# Testing of parameters for a biologically accurate brain membrane and molecular dynamics simulations exploration of membrane interactions and conformational changes exhibited by p110α and its oncogenic mutants

**William A. Irvine**

A dissertation submitted in fulfilment of the requirements for the degree of Masters of Science in Biochemistry.

Massey University

March 2014

**Abstract**

Phosphatidylinositide 3-kinases are a family of enzymes which are involved in the regulation of cell growth and proliferation via signalling pathways. This, in turn, means they are linked with cancer development through mutations borne by the genes which encode them. One of these oncogenes, PIK3CA, encodes the catalytic subunit of p110α. This study will focus on p110α's interaction with a phospholipid bilayer, using computational techniques, in an effort to better understand this protein and the effect the cancer-related mutations have on its activity.

In order to model the phospholipid bilayer in a biologically and physiologically accurate manner, with all key components present in their correct proportions, model parameters for the components had to be produced and tested in small binary systems. The components of the membrane used include the phospholipids POPC, POPE, POPS and $PIP_2$, as well as sphingomyelin and cholesterol.

Using these new parameters for the components of a phospholipid bilayer, molecular dynamics simulations were run of the activated p110α subunit and two of its oncogenic mutants (E545K, H1047R) in the presence of a realistic brain lipid membrane. The results will pave the way to the development of drugs which will serve to inhibit the pathway when necessary, in an effort to control and reduce the incidence of cancer.

**Acknowledgements**

First and foremost, I would like to thank my mentor and supervisor, Dr. Jane Allison, for granting me the opportunity to carry out this research and guiding me every step of the way. This project would also not have been possible without the help of Dr. Jack Flanagan, who offered priceless advice with regards to the biochemistry side of things. On a similar note, I'd like to offer my gratitude to Dr. Thomas Huber who was a major reference point when it came to the computational side of things.

Finally, I am of course indebted to my family and friends, who have been my primary source of motivation and inspiration over the years.

Table of Contents

## 1. Introduction

### 1.1 PI3K

Phosphatidylinositide 3-kinases (PI3K) are proteins that participate in signalling pathways regulating factors such as cell growth, proliferation, and survival (Williams et al.)[i] (Figure 1). These signalling pathways are triggered by the phosphorylation of the hydroxyl group at the 3-position of the inositol ring of phosphatidylinositols (PIs), hence the name given to the enzymes involved (Foster et al.)[ii].

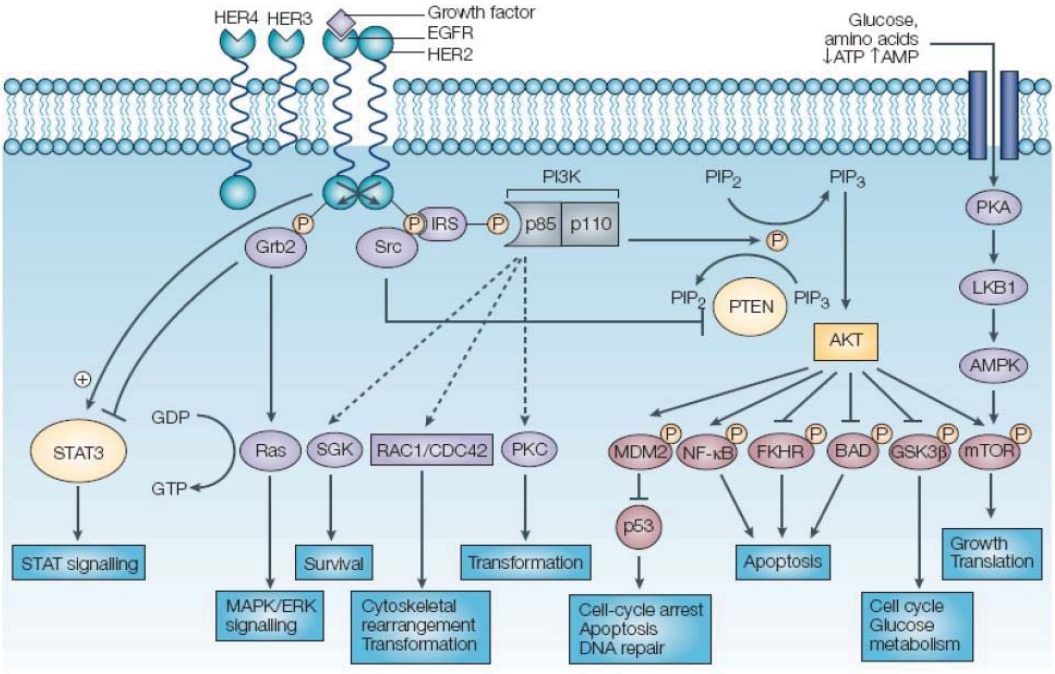


**Figure 1**: *Chart illustrating the complete PI3K signalling pathway* (*Cully et al.*)[iii]*. PI3K is shown in grey, and the ultimate downstream effects of PI3K signalling are shown in cyan boxes.*

Considering the factors PIs regulate (Figure 1), there is little surprise that PI3Ks are implicated in cancer development (Cleary et al.)[iv], and that the genes encoding PI3Ks have been identified as oncogenes (Arcaro et al.)[v].

One of these, PIK3CA, encodes the catalytic subunit of the alpha isoform of PI3K (p110α) (Chomczyk et al.)[vi], the protein that is a key focus of this study. Two key oncogenic mutations observed for PI3Ka include E545K and H1047R (Kang et al.)[vii]. Point mutations at residues 545 (Lysine replaces Glutamic Acid) and 1047 (Arginine replaces Histidine) both increase p110α activity via different methods (Bader et al.)[viii], resulting in enhanced downstream signalling, as outlined below. The locations of these two mutations in the three-dimensional structure of p11oα are highlighted and enlarged in Figure 2.

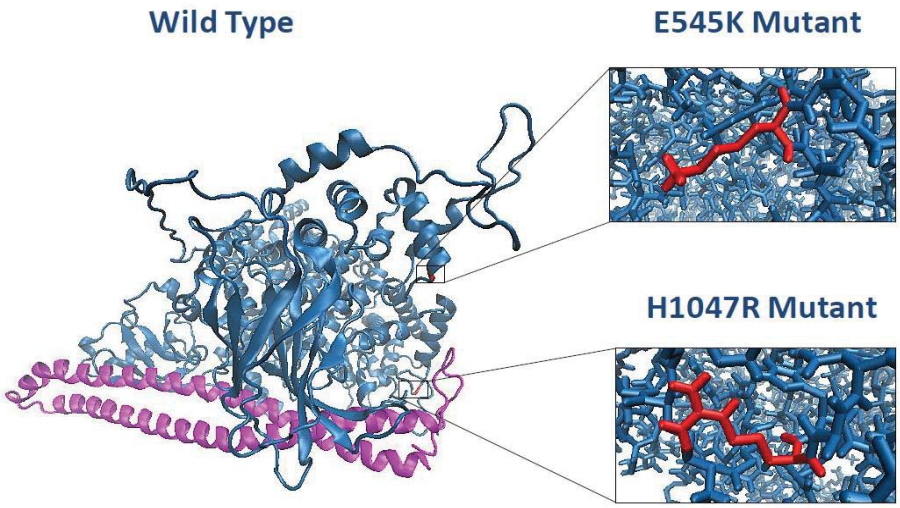


**Figure 2**: *Image showing wild type p110α (blue) bound to the iSH2 domain of p85α (purple) and the location of each mutation (red).*

The p110α subunit consists of an adaptor-binding domain (ABD), Ras-binding domain (RBD), C2 domain, helical domain and kinase domain (Amzel et al.)[ix] (Figure 3). This subunit normally works in tandem with the p85α subunit, which controls p110α activity (Gabelli et al.)[x], and consists of an Src homology (SH)3 domain, a Bcl-2 homology (BH) domain, and three different SH2 domains found on the N-terminus side, C-terminus side and a coiled coil in between both, called the interSH2 (iSH2) domain (Huang et al.)[xi] (Figure 3). In its most active state, the ABD of the p110α subunit is bound with the iSH2 domain of the p85α subunit (Backer)[xii] (Figure 2).

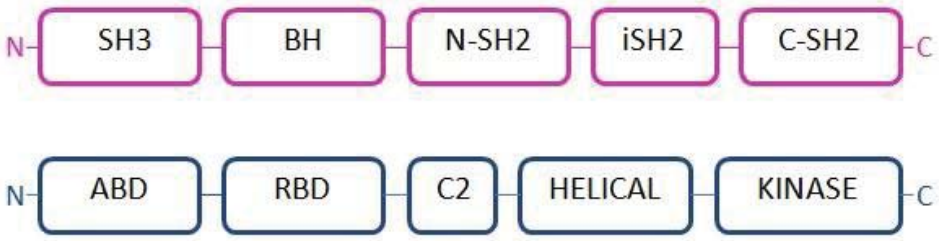


**Figure 3**: *A linear representation of p110α (blue) and p85α (purple), showing the order of domains.*

The E545K mutation occurs in the helical domain of p110α and is thought to enhance activity by initiating a conformational change, which results in the enzyme no longer requiring the p85α subunit for membrane binding (Hao et al.)[xiii]. The H1047R mutation occurs in the C-terminal lobe of the kinase domain and appears to enhance activity by initiating a conformational shift in specific kinase domain loops (residues 864-874 and 1050-1062), which in turn circumvents membrane-binding steps involving the RBD (Mandelker at al.)[xiv], accelerating the activation process. Previous studies have shown that these two mutations act independently (Burke et al.)[xv]. This is suggested by the synergy of both mutations, as well as structural and functional studies (Zhao et al.)[xvi].

Although these experiments have given us some insight, they are unable to provide a detailed view of the structural changes taking place. To better understand the effect that cancer-related mutations have on p110α activity, the aim of this study was to characterise the structural changes that take place upon interaction of wild type and oncogenic mutants of p110α, activated by binding of the iSH2 domain of p85α, with the membrane using a computational technique known as molecular dynamics (MD).

### 1.2 Molecular Dynamics

MD involves the study and observation of molecules through computer simulations. A system comprising many different molecules can be simulated, once they have been parameterised. For biomolecular systems such as those studied here, atomic-level models in which each atom is represented explicitly are typically used.

The atoms interact with each other according to rules laid out in a force field. This contains parameters that define each atom type, in terms of their size, mass and charge. Additionally, it contains potential functions describing any pre-defined interactions between pairs (bond lengths), triplets (bond angles) or quartets (dihedral angles) of atoms joined by one or multiple bonds. These interactions also include the so-called "van der Waals" and electrostatic interactions between pairs of non-bonded atoms. These interaction functions and their parameters have been carefully calibrated over the years, by comparing quantities calculated from MD simulations and quantities measured experimentally.

Each atom of each molecule is assigned *xyz* coordinates describing its position in three-dimensional space, and initial velocities are drawn from a Maxwell-Boltzmann distribution. The behaviour of the molecules is then followed over the course of the simulation by numerically integrating Newton's equations of motion for all atoms of the system, according to the scheme shown in Figure 4.

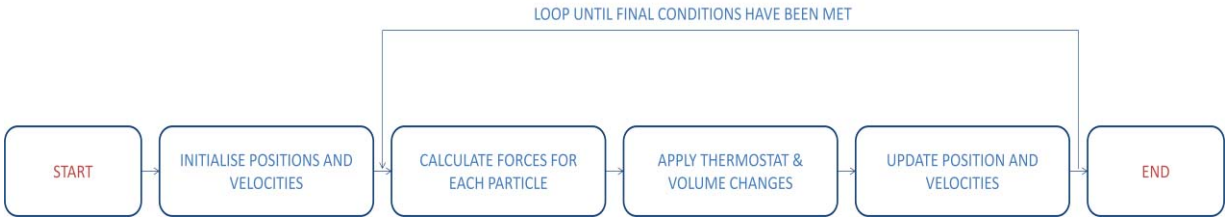


**Figure 4**: *Flow chart depicting the order of operations that take place during a molecular dynamics simulation. Final conditions could refer to system convergence or a point in the simulation at which the process of interest has been completed.*

While parameters for most commonly simulated biomolecules already exist, this study aimed to investigate the interaction of p110α with a brain membrane to mimic previous experiments carried out by Burke et al.[xvii] As such a complex membrane system had not been simulated before, parameters for the different types of lipid found in a brain membrane first had to be assembled and tested.

## 1.3 Overview of this study

This study features the first ever simulation of a phospholipid bilayer containing all the components found in the biological form of the membrane. Specifically, the focus was on mimicking the composition of brain lipid membranes, in which the phospholipids phosphatidylethanolamine (PE), phosphatidylcholine (PC) and phosphatidylserine (PS) are found along with sphingomyelin (SGML) and cholesterol (CHOL). The parameters for palmitoyl oleoyl (PO)PE and POPC gathered from previous studies were combined with parameters for POPS, SGML and CHOL determined here by analogy to previously parameterised similar molecules. The compatibility and appropriateness of the parameters were tested by carrying out simulations of two-component systems for which quantitative experimental data are available. Finally, a lipid bilayer representative of a brain membrane was constructed and simulated. The composition of the bilayer was matched to those used experimentally in terms of the relative proportion of each type of phospholipid and the presence of a lipid raft in the form of an aggregation of sphingomyelin and cholesterol molecules.

The behaviour of this realistic brain phospholipid bilayer was assessed by calculating a variety of different properties and, where possible, comparing these with experimentally measured values. Seeing very good agreement combined with the lack of anomalies during the simulation, suggested that the force field parameters were viable for simulations of more complex systems.

Simulations were henceforth run involving the wild type or oncogenic mutants of activated p110α along with a patch of brain membrane, with the addition of phosphatidylinositol 4,5-biphosphate (PIP2), each in separate equilibrated, solvated boxes. The proteins were also simulated with a more commonly used pure DPPC membrane as a control. The approach of p110α to the membrane was compared between the simulations in terms of time and direction of approach, as well as the part of the protein that interacted with the membrane. This was done in order to elucidate any differences between the wild type enzyme and its mutants, and the changes due to the presence of either bilayer. Additionally, the solvent accessible surface area (SASA) calculated from the MD simulations was compared to Hydrogen-Deuterium exchange (HDx) rates measured experimentally by Burke et al.[xvii] In this way, the simulations were validated and atomic-level insight into the structures and dynamics giving rise to experimental measurements was obtained.

## 2. Methodology

All simulations were prepared, performed and analysed using the GROMACS molecular dynamics package and its associated tools (Pronk et al.)[xviii]. The parameters described are based on the GROMOS 54A7 force-field (Schmid et al.)[xix], slightly modified for this study as described below.

### 2.1 Building the Brain Membrane

#### 2.1.1 Parameter Design

Prior to building the Brain Membrane, coordinates and parameters for each individual lipid needed to be defined. The lipids used were POPE, POPC, POPS, PIP2, SGML and CHOL (Figure 5).

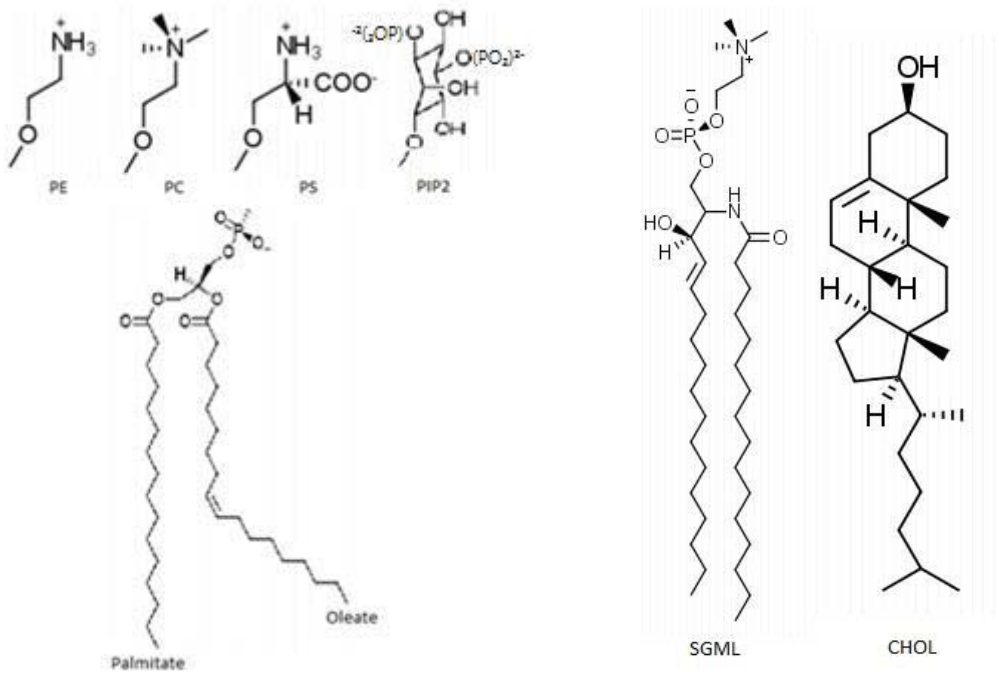


**Figure 5**: *Image showing the full structures (headgroups and tails) of each individual lipid used in the Brain Membrane. The various phospholipids are formed by joining the different headgroups (PE, PC, PS, PIP2), shown on the top left, to the phosphate group connected to the glycerol backbone to which the palmitoyl oleoyl tail is also attached (bottom left).*

**Coordinates**

While coordinates could be constructed manually based on known bond lengths, preferred bond and torsion angle values and steric considerations, it is preferable to use either experimentally determined coordinates, such as those obtained by X-ray crystallography, or coordinates generated by equilibrated simulation in realistic conditions and using a good quality force field.

Coordinates for single molecules can be extracted from the coordinate files of larger systems containing that molecule. As such, the coordinates for POPE, POPC and CHOL were taken from a membrane of those components, simulated for 200 ns (Wennberg et al.)[xx].

The coordinates for POPS were manually generated by superimposing the coordinates of POPE, from above, onto the coordinates of a dioleoylphosphatidylserine (DOPS) molecule (Jämbeck et al.)[xxi]. Once superimposed, the third hydrogen on the second carbon of POPE was replaced by the carboxyl group of DOPS.

The coordinates for PIP2 were manually generated using a combination of the coordinates of a POPC molecule and those of a DPP2 molecule, obtained from a DPPC-DPP2 membrane simulated for 200ns (Holdbrook et al.)[xxii]. The two coordinates of both lipids were superimposed, using the tails from POPC and the headgroup of DPP2.

The coordinates for SGML were taken from a bilayer of 100 SGML molecules equilibrated over the course of a 50 ns simulation (Chiu et al.)[xxiii].

**Parameters**

The parameters describing the atoms comprising each lipid and how they interact, including a description of the bonded structure of the molecule, are defined by the force field. A force field comprises a set of atom types and associated 'bonded' and 'non-bonded' interactions between them. Force fields typically include 'molecular building blocks' that define the set of atom types and interactions required to describe a specific molecule. In the GROMACS software, it is also possible to create so-called "itp" files, effectively molecular building blocks that exist in addition to the standard building blocks distributed with the force field. When doing so, it is preferable to use the existing atom and interaction types included in the force field, as these have already been subjected to careful parameterisation and testing.

While this was for the most part possible here, in certain cases new parameters/interaction types had to be defined. Lists of the new interaction types and the itp files are available in the Appendix.

The GROMOS 54A7 force field that was used here has only one lipid building block, that of DPPC. Additionally, an itp file for POPC was available from Gromacs' user contributions, which could be used directly without any changes. However, itp files had to be constructed for the remainder of the lipids.

For POPE, constructing the itp file was straightforward as it was virtually identical to that of POPC, with simple hydrogen-nitrogen interactions replacing the methyl-nitrogen interactions found in the headgroup. This involved changing all affected parameters- bond lengths, angles and dihedral angles. Refer to interactions involving atoms 1 to 4 in the POPE itp file, given in the appendix.

For POPS, constructing the itp file consisted simply of adding in the presence of the new carboxyl group and redefining the bond lengths, angles, and dihedrals for the affected interactions. Wherever there had been an interaction involving the third hydrogen on the second carbon, the values were changed to represent a carbon from a carboxyl group. Refer to interactions involving atoms 53 to 55 in the POPS itp file, given in the appendix.

For PIP2, constructing the itp file required the listing of all atoms found in the coordinate file and defining the interactions using parameters which already existed in the force field files. The key interaction additions in comparison to previously mentioned itp files included those of the inositol ring and the additional phosphate groups. These were chosen based on similarity to other molecules for which molecular building blocks already existed. Refer to interactions involving atoms 49 to 68 in the PIP2 itp file, given in the appendix.

For SGML, constructing the itp file from scratch consisted of listing parameters defining all the atoms found in the coordinate file and their interactions. As done for PIP2, these were chosen based on similarity to other molecules for which molecular building blocks already existed. Refer to the SGML itp file, given in the appendix.

