Membrane-binding mechanism of the EEA1 FYVE domain revealed by multi-scale molecular dynamics simulations

1

Andreas Haahr Larsen*, Lilya Tata*, Laura John & Mark S.P. Sansom**

*contributed equally

Department of Biochemistry, University of Oxford, Oxford, United Kingdom, OX1 3QU

**contact e-mail: mark.sansom@bioch.ox.ac.uk

For submission to: *PLoS Comp Biol.*

Abstract

Early Endosomal Antigen 1 (EEA1) is a key protein in endosomal trafficking and is implicated in both autoimmune and neurological diseases. The C-terminal FYVE domain of EEA1 binds endosomal membranes, which contain phosphatidylinositol-3-phosphate (PI(3)P). Although it is known that FYVE binds PI(3)P specifically, it has not previously been described of how FYVE attaches and binds to endosomal membranes. In this study, we employed both coarse-grained (CG) and atomistic (AT) molecular dynamics (MD) simulations to determine how FYVE binds to PI(3)P-containing membranes. CG-MD showed that the dominant membrane binding mode resembles the crystal structure of EEA1 FYVE domain in complex with inositol-1,3-diphospate (PDB ID 1JOC). FYVE, which is a homodimer, binds the membrane via a hinge mechanism, where the C-terminus of one monomer first attaches to the membrane, followed by the C-terminus of the other monomer. The total binding energy is 70 kJ/mol, of which 50-60 kJ/mol stems from specific PI(3)P-interactions. By AT-MD, we could partition the binding mode into two types: (i) adhesion by electrostatic FYVE-PI(3)P interaction, and (ii) insertion of amphipathic loops. The AT simulations also demonstrated flexibility within the FYVE homodimer between the C-terminal heads and coiled-coil stem, allowing binding via a mechanism resembling that of a suction cup connected to a locally rigid stem via a flexible hinge.

Keywords: EEA1 FYVE domain, molecular dynamics simulations, coarse-grained, atomistic, multi-scale, PI(3)P, zinc finger, endosome, endocytosis, endosomal trafficking

Introduction

Early Endosomal Antigen 1 (EEA1) acts as a connecter between the early endosomal membrane and Rab GTPase effector proteins [1,2]. Thereby, it plays a key role in endosomal trafficking. Autoantibodies to EEA1 have been implicated in several autoimmune diseases, such as subacute cutaneous systemic lupus erythematosus, polyarthritis and rheumatoid arthritis [3,4]. EEA1 is also implicated in neurological diseases [5] and therefore provides a potential drug target.

EEA1 is homodimeric, composed of a FYVE (Fab1, YOTB, Vac1, EEA1) domain located at the C-terminus of a ~200 nm long coiled-coil. At the N-terminus of the coiled coil, there is a Rab5-binding domain [6–8] (Figure 1). The FYVE domain interacts with the cytosolic face of early endosomal membranes, which contain phosphatidylinositol-3-phosphate (PI(3)P) [9]. Dimeric FYVE domain binds PI(3)P [9–11]. The binding site was identified determined by X-ray crystallography of $Ins(1,3)P_2$ bound to dimeric FYVE (PDB: 1JOC) [7] and by NMR of monomeric FYVE in a complex with $Ins(1,3)P_2$ (PDB: 1HYI) [6,12,13].

The interaction of a FYVE domain with a membrane containing PI(3)P molecules has previously been investigated with molecular dynamics (MD) simulations [14], demonstrating the short timescale (30 ns) stability of monomeric domain docked onto a membrane in an orientation consistent with experimental results [6,7,12]. Improvements in MD simulations for biomembranes, both at the atomistic [15] and coarse-grained (CG) level [16–18] enabled us to simulate binding of a much larger dimeric construct of FYVE in 1.5 μs coarse-grain simulations followed by a 500 ns atomistic simulation to refine the bound pose. In these simulations the dimeric FYVE domain was positioned in different orientations at least 10 nm above the membrane, thus avoiding bias in sampling of possible binding modes of the protein. Analysis of these simulations provides a detailed and dynamic model of how dimeric FYVE binds to target membranes.

Results and Discussion

CG-MD Simulations of FYVE Binding to Membranes

Coarse-grained MD simulations were performed to investigate the binding of the dimeric EEA1 FYVE domain to a simple model of an endosomal membrane composed of PC (95%) plus 5% PI(3)P which is known to bind FYVE. A truncated version of the crystal structure of EEA1 FYVE (PDB 1JOC) was used, with a shorter coiled-coil stalk (see Methods for details). Fifteen repeats of the

simulation, each of duration 1.5 µs, were performed with different randomly selected starting orientations of the FYVE with respect to the membrane, and with a minimum distance of 10 nm between protein and membrane. The minimum distance between the protein and the bilayer was monitored as a function of time for each simulation (Figure 2AB). In 8/15 repeats, the FYVE domain bound to the membrane within the 1.5 µs of the simulations. In the remaining 7/15 repeats, the FYVE domain did not approach close enough to the bilayer to experience an electrostatic interaction, i.e. the minimum distance remained larger than the cutoff distance (see Methods), and so no binding event occurred. For 6 of the binding events, the FYVE domain bound in a 'canonical' manner, with the C-termini of each chain bound to the membrane (Figure 2BC). Intriguingly, we observed a two-stage mechanism of binding, where FYVE first attached to the membrane via the C-terminus of one chain, then made a rotation around the point-of-contact such that the second C-terminus bound as well (Figure 2C). Once this canonical binding mode was established, the FYVE dimer remained in this pose for the duration of the simulation. In control simulations using a PC bilayer with no PI(3)P present, none of the simulations resulted in a canonical interaction. Of the 15 PC-only repeats, 8 showed non-canonical binding, of which 3 subsequently dissociated from the membrane (SI Figure S1).

Returning to the FYVE dimers bound to the PI(3)P membrane in a canonical pose, these often revealed between 5 to 10 PI(3)P molecules interacting with the FYVE dimer in this binding pose, in contrast to two $Ins(1,3)P_2$ molecules bound in the dimeric crystal structure (PDB: 1JOC; one bound per monomer). The NMR structure of the FYVE monomer also had just one diC_4-PI(3)P molecule bound (PDB: 1HYI) [13]. We suggest that multiple PI(3)P lipids are likely to interact with FYVE *in vivo*, but only one binding site / monomer has sufficiently high affinity to retain bound PI(3)P in the absence of a membrane. This is in line with previous studies of other PIP-binding domains, e.g. PH [19] and C2 [20], which interact with multiple PIPs and other anionic lipids.

