Published in final edited form as:

J Struct Funct Genomics. 2014 December ; 15(4): 209–214. doi:10.1007/s10969-014-9185-y.

Solution NMR structures of Immunoglobulin-like domains 7 and 12 from Obscurin-like protein 1 contribute to the structural coverage of the human cancer protein interaction network

Surya VSRK Pulavarti,

Department of Chemistry, The State University of New York at Buffalo, and Northeast Structural Genomics Consortium, Buffalo, NY 14260, USA

Yuanpeng J. Huang,

Center of Advanced Biotechnology and Medicine and Department of Molecular Biology and Biochemistry, Rutgers, The State University of New Jersey and Northeast Structural Genomics Consortium, Piscataway, NJ 08854, USA

Kari Pederson,

Complex Carbohydrate Research Center, University of Georgia, and Northeast Structural Genomics Consortium, Athens, GA 30602, USA

Thomas B. Acton,

Center of Advanced Biotechnology and Medicine and Department of Molecular Biology and Biochemistry, Rutgers, The State University of New Jersey and Northeast Structural Genomics Consortium, Piscataway, NJ 08854, USA

Rong Xiao,

Center of Advanced Biotechnology and Medicine and Department of Molecular Biology and Biochemistry, Rutgers, The State University of New Jersey and Northeast Structural Genomics Consortium, Piscataway, NJ 08854, USA

John K. Everett,

Center of Advanced Biotechnology and Medicine and Department of Molecular Biology and Biochemistry, Rutgers, The State University of New Jersey and Northeast Structural Genomics Consortium, Piscataway, NJ 08854, USA

James H. Prestegard,

Complex Carbohydrate Research Center, University of Georgia, and Northeast Structural Genomics Consortium, Athens, GA 30602, USA

Gaetano T. Montelione, and

Center of Advanced Biotechnology and Medicine and Department of Molecular Biology and Biochemistry, Rutgers, The State University of New Jersey and Northeast Structural Genomics Consortium, Piscataway, NJ 08854, USA

Thomas Szyperski

Department of Chemistry, The State University of New York at Buffalo, and Northeast Structural Genomics Consortium, Buffalo, NY 14260, USA

Thomas Szyperski: szypersk@buffalo.edu

Abstract

High-quality solution NMR structures of immunoglobulin-like domains 7 and 12 from human Obscurin-like protein 1 were solved. The two domains share 30 % sequence identity and their structures are, as expected, rather similar. The new structures contribute to structural coverage of human cancer associated proteins. Mutations of Arg 812 in domain 7 cause the rare 3-M syndrome, and this site is located in a surface area predicted to be involved in protein-protein interactions.

Keywords

PF07679; Obscurin-like protein 1; cytoskeletal adapter; Ig-like domain; structural genomics

Introduction

Cytoskeletal adaptor proteins function by linking cytoskeletal proteins to maintain the structural integrity of cells. Obscurin-like protein 1 (OBSL1; UniProtKB accession number O75147) represents such an adapter protein and constitutes an integral part of the muscle sarcomere, where it is localized at both M- and Z-bands. It interacts with the proteins titin and myomesin [1, 2] and functions in some respects similarly to the protein obscurin (UniProt Q5VST9; hence its name), another very large cytoskeletal adaptor protein [1, 2]. Moreover, OBSL1 controls Golgi and dendrite morphogenesis through regulation of Cullin7 (CUL7) and Fbox/WD repeat containing protein 8 (FBXW8) via ubiquitin ligase signaling [3]. Mutations of OBSL1 [4, 5], CUL7 [6] and the coiled-coil domain containing protein 8 (CCDC8) [7] have been associated with the rare ‘3-M syndrome’ [8] which is manifested in postnatal growth restriction, recognizable facial features and skeletal abnormalities [5, 9]. OBSL1 also interacts with Mitogen-activated protein kinase 11 (MAPK11) and MAPK14 [10]. Both MAPK11 and MAPK14 are involved in cancer associated MAPK and TLR (Toll-like receptor) pathways.

