

Solution Structure of the Cytoplasmic Domain of Syndecan-3 by Two-dimensional NMR Spectroscopy

In Young Yeo, Bonkyung Koo, Eok-soo Oh,[†] Inn-Oc Han,[‡] and Weontae Lee*

Department of Biochemistry and Protein Network Research Center, Yonsei University, Seoul 120-749, Korea

*E-mail: wlee@spin.yonsei.ac.kr

[†]Department of Life Sciences, Division of Life and Pharmaceutical Sciences and the Center for Cell Signaling & Drug Discovery Research, Ewha Womans University, Seoul 120-750, Korea

[‡]Department of Physiology and Biophysics, Inha University, Incheon 402-751, Korea

Received December 17, 2007

Syndecan-3 is a cell-surface heparan sulfate proteoglycan, which performs a variety of functions during cell adhesion process. It is also a coreceptor for growth factor, mediating cell-cell and cell-matrix interaction. Syndecan-3 contains a cytoplasmic domain potentially associated with the cytoskeleton. Syndecan-3 is specifically expressed in neuron cell and has related to neuron cell differentiation and development of actin filament in cell migration. Syndecans each have a unique, central, and variable (V) region in their cytoplasmic domains. And that region of syndecan-3 may modulate the interactions of the conserved C1 regions of the cytoplasmic domains by tyrosine phosphorylation. Cytoplasmic domain of syndecan-3 has been synthesized for NMR structural studies. The solution structure of syndecan-3 cytoplasmic domain has been determined by two-dimensional NMR spectroscopy and simulated-annealing calculation. The cytoplasmic domain of the syndecan proteins has a tendency to form a dimer conformation with a central cavity, however, that of syndecan-3 demonstrated a monomer conformation with a flexible region near C-terminus. The structural information might add knowledge about the structure-function relationships among syndecan proteins.

Key Words : Syndecan-3, Proteoglycan, NMR

Introduction

The syndecans are transmembrane heparan sulfate proteoglycans (HSPGs) that are present on most cell types. HSPGs have been known for some time to regulate a variety of biological processes, ranging from coagulation cascades, growth factor signaling, lipase binding and activity, cell adhesion to ECM and subsequent cytoskeletal organization, to infection of cells with microorganisms.

The syndecan family in vertebrates comprises four members, syndecan-1, syndecan-2 (fibroglycan), syndecan-3 (N-syndecan), and syndecan-4 (amphiglycan, ryudocan)¹ and they have been identified by molecular cloning from mammalian cells. In addition, syndecan proteoglycan has homolog identified in *Drosophila* and *C. elegans*.² All syndecans have a common structure that possess a large extracellular domain with a signal peptide and varying amounts of glycosaminoglycan (GAG) attachment sites, a one transmembrane domain and a very short cytoplasmic domain. They have high sequence homology in the transmembrane and C-terminal domains. Syndecans play an important role in tissue morphogenesis and differentiation by virtue of their ability to bind a number of extracellular adhesive proteins and growth factors. Also, an essential feature of the functional activity of syndecans is their ability to associate with actin filaments.

The cytoplasmic domain of syndecans consists of conserved regions (C1 and C2), with an intervening unique region (termed the variable (V) region). The V regions are

conserved between species but differ between individual syndecans (syndecans 1-4).

All syndecans oligomerize to form higher-order complexes, and oligomerization may underlie their biological roles. The self-interaction of the cytoplasmic domains has only been investigated for syndecan-4 so far. Short synthetic peptides having the sequence of syndecan-4 V region form dimers with a parallel twisted clamp structure.^{4,5} Chemical cross-linking of syndecan-3 core proteins expressed on the surface of transfected cells reveals that most of the core proteins are present in oligomeric form. Self-association of N-syndecan core protein to form stable, noncovalent multimeric complexes requires the presence of the transmembrane domain as well as a short extracellular domain sequence that is conserved in syndecan core protein sequences.⁶

Syndecan 3 is found in neuronal tissues and developing musculoskeletal cells. It is also termed N-syndecan due to its abundant in the nervous system was first cloned in rat from a schwann cell and a rat brain library.^{7,8} Its expression strongly correlates with the differentiation of oligodendrocytes and myelination in the central and peripheral nervous systems.⁹ This syndecan, unlike the others, appears to bind ECM components poorly. Indeed, it may function more in growth factor control of adhesion. Its heparan sulfate chains can bind HB-GAM, which results in axonal extension.^{2,4,5} Binding results in axonal extension, concomitant with an interaction of the C1 region of its cytoplasmic domain with a complex that includes c-src and the syndecan-3 substrate cortactin and tubulin.¹⁰

In the study of syndecan-3V structure, we provide the characteristic of syndecan-3V structure and indicate the position of specific binding of ligand for syndecans. This data might add information about the structure-function relationships and provide a structural background for further study. In this report, we present the solution structure of the syndecan-3V peptide to analyze its local conformation using two dimensional NMR spectroscopy and dynamical simulated-annealing calculations.

