Journal of Biomolecular NMR **29:** 91–92, 2004.

Letter to the Editor: Sequence-specific backbone NMR assignments for the C-terminal globular domain of EMILIN-1

Giuliana Verdone[a], Simon A. Colebrooke[b], Jonathan Boyd[b], Paolo Viglino[a], Alessandra Corazza[a], Roberto Doliana[c], Gabriella Mungiguerra[c], Alfonso Colombatti[a,c,*], Gennaro Esposito[a,*] & Iain D. Campbell[b]

[a]*Dipartimento di Scienze e Tecnologie Biomediche- MATI Centro di Eccellenza, Università di Udine, P.le Kolbe 4, 33100 Udine, Italy;* [b]*Department of Biochemistry and Oxford Centre for Molecular Sciences, University of Oxford, South Parks Road, Oxford OX1 3QU, UK;* [c]*Divisione di Oncologia Sperimentale 2, Centro di Riferimento Oncologico di Aviano, 33081 Aviano, Italy*

Received 8 August 2003; Accepted 6 October 2003

Key words: C1q-like domain, elastic fibres, EMILIN, heteronuclear NMR, partially deuterated proteins, triple resonance experiments, TROSY

Biological context

The protein EMILIN-1 is a component of elastic fibres which constitute the extracellular matrix (ECM) of various organs, particularly abundant in blood vessels (Doliana et al., 1999). EMILIN-1 is commonly localized in regions of the ECM where Elastin and microfibrils are in close contact, hence the name Elastin Microfibril Interface Located proteIN (EMILIN). Amino acid sequence analysis reveals that the protein possesses a composite architecture formed by several distinct domains (Doliana et al., 1999).

Components of elastic fibres have been found to be defective in some heritable human syndromes, the most common being the Marfan syndrome. Since most of those syndromes for which a clear alteration of elastic fibres is observed arise from the consequences of structural and functional changes in the ECM proteins, an investigation of the EMILIN-1 structure is relevant in the context of the study of elastic fibres' diseases.

In this study we present the backbone chemical shift assignment of the C-terminal globular domain of the human (h-)EMILIN-1, a domain which presents sequence homology with proteins of the C1q family, named after the complement C1q protein. Single crystal diffraction data are currently available for two C1q-like domains (Shapiro and Scherer, 1998; Bogin et al., 2002). However, no solution structures or chemical shift assignments have been published so far for any of the domains of the EMILIN family members. In the experimental conditions of the present analysis, the C1q-like domain of h-EMILIN-1 is a homotrimer (MW∼51 kDa) composed of three chains each having MW∼17 kDa. It has been recently proposed that this domain has a fundamental role in promoting the trimerization of h-EMILIN-1 which occurs from the C-terminus of the protein, together with a specific cell-adhesion promoting function (Doliana et al., 1999; Mongiat et al., 2000; Spessotto et al., 2003).

*To whom correspondence should be addressed. E-mail: gesposito@mail.dstb.uniud.it

Methods and experiments

The recombinant C1q-like domain of EMILIN-1 was expressed as (His)6-tagged protein and purified as previously described (Mongiat et al., 2000). The U-^{13}C, U-^{15}N, 80% ^{2}H labelled protein was produced and purified by ASLA (Riga, Latvia) according to the protocol optimised by Colombatti and co-workers.

NMR measurements were performed at 310 K on home-built/GE Omega spectrometers operating at proton frequencies of 750 MHz, 600 MHz or 500 MHz. All data were recorded on a 1.8 mM protein sample in aqueous buffer containing 20 mM NaH_2PO_4, 100 mM NaCl, 0.1% (w/v) NaN_3, and 5% (v/v) D_2O at pH 7.5, introduced into a Shigemi NMR tube. Proton chemical shifts were referenced to dioxane, whose resonance was set to 3.750 ppm. ^{13}C and ^{15}N chemical shifts were referenced indirectly to dioxane, using the absolute frequency ratio (Wishart et al., 1995).

Sequence-specific assignment of the backbone resonances was achieved using ^{1}H,^{15}N-HSQC experiments together with 3D triple-resonance experiments. Due to the size of the observed species, a trimer of 50,473 kDa, the performance of the NMR experiments was greatly improved by using the TROSY approach (Pervushin et al., 1997) at the highest magnetic field available. The pulse sequences employed were based on the experiments published by Salzmann and co-

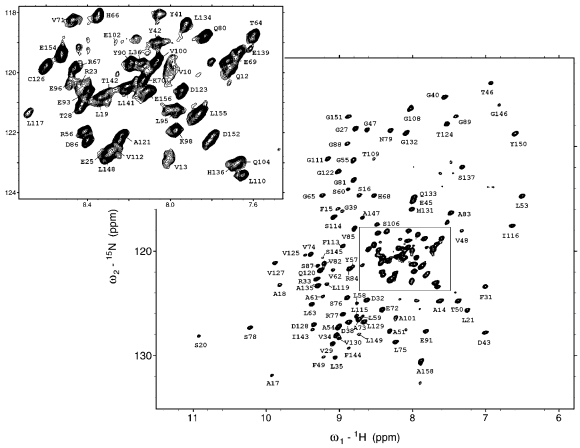

Figure 1. TROSY ^{1}H,^{15}N-HSQC spectrum recorded on a 1.8 mM U-^{13}C, U-^{15}N, 80% ^{2}H labelled C1q sample at 750 MHz and 310 K. Assignments of the central region of the spectrum are reported in the insert.