The genes encoding OBSL1 and obscurin are considered to belong to a novel family of homologous genes derived from the Unc-89/OBSL family [1, 2]. OBSL1 is expressed in wide range of tissues as three major isoforms: A (~210 kDa), B (~110 kDa) and C (~150 kDa). While isoforms A and C are present in non-cardiac tissues, isoform B is highly expressed in heart and placental tissues [1]. According to UniProt entry O75147, the full length OBSL isoform A comprises 14 immunoglobulin-like (Ig-like) domains and a single fibronectin type III domain inserted after the fourth Ig-like domain [1]. However, a multiple sequence alignment (Supplementary Material) shows that UniProt Ig-like domain 12 (residues 1357–1534) actually comprises of two Ig-like domains. Furthermore; the NMR structure of a construct comprising residues 615–713 (Protein Data Bank ID: 2E6Q) reveals the presence of another Ig-like domain not listed in UniProt entry O75147. Taken together, protein OBSL1 thus comprises 16 Ig-like domains.

All Ig-like domains belong to Pfam [11] protein family PF07679 ('Intermediate-set'), which currently contains 1617 members from bacteria and eukaryotic organisms participating in 3030 different unique domain organizations (architectures). Remarkably, the sequence identity among the 16 Ig-like domains varies extensively (between 9 to 67 %, Supplementary Material). Previously, structures were obtained for Ig-like domains 1, 5, 6, and 8 by solving, respectively, the X-ray crystal structure of OBSL1(1–106) in complex with Titin- (M10) (PDB [12] IDs: 2WP3, 2WWK, 2WWM, 3KNB) [13, 14] and the three solution NMR structures of OBSL1(615–713) (PDB ID: 2E6Q), OBSL1(714–804) (PDB ID: 2E6P) and OBSL1(893–992) (PDB ID: 2CPC). These domains exhibit the expected Ig-like fold [15] in which ~8 antiparallel β -strands are arranged in two β -sheets forming a β -sandwich [15].

OBSL1 Ig-like domains 7 and 12 share 30 % sequence identity and were selected by the Northeast Structural Genomics (NESG) Consortium (<http://www.nesg.org>; target IDs: HR8578K, and HR8578D, respectively) for structure determination with the aim to provide extensive structural coverage for human cancer-associated proteins and protein complexes constituting the 'Human Cancer Protein Interaction Network' (HCPIN) [16]. Here we report the high-quality solution NMR structures of OBSL1(805–892) and OBSL1(1277–1357) which comprise Ig-like domains 7 and 12, respectively.

Material and methods

OBSL1(805–892) and OBSL1(1277–1357) were cloned, expressed, and purified following protocols [17–19] established by NESG (see Supplementary Material for details). The corresponding pET expression vectors (NESG HR8578K-805–892-Av6HT.12 and HR8578D-1277–1357-Av6HT.12) have been deposited in the PSI Materials Repository (<http://psimr.asu.edu/>). [U - ^{13}C , ^{15}N]-labeled protein samples were prepared at ~1 and ~0.6 mM concentrations for OBSL1(805–892) and OBSL1(1277–1357) in 90% H_2O /10% D_2O containing, respectively, 20 mM MES, along with 100 mM NaCl, 5 mM CaCl_2 , 10 mM DTT, 50 μM DSS and 0.02% NaN_3 at pH 6.5 and 10 mM Tris-HCl, 100 mM NaCl, 5 mM DTT, 0.02 % NaN_3 , 50 mM DSS at pH 7.5. [5% ^{13}C , U - ^{15}N]-labeled samples enabled stereospecific assignments of the methyl groups of Val and Leu residues [20]. Isotropic overall rotational correlation times around 6 ns were inferred for both proteins from backbone ^{15}N spin relaxation times for both proteins (see Supplementary Material; <http://www.nmr2.buffalo.edu/nesg.wiki>), indicating that the two proteins are monomeric in solution. This finding was confirmed by static light scattering (Supplementary Figs. S1 and S2).

NMR data were acquired at 25 °C on Varian INOVA 600 and 750 spectrometers equipped with cryogenic $^1\text{H}\{^{15}\text{N}, ^{13}\text{C}\}$ probes. The total measurement times for OBSL1(805–892) and OBSL1(1277–1357) were ~8 days for both proteins. Nearly complete sequence-specific ^1H , ^{13}C and ^{15}N resonance assignments (Table 1; Supplementary Figs. S3 and S4) were obtained from conventional triple-resonance NMR experiments (see Supplementary Material for details) using the programs AutoAssign [21, 22], followed by manual assignment of side-chain resonances. Resonance assignments were validated using AVS software suite [23]. For OBSL1(805–892), backbone ^{15}N - ^1H RDCs were measured for two

alignment media (Supplementary Material). Chemical shifts, NOESY peak lists and all time domain NMR data were deposited in the BioMagResBank [accession numbers 18563 and 18511 for OBSL1(805–892) and OBSL1(1277–1357), respectively].