For CHOL, new interactions needed to be added to the force field to describe the presence of carbon in a non-aromatic ring. The added parameters included the ideal bond lengths and bond angles for single and double bonded carbon-carbon interactions present in 5-membered and 6-membered hydrocarbon rings. A starting point for the values was obtained from the modified 43A1-S3 force field specially developed for simulating lipids (Pandit et al.)[xxiv]. The necessary values were adopted and added to the lists of bond length and bond angle parameters of the 54A7 force field. CHOL itp file, given in the appendix, as well as a list of any alterations made to the force field files.

### 2.1.2 Parameter Testing

Each individual lipid was subjected to 1000 steps of the steepest descent energy minimisation in the modified 54A7 force-field[xix]. To test the parameters, systems comprising pairs of lipid types were built and subjected to 10 ns MD simulations at 298K to ensure properties of the membrane components matched previous experimental and computational data. Two binary systems were tested: POPC/CHOL and a lipid raft (SGML/CHOL). These particular systems were built by randomly inserting each lipid component into a bilayer, located in a solvated box. These membrane systems were chosen, as previous simulation and experimental data exist for comparison.

A Brain Membrane bilayer system mimicking that used by Burke et al.[xvii] was then constructed and simulated. A small leaflet was built from the energy minimised lipids to correspond with specific composition matching a brain membrane (60% POPE, 20% POPC, 20% POPS). This leaflet was mirrored to form a bilayer before being multiplied with to form two larger rectangular sections of different dimensions (Figure 6). The same procedure was followed to build a lipid raft comprising 66% sphingomyelin and 34% cholesterol. The final system (51% POPE, 15% POPC, 15% POPS, 13% SGML, 6% CHOL) was constructed by placing the lipid raft in the centre of a large box, and surrounding it with the rectangular sections of phospholipids as shown in Figure 6 to form a box of size 25.337 x 21.114 x 6.757 nm.

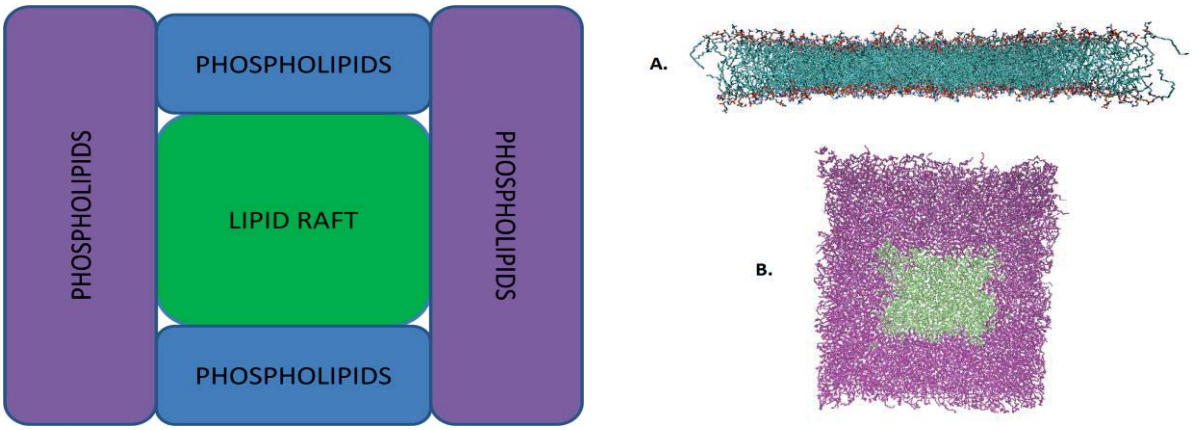


**Figure 6**: *Diagram on the left showing the arrangement of the different components of the Brain Membrane: (purple and blue) mixed phospholipids (60% POPE, 20% POPC, 20% POPS) and (green) lipid raft (66% SGML, 34% CHOL). A snapshot of the complete Brain Membrane after 100 ns of MD simulation is included on the right. Image A shows the bilayer from the side, clearly showing the headgroups and lipid tails. Image B shows the bilayer from above, with the phospholipids and lipid raft in the same colour arrangement as the diagram. Water molecules were removed for clarity.*

Each complete lipid bilayer system was subjected to 1000 steps of steepest descent energy minimisation in the modified 54A7 force field. It was then solvated in a rectangular box, temporarily increasing the van der Waals radius for carbon to prevent insertion of solvent molecules into the lipid bilayer. The simple point charge (SPC) (Berendsen et al.)[xxv] water model was used (spc216), in the form of pre-equilibrated blocks of 216 water molecules, and periodic boundary conditions (PBC) were applied. In the case of the Brain Membrane, the system was neutralised by addition of 240 cations ($Na^+$) to give a final composition of 800 POPE, 240 POPC, 240 POPS, 198 sphingomyelin, 90 cholesterol, 58995 solvent and 240 sodium molecules.
The simulation was initiated with the following equilibration scheme. Firstly, the initial velocities were randomly generated from a Maxwell-Boltzmann distribution at 50 K. The system was then heated to 300 K over the course of 100 ps in the NVT ensemble. The temperature was controlled using the Berendsen thermostat (Berendsen et al.)[xxvi] with a temperature coupling constant ($\tau_T$) of 0.1 ps. The system was further equilibrated for 400 ps at 300 K in the NPT ensemble. The pressure was controlled using the Berendsen barostat[xxvi] with a pressure coupling constant ($\tau_p$) of 0.5 ps and an isothermal compressibility of $4.5 \times 10^{-5}$ $bar^{-1}$.

The final structure was used as the starting configuration for a 100 ns production run at 300 K, with structures saved every 200 ps. The LINCS algorithm (Hess et al.)[xxvii] was used with an order of 4 to constrain bond lengths and water bond angles, allowing for an integration time step of 2 fs. The centre of mass motion was removed every 100 ps. Non-bonded interactions were calculated using a grid cut-off scheme. The non-bonded interactions within a cut-off distance of 0.9 nm were calculated at every step from a pair list that was updated every fifth time step. At this point, interactions between atoms within 1.4 nm were also calculated and were kept constant between updates. Electrostatic interactions were calculated using particle mesh Ewald (PME) summation (Essmann et al.)[xxviii], with a cut-off distance of 0.9 nm. The system following simulation is also shown in Figure 6.

### 2.1.3 Analysis Methods

The properties calculated for the binary test systems included the area per molecule, membrane thickness and lateral diffusion constants.

As there were only two components in each of the test systems, the area per molecule was accurately calculated using the following equation (Hofsäß et al.)[xxix]:

Eq.1 $A_{SGML/POPC} = [2\ A_{BOX} / (1 - x)\ N_{LIPIDS}]\ [1 - (x.N_{LIPIDS}.V_{CHOL} / V_{BOX} - N_{WATER}V_{WATER})]$

Eq.2 $A_{CHOL} = [2\ A_{BOX}.V_{CHOL}] / [V_{BOX} - N_{WATER}.V_{WATER}]$

Where A = Area, N = Number, V = Volume, x = amount of cholesterol as a decimal.

This was possible using the known volumes of a water molecule (0.03 $nm^3$) and a cholesterol molecule (0.593 $nm^3$) (Berkowitz)[xxx]. The number of lipids found in the lipid raft was 360 in a ratio of 252 sphingomyelin to 108 cholesterol molecules; there were also 18443 water molecules. As for the POPC/CHOL membrane, there were 128 lipids in a ratio of 120 POPC to 8 cholesterol molecules; there were also 2225 water molecules.

The membrane thickness of the test systems was calculated using APLvoro (Lukat et al.)[xxxi], using an average of the distances between key atoms near the edge of either leaflet (hydroxyl group for cholesterol, phosphate group for sphingomyelin and the phospholipids) over the course of the simulation.

The lateral diffusion constants of each component in the smaller systems were calculated using the Gromacs program *g_msd*. This program calculates the mean square displacement (MSD) of all the atoms present in the system from their initial positions, and then estimates the diffusion coefficient (D) according to the Einstein relation by carrying out a least-squares fit of a straight line MSD(*t*) = D.*t* + *c*, where *t* refers to the simulation time and *c* is a constant.

The properties tested for the Brain Membrane system included the area per molecule, volume per molecule, membrane thickness, electron density and lateral diffusion constants.

The area per molecule of the Brain Membrane components was calculated using Voronoi tessellation as implemented in APLvoro. The basis of Voronoi tessellation is to project the centre of mass of each molecule onto a plane, before tessellating the area of the plane by bisecting the lines that join the centres of mass (Edholm et al.)[xxxii].

The average membrane thickness was calculated as the average distance between key atoms in each leaflet, as done for the test systems. More specific values, such as thickness of the lipid raft or the individual lipids, were calculated using the selection tool in APLvoro.

The volume per lipid was calculated by taking the average area per lipid and multiplying it by half the average thickness of the membrane. More specific values were calculated by taking the average area per specific lipid group and multiplying it by half the thickness calculated for that section of the bilayer.

The electron density profile of the system was plotted based on values calculated by the Gromacs program *g_density*. This program calculates the partial electron densities of all molecules across the simulation box, using the known number of electrons for each type of atom present.

The lateral diffusion constants for the lipids in the system were calculated using the Gromacs program *g_msd*, as done for the test systems.

## 2.2 Building the DPPC-Protein System

### 2.2.1 Simulation Methods

Initial co-ordinates for activated p110α (in complex with iSH2 of p85α) were obtained from the protein data bank (4A55), and the topology was generated automatically upon conversion to Gromacs format using the program *pdb2gmx*. Point mutations were introduced manually at positions 545 (E→K) and 1047 (H→R). Coordinates for a standard patch of a DPPC bilayer containing 64 molecules were obtained for this system (Tieleman et al.)[xxxiii].

A bilayer of the appropriate size was generated using *genconf*, multiplying x and y dimensions to approximately mirror those of the size of the box of water required to adequately solvate the protein. It was then subjected to 1000 steps of steepest descent energy minimisation using the (unmodified) Gromos54A7 force field. The dimensions of the DPPC box were elongated along the z axis to make space for the protein. Having properly centred the DPPC components and protein components in their respective boxes, the coordinates were then combined, resulting in one file. The topologies of both components had to also be combined into one file, prior to subjecting the new system to a further 1000 steps of steepest descent energy minimisation.

It was then solvated in a rectangular box, using *genbox*, temporarily increasing the van der Waals radius for carbon to prevent insertion of solvent molecules into the lipid bilayer. The simple point charge (SPC) water model was used (spc216), and periodic boundary conditions (PBC) were applied. The system did not require neutralisation, so no ions were added. The final composition of the DPPC-protein system included the p110α chain, the p85α chain, a bilayer patch of 474 DPPC molecules and 73982 solvent molecules.

The simulation was initiated with an equilibration scheme identical to the Brain Membrane system, using the same algorithms, running instead for 50 ns after equilibration.

### 2.2.2 Analysis Methods

To determine the nature and the points of interaction between the protein and the membrane, a number of tests were carried out.

The fluctuation in residue position and its overall motion was determined using the Gromacs program *g_rmsf*. This program computes the root mean square fluctuation of the protein atomic positions over the course of the simulation.

The variation in the total solvent accessible surface area (SASA) of the protein during the simulations and the SASA of each residue averaged over certain parts of the simulations were calculated using the Gromacs program *g_sas*.

A homemade Python script was run to convert the per-residue SASA of the protein during its interaction with the membrane over blocks of five residues. This allowed direct comparison with the results of the Hydrogen-Deuterium exchange experiments.

## 2.3 Building the Brain Membrane-Protein System

### 2.3.1 Simulation Methods

Initial coordinates and topology for the protein were taken from the DPPC-Protein system. Initial coordinates and topology for the Brain Membrane components were taken from the Brain Membrane system. As the Brain Membrane system itself was too large for a simulation including a protein as well, new, smaller bilayer patches were created using the same proportional composition as beforehand except with the addition of PIP2 molecules. SGML and CHOL molecules were randomly inserted, as opposed to being introduced in the form of a lipid raft, as this Brain Membrane system was not large enough to accommodate a separate lipid raft.

As done for the DPPC-Protein system, these bilayer patches were randomly combined and enlarged to match the dimensions of the protein. The combined Brain Membrane-Protein system was generated as before, and neutralised by the addition of 216 cations ($Na^+$). The final composition of the Brain Membrane-Protein system included the p110α chain, the p85α chain, a Brain Membrane bilayer patch of 196 POPE, 69 POPC, 89 POPS, 22 PIP2, 48 CHOL, 22 SGML molecules and finally 73982 solvent and 216 sodium molecules.

The equilibration scheme for initiation of this simulation remained consistent with that of the Brain Membrane and DPPC-Protein systems. This simulation ran for a total of 50 ns after equilibration.

### 2.3.2 Analysis Methods

The analysis carried out on the Brain Membrane-Protein system was exactly the same as that which was carried out on the DPPC-Protein system.

## 3. Results and Discussion

### 3.1 Brain Membrane Parameters

#### 3.1.1 Test Systems Analysis

**Area per molecule**

Area per lipid is one of the key parameters used to access the accuracy of simulations of lipid bilayers. The average area of both components of the lipid raft was slightly higher than the values reported previously by Khelashvili et al.[xxxiv] Specifically, the average area of sphingomyelin in the lipid raft was determined to be 0.54 $nm^2$ using Eq.1 (see Methods), whereas Khelashvili et al. obtained a slightly lower value of 0.51 $nm^2$.

The average area of cholesterol in the lipid raft was determined to be 0.28 $nm^2$ using Eq.2 (see Methods), in comparison to a value of 0.27 $nm^2$ obtained by Khelashvili et al.

These differences are negligible as the simulations were slightly different in terms of number of water molecules, temperature, and molecular ratio of sphingomyelin to cholesterol. Khelashvili et al.'s system consisted of 266 SGML molecules, 122 CHOL molecules and 11861 water molecules at 293 K, while this system consisted of 252 SGML molecules, 108 CHOL molecules and 18443 water molecules at 298 K.

For the POPC/CHOL membrane, the average area of POPC was determined to be 0.64 $nm^2$ using Eq.1, and the average area of cholesterol was determined to be 0.27 $nm^2$ using Eq.2. Previous studies on a POPC/CHOL membrane only reported the overall average area per molecule of 0.62 $nm^2$ (Chiu et al.)[xxxv]; an identical value is obtained for the system simulated here.

There is little variation in the area of cholesterol between the two smaller systems simulated here, which corresponds well with the known lack of variation in the volume and size of cholesterol across similar membranes. However, the proportion of cholesterol present will alter the size and proximity of the other lipids in a membrane (de Meyer et al.)[xxxvi], resulting in them exhibiting a wider range of areas, as is seen for the full Brain Membrane reported below.

**Membrane thickness**

Bilayer thickness is another key quantity often used to ascertain the quality of a membrane simulation. As determined with APLvoro, the average thickness of sphingomyelin molecules across the lipid raft was 2.67 nm, while the average thickness of cholesterol molecules was 2.36 nm. Overall, the average thickness of the lipid raft over the course of the simulation was seen to be 2.58 nm. This is considerably smaller than the thickness estimated from the electron density profile produced by Khelashvili et al.[xxxiv], determined to be just under 4 nm. Both their method of determining thickness and the composition of their system were different. However, these differences cannot fully explain the variation, which is better explained by the original coordinates for sphingomyelin being from different sources- the conformation of SGML molecules used here involves tails which are less extended.

As determined with APLvoro, the average thickness of POPC molecules across the POPC/CHOL membrane was approximately 3.52 nm, while the average thickness of cholesterol molecules was approximately 2.90 nm. Overall, the average thickness of the POPC/CHOL membrane over the course of the simulation was seen to be 3.49 nm. These values are in agreement with the thickness estimate from the electron density profile from the work of Chiu et al.[xxxv]

The POPC/CHOL membrane is more loosely packed and less aggregated than the lipid raft, due to the less saturated tails of POPC compared to sphingomyelin. Lipid rafts display a close interaction between sphingomyelin and cholesterol which results in a more aggregated membrane (Pike)[xxxvii].

**Lateral diffusion constants**

The lateral diffusion constant refers to the mean square displacement of a particular molecule or group of molecules throughout the simulation. A higher lateral diffusion constant would indicate more lateral movement and therefore a higher fluidity. Again, this is a key value for evaluating the degree to which a simulation reproduces the true behaviour of a lipid membrane.

The values of the lateral diffusion constant for the lipid raft components were seen to be in the $10^{-11}$ $m^2/s$ range; sphingomyelin molecules 2.89 (+/- 0.50) $x10^{-11}$ $m^2/s$, cholesterol molecules 2.31 (+/- 0.50) $x10^{-11}$ $m^2/s$, and overall 2.78 (+/- 0.50) $x10^{-11}$ $m^2/s$. Lateral diffusion was not one of the factors determined by Khelashvili et al.[xxxiv]

The values for the POPC/CHOL membrane components were an order of magnitude higher, in the $10^{-10}$ $m^2/s$ range; POPC molecules 4.31 (+/- 0.17) $x10^{-10}$ $m^2/s$, cholesterol molecules 1.38 (+/- 0.3) $x10^{-10}$ $m^2/s$, and overall 3.32 (+/- 0.17) $x10^{-10}$ $m^2/s$. These values are in good agreement with Chiu's work[xxxv], and reflect the more loosely packed nature of the POPC/CHOL system as noted above.

As expected, the values for the phopholipid components are higher than those of the membrane raft components. This would indicate that the lipid raft is a less fluid bilayer with more rigid components as confirmed by previous studies (Niemelä et al.)[xxxviii].

Overall, the values calculated for the physical properties of the smaller systems were quite similar to those seen in previous computational studies, allowing the assumption that the parameters were indeed sound and viable for use in a larger system. The main parameters that needed to be tested were those of sphingomyelin and cholesterol, as they were the lipid components that involved new force field definitions. As explained in the methods, parameters for POPC already existed, while parameters for POPE and POPS were similar and involved existing force field definitions.

### 3.1.2 Brain Membrane System Analysis

**Area per molecule**

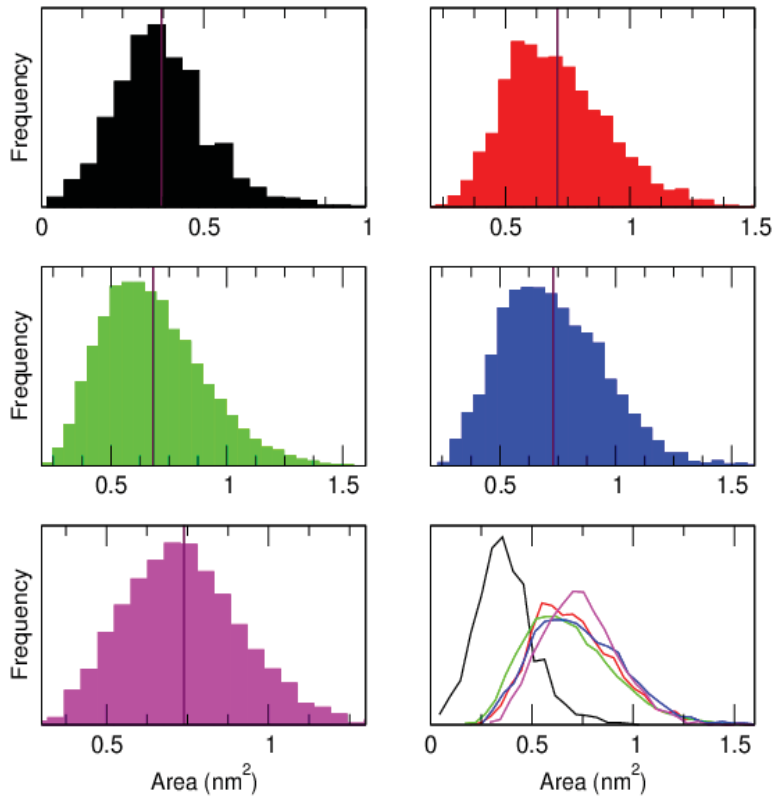


**Figure 7**: *Histogram showing area distribution of the Brain Membrane components. The mean is indicated by a maroon line. CHOL (black), POPC (red), POPE (green), POPS (blue), SGML (pink).*

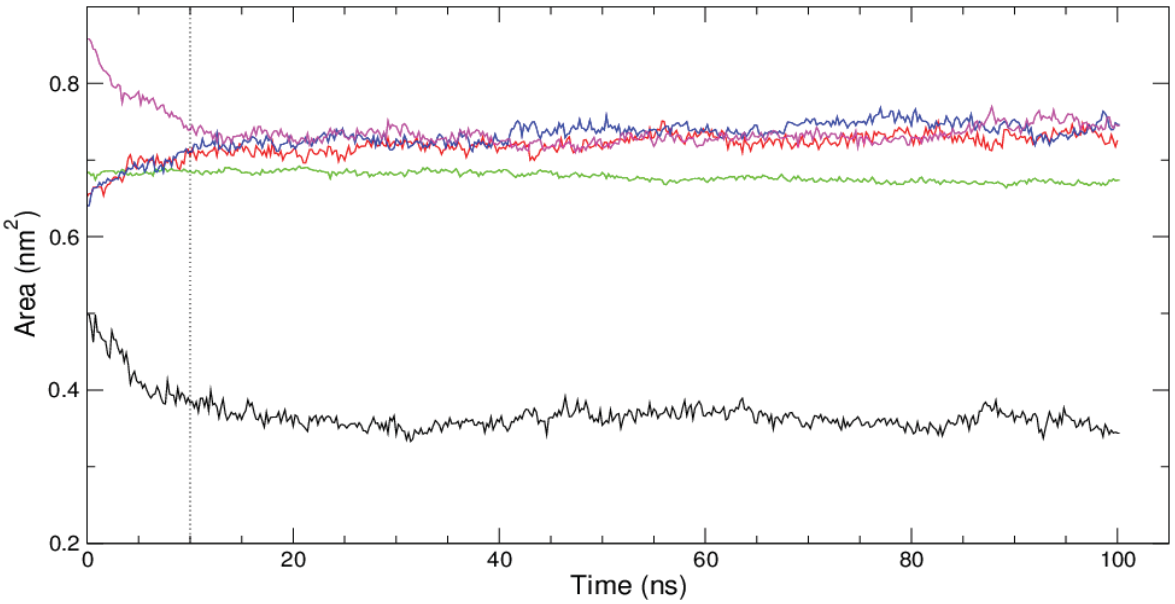


**Figure 8**: *Time series graph showing variation in areas of the Brain Membrane components averaged over all lipids of that type throughout the simulation. CHOL (black), POPC (red), POPE (green), POPS (blue), SGML (pink). The dotted line indicates the 10 ns cut-off used to calculate overall averages from the whole simulation.*

Due to the more complex nature of the brain membrane, it was not possible to use Eq. 1 and 2 to calculate the area per lipid, as were used for the simpler two-component test systems. Instead, the APLvoro program was used, which utilises the Voronoi tessellation method to calculate the area per lipid. Histograms showing the range of areas of each type of phospholipid throughout the simulation are shown in Figure 7; while the time series graph (Figure 8) shows the fluctuation in the average area of each type of lipid throughout the whole simulation. Overall averages of the area and volume per lipid are given in Table 1. These averages were calculated from 10 ns onwards, as the behaviour (in terms of average area per lipid) only stabilised after this point (Figure 8); from this point onwards in the simulation, the magnitude of the variations remained consistent.

