Universidade de Évora - Escola de Ciências e Tecnologia

Mestrado em Bioquímica

Dissertação

Interaction of HIV fusion inhibitor peptides, of the MT-SC familly, with biological membrane models: a molecular dynamics study

João Pedro Maneta Courela

Orientador(es) | António Manuel Teixeira Martins do Canto

Évora 2024

Universidade de Évora - Escola de Ciências e Tecnologia

Mestrado em Bioquímica

Dissertação

Interaction of HIV fusion inhibitor peptides, of the MT-SC familly, with biological membrane models: a molecular dynamics study

João Pedro Maneta Courela

Orientador(es) | António Manuel Teixeira Martins do Canto

Évora 2024

A dissertação foi objeto de apreciação e discussão pública pelo seguinte júri nomeado pelo Diretor da Escola de Ciências e Tecnologia:

Presidente	Ana Vitória Dordio (Universidade de Évora)
Vogais	Alfredo Jorge Palace Carvalho (Universidade de Évora) (Arguente) António Manuel Teixeira Martins do Canto (Universidade de Évora) (Orientador)

Évora 2024

Acknowledgments

I am extremely grateful to my supervisor, Professor António do Canto, for all the hours of work he put into guiding me through this stage of my education. This work truly would not have been possible without his guidance, from teaching me about the intricacies of the field of molecular dynamics that was novel to me, to the writing and editing process of the dissertation. He constantly made sure that everything was in order, while simultaneously making the entire process enjoyable. For this I am profoundly grateful.

Special thanks to Professor Alfredo Carvalho for his hard work in the maintenance of the computers in which I was working on, for always keeping them running and for being available to solve any issues that came up. He made this entire process much easier, and I am grateful for that.

I would also like to thank all the teachers and colleagues who have accompanied me during my academic path, for providing me with everything that I needed to reach this goal.

Thank you to my family for their continuous support. Special thanks to my mother Carla and to my grandmother Maria Joana who have always been by my side. They constantly gave me motivation and supported me whenever I felt overwhelmed or undeserving. I would not be where I am today if it were not for them.

Finally, I would like to thank my dog Sasha for the emotional support she provided me with for twelve long years. I loved being with her, and she was always happy to see me. She was my best friend and I miss her dearly.

Resumo

Interação de Péptidos inibidores de fusão do VIH, da família MT-SC, com modelos de membranas biológicas: um estudo por dinâmica molecular

O VIH infeta células CD4+ através do reconhecimento de recetores CCR5 e CXCR4 á sua superfície num processo mediado por duas glicoproteínas do seu envelope, gp41 e gp120. Péptidos inibidores de fusão são fragmentos de regiões da gp41, que inibem a fusão de VIH com as células, impedindo a formação da sua estrutura de fusão, o *6-Helix Bundle*. Existem estirpes de VIH resistentes aos péptidos disponíveis atualmente, pelo que é necessário desenvolver outros mais eficazes. Este trabalho visa comparar o desempenho de péptidos com adição de dois resíduos na região N-terminal (Metionina e Treonina), com o desempenho dos péptidos de gerações anteriores (Pronto, 2019), para averiguar o impacto do *M-T Hook* na atividade inibitória. Foram feitos testes em três modelos de membrana: POPC, POPC:CHOL(4:1) e POPC:CHOL(1:1). Os resultados obtidos sugerem um impacto positivo do *M-T Hook* em membranas mais fluidas, mas nenhum benefício aparente em membranas mais rígidas. Os péptidos com atividade mais promissora foram MT-C34, MT-SC34 e MT-SC34EK.

Palavras-chave: VIH; SIDA; Péptidos inibidores de fusão; Membranas biológicas; Dinâmica molecular; Interações lípido-péptido; Simulação de biomembranas

Abstract

Interaction of HIV fusion inhibitor peptides, of the MT-SC familly, with biological membrane models: a molecular dynamics study

HIV recognizes CD4+ cell receptors (CCR5 and CXCR4) through the action of two envelope glycoproteins, gp41 and gp120. Fusion inhibitor peptides mimic portions of gp41 and inhibit its fusion to the target cell, by blocking the formation of the 6-Helix Bundle. Certain strains of HIV have acquired resistance to the peptides available nowadays, therefore its necessary to develop better and more efficient peptides to solve this problem. In this work, we seek to compare the behavior of peptides with the addition of Methionine and Threonine residues in the N-terminus of the peptide, with the activity of their parent peptides (Pronto, 2019). Tests were made in three membrane models: POPC, POPC:CHOL (4:1) and POPC:CHOL (1:1). The results suggest that the M-T Hook has a positive impact in more fluid membranes but has no apparent benefits in more rigid membranes. The peptides with the most promising activity were MT-C34, MT-SC34 and MT-SC34EK.

Keywords: HIV; AIDS; Fusion inhibitor peptides; Biological membranes; Molecular dynamics; Lipid-peptide interaction; Biomembrane simulations

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ARV – AIDS-Associated Retrovirus

AIDS – Acquired ImmunoDeficiency Syndrome

Cap – Capsid

CDC – Center for Disease Control and Prevention

CPSF6 – Polyadenylation Specificity Factor-6

CypA – Cyclophilin A

DIS – Dimerization Initiation Site

DNA – Deoxyribonucleic Acid

dsDNA – double stranded DNA

gRNA – genomic RNA

HIV – Human Immunodeficiency Virus

HTLV – Human T-cell Leukemia Virus

IDAV – Immune Deficiency-Associated Virus

IN – Integrase

LAV – Lymphadenopathy-Associated Virus

LTR – Long Terminal Repeat

MA – Matrix

MHC – Major Histocompatibility Complex

mRNA – messenger RNA

NC – Nucleocapsid

PBS – Primer Binding Site

PIC – Pre-Integration Complex

POL II – Polymerase II

PPT – Purine Rich Sequence

PR – Protease

P-TEFb – Positive Transcription Elongation Factor b

RT – Reverse Transcriptase

RNA – Ribonucleic Acid

ssRNA – single stranded RNA

TAR – Transactivation Response Element

TNPO3 – Transportin 3

UTR – Untranslated Region

Problem Statement

HIV drug resistance has been a problem due to the mutations that the virus underwent over the years. There are many anti-HIV drug classes that block different steps of the HIV life cycle, preventing the disease from progressing. The drug class discussed in this study are the HIV peptide fusion inhibitors. Peptide fusion inhibitors prevent HIV entry into the cell by blocking the fusion of the viral envelope to the membrane. (this process will be described thoroughly ahead). The only FDA approved fusion inhibitor on the market is enfuvirtide (T-20). It is complementary to a region of gp41, an HIV envelope glycoprotein, but as the virus mutated it lost some of its inhibitory activity (Greenberg & Cammack, 2004).

The development of novel HIV fusion inhibitors is therefore very important for improving the treatment of HIV. Previous studies on other peptides complementary to gp41, such as C34, SC22EK, SC29EK, SC34, SC34EK and SC35EK have shown more potent inhibitory activity than T-20 (Pronto, 2019). All peptides above were developed from the sequence of C34 with some modifications to improve their helicity, solubility, and other important parameters (the development of C34 and its analogous fusion inhibitors will be described in detail further ahead).

The study will be performed for a fusion inhibitor family that was developed from the one mentioned previously. The “MT-SC” family includes MT-C34, MT-SC22EK, MT-SC29EK, MT-SC34, MT-SC34EK and MT-SC3EK. These fusion inhibitor peptides are developed from the ones mentioned above, with the addition of methionine and threonine at the N-terminus of the sequence. These residues form a hook like structure that has been proven in previous studies to be critical for the stability and activity of HIV fusion inhibitors of other peptides (Chong et al., 2012a). Molecular dynamics simulations performed for these peptides in different membrane systems will help evaluate their inhibitory activity.

Objectives

The goal of this study is to find out how the addition of Methionine and Threonine residues in the N-terminus of the sequence of the fusion inhibitor peptides (MT-C34, MT-SC22EK, MT-SC29EK, MT-SC34, MT-SC34EK and MT-SC35EK) impacts their inhibitory activity and interaction with membranes of different compositions. For this, measurements of certain parameters are crucial. Membrane properties, such as thickness and area per lipid can indicate the impact peptide adsorption has on membrane order. For the peptides, parameters like the quantification of hydrogen bonds between them and the membrane, lateral and rotational dynamics, positions of the alpha carbons and distance from their center of mass relative to the membrane are all necessary to visualize the steps the peptide takes towards adsorption. By comparing the measurements obtained for the MT-SC peptide family with the results obtained for their parent peptides, we can infer on the role of the M-T hook region.

1. Introduction

HIV and Aids

1.1.1. History and Global Importance of HIV and AIDS

The Acquired Immunodeficiency Syndrome (AIDS) pandemic was first recognized in June of 1981 by the Centers for Disease Control and Prevention (CDC) when five cases of pneumonia were reported in gay men from Los Angeles (Gottlieb et al., 1981). It was characterized as a long-lasting disease that was commonly accompanied by a deep state of immunosuppression. This made it easy for opportunistic infections to settle in patients with this condition (Montagnier, 2002). At first, people were skeptical of the impact this disease could have on the world, but it quickly became a worldwide public health issue that is still ongoing. In 2022, 1,3 million people were infected with HIV and 39 million are currently living with the virus. Furthermore, 630 thousand people died of AIDS-related causes in 2021 alone, bringing the death total since the first case in 1981 to 40.4 million (UNAIDS, 2021). There is still no definitive cure for HIV and no vaccine to prevent infection, but it can be treated. The treatment can't fully cure HIV, but it can reduce a person's viral load to the point where it is undetectable by viral load tests. This allows people with HIV to live longer and healthier lives while considerably reducing the risk of transmission of the virus (U.S. Department of Health and Human Services; National Institutes of Health, 2022). At the beginning of the AIDS epidemic, its causative agent was unknown.

In 1972, Luc Montagnier created a research unit in the virology department of the Pasteur Institute. It focused on the study of retroviruses and was given the name viral oncology unit by its creator, as he believed, along with many other biologists at the time, that they could cause cancer. At first, the program failed its task to relate retroviruses to cancer because of difficulties in the isolation process (Montagnier, 2002). It wasn't until 1977 that some advances were made when researchers decided to treat isolated mouse cells with low doses of antiserums to mouse interferon. The interferon was inhibiting the retroviruses' expression, thereby not allowing their observation (Barré-Sinoussi et al., 2004). After experimenting with mouse cells, the investigation moved on to human cells, particularly cells from patients affected by lymphocytic leukemias and breast cancers (R. C. Gallo & Montagnier, 2003).

Robert Gallo and his group in the United States shared the same beliefs as Montagnier’s group in Paris. They discovered a technique in their laboratory that enabled scientists to make T lymphocyte cultures using interleukin 2 (called “T-cell growth factor” at the time) (R. C. Gallo & Montagnier, 2003). In 1982 Michel Crépin was able to obtain a sequence of DNA from a human breast tumor that was very similar to another sequence present in the mouse oncogenic retrovirus (Crépin, 1984). By this time, scientists all over the world were already aware of the existence of AIDS. Gallo’s idea was that this disease was caused by a retrovirus, but back then the only retroviruses known were the human T-cell leukemia viruses (HTLV), HTLV-1, and HTLV-2, discovered by himself and his group. These viruses were somewhat like the causative agent of AIDS, but researchers quickly realized that the two were different. Efforts went on at the Pasteur Institute and the National Institutes of Health in Bethesda to try to identify new lymphotropic retroviruses (R. C. Gallo & Montagnier, 2003).

Early assays made in Bethesda using molecular and immunological probes pointed to the virus being a variant of HTLV-1, however, this turned out to be wrong (R. C. Gallo & Montagnier, 2003). In the supernatant of a culture of T cells derived from a lymph node biopsy of a homosexual patient with lymphadenopathy in the neck, Françoise Sinoussi from the Pasteur Institute detected reverse transcriptase (RT) activity. This indicated the presence of a retrovirus. Tests were carried out using antibodies against HTLV developed in Gallo’s laboratory. Interestingly, the sample didn’t precipitate with the use of the antibodies, it only precipitated with the patient’s serum. Additionally, the virus could infect T cells from healthy donors but could not do so to B cells and fibroblasts, for example. In other samples, including blood samples from patients with full-blown AIDS, viruses capable of quickly killing cultured T lymphocytes were isolated. The viruses isolated from the lymphadenopathy patient were less aggressive than those isolated from patients with full-blown AIDS. The former isolates were designated lymphadenopathy-associated viruses (LAV), while the latter were called immune deficiency-associated viruses (IDAV) by Montagnier himself (Montagnier, 2002; Schmidt, 2018).

Electron microscopy technology allowed for the visual analysis of the new retrovirus. Upon observation, this virus looked similar to some animal lentiviruses, which convinced

Montagnier that they were dealing with a new agent different from HTLV (Montagnier, 2002).

Between 1983 and 1984, Luc Montagnier’s group made two crucial observations. First, patients with lymphadenopathy produced large quantities of antibodies against the virus. Second, the virus seemed to favor infection of CD4-positive T lymphocytes. These results were later replicated by Gallo and his group. They also published more evidence suggesting a correlation between this new agent, which they called HTLV-III, and AIDS by isolating it from 48 subjects including people suffering from AIDS and pre-AIDS and seemingly healthy individuals (Robert C. Gallo, Syed Z. Salahuddin, Mikulas Popovic, Gene M. Shearer et al., 2019). Jay Levy and his group strengthened this connection, as they were able to isolate the virus, which they called AIDS-associated retrovirus (ARV), from 22 out of 45 randomly selected patients with AIDS. They also detected antibodies to ARV in all 86 AIDS patients selected for the study (Levy, J.; Hoffman, A.; Kramer, S.; Landis, J.; Shimabukuro, J.; Oshiro, 1984). These terms were subsequently changed to Human Immunodeficiency Virus (HIV-1) in 1986 (Schmidt, 2018).

Many more discoveries about HIV were made in the following years after its isolation. These include the identification of T cell CD4 as its main receptor (7, 8) and the identification of large surface glycoprotein gp160 (Allan et al., 1985). The connection between HIV and AIDS was also established as more and more isolates of the virus were obtained from patients with different origins affected by the disease. Blood tests were developed and commercially available, which allowed for a drastic reduction in the rate of transmission of AIDS, in developed countries mainly. These tests also made the molecular characterization of HIV possible, which helped prove that it was a lentivirus. In 1986, another strain of HIV, HIV-2, was isolated from patients in West Africa (Clavel et al., 1986; R. C. Gallo & Montagnier, 2003; Montagnier, 2002).

Transmission of HIV generally occurs at mucosal surfaces or through percutaneous inoculation (Shaw & Hunter, 2012). This virus has various invasion sites, which include parts of the male and female genital tracts, parts of the intestinal tract and the placenta, and finally the bloodstream (Hladik & McElrath, 2008). It is found in bodily fluids like blood, semen, cervicovaginal and rectal secretions, and breast milk. Therefore it can be transmitted through penetrative sex (anal or vaginal), blood transfusion, drug injection, and sharing of contaminated needles, and from mother to child during pregnancy, childbirth, or breastfeeding (UNAIDS, n.d.).

The likelihood of infection through sex and from mother to child can vary depending on some risk factors. Individuals with genital ulcers are more likely to transmit the virus during sex than those with healthy genitals. This also happens in people with higher viral loads, as they have an increased risk of transmission through sex or from mother to child. Treatment with antiretrovirals suppresses this load making the risk of infection almost negligible in these cases. The route of exposure to the virus is also proven to affect HIV transmission, with higher rates of transmission through anal sex, compared to vaginal sex. Heterosexual transmissions represent 70% of infections in the world, while maternal-infant infection, intravenous drug use, and homosexual relations between men make up most of the remaining 30% (Hladik & McElrath, 2008; McElrath et al., 2008; UNAIDS, n.d.).

1.1.2. HIV Genome and Structure

HIV-1 has two copies of its single-stranded RNA genome packaged into the viral particle. These copies are bound via the dimerization initiation site (DIS), forming a dimer. This conformation is thought to be essential for genome packaging and evolution (Ye et al., 2022). HIV-1 is part of the *Orthoretrovirinae* subfamily of *Retroviridae*, which houses all retroviruses. Orthoretroviruses are enveloped by a membrane and acquire a spherical shape. They are also capable of budding from the plasma membrane of infected cells. (Ganser-Pornillos et al., 2012). Each RNA strand of the HIV-1 genome has roughly 9800 nucleotides and consists of nine genes that encode 16 proteins that aid the virus in the completion of its life cycle. Some genes code for structural proteins, while others code for regulatory and accessory proteins (Li & Clercq, 2016).

The functional genes are flanked by a capped 5’ and a polyadenylated 3’ untranslated region (UTR). After reverse transcription, 5’ UTR acquires a new U3 region and 3’ UTR a new U5 region, becoming the 5’ and 3’ long terminal repeats (LTR), respectively. These are the most conserved regions of the genome, and they play key roles in the life cycle of the virus. 5’ UTR contains several structures, such as a trans-activation region, a poly(A) domain, a primer binding site, the previously mentioned DIS, a packaging signal, and the start codon for *gag*. 5’ LTR, compared to 5’ UTR, has extra binding sites for host cell factors that regulate transcription of the integrated provirus. 3' UTR grants stability to the viral RNA given its poly-adenylated end and regulates the function of the genomic RNA. (gRNA). 3' LTR mediates the expression of the gRNA, while also being able to stop its transcription. In case the function of the 5' LTR is compromised, 3'LTR can take over its functions. The role of these structures and their components will be explained in further detail in the HIV-1 life cycle (Ye et al., 2022; J. Zhang & Crumpacker, 2022).

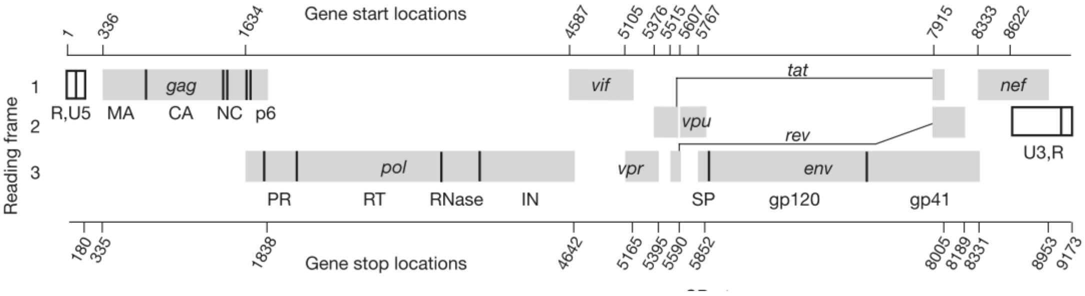


Figure 1 – Structure and organization of the HIV-1 genome. The genes (represented by the grey boxes) are flanked by the 5' and 3' UTR regions. Gene start and end locations are indicated in the above and below lines. Image adapted from Watts et. al (Watts et al., 2009).

The HIV-1 genome has three major genes: *gag*, *pol,* and *env*, that code for structural proteins and enzymes. The *gag* gene originates the Gag polyprotein, which is a precursor for the matrix (MA), capsid (Cap), nucleocapsid (NC), and p6 proteins. It also has two spacer peptides, SP1 and SP2, between the Cap-NC and NC-p6 domains. Pol is only synthesized as a Gag-Pol fusion protein because *pol* lacks an initiation codon. Gag-Pol is a precursor for the viral enzymes protease (PR), reverse transcriptase (RT), and integrase (IN), as well as for many of the Gag polyprotein products. The *env* gene codes for a heavily glycosylated protein termed gp160. This protein has an N-terminal and a C-terminal domain labeled gp120 and gp41, respectively. These domains are separated by a small cleavage site for PR. All other genes are non-structural accessory genes and originate two regulatory proteins, Tat and Rev, and four accessory proteins, Nef, Vif, Vpu/Vpx, and Vpr. Vpu and Vpx are present in HIV-1 and HIV-2 respectively. Although the HIV genome only codes for 16 proteins, these form pairwise interactions with one another, forming complex mechanisms that secure viral replication (Engelman & Cherepanov, 2012; Li & Clercq, 2016; Richter et al., 2008).

The proteins derived from the cleavage of the Gag polyprotein make up the protein core of the mature virion and account for roughly 50% of its mass. MA is involved in particle assembly, incorporation of the envelope into virions, and RNA targeting to the plasma membrane. Cap constitutes the capsid, which houses the viral genome and interacts with multiple host cell factors, facilitating the infection process. NC is the major constituent of the nucleocapsid, and it can recognize and associate itself with the RNA genome, which makes it crucial for the structural assembly of the immature virus by packaging it inside the capsid cavity. Lastly, p6 can interact with host cell machinery needed to complete the budding of immature virions (Bell & Lever, 2013).

The Gag-Pol fusion protein originates the previously mentioned viral enzymes, which catalyze many crucial steps in the viral replication of HIV-1. PR catalyzes the Gag and Gag-Pol precursor protein processing during virion maturation. RT is a multifunctional enzyme that helps in the transformation of the single-stranded RNA (ssRNA) chains of HIV-1 into double-stranded DNA (dsDNA). It also has a nuclease active site that cleaves and removes the RNA template. IN, as the name suggests, catalyzes the integration of the dsDNA produced by RT into the host cell genome (Goodsell, 2002; Hill et al., 2005).

Proteins gp120 and gp41 are spread across the membrane of the virus. These subunits mediate the process of viral and cell membrane fusion. gp120 is a surface protein that recognizes the receptor and co-receptors of the target, while gp41 is a transmembrane protein that mediates the fusion process of the virus and cell membranes (Caffrey, 2011; Ross & Klotman, 2008). Other aspects relative to these protein subunits and how they affect the fusion process will be explained in detail in another section.

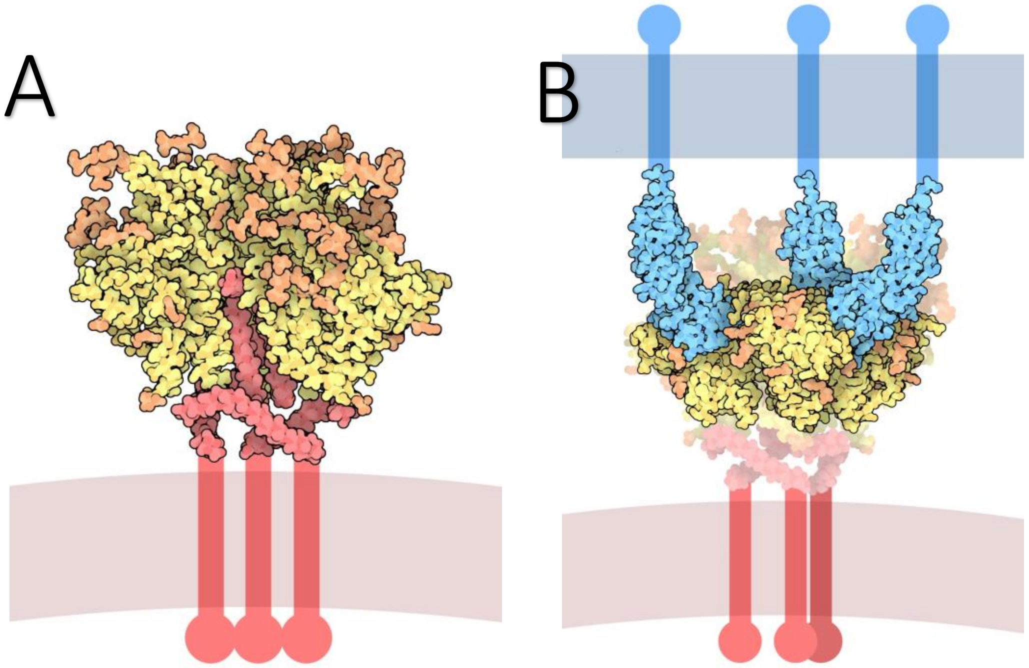


Figure 2 – (A) Partially solved structure of the HIV Envelope Glycoprotein containing gp120 (yellow), and outer domain of gp41 (red) and flexible carbohydrates (orange). (B) HIV Envelope Glycoprotein during attachment to CD4 proteins (blue). The complete structure of the HIV envelope glycoproteins is difficult to obtain due to their flexibility and flexible carbohydrate cover. Image adapted from (Goodsell, 2014).

The regulatory proteins of HIV, Tat, and Rev are simpler than the ones coded by the three major genes, but their role of transcriptional and posttranscriptional regulation is crucial to guarantee that gene expression occurs normally. They are the first to be produced after HIV-1 infection and both have very complex mechanisms that make viral replication possible. Tat can interact with RNA targets present in the R section of the 5’ LTR region of the genome, hereby stimulating elongation of the mRNA chains. Rev secures transportation of the later synthesized partially and fully unspliced mRNAs, which encode the accessory and structural proteins, from the nucleus to the cytoplasm where they are translated (Karn & Stoltzfus, 2012; Rosen, 1991).

Lastly, the accessory proteins were thought to be irrelevant at the beginning of their study, but they proved to be crucial for viral replication as promoters of disease progression and pathogenesis of HIV (Miller & Sarver, 2012). They secure their function by altering cellular pathways by interacting with multiple proteins. Nef acts as a positive viral factor by keeping infected cells intact and alive for longer. It also downmodulates levels of major histocompatibility complex-I and II (MHC I and MHC II) and CD4 present at the immune synapses. Vif, Vpu (exclusive of HIV-1), and Vpx (exclusive of HIV-2) control cellular restriction factors by either redirecting their function, mediating their degradation, or sequestering them. Finally, Vpr contains a co-factor, DCAF1, through which it can alter the expression of many host proteins. This protein can induce cellular apoptosis and G2 cycle arrest (Das & Jameel, 2005; Lubow & Collins, 2020; Strebel, 2013).

The mature HIV virion structure can be seen in Fig 3. The viral genome is complexed with the NC and p6 proteins forming the nucleocapsid. This structure protects the genetic content of the virion from degradation by host-cell nucleases. Around the nucleocapsid is a cone-shaped structure made up of copies of the Cap protein, the capsid. Inside it, we can find two of the HIV-1 enzymes, RT and IN. The capsid is further enveloped by the matrix, which is composed of copies of the MA protein. This conformation protects the genome of the viral particle and helps keep it intact so it can be released into the target cell. The PR enzyme is located between the capsid and the matrix. The outermost layer of the virus, the envelope, is composed of the plasma membrane of the cell from which it originated, and the envelope proteins that are scattered throughout the membrane (Ganser-Pornillos et al., 2008; Watson, 2009).

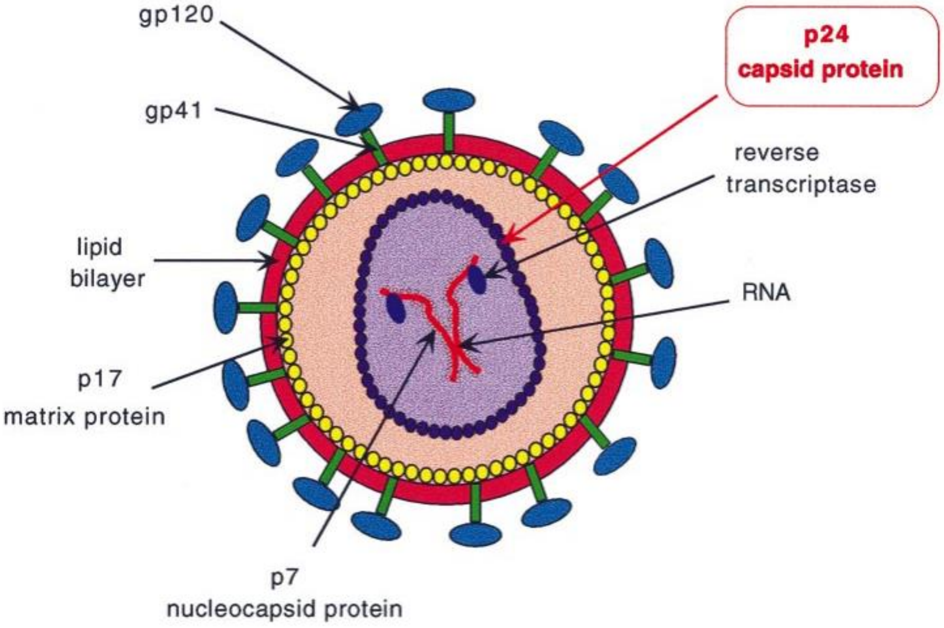


Figure 3 - Cross-section of a mature HIV-1 virion. Surface gp120 and transmembrane gp41 are shown in blue and green, respectively. The red layer corresponds to the plasma membrane of the virion taken from the host cell from the budding process. The yellow and purple circles represent gag proteins MA and Cap, respectively.

Cap proteins form a cone-shaped cavity that encases the genome. NC is complexed with the RNA genome along with p6, viral enzymes IN and RT, and accessory protein Vpr. PR is located between the matrix and the capsid. Image adapted from Carmen et. Al (1999) (Berthet-Colominas et al., 1999).

1.1.3. HIV Life Cycle

The HIV-1 life cycle can be roughly divided into early and late-phase replication. Early-phase replication encompasses all steps from viral attachment to the integration of the proviral dsDNA into the host cell genome. Late-phase replication starts with the transcription of the previously mentioned dsDNA and ends when the resulting mature, infectious virions leave the cell (Kirchhoff, 2016).

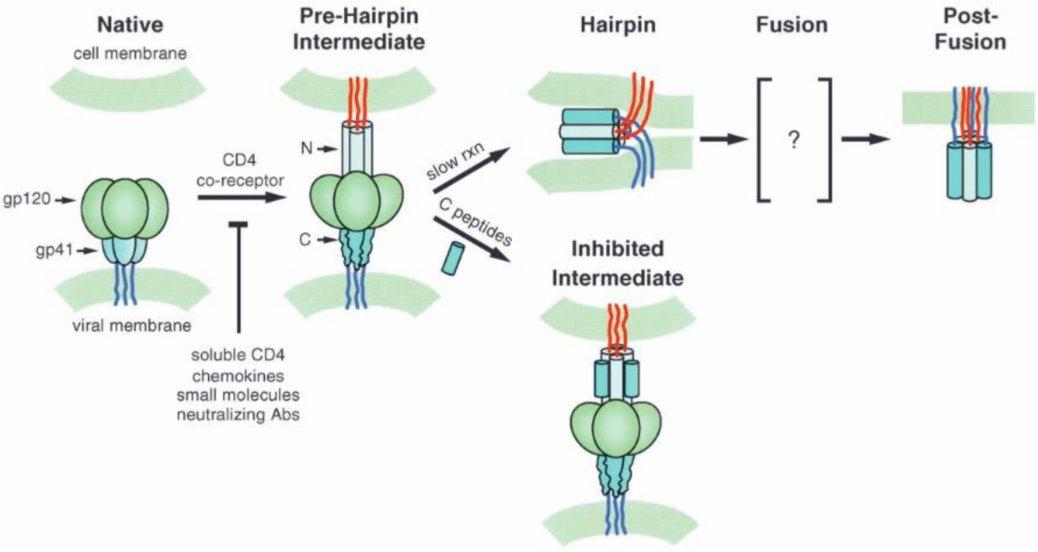


Figure 4 – Envelope-mediated fusion model of HIV. Image adapted from (S. A. Gallo et al., 2003).

Firstly, binding to the target cell is mediated by envelope protein gp120 which recognizes the main receptor CD4 and co-receptors CCR5 and CXCR4. This ultimately allows gp41 to interact directly with the cell membrane, bringing it close to the viral membrane. Once the binding is complete, the membranes of both cells fuse, and the content of the virion is released onto the cell. Membrane fusion is followed by reverse transcription of the genetic material of the virion. This step is coordinated by RT, which can perform the transcription of the viral ssRNAs into dsDNAs. RT has a p51 and a p66 domain, that correspond to the DNA polymerase and RNase H domains, respectively. HIV-1 uses a host cell tRNA, Lys3, as a primer for reverse transcription. A segment of 5'LTR, the primer binding site (PBS), recognizes and pairs to the 3' end region of Lys3. Nucleotides start being added to this region creating the first minus-strand DNA that will be transferred to the 3’ end of the viral RNA to continue elongation. As elongation of the minus-strain continues, the purine-rich sequence (PPT) adjacent to the U3 region of 3’LTR acts as a primer for the formation of a new plus-stranded DNA segment. After primer degradation, this segment will then be transferred to the 5’ end of the viral RNA to continue elongation in the opposite direction of its minus-strand counterpart (Hu & Hughes, 2012). DNA synthesis is followed by RNA degradation. After the genetic material has been converted, the process of uncoating and nuclear entry begins. The capsid of the virion is deconstructed and the pre-integration complex (PIC) containing the newly formed DNA will be transported to the cell’s nucleus. This transport mechanism remains unclear, but PIC is likely to be actively transported by host cell cofactors, such as transportin 3 (TNPO3), cyclophilin A (CypA) and cleavage and polyadenylation specificity factor-6 (CPSF6) since the whole structure is unable to diffuse through the nuclear pores due to its size (Hilditch & Towers, 2014). Integration begins moments before nuclear entry is complete. The integration of the viral DNA into the genome of the infected cell is coordinated by the viral enzyme IN. Before nuclear entry, a tetramer of integrase engages the viral DNA segments present in the PIC. Integrase clips two nucleotides from the 3’ LTR region using water molecules, leaving two hydroxyl groups at the extremities. Inside the nucleus acceptor sites for the viral DNA are recognized. After the target site is detected, the 3’ hydroxyl groups of the viral DNA induce cuts in both strands of the target DNA. Simultaneously, the hydroxyl groups connect to the 5’ phosphates of the cut. This leaves the 3’ ends of the viral and target DNAs connected, but the 5’ ends remain free. Through a process of gap repair, the DNA recombinant becomes the integrated provirus (Vandegraaff & Engelman, 2007). After genome integration, HIV-1 has a latency period that can last from a few months to possibly decades. This process remains a big obstacle to the total eradication of HIV infection (Colin & Van Lint, 2009). When integration is accomplished, transcription of the proviral DNA begins. Host RNA polymerase II (Pol II) is the enzyme that transcribes the integrated provirus. However, it is not able to transcribe the full sequence causing it to pause, resulting in short viral transcripts that are unable to support replication. This limitation is compensated by Tat, who can usurp the function of the positive transcription elongation factor b (P-TEFb). First, Tat binds to the transactivation response element (TAR), a conserved region near the 5’ end of the transcripts. Next, Tat recruits CycT1 and CDK9, two P-TEFb core subunits, to the TAR element's loop region. Finally, CDK9 phosphorylases residues in the C-terminal domain (CTD) of Pol II and negative elongation factors NELF and DSIF. After phosphorylation Pol II activity resumes, producing the full pre-mRNAs needed for replication (Ott et al., 2011; Peterlin & Price, 2006; Razooky & Weinberger, 2011). Upon completion, the pre-mRNAs are spliced into the three main classes of mRNAs necessary to originate the viral components and genome. These classes are divided according to their degree of splicing: (i) unspliced full-length mRNAs, that encode for Gag and Gag/Pol polyproteins; (ii) single spliced mRNA precursors for Env and accessory proteins Vif and Vpu; (iii) fully spliced mRNAs, encoding for regulatory proteins, Tat and Rev, and the remaining accessory proteins, Vpr and Nef (Ohlmann et al., 2014). Besides coding for Gag and Gag/Pol, the full-length mRNAs originate the genomic RNA that is later packaged into the virion progeny. Following provirus transcription comes the translation of the resulting pre-mRNA sequences. These sequences leave the nucleus and head to the cytoplasm of the cell in a process assisted by Rev. All viral components are synthesized using host cell machinery. Different mRNAs are translated by different organelles. Env is synthesized in the endoplasmic reticulum, while Gag and Gag-Pol synthesis is done by free ribosomes in the cytoplasm of the cell (Watson, 2009). There are two possible mechanisms of mRNA translation initiation: cap-dependent or cap-independent. In the main cap-dependent mechanism, specific eukaryotic initiation factors (EIFs) along with other cell factors bind to the 5’ cap and the 3’ poly-A structures of the mRNA and bring them into proximity. The resulting structure is recognized by a 40s ribosomal subunit complexed with an initiator tRNA and subsequently attached to the translation machinery of the host cell (Ohlmann et al., 2014; Sonenberg & Hinnebusch, 2009). The cap-independent mechanism via internal ribosome entry sites (IRES) remains unclear, but full-length mRNAs of HIV-1 contain two of these sequences. Studies suggest that translation can switch between the two mechanisms and that the IRES-dependent process is favored in cellular and environmental stress situations (Ohlmann et al., 2014). Upon translation, the viral components gather near the cell’s plasma membrane. Here the various domains of HIV-1 Gag mediate all necessary interactions for virion assembly, from packaging the genomic RNA to binding the plasma membrane and giving it its spherical shape. The resulting immature virion is released from the cell via budding using host ESCRT (Endosomal Sorting Complexes Required for Transport, also recruited by the p6 domain of Gag) (Sundquist & Kra, 2012). Proteolytic maturation of HIV-1 Gag is crucial and it is conducted by PR. Cleavage sites for this enzyme are located between the various Gag domains. After processing, all the previously mentioned components of Gag are released and the virion acquires its mature infectious form (Kleinpeter & Freed, 2020).

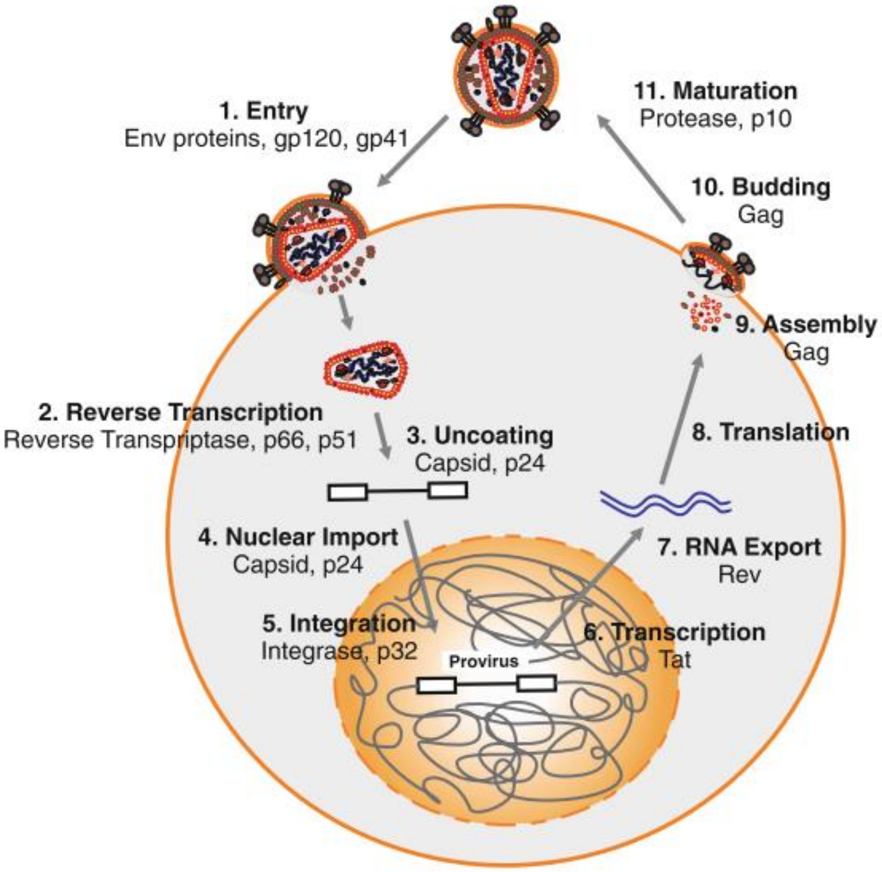


Figure 5 – Illustration of HIV replication. The process begins with (1) cell entry mediated by gp120 and gp41, followed by (2) reverse transcription of the genetic material of the virus. Next, the viral capsid deteriorates in the (3) uncoating process, allowing for (4) nuclear import of the newly synthesized DNA. This DNA will (5) integrate the host cell genome to be (6) transcribed by cell machinery. The resulting RNAs are subsequently (7) exported to the cytoplasm where they are (8) translated. Near the plasma membrane the newly synthesized precursor proteins group for (9) assembly of the immature virion. This immature virion will release itself from the cell by (10) budding. Finally, protease leads (11) virion maturation by cleaving the protein precursors into their fully functional forms. Image adapted from (Kirchhoff, 2016).

1.1.4. HIV Envelope Glycoprotein Structure and Function

Precursor gp160 is coded by the Env gene and has a length of around 856 amino acid residues. After cleavage, it gives rise to two other proteins, gp41 and gp120. gp41 contains roughly 346 amino acid residues and is more conserved than gp120. It can be divided into three domains: the ectodomain, the transmembrane region (TM), and the cytoplasmic tail. The ectodomain is further divided into smaller sections. It contains the N and C-terminal heptad repeats (NHR and CHR, respectively), an N-terminal fusion peptide (FP), a disulfide bridged loop, and a membrane-proximal external region (MPER) bordering the transmembrane region. NHR and CHR are very α-helical regions that will form the fusionready structure of HIV, the six-helix bundle (6-HB) (A. Pantophlet, 2010). gp41 and gp120 are noncovalently linked forming a trimer made up of gp120 and gp41 heterodimers. gp120 is roughly 483 amino acid residues in length and is heavily glycosylated. About 40% to 50% of its molecular weight is solely carbohydrates. Its structure contains five constant regions, C1 to C5 (which, despite their name, can have considerable variations), and five more variable regions, V1 to V5. Early studies of the gp120 structure suggested that the conserved regions composed its central core, while the variable regions formed loops located at the protein's surface (Burton & Montefiori, 1997). This was not completely accurate. Not long after this observation, researchers found they could achieve a clearer view of gp120 by complexing the protein with antibodies. Summarily, the gp120 core consists of 5 α-helices, 25 β-strands, and 10 loop segments. The core has an inner and outer domain that connect via the "bridging sheet" – a four-stranded sheet. Finally, the inner domain is more conserved than the outer domain (Kwong et al., 1998; Poignard et al., 2001).

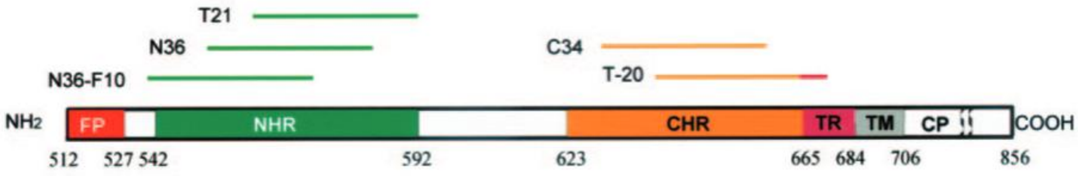


Figure 6 - Diagram representation of the envelope protein gp41. The numbers indicate the positions of the amino acid residues that delimit the various domains. The fusion peptide (FP) is followed by the N-terminal and C-terminal heptad repeats (NHR and CHR) separated by the disulfide bridge loop. After CHR, there is a tryptophan-rich area, the membrane-proximal external region (MPER). The C-terminal end of the sequence is comprised of the transmembrane region (TM) and the cytoplasmic tail (CP). The green and orange lines represent some NHR and CHR-derived fusion inhibitor peptides (FIPs). C34 is derived completely from CHR, while the sequence of T-20 overlaps slightly with the MPER region. Image adapted from (Liu et al., 2005).