Materials and Methods

Peptide synthesis. Syndecan-3V (Ac-Leu-Glu-Glu-Pro-Lys-Gln-Ala-Ser-Val-Thr-Tyr-Gln-Lys-NH₂) corresponding to residues Leu⁴²¹ to Lys⁴³³ of the syndecan-3 cytoplasmic domain was synthesized commercially (Anygen Co., Kwangju, Korea). The synthetic peptide was purified by reverse-phase liquid chromatography using a Shim-Pack C18 column on a Water Delta Prep 4000 system. The peptide was purified using a reverse-phase C18 column with 0.1% trifluoroacetic acid (TFA) in water and developing with a linear gradient of acetonitrile. The purified peptides were finally characterized by combined use of high-performance liquid chromatography (HPLC) and mass spectrometry (MS). The theoretical mass expected for syndecan-3V was 1562.9. The peptide sample for NMR experiments was prepared by dissolving in 100% D₂O and 90% H₂O/10% D₂O. Solution is at pH value of 7.0 with 50 mM potassium phosphate buffer. The final peptide concentration was 2.7 mM in 0.5 mL of buffer solution.

Circular dichroism spectroscopy. CD spectra of 50 μ M syndecan-3V were measured in 50 mM potassium phosphate buffer for designed various pH, temperatures and micelle conditions using a Jasco J-810 spectropolarimeter. CD spectra were recorded from 190 to 250 nm at a scanning rate 50 nm·min⁻¹ with a time constant of 0.5 sec. Each CD spectrum was obtained from an average of 10 accumulations with step resolution of 0.2 nm and bandwidth of 2.0 nm in cells with 0.1 mm path length.

NMR spectroscopy. NMR spectra were acquired at 278 K on Bruker DRX-500 spectrometers equipped with a triple-resonance probe with x, y, z-shielded pulsed-field gradient coil. Two-dimensional (2D) NMR spectra were recorded in phase-sensitive mode using time-proportional phase incrementation (TPPI)¹¹ for quadrature detection in the t₁ domain. The 2D experiments such as total correlation spectroscopy (TOCSY)¹² using a MLEV-17 spin-lock pulse sequence with a mixing time of 69.8 ms and nuclear overhauser enhancement spectroscopy (NOESY)¹³ with mixing times of 200 ms-600 ms were performed. Solvent suppression for TOCSY and NOESY experiments was achieved using a WATERGATE pulse sequence¹⁴ combined with pulsed-field gradient pulse. All NMR spectra were acquired with 2048 complex data points in t₂ and 256 increments in the t₁ dimension, with 64 scan per increment. All NMR data were processed on a Silicon Graphics Indigo II workstation using nmrPipe/nmrDraw or XWIN-NMR (Bruker Instruments, Karlsruhe,

Germany) software and analyzed using the Sparky 3.60 program.¹⁵ The proton chemical shifts were expressed relative to the methyl resonance of the internal sodium 2,2-dimethyl-2-silapentane-5-sulfonic acid (DSS).

Structural restraints and structure calculations. Distance restraints were derived from the NOESY spectra in 90% H₂O/10% D₂O solution. A total of 170 NOE constraints were used for structure calculations. Cross-peak volumes were divided as strong, medium and weak, corresponding to upper bound interproton distance restraints of 2.7, 3.3 and 5.0 Å, respectively. The structures were calculated using the hybrid distance geometry and dynamical simulated annealing (SA) protocol, as previously described,¹⁶⁻¹⁹ using the CNS 1.1 program on a Linux workstation. Final structures were analyzed using PROCHECK²⁰ and displayed using Insight II (Biosym/Molecular Simulations, Inc.) and MOLMOL programs.²¹

Results and Discussion

Circular dichroism spectroscopy. The CD spectrum of the syndecan-3V peptide (sequence LEEPQASVTYQK) at pH 7.0 and 25 °C does not show any standard secondary structures (data not shown).²² However, the structure of the syndecan-3V based on CD data demonstrated as a loop conformation with several turns as previously reported 4V peptide.

