}-free, isolated EGF 3 module (pS EGF 3). Both proteins were folded to a native conformation.

2. Materials and methods

The production of pS EGF 3–4 and pS EGF 3 has been described previously [11,12]. In brief, pS EGF 3–4 was expressed in *Escherichia coli* BL21(DE3) pLysS cells as a fusion protein with a His tag. The recombinant protein was reduced and refolded in vitro, the correctly folded material was purified and the His tag removed. Ca^{2+}-free pS EGF 3–4 was lyophilised and characterised by SDS–PAGE, reverse-phase high-performance liquid chromatography (HPLC), amino acid composition, and immunoblotting using a monoclonal antibody that recognises a conformation-dependent epitope. Agarose gel electrophoresis in the presence and absence of Ca^{2+} was used to detect correctly folded Ca^{2+}-binding pS EGF 3–4. N-terminal sequencing and mass spectroscopy indicated a homogeneous protein ($M_W = 9843$ Da). The sequence of the final product comprised residues 159–245 of human protein S with the addition of an N-terminal dipeptide (His–Met) derived from the His tag.

The EGF 3 module, residues Lys-159–Ile-203, was synthesised using Fmoc chemistry on a Milligen/Biosearch 9050 peptide synthesiser. The peptide was deprotected, cleaved from the resin, extracted and dried. Thereafter the peptide was purified by HPLC performed under reducing conditions. Fractions with the correct sequence and amino acid composition were refolded as previously described [11]. Finally, the refolded peptide was purified by reversed-phase HPLC. Amino acid analysis, N-terminal sequencing and electrospray mass spectrometric analysis gave the expected results.

*Corresponding author. Fax: (46)-46-222 45 43.
E-mail: andreas.muranyi@fkem2.lth.se

[1] Present address: Departments of Medical Genetics and Biochemistry, Medical Sciences Building, University of Toronto, Toronto, Ont., Canada M5S 1A8.
[2] Present address: BioInvent Therapeutic AB, Sölvegatan 41, S-223 70 Lund, Sweden.

0014-5793/00/$20.00
PII: S0014-5793(00)01636-7

Samples for nuclear magnetic resonance (NMR) spectroscopy were prepared by dissolving unlabelled, ^{15}N-labelled, and ^{15}N,^{13}C-labelled pS EGF 3–4 and unlabelled pS EGF 3 in 90%/10% H_2O/D_2O (or 100% D_2O), 0.1 mM NaN_3, and 1 mM DSS. Protein concentrations were in the range 0.5–2 mM. One equivalent $CaCl_2$ was added to the pS EGF 3–4 samples, leading to saturation of the high-affinity Ca^{2+}-binding site in EGF 4, while pS EGF 3 samples were Ca^{2+}-free [13].

All NMR experiments were performed at 36°C on a Varian Unity Inova with a ^1H frequency of 599.89 MHz using inverse or triple-resonance probes equipped with pulsed field gradient coils in the *z*-direction.

Using the unlabelled samples of pS EGF 3–4 and pS EGF 3 we recorded NOESY spectra [14], with mixing times of 80 and 150 ms, and TOCSY spectra [15], with 40 and 110 ms DIPSI isotropic mixing times. On the ^{15}N-labelled pS EGF 3–4 we acquired the following spectra for assignment purposes: ^1H,^{15}N HSQC, ^1H,^{15}N TOCSY-HSQC, with a DIPSI mixing sequence (40–110 ms), and ^1H,^{15}N NOESY-HSQC, with mixing times of 80 and 150 ms. Several ^{15}N spin relaxation experiments were also run, for which the details have been reported previously [13]. To determine $^3J_{HNH\alpha}$ coupling constants HMQC-*J* experiments were run and the peak heights were fitted as a function of dephasing time to equation 4 of Kuboniwa et al. [16]. HNCA [17], HN(CO)CA [17,18], HNCACB [19–21], and HNCO [19,20,22,23] experiments were carried out on the ^{15}N,^{13}C-labelled pS EGF 3–4 to obtain an unambiguous assignment of backbone and C^β resonances. All double- and triple-resonance experiments, except the HMQC-*J*, were performed with gradient selection of coherences and sensitivity enhancement [24]. Quadrature detection in F1 was achieved by the States-TPPI method [25].