Initially, gp120 recognizes the main receptor CD4 in the target cell and binds to it. This process is thought to be mediated more by the constant regions than by the variable regions, as previous studies indicate that core binding activity remained normal despite removing the V1/V2 domain and the V3 loop (James et al., 1998). After CD4 binding, the subunit will also bind to co-receptors CCR5 or CXCR4. The V3 loop appears to have a key role in this process, as some of its highly conserved residues allow the structure to interact with said co-receptors (Sirois et al., 2005). When this step is complete, the gp120/gp41 complex suffers a conformational change that leads to the exposure of the FP region of gp41.

FP will then insert itself into the cell membrane of the target cell. At this point, the cell and the virus are connected by the FP and TM regions of gp41. Subsequently, the disulfide bridge loop will fold, bringing NHR and CHR into proximity and forming the 6-HB. This structure, in turn, brings the cell and viral membranes closer together leading to the formation of a pore that will allow the virus to transfer its contents into the host cell (Top & Chem, 2016).

1.1.5. HIV Fusion Inhibitor Peptides

There are currently nine drug classes of HIV medicines available for antiretroviral therapy (ART). This treatment consists of taking a combination of anti-HIV drugs from different classes every day. In general, different classes disrupt specific steps in the virus's life cycle. Nucleoside Reverse Transcriptase Inhibitors (NRTIs) inhibit genomic RNA transcription by blocking RT. Non-Nucleoside Reverse Transcriptase Inhibitors (NNRTIs) also inhibit RT-mediated transcription by binding to it and altering its function. Protease inhibitors (PIs) block PR function, interfering with the virions' maturation. Fusion inhibitors block viral membrane fusion, denying the release of the genetic material of the virus into CD4 T lymphocytes. CCR5 antagonists interrupt co-receptor binding by gp120 by blocking the CCR5 co-receptors. Integrase Strand Transfer inhibitors (INSTIs) disturb the integration process by blocking IN action. Attachment and Post-Attachment Inhibitors block viral entry by binding to gp120 and by blocking the CD4 receptors, respectively. Finally, Pharmacokinetic Enhancers, as the name suggests, enhance the effects of other anti-HIV drugs used in a treatment regimen (*FDA-Approved HIV Medicines | NIH*, n.d.).

Fusion inhibitors can be classified based on other aspects, such as their composition and where they bind. The four classes of HIV fusion inhibitors are: (i) peptides binding to NHR (or HR1) of gp41; (ii) peptides binding to CHR (or HR2) of gp41; (iii) peptide-mimetic fusion inhibitors; (iv) non-peptide fusion inhibitors (Qadir & Malik, 2010). All compounds from these classes function by blocking the formation of 6-HB. They do so by binding to the CHR and NHR regions of gp41. This stops the domains from being brought into proximity, effectively denying membrane fusion. The only FDA-approved fusion inhibitor today is Enfuvirtide (T-20), a 36 amino acid peptide derived from the CHR region of gp41 that binds to the NHR region of gp41 (*FDA-Approved HIV Medicines | NIH*, n.d.; Greenberg & Cammack, 2004). Many other fusion inhibitors have been developed over the years but are still under research.

C34 is a fusion inhibitor peptide (FIP) derived from a more N-terminal region of CHR, comparatively to T-20. It is composed of residues 628-661, while T-20 is made up of residues 638-673. Therefore, C34 is a shorter peptide containing 34 amino acid residues, compared to the 36 residues of T-20, but both peptides act similarly (D. Zhang et al., 2015). Research shows that C34 has higher inhibitory activity than T-20 in in-vitro assays. The structure of CHR, C34, and C34 derivatives can be seen in Fig. 4. Amino acids in positions a, d, and e constitute the interactive sites of the peptide that will interact with NHR to form the 6-HB. Positions b, c, f, and g are solvent-accessible sites that do not interact with NHR but are responsible for the stability of the peptide in solution. C34 serves as a template for many other fusion inhibitor peptides. Some C34 derivatives include SC34, SC34EK, SC35EK, SC29EK and SC22EK. These were derived from the sequence of C34 by substituting amino acids in the solvent-accessible sites of the peptides with glutamate (E) and lysine (K). No substitutions were made in the interactive sites to avoid harming the activity of the peptides. Glutamate residues were introduced in positions i, i+1 and Lysine residues were introduced in positions i+4, i+5. This leads to the repetition of a Z-EE-ZZ-KK pattern in the sequence, where Z represents the original amino acid residue. SC34 and SC34EK have three incomplete Z-EE-ZZ-KK motifs and have a Norleucine (X) instead of a Methionine (M) in the second position of their sequence. SC34 differs from SC34EK because it contains one KE motif (instead of the usual EK) to resemble the original sequence of C34. SC35K repeats the regular Z-EE-ZZ-KK motif throughout its entire sequence, and it contains an extra Lysine residue at the end of it, making it the longest of the C34-derived peptides with 35 amino acid residues. SC29EK and SC22EK were synthesized from SC35EK by removing some of the previously mentioned motifs. This was done to test how many of these were needed for the peptide to maintain its inhibitory activity. SC34EK exhibited the highest inhibitory activity of all the peptides, stronger than the parent peptide C34. SC35EK and SC34 followed with activity comparable to that of C34. SC29EK activity was slightly lower but still at the level of the other peptides. SC22EK's activity was significantly reduced to very low levels of inhibition. With these results, researchers concluded that peptides needed at least four Z-EE-ZZ-KK motifs to maintain their effect. The C34-derived peptides were more soluble than the parent peptide and more effective at stopping membrane fusion of wild-type (WT) HIV-1 and T-20-resistant strains of HIV-1 (Izumi et al., 2009; Miyamoto & Kodama, 2012; Naito et al., 2009; Okata et al., 2002).

Figure 7 – The structure of the heptad repeat can be seen in the first row (letters a to g), and the amino acid sequence of CHR, from which C34 originates, is indicated below it. Sequences of C34 and its derivates and their inhibitory activity relative to T-20 are also shown below. The greyed areas (a, d, and e) represent the positions of the interactive sites of the peptide. The remaining ones (b, c, f, and g) correspond to the solvent-accessible sites. SC34 and SC34EK contain a Norleucine residue in the second position (indicated in purple) instead of the original Methionine. The complete and incomplete EK motifs can be seen across the sequence of all derivates. Glutamate residues are indicated in pink and Lysine residues are indicated in blue. Image adapted from (Miyamoto & Kodama, 2012).

The addition of Methionine (M) and Threonine (T) residues at the beginning of the sequence results in the formation of a hook-like structure. This structure is termed the M-T hook, and research indicates that it contributes positively to the thermostability and inhibitory activity of the FIPs. There is information regarding MT-C34 (C34 with Methionine and Threonine residues to its N terminus region) that suggests it binds more strongly to the 6-HB, and that the structure is more thermostable. It also displayed several-fold more potent inhibitory activity against WT HIV-1 and T-20-resistant strains (Chong et al., 2012b). Information about the other C34-derived peptides containing the M-T hook region is lacking, but the MT-SCs are an interesting family of FIPs that can harbor many interesting properties.

Cell Membranes and Lipids

1.2.1. Fundamental Structure and Function of Cell Membranes

The cell membrane (CM), or plasma membrane, outlines and protects the cell and its components. It also separates the intracellular space from the extracellular space. This allows for the establishment of two environments with different chemical compositions, which is key, as the cell can only function and survive under a highly restricted range of conditions. The intracellular environment of the cell is also compartmentalized via various membranes that envelop the different organelles present in the cytoplasm. This compartmentalization is also essential because it maximizes the efficacy of the cell processes by providing the ideal conditions for the organelles to function. The core structure of the organelle membranes is the same, but their composition can vary slightly. The CM is responsible for controlling the chemical gradients inside and outside of it. Nutrients, gases, and various other solutes can enter the cell through incorporated transport mechanisms mediated by the CM. The membrane also allows the removal of the products of cell digestion by fusing with the membrane of digestive vesicles called lysosomes. As such, it must be capable of recognizing and differentiating these different substances and structures. Therefore, we say that the CM is selectively permeable (Yeagle, 2016).

The basic structure of biological membranes is the lipid bilayer. These structures are fluid and are formed through a self-assembling process, due to the amphipathic nature of the molecules that comprise it. Besides lipids, they contain proteins and carbohydrates. These substances interact with membrane lipids and mediate many processes, like transmembrane transport, cell signaling, energy storage, etc. The three main classes of lipids that make up the bulk of the membrane are: phospholipids, glycolipids, and sterols. These different lipid classes are present in different quantities depending on the membrane type (Ding, 2017).

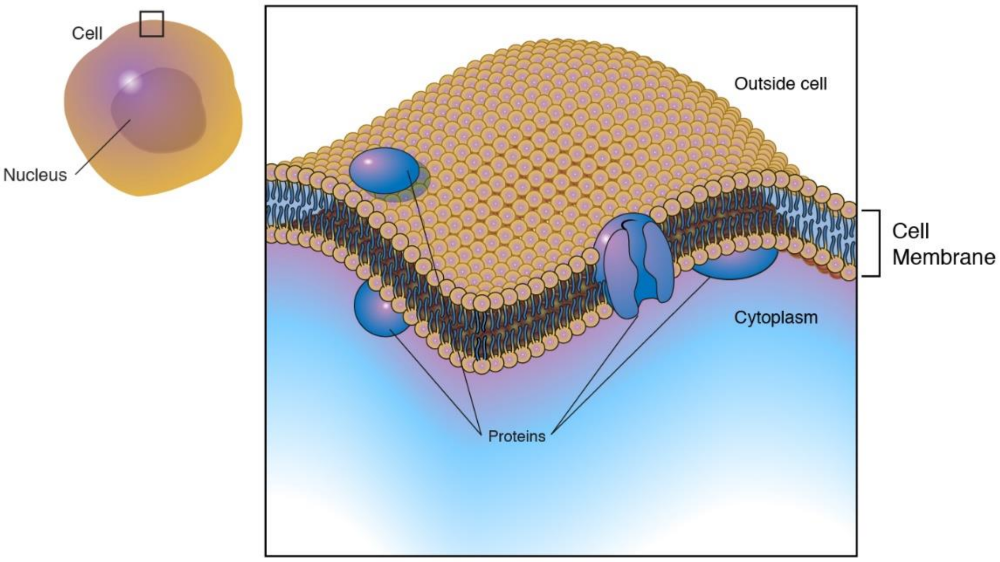


Figure 8 – Lipid bilayer model showing fluid nature of the membrane, as well as the organized phospholipids, along with surface and transmembrane proteins. This model does not include cholesterol molecules in the intermembrane space. Image adapted from (National Human Genome Research Institute, 2023).

1.2.2. Phospholipids

The core structure of a phospholipid consists of a glycerol molecule with two hydrocarbon chains (fatty acids) esterified to positions C_1 and C_2 and a phosphate group esterified to position C_3. The head group is comprised of a glycerol molecule and a phosphate group. This phosphate group, which carries a negative charge, makes the head group polar and hydrophilic. The hydrocarbon chains are not charged, but they attract each other due to the hydrophobic nature of both. Therefore, phospholipids are amphipathic molecules since they have both hydrophilic and hydrophobic properties. The most important phospholipids of the family and their differences will be listed individually ahead (Yeagle, 2016).

Phosphatidic Acid (PA) is the simplest of all phospholipids and it serves mostly an intermediary for the formation of other more complex membrane or storage lipids. PA has also been shown to be involved in signaling pathways of various regulatory processes in the cell, such as cell growth. It is present in the membrane at very low concentrations, which are maintained by lipid phosphate phosphohydrolases that convert it into the metabolic precursor of many other lipids, diacylglycerol (DAG) (Wang et al., 2006).

Phosphatidylethanolamine (PE) is the second most common phospholipid in mammalian cells, and the most common in bacteria. This is due to bacteria lacking the means to convert PE into phosphatidylcholine (PC) (Stillwell, 2016). It is 15-25% of the total lipids of mammalian cells. PE is a cone-shaped non-bilayer phospholipid. Its shape is due to the polar head group having a small diameter compared to the fatty acid tails. It forms non-bilayer structures, such as the inverted hexagonal phase (H_{II}), due to its tendency of promoting negative curvature in the membrane. This structure is an alternative to the liquid crystalline lamellar phase (L_α) bilayers, which normally originate from cylinder-shaped lipids like PC (Ball et al., 2018; Calzada et al., 2016). Even though PE is a non-bilayer-forming phospholipid, it plays numerous roles in the cell. It acts as a precursor for other lipids and contributes to overall mitochondrial stability by aiding the events of mitochondrial fusion and fission (Calzada et al., 2016).

Phosphatidylcholine (PC) is the most common phospholipid in animal cell membranes, constituting nearly half of the membrane’s lipids (45-50%). PC is obtained mainly via the Kennedy pathway, through the addition of CDP-choline to DAG. But, it can be obtained through other processes, such as the triple N-methylation of PE. It is a zwitterion, meaning it has a partial positive and negative charge, but no net charge. This aspect is important since it eliminates the problem of charge repulsion that is common with other phospholipids. PC also plays a role in cellular signaling and is an important source of signaling molecules. It has also been found that disturbances in the pathways of PC formation can disturb its homeostasis, potentially leading to cell death (Cui & Houweling, 2002; Stillwell, 2016).

Phosphatidylserine (PS) is one of the most functionally diverse phospholipids. It is located exclusively in the inner leaflet of the plasma membrane and gradually accumulates in the outer leaflet as we age. PS is more prevalent in the membranes of neurons. PC and PE can be converted into PS by substituting the choline and ethanolamine groups with serine. PS is involved in the activation of neuronal signaling pathways and neurotransmitter release. Docosahexaenoic acid (DHA) is one of the major constituents of PS, and the decline in DHA content lowers PS levels, which can have some implications for cognitive function (Kim et al., 2014; Stillwell, 2016).

Phosphatidylinositol (PI), much like PS, is present in many different tissues but is present in higher quantities in the brain. It plays a small structural role in the membrane but can be converted into more functionally active phospholipids. Phosphoinositides (PPIns) are derived from PI by phosphorylation and can interact with membrane proteins. The amount and position of these phospholipids are heavily regulated in the cell by a multitude of kinases and phosphatases (Beziau et al., 2020). PPIns can affect cell shape and motility by remodeling the cytoskeleton of the cell, and they are involved in multiple cell-signaling pathways associated with cell growth and death. One of the enzymes responsible for the regulation of PPIns, Phosphatidylinositol 3-Kinase (PI3K), is an oncogene that, if activated, can lead to uncontrolled growth of the cell (Goncalves et al., 2018; Stillwell, 2016).

Phosphatidylglycerol (PG) is one of the lipids with the lowest distribution in mammalian cells. PG has a similar structure to PI, which explains why they may act identically in some cases, and to cardiolipin (CL). PG is found more abundantly in lungs and thylakoid membranes. Due to its low levels of expression, it plays a rather small structural role, even so, PG is necessary as a precursor for CL. In mammalian cells, PG also appears to interact with several proteins and lipids that are essential for processes like PC transfer between membranes, activation of RNA synthesis and more (Furse, 2017). In lung surfactant, PG may be used to evaluate the maturity of a baby’s lungs, and it can also serve as an indicator of lung function (Stillwell, 2016).

Cardiolipin (CL), or diphosphatidylglycerol, is possibly the most unique out of all the phospholipids presented so far. It is almost entirely localized in the inner membrane of mitochondria (where it is also synthesized) and has a very odd structure (Houtkooper & Vaz, 2008). CL is a lipid dimer made up of two PA portions connected to glycerol. Consequently, its head group has two separate negative charges (due to the two phosphate groups from each of the PAs) and four acyl chains connected to it (Stillwell, 2016). CL activity is crucial for ATP synthesis since it modulates the activity of the complexes in the electron transport chain (ETC) and organizes them into their optimal configuration. It also appears to play a role in mitochondrial biogenesis and apoptosis (Paradies et al., 2019).

1.2.3. Cholesterol

Cholesterol is the most recognizable and well-known lipid of the sterol group due to its close connection to cardiovascular disease in humans. Despite this, it is a major structural component of the cell membranes of animals (Stillwell, 2016). It helps maintain the fluidity and integrity of the cell membrane, while also serving as a precursor for many other compounds, such as steroid hormones and vitamin D (Zampelas & Magriplis, 2019).

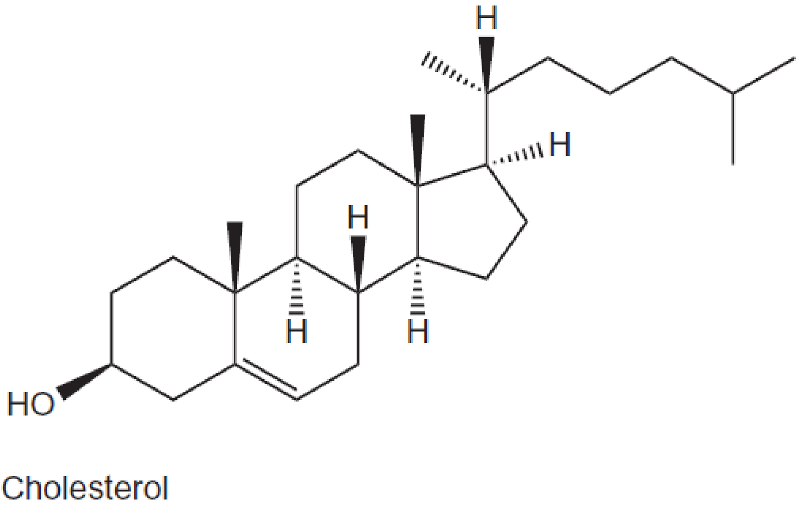


Figure 9 - Chemical structure of cholesterol. The four hydrocarbon rings are linked to a hydroxyl group (OH) at C_3 and a hydrocarbon chain at C_{17}. The ring structure and hydrocarbon tail are buried in the hydrophobic site of the membrane, while OH interacts with the water interface. Image adapted from (Yeagle, 2016).

Like all other sterols, cholesterol has a sterol core, a structure consisting of four hydrocarbon rings. Additionally, it has a small polar head group consisting solely of a hydroxyl group (OH) attached to the third carbon of the ring. It also has a short, saturated hydrocarbon tail attached to the seventeenth carbon of the ring (Fig. 6). OH is polar and, as such, it interacts with the water interface, on the other hand, the four rings and the hydrocarbon tail cannot since they are nonpolar. Removal of the hydrocarbon tail and modifications to the ring structure of cholesterol affects the way the molecule acts (Yeagle, 2016).

As previously mentioned, cholesterol can be catabolized into bile acids. To maintain homeostasis individuals must be able to reestablish the balance of cholesterol levels. This can be achieved in two different ways: uptake through diet or biosynthesis. Cholesterol biosynthesis takes place in the endoplasmic reticulum and most of it heads out to the plasma membrane of that cell. There are also various mechanisms of cholesterol transport between membranes, which explains the differences in cholesterol content of different cell types. A rise in cholesterol levels reduces the passive permeability of the membrane because the membrane lipids will be packed more tightly, which also reduces membrane protein flexibility. The activity of some membrane proteins can also be directly controlled by cholesterol, which means that cholesterol can have large implications for cell growth and function (Yeagle, 2016).

1.2.4. Membrane Dynamics

A CM is a heterogenous and highly complex environment, in which lipids and many proteins can move within it. This leads to the formation of domains with varying compositions that give rise to different membrane phases. Factors like sterol concentration and temperature can affect membrane organization, leading to the formation of many possible lipid phases. The two main phases that are important to understand in this study are the liquid-ordered phase (Lo) and liquid-disordered phase (Ld). The Ld phase occurs at higher temperatures. It is a highly fluid phase of the membrane, where the lipids are not packed very tightly. Lipid mobility is high, but Van der Waals interactions between fatty acid chains are low due to kinks in their structure. These kinks arise due to the movements of these chains caused by the higher temperatures. The Lo phase occurs in membranes containing sterol molecules. This phase is less fluid since lipids are more tightly packed due to the added rigidity of the sterols. It is characterized by its higher rigidity and structural integrity, without compromising lipid lateral diffusion (Ackerman & Feigenson, 2015).

As was previously mentioned, cholesterol is an important regulator of membrane fluidity and phase behavior. Cholesterol has an ordering-disordering effect, where it decreases order at low temperatures (by separating fatty acid chains) and increases order at high temperatures (by providing more rigidity to the membrane and stability to lipids). These effects are proportional to its concentration. This property of cholesterol is crucial to maintain membrane fluidity across a wide range of temperatures, allowing it to function correctly (Dufourc, 2008).

1.2.5. Lipid Rafts

The fluid mosaic model proposed by Singer and Nicholson in 1972 describes the membrane as being laterally heterogenous, dynamic and fluid, and containing embedded proteins. Additionally, the proteins and lipids that constitute the membrane can move around, rather than being fixed in place (Singer & Nicolson, 1972). To this day, this model is the foundation for the study of membranes. Since then, preferential interactions between membrane lipids, such as sphingolipids and cholesterol, have been established. The lipid raft hypothesis extends on the fluid mosaic model, by suggesting that these interactions result in small, ordered membrane domains (Erdinc Sezgin, Ilya Levental, Satyajit Mayor, 2017; Levental et al., 2020). These supposed microdomains also contain membrane proteins, such as glycosylphosphatidylinositol-anchored proteins (GPIs), and appear to serve as platforms for processes of cell signaling, membrane trafficking and protein sorting (Mayor & Pagano, 2007; Mayor & Riezman, 2004; Simons & Toomre, 2000b). Lipid rafts have been a subject of extensive studies for decades, and although there is a great deal of evidence supporting their existence, the exact nature of their composition is unclear. Initially, the only evidence for the existence of these domains was the presence of small regions of the membrane resistant to detergent extraction and mechanical disruption. However, more recent direct observational studies performed in membranes indicated the existence of small, dynamic and selective cholesterol-related heterogeneity in the plasma membrane of cells (Lingwood & Simons, 2010). A model structure of a lipid raft can be seen in Fig. 7.

Currently, two types of rafts have been described: planar lipid rafts and caveolae. Planar lipid rafts are flat, while caveolae are invaginated, bending inward due to the action of the transmembrane protein caveolin-1. Lipid rafts are believed to be ephemeral lasting only from a few seconds to minutes, but their exact duration is still being debated (Lingwood & Simons, 2010). Some evidence even suggests that the domains can be stable under certain conditions (Harder et al., 1998). Their assembly seems to be dependent on the amount of free cholesterol and sphingolipids in the CM. The CM and the lipid rafts exhibit different phases. Due to higher concentrations of cholesterol, raft domains are in a Lo phase, as was previously discussed. In the common CM, the Ld and Lo phases coexist, giving it a high degree of fluidity, while maintaining a good amount of structural integrity. In this conformation, lipids can diffuse more freely along the membrane (Elson et al., 2010; Simons & Toomre, 2000a). The higher fluidity of the CM allows the raft microdomains to move around as a unit along the membrane’s surface.

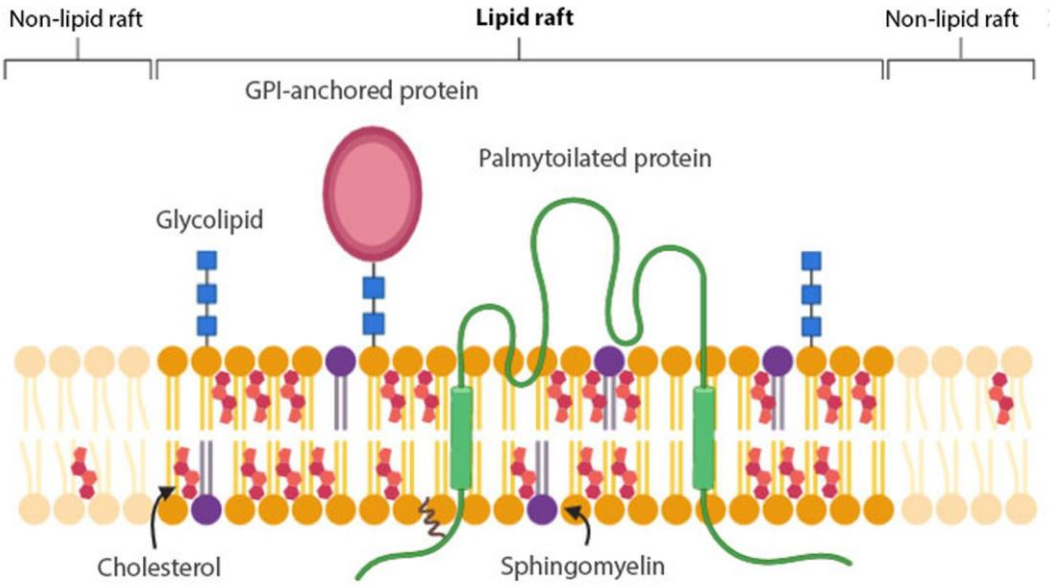


Figure 10- A model structure of a lipid raft and its constituents. Raft domains have higher cholesterol concentrations than non-raft domains, making them more cohesive. Besides cholesterol, glycolipids, sphingomyelin, GPI-anchored proteins, and palmitoylated proteins can be seen. This model structure is likely incomplete since direct observation of lipid rafts has proven very difficult, meaning that there may be elements missing from this representation. Image adapted from (Ripa et al., 2021).

As mentioned above, functions of lipid rafts include protein sorting, membrane trafficking and signal transduction. Lipid rafts appear to be crucial in innate and acquired immunity and are involved in many immune signaling pathways. This happens because lipid rafts can contain Toll-like receptors (TLRs) and c-type lectin receptors (CLRs) that detect pathogens and initiate signaling cascades that lead to their clearance from the cell. These receptors are useful for processes of endocytosis, exocytosis, and cell signaling, but they can also be exploited by pathogens to enter and exit the cells. Viruses, such as HIV-1, influenza, and human herpes virus-6 (HHV-6) use them to enter the cell (Kulkarni et al., 2022). Additionally, enveloped viruses, such as HIV-1 use ESCRT-dependent budding mechanisms to leave the cell. These ESCRT proteins can be present in lipid rafts, so rafts are likely linked to viral budding, suggesting their involvement in pathogen-host interaction (Meng et al., 2015). Rafts can also improve immune response. For example, they can contain many co-receptors responsible for activating and even augmenting B-cell receptor (BCR) activation, which will lead to a better immune response. This is also the case for T-cell receptors (TCRs), proving that rafts may play a role in lymphocyte activation. Adaptor proteins in rafts have also been shown to regulate IgE signaling by promoting or inhibiting the degranulation of mast cells (Rivera & Arudchandran, 2001). Finally, rafts are also involved in autoimmune disorders, such as systemic lupus erythematosus (SLE) and rheumatoid arthritis, which are both characterized by impaired TCR signaling. Disrupting raft activity has been shown to delay the progression of SLE, and a raft adaptor protein responsible for TCR activation is dissociated from rafts in cases of rheumatoid arthritis (Grinnell et al., 2005; Jury & Kabouridis, 2004). This also means that rafts are possible therapeutic targets for these diseases.

Molecular Dynamics Simulations

Molecular Dynamics (MD) simulations are in short, the time integration of Newton's equations of motion. This integration is performed in discrete time-steps. The position and momentum of every particle are calculated as a function time. In MD, Newton's equations are solved at the same time for every atom in a system (Hug, 2013). The first molecular dynamics simulation was performed by Alder and Wainwright, where they simulated a system composed of hard spheres in the late 1950s (Alder & Wainwright, 1957). A protein system wasn't simulated until 20 years later, when McCammon and colleagues ran a simulation of a bovine pancreatic trypsin inhibitor (McCammon et al., 1977). Today, MD simulations are widely used in various fields of research. This rise in popularity was, in part, due to the advances in imaging techniques, such as X-ray crystallography, cryogenic Electron Microscopy, Nuclear Magnetic Resonance spectroscopy and others. These techniques provided novel structures for analysis and increased their resolution significantly. Another reason was the rise in computational power over the last decades (Hollingsworth & Dror, 2018). The improvement in computer hardware (particularly in GPU and CPU power) made it so that MD simulations that previously required a supercomputer, could be performed by regular computers.

Before MD simulations were available, researchers had a trouble understanding how certain systems would evolve. The comparison between initial and final conformations of a protein or other biomolecules already provides a large amount of information, but it would be ideal to watch how they interact with one another, the changes in conformation and many other aspects down to the atomic level. This is exactly what MD simulations do, they use the information regarding the physics governing the system to predict the movement of individual atoms over time (Karplus & McCammon, 2002).

MD can be used to study a variety of qualities of a given system. In this study MD simulations were used to see how peptides adsorbed to the surface of a lipid bilayer and the influence of said peptides on its properties.

1.3.1. Integrator

Atomic trajectories in classical mechanics for a system with N particles, obey Newton’s second law of motion, that can be formulated as the following equation:

$$\vec{F}_i = m_i \ddot{\vec{r}}_i \tag{1}$$

or, also, as two distinct equations:

$$\begin{cases} \dot{\vec{r}}_i = \dfrac{\vec{p}_i}{m_i} \ , \\ \dot{\vec{p}}_i = \vec{F}_i \end{cases} \tag{2}$$

These are second-order and first-order differential equations, respectively. The term $\dot{\vec{r}}_i$ is the derivate of $\vec{r}_i$ with respect to time (t), while m_i represents the mass of the particle. For a three-dimensional system, such as the one used in this study, 3N second-order equations must be solved every time-step, or similarly, 6N first-order equations (Hug, 2013).

Integration schemes are be used to find numerical approximations of ordinary differential equations, such as the ones mentioned previously. There are several conditions that an algorithm must satisfy to function properly. It must be capable of reproducing the differential equations in the limit of the step size. It must solve these differential equations very accurately. It must be efficient in terms of computational power, by allowing, for example, longer step sizes. The algorithm must also be time reversible, like Newton's equations of motion. This means that, by knowing the state of the system at a certain time, t, the integrator should be able to accurately predict the state of that same system in the past or future. Additionally, it should be symplectic, that is, it should be able to maintain the symplectic conditions of phase space. Symplectic integrators can conserve the total energy of the system and maintain its volume from one time-step to another. Symplectic systems are favored in longer simulations since they are more accurate long term. Lastly, the algorithm should be easy to implement (Frenkel & Smith, 2002). The integration scheme used in this study is known as the *leap-frog algorithm* and it is a variant of the original Verlet integration method. This method uses the velocities at time $t - \frac{1}{2}\Delta t$ and positions at time t to update the new positions, based on the force $F(t)$ acting in particles at time t (GROMACS, 2020). It is a second-order algorithm that is easily modifiable for different systems and is fairly accurate in systems with higher degrees of complexity. When very accurate integration is required, it is preferable to use the Verlet algorithm. The equations that constitute the *leap-frog algorithm* are as follows:

$$\vec{v}_i\left(t + \frac{\Delta t}{2}\right) = \vec{v}_i\left(t - \frac{\Delta t}{2}\right) + \frac{\Delta t}{m_i}\vec{F}_i(t) \tag{3}$$

$$\vec{r}_i(t + \Delta t) = \vec{r}_i(t) + \Delta t \vec{v}_i\left(t + \frac{\Delta t}{2}\right) \tag{4}$$

This algorithm estimates the new half step velocities $\left(t + \frac{\Delta t}{2}\right)$ by considering the position and force at a time *t* and the velocity at half a time-step $\left(t - \frac{\Delta t}{2}\right)$. Then by using the velocities from both the previous and next half step, the velocity at a full time-step can be obtained via:

$$\vec{v}_i = \frac{\vec{v}_i\left(t + \frac{\Delta t}{2}\right) + \vec{v}_i\left(t - \frac{\Delta t}{2}\right)}{2}$$

(5)

The explicit velocities obtained from the leapfrog algorithm will allow for the estimation of the coordinates in the next time-step (Hockney et al., 1974; Hug, 2013).

1.3.2. Thermostats and Barostats

In MD simulation experiments it can be beneficial to let the total energy of a system drift to maintain its pressure and temperature constant. The natural ensemble for a MD simulation is the NVE, a microcanonical ensemble where the number of particles, volume and energy of a system are kept constant. Even so, there are alternatives to this ensemble which may be useful in the context of certain experiments. For example, the canonical ensemble (NVT) allows fluctuations in total energy to keep the pressure and temperature constant. The isothermal-isobaric ensemble (NPT) on the other hand, allows for the same temperature and pressure fixation by allowing changes in system volume (Hug, 2013). Thermostats and Barostats are the algorithms that perform this temperature and pressure control, respectively. In this study, the Berendsen Algorithm was used to perform both controls (Berendsen et al., 1984).

In the Berendsen method, a heat bath weakly coupled to the system performs the temperature control. This heat bath senses when the kinetic temperature of the system deviates from the desired temperature, via

$$\frac{dT}{dt} = \frac{T_d - T}{t_T}$$

(6)

where T_d represents the desired temperature and t_T is the time constant. It then corrects them by rescaling the velocities of particles at the needed time-steps (Δt):

$$\chi = \left[1 + \frac{\Delta t}{t_T}\left(\frac{T_d}{T} - 1\right)\right]^{1/2}$$

(7)

This method can cause problems of energy conservation in certain experiments, therefore not generating a canonical ensemble (Berendsen et al., 1984; Harvey et al., 1997).

The pressure control performed by the Berendsen method is similar to its temperature control. First, any changes in pressure that deviate from the desired pressure are detected by

$$\frac{dP}{dt} = \frac{P_d - P}{t_p} \tag{8}$$

where once again, P_d is the desired pressure and t_p the time constant that determines the rate of correction. To correct this variation, the volume of the system is re-scaled in the time-steps of interest:

$$\chi = 1 - \beta_t \frac{\Delta t}{t_p}(P_d - P) \tag{9}$$

where β_t represents the isothermal compressibility. The Berendsen barostat and thermostat are easily appliable together in MD simulations (Berendsen et al., 1984; Hug, 2013).

1.3.3. Periodic Boundary Conditions and Neighbor Lists

Simulation systems can be open, closed or maintained with periodic boundary conditions (PBC). Open systems are suboptimal since the particles will move with no restrictions across an infinite space. In closed systems, particles cannot leave the defined area, which can lead to surface effects. These can affect the behavior of the molecules in the system, altering their regular physical behavior. PBC can be used in various systems and most of the time they are the more natural choice (Hug, 2013).

When PBC are used, images of the original simulation box (central box) are replicated an infinite number of times around it. All the images are completely identical to the central box and so are the movements of their particles. By allowing particles to travel across boundaries, surface effects are greatly diminished. This also makes it so that when a particle leaves the system, that same particle in one of the images enters the central box on the opposite side. Only the properties from the original simulation box are recorded (Miettinen, 2010). Short-ranged forces in PBC systems are estimated using a technique called Minimum Image Convention (MIC), where particles in the system interact with the periodic images closest to them. Even though technically there can be an infinite number of images, the system itself is not infinite. This means, for example, that a particle can interact with its own periodic image (Mandell, 1976). To solve this problem, another MD simulation technique, called Neighbor Lists, is used to define a cutoff distance after which particles do not interact with each other.

Neighbor Lists (NL) are a technique that was developed to reduce computation time and CPU usage in MD simulations. The number of interactions in a system increases exponentially with the number of particles. NLs work by setting a cutoff radius, forming a small circle around every particle in the system. This list is empty initially, but it will search for neighboring particles within a pre-defined cutoff radius. Particles outside this radius are not considered in potential calculations. The cutoff radius must be selected carefully so that the particle does not interact with its own periodic image and to avoid unnecessary long-range interactions (Frenkel & Smith, 2002; Verlet, 1967). NLs need to be updated because the particles of the system will constantly be moving in and out of the circle enveloping the particle. This updating is triggered via a pre-determined time interval. This interval is normally 10-25 time-steps to achieve a balance between computational costs and accuracy in force calculations. The Verlet NL that was used throughout this study updated via the analysis of particle displacement. In this method, a cutoff radius (r_{cut}) is defined. After this a "skin depth" (r_v) (slightly larger than the cutoff radius, r_{cut}) is setup. Particles located within these radiuses constitute the NL for that particle (Frenkel & Smith, 2002).

1.3.4. Constraints

Applying constraint techniques to a system is another way to reduce computation time in MD simulations. Bond vibrations occur at high frequencies, requiring the use of way smaller time-steps. To reduce the complexity of the systems, holonomic constraints are applied to bond properties, like their length and angle. This way, the evolution of the system could be evaluated solely based on the coordinates of the particles (and perhaps their velocities) without requiring information about their acceleration, which would make it impossible to integrate the constraints into the equations of motion. This makes it possible to use slightly larger time-steps, while also providing a good representation of the system (Hug, 2013).

In this study the constraint algorithm used was the Linear Constraint Solver (LINCS), a Lagrange multiplier-based method. This constraint algorithm was developed to be used with Verlet-type integrators, including the leapfrog algorithm used in this study. LINCS is very stable, as it resets the constraints instead of the derivates of the constraints, hereby eliminating drift (Hess et al., 1997). For this method, the positions of the *N* particles in the system are contained within a 3*N* vector $\vec{r}(t)$. The matrix formulation of Newton's equation of motion for a system of *N* particles is as follows:

$$\frac{d^2\vec{r}}{dt^2} = M^{-1}f \tag{10}$$

In this equation $\vec{r}$ is the previously mentioned vector, containing the Cartesian coordinates of the particles of the system. f is the 3*N* force vector and *M* is a $3N \times 3N$ matrix that contains the masses of the particles. Furthermore, K time-independent constraint equations are applied to the system:

$$g_i(r) = 0 \qquad\qquad i = 1, \dots, K \tag{11}$$

The constrained system can also be described by $3N$ second-order differential equations. The constrained equations of motions, assuming that $T = M^{-1}B^T(BM^{-1}B^T)^{-1}$, are:

$$\frac{d^2\vec{r}}{dt^2} = (I - TB)M^{-1}\vec{F} - T\frac{dB}{dt}\frac{d\vec{r}}{dt} \tag{12}$$

T is a $3N \times K$ transformation matrix that transforms the motions of the constrained particles into motions, but in Cartesian coordinates. $I - TB$ is a projection matrix that sets the constrained coordinate to zero. B is a $K \times 3N$ matrix containing the directions of the constraint of interest. $BM^{-1}\vec{F}$ is a K vector that contains the second derivates of the bond lengths in the direction of the bonds. Lastly, the final term $\frac{dB}{dt}\frac{d\vec{r}}{dt}$ describe the centripetal forces originated from rotating bonds (Hess et al., 1997).

LINCS is faster than other constraint algorithms, such as SHAKE and RATTLE, but can only be used for isolated angle constraints. The other algorithms mentioned should be used for other angle constraints.

1.3.5. Force Fields and Interatomic Interactions

Force fields (FFs) consist of mathematical expressions that describe how atoms and molecules interact with each other (Duffield et al., 2011). The parameters in the force fields are derived from a combination of experimental data and quantum mechanical calculations. Today, many different FFs are used in MD simulations to study biological systems, but Eq. 1.1 shows a mathematical expression for a typical FF:

$$\mathrm{U} = \sum_{bond}\frac{1}{2}k_b\,(r-r_0)^2 + \sum_{angles}\frac{1}{2}k_a\,(\theta-\theta_0)^2 + \sum_{torsions}\frac{V_n}{2}[1-\cos(n\phi-\delta)] + \sum_{improper}V_{imp} + \sum_{LJ}4\in_{ij}\left(\frac{\sigma_{ij}^{12}}{r_{ij}^{12}}-\frac{\sigma_{ij}^{6}}{r_{ij}^{6}}\right) + \sum_{elec}\frac{q_iq_j}{4\pi\in_0 r_{ij}} \tag{13}$$

The first four terms describe intramolecular contributions to total energy. The first term is related to the deformation (compression/stretching) of chemical bonds. This term is normally defined by a simple harmonic function; therefore, it is impossible to account for the breaking of chemical bonds. Compressing or stretching a bond requires a lot of energy, so the value of their force constant is usually very high. The second term is associated with deformation of angle geometry. In this case, their force constant is lower since it’s easier to deviate an angle from its reference point compared to deforming a bond. The third term involves lower levels of energy compared to the previous two and is related to the rotation of a molecule about specific dihedral angles and torsional motions. Torsional terms are used for molecules with four or more atoms in a row. The first two terms can be replaced by rigid approximations, but the torsional term is necessary to properly reproduce the conformational changes in molecules. The fourth term provides a sort of penalty for bending of certain groups and is used to preserve their planarity and chirality. The last two are the “non-bonded” terms. They account for electrostatic interactions and Van der Waals forces. The van der Waals term (Lennard-Jones potential) contains repulsive and attractive factors. Repulsion between atoms normally occurs at distances below 0.3 nm, when electron clouds of atoms overlap. Attraction arises from interactions between induced dipoles. Finally, the electrostatic potential is measured via two different techniques. Short-range interactions are described using Coulomb’s law, while long-range interactions require the use of specific techniques, like the Particle-Mesh Ewald (PME). In this method, charges are assigned to atoms based on their electronic structures, and the simulation box is divided into a three-dimensional grid. After distributing the charges through this grid via Fast Fourier Transforms (FFTs) and Ewald summation methods, long-range electrostatic forces are calculated (Darden et al., 1993). Most FFs nowadays are quite similar in their function and parameters, which makes them much more convenient. This makes it so that, for example, certain parts of a FF can be fully adapted from a different one (Monticelli & Tieleman, 2013).

FFs can be roughly classified as atomistic and coarse-grained. As the names suggest, atomistic FFs describe the interactions between every single atom in the system and coarse-grained FFs group together atoms while attempting to maintain the fundamental physical properties of the particles. That aspect of coarse-grained FFs is what makes them viable for use in systems with more atoms, at longer time scales and length scales of molecules (Hug, 2013). Therefore, atomistic FFs require much more computational power than their coarse-grained counterparts, however, atomistic FFs provide more accurate results. Course-grained FFs should be used in situations where time or computational power are limited. If this is not the case, atomistic FFs are usually the better alternative. The GROMOS-96 FF was used in this study. This FF was developed from the GROMOS-87 FF, through some modifications to its parameters. It requires a minimum cut-off range of 1.4 nm, since it was parameterized using this distance. GROMOS-96 is more recent and has improvements in the overall treatment of bonded interactions and angle geometry (GROMACS, 2021a; Schuler et al., 2001). It is better used in united-atom (UA) setups, like the one used in this study. This setup incorporates atomistic and coarse-grained strategies by analyzing atoms individually, while simultaneously considering some groups as single units (pseudo-atoms). This greatly reduces computation time, while maintaining a fair amount of resolution (Chen et al., 2006). To save even more computation time and allow for larger simulations, electrons are not treated explicitly in classical MD. Instead, the Born-Oppenheimer approximation separates the electronic and nuclear motions, that occur at different time scales. The motion of nuclei is treated classically, while the electronic structure is solved through quantum mechanical approaches (Hug, 2013).