Strength of FYVE/PI(3)P Interactions

To investigate the strength of the predicted FYVE/PI(3)P interactions, we undertook both potential of mean force (PMF) and free energy perturbation (FEP) calculations to estimate the free energy of PI(3)P interactions with the protein as estimated by the CG forcefield [21,22]. For the energy calculations, a representative frame of the bound structure in a canonical pose, with six PI(3)P molecules interacting with the dimeric FYVE domain was used. A PMF, using the bilayer to protein centre of mass distance as a reaction coordinate, was estimated by umbrella sampling (Figure 2D).

The depth of the energy well, corresponding to the membrane bound FYVE dimer interacting with six PI(3)P molecules, was -67 ± 2 kJ/mol. The free energy associated with the headgroup of each bound PI(3)P molecule was calculated via FEP calculations [21], whereby each of the six bound PI(3)P headgroups were changed into PC headgroups (see SI Fig. S2AB). The free energy of this perturbation varied between 6 and 12 kJ/mol for each PI(3)P molecule converted to PC, summing to a total perturbation energy of +55 ± 2 kJ/mol (Table S1). Thus, the InsP headgroups comprise about 80% of the total membrane binding free energy as estimated by the PMF. Furthermore, the FEP calculations suggest a twofold range in interaction free energies for the different PI(3)P sites.

The dissociation constant (K_d) of a EEA1 FYVE monomer with a PI(3)P-containing membrane has been estimated experimentally to ~50 nM [23] at room temperature, corresponding to a binding free energy of -42 kJ/mol. If we make the simple assumption that the binding energies are additive, when going from monomer to dimer, this corresponds to -84 kJ/mol for the dimer, i.e. slightly greater than the PMF from the simulations. A later study has shown that K_d is pH-dependent: ~50 nM (42 kJ/mol) at pH 6.0; ~500 nM (36 kJ/mol) at pH 7.4; and ~1200 nM at pH 8.0 (34 kJ/mol) [24]. Again, these numbers are for monomeric FYVE. The protonation states in our simulations are designed to match the conditions at neutral pH. We note that the PMF well depth of -67± 2 kJ/mol is consistent with two times the value reported for the monomer at pH 7.4 (-72 kJ/mol) [24].

We further analysed the binding sites using the program PyLipID [25], which can identify protein-lipid binding sites based on residence time in CG-MD simulations. The three main binding sites for PI(3)P on the dimeric FYVE domain were found to be (Figure S2): (i) residues 1348-1352 (KWAED), where residues 1349-1352 (WAED) form a conserved WxxD motif, and where a W1349A mutation has been shown to substantially reduce PI(3)P binding [23]; (ii) a second site made up primarily of residues 1397-1400 (KPVR), where R1400 is part of a conserved RVC motif of FYVE, mutation of which (R1400A) leads to 6-fold decrease in PI(3)P binding affinity [23]; and (iii) a third site which includes residues 1368-1373 (TVRRH) at the core of a conserved motif spanning residues 1371 to 1376 (RRHHCR), and mutations of which lead to impaired PI(3)P affinity [23]. The first and second binding sites are both placed centrally (SI Figure S2CD), just below the stem of the FYVE domain, similar to the binding sites from the crystal structure complexed with Ins(1,3)P2 (PDB ID 1JOC) [7] and the NMR structure, complexed with diC4-PI(3)P (PDB ID 1HYI) [13].

Atomistic MD Simulations

Using a 0.5 µs atomistic MD simulation starting from a representative pose of the membrane bound protein from one of the CG simulations (Figure 3), we analysed the mode of interaction of the FYVE

dimer with the membrane in more detail. The AT-MD simulation showed that the two "heads" (i.e. C terminal domains) of the FYVE dimer did not remain equally close to the membrane throughout the duration of the simulation (Figure 4A). The COM distance for each head was monitored and head 1 (C terminus of the first monomer) was systematically further away from the membrane than head 2 (Figure 4A). Intriguingly, head 1, which was further away, formed more H-bonds to PI(3)P lipids than head 2 (Figure 4B). Hence, we suggest two modes of interaction for the FYVE heads: (i) mode 1, where the head is relatively further from the membrane, forming multiple H-bonds to PIP(3)P lipids and (ii) mode 2, where the head is closer to the membrane, with a loop inserted into the bilayer and fewer H-bonds with PIP(3)P (Figure 3A-C). In the simulation, head one bind mostly via binding mode 1, and head two binds more via binding mode 2. However, as FYVE is symmetric, we expect, with sufficient sampling, that the heads will shift between the modes. Moreover, the heads can be in the same mode, as observed at ~30 ns and ~240 ns (Figure 4AB).

In the coarse-grain simulations, secondary and tertiary structure was restrained with an elastic network (see Methods). This was not the case in the AT simulations, and so possible large scale conformational changes could thus be observed. From analysis of RMSDs (SI Figure S3) we noted that the heads and likewise the coiled coil (stalk) behaved individually as semi-rigid bodies, as reflected in Cα RMSDs of 0.24 ± 0.02 nm for the heads and 0.30 ± 0.03 nm for the stalk (Figure S3). However, there was substantial hinge-bending flexibility between heads and stalk.

The hinge-bending motion could be observed via visual inspection (Figure 3). By calculating the angle between the stalk axis and the plane of the bilayer (Figure 4C) we can see the local consequence of these fluctuations. As the coiled coil in the full-length EEA1 protein is more than 200 nm long, and will be locally rigid, a tilt of up to ~50° relative to the bilayer normal would result in a later displacement of the Rab5 binding domain at the opposite end of nearly 2000 nm. This could be even greater if one allows for bending of the coiled coil, and indeed recent evidence [26] suggests that binding of Rab5:GTP may increase the flexibility of the EEA1 coiled coil. Thus, the flexibility of the link of early endosome PI(3)P-bound FYVE dimer to the coiled coil would provide a search area of at least 0.12 μm^2 at the other end of the EAA1 molecule for a Rab5 molecule attached to a second early endosome (Figure 5). This emphasizes the importance of considering molecular flexibility when relating static domain structures to dynamic tethering within a cell. We conjecture that mutations in this hinge region could modulate the flexibility, and thus alter the ability of early endosome PI(3)P-bound EEA1 to capture adjacent endosomes.

Conclusions

We have investigated the membrane-binding model for FYVE proposed by e.g. Dumas et al. [7]. Using MD simulations, we have verified that the dominant binding mode is with the C-terminus of each chain of the homodimeric FYVE domain facing the membrane. We have described different modes of binding for the FYVE domain, either driven by electrostatic interactions with PI(3)P headgroups, or driven by insertion of loops in the hydrophobic bilayer. We have shown that FYVE binds multiple PI(3)Ps, albeit more weakly to some of them. Moreover, we have shown that there is hinge-bending flexibility in the linker between the PI(3)P-bound FYVE domain heads and the coiled-coil stalk, thus facilitating search and capture of adjacent early endosomal membranes. A future challenge will be to build a (CG) model of a complete EAA1 molecule linking two vesicles to more fully explore the role of hinge-bending and coiled-coil flexibility in vesicle tethering. This would be a large simulation but not impossible (see e.g. [27,28]).

Methods

Coarse-grained simulations

The crystal structure of the EEA1 FYVE domain in complex with 1,3-diphosphate (PDB ID 1JOC) was used to generate the initial structure for the simulations (Figure 1). This crystal structure is positively charged at its N-terminal region (i.e. the start of the coiled coil stalk, residues 1289-1322 of EEA1), which resulted in artefactual binding of the N-terminus to the anionic lipid bilayer membrane. Therefore, we used a truncated version of the protein with a shorter coiled-coil stalk consisting of residues 1323-1411.

The truncated construct was coarse-grained using the Martinize script [29] to a construct compatible with the Martini2.1 force field [30]. Zinc ions and bound 1,3-diphosphate were removed during the coarse-graining of the initial structure. An elastic network [31] was applied to conserve secondary and tertiary structure, which maintained the CG protein fold in the absence of explicit zinc ions.

A simulation box with the coarse-grained protein, water, ions and a lipid membrane was constructed using the insane script [32]. The initial box size was 16x16x32 nm^3 and had periodic boundary conditions. 10% anti-freeze water were added to avoid frozen water [30], and 0.15 M NaCl ions was included. The protein was placed with a minimum distance of 10 nm from the membrane to avoid any initial binding bias (Figure 2). The membrane was symmetric and had a lipid composition of 95% 1-palmitoyl-2-oloyl-sn-glyecero-3-phosphocholine (POPC) [32] and 5% 1-palmitoyl-2-oloyl-sn-glyecero-3-phosphatidylinositol-3-phosphate (POP(3)P) [18], giving a lipid composition of 418:22 (POPC:POP(3)P) in each leaflet.

Simulations were performed in GROMACS 2018 [33]. The protein was first rotated randomly, then minimised with the steepest decent method, equilibrated for 300 ps and finally simulated for 1500 ns. Both equilibration and production were in the NPT ensemble with a semi-isotropic Parinello-Rahman barostat [34] with 12 ps time constant. Lateral and normal compressibility of $3 \cdot 10^{-4}$ bar^{-1} was used to keep pressure at 1.0 bar, and a v-rescale thermostat with a 1 ps time constant kept the temperature at 323 K. The integration timestep was 30 fs. Long range electrostatics were calculated using particle mash Ewald (PME) [35] with a cut-off of 1.1 nm. Non bonded van der Waals bonds likewise had a 1.1 nm cut-off. The initial position of the protein had a minimum protein-lipid distance of 10 nm, i.e. multiple times the cutoff-distances. 15 repeats were made, each with individual rotation, minimisation and equilibration. The minimum protein-membrane distance was monitored using built-in GROMACS tool *gmx mindist* [36]. Lipid contacts were analysed using PyLipID [25]. Contacts were measured from their residence time [37].

Potential of mean force estimation

A frame from the coarse-grained simulations was selected for calculations of free energies, corresponding the dominant membrane binding mode and with 6 bound PIP(3)P molecules. For the PMF calculations, the protein was pulled away from the membrane using steered MD and a pulling force of 1000 kJ/mol/nm, along a reaction coordinate perpendicular to the membrane plane, employing the GROMACS pull code [38]. The reaction coordinate was the centre-of-mass distance between the lipid bilayer and the protein. Position restraints were applied to PI(3)P heads to prevent these from being pulled out of the membrane. Snapshot were retrieved every 0.5 nm along the reaction coordinate. These were used for umbrella sampling, consisting of 1000 ns simulations with an applied position restraint of 1000 kJ/mol/nm on the protein. From the umbrella sampling, a 1D energy profile was calculated using the built-in weighted histogram method *gmx wham* [39], employing 200 rounds of bootstrapping.

Free energy perturbation

In six independent simulations, one bound PI(3)P headgroup was alchemically transformed into a PC headgroup, by changing the CG bead properties. First the Coulombic interactions were perturbed in 20 steps of 100 ps, and in 20 subsequent steps, the Lennard Jones properties were perturbed. One simulation was also run with free PI(3)P being transformed into PC, and free energy was calculated as the difference between the perturbation energy of bound and free PI(3)P [21], using the alchemical analysis package [40].

