

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

**IDENTIFICATION OF NOVEL INHIBITORS
TARGETING PUTATIVE DANTROLENE BINDING SITE
FOR RYANODINE RECEPTOR 2**

M.Sc. THESIS

Cemil Can SAYLAN

Department of Computational Science and Engineering

Computational Science and Engineering Programme

JUNE 2022

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

**IDENTIFICATION OF NOVEL INHIBITORS
TARGETING PUTATIVE DANTROLENE BINDING SITE
FOR RYANODINE RECEPTOR 2**

M.Sc. THESIS

**Cemil Can SAYLAN
(702191003)**

Department of Computational Science and Engineering

Computational Science and Engineering Programme

Thesis Advisor: Assist. Prof. Dr. Sefer BADAY

JUNE 2022

**RYANODİNE RESEPTÖRÜ İÇİN
DANTROLENE BAĞLANMA BÖLGESİ HEDEF ALINRAK
YENİ İNHİBİTÖR TANIMLAMA**

YÜKSEK LİSANS TEZİ

**Cemil Can SAYLAN
(702191003)**

Hesaplamalı Bilim ve Mühendislik Anabilim Dalı

Hesaplamalı Bilim ve Mühendislik Programı

Tez Danışmanı: Assist. Prof. Dr. Sefer BADAY

June 2022

Cemil Can SAYLAN, a M.Sc. student of ITU Graduate School student ID 702191003 successfully defended the thesis entitled “IDENTIFICATION OF NOVEL INHIBITORS TARGETING PUTATIVE DANTROLENE BINDING SITE FOR RYANODINE RECEPTOR 2”, which he/she prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor : **Assist. Prof. Dr. Sefer BADAY**
Istanbul Technical University

Jury Members : **Assist. Prof. Dr. Sefer BADAY**
Istanbul Technical University

Prof. Dr. Fethiye Aylin SUNGUR
Istanbul Technical University

Assoc. Prof. Dr. Nurcan TUNCBAG
Koc University

Date of Submission : **27 May 2022**

Date of Defense : **13 June 2022**





To my family,



FOREWORD

First of all, I would like to express my gratitude to my thesis advisor Asst. Prof. Sefer BADAY for priceless advices and guidance. I would like to thank him for continuous support and understanding throughout this study.

Furthermore, I would like to thank my friends, Taha Yasin GÜRCAN and Alp Tegin ŞAHİN. Best Friends make good times better and hard times easier.

Lastly and most of all, I would like to express my gratitude to my family, Nurbaki SAYLAN, Tülin SAYLAN, Abdullah SAYLAN and Esra SAYLAN, for their priceless love and support throughout my master studies.

June 2022

Cemil Can SAYLAN

TABLE OF CONTENTS

	Page
FOREWORD	ix
TABLE OF CONTENTS	xii
ABBREVIATIONS	xiii
SYMBOLS	xv
LIST OF TABLES	xvii
LIST OF FIGURES	xxi
SUMMARY	xxiii
ÖZET	xxv
1. INTRODUCTION	1
1.1 Overview of Ryanodine Receptors	1
1.2 Structural information of the RyRs	2
1.3 Diseases Association of the RyR2	5
1.4 Regulation of RyR2	6
1.4.1 Ca^{2+} , Caffeine, Mg^{2+} , and ATP	7
1.4.2 CaM and FKBP	8
1.4.3 Dantrolene	9
1.5 Purpose of Thesis	10
2. THEORY AND METHODS	13
2.1 Molecular Modeling	13
2.1.1 Physical-based methods	14
2.1.2 Knowledge-based methods	14
2.1.2.1 AlphaFold 2	15
2.1.2.2 trRosetta (transform-restrained Rosetta)	17
2.1.2.3 FastRelax	18
2.1.2.4 Rosetta Score Function	19
2.1.3 Modeling of RyR2 homotetramer	19
2.2 Molecular Docking	20
2.2.1 Conformational Search for small molecules	21
2.2.2 Scoring Functions	22
2.2.3 Docking methodology	23
2.2.3.1 Autodock Vina	23
2.2.3.2 Glide	24
2.2.3.3 LeDock	26
2.2.4 Dantrolene Binding Pose Prediction	27
2.2.5 Identification of potential inhibitor for RyR2	28
2.3 Molecular Dynamic Simulations	29
2.3.1 Molecular Dynamics Algorithms	29
2.3.2 Force fields	32
2.4 Free Energy Calculations	32
2.4.1 MMPBSA	32
3. RESULTS AND DISCUSSION	35
3.1 Modeling of the RyR2	35
3.2 Dantrolene binding pose prediction	43
3.3 MMPBSA	59
3.4 High Throughput Screening	60
4. CONCLUSIONS	67
4.1 Modeling of the RyR2	67
4.2 Dantrolene binding pose prediction	68
4.2.1 Docking results	68
4.2.2 MMPBSA	68
4.3 High Throughput Screening	69
4.4 Comment and Future Perspective	69

REFERENCES **71**
APPENDICES **79**
 APPENDIX A : Tables 81
 APPENDIX B : Figures 83
CURRICULUM VITAE..... **85**



ABBREVIATIONS

MD	: Molecular Dynamic
RyRs	: Ryanodine Receptors
ATP	: Adenosine triphosphate
Ca²⁺	: Calcium ion
Mg²⁺	: Magnesium ion
FKBP	: FK506 binding protein
PKA	: Protein kinase A
SPRY	: SplA/ryanodine receptor domain
FDA	: Food and Drug Administration
Cryo-EM	: Cryo Electron Microscopy
GPU	: Graphics processing unit
μM	: Micro Mollar
CaMBD	: Calmodulin Binding Domain
ns	: Nano second



SYMBOLS

m	: mass
F	: Force
a	: Acceleration
t	: Time
r	: Displacement
v	: Velocity
G	: Gibbs Free Energy
E	: Enthalpy
S	: Entropy
T	: Temperature
E_{tot}	: Total Energy
SASA	: Solvent Accessible Surface Area
Å	: Angstrom
d	: Atomic Distance
ω, θ	: Planer angles
ϕ	: Dihedral



LIST OF TABLES

	<u>Page</u>
Table 2.1 : Rosetta Scoring terms parameters, default weights and descriptions..	19
Table 2.2 : AutoDock Vina scoring algorithm scoring weights.	24
Table 3.1 : Docking results with Score and Ranking	45
Table 3.2 : Drug Candidates human oral absorption and docking score results ..	63
Table A.1 : Dantrolene clusters for Cryo-EM model docking pose 1	81
Table A.2 : Dantrolene clusters for Cryo-EM model docking pose 2	81
Table A.3 : Dantrolene clusters for AF2 model docking pose 1	81
Table A.4 : Dantrolene clusters for AF2 model docking pose 2	81
Table A.5 : Dantrolene clusters for Cryo-EM model docking pose 1 with FKBP12.6	81
Table A.6 : Dantrolene clusters for Cryo-EM model docking pose 2 with FKBP12.6	82
Table A.7 : Dantrolene clusters for Cryo-EM model docking pose 3 with FKBP12.6	82
Table A.8 : Dantrolene clusters for AF2 model docking pose 1 with FKBP12.6..	82



LIST OF FIGURES

	<u>Page</u>
Figure 1.1 : An illustration of the Ca^{2+} release mechanism. The figure is adapted from [1].	1
Figure 1.2 : An illustration of the sequence and domain information of the RyR2. The figure is adapted from [2].	2
Figure 1.3 : A structural representation and orientation of RyRs.	3
Figure 1.4 : An illustration domain placement and α -solenoid repeats in the structure of RyRs.	4
Figure 1.5 : Closer look of the U-ring, O-ring, and Ca^{2+} , Caffeine and ATP binding sites. The figure is adapted from [3,4].	4
Figure 1.6 : The I4868 and Q4864 residues structural transition during the pore open-closed mechanism. The figure is adapted from the [2].	5
Figure 1.7 : The point mutations that associated with cardiac disorders such as arrhythmia, and catecholaminergic polymorphic ventricular tachycardia. The figure is adapted from [5].	6
Figure 1.8 : The CaM binding sites. The figure is adapted from [6].	8
Figure 1.9 : The FKBP12.6 binding site for RyR2. The structure PDB ID 6WOV [7].	9
Figure 1.10 : Dantrolene chemical structure	10
Figure 1.11 : The sequence of putative dantrolene binding site for RyR1 and RyR2 [5].	10
Figure 1.12 : The representation of the binding site of Dantrolene in PDB 5GO9 [2].	10
Figure 2.1 : The AF2 deep learning architecture representation adapted from [8].	15
Figure 2.2 : Schematic representation of the GlideXP workflow.	25
Figure 2.3 : High Throughput Screening workflow	28
Figure 2.4 : The grid box used in HTS.	29
Figure 2.5 : Schematic representation of periodic boundary condition.	31
Figure 2.6 : CaFE workflow	33
Figure 3.1 : Sequence comparison of RyR2 Cryo-EM structures	35
Figure 3.2 : The 3D representation of a protomer taken from available RyR2 structures; PDB 6JIY (blue), 5L1D (red), 5GO9 (green) and 6WOV (yellow).	36
Figure 3.3 : AF2 Model results for at 1311-1424 residues.	37
Figure 3.4 : SPRY3 doming for 5GO9 and 6WOV PDB structures	37
Figure 3.5 : Closer look at at 1576-1596 residues near the dantrolene binding sequence. AF2 model represented as blue, and 5GO9 represented as green color.	38
Figure 3.6 : Dantrolene binding sequence conformation in 5GO9 and 6WOV PDB structures.	39

Figure 3.7 : AF2 and trRosetta model results for HD2 domain and HD2 region structural shifts relative to 5GO9 Cryo-EM structure.	40
Figure 3.8 : Structural models for 1940-2000 residues.	41
Figure 3.9 : RyR2 homotetramer full-atom model, and model membrane complex structure.	42
Figure 3.10 : Pore radius of RyR2 full atom model.....	43
Figure 3.11 : RMSD and total energy change during fullatom RyR2-membrane complex. Individual protomer RMSD change during MD simulations	44
Figure 3.12 : RMSD changes of the Domains for Protomer A.....	45
Figure 3.13 : RyR2 homotetramer conformation after 35ns simulation	46
Figure 3.14 : Average pore radius change during the MD simulation.....	47
Figure 3.15 : Pore change and Q4864 and I4868 positions after 35ns MD simulation.	47
Figure 3.16 : The blind docking results. The left image for AF2 model, and right image for Cryo-EM model.	48
Figure 3.17 : Cryo-EM Model docking results. Dantrolene binding sequence is represented as red. Glide, Vina and Ledock poses are represented as blue, green and pink, respectively. a) Pose 1 for Cryo-EM Model, b) Pose 2 for Cryo-EM Model, and c) Pose 3 for Cryo-EM Model	49
Figure 3.18 : AF2 Model docking results. Glide, Vina and Ledock poses are represented as blue, green and pink, respectively. d) Pose 1 for AF2 Model, e) Pose 2 for AF2 Model.....	50
Figure 3.19 : RMSF results of the Cryo-EM Model docking MD simulations.	51
Figure 3.20 : RMSF results of the AF2 Model docking MD simulations.....	52
Figure 3.21 : RMSD results of the MD simulations for docking poses.	53
Figure 3.22 : The top cluster postions for Dantrolene. a) Obtained from Pose 1 for Cryo-EM Model simulation, b) from Pose 2 for Cryo-EM Model simulation, c) from Pose 1 for AF2 Model simulation, and d) from Pose 2 for AF2 Model simulation	54
Figure 3.23 : Comparison of RMSF results for FKBP12.6 present and not included systems for Cryo-EM model docking simulations.....	56
Figure 3.24 : Comparison of RMSF results for FKBP12.6 present and not included systems for AF2 model docking simulations	57
Figure 3.25 : RMSD change of the dantrolene and RYR2 truncated system in presence of FKBP12.6	58
Figure 3.26 : The top cluster postions for Dantrolene in FKBP12.6 present system. a) Obtained from Pose 1 for Cryo-EM Model simulation, b) from Pose 2 for Cryo-EM Model simulation, c) from Pose 1 for AF2 Model simulation, and d) from Pose 2 for AF2 Model simulation.	59
Figure 3.27 : MMPBSA energy changes during the simulations	60
Figure 3.28 : Initial HTS score distribution and density plots.	61
Figure 3.29 : Second step of HTS score distributions.....	61
Figure 3.30 : Drug Candidates for RyR2 inhibition	62
Figure 3.31 : RMSD results for drug candidates (<i>cont.</i>).....	64

Figure B.1 : Restraint locations in the structure are represented as red.	83
Figure B.2 : FKBP12.6 binding location. Pink is FKBP12.6 and red represents dantrolene binding sequence.	83
Figure B.3 : Dantrolene-RyR2 interaction plot for wo FKBP12.6 system.....	84
Figure B.4 : Dantrolene-RyR2 interaction plot for with FKBP12.6 system	84





IDENTIFICATION OF NOVEL INHIBITORS TARGETING PUTATIVE DANTROLENE BINDING SITE FOR RYANODINE RECEPTOR 2

SUMMARY

Ryanodine receptors (RyRs) are large (around 20 MDa) homotetrameric intracellular ion channels. RyRs are located in the membrane of the sarcoplasmic reticulum (SR). RyRs play a central role in the excitation-contraction coupling by regulating Ca^{2+} release from the SR to the cytosol. Three isoforms are sharing 70% sequence similarity; RyR1, RyR2 and RyR3 are predominantly expressed in skeletal, cardiac muscles and neurons, respectively. In all isoforms, the open and close state transition occurs by rigid-body shifts of domains that provide the overall breathing motion of the cytoplasmic region. Dysregulation of RyRs leads to abnormal cellular activity. More than 300 mutations have been associated with muscle and neuronal diseases. Today, there have been identified several heart diseases caused by aberrant RyR2 activity such as catecholaminergic polymorphic ventricular tachycardia, cardiomyopathies, and cardiac arrhythmias. This unwanted RyR2 activity can be modulated by drugs. Dantrolene is an approved muscle relaxant used for the treatment of malignant hyperthermia that occurs by dysregulation of RyR1. Previously, a dantrolene binding sequence was suggested, and this sequence is conserved for all RyRs. However, dantrolene has poor water solubility and dantrolene exerts its effect on RyR2 only in the presence of regulators such as Mg^{2+} and calmodulin. In addition to this, although the amino acids 590-609 in RyR1 (601-620 RyR2 equivalent) were identified as dantrolene binding sequences, dantrolene bound complex structure has not been elucidated yet. Here, we aimed three things; 1) modelling of full atom structure of RyR2 with membrane, 2) predicting dantrolene binding orientation and 3) identifying novel inhibitors targeted to the putative binding sequence of dantrolene to regulate RyR2 function. While most of the structure of RyR2 has been solved recently, some of the regions are still missing. To predict missing regions, we used trRosetta and AlphaFold2 which is a state-of-art deep-learning-based method for protein modelling. Each missing segment was modelled separately and combined at the end. The final model, then, was optimized by 35ns MD simulation. Subsequently, to predict the dantrolene binding pose, the putative dantrolene binding site was searched using three different docking programs; Vina, LeDock and Glide. There was a distinct difference around the dantrolene binding site between the AF2-based model and Cryo-EM based model. Thus, the docking was performed for either structure. The dantrolene population was obtained predominantly in a particular cavity formed by 6 domains. Among docking results, five binding poses (3 of cryo-em, 2 for AF2 model) were selected regarding the affinity scores and pose similarity. These poses were used in 200ns MD simulation (298K, NPT) to address the pose binding behaviour. In the simulations truncated system was used, and restraints were introduced at regions where we split the structure. After 200ns, four orientations of dantrolene (2 for each) remained in its interaction

with the dantrolene binding sequence. To calculate the binding free energy of the binding poses, MMPBSA analysis was performed using MD trajectories. Besides this, we also investigated the FKBP12.6 binding effect on dantrolene binding using the docking structures. Here, all proceeded MD runs with dantrolene were replicated with FKBP12.6 bound conditions. Structural clustering of MD simulations together with MMPBSA results showed that particular dantrolene orientation showed the highest binding capability to around R606, E1649, and L1650 residues. However, we could not identify a significant effect of FKBP12.6 on dantrolene binding. Next, for the identification of novel inhibitors, we focused on the dantrolene binding region and the high-throughput screening of 3.5 million molecules retrieved from the ZINC15 database was applied. The molecules were selected based on molecular weights (<450 Da) and logP values (<3.5). Virtual screening has proceeded with 3-step gradual filtering. The initial step includes 3.5 million molecule screening using AutoDock Vina with 8 exhaustiveness. This step was followed by two screening procedures in which the top-ranked 200K molecules selected from the previous step were filtered with LeDock and Vina (with 24 exhaustiveness). Molecules shared in top-10K for both Vina and LeDock results were used for the third and last screening with GlideXP. Subsequently, among the top-100 molecules, top-20 was selected as the final candidate list. According to >70% human oral absorption conditions, the best 11 molecules have proceeded for MD simulations. 200ns MD simulation was carried out using Desmond. The seven molecules remained in their interaction with the dantrolene binding sequence. These were suggested as candidates that might regulate RyR2 activity. Particularly, two molecules showed significant stability at the binding site at around 1Å. These molecules will be tested experimentally by our experimental collaborators.

RYANODİNE RESEPTÖRÜ İÇİN DANTROLENE BAĞLANMA BÖLGESİ HEDEF ALINRAK YENİ İNHİBİTÖR TANIMLAMA

ÖZET

Ryanodine reseptörleri büyük (yaklaşık 20 MDa) homotetramer kalsiyum iyon kanallarıdır. Bu reseptörler kas hücrelerinde sarkoplazmik retikulum üzerinde t-kanalına yakın oryantasyonda bulunurlar. Bu reseptörler kas kasılıp gevşemesi olayında merkezi rolü üstlenmektedir. Bilinen 3 isoformu bulunuyor. Bunlardan RyR1 daha çok iskelet kaslarında, RyR2 ise kalp kasında çoğunlukla bulunmaktadır. RyR3 ise daha çok nöronlarda görülmektedir. Bu isoformlar 70% oranında sekans benzerliği göstermektedir. Daha önceki çalışmalarda, kalsiyum salınımı kalp ve iskelet kasları için iki şekilde tanımlanmıştır. Bunlarda RyR1 yani iskelet kaslarında hücre zarına gelen voltaj değişimin etkisiyle kalsiyum voltaj kanalları aktifleşir ve bu kanallar ise direk olarak RyR1 açılmasını tetikler. Bu sayede kalsiyum salınımı gerçekleşmiş olur. RyR2 için ise bu salınım kalsiyum voltaj nalları aktifleştikten sonra hücreler arası sıvıda bulunan kalsiyum hücre içine alınımı yapılır ve bu kalsiyum konsantrasyon artışı RyR2 aktifleştirerek kalsiyum salınımı gerçekleştirilir. Bu olaya kalsiyum uyaranlı kalsiyum salınımı denir. Her iki kanal açılıp kapanma olayı protein domainlerinin gövde kayması şeklinde gerçekleşir. Bu salınım her hangi bir mutasyon veya uyaran ile bozulduğunda normal olmayan hücresel aktivitelere neden olmaktadır. Bu reseptörler üzerinde görülen 300 den fazla mutasyon kas hastalıkları ile ilişkilendirilmiştir. Günümüzde RyR2 ile direk ilişkilendirilen ritim bozukluğu gibi bir çok kalp rahatsızlığı tanımlanmıştır. Bu bozuklukların giderilmesi için çeşitli ilaçlar öne sürülmüştür. Dantrolen bunlardan bir tanesidir ve malign hipertermi için tek onaylanmış ilaçtır. Bu rahatsızlık RyR1 aktivitesinin bozukluğunda görülmektedir. Fakat yapılan çalışmalarda dantrolenin RyR2 üzerinde de etkisinin olduğu görülmüştür. Ayrıca, dantrolen için öne sürülen bağlanma sekansı bütün Ryanodine reseptörleri için korunmuş olduğu görülmüştür. Fakat dantrolen kalp rahatsızlıkları için kullanılması için uygun bir ilaç değildir. Bu durumunun nedenini dantrolenin kötü bir su çözünmesi ve ek regülatörlere ihtiyaç duymasına bağlanıyor (Magnezyum, FKBP ve kalmodulin gibi). Dantrolen bağlanma sekansı daha önce önerilmiş olmasına rağmen her hangi bir yapısal bilgi mevcut değil. Biz bu çalışmada üç durumu elde etmeyi amaç edindik. 1) RyR2 için tam atom membrane kompleks oluşturmak, 2) dantrolenin bağlanma pozunun tahmin edilmesi, ve 3) yeni ilaç molekül adaylarının tanımlanması. RyR2 daha önceden cryo-em kullanılarak çıkarılmış yapıları bulunmaktadır. Fakat, bu yapılardan bazı amino asit sekansının tamamen, bazılarının ise sadece ana zinciri bulunamamış ya da iyi bir şekilde görüntülenememiştir. RyR2 yapıları incelendiğinde küçük sekans bölgelerinin yanında, büyük sekans bölgelerinin çözülmediği görülmüştür. Bu çalışmada AlphaFold2 (AF2) ve trRosetta kullanarak RyR2 yapısı çıkarılmıştır.

RyR2 yapı tahminleri parça parça olacak şekilde tahmin edildi. Sonra bu model parçaları son bir modelde birleştirildi. AF2 ve Cryo-EM modelleri incelendiğinde özellikle HD2 domaininde yapısal farklar olduğu gözelemlendi ve bu durumun proteinin aktivasyonunda etkisi olabileceği öngörülmüştür. Oluşturulan model yapısının optimizasyonu 35ns MD simülasyon ile yapıldı. Literatürde dantrolen molekülünün Ryanodin reseptörleri üzerindeki bağlanma bölgesi amino asit sekansı 590–609 kalıntıları RyR1 için, 601–620 kalıntıları RyR2 için olarak verilmiş olsa da deneysel olarak 3-boyutlu yapısı bulunmamaktadır. Ryanodin reseptörünün dantrolen molekülü ile kompleks yapılarının potansiyel bağlanma pozlarının tahmini ve simülasyonları için dantrolenin Ryanodin reseptörünün kenetlenme (docking) hesapları gerçekleştirilmiştir. Bunun için üç farklı program kullanılmıştır. Bunlar Glide XP, AutoDock VINA ve LeDock programlarıdır. Yukarıda belirtilen model yapılarındaki değişikliğe ek olarak bir diğer yapısal model farklılığın dantrolene bağlanma sekansına çok yakın bir bölgedeki amino asit sekansında olduğu gözlenmiştir. Bu farklılıktan kaynaklı oluşabilecek farklılıkları görebilmek için kenetlenme çalışması, yapı modellemesi sonuçlarından elde edilen iki farklı protein modeli üzerinde (AF2-temelli ve Cryo-EM-temelli) uygulandı. Bağlanma sekansının tamamı farklı merkez ve büyüklükler ile oluşturulan grid kutuları ile tarandı. Sonuç olarak dantrolen popülasyonu, ağırlıklı olarak 6 alan tarafından oluşturulan belirli bir lokastonda elde edildi. Yerleştirme sonuçları arasında, afinite puanları ve poz benzerliği açısından beş bağlama pozu (Cryo-EM için 3, AF2 modeli için 2) seçildi. Bu pozlar, poz bağlama davranışını ele almak için 200ns MD simülasyonunda (298K, NPT) kullanıldı. Simülasyonlarda RyR2'den alınan kesilmiş sistem kullanılmış ve yapıyı böldüğümüz bölgelerde kısıtlamalar tanımlanmıştır. 200ns'den sonra, dört dantrolen bağlanma pozunun (her biri için 2) bağlanma bölgesinde kaldığı ve dantrolen bağlanma sekansı ile belirli bir etkileşimde olduğu görüldü. Bağlanma pozlarının bağlanma serbest enerjisini hesaplamak için MD yörüngeleri kullanılarak MMPBSA analizi yapıldı. Bunun yanı sıra, FKBP12.6 bağlamasının dantrolen bağlanmasındaki etkisinin araştırılması için tekrardan kenetlenme sonucunda elde edilen yapılar kullanıldı ve MMPBSA ile bağlanma serbest enerjileri hesaplandı. MMPBSA sonuçlarıyla birlikte MD simülasyonlarının yapısal kümelenmesi sonuçları baz alındığında, belirli dantrolen oryantasyonunun, yaklaşık R606, E1649 ve L1650 kalıntılarına en yüksek bağlanma kabiliyetini gösterdiği görüldü. Fakat, dantrolen bağlanmasında AF2 model ve Cryo-EM model arasında anlamlı bir fark görülmedi. Bu çalışma dantrolene bağlanmasını hesapsal yolla inceleyen ilk çalışma olup daha sonralarında bu konu üzerinde yapılacak olan çalışmalarla dantrolen bağlanması daha kapsamlı olarak incelenecektir. Bunlara ek olarak, yeni inhibitörlerin tanımlanması için dantrolen bağlanmasının yoğunlukta olduğu bölgeye odaklanıldı ve ZINC15 veri tabanından alınan 3.5 milyon molekül yüksek verimli bir tarama ile incelendi. Moleküller, moleküler ağırlıklarına (<450 Da) ve logP değerlerine (<3.5) göre seçildi. 3 aşamalı kademeli filtreleme ile sanal taramaya geçilmiştir. İlk adım, AutoDock Vina kullanılarak 3,5 milyon molekülün taranmasını içermektedir. Burada exhaustiveness parametresi 8 olarak alınmıştır. Bu adımı, önceki adımdan seçilen en üst sıradaki 200K moleküllerinin LeDock ve Vina ile filtrelendiği (Vina için 24 exhaustiveness uygulandı) iki tarama prosedürü izlendi. Hem Vina hem de LeDock sonuçları için ilk 10K'da paylaşılan moleküller, GlideXP ile üçüncü ve son tarama için kullanıldı. Ardından, ilk 100 molekül arasından ilk 20, nihai aday listesi

olarak seçildi. >%70 insan oral absorpsiyon durumuna göre, MD simülasyonları için en iyi 11 molekül seçildi ve 200ns MD simülasyonu Desmond kullanılarak moleküllerin bağlanma davranışlarının incelenmesi gerçekleştirilmiştir. Moleküllerin 7'si, dantrolen bağlanma sekansı ile etkileşimlerini sürdürdüğü görülmüştür. Bunlar, RyR2 aktivitesini düzenleyebilecek adaylar olarak önerildi. Özellikle, iki molekül bağlanma stabilitesinin çok yüksek olabileceği öngörüldü bu moleküllerin RMSD değerlerinin 1Å civarında olduğu görülmüştür. Bu moleküllerin inhibasyon etkilerin deneysel işbirlikçilerimiz tarafından deneysel olarak test edilecektir.





1. INTRODUCTION

1.1 Overview of Ryanodine Receptors

Ca^{2+} is a crucial ion for muscle contraction, as well as signal transduction pathways. Ca^{2+} has stored mainly endoplasmic reticulum (or sarcoplasmic reticulum for muscle) [9]. The Ca^{2+} hemostasis is controlled by Ca^{2+} ion channels such as sarco/endoplasmic reticulum Ca^{2+} -ATPase or Ryanodine Receptors (RyRs). There are three isoforms for RyRs in mammals, RyR1, RyR2 and RyR3. RyR1 is most abundant in skeletal muscle, RyR2 is in cardiac muscle and RyR3 is found in the brain but its function is less known among other isoforms [5]. RyR1 and RyR2 are large homotetrameric intracellular Ca^{2+} ion channels located at the sarcoplasmic reticulum membrane. They regulate the excitation-contraction coupling by releasing stored Ca^{2+} from the sarcoplasmic reticulum.

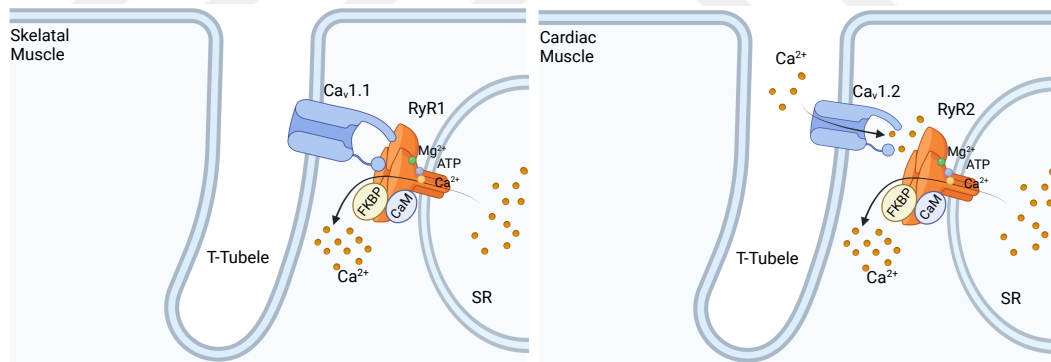


Figure 1.1 : An illustration of the Ca^{2+} release mechanism. The figure is adapted from [1].

RyRs have proximity to transverse tubules (t-tubules) and L-type Ca^{2+} channel ($\text{Ca}_{v1.1}$ in skeletal and $\text{Ca}_{v1.2}$ in cardiac muscles) [10]. Previously, it is shown that RyR1 has mechanical coupling with $\text{Ca}_{v1.1}$. Ca^{2+} release is provided in two ways [1]. The voltage change across the t-tubules activates the $\text{Ca}_{v1.1}$ which triggers RyR1 by direct coupling [1]. This is observed in skeletal muscle contractions. Likewise, in cardiac muscle, the depolarization on t-tubules initiates Ca^{2+} influx from extracellular matrix to

intracellular environment via $\text{Ca}_{v1.2}$. An increase Ca^{2+} concentration in the cytoplasm activates RyR2 to release stored Ca^{2+} in SR. This is known as Ca-induced Ca^{2+} release [11]. This Ca^{2+} releasing is proceeded until Ca^{2+} concentration reaches the $10 \mu\text{M}$ Ca^{2+} concentration level [11]. In either mechanism, Ca^{2+} concentration is restored by sarcoplasmic reticulum Ca^{2+} ATPase (SERCA) via pumping back to the SR [12]. The overview of the mechanism is shown in Figure 1.1.