Finally, for FF methods to work, two fundamental assumptions are crucial: additivity and transferability. Additivity is the idea that the (potential) energy of a system can be written as the sum of the potential energy of its atoms or groups. Transferability, states that functions developed on small sets of molecules are applicable to other groups of molecules with similar chemical compositions (Monticelli & Tieleman, 2013).

2. Methodology

The model peptides of the MT-SC family (MT-C34, MT-SC22EK, MT-SC29EK, MT-SC34, MT-SC34EK and MT-SC35EK) were built using the Discovery Studio software, developed by Dassault Systèmes. All of the peptides were built in a π-helical structure, similar to an α-helix but coiled more loosely (4.4 residues per turn, instead of 3.6 like the α-helix) (Riek & Graham, 2011). Initially the peptides were simulated in an aqueous environment consisting of SPC (simple point charge) water molecules (Berendsen et al., 1981) and ions to neutralize the charge of the system. This was crucial to verify that the peptides behaved properly in solution. Firstly, cubic boxes were built around the peptides with its walls at a minimum distance of 1.5 nm from them. This delimited the system and allowed for solvation. After the water molecules were added, some were replaced with positive ions (Na^+) to nullify the net charge of the system. The number of ions differed from peptide to peptide.

The model membranes were designed with the help of tools from the GROMACS MD simulation package. These membranes contained POPC and cholesterol in different concentrations. One of the membranes was entirely made of POPC, while the other two had POPC to Cholesterol proportions of 4:1 and 1:1. First, the membranes were placed within a simulation box, leaving enough space along the Z axis to accommodate the peptides. Next, the peptides were placed within the simulation box, at 4-6 nm from the membranes. This allowed them to detect the membrane lipids and slowly adjust their structure as needed to complete adsorption. After placing the peptide in the box, the system was once again solvated and neutralized with Na^+ ions.

All the simulations were conducted using the same united atom FF, a modified version of the GROMOS87 FF present in the GROMACS MD simulation package (GROMACS, 2021b). Each peptide, membrane and system were simulated three separate times. The first run was a 1000 step steepest-descent energy minimization of the structure, with a step size of 0.001 (1 fs) and no constraints. The second run was a MD simulation that ensured the water molecules adjusted themselves to the protein and lipids. This simulation was performed under the same conditions as the final run, but with way less time. Finally, the production MD simulation was performed. This simulation had 250 million integration steps and a step size of 0.002 (2 fs). Temperature coupling was made using the v-rescale algorithm (a modified version of the common velocity rescaling method), while pressure coupling was made via the Berendsen barostat. Pressure coupling was isotropic for peptides in water and semi-isotropic for the peptide and membrane systems. These maintained the temperature and pressure conditions within the simulation box constant, at 300K and 1 bar, respectively. Neighbor searching was performed using the Verlet cutoff scheme, using a cutoff range of 1 nm. Electrostatic interactions were calculated via the Fast smooth Particle-Mesh Ewald (PME). Constraints were applied to water molecules via the SETTLE algorithm and to all other bonds in the system using the relatively fast LINCS algorithm.

The *umbrella sampling* technique was also performed for every system, to obtain the free energy profile (ΔG) associated with peptide adsorption to the membrane. 40 frames of the peptide were generated via gradual separations from the membrane of 0.05 nm along the z axis. The last image of the peptide was 2 nm away from the membrane. Each frame underwent four different types of simulations. First, two steepest-descent minimizations were performed with no temperature of pressure constraints and a relatively small step size of 10,000 and 50,000, respectively. The second was more complex, since it employed neighbor searching and treatment of long-range interactions through PME. Next, an NPT (constant number of particles, pressure, and temperature) equilibration was performed before the final *umbrella sampling* simulation. Both runs were performed under the same conditions, but the NPT equilibration had a smaller number of steps, of 50,000, compared to the 5 million steps of the umbrella simulation. The same leap-frog algorithm was used for integration, with constraints applied to all bonds via the LINCS algorithm. Long range interactions were calculated via PME. Temperature and pressure values were maintained at 300K and 1 bar via the Nose-Hoover thermostat and Parrinello-Rahman barostat, respectively. The cutoff range used was 1.4 nm. In these simulations, a biasing potential is administrated to keep the different frames of the peptide within a certain region of the phase-space to properly sample all the possible states of the system within those regions. From these simulations we obtain the biased probability distribution of the system, from which we can obtain its free energy profile (also called the Potential of Mean Force). From these curves we calculate the value of the Gibbs free energy associated with the binding process of the peptide to the membrane (ΔG_{bind}). The reaction coordinate chosen was the distance of the peptide to the membrane along the Z axis.

The comparison between the values obtained from membrane simulations and the systems allow us to derive conclusions about the effects the peptide has on membrane organization, how well it attaches to the membrane, how strongly the peptide interacts with the membrane, etc.

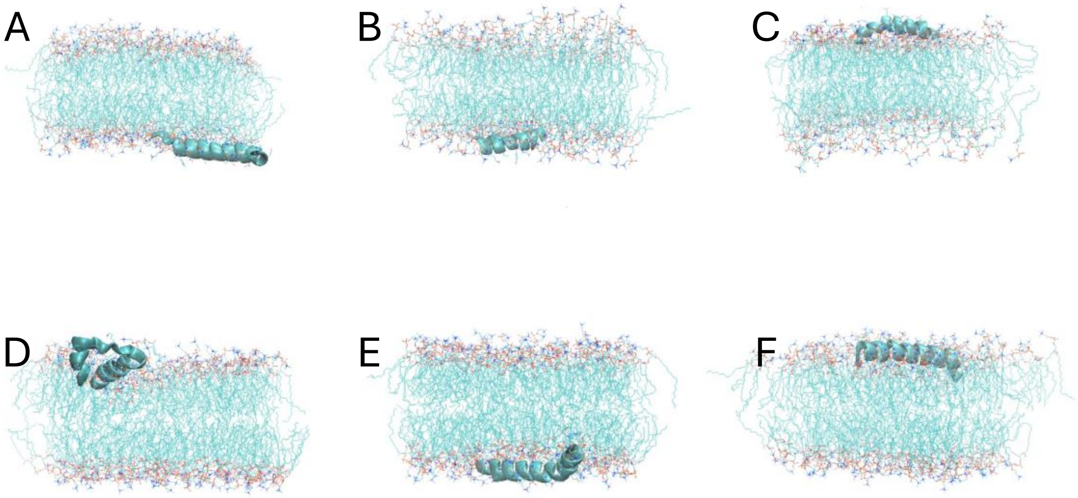


Figure 11 – Final structures of the POPC membrane systems: (A) MT-C34, (B) MT-SC22EK, (C) MT-SC29EK, (D) MT-SC34, (E) MT-SC34EK and (F) MT-SC35EK.

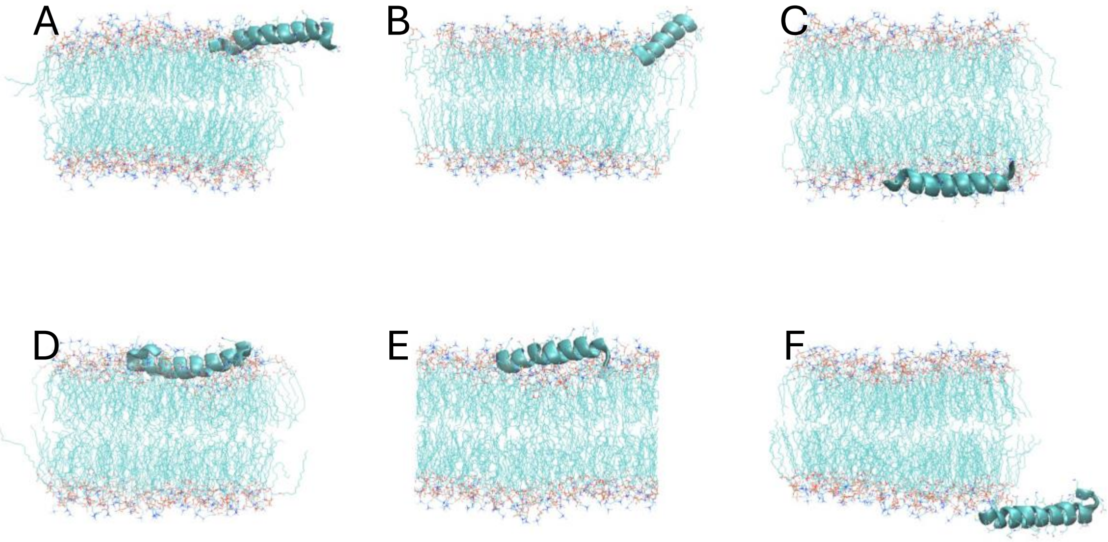


Figure 12 – Final structures of the 20% cholesterol membrane systems: (A) MT-C34, (B) MT-SC22EK, (C) MT-SC29EK, (D) MT-SC34, (E) MT-SC34EK and (F) MT-SC35EK.

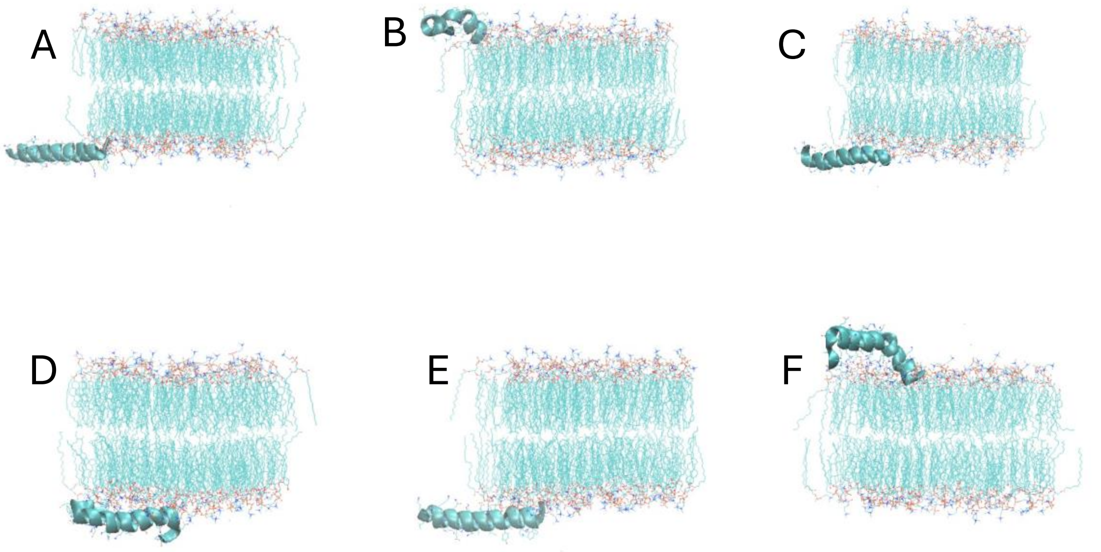


Figure 13 – Final structures of the 50% cholesterol membrane systems: : (A) MT-C34, (B) MT-SC22EK, (C) MT-SC29EK, (D) MT-SC34, (E) MT-SC34EK and (F) MT-SC35EK.

3. Results and Discussion

System Equilibration

3.1.1. Area per lipid and Membrane Thickness

These parameters were calculated to analyze the equilibrium state of the system and to evaluate how the interaction between peptides and membranes affected said state. The lipid area was calculated for POPC (ApPOPC) and cholesterol (ApCHOL), when possible. Calculations were performed for each frame of the trajectory and for the total 500 ns of simulation time. The method used was the same as in (Hofsäß et al., 2003).

$$A_{POPC} = \frac{2A_{box}}{(1-x)N_{lipid}}\left[1 - \frac{x\, N_{lipid}\, V_{chol}}{V_{box} - N_w\, V_w}\right] \tag{14}$$

$$A_{CHOL} = \frac{2\, A_{box}\, V_{chol}}{V_{box} - N_w\, V_w} \tag{15}$$

In the equations shown above, A_{POPC} and A_{CHOL} represent the surface of the POPC and cholesterol molecules, respectively. A_{box} is the area of the xy plane of our three-dimensional rectangular box and V_{box} is its volume. x represents the fraction of cholesterol in the membranes ($x = 0.00$ in the POPC bilayer, 0.20 in the 20% cholesterol bilayer and 0.50 in the 50% cholesterol bilayer). N_{lipid} is the total number of lipids in the system. V_{chol} is the volume of a cholesterol molecule, which was assumed to be 0.593 nm^3. N_w and V_w are the number solvent molecules in the system and their individual volume, respectively. The volume considered for each water molecule was 0.030 nm^3, which is its approximate value under normal pressure and temperature conditions (Hofsäß et al., 2003).

Membrane thickness (MT) calculations were also performed for the total 500 ns of the trajectories. MT is the distance between P8 atoms from opposing bilayers and therefore it is easier to calculate.

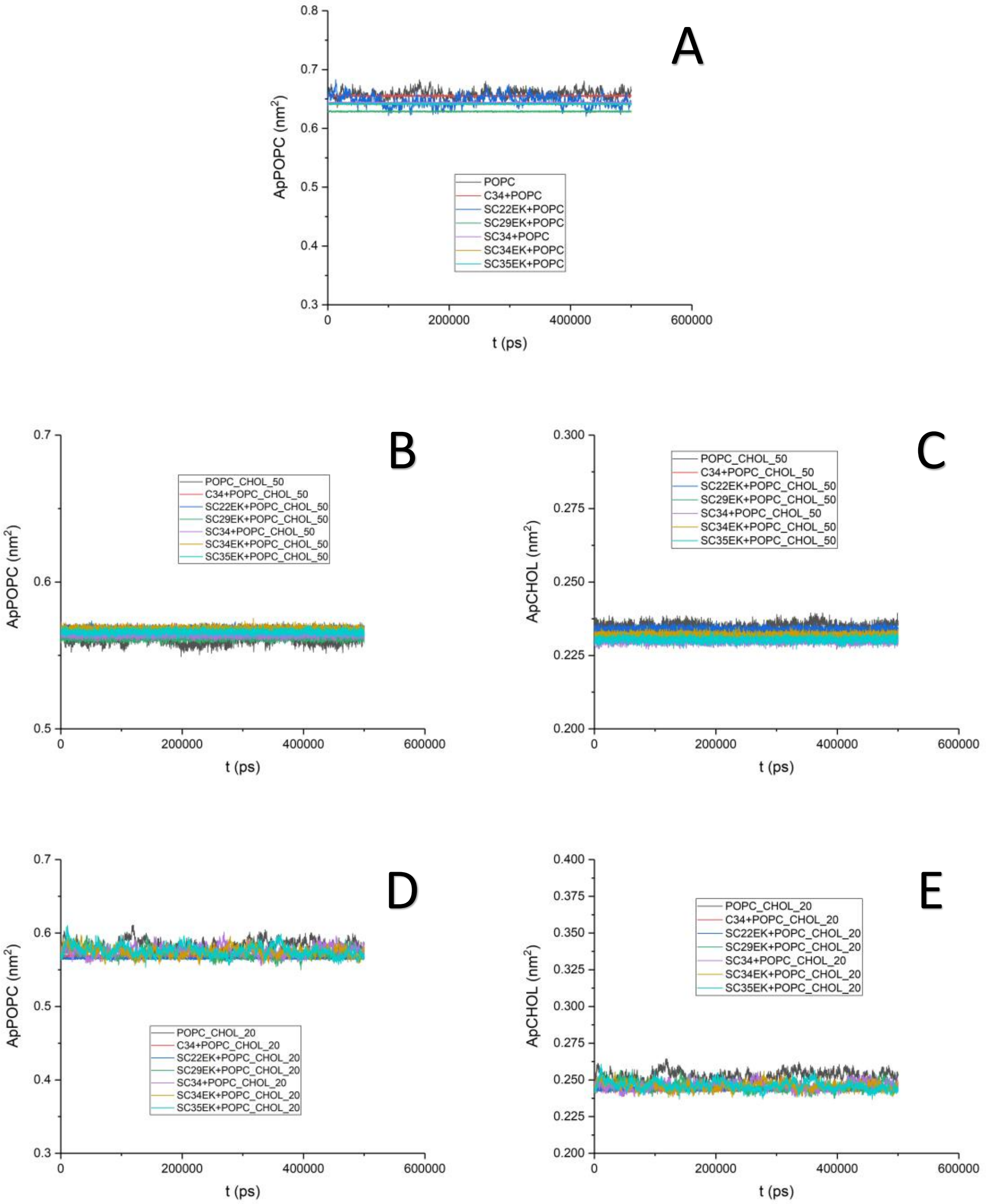


Figure 14 - (A) Area per POPC molecule of the POPC membrane and systems along run time. (B) Area per POPC molecule of the POPC-CHOL-50% membrane and systems along run time. (C) Area per cholesterol molecule of the POPC-CHOL-50% membrane and systems along run time. (D) Area per POPC molecule of the POPC-CHOL-20% membrane and systems along run time. (E) Area per cholesterol molecule of the POPC-CHOL-20% membrane and systems along run time.

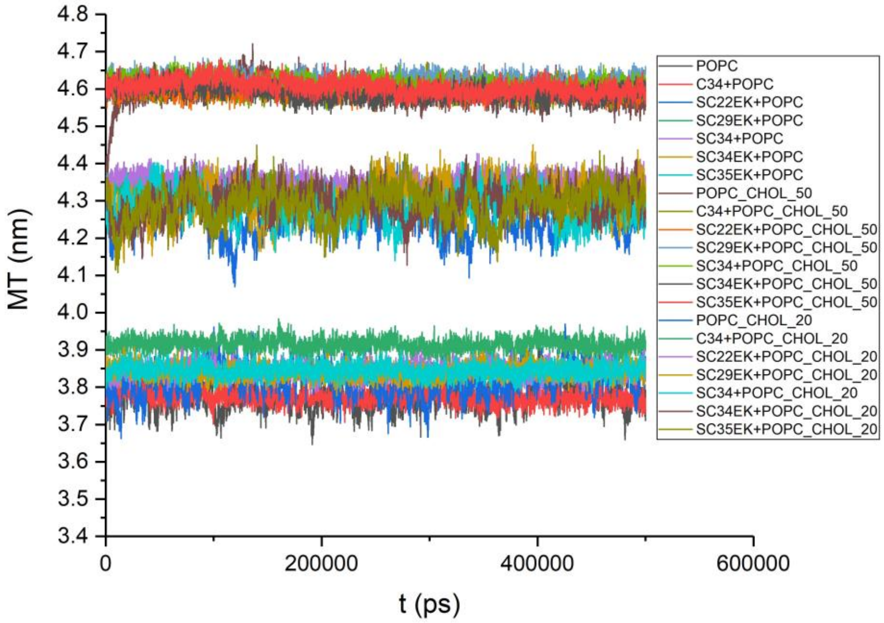


Figure 15 – Membrane thickness (MT) of all POPC and POPC-CHOL systems compared to thickness of standard POPC and POPC-CHOL membranes.

From the analysis of Figure 14, we can see that fluctuations in POPC area are bigger in POPC systems than in POPC-CHOL systems. Even so, for every instance, the area of POPC molecules is bigger in POPC systems than in POPC-CHOL systems. Figure 15 shows that POPC-CHOL membranes are thicker than POPC membranes. Additionally, the more cholesterol a membrane has, the thicker it is. The POPC systems and membrane are located at the bottom, while the 50% cholesterol systems and membrane are at the top. The 20% cholesterol intermediates are in the middle. Cholesterol molecules have an ordering effect on the hydrocarbon chains of POPC that effectively increases their length, increasing membrane thickness. Fluctuations in MT also appear to decrease with higher cholesterol concentration, with the 50% systems showing stabler values. Based on these results, we can assume that cholesterol makes a difference for both parameters by providing additional stability.

Peptide Behavior

3.2.1. Distance between the Membrane and the peptide’s Center of Mass

Initially, the peptides were placed at distances of around 4-6 nm from the membrane in all systems. This simulation was performed for the full trajectory to visualize the approximation of the peptides to the membrane.

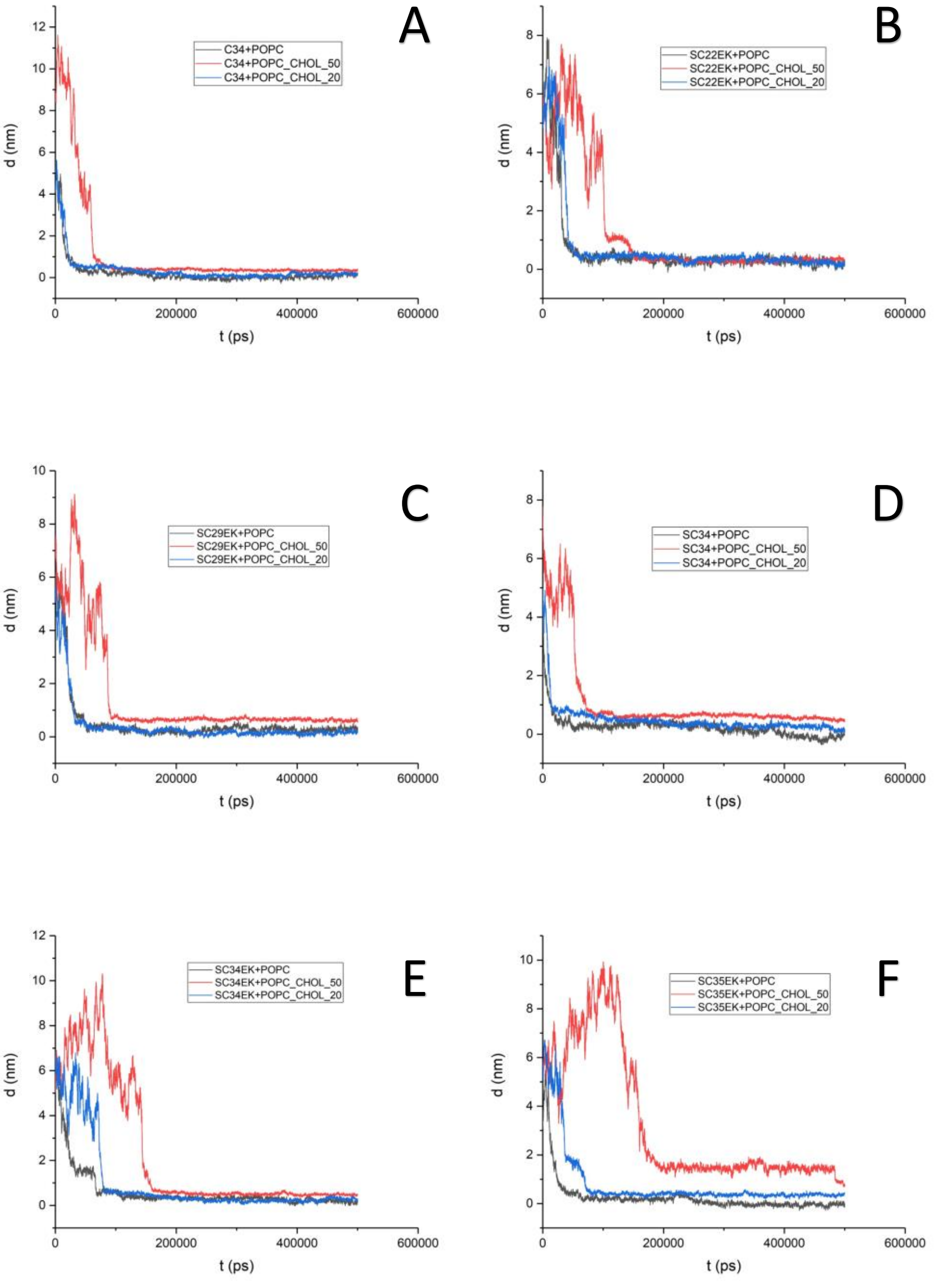


Figure 16 – (A) Distance between MT-C34's center of mass and the POPC and POPC-CHOL membranes. (B) Distance between MT-SC22EK's center of mass and the POPC and POPC-CHOL membranes. (C) Distance between MT-SC29EK's center of mass and the POPC and POPC-CHOL membranes. (D) Distance between MT-SC34's center of mass and the POPC and POPC-CHOL membranes. (E) Distance between MT-SC34EK's center of mass and the POPC and POPC-CHOL membranes. (F) Distance between MT-SC35EK's center of mass and the POPC and POPC-CHOL membranes.

From Figure 16, we see that most of the peptides stabilized their positions in under 100 ns of simulations. The process of adsorption took longer for MT-SC34EK and MT-SC35EK in the POPC-CHOL-50% systems. It is possible that MT-SC35EK could not complete its adsorption process, since its center of mass was still heading towards the membrane at the end of the 500 ns of simulation. There also appears to be a pattern between the cholesterol concentration of a membrane and the relative speed of peptide interaction. Interaction occurs faster in POPC systems and slower is POPC-CHOL systems, with the 50% systems displaying slower interactions. Once again, the 20% systems act as an intermediate. This is probably due to differences in the number of POPC molecules between the systems. In the introduction it was said that these peptides establish significantly more hydrogen bonds with POPC molecules than with cholesterol. The number of POPC molecules is 200, 160 and 112 for the POPC, POPC-CHOL-20% and POPC-CHOL-50% systems, respectively. Therefore, it is to be expected that peptides are faster to detect membranes with more POPC molecules.

In some of the peptide profiles we see that they go away from the membrane initially before closing in on it. This happens because, in the beginning, peptides do not interact with the membrane and instead they diffuse randomly. After the peptides sense the membrane and start interacting with it, the distance between them starts to drop.

3.2.2. Parameter Density

The density of the components of the system was calculated for the last 100 ns of the trajectory to ensure that all peptides were anchored to the membrane. This was done for the POPC and POPC-CHOL membranes separately, and for the peptide membrane systems so we could see the impact that peptide anchoring has on POPC and cholesterol density. The results obtained for each system are shown below alongside normal POPC and POPC-CHOL density of the sole membranes for comparison.

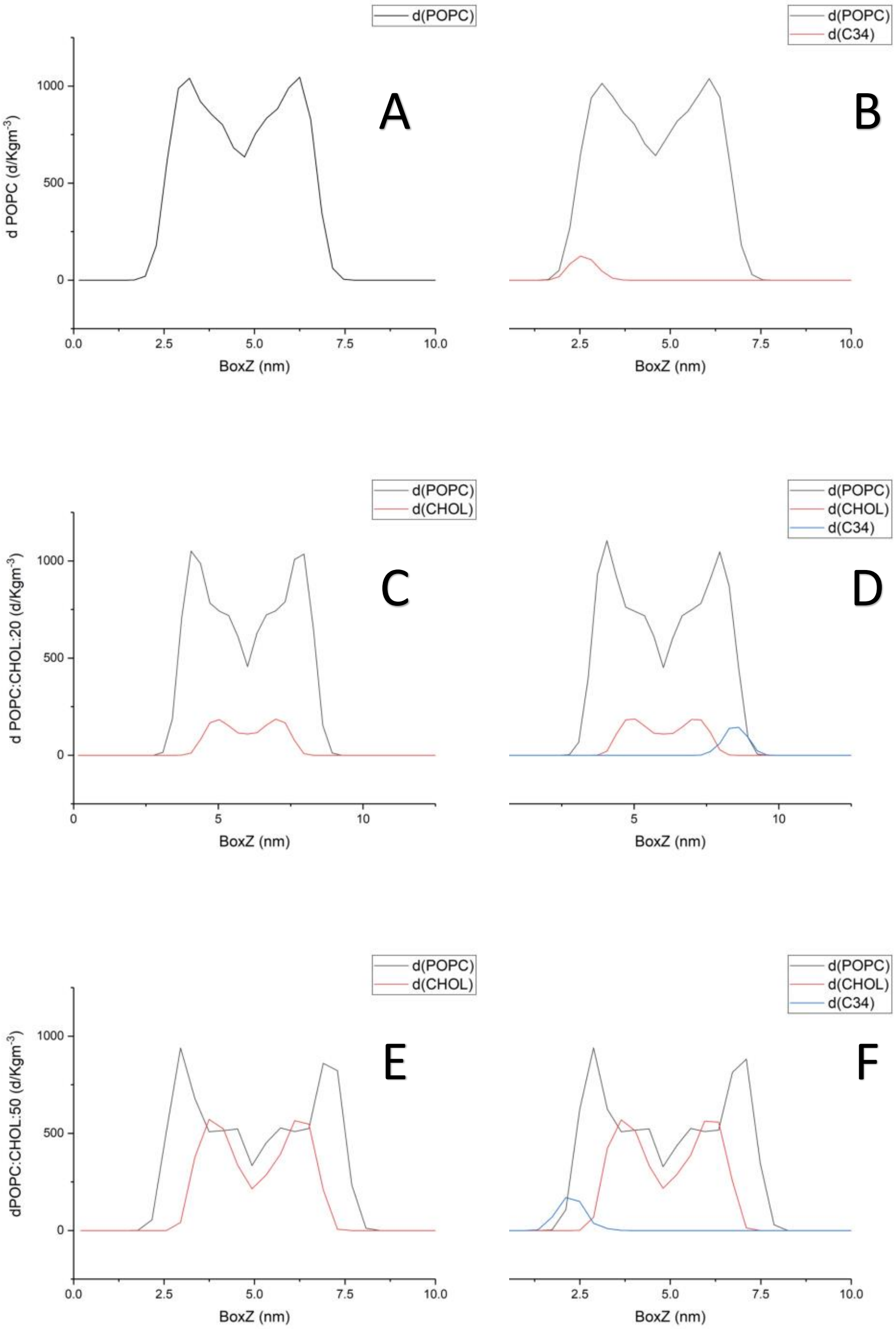


Figure 17 – Parameter density profiles: (A) POPC membrane density; (B) POPC+C34 system density; (C) POPC_CHOL_20% membrane density; (D) POPC_CHOL_20% + C34 system density; (E) POPC_CHOL_50% membrane density; (F) POPC_CHOL_50% + C34 system density

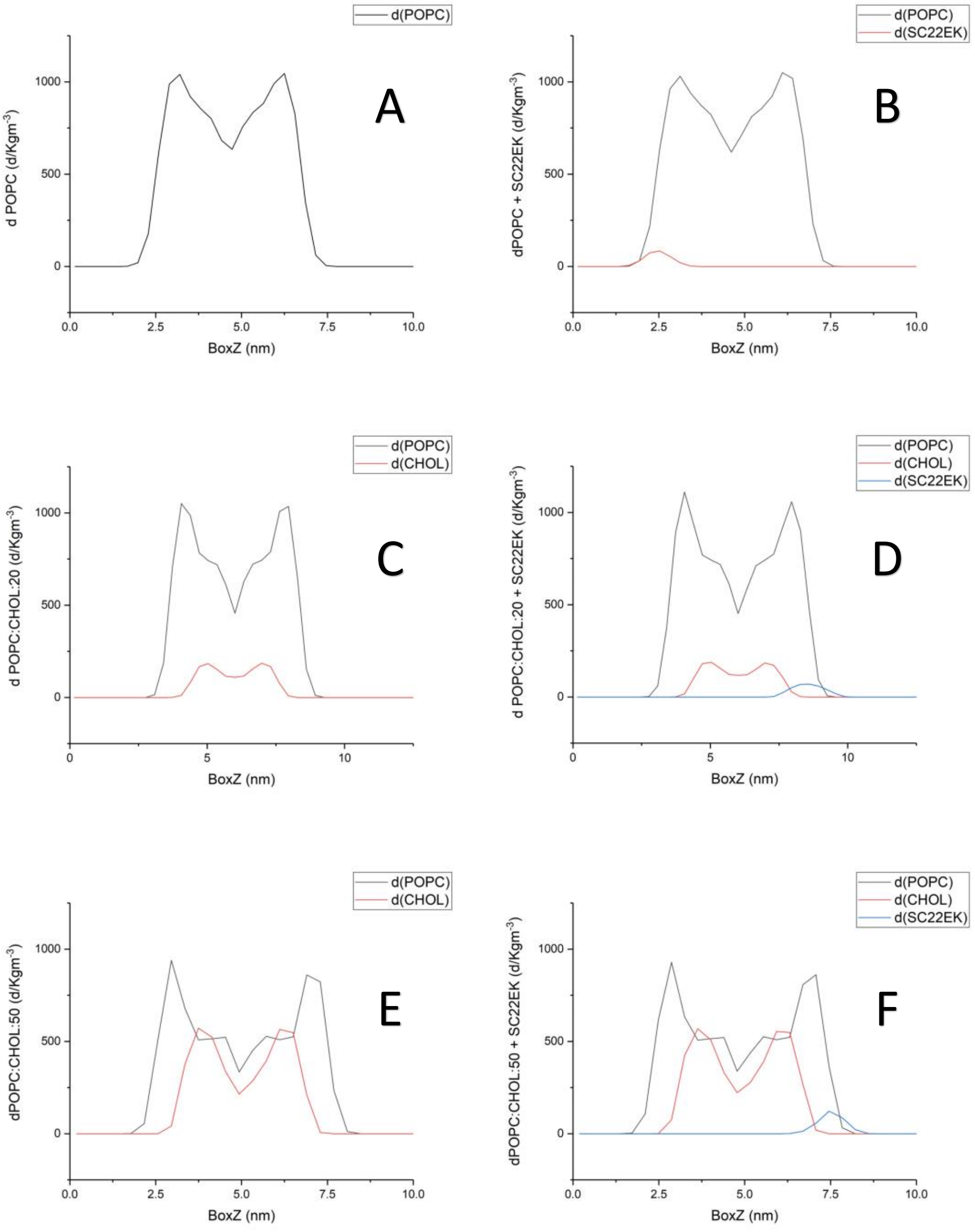


Figure 18 – Parameter density profiles: (A) POPC membrane density; (B) POPC+SC22EK system density; (C) POPC_CHOL_20% membrane density; (D) POPC_CHOL_20%+SC22EK system density; (E) POPC_CHOL_50% membrane density; (F) POPC_CHOL_50%+SC22EK system density

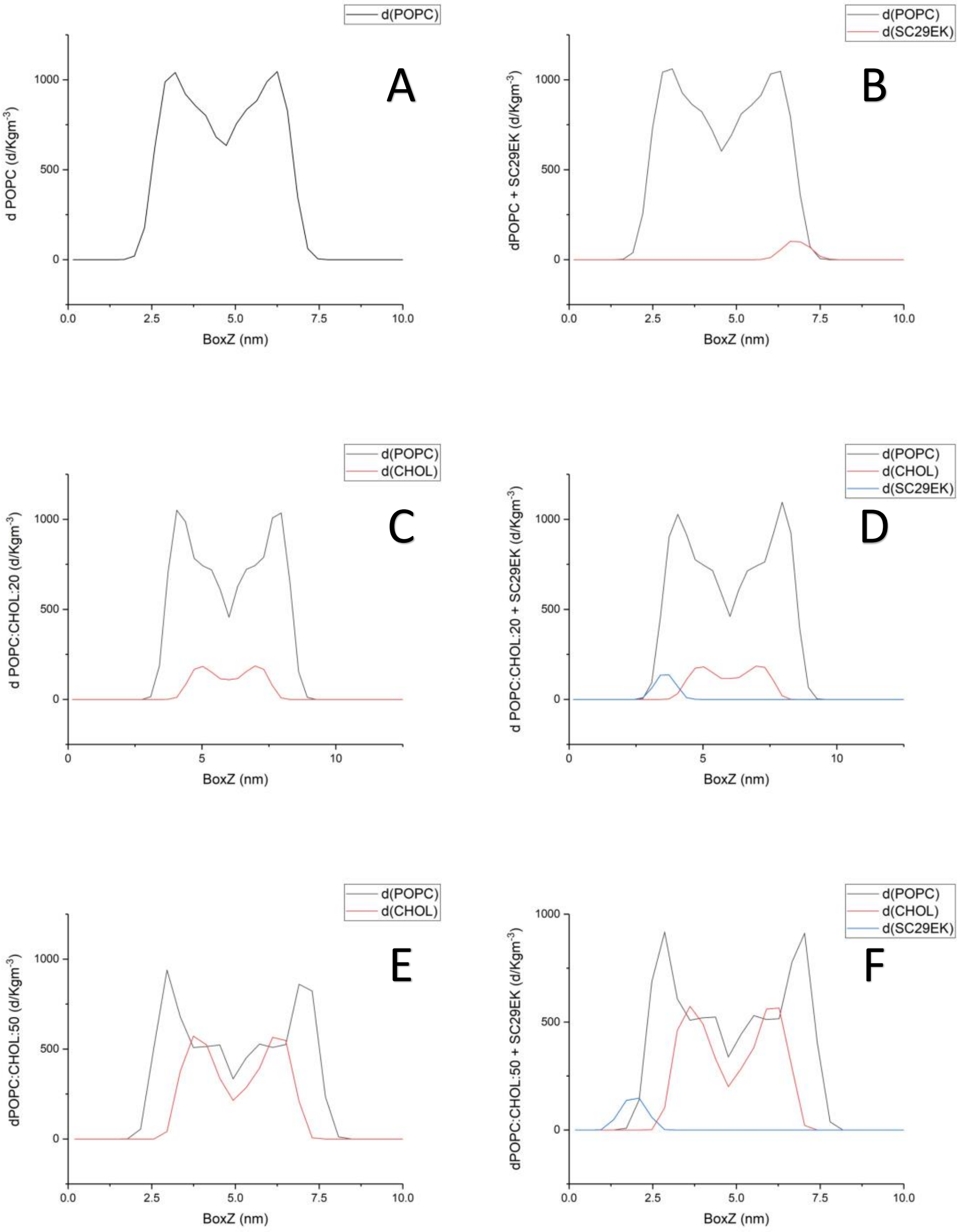


Figure 19 – Parameter density profiles: (A) POPC membrane density; (B) POPC+SC29EK system density; (C) POPC_CHOL_20% membrane density; (D) POPC_CHOL_20%+SC29EK system density; (E) POPC_CHOL_50% membrane density; (F) POPC_CHOL_50%+SC29EK system density

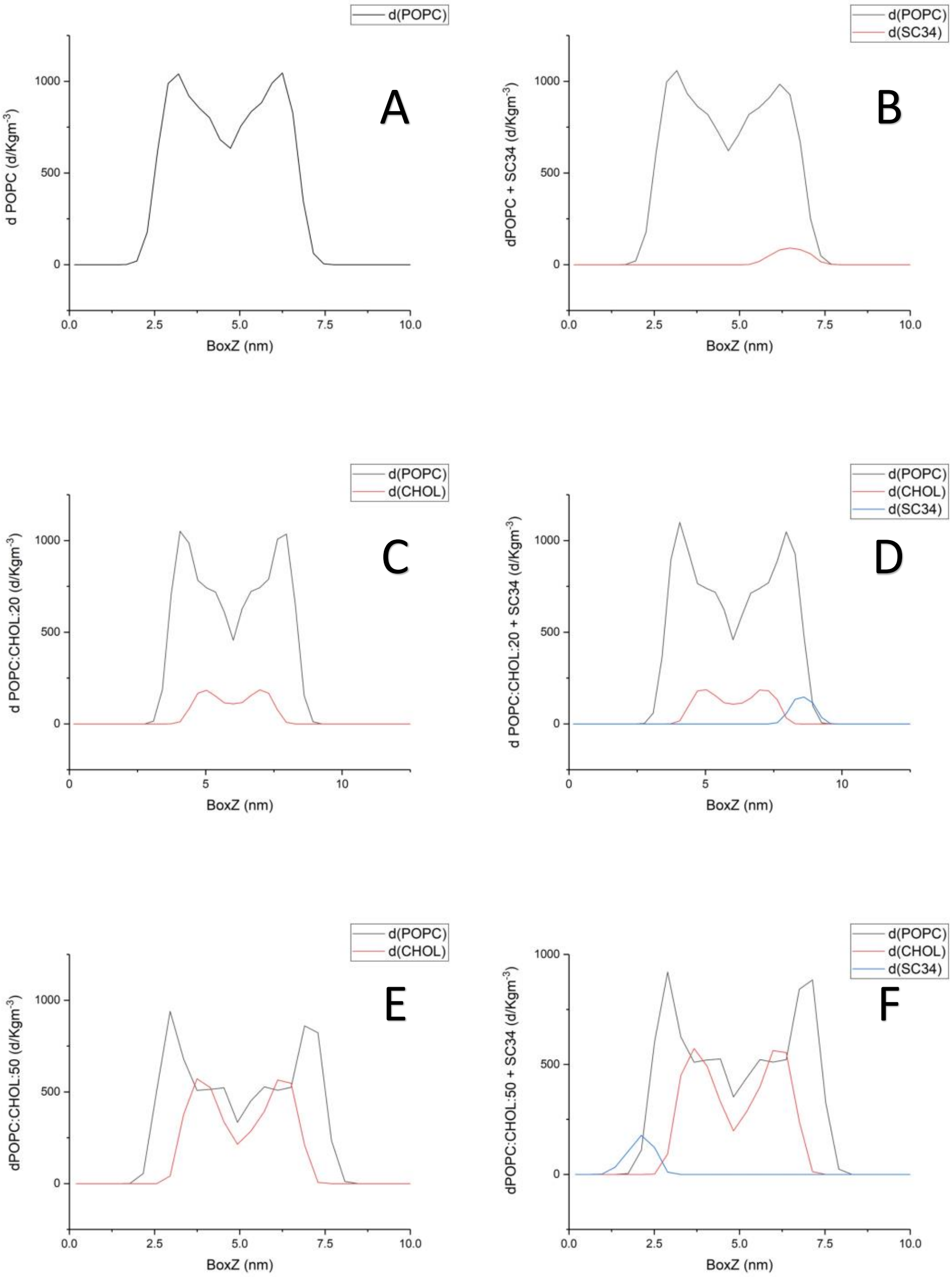


Figure 20 – Parameter density profiles: (A) POPC membrane density; (B) POPC+SC34 system density; (C) POPC_CHOL_20% membrane density; (D) POPC_CHOL_20%+SC34 system density; (E) POPC_CHOL_50% membrane density; (F) POPC_CHOL_50%+SC34 system density