The overall average areas of the phospholipids were approximately 0.7 $nm^2$; POPC 0.72 $nm^2$, POPS 0.73 $nm^2$, POPE 0.68 $nm^2$. These values are somewhat different to those seen in the previous studies done by Chiu[xxxv] and Hofsäß[xxix], who obtained values for area per phospholipid in the 0.65 $nm^2$ range. However, as seen in the histograms in Figure 7, the highest frequency peak (median) is approximately 0.65 $nm^2$, but the mean is larger due to the skew of the distributions towards larger values.

The Voronoi tessellation also gives the overall average area of sphingomyelin as 0.74 $nm^2$, and cholesterol as 0.37 $nm^2$, which show significant increases from the values calculated for smaller system using Eq.1 and Eq.2. The discrepancies between the area per lipid values calculated for the complete brain membrane using the Voronoi tessellation method and those calculated from previous simulations may therefore be due to the method rather than a problem with the simulations themselves. Indeed, it is a known drawback of the Voronoi tessellation method that the size of smaller molecules can be exaggerated (Pandit et al.)[xxiv].

The area calculated using this method also fluctuates significantly depending on which level of the membrane along the z axis the 'surface' is deemed to be (Edholm et al.)[xxxii]. For this study, the invisible line to determine area and thickness was drawn through the phosphate groups, considering them as the key components of the phospholipids most likely situated close to the solvent. For cholesterol, the invisible line was drawn through the hydroxyl group. Had the invisible line been drawn at a different *z* coordinate, results could have varied, as there is a *z*-dependence to molecular areas when examined at an atomic level (Falck et al.)[xxxix].

**Volume per lipid**

As seen in previous studies (Orsi et al.)[xl], volumes for phospholipid components of a membrane are approximately 1.3 $nm^3$, while the known volume of a cholesterol molecule as mentioned in the analysis methods is approximately 0.593 $nm^3$.

The slight variation in the phospholipid volumes reported in Table 1 is easily explained. In comparison to POPE, POPC has an extra three methyl groups while POPS has an extra carboxyl group, resulting in slightly larger volumes. SGML has the largest headgroup, however its straighter, more rigid tails result in a smaller volume than the phospholipids.

**Membrane Thickness**

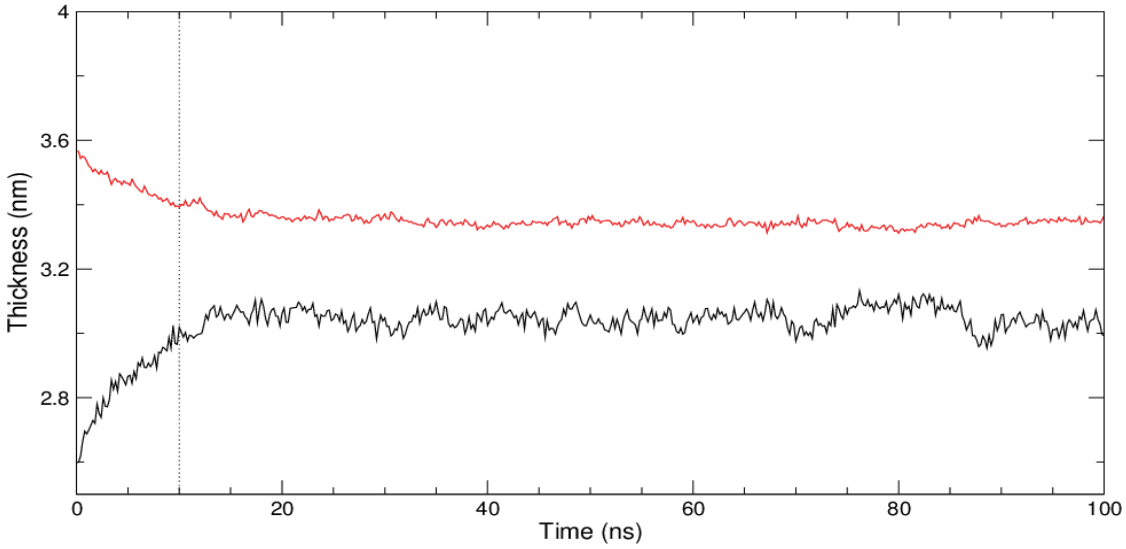


**Figure 9**: *Time series graph showing variation in the thickness of the phospholipid (red) and lipid raft (black) components of the Brain Membrane over the course of the simulation. The dotted line indicates the 10 ns cut-off used to calculate averages.*

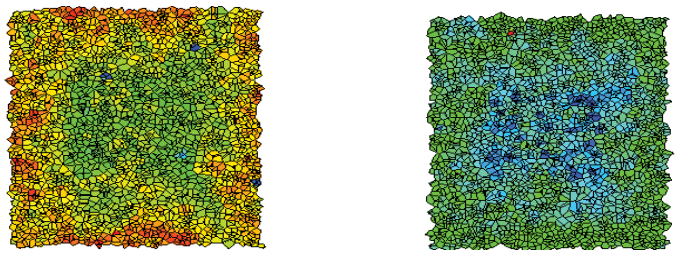

**Figure 10**: *Images showing the Voronoi tessellation at two different points in the simulation of the Brain Membrane highlighting membrane thickness using a cold to hot colour scale (blue being the thinnest, red being the thickest).*

The membrane thickness was also calculated using APLvoro, with overall averages (from 10 ns onwards, to allow for equilibration) given in Table 1, and a time series of the average thickness of the phospholipids and of the lipid raft shown in Figure 9. As expected, the thickness of the brain membrane is smaller in and around the lipid raft and greater on the edges where the phospholipids reside. These results are similar to what was seen for the smaller systems, where the two components were simulated separately. As observed from the general trend of thickness over time, it was deduced that the membrane "breathes" in an undulating fashion. This was confirmed by viewing the trajectory of atomic coordinates saved at 200 ps intervals throughout the simulation (available upon request). The membrane on a whole expands and contracts following a similar pattern (see Figure 10). The lipid raft always remains thinner than the phospholipid section; however the entire membrane becomes thicker and thinner as a unit. Each small sector of the membrane exhibits a pattern within itself resulting in an uneven wavy surface along the membrane at times. As seen in the time series graph (Figure 10), at various points in the simulation, troughs and peaks indicating membrane thickness are synchronised within the lipid raft and phospholipid sections, so that when the lipid raft becomes thinner, the phospholipid regions become thicker and vice versa.

**Lateral diffusion constants**

The lateral diffusion constants are lower for the components of the lipid raft than for the phospholipids, indicating that sphingomyelin and cholesterol are the most anchored molecules, while the phospholipid components are slightly more mobile. Thus, as expected, the lateral diffusion constants reveal less fluidity for the molecules in the lipid raft, as they remain more tightly packed and closely aggregated than the rest of the bilayer. This behaviour was also seen for the test systems.

**Table 1**: *Calculated physical properties of the Brain Membrane, averaged over all lipids of that type and over the final 90 ns of the simulation. The standard deviation of the lateral diffusion constants is also given; for the other properties, the variation is shown in the distribution plot (Figure 7) and time series plots (Figures 8 and 9).*

| | Area per Lipid ($nm^2$) | Volume per Lipid ($nm^3$) | Lateral Diffusion ($x10^{-12}$ $m^2/s$) | Thickness (nm) |
|---|---|---|---|---|
| **POPE** | 0.68 | 1.15 | 4.22 +/- 0.43 | 3.37 |
| **POPC** | 0.72 | 1.23 | 3.91 +/- 0.07 | 3.42 |
| **POPS** | 0.73 | 1.24 | 4.22 +/- 0.51 | 3.41 |
| **SGML** | 0.74 | 1.15 | 1.78 +/- 0.03 | 3.11 |
| **CHOL** | 0.37 | 0.52 | 1.00 +/- 0.07 | 2.83 |
| | | | | |
| **Lipid Raft** | 0.62 | 0.94 | 1.55 +/- 0.03 | 3.02 |
| **Phospholipids** | 0.70 | 1.19 | 4.16 +/- 0.38 | 3.39 |
| | | | | |
| **Membrane** | 0.69 | 1.15 | 3.97 +/- 0.39 | 3.32 |

**Electron density**

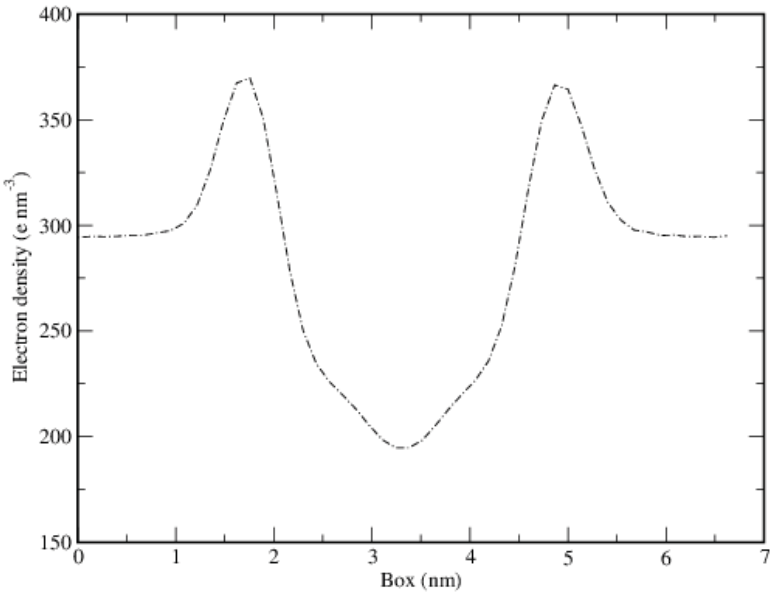


**Figure 11**: *Electron density profile of the Brain Membrane averaged over all lipid types over the final 90 ns of the simulation.*

The electron density profile (Figure 11) of the Brain Membrane exhibits two major peaks as well as a major trough. The peaks refer to the electron-rich phosphate groups in the headgroups of the lipids in either leaflet. The trough refers to the gap between the two leaflets, while the gradient rises moving up the lipid tails. On the outside of both peaks, there exists tiny interference from the solvent molecules of the system.

The shape of the graph, as well as the distance between both peaks compares favourably with another multi-component membrane (Pandit et al.)[xxiv], as well as the general shape expected for a lipid bilayer.

### 3.2 DPPC-Protein System Analysis

Having developed suitable parameters for the various components of a brain membrane and confirmed their validity, it was possible to return to the original aim of this project: simulation of activated p110α in the presence of a lipid bilayer. This was stimulated by the observation of differences in the hydrogen-deuterium exchange rate (HDx) between different parts of the protein upon interaction with a brain membrane by Burke et al.[xvii] MD simulations provide the opportunity to obtain atomic-level insight into the structural changes giving rise to these experimental observations. Two different protein-membrane systems were simulated: firstly, with a pure DPPC bilayer, as is commonly used in simulations of protein-membrane systems, and secondly, with a realistic brain membrane modelled using the parameters developed here. The DPPC bilayer provides a means of assessing how important a realistic representation of the membrane components is to accurately describing the membrane interaction behaviour of p110α. For simplicity, all previous experimental data mentioned from here on refers to the HDx data of Burke et al (Supporting Information Figure 3), unless otherwise stated.

#### 3.2.1 Approach of Protein to Membrane

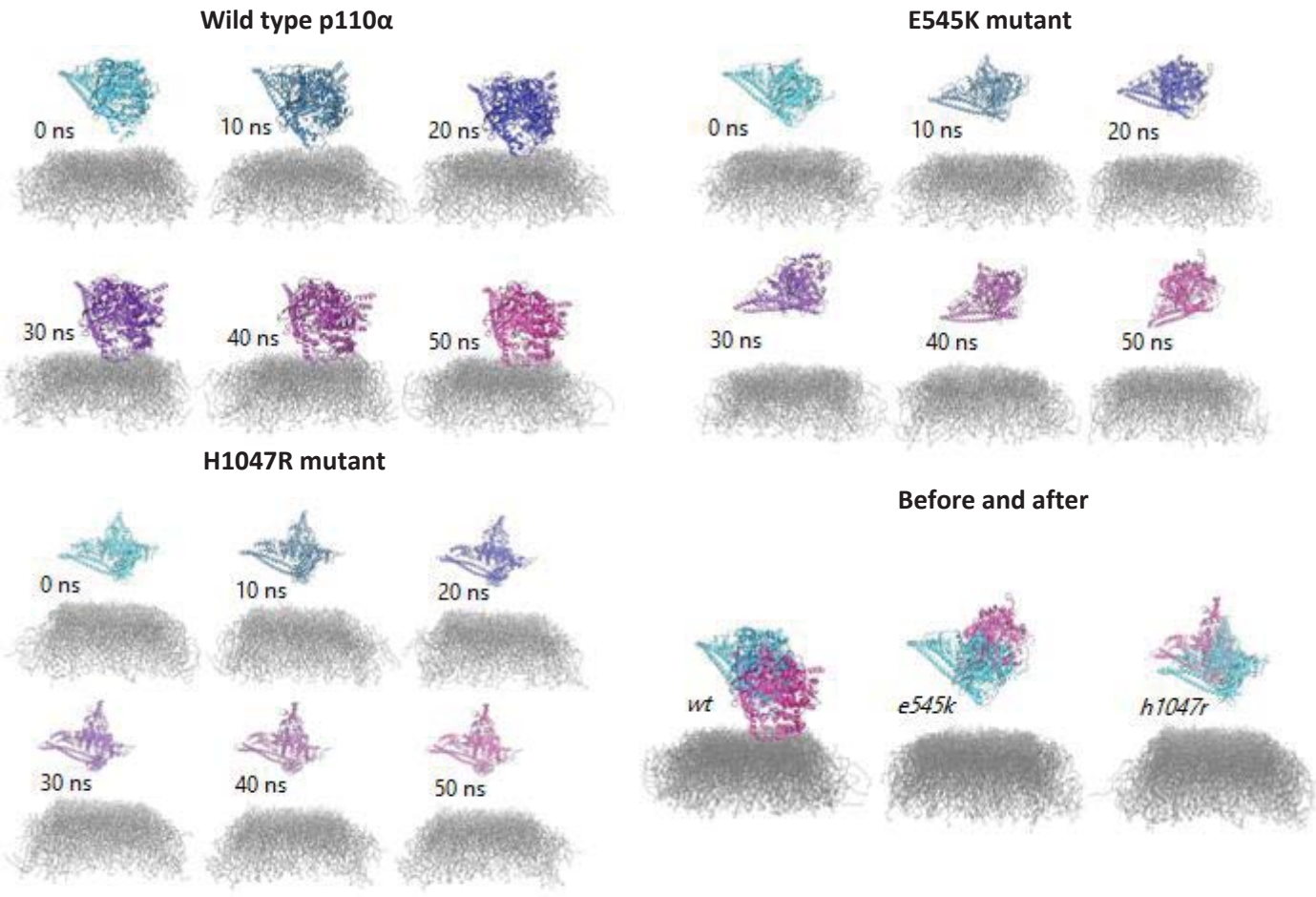


**Figure 12**: *Snapshots of the DPPC membrane + protein systems throughout the MD simulations. Each system is shown at 10 ns intervals. Before (0 ns - cyan) and after (50 ns - magenta) comparisons of each system are also shown. Water molecules have been removed for clarity.*

Each protein was originally placed 2 nm away from the lipid bilayer in all directions, so as to avoid biasing it towards any particular membrane interaction. The snapshots from throughout the 50 ns MD simulation of each protein and a pure DPPC membrane (Figure 12) reveal only the wild type p110α exhibiting a steady approach towards the bilayer, before direct contact occurred at approximately 20 ns. Neither the E545K mutant nor the H1047R mutant showed any significant signs of approaching the membrane, although the E545K mutant does reorient significantly. This is further evidenced by the "before and after" snapshots of each system.

### 3.2.2 Structural Fluctuations

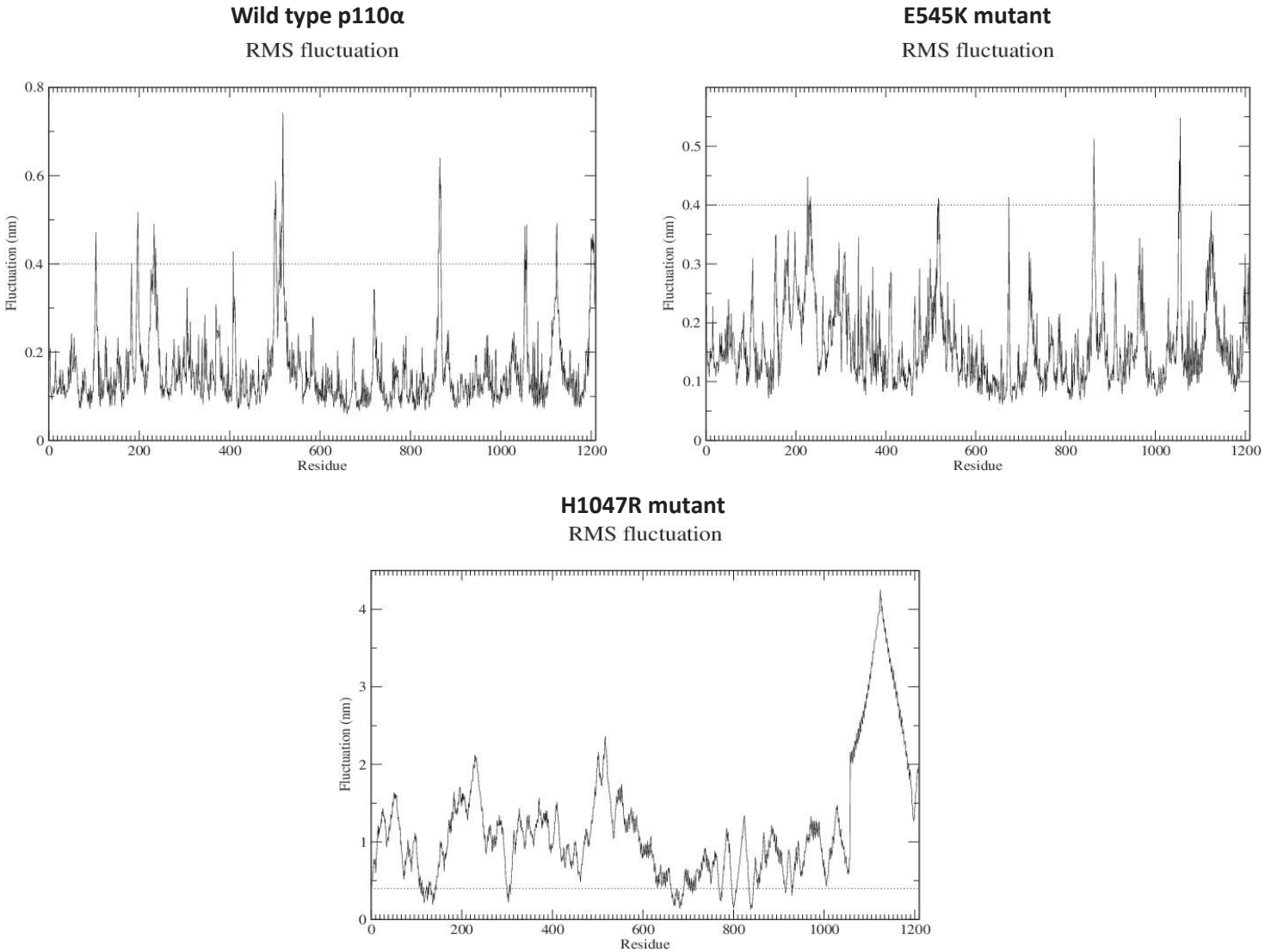


**Figure 13**: *RMS fluctuation of each residue in each protein (wild type, E545K mutant, H1047R mutant) over the course of the simulation of each in the presence of a DPPC bilayer.*

To get a good understanding of which residues were influential in the approach of the protein to the membrane, the root mean square (RMS) fluctuation in the position of each residue was calculated throughout the 50 ns simulation. An arbitrary value of 0.4 nm was chosen as the cut-off point for significant fluctuation, represented by the dotted lines in Figure 13. Any residues with peaks beyond 0.4 nm were considered to be undergoing large-scale conformational change as the protein approached and interacted with the membrane. The large and sudden increase in the value of the RMS fluctuation by around 3 nm at residue 1096 of the H1047R mutant is due to the iSH2 domain of the p85α subunit (residues 1096 to 1208) being situated on the opposite side of the simulation box to the p110α subunit, once the periodicity had been accounted for. Despite the scale of the graph having been affected, the shape and pattern remain correct.