1.2 Structural information of the RyRs

RyRs are large protein, thus the structural identification procedures are challenging. In last two decades, improvements in Cryo-EM technologies have brought great performance in structural determination of RyRs. Even though the initial experiments had resulted $11\text{\AA}$ resolution structures [13], the latest structural studies succeeded to obtain $3\text{-}5\text{\AA}$ resolutions for either RyR1 and RyR2 [2,7,14].

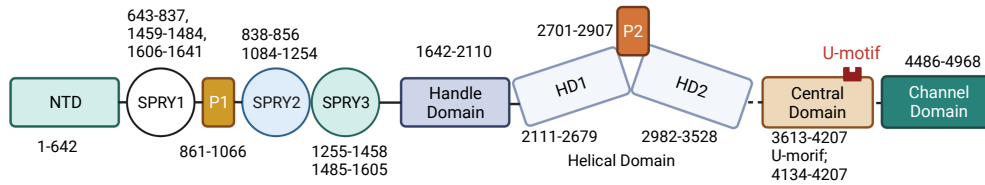


Figure 1.2 : An illustration of the sequence and domain information of the RyR2.
The figure is adapted from [2].

There are about 70% sequence similarity between RyR1 and RyR2. Thus, as expected, the overall quaternary structures are similar. The shape is mushroom-like and consists of large cytoplasmic head and transmembrane channel part. The homotetrameric structure is represented into three main parts. These are Clamp region, Handle and Central rim (also known as Central Tower). There are ten distinct domain in each protomer (chain, monomer). These are an N-terminal domain (NTD), three SPRY domains, two RyR repeat domains (Repeat 1-2, Repeat 3-4), a handle domain, a helical domain (two subdomains, HD1 and HD2), a central domain, and channel domain which includes the transmembrane domain and the C-terminal domain (CTD).

The detailed sequence order based representation is shown in Figure 1.2. It can be seen that Repeat 1-2 intervenes between SPRY1 and SPRY2 domains. There are three

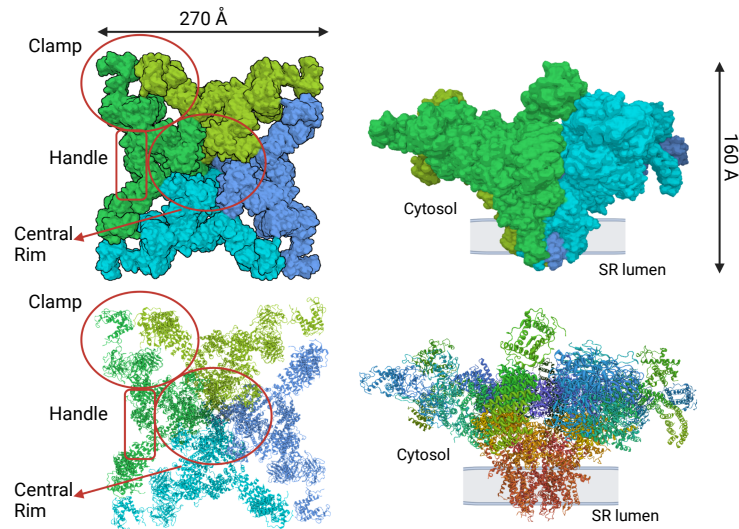


Figure 1.3 : A structural representation and orientation of RyRs.

SPRY domains in RyRs. These domains are structurally similar. Previously, it has been suggested that SPRY2 binds to dihydropyridine receptor (DHPR) in the skeletal muscle (RyR1). [15]–[18]. However, there is no information for RyR2.

The RyRs structures are also represented based on the four α -solenoid repeats (armadillo repeats). These are the core solenoid (CSol), the bridging (BSol), N-terminal solenoid (NSol or NTD-C) and junctional solenoid (JSol). CSol corresponds to the central domain, whereas BSol corresponds to the helical domains [3]. The α -solenoid repeats are flexible and predominantly determined as a role in protein-protein interaction surfaces [19]. These repeats provide the Inter-protomer interactions between SPRY2-HD2 of adjacent protomer (corresponds to clamp domain), and NTD-HD1 of adjacent protomer. This coupling allows the allosteric conformational changes across the channel [3].

NTD consists of three subdomains and it constructs a central tower together with a central domain and channel domain. Besides this, the NTD is identified as one of the mutations hot-spots [20]. The helical domain is a long curved domain formed by two armadillo repeats. These two parts are interrupted by the Repeat 3-4 domain. Essentially, the second half of the HD2 curved towards the central domain. Moreover, because of the higher flexibility [7], HD2 is not resolved well in the RyR2. Even though there has been no structure elucidated between the HD2 and central domain,

located in the cytoplasmic region [2,7]. At the end of the cytoplasmic region, CTD forms a zinc-finger domain. Moreover, together with S2 and S3 parts, CTD creates an “O-ring”. The central domain has a U-motif (residues 4134-4207 for RyR2) which clamps the O ring. During the channel opening or closing, these two structures show coupled motion [4]. Besides this, the central domain interface of the CTD has the binding site of Ca^{2+} , ATP and Caffeine known as channel activators. RyR2 gating is provided by the transmembrane part of the S6 segment (F4854 and later residues). In the close state of the RyR2, the I4868 restricts Ca^{2+} transition in the pore ($<1\text{\AA}$ radius). Upon RyR2, I4868 shifted towards outward of the pore and Q4864 turns the pore interface which allows to Ca^{2+} transition (around $2\text{\AA}$ radius) (Figure 1.6) [2].

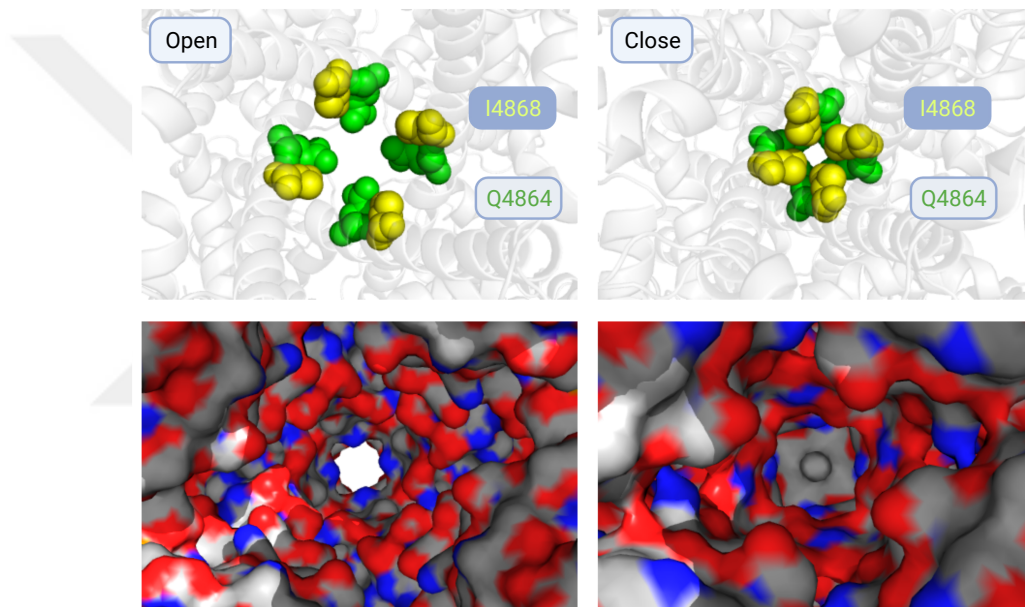


Figure 1.6 : The I4868 and Q4864 residues structural transition during the pore open-closed mechanism. The figure is adapted from the [2]

1.3 Diseases Association of the RyR2

Disruption of Ca^{2+} releasing mechanism by any mutation is resulting in devastating and fatal diseases such as atypical periodic paralyses, myopathies and heart failure. Since RyRs are crucial for life, many studies have been conducted to understand their mechanism and disease relations. Previously, over 300 pathogenic mutations have been identified in RyR2 that is linked to life-threatening cardiac disorders such as arrhythmia, and catecholaminergic polymorphic ventricular tachycardia (CPVT)

(Figure 1.7) [21]. These mutations disrupt the contraction-relaxation cycle during heartbeat causing dysfunctional RyR2 which leads to Ca^{2+} leakage during the diastolic state. In normal conditions, the concentration of free Ca^{2+} oscillates between diastole and systole [22]. Moreover, a great deal of these mutations are located as four mutation hot-spots which are the N-terminal domain, helical domain, central domain, and channel domain [1,23,24].

CPVT is an inherited autosomally dominant ventricular arrhythmias. CPVT is identified as stress-induced ventricular arrhythmias. It can cause sudden death under emotional or physical stress conditions, although the heart is structurally normal. Previously, there have been identified mutations in the *RYR2*, *CASQ2*, *KCNJ2*, *TRDN* and *CALM1* as a CPVT caused genes [25]. RyR2 is to far greater extent among these genes. The known mutations in RyR2 enhance the open probability of the channel which leads to SOICR [26]. S2246L, D3291V, N4104K, R4496C, and N4895D, for instance, are associated with CPVT1 and SOICR [27,28].

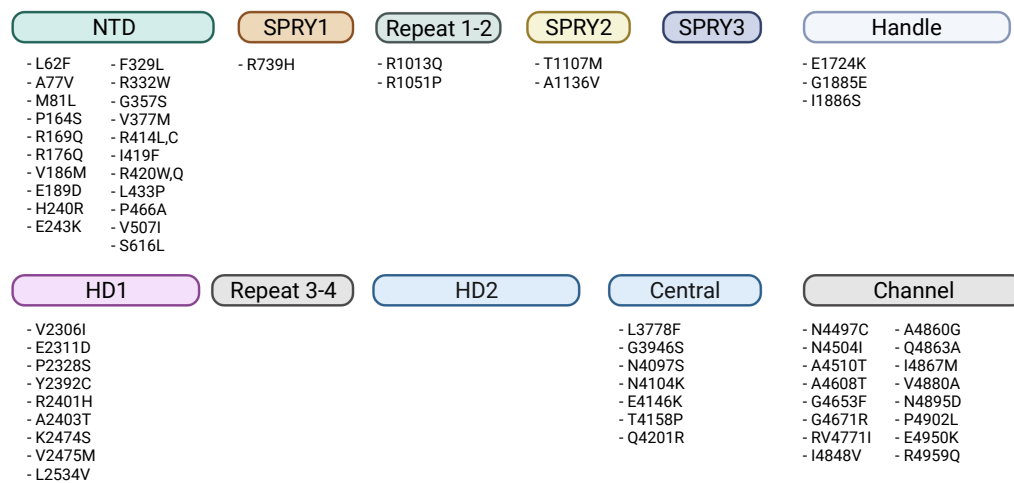


Figure 1.7 : The point mutations that associated with cardiac disorders such as arrhythmia, and catecholaminergic polymorphic ventricular tachycardia. The figure is adapted from [5]

1.4 Regulation of RyR2

The open-closed transition of RyRs is taken place by the rigid-body shifts of the domains [2]. This cycle is regulated by the different stimulants. Ca^{2+} , Mg^{2+} , Caffeine and ATP as well as Calmodulin (CaM), FKBP and Dantrolene can be given as example

[3]. Some of those stabilize the close state conformation of RyRs, whereas some activate and increase the open state probability of RyRs.

1.4.1 Ca^{2+} , Caffeine, Mg^{2+} , and ATP

Ca^{2+} , Caffeine, Mg^{2+} , and ATP are the most studied molecules. Essentially, Ca^{2+} has a primary effect on the RyRs. Either cytosolic or lumen Ca^{2+} can regulate the RyRs activity. Basically, the increased concentration of cytosolic Ca^{2+} activates RyRs to greater extent until it reaches certain level of free Ca^{2+} concentration in cytosol [29]. In the diastolic state of heart beating, cytosolic free Ca^{2+} oscillates around 0.2 μM and the open probability of RYRs are sufficiently low, whereas in systole this value rises to 200–400 μM in Ca-induced Ca^{2+} release [22]. There have been suggested that there might be two Ca^{2+} binding sites with differing affinities because of the bell-shaped concentration vs open probability curve of Ca^{2+} [30]. It is believed that Ca^{2+} can occupy the low-affinity binding site which causes the inhibition at high Ca^{2+} concentration [3].

ATP and caffeine are other known activators of RyRs [31]. Besides this, Ca^{2+} can work with ATP and caffeine synergistically due to the proximity of their binding sites [1]. The highest open probability is observed when all of them are present. Caffeine is an allosteric agonist of Ca^{2+} , and in presence of Ca^{2+} /ATP/Caffeine causes around 90% open probability in the channel [32,33]. The binding forms of Ca^{2+} /ATP and Ca^{2+} /Caffeine have very similar conformational changes on the CTD and consequently, it has the same effects on the S6 segment at the pore region [2].

Mg^{2+} is proposed as an inhibitor for RyRs [34]. Three possible mechanisms have been suggested [3]. One is that Mg^{2+} may antagonize for the Ca^{2+} activation site which prevent the Ca^{2+} binding. Likewise, Mg^{2+} may bind to low-affinity Ca^{2+} sites which leads to inhibition of Ca^{2+} activity [35,36]. In last mechanism, Mg^{2+} has inhibition effect by lowering free ATP, since Mg^{2+} /ATP are known complex in the cell [3].

1.4.2 CaM and FKBP

CaM and FKBP prevent unwanted diastolic RyR2 activity [1]. Besides this, CaM regulates the $\text{Ca}_v1.1$ and $\text{Ca}_v1.2$ by binding directly [37]–[39]. Previously, it is shown that CaM stabilizes the RyR2 closed state and prevents Ca^{2+} release which is called CaM-dependent inactivation [40,41]. Moreover, the R2474S mutation in the CPVT-associated mouse model PKA phosphorylation decreases the CaM binding to RyR2 which causes the spontaneous Ca^{2+} release leading to lethal arrhythmias [42]. There have been three CaM binding sites in the RyR2; CaMBD1 at residues 1940–1965, CaMBD2 at residues 3580–3611, and CaMBD3 at residues 4246–4275 [6,29,43].

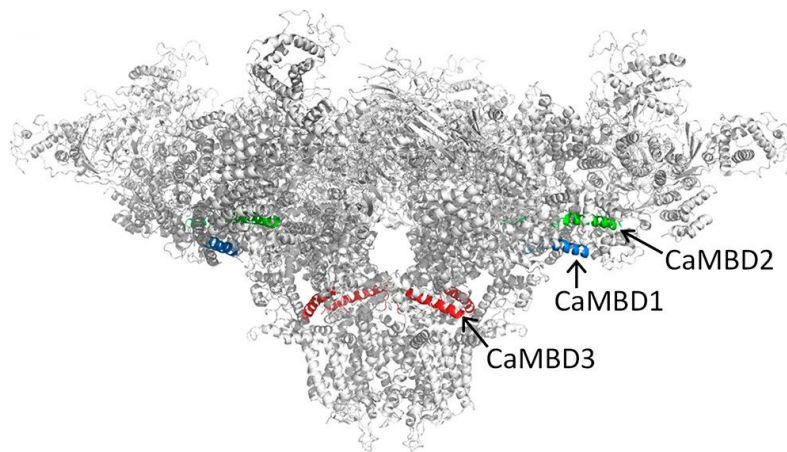


Figure 1.8 : The CaM binding sites. The figure is adapted from [6]

Similarly, FKBP is proposed as one of the stabilizers for the closed state of RyR2 [7]. There are two FKBP; FKBP12 binds the RyR1, and FKBP12.6 binds to RyR2 specifically. There have been many comprehensive studies carried out about the FKBP12.6 regulation on RyR2, but still, there are debates about its function. There are studies which claim that FKBP12 or FKBP12.6 have no effects on the open probability of RyR2. In contrast, recently, it is shown that both FKBP12.6 and FKBP12 reduce the store-overload-induced Ca^{2+} release (SOICR) [4]. Besides this, the FKBP is used to resolve the HD2 region in RyR2 by stabilizing the closed state of RyR2 during the Cryo-EM studies [7]. The binding site of the FKBP12 and FKBP12.6 has been mapped shown in Figure 1.9 [7]. FKBP12.6 is located at the clamp domain, specifically at the cleft formed by NTD-C, SPRY1 and SPRY2.

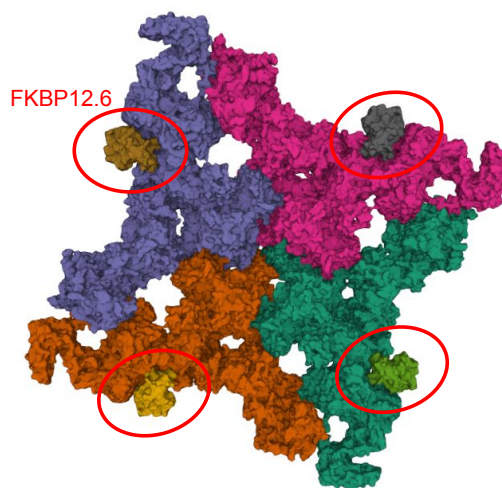


Figure 1.9 : The FKBP12.6 binding site for RyR2. The structure PDB ID 6WOV [7].

1.4.3 Dantrolene

Dantrolene ($C_{14}H_{10}N_4O_5$, Figure 1.10) is a FDA approved drug for malignant hyperthermia. The dantrolene binding sequence is defined at residues 590-609 in the RyR1 based on the evidence of photoaffinity labelling experiments, corresponding residues at 601-620 were defined for RyR2 as dantrolene binding sequence [44,45]. The defined sequence is conserved for all RyRs (Figure 1.11) [5]. Thus, its activity has been evaluated by experimental studies. Previous experiments showed that dantrolene is capable of limiting atrial fibrillation by inhibiting the RyR2 [46]. Moreover, dantrolene reduces the calcium waves and time spent after depolarizations in the failing heart [47]. In addition, it is shown that dantrolene provides an improvement in CPVT patients' health conditions. Recent studies claimed that it requires the CaM and Mg^{2+} [48,49]. Although dantrolene has been proposed as a potential drug [50,51], its chemical properties are not suitable for cardiac disease. Dantrolene is a poor water-soluble drug, therefore the therapeutic plasma level after oral administration can be reached by high dose usage [3]. Hence, potentially it causes liver damage due to the high dose. Besides this, it is suggested that dantrolene is a non-selective inhibitor for RyRs [52].

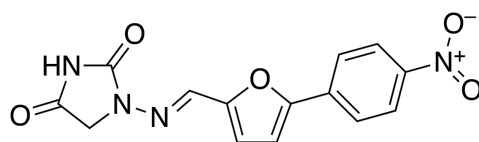


Figure 1.10 : Dantrolene chemical structure

Dantrolene binding sequence	590	LDKHGRN	NHKVLDVLC	SLCVC	609
skeletal muscle (RyR1)	590	LDKHGRN	NHKVLDVLC	SLCVC	NGVAVRSNQDLITENLLPG628
cardiac muscle (RyR2)	601	LDKHGRN	NHKVLDVLC	SLCVC	HGVAVRSNQHLICDNLPG639

Figure 1.11 : The sequence of putative dantrolene binding site for RyR1 and RyR2 [5].

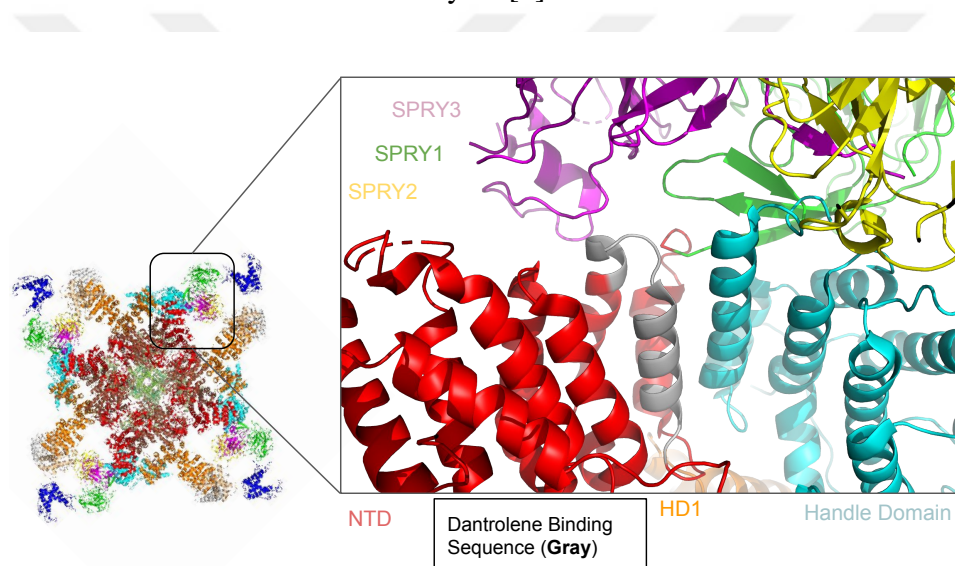


Figure 1.12 : The representation of the binding site of Dantrolene in PDB 5GO9 [2].

The binding sequence is located at the intersection of 6 different domains (Figure 1.12) and has close vicinity to FKBP12.6 and CaM. However, the Dantrolene-RyRs bound confirmation has not been elucidated yet.

1.5 Purpose of Thesis

In summary, specifically for RyR2, a large variety of the mutations are associated with a different types of cardiac arrhythmias. Known mutations mainly caused the Ca^{2+} leakage during the diastolic state in heart beating. The Ca^{2+} concentration is the

primary regulator for the activity of RyR together with other stimulants; ATP, Caffeine, Mg^{2+} , CaM and FKBB12.6. In this thesis, we have three objectives to achieve; 1) modelling of the full-atom structure of RyR2 using current state-of-art methods, 2) Dantrolene binding orientation determination using computational techniques, and 3) identification of novel drug candidates for cardiac disorders by targeting putative Dantrolene binding sequence.





2. THEORY AND METHODS

2.1 Molecular Modeling

Protein structure is crucial for understanding functional protein networks and disease pathways as well as developing possible treatment approaches. Thus, in the last couple of decades, molecular structure determination has been taken into interest by many scientists. Although the latest methodologies are significantly improved, like the Cryo-EM revolution, the experimental procedure for protein structure is challenging. Besides this, these methods are not sufficient for some protein structures that scientists are still working to predict. Therefore, computational studies become more relevant for saving money and time. In the past decade, computational methods for the prediction and design of protein structures have advanced dramatically and highly accurate protein prediction results may be obtained [53]. Historically, predicting the protein folding energy landscape is the foundation of today's prediction algorithms. The protein energy landscape is often described as a funnel consisting of dozens of local minima along with a global minimum which corresponds to the native form of the protein. By considering the thermodynamic or kinetic of proteins together with sampling and energy functions, these algorithms are developed to find these global and local minima [53]. Moreover, traditional protein modelling algorithms time consuming to design and demand large computational power. Gradient-based and Monte-Carlo sampling approaches are the two that are used commonly in algorithms. These algorithms are heavily dependent on system size, degree of freedoms (torsions, angels of atomic interactions) and computational power. Recently integrating evolutionary perspective, advances in sequencing, increases in the collection of protein structures via database, and the availability of GPUs allowed to development of efficient approaches using machine learning and deep learning algorithms which opened a new era in protein structure prediction. Current methods are separated into mainly two major approaches; template-based and template-free. Since sequence similarities of

the proteins provide closely related structural features like homology, template-based methods are designed to predict models using these known structures together with defined energy functions. Whereas, template-free methods do not necessarily need previously known structures. Ab-initio and de novo protein design methods can be given as an example of these approaches [54]. The template-free methods are designed as physical-based and knowledge-based [53].

2.1.1 Physical-based methods

These methods aims to construct folded protein using molecular mechanics force fields. Some of the well-know force fields can be given as AMBER [55], CHARMM [56], OPLS [57] and GROMOS [58]. The protein folding can be achieved by employing molecular dynamic simulations with sufficient time. Previously, this was achieved for up to 80 residues protein sequence in 2011 [59]. However, for moderate-sized proteins the simulation timescale often does not enough to capture folded protein. Moreover, the using force fields in the methods are also combination of the parametrization of atomic features and bonding interactions. Although these force fields are efficient for protein dynamic studies, still convergence of the force fields are issue for some cases. Likewise, bond breaking and formation can not be observed in the molecular dynamics simulations. Because of this, the disulfide bonds formation can not be mimic in the folding process. Since the disulfide bond formation is one of the key component for folding [59], this issue makes the methods unrealistic relative the natural folding process. Even though, these methods may produce sufficiently accurate models of protein physics in theory, these methods are highly challenging owing to computational cost as well as implementations difficulty on more complex proteins.

2.1.2 Knowledge-based methods

These methods are also called data-driven methods in which protein 3D conformations are predicted using the information coming from experimental high-resolution structural data, co-evolution or sequence alignments. The knowledge-based methods are surpassed the physics-based in terms of accuracy, availability and computational

demands [54]. Particularly, knowledge-based methods aim to approximate the atomic interactions, residue-residue and inter-residue distance and contact maps [54]. Over the past two decades, there has been a growing trend to use machine learning and deep learning algorithms in protein modelling. These can be seen recent Critical Assessment of Structure Prediction (CASP) contest which brings both experimentalists and computationalists to determine the state of the art methods in protein structure modelling. Deep learning methods dominated the last couple of years in CASP with outstanding prediction accuracy [60]. Google AlphaFold 2 (AF2) [8], RoseTTAFold [61] and trRosetta [62] are the leading algorithms among other methods.

2.1.2.1 AlphaFold 2

The AlphaFold 2 is recently release the state-of-art method for protein structure prediction. At the time of writing, after publishing source code and paper, in just 1 year, the paper has been cited by over 1600+ different scientific studies, and reviews. As they showed in CASP14 assessment, the AF2 is able model the given protein sequence's structure with high accuracy. Remarkably, it could predict some protein with 0.96 TM-Score and 0.59 RMSD value relative to experimental structures [8].

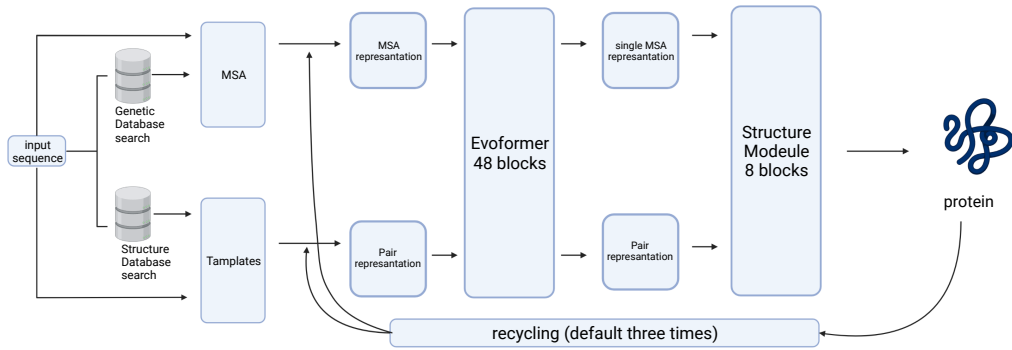


Figure 2.1 : The AF2 deep learning architecture representation adapted from [8].

AF2 architecture starts with preprocessing pipeline including construction of a multiple sequence alignment (MSA) data and an initial representation of the structure from the similar structures in database (Figure 2.1). This structural template is called “pair representation” in the architecture. In the MSA, AF2 generates large aligned similar sequences coming from the other organisms to capture the correlations,

mutations and especially co-evolution. The idea behind the co-evolution is that proteins tend to preserve the structure especially for catalytic sites, therefore the mutations in one amino acids highly possible to be followed by other mutation of closely contacted amino acid. Likewise, protein structures have tendency to maintain similar shape, even though they might have insufficient sequence similarity. It is due to the proteins mutations and evolution prone to protect functionality of the protein as well as protein-protein interaction network. These concepts are not new and unique for AF2, they have been used other protein modeling methods in last decade. The AF2 speciality comes from the its deep neuronal network in which two main stages have been constructed [8].

The first large network is the “Evoformer” which was built based on “a graph inference problem in 3D space” as stated in the original paper of the AF2. This network is designed to allow information exchange between MSA and pair representation to associate the evolutionary information with spatial perspective. This information exchange oscillates throughout to network continuously [8]. The overarching idea is that once the pair of amino acids are captured as they are in a close relationship, there may be another pair of amino acids with a close relationship relative to the previous pair of amino acids. By oscillating the information forth and back, the assumptions (relation of two amino acids) are adjusted based on the aforementioned idea to hypothesis new assumptions [8]. The MSA and pair representations are processed in two “transformer”, one for MSA and the other for pair representation. The MSA transformer is built to compute attention (in this case the informative sequences for the neural network) by employing the operations in a row-wise direction to detect the informative residue pairs and a column-wise direction to decide informative sequences. The other key feature in AF2 is that row-wise operation includes the information that comes from the pair representation which allows for identifying the residue pair interactions. For the second transformer, the pair representation is used to identify attention in three-dimensional space. They used “triangle multiplicative update” and “triangle self-attention” to manage to incorporate distance distribution information in 3D space. The procedure is implemented iteratively with 48 blocks (number of iterations) as described in the original paper. In the end, Evoformer produces

two representative data one comes from the MSA including the sequence variation and the other comes from pair representation which contains the residual interaction information [8]. This stage, then, is followed by a structure module to combine the data produced in Evoformer to construct the protein 3D structure. Here, the protein backbone for each residue (C-C α -N) is treated as a triangle which they called “residue gas”. This residue gas is represented as translations and rotations in affiliate matrices. Iteratively, these translations and rotations are updated using “Invariant Point Attention”. Basically, in this operation, local frame attributes of residues are projected in the global frame, and then this projection is represented in the local frame again. This procedure is applied three times as described in the paper [8]. In the end, the protein model with side chains is built.