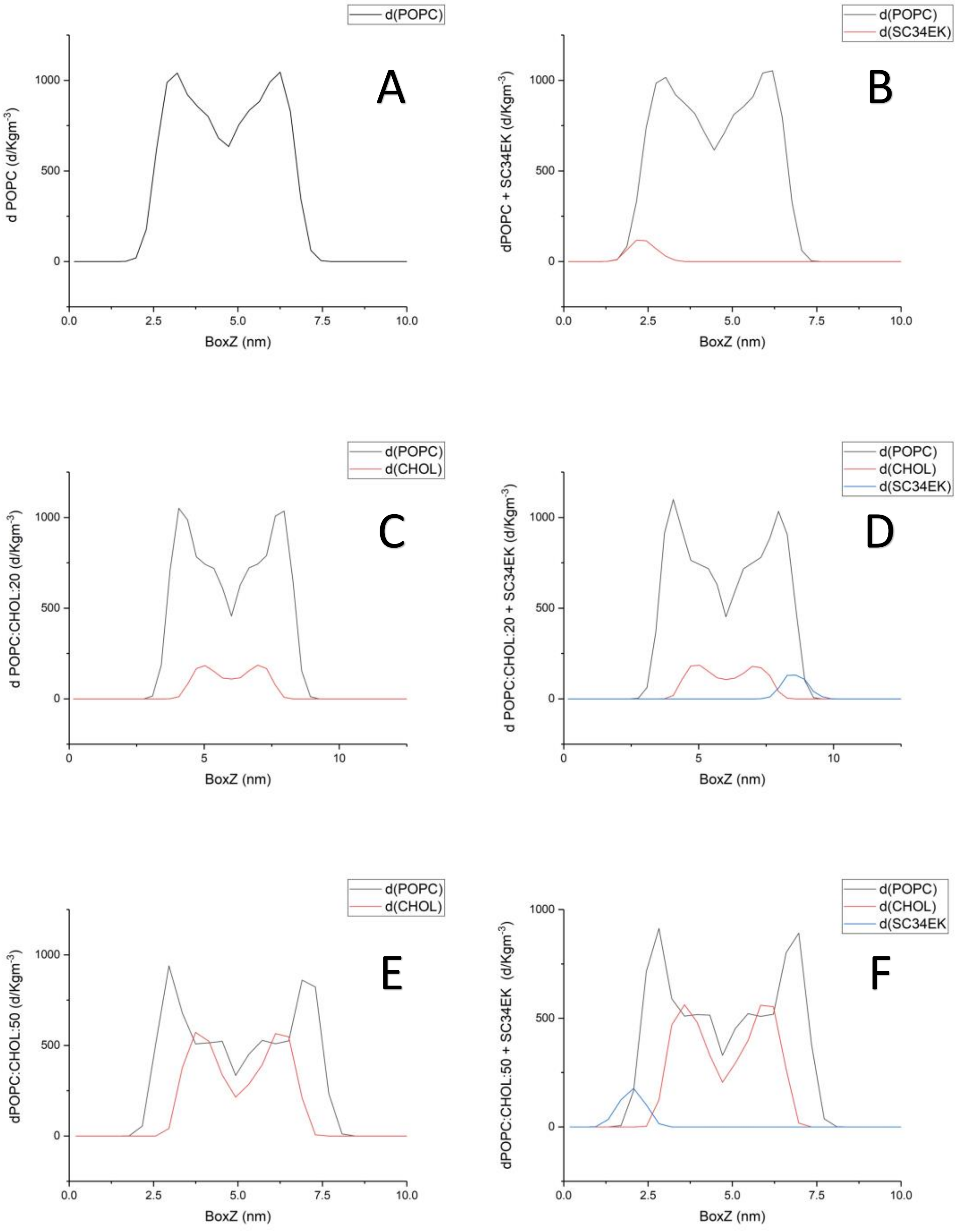


Figure 21 – Parameter density profiles: (A) POPC membrane density; (B) POPC+SC34EK system density; (C) POPC_CHOL_20% membrane density; (D) POPC_CHOL_20%+SC34EK system density; (E) POPC_CHOL_50% membrane density; (F) POPC_CHOL_50%+SC34EK system density

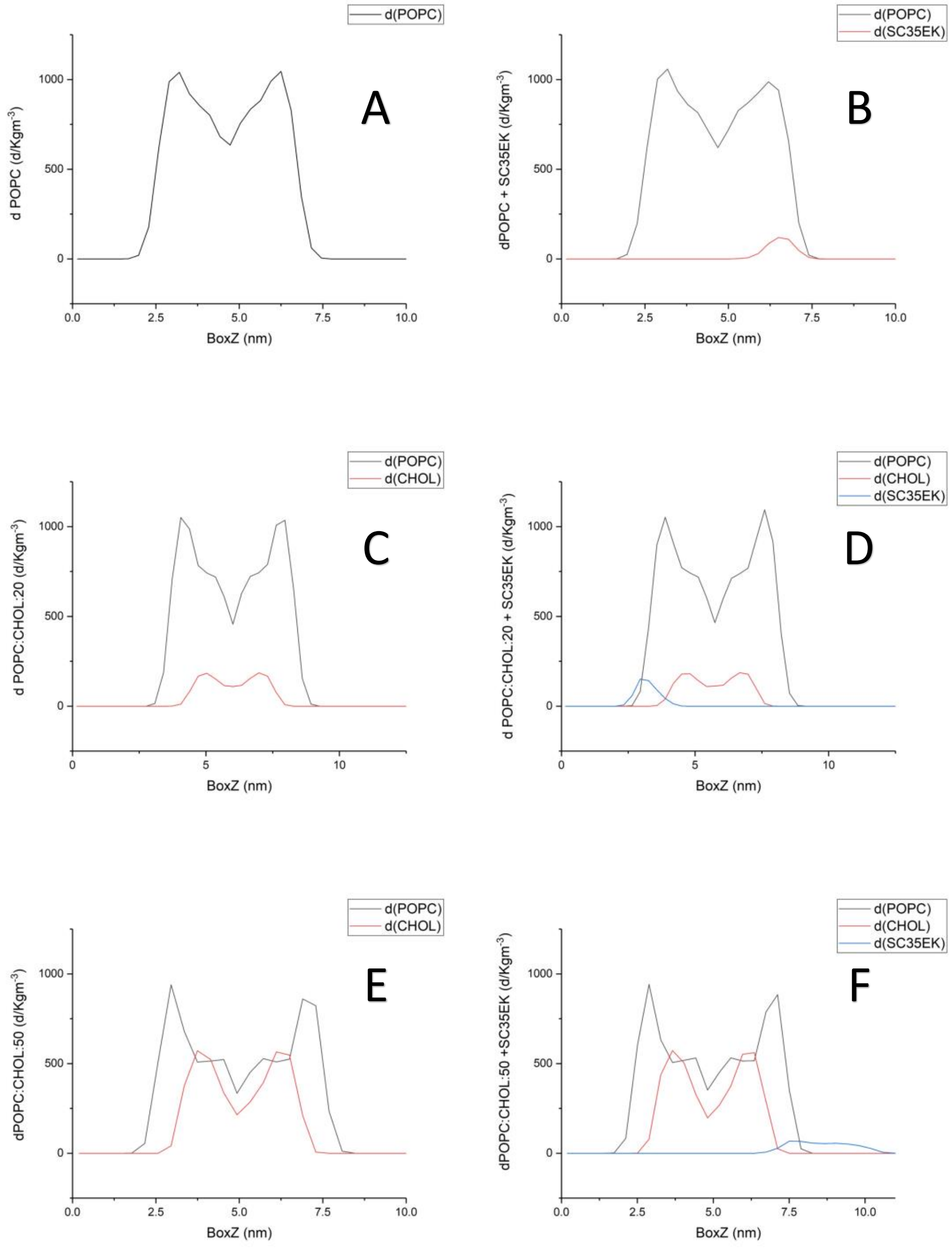


Figure 22 – Parameter density profiles: (A) POPC membrane density; (B) POPC+SC34EK system density; (C) POPC_CHOL_20% membrane density; (D) POPC_CHOL_20%+SC34EK system density; (E) POPC_CHOL_50% membrane density; (F) POPC_CHOL_50%+SC34EK system density

The spikes in POPC density correspond to the surface of each monolayer. By analyzing the density profiles of membranes and systems (Figure 17 – Figure 22), we can see a trend in the results that is consistent for every system. Wherever the peptide is anchored, there is a visible reduction in the density of POPC for that monolayer. This change is caused by the interaction of the peptide with the membrane. Even though the density of POPC is altered, cholesterol density remains unaffected.

3.2.3. Peptide Secondary Structure

The analysis of the secondary structure of the peptides was made via DSSP (Kabsch & Sander, 1983), an algorithm incorporated in the GROMACS MD simulation package. This algorithm detects specific patterns of hydrogen bonds, which are attributed to a specific type of secondary structure. The percentages associated with each structure were calculated for the last 100 ns of the trajectories.

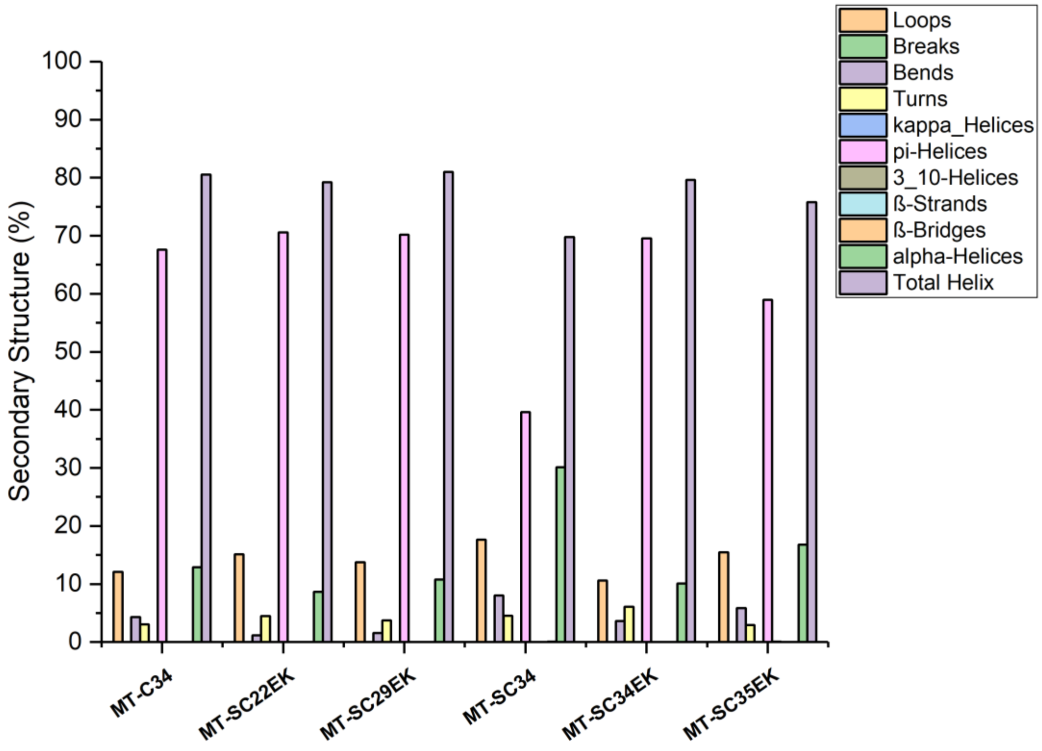


Figure 23 – Percentages of secondary structure for the peptides in water, as well as the total percentage of helical structures.

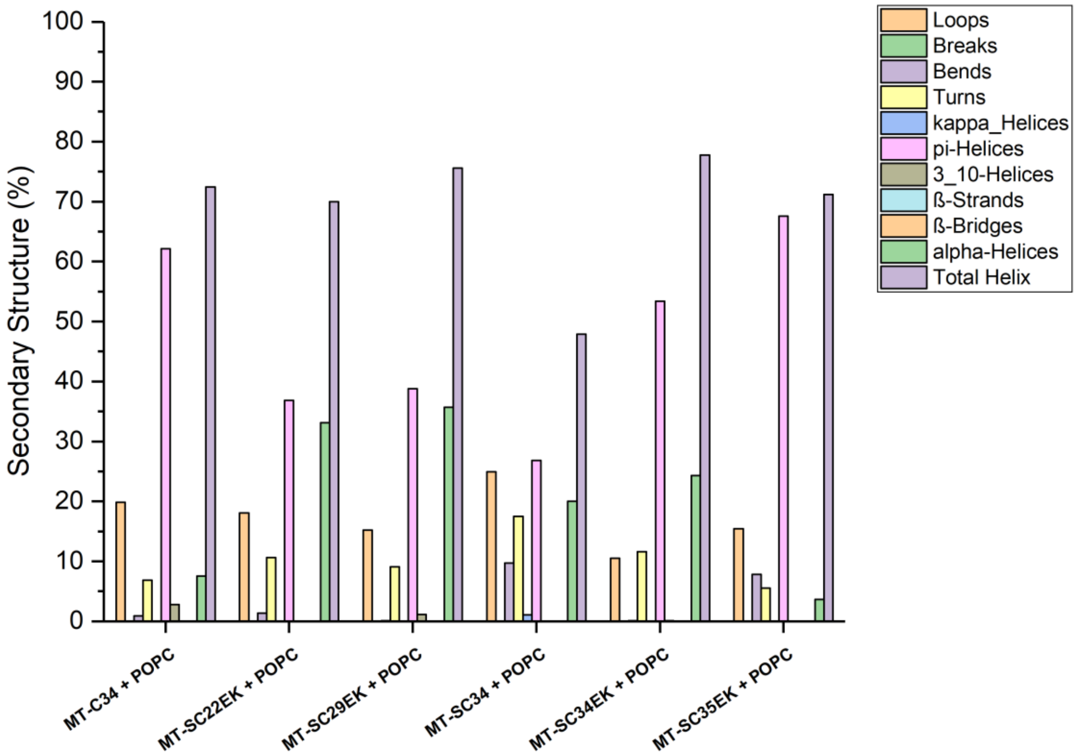


Figure 24 – Percentages of secondary structure for peptides in POPC systems, as well as the total percentage of helical structures.

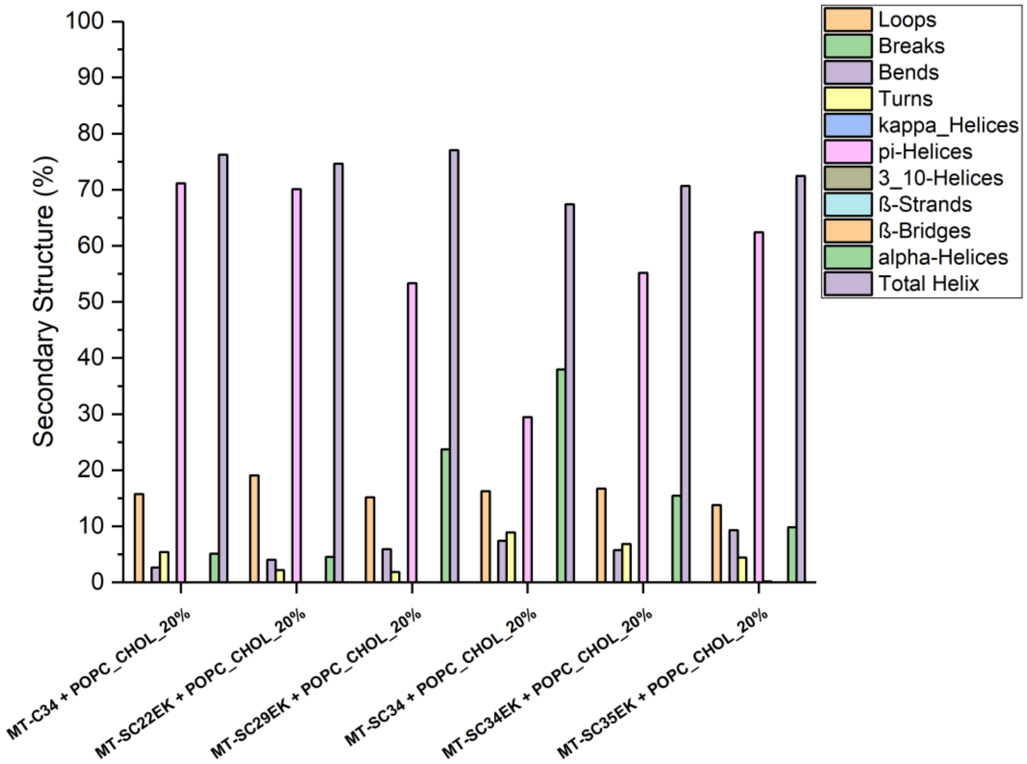


Figure 25 – Percentages of secondary structure for peptides in the 20% cholesterol systems, as well as the total percentage of helical structures.

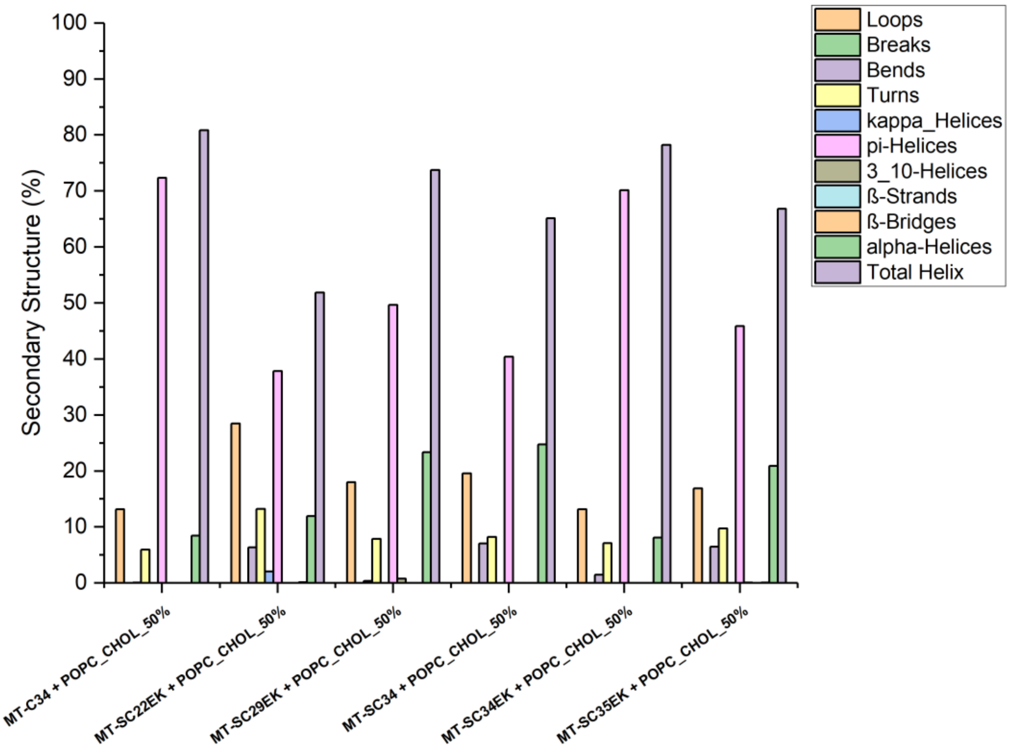


Figure 26 – Percentages of secondary structure for peptides in the 50% cholesterol systems, as well as the total percentage of helical structures

Figures 23-26, show percentages of various possible types of secondary conformations that the peptides can be in. These include various types of helical conformations (α-helix, π-helix, 3_{10} helix and K-helix), β-sheets and strands, breaks, bends, loops and more. A loop in considered by DSSP to be a part of the structure with no special secondary structure designation. As mentioned previously, it is important that peptides can maintain highly helical conformations to facilitate the interaction with the NHR region of gp41, to disrupt the formation of the 6-HB.

The peptides are mostly in their original π-helical conformations in every system. All the peptides in water have a total helix content of over 70%. In this environment, the most helical peptides of the six are MT-C34, MT-SC22EK, MT-SC29EK and MT-SC34EK. In the POPC systems, the total helix content dips slightly in every system, except in MT-SC34 + POPC, where it drops to less than 50%. All other peptides, especially MT-SC34EK and MT-SC29EK, manage to maintain a high helicity. There is an increase in disorder, which can be seen through the increase in loops and turns. Some peptides also adopt more α-helical conformations. In the 20% cholesterol systems, the secondary structure values are closer to the peptides in water. All systems have total helix values over 70%, but there is a slightly higher loop, bend, and α-helix content. Finally, in the 50% cholesterol systems the total helicity of MT-SC22EK drops to slightly over 50%, while MT-SC34 and MT-SC35EK fall to under 70%. MT-C34 is the most helical peptide, with a total helicity of over 80%, followed by MT-SC34EK with just under 80% total helicity.

3.2.4. Positioning of the alpha-Carbons (Cα)

The positioning of the Cα was calculated to further characterize the configuration of the peptides adsorbed to the membranes. The average position of these atoms was calculated for the last 100 ns of the trajectory.

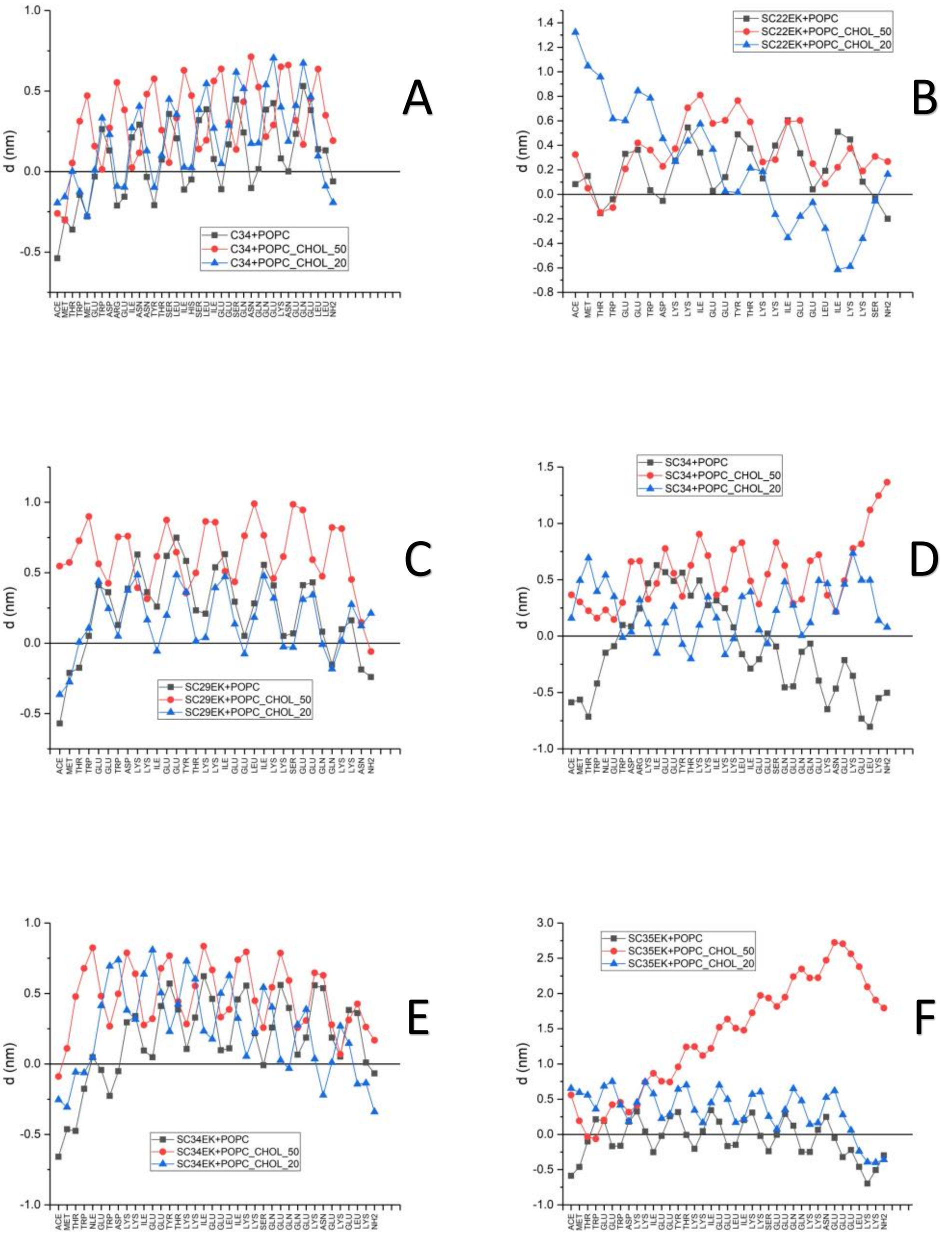


Figure 27 – (A) Average Cα position of the MT-C34 peptide along the 100 ns of simulation. (B) Average Cα position of the MT-SC22EK peptide along the 100 ns of simulation. (C) Average Cα position of the MT-SC29EK peptide along the 100 ns of simulation. (D) Average Cα position of the MT-SC34 peptide along the 100 ns of simulation. (E) Average Cα position of the MT-SC34EK peptide along the 100 ns of simulation. (F) Average Cα position of the MT-SC35EK peptide along the 100 ns of simulation.

Throughout every one of the systems, we can see that the variations in the positions of the Cα are somewhat consistent. This is to be expected due to their helical conformations, meaning that every 3 to 4 amino acid residues they will switch from approximating to the membrane to moving away from it and vice-versa. Figure 27 shows that, in general, position variations in POPC systems are more abrupt than in POPC-CHOL systems. Also, the Cα of peptides in the POPC systems are located below the phosphorus line of the membrane (y=0) more often than their POPC-CHOL counterparts. The degree to which they cross this line is inversely proportional to how much cholesterol the membrane has. In the 20% cholesterol systems we see that the peptides are still quite close to the membrane, while in the 50% systems they are further away from it. These results can be attributed to differences in membrane fluidity. The higher fluidity and free volume of the POPC membrane makes it easier to accommodate the peptide more deeply. On the other hand, the higher rigidity and decreased free volume of the POPC-CHOL membranes makes peptide accommodation more difficult. Overall, the results suggest that there was insertion of the peptides in the membrane and that this insertion was more superficial in membranes with higher cholesterol concentrations, being virtually inexistent in 50% cholesterol membranes.

3.2.5. Hydrogen Bonds between Peptide and Membrane

The number of hydrogen bonds that the peptide establishes with the other components of the system can tell us a lot about how the system evolved throughout the simulation. Hydrogen bonds were registered whenever a donor/acceptor/hydrogen triad was formed, that is, when the components were at a distance below 0.35 nm of each other and formed an angle of about 30°. In POPC systems, the interactions that were analyzed were between the peptide and POPC (PEP-POPC), and between peptide and solvent (PEP-SOL). The same was done for POPC-CHOL systems, but with an added parameter of interaction between peptide and cholesterol (PEP-CHOL).

Initially, the peptide is surrounded by water and far from the membrane. As the simulation progresses the peptide will slowly approach the membrane and adsorb to it. Therefore, assuming the system progresses normally, it is expected that PEP-SOL interactions will be higher in the beginning and slowly start to decline as the simulation progresses. The opposite should happen for PEP-POPC interactions and PEP-CHOL interactions to some degree.

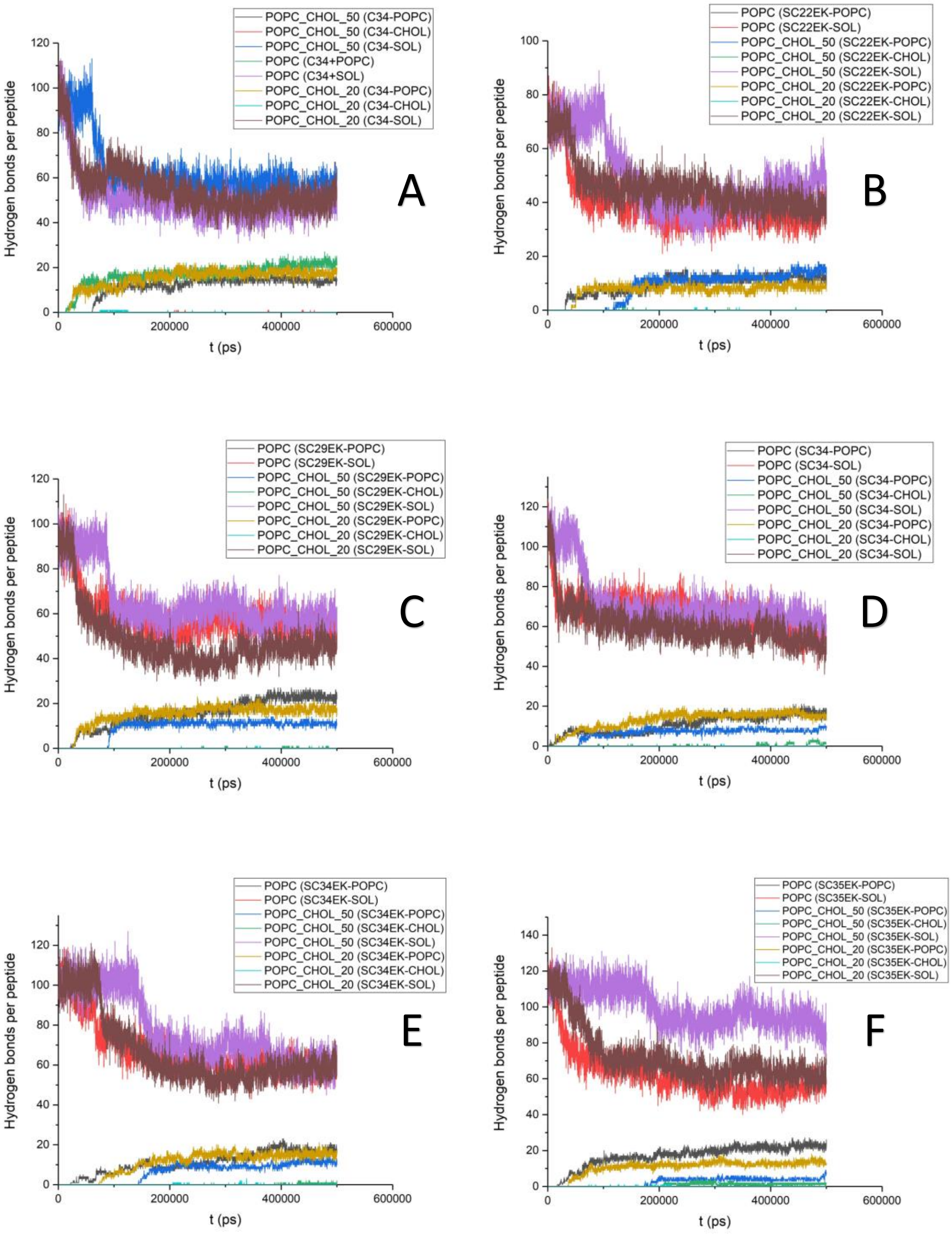


Figure 28 – (A) Number of hydrogen bonds between MT-C34 and the components of the POPC and POPC-CHOL systems. (B) Number of hydrogen bonds between MT-SC22EK and the components of the POPC and POPC-CHOL systems. (C) Number of hydrogen bonds between MT-SC29EK and the components of the POPC and POPC-CHOL systems. (D) Number of hydrogen bonds between MT-SC34 and the components of the POPC and POPC-CHOL systems. (E) Number of hydrogen bonds between MT-SC34EK and the components of the POPC and POPC-CHOL systems. (F) Number of hydrogen bonds between MT-SC35EK and the components of the POPC and POPC-CHOL systems.

Figure 28 confirms what was stated above. From around 100 ns to 200 ns of simulation, the number of interactions between the peptide and the solvent molecules started to decrease and stabilize. This decrease overlaps with an increase of interactions between the peptide, POPC and cholesterol molecules. Even so, interactions between the peptides and cholesterol are almost null.

Table 1 – Average number of hydrogen bonds between PEP-SOL, PEP-POPC and PEP-CHOL, with their respective standard deviations for the last 100 ns of the trajectory.

		Number of hydrogen bonds per peptide		
		PEP-SOL	**PEP-POPC**	**PEP-CHOL**
WATER	MT-C34	94.830 ± 5.711	-	-
	MT-SC22EK	69.975 ± 4.881	-	-
	MT-SC29EK	90.921 ± 5.620	-	-
	MT-SC34	106.327 ± 6.386	-	-
	MT-SC34EK	102.692 ± 5.865	-	-
	MT-SC35EK	112.449 ± 5.977	-	-
POPC	MT-C34	46.860 ± 3.929	21.402 ± 1.535	-
	MT-SC22EK	34.596 ± 3.748	12.247 ± 1.232	-
	MT-SC29EK	54.066 ± 4.106	22.944 ± 1.362	-
	MT-SC34	55.994 ± 5.240	16.336 ± 1.581	-
	MT-SC34EK	59.256 ± 4.430	17.294 ± 1.718	-
	MT-SC35EK	55.478 ± 4.582	22.255 ± 1.610	-
POPC-CHOL_20%	MT-C34	50.258 ± 4.373	17.519 ± 1.396	0
	MT-SC22EK	38.883 ± 3.748	8.877 ± 1.081	0
	MT-SC29EK	45.510 ± 3.952	17.090 ± 1.451	0.001 ± 0.0312
	MT-SC34	52.963 ± 4.972	16.094 ± 1.393	0
	MT-SC34EK	58.439 ± 4.589	15.379 ± 1.503	0.013 ± 0.113
	MT-SC35EK	62.597 ± 4.621	13.242 ± 1.298	0.001 ± 0.032
POPC_CHOL_50%	MT-C34	55.360 ± 4.504	14.346 ± 1.149	0.011 ± 0.104
	MT-SC22EK	46.409 ± 4.675	13.764 ± 1.680	0.006 ± 0.077
	MT-SC29EK	58.129 ± 4.638	10.966 ± 0.902	0.076 ± 0.265
	MT-SC34	64.202 ± 5.262	8.095 ± 0.840	0.788 ± 0.996
	MT-SC34EK	60.222 ± 5.351	11.177 ± 1.244	0.320 ± 0.510
	MT-SC35EK	91.699 ± 6.118	4.184 ± 0.922	1.435 ± 0.601

We observe from Table 1 that peptide interactions with the solvent decrease in the systems, compared to the peptides in water. During peptide adsorption, the peptide ceases interactions with the solvent to adhere to the components of the membrane. There are more PEP-POPC interactions in POPC systems than in POPC-CHOL systems. The number of these interactions is inversely proportional to the amount of cholesterol in the membrane. An increase in cholesterol content is followed by a decrease in the number of interactions between the peptide and POPC (for every system except the MT-SC22EK systems). The higher fluidity of the POPC membrane allows for a closer connection between POPC and peptide amino acids. MT-SC22EK values can be attributed to its negligible interaction with cholesterol. This possibly means the amino acid residues inserted in the membrane interacted with POPC instead of cholesterol. The lower PEP-POPC interactions in POPC-CHOL systems is likely caused by the greater distances between the peptides and the membrane. Rigid membranes undergo less conformation changes than fluid membranes, resulting in more distant and, therefore, weaker connections with the peptides. This is likely what happened to MT-SC34 and MT-SC35EK in the 50% cholesterol systems. The peptide C_α profiles show that MT-SC34 is located completely above the phosphorus axis of POPC in the 50% systems, but its residues are relatively close to it, suggesting that they are close enough to interact with POPC. On the other hand, MT-SC35EK has two residues connected to the POPC axis, while the most of its residues are too distant to interact with the POPC molecules. Therefore, MT-SC35EK + POPC_CHOL_50% has the least PEP-POPC interaction.

3.2.6. Interaction Energy between Peptides and Bilayers

Interaction energy between components of the system is also a great parameter to predict the evolution of a system over time. Variations in Lennard-Jones (LJ) potential and in Coulomb (Coul) forces were calculated for interactions between specific components. The interactions that were studied were: peptide and POPC (PEP-POPC); peptide and cholesterol (PEP-CHOL) and peptide and solvent (PEP-SOL).

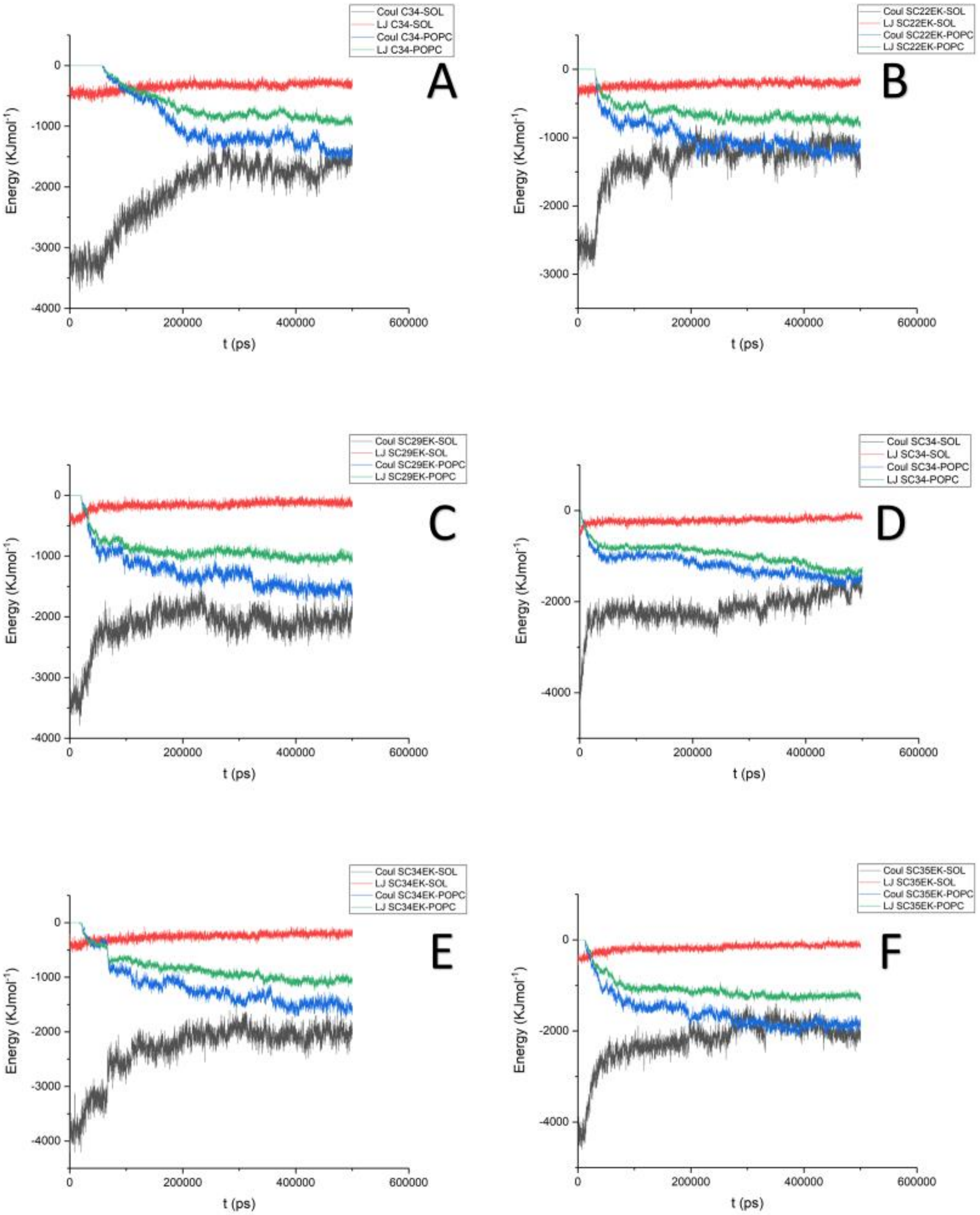


Figure 29 - (A) Variation of the LJ and Coulomb potentials for the POPC + MT-C34 system. (B) Variation of the LJ and Coulomb potentials for the POPC + MT-SC22EK system. (C) Variation of the LJ and Coulomb potentials for the POPC + MT-SC29EK system. (D) Variation of the LJ and Coulomb potentials for the POPC + MT-SC34 system. (E) Variation of the LJ and Coulomb potentials for the POPC + MT-SC34EK system. (F) Variation of the LJ and Coulomb potentials for the POPC + MT-SC35EK system.

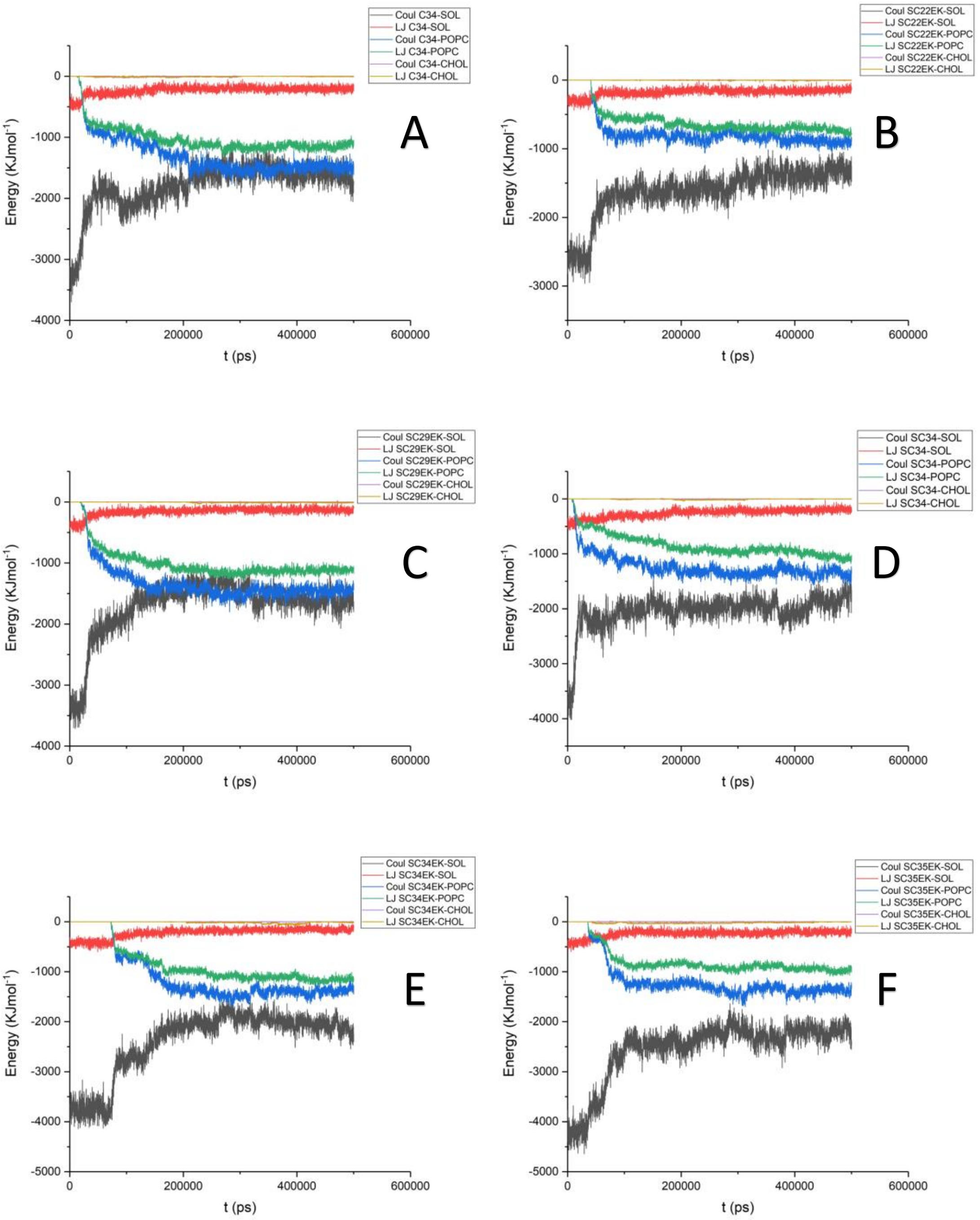


Figure 30 - (A) Variation of the LJ and Coulomb potentials for the POPC_CHOL_20% + MT-C34 system. (B) Variation of the LJ and Coulomb potentials for the POPC_CHOL_20% + MT-SC22EK system. (C) Variation of the LJ and Coulomb potentials for the POPC_CHOL_20% + MT-SC29EK system. (D) Variation of the LJ and Coulomb potentials for the POPC_CHOL_20% + MT-SC34 system. (E) Variation of the LJ and Coulomb potentials for the POPC_CHOL_20% + MT-SC34EK system. (F) Variation of the LJ and Coulomb potentials for the POPC_CHOL_20% + MT-SC35EK system.