As shown in Figure 13, there are significant peaks in the RMS fluctuation which occur for all three proteins. These peaks are exhibited by blocks of residues, the highest of which occur in the residue regions 230 to 245 (loop in the RBD domain), 500 to 510 (loop between the C2 and helical domains), 860 to 875 (loop in the kinase domain), 1050 to 1065 (loop in the kinase domain). This shows good agreement with the HDx data, which represents high relative deuteration, representative of increased solvent exposure, in these areas. The two loops located in the kinase domain have previously been found to interact with the membrane (Mandelker et al.)[xiv]. This suggests that the other loops are implicated in conformational changes either during the protein's approach to or its contact with the membrane.

### 3.2.3 Solvent Accessible Surface Area

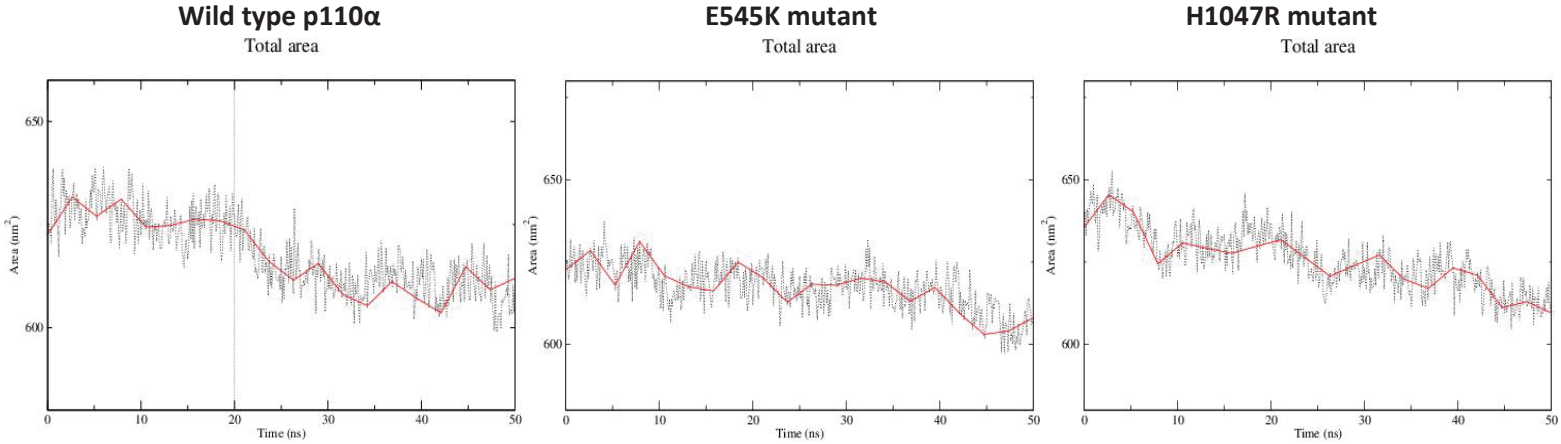


**Figure 14**: *Total solvent accessible surface area of each protein over the course of the 50 ns simulations in the presence of a DPPC bilayer. The general trend is represented by the red line, and the time of protein-membrane contact is shown as a dashed vertical line (if contact occurred).*

The total solvent accessible surface area (SASA) of each protein over the course of the simulation (Figure 14) was calculated to give an indication of the time point at which the protein began to interact significantly with the membrane, as defined by any major decrease in the SASA.

To distinguish from general fluctuation, a major change was defined as a decrease in total area of 10 $nm^2$ in the space of 5 ns of simulation. This would indicate that the protein had come into contact with another molecule, other than solvent, at that point in time.

These graphs were used in association with viewing the trajectory of atomic coordinates saved at 25 ps intervals throughout each simulation, exhibited roughly here in the form of snapshots (Figure 12), to determine an accurate time frame for which initial contact occurred between the protein and the bilayer. This value was then represented as a dashed vertical line on each graph. However, as seen in both Figures 12 and 14, the only protein which initiated contact with the DPPC membrane over the course of the simulation was wild type p110α. Neither the E545K nor the H1047R mutant came into contact with the bilayer according to the SASA analysis reflecting the data provided by their respective simulation snapshots (Figure 12). Monitoring the changes in the total SASA of the protein, wild type p110α was deemed to have come into contact with the membrane at approximately 20 ns, which is in good agreement with its snapshot series (Figure 12).

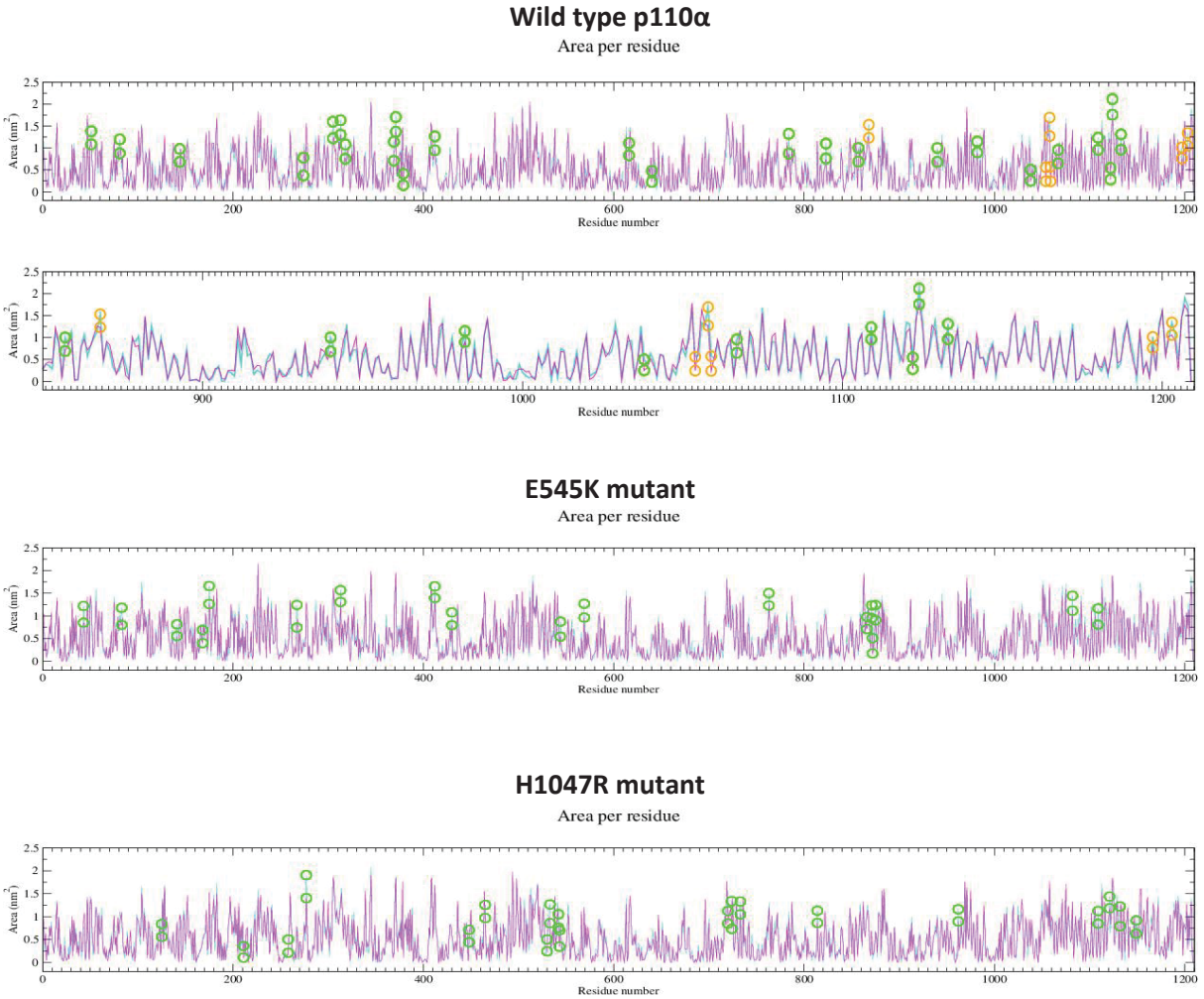


**Figure 15**: *SASA of each protein residue before (cyan) and after (magenta) coming into contact with the DPPC bilayer. Residues that exhibit considerable change in SASA are highlighted in green (residues that do not contact the membrane) and orange (residues that do contact the membrane). In the case of interaction with the membrane (wild type p110α), a further graph focusing on the residues in contact with the membrane (residues 800 to 1208) is also shown.*

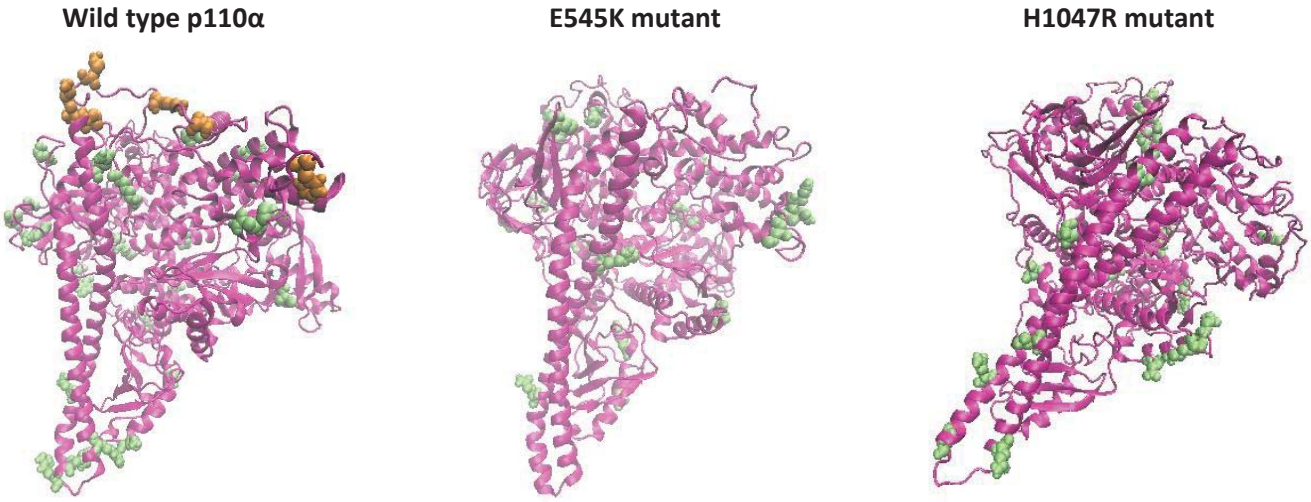


**Figure 16**: *Images of each protein (magenta) depicting the residues that undergo significant change in SASA (green and orange) from Figure 15, showing the sections of the protein where these changes take place.*

To pinpoint which of the groups of residues involved in conformational changes during the course of the simulation were actually involved in the interaction of the protein with the DPPC membrane, the solvent accessible surface area (SASA) of each residue was calculated before (5-20 ns) and after (30-45 ns) the protein came in contact with the membrane (Figure 15). As determined from the snapshots and the analysis of the total SASA over the course of the simulation (Figures 12 and 14), only the wild type p110α appeared to initiate contact with the DPPC membrane. The time points chosen to represent before and after membrane interaction were therefore based on this system. Despite the mutants seemingly not having approached the membrane, the same SASA analysis was carried out on their residues, to allow comparison with the changes observed for wild type p110α. The location of the residues that exhibit significant changes in SASA in the p110α are shown in Figure 16.

In the case of wild type p110α, the change in SASA per residue (Figure 15) confirmed the conclusions drawn from the RMS fluctuation data (Figure 13), as well as giving further insight into the causes of each region of high RMS fluctuation. Upon interaction with the membrane, the SASA of blocks of residues (860 to 875 and 1050 to 1065) decreases significantly, indicating that these loops in the kinase domain were involved in direct contact with the DPPC bilayer. In addition, the SASA of a block of residues (1990 to 1205) also decreases, indicating that this loop on the edge of the iSH2 domain of p85α was also involved in direct contact with the membrane. The high RMS fluctuation values of these regions are therefore due to conformational changes as they interact with the DPPC bilayer. The remaining residues exhibiting high RMS fluctuation values, namely residues in regions 100 to 115 (loop linking ABD to RBD domain), 290 to 320 (loop linking RBD to C2 domain), 340 to 355 (loop in the C2 domain) and 1100 to 1110 (loop in the iSH2 domain of p85α) were therefore involved in conformational changes not directly related to membrane interaction. This is consistent with the HDx analysis data, as well as visual representations provided by Burke et al (Supporting Information Figure 4)[xvii].

### 3.3 Brain Membrane-Protein System Analysis

#### 3.3.1 Approach of Protein to Membrane

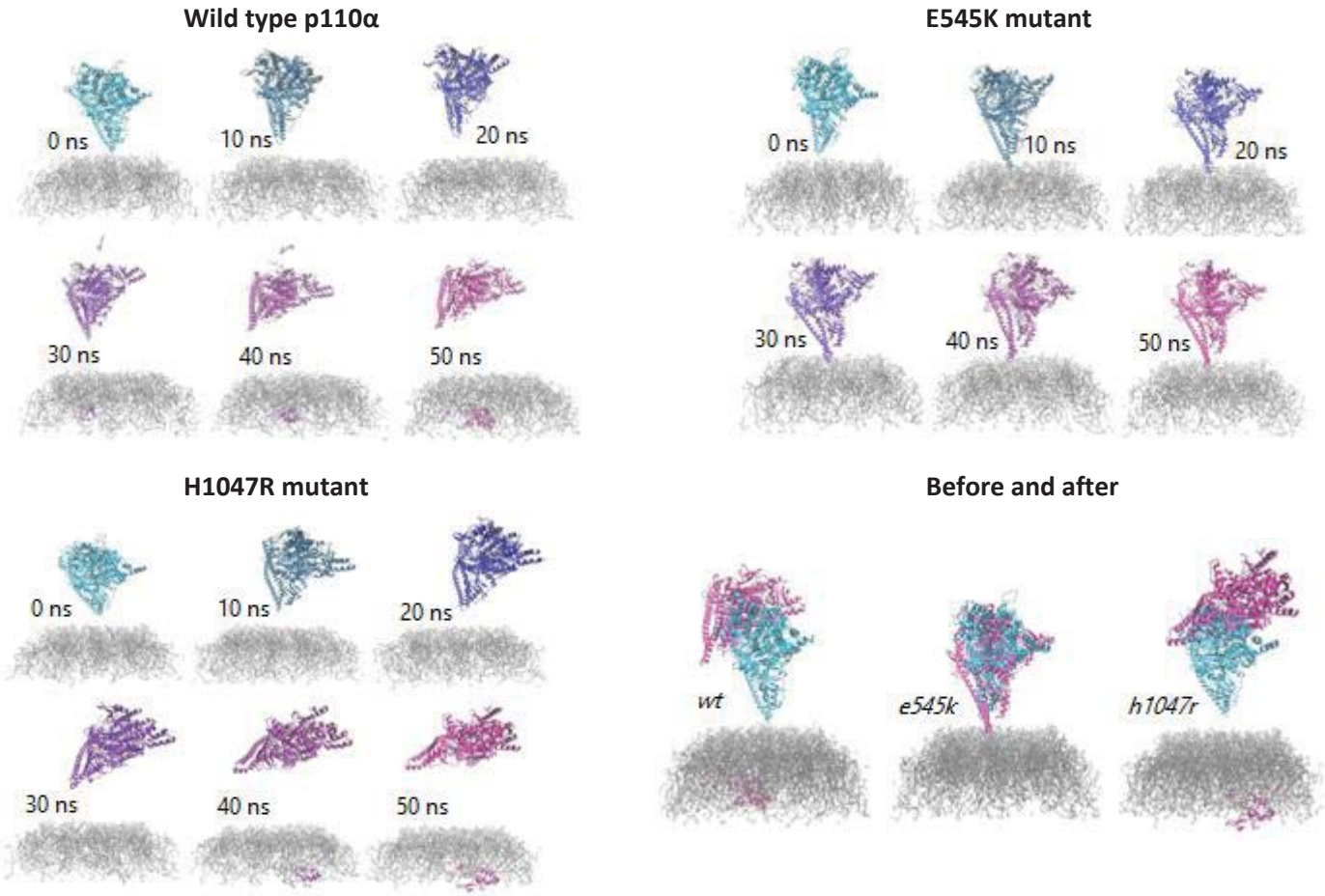


**Figure 17**: *Snapshots of the Brain Membrane + protein systems throughout the MD simulations. Each system is shown at 10 ns intervals. Before (0 ns - cyan) and after (50 ns - magenta) comparisons of each system are also shown. Water molecules have been removed for clarity.*

The simulations of each protein with a realistic Brain Membrane were set up similarly to those with a DPPC bilayer, with each protein placed 2 nm from the brain lipid bilayer to avoid biasing it towards interacting with the membrane. However the initial orientation of each protein was rotated -90° around the X axis relative to the initial orientation for the DPPC-protein systems, simply to allow for a smaller box and therefore a faster simulation time. Due to this, the loops thought to be involved in membrane interaction (860 to 875 and 1050 to 1065) were originally pointed away from the bilayer.

The snapshots of the 50 ns MD simulation of each protein and the newly parameterised brain membrane (Figure 17) reveal both the wild type p110α and the H1047R mutant exhibiting a steady approach towards the bilayer, before direct contact occurred at approximately 30 ns. While the snapshots show the protein approaching the opposite side of the bilayer compared to in the DPPC-protein simulations, this is in fact not a true difference, as both directions are equivalent due to the period boundary conditions. On the other hand, the E545K mutant showed no significant signs of approaching the membrane, remaining in the surrounding solvent and undergoing little reorientation. This is further evidenced by the “before and after” snapshots of each system.