2.1.2.2 trRosetta (transform-restrained Rosetta)

trRosetta is a highly effective deep learning-based protein prediction method. According to the benchmark test on the CASP13 test sets, it had outstanding results among the methods in that time [62]. trRosetta could manage to predict accurately the protein sequence structures which have no co-evolution information [62]. Besides this, it can predict the amino acid changes which are meaningful in a biophysical manner. Hence, it is a highly efficient method that can be integrated into different methods such as denovo protein design. Deep network hallucination, for instance, has been a recently released method based on the trRosetta architecture to design proteins which are not the naturally occurring [63].

The trRosetta is constructed with two stages. The first stage consists of a deep residual-convolutional network architecture which is designed to predict inter-residue orientations and distance from MSA data. As they stated in the paper, initial features were extracted from the MSAs which are “encoded amino acid sequence”, “position-specific frequency matrix” and “positional entropy” along with coupling matrices and average product correction score. After a simultaneous 2D convolution network, inter-residue distance matrices are produced. The inter-residue geometry is represented with 6 parameters which are d , ω , θ_{12} , ϕ_{12} , θ_{21} , and ϕ_{21} . From those, θ_{12} , and θ_{21} are planer angles, and ω , θ_{12} , and ϕ_{12} are dihedrals to represent the

orientation of a pair of amino acid. The last parameter is the d which represents the distance between the $C\beta-C\beta$ pairs [62]. In the second stage, these contacts are converted as Rosetta restraints followed by a constrained minimization procedure incorporating produced restraint subsets to generate protein 3D structure. The torsional quasi-Newton-based energy minimization is applied on a coarse-grained level with 3 consecutive steps. The short-range (distance $<12\text{\AA}$), medium-range ($12\text{\AA}$ distance $<24\text{\AA}$) and long-range (distance $24\text{\AA}$) restraints are applied respectively, with Rosetta energy terms. Full-atom relaxation with FastRelax [64] in Rosetta with the ref2015 scoring function, then, is performed using top-ranked 10 model. In the end, the final model is produced. Normally, trRosetta is not able to process the structure information coming from the templates like homologous structures [62]. Recently, a modified trRosetta was released in which template structure pairwise distances and orientations were incorporated into the original trRosetta network. The 2D feature matrix is generated by 2D convolutions, and this matrix is joined with the MSA features. According to their results, the trRosetta performance was enhanced for protein structure prediction. However, as in the AF2, in either original trRosetta or modified trRosetta, the method is not sufficient for intrinsically disordered regions structures and sequences having relatively low MSA input.

2.1.2.3 FastRelax

FastRelax is standard structure refinement based on the Rosetta Score function [64]. It does not perform comprehensive minimization it finds local minima near the starting position by optimizing the backbone and side-chain conformations [64]. The FastRelax operates several times (default is 5), and in each round lowest-score structure is returned. The optimization is performed based on the gradient-based minimization together with side-chain packing (fixes the lowest rotamer). This minimization is applied with torsional degrees of freedom. The key point in the relaxation protocol is the contribution of repulsive energy weight in the score function. The weight contribution is gradually increased in each round until it reaches the normal value. Typically, the contribution rate changes as 2%, 25%, 55% and 100% respectively [64].

2.1.2.4 Rosetta Score Function

One of the distinguished features of the Rosetta is "Scoring Function". The function combines statistical and physical energy terms within the linear equation system [65]. The overview equation (Eq 2.1) includes van der Waals energies, hydrogen bond terms, electrostatics interaction terms, disulfide bonds energy, solvation energy, backbone torsion angles energy, rotamer energies of side chains, and reference energy. These energy terms are weighted based on statistical, and physical assumptions as well as experimental studies. There are different score weights, and it is still in development. However, the most recent scoring function is *ref2015* that is using weights shown in Table 2.1 with descriptions as stated in the Rosetta manual.

$$E_{total} = E_{vdW} + E_{hbond} + E_{elec} + E_{disulf} + E_{solv} + E_{BBtorsion} + E_{rotamer} + E_{ref} \quad (2.1)$$

Table 2.1 : Rosetta Scoring terms parameters, default weights and descriptions.

Term	Weight	Description
fa_atr	1	Lennard-Jones attractive between atoms in different residues
fa_rep	0.55	Lennard-Jones repulsive between atoms in different residues
fa_sol	0.9375	Lazaridis-Karplus solvation energy
fa_intra_rep	0.005	Lennard-Jones repulsive between atoms in the same residue
fa_elec	0.875	Coulombic electrostatic potential with a distance-dependent dielectric
pro_close	1.25	Proline ring closure energy and energy of psi angle of preceding residue
hbond_sr_bb	1.17	Backbone-backbone hbonds close in primary sequence
hbond_lr_bb	1.17	Backbone-backbone hbonds distant in primary sequence
hbond_bb_sc	1.17	Sidechain-backbone hydrogen bond energy
hbond_sc	1.1	Sidechain-sidechain hydrogen bond energy
dslf_fa13	1.25	Disulfide geometry potential
rama	0.25	Ramachandran preferences (with separate lookup tables for pre-proline positions and other positions)
omega	0.625	Omega dihedral in the backbone. A Harmonic constraint on planarity with standard deviation of 6 deg.
fa_dun	0.7	Internal energy of sidechain rotamers as derived from Dunbrack's statistics
p_aa_pp	0.4	Probability of amino acid, given torsion values for phi and psi
yhh_planarity	0.625	A special torsional potential to keep the tyrosine hydroxyl in the plane of the aromatic ring
ref	1	Reference energy for each amino acid. Balances internal energy of amino acid terms. Plays role in design

2.1.3 Modeling of RyR2 homotetramer

The divide and conquer principle was adopted for RyR2 modelling. Each unresolved region was modelled individually and combined in the finale model. Up to date, RyR2 homotetrameric structures have been released under different conditions. These published structures were compared and the common unresolved regions were identified. Available structures contain similar unresolved regions consisting of several loops together with the five largest regions. Two of those regions are located in the HD2

domain, and one is in between HD2 and channel domains where the handle domain has direct interaction. One of the largest unresolved segment is located at the channel domain with a 280aa long which correspond to the 4206-4482 residues. We excluded this part for modelling because of disordered behaviour. Therefore, 4233 and 4472 residues were connected during modelling. A similar approach was performed by Heinz et.al about the RyR1 in silico assessment [66]. The last unresolved segment is in 1311-1424 at the SPRY3 domain it is most likely disordered. All regions were modelled by employing AF2. Besides this, trRosetta was used for 1311-1424 and 2905-2990 unresolved regions modelling. Among unresolved regions, AF2 was not sufficient to model the disordered region 1311-1424. After performing trRosetta for this region, SPRY3 and modelled segment were optimized by 500ns MD simulation with SPRY3 position restraints. This simulated system, then, was added to the final model. All models were combined using 5GO9 and 6WOV Cryo-EM structures. We manually aligned the overlapped segments by PyMOL and Rosetta FastRelax were performed on model intersections and 20Å around to minimize steric clashes. The obtained model of RyR2 was optimized in the membrane by employing 35ns MD simulation using NAMD [67]. Membrane building was carried out with the Charmm-gui web-server “Membrane Builder” plugin [68]. Membrane bilayers were constructed POPC and Charmm-gui default membrane melting steps were followed. 6 equilibration steps were performed (250 ps for 1-3 steps, 500ps for 4-6 steps). In the equilibration steps, protein restraints were applied with 10, 5, 2.5, 1.0, 0.5 ve 0.1 kcal mol⁻¹Å⁻² force constants respectively along with lipid head groups restraints with force constants 2.5, 2.5, 1.0 ve 0.5, 0.1 ve 0.0 kcal mol⁻¹Å⁻². The system was neutralized by adding counter ions which were restrained in the first step of equilibrium with 10 kcal mol⁻¹Å⁻². In the final step, only protein heavy atoms were harmonically constrained and under NPT conditions (298K and 1 bar), MD simulation was performed to optimize the system.

2.2 Molecular Docking

Molecular docking is a well-known computational approach for assessing protein-small molecule and protein-protein interactions, bound conformations and

binding affinity. The last couple of decades have seen a trend toward docking studies due to its easy operation and moderate binding affinity predictions for small molecules. Even though docking studies achieved great success in drug design, it is also possible that it can produce incorrect predictions. Besides this, the binding affinities are by far away from the predicted experimental binding affinities. It is a practical and beneficial use for identifying active small molecules by screening large chemical libraries [69].

There have been developed different docking algorithms to get better scoring functions and searching algorithms. Traditionally, docking algorithms are designed with two main phases; a conformational search and scoring based on a function [70].

2.2.1 Conformational Search for small molecules

Most of the molecular docking algorithms are designed to search ligand binding conformations within a defined searching space which is called grid box. Docking considers the flexibility of the receptor and ligand. Generally, two methods are common which are rigid receptor-flexible ligand and flexible receptor-flexible ligand. In receptor flexibility, either side-chains or backbone for amino acids are considered as flexible and translation, rotations of bonds are modified. The current limitations of these methods are that upon increasing rotatable bonds and amino acids, calculation needs are exponentially increased. There are three most common types of searching approaches; systematic search, stochastic search, and MD simulation based [70].

Systematic search, also called exhaustive, is an algorithm in which all possible rotational and translational parameters of ligand are systematically modified through the whole conformational space. The number of rotational bonds and size of the conformational space are the main limiting factors for this algorithm. Thus, some of the algorithms are designed to divide subsets to search whole conformational space [70].

Stochastic search is based on randomness, and it is also called Random search. Basically, random changes are applied on ligands until the algorithm reaches the defined criteria. There are two common examples; Monte-Carlo, and genetic algorithms. In Monte-Carlo algorithms, initial conformations of small molecules are taken as random.

Small changes are applied on these conformations. If the calculated energy score is improved, that conformation is retained, otherwise the previously defined Metropolis criterion is performed to accept or reject the conformation. This Metropolis criterion is commonly calculated using a Boltzmann-based probability function. Besides this, different docking algorithms have designed their own specific Metropolis criteria for conformational search. The genetic algorithm is one of the well-known and most used algorithms. The idea comes from biological evolution in which mutations and crossover events take place. Basically, the parameters of small molecules such as translation, rotation, and conformations are considered as state variables (also called gene). These state variables, then, are assigned into a chromosome with a given fitness function. In each iteration, mutations and crossovers are performed, and mutations and crossovers with a lower fitness score are inherited to the offspring of the systems [70].

2.2.2 Scoring Functions

Scoring functions are the calculations for the approximate binding free energy of ligands or peptides. Reliable scoring functions are critical for docking algorithms. Some of the well-predicted binding poses might not be ranked because of scoring functions. Because, even if a highly accurate binding pose is detected by a searching algorithm, it could not be selected as a candidate due to the low energy score. Thus, the scoring functions are one of the bottlenecks for docking calculations. Therefore, the binding pose and scores of a ligand should be assessed by different docking algorithms to evaluate binding interactions. There are three common types of scoring functions which are force field-based scoring, knowledge-based scoring and empirical scoring functions [70].

Non-bonding interactions such as coulombic interactions, van der Waals and hydrogen bonds are parametrized in the force field-based scoring. The idea is similar to molecular dynamic simulations. The interaction energies between small molecules and the receptor are calculated and summed using these parameters. The details of force-field and bonding parameters are described in the Molecular Dynamic Simulations section [70].

In the knowledge-based scoring, statistical analyzes that approximate experimental data are used to calculate binding energy. Intermolecular contacts are evaluated statistically based on the structural information available from databases. The idea is that defined type of interactions frequency should have similar contributions to binding energy [70].

In the empirical scoring, non-bonding interactions of ligand-receptor complexes are defined using physical and chemical terms and coefficients. These coefficients are fit to experimental data to approximate the binding free energy. AutoDock Vina is one of the docking algorithms which it combines knowledge-based potentials and empirical-based scoring to estimate binding energy [70].

2.2.3 Docking methodology

In the literature, the putative binding site in the RyRs of the dantrolene is given as 590-609 for RYR1 and 601-620 residues for RYR2. However, there is no experimental 3-D structure. Dantrolene potential binding poses were assessed using three different programs which are Glide extra-precision (XP) mode [71], AutoDock VINA [72], and Ledock [73]. A recent benchmark study showed that these programs have reliable results among other docking programs.

2.2.3.1 Autodock Vina

Autodock Vina is a user friendly, robust and fast docking algorithm [72]. Vina is widely used that the original paper of Vina has been cited in over 17k studies according to Google Scholar citation information. In Vina, the binding affinities are predicted using a score function in which the contribution of important energy terms for binding are summed up based on atom pair distance (Eq 2.2). Each energy term is normalized with a weight, and depending on atom type hydrophobic and hydrogen bond energy terms are including the energy prediction calculation (Eq 2.3). The weights are listed in Table 2.2. Conventional docking algorithms calculate the binding affinities using physics-based terms including the 6–12 van der Waals interactions and electrostatic energies. Instead of using these term functions, three energetic terms, which are gauss₁, gauss₂ and repulsion are designed

to approximate the Lennard-Jones interaction graph for steric effects. Previously, AutoDock4 uses Lamarckian Genetic Algorithm as a searching algorithm. In Vina, Iterated Local Search global optimizer is performed in which mutations and Broyden-Fletcher-Goldfarb-Shanno local optimization are performed in each step. The acceptance and rejections are evaluated according to the Metropolis criterion.

$$E = \sum e_{pair}(d) \quad (2.2)$$

$$e_{pair}(d) = w_1 * Gauss_1(d) + w_2 * Gauss_2(d) + w_3 * Repulsion(d) + w_4 * Hydrophobic(d) + w_5 * Hbond(d) \quad (2.3)$$

Table 2.2 : AutoDock Vina scoring algorithm scoring weights.

Weight	Term
-0.0356	gauss ₁
-0.516	gauss ₂
0.840	repulsion
-0.351	hydrophobic
-0.587	hbond

2.2.3.2 Glide

Glide is another rapid and precise docking algorithm. Essentially, two modes are available in Glide; standard-precision (SD) and extra-precision (XP) for employing scoring functions. Ligand poses are generated from the series of hierarchical filters established in both methods. Particularly, XP is an expanded version of SD in which the XP filter consists of a more sophisticated scoring function. XP is designed to avoid false-positive results and provide favourable score-pose correlation by applying penalties based on physical chemistry [71].

Glide docking start with ligand conformation generation that feeds the docking protocol funnel depicted in Figure 2.2. Ligand conformation generation is essential and advantageous to eliminate unsuitable conformers by heuristic screen. For these, each ligand is divided into two fragments; core which is the mostly ring conformations

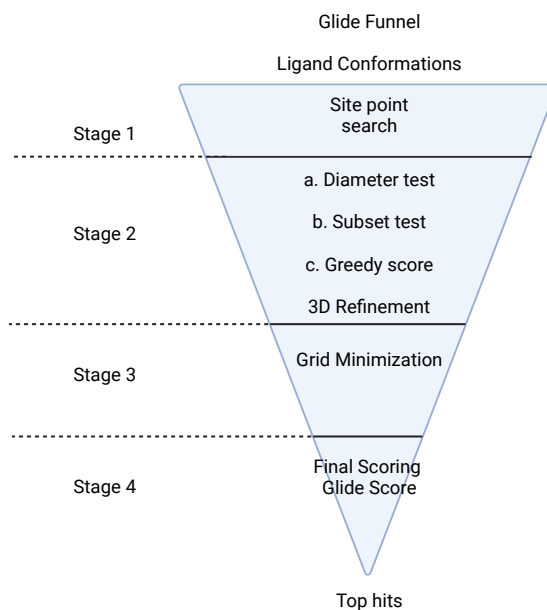


Figure 2.2 : Schematic representation of the GlideXP workflow.

and rotatable bonds. Core regions can be described as what is left if all rotatable fragments are removed and represented by couple of different parameters. Whereas rotatable bonds are described as the bonding interaction has ability to rotate and change the overall ligand conformation. Example of the ligand core region and rotatable bonds are represented in Figure. These ligands, then, are used in docking stages of Glide [71].

Produced ligand conformations are filtered by performing exhaustive search in defined active site. In the Stage 1, the active site is represented with “site-points” which are set of points spaced with $2\text{\AA}$ grids. The site-points includes “bins” which are sorted values of the $1\text{\AA}$ width distance between protein surface and site-point for a predefined specified directions. Similarly, for each ligand, “bins” ($1\text{\AA}$ width) are calculated and sorted using distance between midpoint and surface of ligand. These bins are compared for each site-points to find good matched. In the Stage 2, there are three sub-steps which are diameter test, subset test and greedy score. In diameter test, the orientation of the atoms in the previously defined conformations is decided based on the diameter calculated as largest distance of two width atoms. Next, ligands with moderate steric clashes are selected for next substep. After this step, in the subset test, selected orientations from previous steps are rotated and pre-scoring is applied on

atoms which have capability to hydrogen bond and metal interaction. Among those, if the score is considered as good enough, all other interactions are scored at greedy score step. Scoring of these steps are carried out using ChemScore empirical scoring function. Best greedy scores are, then, used to re-score along with 1 Å rigid movement is allowed. In the third stage, the best poses coming from the stage 2 are minimized using pre-calculated OPLS long range nonbonded interactions (vdW and Coulombic interactions). After minimization, these poses are re-scored using GlideScore function. Unlike the ChemScore, the steric-clash terms are incorporated in the GlideScore. The GlideScore is shown in Eq 2.4.

$$GScore = 0.05*vdW + 0.15*Coul + Lipo + Hbond + Metal + Rewards + RotB + Site \quad (2.4)$$

GlideScore function is an empirical function that is linear combination of electrostatic, van der Waals, hydrogen bond energy contributions together with the rewarding or penalizing interactions occurred in ligand binding as well as lipophilic and metal ligand interaction energy terms. Two scoring function is used in the Glide that one is for “Standard Precision” (SP) and other one is for “Extra precision” (XP) GlideScore. The difference is between those two is that GlideXP includes additional energy terms in the “Rewards” which can be beneficial or add penalty. The “Rewards” terms includes the features that is mostly desirable for ligand binding such as buried polar groups and hydrophobic enclosure (included in the GlideXP Score function).

2.2.3.3 LeDock

Unlike the AutoDock Vina and Glide, LeDock is not well documented. As the developer claimed the docking algorithm is designed based on simulated annealing together with genetic algorithm optimization to search the orientation and bonds for the ligand and it uses the hybrid scoring function to calculate binding affinity [73], although the detailed description has not been available today. However, the recent benchmark studies represent that LeDock outperformed other docking programs [74]. Besides this, other studies suggested that LeDock is fast and has high accuracy.

2.2.4 Dantrolene Binding Pose Prediction

Although the dantrolene binding site is given as residues 590-609 in RyR1 and 601-620 in RyR2, the dantrolene-RyR bound complexes have not been elucidated yet. Herein, dantrolene RyR2 bound orientation was predicted using three different docking programs namely AutoDock VINA, LeDock and Glide for increasing correct binding pose estimation. Due to the disagreement on sequence order between AF2 model and Cryo-EM structures for SPRY3 domain, we applied further protocols for either truncated AF2 model and Cryo-EM structure taken from PDB 5GO9. Truncated structures were generated from residues 410-1920 except Repeat 1-2 domain corresponds to 855-1082 residues. The missing residues in truncated 5GO9 were modeled using trRosetta. The exact dantrolene binding location is unknown. Therefore, several grids with 30Å in size and different centers were generated to screen the entire area around the residues 601-620. These grids were used in rigid docking protocols of three programs. The densities of docked dantrolene positions were evaluated and the most populated region was selected as a putative dantrolene binding site for following steps. Dantrolene parameters which were generated in our group were used (unpublished data). The docking score rankings and confirmations were compared for each docking program. We selected common structures with higher ranks among docked confirmations. To assess conformations binding behavior, MD simulations were carried out using NAMD. In all MD simulations, we used new generated dantrolene parameters by our group (unpublished data). Large systems demand high computational power to run simulations. To overcome this issue, MD simulations were performed using truncated systems. The solvated dantrolene-RyR2 bond complex systems were generated using the Charmm-gui web server solution builder. During the simulations, position restraints were applied on residues where there is an interaction with other than truncated structure and other protomer domains to mimic RyR2 dynamics. All simulations were run 200ns under 298K and 1atm NPT conditions.

2.2.5 Identification of potential inhibitor for RyR2

The candidate inhibitors were selected after series of docking protocol which are shown in Figure 2.3. In the scope of this study, 3.5 million molecules available at ZINC15 database were screened [75]. The screening library was constructed with <425Da MW, <3.5 logP molecules. The filtering was performed in three steps in which 20Å grid box is generated with a center of R606 and E1634 residues center of mass to cover all binding area. The box size and location was depicted in Figure 2.4. The box center was selected based on the dantrolene-truncated RyR2 complex MD simulation results.

Using Vina with 8 exhaustiveness, first screening was performed on 3.5 million molecules. Next two screening was employed on top 200k molecules with LeDock and Vina (24 exhaustiveness). The common molecules within the top 10k molecules were filtered and used in third step (GlideXP docking procedure). In this step 517 molecules were used and top 20 molecules “Percent Human Oral Absorption” properties were inspected using Maestro (v2020.1).

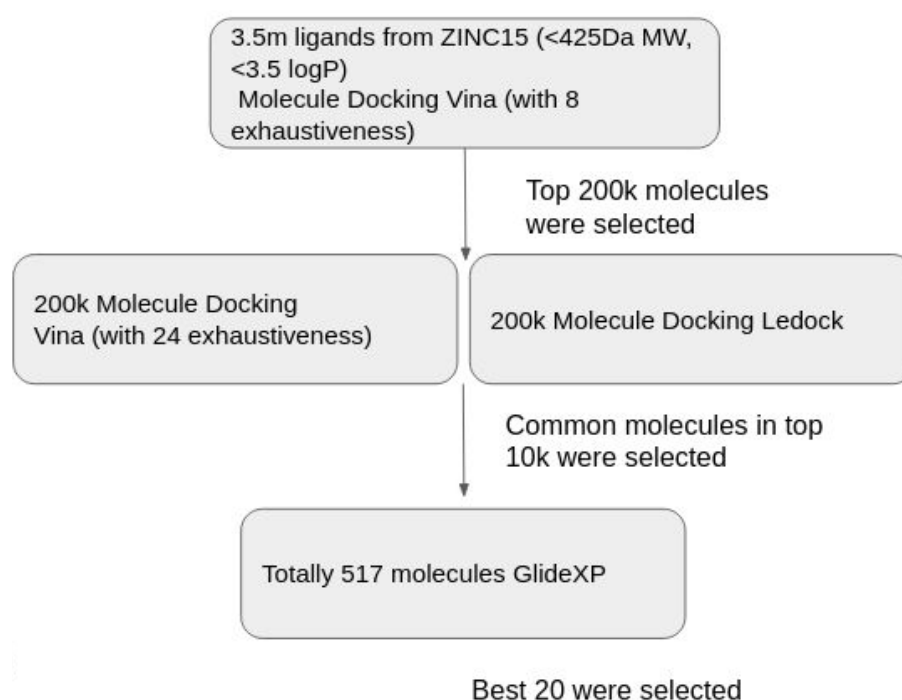


Figure 2.3 : High Throughput Screening workflow

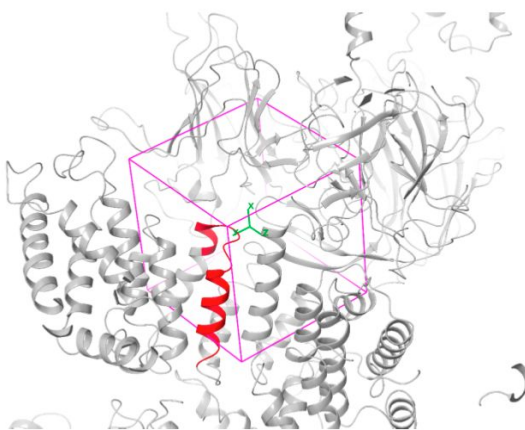


Figure 2.4 : The grid box used in HTS

2.3 Molecular Dynamic Simulations

Molecular dynamic (MD) simulations are a highly powerful approach to investigating protein functional dynamic, protein-ligand binding, and protein-protein interactions studies [76]. It is based on Newton's law in which atomic accelerations are calculated based on the force applied. With increased computational power, accessibility of high-performance computers and integration of GPUs, MD simulations become more popular and standard routine procedures in scientific studies. MD systems can be represented in either atomistic or coarse-grained and in each representation the systems are parametrized termed as a force-field. The idea for MD simulation is simple, however, it highly depends on the force field. Thus, scientific committees continue to improve and develop force fields to obtain accurate and reproducible simulation systems.

2.3.1 Molecular Dynamics Algorithms

Basically, MD simulations follow the Newton's second law (Eq 2.5) [77]. This is a well-known principle that force on a particle equals the mass of that particle and its acceleration [77]. Since the force can be calculated from the potential energy relative to the coordinates of the particle, we can calculate the trajectories of the particle depending on time. Today, the Verlet integration equation derived from a Taylor expansion (Eq 2.6, Eq 2.7, Eq 2.8) is widely used to solve the motion of molecules

in a simulation system [77].

$$F_i = m_i \times a_i = m_i \times \frac{dv_i}{dt} = m_i \times \frac{d^2 r_i}{dt^2} \quad (2.5)$$

The new position of an atom $t + dt$ is calculated from its previous position $t - dt$ together with its acceleration [77]. However, the velocity is not directly integrated into the system, thus temperature changes are not included in the equation explicitly. To solve this issue, variants of this equation have been developed; the most common two integrations are the leap-frog and velocity Verlet algorithms. In the leap-frog, the velocity is calculated at $(t + 1)/2dt$ for defined position, and then the new position at $t + dt$ is calculated based on the velocity (Eq 2.9, Eq 2.10). Unlike the leap-frog, velocity Verlet is calculated velocity at the same $t + dt$ (Eq 2.11, Eq 2.12).

$$\vec{r}(t + \delta t) = \vec{r}(t) + \delta t \vec{v}(t) + \frac{1}{2} \delta t^2 \vec{a}(t) + \dots \quad (2.6)$$

$$\vec{r}(t - \delta t) = \vec{r}(t) - \delta t \vec{v}(t) + \frac{1}{2} \delta t^2 \vec{a}(t) - \dots \quad (2.7)$$

$$\vec{r}(t + \delta t) = 2\vec{r}(t) - \vec{r}(t - \delta t) + \delta t^2 \vec{a}(t) \quad (2.8)$$

the systems are controlled in terms of temperature, pressure, and number of particle [76]. There are several different approaches called thermodynamical ensembles to control system. NVE, NVT and NPT are the most used ensembles. NVE is known as “Microcanonical ensemble” in which number of atoms (N), volume of the system (V) and Energy are fixed as constant value. However, very small fluctuations can be seen during the simulation. Canonical Ensemble which is represented as NVT is the another ensemble that number of atoms (N), volume of the system (V) and temperature of the system (T) are fixed. The NPT termed as Isobaric-Isothermal Ensemble is the highly used ensemble in which number of atoms (N), pressure of the system (V) and temperature of the system (T) are conserved.

$$\mathbf{r}(t + \delta t) = \mathbf{r}(t) + \mathbf{v}(t + \frac{1}{2} \delta t) \delta t \quad (2.9)$$

$$\mathbf{v}(t + \frac{1}{2} \delta t) = \mathbf{v}(t - \frac{1}{2} \delta t) + \mathbf{a} \delta t \quad (2.10)$$

$$\mathbf{r}(t + \delta t) = \mathbf{r}(t) + \mathbf{v}(t) \delta t + \frac{1}{2} \mathbf{a} \delta t^2 \quad (2.11)$$

$$\mathbf{v}(t + \delta t) = \mathbf{v}(t) + \frac{1}{2}[\mathbf{a}(t) + \mathbf{a}(t + \delta t)]\delta t \quad (2.12)$$

In each ensemble, system volume and atom numbers are kept constant by employing periodic boundary conditions. Here, the systems are produced with imaginary replicas that if any particle escapes from the defined simulation box dimensions, the counterpart of the particle in the imaginary simulation box goes into the real simulation box (Figure 2.5). Hence, it allows to generate continuous system and keep to system volume and number of particle constant.

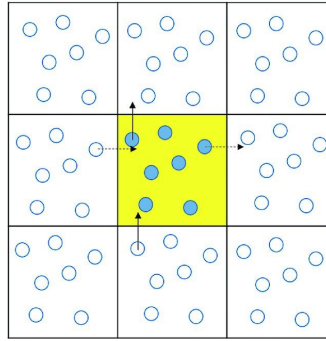


Figure 2.5 : Schematic representation of periodic boundary condition

Moreover, different methods and algorithms have been designed for temperature control. The most common algorithms can be given as the Nosé-Hoover thermostat, Langevin thermostat and Berendsen thermostat. The idea behind the Nosé-Hoover thermostat is the virtual variable which can increase or decrease the temperature is introduced to the system to reach the target temperature. Hence, temperature oscillates around the target temperature with a very small change. Besides this, the Nosé-Hoover thermostat can be integrated into the velocity Verlet. For the Langevin thermostat, the system is exposed to random virtual collisions with various magnitudes at a defined time frequency to control the temperature. Due to the collisions, the temperature is dynamic that it helps to approximate real-life dynamics and cross the energetic barriers that we may not be seen during the time scale of simulations. In the Berendsen thermostat, a coupling parameter, τ , is defined and applied at every dt . This parameter is recalculated and rescaled according to the target temperature during the simulation.