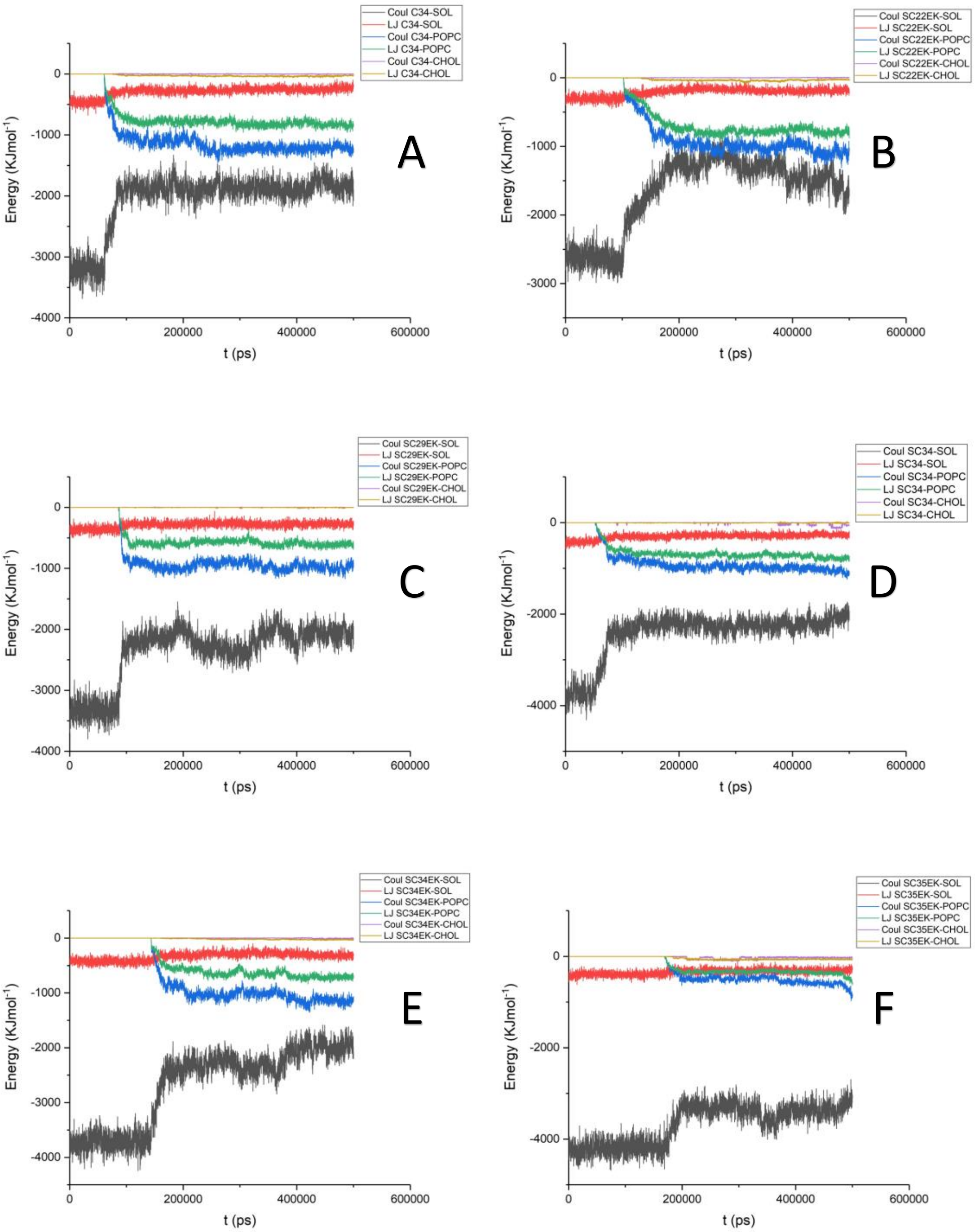


Figure 31 - (A) Variation of the LJ and Coulomb potentials for the POPC_CHOL_50% + MT-C34 system. (B) Variation of the LJ and Coulomb potentials for the POPC_CHOL_50% + MT-SC22EK system. (C) Variation of the LJ and Coulomb potentials for the POPC_CHOL_50% + MT-SC29EK system. (D) Variation of the LJ and Coulomb potentials for the POPC_CHOL_50% + MT-SC34 system. (E) Variation of the LJ and Coulomb potentials for the POPC-CHOL_50% + MT-SC34EK system. (F) Variation of the LJ and Coulomb potentials for the POPC_CHOL_50% + MT-SC35EK system.

The results from the POPC and POPC-CHOL systems are highly consistent and show similarities with one another. From around 10 to 30 ns the interaction energy between the peptide and the solvent starts to decrease in POPC and POPC_CHOL_20% systems. This is accompanied by an increase of interaction energy between the peptide and POPC. The same thing happens in the POPC_CHOL_50% systems, but slightly later, from around 50 to 180 ns. This happens because, initially, the peptide is surrounded by solvent molecules and too far from the membrane to interact with the POPC molecules. As the simulation progresses and the peptide gets closer to the membrane, some of its residues will begin interacting with POPC and cease interacting with the solvent. Also, the interaction energy of the peptides with cholesterol is very low, remaining close to zero throughout the trajectory.

Table 2 - Average and standard deviation of the Lennard-Jones and Coulomb potentials in POPC and POPC-CHOL systems for the different peptide interactions along the runtime.

Lennard-Jones Potential ($KJmol^{-1}$)					Coulomb Potential ($KJmol^{-1}$)		
		PEP-SOL	PEP-POPC	PEP-CHOL	PEP-SOL	PEP-POPC	PEP-CHOL
POPC	**MT-C34**	-302.32 ± 63.52	-918.24 ± 296.92	-	-1723.49 ± 452.63	-1283.54 ± 405.99	-
	MT-SC22EK	-222.36 ± 47.39	-611.41 ± 192.99	-	-1351.49 ± 387.43	-947.94 ± 309.49	-
	MT-SC29EK	-160.88 ± 72.90	-877.14 ± 240.61	-	-2153.81 ± 351.24	-1223.22 ± 377.61	-
	MT-SC34	-215.75 ± 62.15	-967.48 ± 239.51	-	-2149.93 ± 319.57	-1196.21 ± 266.88	-
	MT-SC34EK	-256.10 ± 71.21	-827.13 ± 271.87	-	-2315.36 ± 500.71	-1146.84 ± 420.39	-
	MT-SC35EK	-174.87 ± 81.61	-1069.63 ± 278.24	-	-2223.54 ± 506.66	-1565.70 ± 433.89	-
POPC_CHOL_20%	**MT-C34**	-230.92 ± 71.60	-1012.65 ± 255.70	-7.59 ± 5.51	-1790.41 ± 382.85	-1279.23 ± 358.51	-1.20 ± 4.42
	MT-SC22EK	-177.50 ± 56.89	-594.56 ± 202.26	-2.31 ± 2.52	-1618.04 ± 354.90	-752.44 ± 247.64	0.04 ± 1.33
	MT-SC29EK	-158.51 ± 71.34	-988.98 ± 278.86	-6.66 ± 5.85	-1717.17 ± 474.49	-1295.17 ± 367.24	-0.18 ± 1.19
	MT-SC34	-260.50 ± 78.82	-836.51 ± 219.19	-6.51 ± 6.91	-2046.84 ± 310.24	-1221.13 ± 255.01	-0.30 ± 1.51
	MT-SC34EK	-223.73 ± 101.21	-856.01 ± 398.61	-12.49 ± 15.40	-2390.99 ± 645.77	-1085.13 ± 516.86	-0.50 ± 2.64
	MT-SC35EK	-240.19 ± 75.34	-790.56 ± 275.87	-19.46 ± 12.41	-2549.54 ± 604.15	-1152.36 ± 411.98	0.16 ± 2.75
POPC_CHOL_50%	**MT-C34**	-289.58 ± 81.47	-685.05 ± 272.79	-26.81 ± 13.84	-2060.72 ± 469.53	-1006.89 ± 403.62	-0.34 ± 2.52
	MT-SC22EK	-212.73 ± 63.21	-563.46 ± 320.23	-26.16 ± 17.94	-1676.81 ± 542.48	-740.86 ± 420.98	0.55 ± 2.35
	MT-SC29EK	-291.66 ± 55.80	-479.67 ± 227.69	-1.65 ± 1.51	-2368.24 ± 471.07	-781.58 ± 370.16	-0.29 ± 1.96
	MT-SC34	-302.56 ± 66.44	-612.19 ± 236.36	-4.26 ± 3.39	-2433.71 ± 511.37	-826.32 ± 324.49	-10.97 ± 25.08
	MT-SC34EK	-337.45 ± 77.32	-447.85 ± 300.01	-13.86 ± 12.18	-2677.31 ± 704.35	-722.89 ± 481.39	-0.86 ± 4.68
	MT-SC35EK	-341.81 ± 66.25	-225.69 ± 168.57	-41.09 ± 32.38	-3665.34 ± 434.32	-338.43 ± 258.37	-21.13 ± 25.95

Table 2 shows the average and standard deviation of the LJ and Coulomb potentials for every system and interaction. In general, the Coulomb potential is more intense than the LJ potential. MT-SC22EK has the weakest PEP-POPC interaction out of all peptides in the POPC systems. This can be due to its smaller size, which means there are less amino acid residues to interact with the membrane. This aspect can greatly reduce interaction energy. In the POPC_CHOL_50% systems, MT-SC35EK was the peptide that registered the lowest PEP-POPC interaction energy, while simultaneously having the greatest PEP-CHOL interaction energy out of all other peptides. This can be attributed to the position that this peptide is in relative to the membrane. MT-SC35EK is the largest peptide that was studied, but as we saw in the Cα graphs, only a small part of it is connected to the phosphorus axis of the membrane while the rest is directed upwards away from the membrane. This means that not many amino acid residues are close enough to interact with POPC, greatly reducing PEP-POPC interaction energy. Additionally, the results on hydrogen bonds between the peptides and system components show us that MT-SC35EK establishes the most bonds with cholesterol of all the peptides. The higher number of hydrogen bonds and peptide positioning explains the increased interaction energy values.

Looking at all the systems simultaneously we can see that, from the POPC systems to the 20% cholesterol systems, interaction energies between the peptides with the solvent and POPC remain relatively close. But the 50% cholesterol systems (except the MT-SC22EK systems) show a large decrease in peptide-POPC interaction energy, while simultaneously having the highest peptide-solvent energy values. Once again, the rigidity of the 50% cholesterol membrane led to the peptides being positioned further from the membrane. This was observed in the C_α graphs and caused a decrease in the number of hydrogen bonds established with POPC, explaining the lower energy values.

The Gibbs free energy (ΔG) of each peptide interaction was calculated via the method of *umbrella sampling*. The outputs generated from the umbrella sampling simulations, allowed us to extract the Potential of Mean Force curve associated with each system. ΔG was calculated by subtracting the average of PMF of the last 1 nm of the curve to the minimum value of PMF. The PMF curves of the systems can be seen below, together with the table containing the ΔG values.

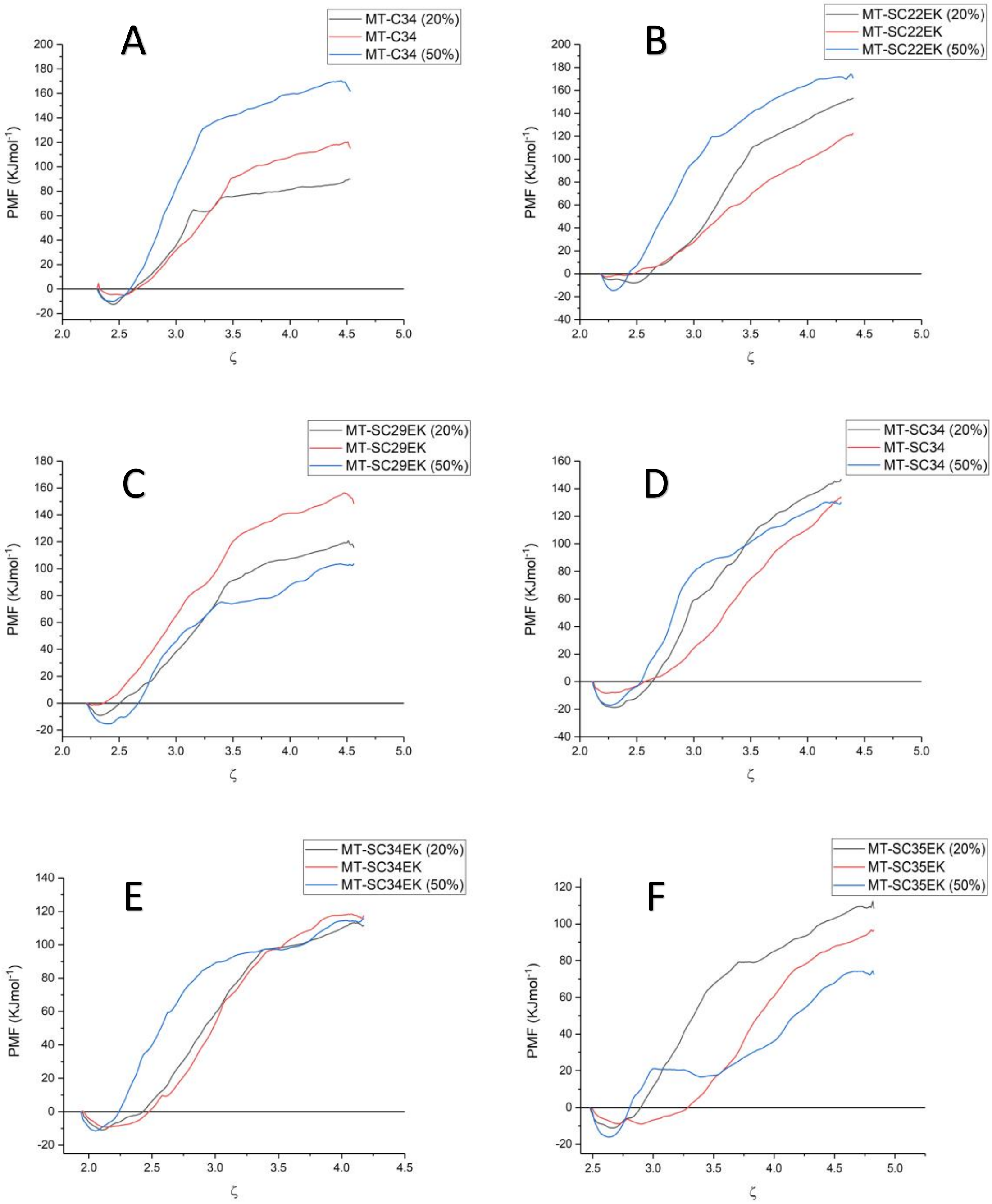


Figure 32 – (A) Potential of Mean Force (PMF) curve for MT-C34 dissociation in POPC, 20% cholesterol and 50% cholesterol membranes. (B) Potential of Mean Force (PMF) curve for MT-SC22EK dissociation in POPC, 20% cholesterol and 50% cholesterol membranes. (C) Potential of Mean Force (PMF) curve for MT-SC29EK dissociation in POPC, 20% cholesterol and 50% cholesterol membranes. (D) Potential of Mean Force (PMF) curve for MT-SC34 dissociation in POPC, 20% cholesterol and 50% cholesterol membranes. (E) Potential of Mean Force (PMF) curve for MT-SC34EK dissociation in POPC, 20% cholesterol and 50% cholesterol membranes. (F) Potential of Mean Force (PMF) curve for MT-SC35EK dissociation in POPC, 20% cholesterol and 50% cholesterol membranes.

Table 3 – Value of the Gibbs free energy (ΔG) for every system, with the respective standard deviations.

Membranes	**Variation in Gibbs free energy ($\Delta G/Kjmol^{-1}$)**	
POPC	**MT-C34**	-113.58 ± 7.09
	MT-SC22EK	-97.52 ± 16.23
	MT-SC29EK	-143.19 ± 8.79
	MT-SC34	-105.30 ± 22.37
	MT-SC34EK	-113.20 ± 12.85
	MT-SC35EK	-82.21 ± 18.37
POPC_CHOL _20%	**MT-C34**	-93.59 ± 3.37
	MT-SC22EK	-136.58 ± 16.14
	MT-SC29EK	-116.13 ± 7.69
	MT-SC34	-139.44 ± 18.49
	MT-SC34EK	-111.97 ± 8.35
	MT-SC35EK	-106.65 ± 10.07
POPC_CHOL_50%	**MT-C34**	-168.19 ± 8.35
	MT-SC22EK	-173.04 ± 11.70
	MT-SC29EK	-104.48 ± 10.61
	MT-SC34	-131.99 ± 11.59
	MT-SC34EK	-115.34 ± 7.48
	MT-SC35EK	-66.65 ± 17.09

In table 3, we see that the value of ΔG is negative in every system, therefore the binding of the peptide to the membrane is spontaneous. No specific pattern can be observed between the values of the different membrane systems. Reversibility, as we mentioned previously, is one of the most fundamental properties of a FIP. But, ΔG alone is not enough to determine whether an interaction is reversible or not (other aspects, such as the equilibrium conditions, the net charge of the system, and other external conditions play a role in determining the reversibility of an interaction), but it can provide a rough estimate (Greaves et al., 1992).

The peptides maintain highly negative ΔG value in the 50% cholesterol systems that resemble the HIV membrane more closely. The lowest values of ΔG are seen in the 50% cholesterol systems, for MT-C34 and MT-SC22EK peptides. The ΔG value of MT-SC34 is also quite low in the 50% and 20% cholesterol systems. The ΔG values of MT-SC34EK in all systems remain close, with the value of the 50% cholesterol system being the highest. ΔG of MT-SC29EK becomes less negative as cholesterol concentration increases, which can be detrimental to its effectiveness as an FIP. MT-SC35EK has the lowest ΔG value of all the peptides, with the 50% cholesterol system being the less negative one. In previous sections, we saw that MT-SC35EK in the 50% cholesterol system was positioned more vertically than other peptides, relative to the membrane. We thought that this might facilitate its interaction with the 6-HB, and ultimately make it a great candidate for use as an FIP. Perhaps these values deny that possibility. It is possible that this positioning was caused by artifacts occurring throughout the simulation or by an incomplete binding, and not by the regular behavior of the peptide.

These values allow us to make assumptions as to how effectively the peptides partition towards membranes. The moderate negative values that we see in every system are indicative of a positive partition coefficient, meaning that the peptides will partition efficiently towards these structures. This means that peptides will have a higher chance to be adsorbed closer to the sites where they are bound to act.

3.2.7. Peptide Lateral Diffusion

The lateral diffusion coefficient (D_{lat}) was calculated from the mean square displacement (MSD) using the Einstein relation:

$$D_{lat} = \frac{1}{2d} \lim_{t \to \infty} \frac{dMSD(t)}{dt} \tag{16}$$

MSD is defined as:

$$MSD = < \| \vec{R}_i(t + t_0) - \vec{R}_i(t_0) \|^2 > \tag{17}$$

Here, $\vec{R}_i$ represents the coordinates of the center of mass of the atoms, i, of the peptide. MSD was calculated for the last 100 ns of simulation, while D_{lat} was obtained by doing a linear regression of the MSD curve from 20 ns to 60 ns for peptides in water and for the peptides in membrane systems.

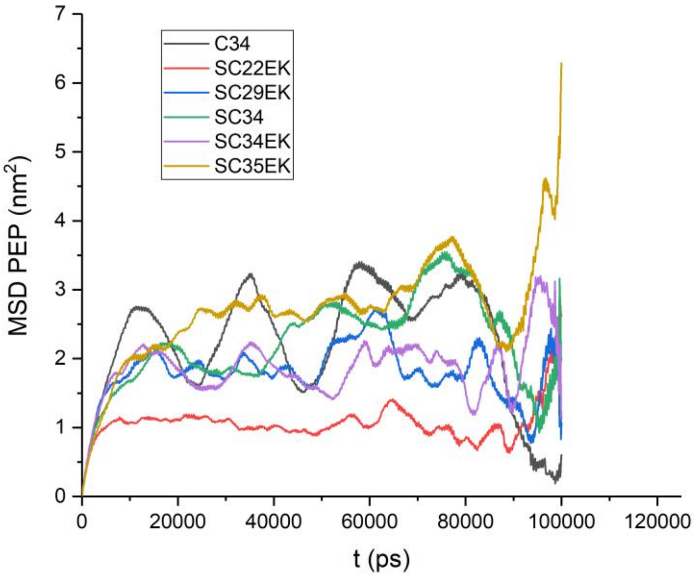


Figure 33 – Mean Square Displacement (MSD) of all the peptides in water in the last 100 ns of the trajectory.

Due to only having one peptide in each system the MSD profiles lacked sampling, resulting in unstable and highly variable profiles. As a result, some of the peptides displayed negative D_{lat} values. To fix this issue, the intervals for linear regression of MSD were changed in those specific systems. In the MT-SC22EK system in water and in the MT-SC29EK systems in the 20% and 50% cholesterol membranes, the interval chosen was from 10 ns - 50 ns.

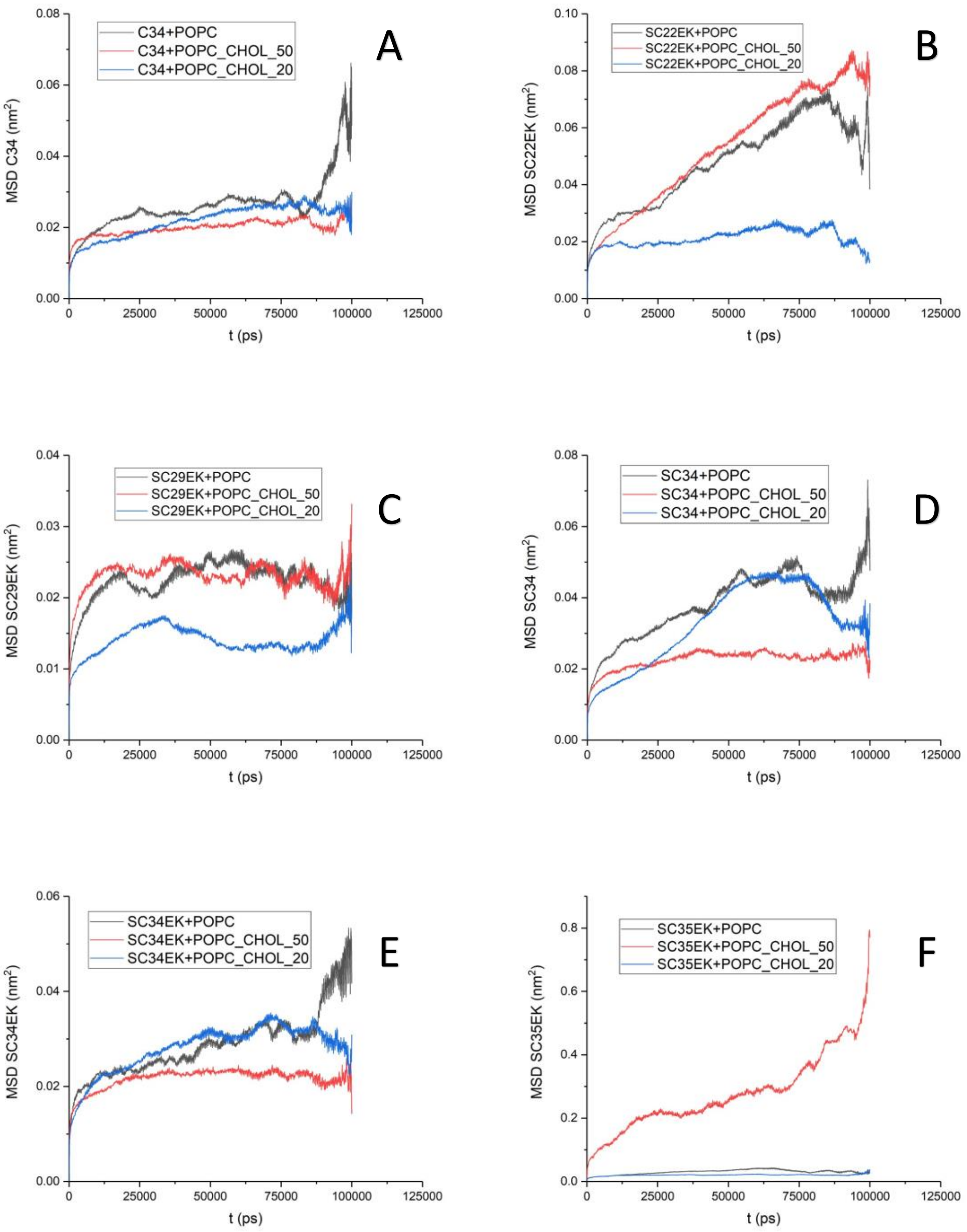


Figure 34 - (A) Mean Square Displacement (MSD) values of MT-C34 in POPC and POPC-CHOL systems. (B) Mean Square Displacement (MSD) values of MT-SC22EK in POPC and POPC-CHOL systems. (C) Mean Square Displacement (MSD) values of MT-SC29EK in POPC and POPC-CHOL systems. (D) Mean Square Displacement (MSD) values of MT-SC34 in POPC and POPC-CHOL systems. (E) Mean Square Displacement (MSD) values of MT-SC34EK in POPC and POPC-CHOL systems. (F) Mean Square Displacement (MSD) values of MT-SC35EK in POPC and POPC-CHOL systems.

Table 4 -Lateral diffusion coefficient values for all peptides of the water, POPC and POPC-CHOL systems from 20 ns to 60 ns.

D_{lat} (10^{-7} cm^2s^{-1})						
Systems	MT-C34	MT-SC22EK	MT-SC29EK	MT-SC34	MT-SC34EK	MT-SC35EK
Water	1.84 ± 0.61	6.52 ± 0.08	1.17 ± 0.24	2.76 ± 0.38	0.140 ± 0.23	0.51 ± 0.13

D_{lat} (10^{-9} cm^2s^{-1})						
Systems	MT-C34	MT-SC22EK	MT-SC29EK	MT-SC34	MT-SC34EK	MT-SC35EK
POPC	1.26 ± 0.00	6.62 ± 0.01	1.41 ± 0.00	4.42 ± 0.01	1.82 ± 0.00	3.52 ± 0.00
POPC_CHOL_20%	1.97 ± 0.00	1.46 ± 0.00	1.38 ± 0.00	7.08 ± 0.01	1.82 ± 0.00	0.43 ± 0.00
POPC_CHOL_50%	0.61 ± 0.00	7.78 ± 0.01	0.48 ± 0.00	0.79 ± 0.00	0.40 ± 0.00	20.50 ± 0.03

From Table 4 we can see that peptides in water have higher D_{lat} values than the peptides from the membrane systems. This is expected since adhesion of the peptides to the membrane surface will impair their mobility. Lateral diffusion is faster in the POPC systems than in the 50% cholesterol systems, except for MT-SC35EK. As we saw previously from the Cα profiles of the peptides, the connection of MT-SC35EK to the membrane in the 50% cholesterol system is very superficial. Few amino acid residues on the 3' end of the peptide are in contact with the membrane. The rest of the chain is drifting in the solvent, which results in more peptide lateral diffusion. This might also have implications on rotational dynamics. MT-SC22EK in the 50% cholesterol system also displayed an unusually high lateral diffusion. Due to the size of the peptide, it lacks many interactive sites with the membrane, leading to weaker interactions with the membrane, which can accelerate its lateral diffusion. The results of the 20% cholesterol systems are not as clear, with MT-C34 and MT-SC34 in these systems having higher lateral diffusion values than in the POPC systems. The values are close, so it is possible that this concentration of cholesterol is low enough to maintain a relatively high membrane fluidity, which is not possible with a 50% cholesterol concentration.

Almost all peptides form more hydrogen bonds in the POPC systems than in POPC-CHOL systems. Interaction energy values are also higher in POPC systems, so these values do not appear to have a considerable influence on lateral diffusion. The discrepancy between systems is likely due to differences in membrane fluidity. The higher fluidity of the POPC membranes results in more lipid movement. In turn, this translates into more peptide movement after adsorption. On the other hand, lipid diffusion across the POPC-CHOL membranes is limited. This will likely decrease all peptide movement, including lateral diffusion.

3.2.8. Peptide Rotational Dynamics

Rotational dynamics studies were performed using the rotation autocorrelation function, C(t). The function is as follows:

$$C(t) = P_2\big(\cos\theta\,(\xi)\big) \tag{18}$$

where $\theta(\xi)$ represents the angle along the axis of the peptide at times ξ and $t+\xi$, and

$$P_2(x) = \frac{3x^2 - 1}{2} \tag{19}$$

is a Legendre polynomial of the second kind. The average performed over ξ is a good approximation to the average of the sample, assuming that the trajectory is ergodic.

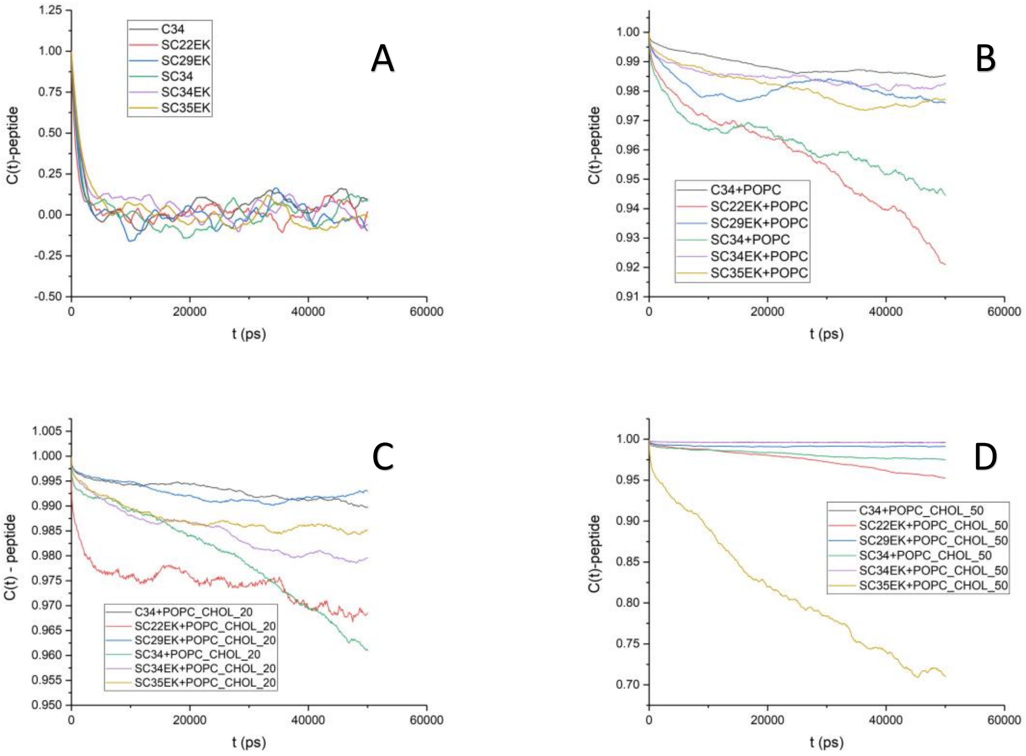


Figure 35 – Rotational autocorrelation function (Rotacf) for: (A) peptides in solution, (B) POPC systems, (C) POPC_CHOL_20% systems and (C) POPC-CHOL_50% systems.

The results in Figure 35 show that the peptides in solution have more freedom of rotation than those in the membrane systems, since C(t) drops to values close to 0 almost instantly. In the POPC and POPC-CHOL systems, C(t) values take longer to converge and hover relatively close to 1. This convergence is more pronounced in the POPC systems and more stable in POPC-CHOL systems, meaning that peptides have more freedom of rotation in the former than in the latter. MT-SC35EK in the 50% cholesterol system is the only exception, with C(t) values dropping to around 0.70, which indicates that the peptide can rotate more freely than the others. As was previously mentioned, MT-SC35EK has the weakest interaction energy with POPC in the POPC-CHOL systems. But it also has the highest interaction energies with the solvent and cholesterol. This together with its positioning relative to the membrane will give the peptide more freedom of rotation, hence the results. This provides more evidence for the fact that the simulation of MT-SC35EK in this system was incomplete and that the peptide did not complete the process of adsorption to the membrane.

Membrane Properties

3.3.1. Area per Lipid and Membrane Thickness

The calculation of the area per lipid in separate monolayers of the membrane systems was made via a custom software that sampled the simulation box, looking for previously defined reference atoms. For POPC, the reference atoms were P8, C6 and N4. For cholesterol, the reference atoms were H7 and C6. The monolayers were obtained by referencing the atoms in separate index files. Sampling is made through random points in the system and rules of three are made between the area of the simulation box, the total number of points sampled and the number of those points that corresponded to a reference atom. The resulting output will be the area per the chosen lipid.

Membrane thickness calculations were made in the same way as in the first section titled: "System Equilibration: Area per Lipid and Membrane Thickness".

Table 5 – Average area per POPC (ApPOPC) and cholesterol (ApCHOL), and membrane thickness (MT) values of the peptide membrane systems and sole membranes.

Systems		**ApPOPC (nm^2)**	**ApCHOL (nm^2)**	**MT (nm)**
POPC	Both	0.660 ± 0.007	-	3.77 ± 0.04
MT-C34+POPC	Upper	0.655 ± 0.001	-	3.77 ± 0.02
	Lower	0.655 ± 0.001	-	
MT-SC22EK+POPC	Upper	0.643 ± 0.007	-	3.84 ± 0.04
	Lower	0.643 ± 0.008	-	
MT-SC29EK+POPC	Upper	0.629 ± 0.001	-	3.91 ± 0.02
	Lower	0.629 ± 0.001	-	
MT-SC34+POPC	Upper	0.644 ± 0.001	-	3.85 ± 0.02
	Lower	0.644 ± 0.001	-	
MT-SC34EK+POPC	Upper	0.641 ± 0.001	-	3.84 ± 0.02
	Lower	0.641 ± 0.001	-	
MT-SC35EK+POPC	Upper	0.642 ± 0.001	-	3.84 ± 0.02
	Lower	0.642 ± 0.001	-	

POPC_CHOL_20%	Both	0.593 ± 0.009	0.214 ± 0.018	4.26 ± 0.05
MT-C34 + POPC_CHOL_20%	Upper	0.577 ± 0.005	0.217 ± 0.015	4.33 ± 0.02
	Lower	0.581 ± 0.004	0.194 ± 0.017	
MT-SC22EK + POPC_CHOL_20%	Upper	0.576 ± 0.004	0.204 ± 0.017	4.35 ± 0.02
	Lower	0.576 ± 0.004	0.205 ± 0.017	
MT-SC29EK + POPC_CHOL_20%	Upper	0.583 ± 0.008	0.211 ± 0.017	4.31 ± 0.04
	Lower	0.582 ± 0.009	0.214 ± 0.019	
MT-SC34 + POPC_CHOL_20%	Upper	0.581 ± 0.008	0.224 ± 0.018	4.27 ± 0.04
	Lower	0.587 ± 0.008	0.203 ± 0.017	
MT-SC34EK + POPC_CHOL_20%	Upper	0.580 ± 0.008	0.225 ± 0.020	4.30 ± 0.04
	Lower	0.596 ± 0.008	0.203 ± 0.018	
MT-SC35EK + POPC_CHOL_20%	Upper	0.585 ± 0.008	0.209 ± 0.017	4.31 ± 0.04
	Lower	0.582 ± 0.010	0.222 ± 0.022	
POPC_CHOL_50%	Both	0.525 ± 0.009	0.269 ± 0.008	4.62 ± 0.02
MT-C34 + POPC_CHOL_50%	Upper	0.529 ± 0.009	0.270 ± 0.009	4.60 ± 0.02
	Lower	0.539 ± 0.008	0.260 ± 0.008	
MT-SC22EK + POPC_CHOL_50%	Upper	0.526 ± 0.009	0.276 ± 0.009	4.59 ± 0.02
	Lower	0.526 ± 0.008	0.275 ± 0.008	
MT-SC29EK +POPC_CHOL_50%	Upper	0.524 ± 0.007	0.268 ± 0.007	4.62 ± 0.02
	Lower	0.508 ± 0.009	0.285 ± 0.010	
MT-SC34 + POPC_CHOL_50%	Upper	0.526 ± 0.008	0.268 ± 0.008	4.60 ± 0.02
	Lower	0.519 ±0.008	0.275 ± 0.008	
MT-SC34EK + POPC_CHOL_50%	Upper	0.529 ± 0.010	0.271 ± 0.010	4.58 ± 0.02
	Lower	0.515 ± 0.007	0.285 ± 0.007	
MT-SC35EK + POPC_CHOL_50%	Upper	0.529 ± 0.016	0.284 ± 0.013	4.60 ± 0.02
	Lower	0.531 ± 0.008	0.265 ± 0.008	

Table 5 shows a gradient of ApPOPC and MT values across the various membranes. POPC membranes have the highest ApPOPC values, followed by the 20% cholesterol membranes and finally the 50% cholesterol membranes. The monolayers of the POPC peptide systems show decreased values of ApPOPC compared to the reference membrane. The decrease is the same for both monolayers. On the other hand, MT values increase slightly in every system except for MT-C34 + POPC. These changes are more noticeable due to the fluidity of the membrane. Peptide adsorption likely disrupts the order of the membrane by compressing POPC molecules, reducing their area. Additionally, as the peptides get closer to the membrane surface, they start inserting themselves in the membranes. It is possible that some membrane lipids expand outwards to cover the side chains of hydrophobic amino acids, increasing membrane thickness. POPC-CHOL systems mostly show different ApPOPC and ApCHOL values for each monolayer. Only MT-SC22EK in both systems remains equal. Every other system, except for MT-C34 + POPC_CHOL_50%, shows the same pattern. The monolayer in which the peptide is adsorbed has a decreased ApPOPC value and an increased ApCHOL value. These differences between monolayers are possible due to the resistance of these membranes to conformation changes due to cholesterol. Even so results suggest that peptide adsorption compresses POPC molecules and affects their ordering. Consequently, reduced protection of the cholesterol molecules by POPC means that cholesterol will be more exposed to the solvent. Solvent and cholesterol interaction cause ApCHOL values to increase. MT values for the 20% and 50% cholesterol systems were larger compared to the values of POPC systems. In the 20% systems MT values increased relative to the membrane. But in the 50% cholesterol systems, MT remained the same or decrease slightly. The 20% cholesterol membrane allowed significant insertion of the peptides, so the lipids probably expanded outwards like in the POPC membrane. On the other hand, the 50% cholesterol membrane, which did not allow for significant insertion of the peptides, got compressed locally, leading to decreased MT values.

3.3.2. Hydrogen bonds between POPC, cholesterol and solvent molecules

The interactions between POPC and cholesterol with solvent molecules were obtained via the same procedure that was previously mentioned for peptide interactions. In POPC systems, only the interaction between POPC and solvent (POPC-SOL) was registered. POPC-CHOL systems had the added interactions between cholesterol and POPC (CHOL-POPC), and cholesterol and solvent (CHOL-SOL). This was done for the last 100 ns of the trajectory

Table 6 – Average number of hydrogen bonds between POPC-SOL, CHOL-POPC and CHOL-SOL, with their respective standard deviations for the last 100 ns of the trajectory.

