JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

pubs.acs.org/jcim Article

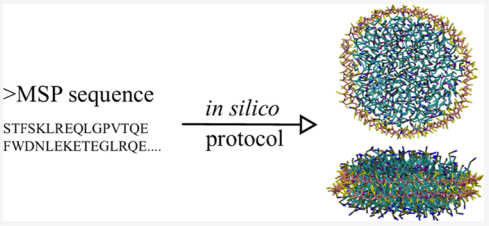
General Protocol for Constructing Molecular Models of Nanodiscs

Lisbeth R. Kjølbye, Leonardo De Maria, Tsjerk A. Wassenaar, Haleh Abdizadeh, Siewert J. Marrink, Jesper Ferkinghoff-Borg,* and Birgit Schiøtt*

Cite This: *J. Chem. Inf. Model.* 2021, 61, 2869−2883 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Nanodisc technology is increasingly being applied for structural and biophysical studies of membrane proteins. In this work, we present a general protocol for constructing molecular models of nanodiscs for molecular dynamics simulations. The protocol is written in python and based on geometric equations, making it fast and easy to modify, enabling automation and customization of nanodiscs *in silico*. The novelty being the ability to construct any membrane scaffold protein (MSP) variant fast and easy given only an input sequence. We validated and tested the protocol by simulating seven different nanodiscs of various sizes and with different membrane scaffold proteins, both circularized and noncircularized. The structural and biophysical properties were analyzed and shown to be in good agreement with previously reported experimental data and simulation studies.



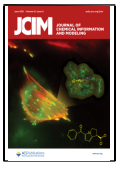
Downloaded via UNIV OF ALBERTA on June 1, 2022 at 21:08:26 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

INTRODUCTION

Recently, nanodisc technology (NDT) has gained increased attention as a valuable tool for biochemical and biophysical functional studies of membrane proteins, as well as for isolation, purification, and solubilization of membrane proteins.[1] Several nanoparticle technologies exist, the difference being how the lipids are stabilized by either amphipathic polymers such as styrene maleic acid copolymers (SMALPS),[2,3] or by amphipathic proteins such as Saposin A (Salipro)[4] or membrane scaffold proteins (MSPs).[5,6] In this work, we focus only on the last mentioned technology, the nanodisc technology. A nanodisc (ND) is a noncovalent assembly of phospholipids and two MSPs.[5,6] The MSPs are re-engineered versions of the human apolipoprotein A-I (apo A-I), which is a 243-residue-long protein weighing around 160 kDa. Natively, Apo A-I is found in high-density lipoprotein (HDL) particles, which transport cholesterol, phospholipids, and triglycerides in the body.[7,8] Apo A-I consists of 10 tandem, (11- or 22-residue) repeats or constituent helices punctuated by either proline or consecutive glycine residues, enclosing the lipid-binding domain,[1] see Figure 1. Several versions of MSPs have been generated by truncating and inserting these tandem repeats,[9] see Figure 1 Panel B. By varying the length of the MSPs, the diameter of the ND can be controlled.[1] Newer versions of NDs have appeared, where the N- and C-terminals are covalently linked thereby circularizing the MSPs.[10,11] This has been shown to increase the stability of the ND.[9,12] Along with circularizing, a high abundance of negative charges have been introduced for higher expression and yield.[10]

However, currently there are only a limited number of three-dimensional (3D) structures of Apo A-I and hence nanodiscs available, thereby limiting the full potential of the technology. Two X-ray crystal structures of the Apo-A-I protein (PDB: 1AV1[13] and 3R2P[14]) in the absence of lipids were initially published, followed by an NMR-derived structure of a truncated version of the Apo-A-I dimer termed MSPΔH5 (PDB: 2N5E[7]), since the tandem sequence identified as helix 5 is removed.[7,13,14] Common for these 3D structures is a more than 80% helical conformation,[15] indicating lipid bound structures, since it has been shown that a lower helical propensity is present without lipids.[16] The helical content of the Apo-A-I structure was determined from circular dichroism (CD) data, to be about 50% in the absence of lipids and about 80% when lipids are present.[17] Before the NMR-derived structure was published, two models were proposed describing how the two MSP proteins could be arranged in a nanodisc known as the picket-fence model[15] and the double-belt model.[18] The double-belt model is now the widely accepted model and the one confirmed in the NMR-derived structure. In the double-belt model, the two MSPs are oriented perpendicular to the lipid tails in an antiparallel fashion with a left to left (LL) interface,[18] see Figure 1 Panels A and C. The MSPs are proposed to orient to optimize the intramolecular salt bridge network; this orientation is believed to be the LL5/5 orientation.[18] The LL5/5 stands for a left to left orientation of helix 5 in one MSP monomer placed opposite to helix 5 in the other MSP monomer in an antiparallel arrangement,[18] see Figure 1 Panels A and C. However, in the NMR-derived model, helix 5 is removed, hence the NMR-identified interface

Received: February 10, 2021
Published: May 28, 2021



ACS Publications

© 2021 The Authors. Published by American Chemical Society

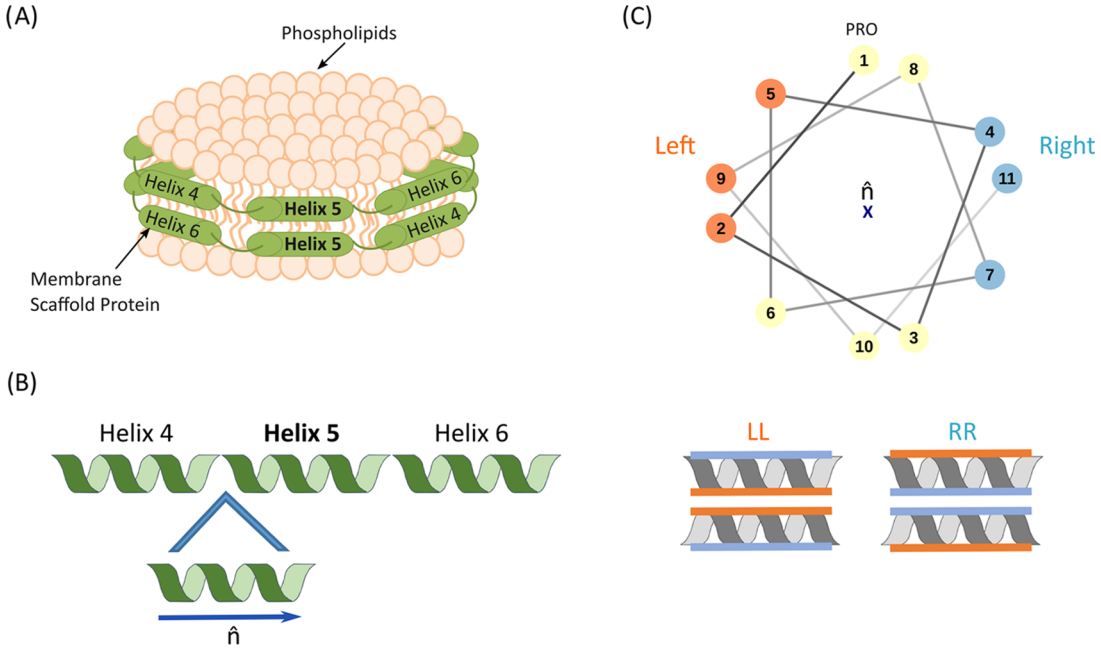


Figure 1. (A) Schematic representation of a nanodisc with an LL5/5 interface. (B) Schematic representation of tandem repeats, which are either inserted or removed from the MSP sequence to increase or decrease the size of the resulting nanodisc. (C) MSP dimer in the nanodisc can either have a left−left (LL) or right−right (RR) interface. Here, each tandem in the MSP is represented by a 11-residue repeat. Proline indicates the beginning of each tandem helix repeat, with residues at positions 4, 7, and 11 defined as the 'right' residues and residues at positions 2, 5, and 9 as the 'left' residues.

is RR9/9, a right to right interface with helix 9 opposite to helix 9 in the two MSP monomers, suggesting the possibility of other interfaces.[7] The interfaces follow the same definition as described in Bibow et al.[7]

Several molecular dynamics (MD) studies have been performed on NDs and related lipid particles, analyzing their self-assembly and structural properties.[16,19−34] However, setting up simulation systems of NDs is not a trivial task. In 2015, Siuda and Tieleman published a potential protocol for constructing NDs *in silico*, based on the crystal structure with PDB ID 1AV1.[16] The authors used coarse grained (CG) methods and an elastic network to maintain a helix conformation, thereby folding the MSPs into a circle, followed by the possibility of adding or removing tandem repeats, hence either increasing or decreasing the size of the ND.[16]

Later CHARMM-GUI introduced a user-friendly ND Builder with a preselected range of different MSPs all assembled with the LL5/5 interface.[35] However, accessing this tool is restricted to the browser interface relying on CHARMM software and other CHARMM-GUI modules,[36] which prevents scripted automation. Furthermore, the CHARMM-GUI ND builder implements a Generalized Born with Simple SWitching (GBSW) minimization for generating the MSP monomers, which becomes increasingly slow with long MSP chains. In this work, we suggest a possible solution to these inconveniences by constructing the MSPs using geometrical equations implemented in python, thereby avoiding the long minimization step in constructing the MSP monomer and neither relying on CHARMM software nor the CHARMM-GUI Modules. The novelty of this work is therefore in the ability to construct the MSP monomers for any sequence and length, as the first critical step in an *in silico* representation of NDs. The tool presented herein thereby enables construction of circularized NDs[9,12] and NDs with a higher abundance of negative charges in the sequence,[10] along with the possibility of testing different interfaces between the two MSP monomers, LL or RR with a user-defined registry. Coupled with the tool Insane.py,[37] the protocol makes it possible to construct CG NDs *in silico* rapidly and easily while still controlling parameters such as the MSP sequence, interface, circularization of the MSP, and bilayer specifications such as lipid types, ratio, and number of lipids per leaflet. The protocol consists of two python scripts and two bash scripts. The two python scripts construct the MSP monomers and assemble the MSP dimer. The atomistic MSP dimer can then be passed to any simulation or building software, applying any force field. In this work, we used the Gromacs suite, being one of the mainly used simulation packages, for which a 'wrapper' bash script was constructed, for automating the process of building the ND, adding the phospholipids, solvation, neutralization, minimization, and equilibration. The last bash script is used for correcting the parameter files for circularized MSPs in Gromacs software, however, any simulation software can be used.[38,39] See Figure 2 for an overview of the protocol. The scripts and examples on complex lipid compositions and/or embedded membrane proteins have been made available in github at https://github.com/LHRK/Nanodisc-Builder/tree/master.