NMR resonance assignments. There are one serine, tyrosine and alanine residues in the syndecan-3V peptide sequence. The one alanine residue was identified from connectivities in TOCSY spectra by its characteristic methyl resonance and the lone tyrosine residue was easily assigned by combined use of TOCSY and NOESY spectra. The single unique serine residue was also assigned from the TOCSY spectrum. These preliminary resonance assignments served as starting points for the sequence specific assignment procedure.²³

Complete proton resonance assignment for syndecan-3V was achieved using the standard sequential resonance assignment procedure.²³ Once the individual spin systems had been classified, backbone sequential resonance assignment was easily completed by $d_{\alpha N}(i, i+1)$ NOE connectivities in the 2D-NOESY spectra. The side chain assignment was performed with TOCSY connectivities (Fig. 2A). Figure 1 summarizes the observed sequential and medium range NOE connectivities for the syndecan-3V. Figure 2B shows NH/NH region of NOESY spectrum to demonstrate d_{NN} NOEs. In the NOESY spectra, a number of d_{NN} cross-peaks from the amide noe region were observed but it can not prove the presence of α -helix. Other NOEs confirm that standard regular secondary structure does not exist. Sequential NOE region was assigned in αN region of NOESY spectra. It does not presented here for overlapping. The aliphatic proton resonance region of the two-dimensional NOESY spectrum of syndecan-3V display NOEs assigned with intramolecular NOEs as well as intermolecular NOEs (In this report, the data was not displayed). All NOEs were

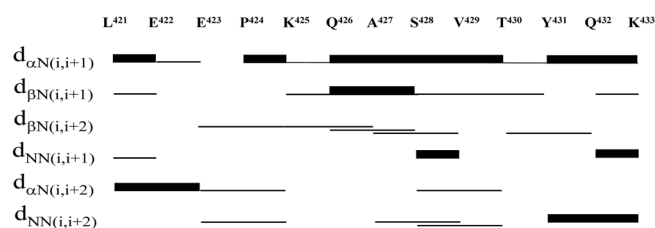


Figure 1. Summary of the NOE connectivities identified in syndecan-3V at 25 °C, pH 7.0. The observed NOE intensities are classified by the thickness of the lines.

observed at mixing times of 600 ms.

The proton chemical shifts for the individual residues in syndecan-3V peptide are in excellent agreement with the corresponding random coil values.²⁴ As expected for a highly flexible peptide, NH-CαH coupling constants for all the residues are of the order of 6.5–8.0 Hz, suggesting retention of significant conformational flexibility. In addition, the sequential NOE intensities displayed a form that has not a secondary structure.

Solution structure of the syndecan-3V. The NMR structures for syndecan-3V were calculated using the experimental constraints derived from both 2D-TOCSY and 2D-NOESY spectra. A total of 100 starting substructures were calculated in the initial simulated annealing stage. After two cycles of the simulated anneal protocol, 95 structures which showed no constraint violations greater than 0.5 Å for distances and 5° for torsion angles were identified. Among 95 structures, the 18 lowest-energy structures (<SA>k) were selected for structural analysis (Table 1). The structures are well defined with a root-mean-square-deviation between backbone atom coordinates of 0.29 Å for residues Glu2-Val 9. The average structure was generated from the geometrical average 18 structure coordinates and was subjected to restrained energy minimization to correct bond length and angle distortions. This average structure exhibited 0.68-Å root-mean-square deviation for backbone atoms with respect to 18 (<SA>k) structures. A best-fit superposition of all final structures and the backbone conformation for the average restrained energy minimized structure (<SA>k_r) are displayed in Figure 3(A). The atomic average root-mean-square deviations of the final structures with respect to the average energy-minimized (REM) structure are shown in Figure 3(B). The Ramachandran plot indicates that Φ, Ψ torsion values of all 18 final NMR structures are distributed properly in energetically acceptable regions.²⁵