Data were processed using FELIX (Molecular Simulations Inc., San Diego, CA, USA). ^1H chemical shifts were referenced to DSS at 0.00 ppm, and ^{15}N and ^{13}C shifts were indirectly referenced using the frequency ratio $\omega_N/\omega_H = 0.101329118$ and $\omega_C/\omega_H = 0.251449530$, respectively [26]. Spectral display and peak picking were performed in FELIX or XEASY [27]. Sequential assignments were made in the 3D spectra using functionalities in XEASY and with the aid of a program for amino acid identification based on C^α and C^β chemical shifts [28]. Peaks in the 3D ^1H,^{15}N NOESY-HSQC spectra were integrated with SPSCAN [29] and calibrated using DYANA [30]. Weighted average secondary chemical shifts (WASS) were calculated using the program GENXPK [31], employing the following weighting factors: H^α −0.1, C^α +0.7, C^β -0.7, C' +0.5, Pro C^α +4.0, Pro C^β −4.0, and Pro C' +4.0.

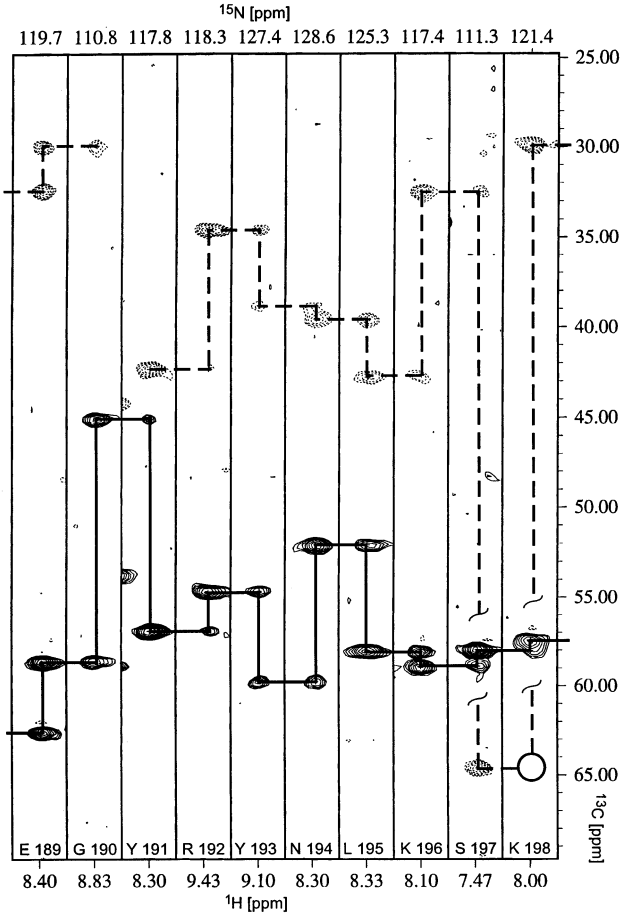

Fig. 1. Strips from the HNCACB experiments showing C^α (solid line) and C^β (broken line) connectivities from Pro-188–Lys-198. In this section the only missing connectivity is Ser-197 C^β–Lys-198 C^β (indicated by a circle).

3. Results and discussion

The assignment process was more complicated than anticipated. Extensive line broadening due to chemical exchange was observed for several resonances, particularly in the N-terminal part of EGF 3. This rendered signals very weak for a number of spin systems, most notably Phe-183–Glu-186 in the second strand of the first anti-parallel β-sheet. Very slow exchange caused by *cis-trans* isomerisation of the Lys-167–Pro-168 peptide bond gave rise to extensive heterogeneity in all spectra. In a ^1H,^{15}N HSQC spectrum of pS EGF 3–4 with one equivalent of Ca^{2+}, 124 peaks instead of the expected 99 were observed. For 23 residues in EGF 3 two separate sets of spin systems, originating from the *cis* and *trans* forms, respectively, were assigned. Finally, an estimate of the correlation time of rotational diffusion (τ_c) suggested that pS EGF 3–4 probably was present as a dimer with $M_W \sim 20$ kDa at 0.5 mM protein concentration. These findings have been described in detail elsewhere [13].