In the following, we will first describe the ND building tool in detail and subsequently demonstrate that this protocol provides good starting conformations of NDs. To validate the building tool, the first step is to build the disc MSPΔH5 (named MSPΔH5_model) and compare it to simulations using the experimental determined structure directly (named MSPΔH5_NMR). Subsequently, we construct and simulate commonly used NDs[10−12,40−47] and analyse and compare the lipid properties to previous studies as well as to a simulated lipid bilayer. The six selected NDs are three noncircularized discs named 1D1, 1E3D1, and 2N2[41−45] and three circularized

https://doi.org/10.1021/acs.jcim.1c00157
J. Chem. Inf. Model. 2021, 61, 2869−2883

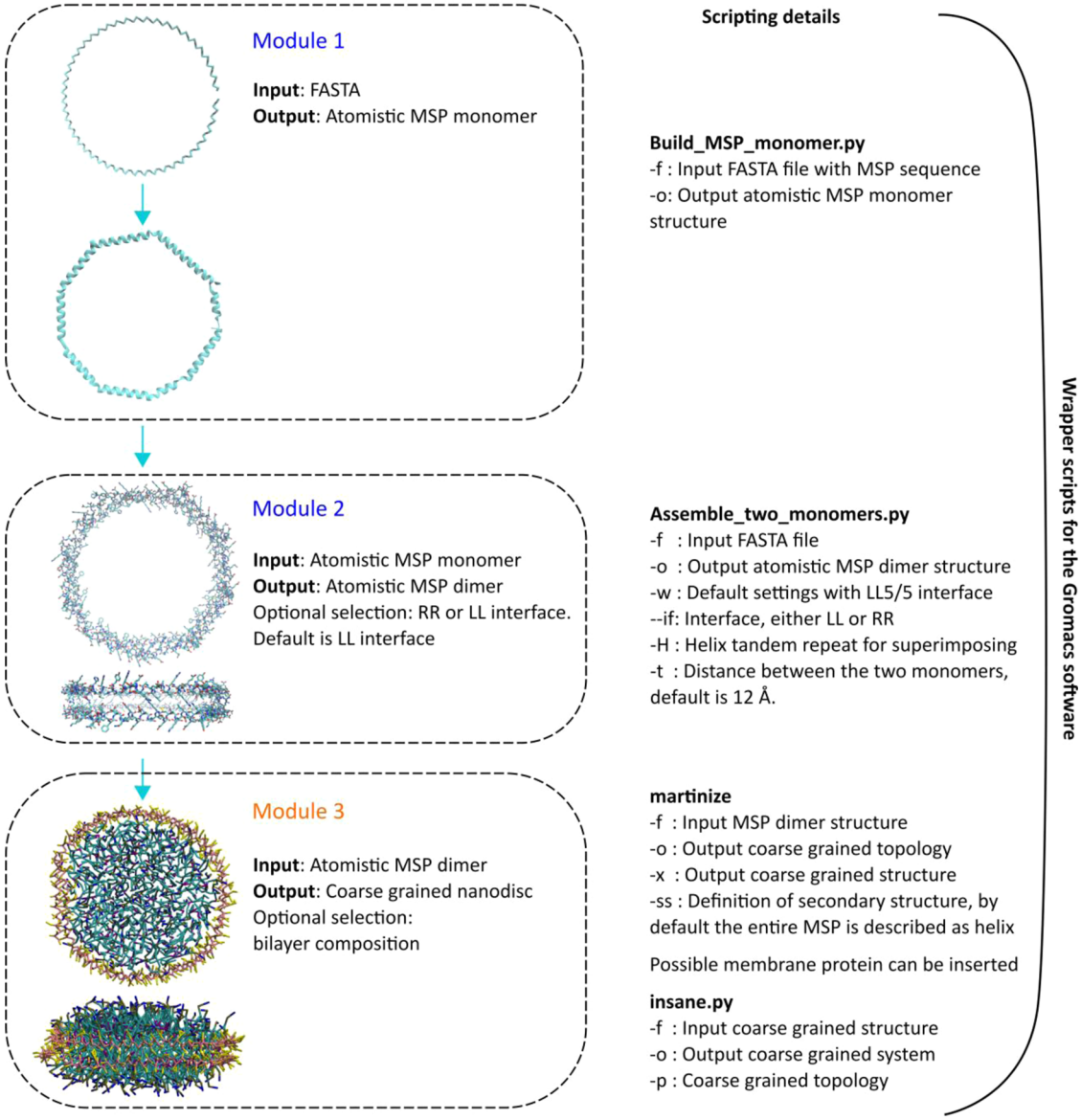


Figure 2. Representation of the protocol workflow for building molecular models of nanodiscs. The workflow consists of three modules. (1) Construct the MSP monomer, (2) assemble the MSP dimer, and (3) coarse grain and insert lipids along with minimization and equilibration. Modules 1−3 are implemented as separate python scripts, with a "master" bash script as a wrapper. Details regarding the scripts are given on the right side.

Table 1. Overview of Systems[a]

system name	diameter (nm)	#lipids	#Cl^-	#Na^+	box size (nm)	simulation length
		Coarse Grained				
^{ls}MSPΔH5_Model	∼8	102 DMPC	291	301	16 × 16 × 16	∼5 μs
^{ls}MSPΔH5_NMR	∼8	100 DMPC	291	302	16 × 16 × 16	∼5 μs
^{ls}1D1	∼9.7	138 POPC	286	296	15 × 15 × 15	∼5 μs
^{ls}1E3D1	∼12.9	268 POPC	673	693	20 × 20 × 20	∼5 μs
^{ls}2N2	∼16	604 POPC	1300	1322	25 × 25 × 25	∼5 μs
^{cs}NW9	∼9	116 POPC	696	730	20 × 20 × 20	∼5 μs
^{cs}NW11	∼11	146 POPC	687	727	20 × 20 × 20	∼5 μs
^{cs}NW13	∼13	302 POPC	650	708	35 × 35 × 35	∼5 μs
POPC LB		660 POPC	74	74	15 × 15 × 7	∼1 μs
		All-Atom				
^{ls}MSPΔH5_Model_AA	∼8	102 DMPC	203	213	15 × 15 × 10	∼200 ns

[a]LB refers to a lipid bilayer and ls and cs in superscript refer to noncircularized and circularized discs, respectively. The lipids used are either 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) or 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC).

https://doi.org/10.1021/acs.jcim.1c00157
J. Chem. Inf. Model. 2021, 61, 2869−2883

discs named NW9, NW11, and NW13[10−12,40,46,47] all of varying sizes, see Table 1.

THEORY

The protocol behind the tool for constructing an ND *in silico* consists of 3 modules, see Figure 2.

Module 1. The first module involves the construction of an MSP monomer given a FASTA[48] sequence as the input provided by the user. The MSP is a circular α-helix punctuated by either a proline or two consecutive glycine residues.

The first challenge is thus to construct the circular α-helix without introducing an unnecessary strain in the structure. It is possible to describe an α-helix using the following geometric equation[49]

$$\begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} + \begin{bmatrix} r\cdot\sin(t\cdot i) \\ [r\cdot\cos(t\cdot i)] \\ p\cdot i \end{bmatrix} \quad (1)$$

where $[x_i, y_i, z_i]$ is the vector identifying the position of the ith C_α atom of the helix, $[x_0, y_0, z_0]$ indicates the origin of the helix system, and the helical axis $\hat{\mathbf{n}}$ is parallel to the z-axis. The three parameters are the radius of the helix (r), translation or pitch (p), and turn angle (t), respectively.[49] For simplicity, transformations are applied to describe the helix in the xy-plane, see the Supporting Information (SI) for these transformations. Once the helical axis is parallel to the y-axis, the helix axis ($\hat{\mathbf{n}}$) can be updated after every turn (four residues) along a circle, as shown in Figure 3, thereby

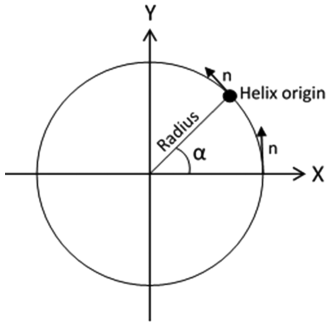


Figure 3. Helical axis ($\hat{\mathbf{n}}$) can be updated for every turn in the helix along a circle with a certain radius.

obtaining eq 2, describing the positions for each ith C_α atom in the input sequence, constructing a circular α-helix trace with a radius R_0.

$$\begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix} = \begin{bmatrix} R_0\cos\alpha \\ [R_0\sin\alpha\,] \\ 0 \end{bmatrix} + \begin{bmatrix} r\cdot\sin(t\cdot i)\cdot\cos\alpha - p\cdot i\cdot\sin\alpha \\ [r\cdot\sin(t\cdot i)\cdot\sin\alpha + p\cdot i\cdot\cos\alpha] \\ -r\cdot\cos(t\cdot i) \end{bmatrix} \quad (2)$$

The three parameters, radius of the helix (r), pitch (p), and turn angle (t) are chosen to be 2.314 Å, 1.516 Å, and 100.1°, respectively, since these are the values for an ideal α-helix.[50] The angle α is chosen to be 2π/number of turns, where the number of turns is estimated from the length of the input sequence. Finally, R_0 is the radius of the circle. After constructing the Cα-trace from the input sequence, a straight α-helix is built using the python module PeptideBuilder.[51] The helix is then subdivided into helix pieces or tandems with the ends defined where either a proline residue or two consecutive glycine residues are located. Each tandem is then aligned with the constructed Cα-trace in Module 1, such that the full straight helix is bent into a circle. Subsequently, for each tandem a hydrophobic vector is defined from the center-of-mass of the tandem to the center-of-mass of the side chains of the residues alanine, valine, isoleucine, leucine, methionine, phenylalanine, tyrosine, tryptophan, and proline. Each tandem is then rotated such that the corresponding hydrophobic vector points toward the center of the circle. By doing so, we make sure that the hydrophobic side of each tandem faces the phospholipids in the ND.

Module 2. The second module involves assembling two identical MSP monomers with a user-defined interface. The interface can be either an LL or an RR interface, but default is an LL interface, see Figure 1 Panels A and C. The interfaces follow the definition described in Bibow, et al.[7] The user will also define which tandem or residue(s) to use for superimposing the two MSP monomers opposite in an antiparallel arrangement along with the distance separating the two MSP monomers. By default helix 5 will be used if present in the input sequence.

Module 3. This module involves coarse graining the MSP dimer based on the Martini force field,[52] using the Martinize script,[52] followed by adding the desired phospholipids using the Insane.py script.[37] The now built ND can then be minimized and equilibrated followed by either a production run or by back-mapping the ND to an all-atom (AA) representation. An atomistic topology file for the MSP dimer can be generated for back-mapping using either the Backward.py or Initram.sh script automatically, if desired.[53]

System Setup and MD Protocols. A total of nine CG and one AA systems were constructed and each simulated in three replicas, using Gromacs version 2018.2 software,[39] see Table 1 for an overview of the systems. The systems named 1D1, 1E3D1, 2N2, NW9, NW11, and NW13 contain three noncircularized (1D1, 1E3D1, 2N2) and three circularized (NW9, NW11, NW13) discs in increasing size, respectively. The sequences for the noncircularized discs are from Sligar's group[9] and the sequences for the circularized discs are from Nasr et al.[11] and from Johansen et al.[10]

All systems, except the MSPΔH5_NMR, MSPΔH5_Model_AA, and POPC LB systems, are constructed using the described building protocol. The MSPΔH5_NMR was constructed directly from the NMR-derived PDB structure (PDB ID 2N5E), chain A and B, and coarse grained using the Martinize.py script[52] in combination with DSSP 2.0.4 (define secondary structure of proteins) software[54,55] for determining the secondary structure. Ten conformations were resolved using NMR, and the first conformation was chosen for this study since the difference in root-mean-square deviation (RMSD) between the conformations was only ∼0.45 nm, mainly caused by the movement of the terminals. The sequence for the MSP in the systems MSPΔH5_Model and MSPΔH5_Model_AA was extracted from chain A using PDB ID 2N5E.