Because of this, the temperature oscillates around the target temperature that which makes the system more dynamic and realistic.

2.3.2 Force fields

Force field is the description of energy parameters such as bonded interactions and nonbonded interaction terms within the protein [76]. The total potential energy of a particle is represented as a linear combination of these terms (Eq 2.13). The bonded terms include bond stretch, angle bending and dihedrals, whereas the non-bonded terms consist of electrostatic, and van der Waals energy terms. In addition to these improper torsions, energy terms have also added the equation [76].

$$E_{Total} = E_{Bonded} + E_{Non-Bonded} \quad (2.13)$$

$$E_{Bonded} = E_{bond} + E_{angle} + E_{dihedral} \quad (2.14)$$

$$E_{Non-Bonded} = E_{elec} + E_{vdW} \quad (2.15)$$

2.4 Free Energy Calculations

Experimental calculations for binding free energy between ligand-receptor are time-consuming together with their high costs. Therefore computational methods become more important to estimate the binding free energy of desired molecules [78]. There have been different methods developed such as the docking score algorithms, the linear interaction energy, MM-GBSA and MM-PBSA, and alchemical free energy calculations like Free Energy Perturbations and thermal integration approaches. The current methods aim to predict binding energy with computationally efficient and accurate. Among those methods, The MM/PBSA and MM/GBSA methods provide sufficient results with low time need and computational cost which make them highly popular in the computational drug design research field [78].

2.4.1 MMPBSA

The free energy calculation is performed as a combination of standard MM energy terms including bonding interactions, electrostatic term, van der Waals term, and

solvation terms as well as an entropic term (Eq 2.16) [79]. Solvation is represented as two terms; polar (G_{pol}) and non-polar (G_{np}) solvation terms. Basically, in the MM/PBSA, the polar solvation term is calculated from the Poisson–Boltzmann (PB) (Eq 2.18). The non-polar term comes from the solvent accessible surface area. TS is obtained from the vibrational frequencies normal-mode analysis. These terms are computed for either complex, protein and ligand individually and from the equation, the binding free energy of ligand is obtained (Eq 2.20) [79].

$$\Delta G = \Delta H - T\Delta S = \Delta E_{MM} + \Delta G_{Sol} - T\Delta S \quad (2.16)$$

$$\Delta E_{MM} = \Delta E_{int} + \Delta E_{ele} + \Delta E_{vdW} \quad (2.17)$$

$$\Delta G_{sol} = \Delta G_{PBGB} + \Delta G_{SA} \quad (2.18)$$

$$\Delta G_{SA} = \gamma SASA_{int} + \beta \quad (2.19)$$

The force field, solvation method, sampling and conformational entropy are crucial for accurate MMPBSA results. Besides this, the charge model and dielectric constant are other methods that MMPBSA is highly dependent [79].

$$\Delta G_{bind} = \Delta G_{complex} - \Delta G_{protein} - \Delta G_{ligand} \quad (2.20)$$

Dantrolene binding affinities are calculated cafe1.0 VMD plugin [80]. The workflow

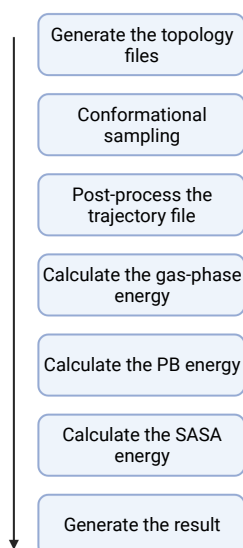


Figure 2.6 : CaFE workflow

for cafe1.0 is given in Figure 2.6. The conformational sampling is provided from the trajectories obtained from the MD simulations. Total 2000 frames had been collected from simulation and the 200 frames were extracted with 10 frame intervals. These 200 frames were used for rest of the energy calculation. The PB energy term is obtained using APBS. Unlike the default protocol, we calculated the binding energy for each frame to observe the energy change.



3. RESULTS AND DISCUSSION

3.1 Modeling of the RyR2

Before starting the modelling procedure, PDB 5GO9, 6IJY, 5L1D and 6WOV Cryo-EM structures were compared to determine the common missing segments. Among those, 6WOV structure has been published most recently and it is only RyR2 structure having well resolved HD2 domain in a large-scale [7]. Likewise, the HD2 domain was resolved in 5L1D structure, yet these are mostly secondary structure backbones with unknown residues and the resolution of the structure is not sufficient [13]. Overall 3D available RyR2 structures are agreed each other and have common missing segments (Figure 3.1, Figure 3.2). However, after residue number 1944, we observed that there are differences in PDB structures together with their sequence order even if secondary structures are similar (like in Figure 3.8.a). Mostly these regions are unresolved, low resolution or higher B-factor regions such as HD2.

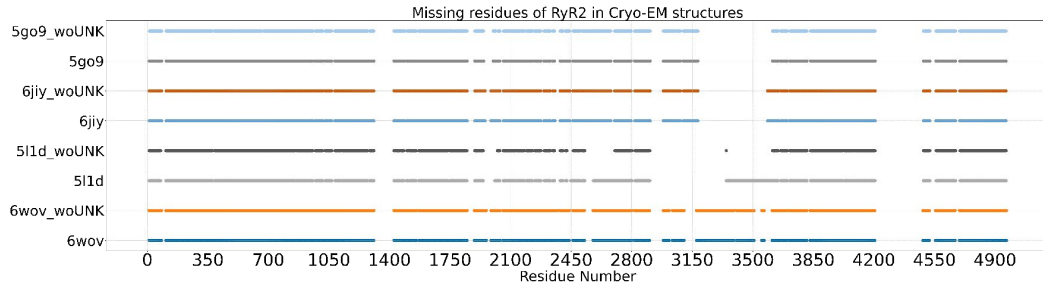


Figure 3.1 : Sequence comparison of RyR2 Cryo-EM structures

We performed state-of-art method AF2 and trRosetta to model the full structure of RyR2. There is a known drawback for the AF2 algorithm that it cannot sufficiently model if the target sequence has low MSA or it is disordered [8]. Previously, the 1311-1424 residues which have proximity to the dantrolene binding sequence have been considered disordered. This sequence is located in the SPRY3 domain and it has not been resolved yet. Moreover, the RyR1 counterpart (1295-1430 residues) for

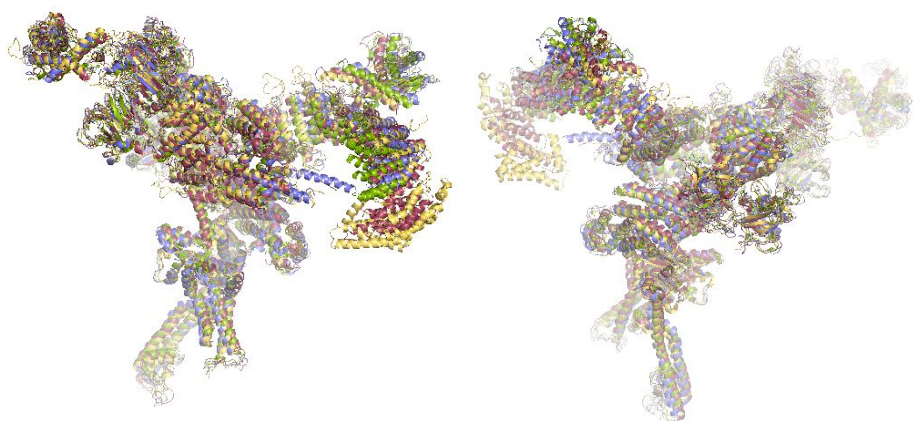


Figure 3.2 : The 3D representation of a proteomer taken from available RyR2 structures; PDB 6JIY (blue), 5L1D (red), 5GO9 (green) and 6WOV (yellow).

this region has a 35.6% similarity (according to P21817 RyR1 Human and Q92736 RyR2 Human UniProt Query IDs). AF2 was applied for this region using ColabFold with default parameters. We obtained an unfolded model with 17 MSA (Figure 3.3). Therefore, trRosetta was used to model this region with and without SPRY3. The stability of the model was reviewed by employing 500ns MD simulation. In both systems, especially 1350-1385 residues are prone to interact. Even if the trRosetta is not efficient for disordered regions [62], these results are enough to proceed with this model. Hence, the trRosetta model was integrated into the final model. Besides this, the SPRY3 domain has proximity to the dantrolene binding sequence which corresponds to 601-620 residues. SPRY3 domain is resolved in all available structures with different resolutions and quality. For 6WOV and 5GO9 comparison, beta-sheets are poorly formed in the 6WOV, whereas these secondary structures are sufficiently modelled in 5GO9 shown in Figure 3.4. Besides this, the MolProbity score [81] for these domains is obtained as 2.07 (72nd) for 5GO9 and 3.15 (18th) for 6WOV. While the 1311-1424 residues were modelled with the SPRY3 domain using AF2, we observed sequence order shifts in 1-2 residues for Cryo-EM structures without structural change around 1576-1596 residues (Figure 3.5). AF2 model is highly similar to RyR1 SPRY3 domain sequence order for this interval. Because of the proximity to the dantrolene binding sequence, the change in this region might affect

the dantrolene binding orientation. Thus, we proceeded with these two models for docking calculations.

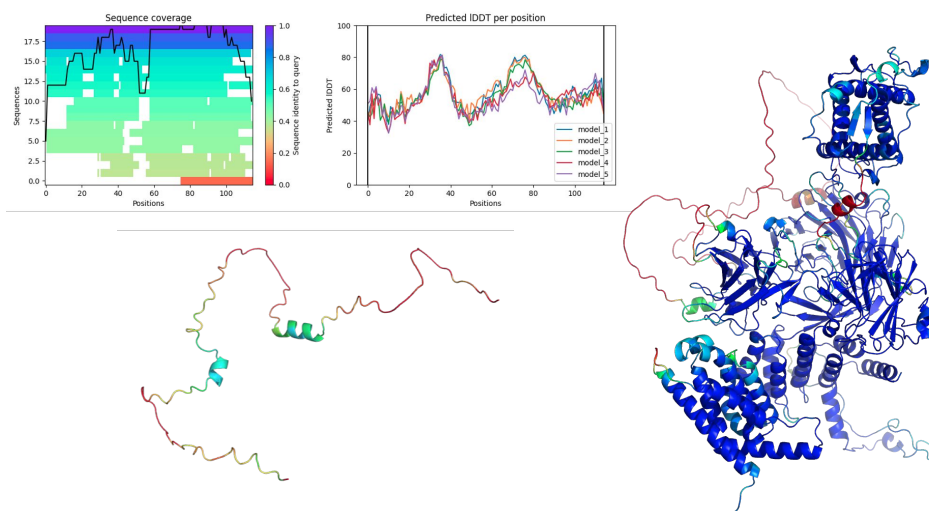


Figure 3.3 : AF2 Model results for at 1311-1424 residues.

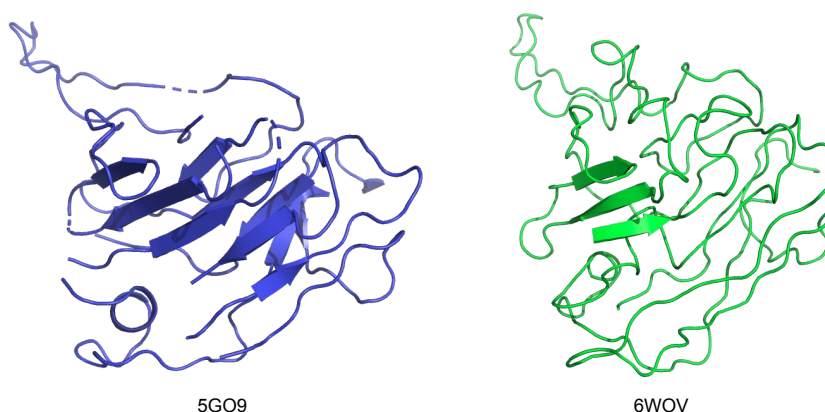


Figure 3.4 : SPRY3 doming for 5GO9 and 6WOV PDB structures

Besides the AF2 model, there is a distinct difference between the 6WOV PDB structure and other available structures at the dantrolene binding sequence. The 601-620 sequence is commonly seen as an alpha helix between H608-C620 together with a small loop at L601-N607. This structure orientation is seen in all RyR2 homotetramer Cryo-EM structures except 6WOV. It is seen that K609 bends towards C620 and forms crab clamp-like conformation shown in Figure 3.6. We suspected that this might be an artefact due to the Cryo-EM model production. Therefore, the sequence was assessed by examining the Ramachandran plot and electron density consistency. ChimeraX

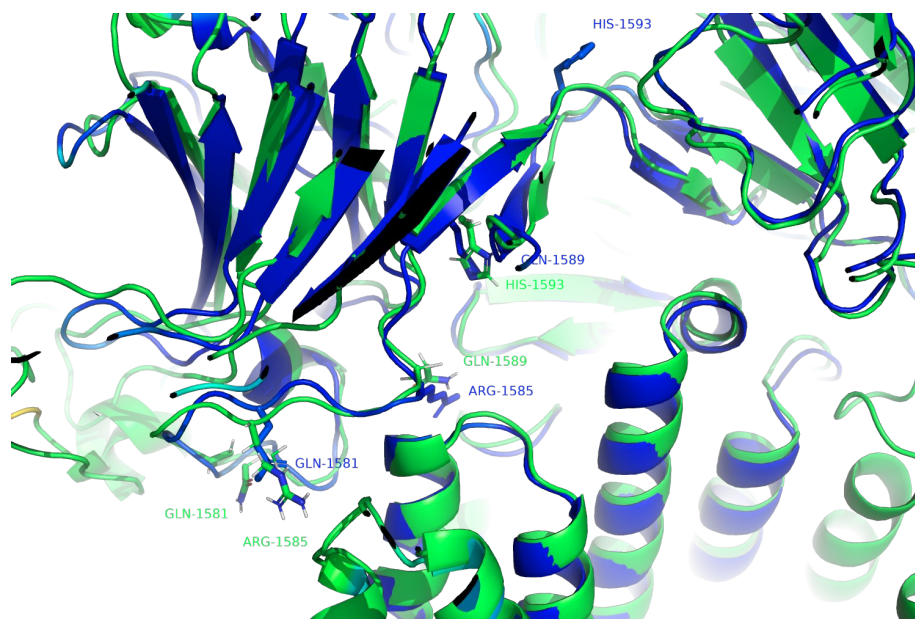


Figure 3.5 : Closer look at at 1576-1596 residues near the dantrolene binding sequence. AF2 model represented as blue, and 5GO9 represented as green color.

ISOLDE [82] software was used for calculations and density fit for structure. The density of the structure was taken from the EMDDataBase with 21862 ID [7]. According to the Ramachandran plot for 590-630 sequences, there is an outlier at N607 in 6WOV and cis-trans transition may cause the sequence bending that is seen in the structure. Moreover, it was observed that the sequence in the 6WOV could not fit well with electron density. Considering these results, the 6WOV structure for these sequences was not included in the subsequent steps.

Two helical domains; HD1 (2110-2679 residues) and HD2 (2982-3390 residues) form a α -solenoid 2 structure. Previously, it is shown that this region have higher flexibility [2,7]. Consequently, resolving the full structure of HD2 domain is challenging. The models for the missing segments are located at 2903-2982 and 3103-3174 residues according to 6WOV. These regions were constructed using AF2. Unlike the available structures, these missing segments were observed as a part of the α -solenoid structure (Figure 3.7). Particularly, These segments suggested as HD2 domain helices. This integration to HD2 changed the α -solenoid sequence and structural orientation and caused the structural shift through the HD2 domain (Figure 3.7.b). Similar outcome

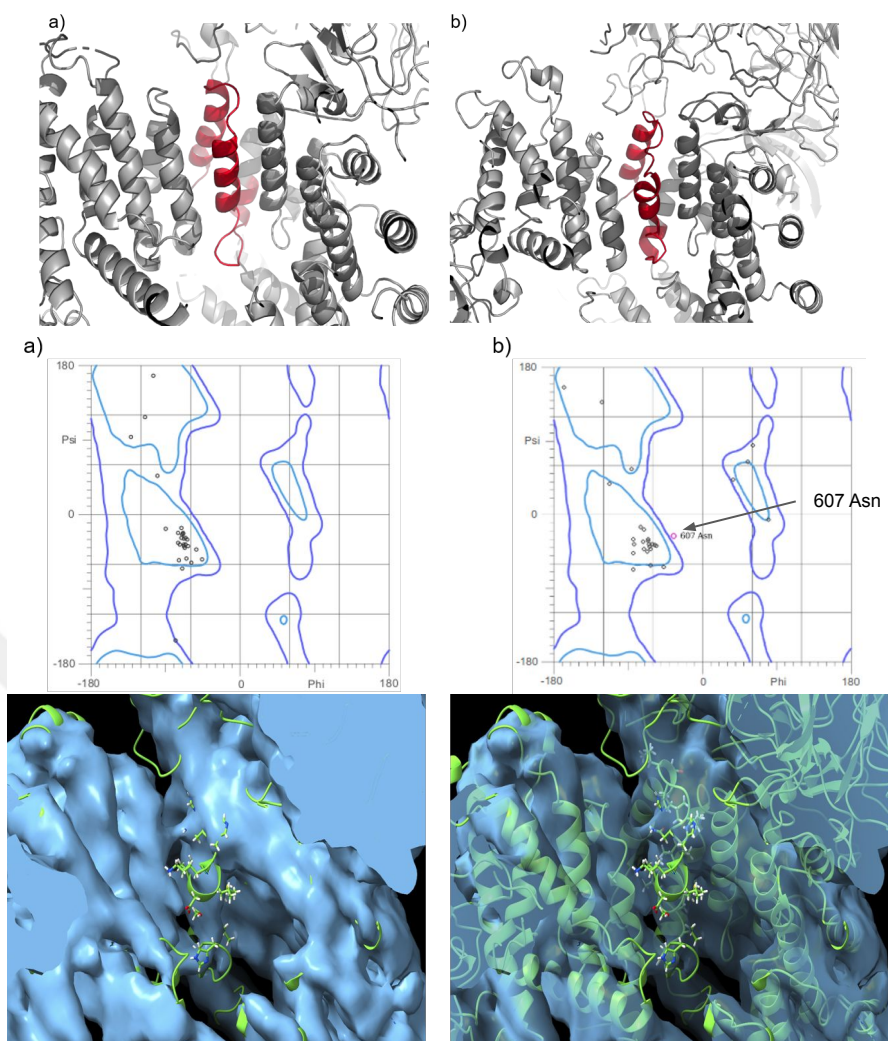


Figure 3.6 : Dantrolene binding sequence conformation in 5GO9 and 6WOV PDB structures.

has been obtained, when trRosetta was applied to truncated sequence to model 2903-2982 residues (Figure 3.7).

HD2, and Handle domain together with central domain form a cleft that the residues 3583-3601 making a bridge between HD2 and central domain. Previously, this sequence was identified as the calmodulin-binding target and the interaction between calmodulin and this sequence has been determined in several X-Ray and Cryo-EM structures (PDB ID 6JIY) [83]. The predicted AF2 model for this region was not well defined, probably due to the disordered behaviour of the sequence (Figure 3.8). It is highly possible that upon calmodulin-binding, it folds as an alpha helix as in the published structures. Besides this, the central domain side of 3583-3601 residues

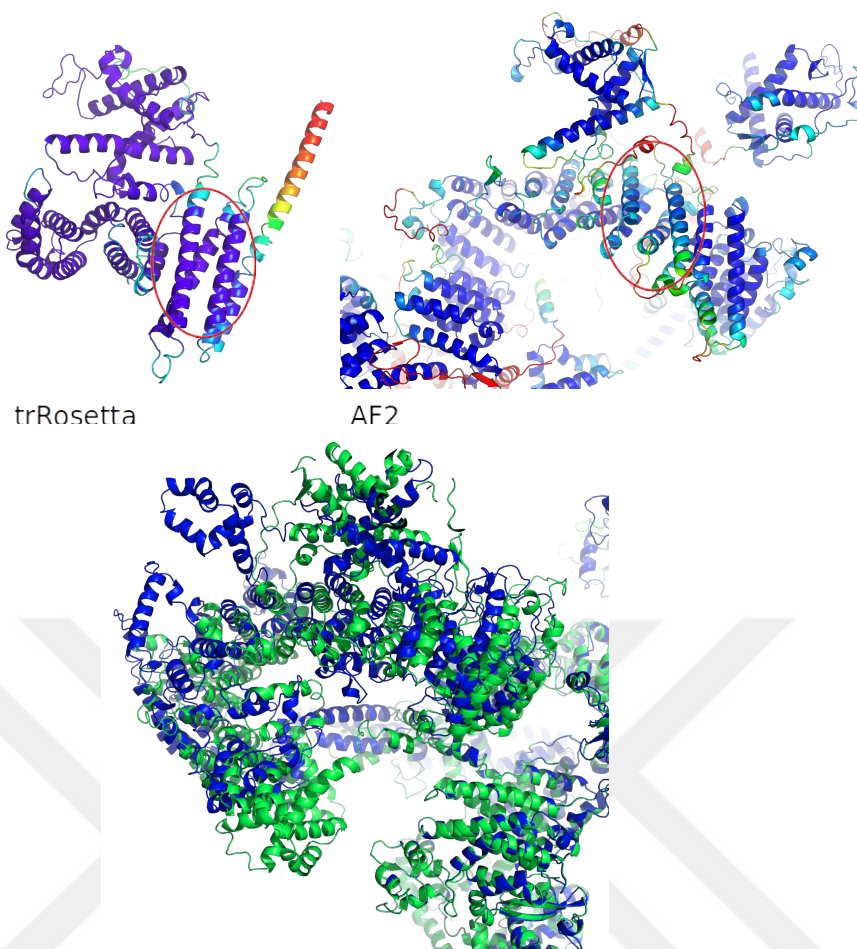


Figure 3.7 : AF2 and trRosetta model results for HD2 domain and HD2 region structural shifts relative to 5GO9 Cryo-EM structure.

is surrounded by the Handle domain where the 1940-2000 unresolved residues are positioned. In the 6WOV and 6JIY structures, there are resolved segments within the 1940-2000, yet these structures are inconsistent (Figure 3.8). AF2 model for the 1940-2000 residues was almost identical with RyR1 (PDB ID 5T15) (Figure 3.8).

All models, then, were gathered in a final model. The all structures were modeled individually, that is why the ensembling all models in one can be challenging due to the variety of the models. Therefore, first one protomer was formed using 5GO9 PDB structure as template. Here, the 5GO9 structure was chosen due to its apo structure, it does not include other modulators like Ca^{2+} , ATP and Caffeine. Combination of models were performed by aligning the overlap structures and residues obtained in models together with 5GO9 structure. The Rosetta FastRelax protocol was applied to minimize the connected bonds and overall structure. Next, the constructed protomer

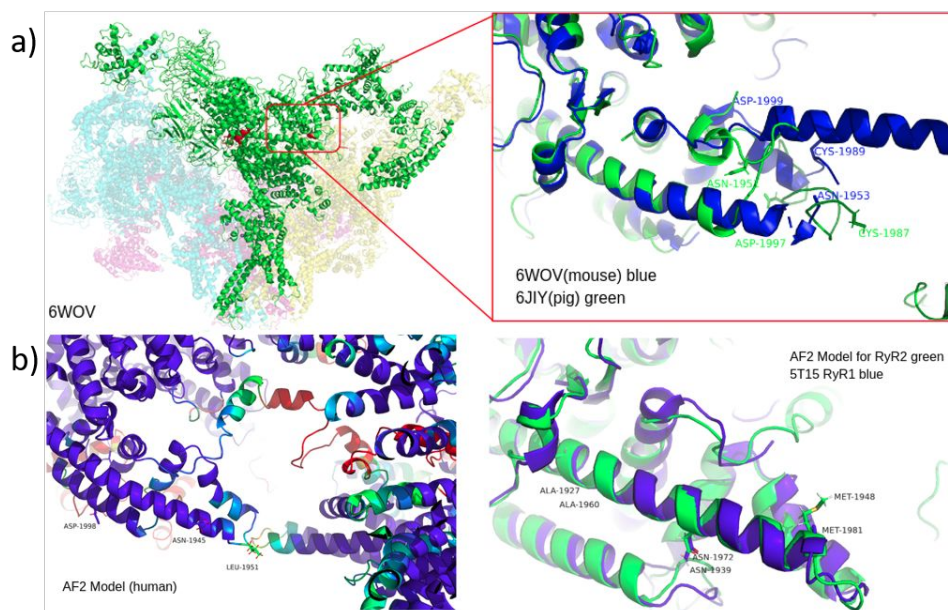


Figure 3.8 : Structural models for 1940-2000 residues.

was duplicated and using the 5GO9 as template, homotetrameric full RyR2 structure was generated. The final structure can be seen in Figure 3.9. The channel pore radius was critical for RyR2 activity, therefore the pore radius were measured using MDAnalysis python package together with HOLE software [84]. The minimum radius of the pore was obtained as $0.83\text{\AA}$ which is below the $1\text{\AA}$ and represents the closed state of the RyR2 (Figure 3.10). The minimum radius was at the pore location where Q4864 positioned. As described in the Peng et. al (2016), although the I4868 and Q4864 occupies the pore when the channel is closed state, the bottle neck of the pore is formed by I4868 (Figure 1.6). To optimize the full structure and see the pore cavity, MD simulation was carried out using RyR2-membrane complex.

The membrane was generated using Charmm-gui webserver “Membrane Builder” integration. The membrane orientation was created using the OPM server within the Charmm-gui. The membrane composition was produced with POPC type lipid and the system was solvated as default simulations parameters. The system includes 3,264,907 atoms. In the MD simulation, the lipid tail melting was performed by applying the default Charmm-gui protocol containing 6 steps for equilibration. The protocol steps were described in the Methodology.

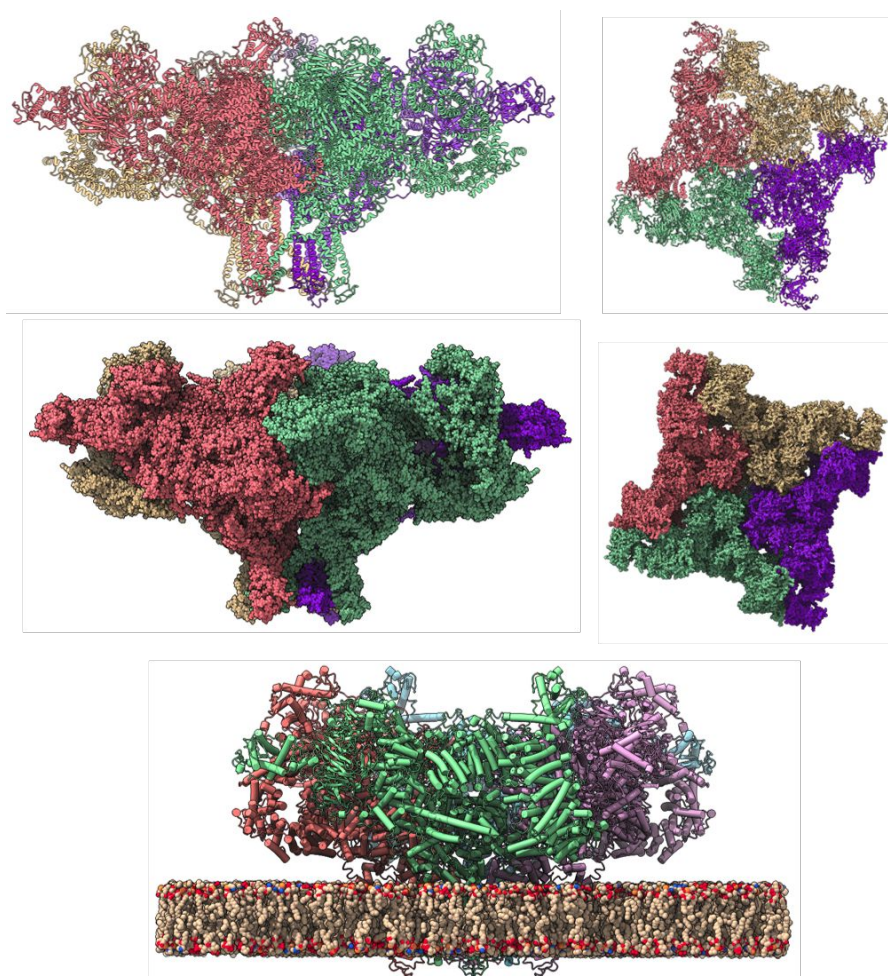


Figure 3.9 : RyR2 homotetramer full-atom model, and model membrane complex structure.