The triple-resonance experiments were used to sequentially assign EGF 4 and most parts of EGF 3. An example of strips from the HNCACB spectrum is shown in Fig. 1. The assignment was confirmed by sequential NOEs in the ^1H,^{15}N NOESY-HSQC spectrum. However, in EGF 3 several signals were missing in the triple-resonance spectra. The lack of observable signals in the triple-resonance spectra coincides with a region of the protein exhibiting extensive ^{15}N line broadening (R_{ex}) (Fig. 2) [13]. A region (residues 228–235) in the C-terminal part of the protein also exhibits extensive ^{15}N line broadening but this does not have the same deleterious effect on the triple-resonance spectra. However, in the N-terminal part the ^{15}N R_{ex} terms are accompanied by ^1H line broadening to some extent. Thus it is likely that the lack of signals from residues in the N-terminal part is the result mainly of the combination of ^1H and ^{15}N line broadening. Additional loss of signal intensity is caused by the fact that many spin systems in EGF 3 (residues 159–200) are doubled due to the *cis-trans* isomerisation of the Lys-167–Pro-168 peptide bond. Parts of EGF 3 was therefore assigned relying only on the ^1H,^{15}N TOCSY-HSQC and NOESY-HSQC spectra.

The secondary structure was inferred from a combination of WASSs [32], $^3J_{HNH\alpha}$ scalar coupling constants [33], and characteristic NOEs [34] (Fig. 3). The data for the two EGF modules are similar and consistent with the classical fold of EGF modules [35]. It consists of a centrally located major (residues 173–179 and 182–188 in EGF 3, and 213–218 and 222–227 in EGF 4) and a C-terminal minor (residues 190–192 and 199–201 in EGF 3, and 232–235 and 241–244 in EGF 4) anti-parallel β-sheet. In addition, there are short, helical turns centred around the first cysteine residue in each module (res-

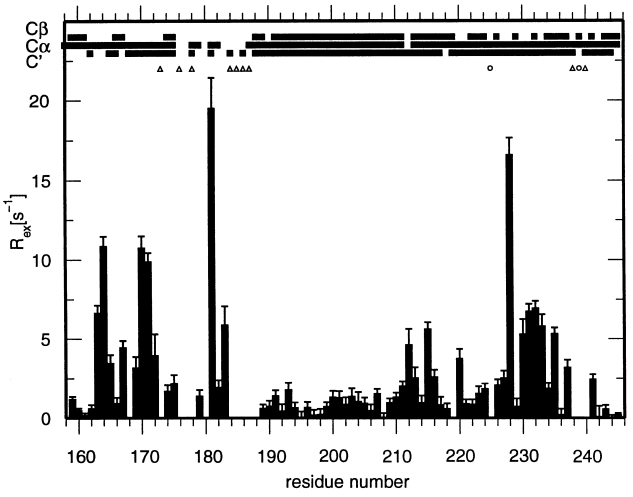

Fig. 2. The residue-specific chemical exchange contributions calculated as $R_{ex} = R_2 - [\eta_{xy}/\eta_z(R_1 - 1.249\sigma_{NH}) + 1.079\sigma_{NH}]$, where R_1 and R_2 are the longitudinal and transverse ^{15}N relaxation rate constants, respectively, η_z and η_{xy} are the longitudinal and transverse cross-correlation relaxation rate constants resulting from ^1H-^{15}N dipolar and ^{15}N chemical shift anisotropy interference, and σ_{NH} is the heteronuclear cross-relaxation rate [13,37]. Residues for which no reliable η_z or η_{xy} could be measured due to ^1H line broadening or spectral overlap are indicated by a triangle or a circle, respectively. At the top of the figure assignments of the chemical shifts of C^β, C^α, and C′ in triple-resonance experiments are indicated by a square.

idues 163–165 and 205–207 in EGF 3 and EGF 4, respectively).