A ratio of 1:50 of MSP/DMPC lipids was used in the MSPΔH5_NMR, MSPΔH5_Model, and MSPΔH5_Model_AA systems to compare with the NMR study.[7]

The number of POPC lipids selected for systems 1D1, 1E3D1, 2N2, NW9, NW11, and NW13 are based on eq 3

$$M = 2(\pi r + \sqrt{\pi NS})/L \quad (3)$$

https://doi.org/10.1021/acs.jcim.1c00157
J. Chem. Inf. Model. 2021, 61, 2869−2883

Here, M represents the number of residues in the MSP sequence, S the mean area per lipid (70 Å^2 for POPC inside an ND),[56,57] r the radius of the MSP helix (5.5 Å),[9] L the helical pitch per residue (1.5 Å),[9] and N the number of lipids per leaflet in the disc. The equation is derived geometrically and validated from small-angle X-ray scattering (SAXS) and high-performance liquid chromatography (HPLC) experiments combined.[9,56]

Coarse Grained Systems, Construction, and the Simulation Protocol. The CG systems, except for the POPC bilayer, were solvated and neutralized using the Insane.py script[37] with a salt concentration of 0.15 M. For the systems, MSPΔH5_NMR and MSPΔH5_Model, 10% antifreeze water particles were added because freezing of the water was observed initially in these systems.[58] Martini 2.2[52] was applied in the coarse grained systems, without an elastic network applied on the MSPs. All of the CG systems were treated using the same protocol for both minimization, equilibration, and the production run. The protocol starts with a steepest descent minimization with 5000 steps followed by one NVT equilibration stage and three NPT equilibration stages. In the NVT stage, a timestep of 0.005 ps was used for a total of 500 ns and with the temperature held constant at 310 K using a Berendsen thermostat with τ = 1.0 ps.[59] In the three NPT stages, timesteps of 0.005, 0.01, and 0.02 ps were applied with a constant temperature and pressure at 310 K and 1 bar, using the Berendsen thermostat and barostat[59] with τ = 1.0 and 12 ps, respectively. The compressibility was set to 3×10^{-4} bar^{-1} for all of the steps. The MSP, lipids, and solvent were coupled independently. During the production run a 20 fs timestep is used with frames saved every 500 ps, in three independent replicas with different starting conformations from the last NPT stage and randomized velocities. The temperature and pressure were held constant at 310 K and 1 bar using a V-rescale thermostat[60] with τ = 1.0 ps and Parrinello−Rahman barostat[61] with τ = 12.0 ps, and compressibility at 3×10^{-4} bar^{-1}, respectively. The electrostatics were treated using the reaction-field method[62] and the van der Waals interactions were truncated after 1.1 nm ("New-RF" Martini settings[63]) using the potential-shift Verlet scheme.[64]

The POPC bilayer system was constructed using the CHARMM-GUI Module Membrane builder.[65] The system was minimized for 5000 steps of the steepest descent, followed by five stages of equilibration with timesteps of 0.002, 0.005, 0.010, 0.015, and 0.020 ps. Position restraints were applied on the lipid headgroups during equilibration with a force constant of 200, 100, 50, 20, and 10 $\frac{\text{kJ}}{\text{mol}\cdot\text{nm}}$, respectively, for the four first stages. The simulation settings are the same as for the above mentioned nanodisc systems, with the exception of semi-isotropic pressure coupling. Three replicas of 1 μs each were run.

All-Atom System Construction and the Simulation Protocol. From the MSPΔH5_Model system, the ND extracted from the first frame after equilibration for all three replicas was converted to AA using the CHARMM-GUI tool based on the Backward.py script,[66−68] and resolvated and neutralized using Gromacs tools solvate and genion,[39] with the TIP3P water model and 0.15 M NaCl. The force field used was CHARMM36.[69,70] The systems were minimized with the steepest descent for 500 steps followed by an NVT equilibration stage for 1 ns with a timestep of 1 fs and the temperature maintained at 310 K using the V-rescale thermostat (τ = 1.0 ps).[60] The production run was performed with a timestep of 2 fs and frames saved every 10 ps in three independent replicas each around 200 ns long. The LINCS[71] algorithm was applied to all of the bonds including hydrogens. The temperature and pressure were held constant at 310 K and 1 bar using the V-rescale thermostat[60] with τ = 1.0 ps and the Parrinello−Rahman barostat[61] with τ = 1.0 ps, respectively. The compressibility was set to 4×10^{-5} bar^{-1}. The electrostatics were treated using particle mesh Ewald (PME),[71] and the van der Waals interactions were truncated after 1.2 nm using a switching function from 1.0 nm.

Calculating Chemical Shifts of Cα Atoms. Chemical shifts of the Cα atoms in the atomistic system were calculated using SPARTA+ software[72] through the MDtraj module in python.[73] The chemical shifts were calculated every 100 frames, corresponding to every 1 ns and averaged across frames and replicas providing an ensemble average. The secondary structure chemical shifts were calculated by subtracting a predicted random coil chemical shift from the calculated shifts from SPARTA+ software. SPARTA+ software has an uncertainty of 0.92 ppm for predicting chemical shifts of Cα-atoms.[72] The random coil chemical shift was obtained from the database at the University of Copenhagen (http://www1.bio.ku.dk/english/research/bms/research/sbinlab/groups/mak/randomcoil/script/), using a perdeuterated protein and a temperature at 42 °C and a pH of 7.4.

Lipid Properties. The lipid order parameters are defined as

$$S_i = \frac{1}{2}\cdot\langle 3\cdot\cos^2(\theta) - 1\rangle \tag{4}$$

where θ is the angle between the ith bond in question and the bilayer normal. For the CG systems the bonds applied in the calculations were defined from the beads of the lipid tails.[74] An in-house python script and the tool *g_ordercg* were used for calculating the order parameter of the phospholipids over a grid with a 2 Å × 2 Å bin size and averaged across frames and lipids for each bin.[75] The mean-square displacement (MSD) was calculated with the Gromacs tool *msd*, using only the phosphate headgroups of the lipids.[39] The diffusion coefficients were obtained through the fit of the MSD curves. The fit is based on eq 5 obtained from "A note on confined diffusion" by Bickel.[76]

$$\text{MSD} = a^2\left(1 - 8\cdot\sum_{i=1}^{n}\frac{\exp\left(-\alpha_i^2\cdot\frac{t}{\tau}\right)}{\alpha_i^2\cdot(\alpha_i^2 - 1)}\right) \tag{5}$$

where $\tau = a^2/D$, D is the diffusion coefficient, and a represents the radius of the disc. The coefficients α_i are the ith positive root of the derivative of the Bessel function of the first kind, $J_1'(\alpha_i) = 0$. In principal, the calculation involves all possible roots α_i, however, in practice it is sufficient with the first few terms only, since the behavior of the system at large times ($t \sim \tau$) is dictated by these (specifically we have used n = 3). The difference between the diffusion coefficients across the systems was accessed with a two-sided t-test, using the function *ttest_ind* in Scipy.[77] An estimate of the curvature of the lipid bilayer patch in the NDs was obtained as the averaged difference in the z-height from the center-of-mass of the ND and the rest of the lipid headgroups in the disc.

Protein Properties. The RMSD of the MSPs was calculated using the Gromacs tool *rms*.[78] For the CG systems, the calculation was based on the backbone beads while the Cα

https://doi.org/10.1021/acs.jcim.1c00157
J. Chem. Inf. Model. 2021, 61, 2869−2883

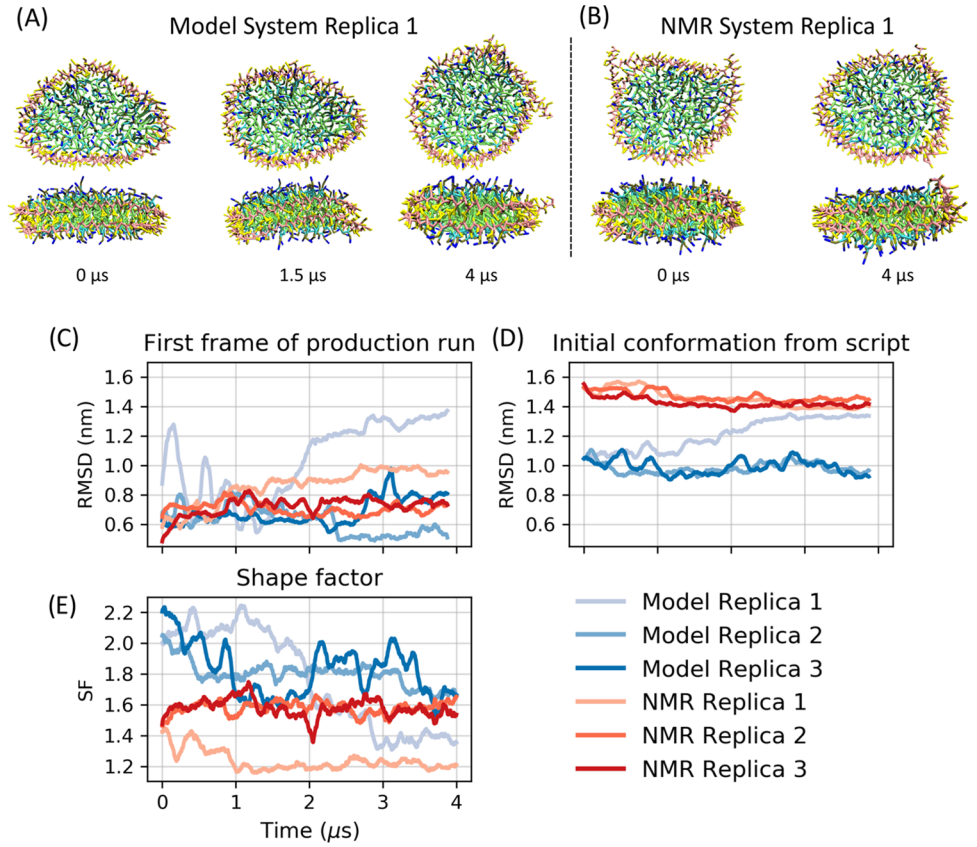


Figure 4. (A + B) Snapshots of the ND from replica 1 of both the MSPΔH5_Model and MSPΔH5_NMR systems. The MSPs are shown in yellow and orange, while the DMPC lipids are shown in green and blue. (C) RMSD plotted as a running averaged across 125 ns of the position of the backbone beads of the MSP in the two systems MSPΔH5_Model and MSPΔH5_NMR, referred to as Model and NMR, respectively. The first frame of the production run is used as the reference conformation. (D) RMSD plotted as a running averaged across 125 ns of the position of the backbone beads of the MSP for the Model and NMR systems, using the initial conformation from the protocol as the reference conformation. (E) Shape factor of the MSP structure in the Model and NMR systems. Plotted as a running average over 125 ns. The shape factor equals 1 for a perfect circle.

atoms were used in the atomistic system. The difference between RMSD across systems was accessed using either a one-way ANOVA or with a two-sided *t*-test using, respectively, the function *f_oneway* or *ttest_ind* in Scipy.[77] The secondary structure of the MSPs in the atomistic system was calculated using the Gromacs tool *do_dssp*.[54,78] The solvent accessible surface area (SASA) was determined using the Gromacs tool *sasa*, using a probe of 0.14 nm and 500 points.[78] The NDs were clustered with the Gromacs tool *cluster*,[39] applying the clustering algorithm Gromos.[79] The backbone beads of the MSPs were used for the clustering with cutoff values of 0.6, 0.9, 1.0, 0.6, 0.4, and 0.6 nm for the discs 1D1, 1E3D1, 2N2, NW9, NW11, and NW13, respectively. The cutoff values were estimated from the RMSD distribution of the backbone in the protein structures in the different discs. The cutoff was selected based on the lowest RMSD value, where the majority of the variance in the data was described. The shape factors were calculated using either the Cα-atoms or the backbone beads of the MSPs for the AA or CG-simulations, respectively, using the python module MDAnalysis.[80] The shape factor was calculated every 5 ns.