Structure calculations were performed using standardized methods of the NESG consortium [24, 25] (<http://www.nmr2.buffalo.edu/nescg.wiki>) and consensus analysis of automated NOESY cross peak assignments provided by the programs CYANA [26, 27] and AutoStructure [28]. This analysis was based on ^1H - ^1H NOE-derived upper limit distance constraints, and back-bone dihedral angle constraints derived from chemical shifts using the program TALOS+ [29] for residues located in well-defined regular structural elements; orientational constraints were derived for a similar set from polypeptide backbone ^{15}N - ^1H RDCs for OBSL1(805–892) [30]. Stereospecific assignments for methylene protons were performed with the GLOMSA module of CYANA. The final structure calculations were performed on CYANA, followed by refinement in an ‘explicit water bath’ using the program CNS [31]. Structure validation was accomplished using the Protein Structure Validation Software (PSVS) server 1.4 [32], and agreement of structures and NOESY peak lists was verified using the AutoStructure RPF 2.2.2 package [24].

Results and discussion

High-quality solution NMR structures of OBSL1(805–892) and OBSL1(1277–1357) containing OBSL1 Ig-like domains 7 and 12, respectively, were obtained (Table I) and the coordinates were deposited in the PDB [12] on 06/30/2012 (accession code 2LVC) and 06/08/2012 (accession code 2LU7). The domain structures exhibit the expected Ig-like fold in which two antiparallel β -sheets form a sandwich. In both domains, sheet I contains strands $\beta 1$, $\beta 2$, $\beta 5$, and $\beta 4$, and sheet II contains strands, $\beta 3$, $\beta 6$ and $\beta 7$) (Fig. 1b, 1g) In addition to these β -strands, domain 7 exhibits a β -strand $1'$ (808–810) contributing to β -sheet-II as well as a 3_{10} -helix (871–873) located in the loop connecting β -strands 5 and 6, and domain 12 exhibits an additional β -strand 8 (Fig. 1b, 1g) added to β -sheet-I. Strand $1'$ in domain 7 arises from a β -bulge formed by Arg 812 and Glu 813, while a twist at *cis*-Pro 811 results in the two short strands $\beta 1$ and $\beta 1'$ participating in different sheets.

For domain 12, two distinct sets of signals with approximately equal intensities are observed in 2D [^{15}N , ^1H] HSQC (Fig. S4) for the polypeptide backbone NH moieties of residues 1282, 1294–1302, 1319–1320, 1323–1325, 1340–1351 and 1354–1355 (Fig. 1f). The lack of exchange broadening of these signals reveals that the two conformations are in slow exchange on the chemical shift time scale, and the analysis of sequential NOEs between Gly 1298 and Pro 1299 revealed that this is due to the Pro being present in both *cis* and *trans* conformation (note that the structure in the PDB contains Pro 1299 in the *cis* conformation). As expected for 30 % sequence identity, the two structures are rather similar: the root mean square deviation (RMSD) calculated between the mean coordinates of the backbone heavy atoms N, C^α , and C' of the β -strands present in both domains is 1.9 Å.

Since Ig-like domains are highly abundant in vertebrates [15], a search of the PDB for structurally similar proteins using the DALI server [33] yielded a very large number of hits spanning a quasi-continuum of Z-scores > 4.0 (i.e., > 953 and > 984 for domains 7 and 12,

respectively), which included the structures of OBSL1 Ig-like domains 1, 5, 6 and 8 with Z-scores ranging from 8 to 14 and corresponding backbone RMSD values between 1.4 and 2.5 Å (Table SI). Electrostatic potential calculations show that domains 7 and 12 both exhibit a mixed, albeit rather different surface charge distribution (Fig. 1d, 1i). This observation indicates that the two domains interact with different epitopes on other cytoskeletal proteins. We thus used the PredUs server [34, 35] to predict potential protein-protein interaction sites. Inspection of Figs 1e, 1j shows that indeed quite different regions were identified for the two domains. Consistent with the prediction for domain 7, mutations of Arg 812 causing the 3-M syndrome [4] are located in the predicted protein-protein interaction site. Inspection of OBSL1 Ig-like domain 1 in a complex with titin Ig-like domain M10 [14] reveals that different surface regions are involved in the protein-protein complex formation for this domain. Overall, it might well be that each of the OBSL1 Ig-like domains is involved in distinctly different protein-protein interactions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the National Institutes of Health, grant number: U54 GM094597 (T.S., J.H.P. and G.T.M.).