In the paper describing the structure of the module pair of Ca^{2+}-binding EGF modules 32–33 from fibrillin-1, a hydrophobic cluster linking the two EGF modules was described [36]. In pS EGF 3–4 Tyr-191, Pro-219 and Gly-220 correspond to the residues involved in this hydrophobic module–module contact in fibrillin-1. In order to probe whether a similar hydrophobic cluster was present in pS EGF 3–4, we compared the chemical shifts of the EGF 3 module in pS EGF 3–4 with those of pS EGF 3. Shift differences were indeed found in the expected region but they were relatively small. The two C-terminal residues of pS EGF 3 (Asp-202 and Ile-203) displayed shift differences of about 0.3 ppm, which is most likely an end effect. Thereafter the largest shift difference was observed for the Arg-192 H^ε (0.22 ppm). Shift differences between 0.10 and 0.16 ppm were seen for Gly-190, Tyr-191, Arg-192, Tyr-193, Asn-194, Lys-196, and Cys-200. Upon Ca^{2+} binding to the proximal, high-affinity site in EGF 4 we observed ^1H shift changes for the latter residues in the range 0.04–0.08 ppm, which, although small, were significant [13]. These relatively small shift differences can hardly be taken as proof that a hydrophobic cluster, similar to the one observed in the fibrillin EGF module pair, exists in pS EGF 3–4, but are certainly consistent with the presence of such a contact.

4. Data deposition

Chemical shifts of the ^1H resonances from pS EGF 3 and the ^1H, ^{15}N, $^{13}C^\alpha$, $^{13}C^\beta$, and $^{13}C'$ resonances of pS EGF 3–4 have been deposited in the BioMagResBank database under

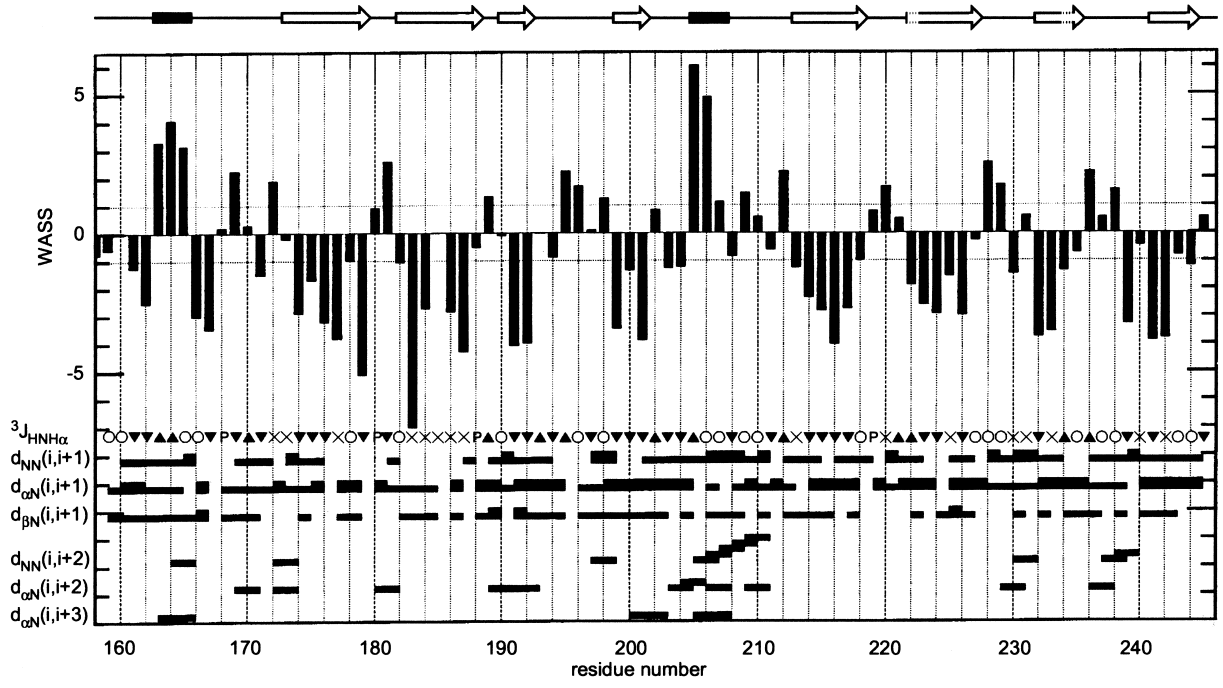

Fig. 3. The WASS of H^α, C^α, C^β and C′ nuclei, $^3J_{HNH\alpha}$ coupling constants ($^3J_{HNH\alpha} < 6$ Hz, ▲; $^3J_{HNH\alpha} > 8$ Hz, ▼; 6 Hz $< ^3J_{HNH\alpha} < 8$ Hz, ○; $^3J_{HNH\alpha}$ that could not be measured, ×; and proline residue, P) and characteristic NOEs of regular secondary structure (the width of the bar indicates the strength of the NOE). At the top the inferred secondary structure is indicated by filled rectangles (helices) and arrows (β-sheets). For residues Tyr-222 and Leu-234 the WASS indicates β-sheet, while $^3J_{HNH\alpha}$ is 5.1 Hz, which is not consistent with a β-sheet. To be able to pair the strands these residues were included in β-sheet structure, although the symbols at the top were drawn with a broken line to indicate this slight inconsistency.

the accession numbers 4728 and 4729, respectively. For both proteins the shifts of the *trans* as well as *cis* Lys-167–Pro-168 forms are deposited.

Acknowledgements: The pulse sequences for the triple-resonance experiments were made available to us by Professor Lewis Kay (University of Toronto), which we gratefully acknowledge. This work was supported by grants from the Swedish Foundation of Strategic Research, the Swedish Medical Research Council and the European Union Biotechnology Program (Contract BIO4-CT96-0662). Some of the NMR experiments were run at the Swedish NMR Centre.