The shape factor SF is defined as

$$\mathrm{SF} = \lambda_1/\lambda_2 \tag{6}$$

where λ_1, λ_2, and λ_3 are the moments of inertia calculated as the eigenvalues of the covariance matrix of the backbone beads or Cα-atoms for CG and AA, respectively. The eigenvalues follow a descending order such that $\lambda_1 \geq \lambda_2 \geq \lambda_3$. The two largest eigenvalues represent the two largest principal axes in the ND plane, which in this work is the *xy*-plane. The shape factor equals 1 for a perfect circle, and hence the deviation from 1 provides an indication of the extent of anisotropy. As we could not assume normality for the SF statistics, differences were studied via a one-sided Mann Whitney U test using the function *mannwhitneyu* in Scipy.[77] All plots and figures were made with the python module matplotlib[81] and visualization software VMD.[82]

RESULTS AND DISCUSSION

The focus of the analysis is to first verify that the building tool produces reasonable starting conformations of NDs near equilibrium. The CG systems named MSPΔH5_Model and MSPΔH5_NMR are therefore constructed. The model system is constructed using the protocol, while the NMR system is constructed using the MSP structure directly from PDB 2N5E. The lipid properties in these systems are thereafter analyzed and compared, followed by studying the protein properties in the system MSPΔH5_Model_AA. The MSPΔH5_Model_AA system is the ND from the MSPΔH5_Model system converted to an all-atom resolution. More specifically, the lipid order parameters, bilayer thickness, and MSD of the lipids are investigated for the two systems, MSPΔH5_Model and

https://doi.org/10.1021/acs.jcim.1c00157
J. Chem. Inf. Model. 2021, 61, 2869−2883

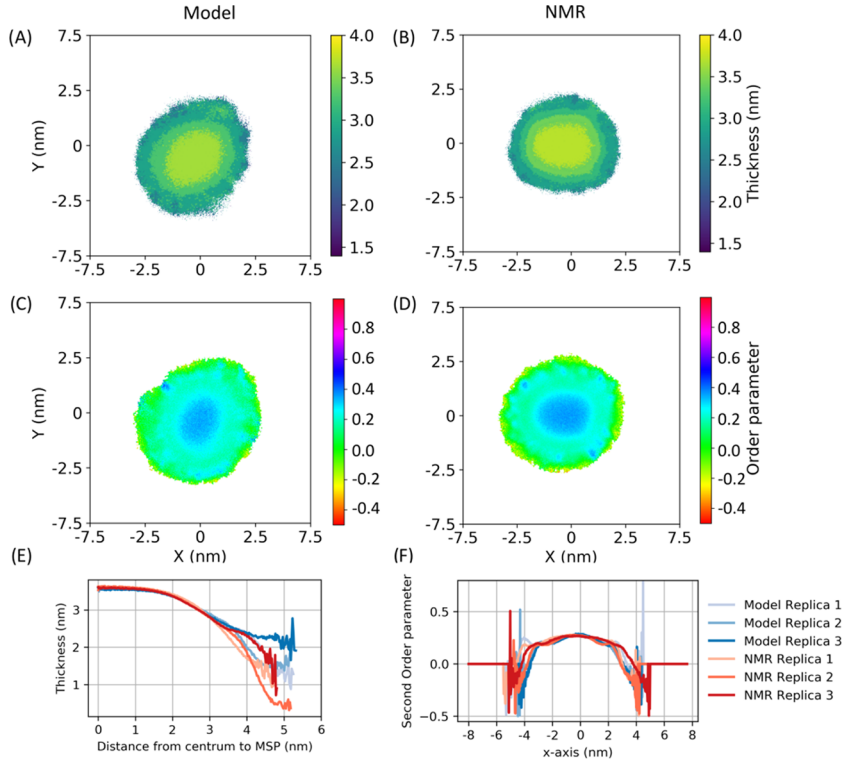


Figure 5. Thickness and order parameter of the lipids in the systems MSPΔH5_Model and MSPΔH5_NMR, referred to as Model and NMR, respectively. (A + B) Thickness and (C + D) secondary order parameters are plotted along both the *x*- and *y*-axes as a top view of replica 1 for both systems. (E) Thickness and (F) order parameter profiles. The thickness profile is plotted as a function of the distance to the MSPs, while the order parameter profile is plotted as a function of the *x*-axis coordinates.

MSPΔH5_NMR, while the RMSD, secondary structure, and chemical shifts of the Cα atoms are calculated for the MSPΔH5_Model_AA system for comparison with experimental data. From these comparisons, we show that the protocol can construct stable and meaningful NDs. After validation of the building tool, three different sized non-circularized (1D1, 1E3D1, 2N2) and three different sized circularized (NW9, NW11, NW13) NDs were constructed, analyzed, and compared across discs and with a simulated lipid bilayer for benchmarking and testing the tool.

Protocol Validation. The stability of the discs in the systems MSPΔH5_Model and MSPΔH5_NMR is inspected through the RMSD of MSPΔH5. As references, both the initial conformation directly from the protocol and the first frame of the production run were used, see Figure 4 Panels C and D.

The overall RMSD, calculated using the first frame from the production run as the reference, remains stable for both systems (Model and NMR) with the replicas fluctuating between 0.5 and 1 nm in RMSD. The observed increase in RMSD for replica 1 of the Model system is caused by a change in the shape of the disc, from an initial more elongated shape after the equilibration to a final circular shape, see Figure 4 Panels A and B. However, overall, the discs remain stable. To study this further, the shape factor was calculated. Figure 4 (Panel E) shows the shape factor as a function of simulation time. After approximately 3.5 μs, replicas 2 and 3 of the Model system reach the same values as replicas 2 and 3 of the NMR system, while replica 1 from both the Model and NMR system equilibrates around 1.2−1.4. These observations suggest similar shapes for both the Model and NMR system, which can also be observed visually from the snapshots in Figure 4 (Panels A and B). However, it could appear that two metastable states are present, one with a shape factor around 1.6 and another between 1.2 and 1.4. Hence, it is possible that both circular and more elliptical shapes of the nanodisc are present in the ensemble. This has previously been suggested, based on SAXS, NMR, and MD data.[83−85]

Regarding the RMSD values calculated using the initial conformation from the protocol as the reference, the RMSD for the NMR system begins at ∼1.5 nm and decreases to ∼1.4 nm. The Model system begins at ∼1.1 nm in RMSD and replicas 2 and 3 end with an RMSD value of ∼1 nm while the first replica has an RMSD value of ∼1.3 nm after 4 μs. The fluctuation in the RMSD is quite low, suggesting stability. However, the Model system takes around 3 μs to reach the same RMSD values and the NMR system indicating a small bias toward the initial conformation.

The next step is to assess the lipid properties by calculating the order parameters, see eq 4, of the phospholipid tails, the thickness of the bilayer, as well as the lateral diffusion of the phospholipids. The thickness and order parameter profiles for the two systems are shown in Figure 5.

In the center of the disc, the thickness and order parameter are approximately 3.6 nm and 0.24 for both the Model and NMR systems, respectively (Figure 5). Interestingly there seems to be zones in the disc where the lipid properties differ. From the thickness and order parameter profiles of the *xy* plane in Figure 5 Panels E + F and the two-dimensional (2D) profiles in Panels A−D, the disc can roughly be divided into three zones. The first zone starts at the center and ends after ∼2 nm, followed by a zone with a decrease in the thickness and order parameter, between 2 and 4 nm from the center. The third zone is between 4 and 5 nm from the center, where the thickness and order parameter fluctuates before decreasing to

https://doi.org/10.1021/acs.jcim.1c00157
J. Chem. Inf. Model. 2021, 61, 2869−2883

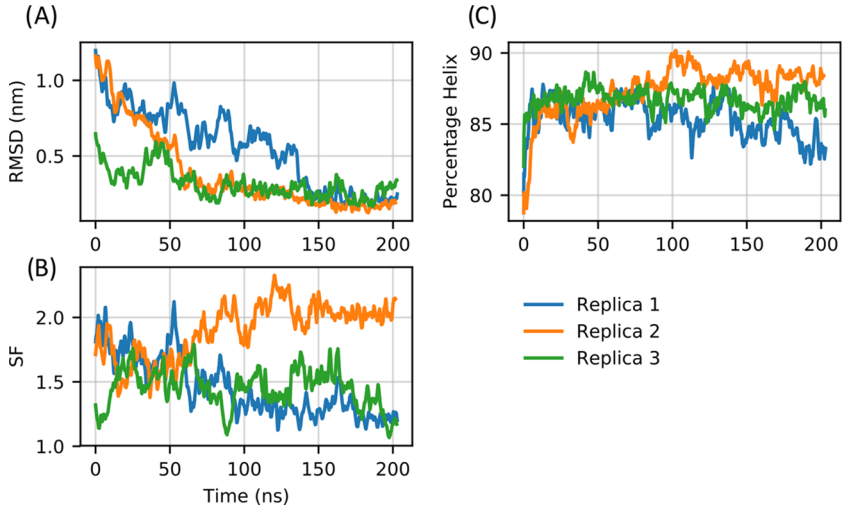


Figure 6. Three observables (RMSD, helicity, and shape factor) are plotted as a running average over 1 ns for the system MSPΔH5_Model_AA. (A) RMSD of the MSPs using the averaged structure from the last 50 ns of the production run as the reference conformation. (B) Shape factor. (C) Fraction of helicity in the MSP structure.

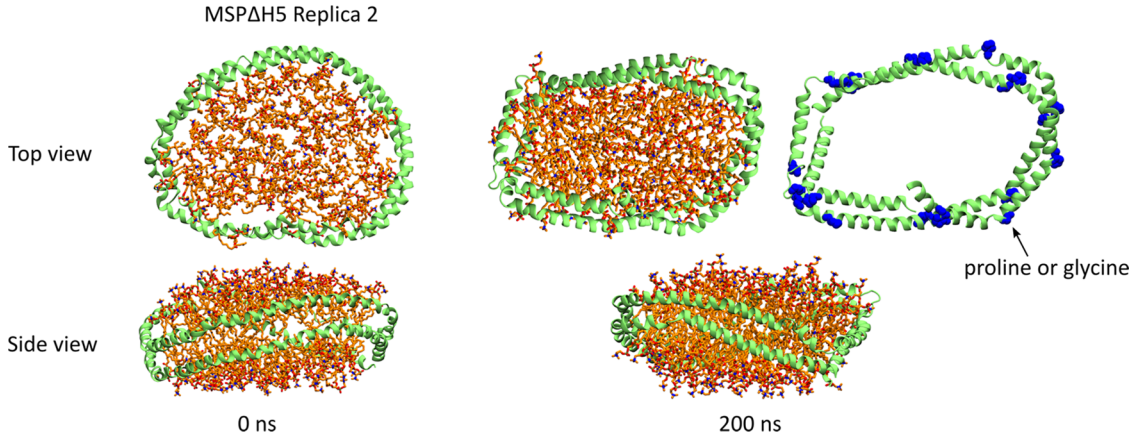


Figure 7. Snapshots of the ND in the system MSPΔH5_Model_AA replica 2 after 0 and 200 ns, The MSP is shown in green and the lipids are shown in orange. Kinks, where either proline or glycine residues are present, become more pronounced resulting in a more elliptic shaped disc at 200 ns. The proline and glycine residues are shown in blue.

low values at the rim of the disc. Based on visual inspections, the reason for this appears to be that the peripheral phospholipids tilt away from the center, toward the scaffold proteins due to interactions between the lipid headgroups and the hydrophilic side chains of the proteins in the rim of the disc. This phenomenon has previously been observed for the same disc, simulated at atomistic resolution.[85] These zones have also been observed in other MD studies of NDs.[16,27,28,85,86] Finally, the lateral diffusion was investigated by calculating the MSD of the phosphate headgroups, with the diffusion coefficients obtained using eq 5 (see Figure S1). The diffusion coefficients are determined to be 49.2 ± 0.3 and 50.0 ± 0.5 nm^2/s for the Model and NMR systems, respectively, averaged across replicas. A two-sided *t*-test showed no significant difference between the two diffusion coefficients (*p*-value 0.13). Overall, the Model system appears to reproduce the same lipid properties as the NMR system, suggesting an accurate model. In the following, we will assess the quality of the model from the perspective of its protein properties.

Validating the Protein Properties for the System MSPΔH5_Model_AA. First, the RMSD of the Cα atoms in the MSPΔH5 was calculated for the system MSPΔH5_Model_AA, see Figure 6 Panel A. The RMSD was calculated using an averaged structure of the last 50 ns of the production run. Replicas 2 and 3 reach a plateau after ∼50 ns while replica 1 continues to decrease in the RMSD value until around 150 ns. This indicates stable nanodiscs in the three replicas.

The stability was also estimated from the fluctuation of the shape factor (eq 6) in Figure 6 Panel B. After around 50−100 ns all of the replicas seem to have found a stable shape. Interestingly, two replicas end with a factor of ∼1.2, while the last replica has a factor of ∼2.0, indicating two states, a circular and an elliptical shape. One example is replica 1, which starts around 2 and ends with a factor of ∼1.2, indicating a change from an elliptical to a circular shape. In general, the kinks, where proline or glycine residues are located in the MSP, become more pronounced visually, see Figure 7. The shape of the discs is still debated in the literature.[83−85] Two papers by Arleth et al. based on small-angle neutron and X-ray scattering have shown that the NDs obtain a more elongated and elliptical shape,[83] whereas other groups report more circular forms.[84] Our simulations hint to a situation where the shape of the discs seems to fluctuate between elliptical and circular, with

https://doi.org/10.1021/acs.jcim.1c00157
J. Chem. Inf. Model. 2021, 61, 2869−2883

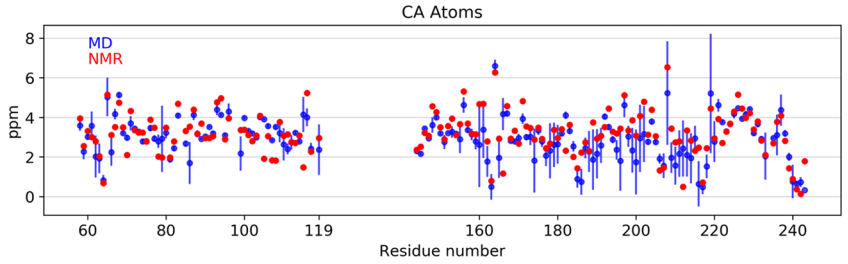


Figure 8. Secondary structure chemical shift of the Cα atoms in the MSP structure. The averaged shift across frames and standard deviation across replicas from the MSPΔH5_Model_AA ensemble are shown in blue. The experimental measured shift for each residue is shown in red. Note residues 121−142 correspond to helix 5, which is deleted in this mutant.

