Recombinant Expression and Purification of Cytoplasmic Domain of Syndecan-2 Proteoglycan

Young Kee Chae[†] and Weontae Lee[*]

Department of Biochemistry, Yonsei University, Seoul 120-749, Korea. [*]E-mail: wlee@spin.yonsei.ac.kr
[†]Department of Chemistry and Recombinant Protein Expression Center, Sejong University, Seoul 143-747, Korea
Received October 22, 2008

The cytoplasmic domain of syndecan-2, a type I transmembrane heparan sulfate proteoglycan, was over-expressed as a fused form with the ubiquitin molecule in Rosetta2(DE3)pLysS, a special strain of *Escherichia coli*, and the fusion protein was purified using immobilized metal affinity chromatography (IMAC). The cytoplasmic domain was released from its fusion partner by using yeast ubiquitin hydrolase (YUH), and subsequently purified by reverse phase chromatography. The integrity of the resulting peptide fragment was checked by MALDI-TOF mass and NMR spectroscopy. The final yields of the target peptide were around 2 and 1.5 mg per liter of LB and minimal media, respectively. The recombinant expression and purification of this domain will enable its structural and functional studies using multidimensional NMR spectroscopy and X-ray crystallography.

Key Words : Syndecan-2, Proteoglycan, NMR, Recombination

Introduction

The syndecans are members of a family of type I transmembrane heparan sulfate proteoglycan (HSPG), and are involved in the regulation of many cellular processes.[1-3] Four sub-types of mammalian syndecans have been reported and among them, syndecan-2 plays a role especially in the cancer development.[4,5] For example, syndecan-2 can affect the basal and chemotherapy-induced apoptosis in osteosarcoma.[4] It can also suppress matrix metalloproteinase-2 activation, suppressing metastasis.[5]

Syndecans have three distinct regions (C1, C2, and V) in their cytoplasmic domain.[3] The C1 region is located right next to the transmembrane domain, then the V region, and the C2 region at the carboxy terminus. The C1 region is highly homologous to C2 region. The C2 region of all four mammalian syndecans has been reported to be able to interact with several PDZ domain-containing proteins like syntenin,[6] CASK,[7] or GIPC (synectin/SEMCAP-1).[8] These interactions may be involved in trafficking and/or establishing a network of submembranous signaling complexes.[9,10] The V region, in between C1 and C2 domains, is unique to each syndecan. Unlike its close relative, syndecan-4, which can form a compact homodimer of the cytoplasmic domain and interact with phosphatidylinositol 4,5 bisphosphate (PIP_2), syndecan-2 cytoplasmic domain does not bind PIP_2 nor readily form dimers.[11-15]

The three dimensional structures of the cytoplasmic domain of syndecan-4 were solved by NMR spectroscopy.[13,16] Although many have been reported on its function, the detailed structural work on syndecan-2 has not been done yet.[17] To better understand its function, it is necessary to get the structural information. For the structural studies, a large amount of sample is needed, and for NMR spectroscopy, it also needs to be labeled, which requires the recombinant expression of the peptide in a suitable host. Here we present our method of producing and purifying recombinant cytoplasmic domain of syndecan-2 (2L) by using the ubiquitin fusion system in Rosetta2(DE3)pLysS, a specialized strain of *Escherichia coli*.

Materials and Methods

Construction of the 2L expression plasmid with $(His)_6$-tagged ubiquitin as a fusion partner. The gene coding for 2L was amplified by PCR. The primers were synthesized by CoreBio, Inc. (Seoul, Korea). The sense primer was 5'-ggg ccc gga tcc cgt atg cgt aaa aaa gat gaa-3', and the antisense, 5'-ggg ccc ctc gag tta agc ata aaa ttc ttt agt-3'. The amino acid sequence of 2L is RMRKK DEGSY DLGER KPSSA AYQKA PTKEF YA (32mer), but two additional residues, Gly and Ser, were added at the amino terminus due to the BamHI restriction site. The amplified PCR product was digested by BamHI and XhoI, and inserted into the vector pET-28a/ubi[18] which was previously digested with BamHI and SalI. The resulting plasmid was named pET-28a/ubi/2L.