References

[1] Davie, E.W., Fujikawa, K. and Kisiel, W. (1991) Biochemistry 30, 10363–10370.
[2] Dahlbäck, B. (1995) Thromb. Res. 77, 1–43.
[3] Lundwall, A., Dackowski, W., Cohen, E., Shaffer, M., Mahr, A., Dahlbäck, B., Stenflo, J. and Wydro, R. (1986) Proc. Natl. Acad. Sci. USA 83, 6716–6720.
[4] Dahlbäck, B., Lundwall, A. and Stenflo, J. (1986) Proc. Natl. Acad. Sci. USA 83, 4199–4203.
[5] Long, G.L., Lu, D., Xie, R.L. and Kalafatis, M. (1998) J. Biol. Chem. 273, 11521–11526.
[6] Rees, D.J., Jones, I.M., Handford, P.A., Walter, S.J., Esnouf, M.P., Smith, K.J. and Brownlee, G.G. (1988) EMBO J. 7, 2053–2061.
[7] Handford, P.A., Mayhew, M., Baron, M., Winship, P.R., Campbell, I.D. and Brownlee, G.G. (1991) Nature 351, 164–167.
[8] Selander-Sunnerhagen, M., Ullner, M., Persson, E., Teleman, O., Stenflo, J. and Drakenberg, T. (1992) J. Biol. Chem. 267, 19642–19649.
[9] Mayhew, M., Handford, P., Baron, M., Tse, A., Campbell, I. and Brownlee, G. (1992) Protein Eng. 5, 489–494.
[10] Stenberg, Y., Linse, S., Drakenberg, T. and Stenflo, J. (1997) J. Biol. Chem. 272, 23255–23260.
[11] Stenberg, Y., Julenius, K., Dahlqvist, I., Drakenberg, T. and Stenflo, J. (1997) Eur. J. Biochem. 248, 163–170.
[12] Stenberg, Y., Muranyi, A., Steen, C., Thulin, E., Drakenberg, T. and Stenflo, J. (1999) J. Mol. Biol. 293, 653–665.
[13] Muranyi, A., Evenäs, J., Stenberg, Y., Stenflo, J. and Drakenberg, T., submitted.
[14] Macura, S. and Ernst, R.R. (1980) Mol. Phys. 41, 95–117.
[15] Braunschweiler, L. and Ernst, R.R. (1983) J. Magn. Reson. 53, 521–528.
[16] Kuboniwa, H., Grzesiek, S., Delaglio, F. and Bax, A. (1994) J. Biomol. NMR 4, 871–878.
[17] Yamazaki, T., Lee, W., Revington, M., Mattiello, D.L., Dahlquist, F.W., Arrowsmith, C.H. and Kay, L.E. (1994) J. Am. Chem. Soc. 116, 6464–6465.
[18] Yamazaki, T., Lee, W., Arrowsmith, C.H., Muhandiram, D.R. and Kay, L.E. (1994) J. Am. Chem. Soc. 116, 11655–11666.
[19] Muhandiram, D.R. and Kay, L.E. (1994) J. Magn. Reson. B 103, 203–216.
[20] Kay, L.E., Keifer, P. and Saarinen, T. (1992) J. Am. Chem. Soc. 114, 10663–10665.
[21] Wittekind, M. and Mueller, L. (1993) J. Magn. Reson. B 101, 201.
[22] Grzesiek, S. and Bax, A. (1992) J. Magn. Reson. 96, 432.
[23] Kay, L.E., Xu, G.Y. and Yamazaki, T. (1994) J. Magn. Reson. A 109, 129–133.
[24] Zhang, O., Kay, L.E., Olivier, J.P. and Forman-Kay, J.D. (1994) J. Biomol. NMR 4, 845–858.
[25] States, D.J., Haberkorn, R.A. and Ruben, D.J. (1982) J. Magn. Reson. 48, 286–292.
[26] Wishart, D.S., Bigam, C.G., Yao, J., Abildgaard, F., Dyson, H.J., Oldfield, E., Markley, J.L. and Sykes, B.D. (1995) J. Biomol. NMR 6, 135–140.
[27] Bartels, C., Xia, T.-H., Billeter, M., Güntert, P. and Wüthrich, K. (1995) J. Biomol. NMR 5, 1–10.
[28] Grzesiek, S. (1993) J. Biomol. NMR 3, 185–204.
[29] Glaser, R.W. and Wüthrich, K. http://www.mol.biol.ethz.ch/wuthrich/software/spscan/.
[30] Güntert, P., Mumenthaler, C. and Wüthrich, K. (1997) J. Mol. Biol. 273, 283–298.
[31] Gippert, G.P. (1995) New Computational Methods for 3D NMR Data Analysis and Protein Structure Determination in High-Dimensional Internal Coordinate Space, Ph.D. Thesis, Scripps Research Institute, La Jolla, CA.
[32] Wishart, D.S. and Sykes, B.D. (1994) Methods Enzymol. 239, 363–392.
[33] Pardi, A. and Billeter, M. (1984) J. Mol. Biol. 180, 741–751.
[34] Wüthrich, K. (1986) NMR of Proteins and Nucleic Acids, John Wiley and Sons, New York.
[35] Campbell, I.D. and Bork, P. (1993) Curr. Opin. Struct. Biol. 3, 385–392.
[36] Downing, A.K., Knott, V., Werner, J.M., Cardy, C.M., Campbell, I.D. and Handford, P.A. (1996) Cell 85, 597–605.
[37] Kroenke, C.D., Loria, J.P., Lee, L.K., Rance, M. and Palmer, A.G. (1998) J. Am. Chem. Soc. 120, 7905–7915.