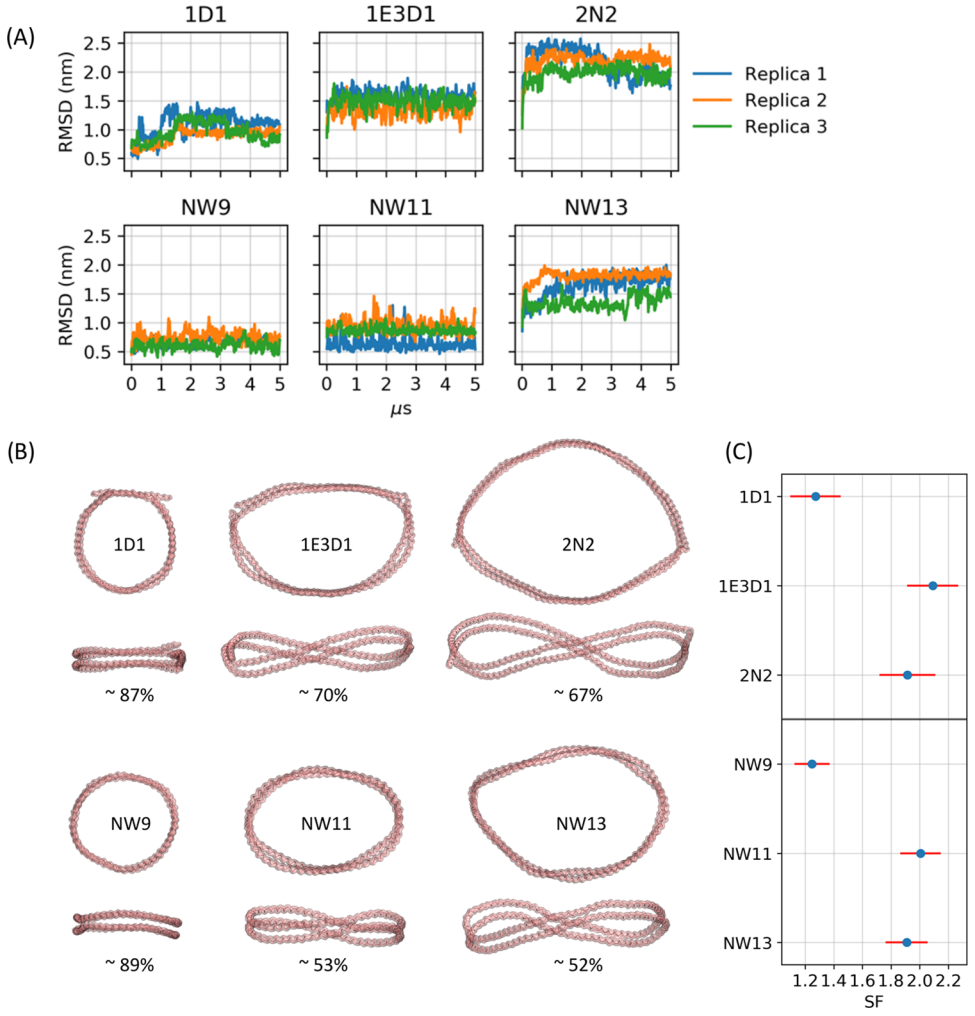


Figure 9. (A) RMSD of the backbone beads in the MSP for the noncircularized discs 1D1, 1E3D1, and 2N2 and of the circularized discs NW9, NW11, and NW13. Plotted as a running average over 20 ns. The first frame of the production run was used as the reference conformation. (B) Representative structures of the largest clusters of each ND. Only the traces of the backbone beads are shown in the top and side views. The population of the main cluster is noted below the extracted representative snapshot. (C) Shape factor for the noncircularized discs 1D1, 1E3D1, and 2N2 and of the circularized discs NW9, NW11, and NW13, with the average across frames and replicas shown in blue, and the standard deviation between replicas shown in red.

the most pronounced change occurring in replica 2, where the disc moves further away from the circle and into a more elongated shape, as depicted in Figure 7. These results support the hypothesis that two metastable states exist, where the nanodisc fluctuates between circular and elliptical shapes.

From CD experiments it is known, that the MSP structure increases in helicity to above 80% in the presence of phospholipids.[15] The secondary structure of the MSP was hence calculated for the three replicas. The fraction of helicity in the MSP as a function of the simulation time is plotted in Figure 6 Panel C. All replicas display fluctuations of the helicity fraction in the range between 82 and 90% throughout the simulation, beginning at around 80%, demonstrating a high stability of the two MSPΔH5 proteins.

To assess the hydrophobic mismatch between the MSPs and the lipids, the SASA of the hydrophobic residues in the MSPs

https://doi.org/10.1021/acs.jcim.1c00157
J. Chem. Inf. Model. 2021, 61, 2869−2883

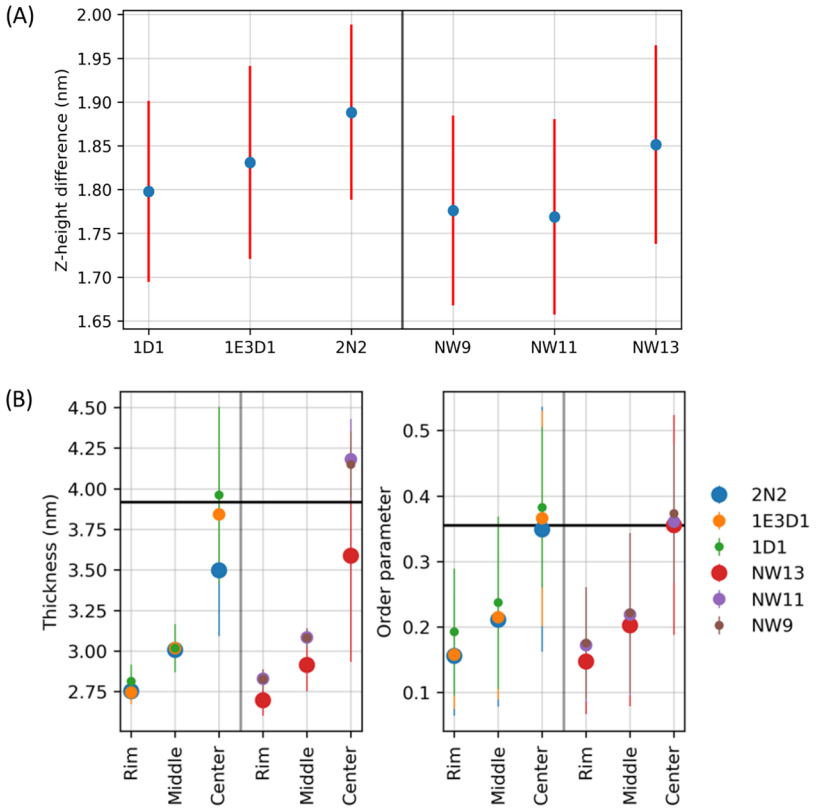


Figure 10. (A) Difference in the height along the *z*-axis from the center-of-mass of the disc and to the rest of the bilayer patch in the ND. The standard deviation between replicas is shown in red. (B) Thickness and order parameters for the three zones (rim, middle, and center) in the NDs. The averages and standard deviations are shown as circles and vertical lines, respectively. The sizes of the circles represent the diameter of the corresponding ND. The averaged order parameters were calculated over the carbon beads in the two tails across frames and lipids. The black horizontal lines are the averages for the simulated CG POPC bilayer across lipids, frames, and replicas.

was calculated. A decrease in SASA indicates that the hydrophobic residues have oriented themselves toward the lipid tails to minimize the hydrophobic mismatch. The SASA of the starting structure was compared to the averaged SASA of the production run for the three replicas. The starting SASA amounts to 220 nm^2 while the averaged SASA values for the three replicas are 195 $\pm$ 2 nm^2, implying a significant decrease of the SASA as part of the equilibration. From these results, it can be concluded that the hydrophobic mismatch between the scaffold proteins and the phospholipid tails is quickly removed during the equilibration and beginning of the production run.

Instead of simply studying the helical propensity in the MSP structure, a more detailed validation of the protein secondary structure is to calculate the chemical shifts and compare with the corresponding measured chemical shifts from the literature.[7] The chemical shifts depend primarily on the type of amino acid and to a minor extent on the secondary structure as well as the environment. The secondary structure chemical shifts were therefore chosen to be the focus for this study.[87] The MD and NMR ensemble of secondary structure chemical shifts for Cα atoms in the MSPΔH5 proteins were obtained by subtracting predicted random coil chemical shifts for the same sequence from, respectively, predicted and recorded chemical shifts. An average across frames and replicas was calculated to represent the MD ensemble. The uncertainty in the NMR ensemble was not taken into account herein.

In Figure 8, the secondary structure chemical shifts for each residue are shown for the MSPΔH5_Model_AA system alongside the experimental data for the Cα atoms. Positive values indicate the presence of a fully formed α-helix. The MSPΔH5_Model_AA ensemble seems, indeed, to reproduce the experimentally measured chemical shifts. We note that a few residues show a deviation from the corresponding experimental chemical shift for the secondary structure. Two sample *t*-tests were therefore performed to assess the *p*-values associated with the differences. These were found to equal 0.072, which is not typically considered statistically significant, in particular not taking into account that no multiple correction was applied for the number of chemical shifts.

To further test whether the MSPΔH5_Model_AA system reproduces the experimental data, a weighted linear regression was performed on the averaged secondary structure chemical shifts from the MSPΔH5_Model_AA system vs experimental secondary structure chemical shifts. The regression was performed with the weights either being the variance between the three MD replicas or the uncertainty of the predictor SPARTA+ (0.92 ppm) plus the variance between the three MD replicas, see Figures S2, S4, and S6 in the SI for the regression plots. The correlation coefficients (R^2-value) obtained from the two regressions were 0.61 and 0.57, respectively. At first glance, this means the MSPΔH5_Model_AA ensemble only reproduces around 60% of the variance in the experimental NMR ensemble. However, this does not include the uncertainty of the experimental data. When studying the residuals for both regressions, all of the secondary structure chemical shifts with the exception of three, are reproduced within an uncertainty of 2 ppm, which is lower than the total noise in the MSPΔH5_Model_AA system (see

https://doi.org/10.1021/acs.jcim.1c00157
J. Chem. Inf. Model. 2021, 61, 2869−2883

Figures S2, S5, and S7 in the SI). This strongly indicates that the unbiased MD simulations indeed do reproduce the experimental values. As a final check, the hypothesis for the slope of the regression being equal to 0 was tested, meaning no correlation. The resulting p-value was $< \varepsilon$, where ε is the computer precision, hence the hypothesis could be rejected. In summary, the simulated protein properties generally agree with the experimentally known protein properties.

Commonly Used Nanodiscs—CG systems. In experimental studies of membrane proteins where the NDT is applied, the most commonly used NDs are the 1D1, 1E3D1, and sometimes the 2N2.[41−45] The naming of these, refer to the different tandems, which are removed or added to the MSP sequence generating three differently sized discs of ∼9.7, 12.9, and 16 nm in diameter. Circularized versions of NDs have subsequently been presented and named NW9,[11] NW11,[11] and NW13,[10] where the naming refers to the approximate diameter of the disc, in this case 9, 11, and 13 nm in diameter, respectively.[11] These circularized discs have shown higher stability compared to the noncircularized versions.[10−12] CG-simulations of the six different discs are performed and properties of each are compared across and with a simulated POPC bilayer to benchmark and validate the building tool.