Atomistic simulations

8

A frame from the coarse-grained simulations corresponding the dominant membrane binding mode and with 6 bound PIP(3)P molecules (as used for energetic calculations – see above) was converted using CG2AT [41] to atomistic resolution and prepared for simulations in the CHARMM36m force field with TIP3P water. Zn^{2+} ions were inserted in the zinc-finger sites in accordance with the crystal structure (PDB: 1JOC). The four cysteines in each zinc fingers were converted to deprotonated cysteine (topology CYM). Zinc ions were restraint with 1000 kJ/mol/nm^2 bonds to the neighbouring cysteines using the pull code in GROMACS. The system was neutralised by converting four Na^+ ions to Cl^-. The system was equilibrated for 10 ns in the NVT ensemble and for 20 ns in the NPT ensemble, then simulated for 500 ns. Both equilibration and production run were kept at 300 K using a v-rescale thermostat with a 0.1 ps time constant. NPT equilibration and production run were kept at 1.0 bar using a semi-isotropic Parrinello-Rahmen barostat [34] with a 5 ps time constant and $4.5 \cdot 10^{-4}$ bar^{-1} isothermal compressibility. LINCS constraints were applied on the H-bonds, so the simulation could be run with 2 fs time steps with a leap-frog integrator. Electrostatics were included using particle mesh Ewald (PME) [35] for long range electrostatics, and a cut-off of 1.2 nm. The van der Waals cut-off for non-bonded interactions was also 1.2 nm. H-bonds were analysed using the *gmx hbonds* tool in GROMACS. COM distances were calculated using *gmx distance*, with the distance defined between the lipids and the protein as separate groups.

Visualisation of CG simulations used VMD [42], and of atomistic simulations used PyMOL (The PyMOL Molecular Graphics System, Version 1.2, Schrödinger, LLC).

References

[1]   I.G. Mills, A.T. Jones, M.J. Clague, Involvement of the endosomal autoantigen EEA1 in homotypic fusion of early endosomes, Curr. Biol. 8 (1998) 881–884. https://doi.org/10.1016/S0960-9822(07)00351-X.

[2]   A. Mishra, S. Eathiraj, S. Corvera, D.G. Lambright, Structural basis for Rab GTPase recognition and endosome tethering by the C2H2 zinc finger of Early Endosomal Autoantigen 1 (EEA1), Proc. Natl. Acad. Sci. U. S. A. 107 (2010) 10866–10871. https://doi.org/10.1073/pnas.1000843107.

[3]   R.L. Waite, J.W. Sentry, H. Stenmark, B.H. Toh, Autoantibodies to a novel early endosome antigen 1, Clin. Immunol. Immunopathol. 86 (1998) 81–87. https://doi.org/10.1006/clin.1997.4455.

[4]   S. Selak, R.C. Woodman, M.J. Fritzler, Autoantibodies to early endosome antigen (EEA1) produce a staining pattern resembling cytoplasmic anti-neutrophil cytoplasmic antibodies (C-ANCA), Clin. Exp. Immunol. 122 (2000) 493–498. https://doi.org/10.1046/j.1365-2249.2000.01390.x.

[5]   X. Xu, L. Pozzo-Miller, EEA1 restores homeostatic synaptic plasticity in hippocampal neurons from Rett syndrome mice, J. Physiol. 595 (2017) 5699–5712. https://doi.org/10.1113/JP274450.

[6]   T.G. Kutateladze, Mechanistic similarities in docking of the FYVE and PX domains to phosphatidylinositol 3-phosphate containing membranes Tatiana, Prog. Lipid Res. 46 (2007) 315–327.

[7]   J.J. Dumas, E. Merithew, E. Sudharshan, D. Rajamani, S. Hayes, D. Lawe, S. Corvera, D.G. Lambright, Multivalent endosome targeting by homodimeric EEA1, Mol. Cell. 8 (2001) 947–958. https://doi.org/10.1016/S1097-2765(01)00385-9.

[8]   J. Callaghan, A. Simonsen, J.M. Gaullier, B.H. Toh, H. Stenmark, The endosome fusion regulator early-endosomal autoantigen 1 (EEA1) is a dimer, Biochem. J. 338 (1999) 539–543. https://doi.org/10.1042/0264-6021:3380539.

[9]   J.M. Gaullier, A. Simonsen, A. D'Arrigo, B. Bremnes, H. Stenmark, R. Aasland, FYVE fingers bind PtdIns(3)P [5], Nature. 394 (1998) 432–433. https://doi.org/10.1038/28767.

[10]   C.G. Burd, S.D. Emr, Phosphatidylinositol(3)-phosphate signaling mediated by specific binding to RING FYVE domains, Mol. Cell. 2 (1998) 157–162. https://doi.org/10.1016/S1097-2765(00)80125-2.

[11]   V. Patki, D.C. Lawe, S. Corvera, J. V. Virbasius, Anil Chawla, A functional PtdIns(3)P-binding motif, Nature. 394 (1998) 433–434.

[12]   R. Brunecky, S. Lee, P.W. Rzepecki, M. Overduin, G.D. Prestwich, A.G. Kutateladze, T.G. Kutateladze, Investigation of the binding geometry of a peripheral membrane protein, Biochemistry. 44 (2005) 16064–16071. https://doi.org/10.1021/bi051127+.

[13]   T. Kutateladze, M. Overduin, Structural mechanism of endosome docking by the FYVE domain, Science (80-.). 291 (2001) 1793–1796. https://doi.org/10.1126/science.291.5509.1793.

[14]   E. Psachoulia, M.S.P. Sansom, PX- and FYVE-mediated interactions with membranes: Simulation studies, Biochemistry. 48 (2009) 5090–5095. https://doi.org/10.1021/bi900435m.

[15]   J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B.L. de Groot, H. Grubmüller, A.D. MacKerell, CHARMM36m: An Improved Force Field for Folded and Intrinsically Disordered Proteins, Nat. Methods. 14 (2017) 71–73. https://doi.org/10.1038/nmeth.4067.

[16]   D.H. De Jong, G. Singh, W.F.D. Bennett, C. Arnarez, T.A. Wassenaar, L. V. Schäfer, X. Periole, D.P. Tieleman, S.J. Marrink, Improved parameters for the martini coarse-grained protein

force field, J. Chem. Theory Comput. 9 (2013) 687–697. https://doi.org/10.1021/ct300646g.

[17]   M. Bulacu, N. Goga, W. Zhao, G. Rossi, L. Monticelli, X. Periole, D.P. Tieleman, S.J. Marrink, Improved angle potentials for coarse-grained molecular dynamics simulations, J. Chem. Theory Comput. 9 (2013) 3282–3292. https://doi.org/10.1021/ct400219n.

[18]   C.A. López, Z. Sovova, F.J. Van Eerden, A.H. De Vries, S.J. Marrink, Martini force field parameters for glycolipids, J. Chem. Theory Comput. 9 (2013) 1694–1708. https://doi.org/10.1021/ct3009655.