In the system building, the membrane building placed the lipids with undesirable large gaps and not packed well around the protein. These gaps disrupt the periodic condition of the system with a large vacuum which makes the unstable and unrealistic. After lipid tail melting and 35ns MD simulation, the membrane become denser and was packed well around protein which can be seen in Figure 3.13. For RMSD calculations, each frame was aligned to the first frame regarding all backbone atoms, and results were given in Figure 3.11. Although the overall structure for full homotetramer RyR2 seems to be conserved as in the Cryo-EM, the backbone RMSD result was constantly increasing when all residues are taking into account. However, when the protomers were examined individually, RMSD values tend to be stable around 6Å. In addition for chain A, RMSD variations for each domain were measured separately. P1 and P2 (called Repeat 1-2 and Repeat 3-4 domains respectively) were domains with the

highest RMSD value among others, whereas Channel and Central Domains have lower RMSD values. Besides this, according to the RMSD fluctuation, the HD2 and P2 domains showed high fluctuation.

Moreover, the overall change in pore radius was depicted in Figure 3.14. As shown in Figure 3.15, the minimum radius is measured as $0.80\text{\AA}$ around I4868. After the 35ns simulation, the Q4864 side chains slightly turned outside of the pore opening, whereas the I4868 side chains plugged the opening. It might be claimed that these residues' positions are correlated with proposed closed state positions by Peng et. al (2016). Together with the RMSD results, we suggest that the model structure represents the apo-RyR2 closed state. However, we might extend the simulation time to obtain better optimized system.

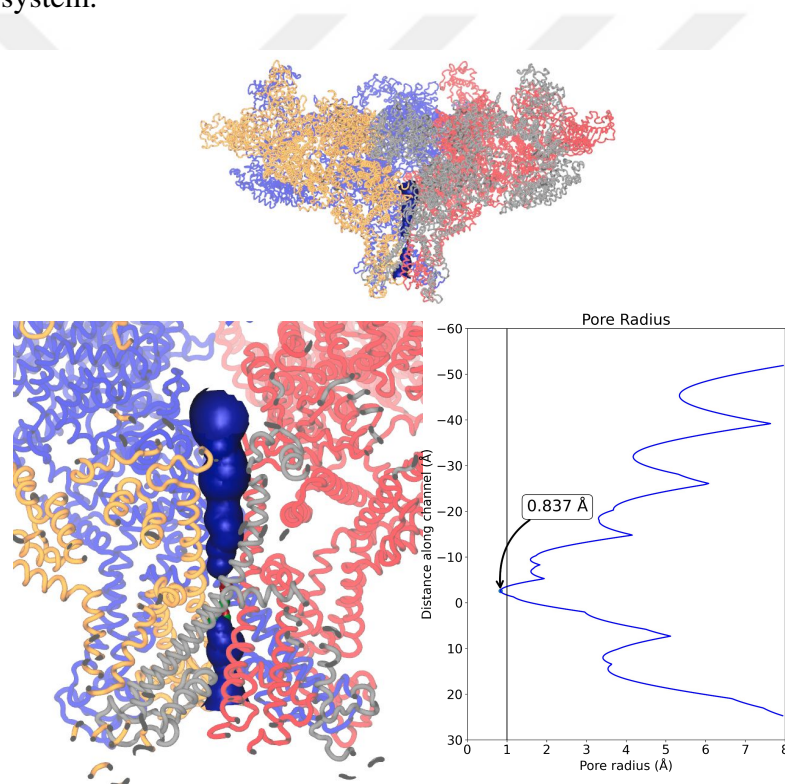


Figure 3.10 : Pore radius of RyR2 full atom model

3.2 Dantrolene binding pose prediction

Before the predict the dantrolene bond orientation on the RyR2, the potential binding sites were determined using Glide, Vina and Ledock docking approaches. By employing blind docking with different grid boxes, we assessed the entire area around

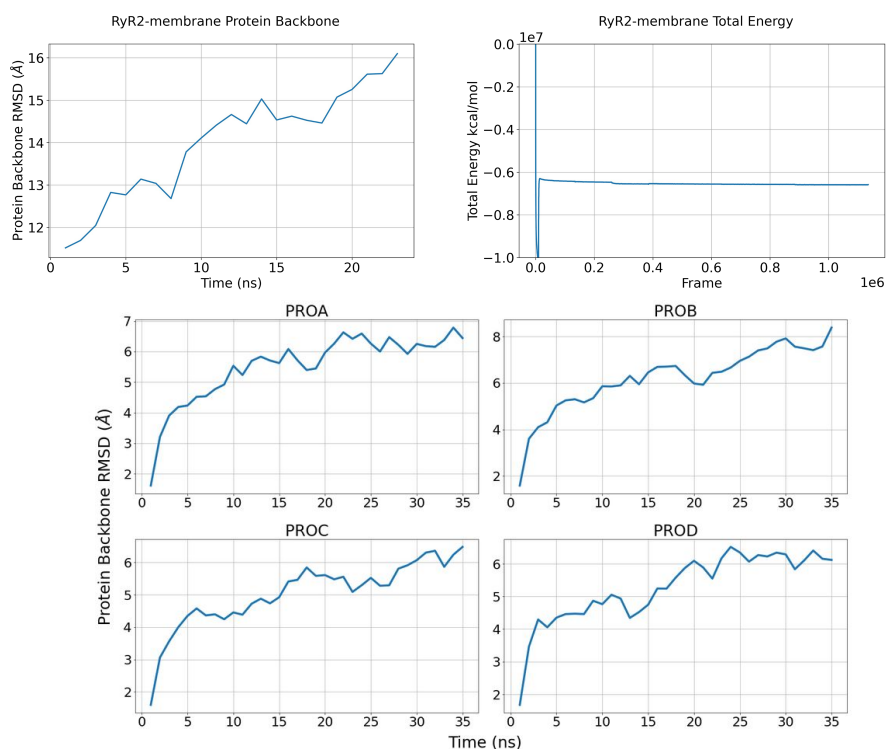


Figure 3.11 : RMSD and total energy change during fullatom RyR2-membrane complex. Individual protomer RMSD change during MD simulations

the 601-629 sequences for RyR2. In either the AF2 model or Cryo-EM model, the dantrolene conformations populated at cavity formed by NTD, Handle domain, and SPRY 1-3 domains (Figure 3.16).

The residues; H604, G605, R606 and N607 construct the base of the cavity and the dantrolene laid on these residues. The highest docking affinities are obtained in the cavity compared to the other conformations, yet overall scores are relatively poor (Table 3.1). The docking poses were diverse and there is a probability that dantrolene might not have a particular bond orientation. Besides these, the rigid body movement of domains leads to the open to close transition of RyR2 [2]. Most of the pathogenic mutations are located at the intersection of domains [5]. Thus, the stabilizing of domain movements can be crucial to restoring mutation effects. According to these, it is suggested that this cavity might be a potential binding site for the dantrolene. Therefore, scores obtained from three programs were ranked separately to determine the dantrolene binding pose. We speculated that the favourable bond conformations for dantrolene should be detected by each docking protocol after screening. Thus,

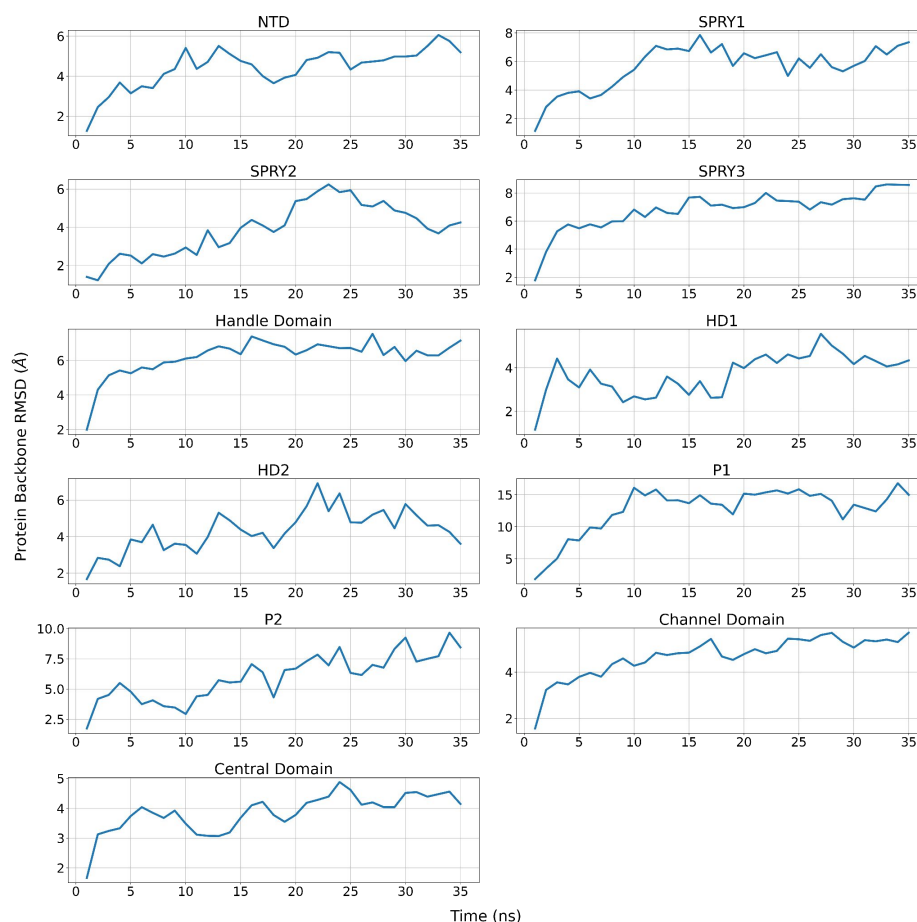


Figure 3.12 : RMSD changes of the Domains for Protomer A

the docking results were compared for each docking program, and to increase the probability, the different common poses were selected as candidate pose for the dantrolene binding. The selected poses were shown in Figure 3.17, and Figure 3.18, and listed with rank and score in Table 3.1.

All poses have proximity to H604, G605 and R606. Among those, pose 2 orientation was the highest-ranked in the Vina and Glide results for Cryo-EM based model. In pose 2, the dantrolene was on the R606 which interacts with the furan ring and the

Table 3.1 : Docking results with Score and Ranking

Model Method	Glide	Vina	LeDock
Pose 1 for Cryo-EM Model	2. rank / -3.0 GScore	2. rank / -7.8 kcal/mol	-
Pose 2 for Cryo-EM Model	1. poz / -3.4 GScore	1. rank / -7.9 kcal/mol	-
Pose 3 for Cryo-EM Model	-	9. rank / -6.8 kcal/mol	1. rank / -5.81 kcal/mol
Pose 1 for AF2 Model	5. rank / -2.23 GScore	1. rank / -7.5 kcal/mol	-
Pose 2 for AF2 Model	2. rank / -3.22 GScore	-	4. rank / -5.81 kcal/mol

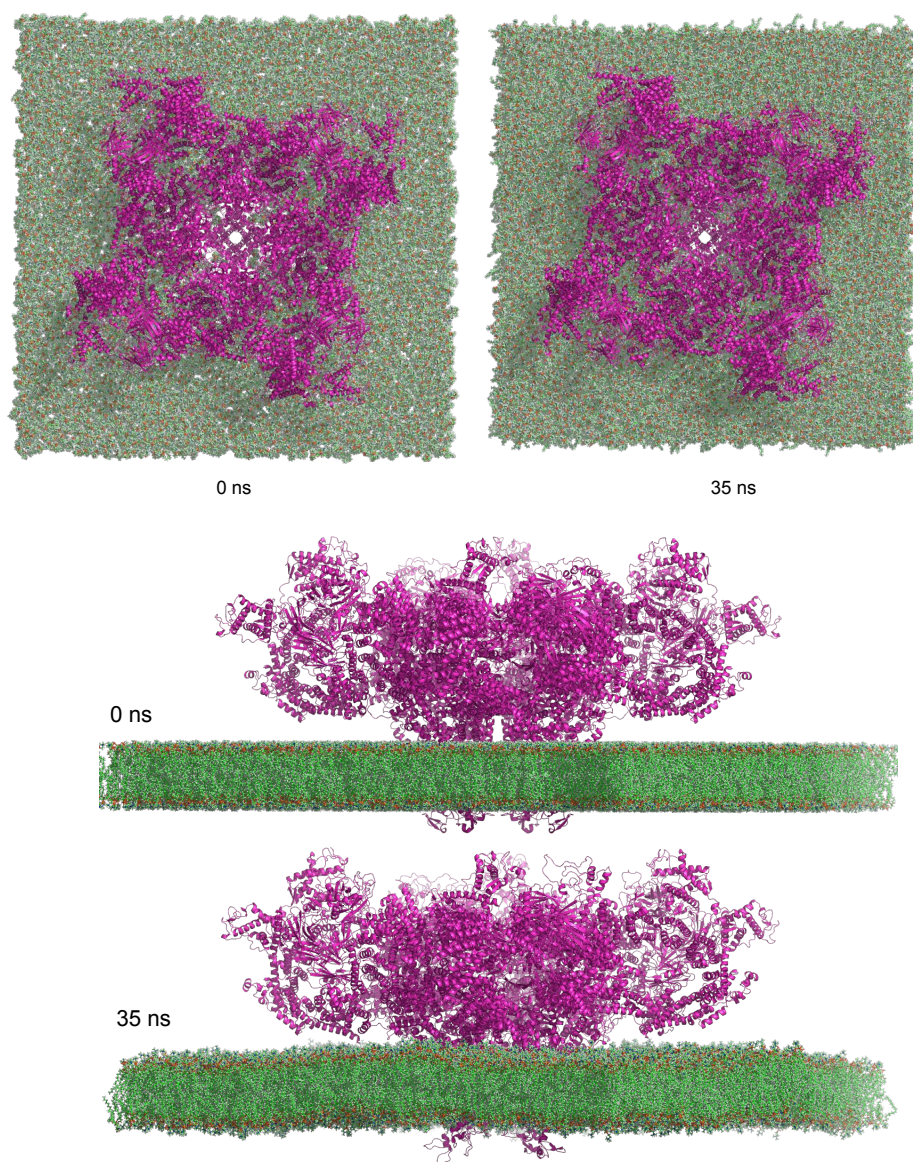


Figure 3.13 : RyR2 homotetramer conformation after 35ns simulation

imidazolidine-2,4-dione ring of dantrolene was positioned towards the H608. Besides this, pose 1 was the second-ranked pose and the imidazolidine-2,4-dione ring was towards the H604. Pose 3, however, was the best-ranked pose in the LeDock and the orientation was opposite direction regarding poses 1 and 2. For the AF2 model, two poses were selected. Pose 1 was almost identical to the Glide (5. rank) and Vina (1. rank) results. In pose 2, even though pose rank was not sufficiently high, the imidazolidine-2,4-dione position was common in all docking programs. In this pose,

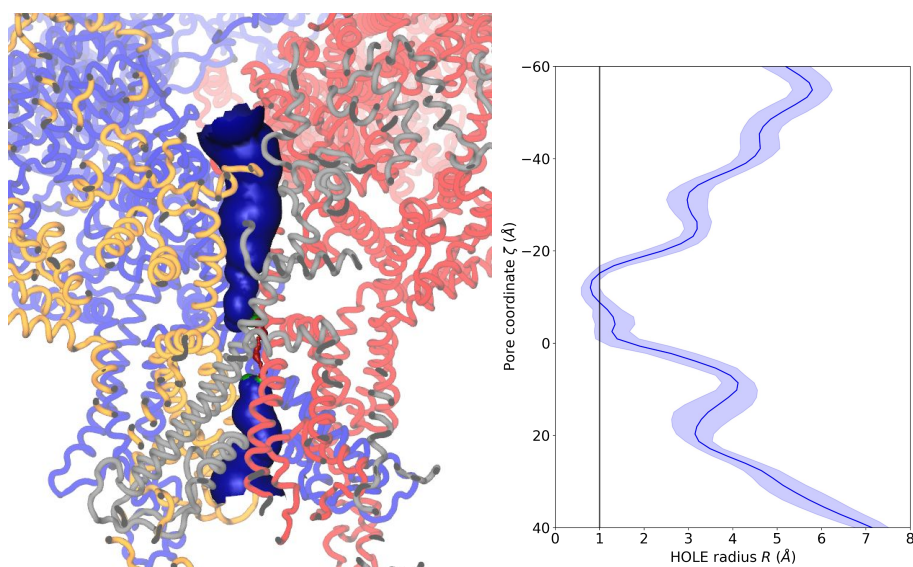


Figure 3.14 : Average pore radius change during the MD simulation.

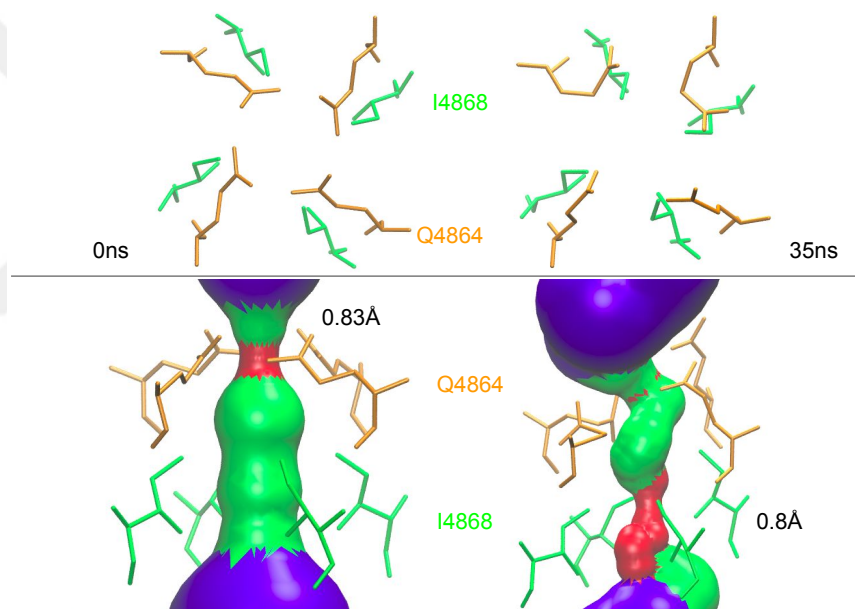


Figure 3.15 : Pore change and Q4864 and I4868 positions after 35ns MD simulation.

the dantrolene made a bridge between the H604-H608 to the adjacent L565-S568 helix. The poses were depicted in Figure 3.17 and Figure 3.18.

MD simulations were carried out using selected poses to assess the binding behaviour of dantrolene. A truncated RyR2 structure was used in all calculations to decrease computational demand. Essentially, residues between 410-1920 numbers were used in the simulation, and the Repeat 1-2 domain which corresponds to 854-1083 residues and 1847-1893 residues were removed from the system. By employing this, the system size

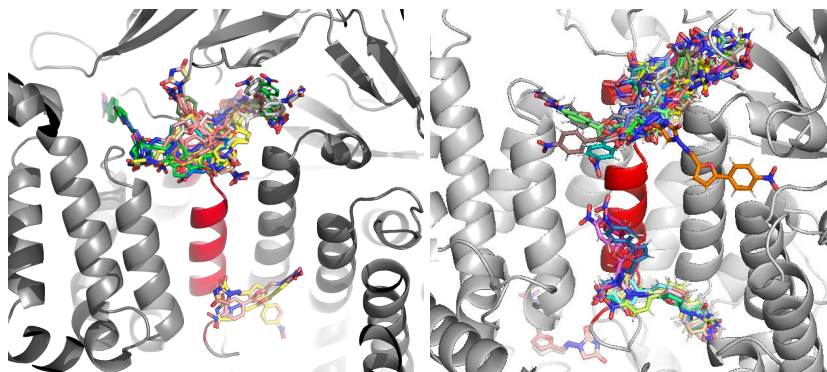


Figure 3.16 : The blind docking results. The left image for AF2 model, and right image for Cryo-EM model.

was decreased to around 100k atoms. In total, simulation systems consist of 173,883 atoms. The position restraints, then, were applied to residues 410-439, 471-489, 625-628, 670-681, 848-854, 1083-1090, 1100-1109, 1709-1721, 1832-1847, and 1893-1920. The spatial locations of these residues were depicted in Appendix Figure B.3. These restraints were decided regarding whether the residues have interaction with other protomers or segments. If a residue has an interaction with any other residues which were not included in the system, the restrain was employed on that residue. In addition, the restraints were applied to the location where the cutting edges of the Repeat 1-2 and 1847-1893 residues.

The protonation states of amino acids were determined based on the PropKa assumptions. For either H604 and H608 residues within the dantrolene binding sequence were predicted as deprotonated. With these parameters, each system was run 200ns. The RMSF of the systems were shown in Figure 3.19 and Figure 3.20. It is apparent from RMSF results that the SPRY3 domain was highly flexible among other domains for all systems, especially for the 1311-1424 disordered segments. The restraints within the other domains likely increased the overall stability. However, it can be clearly stated that compared with the Cryo-EM model, the fluctuation of the SPRY3 except the 1311-1424 segment was lower in the AF2 model. Moreover, as expected, the highest fluctuations were at loop-like structures in which residues have

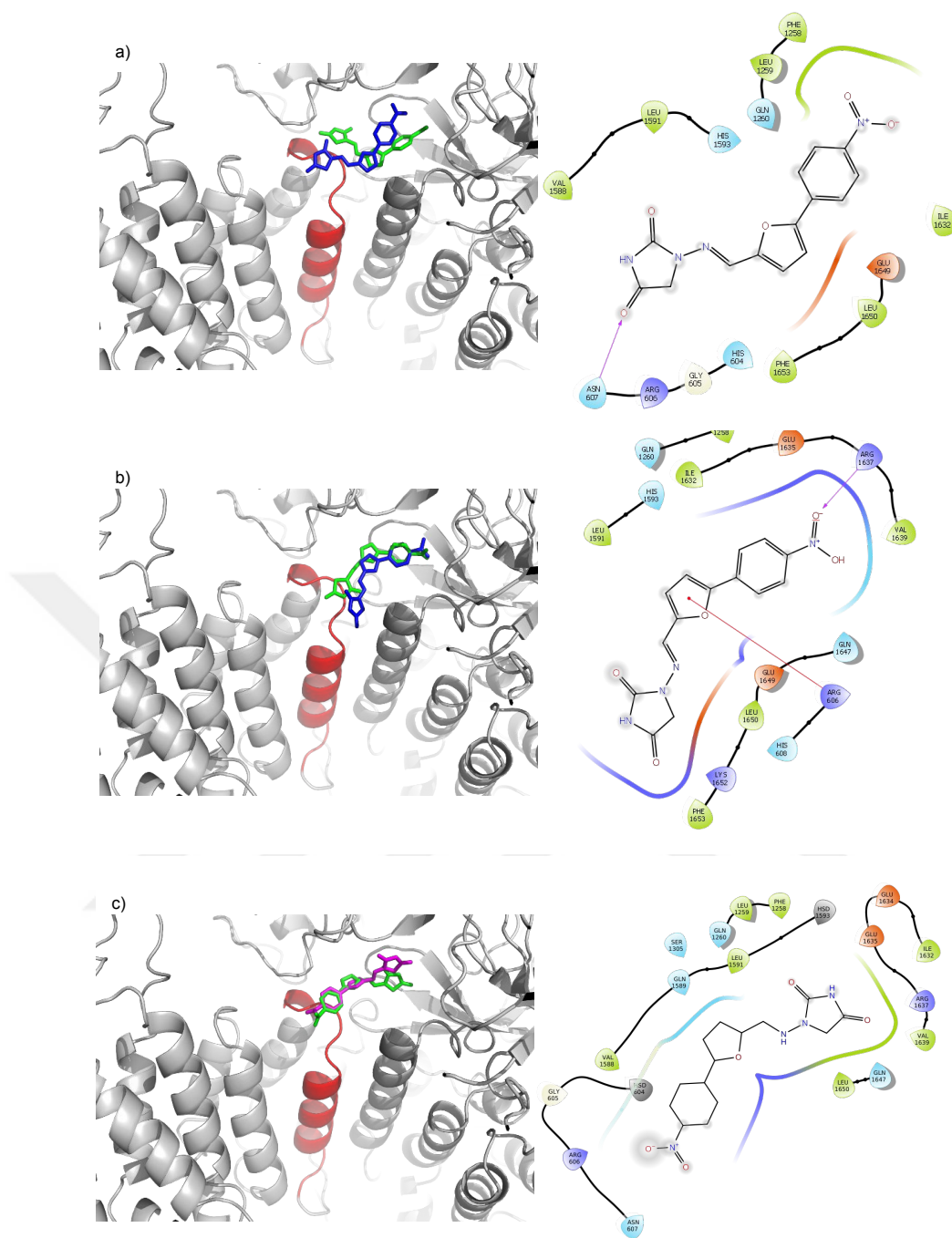


Figure 3.17 : Cryo-EM Model docking results. Dantrolene binding sequence is represented as red. Glide, Vina and Ledock poses are represented as blue, green and pink, respectively. a) Pose 1 for Cryo-EM Model, b) Pose 2 for Cryo-EM Model, and c) Pose 3 for Cryo-EM Model

3Å or higher fluctuations. For instance, 804-832 residues were loop structures located at SPRY1. Besides this, there is a significant difference between Cryo-EM and AF2 at the aforementioned interval where sequence order shift takes place in the SPRY3.

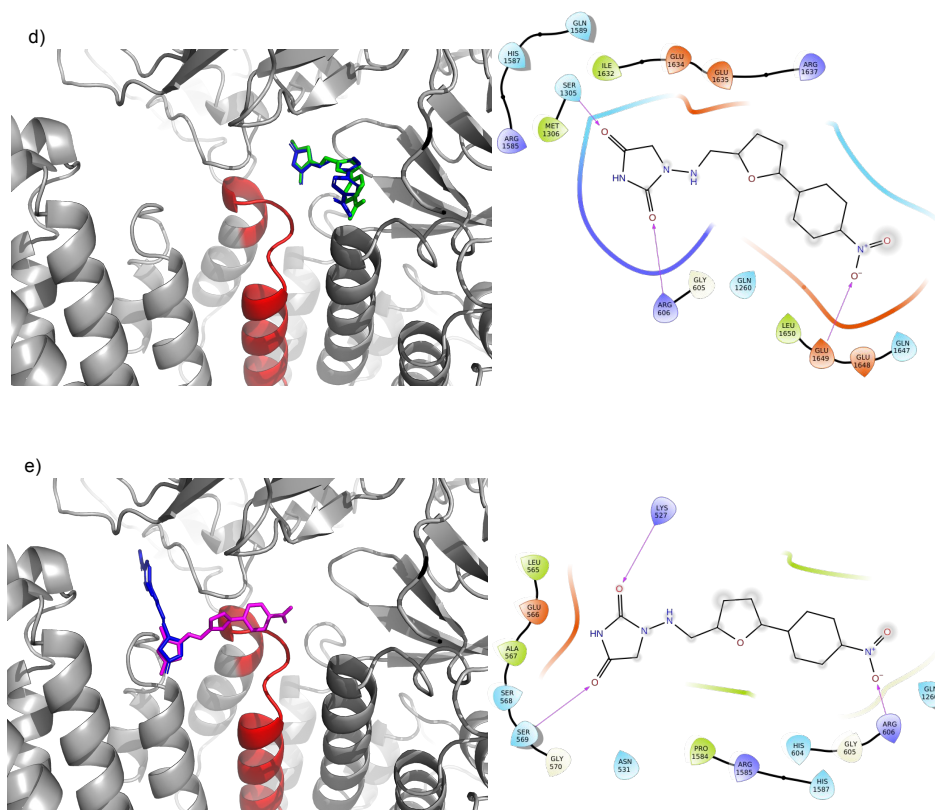


Figure 3.18 : AF2 Model docking results. Glide, Vina and Ledock poses are represented as blue, green and pink, respectively. d) Pose 1 for AF2 Model, e) Pose 2 for AF2 Model.

AF2 model for this region showed explicitly lower fluctuation. For modelling, the AF2 model can be likely the exact sequence order, yet further investigation might need to approve it. Besides this, there was no drastic change observed in the RMSF results instead of loop regions upon binding of the dantrolene. The highest loop fluctuation was observed for the 1722-1779 residues. This region was positioned far away from the dantrolene binding sequence and there were several position restraints near this position. Therefore, the 1722-1779 regions were discarded from the RMSD calculations to approximate the more accurate RMSD values.