First, the RMSD of the backbone beads in the MSPs is analyzed. Figure 9 Panel A shows the RMSD for the three replicas for each system during the simulations. For both noncircularized (1D1, 1E3D1, 2N2) and circularized discs (NW9, NW11, NW13), a trend of increasing RMSD values with increasing diameter is observed (p-value $< \varepsilon$ using ANOVA across systems and replicas), which correlates with an increase of the twist in the MSP structure with increasing diameter, see Figure 9 Panel B. An alternative approach to study the twist and dynamics of the discs is to calculate the shape factor. Panel C in Figure 9 shows the averaged shape factor for each system with the standard deviation between replicas shown in red. Not surprisingly, the two small discs 1D1 and NW9 are more circular compared to the larger discs. Interestingly, the medium-sized discs 1E3D1 and NW11 are the least circular discs among the three sizes. However, 1E3D1 and NW13 are approximately the same size, indicating that the circularized discs are indeed more circular compared to the corresponding noncircularized discs. The shape factors are all significantly different across discs, with the exception of NW13 and 2N2 (p-value 0.1) and NW9 and 1D1 (p-value 0.3). See Table S1 in the SI for all p-values.

To further study the shapes sampled for both the circularized and noncircularized discs during the simulations, the structures of the MSPs in the NDs were extracted and clustered. The representative structures for the largest cluster for each disc are shown in Figure 9 Panel B. The smaller-sized discs 1D1 and NW9 are found to be mostly circular and planar, while the larger discs become less circular and visual inspection reveals that more curvature or twist is introduced in the MSP with increasing size. As mentioned, there has been a discussion in the literature whether NDs are generally circular or elliptical shaped.[11,83,88,89] The conclusion seems to be that it depends on several factors with the most important aspect being the number of lipids in the loaded ND.[83,88] For a fully loaded ND the shape will be circular. However, it is experimentally difficult, if even possible, to fully load the NDs.

Following, the lipid bilayer properties were investigated in the different discs. First, an estimate of the curvature of the discs was calculated. Figure 10 Panel A shows the averaged difference in the z-height from the center-of-mass of the ND and to the rest of the lipid headgroups in the disc, as a measure of the curvature. For both circularized and noncircularized discs, the averaged curvature increases with increasing size. Though the trend is only indicative and not significant from a statistical perspective, given the large fluctuations across replicas, it correlates well with the clustering results depicted in Figure 9 Panel B. Hence, the bilayer patch in the ND becomes less planar with increasing size. As observed in Figure 5, zones exist in the ND where the lipid properties such as thickness and order parameters differ when moving from the edge of the disc toward the center. Three zones were defined based on Figure 5 Panel C and referred to as the rim, middle, and center. The rim zone includes all of the lipids within 10 Å of the MSPs, the middle zone contains lipids placed within 10−20 Å of the MSPs, and the center zone includes all of the lipids further away than 20 Å from the MSPs. Figure 10 Panel B shows how the bilayer thickness and packing of the lipids changes from the rim of the disc toward the center for all of the discs. The lipids in the rim zone are perturbed, in the sense of a lower ordering, resulting in a lower thickness. In the center of the discs, there is a trend of increasing order and hence thickness with decreasing size of the disc. In the center, the small- and medium-sized noncircularized (1D1 and 1E3D1) and circularized (NW9 and NW11) discs show a higher order of the lipids compared to a bilayer, resulting in a slightly higher thickness compared to the simulated bilayer. These observations have previously been reported in other MD studies.[16,65] The circularized and noncircularized discs do not seem to be considerably different when comparing the lipid and protein properties. The circularized discs have shown to maintain stability for longer periods of time and over a larger range of temperatures and pH values.[12] However, in the MD simulations the discs are only simulated at a constant temperature and pressure, and hence not stressed similar to the experiments. Furthermore, the force field Martini 2.2 has previously been shown to be too "sticky", with the tendency to over stabilize the protein−protein interactions.[90,91] It can therefore be speculated that the interface between the two MSPs is overstabilized in the Martini 2.2 force field, leading to a too high stability of the discs. This means that the differences between the different discs would only become apparent at e.g., higher temperature or pressure. The forthcoming release of the improved Martini force field (3.0)[92] could remedy this problem.

CONCLUSIONS

We have developed a generic protocol for building molecular models of NDs, where the sequence of the MSPs, the lipid composition, and the interface between the two MSP monomers are user defined. The protocol was first validated using an NMR-derived ND-structure with PDB ID 2N5E and afterwards tested by constructing and simulating commonly used NDs for experimental studies of membrane proteins. Overall, the protocol produces models with structural and lipid properties that are in fine agreement with previously reported computational and experimental studies.[16,27,83,88] We showed that the protocol is not restricted by the size of the ND nor the variant: circularized or noncircularized. Also, modifying the sequence of the MSPs, e.g., introducing more charges as done experimentally,[10] is easy and straightforward with the protocol. The novel component of the protocol is the method for constructing the MSP monomers and assembling the MSP

https://doi.org/10.1021/acs.jcim.1c00157
J. Chem. Inf. Model. 2021, 61, 2869−2883

dimer with any user-defined sequence and registry. This will greatly assist research groups working with the ND technology, where different variants might need to be screened for desirable properties.

In this work, the phospholipids DMPC and POPC were used; however, any single or mixed phospholipid composition can be applied through the Insane.py script[37] though they have not been validated in this work. Other tools can be used for inserting the lipids, such as Packmol.[93] It is furthermore possible to easily modify the protocol and, if desired, not to coarse grain the system for minimization and equilibration. A bash script for converting the ND to atomistic resolution within the Gromacs suite is available at the provided github page. For embedding a membrane protein, the python tools Insane[37] or Alchembed[94] are available, or the Gromacs g_membed tool can be applied.[95]

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.1c00157.

Geometric equations and transformations for the ND builder; MSD plots for the two CG systems (Model + NMR); additional plots for regression and statistical test of the secondary structure chemical shifts; and *p*-values for accessing the significance between the shape factor of different NDs (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Jesper Ferkinghoff-Borg − *Novo Nordisk A/S, 2760 Måløv, Denmark*; Email: jfgb@novonordisk.com

Birgit Schiøtt − *Department of Chemistry, Aarhus University, 8000 Aarhus C, Denmark; Interdisciplinary Nanoscience Center, Aarhus University, 8000 Aarhus C, Denmark;* orcid.org/0000-0001-9937-1562; Email: birgit@chem.au.dk

Authors

Lisbeth R. Kjølbye − *Department of Chemistry, Aarhus University, 8000 Aarhus C, Denmark; Novo Nordisk A/S, 2760 Måløv, Denmark*

Leonardo De Maria − *Novo Nordisk A/S, 2760 Måløv, Denmark;* Present Address: AstraZeneca Gothenburg Research and Early Development, Respiratory & Immunology BioPharmaceuticals R&D AstraZeneca Pepparedsleden1, SE-432 83 Molndal, Sweden.

Tsjerk A. Wassenaar − *Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, 9747 AG Groningen, The Netherlands;* orcid.org/0000-0002-6345-0266

Haleh Abdizadeh − *Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, 9747 AG Groningen, The Netherlands*

Siewert J. Marrink − *Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, 9747 AG Groningen, The Netherlands;* orcid.org/0000-0001-8423-5277

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jcim.1c00157

Notes

The authors declare the following competing financial interest(s): co-author Leonardo de Maria works at AstraZeneca in Gothenburg, Sweden, and has stock ownership and/or stock options or interests in the company.

All of the scripts for the described protocol along with input FASTA sequences for the NDs studied herein are available under a public license at github: https://github.com/LHRK/Nanodisc-Builder. The molecular structures for all of the NDs studied, as well as examples with input and output files along with documentation, are also available on the given github page.

■ ACKNOWLEDGMENTS

We are grateful for the simulation time at CSCAA (Centre for Scientific Computing Aarhus) and to the Innovation Fund Denmark for financial support.

■ ABBREVIATIONS

AA, all-atom; CD, circular dichroism; CG, coarse grained; DMPC, 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine; LB, lipid bilayer; LL, left−left; MD, molecular dynamics; MSD, mean-square displacement; MSP, membrane scaffold protein; ND, nanodisc; NDT, nanodisc technology; NMR, nuclear magnetic resonance; POPC, 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine; RMSD, root-mean-square deviation; SASA, solvent accessible surface area

■ REFERENCES

(1) Denisov, I. G.; Sligar, S. G. Nanodiscs in Membrane Biochemistry and Biophysics. *Chem. Rev.* **2017**, *117*, 4669−4713.

(2) Postis, V.; Rawson, S.; Mitchell, J. K.; Lee, S. C.; Parslow, R. A.; Dafforn, T. R.; Baldwin, S. A.; Muench, S. P. The Use of SMALPs as a Novel Membrane Protein Scaffold for Structure Study by Negative Stain Electron Microscopy. *Biochim. Biophys. Acta, Biomembr.* **2015**, *1848*, 496−501.

(3) Knowles, T. J.; Finka, R.; Smith, C.; Lin, Y. P.; Dafforn, T.; Overduin, M. Membrane Proteins Solubilized Intact in Lipid Containing Nanoparticles Bounded by Styrene Maleic Acid Copolymer. *J. Am. Chem. Soc.* **2009**, *131*, 7484−7485.

(4) Frauenfeld, J.; Löving, R.; Armache, J. P.; Sonnen, A. F. P.; Guettou, F.; Moberg, P.; Zhu, L.; Jegerschöld, C.; Flayhan, A.; Briggs, J. A. G.; Garoff, H.; Löw, C.; Cheng, Y.; Nordlund, P. A Saposin-Lipoprotein Nanoparticle System for Membrane Proteins. *Nat. Methods* **2016**, *13*, 345−351.

(5) McLean, M. A.; Gregory, M. C.; Sligar, S. G. Nanodiscs: A Controlled Bilayer Surface for the Study of Membrane Proteins. *Annu. Rev. Biophys.* **2018**, *47*, 107−124.

(6) Hagn, F.; Nasr, M. L.; Wagner, G. Assembly of Phospholipid Nanodiscs of Controlled Size for Structural Studies of Membrane Proteins by NMR. *Nat. Protoc.* **2018**, *13*, 79−98.

(7) Bibow, S.; Polyhach, Y.; Eichmann, C.; Chi, C. N.; Kowal, J.; Albiez, S.; McLeod, R. A.; Stahlberg, H.; Jeschke, G.; Güntert, P.; Riek, R. Solution Structure of Discoidal High-Density Lipoprotein Particles with a Shortened Apolipoprotein A-I. *Nat. Struct. Mol. Biol.* **2017**, *24*, 187−193.

(8) Kwiterovich, P. O. The Metabolic Pathways of High-Density Lipoprotein, Low-Density Lipoprotein, and Triglycerides: A Current Review. *Am. J. Cardiol.* **2000**, *86*, 5−10.

(9) Grinkova, Y. V.; Denisov, I. G.; Sligar, S. G. Engineering Extended Membrane Scaffold Proteins for Self-Assembly of Soluble Nanoscale Lipid Bilayers. *Protein Eng. Des. Sel.* **2010**, *23*, 843−848.