Abbreviations

CUL7	Cullin7
CCDC8	Coiled-coil domain containing protein 8
DSS	4,4-dimethyl-4-silapentane-1-sulfonate sodium salt
DTT	Dithiothreitol
FBXW8	Fbox and WD repeat containing protein 8
HCPIN	Human Cancer Pathway Interaction Network
Ig-like	Immunoglobulin-like
MES	2-(<i>N</i> -morpholino)ethanesulfonic acid
MAPK	Mitogen-activated protein kinase
NESG	Northeast Structural Genomics Consortium
NOESY	Nuclear Overhauser enhancement spectroscopy
OBSL1	Obscurin-like protein 1
PDB	Protein Data Bank
RDC	Residual dipolar coupling

RMSD	Root mean square deviation
TLR	Toll-like receptor
Tris	Tris(hydroxymethyl)aminomethane

References

1. Geisler SB, Robinson D, Hauringa M, Raeker MO, Borisov AB, Westfall MV, Russell MW. Obscurin-like 1, OBSL1, is a novel cytoskeletal protein related to obscurin. *Genomics*. 2007; 89:521–531. [PubMed: 17289344]
2. Fukuzawa A, Lange S, Holt M, Vihola A, Carmignac V, Ferreira A, Udd B, Gautel M. Interactions with titin and myomesin target obscurin and obscurin-like 1 to the M-band – implications for hereditary myopathies. *J Cell Sci*. 2008; 121:1841–1851. [PubMed: 18477606]
3. Litterman N, Ikeuchi Y, Gallardo G, O'Connell BC, Sowa ME, Gygi SP, Harper JW, Bonni A. An OBSL1-Cul7 Fbxw8 ubiquitin ligase signaling mechanism regulates golgi morphology and dendrite patterning. *PLoS Biol*. 2011; 9:e1001060. [PubMed: 21572988]
4. Huber C, Fradin M, Edouard T, Le Merrer M, Alanay Y, Da Silva DB, David A, Hamamy H, van Hest L, Lund AM, Michaud J, Oley C, Patel C, Rajab A, Skidmore DL, Stewart H, Tauber M, Munnich A, Cormier-Daire V. OBSL1 mutations in 3-M syndrome are associated with a modulation of IGFBP2 and IGFBP5 expression levels. *Hum Mutat*. 2010; 31:20–26. [PubMed: 19877176]
5. Hanson D, Murray PG, Sud A, Temtamy SA, Aglan M, Superti-Furga A, Holder SE, Urquhart J, Hilton E, Manson FDC, Scambler P, Black GCM, Clayton PE. The primordial growth disorder 3-M syndrome connects ubiquitination to the cytoskeletal adaptor OBSL1. *Am J Hum Genet*. 2009; 84:801–806. [PubMed: 19481195]
6. Sasaki K, Okamoto N, Kosaki K, Yorifuji T, Shimokawa O, Mishima H, Yoshiura Ki, Harada N. Maternal uniparental isodisomy and heterodisomy on chromosome 6 encompassing a CUL7 gene mutation causing 3M syndrome. *Clin Genet*. 2011; 80:478–483. [PubMed: 21166787]
7. Hanson D, Murray Philip G, O'Sullivan J, Urquhart J, Daly S, Bhaskar Sanjeev S, Biesecker Leslie G, Skae M, Smith C, Cole T, Kirk J, Chandler K, Kingston H, Donnai D, Clayton Peter E, Black Graeme CM. Exome sequencing identifies CCDC8 mutations in 3-M syndrome, suggesting that CCDC8 contributes in a pathway with CUL7 and OBSL1 to control human growth. *Am J Hum Genet*. 2011; 89:148–153. [PubMed: 21737058]
8. Clayton PE, Hanson D, Magee L, Murray PG, Saunders E, Abu-Amro SN, Moore GE, Black GCM. Exploring the spectrum of 3-M syndrome, a primordial short stature disorder of disrupted ubiquitination. *Clin Endocrinol*. 2012; 77:335–342.
9. Hanson D, Murray PG, Black GCM, Clayton PE. The genetics of 3-M syndrome: unravelling a potential new regulatory growth pathway. *Horm Res Paediatr*. 2011; 76:369–378. [PubMed: 22156540]
10. Varjosalo M, Keskitalo S, Van Drogen A, Nurkka H, Vichalkovski A, Aebersold R, Gstaiger M. The protein interaction landscape of the human CMGC kinase group. *Cell Reports*. 2013; 3:1306–1320. [PubMed: 23602568]
11. Finn RD, Mistry J, Tate J, Coghill P, Heger A, Pollington JE, Gavin OL, Gunasekaran P, Ceric G, Forslund K, Holm L, Sonnhammer ELL, Eddy SR, Bateman A. The Pfam protein families database. *Nucleic Acids Res*. 2010; 38:D211–D222. [PubMed: 19920124]
12. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. *Nucleic Acids Res*. 2000; 28:235–242. [PubMed: 10592235]
13. Sauer F, Vahokoski J, Song Y-H, Wilmanns M. Molecular basis of the head-to-tail assembly of giant muscle proteins obscurin-like 1 and titin. *EMBO Rep*. 2010; 11:534–540. [PubMed: 20489725]
14. Pernigo S, Fukuzawa A, Bertz M, Holt M, Rief M, Steiner RA, Gautel M. Structural insight into M-band assembly and mechanics from the titin-obscurin-like-1 complex. *Proc Natl Acad Sci U S A*. 2010; 107:2908–2913. [PubMed: 20133654]