It has been reported that the central sequence of syndecan family plays an important role in oligomerization process. The variable domain of syndecan-4 binds and activates PKC through dimerization of the syndecan-4 cytoplasmic domain.²⁶ Serine phosphorylation in the cytoplasmic domain inhibits both dimerization and PKC activation.^{27,28} But preceding study we know that syndecan-3 has a dimerization in core protein. This is proved by glutaraldehyde cross-linking experiment. For dimerization of syndecan 3, trans-membrane region must combine with cytoplasmic domain. Also the four residues in the forward of ecto domain have an

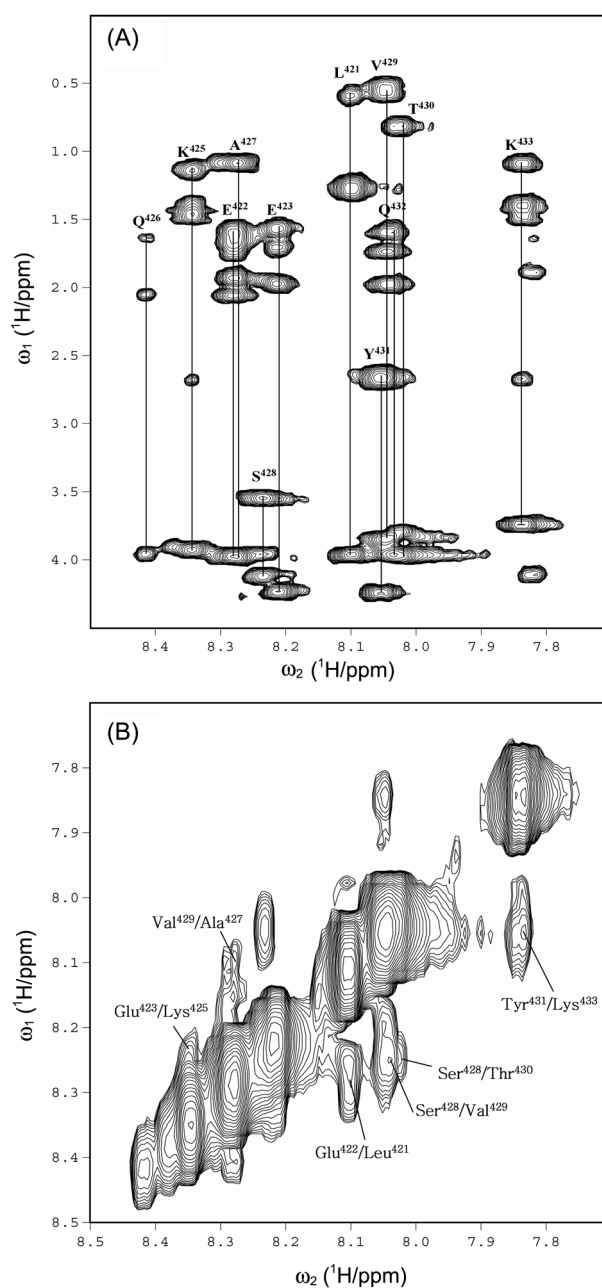


Figure 2. (A) Fingerprint region of 2D-TOCSY spectrum for syndecan-3V with mixing time of 69.8 ms at 25 °C. (B) Amide region of 2D-NOESY spectrum for syndecan-3V with a mixing time of 600 ms in 90% H₂O/10% D₂O solution at 25 °C and pH value of 7.0.

important role in dimerization. For the reasons, we thought that syndecan-3V was monomer condition and calculated structure of the only v region in cytoplasmic domain that makes use to synthesized syndecan-3V in solution.

Structure of the syndecan-3V has a high sequence homology with syndecan-1V. It was only different to serine, threonine, and valine residues in central position of v region. Then, we suggest that the residues play an important role in binding of the signaling molecules. Methyl group of one valine and alanine residue propose to hydrophobic interaction. The serine, tyrosine and threonine residues were

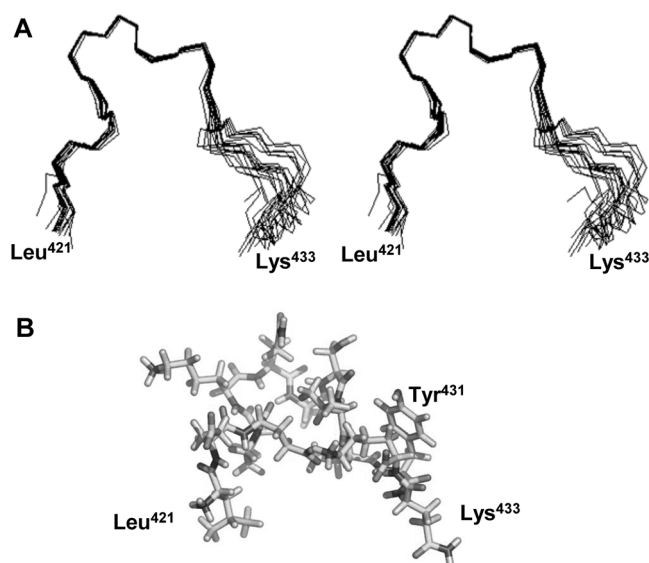
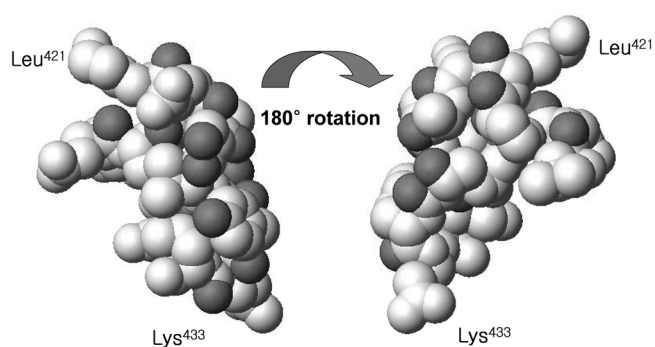
Table 1. Structural statistics for the 18 final simulated-annealing structures (SA) structures of syndecan-3V

	$\langle SA \rangle_k$	$\langle \overline{SA} \rangle_{kr}$
(A) rmsd from experimental distance restraints (Å)		
all (170)	0.0430	0.0418
sequential ($ i - j = 1$) (59)	0.0059	0.0061
inter residue (85)	0.0595	0.0579
intra residue (85)	0.0033	0.0034
(B) rms deviation of structural segment for $\langle SA \rangle_k$ (Gln ² - Val ⁹) (Å)		
backbone	0.29 ± 0.15	
heavy atoms	0.96 ± 0.28	
(C) Energies		
E_{total} (kcal.mol ⁻¹)	39.65	38.13
E_{NOE} (all) (kcal.mol ⁻¹)	15.781	14.862
E_{tor} (kcal.mol ⁻¹)	18.873	18.576
E_{L-J} (kcal.mol ⁻¹) ^a	-6.578	-4.961
(D) Deviations from idealized covalent geometry		
bonds (Å)	0.0029	0.0028
angles (deg)	0.5592	0.5549
impropers (deg)	0.2808	0.2651

^a E_{L-J} is the Lennard-Jones/Van der Waals potential calculated using CHARMM empirical energy function.

located to one surface to optimize molecular interaction. We suppose that their hydroxyl groups might mediate hydrogen bond interactions for protein-protein interaction (Fig. 4). Although syndecans do not themselves possess any cytoplasmic enzymatic activity, they are known to bind potential cytoplasmic signaling molecules. Syndecans have four conserved cytoplasmic tyrosine residues, which can be phosphorylated.²⁹⁻³¹

Tyrosine phosphorylation can be blocked by inhibitor of

**Figure 3.** (A) A stereo view of the backbone superposition of the energy-minimized average structure ($\langle \overline{SA} \rangle_{kr}$) over the family of 18 final simulated annealing structures ($\langle SA \rangle_k$). (B) REM structure of the syndecan-3V. The figure was generated using the pyMOL.**Figure 4.** CPK model of syndecan-3V. The hydroxyl group of serine, threonine and tyrosine and oxygen atoms are presented in grey color.

src family kinase³¹ Syndecan could thus regulate cytoplasmic kinase activities and the activation may depend on ligand induced oligomerization of syndecan core proteins. As the structure calculation, syndecan-3V has the clamp shape that was similar to structure of syndecan-1V, 4 V. On the structure, it was fitting well from residue Glu⁴²² to Val⁴²⁹. However, it was not fitting and has flexible form in the c-terminal region. We are proposed that the flexible site helps the signaling molecules to binding at the syndecan. In this report, we provided the characteristic of its structure and indicate the position of specific binding of ligand for syndecans. This data could serve as an important information about the structure-function relationships, providing structural background for further biological functions of syndecan family proteins.

Acknowledgements. This work was supported by Korea Research Foundation Grant (KRF-2002-CP0327 to E.S.O. & W. Lee).

References

- Carey, D. J. *Biochem J.* **1997**, 327, 1-16.
- Tumova, S.; Woods, A.; Couchman, J. R. *Int J Biochem Cell Biol.* **2000**, 32(3), 269-288.
- Spring, J.; Paine-Saunders, S.; Hynes, R.; Bernfield, M. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, 91, 3334-3338.
- Couchman, J. R.; Chen, L.; Woods, A. *Int. Rev. Cytol.* **2001**, 207, 113-150.
- Woods, A.; Couchman, J. R. *J. Biol. Chem.* **2000**, 275, 24233-24236.
- Asundi, V. K.; Carey, D. J. *J. Biol. Chem.* **1995**, 270, 26404-26410.
- Carey, D. J.; Evans, D. M.; Stahl, R. C.; Asundi, V. K.; Conner, K. J.; Garbes, P.; Cizmeci-Smith, G. *J. Cell. Biol.* **1992**, 117(1), 191-201.
- Carey, D. J.; Conner, K.; Asundi, V. K.; O'Mahony, D. J.; Stahl, R. C.; Showalter, L.; Cizmeci-Smith, G.; Hartman, J.; Rothblum, L. I. *J. Biol. Chem.* **1997**, 272(5), 2873-2879.
- Carey, D. J. *Perspect. Dev. Neurobiol.* **1996**, 3, 331-346.
- Kinnunen, T. et al. *J. Biol. Chem.* **1998**, 273, 10702-10708.
- Marion, D.; Wuthrich, K. *Biochem. Biophys. Res. Commun.* **1983**, 113, 967.
- Davis, D. G.; Bax, A. *J. Am. Chem. Soc.* **1985**, 107, 2820.
- Jeener, J.; Meier, B. H.; Bachman, P.; Ernst, R. R. *J. Chem. Phys.* **1979**, 71, 4546.

14. Piotto, M.; Saudek, V.; Sklenar, V. *J. Biomol. NMR* **1992**, *2*, 661.
 15. Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277.
 16. Driscoll, P. C.; Gronenborn, A. M.; Beress, L.; Clore, G. M. *Biochemistry* **1989**, *28*, 2188.
 17. Nilges, M.; Clore, G. M.; Gronenborn, A. M. *FEBS Lett.* **1988**, *229*, 317.
 18. Nilges, M.; Clore, G. M.; Gronenborn, A. M. *FEBS Lett.* **1988**, *239*, 129.
 19. Nilges, M.; Gronenborn, A. M.; Brünger, A. T.; Clore, G. M. *Protein Eng.* **1988**, *2*, 27.
 20. Lakovski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. *J. Appl. Crystallogr.* **1993**, *26*, 283.
 21. Koradi, R.; Billeter, M.; Wüthrich, K. *J. Mol. Graphics* **1996**, *14*, 51.
 22. Yang, J. T.; Wu, C.-S. C.; Martinez, H. M. *Methods Enzymol.* **1996**, *130*, 208-269.
 23. Wüthrich, K. *NMR of Proteins and Nucleic Acids*; Wiley: New York, 1986.
 24. Wishart, D. S.; Sykes, B. D.; Richards, F. M. *J. Mol. Biol.* **1991**, *222*, 311-333.
 25. Morris, A. L.; MacArthur, M. W.; Hutchison, E. G.; Thornton, J. M. *Proteins* **1992**, *12*, 345-364.
 26. Oh, E. S.; Woods, A.; Couchman, J. R. *J. Biol. Chem.* **1997**, *272*(18), 11805-11811.
 27. Horowitz, A.; Simons, M. *J. Biol. Chem.* **1998**, *273*(40), 25548-25551.
 28. Horowitz, A.; Simons, M. *J. Biol. Chem.* **1998**, *273*(18), 10914-10918.
 29. Reiland, J.; Ott, V. L.; Lebakken, C. S.; Yeaman, C.; McCarthy, J.; Rappaegeer, A. C. *Biochem. J.* **1996**, *319*(Pt 1), 39-47.
 30. Asundi, V. K.; Carey, D. J. *Biochem. Biophys. Res. Commun.* **1997**, *240*(2), 502-506.
 31. Ott, V. L.; Rappaegeer, A. C. *J. Biol. Chem.* **1998**, *273*(52), 35291-35298.
-