workers (Salzmann et al., 1999). Firstly, patterns of sequentially linked spin systems were identified using the 3D [^{15}N, ^{1}H]-TROSY-HNCA/[^{15}N, ^{1}H]-TROSY-HN(CO)CA pair of experiments. Secondly, sequential resonances where matched against the primary sequence of the protein using the 3D [^{15}N, ^{1}H]-TROSY-HNCACB/[^{15}N, ^{1}H]-TROSY-HN(CO)CACB pair of experiments, which allowed the identification of the amino acid types. Finally, the 3D [^{15}N, ^{1}H]-TROSY-HNCO/[^{15}N, ^{1}H]-TROSY-HN(CA)CO pair of experiments provided an alternative route for sequential assignment and was used to confirm and extend the previous attributions.

A 3D ^{15}N-NOESY-HSQC with a mixing time of 150 ms was also used to identify NH(i)-NH(i+1) through-space nuclear Overhauser connectivities.

The NMR spectra were processed using the program FELIX 2.3 (MSI) and analysed using XEasy 1.3.13 (Bartels et al., 1995).

Extent of assignment and data deposition

A TROSY ^{1}H,^{15}N-HSQC spectrum of the C1q domain of EMILIN-1 is shown in Figure 1. Backbone amide ^{1}H, ^{15}N resonances were unambiguously assigned for 132 residues out of the 137 expected (151 amino acids without the N-terminal Met residue and the (His)6 tag, minus 14 proline residues), giving an assignment percentage of 96%. No resonances could be observed for the backbone amide groups of residues Ala8, Asn37, Asn97, Ser103 and His157 due to fast hydrogen exchange with the solvent. Almost complete assignments were obtained for the $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ nuclei, where the only missing resonance pair belongs to Ala8, i.e., the residue that follows the (His)6 tag and is followed by a proline. No assignment was proposed for the initial Met and the six His residues owing to either absence or degeneracy of the corresponding signals. Three weak peaks, possibly belonging to His residues and absent in the triple resonance spectra, remained unassigned in the ^{1}H,^{15}N-HSQC spectrum. Finally, the identification of $^{13}C'$ chemical shifts was obtained for 147 amino acids (97%). Side-chain carboxyamide ^{13}C, ^{15}N and ^{1}H chemical shifts were assigned for two Asn and three Gln residues (Asn79, Asn97, Gln80, Gln104 and Gln120).

The chemical shift information obtained for the $^{13}C^{\alpha}$, $^{13}C^{\beta}$ and $^{13}C'$ nuclei was used to identify secondary structure elements in the protein according to the CSI method (Wishart and Sykes, 1994). Based on this analysis, eleven segments of β strands could be identified. These secondary structure elements in the EMILIN-1 C1q domain closely match the β strand pattern observed in the crystal structures of the two homologous domains from ACRP-30 and Collagen X (Shapiro and Scherer, 1998; Bogin et al., 2002).^{1}H, ^{13}C and ^{15}N chemical shifts have been deposited in the BioMagResBank (http://www.bmrb.wisc.edu) under the BMRB accession number 5882.

Acknowledgements

The authors would like to thank the Italian Ministry of Education (FIRB, COFIN), EMBO and JIF, for funding. Helpful discussions with Dr C.D. Blundell and the kind assistance of Dr A. Makek are gratefully acknowledged.

References

Bartels, C., Xia, T.H., Billeter, M., Guntert, P. and Wüthrich, K. (1995) *J. Biomol. NMR*, **6**: 1–10.

Bogin, O., Kvansakul, M., Rom, E., Singer, J., Yayon, A. and Hohenester, E. (2002) *Structure*, **10**: 165–173.

Doliana, R., Mongiat, M., Bucciotti, F., Giacomello, E., Deutzmann, R., Volpin, D., Bressan, G.M. and Colombatti, A. (1999) *J. Biol. Chem.*, **274**: 16773–16781.

Mongiat, M., Mungiguerra, G., Bot, S., Mucignat, M.T., Giacomello, E., Doliana, R. and Colombatti, A. (2000) *J. Biol. Chem.*, **275**: 25471–25480.

Pervushin, K., Riek, R., Wider, G. and Wüthrich, K. (1997) *Proc. Natl. Acad. Sci. USA*, **94**: 12366–12371.

Salzmann, M., Wider, G., Pervushin, K., Senn, H. and Wuthrich, K. (1999) *J. Am. Chem. Soc.*, **121**: 844–848.

Shapiro, L. and Scherer, P.E. (1998) *Curr. Biol.*, **8**: 335–338.

Spessotto, P., Cervi, M., Mucignat, M.T., Mungiguerra, G., Sartoretto, I., Doliana, R. and Colombatti, A. (2003) *J. Biol. Chem.*, **278**: 6160–6167.

Wishart, D.S. and Sykes, B.D. (1994) *J. Biol. NMR*, **4**: 171–180.

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.