15. Bork P, Holm L, Sander C. The immunoglobulin fold: structural classification, sequence patterns and common core. *J Mol Biol.* 1994; 242:309–320. [PubMed: 7932691]
16. Huang YJ, Hang D, Lu LJ, Tong L, Gerstein MB, Montelione GT. Targeting the human cancer pathway protein interaction network by structural genomics. *Mol Cell Proteomics.* 2008; 7:2048–2060. [PubMed: 18487680]
17. Xiao R, Anderson S, Aramini J, Belote R, Buchwald WA, Ciccocanti C, Conover K, Everett JK, Hamilton K, Huang YJ, Janjua H, Jiang M, Kornhaber GJ, Lee DY, Locke JY, Ma L-C, Maglaqui M, Mao L, Mitra S, Patel D, Rossi P, Sahdev S, Sharma S, Shastry R, Swapna GVT, Tong SN, Wang D, Wang H, Zhao L, Montelione GT, Acton TB. The high-throughput protein sample production platform of the Northeast Structural Genomics Consortium. *J Struct Biol.* 2010; 172:21–33. [PubMed: 20688167]
18. Acton TB, Xiao R, Anderson S, Aramini J, Buchwald WA, Ciccocanti C, Conover K, Everett J, Hamilton K, Huang YJ, Janjua H, Kornhaber G, Lau J, Lee DY, Liu GH, Maglaqui M, Ma LC, Mao L, Patel D, Rossi P, Sahdev S, Shastry R, Swapna GVT, Tang YF, Tong SC, Wang DY, Wang H, Zhao L, Montelione GT. Preparation of protein samples for NMR structure, function, and small-molecule screening studies. *Methods Enzymol.* 2011; 493:21–60. [PubMed: 21371586]
19. Acton TB, Gunsalus KC, Xiao R, Ma LC, Aramini J, Baran MC, Chiang Y-W, Climent T, Cooper B, Denisova NG, Douglas SM, Everett JK, Ho CK, Macapagal D, Rajan PK, Shastry R, Shih LY, Swapna GVT, Wilson M, Wu M, Gerstein M, Inouye M, Hunt JF, Montelione GT. Robotic cloning and protein production platform of the Northeast Structural Genomics Consortium. *Methods Enzymol.* 2005; 394:210–243. [PubMed: 15808222]
20. Neri D, Szyperski T, Otting G, Senn H, Wuthrich K. Stereospecific nuclear magnetic resonance assignments of the methyl groups of valine and leucine in the DNA-binding domain of the 434 repressor by biosynthetically directed fractional ¹³C labeling. *Biochemistry.* 1989; 28:7510–7516. [PubMed: 2692701]
21. Moseley HNB, Monleon D, Montelione GT. Automatic determination of protein backbone resonance assignments from triple resonance nuclear magnetic resonance data. *Methods Enzymol.* 2001; 339:91–108. [PubMed: 11462827]
22. Zimmerman DE, Kulikowski CA, Huang YP, Feng WQ, Tashiro M, Shimotakahara S, Chien CY, Powers R, Montelione GT. Automated analysis of protein NMR assignments using methods from artificial intelligence. *J Mol Biol.* 1997; 269:592–610. [PubMed: 9217263]
23. Moseley HNB, Sahota G, Montelione GT. Assignment validation software suite for the evaluation and presentation of protein resonance assignment data. *J Biomol NMR.* 2004; 28:341–355. [PubMed: 14872126]
24. Huang YJ, Powers R, Montelione GT. Protein NMR recall, precision, and F-measure scores (RPF scores): structure quality assessment measures based on information retrieval statistics. *J Am Chem Soc.* 2005; 127:1665–1674. [PubMed: 15701001]
25. Liu G, Shen Y, Atreya HS, Parish D, Shao Y, Sukumaran DK, Xiao R, Yee A, Lemak A, Bhattacharya A, Acton TA, Arrowsmith CH, Montelione GT, Szyperski T. NMR data collection and analysis protocol for high-throughput protein structure determination. *Proc Natl Acad Sci U S A.* 2005; 102:10487–10492. [PubMed: 16027363]
26. 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. [PubMed: 9367762]
27. Herrmann T, Güntert P, Wüthrich K. Protein NMR Structure Determination with Automated NOE Assignment Using the New Software CANDID and the Torsion Angle Dynamics Algorithm DYANA. *J Mol Biol.* 2002; 319:209–227. [PubMed: 12051947]
28. Huang YJ, Tejero R, Powers R, Montelione GT. A topology-constrained distance network algorithm for protein structure determination from NOESY data. *Proteins.* 2006; 62:587–603. [PubMed: 16374783]
29. Cornilescu G, Delaglio F, Bax A. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J Biomol NMR.* 1999; 13:289–302. [PubMed: 10212987]
30. Liu Y, Prestegard JH. Measurement of one and two bond N-C couplings in large proteins by TROSY-based J-modulation experiments. *J Magn Reson.* 2009; 200:109–118. [PubMed: 19581113]

31. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr Sect D Biol Crystallogr*. 1998; 54:905–921. [PubMed: 9757107]
32. Bhattacharya A, Tejero R, Montelione GT. Evaluating protein structures determined by structural genomics consortia. *Proteins*. 2007; 66:778–795. [PubMed: 17186527]
33. Holm L, Sander C. Dali: a network tool for protein structure comparison. *Trends Biochem Sci*. 1995; 20:478–480. [PubMed: 8578593]
34. Zhang QC, Deng L, Fisher M, Guan J, Honig B, Petrey D. PredUs: a web server for predicting protein interfaces using structural neighbors. *Nucleic Acids Res*. 2011; 39:W283–W287. [PubMed: 21609948]
35. Zhang QC, Petrey D, Norel R, Honig BH. Protein interface conservation across structure space. *Proc Natl Acad Sci U S A*. 2010; 107:10896–10901. [PubMed: 20534496]

Fig. 1.

a Stereo view of the 20 conformers representing the solution structure of OBSL1(805-892) obtained after superposition of the C α atoms of β -strand residues for minimal RMSD. Terminal, flexibly disordered residues 803-805 and 892 were omitted for clarity, and the termini are labeled as 'N' and 'C'. **b** Ribbon diagram of the lowest-energy conformer of OBSL1(805-892) containing Ig-like domain 7. The location of secondary structure elements are: β -strand 1' (808-810), β 1(815-818), β 2(825-831), β 3(839-842), β 4(854-858), β 5(861-866), β 6(876-880), β 7(885-891). β -strands are depicted in cyan, the

3_{10} α -helix is shown in red and yellow, and other polypeptide segments are in gray. **c** Sausage representation of backbone and superposition of the conformation of the best defined side chains (Table 1). A spline curve was drawn through the mean positions of C^α atoms of residues 806–891 with the thickness proportional to the mean global displacement of C^α atoms in the 20 conformers superimposed in **(a)**. **d** The electrostatic surface potential of lowest energy NMR conformer. Positively and negatively charged regions are colored in blue and red, respectively, and the neutral regions are colored in white. **e** Surface and space filling representations of the lowest-energy conformer of OBSL1(805–892) colored according to scores from PredUs prediction of regions potentially involved in protein–protein interactions; the default PredUs color scheme for residues with scores larger than zero are shown from light red to red with increasing score. **f–j** Show the same as in **a–e** but for OBSL1(1277–1357) containing Ig-like domain 12. Terminal, flexibly disordered residues 1277–1279 and 1357 were excluded for clarity. The location of secondary structure elements are: β 1(1279–1281), β 2(1287–1290), β 3(1301–1304), β 8(1307–1308), β 5(1315–1319), β 6(1324–1328), β 7(1337–1342), and β 8(1348–1354). Due to the presence of both the cis- and trans-conformation for Pro 1299, residues 1282, 1294–1302, 1319–1320, 1323–1325, 1340–1351 and 1354–1355 exhibit two sets of signals in 2D [^{15}N , ^1H] HSQC (Fig. S4). These polypeptide segments are highlighted in grey.in **(f)**