Expression and purification of ubiquitin-2L fusion protein from an LB medium. The pET-28a/ubi/2L plasmid was brought into the expression host, BL21(DE3)pLysS or Rosetta2(DE3)pLysS (Novagen, Madison, WI). A single colony was used to inoculate a 100 mL LB medium supplemented with 50 μg/mL kanamycin and 34 μg/mL chloramphenicol, and grown overnight in a shaking incubator at 37 °C. The fully grown culture was used as an inoculum for a fresh one liter LB medium with the same antibiotics the next morning. The culture was grown at 37 °C, and IPTG was added to a final concentration of 0.5 mM when the optical density at 600 nm reached 1.0. The culture was harvested 3 hours later and the cells were resuspended in 30 mL of 10 mM TrisHCl pH 8.0. The cells were lysed by

'freeze-and-thaw' and the DNA was fragmented by ultrasonication. The soluble fraction was retained after centrifugation at 18,000 rpm for 20 min, and loaded onto HiTrap Chelating HP column (5 mL). The imidazole gradient of 0 to 0.4 M was applied to the column on the ÄKTA FPLC system (GE Healthcare, Piscataway, NJ, USA). The fractions containing ubiquitin-2L were pooled and concentrated down to 5 mL. The amount of protein in the pooled fractions was measured by using Protein Assay kit (Bio-Rad, Hercules, CA, USA).

Expression and purification of ubiquitin-2L fusion protein from a minimal medium. BL21(DE3)pLysS or Rosetta2(DE3)pLysS containing the pET-28a/ubi/2L was grown at 37 °C in a 5 ml LB medium inoculated from a single colony. 1 mL of the fully grown culture was used as an inoculum for a 100 mL of the minimal medium and grown overnight at 37 °C. The fully grown culture was used in turn as an inoculum for a 0.9 liter minimal medium and the culture was grown at 37 °C. For the uniform [^{15}N]-labeling, 1 g of $^{15}NH_4Cl$ per liter culture was provided as a sole nitrogen source. The production and purification steps were the same as the previous section.

Purification of 2L. To the ubiquitin-2L fusion protein, β-mercaptoethanol and YUH were added to the final concentrations of 1 mM and 2 mg/mL, respectively. The mixture was incubated at room temperature overnight. The reaction mixture was directly loaded onto a Resource RPC column (GE Healthcare, Piscataway, NJ, USA), and an acetonitrile gradient of 20 to 60% was applied using the HP1100 HPLC system (Agilent Technologies, Palo Alto, CA, USA). The 2L fraction was pooled and lyophilized. The final product was checked by MALDI-TOF. The mass of the peptide was measured directly after lyophilization.

NMR experiments. The NMR sample contained 0.05 mM [^{15}N]-2L in 10 mM sodium phosphate buffer pH 7.0 and 10% D_2O. The ^{1}H-^{15}N heteronuclear single quantum coherence (HSQC) spectrum was collected at 25 °C on a Bruker Avance 500 MHz spectrometer. The raw data contained 2048 and 256 complex points in t_2 and of t_1, respectively. The data was processed using NMRPIPE software package.[19] The final spectrum contained 1024 and 256 real points in t_2 and of t_1, respectively.

Results and Discussion

Construction of expression plasmids. The gene coding for ubiquitin-2L was inserted into $(His)_6$-Tag containing vectors to facilitate the purification of the desired proteins. The ubiquitin fusion system was chosen according to the work done by Moon *et al.*[18] This system was chosen also because the ubiquitin was small enough to be refolded *in vitro* relatively easily in case the fusion protein was expressed as an inclusion body.

Expression and purification of ubiquitin-2L fusion protein. The expressed proteins appeared as 2 bands whose sizes corresponded to ubiquitin alone and ubiquitin/2L fusion protein as shown in lanes 2 and 4 of Figure 1A. The band corresponding to ubiquitin/2L fusion protein appeared to be the weaker of of the two, and this could be due to the presence of protease in the *E. coli* cells as in the case of syndecan-4L.[20] Curiously, Rosetta2(DE3)pLysS showed a better expression of the fusion protein than the BL21(DE3)-pLysS although the genotypes of the two strains are virtually identical, and there was no rare codon in the syndecan-4L sequence. The only difference between those two is that Rosetta2(DE3)pLysS has the ability to make additional tRNAs for rare codons such as AGG, AGA, AUA, CUA, CCC, and GGA. The fusion protein was eluted out around at the imidazole concentration of 200 mM (Fig. 2A). The purity of the fusion protein was examined by SDS-PAGE as shown in Figure 1B. The final yield of the fusion protein was around 20 or 15 mg per liter of LB or minimal medium, respectively. However, this amount included not only the ubiquitin with full-sized 2L, but also the ubiquitin molecule alone. The ratio between the full fusion protein and ubiquitin molecule was measured to be 0.7 with the help of ImageJ software (NIH, USA). Thus, it would be safe to conclude that amount of the ubiquitin/2L fusion protein was around 40% of the calculated mass.