(10) Johansen, N. T.; Tidemand, F. G.; Nguyen, T. T. T. N.; Rand, K. D.; Pedersen, M. C.; Arleth, L. Circularized and Solubility-Enhanced MSPs Facilitate Simple and High-Yield Production of

Stable Nanodiscs for Studies of Membrane Proteins in Solution. *FEBS J.* **2019**, *286*, 1734−1751.
(11) Nasr, M. L.; Baptista, D.; Strauss, M.; Sun, Z. Y. J.; Grigoriu, S.; Huser, S.; Plückthun, A.; Hagn, F.; Walz, T.; Hogle, J. M.; Wagner, G. Covalently Circularized Nanodiscs for Studying Membrane Proteins and Viral Entry. *Nat. Methods* **2016**, *14*, 49−52.
(12) Nasr, M. L.; Wagner, G. Covalently Circularized Nanodiscs; Challenges and Applications. *Curr. Opin. Struct. Biol.* **2018**, *51*, 129−134.
(13) Borhani, D. W.; Rogers, D. P.; Engler, J. A.; Brouillette, C. G. Crystal Structure of Truncated Human Apolipoprotein A-I Suggests a Lipid-Bound Conformation. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 12291−12296.
(14) Mei, X.; Atkinson, D. Crystal Structure of C-Terminal Truncated Apolipoprotein A-I Reveals the Assembly of High Density Lipoprotein (HDL) by Dimerization. *J. Biol. Chem.* **2011**, *286*, 38570−38582.
(15) Phillips, J. C.; Wriggers, W.; Li, Z.; Jonas, A.; Schulten, K. Predicting the Structure of Apolipoprotein A-I in Reconstituted High-Density Lipoprotein Disks. *Biophys. J.* **1997**, *73*, 2337−2346.
(16) Siuda, I.; Tieleman, D. P. Molecular Models of Nanodiscs. *J. Chem. Theory Comput.* **2015**, *11*, 4923−4932.
(17) Lund-Katz, S.; Phillips, M. C. High Density Lipoprotein Structure-Function and Role in Reverse Cholesterol Transport. *Subcell. Biochem.* **2010**, *51*, 183−227.
(18) Segrest, J. P.; Jones, M. K.; Klon, A. E.; Sheldahl, C. J.; Hellinger, M.; De Loof, H.; Harvey, S. C. A Detailed Molecular Belt Model for Apolipoprotein A-I in Discoidal High Density Lipoprotein. *J. Biol. Chem.* **1999**, *274*, 31755−31758.
(19) Augustyn, B.; Stepien, P.; Poojari, C.; Mobarak, E.; Polit, A.; Wisniewska-Becker, A.; Róg, T. Cholesteryl Hemisuccinate Is Not a Good Replacement for Cholesterol in Lipid Nanodiscs. *J. Phys. Chem. B* **2019**, *123*, 9839−9845.
(20) Shih, A. Y.; Freddolino, P. L.; Arkhipov, A.; Schulten, K. Assembly of Lipoprotein Particles Revealed by Coarse-Grained Molecular Dynamics Simulations. *J. Struct. Biol.* **2007**, *157*, 579−592.
(21) Pan, L.; Segrest, J. P. Computational Studies of Plasma Lipoprotein Lipids. *Biochim. Biophys. Acta, Biomembr.* **2016**, *1858*, 2401−2420.
(22) Catte, A.; Patterson, J. C.; Jones, M. K.; Jerome, W. G.; Bashtovyy, D.; Su, Z.; Gu, F.; Chen, J.; Aliste, M. P.; Harvey, S. C.; Li, L.; Weinstein, G.; Segrest, J. P. Novel Changes in Discoidal High Density Lipoprotein Morphology: A Molecular Dynamics Study. *Biophys. J.* **2006**, *90*, 4345−4360.
(23) Sheldahl, C.; Harvey, S. C. Molecular Dynamics on a Model for Nascent High-Density Lipoprotein: Role of Salt Bridges. *Biophys. J.* **1999**, *76*, 1190−1198.
(24) Klon, A. E.; Segrest, J. P.; Harvey, S. C. Molecular Dynamics Simulations on Discoidal HDL Particles Suggest a Mechanism for Rotation in the Apo A-I Belt Model. *J. Mol. Biol.* **2002**, *324*, 703−721.
(25) Catte, A.; Patterson, J. C.; Bashtovyy, D.; Jones, M. K.; Gu, F.; Li, N.; Rampioni, A.; Sengupta, D.; Vuorela, T.; Niemela, P.; Karttunen, M.; Marrink, S. J.; Vattulainen, I.; Segrest, J. P. Structure of Spheroidal HDL Particles Revealed by Combined Atomistic and Coarse-Grained Simulations. *Biophys. J.* **2008**, *94*, 2306−2319.
(26) Maingi, V.; Rothemund, P. W. K. Properties of DNA- and Protein-Scaffolded Lipid Nanodiscs. *ACS Nano* **2021**, *15*, 751−764.
(27) Debnath, A.; Schäfer, L. V. Structure and Dynamics of Phospholipid Nanodiscs from All-Atom and Coarse-Grained Simulations. *J. Phys. Chem. B* **2015**, *119*, 6991−7002.
(28) Schachter, I.; Allolio, C.; Khelashvili, G.; Harries, D. Confinement in Nanodiscs Anisotropically Modifies Lipid Bilayer Elastic Properties. *J. Phys. Chem. B* **2020**, *124*, 7166−7175.
(29) Shih, A. Y.; Denisov, I. G.; Phillips, J. C.; Sugar, S. G.; Schulten, K. Molecular Dynamics Simulations of Discoidal Bilayers Assembled from Truncated Human Lipoproteins. *Biophys. J.* **2005**, *88*, 548−556.
(30) Vestergaard, M.; Kraft, J. F.; Vosegaard, T.; Thøgersen, L.; Schiøtt, B. Bicelles and Other Membrane Mimics: Comparison of Structure, Properties, and Dynamics from MD Simulations. *J. Phys. Chem. B* **2015**, *119*, 15831−15843.
(31) López, C. A.; Swift, M. F.; Xu, X. P.; Hanein, D.; Volkmann, N.; Gnanakaran, S. Biophysical Characterization of a Nanodisc with and without BAX: An Integrative Study Using Molecular Dynamics Simulations and Cryo-EM. *Structure* **2019**, *27*, 988−999e4.
(32) Vuorela, T.; Catte, A.; Niemelä, P. S.; Hall, A.; Hyvönen, M. T.; Marrink, S. J.; Karttunen, M.; Vattulainen, I. Role of Lipids in Spheroidal High Density Lipoproteins. *PLoS Comput. Biol.* **2010**, *6*, No. e1000964.
(33) Murtola, T.; Vuorela, T. A.; Hyvönen, M. T.; Marrink, S. J.; Karttunen, M.; Vattulainen, I. Low Density Lipoprotein: Structure, Dynamics, and Interactions of ApoB-100 with Lipids. *Soft Matter* **2011**, *7*, 8135−8141.
(34) Xue, M.; Cheng, L.; Faustino, I.; Guo, W.; Marrink, S. J. Molecular Mechanism of Lipid Nanodisk Formation by Styrene-Maleic Acid Copolymers. *Biophys. J.* **2018**, *115*, 494−502.
(35) Qi, Y.; Lee, J.; Klauda, J. B.; Im, W. CHARMM-GUI Nanodisc Builder for Modeling and Simulation of Various Nanodisc Systems. *J. Comput. Chem.* **2019**, *40*, 893−899.
(36) Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: A Web-Based Graphical User Interface for CHARMM. *J. Comput. Chem.* **2008**, *29*, 1859−1865.
(37) Wassenaar, T. A.; Ingólfsson, H. I.; Böckmann, R. A.; Tieleman, D. P.; Marrink, S. J. Computational Lipidomics with Insane: A Versatile Tool for Generating Custom Membranes for Molecular Simulations. *J. Chem. Theory Comput.* **2015**, *11*, 2144−2155.
(38) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, Flexible, and Free. *J. Comput. Chem.* **2005**, *26*, 1701−1718.
(39) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindah, E. Gromacs: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1−2*, 19−25.
(40) Yusuf, Y.; Massiot, J.; Chang, Y. T.; Wu, P. H.; Yeh, V.; Kuo, P. C.; Shiue, J.; Yu, T. Y. Optimization of the Production of Covalently Circularized Nanodiscs and Their Characterization in Physiological Conditions. *Langmuir* **2018**, *34*, 3525−3532.
(41) Hoi, K. K.; Robinson, C. V.; Marty, M. T. Unraveling the Composition and Behavior of Heterogeneous Lipid Nanodiscs by Mass Spectrometry. *Anal. Chem.* **2016**, *88*, 6199−6204.
(42) Marty, M. T.; Hoi, K. K.; Robinson, C. V. Interfacing Membrane Mimetics with Mass Spectrometry. *Acc. Chem. Res.* **2016**, *49*, 2459−2467.
(43) Mitra, N.; Liu, Y.; Liu, J.; Serebryany, E.; Mooney, V.; Devree, B. T.; Sunahara, R. K.; Yan, E. C. Y. Calcium-Dependent Ligand Binding and G-Protein Signaling of Family B GPCR Parathyroid Hormone 1 Receptor Purified in Nanodiscs. *ACS Chem. Biol.* **2013**, *8*, 617−625.
(44) Cai, Y.; Liu, Y.; Culhane, K. J.; DeVree, B. T.; Yang, Y.; Sunahara, R. K.; Yan, E. C. Y. Purification of Family B G Protein-Coupled Receptors Using Nanodiscs: Application to Human Glucagon-like Peptide-1 Receptor. *PLoS One* **2017**, *12*, No. e0179568.
(45) Keener, J. E.; Zambrano, D. E.; Zhang, G.; Zak, C. K.; Reid, D. J.; Deodhar, B. S.; Pemberton, J. E.; Prell, J. S.; Marty, M. T. Chemical Additives Enable Native Mass Spectrometry Measurement of Membrane Protein Oligomeric State within Intact Nanodiscs. *J. Am. Chem. Soc.* **2019**, *141*, 1054−1061.
(46) Nasr, M. L.; Simon, J.; Zhao, Z.; Strauss, M.; Shih, W.; Hogle, J.; Wagner, G. Covalently Circularized Nanodiscs: EM and NMR Applications. *Biophys. J.* **2017**, *112*, No. 305a.
(47) Yeh, V.; Lee, T. Y.; Chen, C. W.; Kuo, P. C.; Shiue, J.; Chu, L. K.; Yu, T. Y. Highly Efficient Transfer of 7TM Membrane Protein from Native Membrane to Covalently Circularized Nanodisc. *Sci. Rep.* **2018**, *8*, No. 13501.
(48) Lipman, D. J.; Pearson, W. R. Rapid and Sensitive Protein Similarity Searches. *Science* **1985**, *227*, 1435−1441.

https://doi.org/10.1021/acs.jcim.1c00157
J. Chem. Inf. Model. 2021, 61, 2869−2883