[19]   E. Yamamoto, A.C. Kalli, K. Yasuoka, M.S.P. Sansom, Interactions of Pleckstrin Homology Domains with Membranes: Adding Back the Bilayer via High-Throughput Molecular Dynamics, Structure. 24 (2016) 1421–1431. https://doi.org/10.1016/j.str.2016.06.002.

[20]   A.H. Larsen, M.S.P. Sansom, Binding of Ca2+-Independent C2 Domains to Lipid Membranes: a Multi-Scale Molecular Dynamics Study Andreas, BioRxiv. (2020) 2020.10.30.361964. https://doi.org/https://doi.org/10.1101/2020.10.30.361964.

[21]   R.A. Corey, O.N. Vickery, M.S.P. Sansom, P.J. Stansfeld, Insights into Membrane Protein-Lipid Interactions from Free Energy Calculations, J. Chem. Theory Comput. 15 (2019) 5727–5736. https://doi.org/10.1021/acs.jctc.9b00548.

[22]   R.A. Corey, P.J. Stansfeld, M.S.P. Sansom, The energetics of protein-lipid interactions as viewed by molecular simulations, Biochem. Soc. Trans. 48 (2020) 25–37. https://doi.org/10.1042/BST20190149.

[23]   J.M. Gaullier, E. Ronning, D.J. Gillooly, H. Stenmark, Interaction of the EEA1 FYVE finger with phosphatidylinositol 3-phosphate and early endosomes. Role of conserved residues, J. Biol. Chem. 275 (2000) 24595–24600. https://doi.org/10.1074/jbc.M906554199.

[24]   J. He, M. Vora, R.M. Haney, G.S. Filonov, C.A. Musselman, C.G. Burd, A.G. Kutateladze, V. V. Verkhusha, R. V. Stahelin, T.G. Kutateladze, Membrane insertion of the FYVE domain is modulated by pH, Proteins Struct. Funct. Bioinforma. 76 (2009) 852–860. https://doi.org/10.1002/prot.22392.

[25]   W. Song, A. Duncan, R. Corey, B. Ansell, PyLipID, (2020). https://github.com/wlsong/PyLipID.

[26]   D.H. Murray, M. Jahnel, J. Lauer, M.J. Avellaneda, N. Brouilly, A. Cezanne, H. Morales-Navarrete, E.D. Perini, C. Ferguson, A.N. Lupas, Y. Kalaidzidis, R.G. Parton, S.W. Grill, M. Zerial, An endosomal tether undergoes an entropic collapse to bring vesicles together, Nature. 537 (2016) 107–111. https://doi.org/10.1038/nature19326.

[27]   J.D. Durrant, S.E. Kochanek, L. Casalino, P.U. Ieong, A.C. Dommer, R.E. Amaro, Mesoscale All-Atom Influenza Virus Simulations Suggest New Substrate Binding Mechanism, ACS Cent. Sci. 6 (2020) 189–196. https://doi.org/10.1021/acscentsci.9b01071.

[28]   A. Yu, A.J. Pak, P. He, V. Monje-Galvan, L. Casalino, Z. Gaieb, A.C. Dommer, R.E. Amaro, G.A. Voth, A Multiscale Coarse-Grained Model of the SARS-CoV-2 Virion, Biophys. J. 120 (2021) 1–8. https://doi.org/10.1016/j.bpj.2020.10.048.

[29]   D.H. De Jong, S. Baoukina, H.I. Ingólfsson, S.J. Marrink, Martini straight: Boosting performance using a shorter cutoff and GPUs, Comput. Phys. Commun. 199 (2016) 1–7. https://doi.org/10.1016/j.cpc.2015.09.014.

[30]   S.J. Marrink, H.J. Risselada, S. Yefimov, D.P. Tieleman, A.H. de Vries, The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations, J. Phys. Chem. B. 111 (2007) 7812–7824.

[31]   X. Periole, M. Cavalli, S.-J. Marrink, M.A. Ceruso, Combining an Elastic Network With a Coarse-Grained Molecular Force Field : Structure, Dynamics, and Intermolecular Recognition, J. Chem. Theory Comput. 5 (2009) 2531–2543.

[32]   T.A. Wassenaar, H.I. Ingólfsson, R.A. Böckmann, D.P. Tieleman, S.J. Marrink, Computational lipidomics with insane: A versatile tool for generating custom membranes for molecular

simulations, J. Chem. Theory Comput. 11 (2015) 2144–2155. https://doi.org/10.1021/acs.jctc.5b00209.

[33]   M.J. Abraham, T. Murtola, R. Schulz, S.S. Páll, J.C. Smith, B. Hess, E. Lindahl, E. Lindah, Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX. 1–2 (2015) 19–25. https://doi.org/10.1016/j.softx.2015.06.001.

[34]   M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method, J. Appl. Phys. 52 (1981) 7182–7190. https://doi.org/10.1063/1.328693.

[35]   T. Darden, D. York, L. Pedersen, Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems, J. Chem. Phys. 98 (1993) 10089–10092. https://doi.org/10.1063/1.464397.

[36]   GROMACS, gmx mindist, (2020). http://manual.gromacs.org/documentation/current/onlinehelp/gmx-mindist.html.

[37]   A.L. Duncan, R.A. Corey, M.S.P. Sansom, Defining how multiple lipid species interact with inward rectifier potassium (Kir2) channels, Proc. Natl. Acad. Sci. U. S. A. 117 (2020) 7803–7813. https://doi.org/10.1073/pnas.1918387117.

[38]   GROMACS, GROMACS pull code, (2020). http://manual.gromacs.org/documentation/current/user-guide/mdp-options.html.

[39]   J.S. Hub, B.L. De Groot, D. Van Der Spoel, g_whams - A Free Weighted Histogram Analysis Implementation Includin Robust Error and Autocorrelation Estimates, J. Chem. Theory Comput. 6 (2010) 3713–3720. https://doi.org/10.1021/ct100494z.

[40]   P. V. Klimovich, M.R. Shirts, D.L. Mobley, Guidelines for the analysis of free energy calculations, J. Comput. Aided. Mol. Des. 29 (2015) 397–411. https://doi.org/10.1007/s10822-015-9840-9.

[41]   O.N. Vickery, P.J. Stansfeld, CG2AT, GitHub. (2020). https://doi.org/10.5281/zenodo.3890163.

[42]   W. Humphrey, A. Dalke, K. Schulten, VMD: Visual Molecular Dynamics, J. Mol. Graph. 14 (1996) 33–38. https://doi.org/10.1016/0263-7855(96)00018-5.
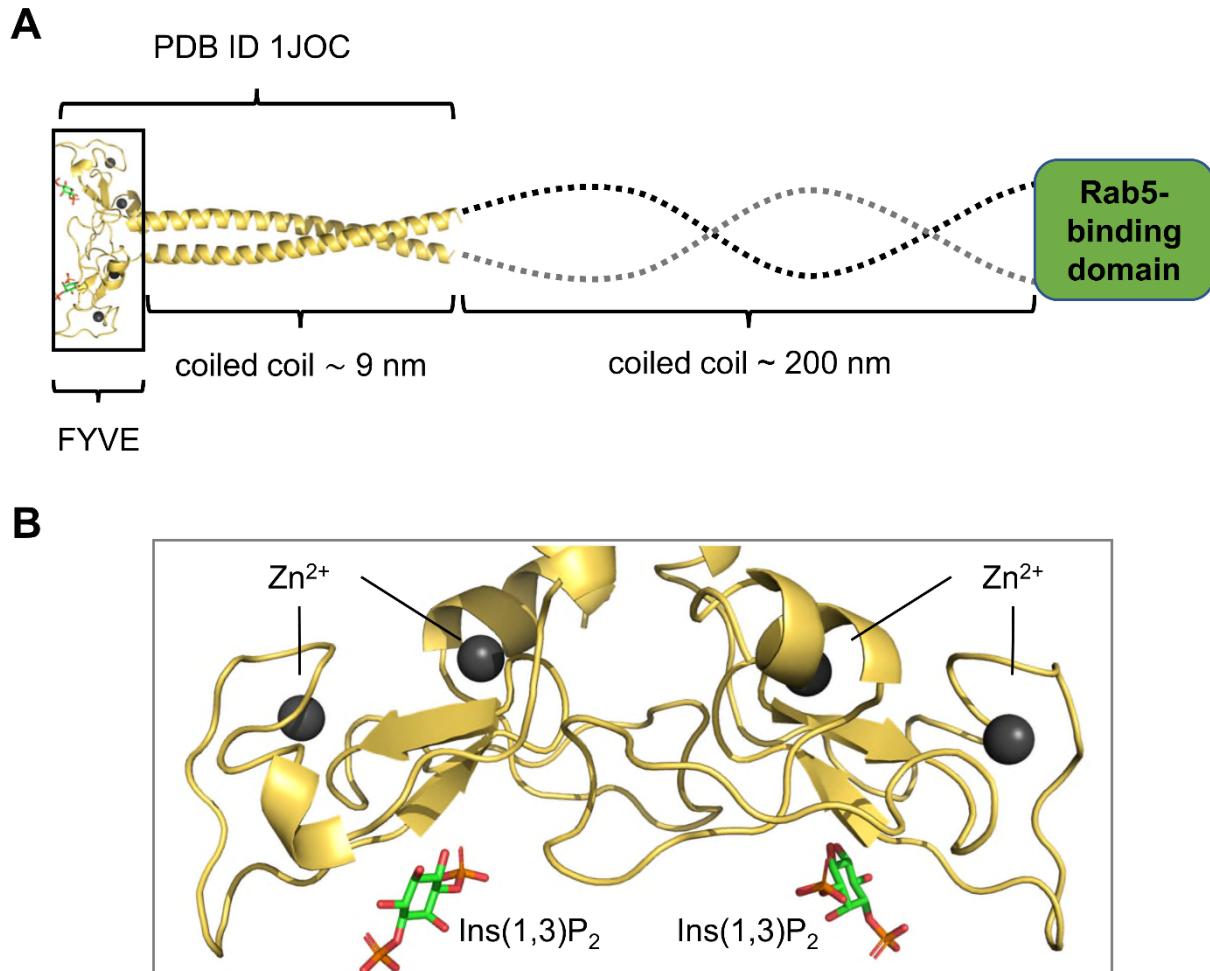
Figures

A

B

Figure 1. Structure of EEA1 and the FYVE domain.

A: Schematic of the full-length structure of dimeric EEA1, with the crystal structure of the FYVE domain dimer plus start of the coiled coil (PDB ID 1JOC) in yellow, followed by a ~200 nm coiled coil (dotted line) and a Rab5-binding domain.

B The FYVE domains of the dimer, with four zinc ions (dark spheres) bound by 4-cysteine zinc fingers. Two bound inositol (1,3)-bisphosphate molecules (red/orange/green) are shown.
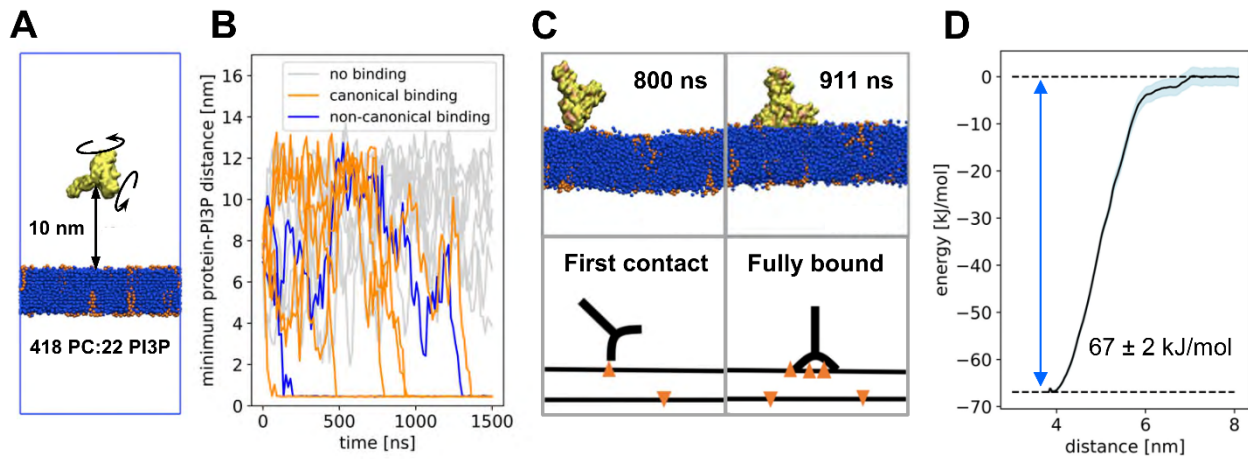
Figure 2. Coarse-grained simulations. (A) Set-up of the CG simulation, with the truncated FYVE dimer (yellow) placed at a minimum distance of 10 nm from the lipid bilayer (blue; orange PI(3)P molecules) and rotated at random. (B) Minimum distance for each of the 15 repeats, showing repeats without binding (grey, 7/15), with canonical binding (similar to the binding mode of the crystal structure, orange, 6/15) and with non-canonical binding (blue, 2/15). (C) Two snapshots from one of the repeats with canonical binding, after 800 and 911 ns. The snapshots show the hinge-mechanism by which FYVE binds to the membrane. Sketches below for clarity. (D) Potential of mean force (PMF) calculations for the binding of FYVE to a lipid membrane containing PC and PI(3)P, in the canonical binding mode, showing mean (dark blue line), standard deviation (shaded area) as well as minimum and maximum values (dotted lines). The well depth is 67 ± 2 kJ/mol.
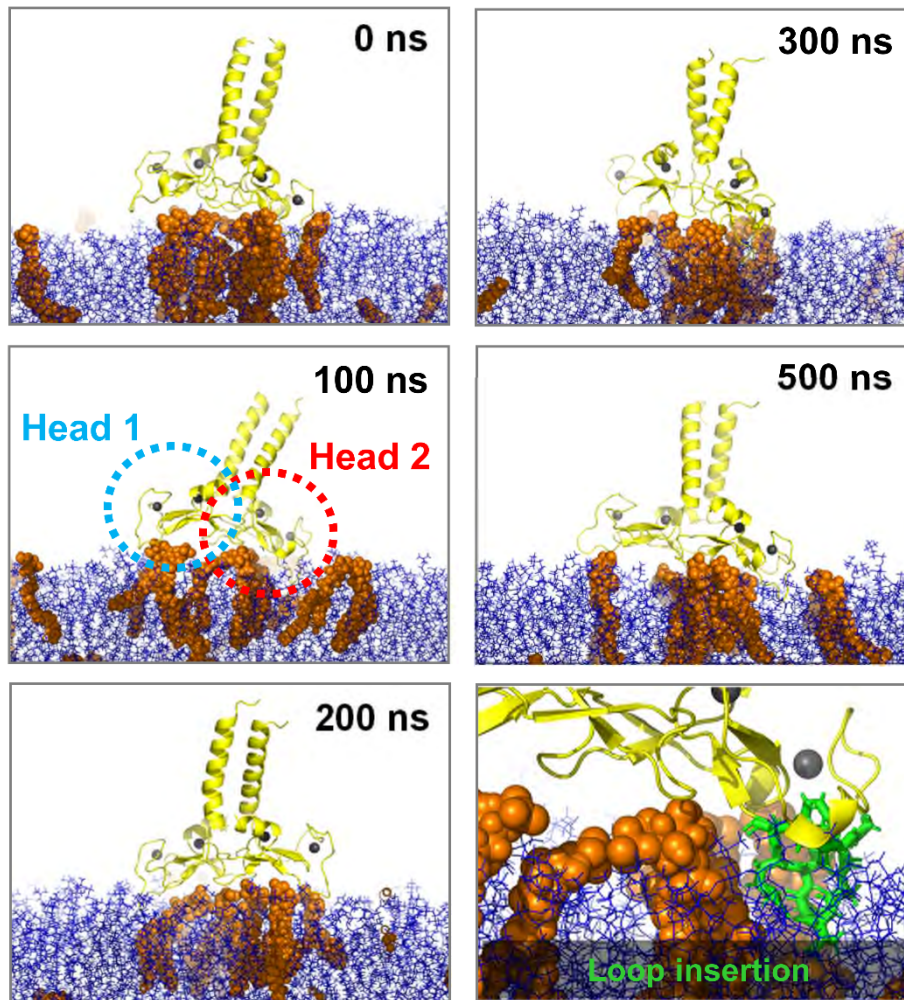
Figure 3. Atomistic simulations. Snapshots from the simulation at 0, 50, 100, 200, 300 and 500 ns. Heads 1 and 2 of the FYVE domain are highlighted (dotted circles) for the 100 ns snapshot. The inserted loop as seen at 100 ns shown in in green stick representation in the final panel.
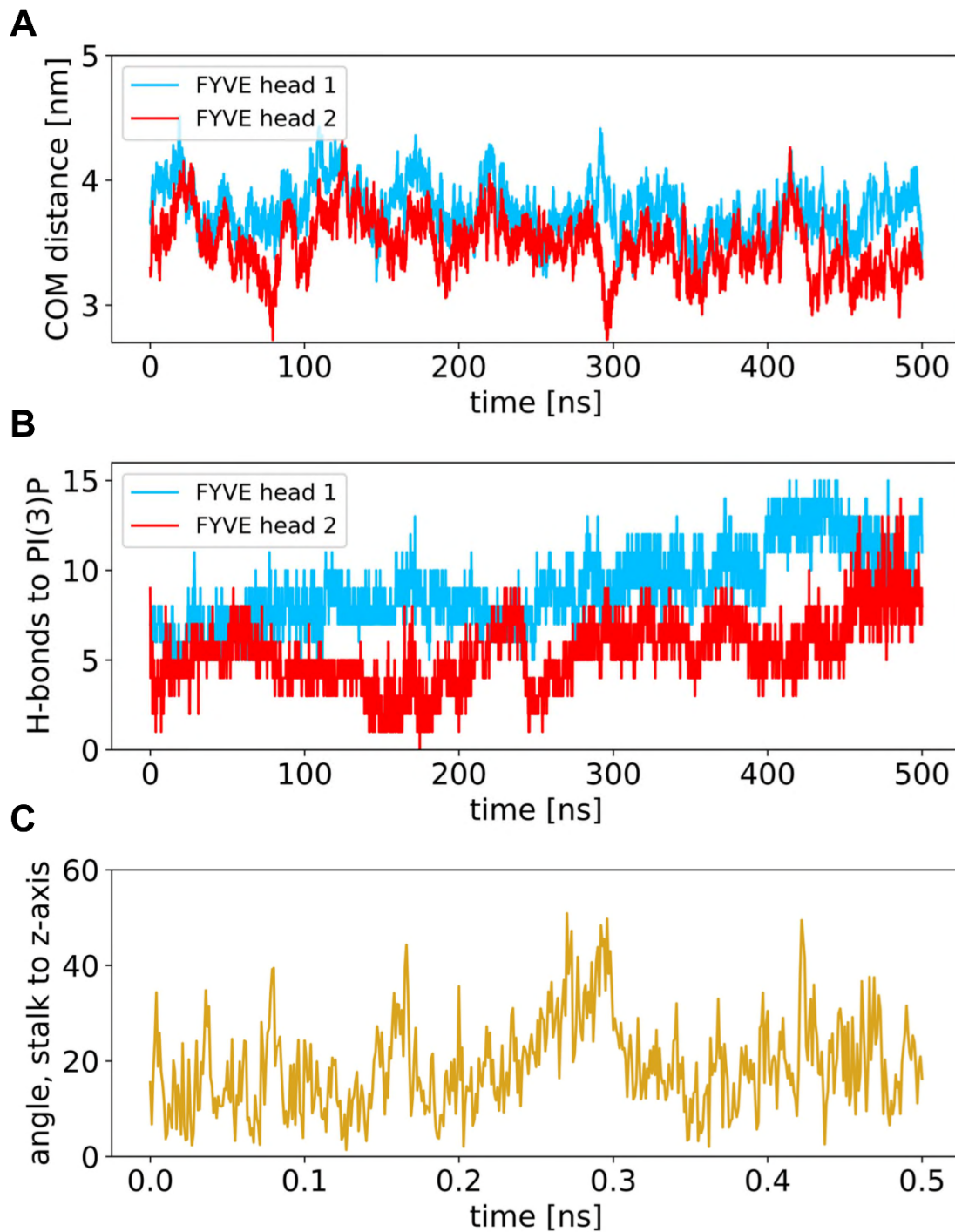
Figure 4. Atomistic simulations – distances, H-bonds and stalk tilt angle.