	Number of hydrogen bonds per lipid		
	POPC-SOL	**CHOL-POPC**	**CHOL-SOL**
POPC	6.471 ± 0.105	-	-
MT-C34 + POPC	6.225 ± 0.095	-	-
MT-SC22EK + POPC	6.206 ± 0.102	-	-
MT-SC29EK + POPC	6.079 ± 0.107	-	-
MT-SC34 + POPC	6.185 ± 0.103	-	-
MT-SC34EK + POPC	6.142 ± 0.097	-	-
MT-SC35EK + POPC	6.127 ± 0.100	-	-
POPC_CHOL_20%	6.198 ± 0.113	0.867 ± 0.052	0.436 ± 0.090
MT-C34 + POPC_CHOL_20%	6.697 ± 0.113	0.847 ± 0.053	0.268 ± 0.059
MT-SC22EK + POPC_CHOL_20%	5.881 ± 0.097	0.868 ± 0.054	0.414 ± 0.085
MT-SC29EK + POPC_CHOL_20%	5.835 ± 0.112	0.859 ± 0.054	0.434 ± 0.090
MT-SC34 + POPC_CHOL_20%	5.830 ± 0.118	0.843 ± 0.060	0.463 ± 0.087
MT-SC34EK + POPC_CHOL_20%	5.861 ± 0.106	0.859 ± 0.060	0.464 ± 0.097
MT-SC35EK + POPC_CHOL_20%	5.809 ± 0.100	0.843 ± 0.057	0.427 ± 0.085
POPC_CHOL_50%	6.464 ± 0.136	0.771 ± 0.038	0.687 ± 0.062
MT-C34 + POPC_CHOL_50%	6.296 ± 0.124	0.750 ± 0.034	0.708 ± 0.058
MT-SC22EK + POPC_CHOL_50%	6.236 ± 0.120	0.786 ± 0.031	0.662 ± 0.055
MT-SC29EK + POPC_CHOL_50%	6.235 ± 0.132	0.770 ± 0.039	0.721 ± 0.059
MT-SC34 + POPC_CHOL_50%	6.208 ± 0.122	0.758 ± 0.034	0.708 ± 0.062

MT-SC34EK + POPC_CHOL_50%	6.244 ± 0.126	0.765 ± 0.031	0.711 ± 0.054
MT-SC35EK + POPC_CHOL_50%	6.271 ± 0.127	0.770 ± 0.034	0.667 ± 0.060

The total number of hydrogen bonds formed in the POPC-CHOL systems is higher than in POPC systems because of the added interactions with cholesterol. Table 6 shows a reduction in hydrogen bonds between POPC and solvent when peptides are adsorbed to the membrane (except in the MT-C34 20% cholesterol system). This happens for all membrane systems. In the 20% and 50% cholesterol systems, the number of hydrogen bonds formed between cholesterol and POPC decreases relative to the membrane, while interactions between cholesterol and solvent increase slightly. These changes can be attributed to the effects that peptide adsorption has on the membranes. As the peptide approaches the membrane it will start interacting with POPC and some cholesterol molecules located closer to the surface. As a result, POPC-SOL and CHOL-POPC interactions decrease since these molecules will cease interactions with each other to interact with the peptide. Finally, changes in membrane conformation caused by the peptide will disrupt the umbrella effect that the phospholipids have on cholesterol molecules. Therefore, these molecules will be more exposed to the solvent. This raises CHOL-SOL interactions. In the MT-SC22EK 50% cholesterol system, this trend is inverted. This can be due to the smaller size of MT-SC22EK, which causes less CHOL-POPC interactions to be broken for these molecules to interact with the peptide. Less conformation changes will also mean less cholesterol exposure, reducing CHOL-SOL interactions. MT-SC35EK in the 50% cholesterol system also has very little CHOL-SOL interactions. As mentioned previously, in this system, only a small section of MT-SC35EK interacts with the membrane, inducing little changes in conformation. This leads to less cholesterol exposure and less CHOL-SOL interactions.

3.3.3. POPC and Cholesterol Lateral Diffusion

The lateral diffusion coefficient (D_{lat}) was calculated from the mean square displacement (MSD) using Einstein's relation (d=2 for bilayer systems), as described previously for the peptides. D_{lat} was calculated for each monolayer separately in the case of systems, between 20 ns and 60 ns.

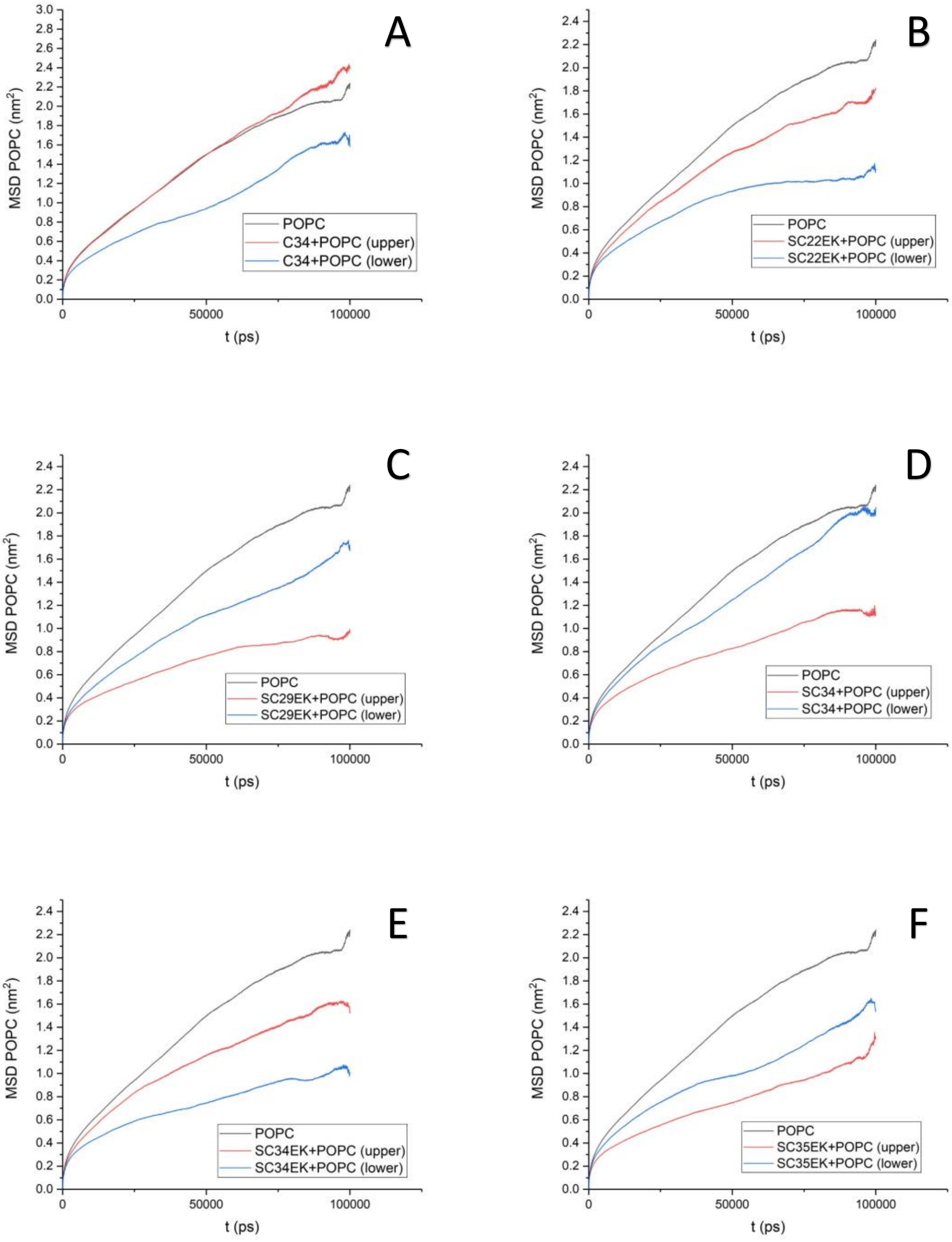


Figure 36 – Mean square displacement of POPC (MSD_{POPC}) of the system's monolayers, compared to MSD_{POPC} of the POPC bilayer. (A) MSD_{POPC} for the monolayers of the MT-C34+POPC system. (B) MSD_{POPC} for the monolayers of the MT-SC22EK+POPC system. (C) MSD_{POPC} for the monolayers of the MT-SC29EK+POPC system. (D) MSD_{POPC} for the monolayers of the MT-SC34+POPC system. (E) MSD_{POPC} for the monolayers of the MT-SC34EK+POPC system. (F) MSD_{POPC} for the monolayers of the MT-SC35EK+POPC system.

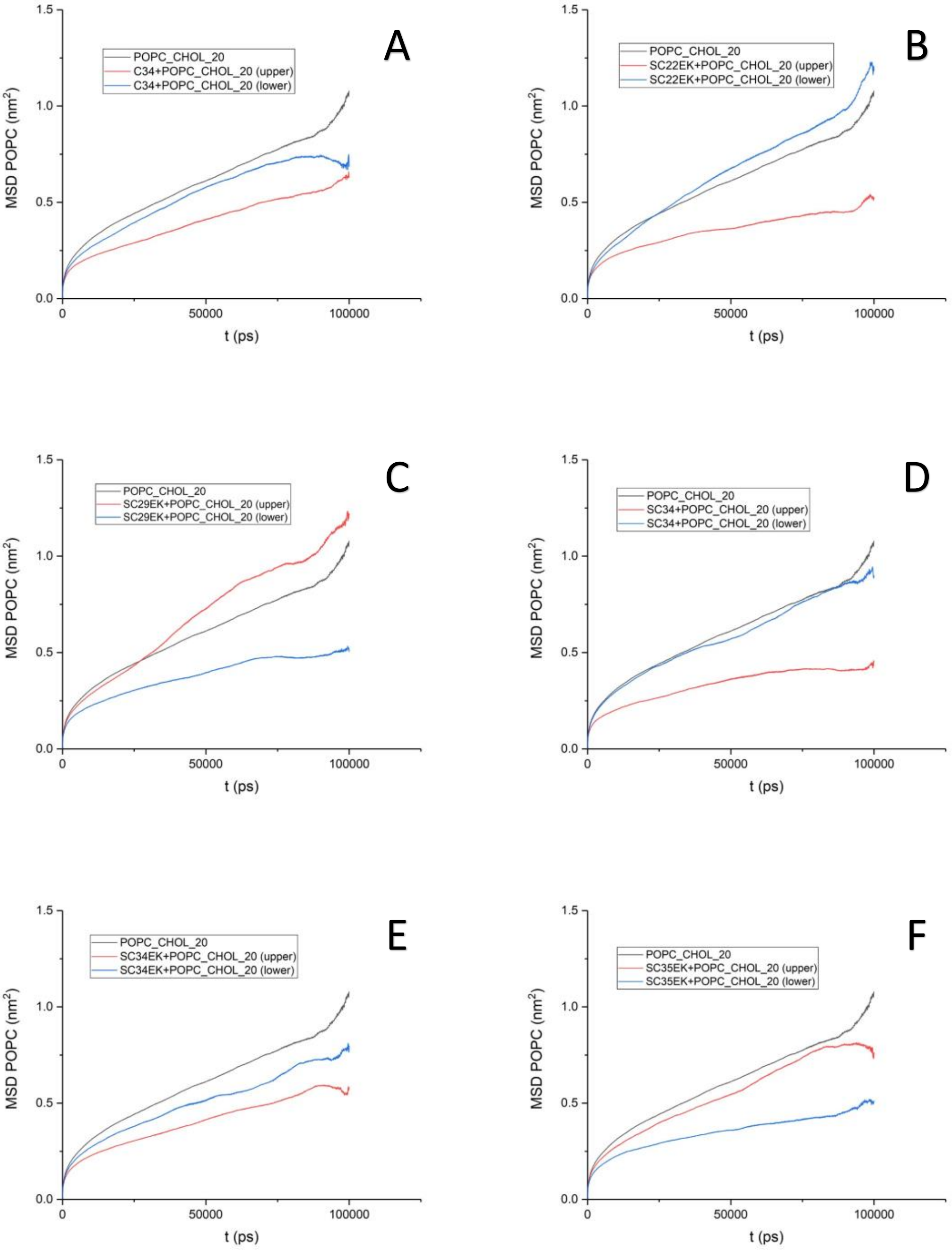


Figure 37 – Mean square displacement of POPC (MSD$_{POPC}$) of the system's monolayers, compared to MSD$_{POPC}$ of the POPC_CHOL_20% bilayer. (A) MSD$_{POPC}$ for the monolayers of the MT-C34+POPC_CHOL_20% system. (B) MSD$_{POPC}$ for the monolayers of the MT-SC22EK+POPC_CHOL_20% system. (C) MSD$_{POPC}$ for the monolayers of the MT-SC29EK+POPC_CHOL_20% system. (D) MSD$_{POPC}$ for the monolayers of the MT-SC34+POPC_CHOL_20% system. (E) MSD$_{POPC}$ for the monolayers of the MT-SC34EK+POPC_CHOL_20% system. (F) MSD$_{POPC}$ for the monolayers of the MT-SC35EK+POPC_CHOL_20% system.

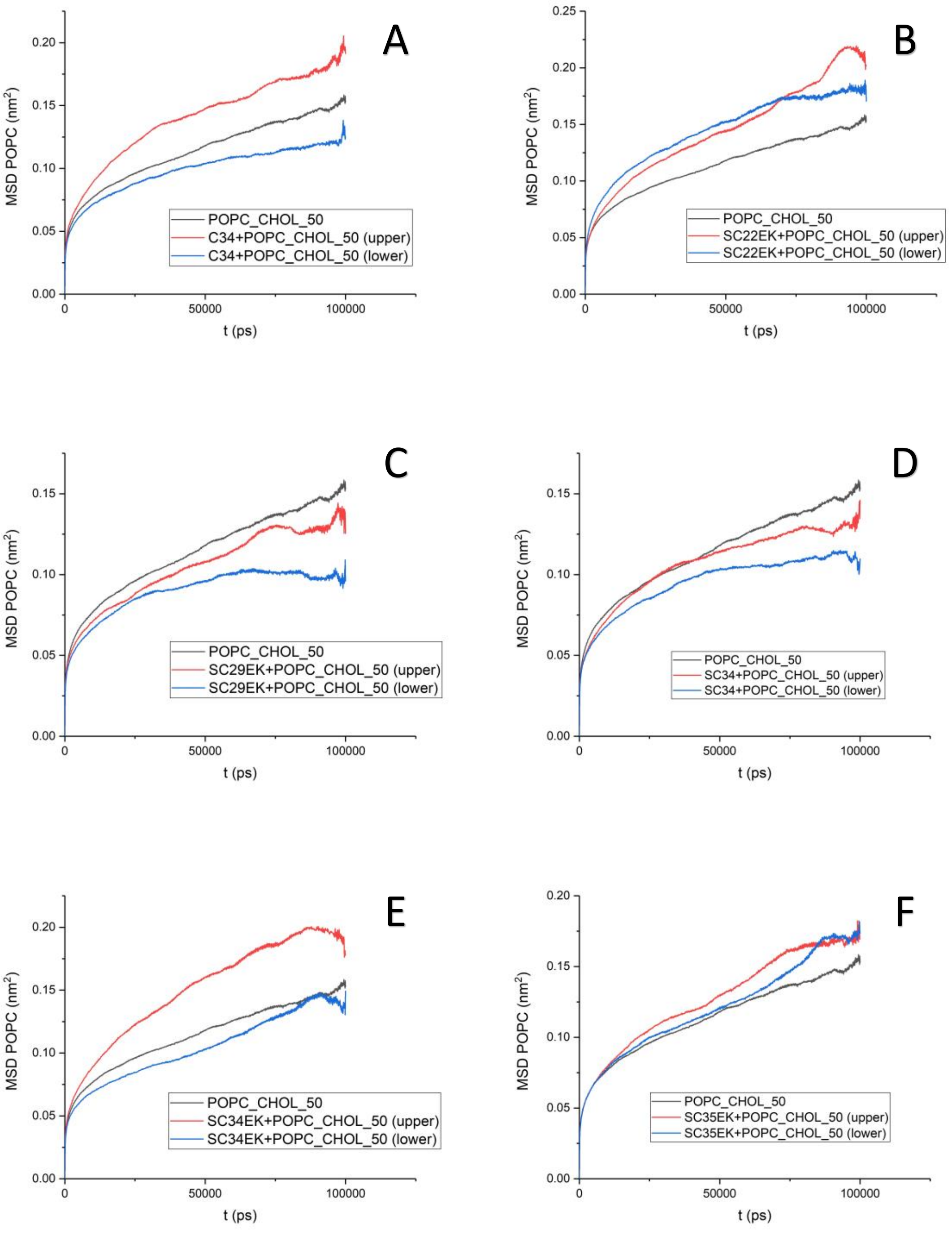


Figure 38- Mean square displacement of POPC (MSD_{POPC}) of the system's monolayers, compared to MSD_{POPC} of the POPC_CHOL_50% bilayer. (A) MSD_{POPC} for the monolayers of the MT-C34+POPC_CHOL_50% system. (B) MSD_{POPC} for the monolayers of the MT-SC22EK+POPC_CHOL_50% system. (C) MSD_{POPC} for the monolayers of the MT-SC29EK+POPC_CHOL_50% system. (D) MSD_{POPC} for the monolayers of the MT-SC34+POPC_CHOL_50% system. (E) MSD_{POPC} for the monolayers of the MT-SC34EK+POPC_CHOL_50% system. (F) MSD_{POPC} for the monolayers of the MT-SC35EK+POPC_CHOL_50% system.

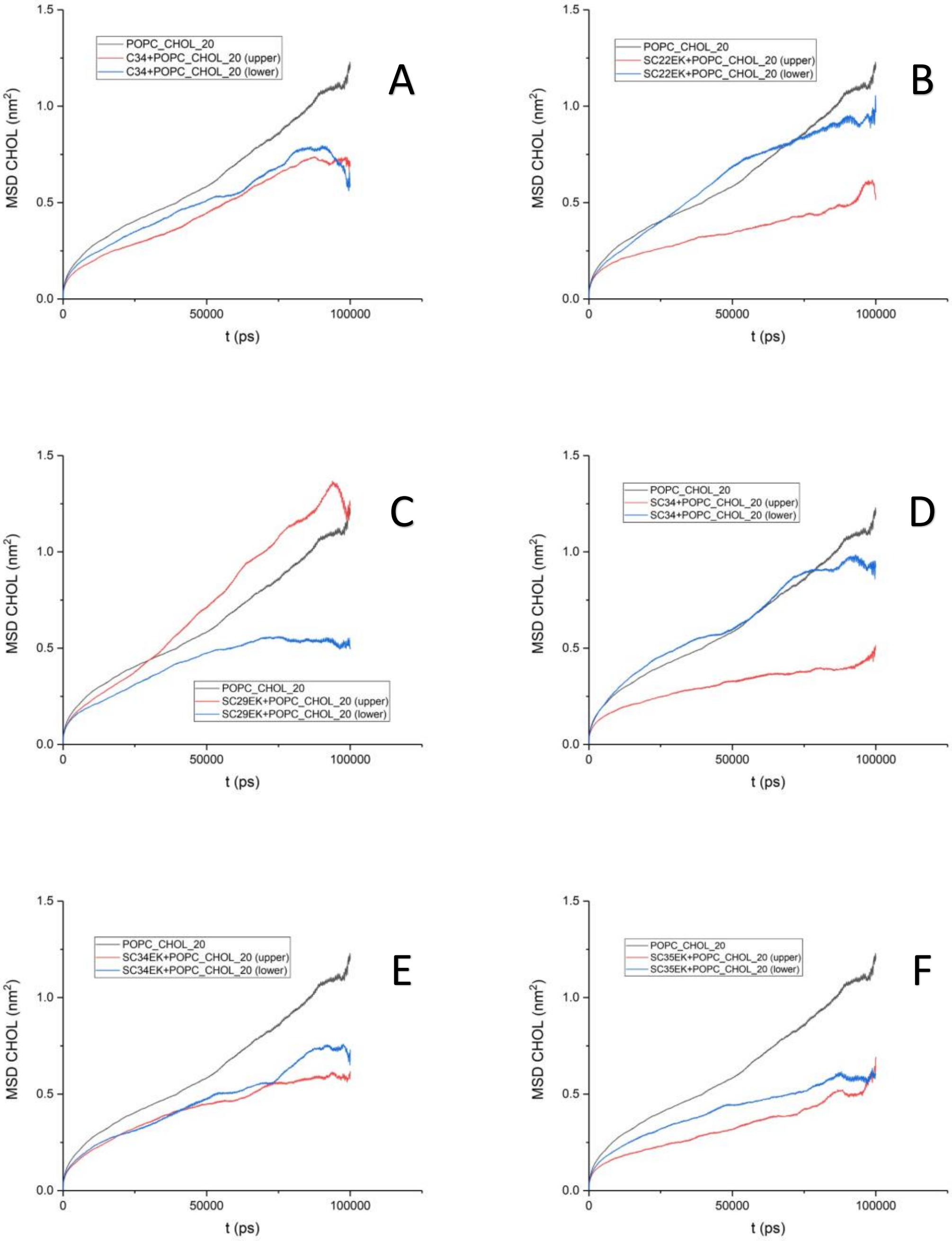


Figure 39 – Mean square displacement of cholesterol (MSD_{CHOL}) of the system's monolayers, compared to MSD_{CHOL} of the POPC_CHOL_20% bilayer. (A) MSD_{CHOL} for the monolayers of the MT-C34+POPC_CHOL_20% system. (B) MSD_{CHOL} for the monolayers of the MT-SC22EK+POPC_CHOL_20% system. (C) MSD_{CHOL} for the monolayers of the MT-SC29EK+POPC_CHOL_20% system. (D) MSD_{CHOL} for the monolayers of the MT-SC34+POPC_CHOL_20% system. (E) MSD_{CHOL} for the monolayers of the MT-SC34EK+POPC_CHOL_20% system. (F) MSD_{CHOL} for the monolayers of the MT-SC35EK+POPC_CHOL_20% system.

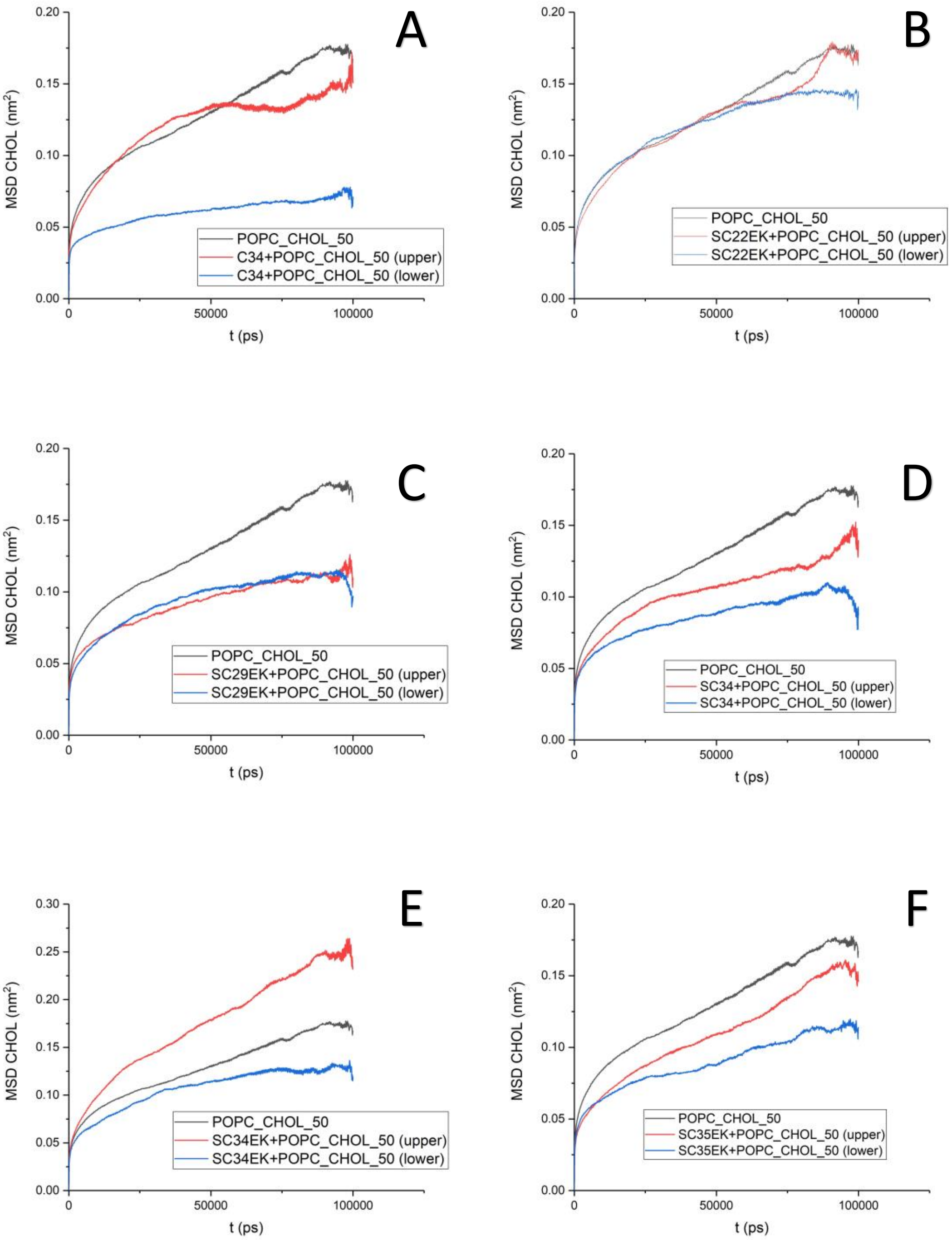


Figure 40 - Mean square displacement of cholesterol (MSD_{CHOL}) of the system's monolayers, compared to MSD_{CHOL} of the POPC_CHOL_50% bilayer. (A) MSD_{CHOL} for the monolayers of the MT-C34+POPC_CHOL_50% system. (B) MSD_{CHOL} for the monolayers of the MT-SC22EK+POPC_CHOL_50% system. (C) MSD_{CHOL} for the monolayers of the MT-SC29EK+POPC_CHOL_50% system. (D) MSD_{CHOL} for the monolayers of the MT-SC34+POPC_CHOL_50% system. (E) MSD_{CHOL} for the monolayers of the MT-SC34EK+POPC_CHOL_50% system. (F) MSD_{CHOL} for the monolayers of the MT-SC35EK+POPC_CHOL_50% system.

Figure 36 shows the MSD_{POPC} values for the POPC systems. We see that MT-C34, MT-SC22EK and MT-SC34EK have lower diffusion values in the bottom monolayer. This happens because the peptides are adsorbed to this monolayer, where they slow lipid diffusion. The contrary happens for the remaining peptides that are adsorbed to the upper monolayer, leading to slower diffusion rates in this layer compared to the other. In the MT-C34 + POPC system, diffusion of the upper monolayer closely resembles that of the sole POPC membrane. All other systems display lower lateral diffusion rates than the POPC membrane. This shows that peptide adsorption slightly hinders this process. MT-SC35EK + POPC appears to have the lowest overall lateral diffusion value of all these systems. This is probably due to its size. MSD_{POPC} values for these monolayers peak at around 0.8 nm^2 for the layer with the peptide, and 1.8 nm^2 for the free monolayer.

In Figure 37 we can also distinguish the peptides that are adsorbed to the upper and lower monolayers by their respective MSD_{POPC} values. Every peptide, except MT-SC29EK and MT-SC35EK is adsorbed to the upper monolayer. Relative to the POPC systems, MSD_{POPC} values for the 20% cholesterol monolayers are lower, peaking at around 0.4 nm^2 for the layer with the peptide adsorbed, and 1 nm^2 for the free monolayer. This is close to half the value of the POPC systems.

The results from the 50% cholesterol systems are less clear. In Figure 38 we see that MSD_{POPC} values of both monolayers are much closer. Only MT-SC22EK and MT-SC35EK are adsorbed to the upper monolayer, even though MT-SC34 also has higher MSD_{POPC} values in the bottom monolayer. The 50% cholesterol systems have the lowest MSD_{POPC} values out of all the systems, peaking at around 0.12 nm^2 for the layer with the peptide adsorbed, and 0.2 nm^2 for the free monolayer.

Figures 39 and 40 show the MSD_{CHOL} values of the 20% and 50% cholesterol systems. Once again, these values are lower in the 50% cholesterol systems than in the 20% cholesterol systems, due to the rigidity of the membrane. There is no clear relationship between where the peptide is adsorbed and MSD_{CHOL} values from observation of the figures alone, but perhaps something can be said from the analysis of the lateral diffusion coefficient.

Overall, MSD_{POPC} in POPC-CHOL systems will tend to be lower than in POPC systems. This is due to the stability that cholesterol gives to the membrane, which will impair its movement. As a result, the slopes of POPC-CHOL graphs are less pronounced.

Table 7 – Lateral diffusion coefficient (D_{lat}) values of POPC and cholesterol for all peptide-membrane systems.

		Dlat (10^{-8} cm^{2}s^{-1})	
Systems	**Layers**	**POPC**	**CHOL**
POPC	Both	21.456 ± 0.250	-
MT-C34 + POPC	Upper	22.023 ± 0.255	-
	Lower	10.798 ± 0.125	-
MT-SC22EK + POPC	Upper	15.897 ± 0.185	-
	Lower	9.808 ± 0.114	-
MT-SC29EK + POPC	Upper	8.357 ± 0.097	-
	Lower	13.618 ± 0.159	-
MT-SC34 + POPC	Upper	8.163 ± 0.095	-
	Lower	1.608 ± 0.186	-
MT-SC34EK + POPC	Upper	12.505 ± 0.145	-
	Lower	6.478 ± 0.075	-
MT-SC35EK + POPC	Upper	7.798 ± 0.091	-
	Lower	8.750 ± 0.103	-
POPC_CHOL_20%	Both	6.77 ± 0.078	7.828 ± 0.091
MT-C34+POPC_CHOL_20%	Upper	4.786 ± 0.055	6.684 ± 0.078
	Lower	7.079 ± 0.082	6.104 ± 0.071
MT-SC22EK+POPC_CHOL_20%	Upper	2.763 ± 0.033	3.293 ± 0.038
	Lower	8.955 ± 0.103	10.844 ± 0.126
MT-SC29EK+POPC_CHOL_20%	Upper	11.738 ± 0.136	13.534 ± 0.157
	Lower	3.702 ± 0043	6.055 ± 0.071
MT-SC34+POPC_CHOL_20%	Upper	3.547 ± 0.041	3.101 ± 0.036
	Lower	5.568 ± 0.065	6.222 ± 0.073
MT-SC34EK+POPC_CHOL_20%	Upper	4.405 ± 0.051	4.485 ± 0.053
	Lower	5.251 ± 0.061	6.255 ± 0.0728
MT-SC35EK+POPC_CHOL_20%	Upper	6.110 ± 0.071	3.785 ± 0.044
	Lower	2.766 ± 0.032	4.622 ± 0.054

POPC_CHOL_50%	Both	0.863 ± 0.010	1.008 ± 0.012
MT-C34+POPC_CHOL_50%	Upper	0.998 ± 0.012	1.026 ± 0.010
	Lower	0.625 ± 0.007	0.258 ± 0.003
MT-SC22EK+POPC_CHOL_50%	Upper	1.145 ± 0.013	1.026 ± 0.012
	Lower	1.103 ± 0.013	0.794 ± 0.009
MT-SC29EK+POPC_CHOL_50%	Upper	0.770 ± 0.009	0.626 ± 0.007
	Lower	0.467 ± 0.006	0.629 ± 0.007
MT-SC34+POPC_CHOL_50%	Upper	0.684 ± 0.008	0.536 ± 0.006
	Lower	0.628 ± 0.007	0.471 ± 0.005
MT-SC34EK+POPC_CHOL_50%	Upper	1.444 ± 0.017	1.643 ± 0.019
	Lower	0.756 ± 0.009	0.751 ± 0.009
MT-SC35EK+POPC_CHOL_50%	Upper	0.971 ± 0.011	0.846 ± 0.010
	Lower	0.851 ± 0.010	0.468 ± 0.006

Table 7 confirms what was previously stated. For POPC systems, MT-C34, MT-SC22EK and MT-SC34EK are adsorbed to the lower monolayer, which translates into reduced lateral diffusion rates in this layer. MT-C34 + POPC has the highest diffusion rate for the upper monolayer of every system (higher than the sole POPC membrane). MT-SC29EK, MT-SC34 and MT-SC35EK are adsorbed to the upper monolayer, and this causes a decrease in lateral diffusion on this specific layer. Therefore, there is, for fluid POPC membranes, a clear connection between peptide adsorption and reduction in lipid lateral movement.

D_{lat} values of POPC in POPC-CHOL systems are lower. The higher membrane stability and rigidity given by cholesterol restricts the movement of the lipids. By looking at the results for the 20% and 50% cholesterol membranes, it seems that the amount of membrane cholesterol and the lateral movement of POPC are inversely proportional. The same can be said to some degree for cholesterol movement, but the results are not so clear. The 20% cholesterol systems behave in the same way as the POPC systems. Peptide adsorption lowers D_{lat} values, which also suggests a connection between adsorption and reduction in lateral movement. Cholesterol molecules do not seem to follow a specific trend.

In the 50% cholesterol systems, MT-C34, MT-SC29EK, MT-SC34 and MT-SC34EK, are adsorbed to the lower monolayer, which lowers the lateral diffusion values of this layer. But this also happens for MT-SC22EK and MT-SC35EK, which are adsorbed to the upper monolayer. It is possible that the effects of peptide adsorption on lipid lateral diffusion are reduced by membrane cholesterol concentration. This can explain why we cannot see a clear trend for D_{lat} values in these systems, like we do for the others.

3.3.4. Order Parameters

Order parameters (-S_{CD}) are calculated for the hydrocarbon chains of the POPC molecules. This parameter analyzes the orientation of two different axis relative to each other. The definition is presented below:

$$S_{ab} = \frac{1}{2} < 3\cos(\theta_a)\cos(\theta_b) - \delta_{ab} > \qquad a, b = x, y, z \tag{20}$$

Here, δ_{ab} represents the Kronecker delta, while θ_a and θ_b are the angles formed between the chosen axis and the normal orientation of the bilayer. Using a *united atoms* force field, the order parameters for the saturated and unsaturated carbons are obtained via the following equations:

$$-S_{CD}^{Sat} = \frac{2}{3} S_{xx} + \frac{1}{3} S_{yy} \tag{21}$$

$$-S_{CD}^{Unsat} = \frac{1}{4} S_{zz} + \frac{3}{4} S_{yy} + \frac{\sqrt{3}}{2} S_{xy} \tag{22}$$

The value of $-S_{CD}$ can vary between 0.5 and -0.25. The first one means complete order along the bilayer normal, while the second means complete order along the plane of the bilayer. A null value of $-S_{CD}$ indicates an isotropic (random) orientation. Order parameters simulations were performed for the last 100 ns of the trajectory.

Table 8 – Order parameters for the sn-1 and sn-2 chains of cholesterol with their respective percentage of variation.

Systems	**sn-1**	**Variation (%)**	**sn-2**	**Variation (%)**
POPC	0.168 ± 0.040	-	0.119 ± 0.050	-
MT-C34 + POPC	0.165 ± 0.037	-1.79	0.118 ± 0.048	-0.84
MT-SC22EK + POPC	0.176 ± 0.038	4.76	0.124 ± 0.047	4.20
MT-SC29EK + POPC	0.187 ± 0.038	11.31	0.136 ± 0.051	14.29
MT-SC34 + POPC	0.177 ± 0.036	5.36	0.129 ± 0.048	8.40
MT-SC34EK + POPC	0.179 ± 0.039	6.55	0.126 ± 0.046	5.88
MT-SC35EK + POPC	0.177 ± 0.039	5.36	0.127 ± 0.049	6.72
POPC_CHOL_20%	0.254 ± 0.044	-	0.183 ± 0.052	-
MT-C34 + POPC_CHOL_20%	0.269 ± 0.044	5.91	0.197 ± 0.053	7.65
MT-SC22EK + POPC_CHOL_20%	0.274 ± 0.044	7.87	0.196 ± 0.055	7.10
MT-SC29EK + POPC_CHOL_20%	0.265 ± 0.045	4.33	0.189 ± 0.054	3.28
MT-SC34 + POPC_CHOL_20%	0.258 ± 0.042	1.58	0.185 ± 0.055	1.09
MT-SC34EK + POPC_CHOL_20%	0.265 ± 0.043	4.33	0.194 ± 0.055	6.01
MT-SC35EK + POPC_CHOL_20%	0.269 ± 0.043	5.91	0.191 ± 0.056	4.37

POPC_CHOL_50%	0.369 ± 0.062	-	0.270 ± 0.074	-
MT-C34 + POPC_CHOL_50%	0.367 ± 0.061	-0.54	0.267 ± 0.073	-1.11
MT-SC22EK + POPC_CHOL_50%	0.362 ± 0.061	-1.90	0.263 ± 0.072	-2.59
MT-SC29EK + POPC_CHOL_50%	0.373 ± 0.061	1.08	0.271 ± 0.075	0.37
MT-SC34 + POPC_CHOL_50%	0.371 ± 0.062	0.54	0.268 ± 0.074	-0.74
MT-SC34EK + POPC_CHOL_50%	0.365 ± 0.061	-1.08	0.262 ± 0.071	-2.96
MT-SC35EK + POPC_CHOL_50%	0.370 ± 0.060	0.27	0.275 ± 0.076	1.85

Figure 41 shows the $-S_{CD}$ profiles of the systems together with the profiles corresponding to the lone membranes for comparison. Table 8 shows the average $-S_{CD}$ values for the systems and the percentage of the variation compared to the membranes with no peptide. In the POPC systems, we can see that $-S_{CD}$ values increased slightly, except in MT-C34 + POPC, where values decrease very lightly. This increase is probably due to the interaction between the hydrophobic side chains of the amino acids and the hydrophobic chains of POPC. This leads to slight increases in order. The POPC_CHOL_20% systems also show increases in membrane order for sn-1 and sn-2 in every system. These increases are lower in some systems and higher in others and are also more consistent than those of the POPC systems. In the POPC_CHOL_50% systems, the magnitude of the variations is much smaller due to the rigidity of the membrane. Half of the systems show increases in sn-1 order, while the other half show a decrease in this parameter. For sn-2, only the MT-SC29EK and MT-SC35EK systems show an increase in order, while the rest decrease. Increases also tend to be smaller than the decreases. These results point to the fact that peptide adsorption disrupts membrane order in the 50% cholesterol systems, contrary to what occurs in POPC and 20% cholesterol membranes. Even so, the level to which this occurs is very slight. It is possible that cholesterol compromises POPC rearrangement by inhibiting its mobility. This will hinder the changes in membrane conformation necessary to accommodate the peptide. Also, as we know, cholesterol has a beneficial effect on membrane order. We can see this by noticing the -S_{CD} values increase as cholesterol concentration increases.

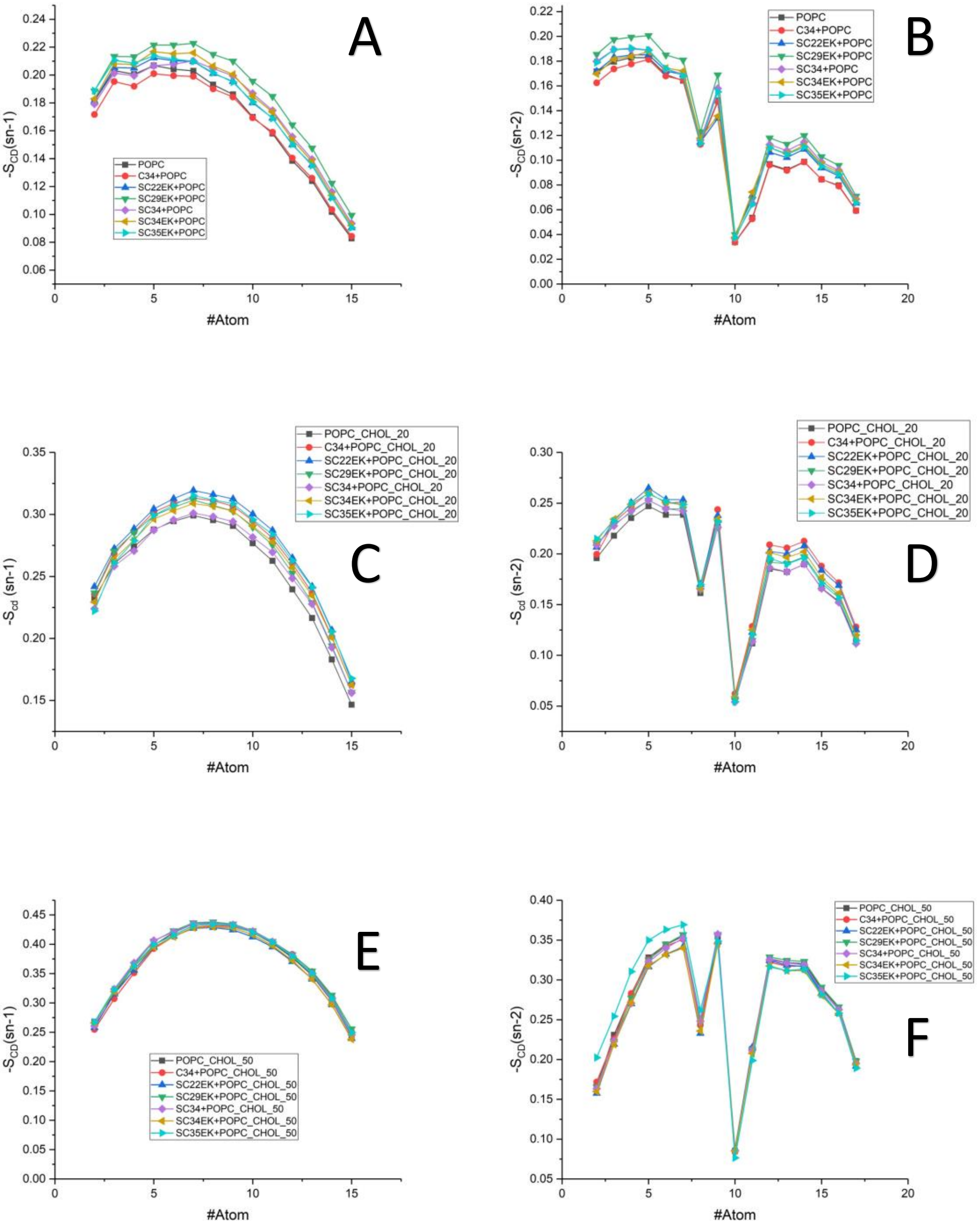


Figure 41 – (A) Order parameters for the sn-1 axis of the POPC membranes. (B) Order parameters for the sn-2 axis of the POPC membranes. (C) Order parameters for the sn-1 axis of the POPC_CHOL_20% membranes.

(D) Order parameters for the sn-2 axis of the POPC_CHOL_20% membranes. (E) Order parameters for the sn-1 axis of the POPC_CHOL_50% membranes. (F) Order parameters for the sn-2 axis of the POPC_CHOL_50% membranes.

3.3.5. POPC and Cholesterol Rotational Dynamics

Rotational dynamics simulations of the membranes were performed using the same method as for the peptides but for different axis. The selected axis for POPC were P-N (P8-N4), sn-1 (C36-C49) and sn-2 (C17-CA1). Only one axis was selected for cholesterol (C5-C16). These axes allow for a good observation of rotational dynamics of the polar head groups and hydrophobic regions of the lipids. The definition of the rotational autocorrelation function was the same as in the previous section of rotational dynamics for peptides.