### 3.3.2 Structural Fluctuations

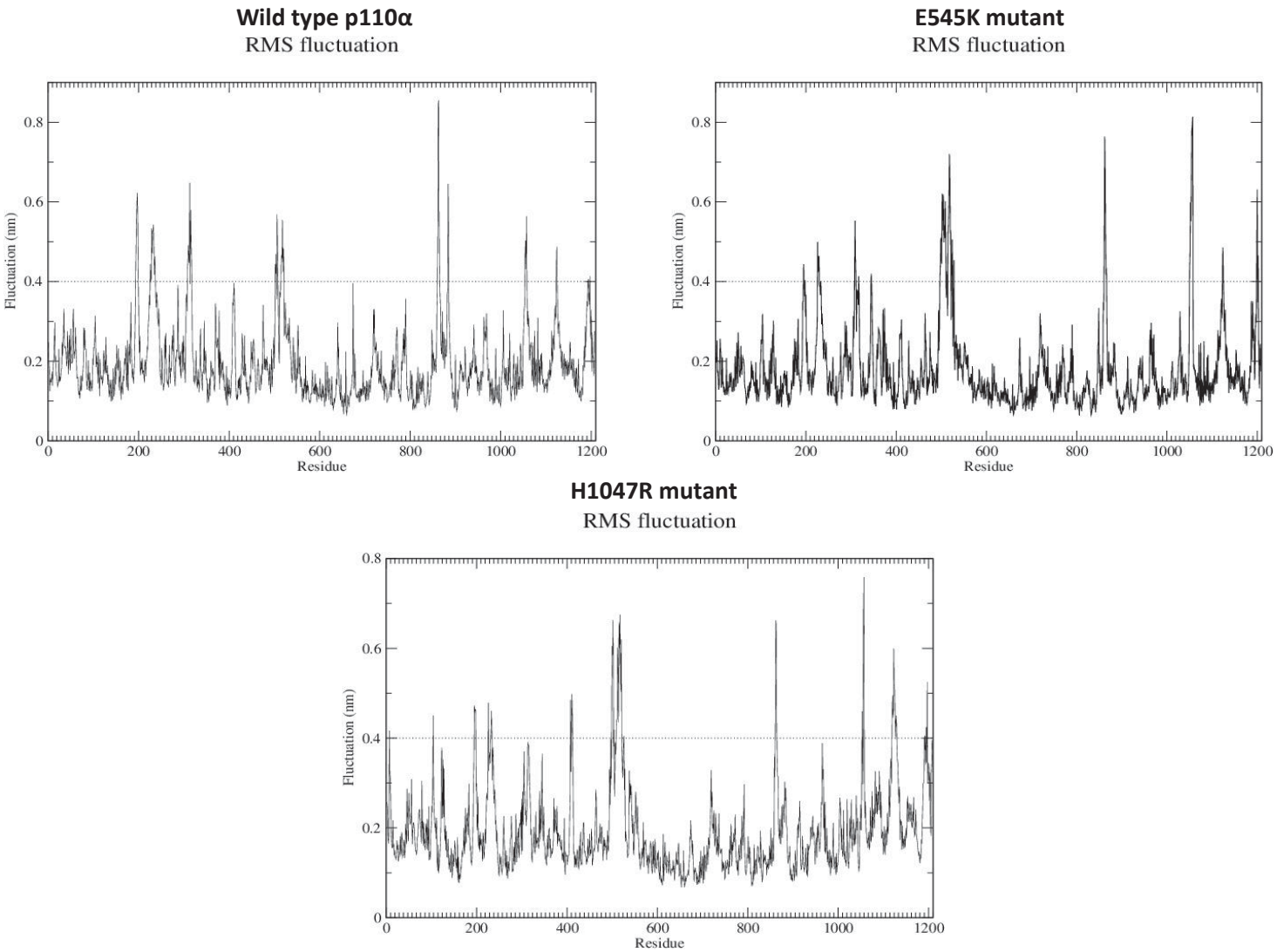


**Figure 18**: *RMS fluctuation of each residue in each protein (wild type, E545K mutant, H1047R mutant) over the course of the simulation of each in the presence of the Brain Membrane.*

As done for the DPPC-protein systems, the structural changes during the course of the simulation were investigated by calculating the per-residue RMS fluctuations, with an arbitrary value of 0.4 nm chosen as the cut-off point for significant fluctuation, represented by the dotted lines in Figure 18. Residues with peaks beyond 0.4 nm were considered to be undergoing large-scale conformational change as the protein approached and interacted with the membrane.

The Brain Membrane-protein systems showed the same peaks in the RMS fluctuation profile as were seen for the DPPC-protein simulations (Figure 18). These peaks are exhibited by blocks of residues, with the highest RMS fluctuation values occurring in the residue regions 230 to 245 (loop in the RBD domain), 500 to 510 (loop between the C2 and helical domains), 860 to 875 (loop in the kinase domain), 1050 to 1065 (loop in the kinase domain). The conformational changes exhibited by these groups of residues are therefore independent of whether a pure DPPC bilayer or a realistic Brain Membrane is used. This is not unexpected in the case of the first two regions, which did not interact

with the bilayer in the protein-DPPC simulations and so are not expected to exhibit different behaviour depending on the nature of the membrane. While different behaviour in the presence of the brain membrane might have been expected for the latter two loops, located in the kinase domain, that are expected to interact with the membrane (Mandelker et al.)[xiv], the RMS fluctuations merely indicate the extent of conformational change, not the nature of it, thus these results alone are not sufficient to determine in detail any changes that might occur in the protein-membrane interaction stemming from use of a more realistic membrane model.

Interestingly, for the Brain Membrane-protein systems, the peak occurring in the residue region 500 to 510 appeared significantly larger in the mutant (E545K and H1047R) systems when compared to the wild type p110α. While overall the RMS fluctuation data shows good agreement with the HDx data, the difference observed at the loop between the C2 and helical domains (residues 500 to 510) was not registered in the HDx data. This could, however, be merely due to experimental error or uncertainty.

### 3.3.3 Solvent Accessible Surface Area

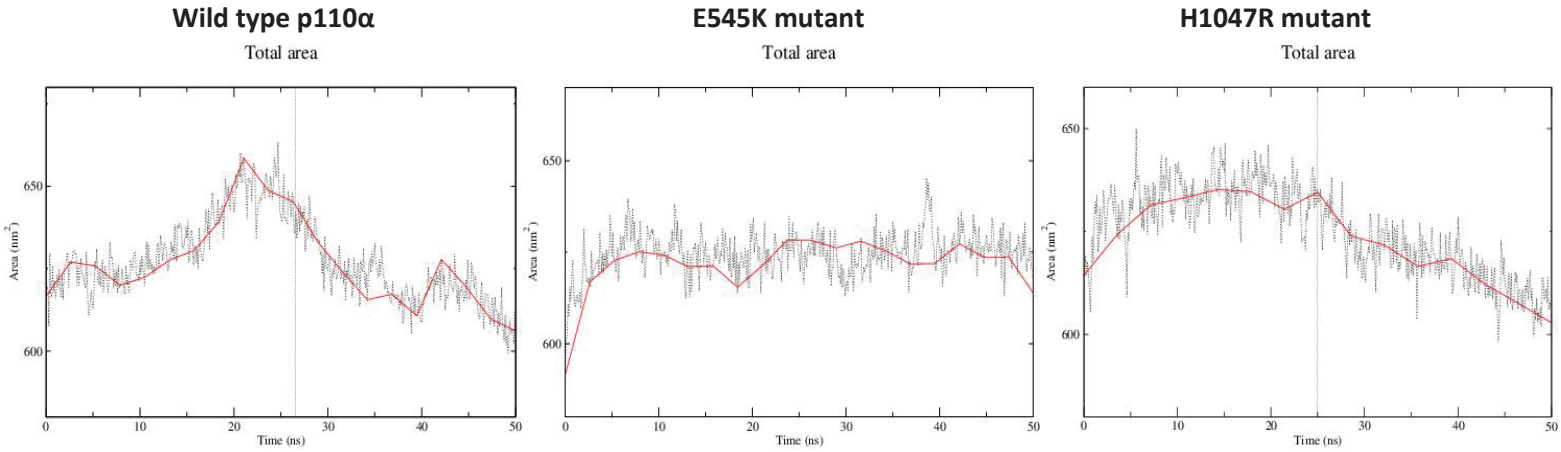


**Figure 19**: *Total solvent accessible surface area of each protein over the course of the 50 ns simulations with the brain membrane. The general trend is represented by the red line, and the time of protein-membrane contact is shown as a dashed vertical line (if contact occurred).*

The total solvent accessible surface area (SASA) of each protein was again calculated over the course of the simulation (Figure 19) to give an indication of the time point at which the protein began to interact significantly with the membrane, as defined by a major decrease in the SASA. As done for the DPPC-protein systems, a major change was defined as a decrease in total area of 10 $nm^2$ in the space of 5 ns of simulation.

These graphs were used in association with viewing the trajectory of atomic coordinates saved at 25 ps intervals throughout each simulation, summarised as snapshots taken every 10 ns (Figure 17), to determine an accurate time frame for which initial contact occurred between the protein and the bilayer. This value was displayed as a dashed vertical line on each graph (Figure 19). As observed in both Figures 17 and 19, both the wild type p110α and the H1047R mutant initiated contact with the bilayer; but, once again, the E545K mutant failed to come into contact with the bilayer. While the

E545K mutant's activity is independent of p85α, it does require the presence of Ras-GTP to initiate membrane interaction (Meyer et al.)[xli]. These deductions were confirmed by their respective simulation snapshots (Figure 17). Based on the change in total SASA of each protein (Figure 19), wild type p110α was estimated to have directly interacted with the membrane at 27 ns, while the H1047R mutant approached 2 ns sooner. The significant decreases in the trend line seen at these time points are a good indication of the sudden decrease in the interaction of the protein with the surrounding solvent. The simulation snapshots (Figure 17) do not allow for any perception of such a small a difference in the time of interaction, but they confirm the general time frame for interaction.

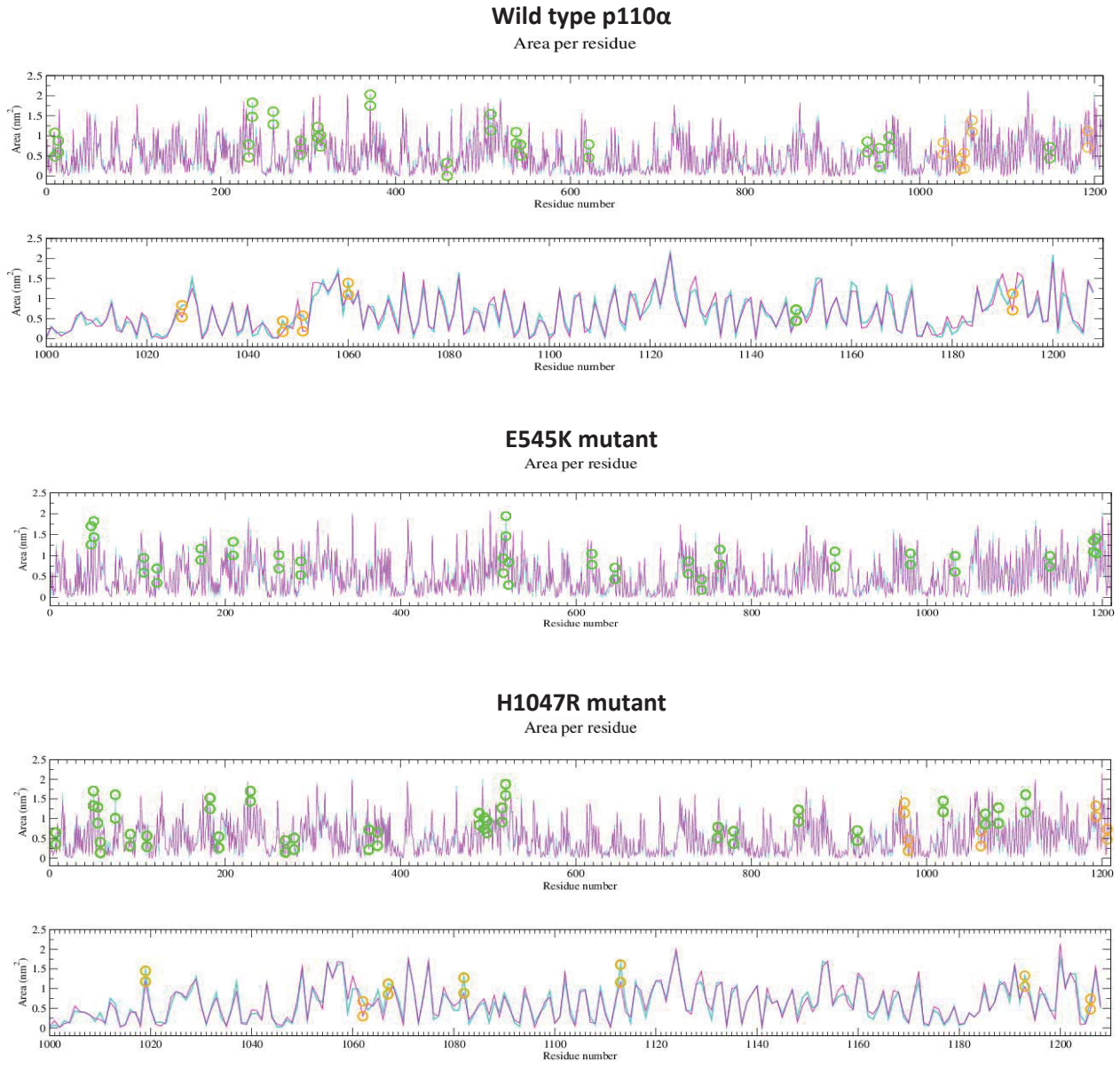


**Figure 20**: *SASA of each protein residue before (cyan) and after (magenta) coming into contact with the brain membrane. Residues that exhibit considerable change in SASA are highlighted in green (residues that do not contact the membrane) and orange (residues that do contact the membrane). In the case of interaction with the membrane (wild type p110α and H1047R mutant), a further graph focusing on the residues in contact with the membrane (residues 800 to 1208) is also shown.*

**Wild type p110α** **E545K mutant** **H1047R mutant**

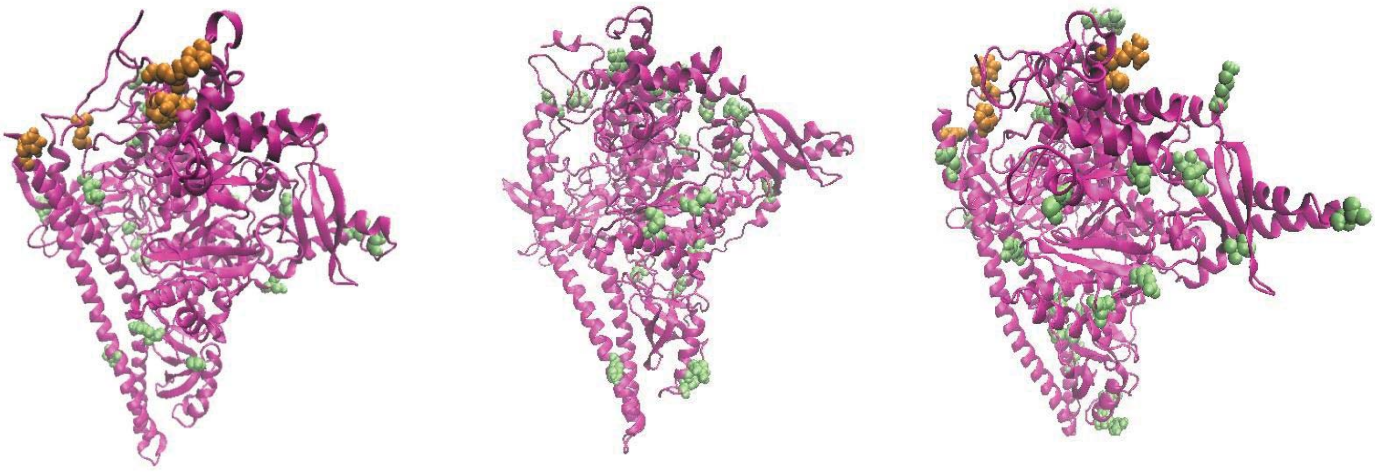

**Figure 21**: *Images of each protein (magenta) depicting the residues that undergo significant change in SASA (green and orange) from Figure 20, showing the sections of the protein where these changes take place.*

As done for the DPPC-protein system, the groups of residues involved in conformational changes during and upon the approach of the protein to the Brain Membrane were pinpointed by calculating the per-residue SASA at points before (10 to 25 ns) and after (35 to 50 ns) interaction with the Brain Membrane occurred. These time intervals were chosen using the times at which interaction occurred estimated from Figure 19. Although the E545K mutant appeared to not have directly interacted with the membrane, identical SASA analysis was carried out on its residues to facilitate comparative analysis.

Combining the SASA per residue data (Figure 20) and the RMS fluctuation data (Figure 18), results in the same blocks of residues as in the DPPC-protein systems being implicated in the interaction of wild type p110α with the Brain Membrane. The locations of the residues that undergo the greatest change in SASA upon interaction of the protein with the membrane are highlighted in Figure 21).

In the case of the H1047R mutant, similar residues were involved in direct contact with the membrane, but there were many more residues implicated in conformational changes in regions that do not directly come into contact with the membrane. These primarily occurred in the residue blocks in and around the 1050-1065 kinase domain loop that does contact the membrane. This would suggest that the H1047R mutation initiates more conformational changes around membrane binding loops to facilitate interaction, in agreement with the greater conformational motion observed in this region by X-ray diffraction (Mandelker et al.)[xiv]. Overall, these findings are once again consistent with the HDx analysis, as well as visual representations provided by Burke et al. (Supporting Information Figure 4)[xvii].

## 4. Conclusions and Future Research

The aim of this study was to use computer simulation to provide atomic-level insight into the structural changes underlying the results of HDx experiments carried out by Burke et al.[xvii], in which changes in the SASA of specific groups of residues upon interaction between p110α and its mutants with a brain phospholipid bilayer were observed. Before this could be done, it was first necessary to assemble and test parameters for the various components of a brain membrane.

Parameters for the phospholipids POPC, POPE, and POPS, and for the additional membrane components cholesterol, sphingomyelin and PIP2 compatible with the GROMOS 54A7 force field were generated. In the case of sphingomyelin and cholesterol, this required slight modification of the force field. Simulations of small binary test systems showed these parameters to be valid, as judged by their ability to reproduce, within the limits of the analysis methods used, key structural and dynamic properties of previously simulated membranes.

A complete brain membrane, including a lipid raft, was then constructed and simulated. This represents the first atomic-level MD simulation of such a system. As well as further validating the lipid parameters, it provided insight into the nanosecond-scale dynamics of such a bilayer, revealing an undulating breathing motion.

The wild type and two oncogenic mutants of p110α in its activated form (bound to the iSH2 domain of p85α) were then simulated in the presence of a pure DPPC membrane and a brain membrane. The simulations with a DPPC membrane allowed the effects of realistically modelling the brain membrane to be evaluated.

Wild type p110α approached and interacted with both the DPPC bilayer and the brain membrane, undergoing similar conformational motion and changes to the SASA in both cases. Of the oncogenic mutants, only H1047R interacted with the membrane, and only in the presence of the brain membrane. While this might suggest that this mutant is more sensitive to the membrane composition, only one simulation of each system was carried out, and the initial orientation of each protein differed between the DPPC and Brain Membrane systems, thus no definitive conclusions can be drawn.

Overall, the regions of each protein that exhibited large-scale conformational change, as shown by peaks in the RMS fluctuation profiles, are in very good agreement with the regions in the protein identified as undergoing increased Hydrogen-Deuterium exchange, indicative of increased conformational motion and solvent exposure, by Burke et al. For those that contacted the membrane, the regions that showed a significant decrease in the SASA upon membrane interaction are also in keeping with the regions identified as contacting the membrane by Burke et al. This good agreement with the HDx data suggests that the simulations are correctly reproducing p110α's approach to and subsequent interaction with the membrane.

While a considerable level of structural and atomic-level insight has been gained into the solution dynamics of wild type p110α and its oncogenic mutants, and the interaction between wild type and H1047R p110α and a realistic brain membrane, the surface has but been scratched. There still remains a substantial amount of testing and analysis to be carried out to fully understand and assess the nature and consequences of this interaction.

Many factors remain untried; these include but are not limited to: multiple simulations of each system with different initial velocities, the starting orientation of the protein, the distance of the protein from the membrane, the length of the simulation, the composition of the membrane, and the effect of countless more mutations which were not in the scope of this study. In terms of analysis, more detailed insight into the nature of the conformational changes that occur in solution and upon membrane binding, including any differences that may occur between wild-type p110α and its oncogenic mutants, and between interaction with pure DPPC compared to with a realistic brain membrane, will be sought.

This research described here has set the scene for these future directions. Ultimately, once the effects of the aforementioned elements of research have been covered and the results of the simulations comprehensively analysed, the scientific community will be another step closer in the fight against cancer.

## Appendix I

## Lipid topology (itp) files

### POPE

```
[ moleculetype ]
; Name   nrexcl
POPE     3

[ atoms ]
;  nr  type  resnr  resid  atom  cgnr  charge    mass    total_charge
    1     H    1    POPE     H1    1    0.400   1.0080
    2     H    1    POPE     H2    1    0.400   1.0080
    3     H    1    POPE     H3    1    0.400   1.0008
    4    NL    1    POPE    NTM    1   -0.500  14.0067
    5   CH2    1    POPE     CA    1    0.300  14.0270      ;  1.000
    6   CH2    1    POPE     CB    2    0.400  14.0270
    7    OA    1    POPE     OA    2   -0.800  15.9994
    8     P    1    POPE      P    2    1.700  30.9738
    9    OM    1    POPE     OB    2   -0.800  15.9994
   10    OM    1    POPE     OC    2   -0.800  15.9994
   11    OA    1    POPE     OD    2   -0.700  15.9994      ; -1.000
   12   CH2    1    POPE     CC    3    0.400  14.0270
   13   CH1    1    POPE     CD    3    0.300  13.0190
   14    OE    1    POPE     OE    3   -0.700  15.9994
   15     C    1    POPE    C1A    3    0.700  12.0110
   16     O    1    POPE     OF    3   -0.700  15.9994      ;  0.000
   17   CH2    1    POPE    C1B    4    0.000  14.0270      ;  0.000
   18   CH2    1    POPE    C1C    5    0.000  14.0270      ;  0.000
   19   CH2    1    POPE    C1D    6    0.000  14.0270      ;  0.000
   20   CH2    1    POPE    C1E    7    0.000  14.0270      ;  0.000
   21   CH2    1    POPE    C1F    8    0.000  14.0270      ;  0.000
   22   CH2    1    POPE    C1G    9    0.000  14.0270      ;  0.000
   23   CH2    1    POPE    C1H   10    0.000  14.0270      ;  0.000
   24   CR1    1    POPE    C1I   11    0.000  13.0190      ;  0.000  ; double bond
   25   CR1    1    POPE    C1J   12    0.000  13.0190      ;  0.000  ; double bond
   26   CH2    1    POPE    C1K   13    0.000  14.0270      ;  0.000
   27   CH2    1    POPE    C1L   14    0.000  14.0270      ;  0.000
   28   CH2    1    POPE    C1M   15    0.000  14.0270      ;  0.000
   29   CH2    1    POPE    C1N   16    0.000  14.0270      ;  0.000
   30   CH2    1    POPE    C1O   17    0.000  14.0270      ;  0.000
   31   CH2    1    POPE    C1P   18    0.000  14.0270      ;  0.000
   32   CH2    1    POPE    C1Q   19    0.000  14.0270      ;  0.000
   33   CH3    1    POPE    C1R   20    0.000  15.0350      ;  0.000
   34   CH2    1    POPE     CE   21    0.500  14.0270
   35    OE    1    POPE     OG   21   -0.700  15.9994
   36     C    1    POPE    C2A   21    0.800  12.0110
   37     O    1    POPE     OH   21   -0.600  15.9994      ;  0.000
   38   CH2    1    POPE    C2B   22    0.000  14.0270      ;  0.000
   39   CH2    1    POPE    C2C   23    0.000  14.0270      ;  0.000
   40   CH2    1    POPE    C2D   24    0.000  14.0270      ;  0.000
   41   CH2    1    POPE    C2E   25    0.000  14.0270      ;  0.000
   42   CH2    1    POPE    C2F   26    0.000  14.0270      ;  0.000
   43   CH2    1    POPE    C2G   27    0.000  14.0270      ;  0.000
   44   CH2    1    POPE    C2H   28    0.000  14.0270      ;  0.000
   45   CH2    1    POPE    C2I   29    0.000  14.0270      ;  0.000
   46   CH2    1    POPE    C2J   30    0.000  14.0270      ;  0.000
   47   CH2    1    POPE    C2K   31    0.000  14.0270      ;  0.000
   48   CH2    1    POPE    C2L   32    0.000  14.0270      ;  0.000
   49   CH2    1    POPE    C2M   33    0.000  14.0270      ;  0.000
   50   CH2    1    POPE    C2N   34    0.000  14.0270      ;  0.000
   51   CH2    1    POPE    C2O   35    0.000  14.0270      ;  0.000
   52   CH3    1    POPE    C2P   36    0.000  15.0350      ;  0.000
; total charge of the molecule:   0.000

[ bonds ]
;  ai   aj  funct   c0         c1
    1    4    2   gb_2
    2    4    2   gb_2
    3    4    2   gb_2
    4    5    2   gb_21
    5    6    2   gb_27
    6    7    2   gb_18
    7    8    2   gb_28
```

```
     8    9    2   gb_24
     8   10    2   gb_24
     8   11    2   gb_28
    11   12    2   gb_18
    12   13    2   gb_27
    13   14    2   gb_18
    13   34    2   gb_27
    14   15    2   gb_10
    15   16    2   gb_5
    15   17    2   gb_23
    17   18    2   gb_27
    18   19    2   gb_27
    19   20    2   gb_27
    20   21    2   gb_27
    21   22    2   gb_27
    22   23    2   gb_27
    23   24    2   gb_27
    24   25    2   gb_10 ; double bond
    25   26    2   gb_27
    26   27    2   gb_27
    27   28    2   gb_27
    28   29    2   gb_27
    29   30    2   gb_27
    30   31    2   gb_27
    31   32    2   gb_27
    32   33    2   gb_27
    34   35    2   gb_18
    35   36    2   gb_10
    36   37    2   gb_5
    36   38    2   gb_23
    38   39    2   gb_27
    39   40    2   gb_27
    40   41    2   gb_27
    41   42    2   gb_27
    42   43    2   gb_27
    43   44    2   gb_27
    44   45    2   gb_27
    45   46    2   gb_27
    46   47    2   gb_27
    47   48    2   gb_27
    48   49    2   gb_27
    49   50    2   gb_27
    50   51    2   gb_27
    51   52    2   gb_27

[ pairs ]
;  ai   aj  funct  ;  all 1-4 pairs but the ones excluded in GROMOS itp
     4    7    1
     5    8    1
     6    9    1
     6   10    1
     6   11    1
     7   12    1
     8   13    1
     9   12    1
    10   12    1
    11   14    1
    11   34    1
    12   15    1
    12   35    1
    13   16    1
    13   17    1
    13   36    1
    14   18    1
    14   35    1
    15   19    1
    15   34    1
    16   18    1
    17   20    1
    18   21    1
    19   22    1
    20   23    1
    21   24    1
    22   25    1
    24   27    1
    25   28    1
    26   29    1
```

```
   27   30    1
   28   31    1
   29   32    1
   30   33    1
   34   37    1
   34   38    1
   35   39    1
   36   40    1
   37   39    1
   38   41    1
   39   42    1
   40   43    1
   41   44    1
   42   45    1
   43   46    1
   44   47    1
   45   48    1
   46   49    1
   47   50    1
   48   51    1
   49   52    1

[ angles ]
;  ai   aj   ak  funct   angle     fc
    1    4    2    2    ga_10
    2    4    3    2    ga_10
    3    4    1    2    ga_10
    1    4    5    2    ga_11
    2    4    5    2    ga_11
    3    4    5    2    ga_11
    4    5    6    2    ga_13
    5    6    7    2    ga_13
    6    7    8    2    ga_26
    7    8    9    2    ga_14
    7    8   10    2    ga_14
    7    8   11    2    ga_5
    8   11   12    2    ga_26
    9    8   10    2    ga_29
    9    8   11    2    ga_14
   10    8   11    2    ga_14
   11   12   13    2    ga_15
   12   13   14    2    ga_13
   12   13   34    2    ga_13
   13   14   15    2    ga_22
   13   34   35    2    ga_15
   14   13   34    2    ga_13
   14   15   16    2    ga_31
   14   15   17    2    ga_16
   15   17   18    2    ga_15
   16   15   17    2    ga_35
   17   18   19    2    ga_15
   18   19   20    2    ga_15
   19   20   21    2    ga_15
   20   21   22    2    ga_15
   21   22   23    2    ga_15
   22   23   24    2    ga_15
   23   24   25    2    ga_27 ; C1H-C1I=C1J
   24   25   26    2    ga_27 ; C1I=C1J-C1K
   25   26   27    2    ga_15
   26   27   28    2    ga_15
   27   28   29    2    ga_15
   28   29   30    2    ga_15
   29   30   31    2    ga_15
   30   31   32    2    ga_15
   31   32   33    2    ga_15
   34   35   36    2    ga_22
   35   36   37    2    ga_31
   35   36   38    2    ga_16
   36   38   39    2    ga_15
   37   36   38    2    ga_35
   38   39   40    2    ga_15
   39   40   41    2    ga_15
   40   41   42    2    ga_15
   41   42   43    2    ga_15
   42   43   44    2    ga_15
   43   44   45    2    ga_15
   44   45   46    2    ga_15
```

```
   45   46   47    2    ga_15
   46   47   48    2    ga_15
   47   48   49    2    ga_15
   48   49   50    2    ga_15
   49   50   51    2    ga_15
   50   51   52    2    ga_15

[ dihedrals ]
; GROMOS improper dihedrals
;  ai   aj   ak   al  funct   angle     fc
   13   14   34   12    2     gi_2 ; asymmetric atom CD
   15   14   16   17    2     gi_1 ; planarity of sn-2 carbonyl
   23   24   25   26    2     gi_1 ; planarity of the double bond
   36   35   37   38    2     gi_1 ; planarity of sn-1 carbonyl

[ dihedrals ]
;  ai   aj   ak   al  funct    ph0      cp     mult
    1    4    5    6    1      gd_29
    4    5    6    7    1      gd_4
    5    6    7    8    1      gd_29
    6    7    8   11    1      gd_20
    6    7    8   11    1      gd_27
    7    8   11   12    1      gd_20
    7    8   11   12    1      gd_27
    8   11   12   13    1      gd_29
   11   12   13   34    1      gd_34
   12   13   14   15    1      gd_29
   12   13   34   35    1      gd_34
   13   14   15   17    1      gd_13
   13   34   35   36    1      gd_29
   14   15   17   18    1      gd_40
   15   17   18   19    1      gd_34
   17   18   19   20    1      gd_34
   18   19   20   21    1      gd_34
   19   20   21   22    1      gd_34
   20   21   22   23    1      gd_34
   21   22   23   24    1      gd_34
   22   23   24   25    1      gd_40
   24   25   26   27    1      gd_40
   25   26   27   28    1      gd_34
   26   27   28   29    1      gd_34
   27   28   29   30    1      gd_34
   28   29   30   31    1      gd_34
   29   30   31   32    1      gd_34
   30   31   32   33    1      gd_34
   34   35   36   38    1      gd_13
   35   36   38   39    1      gd_40
   36   38   39   40    1      gd_34
   38   39   40   41    1      gd_34
   39   40   41   42    1      gd_34
   40   41   42   43    1      gd_34
   41   42   43   44    1      gd_34
   42   43   44   45    1      gd_34
   43   44   45   46    1      gd_34
   44   45   46   47    1      gd_34
   45   46   47   48    1      gd_34
   46   47   48   49    1      gd_34
   47   48   49   50    1      gd_34
   48   49   50   51    1      gd_34
   49   50   51   52    1      gd_34
```

## POPS

```
[ moleculetype ]
; Name   nrexcl
POPS     3

[ atoms ]
;  nr  type  resnr  resid  atom  cgnr  charge    mass    total_charge
    1     H    1    POPS     H1    1    0.400   1.0080
    2     H    1    POPS     H2    1    0.400   1.0080
    3     H    1    POPS     H3    1    0.400   1.0008
    4    NL    1    POPS    NTM    1   -0.500  14.0067
    5   CH1    1    POPS     CA    1    0.300  14.0270      ;  1.000
    6   CH2    1    POPS     CB    2    0.400  14.0270
    7    OA    1    POPS     OA    2   -0.800  15.9994
    8     P    1    POPS      P    2    1.700  30.9738
    9    OM    1    POPS     OB    2   -0.800  15.9994
   10    OM    1    POPS     OC    2   -0.800  15.9994
   11    OA    1    POPS     OD    2   -0.700  15.9994      ; -1.000
   12   CH2    1    POPS     CC    3    0.400  14.0270
   13   CH1    1    POPS     CD    3    0.300  13.0190
   14    OE    1    POPS     OE    3   -0.700  15.9994
   15     C    1    POPS    C1A    3    0.700  12.0110
   16     O    1    POPS     OF    3   -0.700  15.9994      ;  0.000
   17   CH2    1    POPS    C1B    4    0.000  14.0270      ;  0.000
   18   CH2    1    POPS    C1C    5    0.000  14.0270      ;  0.000
   19   CH2    1    POPS    C1D    6    0.000  14.0270      ;  0.000
   20   CH2    1    POPS    C1E    7    0.000  14.0270      ;  0.000
   21   CH2    1    POPS    C1F    8    0.000  14.0270      ;  0.000
   22   CH2    1    POPS    C1G    9    0.000  14.0270      ;  0.000
   23   CH2    1    POPS    C1H   10    0.000  14.0270      ;  0.000
   24   CR1    1    POPS    C1I   11    0.000  13.0190      ;  0.000  ; double bond
   25   CR1    1    POPS    C1J   12    0.000  13.0190      ;  0.000  ; double bond
   26   CH2    1    POPS    C1K   13    0.000  14.0270      ;  0.000
   27   CH2    1    POPS    C1L   14    0.000  14.0270      ;  0.000
   28   CH2    1    POPS    C1M   15    0.000  14.0270      ;  0.000
   29   CH2    1    POPS    C1N   16    0.000  14.0270      ;  0.000
   30   CH2    1    POPS    C1O   17    0.000  14.0270      ;  0.000
   31   CH2    1    POPS    C1P   18    0.000  14.0270      ;  0.000
   32   CH2    1    POPS    C1Q   19    0.000  14.0270      ;  0.000
   33   CH3    1    POPS    C1R   20    0.000  15.0350      ;  0.000
   34   CH2    1    POPS     CE   21    0.500  14.0270
   35    OE    1    POPS     OG   21   -0.700  15.9994
   36     C    1    POPS    C2A   21    0.800  12.0110
   37     O    1    POPS     OH   21   -0.600  15.9994      ;  0.000
   38   CH2    1    POPS    C2B   22    0.000  14.0270      ;  0.000
   39   CH2    1    POPS    C2C   23    0.000  14.0270      ;  0.000
   40   CH2    1    POPS    C2D   24    0.000  14.0270      ;  0.000
   41   CH2    1    POPS    C2E   25    0.000  14.0270      ;  0.000
   42   CH2    1    POPS    C2F   26    0.000  14.0270      ;  0.000
   43   CH2    1    POPS    C2G   27    0.000  14.0270      ;  0.000
   44   CH2    1    POPS    C2H   28    0.000  14.0270      ;  0.000
   45   CH2    1    POPS    C2I   29    0.000  14.0270      ;  0.000
   46   CH2    1    POPS    C2J   30    0.000  14.0270      ;  0.000
   47   CH2    1    POPS    C2K   31    0.000  14.0270      ;  0.000
   48   CH2    1    POPS    C2L   32    0.000  14.0270      ;  0.000
   49   CH2    1    POPS    C2M   33    0.000  14.0270      ;  0.000
   50   CH2    1    POPS    C2N   34    0.000  14.0270      ;  0.000
   51   CH2    1    POPS    C2O   35    0.000  14.0270      ;  0.000
   52   CH3    1    POPS    C2P   36    0.000  15.0350      ;  0.000
   53     C    1    POPS     CF   37    0.600  12.0110
   54    OM    1    POPS     OI   37   -0.700  15.9994
   55    OM    1    POPS     OJ   37   -0.900  15.9994      ; -1.000
; total charge of the molecule:  -1.000

[ bonds ]
;  ai   aj  funct   c0         c1
    1    4    2   gb_2
    2    4    2   gb_2
    3    4    2   gb_2
    4    5    2   gb_21
    5    6    2   gb_27
    5   53    2   gb_3
    6    7    2   gb_18
    7    8    2   gb_28
    8    9    2   gb_24
    8   10    2   gb_24
```

```
     8    11     2    gb_28
    11    12     2    gb_18
    12    13     2    gb_27
    13    14     2    gb_18
    13    34     2    gb_27
    14    15     2    gb_10
    15    16     2    gb_5
    15    17     2    gb_23
    17    18     2    gb_27
    18    19     2    gb_27
    19    20     2    gb_27
    20    21     2    gb_27
    21    22     2    gb_27
    22    23     2    gb_27
    23    24     2    gb_27
    24    25     2    gb_10 ; double bond
    25    26     2    gb_27
    26    27     2    gb_27
    27    28     2    gb_27
    28    29     2    gb_27
    29    30     2    gb_27
    30    31     2    gb_27
    31    32     2    gb_27
    32    33     2    gb_27
    34    35     2    gb_18
    35    36     2    gb_10
    36    37     2    gb_5
    36    38     2    gb_23
    38    39     2    gb_27
    39    40     2    gb_27
    40    41     2    gb_27
    41    42     2    gb_27
    42    43     2    gb_27
    43    44     2    gb_27
    44    45     2    gb_27
    45    46     2    gb_27
    46    47     2    gb_27
    47    48     2    gb_27
    48    49     2    gb_27
    49    50     2    gb_27
    50    51     2    gb_27
    51    52     2    gb_27
    53    54     2    gb_6
    53    55     2    gb_6

[ pairs ]
;  ai    aj  funct  ;  all 1-4 pairs but the ones excluded in GROMOS itp
     4     7     1
     5     8     1
     6     9     1
     6    10     1
     6    11     1
     7    12     1
     8    13     1
     9    12     1
    10    12     1
    11    14     1
    11    34     1
    12    15     1
    12    35     1
    13    16     1
    13    17     1
    13    36     1
    14    18     1
    14    35     1
    15    19     1
    15    34     1
    16    18     1
    17    20     1
    18    21     1
    19    22     1
    20    23     1
    21    24     1
    22    25     1
    24    27     1
    25    28     1
    26    29     1
```

```
   27   30    1
   28   31    1
   29   32    1
   30   33    1
   34   37    1
   34   38    1
   35   39    1
   36   40    1
   37   39    1
   38   41    1
   39   42    1
   40   43    1
   41   44    1
   42   45    1
   43   46    1
   44   47    1
   45   48    1
   46   49    1
   47   50    1
   48   51    1
   49   52    1
   53    7    1
   54    4    1
   54    6    1
   55    4    1
   55    6    1

[ angles ]
;  ai   aj   ak  funct   angle     fc
    1    4    2    2    ga_10
    2    4    3    2    ga_10
    3    4    1    2    ga_10
    1    4    5    2    ga_11
    2    4    5    2    ga_11
    3    4    5    2    ga_11
    4    5    6    2    ga_13
    5    6    7    2    ga_13
    6    7    8    2    ga_26
    7    8    9    2    ga_14
    7    8   10    2    ga_14
    7    8   11    2    ga_5
    8   11   12    2    ga_26
    9    8   10    2    ga_29
    9    8   11    2    ga_14
   10    8   11    2    ga_14
   11   12   13    2    ga_15
   12   13   14    2    ga_13
   12   13   34    2    ga_13
   13   14   15    2    ga_22
   13   34   35    2    ga_15
   14   13   34    2    ga_13
   14   15   16    2    ga_31
   14   15   17    2    ga_16
   15   17   18    2    ga_15
   16   15   17    2    ga_35
   17   18   19    2    ga_15
   18   19   20    2    ga_15
   19   20   21    2    ga_15
   20   21   22    2    ga_15
   21   22   23    2    ga_15
   22   23   24    2    ga_15
   23   24   25    2    ga_27 ; C1H-C1I=C1J
   24   25   26    2    ga_27 ; C1I=C1J-C1K
   25   26   27    2    ga_15
   26   27   28    2    ga_15
   27   28   29    2    ga_15
   28   29   30    2    ga_15
   29   30   31    2    ga_15
   30   31   32    2    ga_15
   31   32   33    2    ga_15
   34   35   36    2    ga_22
   35   36   37    2    ga_31
   35   36   38    2    ga_16
   36   38   39    2    ga_15
   37   36   38    2    ga_35
   38   39   40    2    ga_15
   39   40   41    2    ga_15
```

```
   40   41   42    2    ga_15
   41   42   43    2    ga_15
   42   43   44    2    ga_15
   43   44   45    2    ga_15
   44   45   46    2    ga_15
   45   46   47    2    ga_15
   46   47   48    2    ga_15
   47   48   49    2    ga_15
   48   49   50    2    ga_15
   49   50   51    2    ga_15
   50   51   52    2    ga_15
   54   53   55    2    ga_38
   54   53    5    2    ga_22
   55   53    5    2    ga_22

[ dihedrals ]
; GROMOS improper dihedrals
;  ai   aj   ak   al  funct   angle     fc
   13   14   34   12    2     gi_2 ; asymmetric atom CD
   15   14   16   17    2     gi_1 ; planarity of sn-2 carbonyl
   23   24   25   26    2     gi_1 ; planarity of the double bond
   36   35   37   38    2     gi_1 ; planarity of sn-1 carbonyl
   53    5   54   55    2     gi_1

[ dihedrals ]
;  ai   aj   ak   al  funct    ph0      cp     mult
    1    4    5    6    1      gd_29
    4    5    6    7    1      gd_4
    5    6    7    8    1      gd_29
    6    7    8   11    1      gd_20
    6    7    8   11    1      gd_27
    7    8   11   12    1      gd_20
    7    8   11   12    1      gd_27
    8   11   12   13    1      gd_29
   11   12   13   34    1      gd_34
   12   13   14   15    1      gd_29
   12   13   34   35    1      gd_34
   13   14   15   17    1      gd_13
   13   34   35   36    1      gd_29
   14   15   17   18    1      gd_40
   15   17   18   19    1      gd_34
   17   18   19   20    1      gd_34
   18   19   20   21    1      gd_34
   19   20   21   22    1      gd_34
   20   21   22   23    1      gd_34
   21   22   23   24    1      gd_34
   22   23   24   25    1      gd_40
   24   25   26   27    1      gd_40
   25   26   27   28    1      gd_34
   26   27   28   29    1      gd_34
   27   28   29   30    1      gd_34
   28   29   30   31    1      gd_34
   29   30   31   32    1      gd_34
   30   31   32   33    1      gd_34
   34   35   36   38    1      gd_13
   35   36   38   39    1      gd_40
   36   38   39   40    1      gd_34
   38   39   40   41    1      gd_34
   39   40   41   42    1      gd_34
   40   41   42   43    1      gd_34
   41   42   43   44    1      gd_34
   42   43   44   45    1      gd_34
   43   44   45   46    1      gd_34
   44   45   46   47    1      gd_34
   45   46   47   48    1      gd_34
   46   47   48   49    1      gd_34
   47   48   49   50    1      gd_34
   48   49   50   51    1      gd_34
   49   50   51   52    1      gd_34
    4    5   53   54    1      gd_45
    4    5   53   54    1      gd_42
```

**PIP2**

```
[ moleculetype ]
; Name   nrexcl
PIP2     3

[ atoms ]
;  nr  type  resnr  resid  atom  cgnr  charge    mass    total_charge
    1   CH1    1    PIP2     C1    1    0.400  13.0190
    2    OA    1    PIP2     O1    1   -0.800  15.9994
    3     P    1    PIP2     P1    1    1.700  30.9738
    4    OM    1    PIP2    O11    1   -0.800  15.9994
    5    OM    1    PIP2    O12    1   -0.800  15.9994
    6    OA    1    PIP2    O13    1   -0.700  15.9994      ; -1.000
    7   CH2    1    PIP2    C1C    2    0.400  14.0270
    8   CH1    1    PIP2    C2C    2    0.300  13.0190
    9    OE    1    PIP2    O2C    2   -0.700  15.9994
   10     C    1    PIP2    C1A    2    0.700  12.0110
   11     O    1    PIP2    O1A    2   -0.700  15.9994      ;  0.000
   12   CH2    1    PIP2    C2A    3    0.000  14.0270      ;  0.000
   13   CH2    1    PIP2    C3A    4    0.000  14.0270      ;  0.000
   14   CH2    1    PIP2    C4A    5    0.000  14.0270      ;  0.000
   15   CH2    1    PIP2    C5A    6    0.000  14.0270      ;  0.000
   16   CH2    1    PIP2    C6A    7    0.000  14.0270      ;  0.000
   17   CH2    1    PIP2    C7A    8    0.000  14.0270      ;  0.000
   18   CH2    1    PIP2    C8A    9    0.000  14.0270      ;  0.000
   19   CR1    1    PIP2    C9A   10    0.000  13.0190      ;  0.000  ; double bond
   20   CR1    1    PIP2   C10A   11    0.000  13.0190      ;  0.000  ; double bond
   21   CH2    1    PIP2   C11A   12    0.000  14.0270      ;  0.000
   22   CH2    1    PIP2   C12A   13    0.000  14.0270      ;  0.000
   23   CH2    1    PIP2   C13A   14    0.000  14.0270      ;  0.000
   24   CH2    1    PIP2   C14A   15    0.000  14.0270      ;  0.000
   25   CH2    1    PIP2   C15A   16    0.000  14.0270      ;  0.000
   26   CH2    1    PIP2   C16A   17    0.000  14.0270      ;  0.000
   27   CH2    1    PIP2   C17A   18    0.000  14.0270      ;  0.000
   28   CH3    1    PIP2   C18A   19    0.000  15.0350      ;  0.000
   29   CH2    1    PIP2    C3C   20    0.500  14.0270
   30    OE    1    PIP2    O3C   20   -0.700  15.9994
   31     C    1    PIP2    C1B   20    0.800  12.0110
   32     O    1    PIP2    O1B   20   -0.600  15.9994      ;  0.000
   33   CH2    1    PIP2    C2B   21    0.000  14.0270      ;  0.000
   34   CH2    1    PIP2    C3B   22    0.000  14.0270      ;  0.000
   35   CH2    1    PIP2    C4B   23    0.000  14.0270      ;  0.000
   36   CH2    1    PIP2    C5B   24    0.000  14.0270      ;  0.000
   37   CH2    1    PIP2    C6B   25    0.000  14.0270      ;  0.000
   38   CH2    1    PIP2    C7B   26    0.000  14.0270      ;  0.000
   39   CH2    1    PIP2    C8B   27    0.000  14.0270      ;  0.000
   40   CH2    1    PIP2    C9B   28    0.000  14.0270      ;  0.000
   41   CH2    1    PIP2   C10B   29    0.000  14.0270      ;  0.000
   42   CH2    1    PIP2   C11B   30    0.000  14.0270      ;  0.000
   43   CH2    1    PIP2   C12B   31    0.000  14.0270      ;  0.000
   44   CH2    1    PIP2   C13B   32    0.000  14.0270      ;  0.000
   45   CH2    1    PIP2   C14B   33    0.000  14.0270      ;  0.000
   46   CH2    1    PIP2   C15B   34    0.000  14.0270      ;  0.000
   47   CH3    1    PIP2   C16B   35    0.000  15.0350      ;  0.000
   48   CH1    1    PIP2     C2   36    0.300  13.0190
   49    OA    1    PIP2     O2   36   -0.700  15.9994
   50     H    1    PIP2     H2   36    0.400   1.0080      ;  0.000
   51   CH1    1    PIP2     C3   37    0.300  13.0190
   52    OA    1    PIP2     O3   37   -0.700  15.9994
   53     H    1    PIP2     H3   37    0.400   1.0080      ;  0.000
   54   CH1    1    PIP2     C4   38    0.300  13.0190
   55    OA    1    PIP2     O4   38   -0.800  15.9994
   56     P    1    PIP2     P4   38    0.900  30.9738
   57    OM    1    PIP2    O41   38   -0.800  15.9994
   58    OM    1    PIP2    O42   38   -0.800  15.9994
   59    OM    1    PIP2    O43   38   -0.800  15.9994      ; -2.000
   60   CH1    1    PIP2     C5   39    0.300  13.0190
   61    OA    1    PIP2     O5   39   -0.800  15.9994
   62     P    1    PIP2     P5   39    0.900  30.9738
   63    OM    1    PIP2    O53   39   -0.800  15.9994
   64    OM    1    PIP2    O52   39   -0.800  15.9994
   65    OM    1    PIP2    O51   39   -0.800  15.9994      ; -2.000
   66   CH1    1    PIP2     C6   40    0.300  13.0190
   67    OA    1    PIP2     O6   40   -0.700  15.9994
   68     H    1    PIP2     H6   40    0.400   1.0080      ;  0.000
; total charge of the molecule:  -5.000
```

```
[ bonds ]
;  ai    aj  funct   c0         c1
    1   48    2   gb_26
    1   66    2   gb_26
    1    2    2   gb_18
    2    3    2   gb_28
    3    4    2   gb_24
    3    5    2   gb_24
    3    6    2   gb_28
    6    7    2   gb_18
    7    8    2   gb_27
    8    9    2   gb_18
    8   29    2   gb_27
    9   10    2   gb_10
   10   11    2   gb_5
   10   12    2   gb_23
   12   13    2   gb_27
   13   14    2   gb_27
   14   15    2   gb_27
   15   16    2   gb_27
   16   17    2   gb_27
   17   18    2   gb_27
   18   19    2   gb_27
   19   20    2   gb_10 ; double bond
   20   21    2   gb_27
   21   22    2   gb_27
   22   23    2   gb_27
   23   24    2   gb_27
   24   25    2   gb_27
   25   26    2   gb_27
   26   27    2   gb_27
   27   28    2   gb_27
   29   30    2   gb_18
   30   31    2   gb_10
   31   32    2   gb_5
   31   33    2   gb_23
   33   34    2   gb_27
   34   35    2   gb_27
   35   36    2   gb_27
   36   37    2   gb_27
   37   38    2   gb_27
   38   39    2   gb_27
   39   40    2   gb_27
   40   41    2   gb_27
   41   42    2   gb_27
   42   43    2   gb_27
   43   44    2   gb_27
   44   45    2   gb_27
   45   46    2   gb_27
   46   47    2   gb_27
   48   51    2   gb_26
   48   49    2   gb_20
   49   50    2   gb_1
   51   54    2   gb_26
   51   52    2   gb_20
   52   53    2   gb_1
   54   60    2   gb_26
   54   55    2   gb_18
   55   56    2   gb_28
   56   57    2   gb_24
   56   58    2   gb_24
   56   59    2   gb_24
   60   66    2   gb_26
   60   61    2   gb_18
   61   62    2   gb_28
   62   63    2   gb_24
   62   64    2   gb_24
   62   65    2   gb_24
   66   67    2   gb_20
   67   68    2   gb_1

[ pairs ]
;  ai    aj  funct  ;  all 1-4 pairs but the ones excluded in GROMOS itp
    1   61    1
    1   52    1
    1   54    1
```

```
 1    4    1
 1    5    1
 1    6    1
 2   49    1
 2   67    1
 2   51    1
 2   60    1
 2    7    1
 3   48    1
 3   66    1
 3    8    1
 4    7    1
 5    7    1
 6    9    1
 6   29    1
 7   10    1
 7   30    1
 8   11    1
 8   12    1
 8   31    1
 9   13    1
 9   30    1
10   29    1
10   14    1
11   13    1
12   15    1
13   16    1
14   17    1
15   18    1
16   19    1
17   20    1
18   21    1
19   22    1
20   23    1
21   24    1
22   25    1
23   26    1
24   27    1
25   28    1
29   32    1
29   33    1
30   34    1
31   35    1
32   34    1
33   36    1
34   37    1
35   38    1
36   39    1
37   40    1
38   41    1
39   42    1
40   43    1
41   44    1
42   45    1
43   46    1
44   47    1
48   60    1
48   55    1
48   67    1
49   52    1
49   54    1
49   66    1
51   66    1
51   61    1
51   56    1
52   55    1
52   60    1
54   67    1
54   62    1
54   57    1
54   58    1
54   59    1
55   61    1
55   66    1
56   60    1
60   63    1
60   64    1
```

```
   60   65    1
   61   67    1
   62   66    1

[ angles ]
;  ai   aj   ak  funct   angle     fc
    2    1   48    2    ga_9
    2    1   66    2    ga_9
   48    1   66    2    ga_8
    1    2    3    2    ga_26
    2    3    4    2    ga_14
    2    3    5    2    ga_14
    2    3    6    2    ga_5
    4    3    5    2    ga_29
    4    3    6    2    ga_14
    5    3    6    2    ga_14
    3    6    7    2    ga_26
    6    7    8    2    ga_15
    7    8    9    2    ga_13
    7    8   29    2    ga_13
    9    8   29    2    ga_13
    8    9   10    2    ga_22
    9   10   11    2    ga_31
    9   10   12    2    ga_16
   11   10   12    2    ga_35
   10   12   13    2    ga_15
   12   13   14    2    ga_15
   13   14   15    2    ga_15
   14   15   16    2    ga_15
   15   16   17    2    ga_15
   16   17   18    2    ga_15
   17   18   19    2    ga_15
   18   19   20    2    ga_27
   19   20   21    2    ga_27
   20   21   22    2    ga_15
   21   22   23    2    ga_15
   22   23   24    2    ga_15
   23   24   25    2    ga_15
   24   25   26    2    ga_15
   25   26   27    2    ga_15
   26   27   28    2    ga_15
    8   29   30    2    ga_15
   29   30   31    2    ga_22
   30   31   32    2    ga_31
   30   31   33    2    ga_16
   32   31   33    2    ga_35
   31   33   34    2    ga_15
   33   34   35    2    ga_15
   34   35   36    2    ga_15
   35   36   37    2    ga_15
   36   37   38    2    ga_15
   37   38   39    2    ga_15
   38   39   40    2    ga_15
   39   40   41    2    ga_15
   40   41   42    2    ga_15
   41   42   43    2    ga_15
   42   43   44    2    ga_15
   43   44   45    2    ga_15
   44   45   46    2    ga_15
   45   46   47    2    ga_15
    1   48   49    2    ga_9
    1   48   51    2    ga_8
   49   48   51    2    ga_9
   48   49   50    2    ga_12
   48   51   52    2    ga_9
   48   51   54    2    ga_8
   52   51   54    2    ga_9
   51   52   53    2    ga_12
   51   54   55    2    ga_9
   51   54   60    2    ga_8
   55   54   60    2    ga_9
   54   55   56    2    ga_26
   55   56   57    2    ga_14
   55   56   58    2    ga_14
   55   56   59    2    ga_14
   57   56   58    2    ga_29
   57   56   59    2    ga_29
```

```
   58   56   59    2    ga_29
   54   60   61    2    ga_9
   54   60   66    2    ga_8
   61   60   66    2    ga_9
   60   61   62    2    ga_26
   61   62   63    2    ga_14
   61   62   64    2    ga_14
   61   62   65    2    ga_14
   63   62   64    2    ga_29
   63   62   65    2    ga_29
   64   62   65    2    ga_29
    1   66   60    2    ga_8
    1   66   67    2    ga_9
   60   66   67    2    ga_9
   66   67   68    2    ga_12

[ dihedrals ]
; GROMOS improper dihedrals
;  ai   aj   ak   al  funct   angle     fc
    1   66    2   48    2     gi_2
    1   60   67   66    2     gi_2
    7   29    9    8    2     gi_2
   10    9   12   11    2     gi_1
   31   30   33   32    2     gi_1
   18   19   20   21    2     gi_1
   48    1   49   51    2     gi_2
   51   48   52   54    2     gi_2
   54   55   51   60    2     gi_2
   60   54   61   66    2     gi_2

[ dihedrals ]
;  ai   aj   ak   al  funct    ph0      cp     mult
   66    1    2    3    1      gd_30
    2    1   48   49    1      gd_18
    2    1   48   51    1      gd_17
   66    1   48   49    1      gd_17
   66    1   48   51    1      gd_34
    2    1   66   60    1      gd_17
   48    1   66   60    1      gd_34
   48    1   66   67    1      gd_17
    1    2    3    6    1      gd_20
    1    2    3    6    1      gd_27
    2    3    6    7    1      gd_20
    2    3    6    7    1      gd_27
    3    6    7    8    1      gd_29
    6    7    8   29    1      gd_34
    7    8    9   10    1      gd_29
    7    8   29   30    1      gd_34
    8    9   10   12    1      gd_13
    9   10   12   13    1      gd_40
   10   12   13   14    1      gd_34
   12   13   14   15    1      gd_34
   13   14   15   16    1      gd_34
   14   15   16   17    1      gd_34
   15   16   17   18    1      gd_34
   16   17   18   19    1      gd_34
   17   18   19   20    1      gd_40
   19   20   21   22    1      gd_40
   20   21   22   23    1      gd_34
   21   22   23   24    1      gd_34
   22   23   24   25    1      gd_34
   23   24   25   26    1      gd_34
   24   25   26   27    1      gd_34
   25   26   27   28    1      gd_34
    8   29   30   31    1      gd_29
   29   30   31   33    1      gd_13
   30   31   33   34    1      gd_40
   31   33   34   35    1      gd_34
   33   34   35   36    1      gd_34
   34   35   36   37    1      gd_34
   35   36   37   38    1      gd_34
   36   37   38   39    1      gd_34
   37   38   39   40    1      gd_34
   38   39   40   41    1      gd_34
   39   40   41   42    1      gd_34
   40   41   42   43    1      gd_34
   41   42   43   44    1      gd_34
```

```
   42   43   44   45    1       gd_34
   43   44   45   46    1       gd_34
   44   45   46   47    1       gd_34
    1   48   49   50    1       gd_30
    1   48   51   52    1       gd_17
    1   48   51   54    1       gd_34
   49   48   51   52    1       gd_18
   49   48   51   54    1       gd_17
   48   51   52   53    1       gd_30
   48   51   54   55    1       gd_17
   48   51   54   60    1       gd_34
   52   51   54   60    1       gd_17
   51   54   55   56    1       gd_30
   51   54   60   61    1       gd_17
   51   54   60   66    1       gd_34
   55   54   60   61    1       gd_18
   55   54   60   66    1       gd_17
   54   55   56   57    1       gd_22
   54   60   61   62    1       gd_30
   54   60   66    1    1       gd_34
   54   60   66   67    1       gd_17
   61   60   66    1    1       gd_17
   61   60   66   67    1       gd_18
   60   61   62   63    1       gd_22
   60   66   67   68    1       gd_30
```

**SGML**

```
[ moleculetype ]
; Name            nrexcl
SGML               3

[ atoms ]
;   nr  type  resnr  resid  atom  cgnr  charge    mass    total_charge
     1  CH3p     1    SGML     C1     1    0.400  15.0350
     2  CH3p     1    SGML     C2     2    0.400  15.0350
     3  CH3p     1    SGML     C3     3    0.400  15.0350
     4    NL     1    SGML     N4     4   -0.500  14.0067
     5   CH2     1    SGML     C5     5    0.300  14.0270      ;  1.000
     6   CH2     1    SGML     C6     6    0.400  14.0270
     7    OA     1    SGML    OS7     7   -0.800  15.9994
     8     P     1    SGML     P8     8    1.700  30.9738
     9    OM     1    SGML    OM9     9   -0.800  15.9994
    10    OM     1    SGML   OM10    10   -0.800  15.9994
    11    OA     1    SGML   OS11    11   -0.700  15.9994      ;  0.000
    12   CH2     1    SGML    C12    12    0.250  14.0270
    13   CH1     1    SGML    C13    13    0.150  13.0190
    14     N     1    SGML    N14    14   -0.850  14.0067
    15     H     1    SGML    H15    15    0.450   1.0080      ;  0.000
    16     C     1    SGML    C16    16    0.700  12.0110
    17     O     1    SGML    O17    17   -0.700  15.9994
    18   CH2     1    SGML    C18    18    0.000  14.0270      ;  0.000
    19   CH2     1    SGML    C19    18    0.000  14.0270      ;  0.000
    20   CH2     1    SGML    C20    18    0.000  14.0270      ;  0.000
    21   CH2     1    SGML    C21    19    0.000  14.0270      ;  0.000
    22   CH2     1    SGML    C22    19    0.000  14.0270      ;  0.000
    23   CH2     1    SGML    C23    19    0.000  14.0270      ;  0.000
    24   CH2     1    SGML    C24    20    0.000  14.0270      ;  0.000
    25   CH2     1    SGML    C25    20    0.000  14.0270      ;  0.000
    26   CH2     1    SGML    C26    20    0.000  14.0270      ;  0.000
    27   CH2     1    SGML    C27    21    0.000  14.0270      ;  0.000
    28   CH2     1    SGML    C28    21    0.000  14.0270      ;  0.000
    29   CH2     1    SGML    C29    21    0.000  14.0270      ;  0.000
    30   CH2     1    SGML    C30    22    0.000  14.0270      ;  0.000
    31   CH2     1    SGML    C31    22    0.000  14.0270      ;  0.000
    32   CH2     1    SGML    C32    22    0.000  14.0270      ;  0.000
    33   CH2     1    SGML    C33    23    0.000  14.0270      ;  0.000
    34   CH3     1    SGML    C34    23    0.000  15.0350      ;  0.000
    35   CH1     1    SGML    C35    24    0.325  13.0190
    36    OA     1    SGML   OA36    25   -0.760  15.9994
    37     H     1    SGML    H37    26    0.435   1.0080      ;  0.000
    38   CH1     1    SGML    C38    27    0.000  13.0190      ;  0.000
    39   CH1     1    SGML    C39    27    0.000  13.0190      ;  0.000
    40   CH2     1    SGML    C40    27    0.000  14.0270      ;  0.000
    41   CH2     1    SGML    C41    28    0.000  14.0270      ;  0.000
    42   CH2     1    SGML    C42    28    0.000  14.0270      ;  0.000
```

```
   43   CH2    1    SGML    C43   28    0.000  14.0270      ;  0.000
   44   CH2    1    SGML    C44   29    0.000  14.0270      ;  0.000
   45   CH2    1    SGML    C45   29    0.000  14.0270      ;  0.000
   46   CH2    1    SGML    C46   29    0.000  14.0270      ;  0.000
   47   CH2    1    SGML    C47   30    0.000  14.0270      ;  0.000
   48   CH2    1    SGML    C48   30    0.000  14.0270      ;  0.000
   49   CH2    1    SGML    C49   30    0.000  14.0270      ;  0.000
   50   CH2    1    SGML    C50   31    0.000  14.0270      ;  0.000
   51   CH2    1    SGML    C51   31    0.000  14.0270      ;  0.000
   52   CH3    1    SGML    C52   31    0.000  15.0350      ;  0.000
; total charge of the molecule:   0.000

[ bonds ]
;  ai   aj  funct   c0         c1
    1    4    2   gb_21
    2    4    2   gb_21
    3    4    2   gb_21
    4    5    2   gb_21
    5    6    2   gb_27
    6    7    2   gb_18
    7    8    2   gb_28
    8    9    2   gb_24
    8   10    2   gb_24
    8   11    2   gb_28
   11   12    2   gb_18
   12   13    2   gb_27
   13   14    2   gb_21
   13   35    2   gb_27
   14   15    2   gb_2
   14   16    2   gb_11
   16   17    2   gb_5
   16   18    2   gb_27
   18   19    2   gb_27
   19   20    2   gb_27
   20   21    2   gb_27
   21   22    2   gb_27
   22   23    2   gb_27
   23   24    2   gb_27
   24   25    2   gb_27
   25   26    2   gb_27
   26   27    2   gb_27
   27   28    2   gb_27
   28   29    2   gb_27
   29   30    2   gb_27
   30   31    2   gb_27
   31   32    2   gb_27
   32   33    2   gb_27
   33   34    2   gb_27
   35   36    2   gb_18
   35   38    2   gb_53
   36   37    2   gb_1
   38   39    2   gb_10
   39   40    2   gb_53
   40   41    2   gb_27
   41   42    2   gb_27
   42   43    2   gb_27
   43   44    2   gb_27
   44   45    2   gb_27
   45   46    2   gb_27
   46   47    2   gb_27
   47   48    2   gb_27
   48   49    2   gb_27
   49   50    2   gb_27
   50   51    2   gb_27
   51   52    2   gb_27

[ pairs ]
;  ai   aj  funct  ;  all 1-4 pairs but the ones excluded in GROMOS itp
    1    6    1
    2    6    1
    3    6    1
    4    7    1
    5    8    1
    6    9    1
    6   10    1
    6   11    1
    7   12    1
```

```
    8   13    1
    9   12    1
   10   12    1
   11   14    1
   11   35    1
   12   36    1
   14   36    1
   14   38    1
   15   35    1
   16   35    1
   35   40    1
   36   39    1

[ angles ]
;  ai   aj   ak  funct   angle     fc
    1    4    2    2    ga_13
    1    4    3    2    ga_13
    1    4    5    2    ga_13
    2    4    3    2    ga_13
    2    4    5    2    ga_13
    3    4    5    2    ga_13
    4    5    6    2    ga_15
    5    6    7    2    ga_15
    6    7    8    2    ga_26
    7    8    9    2    ga_14
    7    8   10    2    ga_14
    7    8   11    2    ga_5
    9    8   10    2    ga_29
    9    8   11    2    ga_14
   10    8   11    2    ga_14
    8   11   12    2    ga_26
   11   12   13    2    ga_15
   12   13   14    2    ga_13
   12   13   35    2    ga_13
   14   13   35    2    ga_13
   13   14   15    2    ga_25
   13   14   16    2    ga_31
   15   14   16    2    ga_20
   14   16   17    2    ga_33
   14   16   18    2    ga_19
   17   16   18    2    ga_35
   16   18   19    2    ga_15
   18   19   20    2    ga_15
   19   20   21    2    ga_15
   20   21   22    2    ga_15
   21   22   23    2    ga_15
   22   23   24    2    ga_15
   23   24   25    2    ga_15
   24   25   26    2    ga_15
   25   26   27    2    ga_15
   26   27   28    2    ga_15
   27   28   29    2    ga_15
   28   29   30    2    ga_15
   29   30   31    2    ga_15
   30   31   32    2    ga_15
   31   32   33    2    ga_15
   32   33   34    2    ga_15
   13   35   36    2    ga_15
   13   35   38    2    ga_15
   36   35   38    2    ga_13
   35   36   37    2    ga_12
   35   38   39    2    ga_56
   38   39   40    2    ga_56
   39   40   41    2    ga_55
   40   41   42    2    ga_15
   41   42   43    2    ga_15
   42   43   44    2    ga_15
   43   44   45    2    ga_15
   44   45   46    2    ga_15
   45   46   47    2    ga_15
   46   47   48    2    ga_15
   47   48   49    2    ga_15
   48   49   50    2    ga_15
   49   50   51    2    ga_15
   50   51   52    2    ga_15
```

```
[ dihedrals ]
; GROMOS improper dihedrals
;  ai   aj   ak   al  funct   angle     fc
   13   14   35   12    2     gi_2
   14   16   13   15    2     gi_1
   16   18   14   17    2     gi_1
   35   36   38   13    2     gi_2

[ dihedrals ]
;  ai   aj   ak   al  funct    ph0      cp     mult
    1    4    5    6    1      gd_29
    4    5    6    7    1      gd_4
    4    5    6    7    1      gd_36
    5    6    7    8    1      gd_29
    6    7    8   11    1      gd_20
    6    7    8   11    1      gd_27
    7    8   11   12    1      gd_20
    7    8   11   12    1      gd_27
    8   11   12   13    1      gd_29
   11   12   13   35    1      gd_34
   12   13   35   38    1      gd_34
   12   13   14   16    1      gd_14
   13   14   16   18    1      gd_39
   14   16   18   19    1      gd_4
   16   18   19   20    1      gd_34
   18   19   20   21    1      gd_34
   19   20   21   22    1      gd_34
   20   21   22   23    1      gd_34
   21   22   23   24    1      gd_34
   22   23   24   25    1      gd_34
   23   24   25   26    1      gd_34
   24   25   26   27    1      gd_34
   25   26   27   28    1      gd_34
   26   27   28   29    1      gd_34
   27   28   29   30    1      gd_34
   28   29   30   31    1      gd_34
   29   30   31   32    1      gd_34
   30   31   32   33    1      gd_34
   31   32   33   34    1      gd_34
   35   38   39   40    1      gd_15
   38   35   36   37    1      gd_17
   13   35   38   39    1      gd_40
   38   39   40   41    1      gd_40
   39   40   41   42    1      gd_34
   40   41   42   43    1      gd_34
   41   42   43   44    1      gd_34
   42   43   44   45    1      gd_34
   43   44   45   46    1      gd_34
   44   45   46   47    1      gd_34
   45   46   47   48    1      gd_34
   46   47   48   49    1      gd_34
   47   48   49   50    1      gd_34
   48   49   50   51    1      gd_34
   49   50   51   52    1      gd_34
```

**CHOL**

```
[ moleculetype ]
; Name            nrexcl
CHOL              3

[ atoms ]
;  nr  type  resnr  resid  atom  cgnr  charge    mass    total_charge
    1   CH2r   1    CHOL     C1    1    0.0     14.027      ;  0.000
    2   CH2r   1    CHOL     C2    2    0.0     14.027      ;  0.000
    3   CH2r   1    CHOL     C3    3    0.270   13.019
    4   CH2r   1    CHOL     C4    4    0.0     14.027
    5   CH2r   1    CHOL     C5    5    0.0     13.019
    6   CH2r   1    CHOL     C6    6    0.0     14.027
    7   CH2r   1    CHOL     C7    7    0.0     14.027
    8   CH2r   1    CHOL     C8    8    0.0     13.019
    9   CH2r   1    CHOL     C9    9    0.0     13.019
   10    CH0   1    CHOL    C10   10    0.0     12.011
   11   CH2r   1    CHOL    C11   11    0.0     14.027
   12   CH2r   1    CHOL    C12   12    0.0     14.027
   13    CH0   1    CHOL    C13   13    0.0     12.011
   14   CH2r   1    CHOL    C14   14    0.0     13.019
```

```
   15   CH2r   1    CHOL    C15    15   0.0     14.027
   16   CH2r   1    CHOL    C16    16   0.0     14.027
   17   CH2r   1    CHOL    C17    17   0.0     13.019
   18    CH3   1    CHOL    C18    18   0.0     15.035
   19    CH3   1    CHOL    C19    19   0.0     15.035
   20    CH1   1    CHOL    C20    20   0.0     13.019
   21    CH3   1    CHOL    C21    20   0.0     15.035
   22    CH2   1    CHOL    C22    20   0.0     14.027
   23    CH2   1    CHOL    C23    21   0.0     14.027
   24    CH2   1    CHOL    C24    21   0.0     14.027
   25    CH1   1    CHOL    C25    22   0.0     13.019
   26    CH3   1    CHOL    C26    22   0.0     15.035
   27    CH3   1    CHOL    C27    22   0.0     15.035
   28     OA   1    CHOL    O28    23  -0.675   15.999
   29      H   1    CHOL    H29    24   0.405    1.008      ;  0.000
; total charge of the molecule:   0.000

[ bonds ]
;  ai   aj  funct   c0         c1
    1    2    2   gb_27
    1   10    2   gb_27
    2    3    2   gb_27
    3    4    2   gb_27
    4    5    2   gb_53
    5    6    2   gb_10
    5   10    2   gb_53
    6    7    2   gb_53
    7    8    2   gb_27
    8    9    2   gb_27
    8   14    2   gb_27
    9   10    2   gb_27
    9   11    2   gb_27
   10   19    2   gb_27
   11   12    2   gb_27
   12   13    2   gb_27
   13   14    2   gb_53
   13   17    2   gb_53
   13   18    2   gb_27
   14   15    2   gb_53
   15   16    2   gb_53
   16   17    2   gb_53
   17   20    2   gb_27
   20   21    2   gb_27
   20   22    2   gb_27
   22   23    2   gb_27
   23   24    2   gb_27
   24   25    2   gb_27
   25   26    2   gb_27
   25   27    2   gb_27
    3   28    2   gb_18
   28   29    2   gb_1

[ pairs ]
;  ai   aj  funct  ;  all 1-4 pairs but the ones excluded in GROMOS itp
    1    4    1
    2    5    1
    3   10    1
;   5    8    1
;   6    9    1
    7   10    1
    8   12    1
    9   13    1
   11   14    1
   18   11    1
   18    8    1
   18   15    1
   18   16    1
   19    2    1
   19    4    1
   19    6    1
   19    8    1
   19   11    1
   20   15    1
   28    1    1
   28    5    1
   12   20    1
   13   21    1
```

```
   13   22    1
   14   20    1
   15   20    1
   16   21    1
   16   21    1
   16   22    1
   17   23    1
   18   20    1
   20   24    1
   21   23    1
   22   25    1
   23   26    1
   23   27    1

[ angles ]
;  ai   aj   ak  funct   angle     fc
    1    2    3    2   ga_15
    1   10    5    2   ga_13
    1   10    9    2   ga_13
    1   10   19    2   ga_13
    2    1   10    2   ga_15
    2    3    4    2   ga_13
    3    4    5    2   ga_15
    4    5    6    2   ga_56
    4    5   10    2   ga_55
    5    6    7    2   ga_56
    5   10    9    2   ga_13
    5   10   19    2   ga_13
    6    5   10    2   ga_56
    6    7    8    2   ga_15
    7    8    9    2   ga_13
    7    8   14    2   ga_13
    8    9   10    2   ga_13
    8    9   11    2   ga_13
    8   14   13    2   ga_13
    8   14   15    2   ga_13
    9   10   19    2   ga_13
   10    9   11    2   ga_13
   11   12   13    2   ga_15
   12   13   14    2   ga_13
   12   13   17    2   ga_13
   12   13   18    2   ga_13
   13   14   15    2   ga_55
   13   17   16    2   ga_55
   13   17   20    2   ga_13
   14   13   17    2   ga_55
   14   13   18    2   ga_13
   14   15   16    2   ga_55
   15   16   17    2   ga_55
   16   17   20    2   ga_13
   17   13   18    2   ga_13
   17   20   21    2   ga_13
   17   20   22    2   ga_13
   20   22   23    2   ga_15
   21   20   22    2   ga_13
   22   23   24    2   ga_15
   23   24   25    2   ga_15
   24   25   26    2   ga_13
   24   25   27    2   ga_13
   26   25   27    2   ga_13
    3   28   29    2   ga_12
    2    3   28    2   ga_13
    4    3   28    2   ga_13

[ dihedrals ]
; GROMOS improper dihedrals
;  ai   aj   ak   al  funct   angle     fc
   10    5    6    7    2     gi_1
   14   15   16   17    2     gi_1a
    8   14    7    9    2     gi_2
    9   11   10    8    2     gi_2
   14    8   15   13    2     gi_2
   17   20   13   16    2     gi_2
   20   17   21   22    2     gi_2
    3    4    2   28    2     gi_2
```

```
[ dihedrals ]
;  ai   aj   ak   al  funct    ph0      cp     mult
    1    2    3    4   1      gd_34
    2    3    4    5   1      gd_34
;   3    4    5   10   1      gd_34
    3    4    5    6   1      gd_40
    4    5   10    9   1      gd_34
    2    1   10    9   1      gd_34
   10    1    2    3   1      gd_34
;
    4    5    6    7   1      gd_15
    5    6    7    8   1      gd_40
    6    7    8    9   1      gd_40
    7    8    9   11   1      gd_34
    8    9   10    1   1      gd_34
;
   10    9   11   12   1      gd_34
    9   11   12   13   1      gd_34
   11   12   13   17   1      gd_34
   12   13   14   15   1      gd_34
   13   14    8    7   1      gd_34
;
    8   14   15   16   1      gd_34
;  14   15   16   17   1      gd_34
   15   16   17   20   1      gd_34
   16   17   13   12   1      gd_34
;
   16   17   20   21   1      gd_34
   17   20   22   23   1      gd_34
   20   22   23   24   1      gd_34
   22   23   24   25   1      gd_34
   23   24   25   26   1      gd_34
;
    4    3   28   29   1      gd_26
```

**Appendix II**

**Alterations to force field files (ffbonded.itp)**

New bond length for carbon-carbon bonds in 5-membered and 6-membered hydrocarbon rings.

```
#define gb_53           0.150       0.750E+07
; C,CHn - C,CHn         5-6 mem rings     800
```

New bond angle for carbon-carbon single bonds in 5-membered and 6-membered hydrocarbon rings.

```
#define ga_55           108.0       535.00
; CHn - CHn - CHn 5-6 mem rings
```

New bond angle for carbon-carbon double bonds in 6-membered hydrocarbon rings.

```
#define ga_56           124.0       615.00
; CHn - CH = CH - CHn         6 mem ring
```

New improper dihedral angle for 5-membered planar cholesterol ring.

```
#define gi_1a           0.0         33.5
; planar group          5 mem ring
```

## References


[i] Form and flexibility in phosphoinositide 3-kinases.
R. Williams et al. - *Biochemical Society Transactions* **37**: 615-626 (2009)

[ii] The phosphoinositide (PI) 3-kinase family.
F.M. Foster et al. - *Journal of Cell Science* **116**: 3037-3040 (2003)

[iii] Beyond PTEN mutations: the PI3K pathway as an integrator of multiple inputs during tumorigenesis.
M. Cully et al. - *Nature Reviews Cancer* **6**: 184-192 (2006)

[iv] Development of phosphoinositide 3-kinase pathway inhibitors for advanced cancer.
J.M. Cleary et al. - *Current Oncology Reports* **12(2)**: 87-94 (2010)

[v] The phosphoinositide 3-kinase pathway in human cancer: genetic alterations and therapeutic implications.
A. Arcaro et al. - *Current Genomics* **8(5)**: 271-306 (2007)

[vi] PIK3CA mutations in the most common types of cancer.
M.A. Chomczyk et al. - *Postepy Biochemii* **59(3)**: 280-284 (2013)

[vii] Phosphatidylinositol 3-kinase mutations identified in human cancer are oncogenic.
S. Kang et al. - *Proceedings of the National Academy of Sciences* **102(3)**: 802-807 (2005)

[viii] Cancer-specific mutations in *PIK3CA* are oncogenic *in vivo*.
A.G. Bader et al. - *Proceedings of the National Academy of Sciences* **103(5)**: 1475-1479 (2006)

[ix] Structural comparisons of class I phosphoinositide 3-kinases.
L.M. Amzel et al. - *Nature Reviews Cancer* **8(9)**: 665-669 (2008)

[x] Structural effects of oncogenic PI3Kα mutations.
S. Gabelli et al. - *Current Topics in Microbiology and Immunology* **347**: 43-53 (2010)

[xi] The structure of a human p110α/p85α complex elucidates the effects of oncogenic PI3Kα mutations.
C.H. Huang et al. - *Science* **318(5857)**: 1744-1748 (2007)

[xii] The regulation of class IA PI 3-kinases by inter-subunit interactions.
J.M. Backer - *Current Topics in Microbiology and Immunology* **346**: 87-114 (2010)

[xiii] Gain of interaction with IRS1 by p110α-helical domain mutants is crucial for their oncogenic functions.
Y. Hao et al. - *Cancer Cell* **23(5)**: 583-593 (2013)

[xiv] A frequent kinase domain mutation that changes the interaction between PI3Kα and the membrane.
D. Mandelker et al. - *Proceedings of the National Academy of Sciences* **106(40)**: 16996-17001 (2009)

[xv] Oncogenic mutations mimic and enhance dynamic events in the natural activation of p110α (*PIK3CA*).
J.E. Burke et al. - *Proceedings of the National Academy of Sciences* **109(38)**: 15259-15264 (2012)

[xvi] Helical and kinase domain mutations in p110α (PI 3-kinase) induce gain of function by different mechanisms.
L. Zhao and P.K. Vogt - *Proceedings of the National Academy of Sciences* **105(7)**: 2652-2657 (2008)

[xvii] Dynamic steps in receptor tyrosine kinase mediated activation of class IA PI3K captured by H/D exchange.
J.E. Burke et al. - *Advances in Biological Regulation* **53(1)**: 97-110 (2013)

[xviii] GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit.
S. Pronk et al. - *Bioinformatics* **29(7)**: 845-854 (2013)

[xix] Definition and testing of the GROMOS force field versions 54A7 and 54B7.
N. Schmid et al. - *European Biophysics Journal* **40(7)**: 843-856 (2011)

[xx] Large influence of cholesterol on solute partitioning into lipid membranes.
C.L. Wennberg et al. - *Journal of the American Chemical Society* **134**: 5351-5361 (2012)

[xxi] Another piece of the membrane puzzle: extending Slipids further.
J.P.M. Jämbeck et al. - *Journal of Chemical Theory and Computation* **9(1)**: 774-784 (2013)

[xxii] Stability and membrane orientation of the fukutin transmembrane domain: a combined multiscale molecular dynamics and circular dichroism study.
D.A. Holdbrook et al. - *Biochemistry* **49(51)**: 10796-10802 (2010)

[xxiii] Structure of sphingomyelin bilayers: a simulation study.
S.W. Chiu et al. - *Biophysical Journal* **85**: 3624-3636 (2003)

[xxiv] Sphingomyelin-cholesterol domains in phospholipid membranes: atomistic simulation.
S.A. Pandit et al. - *Biophysical Journal* **87**: 1092-1100 (2004)

[xxv] Interaction models for water in relation to protein hydration.
H.J.C. Berendsen et al. - *Intermolecular Forces* **B. Pullman edition**: 331-342 (1981)

[xxvi] Molecular dynamics with coupling to an external bath.
H.J.C. Berendsen et al. - *Journal of Chemical Physics* **81**: 3684 (1984)

[xxvii] LINCS: A linear constraint solver for molecular simulations.
B. Hess et al. - *Journal of Computational Chemistry* **18(12)**: 1463-1472 (1997)

[xxviii] A smooth particle mesh Ewald method.
U. Essmann et al. - *Journal of Chemical Physics* **103**: 8577-8593 (1995)

[xxix] Molecular dynamics simulations of phospholipid bilayers with cholesterol.
C. Hofsäß et al. - *Biophysical Journal* **84(4)**: 2192-2206 (2003)

[xxx] Detailed molecular dynamics simulations of model biological membranes containing cholesterol.
M.L. Berkowitz - *Biochimica et Biophysica Acta- Biomembranes* **1788**: 86-96 (2009)

[xxxi] APL@Voro: a Voronoi-based membrane analysis tool for GROMACS trajectories.
G. Lukat et al. - *Journal of Chemical Information and Modelling* **53(11)**: 2908-2925 (2013)

[xxxii] Areas of molecules in membranes consisting of mixtures.
O. Edholm and J.F. Nagle - *Biophysical Journal* **89(3)**: 1827-1832 (2005)

[xxxiii] MD simulations of fully hydrated DPPC with different macroscopic boundary conditions and parameters.
D.P. Tieleman and H.J.C. Berendsen - *Journal of Chemical Physics* **105**: 4871 (1996)

[xxxiv] Combined Monte Carlo and MD simulations of hydrated 18:0 sphingomyelin-cholesterol lipid bilayers.
G.A. Khelashvili and H.L. Scott - *Journal of Chemical Physics* **120(20)**: 9841-9847 (2004)

[xxxv] Combined Monte Carlo and MD simulation of hydrated lipid-cholesterol lipid bilayers at low cholesterol concentration.
S.W. Chiu et al. - *Biophysical Journal* **80**: 1104-1114 (2001)

[xxxvi] Effect of cholesterol on the structure of a phospholipid bilayer.
F. de Meyer et al. - *Proceedings of the National Academy of Sciences* **106(10)**: 3654-3658 (2009)

[xxxvii] The challenge of lipid rafts.
L.J. Pike - *Journal of Lipid Research* **50(Suppl)**: S323-S328 (2009)

[xxxviii] Assessing the nature of lipid raft membranes.
P.S. Niemelä et al. - *PLoS Computational Biology* **3(2)**: e34 (2007)

[xxxix] Lessons of slicing membranes: interplay of packing, free area, and lateral diffusion in phospholipid/ cholesterol bilayers.
E. Falck et al. - *Biophysical Journal* **87(2)**: 1076-1091 (2004)

[xl] The ELBA force field for coarse-grain modelling of lipid membranes.
M. Orsi and J.W. Essex - *Public Library of Science ONE* **6(12)**: e28637 (2011)

[xli] Expression of PIK3CA mutant E545K in the mammary gland induces heterogeneous tumours but is less potent than mutant H1047R.
D.S. Meyer et al. - *Oncogenesis* **2**: e74 (2013)