A Centre-of-mass distance between the membrane and each of the heads as function of time.

B Number of H-bonds with PI(3)P headgroups for the two heads of FYVE. (D) Root mean square deviation (RMSD) for, respectively, the head after aligning all structures to the head, the stalk after aligning to the stalk, and the head after alignment to the stalk.

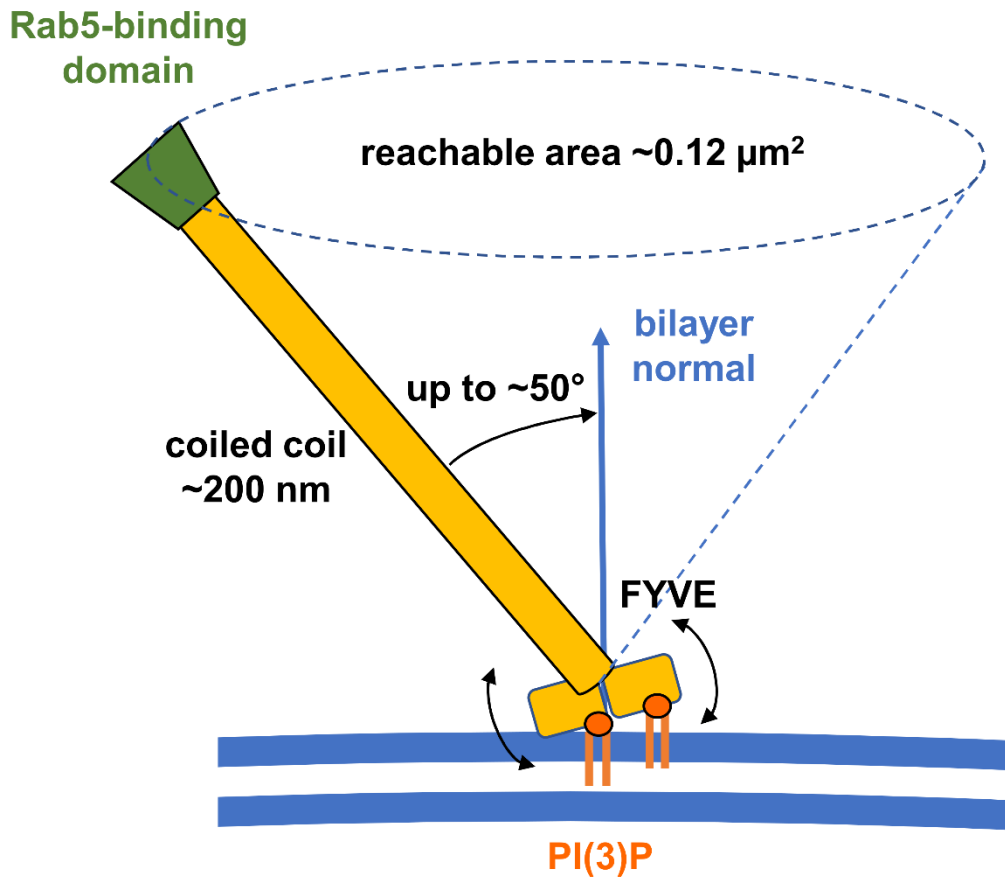**C** Angle between the stalk axis and the bilayer normal as a function of time.

Figure 5. Vesicle tethering and title of EAA1 coiled coil stalk relative to the bilayer normal. The FYVE dimer is shown interacting with two PI(3)P molecules (orange) in a cell membrane (blue lines). Based on the analysis of our simulations this allows the coiled coil to tilt by up to 50° relative to the bilayer normal. Thus, the Rab5-binding domain (green) is able to search an area of ~0.12 nm² for a contact to the target vesicle.