(49) Chen, C.; Shutan, X.; Lincong, W. An Algorithm for Protein Helix Assignment Using Helix Geometry. *PLoS One* **2015**, *10*, No. e0129674.
(50) Berg, M. J.; Tymoczko, L.; John, S. L. *Biochemistry*, 7th ed.; Kate Ahr Parker, 2012.
(51) Tien, M. Z.; Sydykova, D. K.; Meyer, A. G.; Wilke, C. O. Peptidebuilder: A Simple Python Library to Generate Model Peptides. *PeerJ* **2013**, *2013*, No. e80.
(52) De Jong, D. H.; Singh, G.; Bennett, W. F. D.; Arnarez, C.; Wassenaar, T. A.; Schäfer, L. V.; Periole, X.; Tieleman, D. P.; Marrink, S. J. Improved Parameters for the Martini Coarse-Grained Protein Force Field. *J. Chem. Theory Comput.* **2013**, *9*, 687−697.
(53) Wassenaar, T. A.; Pluhackova, K.; Böckmann, R. A.; Marrink, S. J.; Tieleman, D. P. Going Backward: A Flexible Geometric Approach to Reverse Transformation from Coarse Grained to Atomistic Models. *J. Chem. Theory Comput.* **2014**, *10*, 676−690.
(54) Kabsch, W.; Sander, C. Dictionary of Protein Secondary Structure: Pattern Recognition of Hydrogen-bonded and Geometrical Features. *Biopolymers* **1983**, *22*, 2577−2637.
(55) Touw, W. G.; Baakman, C.; Black, J.; Te Beek, T. A. H.; Krieger, E.; Joosten, R. P.; Vriend, G. A Series of PDB-Related Databanks for Everyday Needs. *Nucleic Acids Res.* **2015**, *43*, D364−D368.
(56) Denisov, I. G.; Grinkova, Y. V.; Lazarides, A. A.; Sligar, S. G. Directed Self-Assembly of Monodisperse Phospholipid Bilayer Nanodiscs with Controlled Size. *J. Am. Chem. Soc.* **2004**, *126*, 3477−3487.
(57) Bayburt, T. H.; Grinkova, Y. V.; Sligar, S. G. Self-Assembly of Discoidal Phospholipid Bilayer Nanoparticles with Membrane Scaffold Proteins. *Nano Lett.* **2002**, *2*, 853−856.
(58) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; De Vries, A. H. The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B* **2007**, *111*, 7812−7824.
(59) Berendsen, H. J. C.; Postma, J. P. M.; Van Gunsteren, W. F.; Dinola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(60) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126*, No. 014101.
(61) Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **1981**, *52*, 7182−7190.
(62) Bogusz, S.; Cheatham, T. E.; Brooks, B. R. Removal of Pressure and Free Energy Artifacts in Charged Periodic Systems via Net Charge Corrections to the Ewald Potential. *J. Chem. Phys.* **1998**, *108*, 7070−7084.
(63) De Jong, D. H.; Baoukina, S.; Ingólfsson, H. I.; Marrink, S. J. Martini Straight: Boosting Performance Using a Shorter Cutoff and GPUs. *Comput. Phys. Commun.* **2016**, *199*, 1−7.
(64) Páll, S.; Hess, B. A Flexible Algorithm for Calculating Pair Interactions on SIMD Architectures. *Comput. Phys. Commun.* **2013**, *184*, 2641−2650.
(65) Lee, J.; Patel, D. S.; Ståhle, J.; Park, S.-J.; Kern, N. R.; Kim, S.; Lee, J.; Cheng, X.; Valvano, M. A.; Holst, O.; Knirel, Y. A.; Qi, Y.; Jo, S.; Klauda, J. B.; Widmalm, G.; Im, W. CHARMM-GUI Membrane Builder for Complex Biological Membrane Simulations with Glycolipids and Lipoglycans. *J. Chem. Theory Comput.* **2019**, *15*, 775−786.
(66) Lee, J.; Cheng, X.; Swails, J. M.; Yeom, M. S.; Eastman, P. K.; Lemkul, J. A.; Wei, S.; Buckner, J.; Jeong, J. C.; Qi, Y.; Jo, S.; Pande, V. S.; Case, D. A.; Brooks, C. L.; MacKerell, A. D.; Klauda, J. B.; Im, W. CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J. Chem. Theory Comput.* **2016**, *12*, 405−413.
(67) Allouche, A. Software News and Updates Gabedit—A Graphical User Interface for Computational Chemistry Softwares. *J. Comput. Chem.* **2012**, *32*, 174−182.
(68) Qi, Y.; Ingólfsson, H. I.; Cheng, X.; Lee, J.; Marrink, S. J.; Im, W. CHARMM-GUI Martini Maker for Coarse-Grained Simulations with the Martini Force Field. *J. Chem. Theory Comput.* **2015**, *11*, 4486−4494.
(69) Huang, J.; Mackerell, A. D. CHARMM36 All-Atom Additive Protein Force Field: Validation Based on Comparison to NMR Data. *J. Comput. Chem.* **2013**, *34*, 2135−2145.
(70) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone φ, ψ and Side-Chain X1 and X2 Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257−3273.
(71) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N·log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(72) Shen, Y.; Bax, A. SPARTA+: A Modest Improvement in Empirical NMR Chemical Shift Prediction by Means of an Artificial Neural Network. *J. Biomol. NMR* **2010**, *48*, 13−22.
(73) McGibbon, R. T.; Beauchamp, K. A.; Harrigan, M. P.; Klein, C.; Swails, J. M.; Hernández, C. X.; Schwantes, C. R.; Wang, L. P.; Lane, T. J.; Pande, V. S. MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. *Biophys. J.* **2015**, *109*, 1528−1532.
(74) Piggot, T. J.; Allison, J. R.; Sessions, R. B.; Essex, J. W. On the Calculation of Acyl Chain Order Parameters from Lipid Simulations. *J. Chem. Theory Comput* **2017**, *13*, 5683−5696.
(75) Castillo, N.; Monticelli, L.; Barnoud, J.; Tieleman, D. P. Free Energy of WALP23 Dimer Association in DMPC, DPPC, and DOPC Bilayers. *Chem. Phys. Lipids* **2013**, *169*, 95−105.
(76) Bickel, T. A Note on Confined Diffusion. *Phys. A* **2007**, *377*, 24−32.
(77) Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; van der Walt, S. J.; Brett, M.; Wilson, J.; Millman, K. J.; Mayorov, N.; Nelson, A. R. J.; Jones, E.; Kern, R.; Larson, E.; Carey, C. J.; Polat, İ.; Feng, Y.; Moore, E. W.; VanderPlas, J.; Laxalde, D.; Perktold, J.; Cimrman, R.; Henriksen, I.; Quintero, E. A.; Harris, C. R.; Archibald, A. M.; Ribeiro, A. H.; Pedregosa, F.; van Mulbregt, P.; Vijaykumar, A.; Bardelli, A.; Pietro; Rothberg, A.; Hilboll, A.; Kloeckner, A.; Scopatz, A.; Lee, A.; Rokem, A.; Woods, C. N.; Fulton, C.; Masson, C.; Häggström, C.; Fitzgerald, C.; Nicholson, D. A.; Hagen, D. R.; Pasechnik, D. V.; Olivetti, E.; Martin, E.; Wieser, E.; Silva, F.; Lenders, F.; Wilhelm, F.; Young, G.; Price, G. A.; Ingold, G. L.; Allen, G. E.; Lee, G. R.; Audren, H.; Probst, I.; Dietrich, J. P.; Silterra, J.; Webber, J. T.; Slavič, J.; Nothman, J.; Buchner, J.; Kulick, J.; Schönberger, J. L.; de Miranda Cardoso, J. V.; Reimer, J.; Harrington, J.; Rodríguez, J. L. C.; Nunez-Iglesias, J.; Kuczynski, J.; Tritz, K.; Thoma, M.; Newville, M.; Kümmerer, M.; Bolingbroke, M.; Tartre, M.; Pak, M.; Smith, N. J.; Nowaczyk, N.; Shebanov, N.; Pavlyk, O.; Brodtkorb, P. A.; Lee, P.; McGibbon, R. T.; Feldbauer, R.; Lewis, S.; Tygier, S.; Sievert, S.; Vigna, S.; Peterson, S.; More, S.; Pudlik, T.; Oshima, T.; Pingel, T. J.; Robitaille, T. P.; Spura, T.; Jones, T. R.; Cera, T.; Leslie, T.; Zito, T.; Krauss, T.; Upadhyay, U.; Halchenko, Y. O.; Vázquez-Baeza, Y. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **2020**, *17*, 261−272.
(78) Abraham, M. J.; van der Spoel, D.; Lindahl, E.; Hess, B. *GROMACS User Manual Version 2018*; Royal Institue of Technology and Uppsala University, 2018.
(79) Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. Peptide Folding: When Simulation Meets Experiment. *Angew. Chem., Int. Ed.* **1999**, *38*, 236−240.
(80) Gowers, R.; Linke, M.; Barnoud, J.; Reddy, T.; Melo, M.; Seyler, S.; Domański, J.; Dotson, D.; Buchoux, S.; Kenney, I.; Beckstein, O. MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations, In *Proceedings of the 15th Python in Science Conference*,, 2016; pp 98−105.
(81) Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **2007**, *9*, 90−95.
(82) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14*, 33−38.

https://doi.org/10.1021/acs.jcim.1c00157
J. Chem. Inf. Model. 2021, 61, 2869−2883

(83) Skar-Gislinge, N.; Simonsen, J. B.; Mortensen, K.; Feidenhans'l, R.; Sligar, S. G.; Lindberg Møller, B.; Bjørnholm, T.; Arleth, L. Elliptical Structure of Phospholipid Bilayer Nanodiscs Encapsulated by Scaffold Proteins: Casting the Roles of the Lipids and the Protein. *J. Am. Chem. Soc.* **2010**, *132*, 13713−13722.
(84) Midtgaard, S. R.; Pedersen, M. C.; Arleth, L. Small-Angle X-Ray Scattering of the Cholesterol Incorporation into Human ApoA1-POPC Discoidal Particles. *Biophys. J.* **2015**, *109*, 308−318.
(85) Bengtsen, T.; Holm, V. L.; Kjølbye, L. R.; Midtgaard, S. R.; Johansen, N. T.; Tesei, G.; Bottaro, S.; Schiøtt, B.; Arleth, L.; Lindorff-Larsen, K. Structure and Dynamics of a Nanodisc by Integrating NMR, SAXS and SANS Experiments with Molecular Dynamics Simulations. *eLife* **2020**, *9*, No. e56518.
(86) Stepien, P.; Augustyn, B.; Poojari, C.; Galan, W.; Polit, A.; Vattulainen, I.; Wisnieska-Becker, A.; Rog, T. Complexity of Seemingly Simple Lipid Nanodiscs. *Biochim. Biophys. Acta, Biomembr.* **2020**, *1862*, No. 183420.
(87) Mielke, S. P.; Krishnan, V. V. Characterization of Protein Secondary Structure from NMR Chemical Shifts. *Prog. Nucl. Magn. Reson. Spectrosc.* **2009**, *54*, 141−165.
(88) Skar-Gislinge, N.; Johansen, N. T.; Høiberg-Nielsen, R.; Arleth, L. Comprehensive Study of the Self-Assembly of Phospholipid Nanodiscs: What Determines Their Shape and Stoichiometry? *Langmuir* **2018**, *34*, 12569−12582.
(89) Skar-Gislinge, N.; Arleth, L. Small-Angle Scattering from Phospholipid Nanodiscs: Derivation and Refinement of a Molecular Constrained Analytical Model Form Factor. *Phys. Chem. Chem. Phys.* **2011**, *13*, 3161−3170.
(90) Alessandri, R.; Souza, P. C. T.; Thallmair, S.; Melo, M. N.; De Vries, A. H.; Marrink, S. J. Pitfalls of the Martini Model. *J. Chem. Theory Comput.* **2019**, *15*, 5448−5460.
(91) Javanainen, M.; Martinez-Seara, H.; Vattulainen, I. Excessive Aggregation of Membrane Proteins in the Martini Model. *PLoS One* **2017**, *12*, No. e0187936.
(92) Souza, P. C.; Alessandri, R.; Jonathan Barnoud, S. T.; Faustino, I.; Grunewald, F.; Patmanidis, I.; Abdizadeh, H.; Bruininks, B. M..; Wassenaar, T. A.; Kroon, P. C.; Josef Melcr, V. N.; Corradi4, V.; Khan, H. M.; Domański, J.; Matti Javanainen, H.; Martinez-Seara, H.; Reuter, N.; Best, R. B.; Ilpo Vattulainen, L. M.; Periole, X.; Peter Tieleman, D.; de Vries, A. H.; Marrink, S. J. Martini 3.0: A General Purpose Force Field for Coarse-Grained Molecular Dynamics. *Nat. Methods* **2021**, *18*, 382−388.
(93) Martínez, L.; Andrade, R.; Birgin, E. G.; Martínez, J. M. PACKMOL: A Package for Building Initial Configurations for Molecular Dynamics Simulations. *J. Comput. Chem.* **2009**, *30*, 2157−2164.
(94) Jefferys, E.; Sands, Z. A.; Shi, J.; Sansom, M. S. P.; Fowler, P. W. Alchembed: A Computational Method for Incorporating Multiple Proteins into Complex Lipid Geometries. *J. Chem. Theory Comput.* **2015**, *11*, 2743−2754.
(95) Wolf, M. G.; Hoefling, M.; Aponte-Santamarí, C.; Grubmüller, H.; Groenhof, G. G-Membed: Efficient Insertion of a Membrane Protein into an Equilibrated Lipid Bilayer with Minimal Perturbation. *J. Comput. Chem.* **2010**, *31*, 2169−2174.

https://doi.org/10.1021/acs.jcim.1c00157
J. Chem. Inf. Model. 2021, 61, 2869−2883