RMSD values for the systems were shown in Figure 3.21. As a control, MD simulations were carried out for the AF2 model and Cryo-EM model without dantrolene systems. For Cryo-EM, the overall displacement was 5Å. However, the RMSD value increased after 125ns because of the increased fluctuation of 1311-1424 in the SPRY3 domain. Similarly, this RMSD increase was observed after 90ns for the

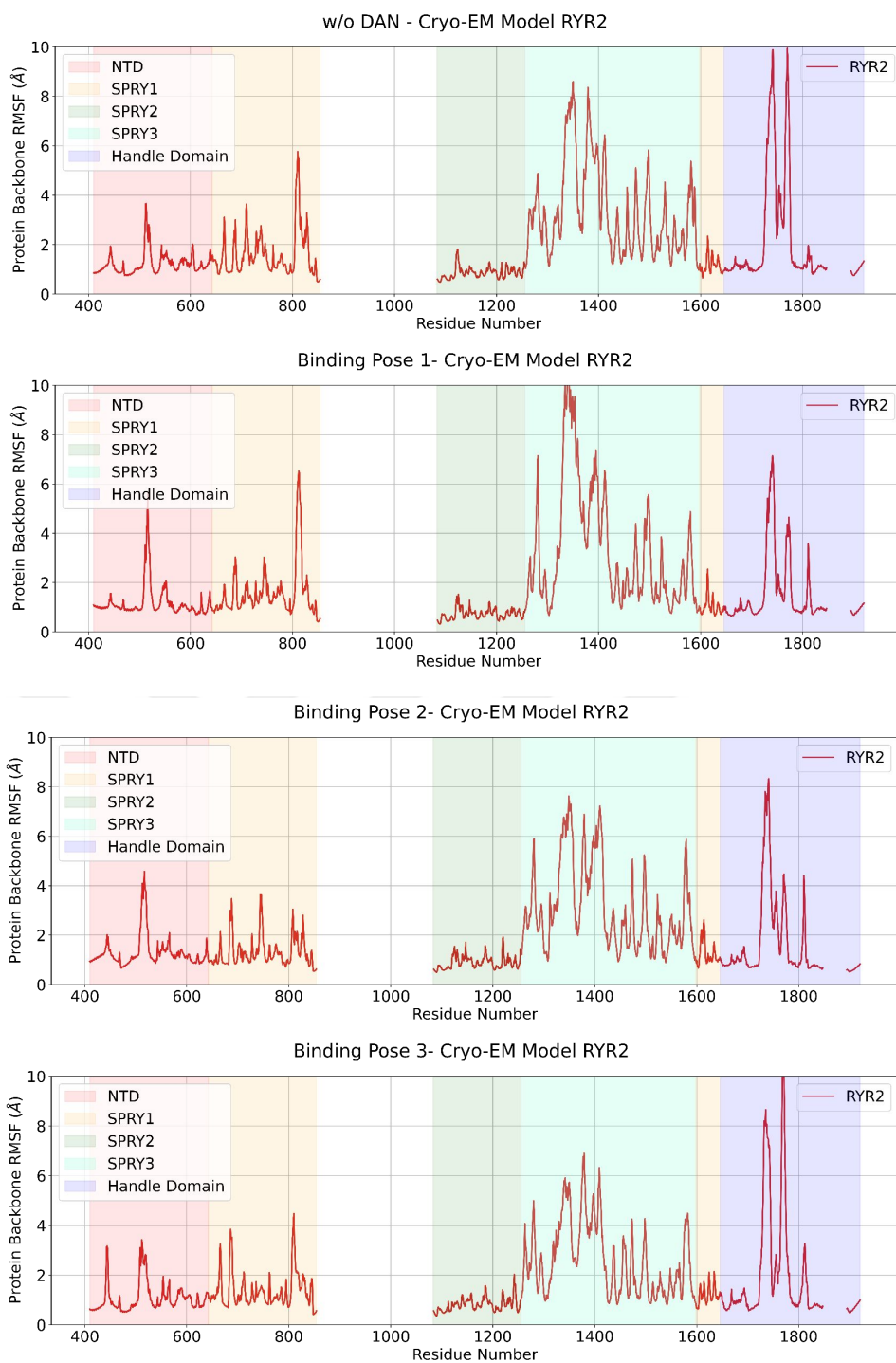


Figure 3.19 : RMSF results of the Cryo-EM Model docking MD simulations.

AF2 model. This region is likely disordered as stated in previous studies [66], hence the segment activity might not be stable in MD simulations and might increase RMSD values. We considered these jumps in the RMSD as false-positive increments.

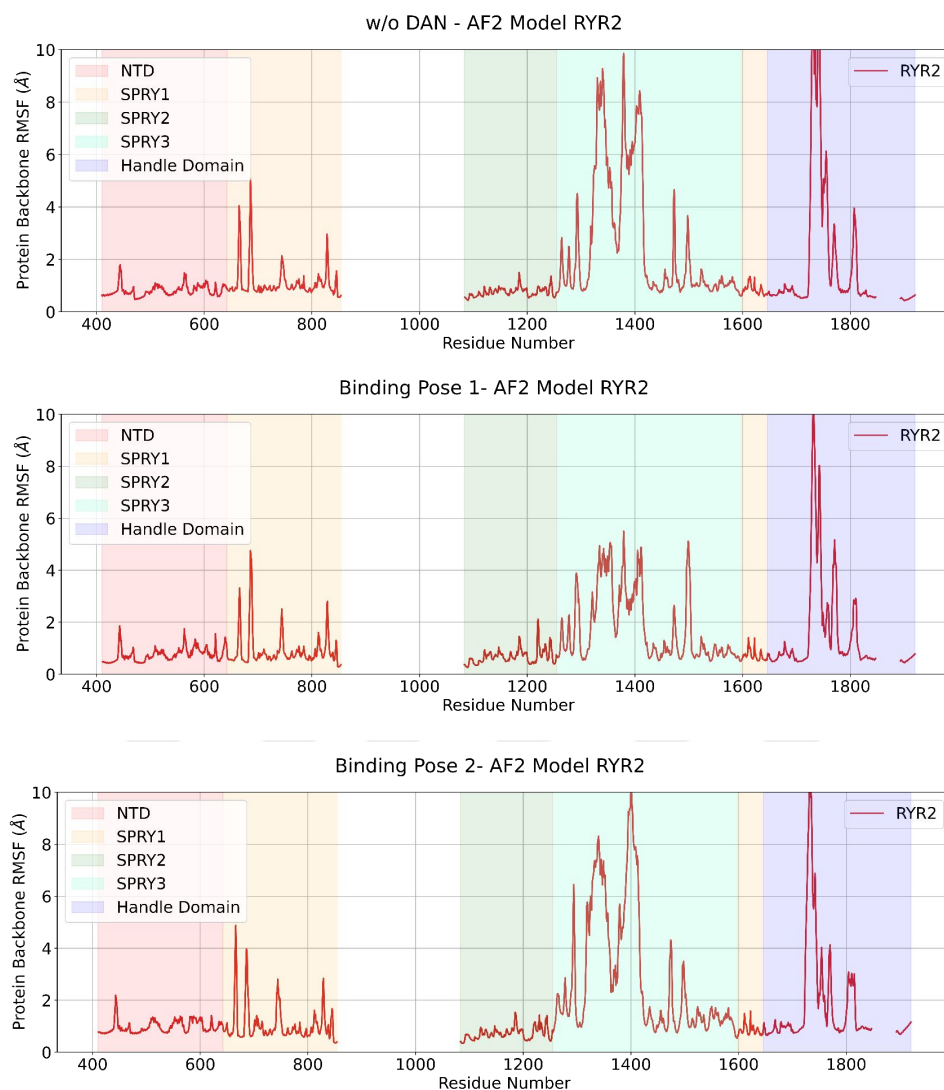


Figure 3.20 : RMSF results of the AF2 Model docking MD simulations.

Since dantrolene exhibited various binding orientations, clustering operation is employed on trajectories based on the dantrolene binding conformation similarities to evaluate binding conformations. We used the hierarchical agglomerative approach with epsilon value as 10 and forced to generate at most 10 clusters with cptraj software [85]. Top-3 clusters were listed in Appendix Table A. Among the Cryo-EM models-dantrolene systems, only the pose 1 remained its interaction after 200ns. In pose 1, the first cluster frequency for dantrolene made up the 50% of the trajectory and the dantrolene was positioned in the binding cavity (Figure 3.22.a). Interestingly, the orientation of this cluster was similar to the one which was obtained from docking pose 2 conformations. According to pose 1 simulation, the imidazolidine-2,4-dione

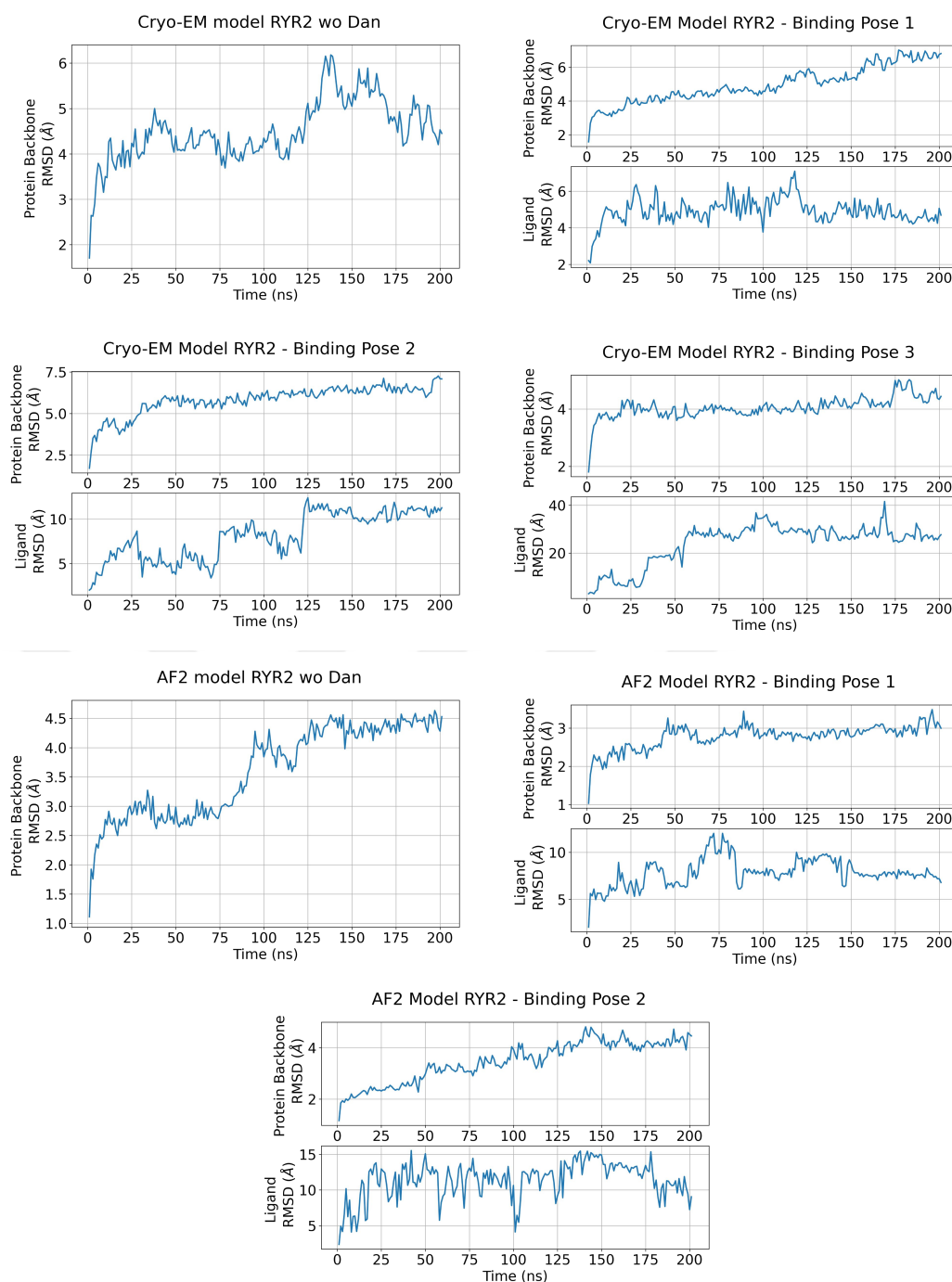


Figure 3.21 : RMSD results of the MD simulations for docking poses.

ring is prone to be near the R606, E1649 and L1650 residues. According to the dantrolene interaction diagram shown in Appendix Table B.3, the majority of the interaction is provided by hydrophobic interaction with L1650. However, dantrolene taken from pose 2 and pose 3 lost their interactions with the binding sequence after

200ns MD simulations. For pose 2, dantrolene was found in a position shown in Figure 3.22.b at least 37% of the trajectory. Lastly, for pose 3, the clustering operation was not applied due to the dantrolene's high RMSD value and lost interactions with the binding sequence. Clusters obtained from MD simulations for pose 1 in the AF2 model showed that dantrolene could not show a direct interaction with the 601-620 sequence for around 70% of simulation time (Figure 3.22.c). Unlike the pose 1 for the Cryo-EM model, the nitro-benzene ring of dantrolene tended towards being an area between R606, E1649 and L1650 residues for 42% of simulation (clusters 0 and 1) of the AF2 model pose 2 (Figure 3.22.d). The majority of the interactions come from the hydrophobic interactions with residues L1595, I1632, R1637, and V1639 shown in Appendix Table B.3.

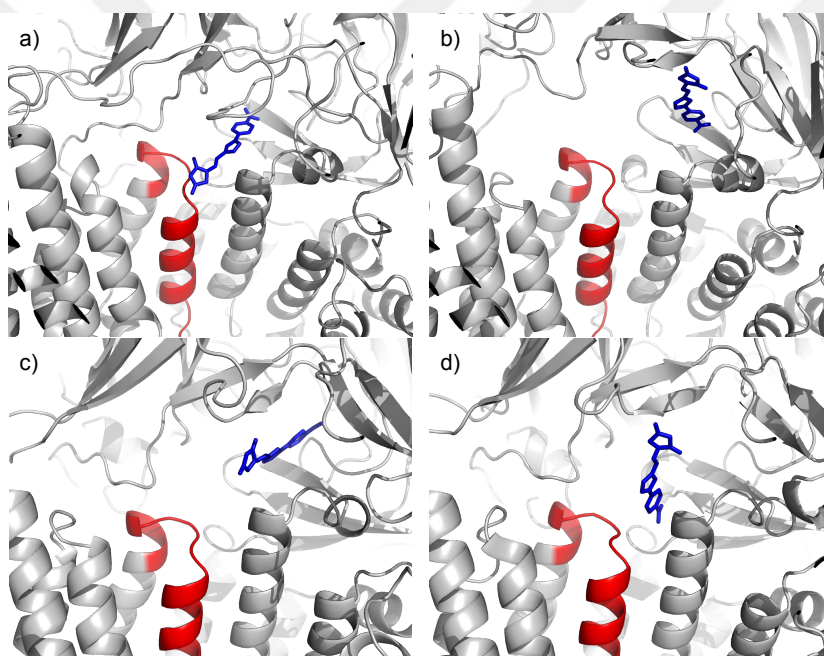


Figure 3.22 : The top cluster positions for Dantrolene. a) Obtained from Pose 1 for Cryo-EM Model simulation, b) from Pose 2 for Cryo-EM Model simulation, c) from Pose 1 for AF2 Model simulation, and d) from Pose 2 for AF2 Model simulation

As in Figure 3.19 and Figure 3.20 show, there is a significant fluctuation in SPRY3 domain. This fluctuation in the SPRY3 likely affects the dantrolene binding. Previously, Laver et al. (2018) claim that FKBP12.6 is essential for dantrolene inhibition together with Calmodulin. FKBP12.6 is a known RyR2 inhibitor and the binding site has been resolved at a location between NTD and SPRY1, very near the SPRY3 as well. Because of the location of the binding site of FKBP12.6 (Figure

B.4), we suspected that it might position dantrolene through the stabilizing SPRY3 domain. Moreover, FKBP12.6 is the only stimulant that we can work in truncated RyR2 system. Thus, the simulations were replicated in the FKBP12.6-RyR2 bond system with the same conditions and run-time. FKBP12.6 bond location was determined from the 51LD PDB structure. FKBP12.6 was integrated manually by aligning the model structure to the 51LD structure, and FastRelax was applied to minimize the bonding interactions.

The comparison RMSF of the systems were shown in Figure 3.23 and Figure 3.24. Upon binding FKBP12.6, there was a significant decrease in the fluctuation of the 1311-1424 disordered region, especially for 1326-1338 segments where FKBP12.6 direct interaction location in the Cryo-EM model pose 1 and 2. Similarly, in AF2 models pose 1 and AF2 model RyR2-apo systems, FKBP12.6 reduced the fluctuations in the similar region with respect to systems without FKBP12.6. However, the overall fluctuation peak points and magnitudes were similar for other systems in either Cryo-EM or AF2 model.

As previously, the 1722-1779 regions were removed from RMSD calculations. The results for FKBP12.6 present systems were shown in Figure 3.25. For all binding poses, the dantrolene starting orientation completely changed after the 200ns MD simulation. However, dantrolene kept its interaction with 601-620 residues and positioned inside of the cavity formed by NTD, Handle domain, and SPRY 1-3 domains. Cluster results for the binding pose 1 showed that the dantrolene located at the particular the dominant conformation (74% of simulation time, Figure 3.26). Unlike the previously conducted simulation without the FKBP12.6 system, the dantrolene position is embedded at the intersection of NTD and Handle Domain. The overall RMSD value for dantrolene was calculated around 6Å. For binding pose 2 starting simulation, dantrolene was obtained between beta-sheet formed by 1255-1264 and 1588-1595 in SPRY3 and the area in between the R606 in NTD and L1650 in Handle Domain. The nitro-benzene ring is located at the intersection of the R606 and L1650, whereas the imidazolidine-2,4-dione ring is positioned towards the beta-sheet. The main interaction is provided hydrophobic interactions with Q1260, E1649, and L1650 (Cryo-EM model residue, Appendix Figure B.4). Similarly, this orientation was

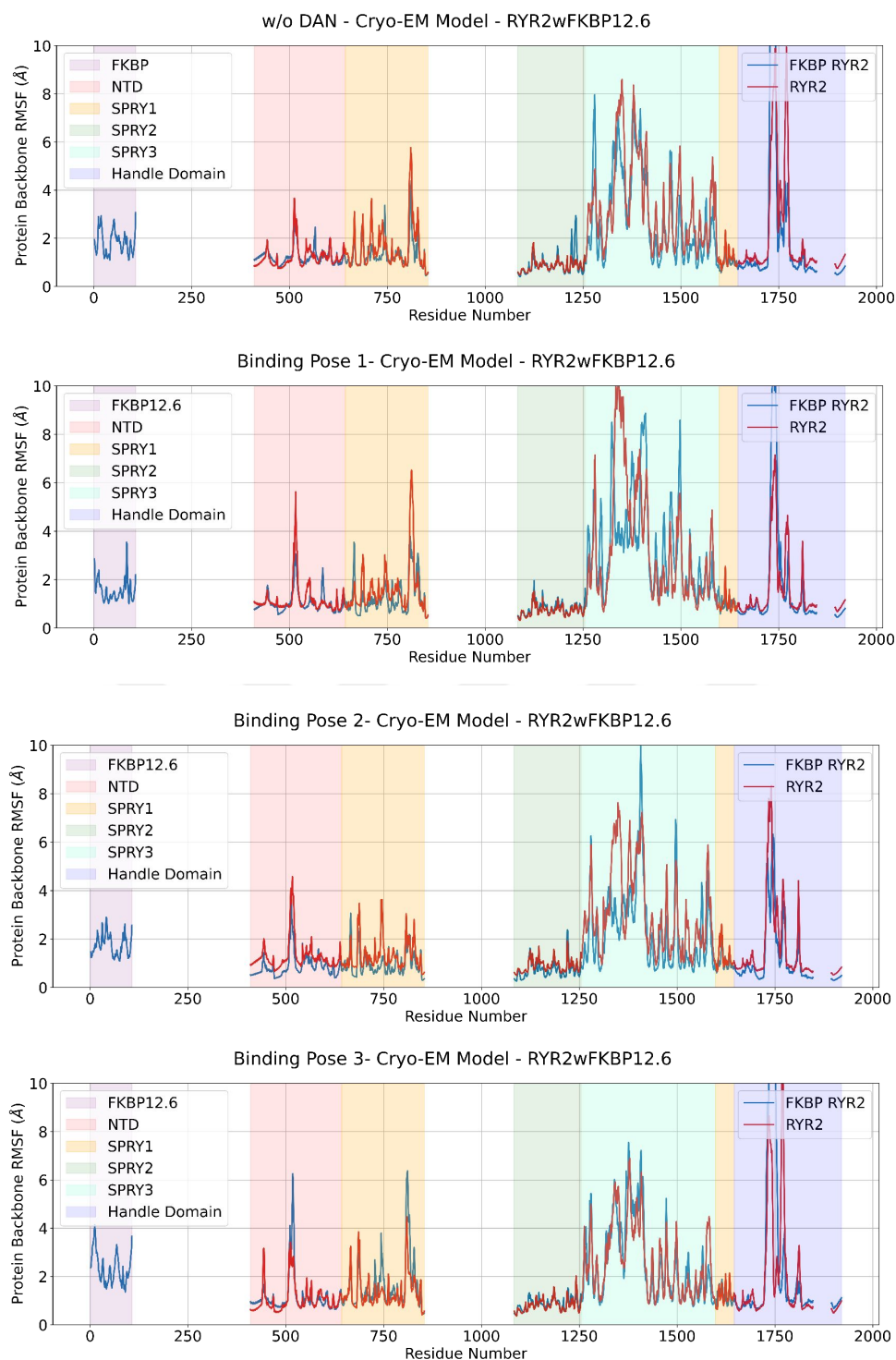


Figure 3.23 : Comparison of RMSF results for FKBP12.6 present and not included systems for Cryo-EM model docking simulations

observed in AF2 models with and without the presence of FKBP12.6 (Figure 3.26). Besides this, this orientation was seen in 84% of the simulation for Cryo-EM model pose 2 with FKBP12.6, 30% for AF2 model pose 2 with FKBP12.6 and 25% for AF2

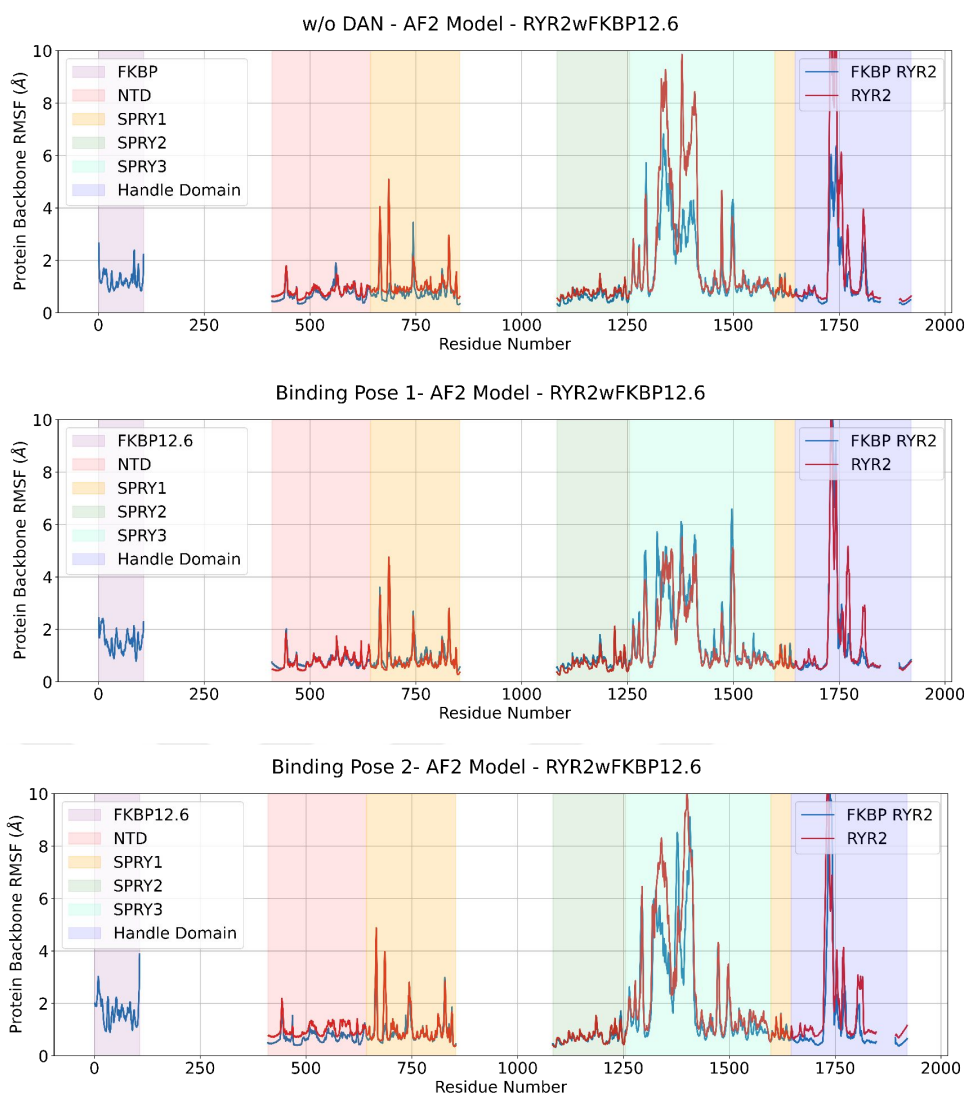


Figure 3.24 : Comparison of RMSF results for FKBP12.6 present and not included systems for AF2 model docking simulations

model pose 2 without FKBP12.6. However, the first 3 clusters which correspond to 66% of trajectory were similar in orientation in the AF2 model pose 2 with FKBP12.6, although the binding structures slightly differ from each other. Likewise, the top 5 clusters binding shapes that cover 78% of the simulation were similar in the AF2 model pose 2 without FKBP12.6, yet postures for these poses were different. These structures were shown in Figure 3.26. Lastly, in the Cryo-EM model pose 3 with FKBP12.6, the dantrolene pose was obtained similar to Cryo-EM model pose 1 without FKBP12.6 top cluster conformation. This structure corresponds to the 54% fraction of the 200ns trajectories.

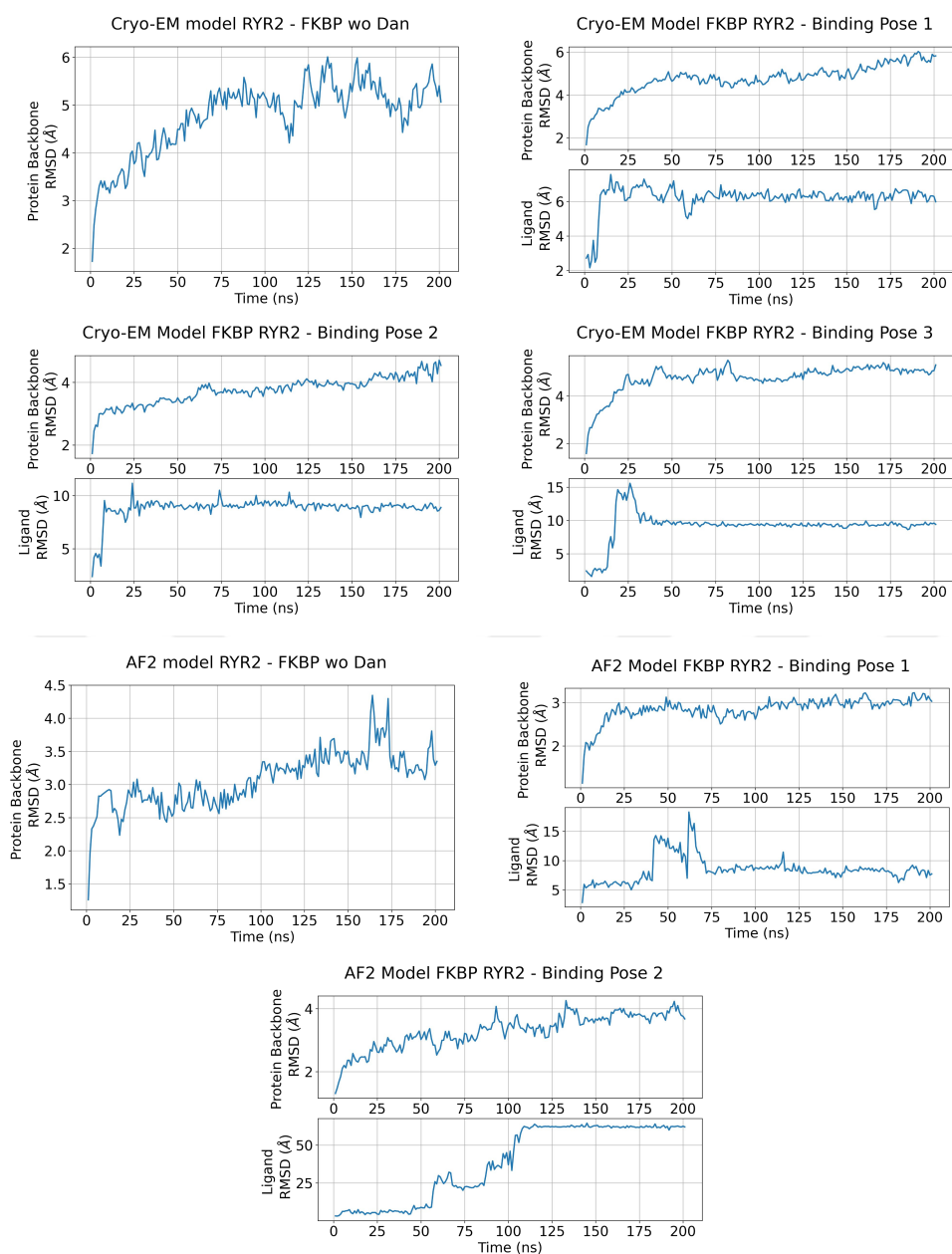


Figure 3.25 : RMSD change of the dantrolene and RYR2 truncated system in presence of FKBP12.6

Based on these results, the starting position of the dantrolene was obtained significantly different from those of structures generated from clustering after 200ns simulation in either system with or without the presence of FKBP12.6. Nonetheless, predominantly the dantrolene was inclined to present around R606, E1649, and L1650 (Appendix Figure B.3 and B.4). Conversely, substructures' exact positioning are debatable. Thus,

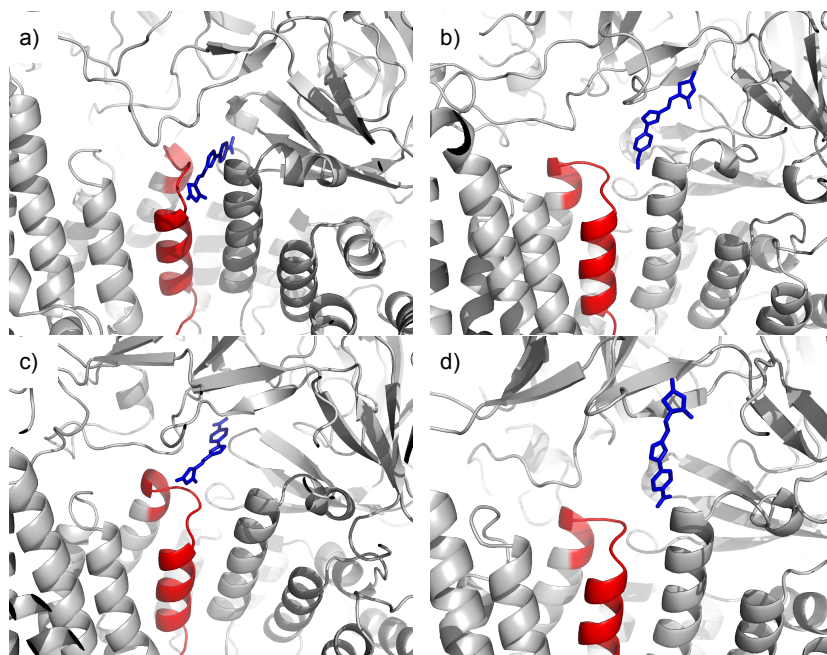


Figure 3.26 : The top cluster positions for Dantrolene in FKBP12.6 present system. a) Obtained from Pose 1 for Cryo-EM Model simulation, b) from Pose 2 for Cryo-EM Model simulation, c) from Pose 1 for AF2 Model simulation, and d) from Pose 2 for AF2 Model simulation.

the MMPBSA approach was performed to calculate and compare binding affinities on the generated trajectories.