Table 1Statistics^a for NMR Structures of OBSL1(805–892) and OBSL1(1277–1357)

	OBSL1 (805–892)	OBSL1 (1277–1357)		
Completeness of resonance assignments ^b				
Backbone (%)	99.5	94.8		
Side chain (%)	99.1	99.4		
Aromatic (%)	92.2	100		
Stereospecific methyl (%)	100	100		
Conformationally-restricting constraints ^c				
Distance constraints				
Total	1380	1270		
intra-residue ($i = j$)	228	170		
sequential ($ i - j = 1$)	389	310		
medium range ($1 < i - j < 5$)	198	173		
long range ($ i - j \geq 5$)	565	617		
Dihedral angle constraints	68	64		
RDC constraints (¹⁵ N- ¹ H)	147			
No. of constraints per residue	15.8	16.9		
No. of long range constraints per residue	6.7	8.2		
Residual constraint violations ^c				
Average no. of distance violations per structure:				
0.1 – 0.2 Å	4.45	13.9		
0.2 – 0.5 Å	0.7	3.85		
> 0.5 Å	0	0		
Average no. of dihedral angle violations per structure:				
1 – 10°	7.6	11.7		
> 10°	0	0		
Model Quality ^c				
RMSD backbone atoms (Å) ^d	0.6	0.5		
RMSD heavy atoms (Å) ^d	1.1	0.9		
RMSD bond lengths (Å)	0.021	0.021		
RMSD bond angles (°)	1.1	1.1		
RMSD heavy atoms of molecular core (Å) ^e	0.9	0.8		
MolProbity Ramachandran statistics ^{c,d}				
Most favored regions (%)	96.5	94.4		
Allowed regions (%)	3.0	5.3		
Disallowed regions (%)	0.6	0.3		
Global quality scores (Raw / Z-score) ^c				
Verify3D	0.32	–2.25	0.43	–0.48

	OBSL1 (805–892)		OBSL1 (1277–1357)	
ProsaII	0.48	−0.70	0.44	−0.87
Procheck (phi-psi) ^d	−0.68	−2.36	−0.85	−3.03
Procheck (all) ^d	−0.45	−2.66	−0.51	−3.02
MolProbity clash score	13.75	−0.83	18.56	−1.66
RPF Scores ^f				
Recall / Precision	0.993	0.876	0.966	0.832
F-measure / DP-score	0.931	0.860	0.894	0.829
<i>BMRB accession number</i>		18563		18511
<i>PDB ID</i>		2LVC		2LU7

^aStructural statistics computed for the ensemble of 20 deposited structures.

^bComputed using AVS software [23] from the expected number of resonances, excluding: highly exchangeable protons (N-terminal, Lys, and Arg amino groups, hydroxyls of Ser, Thr, Tyr), carboxyls of Asp and Glu, non-protonated aromatic carbons, and the C-terminal His₆ tag.

^cCalculated using PSVS 1.4 [32]. Average distance violations were calculated using the sum over r^{-6} .

^dBased on ordered residue ranges [S(phi) + S(psi) > 1.8], Residues 807–891 for OBSL1(805–892) and residues 1278–1356 for OBSL1(1277–1357)

^eCalculated using MOLMOL [35] for residues 808,809,811,815,817,819, 825,827–829,831,840,847,854,855, 864, 865,867,869,871,877,879–881,886, 889–892 for OBSL1(805–892) and residues 1277,1279,1281–1283,1287–1291, 1293,1300,1302–1304,1308–1311, 1315,1317,1324–1326,1328,1331,1336, 1338–1340,1342,1343,1347,1349–1354 for OBSL1(1277–1357)

^fRPF scores [24] reflecting the goodness-of-fit of the final ensemble of structures (including disordered residues) to the NOESY data and resonance assignments.