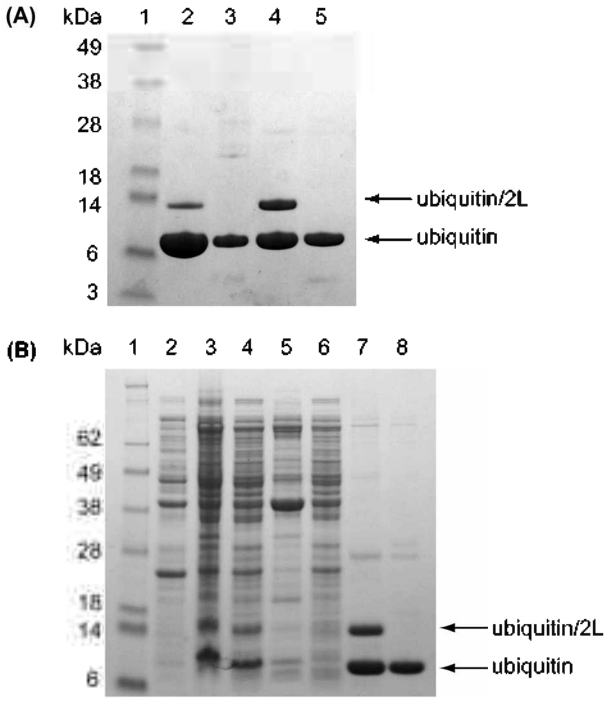

Figure 1. Expression and purification of syndecan 2L. (A) Comparison of expression levels of fusion protein in BL21(DE3)pLysS and Rosetta2(DE3)pLysS. Lane 1, size marker; lane 2, ubiquitin-2L fusion protein from BL21(DE3)pLysS; lane 3, sample of lane 2 after YUH cleavage; lane 4, ubiquitin-2L fusion protein from Rosetta2(DE3)pLysS; lane 5, sample of lane 4 after YUH cleavage (B) Purification of 2L produced from 1 liter LB medium as monitored by 16 % SDS-PAGE. Lane 1, size marker; lane 2 and 3, whole cell lysate before and after IPTG induction; lanes 4 and 5, supernatant and pellet of cell lysate, respectively; lane 6, flow-through fraction of supernatant from HiTrap Chelating HP column; lane 7, fraction that was bound to the HiTrap column; lane 8, sample of lane 7 after YUH cleavage reaction.

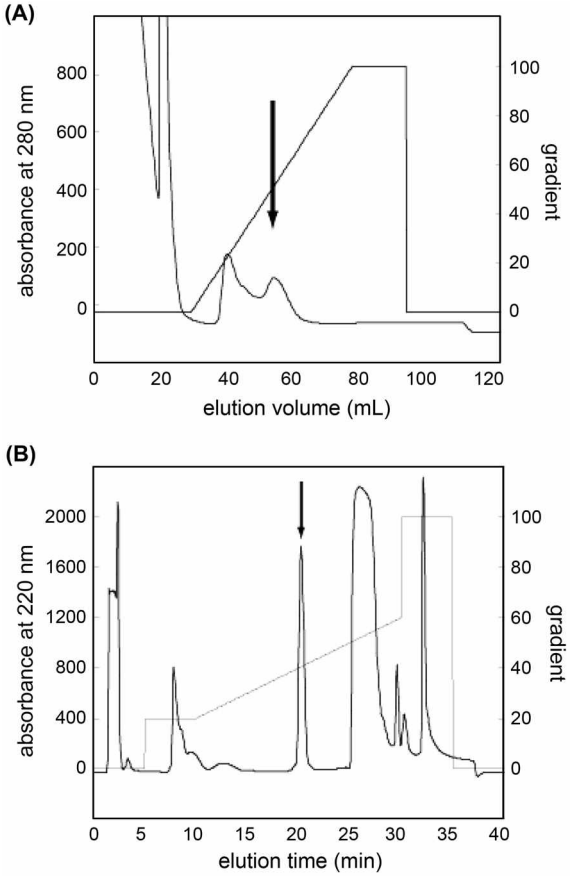

Figure 2. The elution profile from liquid chromatography. (A) HiTrap Chelation HP column on FPLC. The peak corresponding to ubiquitin-2L fusion protein is marked with an arrow. The imidazole gradient is also shown. (B) Resource RPC column on HPLC. The peak corresponding to 2L is marked with an arrow. The acetonitrile gradient is also shown.

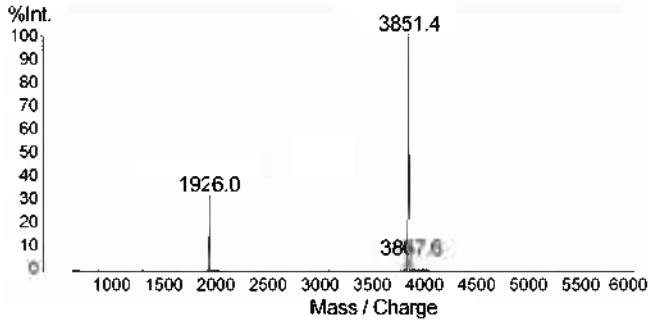

Figure 3. MALDI-TOF spectrum of purified syndecan 2L. Single and double charged molecular ions are seen at m/z = 3851.4 and 1926.0, respectively.

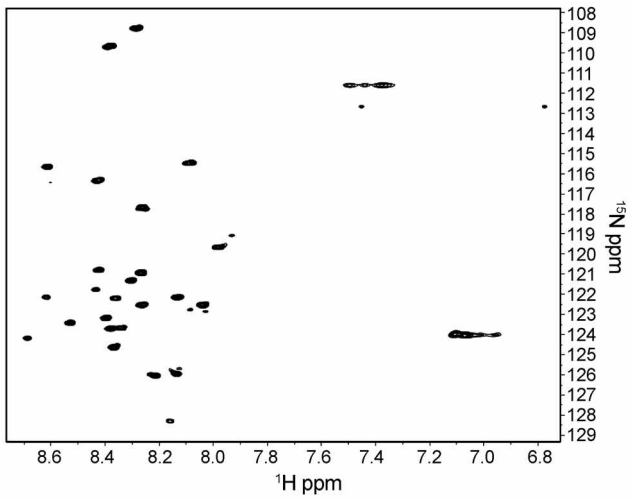

Figure 4. [^{1}H-^{15}N] HSQC spectrum of 2L. The NMR sample contained 0.05 mM [^{15}N]-2L in 10 mM sodium phosphate buffer pH 7.0 and 10% D_2O. The spectrum was collected at 25 °C on a Bruker Avance 500 MHz spectrometer.

Purification of 2L peptide. The 2L peptide was efficiently clipped off of ubiquitin by YUH. Compared to the widely used proteases such as thrombin and TEV protease, very small amount of YUH (50 μg) was enough to cleave 10-20 mg of the ubiquitin fusion protein. Following cleavage, the entire reaction mixture was loaded onto a Resource RPC column in 1 mL aliquots on HP1100 HPLC system. The volume of the cleavage reaction mixture was 5 mL, so the reverse phase chromatography was repeated 5 times. The peptide was eluted around at 28% acetonitrile, and all the 2L-containing fractions were pooled and lyophilized (Fig. 2B). The final yield of 2L was around 2 or 1.5 mg per liter of LB or minimal medium, respectively. MALDI-TOF was used to verify the purified peptide (Fig. 3). MALDI-TOF showed a major peak at 3851.4 Da, which is in good agreement with the theoretical molecular weight of 3853.2 Da.

HSQC spectrum of 15N-labeled 2L peptide. The HSQC spectrum showed 27 strong and sharp signals, some weak ones, and some broad ones (Fig. 4). Considering the theoretical number of resonances, there should be 31 signals from backbone, 2 from the sidechain amide group, and several folded peaks from arginines and lysines. Among these, the peaks with proton frequency greater than 7.8 ppm corresponded to the backbone amide groups. The pair around (^{15}N, ^{1}H) = (112.5 ppm, 6.75/7.45 ppm) was characteristic resonances of the sidechain amide protons of Asn or Gln. There was one Gln but no Asn in 2L, so these peaks could easily be assigned to the sidechain amide group of Gln-23, and this finding provided another assurance of the purified 2L. The broad peaks around (^{15}N, ^{1}H) = (124 ppm, 7.0 ppm) were the folded ones judged by their negative intensities. These peaks came from N^{η}/H^{η} of arginines. Another set of broad peaks around (^{15}N, ^{1}H) = (112 ppm, 7.4 ppm) were the double folded ones judged by their positive intensities and separation from backbone resonances. These could have come from N^{ζ}/H^{ζ} of lysines.