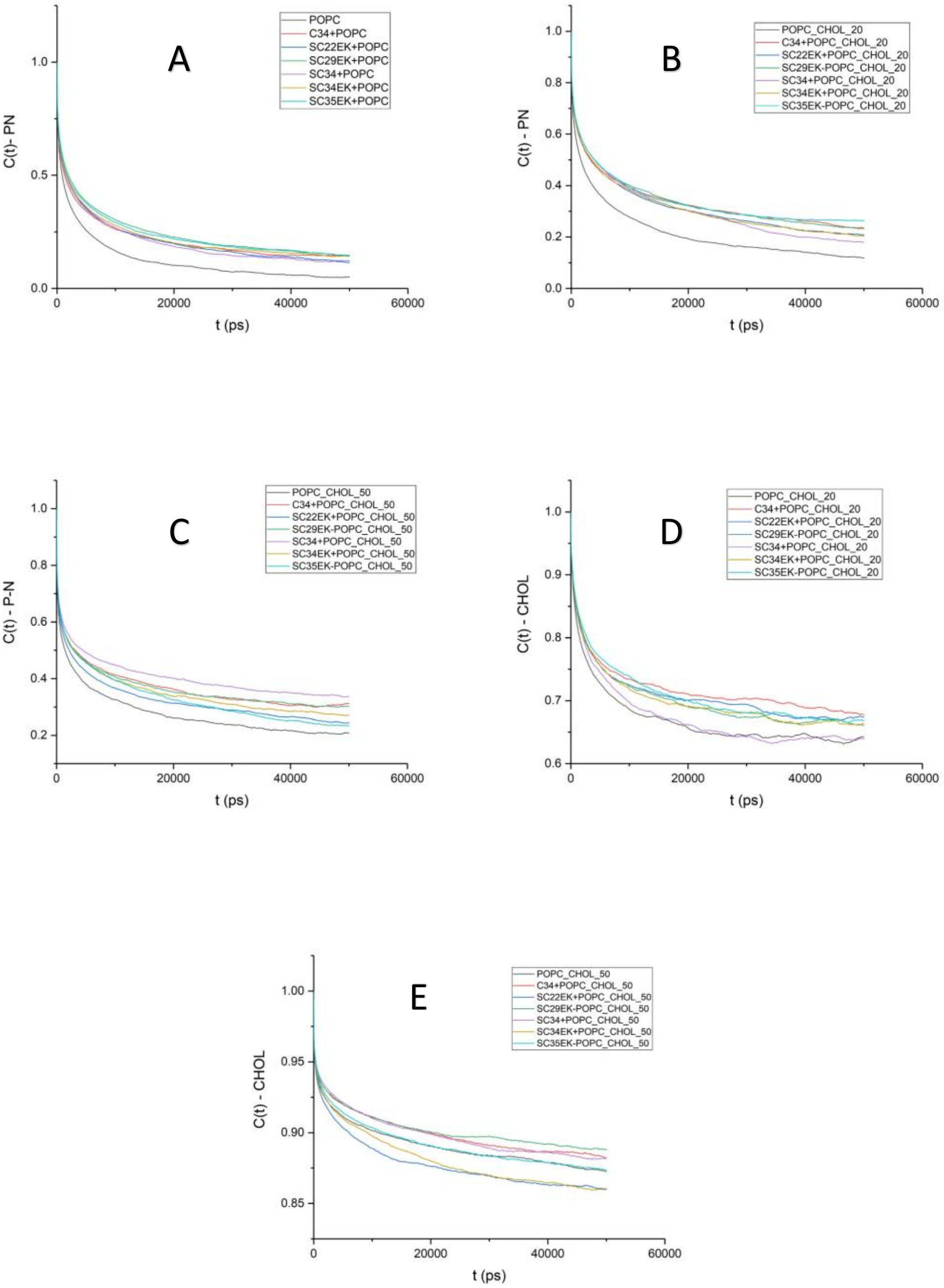


Figure 42 – Rotation autocorrelation function (Rotacf) for the: (A) P8-N4 axis of POPC for the POPC systems; (B) P8-N4 axis of POPC for the POPC_CHOL_20% systems; (C) P8-N4 axis of POPC for the POPC_CHOL_50% systems; (D) C6-C15 axis of cholesterol for the POPC_CHOL_20% systems and (E) C6_C15 axis of cholesterol for the POPC_CHOL_50% systems.

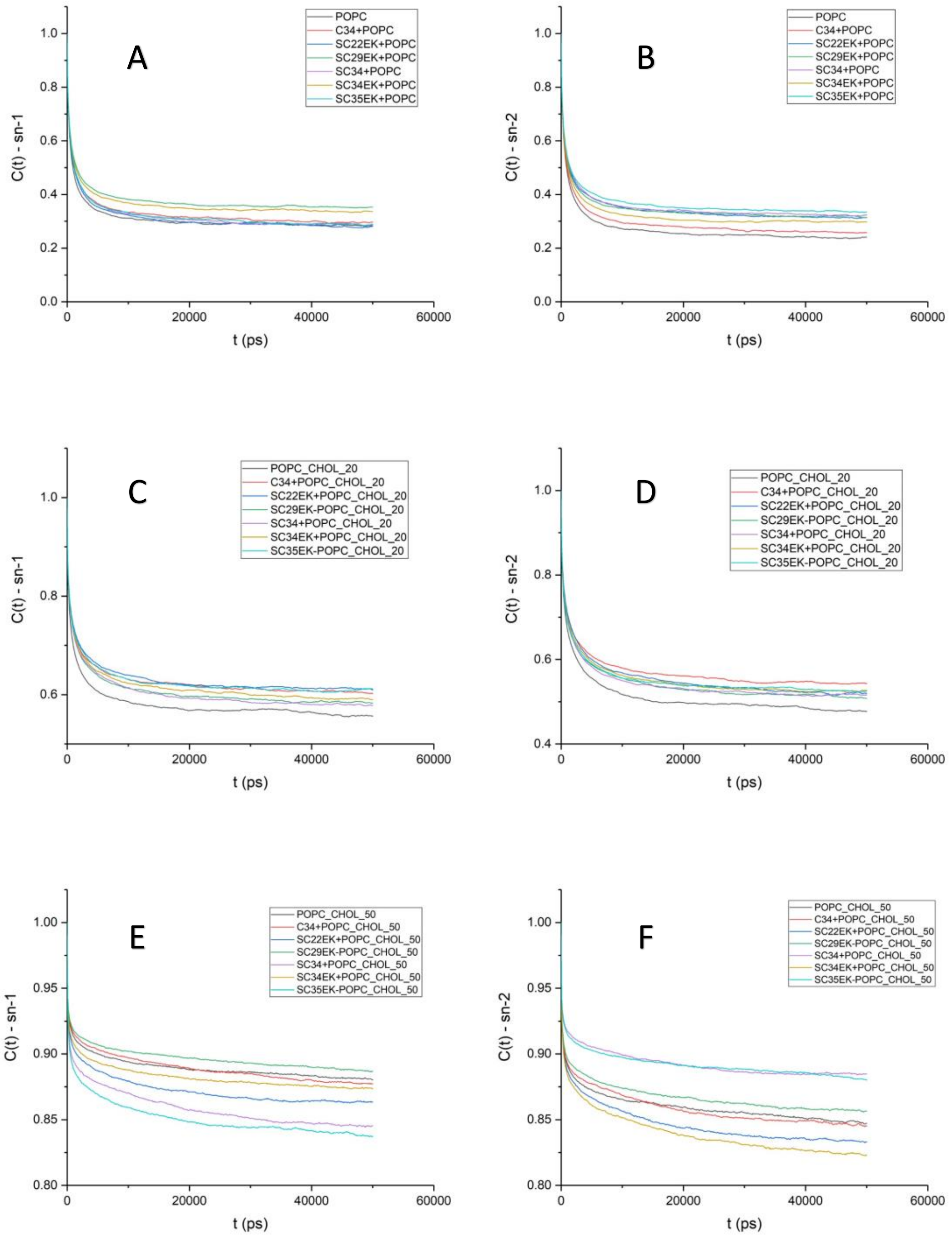


Figure 43 – Rotation autocorrelation function (Rotacf) for the: (A) sn-1 axis of POPC for the POPC systems; (B) sn-2 axis of POPC for the POPC systems; (C) sn-1 axis of POPC for the POPC_CHOL_20% systems; (D) sn-2 axis of POPC for the POPC_CHOL_20% systems. (E) sn-1 axis of POPC for the POPC_CHOL_50% systems and (F) sn-2 axis of POPC for the POPC_CHOL_50% systems.

From the comparison of rotacf values between systems, we can confirm that membrane cholesterol concentration and lipid rotation are inversely proportional. The more cholesterol there is in a membrane, the less membrane lipids will rotate and vice-versa. This can be seen very clearly for all the axes of Figures 42 and 43.

Peptide results showed that peptides in water rotated more freely, displaying values of C(t) close to 0. In membrane systems, peptide adsorption inhibited their rotation leading to a convergence to C(t) values closer to 1. In Figures 42 A-C we see that rotation of the polar head groups of POPC was somewhat affected. Every POPC and POPC-CHOL system had a higher C(t) value for P-N compared to the reference membrane with no peptide adsorbed. This happens because POPC molecules cease interaction with the solvent and with one another to a certain degree, to interact with the peptide. POPC interactions with peptides hinder the movement of the lipids leading to a fewer and slower rotation. The same trend occurs in Figures 43 A-D for the sn-1 and sn-2 axes of POPC and 20% cholesterol systems. Less rotation of the polar head group leads to slower rotation of the hydrophobic tails. Additionally, as we saw previously, the interaction between side chains of amino acids with the hydrophobic tails of POPC increases their order, which leads to less movement and slower dynamics overall.

The values for sn-1 and sn-2 in 50% cholesterol systems are less consistent. Here, rotation is much lower due to the higher cholesterol concentration. This leads to much higher C(t) values compared to the other two membrane systems. In Figure 43 E we see that sn-1 rotation was only slightly slower in the MT-SC29EK system, compared to the regular membrane. MT-SC29EK + POPC_CHOL_50% is also the 50% cholesterol system with the highest increase in sn-1 order, as stated previously. This can explain the slower rotation speeds for this axis in this system. Figure 43 F shows that sn-2 rotation was only significantly slower in the MT-SC34 and MT-SC35EK systems. In the Cα profiles of these peptides we saw that their adsorption was the most superficial of all the peptides in these systems. This leaves the peptides more distant from the membrane where they do not disturb rotational dynamics of the lipids as much.

Figures 42 D and E show the dynamics of the cholesterol molecules in the 20% and 50% cholesterol systems. Peptide adsorption led to a slight decrease in cholesterol dynamics in the 20% cholesterol systems, probably due to the beneficial effect of this process on membrane order. Results of the 50% cholesterol systems vary, with some systems showing increases in the speed of rotational dynamics, while others show decreases. In these systems, adsorption had varying effects on membrane order, this likely explains the variations in the results.

4. Conclusion

This work had the goal of testing the behavior of the MT-SC peptide family and of seeing how the M-T hook structure affects the fusion process. To make better sense of the results, these can be compared to the values obtained from the respective peptides with no M-T hook added (Pronto, 2019).

Tests were run in three different membrane models, POPC, POPC_CHOL_20% and POPC-CHOL_50%. The first membrane is made entirely of POPC, the second has a POPC to cholesterol ratio of 4:1, and the third has POPC and cholesterol in a 1:1 ratio. The first stage of this work was to determine the properties of the membrane with no peptide adsorbed. This was performed for validation and comparison purposes. After the behaviors of the membranes are noted, we can use the results provided from the peptide-membrane systems to infer on the changes the peptides caused. For system equilibration, area per POPC and membrane thickness values were measured. The results showed that the membranes got thicker as cholesterol concentration increased. They also showed that area per POPC was smaller in POPC-CHOL systems. This difference can be attributed to the ordering effect of cholesterol. Cholesterol molecules are packed together in the intermembrane space, promoting the tighter packaging of the POPC molecules at the surface, ultimately reducing the area they occupy.

Parameter density values indicated the monolayers in which the peptides were adsorbed. The distance between the centers of mass of the peptides and membranes, and the Cα profiles were measured to prove the occurrence of peptide adsorption, and to help verify how and when it occurred. The center of mass results showed that adhesion was faster to POPC membranes, followed by the 20% cholesterol membranes and the 50% cholesterol membranes respectively. In turn, the Cα profiles of the peptides showed that adsorption was more superficial in POPC-CHOL systems, and that the depth of the interaction was inversely proportional to cholesterol concentration. Both results point to the fact that adsorption is facilitated by high membrane fluidity. The higher mobility of the lipids in the POPC membrane makes it easier for the membrane to accommodate the peptides. On the other hand, the impairment on mobility caused by cholesterol hampers this process.

Hydrogen bonds between the peptides and POPC decrease as cholesterol concentration increases. This does not favor the use of these peptides as fusion inhibitors since they need to form stable interactions with the membrane to act as a reservoir. The values obtained for the M-T hook peptides in the 50% cholesterol systems (except MT-SC22EK) were lower than the values of the parent peptides under the same conditions (Pronto, 2019). According to these results, it is possible that the M-T hook disrupts the inhibitory activity of longer peptides, but benefits the activity of smaller ones, such as MT-SC22EK. Additionally, interaction energy values between the peptides and POPC significantly decrease in the 50% cholesterol systems, after remaining relatively unchanged in the POPC and 20% cholesterol systems. Once again, there is a noticeable decrease in interaction energy compared with the parent peptides, in the 50% cholesterol systems. However, the interaction energy and hydrogen bond values in the POPC systems are higher for the M-T hook peptides. It is possible that the hook structure is optimal for fixation to highly fluid membranes that allow its insertion below the phosphorus axis of POPC. Less fluid membranes will impair or even block the insertion of the hook region. From the observation of the Cα graphs we saw that peptides appear to be more deeply accommodated in the POPC membrane and that adsorption to the 20% and 50% cholesterol membranes is more superficial. Therefore, the connection with the POPC membrane appears more stable. It could be expected that the higher stability of this connection would cost some of the peptide’s mobility, but this does not seem to be the case. As we saw, peptide rotational dynamics and lateral diffusion values remained high in POPC systems. This means that the peptides can benefit from the higher stability of the connection without sacrificing too much mobility. Peptides in the more rigid POPC-CHOL membranes formed more unstable connections and had their rotation and lateral diffusion reduced (except for MT-SC35EK for the reasons mentioned above). If the membrane enabled the penetration of the hook structure beyond the phosphorus axis of POPC, things could be different. Rotational dynamics and lateral diffusion would likely still be heavily impaired due to fixation by cholesterol molecules, but the connection would be more stable.

MT-C34 and MT-SC22EK in the 50% cholesterol system have the most negative ΔG values of all M-T hook peptides, followed by MT-SC34, which indicates that they have the most spontaneous interaction with the membrane. We saw that MT-SC35EK was in a more vertical position than the other peptides, which could facilitate the interaction with the NHR region of gp41. Hence, this verticality may have been very advantageous to its inhibitory activity, in addition to its size. However, MT-SC35EK has the least negative value of ΔG, which can compromise its function as a fusion inhibitor, due to the possible low spontaneity of the interaction with the membrane. Also, recognizing the effect that size has on inhibitory activity, MT-SC22EK may not be an optimal fusion inhibitor. Higher concentrations of this peptide would be necessary to achieve the same effects of lower concentrations of other peptides, such as MT-C34 or MT-SC34, despite having great characteristics. According to Mavioso and colleagues (Mavioso et al., 2017), the magnitude of the ΔG values of all peptides provides a good balance between affinity to the membrane and to the solution. This makes it possible for a reservoir of peptide to be established at the surface of the membrane, while also ensuring a certain degree of reversibility.

From the results we also saw that the thickness of the POPC and 20% cholesterol membranes rose, while the thickness of the 50% cholesterol membrane stayed the same or decreased slightly. Rotational dynamics of POPC and cholesterol showed a slight decrease in most systems. So did lateral diffusion in the monolayers where the peptide was adsorbed. Even so, peptide adsorption benefited POPC and 20% cholesterol membrane order, while 50% cholesterol membranes were relatively unaffected by it. We can conclude from this that the 50% cholesterol membranes can maintain their overall properties after peptide adsorption. The same cannot be said for the intermediate membrane of 20% cholesterol or the POPC membrane. Even so, in these membranes, peptide adsorption benefits factors like membrane order by a wide margin, while not sacrificing too much lateral diffusion or rotation.

This set of results indicate that the M-T hook structure is beneficial to the interaction of the peptides in more fluid membranes like the POPC and 20% cholesterol membranes. In particular, the fluidity of the 20% cholesterol membrane is closer to the fluidity of a common mammalian plasma membrane. There are no apparent benefits for this structure in the interaction with more rigid membranes, seeing that interaction energy and hydrogen bond values decrease in comparison to the analogous peptides with no M-T hook. The MT-SC peptides with the most promising activity in this study were MT-C34 and MT-SC34, with MT-SC34EK being a close third. MT-SC22EK had some good parameter values, but its size would likely increase its IC_{50} value. It is worth noting that the main purpose of the M-T hook is not to benefit the interaction of the peptides with membranes, with this fact being a positive side-effect. This structure provides considerable stability to the 6-HB formed between the peptide and the NHR region of gp41, increases the affinity of the peptide with that same region and gives the peptides a higher genetic barrier against mutations by HIV (Chong et al., 2012a; Chong, Qiu, et al., 2014; Chong, Yao, et al., 2014). This gives the peptides better characteristics that benefit their inhibitory activity, allow them to stay active for longer periods of time and maintain their efficacy for longer.

To further improve on this study, experiments can be performed on longer time scales. It is possible that the POPC-CHOL systems were not yet in equilibrium and that given enough time, they would keep evolving. Specifically, the 50% cholesterol system with MT-SC35EK is likely not in equilibrium yet and would require further studies. To reduce computational costs, instead of raising the number of steps for its simulation, experiments could be made by changing its orientation and placing it closer to the membrane. Also, experiments can be attempted in membranes with different cholesterol concentrations to better understand the impact cholesterol has on peptide adsorption. More accurate membrane models, mimicking the HIV membrane, can also be designed to be used in the simulations. Lastly, placing peptides further from the membrane may help understand what mechanisms go into recognition of the membrane surfaces, while placing them closer may be an effective strategy to reduce simulation times needed to study the adsorption process.

5. References

A. Pantophlet, R. (2010). Antibody Epitope Exposure and Neutralization of HIV-1. *Current Pharmaceutical Design*, *16*(33), 3729–3743. https://doi.org/10.2174/138161210794079182

Ackerman, G. D., & Feigenson, W. G. (2015). Lipid Bilayers: Clusters, Domains and Phases David. *Essays in Biochemisrty*, *57*(1), 33–42. https://doi.org/10.1042/bse0570033

Alder, B. J., & Wainwright, T. E. (1957). Phase transition for a hard sphere system. *The Journal of Chemical Physics*, *27*(5), 1208–1209. https://doi.org/10.1063/1.1743957

Allan, J. S., Coligan, J. E., Barin, F., Mclane, M. F., Sodroski, J. G., Rosen, C. A., Haseltine, W. A., Lee, T. H., & Essex, M. (1985). Major glycoprotein antigens that induce antibodies in AIDS patients are encoded by HTLV-III. *Science*, *228*(4703), 1091–1094. https://doi.org/10.1126/science.2986290

Ball, W. B., Neff, J. K., & Gohil, V. M. (2018). The role of non-bilayer phospholipids in mitochondrial structure and function. *FEBS Letters*, *592*(8), 1273–1290. https://doi.org/10.1002/1873-3468.12887

Barré-Sinoussi, F., Chermann, J. C., Rey, F., Nugeyre, M. T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vézinet-Brun, F., Rouzioux, C., Rozenbaum, W., & Montagnier, L. (2004). Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Revista de Investigacion Clinica*, *56*(2), 126–129.

Bell, N. M., & Lever, A. M. L. (2013). HIV Gag polyprotein: Processing and early viral particle assembly. *Trends in Microbiology*, *21*(3), 136–144. https://doi.org/10.1016/j.tim.2012.11.006

Berendsen, H. J. C., Postma, J. P. M., Van Gunsteren, W. F., Dinola, A., & Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, *81*(8), 3684–3690. https://doi.org/10.1063/1.448118

Berendsen, H. J. C., Postma, J. P. M., Van Gunsteren, W. F., & Hermans, J. (1981). Interaction Models for Water in Relation to Protein Hydration. In *The Jerusalem Symposia on Quantum Chemistry and Biochemistry* (pp. 331–342). https://doi.org/https://doi.org/10.1007/978-94-015-7658-1_21

Berthet-Colominas, C., Monaco, S., Novelli, A., Sibaï, G., Mallet, F., & Cusack, S. (1999). Head-to-tail dimers and interdomain flexibility revealed by the crystal structure of HIV-1 capsid protein (p24) complexed with a monoclonal antibody Fab. *EMBO Journal*, *18*(5), 1124–1136. https://doi.org/10.1093/emboj/18.5.1124

Beziau, A., Brand, D., & Piver, E. (2020). The role of phosphatidylinositol phosphate kinases during viral infection. *Viruses*, *12*(10), 1–16. https://doi.org/10.3390/v12101124

Burton, D. R., & Montefiori, D. C. (1997). The antibody response in HIV-1 infection. *AIDS (London, England)*, *11 Suppl A*, S87—98.

http://europepmc.org/abstract/MED/9451972

Caffrey, M. (2011). HIV envelope: Challenges and opportunities for development of entry inhibitors. *Trends in Microbiology*, *19*(4), 191–197. https://doi.org/10.1016/j.tim.2011.02.001

Calzada, E., Onguka, O., & Claypool, S. M. (2016). Phosphatidylethanolamine Metabolism in Health and Disease. *International Review of Cell and Molecular Biology*, *321*(1), 29–88. https://doi.org/10.1016/bs.ircmb.2015.10.001

Chen, C., Depa, P., Sakai, V. G., Maranas, J. K., Lynn, J. W., Peral, I., & Copley, J. R. D. (2006). A comparison of united atom, explicit atom, and coarse-grained simulation models for poly(ethylene oxide). *Journal of Chemical Physics*, *124*(23). https://doi.org/10.1063/1.2204035

Chong, H., Qiu, Z., Sun, J., Qiao, Y., Li, X., & He, Y. (2014). Two M-T hook residues greatly improve the antiviral activity and resistance profile of the HIV-1 fusion inhibitor SC29EK. *Retrovirology*, *11*(1), 1–10. https://doi.org/10.1186/1742-4690-11-40

Chong, H., Yao, X., Qiu, Z., Sun, J., Qiao, Y., Zhang, M., Wang, M., Cui, S., & He, Y. (2014). The M-T hook structure increases the potency of HIV-1 fusion inhibitor sifuvirtide and overcomes drug resistance. *Journal of Antimicrobial Chemotherapy*, *69*(10), 2759–2769. https://doi.org/10.1093/jac/dku183

Chong, H., Yao, X., Sun, J., Qiu, Z., Zhang, M., Waltersperger, S., Wang, M., Cui, S., & He, Y. (2012a). The M-T hook structure is critical for design of HIV-1 fusion inhibitors. *Journal of Biological Chemistry*, *287*(41), 34558–34568. https://doi.org/10.1074/jbc.M112.390393

Chong, H., Yao, X., Sun, J., Qiu, Z., Zhang, M., Waltersperger, S., Wang, M., Cui, S., & He, Y. (2012b). The M-T hook structure is critical for design of HIV-1 fusion inhibitors. *Journal of Biological Chemistry*, *287*(41), 34558–34568. https://doi.org/10.1074/jbc.M112.390393

Clavel, F., DENISE GUETARD, FRAN9OISE BRUN-VEZINET, SOPHIE CHAMARET, MARIE-ANNE REY, M. 0. SANTOS-FERREIRA, ANNE G. LAURENT, CHARLuES DAUGUET, CHRISTINE KATLAMA, CHRISTINE RouzIoux, DAVID KLATZMANN, J. L. CHAMPALIMAUD, & Luc MONTAGNIER. (1986). Isolation of a New Human Retrovirus African Patients with AIDS. *Science*, *233*(December 1985), 343–346.

Colin, L., & Van Lint, C. (2009). Molecular control of HIV-1 postintegration latency: Implications for the development of new therapeutic strategies. *Retrovirology*, *6*(3), 1–29. https://doi.org/10.1186/1742-4690-6-111

Crépin, M. . L. R. . C. J. C. . P. P. . M. H. . M. L. (1984). Sequences Related to Mouse Mammary Tumor Virus Genome in Tumor Cells and Lymphocytes from Patients With Breast Cancer. *Biochemical and Biophysical Research Communications*, *118*(1), 324–331.

Cui, Z., & Houweling, M. (2002). Phosphatidylcholine and cell death. *Biochimica et Biophysica Acta*, *1585*, 87–96. https://doi.org/https://doi.org/10.1016/S1388-1981(02)00328-1

Darden, T., York, D., & Pedersen, L. (1993). Particle mesh Ewald: An N·log(N) method for

Ewald sums in large systems. *The Journal of Chemical Physics*, *98*(12), 10089–10092. https://doi.org/10.1063/1.464397

Das, S. R., & Jameel, S. (2005). Biology of the HIV Nef protein. *Indian Journal of Medical Research*, *121*(4), 315–332.

Ding, W. (2017). *Molecular Dynamics Simulation of Biomembrane Systems*.

Duffield, P. A., Holt, R. R. D., & Smith, J. R. (2011). Force fields and molecular dynamics simulations. *Gonzáles, A. M*, 169–200.

Dufourc, E. J. (2008). Sterols and membrane dynamics. *Journal of Chemical Biology*, *1*(1–4), 63–77. https://doi.org/10.1007/s12154-008-0010-6

Elson, L. E., Fried, E., Dolbow, E. J., & Genin, M. G. (2010). Phase separation in biological membranes: integration of theory and experiment. *Annual Review of Biophysics*, *39*, 207–226. https://doi.org/10.1146/annurev.biophys.093008.131238

Engelman, A., & Cherepanov, P. (2012). The structural biology of HIV-1: Mechanistic and therapeutic insights. *Nature Reviews Microbiology*, *10*(4), 279–290. https://doi.org/10.1038/nrmicro2747

Erdinc Sezgin, Ilya Levental, Satyajit Mayor, C. E. (2017). The mystery of membrane organization: composition, regulation and physiological relevance of lipid rafts. *Nat Rev Mol Cell Biol*, *18*(6), 361–374. https://doi.org/10.1038/nrm.2017.16.The

FDA-Approved HIV Medicines | NIH. (n.d.). Retrieved February 12, 2023, from https://hivinfo.nih.gov/understanding-hiv/fact-sheets/fda-approved-hiv-medicines

Frenkel, D., & Smith, B. (2002). *Understanding Molecular simulation* (p. 658).

Furse, S. (2017). Is phosphatidylglycerol essential for terrestrial life? *Journal of Chemical Biology*, *10*(1), 1–9. https://doi.org/10.1007/s12154-016-0159-3

Gallo, R. C., & Montagnier, L. (2003). The Discovery of HIV as the Cause of AIDS. *New England Journal of Medicine*, *349*(24), 2283–2285. https://doi.org/10.1056/nejmp038194

Gallo, S. A., Finnegan, C. M., Viard, M., Raviv, Y., Dimitrov, A., Rawat, S. S., Puri, A., Durell, S., & Blumenthal, R. (2003). The HIV Env-mediated fusion reaction. *Biochimica et Biophysica Acta - Biomembranes*, *1614*(1), 36–50. https://doi.org/10.1016/S0005-2736(03)00161-5

Ganser-Pornillos, B. K., Yeager, M., & Pornillos, O. (2012). Assembly and architecture of HIV. *Advances in Experimental Medicine and Biology*, *726*, 441–465. https://doi.org/10.1007/978-1-4614-0980-9_20

Ganser-Pornillos, B. K., Yeager, M., & Sundquist, W. I. (2008). The structural biology of HIV assembly. *Current Opinion in Structural Biology*, *18*(2), 203–217. https://doi.org/10.1016/j.sbi.2008.02.001

Goncalves, M. D., Hopkins, B. D., & Cantley, L. C. (2018). Phosphatidylinositol 3-Kinase, Growth Disorders, and Cancer. *New England Journal of Medicine*, *379*(21), 2052–2062. https://doi.org/10.1056/nejmra1704560

Goodsell, D. (2002). *PDB-101: Molecule of the Month: HIV Reverse Transcriptase*. Pdb101.Rcsb.Org. https://doi.org/http://doi.org/10.2210/rcsb_pdb/mom_2002_9

Goodsell, D. (2014). *Molecule of the Month: HIV Envelope Glycoprotein*. Pdb101.Rcsb.Org. https://pdb101.rcsb.org/motm/169

Gottlieb, M. S., Schanker, H. M., Fan, P. T., Saxon, A., Weisman, J. D., & Pozalski, I. (1981). Pneumocystis pneumonia—Los Angeles. *Mmwr*, *30*(21), 250–252.

Greaves, R. J., Schlecht, K. D., & Jones, R. F. (1992). Gibbs free energy: The criteria for spontaneity. *Journal of Chemical Education*, *69*(5), 417–419. https://doi.org/10.1021/ed069p417

Greenberg, M. L., & Cammack, N. (2004). Resistance to enfuvirtide, the first HIV fusion inhibitor. *Journal of Antimicrobial Chemotherapy*, *54*(2), 333–340. https://doi.org/10.1093/jac/dkh330

Grinnell, S., Yoshida, K., & Jasin, H. E. (2005). Responses of lymphocytes of patients with rheumatoid arthritis to IgG modified by oxygen radicals or peroxynitrite. *Arthritis and Rheumatism*, *52*(1), 80–83. https://doi.org/10.1002/art.20760

GROMACS. (2020). *THE GLOBAL MD ALGORITHM*. https://manual.gromacs.org/documentation/2020-beta2/reference-manual/algorithms/molecular-dynamics.html#update

GROMACS. (2021a). *Force field*. https://manual.gromacs.org/documentation/2021/reference-manual/functions/force-field.html

GROMACS. (2021b). *GROMACS Documentation - Release 2021*. 2–607. http://manual.gromacs.org/documentation/2019/manual-2019.pdf

Harder, T., Scheiffele, P., Verkade, P., & Simons, K. (1998). Lipid domain structure of the plasma membrane revealed by patching of membrane components. *Journal of Cell Biology*, *141*(4), 929–942. https://doi.org/10.1083/jcb.141.4.929

Harvey, S. C., Tan, R. K., & Iii, T. E. C. (1997). *The Flying Ice Cube : Velocity Rescaling in Molecular Dynamics Leads to Violation of Energy Equipartition*. *19*(7), 726–740.

Hess, B., Bekker, H., Berendsen, H. J. C., & Fraaije, J. G. E. M. (1997). LINCS: A Linear Constraint Solver for molecular simulations. *Journal of Computational Chemistry*, *18*(12), 1463–1472. https://doi.org/10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H

Hilditch, L., & Towers, G. J. (2014). A model for cofactor use during HIV-1 reverse transcription and nuclear entry. *Current Opinion in Virology*, *4*, 32–36. https://doi.org/10.1016/j.coviro.2013.11.003

Hill, M., Tachedjian, G., & Mak, J. (2005). The Packaging and Maturation of the HIV-1 Pol Proteins. *Current HIV Research*, *3*(1), 73–85. https://doi.org/10.2174/1570162052772942

Hladik, F., & McElrath, M. J. (2008). Setting the stage: Host invasion by HIV. *Nature Reviews Immunology*, *8*(6), 447–457. https://doi.org/10.1038/nri2302

Hockney, R. W., Goel, S. P., & Eastwood, J. W. (1974). Quiet high-resolution computer models of a plasma. *Journal of Computational Physics*, *14*(2), 148–158. https://doi.org/10.1016/0021-9991(74)90010-2

Hofsäß, C., Lindahl, E., & Edholm, O. (2003). Molecular dynamics simulations of phospholipid bilayers with cholesterol. *Biophysical Journal*, *84*(4), 2192–2206. https://doi.org/10.1016/S0006-3495(03)75025-5

Hollingsworth, S. A., & Dror, R. O. (2018). Molecular Dynamics Simulation for All. *Neuron*, *99*(6), 1129–1143. https://doi.org/10.1016/j.neuron.2018.08.011

Houtkooper, R. H., & Vaz, F. M. (2008). Cardiolipin, the heart of mitochondrial metabolism. *Cellular and Molecular Life Sciences*, *65*(16), 2493–2506. https://doi.org/10.1007/s00018-008-8030-5

Hu, W., & Hughes, S. H. (2012). HIV-1 Reverse Transcription. *Cold Spring Harbor Perspectives in Medicine*.

Hug, S. (2013). Classical molecular dynamics in a nutshell. *Methods in Molecular Biology*, *924*, 127–152. https://doi.org/10.1007/978-1-62703-17-5_6

Izumi, K., Kodama, E., Shimura, K., Sakagami, Y., Watanabe, K., Ito, S., Watabe, T., Terakawa, Y., Nishikawa, H., Sarafianos, S. G., Kitaura, K., Oishi, S., Fujii, N., & Matsuoka, M. (2009). Design of peptide-based inhibitors for human immunodeficiency virus type 1 strains resistant to T-20. *Journal of Biological Chemistry*, *284*(8), 4914–4920. https://doi.org/10.1074/jbc.M807169200

James, M., Wyatt, R., Desjardins, E., Kwong, P. D., Moore, J. P., Sodroski, J., & V, C. (1998). Analysis of the Interaction of Antibodies with a Conserved , Enzymatically Deglycosylated Core of the HIV Type 1 Envelope Glycoprotein 120. In *AIDS Research and Human Retroviruses* (Vol. 14, Issue 3, pp. 191–198).

Jury, E. C., & Kabouridis, P. S. (2004). T-lymphocyte signalling in systemic lupus erythematosus: A lipid raft perspective. *Lupus*, *13*(6), 413–422. https://doi.org/10.1191/0961203304lu1045rr

Kabsch, W., & Sander, C. (1983). Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, *22*(12), 2577–2637. https://doi.org/10.1002/bip.360221211

Karn, J., & Stoltzfus, C. M. (2012). Transcriptional and Posttranscriptional Regulation of HIV-1 Gene Expression. *Cold Spring Harbor Perspectives in Medicine*, 1–18. https://doi.org/10.1101/cshperspect.a006916

Karplus, M., & McCammon, J. A. (2002). Molecular dynamics simulations of biomolecules. *Nature Structural Biology*, *9*(9), 646–652. https://doi.org/10.1038/nsb0902-646

Kim, H. Y., Huang, B. X., & Spector, A. A. (2014). Phosphatidylserine in the brain: Metabolism and function. *Progress in Lipid Research*, *56*(1), 1–18. https://doi.org/10.1016/j.plipres.2014.06.002

Kirchhoff, F. (2016). Encyclopedia of AIDS. *Encyclopedia of AIDS*, *June*. https://doi.org/10.1007/978-1-4614-9610-6

Kleinpeter, A. B., & Freed, E. O. (2020). HIV-1 maturation: Lessons learned from inhibitors. *Viruses*, *12*(9). https://doi.org/10.3390/v12090940

Kulkarni, R., Wiemer, E. A. C., & Chang, W. (2022). Role of Lipid Rafts in Pathogen-Host Interaction - A Mini Review. *Frontiers in Immunology*, *12*(January), 1–11. https://doi.org/10.3389/fimmu.2021.815020

Kwong, P. D., Wyatt, R., Robinson, J., Sweet, R. W., Sodroski, J., & Hendrickson, W. A.

(1998). Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature, 393*(6686), 648–659. https://doi.org/10.1038/31405

Levental, I., Levental, K. R., & Heberle, F. A. (2020). Lipid Rafts: Controversies Resolved, Mysteries Remain. *Trends in Cell Biology*, *30*(5), 341–353. https://doi.org/10.1016/j.tcb.2020.01.009

Levy, J.; Hoffman, A.; Kramer, S.; Landis, J.; Shimabukuro, J.; Oshiro, L. (1984). Isolation of Lymphocytopathic Retroviruses from San Francisco Patients with AIDS. *Science, 225*(4664), 840–842. https://doi.org/10.1126/science.6206563

Li, G., & Clercq, E. De. (2016). HIV Genome-Wide Protein Associations: a Review of 30 Years of Research. *American Society for Microbiology*, *80*(3). https://doi.org/10.1128/MMBR.00065-15.Address

Lingwood, D., & Simons, K. (2010). Lipid rafts as a membrane-organizing principle. *Science, 327*(5961), 46–50. https://doi.org/10.1126/science.1174621

Liu, S., Lu, H., Niu, J., Xu, Y., Wu, S., & Jiang, S. (2005). Different from the HIV fusion inhibitor C34, the anti-HIV drug fuzeon (T-20) inhibits HIV-1 entry by targeting multiple sites in gp41 and gp120. *Journal of Biological Chemistry*, *280*(12), 11259–11273. https://doi.org/10.1074/jbc.M411141200

Lubow, J., & Collins, K. L. (2020). Vpr Is a VIP: HIV Vpr and infected macrophages promote viral pathogenesis. *Viruses*, *12*(8). https://doi.org/10.3390/v12080809

Mandell, M. J. (1976). On the properties of a periodic fluid. In *Journal of Statistical Physics* (Vol. 15, Issue 4, pp. 299–305). https://doi.org/10.1007/BF01023055

Mavioso, I. C. V. C., de Andrade, V. C. R., Palace Carvalho, A. J., & Martins do Canto, A. M. T. (2017). Molecular dynamics simulations of T-2410 and T-2429 HIV fusion inhibitors interacting with model membranes: Insight into peptide behavior, structure and dynamics. *Biophysical Chemistry*, *228*(June), 69–80. https://doi.org/10.1016/j.bpc.2017.06.012

Mayor, S., & Pagano, R. E. (2007). Pathways of clathrin-independent endocytosis. *Nature Reviews Molecular Cell Biology, 8*(8), 603–612. https://doi.org/10.1038/nrm2216

Mayor, S., & Riezman, H. (2004). Sorting GPI-anchored proteins. *Nature Reviews Molecular Cell Biology, 5*(2), 110–120. https://doi.org/10.1038/nrm1309

McCammon, J. A., Gelin, B. R., & Karplus, M. (1977). Dynamics of folded proteins. *Nature, 267*(1), 585–590. https://doi.org/https://doi.org/10.1038/267585a0

McElrath, M. J., De Rosa, S. C., Moodie, Z., Dubey, S., Kierstead, L., Janes, H., Defawe, O. D., Carter, D. K., Hural, J., Akondy, R., Buchbinder, S. P., Robertson, M. N., Mehrotra, D. V., Self, S. G., Corey, L., Shiver, J. W., & Casimiro, D. R. (2008). HIV-1 vaccine-induced immunity in the test-of-concept Step Study: a case-cohort analysis. *The Lancet*, *372*(9653), 1894–1905. https://doi.org/10.1016/S0140-6736(08)61592-5

Meng, B., Ip, N. C. Y., Prestwood, L. J., Abbink, T. E. M., & Lever, A. M. L. (2015). Evidence that the endosomal sorting complex required for transport-II (ESCRT-II) is required for efficient human immunodeficiency virus-1 (HIV-1) production. *Retrovirology*, *12*(1), 1–15. https://doi.org/10.1186/s12977-015-0197-x

Miettinen, M. (2010). *Computational Modeling of Cationic Lipid Bilayers in Saline Solutions*.

Miller, R. H., & Sarver, N. (2012). Therapeutic Targets. *Therapeutic Targets*, *1*(5), 479–485. https://doi.org/10.1002/9781118185537

Miyamoto, F., & Kodama, E. N. (2012). Novel HIV-1 fusion inhibition peptides: Designing the next generation of drugs. *Antiviral Chemistry and Chemotherapy*, *22*(4), 151–158. https://doi.org/10.3851/IMP1930

Montagnier, L. (2002). Historical essay: A history of HIV discovery. *Science*, *298*(5599), 1727–1728. https://doi.org/10.1126/science.1079027

Monticelli, L., & Tieleman, D. P. (2013). Force fields for classical molecular dynamics. *Methods in Molecular Biology*, *924*, 197–213. https://doi.org/10.1007/978-1-62703-17-5_8

Naito, T., Izumi, K., Kodama, E., Nishikawa, H., Watanabe, K., Sakagami, Y., Kajiwara, K., Sarafianos, S. G., Oishi, S., Fujii, N., & Matsuoka, M. (2009). SC29EK, a peptide fusion inhibitor with enhanced α-helicity, inhibits replication of human immunodeficiency virus type 1 mutants resistant to enfuvirtide. *Antimicrobial Agents and Chemotherapy*, *53*(3), 1013–1018. https://doi.org/10.1128/AAC.01211-08

National Human Genome Research Institute. (2023). *Cell Membrane (Plamsa Membrane)*. https://www.genome.gov/genetics-glossary/Cell-Membrane

Ohlmann, T., Mengardi, C., & López-Lastra, M. (2014). Translation initiation of the HIV-1 mRNA. *Translation*, *2*(2), e960242. https://doi.org/10.4161/2169074x.2014.960242

Okata, A., Nakamura, M., Nameki, D., Kodama, E., Uchiyama, S., Nakamura, S., Nakano, H., Tamamura, H., Kobayashi, Y., Matsuoka, M., & Fuijii, N. (2002). Promising anti-HIV-1 peptide was developed. *Angewandte Chemie*, *41*(16), 2937–2940.

Ott, M., Geyer, M., & Zhou, Q. (2011). The control of HIV transcription: Keeping RNA polymerase II on track. *Cell Host and Microbe*, *10*(5), 426–435. https://doi.org/10.1016/j.chom.2011.11.002

Paradies, G., Paradies, V., Ruggiero, F. M., & Petrosillo, G. (2019). Role of cardiolipin in mitochondrial function and dynamics in health and disease: Molecular and pharmacological aspects. *Cells*, *8*(7). https://doi.org/10.3390/cells8070728

Peterlin, B. M., & Price, D. H. (2006). Controlling the Elongation Phase of Transcription with P-TEFb. *Molecular Cell*, *23*(3), 297–305. https://doi.org/10.1016/j.molcel.2006.06.014

Poignard, P., Saphire, E. O., Parren, P. W. H. I., & Burton, D. R. (2001). GP120: Biological aspects of structural features. *Annual Review of Immunology*, *19*, 253–274. https://doi.org/10.1146/annurev.immunol.19.1.253

Pronto, C. (2019). *Interação de Péptidos inibidores de fusão do HIV com modelos de membranas biológicas: um estudo por dinâmica molecular*. Universidade de Évora.

Qadir, M. I., & Malik, S. A. (2010). HIV fusion inhibitors. *Reviews in Medical Virology*, *20*(1), 23–33. https://doi.org/10.1002/rmv.631

Razooky, B. S., & Weinberger, L. S. (2011). Mapping the architecture of the HIV-1 Tat circuit: A decision-making circuit that lacks bistability and exploits stochastic noise.

Methods, *53*(1), 68–77. https://doi.org/10.1016/j.ymeth.2010.12.006

Richter, S., Frasson, I., & Palu, G. (2008). Strategies for Inhibiting Function of HIV-1 Accessory Proteins: A Necessary Route to AIDS Therapy? *Current Medicinal Chemistry*, *16*(3), 267–286. https://doi.org/10.2174/092986709787002646

Riek, R. P., & Graham, R. M. (2011). The elusive π-helix. *Journal of Structural Biology*, *173*(1), 153–160. https://doi.org/10.1016/j.jsb.2010.09.001

Ripa, I., Andreu, S., López-Guerrero, J. A., & Bello-Morales, R. (2021). Membrane Rafts: Portals for Viral Entry. *Frontiers in Microbiology*, *12*(February). https://doi.org/10.3389/fmicb.2021.631274

Rivera, J., & Arudchandran, R. (2001). A Perspective : Regulation of IgE Receptor-Mediated Mast Cell Responses by a Signaling Complex. *Allergy*, *20892*, 137–141.