3.3 MMPBSA

The binding energy values for each frame were shown in Figure 3.27. The highest average binding energy was obtained from the Cryo-EM model with FKBP12.6 starting pose 2, without FKBP12.6 starting pose 2, AF2 model with FKBP12.6 starting pose 1 and with FKBP12.6 starting pose 2. The energy values are -10.14, -7.32, -7.65 and -8.15 kcal/mol respectively. In these simulations, even if the dantrolene starting pose was different, the binding orientation of dantrolene was often observed in a particular area and conformation. For the Cryo-EM model with FKBP12.6 starting pose 2, the dantrolene binding frequency was 0.84 for the orientation in which nitro-benzene ring of dantrolene tend to bind predominantly to around R606, E1649, and L1650 area shown in Figure 3.26. Likewise, this conformation is observed for pose with the second highest binding energy. Based on the energy and MD simulations, the dantrolene binding orientation may be similar with these conformations.

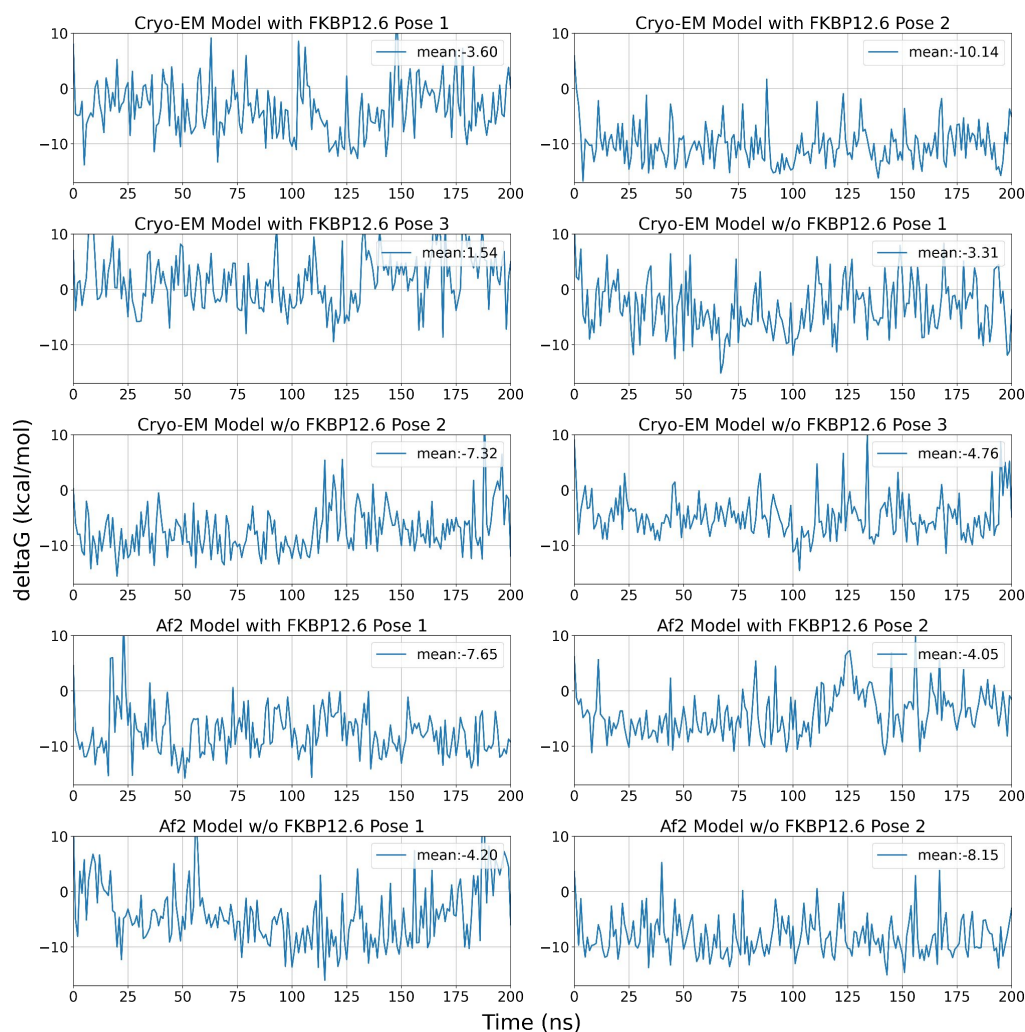


Figure 3.27 : MMPBSA energy changes during the simulations

3.4 High Throughput Screening

We performed the high throughput screening to identify the drug candidates to inhibit RyR2 activity by targeting the dantrolene putative binding site. This procedure was carried out before the release of the AF2 source code, thus in this thesis timeline, we could consider only the Cryo-EM model as the target for drug design and applied the protocols using this model.

First filtering was carried out using Vina with 8 exhaustiveness parameters. Unexpectedly, two molecules were populated in two distinct areas. The first one was a targeted binding cavity near the R606, whereas the second site was a place where FKBP12.6 binds to RyR2. The total score distribution for docking scores and densities

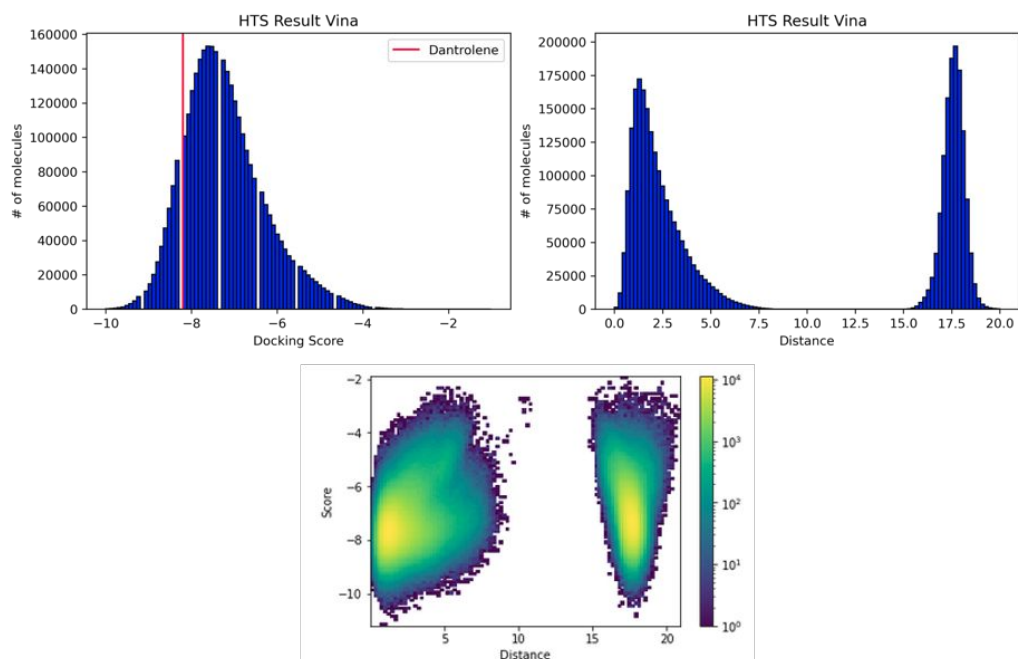


Figure 3.28 : Initial HTS score distribution and density plots.

were shown in Figure 3.28. The peak score for the distribution was observed as -7.7 kcal/mol. The binding score densities for the two sites were similar and predominantly around -8 kcal/mol. The focus of this study targets to grid centre near the R606 and Handle Domain. Thus, The molecules were refined by taking a distance lower than $10\text{\AA}$ in between the molecule centre of mass and grid centre. After filtration, around 2 million molecules remained and the top 200k of those molecules have proceeded for next step.

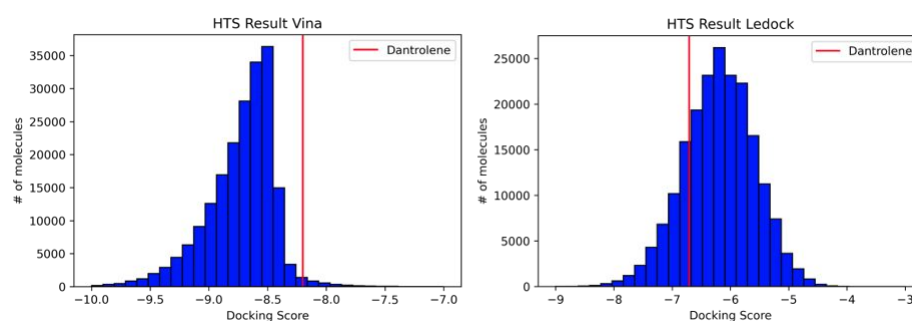


Figure 3.29 : Second step of HTS score distributions.

For second round of the screening, LeDock and Vina (24 exhaustiveness) were applied with same grid box and center used in the previous screening. The normal distribution of the results were shown in Figure 3.29. Vina results trend left side with -8.8 kcal/mol

mean as expected. Unlike the vina results, LeDock showed that scores of the most molecules were obtained around -6.1 kcal/mol. Molecules were refined according to common ranking. If molecules were detected in top 10k among molecules, they were used for next step. By employing these protocol, 517 molecules were determined. These molecules, then, were used in GlideXP to additional docking to rank their binding capabilities.

Similarly, GlideXP was performed with same grid box and center. Besides GlideXP docking, the molecules were clustered using RDKit Python module to determine the similar molecules. We used Butina Daylight fingerprint together with Tanimoto similarity with default parameters. Total 452 clusters were generated using 0.2 similarity cutoff. According to these clusters, the top 20 molecules with the highest scores were selected for Human Oral Absorption and Percent Human Oral Absorption calculations. The molecules were refined based on the >3 Human Oral Absorption and 70% Percent Human Oral Absorption as hit molecules shown as red in Table 3.2.

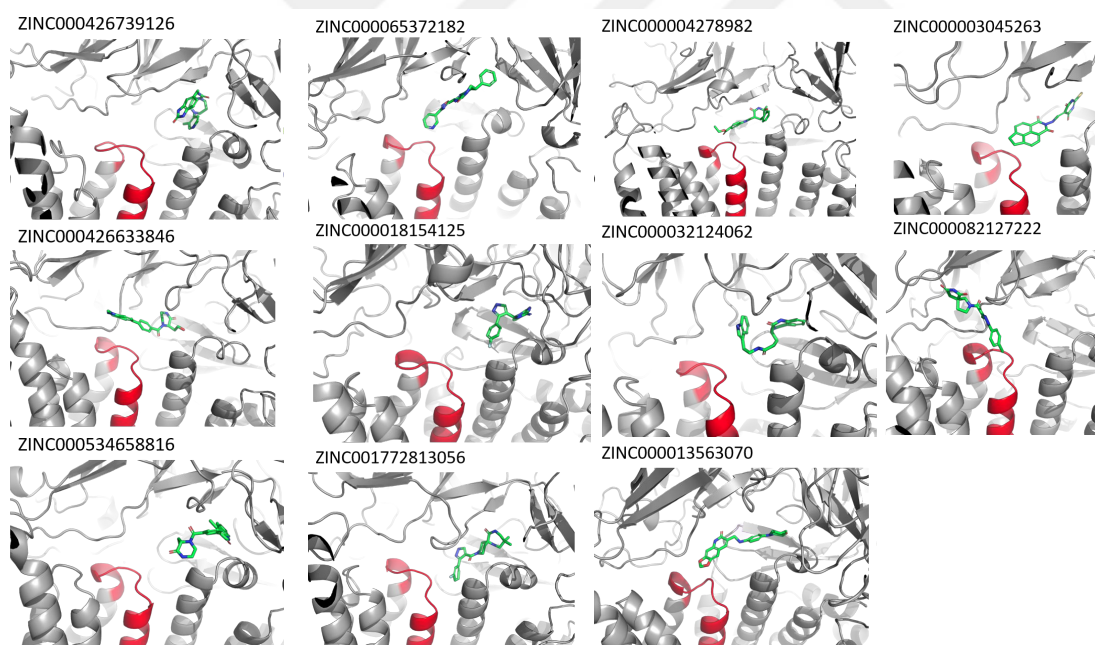


Figure 3.30 : Drug Candidates for RyR2 inhibition

As can be seen from the table, -9.08 kcal/mol GScore was obtained for the best docking score. However, the Percent Human Oral Absorption was 66.874, therefore this molecule was dismissed. Similarly, rank 6, 16, 18, 19 and 20 were included for further

Table 3.2 : Drug Candidates human oral absorption and docking score results

Rank	Molecules	Human Oral Absorption	Percent Human Oral Absorption	GScore
1	ZINC000409016884	3	66.874	-9.08
2	ZINC000008088144	3	92.686	-8.91
3	ZINC000004278982	3	91.286	-8.67
4	ZINC000032124062	3	71.309	-8.48
5	ZINC000013563070	3	94.445	-8.41
6	ZINC000014732594	3	58.113	-8.28
7	ZINC000426739126	3	72.741	-8.04
8	ZINC000426633846	3	74.613	-8.00
9	ZINC000534658816	3	74.413	-7.73
10	ZINC000065372182	3	75.987	-7.71
11	ZINC000018154125	3	80.402	-7.69
12	ZINC001772813056	3	70.863	-7.69
13	ZINC000019236630	3	82.248	-7.66
14	ZINC000408529861	3	94.433	-7.62
15	ZINC000003045263	3	76.016	-7.58
16	ZINC000097757763	2	54.361	-7.57
17	ZINC000082127222	3	82.133	-7.55
18	ZINC000004090173	2	64.539	-7.48
19	ZINC000014536718	2	46.291	-7.43
20	ZINC000426668198	3	67.428	-7.33

studies. Rank 2 and 3 were the best molecules in term of GScore and Percent Human Oral Absorption value. 92.6% and 91.2% were measured for rank 2 and 3 respectively. Moreover, 3 out of 14 molecules were not ready to buy for experiments. Therefore, these were discarded for now from the potential inhibitor list. Taken together, 11 molecules were suggested as hit among 3.5 million molecules. These molecules were shown in Figure 3.30. Even though ZINC000008088144, ZINC000019236630 and ZINC000408529861 may have potential as candidate, we discarded these molecules further steps because of not being able to commercial purchasing.

To evaluate the binding capabilities of selected 11 molecules, MD simulations were carried out. Each molecules were run 200ns with NPT condition at 298K. OPLS2005 force filed was used and default relaxation protocol was followed. The RMSD results of molecules were shown in Figure 3.31. From the results, it was observed that only ZINC000534658816, ZINC000018154125 and ZINC000534658816 molecules could

not interact with the dantrolene binding site at the end of 200ns (Figure 3.31), whereas other molecules stably maintained their docking positions to a large extent.

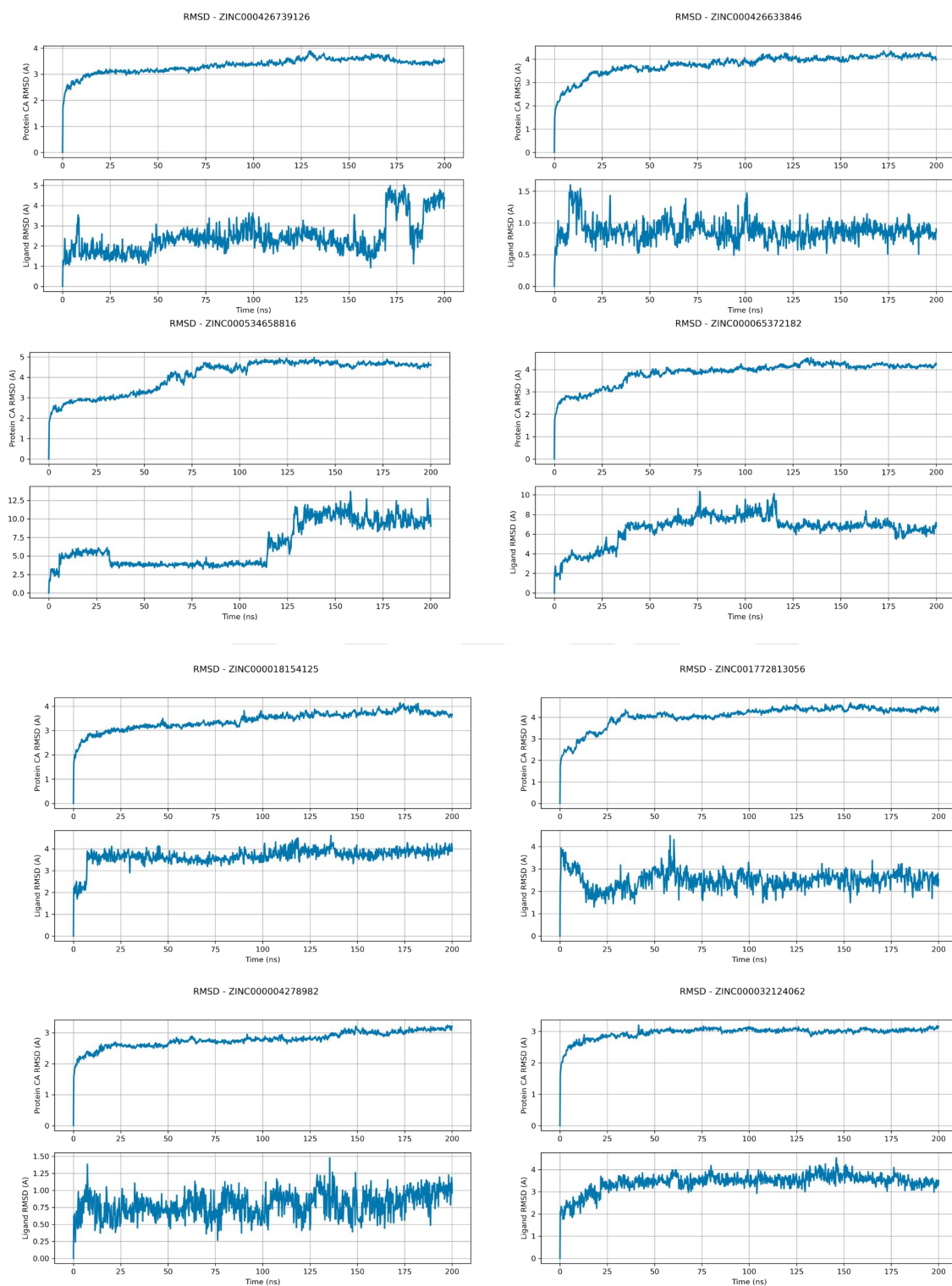


Figure 3.31 : RMSD results for drug candidates (*cont.*)

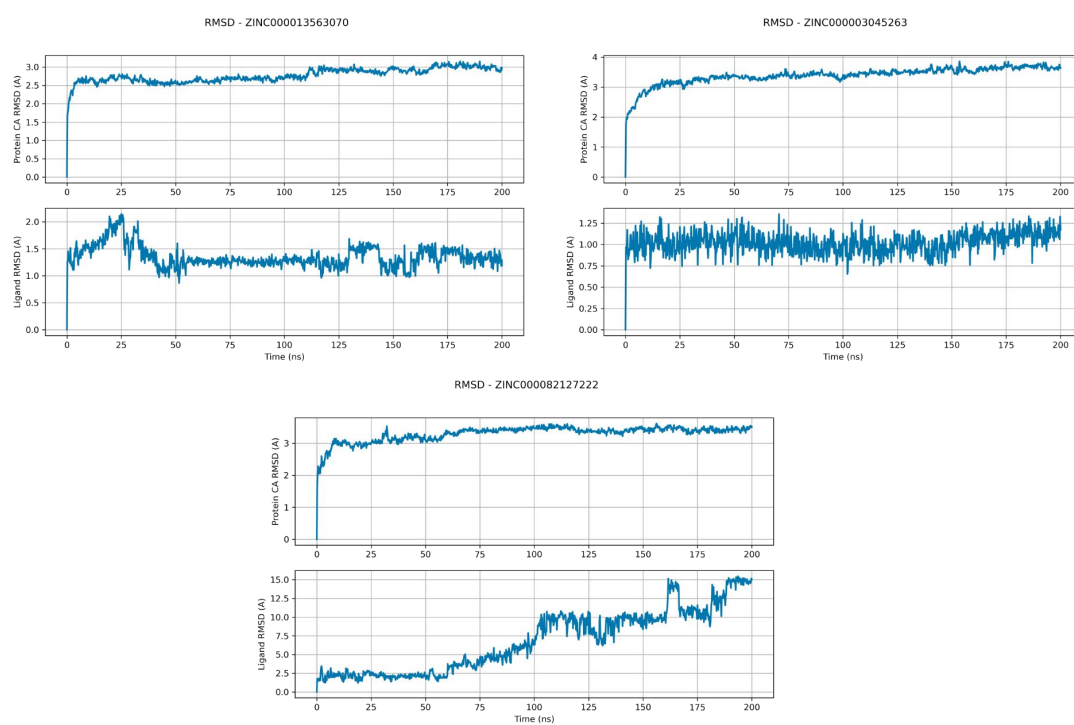


Figure 3.31 : RMSD results for drug candidates



4. CONCLUSIONS

RyR2 is large homotetrameric Ca^{2+} ion channel. It has a central role in Ca^{2+} in cardiac muscle contraction (heart beating), intra-intercellular signalling transduction and cellular regulations. The main mechanism of RyR2 relies on rigid-body shifts of domains ignited by Ca^{2+} -induced Ca^{2+} release which leads to channel opening. This mechanism is regulated by various stimulants such as Ca^{2+} , ATP, Caffeine, Mg^{2+} , CaM, FKBP12.6 and dantrolene. Many studies have been conducted to understand these modulator actions in RyR2 functionality and disease mechanisms. Even though there might be exceptions and direct mechanisms, the common inhibition-activation mechanisms are based on the allosteric interaction and activity triggered by the aforementioned modulators. Any mutation, dysregulations and modifications can disrupt the Ca^{2+} homeostasis and cause cardiac disorders such as CPVT and arrhythmia. Hence the regulation of the pore activity is the central research focus for the broad scientific community. Herein, we investigated the dantrolene binding in RyR2 truncated system with and without the presence of FKBP12.6 along with the full-atom model construction of the RyR2-homotetramer membrane complex. Besides this, we searched for novel drug-like molecules by screening a large database.

4.1 Modeling of the RyR2

The overall structure of RyRs is similar and recent Cryo-EM studies have shown great resolutions for RyR2 structure. However, there are regions which are still missing and hard to resolve in experimental procedures. Using state-of-art methods, we modelled the full-atom structure of the RyR2 homotetramer. The structure optimization procedure was followed by performing 35ns MD simulations. According to individual domain RMSD results, pore radius and composition (I4868 and Q4864), the structure represents the closed state of RyR2. Although these results are promising, it will be more reliable to investigate structure further by extending MD simulation.

Moreover, we showed that there are several regions in the Cryo-EM structures which agreed with our model in terms of backbone structure, but not in sequence order. Especially, a similar issue was observed at a segment in the SPRY3 domain which might affect the dantrolene binding. Thus, we considered both structures throughout the study.

4.2 Dantrolene binding pose prediction

There are experimental studies for dantrolene and its mechanism in various diseases induced by RyR1 or RyR2 dysregulation, but lack of computational studies. The major limitation of this is that there are no dantrolene-bound RyRs complex structures although the binding sequence has been proposed. This project is the first computational investigation of dantrolene bond conformation determination.

4.2.1 Docking results

First, we determined the possible binding sites by screening the area around the putative dantrolene binding sequence using a blind-like docking approach. Based on the ranking and pose compositions, we selected 5 poses; 3 for the Cryo-EM model, and 2 for the AF2 model. These poses were assessed using MD simulations in the presence and absence of FKBP12.6. At the end of MD simulations, 4 of those remained interactions with dantrolene binding sequences in both systems.

4.2.2 MMPBSA

To calculate binding free energy, MMPBSA was performed on a total of 8 poses taken from the previous step. The binding energy was calculated for 200 frames extracted from MD simulation trajectories. The results showed that a particular bond conformation of dantrolene has the highest binding free energy independent from the FKBP12.6 presence and structure model (Figure 3.22.d, Figure 3.26.b, Figure 3.26.d). Thus we suggested that this conformation is highly possible for dantrolene binding orientation.

4.3 High Throughput Screening

The identification of drug candidates was selected among 3.5 million molecules by performing several docking steps. After gradually filtering molecules, 20 of those were selected and the last filtering was applied based on the human oral absorption rates. In the last step, the binding capabilities of 11 molecules were assessed by MD simulations. It was observed that 8 molecules were highly stable at the binding site that we previously suggested. Hence we proposed that these molecules have potential as an inhibitor for RyR2 activity.

4.4 Comment and Future Perspective

The RyR2 is an enormous protein, thus modelling these molecules is challenging. Based on the MD simulations, our model corresponds to the apo-RyR2 closed state. However, it is lack other modulators such as Ca^{2+} or ATP. Hence the effects of these stimulants should be observed in the full-atom system. Besides this, the MD simulations were applied with default protocols. Therefore, the ion compositions and amounts are similar for both cytosolic and lumen sites. These ions ratios should be considered further. Previously, Heinz et. al. (2018) performed an MD simulation on RyR1, and in their study, they introduced restraint at the pore region to obtain closed state RyR2 while constructing their model. In this study, a similar approach can be adapted further. Moreover, dantrolene docking scores are poor and not sufficient. Thus docking approaches can be performed using the clustered structures coming from full-atom RYR2 MD simulations. After that similar approaches described in this study could be followed. For the HTS, we can perform screening on the Cryo-EM model during the timeline of this thesis. Thus, it is needed another screening using the AF2 model structure.



REFERENCES

- [1] **Lanner, J.T., Georgiou, D.K., Joshi, A.D. and Hamilton, S.L.** (2010). Ryanodine receptors: structure, expression, molecular details, and function in calcium release, *Cold Spring Harbor perspectives in biology*, 2(11), a003996.
- [2] **Peng, W., Shen, H., Wu, J., Guo, W., Pan, X., Wang, R., Chen, S.W. and Yan, N.** (2016). Structural basis for the gating mechanism of the type 2 ryanodine receptor RyR2, *Science*, 354(6310), aah5324.
- [3] **Harris, J.R. and Boekema, E.J.** (2018). *Membrane protein complexes: structure and function*, volume 87, Springer.
- [4] **Chi, X., Gong, D., Ren, K., Zhou, G., Huang, G., Lei, J., Zhou, Q. and Yan, N.** (2019). Molecular basis for allosteric regulation of the type 2 ryanodine receptor channel gating by key modulators, *Proceedings of the National Academy of Sciences*, 116(51), 25575–25582.
- [5] **Bauerová-Hlinková, V., Hajdúchová, D. and Bauer, J.A.** (2020). Structure and Function of the Human Ryanodine Receptors and Their Association with Myopathies—Present State, Challenges, and Perspectives, *Molecules*, 25(18), 4040.
- [6] **Yu, Q., Anderson, D.E., Kaur, R., Fisher, A.J. and Ames, J.B.** (2021). The Crystal Structure of Calmodulin Bound to the Cardiac Ryanodine Receptor (RyR2) at Residues Phe4246–Val4271 Reveals a Fifth Calcium Binding Site, *Biochemistry*, 60(14), 1088–1096.
- [7] **Iyer, K.A., Hu, Y., Nayak, A.R., Kurebayashi, N., Murayama, T. and Samsó, M.** (2020). Structural mechanism of two gain-of-function cardiac and skeletal RyR mutations at an equivalent site by cryo-EM, *Science advances*, 6(31), eabb2964.
- [8] **Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A. et al.** (2021). Highly accurate protein structure prediction with AlphaFold, *Nature*, 596(7873), 583–589.
- [9] **Hines, M. and Blum, J.** (1979). Bend propagation in flagella. II. Incorporation of dynein cross-bridge kinetics into the equations of motion, *Biophysical journal*, 25(3), 421–441.