Conclusion

The syndecan-2L peptide was recombinantly expressed and purified successfully, suggesting that the choice of a proper strain (as Rosetta2(DE3)pLysS in this present study) could be a key step in this procedure. The sequences of syndecan-2L and 4L are very similar (sequence identity of

66%), but their expression profiles in BL21(DE3)pLysS exhibited quite differently, and the final product of 2L varied much by different strains. Therefore, the shrewd choice of a suitable expression host was vital in this case. The ubiquitin fusion system was proven to be successful again for recombinant 2L peptide expression. The biggest advantage of the ubiquitin fusion system over the others stems from the small size of ubiquitin, which provides a higher net amount of the target peptides and relatively easier refolding in case of the inclusion body formation. The NMR and X-ray structural studies using this method are now in progress.

Acknowledgments. This work was supported by Sejong University research grant (to Y.K.C), and the KOSEF grant (to W.L) funded by the Korea government (MOST) (R01-2007-000-10161-0). This work was also supported in part by Brain Korea 21(BK21) program.

References

1. Bernfield, M.; Gotte, M.; Park, P. W.; Reizes, O.; Fitzgerald, M. L.; Lincecum, J.; Zako, M. *Annu. Rev. Biochem.* **1999**, *68*, 729-777.
2. Carey, D. J. *Biochem. J.* **1997**, *327*, 1-16.
3. Couchman, J. R.; Woods, A. *J. Cell Sci.* **1999**, *112*, 3415-3420.
4. Orosco, A.; Fromigué, O.; Bazille, C.; Entz-Werle, N.; Levillain, P.; Marie, P. J.; Modrowski, D. *Cancer Res.* **2007**, *67*, 3708-3715.
5. Munesue, S.; Yoshitomi, Y.; Kusano, Y.; Koyama, Y.; Nishiyama, A.; Nakanishi, H.; Miyazaki, K.; Ishimaru, T.; Miyaura, S.; Okayama, M.; Oguri, K. *J. Biol. Chem.* **2007**, *282*, 28164-28174.
6. Grootjans, J. J.; Zimmerman, P.; Reekmans, G.; Smets, A.; Degeest, G.; Durr, J.; David, G. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 13683-13688.
7. Hsueh, Y. P.; Yang, F. C.; Kharazia, V.; Naisbitt, S.; Cohen, A. R.; Weinberg, R. J.; Sheng, M. *J. Cell Biol.* **1998**, *142*, 139-151.
8. Gao, Y.; Li, M.; Chen, W.; Simons, M. *J. Cell Physiol.* **2000**, *184*, 373-379.
9. Grootjans, J. J.; Reekmans, G.; Ceulemans, H.; David, G. *J. Biol. Chem.* **2000**, *275*, 19933-19941.
10. Zimmermann, P.; Tomatis, D.; Rosas, M.; Grootjans, J.; Leenarts, I.; Degeest, G.; Reekmans, G.; Coomans, C.; David, G. *Mol. Biol. Cell* **2001**, *12*, 339-350.
11. Lee, D.; Oh, E. S.; Woods, A.; Couchman, J. R.; Lee, W. *J. Biol. Chem.* **1998**, *273*, 13022-13029.
12. Oh, E. S.; Woods, A.; Couchman, J. R. *J. Biol. Chem.* **1997**, *272*, 11805-11811.
13. Shin, J.; Lee, W.; Lee, D.; Koo, B. K.; Lim, Y.; Woods, A.; Couchman, J. R.; Oh, E. S. *Biochemistry* **2001**, *40*, 8471-8478.
14. Keum, E.; Kim, Y.; Kim, J.; Kwon, S.; Lim, Y.; Han, I.; Oh, E. S. *Biochem. J.* **2004**, *378*, 1007-1014.
15. Couchman, J. R. *Nat. Rev. Mol. Cell Biol.* **2003**, *4*, 926-937.
16. Koo, B. K.; Jung, Y. S.; Shin, J.; Han, I.; Mortier, E.; Zimmermann, P.; Whiteford, J. R.; Couchman, J. R.; Oh, E. S.; Lee, W. *J. Mol. Biol.* **2006**, *355*, 651-663.
17. Essner, J. J.; Chen, E.; Ekker, S. C. *Int. J. Biochem. Cell Biol.* **2006**, *38*, 152-156.
18. Moon, W. J.; Hwang, D. K.; Park, E. J.; Kim, Y. M.; Chae, Y. K. *Protein Expr. Purif.* **2007**, *51*, 141-146.
19. Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277-293.
20. Chae, Y. K.; Lee, H.; Lee, W. *Bull. Kor. Chem. Soc.* **2007**, *28*, 1549-1552.
21. Chae, Y. K.; Cho, K. S.; Chun, W.; Lee, K. *Protein Pept. Lett.* **2003**, *10*, 369-374.