Robert C. Gallo, Syed Z. Salahuddin, Mikulas Popovic, Gene M. Shearer, M., Kaplan, Barton F. Haynes, Thomas J. Palker, Robert Redfield, James Oleske, B. S., & Gilbert White, P. F. and P. D. M. (2019). Frequent Detection and Isolation of Cytopathic Retroviruses (HTLV-III) from Patients with AIDS and at Risk for AIDS. *Science*, *688*(38).

Rosen, C. A. (1991). Tat and Rev: positive modulators of human immunodeficiency virus gene expression. *Gene Expression*, *1*(2), 85–90.

Ross, M. J., & Klotman, P. E. (2008). Pathogenesis of HIV-Associated nephropathy. In *Molecular and Genetic Basis of Renal Disease* (First Edit). Elsevier Inc. https://doi.org/10.1016/B978-1-4160-0252-9.50027-6

Schmidt, S. (2018). The Discovery of HIV-1. *Nature*, *1981*(Milestone 1), 4. https://www.nature.com/articles/d42859-018-00003-x

Schuler, L. D., Daura, X., & Van Gunsteren, W. F. (2001). An improved GROMOS96 force field for aliphatic hydrocarbons in the condensed phase. *Journal of Computational Chemistry*, *22*(11), 1205–1218. https://doi.org/10.1002/jcc.1078

Shaw, G. M., & Hunter, E. (2012). HIV transmission. *Cold Spring Harbor Perspectives in Medicine*, *2*(11), 1–24. https://doi.org/10.1101/cshperspect.a006965

Simons, K., & Toomre, D. (2000a). Lipid rafts and signal transduction. *Nature Reviews Molecular Cell Biology*, *1*(1), 31–39. https://doi.org/10.1038/35036052

Simons, K., & Toomre, D. (2000b). Lipid Rafts and Signal Transduction. *Nature Reviews Molecular Cell Biology*, *1*(1), 31–39. https://doi.org/https://doi.org/10.1038/35036052

Singer, S. J., & Nicolson, G. L. (1972). The Fluid Mosaic Model of the Structure of Cell Membranes. *Science*, *175*(4023), 720–731. https://doi.org/https://doi.org/10.1126/science.175.4023.720

Sirois, S., Sing, T., & Chou, K.-C. (2005). HIV-1 gp120 V3 Loop for Structure-Based Drug Design. *Current Protein & Peptide Science*, *6*(5), 413–422. https://doi.org/10.2174/138920305774329359

Sonenberg, N., & Hinnebusch, A. G. (2009). Regulation of translation initiation in eukaryotes. *Cell*, *136*(4), 731–745. https://doi.org/10.1016/j.cell.2009.01.042.Regulation

Stillwell, W. (2016). An Introduction to Biological Membranes. In *Academic Press* (Vol. 3, Issue 10).

Strebel, K. (2013). HIV accessory proteins versus host restriction factors. *Current Opinion in Virology*, *3*(6), 692–699. https://doi.org/10.1016/j.coviro.2013.08.004

Sundquist, W. I., & Kra, H. (2012). *HIV-1 Assembly , Budding , and Maturation*.

Top, C., & Chem, M. (2016). *HHS Public Access*. *16*(10), 1074–1090.

U.S. Department of Health and Human Services; National Institutes of Health. (2022). *HIV Treatment HIV Treatment : The Basics*. National Institutes of Health. https://hivinfo.nih.gov/understanding-hiv/fact-sheets/hiv-treatment-basics

UNAIDS. (n.d.). *HIV and AIDS - Basic facts | UNAIDS*. Retrieved January 6, 2023, from https://www.unaids.org/en/frequently-asked-questions-about-hiv-and-aids

UNAIDS. (2021). Global HIV Statistics. *Fact Sheet 2021*, *June*, 1–3.

Vandegraaff, N., & Engelman, A. (2007). Molecular mechanisms of HIV integration and therapeutic intervention. *Expert Reviews in Molecular Medicine*, *9*(6), 1–19. https://doi.org/10.1017/S1462399407000257

Verlet, L. (1967). Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennar-Jones Molecules. *Physical Review*, *159*(1), 6. https://doi.org/https://doi.org/10.1103/PhysRev.159.98

Wang, X., Devaiah, S. P., Zhang, W., & Welti, R. (2006). Signaling functions of phosphatidic acid. *Progress in Lipid Research*, *45*(3), 250–278. https://doi.org/10.1016/j.plipres.2006.01.005

Watson, S. J. (2009). Molecular dynamics simulations of HIV-1 protease complexed with saquinavir. *In Department of Infection & Immunity, University College London, London.*

Watts, J. M., Dang, K. K., Gorelick, R. J., Leonard, C. W., Bess, J. W., Swanstrom, R., Burch, C. L., & Weeks, K. M. (2009). Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature*, *460*(7256), 711–716. https://doi.org/10.1038/nature08237

Ye, L., Gribling-Burrer, A. S., Bohn, P., Kibe, A., Börtlein, C., Ambi, U. B., Ahmad, S., Olguin-Nava, M., Smith, M., Caliskan, N., von Kleist, M., & Smyth, R. P. (2022). Short- and long-range interactions in the HIV-1 5′ UTR regulate genome dimerization and packaging. *Nature Structural and Molecular Biology*, *29*(4), 306–319. https://doi.org/10.1038/s41594-022-00746-2

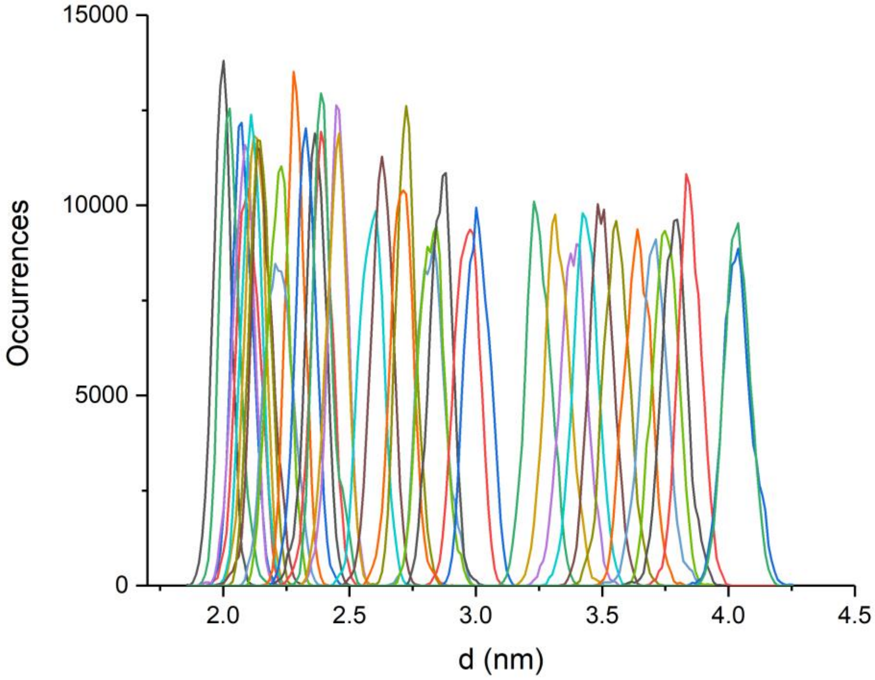
Yeagle, P. L. (2016). The Membranes of Cells, Third Edition. In *The Membranes of Cells, Third Edition*.

Zampelas, A., & Magriplis, E. (2019). New Insights into Cholesterol Functions: A Friend or an Enemy? *Nutrients*, *11*(7). https://doi.org/10.3390/nu11071645

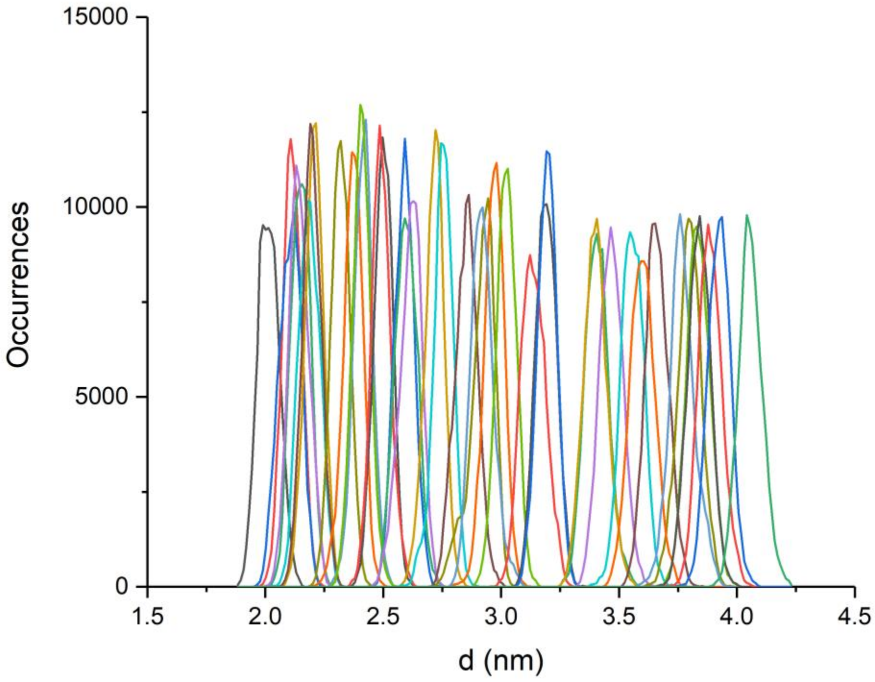
Zhang, D., Li, W., & Jiang, S. (2015). Peptide fusion inhibitors targeting the HIV-1 gp41: A patent review (2009 - 2014). *Expert Opinion on Therapeutic Patents*, *25*(2), 159–173. https://doi.org/10.1517/13543776.2014.987752

Zhang, J., & Crumpacker, C. (2022). HIV UTR, LTR, and Epigenetic Immunity. *Viruses*, *14*(5). https://doi.org/10.3390/v14051084

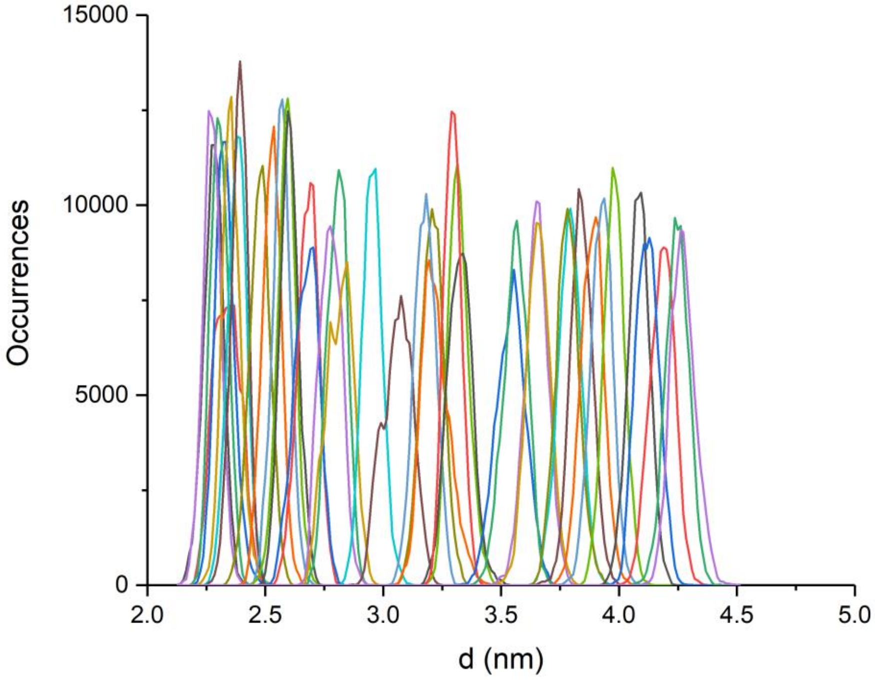
6. Appendix



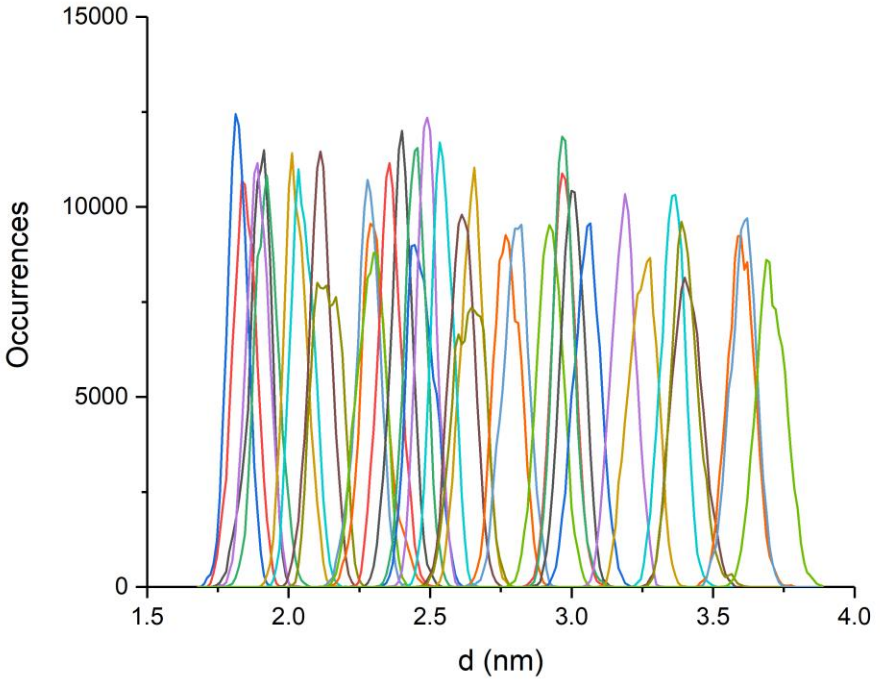
Appendix A – Histogram showing the number of occurrences as a function of the distance in nm for MT-C34 in the POPC system.



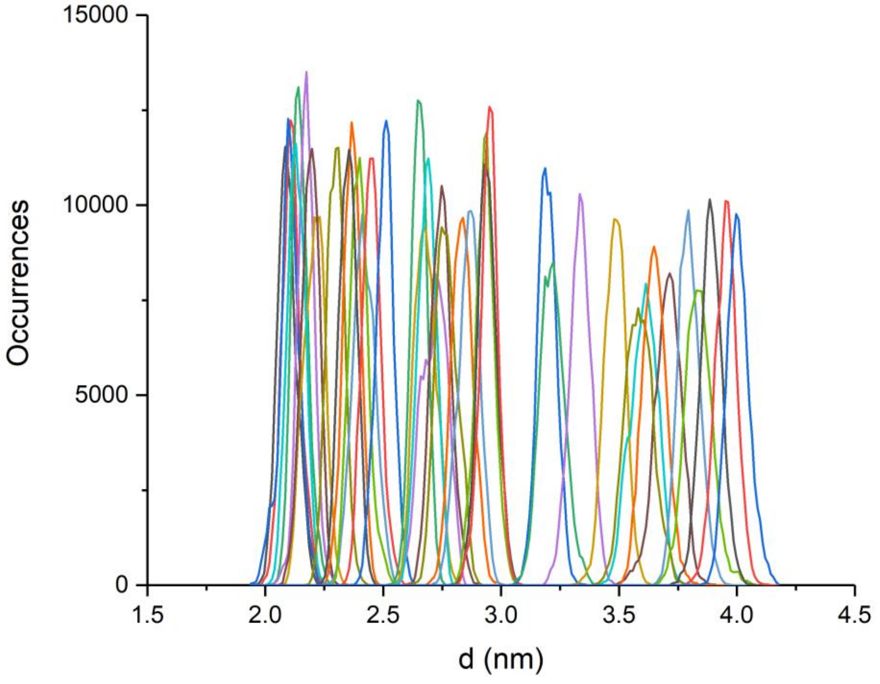
Appendix B - Histogram showing the number of occurrences as a function of the distance in nm for MT-SC22EK in the POPC system.



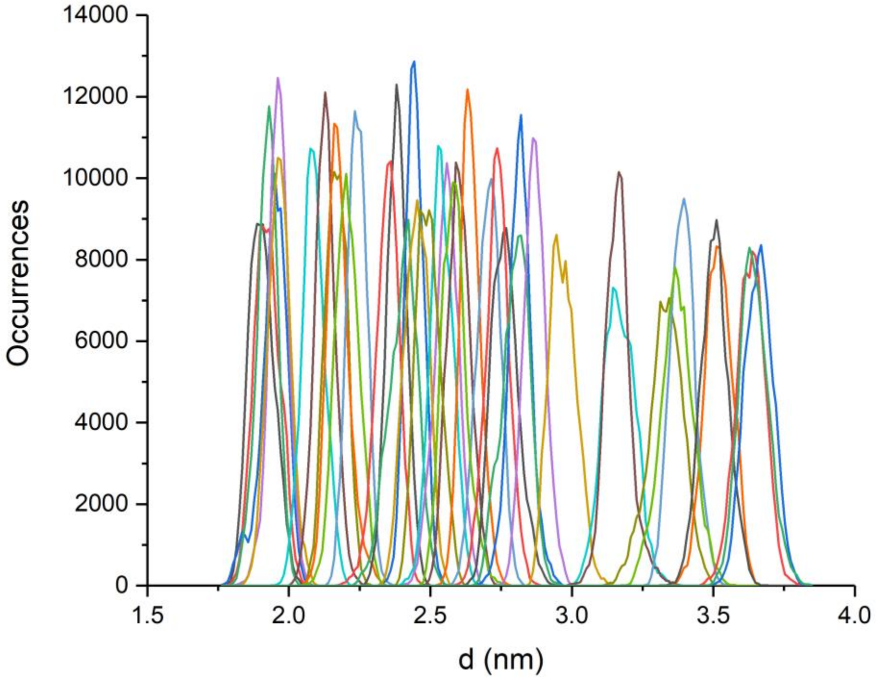
Appendix C - Histogram showing the number of occurrences as a function of the distance in nm for MT-SC29EK in the POPC system.



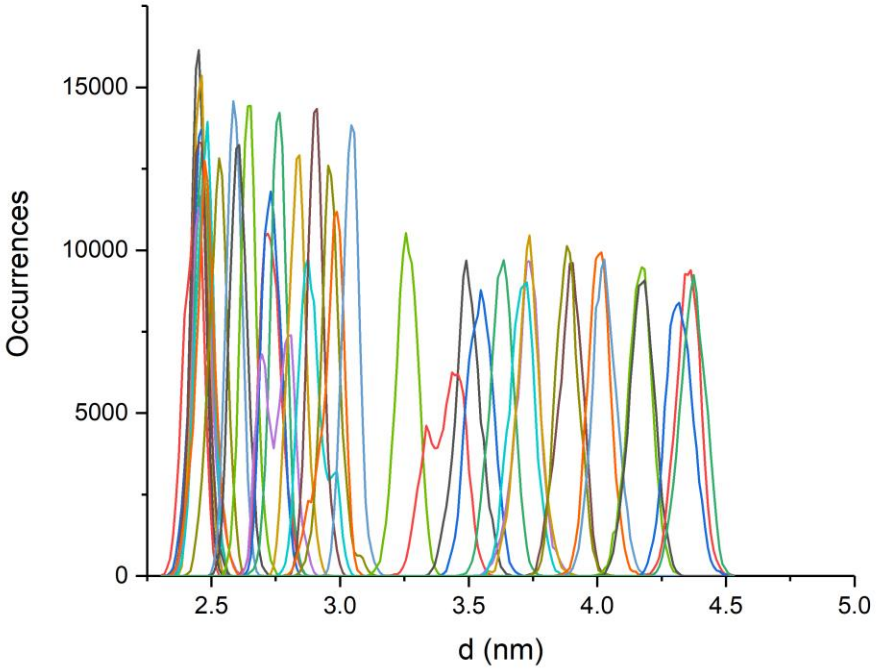
Appendix D - Histogram showing the number of occurrences as a function of the distance in nm for MT-SC34 in the POPC system.



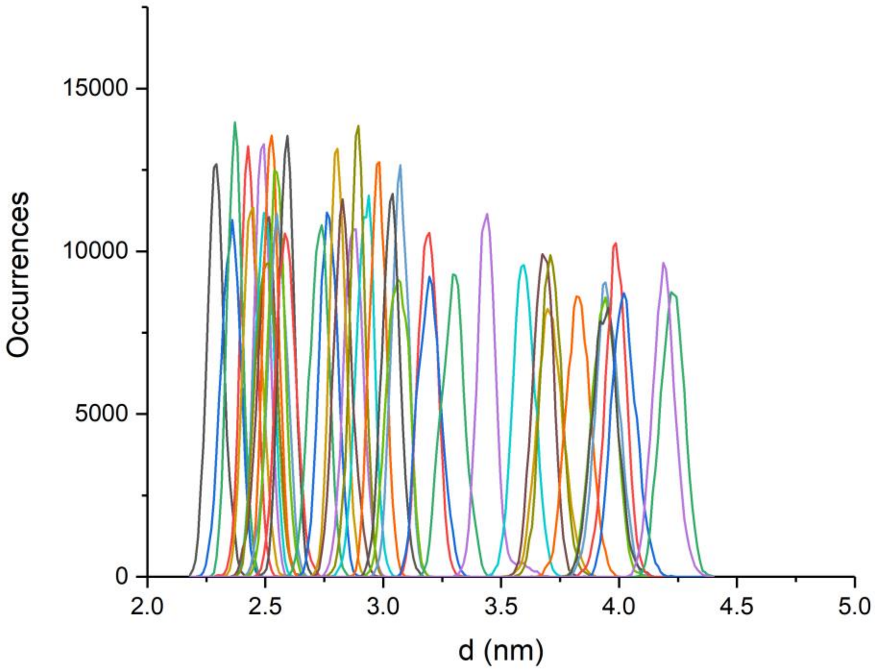
Appendix E - Histogram showing the number of occurrences as a function of the distance in nm for MT-SC34EK in the POPC system.



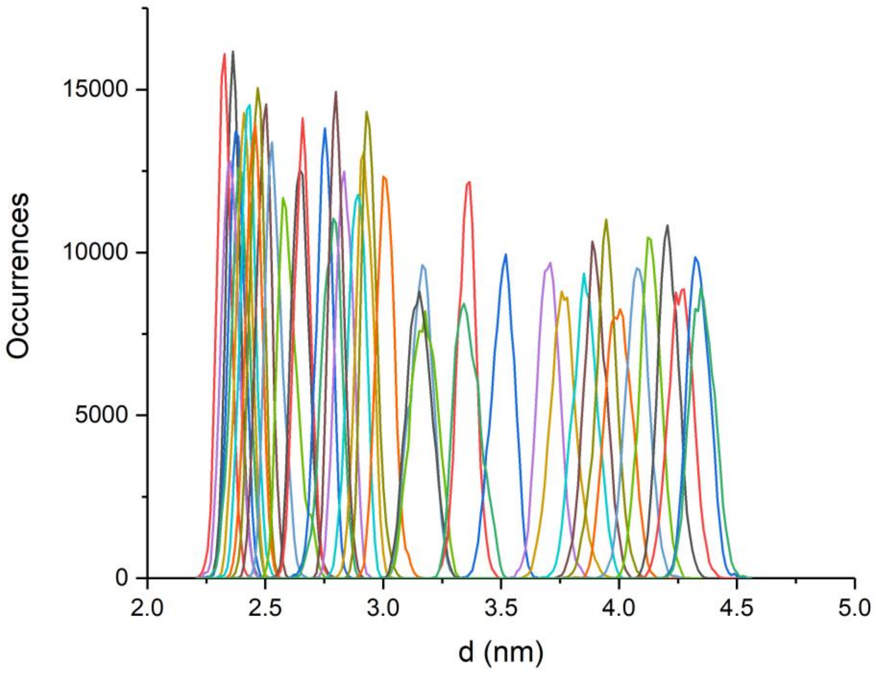
Appendix F - Histogram showing the number of occurrences as a function of the distance in nm for MT-SC35EK in the POPC system.



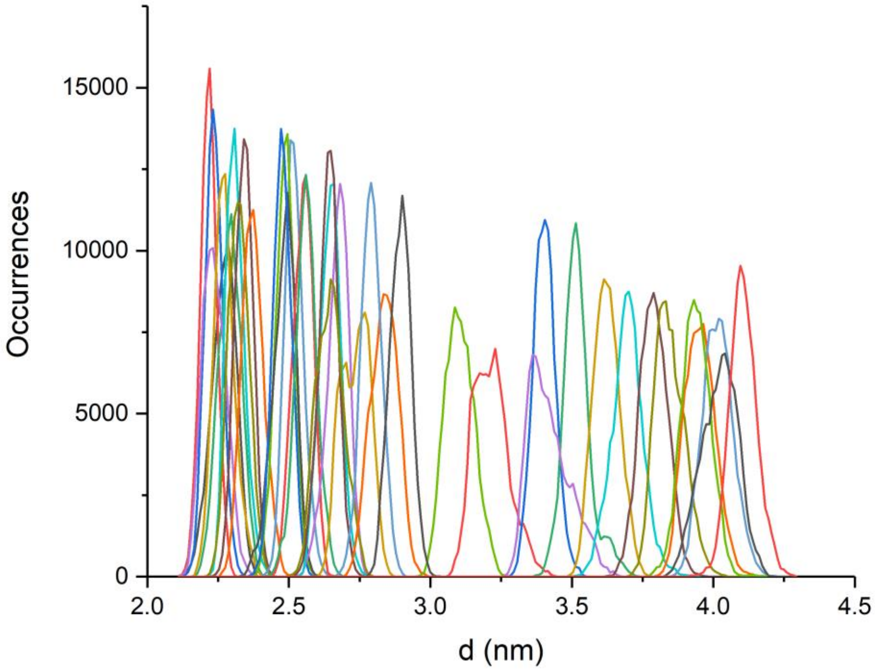
Appendix G - Histogram showing the number of occurrences as a function of the distance in nm for MT-C34 in the 20% cholesterol system.



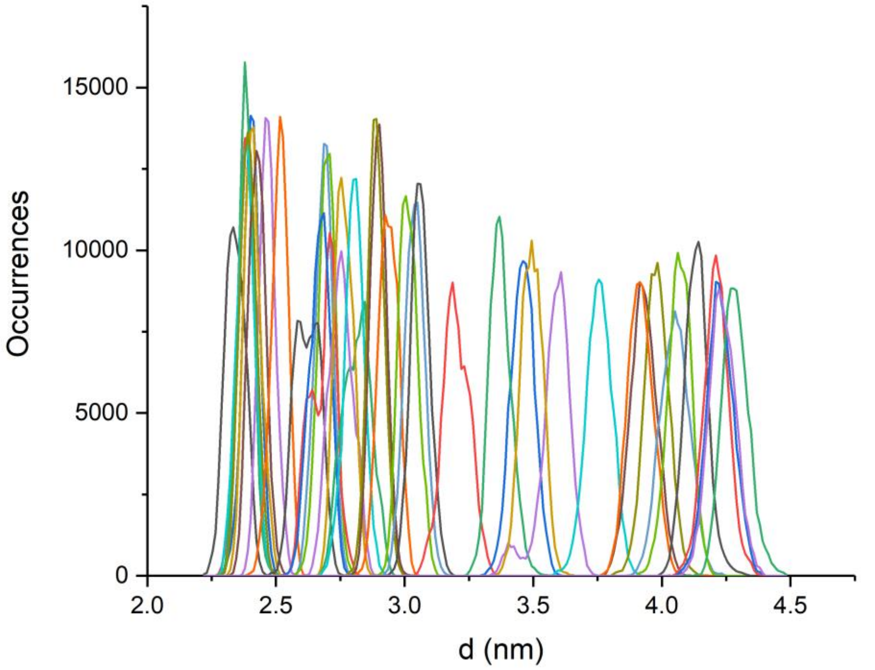
Appendix H - Histogram showing the number of occurrences as a function of the distance in nm for MT-SC22EK in the 20% cholesterol system.



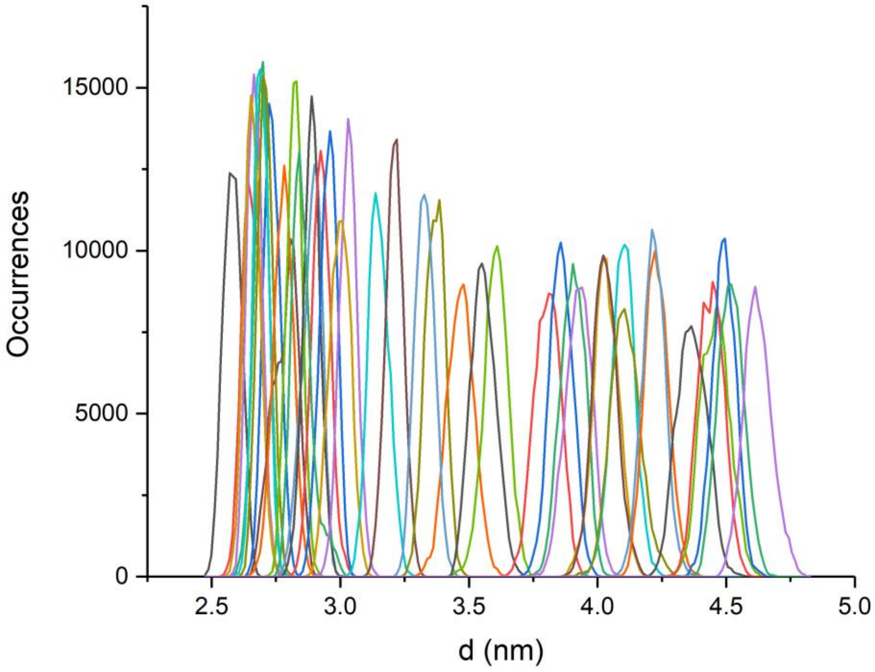
Appendix I - Histogram showing the number of occurrences as a function of the distance in nm for MT-SC29EK in the 20% cholesterol system.



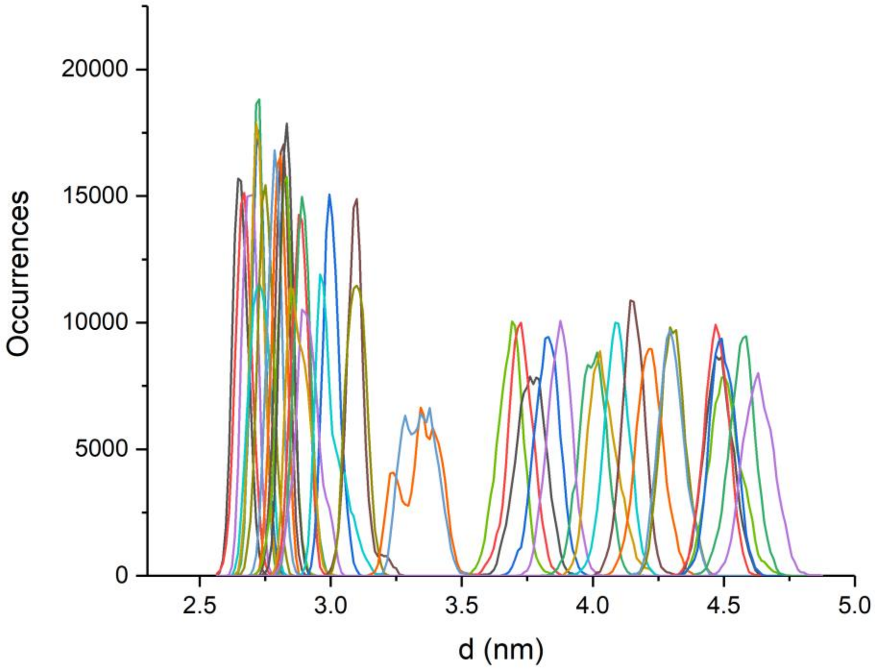
Appendix J - Histogram showing the number of occurrences as a function of the distance in nm for MT-SC34 in the 20% cholesterol system.



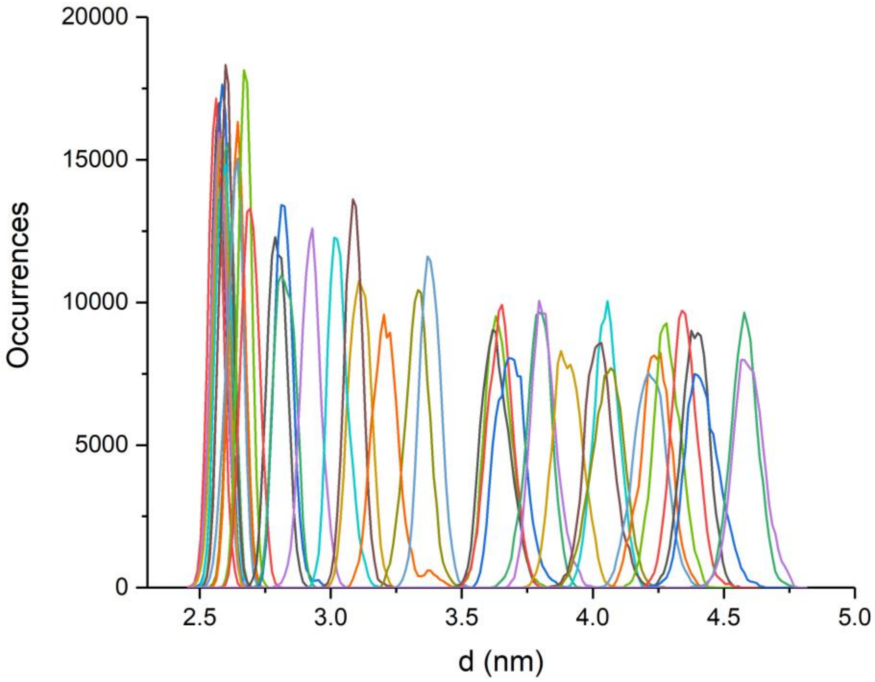
Appendix K - Histogram showing the number of occurrences as a function of the distance in nm for MT-SC34EK in the 20% cholesterol system.



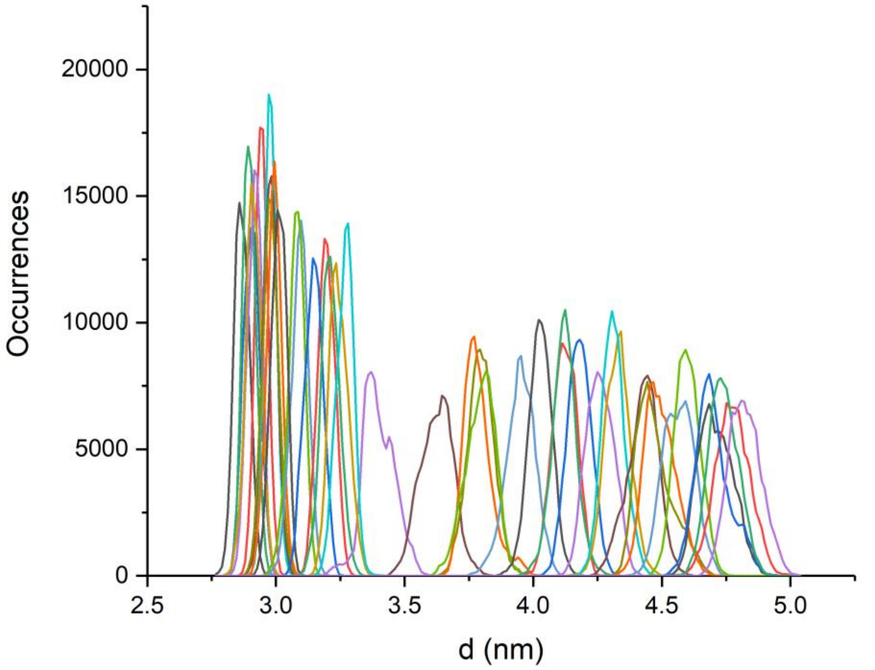
Appendix L - Histogram showing the number of occurrences as a function of the distance in nm for MT-SC35EK in the 20% cholesterol system.



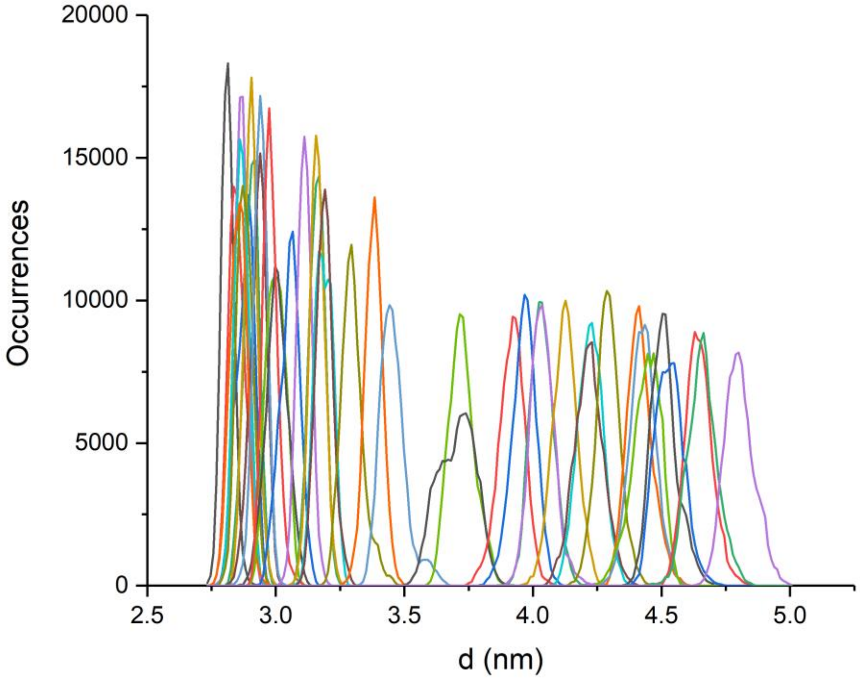
Appendix M - Histogram showing the number of occurrences as a function of the distance in nm for MT-C34 in the 50% cholesterol system.



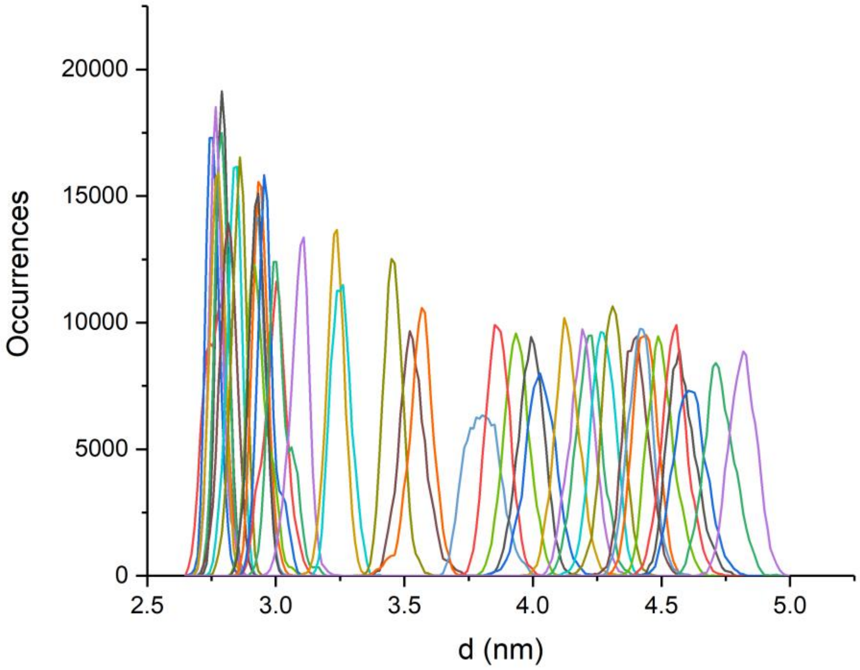
Appendix N - Histogram showing the number of occurrences as a function of the distance in nm for MT-SC22EK in the 50% cholesterol system.



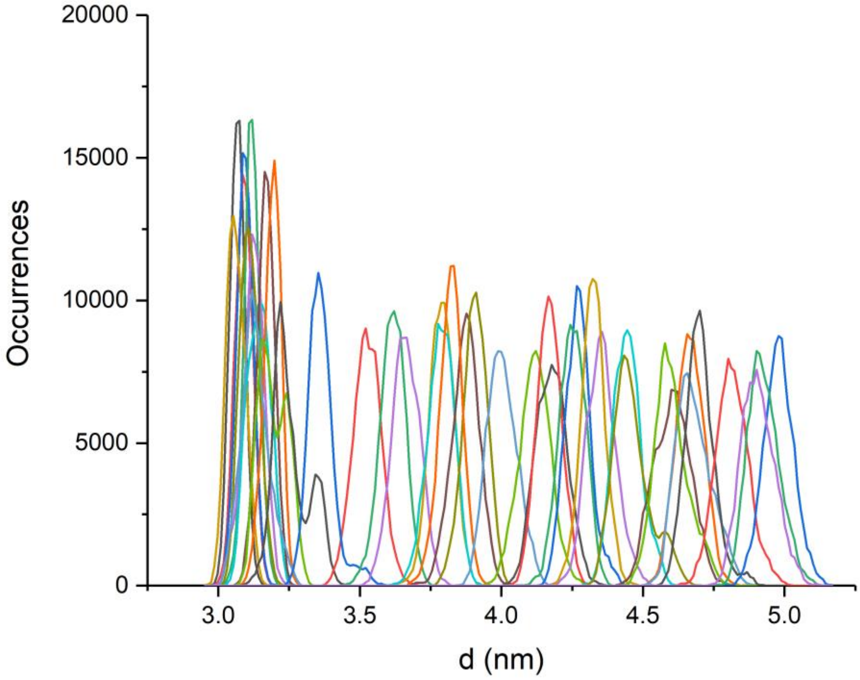
Appendix O - Histogram showing the number of occurrences as a function of the distance in nm for MT-SC29EK in the 50% cholesterol system.



Appendix P - Histogram showing the number of occurrences as a function of the distance in nm for MT-SC34 in the 50% cholesterol system.



Appendix Q - Histogram showing the number of occurrences as a function of the distance in nm for MT-SC34EK in the 50% cholesterol system.



Appendix R - Histogram showing the number of occurrences as a function of the distance in nm for MT-SC29EK in the 50% cholesterol system.