- [10] **George, C.H., Yin, C.C. and Lai, F.A.** (2005). Toward a molecular understanding of the structure-function of ryanodine receptor Ca^{2+} release channels, *Cell biochemistry and biophysics*, 42(2), 197–222.
- [11] **Endo, M.** (1977). Calcium release from the sarcoplasmic reticulum, *Physiological reviews*, 57(1), 71–108.
- [12] **Fill, M. and Copello, J.A.** (2002). Ryanodine receptor calcium release channels, *Physiological reviews*, 82(4), 893–922.
- [13] **Dhindwal, S., Lobo, J., Cabra, V., Santiago, D.J., Nayak, A.R., Dryden, K. and Samsó, M.** (2017). A cryo-EM-based model of phosphorylation-and FKBP12. 6-mediated allosterism of the cardiac ryanodine receptor, *Science signaling*, 10(480), eaai8842.
- [14] **Yan, Z., Bai, X.c., Yan, C., Wu, J., Li, Z., Xie, T., Peng, W., Yin, C.c., Li, X., Scheres, S.H. et al.** (2015). Structure of the rabbit ryanodine receptor RyR1 at near-atomic resolution, *Nature*, 517(7532), 50–55.
- [15] **Cui, Y., Tae, H.S., Norris, N.C., Karunasekara, Y., Pouliquin, P., Board, P.G., Dulhunty, A.F. and Casarotto, M.G.** (2009). A dihydropyridine receptor α 1s loop region critical for skeletal muscle contraction is intrinsically unstructured and binds to a SPRY domain of the type 1 ryanodine receptor, *The international journal of biochemistry & cell biology*, 41(3), 677–686.
- [16] **Casarotto, M.G., Cui, Y., Karunasekara, Y., Harvey, P.J., Norris, N., Board, P.G. and Dulhunty, A.F.** (2006). Structural and functional characterization of interactions between the dihydropyridine receptor II-III loop and the ryanodine receptor., *Clinical and Experimental Pharmacology & Physiology*, 33(11), 1114–1117.
- [17] **Kimura, T., Pace, S.M., Wei, L., Beard, N.A., Dirksen, R.T. and Dulhunty, A.F.** (2007). A variably spliced region in the type 1 ryanodine receptor may participate in an inter-domain interaction, *Biochemical Journal*, 401(1), 317–324.
- [18] **Tae, H., Casarotto, M.G. and Dulhunty, A.F.** (2009). Ubiquitous SPRY domains and their role in the skeletal type ryanodine receptor, *European Biophysics Journal*, 39(1), 51–59.
- [19] **Fournier, D., Palidwor, G.A., Shcherbinin, S., Szengel, A., Schaefer, M.H., Perez-Iratxeta, C. and Andrade-Navarro, M.A.** (2013). Functional and genomic analyses of alpha-solenoid proteins, *PLoS One*, 8(11), e79894.
- [20] **Lobo, P.A. and Van Petegem, F.** (2009). Crystal structures of the N-terminal domains of cardiac and skeletal muscle ryanodine receptors: insights into disease mutations, *Structure*, 17(11), 1505–1514.
- [21] **Gambardella, J., Trimarco, B., Iaccarino, G. and Santulli, G.** (2017). New insights in cardiac calcium handling and excitation-contraction coupling, *Heart Failure: From Research to Clinical Practice*, 373–385.

- [22] **Van Petegem, F.** (2015). Ryanodine receptors: allosteric ion channel giants, *Journal of molecular biology*, 427(1), 31–53.
- [23] **Yano, M., Yamamoto, T., Ikeda, Y. and Matsuzaki, M.** (2006). Mechanisms of Disease: ryanodine receptor defects in heart failure and fatal arrhythmia, *Nature Clinical Practice Cardiovascular Medicine*, 3(1), 43–52.
- [24] **Yano, M., Yamamoto, T., Ikemoto, N. and Matsuzaki, M.** (2005). Abnormal ryanodine receptor function in heart failure, *Pharmacology & therapeutics*, 107(3), 377–391.
- [25] **Priori, S.G., Napolitano, C., Tiso, N., Memmi, M., Vignati, G., Bloise, R., Sorrentino, V. and Danieli, G.A.** (2001). Mutations in the cardiac ryanodine receptor gene (hRyR2) underlie catecholaminergic polymorphic ventricular tachycardia, *Circulation*, 103(2), 196–200.
- [26] **Jiang, D., Xiao, B., Yang, D., Wang, R., Choi, P., Zhang, L., Cheng, H. and Chen, S.W.** (2004). RyR2 mutations linked to ventricular tachycardia and sudden death reduce the threshold for store-overload-induced Ca²⁺ release (SOICR), *Proceedings of the National Academy of Sciences*, 101(35), 13062–13067.
- [27] **George, C.H., Higgs, G.V. and Lai, F.A.** (2003). Ryanodine receptor mutations associated with stress-induced ventricular tachycardia mediate increased calcium release in stimulated cardiomyocytes, *Circulation research*, 93(6), 531–540.
- [28] **Blancard, M., Touat-Hamici, Z., Aguilar-Sanchez, Y., Yin, L., Vaksman, G., Roux-Buisson, N., Fressart, V., Denjoy, I., Klug, D., Neyroud, N. et al.** (2021). A Type 2 Ryanodine Receptor Variant in the Helical Domain 2 Associated with an Impairment of the Adrenergic Response, *Journal of personalized medicine*, 11(6), 579.
- [29] **Brohus, M., Søndergaard, M.T., Wayne Chen, S.R., Van Petegem, F. and Overgaard, M.T.** (2019). Ca²⁺-dependent calmodulin binding to cardiac ryanodine receptor (RyR2) calmodulin-binding domains, *Biochemical Journal*, 476(2), 193–209.
- [30] **Watras, J., Ehrlich, B.E. et al.** (1991). Bell-shaped calcium-response curves of Ins (1, 4, 5) P₃-and calcium-gated channels from endoplasmic reticulum of cerebellum, *Nature*, 351(6329), 751–754.
- [31] **Bai, X.C., Yan, Z., Wu, J., Li, Z. and Yan, N.** (2016). The Central domain of RyR1 is the transducer for long-range allosteric gating of channel opening, *Cell research*, 26(9), 995–1006.
- [32] **des Georges, A., Clarke, O.B., Zalk, R., Yuan, Q., Condon, K.J., Grassucci, R.A., Hendrickson, W.A., Marks, A.R. and Frank, J.** (2016). Structural basis for gating and activation of RyR1, *Cell*, 167(1), 145–157.

- [33] **Copello, J., Barg, S., Sonnleitner, A., Porta, M., Diaz-Sylvester, P., Fill, M., Schindler, H. and Fleischer, S.** (2002). Differential activation by Ca^{2+} , ATP and caffeine of cardiac and skeletal muscle ryanodine receptors after block by Mg^{2+} , *Journal of Membrane Biology*, 187(1).
- [34] **Laver, D.R.** (2018). Regulation of the RyR channel gating by Ca^{2+} and Mg^{2+} , *Biophysical reviews*, 10(4), 1087–1095.
- [35] **Meissner, G., Darling, E. and Eveleth, J.** (1986). Kinetics of rapid calcium release by sarcoplasmic reticulum. Effects of calcium, magnesium, and adenine nucleotides, *Biochemistry*, 25(1), 236–244.
- [36] **Laver, D., Baynes, T. and Dulhunty, A.** (1997). Magnesium inhibition of ryanodine-receptor calcium channels: evidence for two independent mechanisms, *The Journal of membrane biology*, 156(3), 213–229.
- [37] **Tang, W., Sencer, S. and Hamilton, S.L.** (2002). Calmodulin modulation of proteins involved in excitation-contraction coupling, *Front Biosci*, 7, d1583–d1589.
- [38] **Ohrtman, J., Ritter, B., Polster, A., Beam, K.G. and Papadopoulos, S.** (2008). Sequence Differences in the IQ Motifs of $\text{CaV}1.1$ and $\text{CaV}1.2$ Strongly Impact Calmodulin Binding and Calcium-dependent Inactivation, *Journal of Biological Chemistry*, 283(43), 29301–29311.
- [39] **Halling, D.B., Georgiou, D.K., Black, D., Yang, G., Fallon, J.L., Quirocho, F.A., Pedersen, S.E. and Hamilton, S.L.** (2009). Determinants in $\text{CaV}1$ channels that regulate the Ca^{2+} sensitivity of bound calmodulin, *Journal of Biological Chemistry*, 284(30), 20041–20051.
- [40] **Yamaguchi, N., Xu, L., Pasek, D.A., Evans, K.E. and Meissner, G.** (2003). Molecular basis of calmodulin binding to cardiac muscle Ca^{2+} release channel (ryanodine receptor), *Journal of Biological Chemistry*, 278(26), 23480–23486.
- [41] **Balshaw, D.M., Xu, L., Yamaguchi, N., Pasek, D.A. and Meissner, G.** (2001). Calmodulin binding and inhibition of cardiac muscle calcium release channel (ryanodine receptor), *Journal of Biological Chemistry*, 276(23), 20144–20153.
- [42] **Xu, X., Yano, M., Uchinoumi, H., Hino, A., Suetomi, T., Ono, M., Tateishi, H., Oda, T., Okuda, S., Doi, M. et al.** (2010). Defective calmodulin binding to the cardiac ryanodine receptor plays a key role in CPVT-associated channel dysfunction, *Biochemical and biophysical research communications*, 394(3), 660–666.
- [43] **Lau, K., Chan, M.M. and Van Petegem, F.** (2014). Lobe-specific calmodulin binding to different ryanodine receptor isoforms, *Biochemistry*, 53(5), 932–946.

- [44] Paul-Pletzer, K., Yamamoto, T., Bhat, M.B., Ma, J., Ikemoto, N., Jimenez, L.S., Morimoto, H., Williams, P.G. and Parness, J. (2002). Identification of a dantrolene-binding sequence on the skeletal muscle ryanodine receptor, *Journal of Biological Chemistry*, 277(38), 34918–34923.
- [45] Paul-Pletzer, K., Yamamoto, T., Ikemoto, N., Jimenez, L.S., Morimoto, H., Williams, P.G., Ma, J. and Parness, J. (2005). Probing a putative dantrolene-binding site on the cardiac ryanodine receptor, *Biochemical Journal*, 387(3), 905–909.
- [46] Avula, U.M.R., Hernandez, J.J., Yamazaki, M., Valdivia, C.R., Chu, A., Rojas-Pena, A., Kaur, K., Ramos-Mondragón, R., Anumonwo, J.M., Nattel, S. *et al.* (2018). Atrial infarction-induced spontaneous focal discharges and atrial fibrillation in sheep: role of dantrolene-sensitive aberrant ryanodine receptor calcium release, *Circulation: Arrhythmia and Electrophysiology*, 11(3), e005659.
- [47] Hartmann, N., Pabel, S., Herting, J., Schatter, F., Renner, A., Gummert, J., Schotola, H., Danner, B.C., Maier, L.S., Frey, N. *et al.* (2017). Antiarrhythmic effects of dantrolene in human diseased cardiomyocytes, *Heart Rhythm*, 14(3), 412–419.
- [48] Choi, R.H., Koenig, X. and Launikonis, B.S. (2017). Dantrolene requires Mg²⁺ to arrest malignant hyperthermia, *Proceedings of the National Academy of Sciences*, 114(18), 4811–4815.
- [49] Oo, Y.W., Gomez-Hurtado, N., Walweel, K., van Helden, D.F., Imtiaz, M.S., Knollmann, B.C. and Laver, D.R. (2015). Essential role of calmodulin in RyR inhibition by dantrolene, *Molecular pharmacology*, 88(1), 57–63.
- [50] Rebbeck, R.T., Essawy, M.M., Nitu, F.R., Grant, B.D., Gillispie, G.D., Thomas, D.D., Bers, D.M. and Cornea, R.L. (2017). High-throughput screens to discover small-molecule modulators of ryanodine receptor calcium release channels, *SLAS DISCOVERY: Advancing Life Sciences R&D*, 22(2), 176–186.
- [51] Murayama, T., Kurebayashi, N., Ishigami-Yuasa, M., Mori, S., Suzuki, Y., Akima, R., Ogawa, H., Suzuki, J., Kanemaru, K., Oyamada, H. *et al.* (2018). Efficient high-throughput screening by endoplasmic reticulum Ca²⁺ measurement to identify inhibitors of ryanodine receptor Ca²⁺-release channels, *Molecular pharmacology*, 94(1), 722–730.
- [52] Loescher, C.M., Gibson, L.M. and Stephenson, D.G. (2019). Dantrolene sodium increases calcium binding by human recombinant cardiac calsequestrin and calcium loading by sheep cardiac sarcoplasmic reticulum, *Acta Physiologica*, 226(3), e13261.
- [53] Kuhlman, B. and Bradley, P. (2019). Advances in protein structure prediction and design, *Nature Reviews Molecular Cell Biology*, 20(11), 681–697.

- [54] **Pearce, R. and Zhang, Y.** (2021). Toward the solution of the protein structure prediction problem, *Journal of Biological Chemistry*, 297(1).
- [55] **Ponder, J.W. and Case, D.A.** (2003). Force fields for protein simulations, *Advances in protein chemistry*, 66, 27–85.
- [56] **Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S.a. and Karplus, M.** (1983). CHARMM: a program for macromolecular energy, minimization, and dynamics calculations, *Journal of computational chemistry*, 4(2), 187–217.
- [57] **Robertson, M.J., Tirado-Rives, J. and Jorgensen, W.L.** (2015). Improved peptide and protein torsional energetics with the OPLS-AA force field, *Journal of chemical theory and computation*, 11(7), 3499–3509.
- [58] **Scott, W.R., Hünenberger, P.H., Tironi, I.G., Mark, A.E., Billeter, S.R., Fennel, J., Torda, A.E., Huber, T., Krüger, P. and van Gunsteren, W.F.** (1999). The GROMOS biomolecular simulation program package, *The Journal of Physical Chemistry A*, 103(19), 3596–3607.
- [59] **Lindorff-Larsen, K., Piana, S., Dror, R.O. and Shaw, D.E.** (2011). How fast-folding proteins fold, *Science*, 334(6055), 517–520.
- [60] **Ruiz-Serra, V., Pontes, C., Milanetti, E., Kryshtafovych, A., Lepore, R. and Valencia, A.** (2021). Assessing the accuracy of contact and distance predictions in CASP14, *Proteins: Structure, Function, and Bioinformatics*, 89(12), 1888–1900.
- [61] **Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G.R., Wang, J., Cong, Q., Kinch, L.N., Schaeffer, R.D. et al.** (2021). Accurate prediction of protein structures and interactions using a three-track neural network, *Science*, 373(6557), 871–876.
- [62] **Yang, J., Anishchenko, I., Park, H., Peng, Z., Ovchinnikov, S. and Baker, D.** (2020). Improved protein structure prediction using predicted interresidue orientations, *Proceedings of the National Academy of Sciences*, 117(3), 1496–1503.
- [63] **Anishchenko, I., Pellock, S.J., Chidyausiku, T.M., Ramelot, T.A., Ovchinnikov, S., Hao, J., Bafna, K., Norn, C., Kang, A., Bera, A.K. et al.** (2021). De novo protein design by deep network hallucination, *Nature*, 600(7889), 547–552.
- [64] **Conway, P., Tyka, M.D., DiMaio, F., Konerding, D.E. and Baker, D.** (2014). Relaxation of backbone bond geometry improves protein energy landscape modeling, *Protein Science*, 23(1), 47–55.
- [65] **Alford, R.F., Leaver-Fay, A., Jeliaskov, J.R., O’Meara, M.J., DiMaio, F.P., Park, H., Shapovalov, M.V., Renfrew, P.D., Mulligan, V.K., Kappel, K. et al.** (2017). The Rosetta all-atom energy function for macromolecular

modeling and design, *Journal of chemical theory and computation*, 13(6), 3031–3048.

- [66] **Heinz, L.P., Kopec, W., de Groot, B.L. and Fink, R.H.** (2018). In silico assessment of the conduction mechanism of the Ryanodine Receptor 1 reveals previously unknown exit pathways, *Scientific reports*, 8(1), 1–12.
- [67] **Nelson, M.T., Humphrey, W., Gursoy, A., Dalke, A., Kalé, L.V., Skeel, R.D. and Schulten, K.** (1996). NAMD: a parallel, object-oriented molecular dynamics program, *The International Journal of Supercomputer Applications and High Performance Computing*, 10(4), 251–268.
- [68] **Jo, S., Kim, T., Iyer, V.G. and Im, W.** (2008). CHARMM-GUI: a web-based graphical user interface for CHARMM, *Journal of computational chemistry*, 29(11), 1859–1865.
- [69] **Campagna-Slater, V., Therrien, E., Weill, N. and Moitessier, N.** (2014). Methods for docking small molecules to macromolecules: A user's perspective. 2. Applications, *Current Pharmaceutical Design*, 20(20), 3360–3372.
- [70] **Yadava, U.** (2018). Search algorithms and scoring methods in protein-ligand docking, *Endocrinol Int J*, 6(6), 359–367.
- [71] **Friesner, R.A., Murphy, R.B., Repasky, M.P., Frye, L.L., Greenwood, J.R., Halgren, T.A., Sanschagrin, P.C. and Mainz, D.T.** (2006). Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein- ligand complexes, *Journal of medicinal chemistry*, 49(21), 6177–6196.
- [72] **Trott, O. and Olson, A.J.** (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, *Journal of computational chemistry*, 31(2), 455–461.
- [73] **Liu, N. and Xu, Z.** (2019). Using LeDock as a docking tool for computational drug design, *IOP Conference Series: Earth and Environmental Science*, volume218, IOP Publishing, p.012143.
- [74] **Wang, Z., Sun, H., Yao, X., Li, D., Xu, L., Li, Y., Tian, S. and Hou, T.** (2016). Comprehensive evaluation of ten docking programs on a diverse set of protein–ligand complexes: the prediction accuracy of sampling power and scoring power, *Physical Chemistry Chemical Physics*, 18(18), 12964–12975.
- [75] **Sterling, T. and Irwin, J.J.** (2015). ZINC 15–ligand discovery for everyone, *Journal of chemical information and modeling*, 55(11), 2324–2337.
- [76] **Hollingsworth, S.A. and Dror, R.O.** (2018). Molecular dynamics simulation for all, *Neuron*, 99(6), 1129–1143.

- [77] **Hospital, A., Goñi, J.R., Orozco, M. and Gelpí, J.L.** (2015). Molecular dynamics simulations: advances and applications, *Advances and applications in bioinformatics and chemistry: AABC*, 8, 37.
- [78] **Sun, H., Duan, L., Chen, F., Liu, H., Wang, Z., Pan, P., Zhu, F., Zhang, J.Z. and Hou, T.** (2018). Assessing the performance of MM/PBSA and MM/GBSA methods. 7. Entropy effects on the performance of end-point binding free energy calculation approaches, *Physical Chemistry Chemical Physics*, 20(21), 14450–14460.
- [79] **Wang, E., Sun, H., Wang, J., Wang, Z., Liu, H., Zhang, J.Z. and Hou, T.** (2019). End-point binding free energy calculation with MM/PBSA and MM/GBSA: strategies and applications in drug design, *Chemical reviews*, 119(16), 9478–9508.
- [80] **Liu, H. and Hou, T.** (2016). CaFE: a tool for binding affinity prediction using end-point free energy methods, *Bioinformatics*, 32(14), 2216–2218.
- [81] **Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S. and Richardson, D.C.** (2010). MolProbity: all-atom structure validation for macromolecular crystallography, *Acta Crystallographica Section D: Biological Crystallography*, 66(1), 12–21.
- [82] **Croll, T.I.** (2018). ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps, *Acta Crystallographica Section D: Structural Biology*, 74(6), 519–530.
- [83] **Gong, D., Chi, X., Wei, J., Zhou, G., Huang, G., Zhang, L., Wang, R., Lei, J., Chen, S. and Yan, N.** (2019). Modulation of cardiac ryanodine receptor 2 by calmodulin, *Nature*, 572(7769), 347–351.
- [84] **Gowers, R.J., Linke, M., Barnoud, J., Reddy, T.J.E., Melo, M.N., Seyler, S.L., Domanski, J., Dotson, D.L., Buchoux, S., Kenney, I.M. et al.** (2019). MDAnalysis: a Python package for the rapid analysis of molecular dynamics simulations, **Technical Report**, Los Alamos National Lab.(LANL), Los Alamos, NM (United States).
- [85] **Roe, D.R. and Cheatham III, T.E.** (2013). PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data, *Journal of chemical theory and computation*, 9(7), 3084–3095.

APPENDICES

APPENDIX A

APPENDIX B





APPENDIX A : Tables

Table A.1 : Dantrolene clusters for Cryo-EM model docking pose 1

Cluster	Frames	Frac	AvgDist	Stdev	Centroid	AvgCDist
0	1015	0.505	0.735	0.199	507	1.639
1	339	0.169	0.731	0.199	814	1.557
2	228	0.113	0.585	0.197	1778	1.862

Table A.2 : Dantrolene clusters for Cryo-EM model docking pose 2

Cluster	Frames	Frac	AvgDist	Stdev	Centroid	AvgCDist
0	750	0.373	0.958	0.316	1779	1.857
1	473	0.235	0.763	0.229	129	1.756
2	307	0.153	0.748	0.241	743	1.789

Table A.3 : Dantrolene clusters for AF2 model docking pose 1

Cluster	Frames	Frac	AvgDist	Stdev	Centroid	AvgCDist
0	400	0.199	0.803	0.263	1402	1.708
1	358	0.178	0.776	0.228	376	1.571
2	330	0.164	0.625	0.194	1651	1.893

Table A.4 : Dantrolene clusters for AF2 model docking pose 2

Cluster	Frames	Frac	AvgDist	Stdev	Centroid	AvgCDist
0	502	0.250	0.835	0.255	753	1.811
1	346	0.172	0.671	0.188	857	1.857
2	311	0.155	0.770	0.226	1184	1.854

Table A.5 : Dantrolene clusters for Cryo-EM model docking pose 1 with FKBP12.6

Cluster	Frames	Frac	AvgDist	Stdev	Centroid	AvgCDist
0	1501	0.747	0.582	0.161	862	1.502
1	173	0.086	0.609	0.145	280	1.430
2	150	0.075	0.631	0.139	468	1.782

Table A.6 : Dantrolene clusters for Cryo-EM model docking pose 2 with FKBP12.6

Cluster	Frames	Frac	AvgDist	Stdev	Centroid	AvgCDist
0	1695	0.843	0.623	0.170	1581	1.900
1	121	0.060	0.581	0.138	1921	1.756
2	51	0.025	0.714	0.157	41	1.566

Table A.7 : Dantrolene clusters for Cryo-EM model docking pose 3 with FKBP12.6

Cluster	Frames	Frac	AvgDist	Stdev	Centroid	AvgCDist
0	1087	0.541	0.688	0.199	1046	1.733
1	479	0.238	0.732	0.219	1605	1.762
2	112	0.056	0.739	0.234	65	1.862

Table A.8 : Dantrolene clusters for AF2 model docking pose 1 with FKBP12.6

Cluster	Frames	Frac	AvgDist	Stdev	Centroid	AvgCDist
0	604	0.300	0.675	0.189	1093	1.704
1	396	0.197	0.689	0.187	753	1.647
2	344	0.171	0.711	0.196	1585	1.661

APPENDIX B : Figures

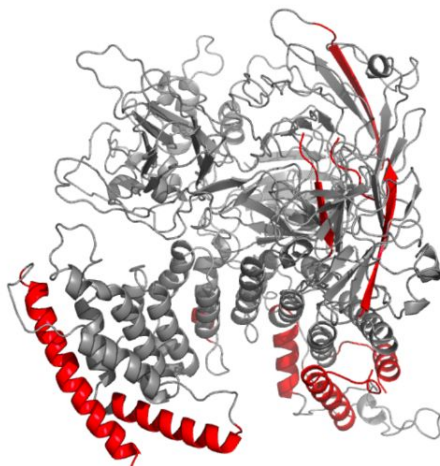


Figure B.1 : Restraint locations in the structure are represented as red.

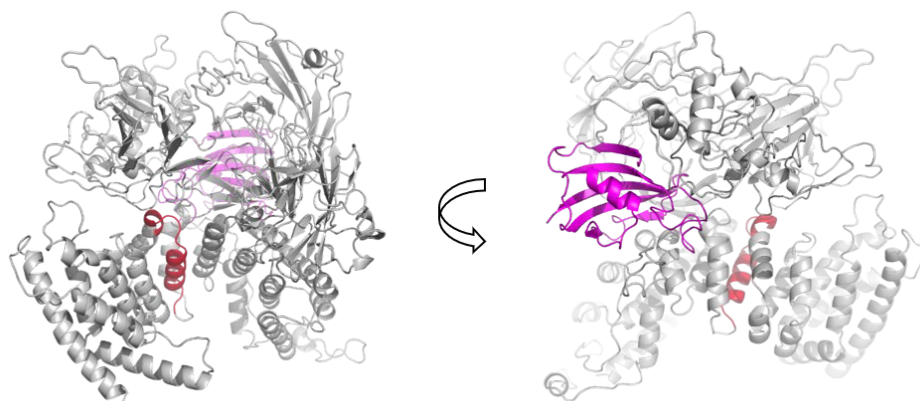


Figure B.2 : FKBP12.6 binding location. Pink is FKBP12.6 and red represents dantrolene binding sequence.

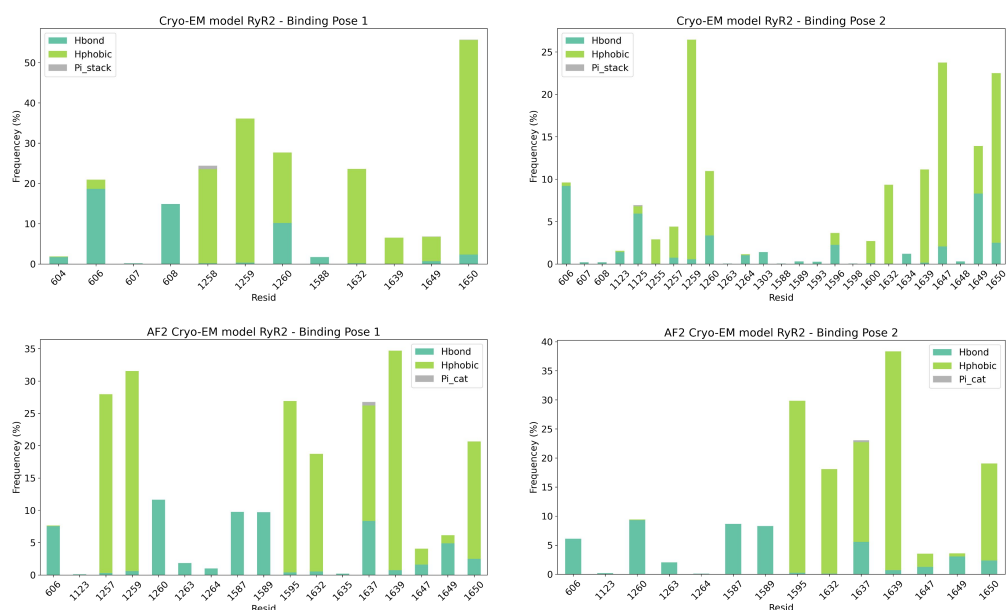


Figure B.3 : Dantrolene-RyR2 interaction plot for wo FKBP12.6 system

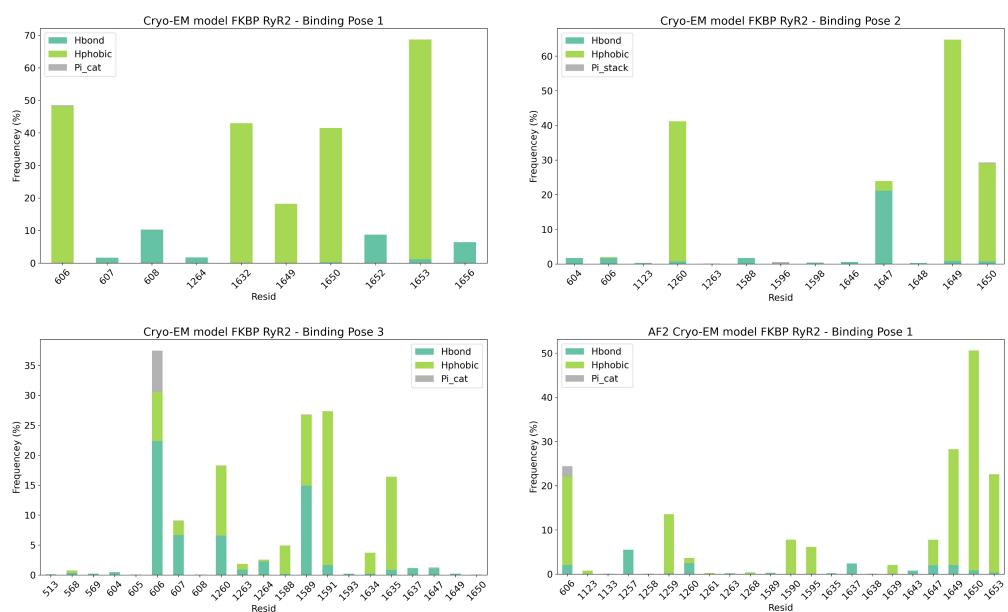


Figure B.4 : Dantrolene-RyR2 interaction plot for with FKBP12.6 system

CURRICULUM VITAE

Cemil Can SAYLAN:

EDUCATION:

- **B.Sc.:** 2019, Istanbul Technical University, Molecular Biology and Genetics Department
- **M.Sc.:** 2022, Istanbul Technical University, Informatic Institute, Computational Science and Engineering Department

PROFESSIONAL EXPERIENCE AND REWARDS:

- 2019-2022 Research Assistant at Istanbul Technical University

PUBLICATIONS, PRESENTATIONS AND PATENTS ON THE THESIS: