

INVESTIGATION OF NOVEL PHARMACOLOGICAL CHAPERONES AGAINST
GAUCHER DISEASE

by

Büge Yilmazer

B.S., Chemical Engineering, Yeditepe University, 2014

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Chemical Engineering
Boğaziçi University
2016

INVESTIGATION OF NOVEL PHARMACOLOGICAL CHAPERONES AGAINST
GAUCHER DISEASE

APPROVED BY:

Prof. Kutlu Ülgen
(Thesis Supervisor)

Assoc. Prof. Elif Özkırımlı Ölmez
(Thesis Co-Supervisor)

Assoc. Prof. Burak Alakent

Assoc. Prof. Serdar Durdağı

Prof. Seyda Malta

DATE OF APPROVAL:

ACKNOWLEDGEMENTS

Firstly, I thank almighty my parents for their constant encouragements. Without their encouragements this dissertation would not be possible. They always believe me and support me in every situation.

Furthermore, I would like to thank my supervisor, Prof. Kutlu Ülgen and co-supervisor, Assoc. Prof. Elif Özkırımlı Ölmez in order to give me opportunity to work on this project. I have to acknowledge to my lab mates, Özge, Deniz and Aziz for their support. Of course, I would like to thank Pınar for her help and guidance.

Of course, I want to express the deepest appreciation to my undergraduate supervisor, Prof. Seyda Malta and my legend instructor, Prof. Ali Mahir Arikol. I always feel the support of them and they taught me how to love this occupation.

Lastly, I would like to thank the Scientific and Technological Research Council of Turkey for financial support to our project (115S208).

ABSTRACT

INVESTIGATION OF NOVEL PHARMACOLOGICAL CHAPERONES AGAINST GAUCHER DISEASE

Beta-Glucocerebrosidase (GBA) is a lysosomal protein which is responsible for the hydrolysis of glycosylceramide into glucose and ceramide. Mutations in GBA cause a dramatic decrease in activity and lead to accumulation of glycosylceramide with the increase in spleen and liver volumes as well as skeletal deformations, and these symptoms are named as Gaucher Disease. Although enzyme replacement therapies and substrate reduction methods are utilized to overcome Gaucher Disease, they are not enough to treat Gaucher Disease patients, because the disease is neuropathic, and most of the drugs like molecules reported so far fail to pass the blood brain barrier.

In this work, we studied pyrrolidine-based compounds with QSAR, Virtual Screening, Glide Docking and MD simulations in order to identify pharmacological chaperones which cross blood brain barrier and stabilize GBA to prevent its degradation in endoplasmic reticulum. New inhibitors PCs were retrieved from ZINC Database using pharmacophore hypotheses of ligand-based and structure-based models. After rigid docking of these molecules at SP and XP modes and filtration with respect to strain energy cutoff and druglikeness properties, top ligands were further examined by induced fit docking. As a result, 7 new compounds that can potentially cross the blood brain barrier were proposed as GBA PCs. Their stabilities were examined by Molecular Dynamic simulations. All compound except compound 6 remained bound to the GBA active site until the end of 50 ns simulation. 3 of these 7 compounds have tricyclic pyrido-thienopyrimidine scaffold and one has dioxino quinolone scaffold. The derivatives of these scaffolds are used as antiallergic agents, antibiotic and anticancer compounds. Now, they may offer a new approach for the treatment of neuropathic Gaucher Disease Type 2 and Type 3.

ÖZET

GAUCHER HASTALIĞINA YÖNELİK YENİ FARMKOLOJİK ŞAPERONLARIN BULUNMASI

Beta-Glikoserebrosidaz (GBA) glikoseramitin glikoz ve seramit moleküllerine hidrolizinden sorumlu olan lizozomal proteindir. GBA proteininin yapısında meydana gelen mutasyonlar glikoseromit birikmesine, böbrekler ve ciğerlerde hacim artmasına ve iskelet yapısında deformasyonlara neden olmaktadır. Gaucher adı verilen bu hastalığın tedavisinde, enzim replasman tedavisi ve substrat azaltma tedavisinden yararlanılmasına rağmen, bu tedaviler Gaucher hastalığının nöropatik bir hastalık olması ve kullanılan ilaçların kan-beyin bariyerini geçememesi nedeniyle hastalarının tedavisinde yetersiz kalmaktadır.

Bu çalışmada, kan-beyin bariyerini geçebilecek ve GBA enzimini stabilize ederek endoplasmik retikulum içerisinde parçalanması önleyecek yeni moleküller bulabilmek amacıyla pirolidin bazlı moleküller ile sayısal aktivite-yapı ilişkisi, sanal görüntüleme, doklama ve moleküler dinamik simulasyon çalışmaları yapılmıştır. Öncelikle, yapı bazlı ve ligand bazlı modelleme ile bulunan hipotezler kullanılarak ZINC databankası filtrelenmiş ve yeni moleküller bulunmuştur. Bu moleküller sırasıyla önce SP daha sonra XP modunda doklandıktan sonra, gerilme enerjisi ve ilaç özelliklerine göre filtrelenerek sonra en iyi moleküller esnek mod ile proteinin aktif bölgesine doklanmıştır. Sonuç olarak, kan-beyin bariyerini geçebilecek en iyi 7 molekül bulunmuş ve bu moleküllerin stabiliteleri moleküler dinamik simulasyonu ile incelenmiştir. Molekül 6 dışındaki bütün moleküller GBA aktif bölgesine simulasyon sonuna kadar bağlı kaldı. En iyi 7 molekülden 3 tanesi üçhalkalı pirido-tieno-pirimidin iskeletine, 1 tanesi ise dioksino-kuinolon iskeletine sahiptir. Bu iskeletine sahip moleküller antialerjik faktör, antibiyotik ve antikanser ilacı olarak kullanılmaktadır. Bu çalışma ile bu yapıları barındıran moleküllerin Gaucher Tip 2 ve Tip 3 nöropatik hastalıklarının tedavisi için de kullanılabilecekleri belirlenmiştir.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	iii
ABSTRACT	iv
ÖZET	v
LIST OF FIGURES	x
LIST OF TABLES	xvi
LIST OF SYMBOLS/ABBREVIATIONS	xxii
1. INTRODUCTION	1
2. THEORETICAL BACKGROUND	3
2.1. Sphingolipids	3
2.2. Beta-Glucocerebrosidase	4
2.3. Pharmacological Chaperones	5
2.4. Previous Works on Glucocerebrosidase and Gaucher Disease	6
3. METHODS	14
3.1. Receptor Structure Preparation	15
3.2. Receptor Grid Generation	15
3.3. Generation of Pharmacophore Hypothesis	16
3.3.1. Structure Based Pharmacophore Modeling	16
3.3.2. Ligand Based Pharmacophore Modeling	17
3.4. Molecular Docking and Virtual Screening Workflow	19
3.5. Induced Fit Docking	20
3.6. Post Docking Applications	20
3.7. Clustering and Substructure Search	22
3.8. Selectivity Analysis	22

3.9. Binding Free Energy Calculations	23
3.10. Molecular Dynamics Simulations	23
4. RESULTS AND DISCUSSION	24
4.1. Structure-Based Modeling	24
4.1.1. Structure-Based Pharmacophore Modeling with E-Pharmacophore.....	24
4.1.2. Molecular Docking with HRH Pharmacophore Hypothesis	26
4.1.3. Post Docking Analysis of Potential Ligands	26
4.1.4. Substructure Search Using Scaffolds of Compounds from Structure- Based Modeling	29
4.1.5. Rigid Docking of New Molecules obtained by Substructure Search	30
4.1.6. Induced Fit Docking for Ligands by Structure-Based Modeling Studies	35
4.2. Ligand-Based Modeling	38
4.2.1. Ligand Based Pharmacophore Modeling	38
4.2.2. Quantitative Structure Activity Relationship (QSAR) Model	42
4.2.3. Pharmacophore Based Virtual Screening and Molecular Docking	46
4.2.3.1. Virtual Screening and Molecular Docking Results for AAADR.305 Hypothesis.....	47
4.2.3.2. Virtual Screening and Molecular Docking Results for ADDHR.125 Hypothesis.....	49
4.2.4. Rigid Docking of Compounds from Ligand Based Modeling	51
4.2.4.1. Rigid Docking of Commercial Compounds.	51
4.2.4.2. Rigid Docking of Experimentally Known Ligands.	55
4.2.5. Substructure Search for Surviving Ligands of Ligand-Based Pharmacophore Hypotheses	58
4.2.5.1. Substructure Search for Compounds of AAADR.305 Hypothesis.	58

4.2.5.2. Substructure Search for Compounds of ADDHR.125 Hypothesis.....	60
4.2.5.3. Rigid Docking of Ligands Obtained by Substructure Search of Cluster's Scaffold for ADDHR.125.....	62
4.2.5.4. Rigid Docking of Ligands Obtained by Substructure Search of Cluster's Scaffold for AAADR.305.....	67
4.2.6. Induced Fit Docking for Ligand Based Modeling.....	71
4.2.6.1. Induced Fit Docking Results of Compounds from ADDHR.125 Hypothesis.....	72
4.2.6.2. Induced Fit Docking Results of Compounds from AAADR.305 Hypothesis.....	72
4.3. Selectivity Analysis Results.....	74
4.4. Binding Free Energy Calculations.....	75
4.5. Molecular Dynamics Simulation Results.....	83
4.5.1. MD Simulations of Apo-GBA Structure.....	83
4.5.2. MD Simulations in the presence of N-nonyl-deoxynojirimycin (NN-DNJ).....	84
4.5.3. MD Simulations in the presence of Zinc00895521.....	89
4.5.4. MD Simulations in the presence of Zinc04838663.....	94
4.5.5. MD Simulations in the presence of Zinc00891180.....	100
4.5.6. MD Simulations in the presence of Zinc35092294.....	106
4.5.7. MD Simulations in the presence of Zinc00017874.....	110
4.5.8. MD Simulations in the presence of Zinc01726382.....	113
4.6 Glycosylation-Stability Analysis of Proposed Compounds.....	120
5. CONCLUSIONS AND RECOMMENDATIONS.....	123
5.1. Conclusions.....	123
5.2. Recommendations for Future Works.....	126

REFERENCES.....	127
APPENDIX A: PHARMACOPHORE MODELING.....	137



LIST OF FIGURES

Figure 2.1. Sphingolipid Pathways	3
Figure 2.2. The structure of GBA in which active site is shown with red circle.....	5
Figure 2.3. Mechanism of action of pharmacological chaperone.	6
Figure 2.4. Pyrazolopyrimidine compound.....	12
Figure 3.1. Workflow for the determination of potential pharmacological chaperones for Beta-Glucocerebrosidase.....	14
Figure 3.3. Grid of GBA generated for rigid and induced fit docking.....	16
Figure 3.4. Steps of ligand-based pharmacophore modeling.....	17
Figure 4.1. Site point locations and types at the active site of GBA.....	25
Figure 4.2. Glide GScore of top 1000 Ligands obtained by XP Docking to GBA.....	26
Figure 4.3. BEI-SEI of ligands obtained by structure-based pharmacophore modeling	27
Figure 4.4. cLogP vs potency for 199 ligands obtained from HRH hypothesis	27
Figure 4.5. LLE vs potency for 199 ligands obtained from HRH hypothesis	28
Figure 4.6. BEI and SEI plots of each cluster for structure-based modeling.....	31
Figure 4.7. Interactions of best 5 compounds obtained by structure-based modeling	37

Figure 4.8.	Canvas clustering results of 182 compounds.....	39
Figure 4.9.	Site point distances of HRH ligand-based hypothesis	46
Figure 4.10.	Glide XP Gscore for 1000 molecules obtained from AAADR.305 hypothesis	47
Figure 4.11.	BEI-SEI values of 218 compounds obtained from AAADR.305 hypothesis	47
Figure 4.12.	cLogP vs Potency of 218 molecules obtained from AAADR.305 hypothesis	48
Figure 4.13.	LLE vs Potency of 218 molecules obtained from AAADR.305 hypothesis	48
Figure 4.14.	Glide XP Gscore for 1000 molecules obtained from ADDHR.125 hypothesis	49
Figure 4.15.	BEI-SEI value for surviving molecules from ADDHR.125 hypothesis	50
Figure 4.16.	cLogP vs Potency of 241 molecules obtained from ADDHR.125 hypothesis	50
Figure 4.17.	LLE vs potency of 241 molecules obtained from ADDHR.125 hypothesis	51
Figure 4.18.	Interaction maps of 5 commercial compounds.....	53
Figure 4.19.	Alignment of NN-DNJ in crystal structure and docked NN-DNJ	58

Figure 4.20.	Average BEI and SEI values of clusters of hypothesis ADDHR.125	63
Figure 4. 21.	Average BEI and SEI values of clusters of hypothesis AAADR.305	68
Figure 4.22.	IFD Interactions between top 6 ligands and GBA.....	77
Figure 4.23.	RMSD properties of apo-GBA C α atoms.....	83
Figure 4.24.	Average RMSF Values of the APO-GBA Structure	84
Figure 4.25.	RMSD properties of GBA C α atoms (top panel) and ligand, NN-DNJ, (bottom panel) after superpositioning based on protein C α coordinates	84
Figure 4.26.	RMSD properties of ligand, NN-DNJ, (bottom panel) after superpositioning based on protein C α coordinates	85
Figure 4.27.	Average RMSF values of the GBA in the presence of NN-DNJ	85
Figure 4.28.	Conformational change in GBA- NN-DNJ complex throughout simulation.	86
Figure 4.29.	Conformations of NN-DNJ throughout the MD Simulation	88
Figure 4.30.	RMSD properties of GBA C α atoms (top panel) and ligand, Zinc00895521, (bottom panel) after superpositioning based on protein C α coordinates.....	89
Figure 4.31.	RMSD properties of GBA C α atoms (top panel) and ligand, Zinc00895521, (bottom panel) after superpositioning based on protein C α coordinates.....	89

Figure 4.32.	Conformational change in GBA-ZINC00895521 complex throughout the simulation.....	90
Figure 4.33.	Average RMSF values of the GBA	91
Figure 4.34.	Conformations of Zinc00895521 throughout the MD Simulation	93
Figure 4.35.	RMSD properties of GBA C α atoms (top panel) and ligand, Zinc04838663, (bottom panel) after superpositioning based on protein C α coordinates.....	94
Figure 4.36.	RMSF values of GBA bound Zinc04838663	95
Figure 4.37.	Conformations of Zinc04838663 throughout the MD Simulation	96
Figure 4.38.	Conformational change in GBA-ZINC04838663 complex throughout the simulation.....	97
Figure 4.39.	The movement of Phe 347 at different time intervals.....	98
Figure 4.40.	Scaffold of Zinc04838663	99
Figure 4.41.	RMSD properties of GBA C α atoms (top panel) and Zinc00891180, (bottom panel) after superpositioning based on protein C α coordinates.....	101
Figure 4.42.	Average RMSF values of residues throughout the simulation.	102
Figure 4.43.	Conformation of Zinc00891180 throughout the simulation.....	103
Figure 4.44	Conformational change in GBA-ZINC00891180 complex throughout the simulation.....	104

Figure 4.45.	Scaffold of Zinc00891180	105
Figure 4.46.	RMSD properties of GBA C _α atoms (top panel) and Zinc35092294 after superpositioning based on protein C _α coordinates.....	106
Figure 4.47.	Average RMSF values of each residues in GBA throughout the simulation	107
Figure 4.48.	Conformations of Zinc35092294 throughout the MD Simulation	108
Figure 4.49.	Conformational change in GBA-ZINC35092294 complex throughout the simulation.....	109
Figure 4.50.	RMSD properties of GBA C _α (top panel) and ligand Zinc00017874 (bottom panel) after superpositioning based on protein.....	110
Figure 4.51.	Average RMSF values of GBA in the presence of Zinc00017874	111
Figure 4.52.	Conformations of Zinc00017874 throughout the MD simulation	112
Figure 4.53.	Conformational change in GBA-ZINC00017874 complex throughout the simulation.....	113
Figure 4.54.	RMSD properties of GBA C _α (top panel) and ligand Zinc01726382 (bottom panel) after superpositioning based on protein.....	114
Figure 4.55.	Average RMSF values of residues of GBA.....	115
Figure 4.56.	Conformations of Zinc01726382 throughout the MD simulation	116

Figure 4.57.	Conformational change in GBA-ZINC01726382 complex throughout the simulation.....	117
Figure 4.58.	RMSD values of ligands bound to GBA active site through MD simulation.....	120
Figure 4.59	RMSF values of residues in the presence of different ligands	121
Figure A.1.	Second and third set of ligands used for QSAR studies.	180
Figure A.2.	Interaction maps of NB-DNJ (Miglustat) (A) and NN-DNJ (B) to GBA active site.....	181
Figure A.3.	IFD XP GScore versus predicted IC50 of compounds obtained by ligand-based model.....	181

LIST OF TABLES

Table 2.1.	Type of mutations in GBA and compounds used to treat symptoms of these mutations.....	13
Table 4.1.	Ranks and scores of the top seven pharmacophore features.....	24
Table 4.2.	Inner-site distances of three pharmacophore sites.	25
Table 4.3.	Clusters involving more than 4 ligands and common substructure of ligands.....	29
Table 4.4.	ZINC ID, XP GScore and strain energy of compounds retrieved by using the scaffold of cluster 1	31
Table 4.5.	ZINC ID, XP GScore and strain energy of compounds retrieved by using the scaffold of cluster 4	32
Table 4.6.	Representative interactions between ligands from different clusters and GBA.	34
Table 4.7.	Rigid docking and IFD results of hit 15 Ligands.....	35
Table 4.8.	Druglikeness properties of top 5 ligands.....	36
Table 4.9.	Pharmacophore sites of known pyrrolidine-based GBA inhibitors.	40
Table 4.10.	Scores of top 5 hypotheses constructed using 30 pyrrolidine based inhibitors.	40

Table 4.11.	Fitness of each compound for different hypotheses and their pIC ₅₀ values	41
Table 4.12.	QSAR results for 182 Ligands	42
Table 4.13.	QSAR results for 111 Ligands	43
Table 4.14.	QSAR results for 49 ligands	43
Table 4.15.	30 Pyrrolidine based ligands chosen for QSAR study.	44
Table 4.16.	QSAR results of 30 ligands.....	45
Table 4.17.	QSAR results for top two hypotheses	45
Table 4.18.	XP Gscore and druglikeness properties of 5 commercial inhibitors of GBA.	52
Table 4.19.	Interactions of commercial 5 ligands with the active site of GBA.	54
Table 4.20.	Interactions of best 10 experimentally known ligands and well-known GBA inhibitors with GBA.	56
Table 4.21.	Cluster results for AAADR.305 hypothesis.....	59
Table 4.22.	Substructure results of clusters obtained from AAADR.305 hypothesis	60
Table 4.23.	Cluster results for ADDHR.125 hypothesis.....	60
Table 4.24.	Substructure results of clusters obtained from ADDHR.125 hypothesis	61

Table 4.25.	ZINC ID, XP GScore and strain values of ligands in cluster 7	64
Table 4.26.	ZINC ID, XP GScore and strain values of ligands in cluster 9	64
Table 4.27.	ZINC ID, XP GScore and strain values of ligands in cluster 13	64
Table 4.28.	Representative interactions of ligands in different clusters of ADDHR.125 hypothesis.....	66
Table 4.29.	ZINC ID, XP GScore and strain values of ligands for cluster 12	68
Table 4.30.	ZINC ID, XP GScore and strain values of ligands for cluster 17	69
Table 4.31.	Representative interactions of ligands in different clusters of ADDHR.305 hypothesis.....	70
Table 4.32.	Induced fit docking scores of well-known ligands of GBA and commercial compounds.....	71
Table 4.33.	Scores and druglikeness properties of proposed 10 ligands of ADDHR.125 hypothesis for IFD.....	72
Table 4.34.	Proposed 27 ligands of AAADR.305 hypothesis for IFD.	72
Table 4.35.	XP Gscore of Ligands bound to Different Proteins.	74
Table 4.36.	Binding Free Energy Values of Selected Ligands	75
Table 4.37.	Predicted IC50 values of selected ligands using IFD XP GScore.	76
Table 4.38.	Average RMSD values of different poses of well-known and best ligands from all hypothesis.	79

Table 4.39.	Average RMSD values of protein-ligand complexes	80
Table 4.40.	Properties of top ligands from ligand-based and structure-based modeling and their derivatives.	81
Table 4.41.	Functions of molecules having scaffold of tricyclic pyrido-thieno-pyrimidine.....	99
Table 4.42.	Functions of Compounds with Dioxino-Quinolione.....	105
Table 4.43.	Binding free energy of NN-DNJ and proposed compounds for whole simulation.....	119
Table 4.44.	Protein RMSD and RMSF values of significant mutated residues and glycosylated Asn 19.....	122
Table 5.1.	Propoerties of existing and proposed compounds for GBA.	125
Table A.1.	Glide GScore and strain energy values of 96 ligands obtained by structure-based modeling.....	137
Table A.2.	ADME properties of survived 96 compounds of HRH hypothesis.....	138
Table A.3.	Clusters involving less than 4 ligands.....	143
Table A.4.	ZINC ID,GScore and strain energy of compounds retrieved by using the scaffold of cluster 6.....	143
Table A.5.	ZINC ID, GScore and strain energy of compounds retrieved by scaffold of cluster 8	144

Table A.6.	ZINC ID, GScore and strain energy of compounds retrieved by scaffold of cluster 10	144
Table A.7.	Interactions of ligands in cluster 1 of structure-based modeling	145
Table A.8.	Interactions of ligands in cluster 4	146
Table A.9.	Interactions between GBA and ligands in cluster 6	149
Table A.10.	Interactions between GBA and ligands in cluster 8	151
Table A.11.	Interactions between GBA and ligands in cluster 10	152
Table A.12.	QLogBB, QLogPo/w and XP GScore of 182 ligands obtained from ChEMBL.....	154
Table A.13.	218 survived compounds from hypothesis AAADR.305	159
Table A.14.	241 survived compounds from hypothesis ADDHR.125	162
Table A.15.	ZINC ID, XP GScore and strain of ligands for cluster 11 for ADDHR.125	165
Table A.16.	ZINC ID, XP GScore and strain values of ligands for cluster 6 for AAADR.305.....	166
Table A.17.	ZINC ID, XP GScore and strain values of ligands for cluster 7 for AAADR.305.....	166
Table A.18.	ZINC ID, XP GScore and strain values of ligands for cluster 16 for AAADR.305	167

Table A.19.	Significant interactions of ligands in cluster 7 of ADDHR.125 with GBA.....	169
Table A.20.	Significant interactions of ligands in cluster 9 of ADDHR.125 with GBA.....	170
Table A.21.	Significant interactions of ligands in cluster 11 of ADDHR.125 with GBA.....	170
Table A.22.	Significant interactions of ligands in cluster 13 of ADDHR.125 with GBA.....	171
Table A.23.	Significant interactions of ligands in cluster 6 of AAADR.305 with GBA.....	171
Table A.24.	Significant interactions of ligands in cluster 7 of AAADR.305 with GBA.....	172
Table A.25.	Significant interactions of ligands in cluster 12 of AAADR.305 with GBA.....	173
Table A.26.	Significant interactions of ligands in cluster 16 of AAADR.305 with GBA.....	175
Table A.27.	Significant interactions of ligands in cluster 17 of AAADR.305 with GBA.....	178
Table A.28.	RMSD values of hit 8 ligands	178

LIST OF SYMBOLS/ABBREVIATIONS

2D	Two dimensional
3D	Three dimensional
ADME	Adsorption, distribution, metabolism and excretion
BBB	Blood brain barrier
BEI	Binding efficiency index
Coul	Coulomb Energy
Covalent	Covalent binding energy
DAB	dideoxy-1,4-imino-D--arabinitol
ER	Endoplasmic reticulum
GBA	Glucocerebrosidase
H _{bond}	Hydrogen binding
HTVS	High-throughput virtual screening
IFD	Induced fit docking
LAB	dideoxy-1,4-imino-L-arabinitol
LEI	Ligand efficiency index
Lipo	Lipophilic term derived from hydrophobic grid potential
LLE	Lipophilic ligand efficiency
Metal	Metal binding term
NB-DNJ	N-butyl-deoxynojirimycin
NN-DNJ	N-nonyl-deoxynojirimycin
MD	Molecular dynamics
MW	Molecular weight
Packing	Pi-pi packing energy
PC	Pharmacological chaperone
PDB	Protein databank
PHOA	Percent Human Oral Absorption
PSA	Polar surface area
QLogPBB	Predicted blood brain barrier coefficient

QLogPo/w	Predicted octanol/water partition coefficient
QSAR	Quantitative structure activity relationships
Rewards	Rewards and penalties for different features
RMSD	Root mean square deviation
RMSE	Root mean square error
RMSF	Root mean square fluctuation
RotB	Penalty for Freezing Rotatable Bonds
SEI	Surface efficiency index
Site	Polar Interactions in the active site
Solv_GB	Generalized Born electrostatic solvation energy
SP	Standart precision
STD	Standart deviation
XP	Extra precision
vdW	Van der Waals energy
VSW	Virtual screening workflow

1. INTRODUCTION

Drug discovery is a process of identifying new molecules using both experimental and computational techniques. The drug is generally an inhibitor or an activator molecule against the target protein, which is not functioning properly. The interaction between the target and putative drug is analyzed using various techniques such as docking and/or molecular dynamics simulations and possible ligands (drugs) are identified in order to overcome the symptoms of the mutated or disfunctioning protein. One attractive area for drug discovery is lysosomal storage diseases which are caused by mutations in lysosomal enzymes. Mutant enzymes are not able to pass the quality control mechanisms of the endoplasmic reticulum leading to reduced lysosomal trafficking and abnormal substrate accumulation.

For lysosomal storage diseases, there are different therapeutic approaches, two of which are substrate reduction and enzyme replacement therapy. The principle of substrate reduction therapy is based on inhibition of the enzyme and prevention of the formation of product, whereas enzyme replacement therapy involves infusion of the manufactured enzyme to the patient[1]. However, enzyme replacement therapy is limited and highly costly. An alternative therapy is to use pharmacological chaperones in order to correct the folding of the mutant enzyme and hence to treat the disease. Pharmacological chaperones are small molecules that bind and stabilize the mutated enzyme and help enzyme pass quality control mechanism of endoplasmic reticulum and dissociate from the enzyme when it reaches to lysosome.

Degradation or deposition of different lysosomal proteins results in different diseases. For instance, Fabry Disease is caused by the mutation in α -galactosidase A protein, Krabbe Disease is caused by degradation in galactosylcerebrosidase enzyme. Gaucher Disease is another lysosomal disease, and it appears when glucocerebrosidase (GBA) enzyme loses its activity, even partially. GBA enzyme is responsible for hydrolyzing sphingolipid glycosylceramide into glucose and ceramide[2]. When GBA does not work properly, glycosylceramide accumulates within the lysosome and cells start to

swell [3]. The purpose of this study is to identify potential ligands or chaperone molecules that will facilitate proper folding of mutated GBA protein and consequently enable it to work in the cell.

In the present thesis, potential beta-Glucocerebrosidase chaperones were determined using both ligand-based and structure-based pharmacophore modeling techniques. ZINC database consisting of 500,000 commercially available molecules were searched by using pharmacophore hypotheses obtained by QSAR and E-pharmacophore studies, in order to obtain ligands of the same properties with these pharmacophore features. Then, each ligand was docked to the active site of beta-Glucocerebrosidase (GBA). Compounds with high docking scores were filtered with respect to their strain energy and efficiency values as well as ADME properties. The remaining compounds were clustered based on their structural similarity and their common scaffolds were used in order to perform a second round of screening from the ZINC Database.

This thesis is organized as follows: Theoretical background section includes structural and biochemical information about GBA protein as well as previous work on the inhibition of the protein. Methods section includes computational tools that are used for all virtual screening, docking and post-docking calculations and stability analysis by molecular dynamic simulations. The computational results and their discussions are given in Result and Discussion part. In conclusion section, significant results are summarized, and suggestions for future work are given.

2. THEORETICAL BACKGROUND

2.1. Sphingolipids

Sphingolipids are made up of eighteen carbon amino-alcohol backbones that are synthesized in the endoplasmic reticulum [4]. The base molecule of the sphingolipid pathway (Figure 2.1) is ceramide. Complex sphingolipids are obtained by addition of charged, neutral, phosphorylated or glycosylated moieties[5]. Variations in these moieties leads to difference in cell functions.

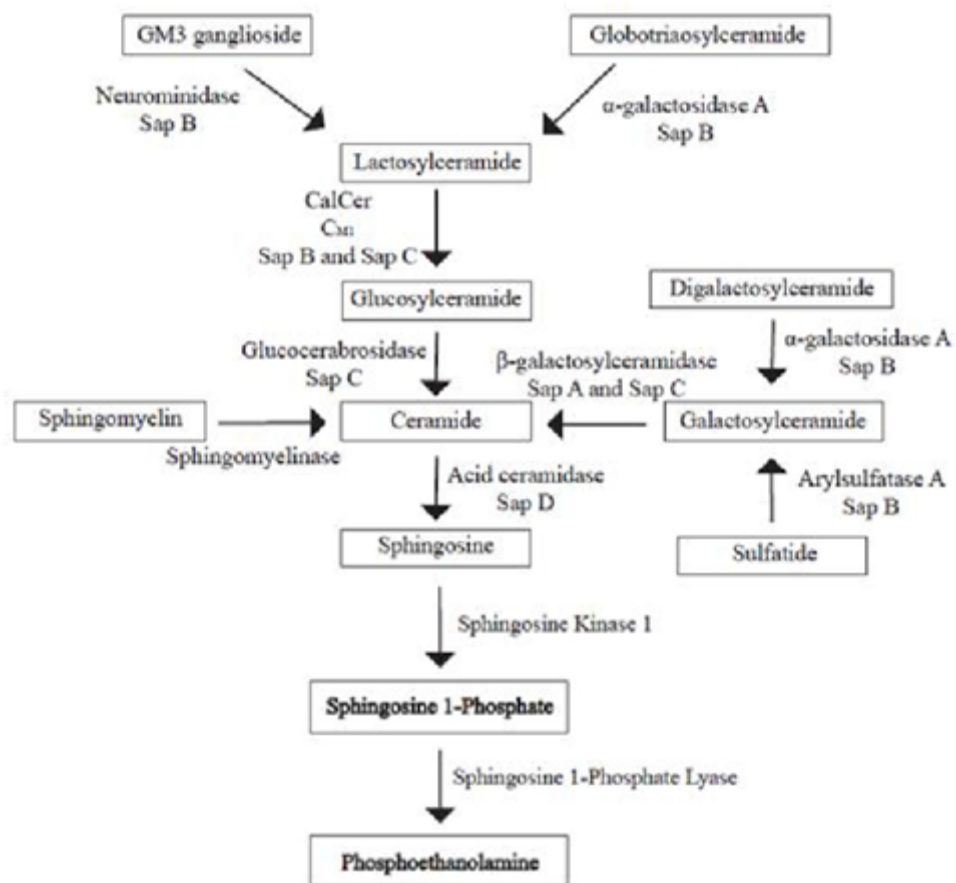


Figure 2.1. Sphingolipid pathways

Sphingolipid metabolic pathway can be defined as an interconnected network where there is a collaboration between enzymes[4]. Any mutation in the proteins, that target these enzymes leads to deficiency in cellular programs. The accumulation of these compounds in cell results to various diseases such as Alzheimer's, Gaucher, Fabry Diseases[1]. In this work, glucocerebrosidase enzyme was studied.

2.2. Beta-Glucocerebrosidase

Beta-glucocerebrosidase (GBA) protein catalyzes the breakdown of sphingolipid glycosylceramide into ceramide and glucose [6]. Mutations in GBA protein lead to accumulation of glycosylceramide in the lysosome, and consequently Gaucher Disease resulting in an increase in spleen and liver volumes and skeletal deformities[1]. Over 250 mutations were identified in GBA structure leading to Gaucher Disease. 203 of these mutations are missense mutations which influence the proper folding of GBA protein [7]. N370S is the most prevalent mutation that is seen in type 1 Gaucher Disease patients, and L444P, G202R, 84GG, F213I, R496H, V15M, S364R and G377S are other examples of these mutations. L444P is the most severe mutation leading to neuropathic disease which is very difficult to treat due to inability of current drugs to cross the blood brain barrier.

There are 23 different crystal structures of GBA in Protein DataBank. The crystal structure of GBA has been determined for the wildtype form (PDB ID: 1OGS[8]), for the different mutant (PDB IDs: 3KE0 and 3KEH) and in the presence of different compounds such as isofagomine, NB-DNJ and NN-DNJ (PDB IDs: 3GXF[9], 2V3D and 2V3E [10]respectively) GBA has two chains and GBA consists of three domains. Domains 1 and 2 involve mostly beta sheets while domain 3 is formed by 3 (α/β)₈ (TIM) barrel containing the catalytic site.

There is a hydrophobic pocket at the entrance of the active site formed by residues of Tyr 313, Leu 314, Ser 345, Glu 325 and Leu 241. The active site residues are Asn 396, Asp 127, Trp 179, Trp 381, Glu 340, and Asn 234 which are aromatic, polar uncharged and negatively charged. Tyr313 stabilizes the alkyl chains while the hydrophobic loops at the entrance of the active site help holding the ligand. The ligand binding is dependent on the

conformational changes of the loops (loop 1:residues 341-350, loop 2:Residues 393-396 and loop 3:Residues 312-319). The conformational change on loop 3 is facilitated by the formation of hydrogen bond between Tyr313 and Glu340, so these residues are important for ligand binding. The presence of water molecules near Glu235 is important for catalytic activity[10]



Figure 2.2. The structure of GBA in which active side is shown with red circle (only chain A, chain B is deleted because GBA protein functions as monomer)

2.3. Pharmacological Chaperones

Pharmacological chaperone (PC) is a small molecule that binds to target protein in the endoplasmic reticulum and provides the stabilization of the protein properly, and allows the protein to traffic through the pathway to lysosome. Most pharmacological chaperones are competitive inhibitors that bind to the target protein reversibly. They bind to the catalytic site of the protein, so that their lysosomal half-life needs to be shorter than the half-life of mutant enzyme. In addition, the binding affinity of chaperone must decrease in lysosome to allow substrate binding[1].

Mutated proteins generally misfold and misfolded proteins are recognized and degraded by endoplasmic reticulum (ER) quality control system. PCs bind to these misfolded proteins in ER and stabilize them allowing to pass through the endoplasmic reticulum and trafficking through the golgi. After glycosylation, the protein with PC is transferred to lysosome. Because of the change in pH, binding affinity of PC toward the protein decreases and leaves the protein in order to allow substrate binding.

The mechanism of action of pharmacological chaperones is shown in Figure 2.3.

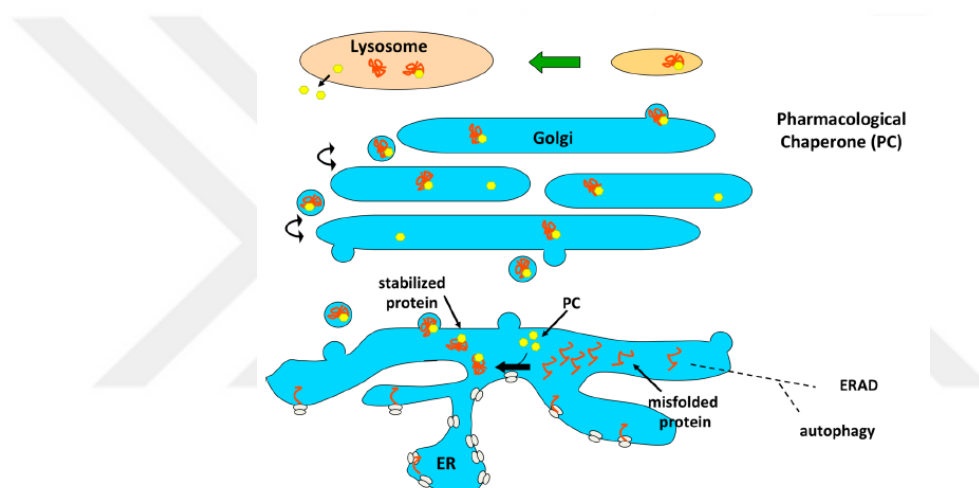


Figure 2.3. Mechanism of action of pharmacological chaperone. [1]

There have been many different types of pharmacological chaperones that were tried to treat Gaucher Disease[1]. However, in addition to important biological properties, PC used for Gaucher Disease treatment needs to have ability to pass the blood brain barrier because the disease is neuropathic. Some of these studies are presented in the next section.

2.4. Previous Works on Glucocerebrosidase and Gaucher Disease

Lysosomal enzymes are responsible for hydrolysis of glycolipids, oligosaccharides and glucoproteins. The mutated enzyme loses its activity, stability and is unable to pass the control mechanism of endoplasmic reticulum. The lysosomal enzyme, Beta-Glucocerebrosidase (GBA), is responsible for the hydrolysis of glycosylceramide. Many different types of molecules (inhibitors or chaperons) are investigated in order to enable

mutant Glucocerebrosidase to work properly and consequently prevent Gaucher Disease. There are three types of Gaucher diseases. In Type 1 Gaucher disease (adult, chronic nonneuropathic) GBA level is between 12 and 44% and GBA accumulates in internal organs. Clinical deficiencies occur in childhood with hepatosplenomegaly, skin pigmentations and orthopedic problems. In Type 2, juvenile form of Gaucher disease (subacute neuropathic), in the second decade of life neurological problems appear, and patient may die. In the infantile form Type 3 Gaucher disease (acute neuropathic) neurological manifestations begin to appear in neonatal period. Improvement of retardation, cranial nerve neuropathies and spasticity occur, and patients usually die under 2 years old.

Recently, there are two therapeutic options for Gaucher Disease which are Enzyme Replacement Therapy and Substrate Reduction Therapy. Enzyme Replacement Therapy utilizes recombinant GBA. This intravenous therapy is based on increasing the GBA level in the Gaucher Disease patient in order to prevent accumulation of glycosylceramide in the cell. However, the method fails in neurological manifestations of Gaucher disease and becomes ineffective[6]. In addition, it is expensive and requires bi-weekly infusion of recombinant GBA. Substrate Reduction Therapy aims to reduce the level of natural ligand, glycosylceramide, rather than increasing the level of deficient enzyme. In contrast to Enzyme Replacement Therapy, it is an oral treatment and iminosugars are used in order to treat Gaucher Patients. Miglustat inhibits the ceramide-specific glycosyltransferase reaction leading to substrate biosynthesis. First treatment of patients with type 1 Gaucher Disease using Miglustat showed that mean liver and spleen volumes decreased by 12% and 19% respectively, while mean hemoglobin level increased [11]. Nevertheless, both enzyme replacement and substrate reduction therapies are effective for nonneuropathic Gaucher disease, and their biodistribution is very unfavorable leading to many side effects[1].

Pharmacological chaperone therapy is a new approach to treat lysosomal diseases. Many different types of mutated enzymes result in lysosomal storage diseases such as Fabry, Pompe and Gaucher Diseases etc. The transport mechanisms of these pharmacological chaperones from endoplasmic reticulum to lysosome were explained in detail in section 2.3. Particularly, it is stated that 80% of Gaucher Disease patients have either N370S or L444P mutation (Table 2.1.), which leads to accumulation of glycosylceramide, natural

ligand of GBA. Because of their similarities to natural ligand, carbohydrate mimetics and non-carbohydrate compounds have been utilized as pharmacological chaperones by highthroughput screening in the treatment of Gaucher Disease [1].

There are many different types of carbohydrate mimetics such as iminosugars, carbasugars, azasugars etc. Well known iminosugar inhibitor is N-(n-butyl) deoxynojirimycin (NB-DNJ, Miglustat). A study revealed that Miglustat is a chemical chaperone[6]. The activities of wild typeGBA and the S364R, N370S, V15M and M123T mutants were found to increase upon incubation with 10 μ M Miglustat. Molecular docking studies were carried out in order to identify binding sites of the protein. Two alternative isomers of Miglustat were docked to GBA protein independently and surface of GBA was searched. Molecular docking of Miglustat to GBA showed that the best-ranked site was the substrate-binding site of the enzyme, suggesting that Miglustat acted as an inhibitor at the binding site[6]. Another effective iminosugar is N-(n-nonyl) deoxynojirimycin (NN-DNJ). Fibroblast culture medium having N370S mutation was treated for 9 days with 10 μ M of NN-DNJ and a 2-fold increase in activity of GBA was observed. The activity of wild type GBA also increased, but GBA L444P mutant was not affected by treatment with NN-DNJ[12]. 20% increase in enzyme activity was observed in the treatment of N188S mutation with 5 μ M of NN-DNJ [13]. Crystallographic studies on the interaction between GBA and NB-DNJ and NN-DNJ showed that both inhibitors bind to the active site of GBA via hydrogen bonding with the imino sugar moiety as well as by hydrophobic interactions with the alkyl chain of iminosugars. The pKa value of NB-DNJ was shown to play an important role in fulfilling its role as chaperone. In endoplasmic reticulum, the pKa value of NB-DNJ is very close to neutral pH, and thus NB-DNJ can be neutral and less soluble, which raises its tendency to bind to the active site of GBA. However, NB-DNJ is charged in the acidic environment of the lysosome leading to an increase in its solubility and a decrease in the free energy of binding, which in turn causes less competition with GBA [10]. Iminosugars such as Miglustat and castanospermine were found to be ineffective in clinical trials because of their lack of selectivity to GBA. Some modifications are needed on that carbomimics to increase their potency. Bicyclic galactonijimycin analogues were also studied for the treatment of Gaucher Disease.

A set of compounds consisting of monosaccharides, heterocyclic cores and exocyclic nitrogen substituents were synthesized in order to determine their inhibitory activities. Comparison of these sp^2 -iminosugars with commercial glycosidase showed that molecules with lipophilic substituents are potential pharmacological chaperones of β -glucocerebrosidase. In addition, they measured the relative activity of sp^2 -iminosugars with respect to NN-DNJ in order to determine ability of compounds to inhibit thermal-induced denaturation of GBA. It was found that amphiphilic- sp^2 iminosugars are more effective than NN-DNJ on stabilizing GBA protein[14].

Isfagomine is an azasugar mimetic, and a competitive inhibitor of GBA. Isfagomine forms hydrogen bonds with GBA at the active site of protein, and increases the stability of protein in wild-type cells as well as pre-treated N370S/N370S and F213/L444P patient fibroblasts[15]. Besides in vitro studies about isfagomine showed that D-isfagomine has IC₅₀ Value of 0.063 μ M, and positively charged D-isfagomines improved the potency towards GBA. In addition, the epimerization of C3 OH group was significantly decreased by over three orders of magnitude in the presence of positively charged D-isfagomines. In contrast, L-isfagomine is not as effective as D-isfagomine with an IC₅₀ value of 8.7 μ M [16].

Carbasugars are also utilized as carbomimetic compounds against Gaucher Disease. N-Octyl-Beta-Valienamine (NOV) was utilized to treat F213I mutations of GBA protein. Incubation of cultured Gaucher Disease fibroblast with NOV at 30 μ M for 4 days resulted in 6-fold increase in activity of F213I/F213I mutant of GBA. Nevertheless, other mutations such as N370S, L444P, 84GG were not influenced by treatment with NOV [17].

Seven-membered carbohydrate-like lactams were also studied to treat Gaucher Disease. A significant result was obtained with ϵ -glucoimidazole which was studied both experimentally and computationally. As a result, all ϵ -glucoimidazole derivatives were found to be more effective than iminosugar, in terms of enhancing the activity of N370S mutated GBA proteins. Especially, ϵ -glucoimidazole having R group ($CH_2)_8CH_3$ increased the activity of N370S mutated GBA up to 2.4 fold. This derivative was docked flexibly to

the binding site of GBA (PDB ID:2V3E). Consequently, 8 hydrogen bonds were formed with E235, N396, S345, D127, W179 and N234 aminoacids of target protein[18].

Glucoimidazoles (GIZ), pyridoazole type pharmacological chaperones have also been utilized for GD. C-2 substituted glucoimidazole derivatives were used in the study of Li et al (2013). After the evaluation of IC₅₀ values of GBA protein at pH=5.2 and pH=7, they concluded that C-2 derivatives of glucoimidazole enhanced the activity of N370S mutated GBA, imiglucerase, 2- to 30- fold in comparison to NN-DNJ. Therefore it is found that lipophilic chain at C-2 is very significant for GBA inhibition. 3,3-dimethyl-N-phenyl-4-amide-1-butyl substituted GIZ was selected for further studies because of its high sensitivity to GBA from almond. According to best docking pose of this ligand, 3,3-dimethyl-N-phenyl-4-amide-1-butyl group formed bent conformation between loop 3 and loop 4. In addition catalytic residues Glu340 and Glu235 were stabilized by hydroxymethyl of sugar moiety. Besides, experimental studies revealed that 3,3-dimethyl-N-phenyl-4-amide-1-butyl substituted GIZ protected recombinant GBA, imiglucerase, from proteolysis, it implies that the ligand is successful at stabilizing the protein[19].

Another class of carbohydrate pharmacological chaperones used in the treatment of Gaucher Disease includes aminocyclitols[7]. Molecular docking, molecular dynamic simulations and interaction energy calculations showed that alkyl chain is the major modulator of affinity of aminocyclitols. The length of linker between the triazole ring and aminocyclitol also plays a significant role in ligand affinity. Moreover, it was found that mobility of the ligand and strength of electrostatic interactions of the triazole and aminocyclitol moieties depended on the alkyl substitution in the ligand[2]. The effect of aminocyclitols on G202R mutation was investigated using computational and experimental techniques. All synthesized aminocyclitols were highly active as chaperones in fibroblasts taken from Gaucher Disease patients having G202R mutation. Similarly, the use of Iminosugars (NB-DNJ and NN-DNJ) and isofagomine also increased the activity in fibroblast cells having same mutation. Nevertheless the molecular dynamics studies were unsuccessful, i.e. the comparison between wild type and G202R mutated enzymes did not give any reliable information about the activity impairment in GBAbearing G202R mutation[20]. In short, the activity of GBA was found to increase with both NB-DNJ and

NN-DNJ in transfected cell lines, whereas only NN-DNJ and an aminocyclitol derivative, which has $C_{10}H_{21}$ alkyl group, increased GBA activity in patient fibroblasts[13].

Fluorous iminoalditol is another pharmacological chaperone in the treatment of Gaucher Disease. Iminoalditols form ionic bonds with carboxyl group of active site residues of GBA. Oligofluoro moieties in fluorous iminoalditols increase hydrophobicity and enhance the activity or selectivity of GBA more than simple lipophilic N-alkyl derivatives. Length of spacer arms and number and distribution of fluorine atoms in fluorous substituent affected binding affinity[21].

Non-carbohydrate molecules such as ambroxol were also considered as chaperones. Maegawa et al. showed that after binding of Ambroxol, amino acid regions, 243–249, 310–312, and 386–400, which are near the active site of GBA, became stable. Furthermore activities of N370S and F213I mutants GBA increased significantly. Treatment of GD patient fibroblasts having homozygous N370S mutation with 10 μ M, 30 μ M and 60 μ M ambroxol showed that both enzyme activity and protein level increased with increasing concentration of ambroxol. As a result of experimental studies, 10 μ M of ambroxol treatment provided 2-fold increase in GBA activity and 1.5-fold increase in protein level. With respect to heterozygote F213I/L444P mutation, treatment with the same amounts of Ambroxol also enhanced both GBA activity and protein level up to 3.5-fold and 2-fold respectively. Nevertheless the severe mutation L444P mutation was not recovered by the addition of Ambroxol. Activity and protein level increase with Ambroxol were lower than those of obtained with Isfagomine[7]. In clinical trials with 12 patients, one patient showed a great response to ambroxol, his spleen and liver volume decreased significantly while the level of hemoglobin and platelet increased dramatically. This result indicated that Ambroxol may be a safe method of treatment in Gaucher Disease[22].

Another non-carbohydrate compounds used in GD treatment is Diltiazem, a L -type Ca^{+2} channel blocker. It was found that Diltiazem partially restored the activity of GBA with N370S and L444P mutations. Diltiazem increased the activity of GBA enzyme in normal cells and in cells with N370S/N370S and F213I/L444P mutations. Also, Diltiazem increased the heat stability and inhibited the GBA protein at neutral pH whereas in

lysosomes at acidic pH it did not show any inhibition or any increase in the heat stability of the protein. Verapamil did not enhance intracellular levels of GBA, while Diltiazem increased the GBA level in cells carrying N370S or F213I mutations, but it is ineffective at enhancing GBA in L444P mutant cells. Co-treatment with Diltiazem and Isofagomine did not increase enzyme activity of mutant cells having N370S and L444P/L444P mutations, whereas nearly 2.5 fold increase in activity was observed in the treatment of F213I/L444P mutant GBA protein with 5 μ M Diltiazem and 10 μ M isofagomine [23].

These pharmacological chaperones are also inhibitors of GBA. Because of their ability to inhibit protein, it is very critical to balance their concentrations in the cell. As an alternative, non-inhibitory chaperones of GBA were also studied in 2012. Both high throughput screening and immunostaining methods were utilized to identify novel non-iminosugar small compounds. Pyrazolopyrimidine is a potential activator of GBA (Figure 2.4), and it is highly GBA selective. Pyrazolopyrimidine has low AC₅₀ values at around 6 μ M for both purified wildtype GBA and N370S enzyme. Compound can activate the hydrolysis of two different ligands because of difference in substrate binding and it can stabilize one particular conformation more than the other conformation. Major advantage of non-inhibitory chaperones is that their removal by the natural substrate in the lysosome is not required for the resumption of catalytic activity[3].

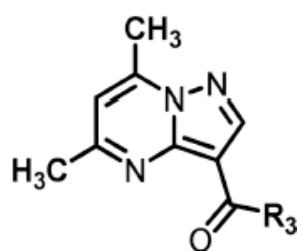


Figure 2.4. Pyrazolopyrimidine compound.

Although, iminosugars such as NN-DNJ, NB-DNJ are accepted as moderate inhibitors of GBA. However they have some limitations because of their pyronose moiety. Pyrrolidine-based iminosugars were studied both experimentally and computationally. α -1-C-alkyl-DAB and LAB compounds with various alkyl chains were synthesized to observe the effect of hydroxyl conformation and length of alkyl chain on activity of GBA. It was concluded that although LAB derivatives enhanced the activity of α -glucocerebrosidase,

they did not show any inhibition against GBA. In addition, only DAB with octyl chain inhibited GBA, and as alkyl chain length increased the activity of GBA increased significantly. It was found that interaction of α -1-C-tridecyl DAB with GBA was similar to that of isofagomine. Nitrogen and the hydroxyl groups of α -1-C-tridecyl-DAB formed hydrogen bonding with Glu235, Glu340, Asp127 and Trp179, which are very significant for ligand binding in GBA. Tridecyl group of DAB is located in hydrophobic pocket and forms hydrophobic interactions which lead to increase in GBA enzyme activity[24]. Multivalent iminosugars have also been worked as prodrugs for the treatment of Gaucher patients. As a result of immunofluorescence staining and in vitro analysis, it was observed that although trivalent iminosugars are effective on Gaucher Disease, tetravalent iminosugars and morpholine derivatives do not have any effect on GBA activity[25].

In conclusion, although new compounds have been proposed to treat Gaucher Disease, most of them fail in clinical trials. Increase in activity of GBA enzyme is not enough because Type 2 and Type 3 of Gaucher Disease is neuropathic and compounds fail to pass the blood brain barrier. Because of that reason, it is very significant to discover a compound with great ability to cross the blood brain barrier. In this regard, we propose novel ligands for treatment of Gaucher Disease by taking the blood brain barrier into consideration during the computational drug discovery study.

Table 2.1. Type of mutations in GBA and compounds used to treat symptoms of these mutations.

Type of Mutation	Compound	Reference
L444P	Diltiazem	[23]
N370S	Isofagomine	[15]
	NN-DNJ	[12]
	NB-DNJ	[6]
	Ambroxol	[7][22]
	Diltiazem	[23]
	Verapamil	[23]
	ϵ -glucoimidazole	[18]
N188S	NN-DNJ	[13]
F213I	Isofagomine	[15]
	Ambroxol	[7]
	N-Octyl-Beta-Valienamine	[17]
G202R	Aminocyclitol	[20]

3. METHODS

Virtual Screening Methods are commonly used in drug discovery studies. In this project, all virtual screening techniques of Schrödinger Suite 2013 were used to investigate the novel pharmacological chaperones for the treatment of Gaucher Disease. Schrödinger Software involves many different tasks such as LigPrep[26], Confgen[27], QikProp [28]etc. These tasks enable researchers to obtain the most reliable conformations of ligand and receptor structure to understand their interaction. In general, the workflow of this study are given in Figure 3.1

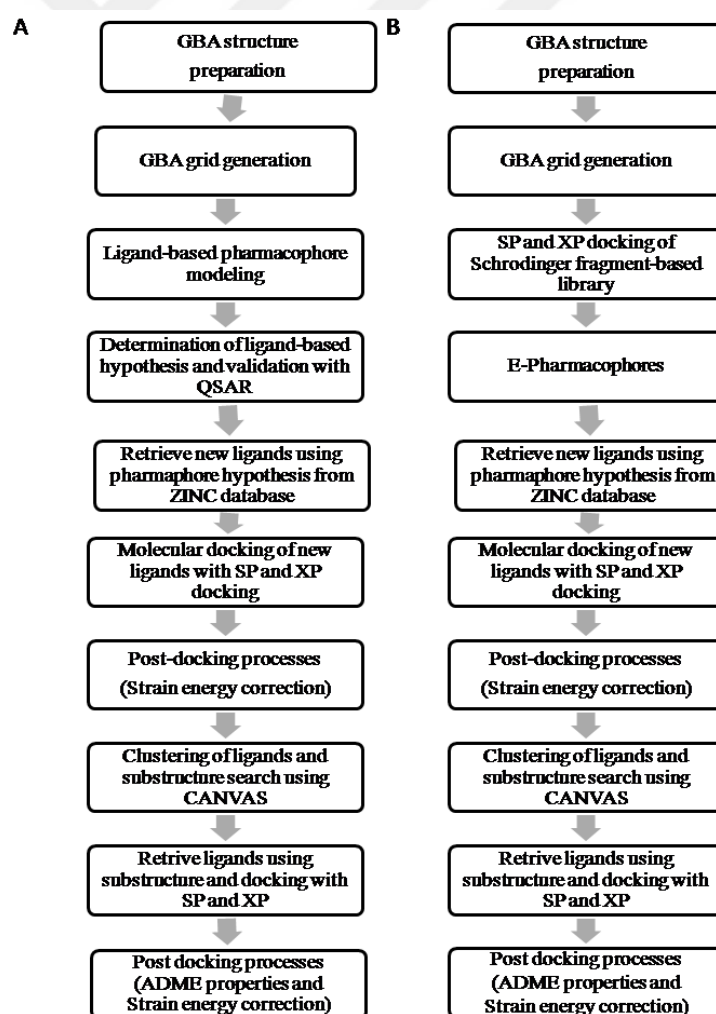


Figure 3.1. Workflow for the determination of potential pharmacological chaperones for Beta-Glucocerebrosidase (A:Ligand based model, B: Structure based model)

3.1. Receptor Structure Preparation

The crystal structure of GBA was downloaded from Protein Data Bank (PDB) (PDB ID: 2V3E). GBA structure with NN-DNJ (PDB ID:2V3E) was utilized for studies instead of wildtype form (PDB ID: 1OGS) because it is known that well-known ligands such as NN-DNJ and isofagomine bind to the active site of GBA, and in the form of 1OGS the structure is apo. Protein Preparation Tool of Schrödinger Suite 2013 was utilized to prepare protein[29][30]. Missing loops and missing residues (Chain A residues: 31, 498, 499, 500, 501, 502, 503 and chain B Residues: -1, 498, 499, 500, 501, 502, 503) of the GBA were completed via Prime module. Then, water molecules were deleted from the system except water molecules around Glu 235. Only chain A of GBA was kept in the workspace because GBA protein functions as monomer.

As a last stage, hydrogen bonding network was optimized while the protonation states of residues were determined at pH=7. Energy of all protein atoms was minimized using the OPLS 2005 Force Field and maximum RMSD of heavy atoms were set to 0.30 Å.

3.2. Receptor Grid Generation

Receptor Grid is the place where ligands are docked to protein, in general it is accepted as active site of the receptor. The electrostatic energy of protein is calculated by these grids and they are used to determine docking score of a ligand. The prepared protein structure of GBA (PDB Code:2V3E) was used to generate grid via Receptor Grid Generation task[31].

Receptor Grid Generation may composed of 5 steps. First, because receptor were rigid in Glide Docking, van der Waals radii of nonpolar atoms with partial charges were declined to allow little flexibility. Default values were used for partial charge and scaling values. Then, the center of grid box was accepted as the place where NN-DNJ binds because it is known that NN-DNJ binds to the active site of protein. The dimensions of the inner box was kept as default value which is 10 Å. No constraint was given to system and

excluded volumes were not utilized for ligand binding. The prepared grid of GBA can be shown as Figure 3.2.

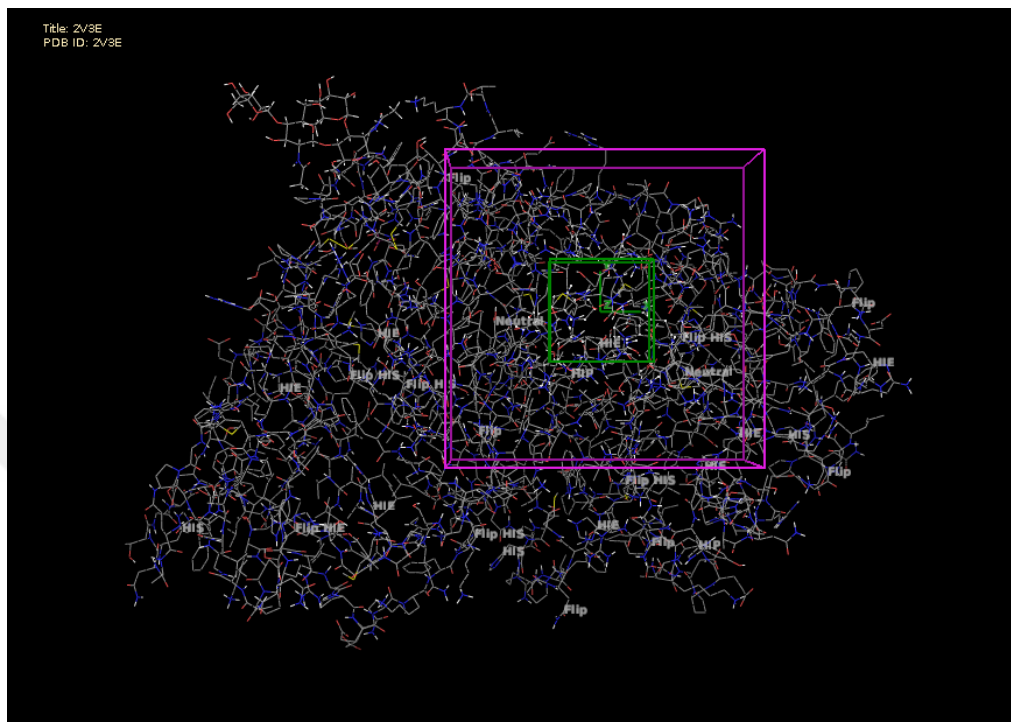


Figure 3.2. Grid of GBA generated for rigid and induced fit docking

3.3. Generation of Pharmacophore Hypothesis

Pharmacophore hypothesis represents the common features of a set of ligands. In order to obtain common pharmacophore hypothesis, there are two types of pharmacophore modeling, that are structure-based and ligand-based modeling, which will be explained below.

3.3.1. Structure Based Pharmacophore Modeling

Glide[31] and Phase[32][33][34] programs were utilized to obtain a structure-based pharmacophore hypothesis. This technique is generally used in the absence of experimentally known ligands for receptor protein. Active site of the enzyme and interactions between active site residues and the ligand were analyzed by these programs. As a result, properties of the active site of the enzyme were defined using different features such as hydrogen bond acceptor (A), hydrogen bond donor (D), hydrophobic group (H),

ring group (R) and negatively charged group (N). The Schrödinger Software has a special library for structure based pharmacophore modeling. This library consists of 667 conformers[31]. These fragments were docked using XP mode in Glide Docking. Then, E-Pharmacophore script was used in order to determine the properties of the active site of receptor protein[35][36]. Common features of each interaction between fragment and protein formed the pharmacophore hypothesis. The hypothesis can involve three to seven pharmacophore sites. HRH Hypothesis were obtained from the docking results of GBA 10,000 ligands with the same features (HRH) were screen from ZINC Database involving 500,000 compounds based on their fitness scores.

3.3.2. Ligand Based Pharmacophore Modeling

Ligand Based Pharmacophore Hypothesis was formed by using experimentally known inhibitors which were extracted from ChEMBL website[37]. ChEMBL is a website of European Bioinformatics Institute including informations such as molecular weight, smiles, inhibition constant, references about experimentally known ligands for various proteins.

Pharmacophore hypotheses were obtained by following 5 steps shown in Figure 3.3.

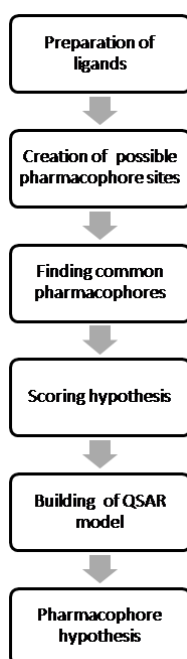


Figure 3.3. Steps of ligand-based pharmacophore modeling

First step of development of common hypothesis was the preparation of ligands. LigPrep [26] was utilized to generate low-energy 3D structures of ligands from 2D structures. Furthermore, ionization state of ligands were determined at pH=7 and a maximum of 32 stereoisomers of each ligand was generated. Afterwards, energies of compounds were minimized and maximum 100 different conformations of each ligand were selected using ConfGen[27].

In the “create sites” part, potential pharmacophore sites of each ligand were determined and pIC50 values of inhibitors were included to the system. Variant list including the number of features was formed with Phase[32] where it was chosen to have three to five features in a hypothesis. Common pharmacophore sites were obtained by using created sites.

With regards to scoring of the hypothesis, reliability of each hypothesis was determined based on the survival score. Survival score combines the alignment score, vector score and volume score. Alignment score shows the root-mean squared deviation (RMSD) of site-points of a pharmacophore from the reference compound. Vector alignment score measures how well the vectors are for acceptor and donors in the structures, that make contributions to related hypothesis while the structures are aligned to pharmacophore. Vector score is the average of cosines between the vectors of reference and aligned ligands. In addition, the overlap of the volume of an aligned ligand with the reference ligand, divided by the total volume occupied by the two ligands is defined with individual volume score. Volume score is measured by taking the average of all individual volume scores[38].

For the hypothesis having the greatest survival score, QSAR Model was built to validate it. In this step, ligands were aligned to pharmacophore sites in the proposed hypothesis and fitness score was measured. Partial Least Squares method was utilized and structural properties that affect activity of a ligand were determined via QSAR Model. 30 active ligands were divided into training set (7 compounds) and test set (23 compounds) randomly. Regression was done by using training set ligands, whereas test set molecules were used to validate the QSAR Model.

As a result, two five-membered pharmacophore hypotheses, AAADR and ADDHR were obtained. The selection of hypothesis depended on alignment of ligands in terms of site, volume, vector and total survival scores. Then, AAADR and ADDHR hypotheses were used to retrieve similar ligands from ZINC Database consisting of 500,000 molecules, top 10,000 molecules were used for further docking process.

3.4. Molecular Docking and Virtual Screening Workflow

Molecular docking can be done in two different ways. One method is called lock and key method or Rigid Docking. Internal geometry of both ligand and target remains stable during rigid docking. Second method is called Induced Fit or Flexible Docking. In this technique, the ligands and side chains of GBA protein located within 5 Å were flexible whereas other residues and backbone of protein remained stable, and binding energies of different conformations were calculated. Induced fit method takes more time than the rigid docking, but it is more reliable.

Glide, Autodock or GOLD can be given as examples of softwares used for docking. In this thesis, Glide was used for docking studies.

The energy function of Glide can be given as;

$$\text{Glide Score} = (0.05 \times \text{vdW}) + (0.15 \times \text{Coul}) + \text{Lipo} + H_{\text{bond}} + \text{Metal} + \text{Rewards} + \text{RotB} + \text{Site} \quad (3.1)$$

There are three docking modes in Glide[31]. High Throughput Virtual Screening (HTVS) mode is suitable to enrich million-compounds while Standard Precision (SP) is more reliable for docking tens to hundreds of thousands of ligands [39]. Extra Precision (XP) scoring function involves many penalties such as desolvation or charge penalty and suitable for ligands for less than 1000. XP Docking is more precise than SP Docking, because their scoring functions are different. Not only water desolvation energy but also different protein-ligand conformations that increase binding affinity are involved in XP mode. Hydrophobic interactions such as neutral-neutral single or hydrogen bonds are also

taken into account in XP mode. XP Docking is especially useful to eliminate false absolute[40].

Virtual Screening Workflow (VSW) was used in order to dock ligands to active site of the receptor protein. VSW enables to dock ligand to protein with SP and XP modes without any intervention. In this work, 1000 ligands with the highest positive SP Score were docked again using XP mode.

3.5. Induced Fit Docking

In Standard docking applications, receptor protein is not allowed to make any conformational change and it is kept rigid. Nevertheless, this may lead to deviation from the real situation. In induced fit docking, active site side chains are allowed to move [41][42].

After XP docking and filtration of ligands based on their ADME properties and strain energy values, surviving ligands were analyzed using QikProp [28] and top 5 ligands were selected for induced fit docking studies. Prime task was used to dock the chosen ligands to active site of the GBA enzyme. First, maximum 20 poses of each ligand and residues located within 5 Å of these 20 poses were generated. GBA protein was prepared with LigPrep[26] and its only active site side chains were allowed to be flexible whereas other parts kept stable. Grid box size of receptor was selected as auto and the center of the grid was set as active site of GBA protein. Then, similar to the rigid docking, top 5 ligands were redocked to the active site of protein with XP mode without constraint.

3.6. Post Docking Applications

Ligands with high docking scores were further analyzed in terms of their ligand strain energy, efficiency indices, ADME and druglikeness properties. Rigid docking results in ligand strain due to lack of receptor flexibility. Strain energies of ligands were measured using Strain-Rescore script and threshold of strain was set to 4 kcal/mol.

Ligand efficiency quantifies the molecular properties of small molecules in order to gain binding affinity to target molecule [43]. After ligands were eliminated with respect to their strain energy values, they were further analyzed according to their ligand efficiencies. Ligand Efficiency Index (LEI) which is the ratio of Gibbs Free energy to number of non-hydrogen atoms in ligand, and Binding Efficiency Index (BEI) based on the ratio of the binding affinity to molecular weight of compound are critical for guiding fragment-based strategies to accelerate drug discovery[44]. Surface Efficiency Index (SEI) is the ratio of the binding affinity of the ligand and the polar surface area (PSA) of the ligand scaled to 100 Å [44].

Calculation of BEI and SEI values were performed using Equations 3.2-3.6

$$LEI = \frac{\Delta G}{\# \text{ of non - hydrogen atoms}} \quad (3.2)$$

$$\Delta G \left(\frac{\text{kcal}}{\text{mol}} \right) = RT \ln K_i \quad (3.3)$$

$$p(K_i) = -\log(K_i) \quad (3.4)$$

$$BEI = \frac{p(K_i)}{\text{Molecular Weight (kDa)}} \quad (3.5)$$

$$SEI = \frac{p(K_i)}{\text{Polar Surface Area}/100 \text{ Å}^2} \quad (3.6)$$

Blood brain barrier partition coefficient was determined using QikProp. It mainly depends on polar surface area and octanol-water partition coefficient.

$$QLogBB = -0.0148PSA + 0.152QLogo/w + 0.139 \quad (3.7)$$

Another critical parameter of drug discovery is ADME for obtaining compounds that are reliable and suitable to function properly in human body. ADME properties, which imply the pharmacokinetic properties for adsorption, distribution, metabolism and excretion were calculated by using QikProp script.

3.7. Clustering and Substructure Search

The ligands obtained by docking were clustered with respect to their structural similarities via Canvas Module[45][46][47]. Hierarchical clustering is utilized to group the different compounds. Similarities between structures are defined by Tanimoto Coefficient. Tanimoto Coefficient is calculated with atom pairs which are the shortest path among the non-hydrogen atoms, and it changes between zero (0) and one (1) while one (1) represents the maximum similarity between compounds.

Structural similarity was utilized to determine the common scaffold found among the structures of ligands in a cluster. Smiles of common scaffolds were written to ZINC Website and the system formed the 2D structures of scaffolds. Using these 2D structures, ZINC Website was searched immediately, and new ligands were found without any commercial concern. ZINC uses ChemAxon's JChemBase to search twenty seven million compounds[48]. This step is very significant for recovering the missed ligands during filtration with pharmacophore hypothesis.

3.8. Selectivity Analysis

Top ligands were examined for their selectivity to GBA Protein. They were redocked with XP mode to the other proteins that are the members of sphingolipid metabolism and close to GBA Protein. Alpha-Glucocerebrosidase (PDB Code:3TOP) and Alpha-Galactosidase (PDB Code:3GXT) receptors were utilized for that analysis. Ligands having positive XP GScore values for these two proteins were further analyzed by Molecular Dynamic Simulations.

3.9. Binding Free Energy Calculations

Binding free energy is defined as the energy difference between the free energy of protein ligands complex and sum of the free energies of protein and the ligand. Relative binding affinities of various compounds were calculated using Prime module of Schrodinger[49]. MM-GBSA (Molecular Mechanics-Generalized Born Solvent Accessible Surface Area) was used to compute relative binding free energies.

Snapshots extracted for the trajectories are utilized to compute the change in energy during the simulation. Binding free energy of protein-ligand complex is computed by summation of two terms which are total enthalpy change and entropy change multiplied by temperature. Total enthalpy change consists of the summation of solvation free energy for transferring the solute from vacuum into solvent and gas entalpy which is determined as the sum of van der Waals, electrostatic, and internal energies in gas phase. Entropy change is the sum of the translational, the rotational, and the vibrational entropies of each species[50].

3.10. Molecular Dynamics Simulations

Molecular Dynamics (MD) Simulations were performed in order to determine the stability of protein-ligand binding via Desmond [51][52][53]module of Maestro [54]using the OPLS 2005 force field. Selective ligands which were chosen based on induced fit docking scores and druglikeness properties such as octanol-water partition coefficient and blood brain barrier partition coefficients were simulated in an orthorombic box of SPC solvent model water molecules. Salt concentration was set to 0.15 M to obtain physiological conditions. The system was relaxed and minimized before starting the simulation. The temperature of simulation was held at 310 K and 1 atm for 50 ns.

4. RESULTS AND DISCUSSION

4.1. Structure-Based Modeling

4.1.1. Structure-Based Pharmacophore Modeling with E-Pharmacophore

Structure-based pharmacophore modeling was performed by using Schrödinger Fragment Library which consists of 667 fragments. Each fragment was docked to binding site of Glucocerebrosidase. The binding modes and energetic terms of the fragments were determined and pharmacophore features were obtained as shown in Table 4.1.

Table 4.1. Ranks and scores of the top seven pharmacophore features

Rank	Feature label	Score	Type
1	H1422	-8.17	H
2	R2304	-4.95	R
3	H1436	-3.59	H
4	R2278	-2.88	R
5	D1026	-2.75	D
6	A900	-2.52	A
7	A688	-2.40	A

The binding site of Glucocerebrosidase enzyme includes Asn 396, Asp 127, Trp 179, Trp 381, Glu 340, Asn234, and its hydrophobic pocket consists of Tyr 313, Leu314, Ser 345, Glu 325 and Leu 241. Both H (hydrophobic) and R (Aromatic ring) features implied these hydrophobic properties. As the top 3 features represented hydrophobic properties and their scores were significantly higher, the pharmacophore sites obtained by E-pharmacophore satisfied the hydrophobic feature of the binding site. The hydrophobic ligands are expected to interact with the hydrophobic pocket near the protein active site.

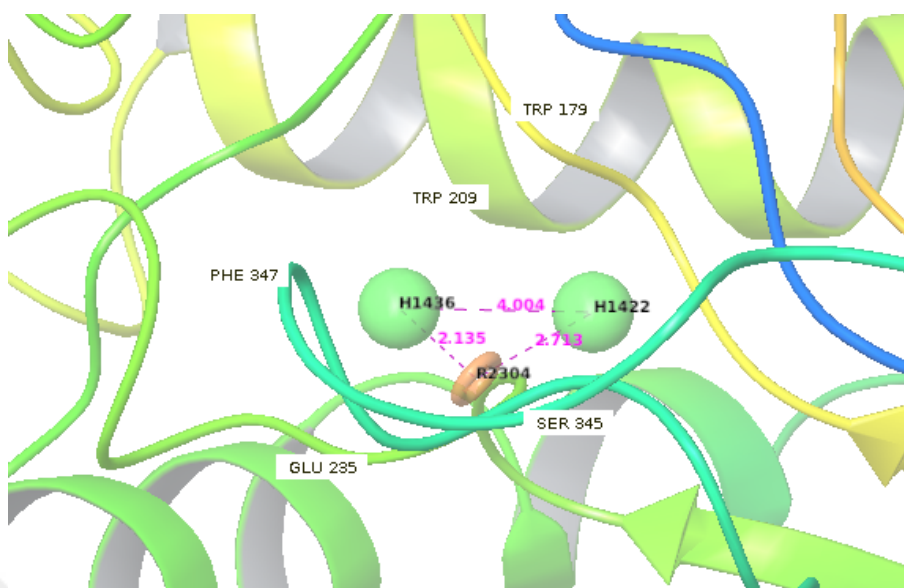


Figure 4.1. Site point locations and types at the active site of GBA

The inner-site distances between pharmacophore features were tabulated in Table 4.2. Potential ligands were desired to be small molecules in order to cross the cell membrane, therefore, small distances between the pharmacophore features were preferred. The number of pharmacophore properties was kept at three to obtain as many compounds as possible after filtering. Top three features, HRH were used to filter a ZINC Database including 500000 lead-like compounds, which were previously selected by filtering ligands of ZINC with molecular weight between 250 and 350 Da, rotatable bond less than 7 and octanol-water partition coefficient less than 3.5. As a result, top 10000 compounds with good fitness scores were obtained.

Table 4.2. Inner-site distances of three pharmacophore sites.

Site 1	Site 2	Distance (Å)
H1436	H1422	4.004
H1436	R2306	2.138
H1422	R2304	2.713

4.1.2. Molecular Docking with HRH Pharmacophore Hypothesis

Docking was carried out using the Virtual Screening Workflow (VSW). First, top 10000 potential ligands with fitness scores between 1.892 and 1.403 retrieved by HRH hypothesis were docked to Glucocerebrosidase active site without any constraint using SP docking. Then, top ranking 1000 hits from SP docking were docked to the GBA active site by XP docking. The Glide GScores of the 1000 ligands were shown in Figure 4.2.

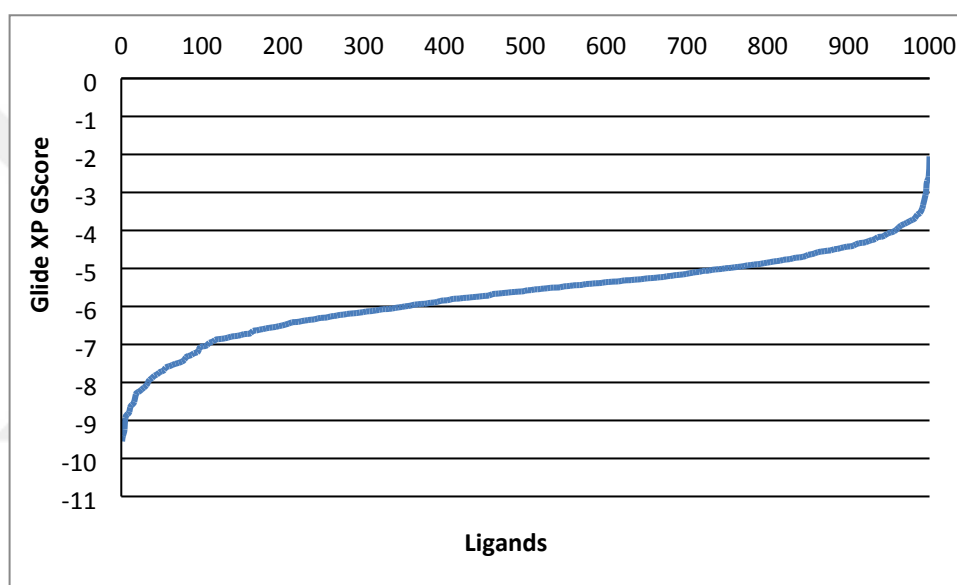


Figure 4.2. Glide GScore of top 1000 Ligands obtained by XP Docking to GBA

The Glide GScore values of 1000 ligands (XP Mode) were in the range of -9.56 and -2.05 kcal/mol. GScores of ligands decrease dramatically up to -6.5 kcal/mol but thereafter the change in GScore occurred very slowly. -6.5 kcal/mol was selected as the cut-off value and 199 ligands remained. These 199 compounds were used in the substructure search from ZINC[48].

4.1.3. Post Docking Analysis of Potential Ligands

Following the docking step, these hit compounds (199 ligands) were analyzed in terms of their strain energy, efficiency values and interacting residues. In addition, ligand efficiency index (LEI), binding efficiency index (BEI) and surface efficiency index (SEI) values of these ligands were estimated.

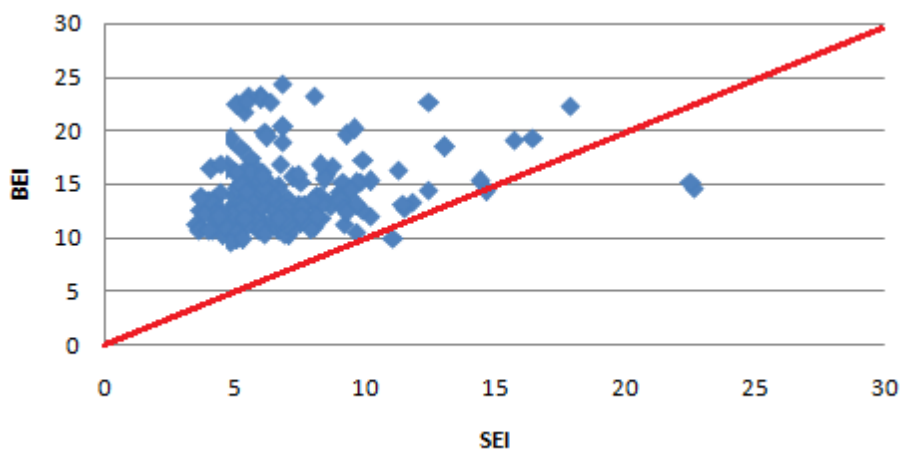


Figure 4.3. BEI-SEI values of ligands obtained by structure-based pharmacophore modeling

The BEI-SEI plot for the 199 ligands shows that most of the ligands had greater BEI values than SEI values suggesting that most of the ligands were polar. However, there were also apolar compounds in the ligand set remaining below trendline.

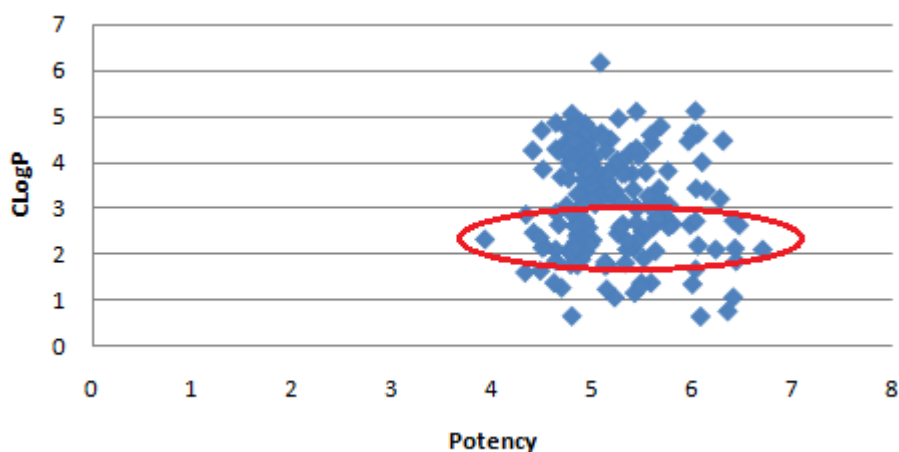


Figure 4.4. cLogP vs potency for 199 ligands obtained from HRH hypothesis

Lipophilicity of the compounds is a significant factor in the discovery of new drugs. Ligand efficiency quantifies the size and lipophilicity of small compounds to determine binding affinity of the drug to its target. Large compounds with high lipophilicity fail in clinical studies[43]. Drug candidates need to have optimum physico-chemical properties such as lipophilicity, and potency and these properties can be optimized by

experimental studies. Octanol-Water Partition Coefficient is helpful to understand where drug molecules distribute in the cell. Molecules with high octanol/water partition coefficient are essentially distributed to hydrophobic areas such as lipid bilayers of cell[55]. For drug molecules of oral administration, average potency and cLogP (octanol-water partition coefficient) values are accepted as 1-10 nM and 2.5-3 respectively[43]. In structure-based modeling studies, about 30 compounds were in the desired range of both cLogP and potency in Figure 4.4.

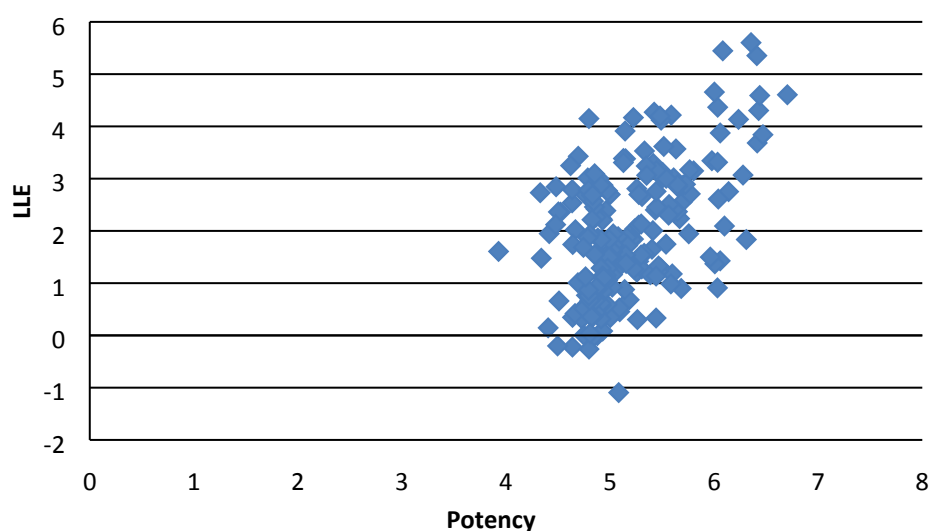


Figure 4.5. LLE vs potency for 199 ligands obtained from HRH hypothesis

Lipophilic ligand efficiency (LLE), which is calculated by taking the difference between potency and cLogP, needs to be between 5-7 for oral drugs[43]. LLE values of all compounds were lower than 6 in Figure 4.5. Nevertheless, it is known that fragments and lead-like molecules may not have drug-like LLE values. High throughput virtual screening (HTVS) method leads to obtain compounds having lower LLE values, and these can be improved by optimising the drug candidate[43]. Lead-like molecules are optimized by adding substituents onto lead template, and binding affinity and selectivity to target enzyme can thus be improved.

Strain energy calculations for all 199 ligands were carried out using Post-Docking Calculations script of Schrödinger. The cut-off value for the strain energy was taken as 4 kcal/mol, and consequently 96 out of 199 ligands survived (Table A.1).

4.1.4. Substructure Search Using Scaffolds of Compounds from Structure-Based Modeling

All 96 ligands were clustered based on their pairwise similarity in terms of Tanimoto Coefficient using Hierarchical Clustering in CANVAS application of Schrödinger. As a result of similarity analysis, 10 clusters were formed, and the common substructure of each cluster with more than 4 ligands was determined via Canvas in order to search for similar ligands in ZINC database. The ligand clusters and common substructure of each cluster were shown in Table 4.3, while the clusters with fewer than 4 ligands were tabulated in Table A.3.

Table 4.3. Clusters involving more than 4 ligands and common substructure of ligands

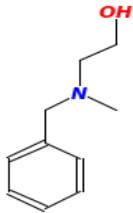
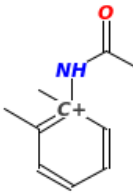
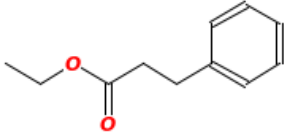
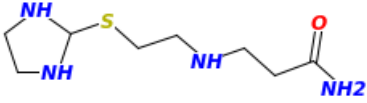
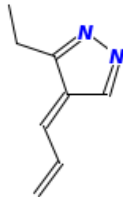
Cluster	Number of Ligands in Cluster	Number of Ligands with this Scaffold	2D View of Selected Substructure
Cluster 1	10	5	
Cluster 4	58	10	
Cluster 6	6	3	
Cluster 8	8	4	

Table 4.3. Clusters involving more than 4 ligands and common substructure of ligands (cont.)

Cluster	Number of Ligands in Cluster	Number of Ligands with this Scaffold	2D View of Selected Substructure
Cluster 10	8	8	

4.1.5. Rigid Docking of New Molecules obtained by Substructure Search

Ligands that contain the scaffolds listed in Table 4.3 were obtained using structural similarity search tool of Zinc Website involving 23 million compounds, and they are prepared with LigPrep module [56]. After 50 different conformations of each ligand in Table 4.3. were generated, and HTVS was performed. Then, top 10000 molecules were docked to the active site of the GBA protein in SP mode and best 1000 compounds were docked to the same site with XP mode. GScore cut off value was kept as -6.5 kcal/mol and strain energy cut off was taken as 4 kcal/mol and compounds were filtered with these constraints.

BEI and SEI values of remained ligands in each cluster were calculated using docking data and ADME properties of each ligand were obtained by QikProp module [28], then arithmetic average of BEI and SEI values for each cluster were estimated as shown in Figure 4.6.

BEI plot showed that all clusters had nearly the same average BEI values suggesting that their druglikeness was similar. Nevertheless there were ligands with high BEI values in cluster 1 and 8 and their ranges were wider than other clusters. With regards to SEI, cluster 1 had higher average SEI value than that of other clusters as well as higher range. Cluster 4 also had good BEI and SEI values. Glide GScore and strain energy values of

ligands in cluster 1 and 4 were shown in Table 4.4 and 4.5, Glide GScore and strain energy values of clusters 6, 8 and 10 were given in Tables A.4-A.6.

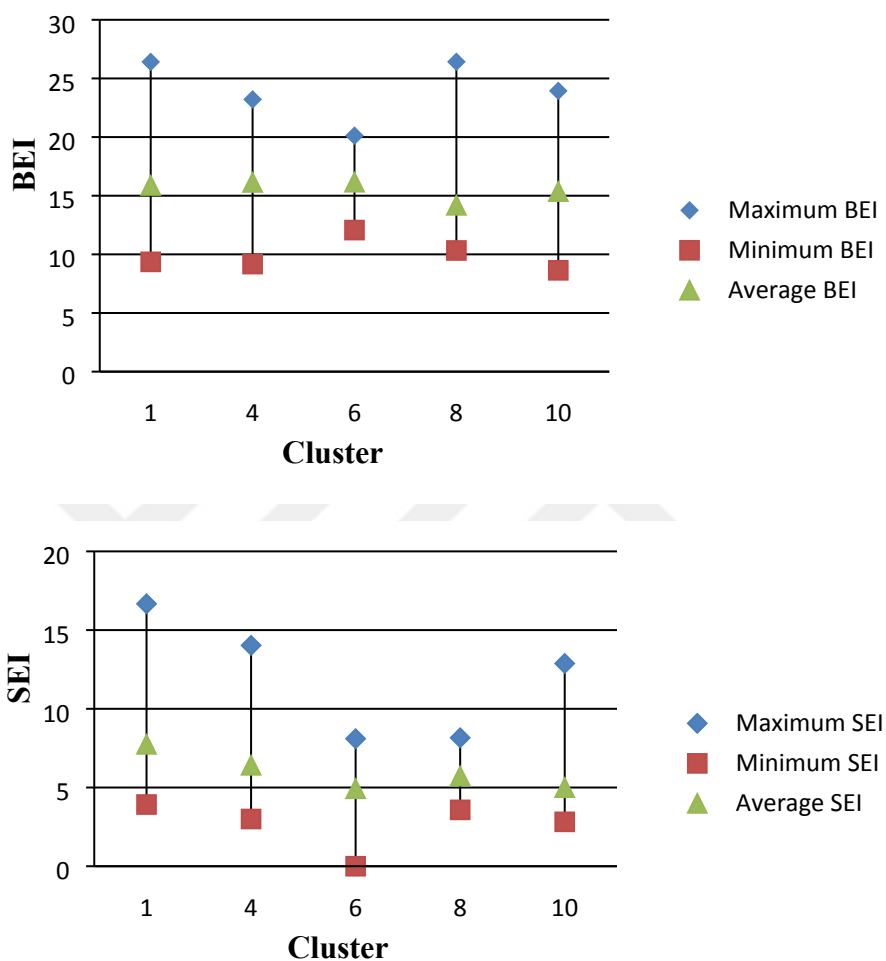


Figure 4.6. BEI and SEI plots of each cluster for structure-based modeling

Table 4. 4. ZINC ID, XP GScore and strain energy of compounds retrieved by using the scaffold of cluster 1

ZINC ID	Glide GScore	Strain Energy (kcal/mol)	ZINC ID	Glide GScore	Strain Energy (kcal/mol)
ZINC00213230	-7.693	0.481	ZINC09447905	-8.231	1.955
ZINC00214263	-7.729	0.783	ZINC12463555	-7.255	1.05
ZINC00279749	-7.704	1.569	ZINC12873329	-7.949	2.009

Table 4.4. ZINC ID, XP GScore and strain energy of compounds retrieved by using the scaffold of cluster 1 (cont.)

ZINC ID	Glide GScore	Strain Energy (kcal/mol)	ZINC ID	Glide GScore	Strain Energy (kcal/mol)
ZINC03683456	-8.227	1.478	ZINC13728463	-8.179	1.135
ZINC04860929	-6.834	1	ZINC13991807	-7.416	1
ZINC05343749	-6.574	1.463	ZINC14879274	-8.917	3.227
ZINC06753198	-6.539	1.494	ZINC19397245	-6.916	1.606
ZINC08996357	-8.206	3.105	ZINC19789949	-6.535	2.491
ZINC09010544	-6.535	2.491	ZINC19841201	-7.221	0.892
ZINC09089201	-8.732	2.049	ZINC20538875	-8.171	1.598
ZINC09241405	-8.639	2.522	ZINC09353855	-7.456	3.403

Table 4.5. ZINC ID, XP GScore and strain energy of compounds retrieved by using the scaffold of cluster 4

ZINC ID	Glide Gscore	Strain Energy (kcal/mol)	ZINC ID	Glide Gscore	Strain Energy (kcal/mol)
ZINC01726382	-10.606	3.584	ZINC00243679	-7.509	2.33
ZINC03818225	-10.11	0.219	ZINC04912920	-7.504	0.049
ZINC04687140	-10.044	3.453	ZINC00629371	-7.5	0.101
ZINC05323884	-9.978	3.631	ZINC00421743	-7.47	1.681
ZINC04728750	-9.882	2.962	ZINC04913783	-7.4	0.103
ZINC04033947	-9.863	0.794	ZINC01004334	-7.388	2.422
ZINC04650245	-9.805	3.339	ZINC00195125	-7.33	0.059
ZINC04581688	-9.525	1.668	ZINC01726383	-7.295	2.422
ZINC04581736	-9.226	2.876	ZINC04903360	-7.197	1.421
ZINC04316723	-8.974	2.345	ZINC01049890	-7.087	0.216
ZINC04743962	-8.838	3.174	ZINC02427877	-7.064	1.321
ZINC05258075	-8.729	3.861	ZINC01049873	-7.058	0.167
ZINC05154911	-8.724	2.731	ZINC00787014	-7.057	0.904
ZINC04687139	-8.673	3.563	ZINC04815668	-7.01	1.962
ZINC04348040	-8.442	3.873	ZINC05160955	-7.004	1.435
ZINC04743969	-8.433	3.992	ZINC03505496	-6.971	0.644
ZINC00924612	-8.302	1.576	ZINC03292318	-6.967	2.397
ZINC01022581	-8.249	0	ZINC00895379	-6.963	1.488
ZINC01717584	-8.149	1.482	ZINC04676095	-6.888	0.77
ZINC04488805	-8.085	0.266	ZINC00886718	-6.875	1.802

Table 4.5. ZINC ID, XP GScore and strain energy of compounds retrieved by using the scaffold of cluster 4 (cont.)

ZINC ID	Glide Gscore	Strain Energy (kcal/mol)	ZINC ID	Glide Gscore	Strain Energy (kcal/mol)
ZINC04524123	-7.97	3.039	ZINC04249559	-6.862	0.061
ZINC01298527	-7.858	0.801	ZINC01172808	-6.829	0.515
ZINC00886837	-7.833	0.044	ZINC01948887	-6.708	3.367
ZINC00784562	-7.782	2.098	ZINC00716655	-6.64	0.714
ZINC01018029	-7.765	0	ZINC04940487	-6.64	3.26
ZINC01070395	-7.754	0.007	ZINC01225992	-6.632	0.581
ZINC00124158	-7.69	0.04	ZINC01379166	-6.612	0.227
ZINC04838248	-7.587	0.528	ZINC04676317	-6.567	3.56
ZINC03434247	-7.583	3.623	ZINC05160559	-6.567	0.622
ZINC00195055	-7.568	0.38	ZINC04868325	-6.564	0.001
ZINC00705368	-7.547	0.065	ZINC00886720	-6.555	2.462

The interactions of the ligands with GBA were examined for each ligand in these clusters using LigProt (Table 4.6. and Tables A.7-A.11). The average XP Gscore of each cluster was very close to each other while cluster 4 had the best average XP Gscore. When the interactions of ligands in cluster 4 were analyzed in detail, there was not any common interaction between the ligands and GBA whereas almost all ligands with good XP score interacted with Asn 234, Trp 381 and Trp 179. Cluster 6 consists of ligands with poor XP GScore values and that rarely made hydrogen bonds with two active site residues, Asp127 and Trp179 residues, but mostly interacted with Ser345 and Lys 346. Cluster 8 compounds interacted with Ser345 and Asn396 and the best ligand in this cluster interacts with Trp 179 and Asp 127. Ligands making hydrogen bonds with Asp 127 had better XP GScore than other ligands interacting with Ser 345 such as Zinc00122581 and Zinc05634057 suggesting that interaction with Asp 127 may be critical for binding.

Table 4.6. Representative interactions between ligands from different clusters and GBA.

Cluster	ZINC ID	Catalytic Residues		Active Site Residues				Hydrophobic Pocket Residues			Other Residues				XP Score of Ligand	Avg XP Score of Cluster	STD
		Glu 235*	Glu 340*	Asn 234	Trp 179	Asp 127	Trp 381	Asn 396	Ser 345	Tyr 313	Phe 347	Tyr 244	Lys 346	Phe 246			
1	Zinc14879274				H(s)	H(s)						π - π			-8.917	-7.618	0.918
	ZINC12873329	H(s)	H (s)		H(s)	H (s)									-7.949		0.234
4	Zinc01726382				H (s)	H (s) (2)	H (s)								-10.606	-7.994	1.850
	ZINC04033947			H(s)	H (s)		H (s)				π - π	π - π			-9.864		1.322
6	Zinc00629284						H (s)						π - π		-7.94	-7.486	0.321
	Zinc 00122581			H(s)					H(s)				H(s)		-7.487		0.000
	Zinc00178860								H(s)				H(s)		-7.343		0.101
8	Zinc90243750	H (s)			H (s)	H (s)									-8.928	-7.249	1.187
	ZINC21236699					H(s)			H(s)				H(s)		-6.838		0.291
10	Zinc05488489	H (s)							H(b)&(s)				H(s)		-7.432	-7.234	0.140
	Zinc05130058	H(s)							H(b)&(s)		π - π				-7.259		0.018
	Zinc05634057	H(s)							H(b)&(s)		π - π				-7.084		0.106
	Zinc05130057								H (b)						-6.707		0.372

4.1.6. Induced Fit Docking for Ligands by Structure-Based Modeling Studies

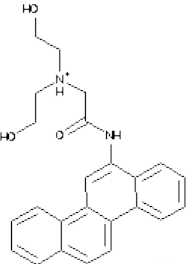
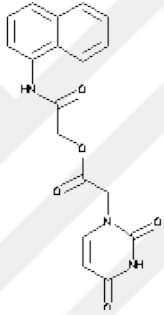
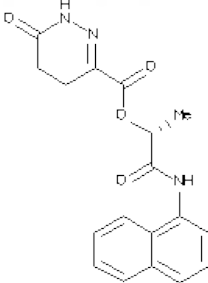
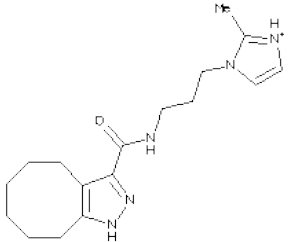
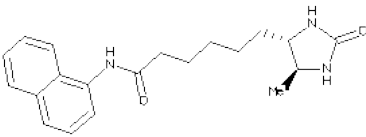
164 surviving ligands obtained from different clusters were further analyzed for their ADME properties, their molecular weight, percent human oral absorption values. Another important factor is the ability to cross the Blood Brain Barrier (BBB). Compounds with QLogBB values greater than 0.3 have been reported to pass the BBB and compounds with QLogBB values between 0.3 and -1 can still enter to Central Nervous System whereas ligands having QLogBB values less than -1 are not distributed to brain[57]. Consequently, compounds with QLogBB values less than -1 were eliminated and 69 ligands survived. Afterwards, top 15 ligands were selected for further studies based on their docking score, polar surface area (PSA), number of acidic and basic atoms, molecular weight and percent human oral absorption values. Proposed 15 ligands were redocked to active site of GBA using induced fit docking to allow side chain movement and ligand flexibility.

Table 4.7. Rigid docking and IFD results of hit 15 Ligands

Cluster	ZINC ID	XP GScore	IFD XP GScore	Total GScore
1	ZINC08099201	-8.732	-8.883	-17.614
1	ZINC09447905	-8.230	-8.405	-16.634
1	ZINC03683456	-8.200	-8.974	-17.174
1	ZINC13728463	-8.179	-8.025	-16.204
4	ZINC01726382	-10.600	-12.362	-22.962
4	ZINC04033947	-9.863	-9.531	-19.394
4	ZINC04316723	-8.974	-8.227	-17.200
4	ZINC04348040	-8.442	-10.364	-18.804
4	ZINC04743969	-8.443	-9.148	-17.590
4	ZINC03818225	-10.11	-10.217	-20.326
4	ZINC05323884	-9.978	-11.414	-21.392
4	ZINC04687140	-10.044	-9.390	-19.434
10	ZINC23428850	-9.428	-9.137	-18.564
10	ZINC23399033	-9.854	-8.052	-17.906
10	ZINC19288139	-8.617	-10.095	-18.712

6 ligands had poor IFD XP GScores than the rigid docking score because of entropy penalties. 5 ligands obtained from cluster 4 and cluster 10 were chosen for selectivity analysis. The druglikeness properties of these ligands were shown in Table 4.8.

Table 4.8. Druglikeness properties of top 5 ligands

ZINC ID	2D Structure	IFD XP Score	QLogPo/w	QLogBB	PHOA	LLE
ZINC01726382		-12.362	2.614	-0.625	84.6	5.17
ZINC05323884		-11.414	1.664	-1.801	73	5.65
ZINC04348040		-10.364	2.656	-1.395	82.3	3.54
ZINC19288139		-10.095	3.079	-0.918	95	2.87
ZINC04743969		-9.148	3.791	-1.611	94.3	2.39

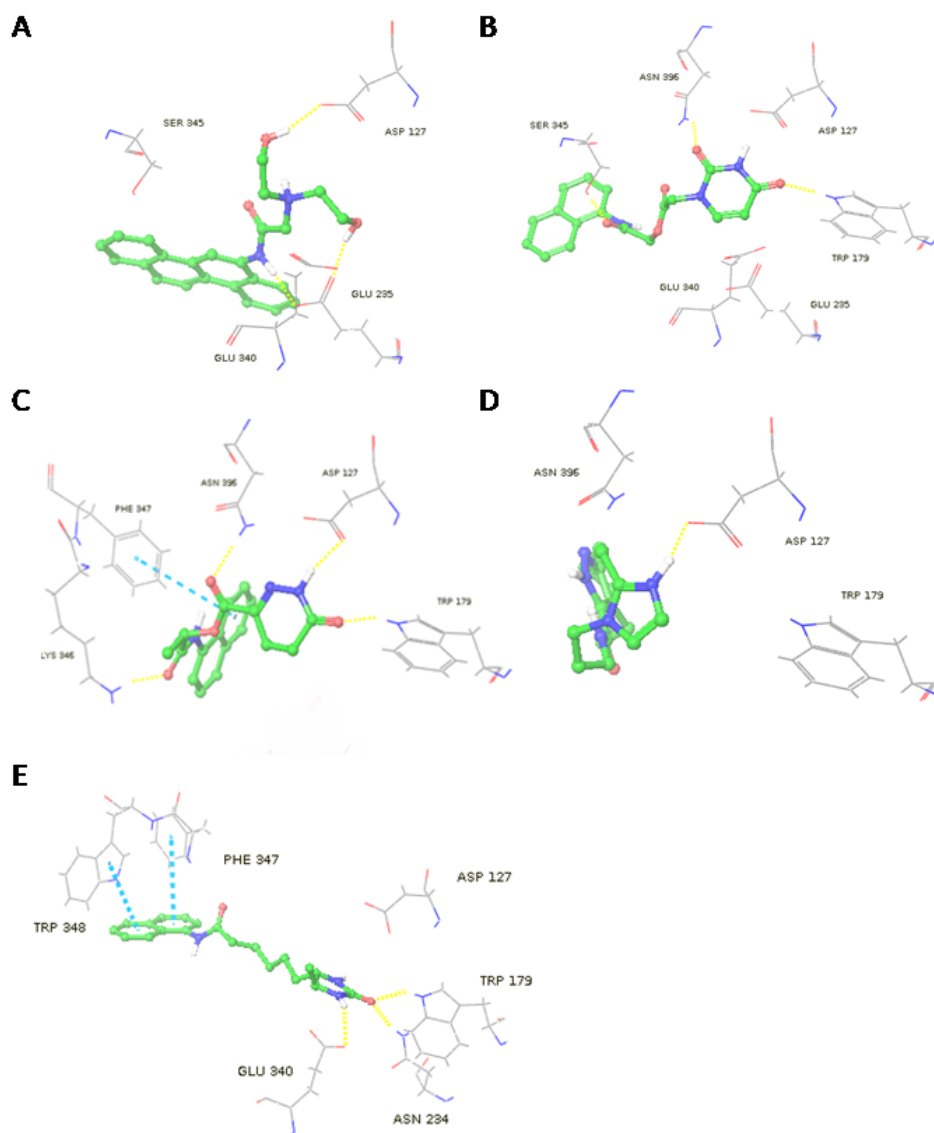


Figure 4.7. Interactions of best 5 compounds obtained by structure-based modeling (A: ZINC01726382, B: ZINC05323884, C: ZINC04348040, D: ZINC19288139, E: ZINC04743969)

4 of proposed 5 compounds modeling consists of two different aromatic ring groups connected by a long tail while best compound, Zinc01726382, have also four-membered ring group and a tail group connected to rings. All ligands had hydrogen bonds with active site residues either Asp 127 or Trp 179. Best ligand bound to GBA from its tail part instead of making interactions with ring groups while other 4 compounds bound to GBA from their rings. Moreover Zinc01726382 made three hydrogen bonds with Asp 127 and catalytic residue Glu 235 which favors ligand binding.

4.2. Ligand-Based Modeling

4.2.1. Ligand Based Pharmacophore Modeling

For GBA enzyme, 354 ligands with their experimentally determined IC₅₀ values were retrieved from ChEMBL[37]. The inhibition constant of a ligand is highly influenced by the experimental conditions such as pH and temperature. For ligands having more than one inhibition constant, the one with the lowest IC₅₀ value was kept in the list, and the others were deleted and consequently 182 unique compounds from different chemical families including iminosugars, isofagomine, pyrrolidine based iminosugars etc. remained. These molecules were clustered based on their structural similarity (Figure 4.8) via Hierarchical Clustering tool in Chemmine website [58].

Clustering indicated that these 182 molecules formed 2 different sets of ligands. One set included ligands with long tail attached to a six membered ring involving oxygen and nitrogen such as pyridine and phenylamine derivatives. The other group included ligands with more than one ring and very complex structures such as triazine derivatives. Pyrrolidine-based inhibitors have high potential to treat Gaucher disease at lower concentrations than known inhibitors such as isofagomine. Their alkyl chain group interacts with hydrophobic pocket while sugar moiety binds to active site residues involving Asp 127, Glu 325 and Glu 340 [24]. Consequently, pyrrolidine-based compounds from 182 ligands were found and utilized to generate pharmacophore hypothesis. These 30 pyrrolidine-based compounds and their pharmacophore sites were tabulated as shown in Table 4.9.

As displayed in Table 4.9, none of the compounds had charged group, so that four features which are hydrophobic group (H), hydrogen bond donor (D), hydrogen bond acceptor (A) and aromatic ring (R) were used for pharmacophore building. After hypotheses were constructed, they were scored and ranked with respect to their survival scores. Top 5 hypotheses were shown in Table 4.10.

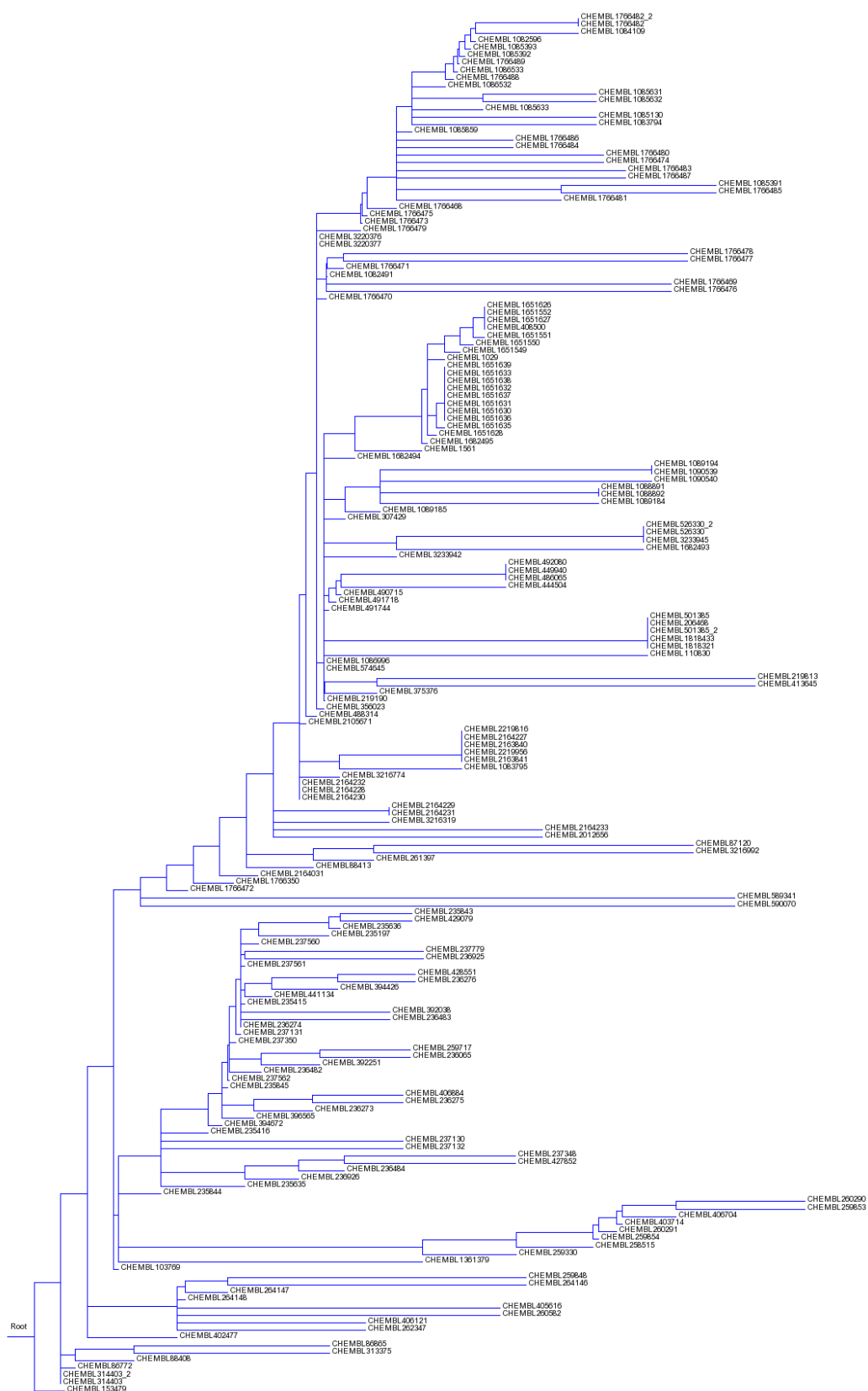


Figure 4.8. Canvas clustering results of 182 compounds

Table 4.9. Pharmacophore sites of known pyrrolidine-based GBA inhibitors.

CHEMBL ID	A	D	H	N	P	R	Total
CHEMBL429079	5	3	3	0	0	2	13
CHEMBL235197	4	2	4	0	0	2	12
CHEMBL237560	5	3	3	0	0	2	13
CHEMBL237779	5	3	3	0	0	2	13
CHEMBL236925	5	3	2	0	0	2	12
CHEMBL236276	5	3	2	0	0	2	12
CHEMBL441134	5	3	2	0	0	2	12
CHEMBL235415	5	3	2	0	0	2	12
CHEMBL392038	4	3	2	0	0	2	11
CHEMBL236483	5	4	1	0	0	2	12
CHEMBL236274	4	3	2	0	0	2	11
CHEMBL237131	6	3	3	0	0	2	14
CHEMBL237350	5	4	2	0	0	2	13
CHEMBL236065	4	3	2	0	0	2	11
CHEMBL392251	4	3	2	0	0	2	11
CHEMBL236482	5	3	2	0	0	2	12
CHEMBL237562	4	3	1	0	0	2	10
CHEMBL235845	5	3	4	0	0	2	14
CHEMBL236275	4	3	2	0	0	2	11
CHEMBL236273	4	3	2	0	0	2	11
CHEMBL396565	5	3	2	0	0	2	12
CHEMBL394672	6	3	3	0	0	2	14
CHEMBL237130	5	3	2	0	0	2	12
CHEMBL237132	5	3	3	0	0	2	13
CHEMBL237348	5	3	2	0	0	3	13
CHEMBL427852	5	4	1	0	0	3	13
CHEMBL236484	5	3	1	0	0	3	12
CHEMBL236926	5	3	1	0	0	3	12
CHEMBL235635	5	3	2	0	0	3	13
CHEMBL237561	5	3	3	0	0	2	13

Table 4.10. Scores of top 5 hypotheses constructed using 30 pyrrolidine-based inhibitors.

Hypothesis	Survival	Site	Vector	Volume	Selectivity
AAADR.94	3.896	1	1	0.9	1.097
AADDR.407	3.892	0.99	1	0.899	1.237
AAADD.73	3.877	0.99	0.999	0.887	0.792
AAADH.60	3.856	0.97	1	0.881	1.176
ADDHR.21	3.854	0.97	0.999	0.889	1.650

The quality of the hypothesis was determined according to its survival score which combines the site, vector and volume scores. The ranges of site, vector and volume scores were between 0-1 while one (1) implied the best alignment of ligands on the hypothesis. In addition, selectivity defined the fraction of any random drug-like molecule to match the hypothesis. As shown in Table 4.10, survival scores of top 5 hypotheses were significantly close to each other differing from 3.896 to 3.854, and selectivity values did not remark a single hypothesis that represents the best hypothesis for these set of ligands. Therefore, quantitative structure-activity relationship (QSAR) studies were performed in order to determine reliable hypothesis and to validate it (section 4.2.2).

Table 4.11. Fitness of each compound for different hypotheses and their pIC50 values

ChEMBL ID	AAADR.94	AADDR.407	AAADD.73	AAADH.60	ADDHR.21	pIC50
CHEMBL429079	3.00	3.00	3.00	3.00	2.95	3.398
CHEMBL235197	2.98	2.98	2.98	2.9	2.89	2.109
CHEMBL237560	2.93	2.92	2.98	2.87	2.86	3.018
CHEMBL237779	2.93	2.93	2.93	2.89	2.82	2.983
CHEMBL236925	2.91	2.91	2.91	2.87	2.84	3.481
CHEMBL236276	2.98	2.98	2.98	2.94	3.00	2.967
CHEMBL441134	2.91	2.9	2.96	2.85	2.87	3.113
CHEMBL235415	2.95	2.95	2.94	2.92	2.95	1.684
CHEMBL392038	2.95	2.95	2.89	2.95	2.94	2.405
CHEMBL236483	2.95	2.95	2.8	2.87	2.95	2.699
CHEMBL236274	2.95	2.95	2.88	2.95	2.95	1.538
CHEMBL237131	2.98	2.98	2.98	2.96	2.92	2.815
CHEMBL237350	2.97	2.97	2.79	2.88	2.92	3.136
CHEMBL236065	2.93	2.93	2.94	2.83	2.91	2.366
CHEMBL392251	2.9	2.9	2.94	2.83	2.88	2.174
CHEMBL236482	2.92	2.92	2.92	2.88	2.88	2.084
CHEMBL237562	2.82	2.82	2.73	2.91	2.82	1.783
CHEMBL235845	2.92	2.92	2.92	2.74	2.78	2.111
CHEMBL236275	2.81	2.81	2.72	2.89	2.81	2.198
CHEMBL236273	2.9	2.9	2.91	2.9	2.91	1.998
CHEMBL396565	2.79	2.78	2.88	2.87	2.79	2.055
CHEMBL394672	2.86	2.86	2.94	2.81	2.84	2.309
CHEMBL237130	2.9	2.9	2.91	2.87	2.85	2.476
CHEMBL237132	2.93	2.93	2.91	2.85	2.81	2.577
CHEMBL237348	2.74	2.68	2.86	2.77	2.67	2.653
CHEMBL427852	2.83	2.83	2.58	2.69	2.77	3.432

Table 4.11. Fitness of each compound for different hypotheses and their pIC50 values
(cont.)

ChEMBL ID	AAADR.94	AADDR.407	AAADD.73	AAADH.60	ADDHR.21	pIC50
CHEMBL236484	2.81	2.81	2.76	2.74	2.8	1.664
CHEMBL236926	2.8	2.79	2.8	2.69	2.79	1.503
CHEMBL235635	2.78	2.77	2.78	2.77	2.74	2.167
CHEMBL237561	2.96	2.95	2.96	2.95	2.88	2.463
Average Combined Fitness	7.10	7.09	7.03	7.00	6.99	-

Fitness score of each compound for the top 5 hypotheses and their pIC50 values were tabulated in Table 4.11. Average combined fitness values were also calculated by multiplying pIC50 value of each ligand with fitness score and taking the arithmetic average. Average combined fitness score of each hypothesis were also similar to each other, eventually QSAR studies were performed using Phase module.

4.2.2. Quantitative Structure Activity Relationship (QSAR) Model

First, all 182 ligands obtained from ChEMBL website were used to determine pharmacophore hypothesis. The results of several runs were tabulated in Table 4.12. Using different test sets, various pharmacophore hypotheses were obtained. The accuracy of these pharmacophore hypotheses was determined by Q^2 and R^2 values. Hypotheses with Q^2 values higher than 0.6 were accepted as reliable hypotheses.

Table 4.12. QSAR results for 182 Ligands

Hypothesis	Q^2	R^2
ADH.10	0.2355	0.5257
ADH.16	0.2572	0.52
ADD.3	0.1772	0.4674
ADD.13	0.1772	0.4675
ADH.12	0.1955	0.4448
AAH.10	0.122	0.3791

When all 182 compounds were used, the Q^2 values were all smaller than 0.6. Second set consisted of 111 ligands (Figure A.1) with acceptable structural similarities according to Chemmine clustering results, and the results of this set were given in Table 4.13.

Table 4.13. QSAR results for 111 Ligands

Hypothesis	Q^2	R^2
AADDP.482	0.3469	0.6025
AADDP.379		
ADDDP.228	0.3014	0.5768
ADDDP.681		
AADP.62	0.2522	0.6005
AADP.123		
AAADP.320	0.2669	0.544
AAADP.245		
AADDP.1300	0.2613	0.5523
ADDDP.1597	0.2574	0.5194
AAADP.91	0.1538	0.4023
ADDD.211	0.3483	0.5991
ADDD.380		
ADDD.201		
ADDD.389		
AADDP.24	0.1113	0.374

Although Q^2 values of the second set were higher than the Q^2 values of the first set, the obtained hypothesis was not sufficient to be used for filtering because Q^2 values were still below 0.6. The third set included structurally more similar ligands and it consisted of 49 ligands (Figure A.1). The results of this third set were tabulated in Table 4.14. However, the hypotheses obtained from the third set were still not reliable to be used because of the low Q^2 values.

Table 4.14. QSAR results for 49 ligands

Hypothesis	Q^2	R^2
DDDP.37	0.3185	0.5668
AADHP.250	0.2490	0.6068
DDDP.285	0.2072	0.4889
AADP.45	0.2336	0.5464
ADDDP1383	0.2978	0.5530
ADDDP.1377	0.2638	0.5269

Pyrollidine-based iminosugars have high potential to treat Gaucher disease[24]. On this basis, ligands with five-membered ring in their structures were taken into consideration and QSAR studies were carried out with these molecules. These compounds were tabulated in Table 4.15.

Table 4.15. 30 Pyrolidine based ligands chosen for QSAR study.

Number of Ligand	Ligand ID	IC ₅₀ (nM)	pIC ₅₀
1	ChEMBL429079	400	3.398
2	ChEMBL235197	7770	2.109
3	ChEMBL237560	960	3.018
4	ChEMBL237779	1040	2.983
5	ChEMBL236925	330	3.481
6	ChEMBL236276	1080	2.967
7	ChEMBL441134	770	3.113
8	ChEMBL235415	20700	1.684
9	ChEMBL392038	3940	2.405
10	ChEMBL236483	2000	2.699
11	ChEMBL236274	28970	1.538
12	ChEMBL237131	1530	2.815
13	ChEMBL237350	730	3.136
14	ChEMBL236065	4310	2.366
15	ChEMBL392251	6690	2.174
16	ChEMBL236482	8250	2.084
17	ChEMBL237562	16470	1.783
18	ChEMBL235845	7750	2.111
19	ChEMBL236275	6340	2.198
20	ChEMBL236273	10040	1.998
21	ChEMBL396565	8800	2.055
22	ChEMBL394672	4900	2.309
23	ChEMBL237130	3340	2.476
24	ChEMBL237132	2650	2.577
25	ChEMBL237348	2220	2.653
26	ChEMBL427852	370	3.432
27	ChEMBL235484	21670	1.664
28	ChEMBL237926	31400	1.503
29	ChEMBL235635	6800	2.167
30	ChEMBL237561	3440	2.463

Different conformations of these 30 ligands were generated using ConfGen for QSAR. Afterwards, IC₅₀ values of these 30 compounds were converted to pIC₅₀ values by taking the logarithm of IC₅₀ values and multiplying them with -1. Training set consisted of 75% of total ligands while other compounds formed the test set. Hence, six of the 30 ligands were selected randomly to form the test set and the remaining ligands were set as the training set. Using 6 different training sets, 14 pharmacophore hypotheses were obtained. 26 different QSAR runs were performed for accuracy of the results, and the hypotheses with Q² values higher than 0.6 were shown in Table 4.16.

Table 4.16. QSAR results of 30 ligands

Set	Hypothesis	Q ²	R ²
1	AAADD.8287	0.6823	0.951
2	AAADD.429	0.7026	0.8877
	ADDHH.1280	0.6428	0.8850
	AAADR.1332	0.6793	0.8739
	AAAHR.614	0.6563	0.8304
3	AADDR.2735	0.6351	0.8461
	AAAHR.600	0.6761	0.8893
4	AADHH.177	0.6607	0.8552
5	ADHHR.1932	0.7641	0.8941
	ADDHH.559	0.6084	0.7943
	AAAHR.680	0.6571	0.8200
6	AAAHR.519	0.7561	0.9685
	ADDHR.125	0.7659	0.9091
	AAADR.305	0.7468	0.8752

As the number of features increases, the selectivity of hypothesis to find similar compounds increases. Hypothesis with 5 features enabled to find more specific ligands to GBA. In this context, ADDHR.125 and AAADR.305 hypotheses were used in order to retrieve compounds from ZINC Database involving 500000 molecules. Despite of high Q² and R² values, AAAHR.519 and ADHHR. 1932 were eliminated because the distances between pharmacophore sites were too large.

Table 4.17. QSAR results for top two hypotheses

Set	Parameters	ADDHR.125	AAADR.305
Training	SD	0.4642	0.4893
	R ²	0.9091	0.8752
	F	14.3	10.7
Test	RMSE	0.2345	0.2467
	Q ²	0.7659	0.7468
	Pearson-R	0.9091	0.8752

In Table 4.17, SD value shows the strength of the fit to training set and in comparison of top two hypothesis the lowest standard deviation value belonged to ADDHR.125 hypothesis. Besides R² value of hypothesis indicates how good the structure-activity relationship between the model and training set. The R² value of ADDHR.125 hypothesis was higher than that of AAADR.305 implying that this model defined structure-activity relationship better than AAADR.305 hypothesis. Moreover, large value

of F implies a more important regression. The comparison of all these parameters shows that ADDHR.125 hypothesis fits slightly to training set better.

A 3-featured hypothesis was also searched using the same 30 ligands in order to compare with the structure-based modeling. Similar to previous QSAR studies, the hypotheses were formed with 3-features, instead of 5-features. As a result, HRH hypothesis was obtained with Q^2 value of 0.7294 and R^2 value of 0.8604. Compared to hypothesis obtained by structure-based modeling, the distances between features in ligand-based modeling were found to be greater. In structure-based modeling, little fragments were used to determine the hypothesis instead of large ligands. The hypothesis was shown in Figure 4.9.

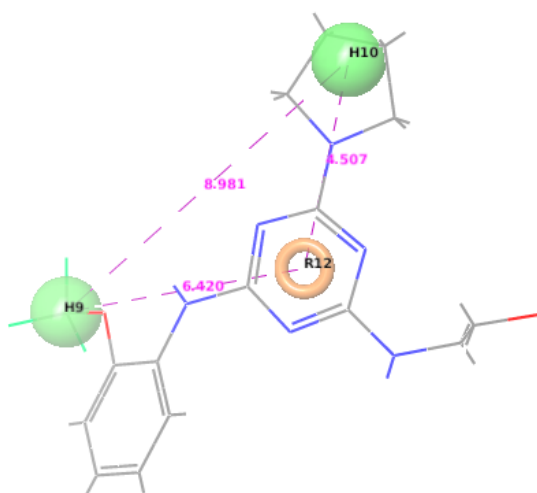


Figure 4.9. Site point distances of HRH ligand-based hypothesis
(Ligand displayed in Figure 4.9 is Chembl236276)

4.2.3. Pharmacophore Based Virtual Screening and Molecular Docking

Using the hypothesis of AAADR.305 and ADDHR.125 for filtering of ZINC database, top 10000 new ligands were considered with respect to their fitness scores. Different conformations of these molecules were generated in order to find the best conformation of the ligand for its binding to the receptor. Then, these molecules were docked to active site of GBA in SP docking and best 1000 molecules in terms of docking energy score were docked to the same site of the enzyme in XP mode.

4.2.3.1. Virtual Screening and Molecular Docking Results for AAADR.305 Hypothesis

Top 1000 molecules obtained by XP docking were eliminated with respect to their XP Gscores and strain energy values. -6.5 kcal/mol and 4 kcal/mol were accepted as docking score and strain energy cut off limits similar to structure-based studies. As a result, 218 compounds were remained for further substructure search studies.

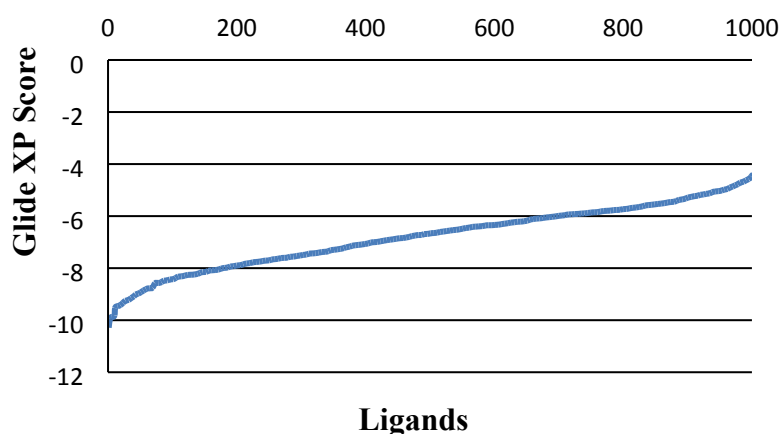


Figure 4.10. Glide XP Gscore for 1000 molecules obtained from AAADR.305 hypothesis

BEI-SEI values of each compound were calculated using docking results, and plotted as shown in Figure 4.11. BEI values of compounds were very high differing between 20 and 60. SEI values were lower than BEI values. Figure 4.11 showed that compounds were very polar which satisfied three aromatic and one ring features of the hypothesis.



Figure 4.11. BEI-SEI values of 218 compounds obtained from AAADR.305 hypothesis

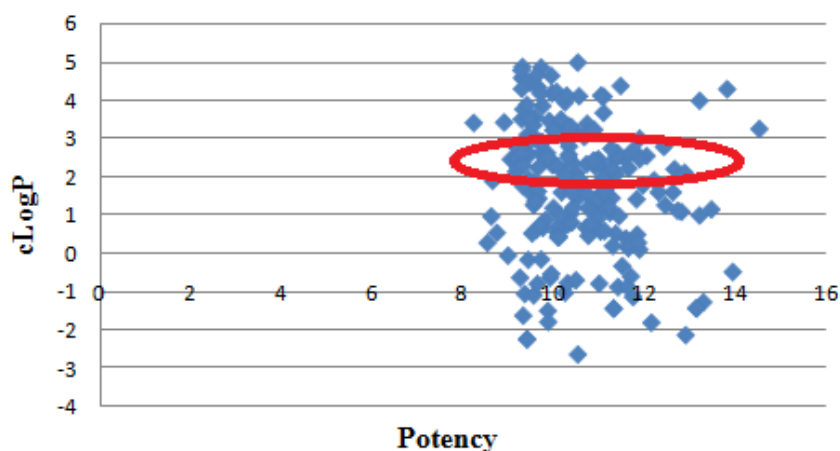


Figure 4.12. cLogP vs Potency of 218 molecules obtained from AAADR.305 hypothesis

Figure 4.12 showed that octanol-water partition coefficient of most compounds were greater than zero implying that they are hydrophobic, and about 14 molecules were hydrophilic. Octanol-water partition coefficient of oral drugs need to be between 2.5-3, in the study with AAADR.305 hypothesis, there were 22 compounds in that range. In addition, potency values of most compounds were greater than the maximum potency that is set for oral drugs. With regards to LLE values shown in Figure 4.13, they were greater than 4 for all compounds; because of that reason compounds do not require any optimization and they can be used directly in the treatment of Gaucher disease.

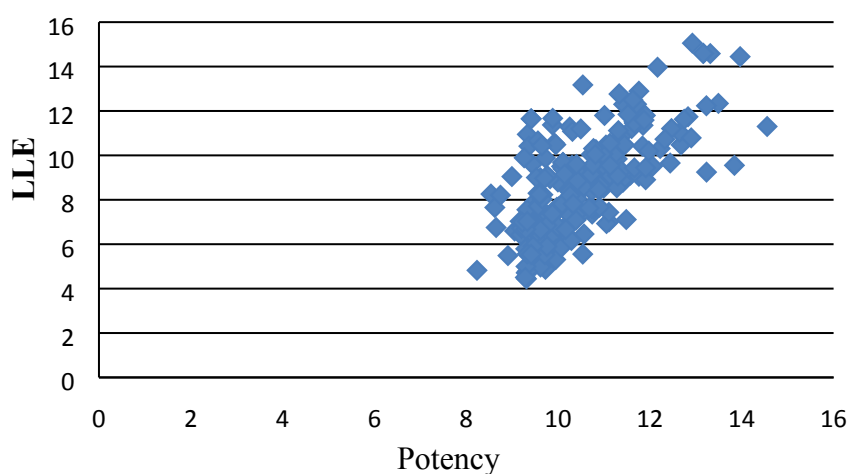


Figure 4.13. LLE vs Potency values of 218 molecules obtained from AAADR.305 hypothesis

4.2.3.2. Virtual Screening and Molecular Docking Results for ADDHR.125 Hypothesis

Similar to AAADR.305 hypothesis, top 1000 molecules obtained from ADDHR.125 hypothesis were further examined for post-docking studies. XP GScores of all molecules were drawn as shown in Figure 4.14. The XP GScore cut off was accepted as -6.5 kcal/mol in order to be consistent with structure-based modeling. Furthermore, compounds were filtered with respect to their strain energy values using strain energy cut off, 4 kcal/mol, as a result 241 compounds survived for ADDHR.125 hypotheses.

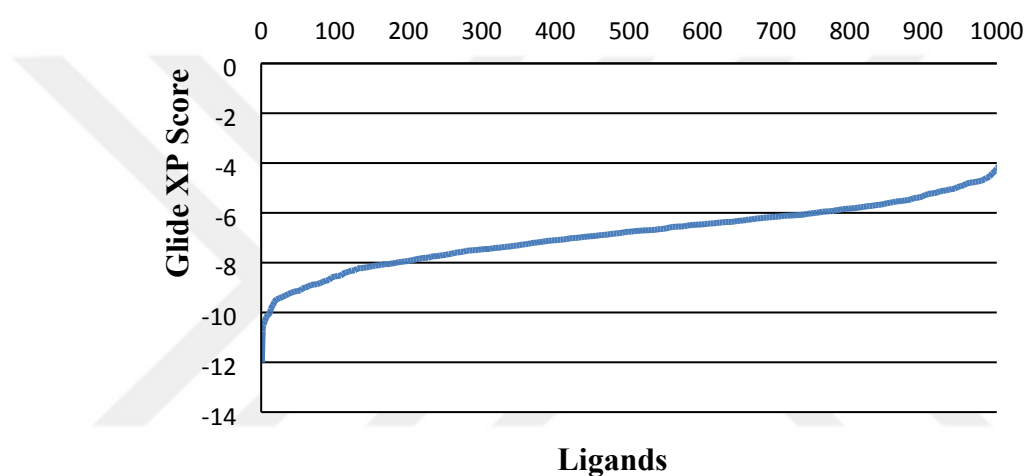


Figure 4.14. Glide XP Gscore for 1000 molecules obtained from ADDHR.125 hypothesis

BEI and SEI values of surviving molecules were calculated using docking results and graphed as displayed in Figure 4.15. The figure implied that all compounds had higher BEI values than SEI values, and they were highly polar as expected, because the hypothesis involves hydrogen acceptor, hydrogen donor, hydrophobic and ring features.

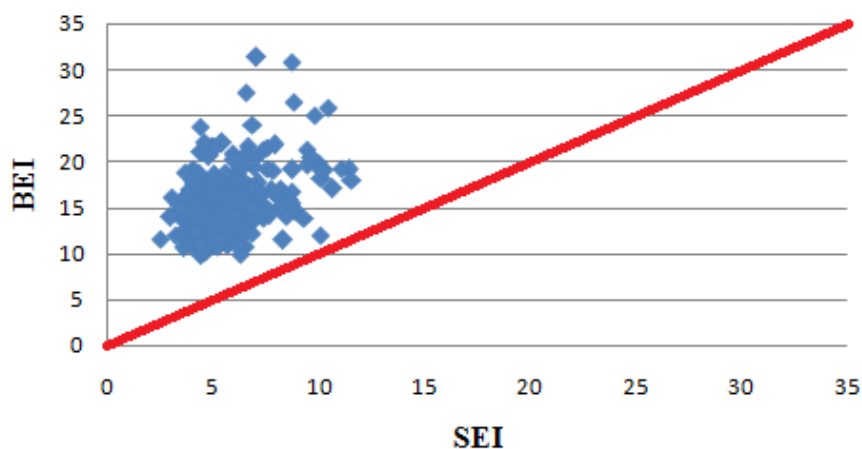


Figure 4.15. BEI-SEI value for surviving molecules from ADDHR.125 hypothesis

In Figure 4.16, potency versus octanol-water partition coefficients of surviving molecules were graphed. For oral drugs, octanol-water partition coefficient is expected to be between 2.5-3. There are 40 molecules, that are in the range of these values, found in ADDHR.125 hypothesis. In addition, all compounds except one had positive octanol-water partition coefficient showing that they are hydrophobic drug candidates and mainly distribute in hydrophobic areas.

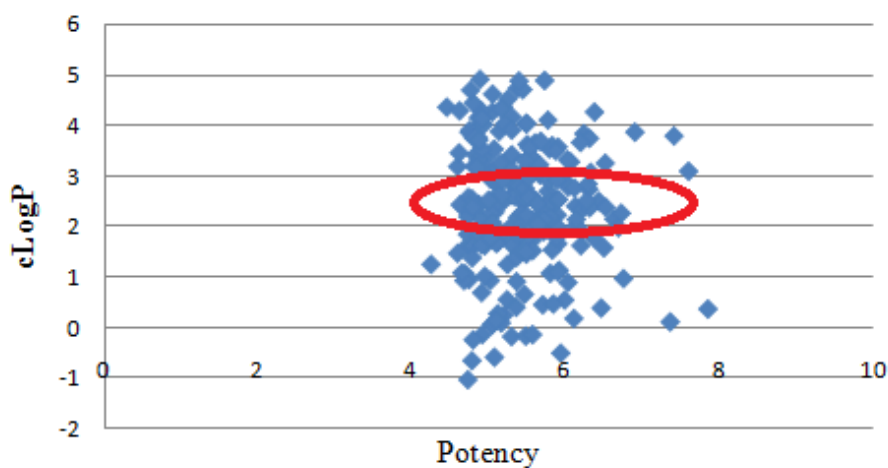


Figure 4.16. cLogP vs Potency values of 241 molecules obtained from ADDHR.125 hypothesis

LLE vs potency graph showed that most of compounds had LLE value below 5, about 20 compounds had LLE value greater than 5. For oral drugs, LLE value must be

greater than 5, so that most of these compounds require optimization by adding side chains or ions. With regards to potency, their potency values were at the optimum value which is between 0 and 10. In comparison with compounds of AAADR.305 hypothesis, the compounds obtained by ADDHR.125 hypothesis had lower potency and LLE values.

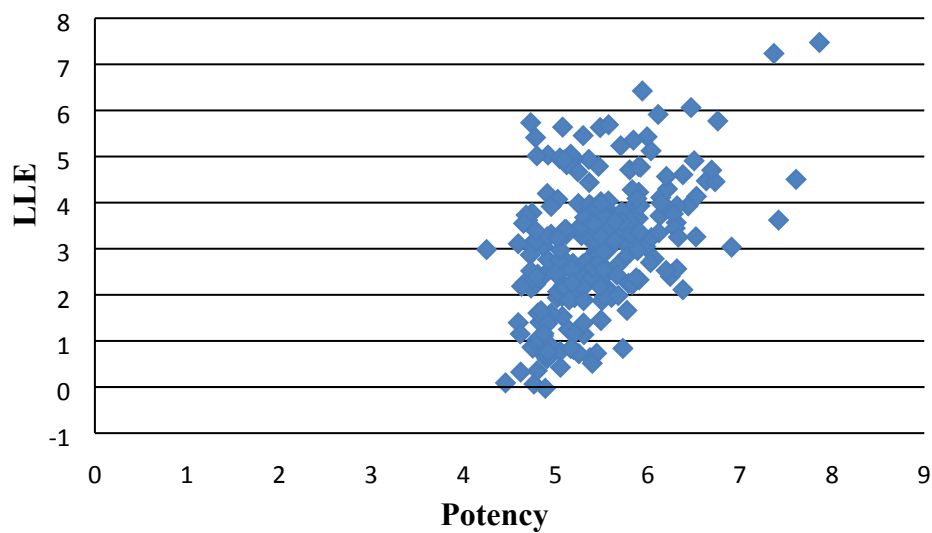


Figure 4.17. LLE vs potency values of 241 molecules obtained from ADDHR.125 hypothesis

4.2.4. Rigid Docking of Compounds from Ligand Based Modeling

Experimentally known ligands and already used commercial ligands were utilized to find novel compounds for GBA. Their XP GScore values and drug properties were used as basis for filtration and selection of ligands.

4.2.4.1. Rigid Docking of Commercial Compounds. Fingolimod targets sphingosine 1-phosphate receptor. Miglustat is already known inhibitor of GBA and spissulosine is activator of phosphatidylinositol 3-kinase. In addition GW4869 is inhibitor of N-SMase (neutral sphingomyelinase) and U18666A is cholesterol synthesis inhibitor. They are docked to GBA active site in order to understand which types of compounds have tendency to bind GBA. The docking result and druglikeness properties of these 5 ligands were tabulated in Table 4.18.

Table 4.18. XP Gscore and druglikeness properties of 5 commercial inhibitors of GBA.

Commercial Name	Commercial Code	Rigid Docking XP GScore	QLogPo/w	QLogBB	Percent Human Oral Absorption
Fingolimod	FTY720	-9.146	2.546	-0.733	87
Miglustat	OGT918	-8.014	-1.145	-0.896	53.23
-	U18666A	-6.463	4.346	0.158	100
Spisulosine	ES-285	-5.421	4.155	-0.995	100
-	GW4869	-5.395	4.764	-1.932	87.2

The interactions of these 5 commercial compounds with GBA were also analyzed with Ligand Interaction module and tabulated as shown in Table 4.19. The interaction maps were shown in Figure 4.18. Table 4.19 displayed that the maximum number of hydrogen bonds with GBA belonged to Miglustat whereas the best XP Score belonged to Fingolimod. It implied that XP Score does not depend on the number of hydrogen bonds but it depends on the strength of the bond between the atoms. GW4869 had the least XP Score and it made interaction with only Ser 345. Consequently, because of having XP Score less than -8 kcal/mol, U19666A, ES-285 and GW4869 were not suitable to be used in GBA studies. Moreover, Fingolimod consists of different ring groups and it can be concluded that potential GBA chaperones should involve aromatic rings.

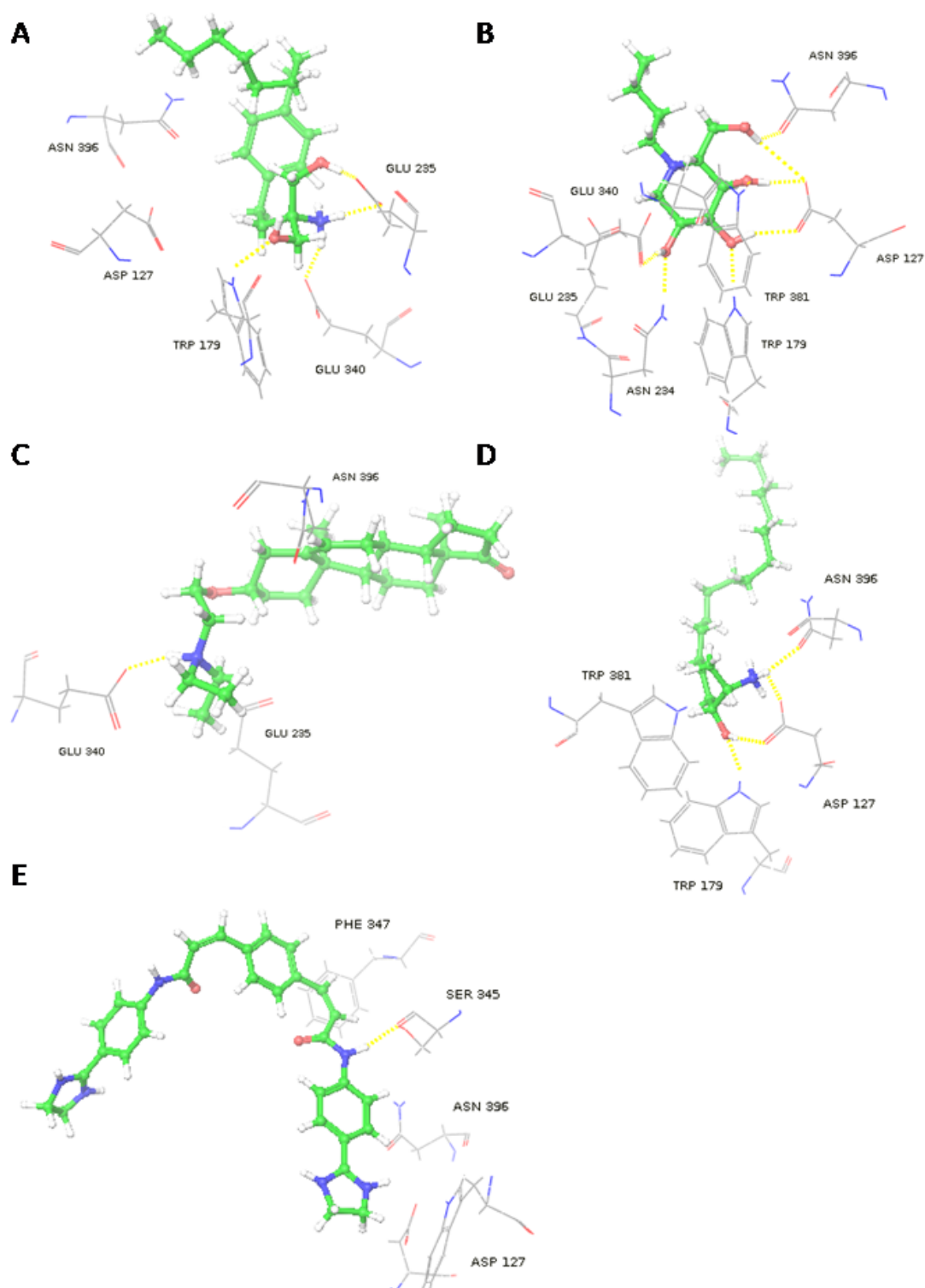


Figure 4.18. Interaction maps of 5 commercial compounds.
(A:FTY 720 B: OGT918 C: U18666A D: ES-285 E: GW4869)

Table 4.19. Interactions of commercial 5 ligands with the active site of GBA.

Compound	Catalytic Residues		Active Site Residues				Hydrophobic Pocket Residues			Other Residues			
	Glu 235	Glu 340	Asn 234	Trp 179	Asp 127	Trp 381	Asn 396	Ser 345	Tyr 313	Phe 347	Tyr 244	Lys 346	Phe 246
FTY 720	H(s) (2)	H(s)		H(s)									
OGT918		H(s)	H(s)	H(s)	H(s) (3)	H(s)	H(s)						
U18666A		H(s)											
ES-285				H(s)	H(s) (2)		H(s)						
GW4869								H(b)					

4.2.4.2. Rigid Docking of Experimentally Known Ligands. Imimosugars, azosugars and other experimentally known 182 inhibitors of GBA from ChEMBL were docked to the active site of the protein in XP mode. XP GScores of 182 compounds were between -11.602 and -3.156, while only 70 of these 182 ligands had XP GScores less than -8.0. In addition, top 50 ligands had blood brain barrier coefficients less than -1 showing that these ligands were not so effective in passing the blood brain barrier, and they are not suitable for the treatment of Type 2 and Type 3 Gaucher Disease.

Interactions of best 10 compounds and well-known ligands (found in these 182 ligands) with GBA were further analyzed in order to determine the binding types of top molecules to GBA active site (Table 4.20). It was found that all well-known ligands bound to the active site of GBA, and most of them interacted with almost all active site residues involving Trp 179, Asn 234, Asp 127 and Trp 381. Besides, 9 of top 10 compounds consisted of a six-membered ring with hydroxyl ions attached to carbon-atoms (named as 1,2,3,4,5-Cyclohexanepentol), and a tail group involving an aromatic ring. The binding type of these top ligands were similar to well-known ligands because of their six-membered ring with hydroxyl ions. That part of the ligand enabled to make hydrogen bonds with active site residues similar to miglustat and isofagomine.

Table 4.20. Interactions of best 10 experimentally known ligands and well-known GBA inhibitors with GBA.

ChEMBL ID	Catalytic Residues		Active Site Residues				Hydrophobic Pocket Residues			Other Residues						Rigid XP Score
	Glu 235	Glu 340	Asn 234	Trp 179	Asp 127	Trp 381	Asn 396	Ser 345	Tyr 313	Phe 347	Tyr 244	Lys 346	Phe 246	Glu 349	Ser 242	
1029 (Miglustat)		H(s)	H(s)	H(s)	H(s) (2)	H(s)	H(s)									-8.71
408500 (NN-DNJ)		H(s)	H(s)	H(s)	H(s) (2)	H(s)	H(s)									-8.39
501385 (Isofagomine)			H(s)		H(s)	H(s)	H(s)									-5.50
153479 (Ambroxol)				H(s)	H(s)											-5.12
1083795 (Aminocyclitol)	H(s)	H(s)	H(s)	H(s)	H(s) (2)	H(s)	H(s)									-9.05
589341			H (s)	H (s)	H (s) (2)	H (s)	H (s)			π - π	H (s)					-11.60
1766488		H (s)	H (s)	H (s)	H (s) (2)	H (s)	H (s)								H (s)	-10.83
1766477	H (s)		H (s)	H (s)	H (s) (2)	H (s)	H (s)			π - π	H (s)					-10.82

Table 4.20. Interactions of best 10 experimentally known ligands and well-known GBA inhibitors with GBA (cont.)

ChEMBL ID	Catalytic Residues		Active Site Residues				Hydrophobic Pocket Residues			Other Residues						Rigid XP Score
	Glu 235	Glu 340	Asn 234	Trp 179	Asp 127	Trp 381	Asn 396	Ser 345	Tyr 313	Phe 347	Tyr 244	Lys 346	Phe 246	Glu 349	Ser 242	
1766469		H (s)	H (s)	H (s)	H (s) (2)	H (s)	H (s)			π - π (2)	π - π					-10.44
1766350	H(s)	H (s)	H (s)	H (s)	H (s) (2)	H (s)	π - π			H (s)						-10.43
1766479	H (s)	H (s)	H (s)	H (s)	H (s) (2)	H (s)				H (s)						-10.42
1766476	H (s)	H (s)	H (s)	H (s)	H (s) (2)	H (s)	π - π (2)			H (s)	π - π					-10.39
1766471		H (s)	H (s)	H (s)	H (s) (2)	H (s)	π - π			H (s)						-10.20
1766478	H (s)		H (s)	H (s)	H (s) (2)	H (s)	π - π				π - π (2)					-10.16

RMSD value between NN-DNJ bound to GBA active site in crystal (PDB ID: 2V3E) and NN-DNJ redocked to GBA active site using Glide was calculated in order to verify the accuracy of docking. The RMSD value between NN-DNJ in crystal structure and docked NN-DNJ was 1.81 Å (less than 2Å) showing that NN-DNJ was docked to active site properly (Figure 4.19).

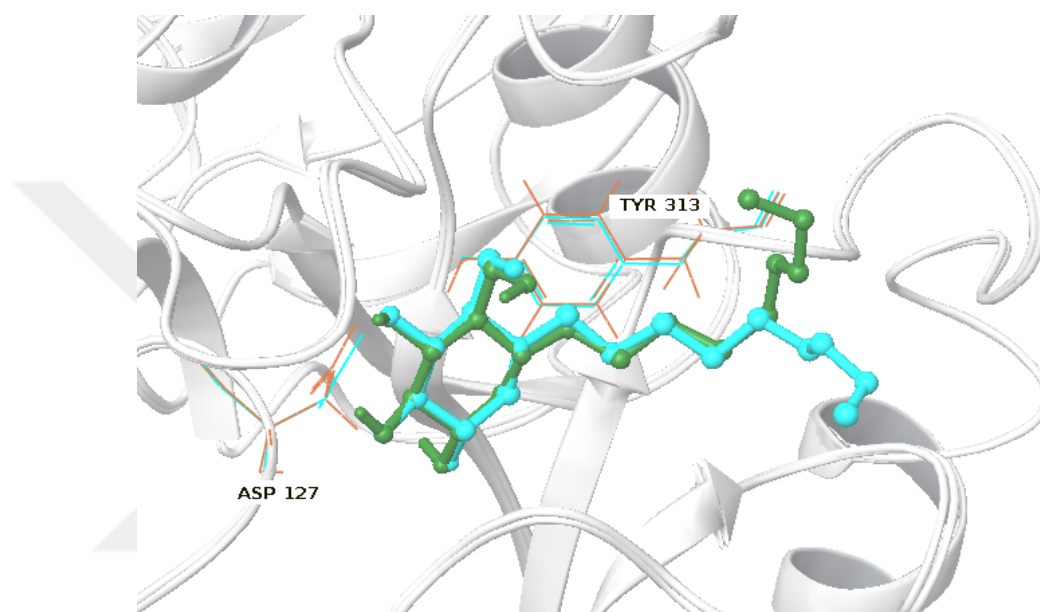


Figure 4.19. Alignment of NN-DNJ in crystal structure and docked NN-DNJ. (NN-DNJ in crystal structure = Turquoise, Docked NN-DNJ:Green)

4.2.5. Substructure Search for Surviving Ligands of Ligand-Based Pharmacophore Hypotheses

Surviving ligands for each hypothesis were clustered using Canvas with respect to their structural similarities. Results of each hypothesis were presented in Tables 4.21-4.24.

4.2.5.1. Substructure Search for Compounds of AAADR.305 Hypothesis. Surviving ligands from AAADR.305 hypothesis were separated with respect to their structural similarities. As a result, 23 different clusters occurred. Nevertheless, the numbers of ligands in 17 clusters were too small in order to use them for substructure search.

Table 4.21. Cluster results for AAADR.305 hypothesis

Number of Cluster	Number of Ligands in Cluster
1	1
2	2
3	1
4	1
5	1
6	18
7	23
8	1
9	4
10	1
11	1
12	30
13	11
14	13
15	3
16	51
17	21
18	14
19	12
20	1
21	1
22	2
23	5

Substructure search was done with 5 clusters containing more than 18 ligands. Different scaffolds were found from the substructure search of each cluster. For one cluster, Canvas gives different scaffolds according to the number of ligands used in substructure search. When more ligands were used in substructure search, more common scaffolds such as a single benzene ring or three carbon atoms were obtained. Therefore, the number of ligands was declined in order to obtain more specific scaffolds. Number of ligands included in substructure search, common substructure of these ligands and 2D views of these substructures were tabulated in Table 4.22.

Table 4.22. Substructure results of clusters obtained from AAADR.305 hypothesis

Cluster	Number of Ligands Included	Common Substructure	2D View of Structure
6	4	<chem>c1cccc(c1n)ncC(O)C(O)C(O)CO</chem>	
7	15	<chem>cncccnC1CC(O)C(O1)CO</chem>	
12	20	<chem>CSc1ncccc(n1)=O</chem>	
16	18	<chem>Oc1cc(C)ccc1O</chem>	
17	9	<chem>O=S(=O)c1ccccc1</chem>	

4.2.5.2. Substructure Search for Compounds of ADDHR.125. Similar to AAADR.305 hypothesis, surviving ligands from the second hypothesis were clustered according to their structural similarities using CANVAS module. As a result, each cluster had different numbers of members, and most of ligands were in cluster 9. Similar to previous substructure search study for AAADR.305 hypothesis, 11 clusters contained small number of the ligands and they were not suitable to do substructure search.

Table 4.23. Cluster results for ADDHR.125 hypothesis

Number of Cluster	Number of Ligand in Cluster
1	1
2	4

Table 4.23. Cluster results for ADDHR.125 hypothesis (cont.)

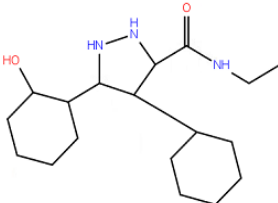
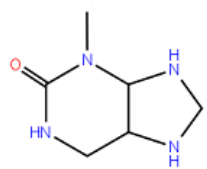
Number of Cluster	Number of Ligand in Cluster
3	1
4	2
5	1
6	6
7	19
8	9
9	143
10	9
11	20
12	2
13	19
14	1
15	4

If large numbers of ligands in a cluster were used for substructure search, general scaffolds were obtained such as three carbon atoms or a single benzene ring. In order to obtain more specific scaffolds, number of ligands used for scaffold search was decreased. The number of ligands used for substructure search in each cluster and scaffold of these ligands were given in Table 4.24.

Table 4.24. Substructure results of clusters obtained from ADDHR.125 hypothesis

Cluster	Number of Ligands Included	Common Substructure	2D View of Structure
7	14	<chem>NCNc(s1)c(C=O)c(c12)CC[N+](C2)C</chem>	
9	15	<chem>c1ccccc1C2NC(=O)NC=C2C=O</chem>	

Table 4.24. Substructure results of clusters obtained from ADDHR.125 hypothesis (cont.)

Cluster	Number of Ligands Included	Common Substructure	2D View of Structure
11	11	<chem>c1ccccc1C(N(C2=O)CC)c(c23)c(nn3)-c4ccccc4O</chem>	
13	15	<chem>c1nc(=O)n(C)c(c12)ncn2</chem>	

Consequently, 2D structures of each scaffold were generated in ZINC website involving 23 million compounds, and the database of this website was searched for these scaffolds in order to retrieve new compounds without any commercial availability concern.

4.2.5.3. Rigid Docking of Ligands obtained by Substructure Search of Cluster's Scaffold for ADDHR.125. Compounds retrieved from ZINC Website after structural search were docked with HTVS, SP and XP mode respectively. Top 1000 compounds obtained from XP mode were filtered using Glide Score cut off (-6.5 kcal/mol) and Strain Energy cut off (4 kcal/mol). Average BEI-SEI values of each cluster were calculated in order to determine the most druglike cluster, and their BEI and SEI values were plotted as shown in Figure 4.20.

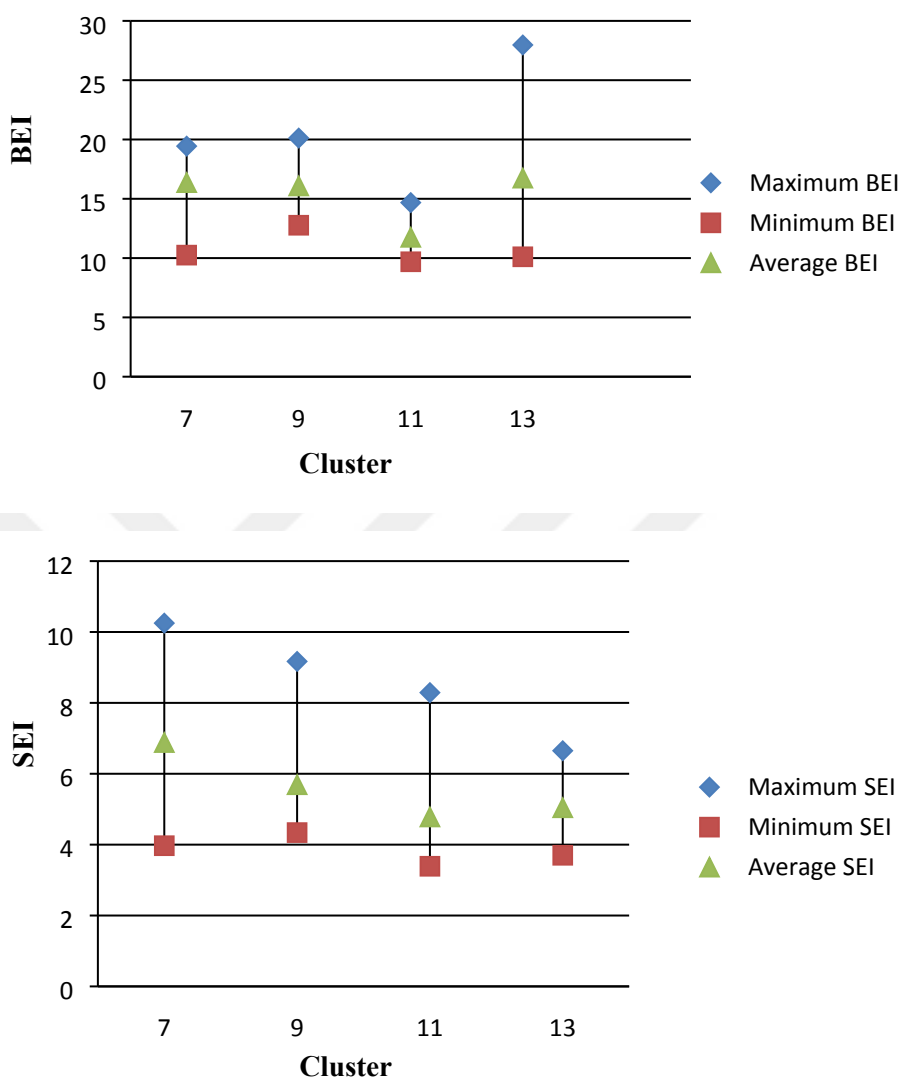


Figure 4.20. Average BEI and SEI values of clusters of hypothesis ADDHR.125

According to Figure 4.20, BEI values of compounds in clusters 7, 9 and 13 were higher than those in cluster 11 suggesting that ligands in clusters 7,9 and 13 were more druglike. The XP GScore and strain energy calculations results were tabulated in Tables 4.25-4.27 for these ligands.

Table 4.25. ZINC ID, XP GScore and strain values of ligands in cluster 7

ZINC ID	XP GScore	Strain Energy (kcal/mol)	ZINC ID	XP GScore	Strain Energy (kcal/mol)
ZINC12469416	-8.664	0.327	ZINC04838663	-8.116	1.643
ZINC00551825	-8.562	0.453	ZINC20135472	-7.956	1.524
ZINC18324763	-8.530	2.638	ZINC18324765	-7.929	1.153
ZINC02488838	-8.361	0.817	ZINC04172301	-7.883	0.379
ZINC20135476	-8.320	0.567	ZINC20135472	-7.787	0.651
ZINC33403498	-8.293	0.241	ZINC20064461	-7.507	0.261
ZINC20135476	-8.241	0.366	ZINC04891189	-7.095	0.218
ZINC20135476	-8.241	0.366	ZINC18324765	-6.889	3.037
ZINC20100881	-8.130	0.165			

Table 4.26. ZINC ID, XP GScore and strain values of ligands in cluster 9

ZINC ID	XP GScore	Strain Energy (kcal/mol)	ZINC ID	XP GScore	Strain Energy (kcal/mol)
ZINC00728611	-8.951	1.699	ZINC00666373	-7.611	1.622
ZINC00812896	-8.498	1.565	ZINC01422345	-7.593	1.77
ZINC02057831	-8.352	1.688	ZINC00036961	-7.580	3.03
ZINC02336211	-8.217	1.69	ZINC00531136	-7.571	1.441
ZINC00817961	-8.208	3.987	ZINC00071726	-7.562	1.763
ZINC00466889	-8.118	0.56	ZINC00115929	-7.536	3.507
ZINC01384119	-8.073	1.09	ZINC00844377	-7.514	1.613
ZINC01423544	-7.923	1.956	ZINC00844343	-7.465	1.381
ZINC02059291	-7.741	1.116	ZINC02057819	-7.385	1.02
ZINC01384109	-7.722	0.522	ZINC00861388	-6.881	2.041
ZINC01422341	-7.665	1.651	ZINC00182707	-6.511	1.587

Table 4.27. ZINC ID, XP GScore and strain values of ligands in cluster 13

ZINC ID	XP GScore	Strain Energy (kcal/mol)	ZINC ID	XP GScore	Strain Energy (kcal/mol)
ZINC00634793	-9.989	1.603	ZINC00287483	-7.635	2.51
ZINC00640019	-9.437	1.393	ZINC00407171	-7.507	2.905
ZINC00485734	-9.158	0.021	ZINC00092990	-7.307	0.829
ZINC00863220	-8.948	2.082	ZINC00531403	-7.281	0.425
ZINC01108051	-8.27	1.021	ZINC00662961	-7.279	0.248
ZINC00924613	-8.027	1.295	ZINC00287536	-7.248	3.236
ZINC01269582	-7.964	0.269	ZINC00266365	-7.226	0

Table 4.27. ZINC ID, XP GScore and strain values of ligands in cluster 13. (cont.)

ZINC ID	XP GScore	Strain Energy (kcal/mol)	ZINC ID	XP GScore	Strain Energy (kcal/mol)
ZINC00824792	-7.961	0.871	ZINC01069211	-7.008	3.484
ZINC00407170	-7.849	1.796	ZINC00049541	-6.814	0
ZINC00518118	-7.76	0.399	ZINC01021695	-6.573	1.911
ZINC00420705	-7.706	1.033			

The hydroxyl group of most ligands made either side chain or backbone hydrogen bond with Ser 345, hydrophobic pocket residue, as shown in Table 4.28. Average GScore of each cluster was slightly different from each other because of having different types of interactions with GBA, and best average XP Score belonged to cluster 13. Nevertheless the maximum standard deviation between the ligand's XP score and cluster's average XP score was calculated for cluster 13 showing that the range of XP Score in that cluster is wider than those of other clusters. In general, ligands in cluster 7 made hydrogen bonds with both Ser 345 and Glu 340 while ligands in cluster 11 made hydrogen bonds with another catalytic residue Glu 235 and Asp 127. Although the average XP GScore of cluster 11 was less than the other clusters, BEI and SEI values were higher than those of other clusters suggesting that ligands making bonds with Glu 235 and Asp 127 were more druglike. Moreover, it implied that Glu 340 contributed more than Glu 235 in binding ability of ligand because ligands interacting with Glu 340 had better XP GScores. Zinc00634793 from cluster 13 having the best GScore for this pharmacophore hypothesis made three hydrogen bonds with active site residues Glu340, Asp127 and Asn 396, which favored ligand binding.

Table 4.28. Representative interactions of ligands in different clusters of ADDHR.125 hypothesis.

Cluster	ZINC ID	Catalytic Residues		Active Site Residues				Hydrophobic Pocket Residues			Other Residues				XP Score of Ligand	Avg XP Score of Cluster	STD
		Glu 235	Glu 340	Asn 234	Trp 179	Asp 127	Trp 381	Asn 396	Ser 345	Tyr 313	Phe 347	Tyr 244	Lys 346	Phe 246			
7	ZINC12469416		H(s)						H(b)						-8.664	-7.975	0.487
7	ZINC00551825		H(s)						H(b)		π - π				-8.562		0.415
7	ZINC18324763		H(s)						H(b)		π - π				-8.530		0.392
9	ZINC00728611								H(s)			H(s)	H(s)	π - π	-8.951	-7.765	0.838
9	ZINC00812896		H(s)						H(s)		π - π		H(s)		-8.498		0.518
11	ZINC05343657	H(s)				H(s)	H(s)		H(b)	π - π	π - π				-6.905	-6.903	0.001
11	ZINC09124654	H(s)				H(s)	H(s)		H(b)	π - π	π - π				-6.870		0.023
13	ZINC00634793		H(s)			H(s)		H(s)	H(b)						-9.989	-8.150	1.300

4.2.5.4. Rigid Docking of Ligands obtained by Substructure Search of Cluster's Scaffold for Hypothesis AAADR.305. Similar to the case with hypothesis ADDHR.125, top 1000 ligands obtained by XP mode were analyzed with respect to Glide GScore and strain energy values. BEI and SEI values of each ligand were calculated in order to find the average BEI and SEI values of each cluster.

As shown in Figure 4.21, clusters with the highest average BEI values were cluster 6 and cluster 7. In cluster 16, there were some ligands with high BEI values as well as ligand with very low BEI values, so that the BEI range of this cluster was too large and its average BEI value was lower than those of other clusters. With respect to SEI results, cluster 12, 16 and 17 had approximately the same value of average SEI. Nevertheless, similar to BEI, the range of cluster 16 was very large. According to analysis of these two graphs, cluster 12 and cluster 17 can be accepted as the most druglike clusters for ADDHR.125 hypothesis.

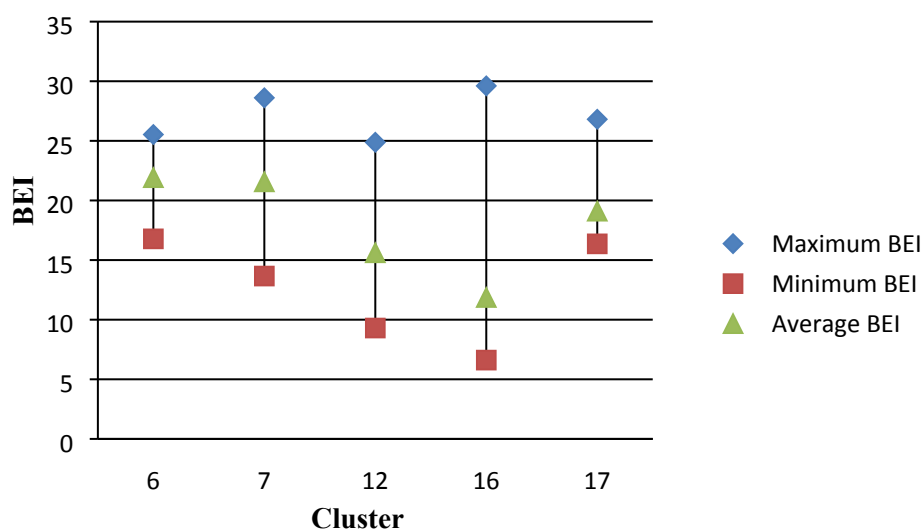


Figure 4.21. Average BEI and SEI values of clusters of hypothesis AAADR.305

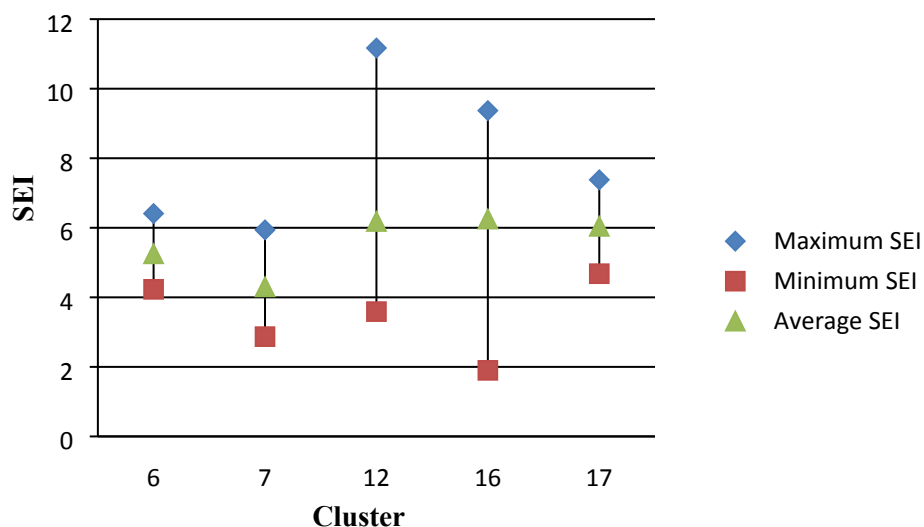


Figure 4.21. Average BEI and SEI values of clusters of hypothesis AAADR.305 (cont.)

Table 4.29. ZINC ID, XP GScore and strain values of ligands for cluster 12

ZINC ID	XP GScore	Strain Energy (kcal/mol)	ZINC ID	XP GScore	Strain Energy (kcal/mol)
ZINC00032841	-7.393	1.876	ZINC00755437	-6.527	1.360
ZINC00106233	-8.127	1.924	ZINC00798421	-7.246	2.760
ZINC00145814	-6.541	1.056	ZINC00798422	-6.848	0.263
ZINC00150595	-7.443	1.527	ZINC00798423	-7.821	0.938
ZINC00205703	-7.554	3.039	ZINC00814515	-6.785	0.556
ZINC00214015	-7.128	0.005	ZINC00925300	-8.457	0.187
ZINC00214017	-7.943	0.081	ZINC00934689	-7.016	1.285
ZINC00214019	-8.165	0.056	ZINC00958606	-7.469	3.730
ZINC00214024	-8.284	0.206	ZINC00959775	-6.630	1.790
ZINC00214031	-7.838	1.470	ZINC00999647	-7.88	0.042
ZINC00214036	-7.884	0.071	ZINC01002600	-7.606	1.339
ZINC00309374	-7.863	0.094	ZINC01015229	-7.276	0.208
ZINC00467804	-7.591	3.279	ZINC01044571	-7.947	0.674
ZINC00467805	-7.571	1.146	ZINC01107680	-7.282	2.731
ZINC00467806	-7.644	3.711	ZINC01107695	-6.599	0.145
ZINC00467807	-7.521	1.543	ZINC01122277	-6.835	3.308
ZINC00467891	-8.233	0.011	ZINC01243125	-6.549	2.733
ZINC00529237	-7.196	1.785	ZINC01243754	-6.948	1.694
ZINC00574173	-6.765	0.421	ZINC01255543	-7.29	2.602
ZINC00608850	-7.256	1.310	ZINC01269540	-7.816	3.247
ZINC00610007	-7.539	2.693	ZINC01303070	-6.675	1.735
ZINC00616788	-8.618	0.458	ZINC01307568	-6.747	3.762

Table 4.29. ZINC ID, XP GScore and strain values of ligands for cluster 12 (cont.)

ZINC ID	XP GScore	Strain Energy (kcal/mol)	ZINC ID	XP GScore	Strain Energy (kcal/mol)
ZINC00619426	-6.577	0.007	ZINC01394651	-7.32	2.346
ZINC00636618	-7.282	1.349	ZINC01394661	-7.384	0.018
ZINC00710357	-7.602	1.079	ZINC01394918	-6.522	2.887
ZINC01407282	-7.442	3.438			

Table 4.30. ZINC ID, XP GScore and strain values of ligands for cluster 17

ZINC ID	XP GScore	Strain Energy (kcal/mol)	ZINC ID	XP GScore	Strain Energy (kcal/mol)
ZINC00003944	-7.849	0.936	ZINC00056389	-8.302	2.201
ZINC00017874	-9.114	0.623	ZINC00109315	-7.900	0.097
ZINC00021604	-8.672	3.618	ZINC00109327	-6.854	0.321
ZINC00022009	-6.947	0.227	ZINC00109388	-8.180	3.232
ZINC00022010	-8.571	2.867	ZINC00116819	-7.732	0.401
ZINC00026427	-7.234	0.000	ZINC00116833	-7.847	0.866

Representative interaction Table 4.31 implied that the best cluster for AAADR.305 hypothesis is cluster 7. Generally, Ligands in cluster 7 made hydrogen bonds with both Asp 127 and Trp 179, which are active site residues. With regards to cluster 16, most of the compounds made two side chain hydrogen bonds with Ser345 and Lys 346, residue that located in Loop 1 and provides electrostatic interaction with ligands [59], and pi-pi stacking bond with Trp 313. In addition, pi-pi stacking bond with Phe 347 was also observed with some ligands. Best compound in cluster 16, Zinc15150733 made five hydrogen bonds with different active site residues such as Trp 179, Glu 340 resulting enhancement of GScore value. In cluster 12, ligands have lower absolute GScore values than the ligands in other clusters. The reason can be lack of interaction with catalytic and active site residues.

Table 4.31. Representative Interactions of Ligands in Different Clusters of ADDHR.305 Hypothesis

Cluster	ZINC ID	Catalytic Residues		Active Site Residues				Hydrophobic Pocket Residues			Other Residues				XP Score of Ligand	Avg XP Score of Cluster	STD
		Glu 235	Glu 340	Asn 234	Trp 179	Asp 127	Trp 381	Asn 396	Ser 345	Tyr 313	Phe 347	Tyr 244	Lys 346	Phe 246			
6	ZINC04715033		H(s)		H(s)	H(s)(2)	H(s)	H(s)(2)	H(b)(2)		π - π (2)				-10.885	-8.888	1.412
6	ZINC05606432	H(s)	H(s)(2)	H(s)				H(s)							-7.467		1.004
7	Zinc04096954			H(s)	H(s)	H(s)			H(s)				H(s)		-11.138	-9.088	1.449
7	Zinc03778859			H(s)	H(s)	H(s)(2)	H(s)	H(s)(2)	H(b)						-10.559		1.040
12	Zinc00616788			H(s)				H(s)			π - π				-8.618	-7.376	0.878
12	Zinc00610007			H(s)					H(s)				H(s)		-7.539		0.115
16	Zinc15150733		H(s)	H(s)	H(s)	H(s)(2)	π - π	H(s)	H(s)	π - π	π - π		H(s)		-9.675	-7.864	1.280
16	Zinc00891187								H(s)	π - π	π - π		H(s)		-8.841		0.690
16	Zinc01335753								H(s)	π - π			H(s)		-8.653		0.557
17	Zinc00017874				H(s)				H(s)		π - π				-9.114	-7.835	0.904

4.2.6. Induced Fit Docking for Ligand Based Modeling

After rigid docking, compounds were examined according to their XP Gscore and druglikeness properties such as human oral absorption, blood brain barrier coefficient and octanol-water partition coefficient.

Table 4.32. Induced fit docking scores of well-known ligands of GBA and commercial compounds.

Compound	IFD XP Score
NB-DNJ	-10.877
NN-DNJ	-8.364
Isofagomine	-7.368
Ambroxol	-6.151
Aminocyclitol	-11.701
Fingolimod	-10.265
U18666A	-7.022
Spisulosine	-7.471
GW4869	-6.223

In the literature, docking studies were performed in order to determine the binding affinity of NB-DNJ to GBA active site. In one study, by using Autodock, the ligand was docked to GBA flexibly while the protein was kept rigid. The binding affinity of NB-DNJ to GBA was found as -8 kcal/mol [60]. Another induced fit docking study was performed using Glide, and the binding affinity of aminocyclitols were estimated in the range of -7.76 and -13.28 kcal/mol. The study revealed that the binding affinities of aminocyclitols differ with respect to R group attached to sugar moiety[2]. In our study, GScore of aminocyclitol with R group C-10 alkyl was -11.701 kcal/mol, which is consistent with the literature.

NN-DNJ, well-known inhibitor of GBA has IFD XP docking score of about -8.0 as given in Table 4.18. Consequently, compounds having lower than -8.0 XP Gscore values were used for induced fit docking.

4.2.6.1. Induced Fit Docking Results of Compounds from ADDHR.125 Hypothesis. 10 potential ligands were redocked to active site of GBA shown in Table 4.33. It is observed that IFD XP GScores of 5 ligands were lower than the rigid docking XP Score. Large molecules can make many interactions with the protein and have greater scores in rigid docking. However, when the entropy penalties are involved, flexible molecules may have lower Gscore than the rigid ones[61].

Table 4.33. Scores and druglikeness properties of proposed 10 ligands of ADDHR.125 hypothesis for IFD.

ZINC ID	Rigid Docking XP Score	IFD XP Score	QLogPo/w	QLogBB	PHOA	LLE
Zinc00640019	-9.437	-7.284	1.956	-0.934	86.2	0.127
Zinc00634793	-9.989	-10.820	1.109	-1.390	51.4	0.68
Zinc00485734	-9.158	-8.086	-1.000	-1.485	56	3.28
Zinc12469416	-8.664	-7.972	2.124	0.037	83.89	0.74
Zinc18324763	-8.530	-7.268	2.116	-0.975	69.3	0.95
Zinc04838663	-8.116	-9.346	2.185	0.201	90.2	0.466
Zinc00863220	-8.948	-10.119	2.542	-1.886	85.4	-0.09
Zinc00551825	-8.562	-10.142	2.796	0.121	89.8	0.24
Zinc33403498	-8.293	-7.728	2.112	-0.959	69	1.129
Zinc20135476	-8.241	-11.420	0.832	-0.591	63.8	2.51

4.2.6.2. Induced Fit Docking Results of Compounds from AAADR.305 Hypothesis. Similar to ADDHR.125, potential 27 compounds from AAADR.305 hypothesis were redocked in XP mode to the active site of GBA flexibly. 8 ligands had rigid XP score better than IFD XP score and they were eliminated from study. In addition, IFD XP scores of ligands of AAADR.305 hypothesis are higher than those of ADDHR.125 hypothesis.

Table 4.34. Proposed 27 ligands of AAADR.305 hypothesis for IFD.

ZINC ID	Rigid Docking XP Score	IFD XP Score	QLogPo/w	QLogBB	PHOA	LLE
Zinc04715033	-10.885	-11.748	-1.828	-2.822	30.2	2.916
Zinc01081233	-10.281	-9.093	-1.717	-1.719	44.48	3.178
Zinc01728740	-10.173	-9.512	-1.835	-1.721	43.85	3.376

Table 4.34. Proposed 27 ligands of AAADR.305 hypothesis for IFD (cont.)

ZINC ID	Rigid Docking XP Score	IFD XP Score	QLogPo/w	QLogBB	PHOA	LLE
Zinc05437232	-10.017	-12.354	-1.879	-2.174	39.7	3.530
Zinc01081238	-9.997	-10.511	-1.566	-1.709	45.56	3.239
Zinc03953883	-9.859	-11.289	-0.545	-1.86	46.09	2.665
Zinc03782862	-9.796	-9.795	-0.013	-1.046	63.61	1.835
Zinc01078624	-9.495	-9.287	-1.928	-1.771	42.6	3.970
Zinc00895521	-9.444	-10.174	4.378	-0.088	91.2	-2.08
Zinc02381519	-9.2	-7.633	-1.572	-1.599	51.29	3.824
Zinc00017874	-9.114	-10.072	1.8	-0.952	82.43	1.552
Zinc01335983	-9.077	-9.401	4.106	-0.353	92.5	-0.31
Zinc35092286	-9.065	-10.017	-0.702	-1.616	59.3	3.071
Zinc03964331	-9.039	-9.682	0.803	-1.802	60.45	1.578
Zinc01329311	-8.867	-10.181	6.318	-0.326	100	-3.810
Zinc00891187	-8.841	-7.648	4.958	-0.042	100	-2.435
Zinc00891180	-8.805	-9.905	5.183	0.13	100	-2.637
Zinc35092294	-8.73	-9.184	-0.174	-1.002	90.3	-2.784
Zinc01335991	-8.673	-8.260	4.499	-0.813	96.59	-1.848
Zinc01335753	-8.673	-11.615	5.058	-0.396	100	-2.424
Zinc00021604	-8.672	-8.590	1.517	-1.584	72.93	1.140
Zinc01638930	-8.654	-10.284	0.19	-1.56	63.52	2.468
Zinc00891392	-8.654	-9.472	4.482	-0.379	100	-0.586
Zinc01335704	-8.64	-11.207	5.135	-0.621	100	-2.476
Zinc00616788	-8.618	-8.874	1.621	-1.341	76.65	4.026
Zinc00896096	-8.616	-10.974	6.241	-0.79	94.87	-3.562
Zinc013355993	-8.552	-6.784	5.199	-0.936	82.56	-2.463

In the literature, docking studies were performed in order to determine the binding affinity of NB-DNJ to GBA active site. In one study, by using Autodock, the ligand was docked to GBA flexibly while the protein was kept rigid. The binding affinity of NB-DNJ to GBA was found as -8 kcal/mol [1]. Another induced fit docking study was performed using Glide, and the binding affinity of aminocyclitols were estimated in the range of -7.76 and -13.28 kcal/mol. The study revealed that the binding affinities of aminocyclitols differ with respect to R group attached to sugar moiety. In our study, GScore of aminocyclitol with R group C-10 alkyl was -11.701 kcal/mol, which is consistent with the literature.

4.3. Selectivity Analysis Results

Selectivity is another factor that must be taken into consideration during drug discovery. Greater binding tendency of ligands to proteins other than GBA is not desired. As a result, Compounds with higher IFD XP Scores were further analyzed in terms of their selectivity. Potential ligands were docked to active site of various proteins in XP mode. Alpha-Glucocerebrosidase (PDB Code:3TOP) and Alpha-Galactosidase (PDB Code:3GXT) proteins are also lysosomal proteins and they are the members of spingholipid pathway. The mutations on these proteins cause Pompe and Fabry Diseases respectively. These proteins were chosen for selectivity analysis.

Table 4.35. XP Gscore of ligands bound to different proteins.

Compound or Hypothesis	ZINC ID	2V3E (Beta- Glucocerebrosidase)	3TOP (Alpha- Glucocerebrosidase)	3GXT (Alpha- Galactosidase A)
Miglustat	Zinc0379471	-8.014	-5.868	-7.373
ADDHR.125	Zinc00634793	-9.989	-6.237	-7.118
	Zinc04838663	-8.116	-5.063	-3.433
	Zinc00863220	-8.948	-6.457	-4.419
	Zinc00551825	-8.562	-6.107	-3.156
	Zinc20135476	-8.241	-5.712	-5.501
	Zinc12469416	-8.664	-5.821	-5.417
AAADR.305	Zinc04715033	-10.885	-9.229	-8.304
	Zinc05437232	-10.017	-7.822	-6.335
	Zinc03953883	-9.859	-7.421	-8.021
	Zinc00895521	-9.444	-8.371	-5.832
	Zinc00017874	-9.114	-7.835	-5.725
	Zinc35092286	-9.065	-8.428	-6.688
	Zinc01329311	-8.867	-7.131	-3.343
	Zinc35092294	-8.73	-8.292	-6.476
	Zinc01335753	-8.673	-6.703	-3.485
	Zinc01638930	-8.654	-5.795	-3.782
	Zinc00891392	-8.654	-6.820	2.749
	Zinc01335704	-8.64	-5.570	-3.066
	Zinc00896096	-8.616	-6.211	-3.685
	Zinc00891180	-8.805	-3.792	-3.102
Structure-Based	Zinc01726382	-10.60	-10.589	-6.673
	Zinc04348040	-8.442	-6.918	-5.276

Table 4.35. XP Gscore of ligands bound to different proteins (cont.)

Structure-Based	Zinc04743969	-8.433	-8.462	-6.276
	Zinc19288139	-8.617	-7.748	-5.674
	Zinc05323884	-9.978	-5.926	-3.924

Consequently, the XP scores of all potential ligands docked to GBA were better than those of docked to alpha-glucocerebrosidase and alpha-galactosidase A implying that all ligands have greater tendency to bind GBA. The greater the difference between XP scores, the more selective ligands to GBA are obtained.

4.4. Binding Free Energy Calculations

Binding free energies of 14 selected ligands with respect to their IFD XP scores were calculated using Prime module of Schrödinger 2016. The defaults were accepted for the simulation and no constraints were given to system. All ligands had negative binding free energies showing that the bindings are spontaneous and highly negative binding free energies imply that ligand-GBA bindings are favorable.

Table 4.36. Binding Free Energy Values of Selected Ligands

ZINC ID	MMGBSA dG Bind	IFD XP Score	QlogBB
ZINC00891180	-81.643	-9.905	0.13
ZINC0017874	-68.882	-10.072	-0.952
ZINC00895521	-112.656	-10.174	-0.088
ZINC20135476	-58.063	-11.420	-0.591
ZINC00896096	-90.083	-10.974	-0.79
ZINC01329311	-80.887	-10.181	-0.326
ZINC01335983	-75.819	-9.401	-0.353
ZINC01726382	-94.466	-12.36	-0.62
ZINC35092294	-46.616	-9.905	-1.002
ZINC00551825	-66.279	-10.142	0.121
ZINC03964331	-89.849	-9.682	-1.802
ZINC18324765	-64.112	-9.343	-0.971
ZINC01638930	-61.341	-10.284	-1.56
ZINC04838663	-64.166	-9.346	0.201

IC₅₀ values of selected ligands were also calculated using IFD XP Gscore values. Because it is known that there is a correlation between binding affinity and inhibition

constant as shown in Equation 4.1. Also, K_d value is approximately equal to IC50 values [62].

$$\Delta G = -RT \ln K_d \quad (4.1.)$$

Table 4.37. Predicted IC50 values of selected ligands using IFD XP GScore.

ZINC ID	IFD XP Score	Predicted IC50 (nM)
ZINC00891180	-9.905	103.85
ZINC0017874	-10.072	79.19
ZINC00895521	-10.174	67.10
ZINC20135476	-11.420	8.87
ZINC00896096	-10.974	18.31
ZINC01329311	-10.181	66.34
ZINC01335983	-9.401	235.37
ZINC01726382	-12.36	1.93
ZINC35092294	-9.905	103.84
ZINC00551825	-10.142	70.68
ZINC03964331	-9.682	149.15
ZINC18324765	-9.343	258.61
ZINC01638930	-10.284	56.13
ZINC04838663	-9.346	257.35

As shown in Table 4.37, predicted IC50 values were less than those of experimental ligands tabulated in Table 4.15. Also, it was investigated that IC50 values of compounds increased exponentially with increasing GScore value (Figure A.3.). Binding free energy values and blood brain barrier coefficients were taken as the major parameters in the selection of potential pharmacological chaperones of GBA protein. Because Type 2 and Type 3 Gaucher Diseases are neurodegenerative, ligands having positive blood brain coefficient or close to zero were further examined. The blood brain barrier coefficient cut off value was accepted as -0.8 which is the BBB coefficient of Miglustat. After the comparison of all selected ligands, ZINC00895521, ZINC04838663, ZINC00891180, ZINC35092294, ZINC0017874 and ZINC01726382 were chosen for further MD simulations because of their good induced fit Gscores and BBB coefficients. Interactions between selected 6 ligands and GBA protein were further analyzed to understand changes upon binding using induced fit docking results.

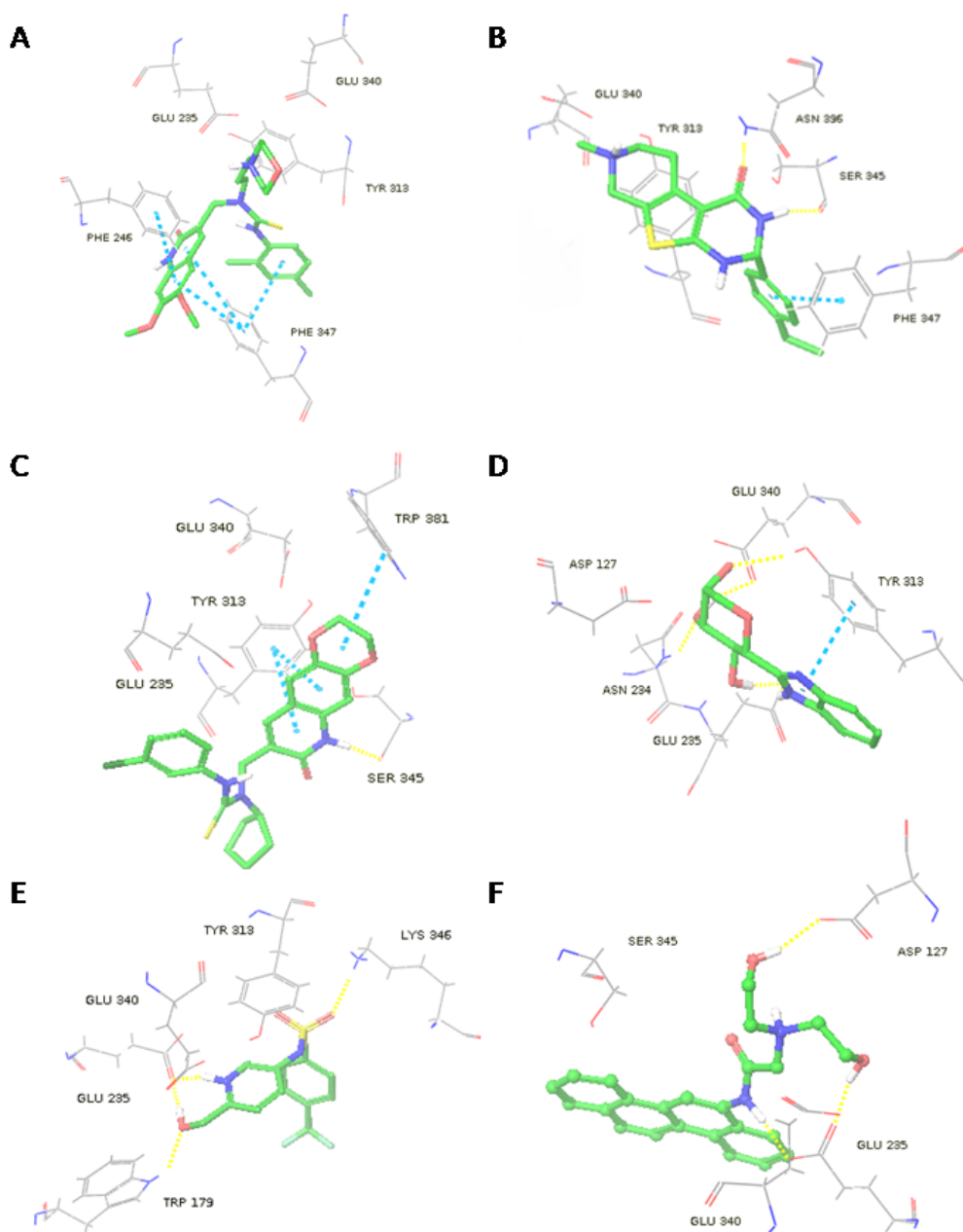


Figure 4.22. IFD Interactions between top 6 ligands and GBA (A: Zinc00895521, B: Zinc04838663, C: Zinc00891180, D: Zinc35092294, E: Zinc00017874, F: Zinc01726382) (Hydrogen bond: Yellow dashed line, Pi-Pi stacking bond: Turquoise dashed line)

In the interactions of Zinc0895521 and GBA protein, two Pi-Pi bonds were formed between veratrol part of the ligand and aromatic residues of protein (Phe 347 and Phe 246) while Phe 347 had another pi-pi stacking bond with benzene of xylene part of the ligand. In addition, there were three salt bridges with catalytic residues (Glu340 and Glu235) and

Asp127. These interactions were completely different than those of Miglustat. Using crystal structure, 2V3E, Miglustat makes hydrogen bonds with Trp 127, Trp 179, Glu 340, Trp 381 and Asn 396.

Zinc04838663 is smaller than Zinc00895521. It made pi-pi stacking interaction between benzene part of the ligand and Phe 347. Ser 345 of the hydrophobic pocket made a backbone hydrogen bond with hexahydropyrimidine portion of the ligand.

The nitrogen atom of Zinc00891180 made a hydrogen bond with Ser 345. In addition Tyr 313 and Trp 381 residues made pi-pi stacking bonds with two aromatic rings. Due to interaction with hydrophobic pocket of GBA, these rings play major role on the binding of Zinc00891180 at the active site of GBA.

In addition, Zinc35092294 is the lightest potential drug in this study. It made hydrogen bonds with catalytic residues, Glu340 and Glu235. Furthermore, single pi-pi stacking bond was formed with Tyr 313 of the hydrophobic pocket.

Zinc00017874 is also a light compound with respect to other potential ligands. It has two aromatic rings connected with a sulfure atom. It made two hydrogen bonds with catalytic residue, Glu 235 and one hydrogen bond with active site residue, Trp 179. Moreover, it also interacted with Lys 346 which lies in the groove between GBA C terminus beta-sheet domain and TIM barrel near loop 1.

With regards to Zinc01726382, it made hydrogen bonds with catalytic residues Glu 235 and Glu 340 as well as active site residue Asp 127. Its binding type was more different than other ligands. It bound to ligand from its tail group in stead of its ring moiety.

Ligand flexibility was also studied by superpositioning of different poses of the ligand without transformation of molecules in order to understand conformational difference between poses having different XP scores. The average RMSD values of all ligands were measured by taking the arithmetic average of RMSD values of top poses of each ligand, and tabulated in Table 4.38 while RMSD values of each pose were added into

Table A.28. First, the protein was deleted from the workspace and the best pose of each ligand was fixed at workspace of Maestro, and other 4 poses were selected one by one and their RMSD values with respect to best pose were calculated while keeping their position and preventing overlapping with each other.

Table 4.38. Average RMSD values of different poses of well-known and best ligands from all hypothesis.

Compound	Average RMSD
NB-DNJ	1.46
NN-DNJ	1.25
Isofagomine	2.02
Aminocyclitol	1.92
Ambroxol	1.57
GW4869	2.00
Sphingosine	5.83
Fingolimod	2.94
Zinc00895521	5.82
Zinc04838663	3.43
Zinc00551825	3.87
Zinc20135476	1.98
Zinc00891180	3.29
Zinc00017874	2.26
Zinc35092294	2.01
Zinc01726382	1.63

Table 4.38 showed that average RMSD values of experimentally known ligands are less than 2 Å implying that there is not major difference among the binding conformations of these ligands. Most of these ligands are very small ligands having a cyclohexane ring moiety. In addition, the alkyl chain attached to ring moiety is not too long. Nevertheless, although proposed compounds have a cyclohexane moiety similar to experimentally known ligands, their tail group involve large rings and they change their conformations leading to increase in average RMSD value. Similarly, the ligand, sphingosine, which has long tail group, has higher RMSD value than those of experimentally known ligands because of the different conformations of long tail group. Zinc00895521 had the maximum

conformational difference between its poses whereas the minimum difference belonged to Zinc01726382. Zinc00895521 is the heaviest molecule in that study while Zinc01726382 is the only molecule obtained from structure-based modeling.

In addition, the conformational difference between GBA and a ligand was also determined by calculating RMSD. This time, protein was also kept in the system. The arithmetic average of RMSD values for different protein-ligand complexes were calculated and tabulated in Table 4.39. Average RMSD values of all complexes were less than 1 Å showing that there was not any major change between the different conformations of protein-ligand complexes.

Table 4.39. Average RMSD values of protein-ligand complexes

Complex	Average RMSD Value (Å)
GBA-Zinc00895521	0.72
GBA-Zinc04838663	0.43
GBA-Zinc00891180	0.36
GBA-Zinc35092294	0.19
GBA-Zinc01726382	0.25
GBA-Zinc00551825	0.42
GBA-Zinc20135476	0.20
GBA-Zinc0017874	0.22

Table 4.40. Properties of top ligands from ligand-Based and structure-based modeling and their derivatives.

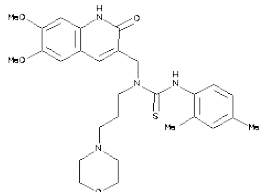
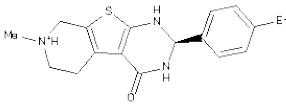
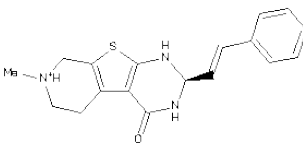
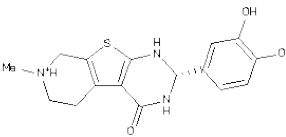
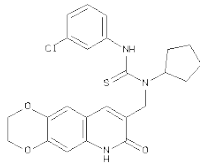
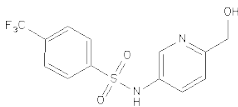
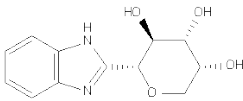
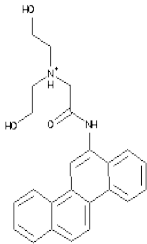
Compound	Hypothesis	ZINC ID	2D Structure	MW (Da)	APO Str. XP GScore	Rigid XP GScore	Induced Fit XP GScore	QlogPo/w	QlogBB	%HOA	LLE
1	AAADR	Zinc00895521		525	-6.206	-9.44	-10.174	4.378	-0.088	91.2	2.32
2	ADDHR	Zinc04838663		327.44	-5.145	-8.12	-9.346	2.185	0.201	90.2	3.54
3	ADDHR	Zinc00551825		325.43	-5.122	-8.56	-10.142	2.796	0.121	89.9	3.15
4	ADDHR	Zinc20135476		331.38	-4.207	-8.24	-11.420	0.832	-0.591	63.8	4.82

Table 4.40. Properties of top ligands from ligand-based and structure-based modeling and their derivatives (cont.)

Compound	Hypothesis	ZINC ID	2D Structure	MW (Da)	APO Str. XP GScore	Rigid XP GScore	Induced Fit XP GScore	QlogPo/w	QlogBB	%HOA	LLE
5	AAADR	Zinc00891180		469.98	-3.505	-8.82	-9.905	5.183	0.13	100	1.27
6	AAADR	Zinc00017874		332.30	-5.003	-9.11	-10.072	1.8	-0.952	82.4	3.84
7	AAADR	Zinc35092294		250.25	-4.158	-8.73	-9.18	-0.174	-1.002	70.3	6.56
8	HRH	Zinc01726382		388.46	-6.917	-10.60	-12.36	2.61	-0.62	84.6	5.17

4.5. Molecular Dynamics Simulation Results

4.5.1. MD Simulations of Apo-GBA Structure

GBA structure of GBA (PDB ID: 1OGS) was downloaded from Protein Data bank and prepared using Protein Preparation Task. After missing residues were completed, it was simulated for 50 ns at 310 K and 1 atm in order to compare the difference between the mobility of GBA in the presence and absence of a ligand as well as to observe the effect of ligand on mutated residues such as Leu 444 and Asn 370. RMSD and RMSF profiles of apo form of GBA were given in Figures 23 and 24.

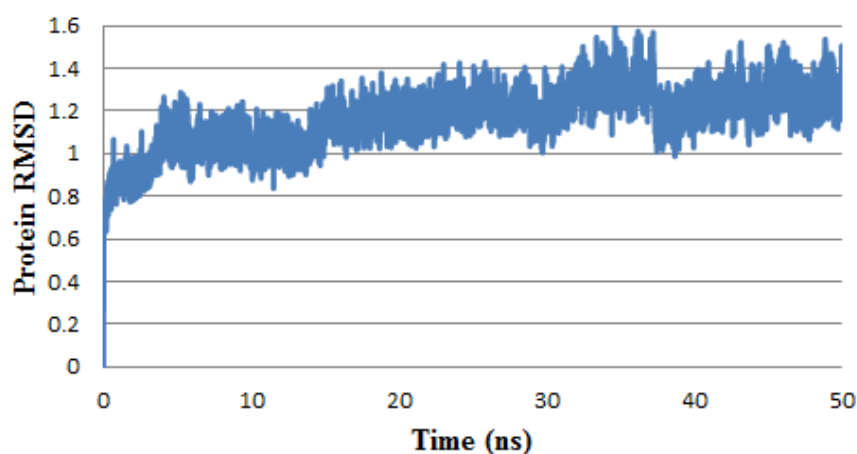


Figure 4.23. RMSD properties of apo-GBA Ca atoms

Average RMSD of apo-GBA structure was about 1.167 Å. As shown in Figure 4.23, RMSD value of GBA increased during the simulation and did not reach to equilibrium in 50 ns. A ligand is needed to stabilize the protein. In Figure 4.24 RMSF values was plotted. The figure showed that most of residues had RMSF values lower than 1 Å. The mutated residues such as Leu 444, Asn 370, Gly 202 also had low RMSF values. Nevertheless, some regions exceeded the RMSF value of 1.5 Å. For instance, Arg 170 had high RMSF value as expected because it is located between the beta-sheet and alpha helix. So it was very mobile because of the change in conformation.

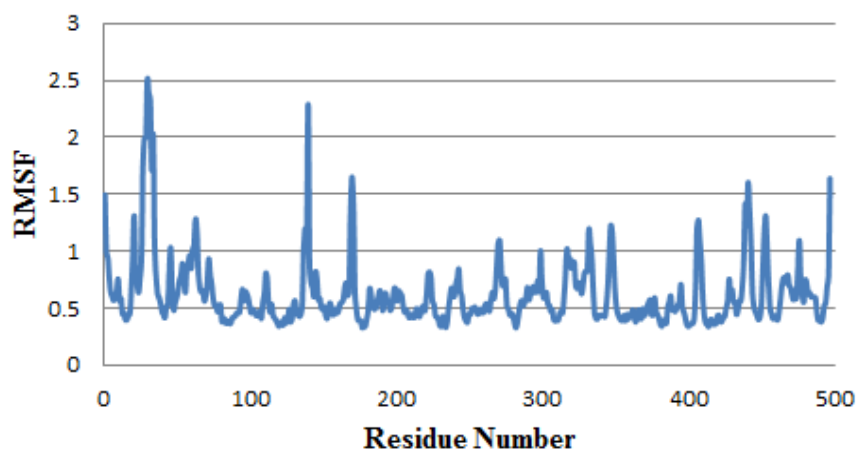


Figure 4.24. Average RMSF Values of the APO-GBA Structure

4.5.2. MD Simulations in the presence of N-nonyl-deoxynojirimycin (NN-DNJ)

GBA structure with N-nonyl-deoxynojirimycin, NN-DNJ, (PDB ID: 2V3E) was retrieved from Protein Databank and prepared for MD simulation in order to use NN-DNJ as base compound. NN-DNJ compound was used for simulation instead of NB-DNJ (Miglustat) because all docking studies were carried out using GBA with NN-DNJ (PDB ID: 2V3E). NB-DNJ and NN-DNJ make exactly the same number of hydrogen bonds with the same residues of GBA active site (Figure A.2.). MD simulation was performed at 310 K and 1 atm for 50 ns. RMSD profiles of both ligand and protein were shown in Figure 4.25.

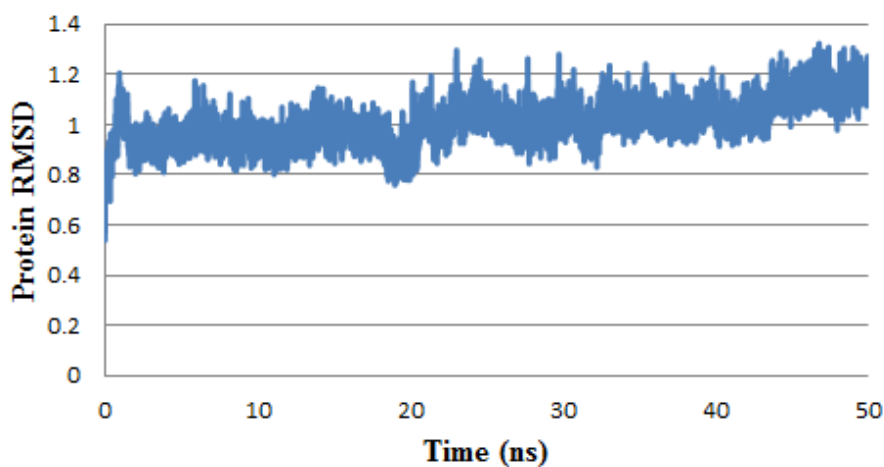


Figure 4.25. RMSD properties of GBA C α atoms (top panel) and ligand, NN-DNJ, (bottom panel) after superpositioning based on protein C α coordinates

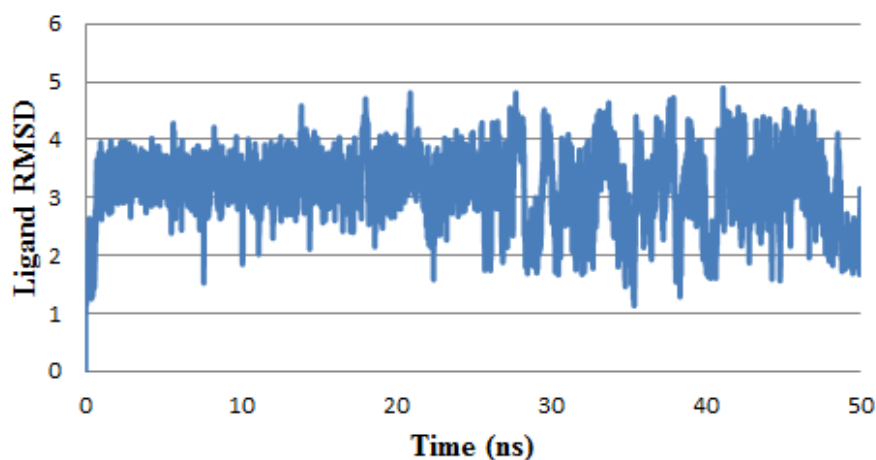


Figure 4.26. RMSD properties of ligand, NN-DNJ, (bottom panel) after superpositioning based on protein C α coordinates

Average RMSD of GBA was around 1 Å implying that NN-DNJ stabilized the GBA while the ligand was more flexible and its RMSD value was calculated as 3.21. As displayed in Figure 26, NN-DNJ was very mobile, and RMSD values fluctuated because of its movement. This fluctuations are caused by the movement of nonyl group attached to a sugar moiety. Nevertheless, the bonds between the protein and NN-DNJ were highly strong and stable preventing the separation of NN-DNJ from GBA active site.

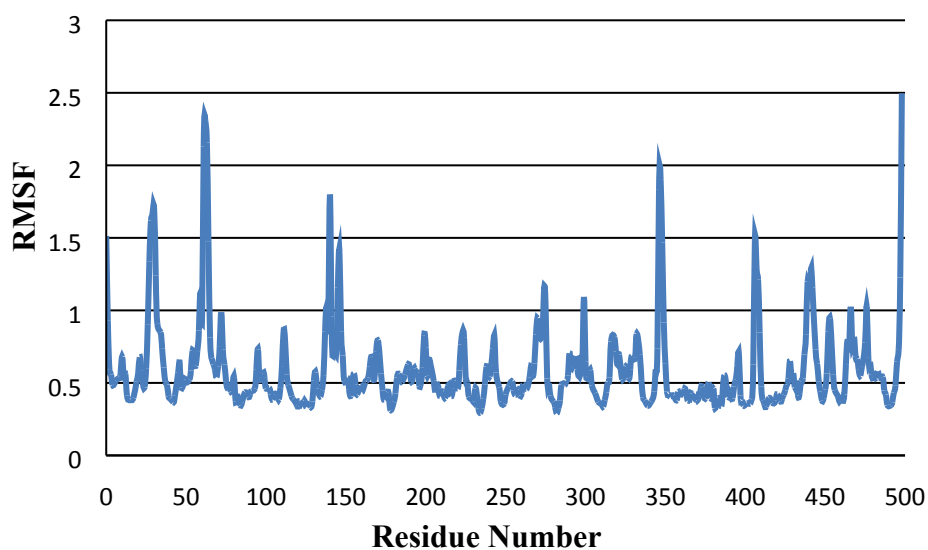


Figure 4.27. Average RMSF values of the GBA in the presence of NN-DNJ

RMSF values of each residue were drawn as displayed in Figure 4.27. Most of residues had RMSF values lower than 1 Å showing that they were almost rigid during the simulation. Besides, residues Phe 346 and Phe 347 had RMSF values greater than 1.5 Å as expected because they are involved in Loop 1. In addition, RMSF values of region 61-63 also exceed 2 Å corresponding to beginning of the beta-sheet. With regards to catalytic residues, Glu 235 and Glu 340, they had RMSF values around 0.3 Å. The mutated residues Leu 444 and Asn 370 were more flexible with RMSF of 0.729 Å and 0.645 Å, respectively. Nevertheless, mutated residue caused to Type 1 Gaucher Disease, Asn 370 was more stable than other mutated residues, its RMSF value was around 0.38 Å.

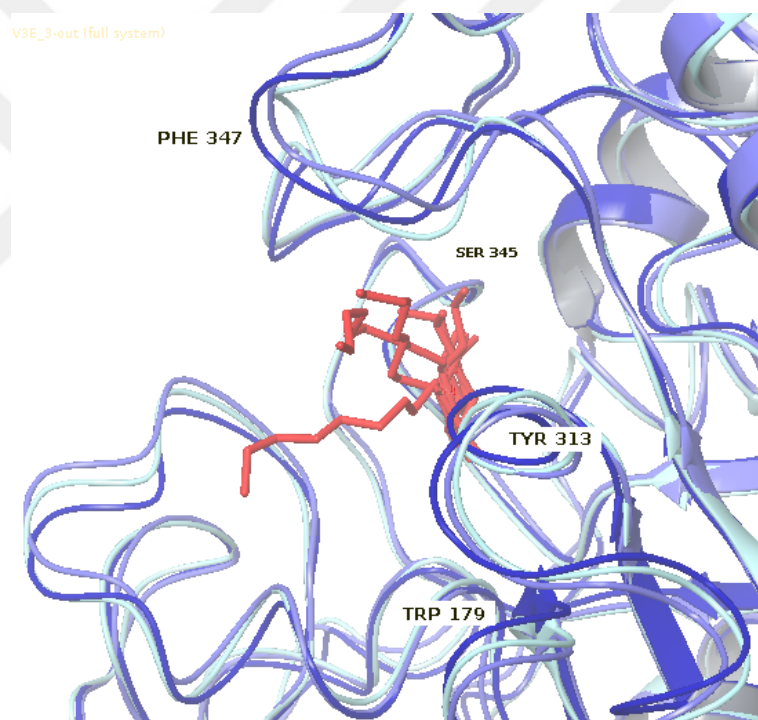


Figure 4.28. Conformational change in GBA-NN-DNJ complex throughout the simulation. (Initial conformation = Dark blue, Conformation at 15 ns = light blue, Conformation at 40 ns = Turquoise)

Conformational changes in the structure of GBA-NN-DNJ complex were examined by superposition of C α atoms of GBA at different time steps. As shown in Figure 4.28, the ligand was very mobile at the active site. Nonyl moiety of NN-DNJ was flexible while sugar moiety was stabilized by the hydrogen bonds occurred with the active site residues. The most flexible loop was Loop 1. It moved away from the ligand. Especially, Phe 347 in

loop 1 was very mobile and Ser 345 involved in hydrophobic pocket made conformational change because of the movement of loop 1. Other hydrophobic residue, Tyr 313 was very rigid throughout the simulation. Similarly, active site residues Trp 179, Asp 127 and Trp 381 were stable and did not make significant conformational changes. Hydrogen bond, occurred between hydroxyl group and residues Asp 127 and Trp 179, increased ligand affinity to GBA.

The ring moiety of NN-DNJ was very rigid during MD simulation because of strong hydrogen bonds with active site residues Asn 234, Trp 179 and Asp 127 etc. In contrast, tail group was very flexible leading to increase in RMSD value. In detailed analysis of interactions shown in Figure 4.29 implied that hydrogen bonds between NN-DNJ and GBA active site maintained until 30 ns. Around 30 ns, one of two hydrogen bonds between NN-DNJ and Asp 127 broke, but stability did not influence. Then, the same hydrogen bonds were observed for the rest of the simulation like at the beginning of the simulation. In addition, the distances between active site residues and NN-DNJ were measured, and it was found that the distance between all active site residues and NN-DNJ changed between 1.85 Å and 2.1 Å showing that it was very close to these residues. Consequently, it was determined that NN-DNJ made very strong hydrogen bonds with GBA active site.

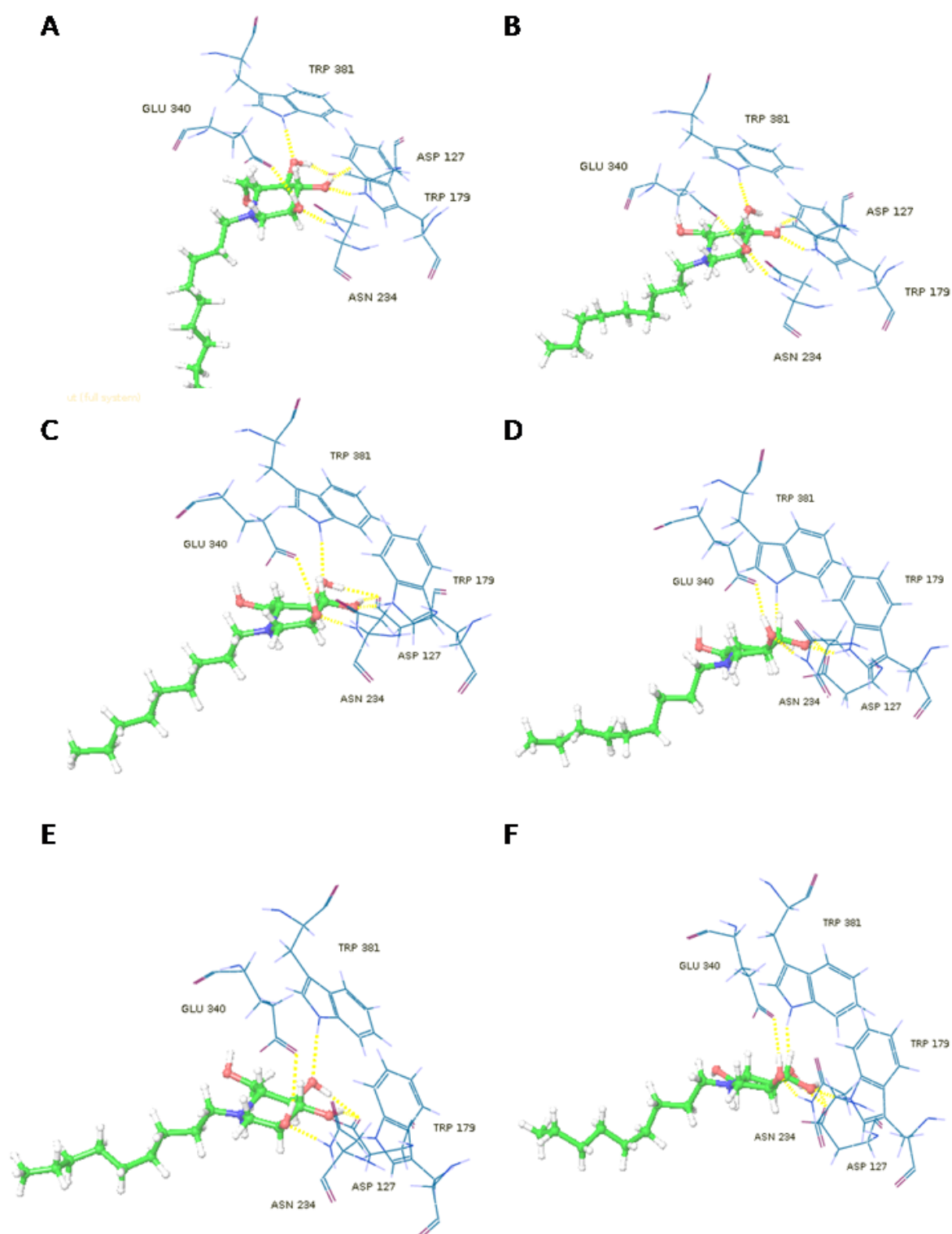


Figure 4.29. Conformations of NN-DNJ throughout the MD Simulation

(A: conformation at 0 ns, B: conformation at 10 ns, C: conformation at 20 ns, D: conformation at 30 ns, E: conformation at 40 ns.) (Hydrogen bond: Yellow dashed line)

4.5.3. MD Simulations in the presence of Zinc00895521

The stabilities of hit ligands were studied by MD simulations using Desmond. The total simulation time was set to 50 ns at 310 K and 1 atm. Root Mean Square Deviation (RMSD) and Root Mean Square Fluctuation (RMSF) graphs of ligands were obtained in order to determine flexible regions of both ligand and protein.

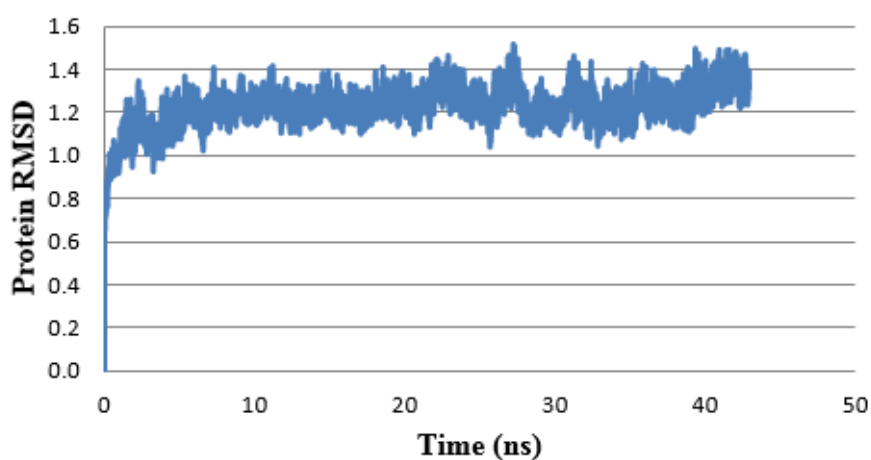


Figure 4.30. RMSD properties of GBA C α atoms (top panel) and ligand, Zinc00895521, (bottom panel) after superpositioning based on protein C α coordinates

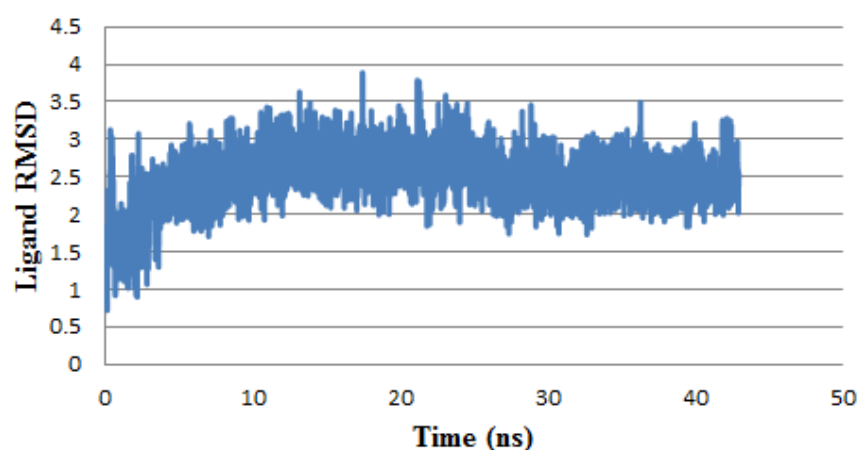


Figure 4.31. RMSD properties of GBA C α atoms (top panel) and ligand, Zinc00895521, (bottom panel) after superpositioning based on protein C α coordinates

RMSD values of both ligand and GBA were measured with respect to their initial positions at the beginning of the simulation (Figure 4.30 and 4.31). RMSD properties showed that ligand and GBA reached to equilibrium around 5-10 ns. RMSD change of the protein was between 1- and 1.4 Å, and the average RMSD value of the GBA was about 1.21 Å. RMSD change of ligand was slightly larger than the protein, it was between 2.2 Å and 3.2 Å while the average RMSD value of the ZINC00895521 compound was nearly 2.49 Å. This large RMSD values is caused by the high molecular weight of ring groups attached to scaffold of Zinc00895521. These ring groups did not make hydrogen bonds with GBA, so that they are flexible and leads to increase in ligand RMSD.

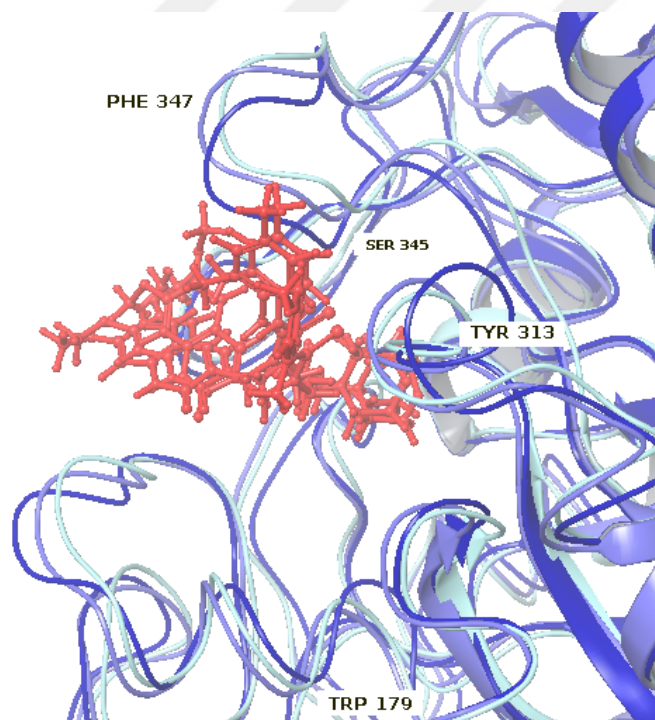


Figure 4.32. Conformational change in GBA-ZINC00895521 complex throughout the simulation. (Initial conformation = Dark blue, Conformation at 15 ns = light blue, Conformation at 40 ns = Turquoise)

ZINC00895521 was more rigid than NN-DNJ, superposition of the ligand at different times displayed that it did not make significant conformational change. This rigidity can be caused by its nitrogen and oxygen atoms, which can make various hydrogen bonds with the active site of GBA. Hydrophobic pocket was more flexible than NN-DNJ. Tyr 313 in loop 3 moved towards the ligand in 15 ns, while loop 1 moved away from the

ligand up to 15 ns. After 15 ns, no dramatic change in conformation was observed. Besides, active site residues such as Trp 179, Asp 127 were very rigid. Purine ring made π - π interaction with Phe 347 which may increase ligand stabilization. Besides, N atom in the morpholine moiety made a hydrogen bond with the active site residue Asp 127, and increased ligand binding affinity to GBA. Also, benzene-purine ring made hydrophobic interaction with Phe 246.

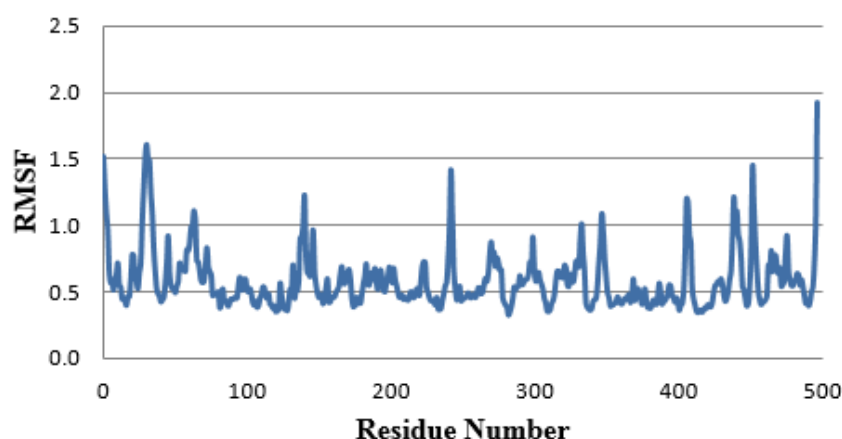


Figure 4.33. Average RMSF values of the GBA

The mobility of the residues was determined by using RMSF results as shown in Figure 4.33. The average RMSF values of residues was about 0.7 Å. The most flexible part of the GBA receptor was the residues between 50-100, 240-250, 340-350 and 445-450. Loops are expected to be more flexible than the other parts of the protein. On that bases, it was observed that residues 341-350 corresponding to loop 1 were more flexible whereas residues 393-396 corresponding to loop 2 were more stable than loop1. This stability may be due to having hydrogen bond between the ligand and active site residue Asn 396. In addition, Residues between 40-50 were highly mobile because there were some highly unconserved residues in that region. In contrast, residues forming secondary structures are more rigid than the other parts of the protein as expected. For instance, residues between 359-371 and 205-221 are alpha-helices and they had the lowest RMSF values.

Asn 370 that had the highest tendency to mutation has RMSF value of 0.51 Å while the Leu 444 causing the most severe disease type of Gaucher has RMSF value of 0.79 Å. Similarly, catalytic and active site residues had less than 1 Å RMSF for the whole

simulation. Low RMSF values indicate that these parts were highly rigid. In addition, Zinc0895521 has good blood brain barrier coefficient enabling it to cross the blood brain barrier. This result is a very hopeful improvement for treatment of Type 2 Gaucher disease.

Throughout the simulation there was not any major change in the loops and active site as can be seen from the RMSF and RMSD results. Ligand was surrounded by different loops involving the catalytic residues of Asn 396, Asp 127 and Trp 381. Hydrogen bond with Asp 127 was maintained during simulation implying that this bond is very stable as shown in Figure 4.34. It is known that bond with Asp 127 is very critical for pharmacological chaperone binding[24]. However, hydrogen bond between Ser345 and the ligand broke after the simulation started and was not observed again for the rest of simulation. Although pi-pi stacking interactions were observed between the ligand and Residues Phe 246 and Tyr 313, the bond between Phe 246 and ligand was lost after 20 ns whereas the bond with Tyr 313 was maintained from 10 ns to end of the simulation. The difference between the interacted residues showed that the ligand was unstable. The reason of this can be the non-interacting ring groups with high molecular weights. These rings increased the mobility of the ligand, and caused to break down of interactions. If the simulation is extended enough, the separation of the ligand from GBA active site may occur.

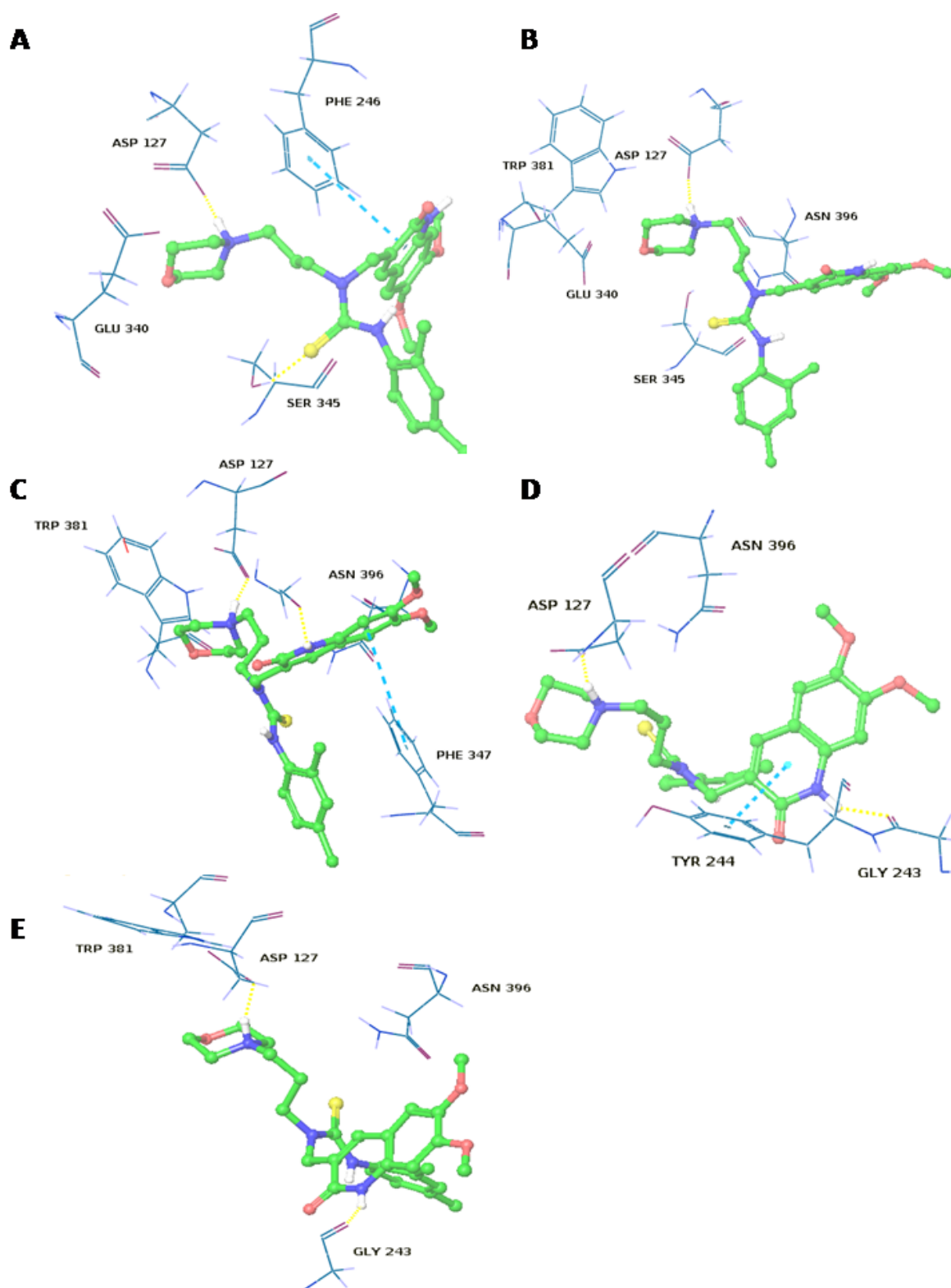


Figure 4.34. Conformations of Zinc00895521 throughout the MD Simulation
 (A: conformation at 0 ns, B: conformation at 10 ns, C: conformation at 20 ns, D: conformation at 30 ns, E: conformation at 40 ns.) (Hydrogen bond: Yellow dashed line, Pi-Pi stacking bond: Green dashed line)

4.5.4. MD Simulations in the presence of Zinc04838663

Zinc04838663 has three aromatic rings enabling to form strong hydrogen bonds with active site residues of GBA. Although Zinc04838663 has lower binding free energy value with respect to other ligands, it is analyzed with MD simulation because of its good IFD XP score and BBB coefficient. 50 ns simulation was carried out in order to analyze protein-ligand stability. Temperature and pressure of simulation box were set to 310 K and 1 atm respectively.

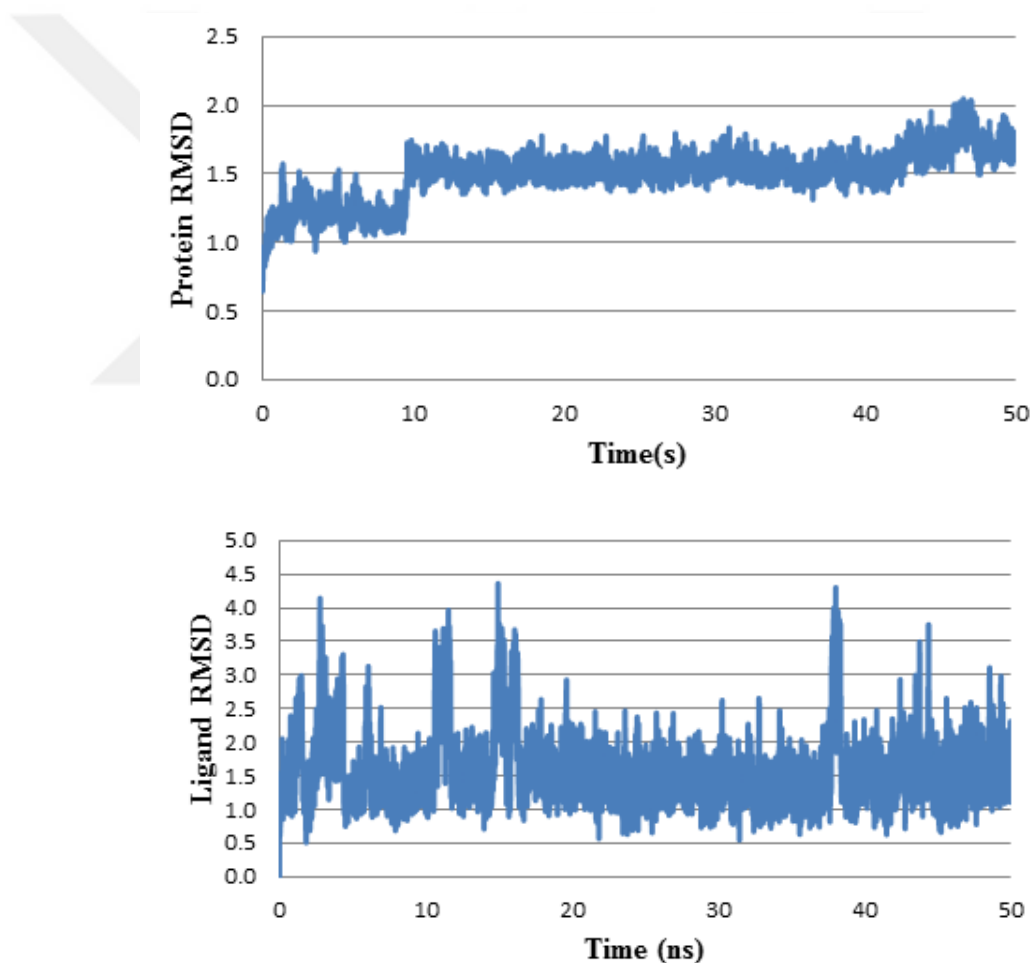


Figure 4.35. RMSD properties of GBA C α atoms (top panel) and ligand, Zinc04838663, (bottom panel) after superpositioning based on protein C α coordinates

RMSD results show that both ligand and protein reached to equilibrium within 20 ns. The average RMSD values of protein and ligand were about 1.45 Å.

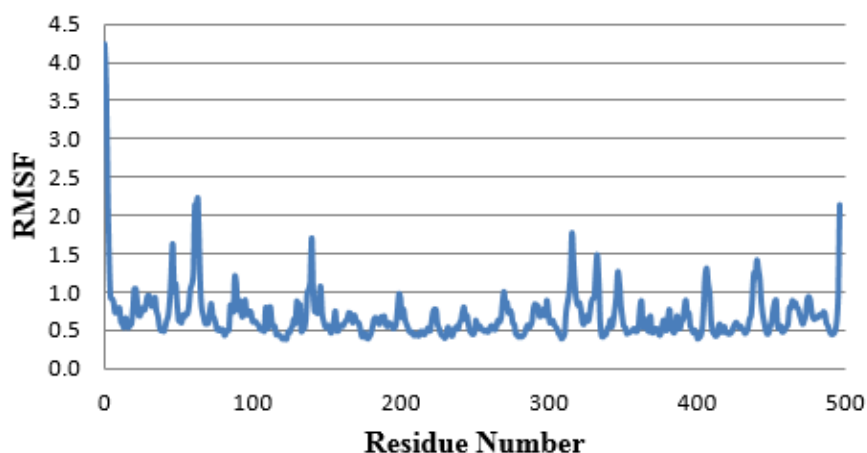


Figure 4.36. RMSF values of GBA bound Zinc04838663

RMSF result of each residue showed the mobility of residues. The catalytic residues Glu235 and Glu 340 had the lowest RMSF values which are 0.43 Å and 0.49 Å respectively. Similarly, active site residues have less than 1 Å RMSF. In addition, the residues which cause Gaucher Disease in the case of mutation were also stabilized, and their RMSF values were very low. For instance, Leu 444 had an RMSF of 0.84 Å whereas Asn 370 has 0.49 Å. In addition, loop 3 was very flexible, Phe 316 in loop 3 had RMSF of 1.71 Å. Loop 1 containing Phe 347 was also flexible leading to movement of head of the ligand, as a result of this protein movement, the ligand started to drift.

Loops consist of the active site residues which are stabilized by the ligand. So that, there were not any significant movement of loops during the simulation. Although the carbon-end of the ligand was very flexible in the simulation, protein-ligand equilibrium was not influenced by this movement, because there was a strong hydrogen bond between the nitrogen atom of piperidine portion of the ligand and Asp 127. Additional hydrogen bond occurred with residue Ser 345, but it was not as stable as hydrogen bond between Asp127 and ligand, it disappeared from time to time. Moreover, Trp 381 and Phe 246 had great tendency to make interaction with the ligand, Pi-Pi stacking bonds formed between the ligand and these residues. Tyr 313 also had favorable interactions with ligand, Zinc04838663, and made contributions to protein-ligand stability as shown in Figure 4.37.

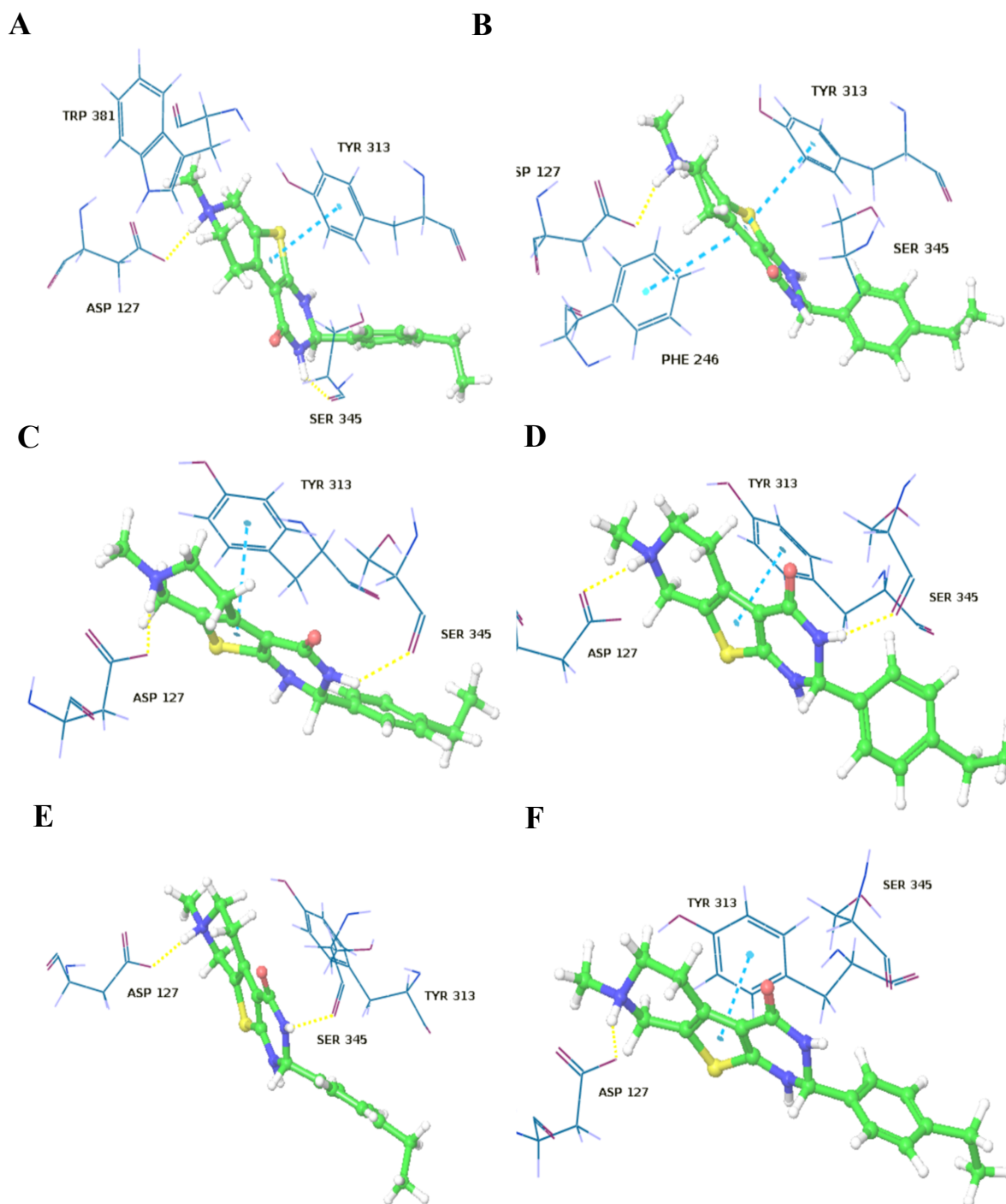


Figure 4.37. Conformations of Zinc04838663 throughout the MD Simulation (A: conformation at 0 ns, B: conformation at 10 ns, C: conformation at 20 ns, D: conformation at 30 ns, E: conformation at 40 ns, F: conformation at 50 ns) (Hydrogen bond: Yellow dashed line, Pi-Pi stacking bond: Turquoise dashed line)

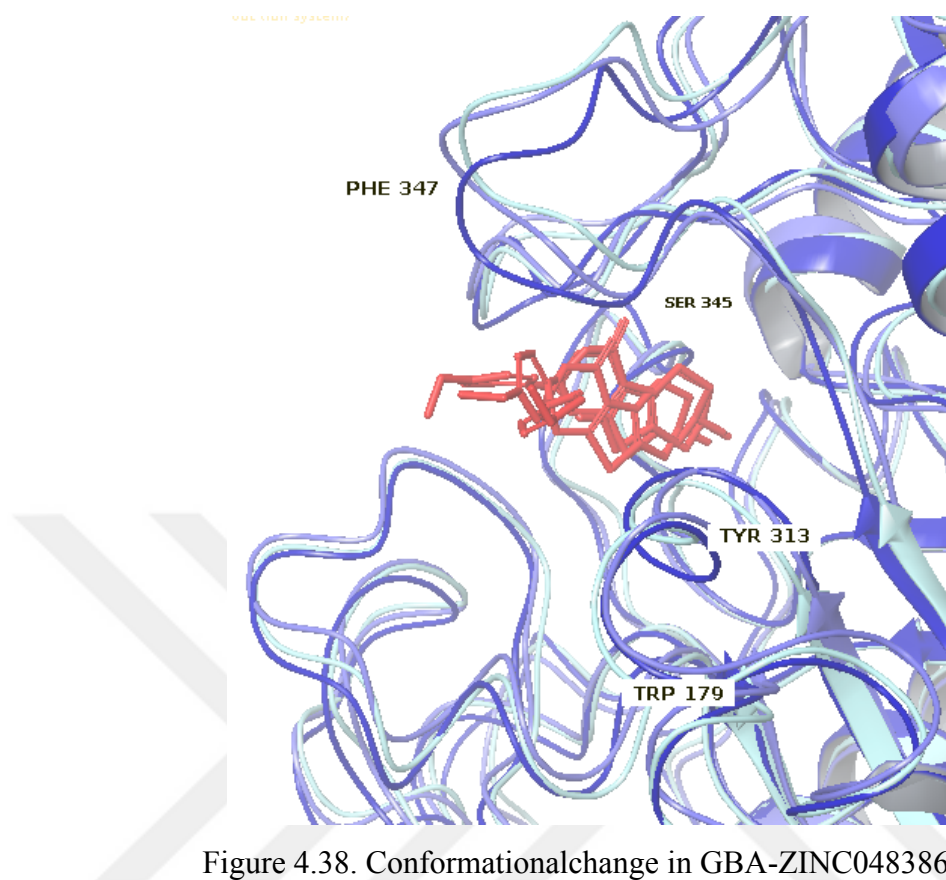


Figure 4.38. Conformational change in GBA-ZINC04838663 complex throughout the simulation. (Initial conformation = Dark blue, Conformation at 15 ns = light blue, Conformation at 40 ns = Turquoise)

Similar to other simulations, the most flexible loop is loop 1 shown in Figure 4.38. It moved away from the active site up to 15 ns, it moved back to its initial position time to time but it did not reach to the same position. The ligand was rigid similar to Zinc00895521 and it mostly interacted with hydrophobic pocket residue Tyr 313. Also the nitrogen atoms in pyrido-thieno-pyridine moiety made strong hydrogen bonds with the active site residues, so that active site residues were stable. Zinc04838663 has an N⁺ atom in the cyclohexane region that may have significantly increased the electrostatic contribution in the binding free energy. In addition, hydrogen bonds with hydrophobic residue, Ser 345 and active site residue Asp 127 enhanced the binding affinity of Zinc04838663. Tetrahydrothiophene moiety made hydrophobic interactions with Tyr 313 and Ser 345. It also made two π - π stacking bonds with Tyr 313 and Phe 246, which may increase stabilization of the ligand.

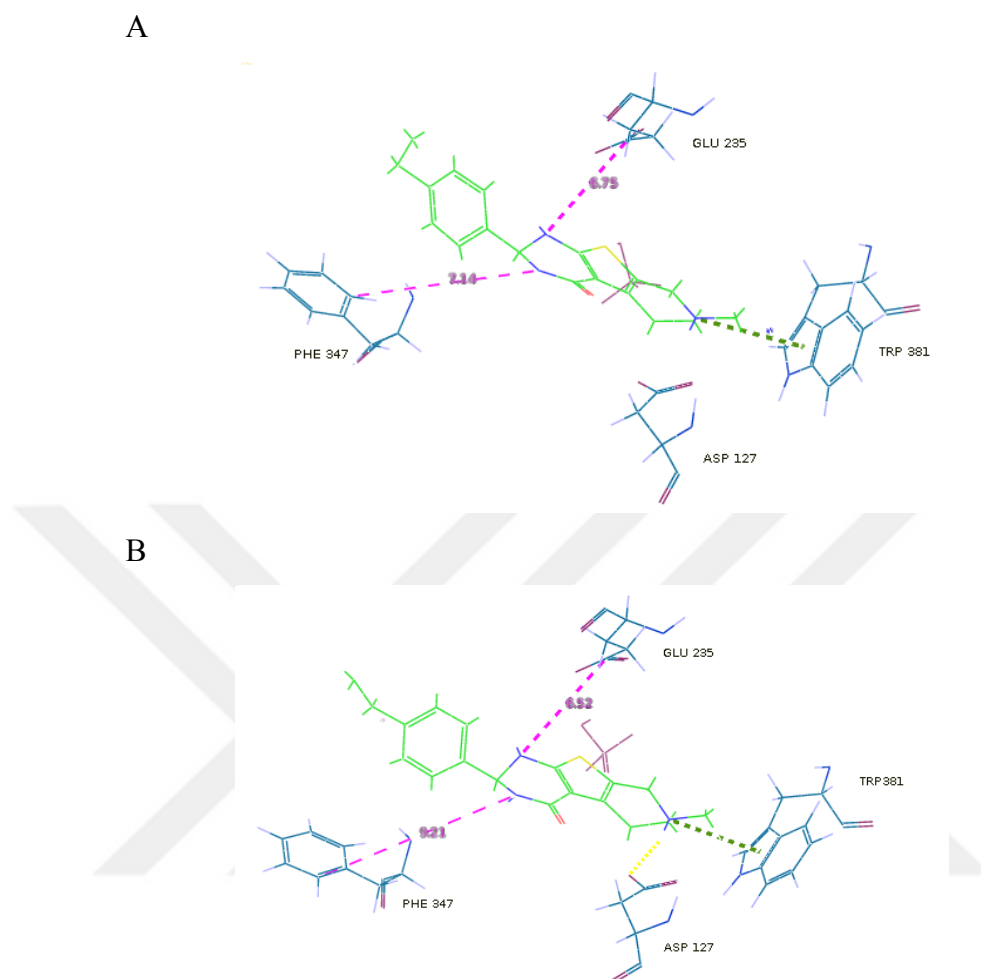


Figure 4.39. The movement of Phe 347 at different time intervals. (A: 25 ns, B:50 ns)

Changes in active site and catalytic residues were further examined. Asp 127 which is one of the most significant residues for ligand binding was not affected by the binding of different ligands. The distance between the crystal structure of GBA and GBA having Zinc04838663 was calculated at different times. At 25 ns, the distance between Asp 127 in crystal structure and Asp 127 in simulated structure was 0.41 Å while at 50 ns, it increased slightly and calculated as 0.78 Å. Movements of catalytic residues were also examined. The distances between Glu 235 in simulated structure and Glu 235 in crystal structure were 6.75 Å and 6.52 at 25 ns and 50 ns, respectively, showing that there was a slight movement of catalytic residues. In addition, the distance between Phe 347 and the ligand started to increase throughout the simulation. At 25 ns, the distance between the nitrogen atom of the ligand and the ring group of Phe 347 was 7.14 Å whereas it became 9.21 Å at 50 ns. This movement can be caused by the pi-pi stacking interaction of Trp 381 and hydrogen

bond between Asp 127 and the ligand as shown in Figure 4.39. Eventually, Zinc04838663-GBA complex was more stable than Zinc00895521-GBA complex. The reason of this can be the strong hydrogen bond between Asp 127 and Zinc04838663 that hold the ligand very closed to the GBA active site.

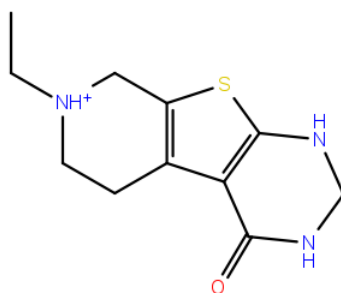


Figure 4.40. Scaffold of Zinc04838663

In the SciFinder, five derivatives of this scaffold are used as target indicators. None of these 5 compounds targets the GBA protein. Their targets are acetylcholine esterase, butylcholine esterase, lactate dehydrogenase, carbonic anhydrase 1,2,9, filamin, protein geranylgeranyltransferase and Ras protein. The functions of different derivatives of tricyclic Pyrido-Thieno-Pyrimidine were tabulated in Table 4.41.

Table 4. 41. Functions of molecules having scaffold of tricyclic pyrido-thieno-pyrimidine

Molecule	Function
Pyrido[4',3':4,5]thieno[2,3- <i>d</i>]pyrimidine-1(2 <i>H</i>)-acetamide, 7-acetyl-3,4,5,6,7,8-hexahydro-2,4-dioxo-3-(2-phenylethyl)-	Anticancer compound which targets RAL GTPase
Pyrido[4',3':4,5]thieno[2,3- <i>d</i>]pyrimidine-1(2 <i>H</i>)-acetamide, 7-acetyl- <i>N</i> -(3-fluoro-4-methylphenyl)-3,4,5,6,7,8-hexahydro-2,4-dioxo-3-(phenylmethyl)-	Antibiotic tolerance inhibitor to treat infections
Imidazo[1,2- <i>a</i>]pyrido[4',3':4,5]thieno[2,3- <i>d</i>]pyrimidin-5(6 <i>H</i>)-one, 2,3,7,8,9,11-hexahydro-8-(phenylmethyl)-	Inhibitor of Acetylcholineesterase
Pyrido[4',3':4,5]thieno[2,3- <i>d</i>]pyrimidine-5-carboxylic acid, 1,2,3,4,5,6,7,8-octahydro-4-oxo-7-(phenylmethyl)-2-thioxo-, methyl ester	Ligand of 5-HT3 receptor
Pyrido[4',3':4,5]thieno[2,3- <i>d</i>]pyrimidin-4(1 <i>H</i>)-one, 2-(chloromethyl)-5,6,7,8-tetrahydro-7-(phenylmethyl)-	Inhibitor of Mcl-1 receptor
Pyrido[4',3':4,5]thieno[2,3- <i>d</i>]pyrimidin-4(1 <i>H</i>)-one, 2-(chloromethyl)-5,6,7,8-tetrahydro-7-(phenylmethyl)-	Treatment of Trypanosoma brucei

Table 4.41. Functions of molecules having scaffold of tricyclic pyrido-thieno-pyrimidine (cont.)

Molecule	Function
Pyrido[4',3':4,5]thieno[2,3- <i>d</i>]pyrimidin-4(3 <i>H</i>)-one, 7-acetyl-3-amino-5,6,7,8-tetrahydro-2-[4-[4-(2-quinolinyl)-1-piperazinyl]butyl]-	Dual 5-HT _{1a} agonists and 5-HT ₃ antagonists for use in treating irritable bowel syndrome
Pyrido[4',3':4,5]thieno[2,3- <i>d</i>]pyrimidine-2,4(1 <i>H</i> ,3 <i>H</i>)-dione, 7-acetyl-5,6,7,8-tetrahydro-3-[2-[4-(2-methoxyphenyl)-1-piperazinyl]ethyl]-	Treatment of cerebral ischemia
Pyrido[4',3':4,5]thieno[2,3- <i>d</i>]pyrimidin-4(3 <i>H</i>)-one, 5,6,7,8-tetrahydro-2-methyl-3-(2-methylphenyl)-7-(phenylmethyl)-, hydrochloride (1:1)	Antiallergic agent

4.5.5. MD Simulations in the presence of Zinc00891180

GBA with Zinc00891180 was also simulated for 50 ns using Desmond. Zinc00891180 is heavier than the other potential ligands and has a greater octanol-water partition coefficient than those of other ligands. Because the GBA protein has hydrophobic pocket, the ligand targeting GBA protein needs to possess hydrophobic properties to dissolve at the surface of the protein. Same simulation conditions used for the simulation of other ligand-protein complex were utilized to observe the stability of Zinc00891180-GBA complex.

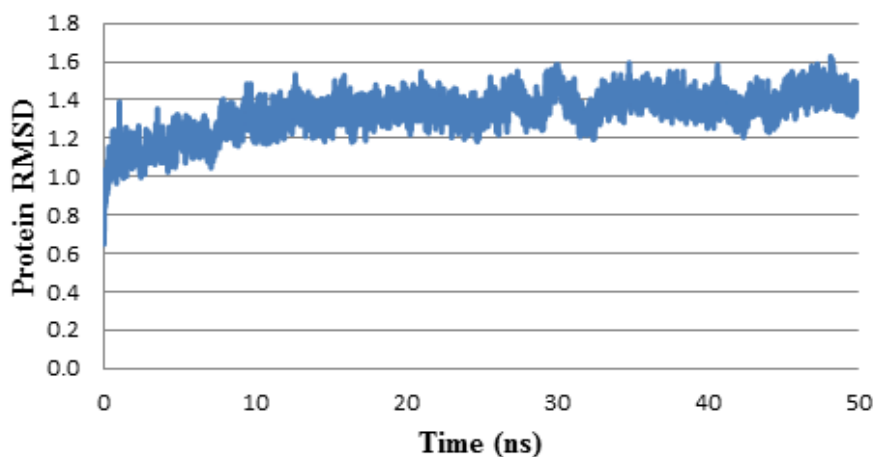


Figure 4.41. RMSD properties of GBA Cα atoms (top panel) and ligand, Zinc00891180, (bottom panel) after superpositioning based on protein Cα coordinates

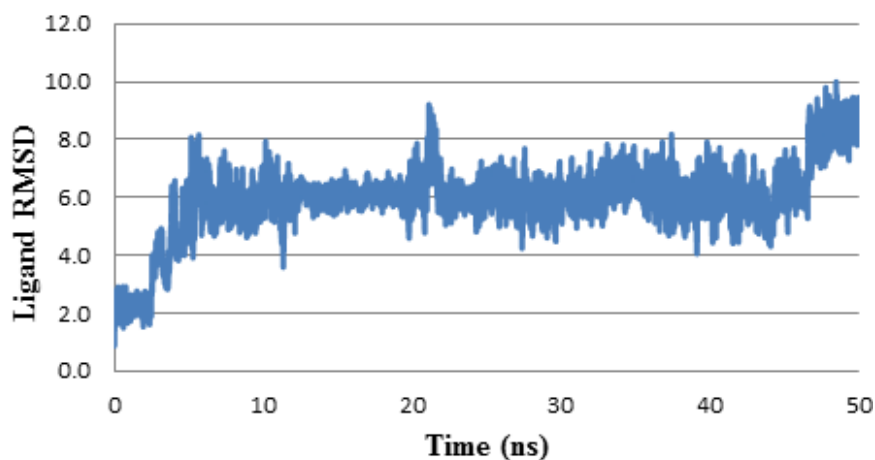


Figure 4.41. RMSD properties of GBA C α atoms (top panel) and ligand, Zinc00891180, (bottom panel) after superpositioning based on protein C α coordinates (cont.)

RMSD value of the protein increased from 0.9 Å to 1.3 Å slightly in 10 ns and reached its equilibrium. The average protein RMSD was about 1.3 Å similar to other RMSD values. However, the ligand was very flexible. The average RMSD value of Zinc00891180 was nearly 6 Å, and this value was significantly higher than those of other ligands. RMSD value of Zinc00891180 increased from 0.4 Å to 6 Å in 5 ns. Besides, there were some high peaks exceeding the average RMSD value throughout the simulation. Moreover, after 47 ns, RMSD value of the ligand further enhanced and changed between 8-9 Å, but the ligand did not separate from the active site of the protein. The reason of this movement can be the immediate change in position of the head group consisting of one cyclopentane and cyclohexane. Suddenly, the head of the ligand turned towards downward leading to deviation in RMSD.

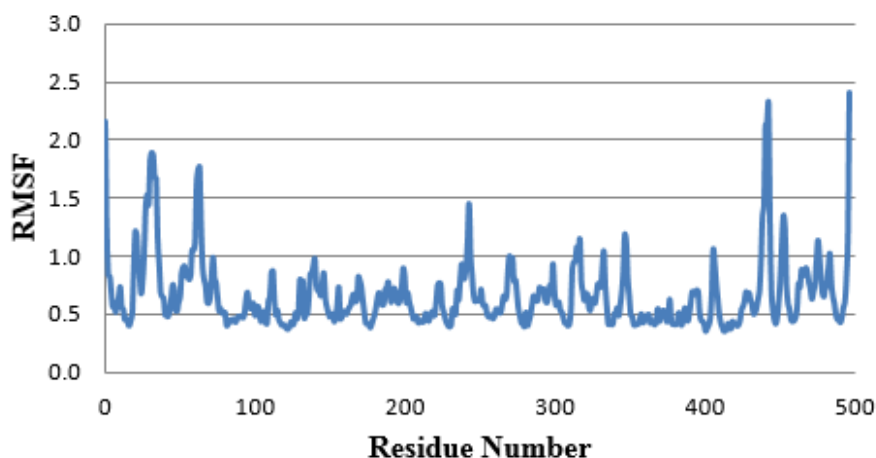


Figure 4.42. Average RMSF values of residues throughout the simulation.

In Figure 4.42, the region of Ser439-Leu444 was highly flexible. Average RMSF value of this region changed between 1.34-2.31 Å. The most mutated residue, Leu 444 is located in that part of GBA and it is known that it can be stabilized by glycosylation of Asn 19 [63]. Other residue, Asn 370 that is exposed to mutation was more rigid than Leu 444, and its RMSF value was measured as 0.437 Å. Besides, residues 26-36 is located in between disulfide bridge and beta strand, and their RMSF value were also higher than those of other parts of the protein.

Zinc00891180 interacted mostly with Residues of Phe 246 and Phe 347. Although H bond with Ser 345 disappeared after the simulation has started, H bond with Phe 347 was maintained until 50 ns. Similarly, despite the bond disappeared time to time, Pi-Pi stacking bond with Phe 246 was kept until the end of the simulation. In the detailed analysis, the distances between the catalytic residues and the nitrogen atom of the main group of the ligand were estimated in order to determine movement of the ligand. As a result, the distance between Phe 347 and N-atom declined from 7.04 Å to 6.25 Å slightly whereas the gap between catalytic residues Glu235 and Glu340 increased almost 4 Å dramatically, and similar tendency was observed in the simulation of GBA with Zinc04838663. Besides, loop 2 consisting of Residues 393-396 was stable than other loops, and the head group of the ligand moved away from the loops residues of 312-319 and 240-250.

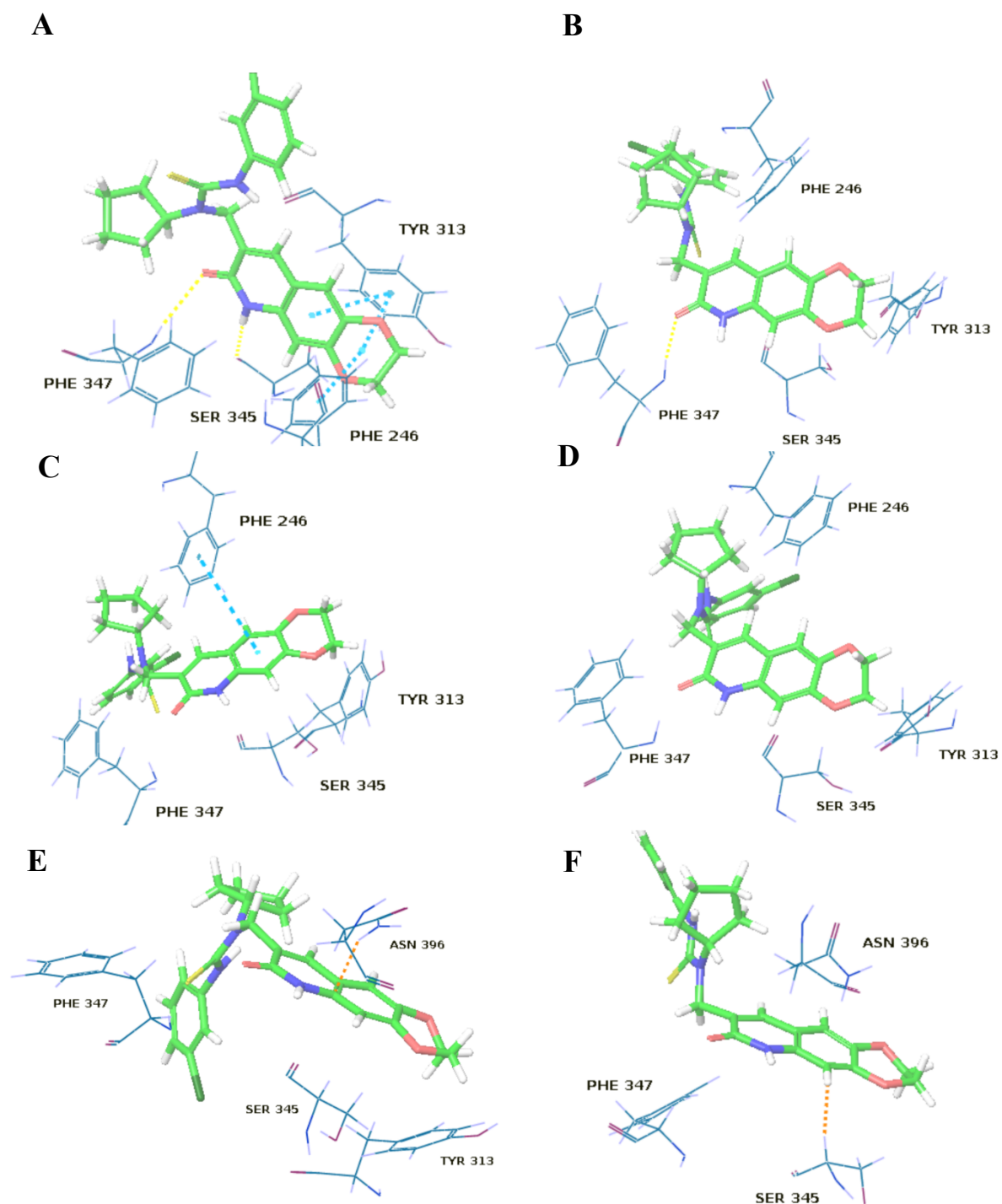


Figure 4. 43. Conformation of Zinc00891180 throughout the simulation (A: conformation at 0 ns, B: conformation at 10 ns, C: conformation at 20 ns, D: conformation at 30 ns, E: conformation at 40 ns, F: conformation at 50 ns) (Hydrogen bond: Yellow dashed line, Pi-Pi stacking bond: Turquoise dashed line, Favorable Contact: Orange dashed line)

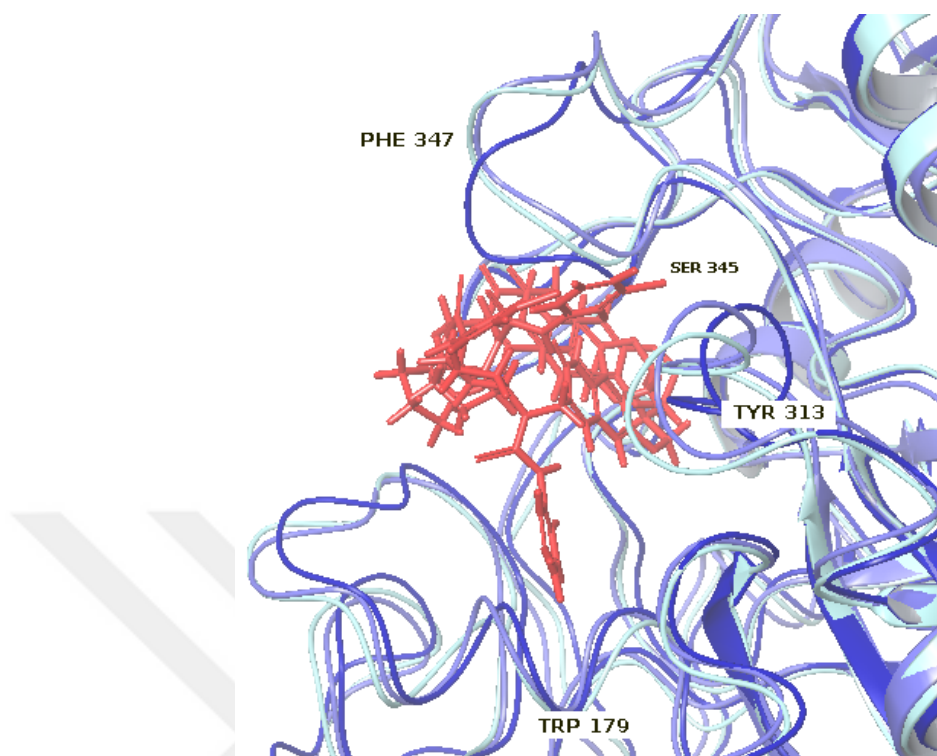


Figure 4.44. Conformational change in GBA-ZINC00891180 complex throughout the simulation. (Initial conformation = Dark blue, Conformation at 15 ns= light blue, Conformation at 40 ns= Turquoise)

As shown in the Figure 4.44, head group of the ligand was very flexible, the ring moiety in head group changed its conformation time to time and this change led to increase in RMSD value. However, head group, Dioxino-quinolione includes nitrogen and oxygen atoms that contribute ligand binding by making hydrogen bonds. The major change in conformation occurred in 15 ns. Loop 1 moved away from the ligand. Phe 347 in loop 1 deviated from its initial position about 4 Å. Similarly, hydrophobic pocket residue Tyr 313 moved about 2 Å and loop 3 moved away from the active site. With regards to active site residues, Trp 179, Glu 235 and Glu 340 were very stable during the simulation. Dioxino-quinolione moiety of the ligand made hydrophobic interactions with Tyr 313 and Ser 345. Also, it made two π - π stacking bonds with Phe 246 and Tyr 313, which enhanced ligand stabilization. Moreover, hydrogen bond between Phe 347 and oxygen atom in dioxino-quinolione moiety increased ligand binding affinity to GBA.

Zinc00891180 has high molecular weight with respect to other ligands, but the main group involving three ring groups is very stable at the active site of the GBA showing that this scaffold can be used to obtain different compounds by replacing the head group with lighter ones[64]. In one study, potential compounds were found against CB2 agonist using high throughput screening. 1,4-diazepane compound was found from the study, but it had poor drug-like parameters such as cLogP of 5.2, low microsomal stability and high molecular weight of nearly 470 amu. The compound was used as base and its scaffold were used to retrieve new molecules by changing alkyl substituents in order to obtain compound with better drug-like parameters[65].

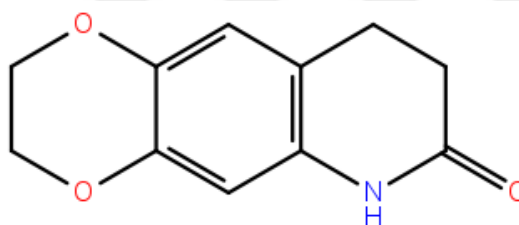


Figure 4.45. Scaffold of Zinc00891180

As a result of search in SciFinder, 139 substances were determined while only 44 of these substances are used as target indicators. These 44 substance target erythropoietin, androgen receptors and checkpoint protein kinase 1.

Table 4.42. Functions of compounds with dioxino-quinolione

Compound	Function
1,4-Dioxino[2,3-g]quinolin-7(6H)-one, 2,3,8,9-tetrahydro-9-(2-thienyl)-	Positive allosteric modulator for melanocortin receptors
1,4-Dioxino[2,3-g]quinolin-7(6H)-one, 9-(2-chlorophenyl)-2,3,8,9-tetrahydro-	MET receptor targeting antimetastatic compound for treatment and prevention of cancer
4H-[1,4]Dioxino[2,3-g]pyrazolo[4,3-c]quinolin-4-one, 5-(3-aminopropyl)-2,5,8,9-tetrahydro-3-methyl-	Inhibitors of checkpoint kinase CHK1 for use in cancer treatment

4.5.6. MD Simulations in the presence of Zinc35092294

Chemical structure of Zinc35092294 is different than the previous ligands studied. It has a five-membered ring with two nitrogen atoms and six-membered ring with an oxygen atom similar to NN-DNJ. It is lighter than the other ligands such as Zinc00895521 and its octanol-water partition coefficient is lower. It was analyzed with MD simulation for 50 ns and GBA-Zinc35092294 complex was maintained at the end of the simulation. The RMSD results are given in Figure 4.46.

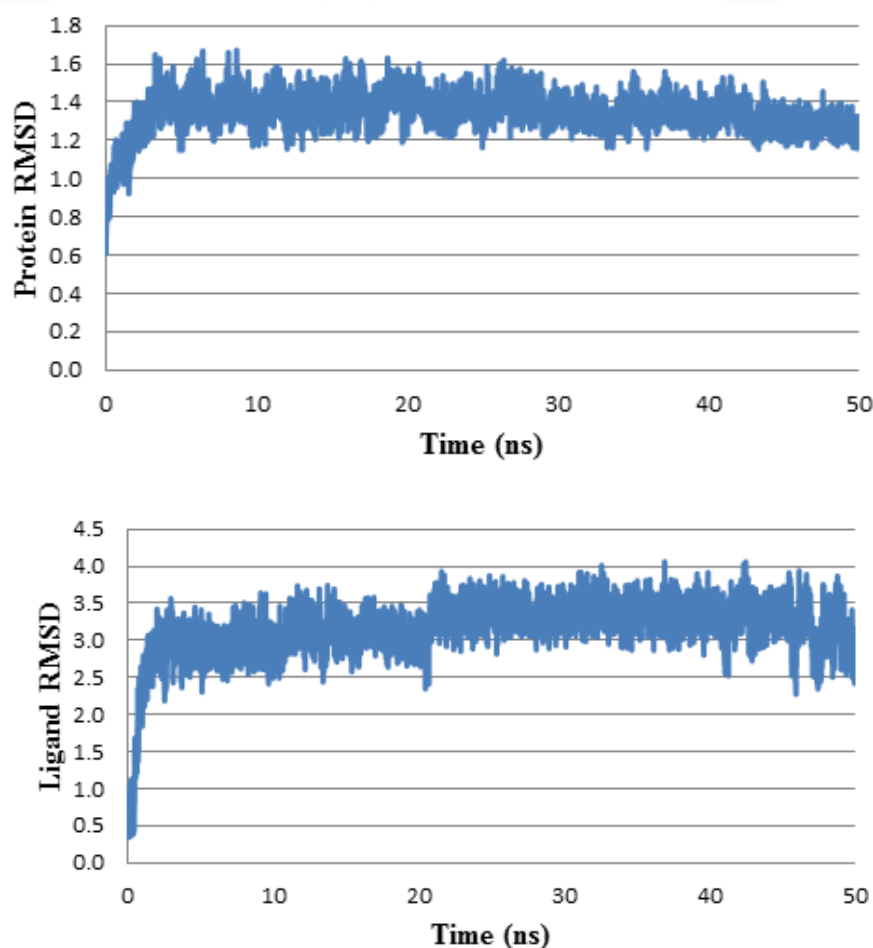


Figure 4.46. RMSD properties of GBA C_{α} atoms (top panel) and the ligand, Zinc35092294 after superpositioning based on protein C_{α} coordinates.

The average RMSD of GBA with Zinc35092294 was about 1.30 Å as shown in Figure 4.46. With respect to ligand, the average RMSD was about 3.16 Å. At about 20 ns

sudden increase occurred in RMSD of the ligand. The reason of this was the movement of tetrahydropyran moiety, but the ligand stayed bound to GBA until the end of simulation.

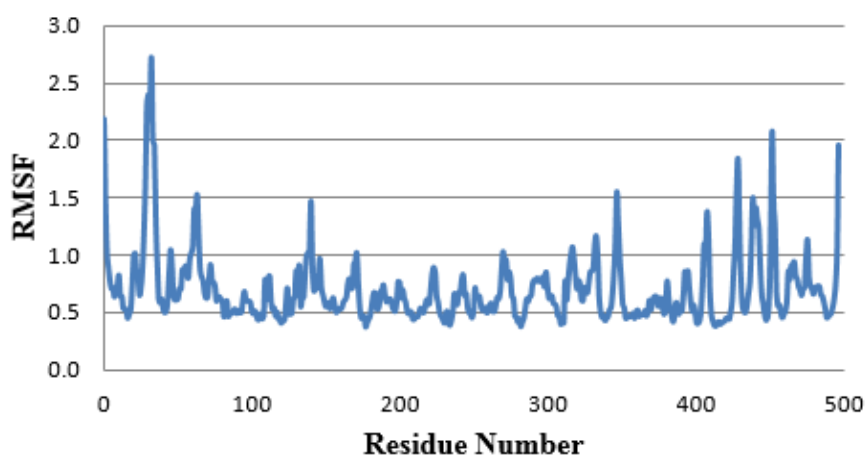


Figure 4.47. Average RMSF values of each residues in GBA throughout the simulation

RMSF values of all active site residues were below 1 Å as shown in Figure 4.47. Catalytic residues Glu 235 and Glu 340 had RMSF values of 0.42 Å and 0.43 Å respectively, these values were similar to other simulation results. Residues having high tendency to mutation are also stationary. Leu 444 had RMSF value of 0.90 Å and RMSF of Asn 370 was about 0.52 Å. There was not any large fluctuation or unexpected result in RMSF results.

Zinc35092294 made two stable hydrogen bonds with catalytic residue, Glu 340 throughout the simulation. The distance between the Glu 340 and the ligand did not change during simulation and remained around 2.7 Å. In addition, hydrogen bond occurred between the active site residue, Asn 234 and oxygen atom of Zinc35092294 at 10 ns and maintained until the end of the simulation while hydrogen bond with another active site residue, Trp 179 disappeared after 10 ns. The loops around the active site did not make any significant movement in 50 ns. However, the distance between Phe 347 in loop 1 and the ligand increased from 7.08 Å to 12.04 Å in 25 ns, but it started to decline after 25 ns and measured as 8.31 Å at the end of simulation. Similar to previous simulations, Phe 347 was the most mobile residue around the active site. The distance between the other residues such as Asn 396, Tyr 313, Glu 340 and the ligand changed around 2 Å.

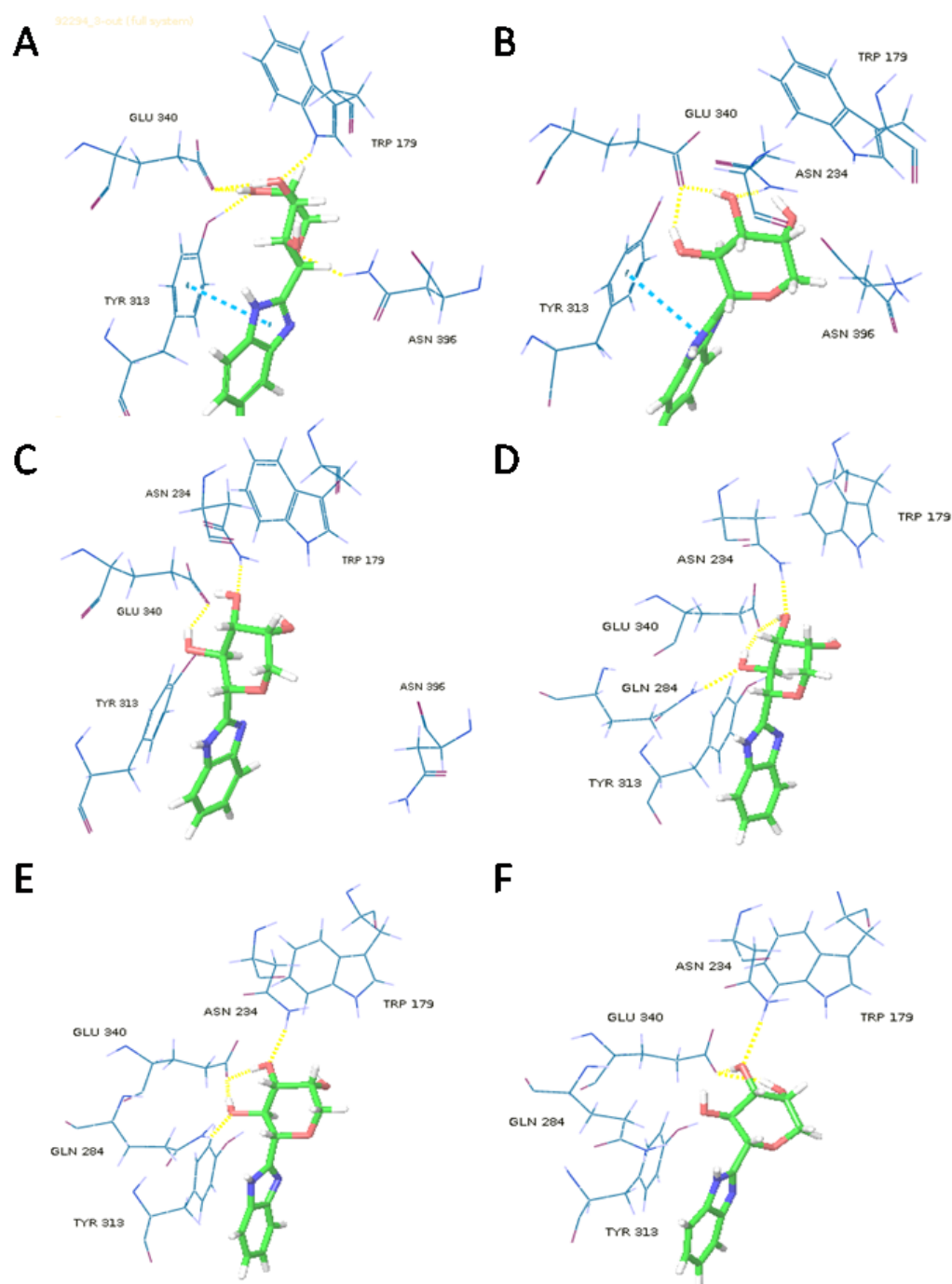


Figure 4. 48. Conformations of Zinc35092294 throughout the MD Simulation (A: conformation at 0 ns, B: conformation at 10 ns, C: conformation at 20 ns, D: conformation at 30 ns, E: conformation at 40 ns, F: conformation at 50 ns) (Hydrogen bond: Yellow dashed line, Pi-Pi stacking bond: Turquoise dashed line)

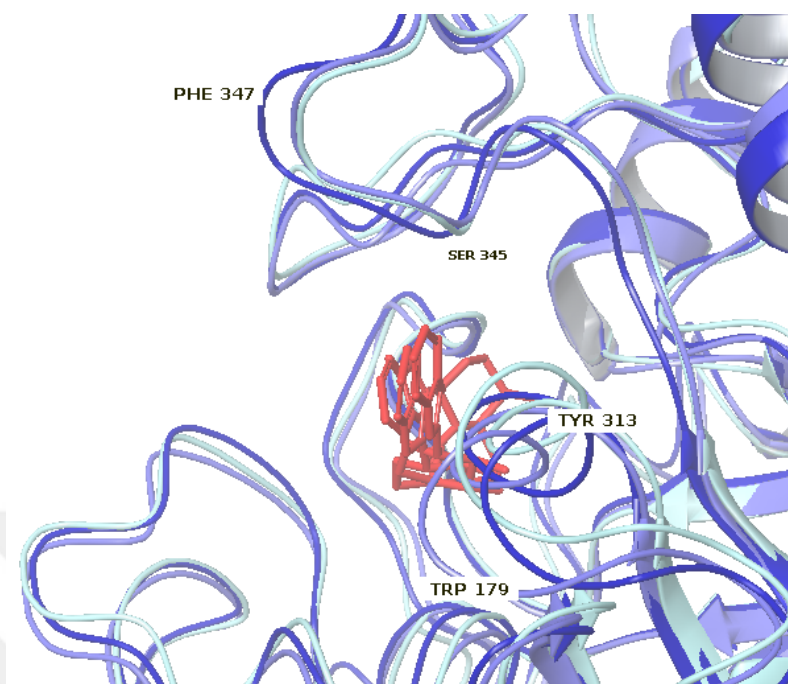


Figure 4. 49. Conformational change in GBA-ZINC35092294 complex throughout the simulation. (Initial conformation = Dark blue, Conformation at 15 ns= light blue, Conformation at 40 ns= Turquoise)

As displayed in Figure 4.49, Zinc35092294 was very rigid at the active site. In head group, it has cyclohexane moiety with three hydroxyl atom similar to that of NN-DNJ. This part of the ligand interacted with the active site residues such as Trp 179 and Glu 340. On the other hand, pyrimidine moiety of the ligand had π - π stacking interactions with the hydrophobic pocket, especially with Tyr 313. These interactions are very similar to those of NN-DNJ. Both loop 1 and loop 3 were flexible for 15 ns. Phe 347 in loop 1 moved away from the ligand in the same way of other simulations whereas Phe 317 in loop 3 got close to ligand about 3 Å. With regards to loop 2, it was more stable than the other loops and there was no further significant movement throughout the simulation. Benzene-purine moiety made hydrophobic interactions with Tyr 313 while sugar moiety made hydrogen bonds with Asn 234 and Glu 340 which made great contributions on ligand binding.

4.5.7. MD Simulations in the presence of Zinc0017874

Insertion of fluorine into a molecule develops the properties of compound. The potency, membrane permeability increase with addition of fluorine while pKa and clearance decreases. In addition, fluor has many advantages like having ability to emit positron[66]. Zinc0017874 has three fluorine atom attached to benzene ring. It was analyzed in order to understand the effect of fluorine compound on GBA. MD simulation was started for 50 ns, but all bonds between the ligand and GBA broke at 9 ns and the ligand moved away from GBA. The RMSD analysis of MD simulation is given in Figure 4.50.

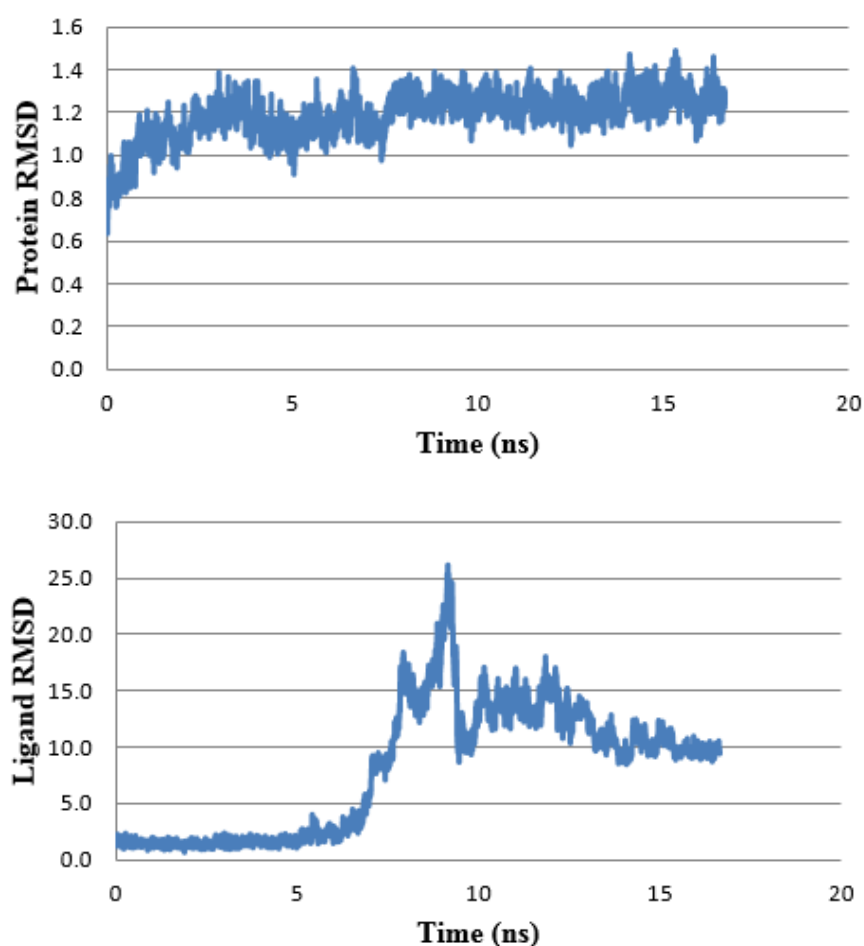


Figure 4.50. RMSD properties of GBA C_α(top panel) and ligand Zinc00017874 (bottom panel) after superpositioning based on protein.

Figure 4.50 showed that the average RMSD value of protein was about 1.14 Å. Nevertheless the average RMSD of the ligand fluctuated dramatically during the simulation. At 8 ns, the ligand left the GBA, and RMSD value raised suddenly up to 15 ns. The ligand floated for a while and moved away from GBA up to 12 ns, then it started to approach to the receptor, but it could not bind to GBA again.

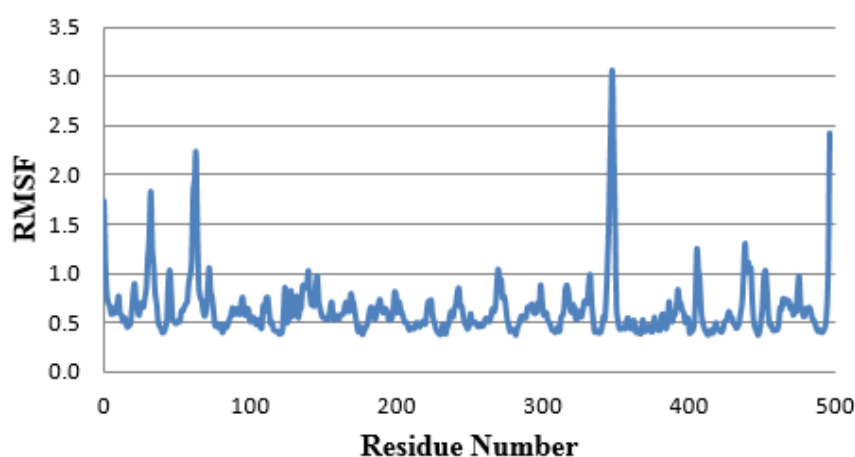


Figure 4.51. Average RMSF values of GBA in the presence of Zinc00017874

The RMSF values of GBA bound to Zinc00017874 was higher than those of GBA bound to previous ligands. Especially the region of Residues 346-350 forming loop 1 was very flexible. With regards to active site residues, their RMSF values were not so high. For instance, Trp 381 had RMSF value of 0.40 Å, and Asp 127 had 0.55 Å. Catalytic residues Glu 340 and Glu 235 had RMSF values of 0.38 Å and 0.46 Å respectively. So that, the separation between the receptor and the ligand may be caused by the flexibility of loop 1.

Interaction analysis displayed in Figure 4.52 implied that the ligand made three hydrogen bonds with active site residues Asn 234, Glu 340 and Ser 345. However, these bonds were not strong enough and broke down. At 5 ns, the ligand interacted with another two active site residues Trp 179 and Asp 127 implying that the ligand started to move away from the initial position. Despite of the interactions with active site residues, the ligand could not stay bound to protein and left GBA.

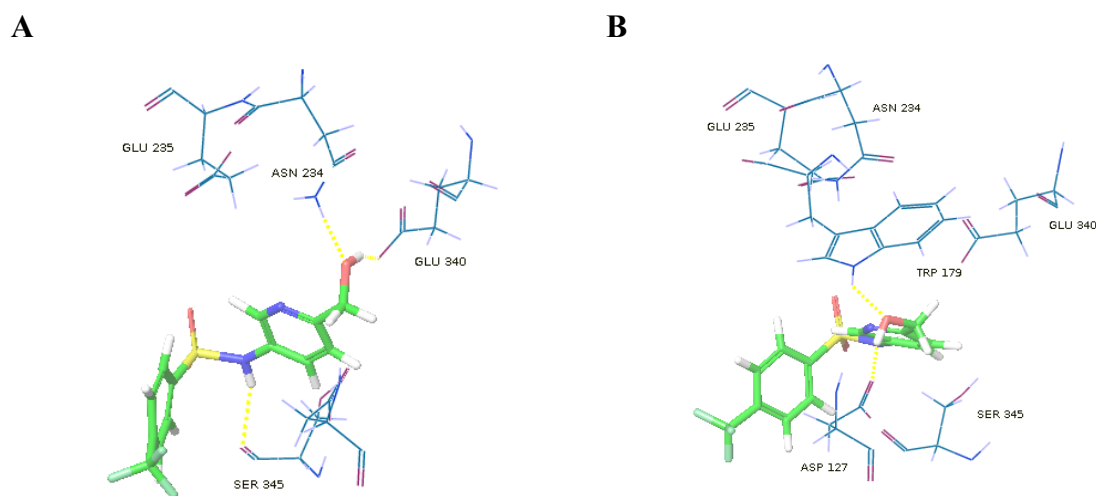


Figure 4.52. Conformations of Zinc00017874 throughout the MD simulation
(A: conformation at 0 ns and B: conformation at 5 ns)

GBA-Zinc00017874 complex at 5 ns and 10 ns were superimposed to initial structure to understand the reason of separation of the ligand from GBA active site. As shown in Figure 4.53, the ligand was very rigid up to 5 ns, the hydroxyl and two nitrogen atoms of Zinc0001787 made hydrogen bonds with Trp 179, Ser 345 and Asp 127 and its binding type was similar to those of NN-DNJ whose sugar moiety made three hydrogen bonds with the active site residues. Separation may be caused by the movement of loop 1. Ser 345 in Loop 1 changed its position about 3 Å and the hydrogen bond between the ligand and GBA broke down. Also Phe 347 in Loop 1 was mobile during the simulation, and the average RMSF of Phe 347 exceeded 3 Å. N atom made hydrophobic interaction with Ser 345. Also hydrogen bond occurred between hydroxyl atom and Asn 234 and Trp 179 at different time steps. Tail group consisting of benzene ring with trifluoromethyl did not make any contribution on the ligand binding by interacting with either hydrophobic pocket or active site residues. Because of that reason, ligand could not be kept bound to the GBA active site.

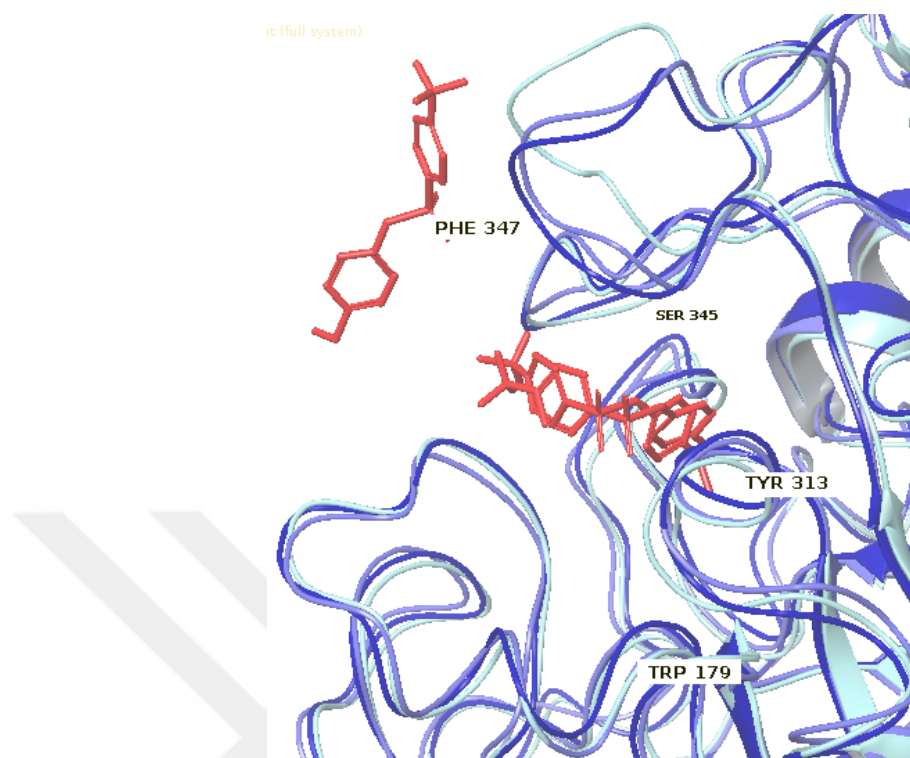


Figure 4.53. Conformational change in GBA-ZINC00017874 complex throughout the simulation. (Initial conformation = Dark blue, Conformation at 5 ns= light blue, Conformation at 10 ns=Turquoise)

4.5.8. MD Simulations in the presence of Zinc01726382

The best compound from structure-based hypothesis is Zinc01726382. The induced fit docking score of this compound is higher than other compounds obtained from the same hypothesis. The stability of this molecule was analyzed with MD simulation. Generally, having a lot of aromatic rings leads to mobility and makes hard to stay bound to the receptor. Nevertheless, although Zinc01726382 has a great deal of aromatic rings, it did not leave the receptor until the end of simulation, and GBA-Zinc01726382 was more stable than other GBA-inhibitor complexes. RMSD properties of the ligand and the receptor were given in Figure 4.54.

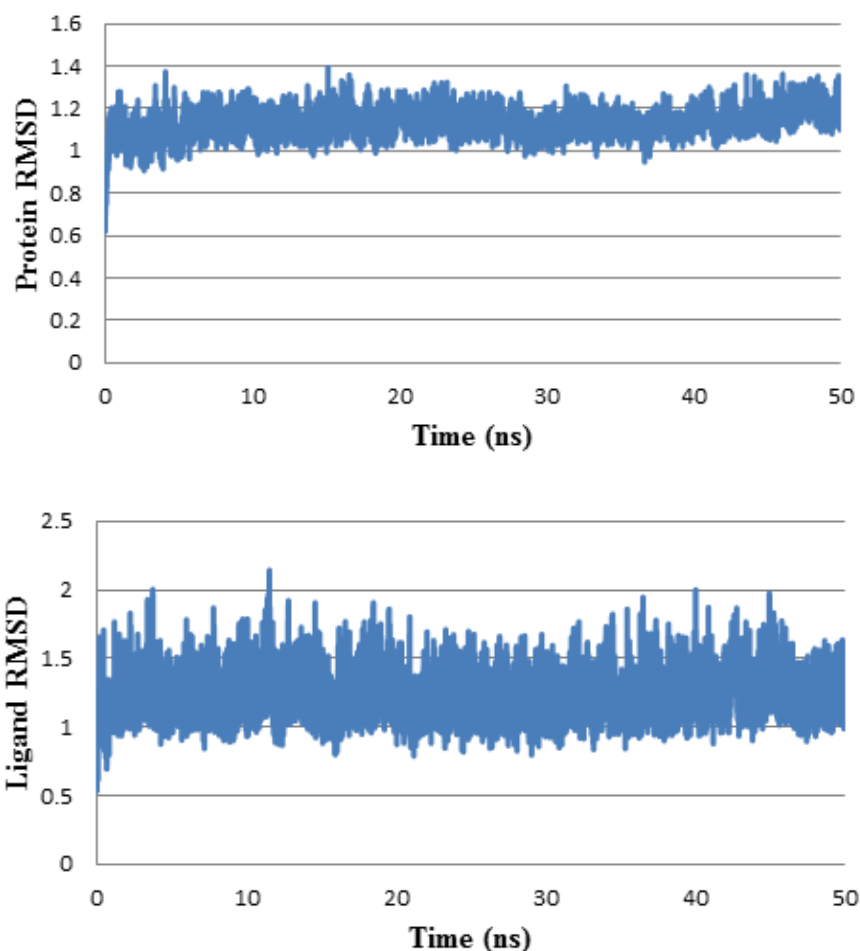


Figure 4.54. RMSD properties of GBA C_{α} (top panel) and ligand Zinc01726382 (bottom panel) after superpositioning based on protein.

The average RMSD of GBA was 1.09 Å while the average RMSD of the ligand was 1.22 Å. According to RMSD results, Zinc01726382 had the lowest average RMSD despite having aromatic rings which do not make any hydrogen bonds with GBA. Furthermore, the profile of RMSD values for both GBA and the ligand is very stable, there was not any sudden increase or decrease in RMSD values implying that GBA-Zinc01726382 complex was very rigid.

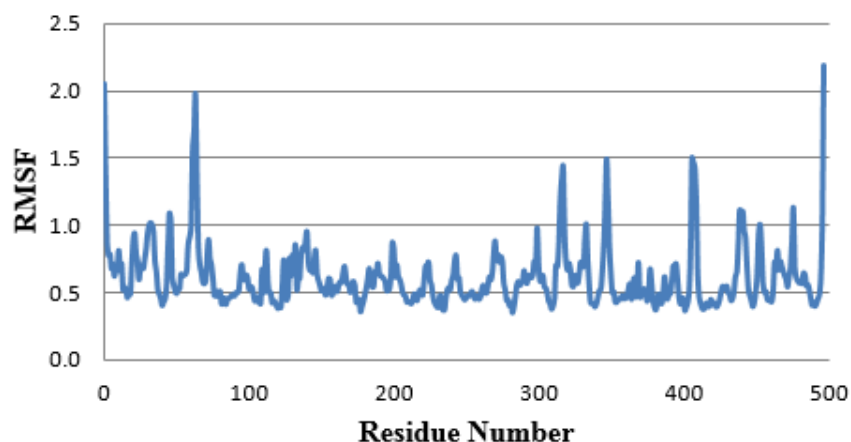


Figure 4.55. Average RMSF values of residues of GBA

Average RMSF values of each residue were analyzed in order to determine the flexible residues. Catalytic residues Glu 340 and Glu 235 have RMSF values of 0.37 Å and 0.35 Å respectively. In addition, RMSF values of active site residues and residues having great tendency to mutate were also below 1, for instance Leu 444 and Asn 370 had RMSF values of 0.69 Å and 0.46 Å respectively. The interactions between the receptor and the ligand were also analyzed by taking the snapshots from the simulation.

At the beginning of the simulation, the ligand made two hydrogen bonds with Asp 127 and catalytic residue, Glu 235. However, the hydrogen bond between Glu 235 and Asp 127 was not strong enough, and broke down whereas hydrogen bond with Asp 127 remained until 50 ns. In contrast to other ligands, pi-pi stacking bond did not occur between the protein and the ligand. Besides, aromatic ring groups did not make any hydrogen bond with GBA, as well as they were very rigid with respect to those of other ligands.

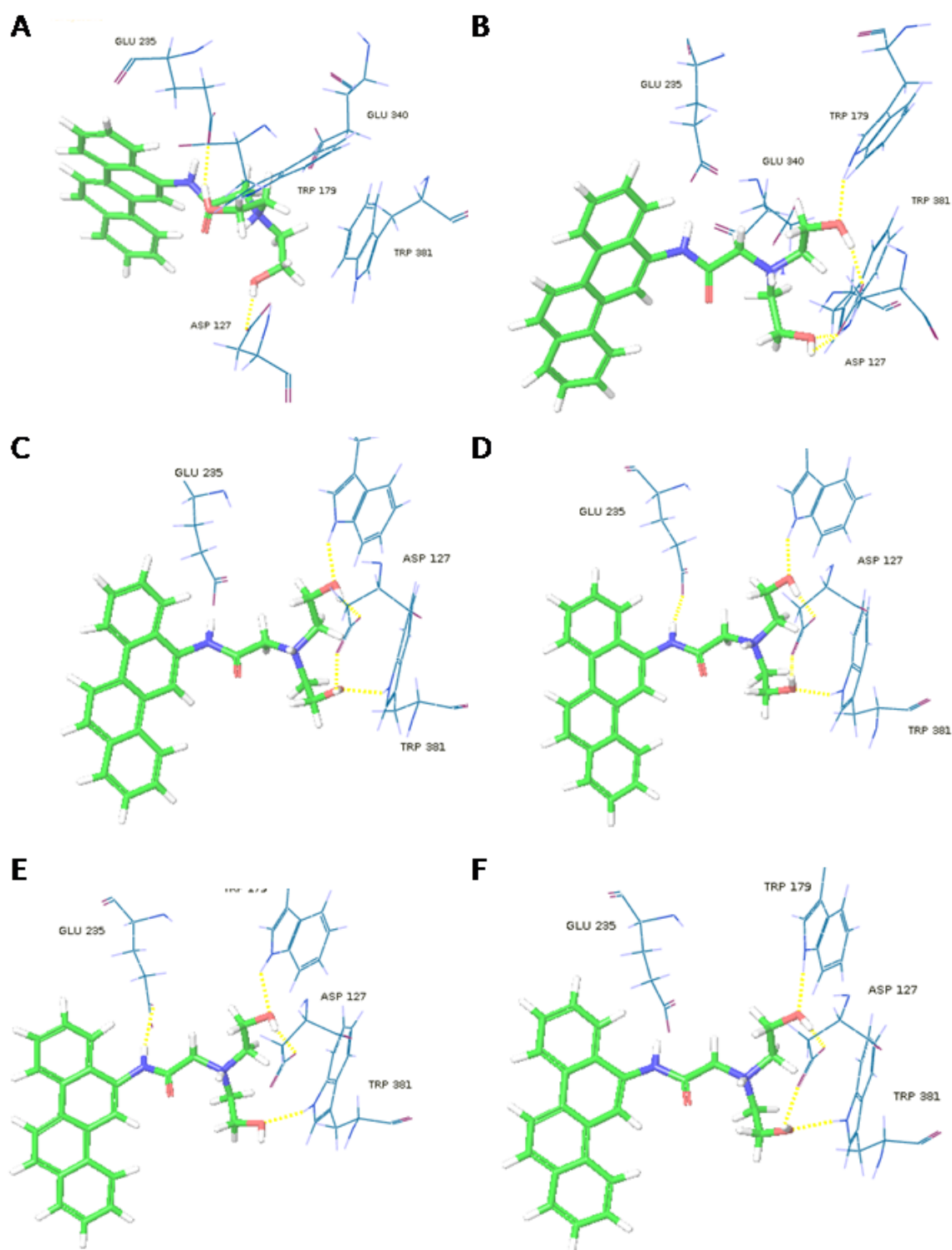


Figure 4.56. Conformations of Zinc01726382 throughout the MD simulation (A: conformation at 0 ns, B: conformation at 10 ns, C:conformation at 20 ns, D:conformation at 30 ns, E: conformation at 40 ns, F:conformation at 50 ns) (Hydrogen bond: Yellow dashed line)

Distances between residues were also analyzed in order to determine the movement of the protein. The distance between the ligand and residues Glu 235, Trp 381 and Asp 127 did not change more than 1 Å. Nevertheless, similar to other simulations, the distance between the ligand and Phe 347 decreased significantly during the simulation. At the beginning of the simulation, the gap between Phe 347 and the ligand was measured as 14.94 Å, then at 25 ns it was measured as 13.74, consequently it decreased to 10.17 Å at 50 ns. A study revealed that aminoacid region 347-422 are critical aminoacids for targeting the lysosome and binding of Limp2, receptor for trafficking of GBA to lysosome. Nevertheless, aminoacid region between 347-398 and 411-422 are internal to the structure, and possibility to interact with Limp 2 was very low[67]. The increase in distance between ligand and Phe 347 can be correlated with these findings.

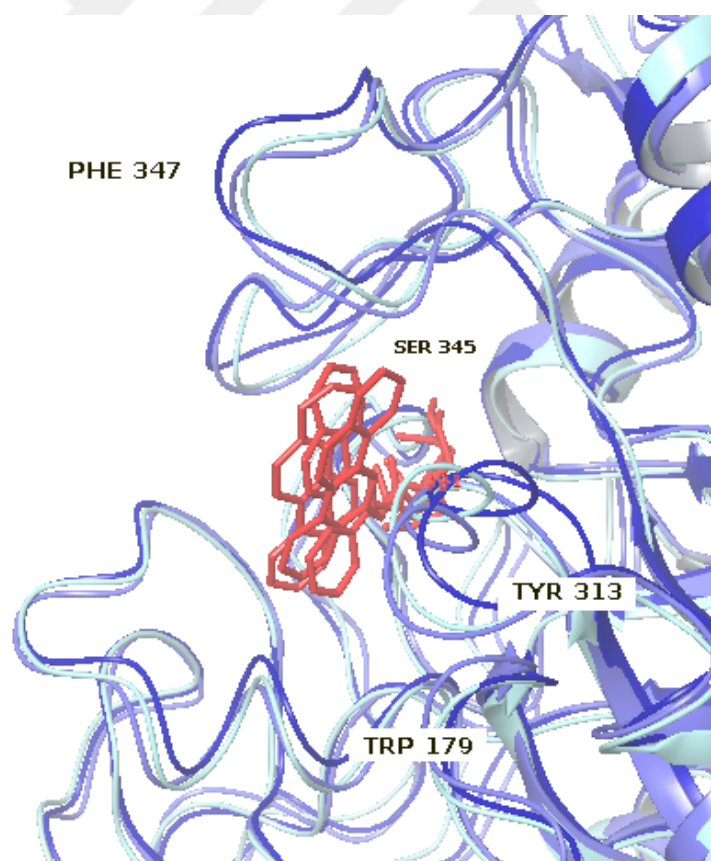


Figure 4.57. Conformational change in GBA-ZINC01726382 complex throughout the simulation. (Initial conformation = Dark blue, Conformation at 15 ns = light blue, Conformation at 40 ns = Turquoise)

Zinc01726382 is the most rigid ligand among all proposed compounds as shown in RMSD values and superposition of various conformations that were taken from different time steps of the simulation. In the head group, it included two oxygen and one nitrogen atom that made hydrogen bonds with the active site residues, while the tail group consisted of large rings which interacted with the hydrophobic pocket. Loop 2 and the active sites were extremely stable during the simulation while loop 3 was more flexible. Loop 3 did not move away or towards the ligand, but it changed its position during the simulation. Similar to other simulations, most flexible loop was loop 1. Ser 345 in loop 1 moved towards the ligand about 2 Å whereas Phe 347 in loop 1 moved away from the active site. Moreover, although the tail group of the ligand stretched out through the active site, it did not affect the ligand binding. Both the well-known inhibitor, NN-DNJ and the natural ligand, glycosylceramide, of GBA have a long tail groups consisting of C-10 and C-18 alkyl chains, respectively. Zinc01726382 has an N⁺ atom, which is attached to dihydronaphthalene moiety, and may contribute to binding free energy in terms of electrostatic interactions. Also nitrogen and hydroxyl atoms of Zinc01726382 made many hydrogen bonds with active site residues Trp 179, Asp 129 and catalytic residues Glu 340 and Glu 235 at different time steps of the simulation and these interactions play significant roles on ligand binding affinity.

Average binding free energy value of each compound was calculated by taking the arithmetic average of frames in 104 frame interval after compounds reached to equilibrium as shown in Table 4.43. Zinc00017874 had the worst energy as expected because it left the GBA active site at about 7 ns while Zinc00895521 has the best binding free energy. In addition Zinc00017874 had positive coulomb energy implying that the force was repulsive, also its solvation energy was negative showing that the solvation mechanism was exothermic and the solution was less soluble at high temperatures. Also, Zinc00895521 and Zinc01726382 had better binding free energy than that of NN-DNJ showing that their bindings to GBA were more favourable than NN-DNJ.

Table 4.43. Binding free energy of NN-DNJ and proposed compounds for whole simulation.

Compound	Time to start calculation (ns)	Coul	Covalent	vdW	Lipo	H-bond	Packing	Solv_GB	Avg. Free Energy (kcal/mol)
NN-DNJ	2	-52.45	1.47	-34.74	-30.49	-3.76	0	48.56	-71.41
Zinc00895521	5	-5.73	2.84	-57.14	-42.74	-0.86	-1.65	4.13	-101.164
Zinc04838663	5	-55.44	2.88	-44.89	-30.73	-1.28	-2.44	64.89	-68.014
Zinc00891180	5	-7.42	2.58	-46.56	-32.42	-0.06	-2.02	18.21	-66.775
Zinc35092294	5	-17.18	1.92	-26.39	-15.98	-2.40	-3.77	22.65	-41.233
Zinc00017874	3	24.37	0.71	-22.21	-13.56	-0.32	-2.614	-17.37	-31.563
Zinc01726382	0	-61.76	2.84	-43.47	-32.18	-3.08	-1.01	56.61	-82.058

4.6 Glycosylation-Stability Analysis of Proposed Compounds

Glycosylation is essential reaction for GBA. However, a study revealed that glycosylation does not affect the GBA activity, it influences the protein structure and flexibility. Also, it was found that Asn 19 glycosylation in GBA leads to stabilization around aminoacid region 438-444 where the severe mutation, L444P, occurs. In addition, Asn 59, Asn 146, Asn 270-glycosylated GBA structures were also examined, as a result, it was found that they affected different regions of the protein. For instance, glycosylation of Asn 59 reduced mobility around residues 49-59 and 16-156. Stabilizing activity analysis showed that two or more oligosaccharides reduce the mobility around residues 442-462. In addition at least one oligosaccharide with Asn-19 glycosylation decline the mobility of residues between 215 and 235. So that, it was found that Asn 19 glycosylation is very crucial for stability [63]. Besides, despite the phosphorylation occurs during the transfer of lysosomal enzymes to lysosome, it was found that GBA does not undergo phosphorylation [68].

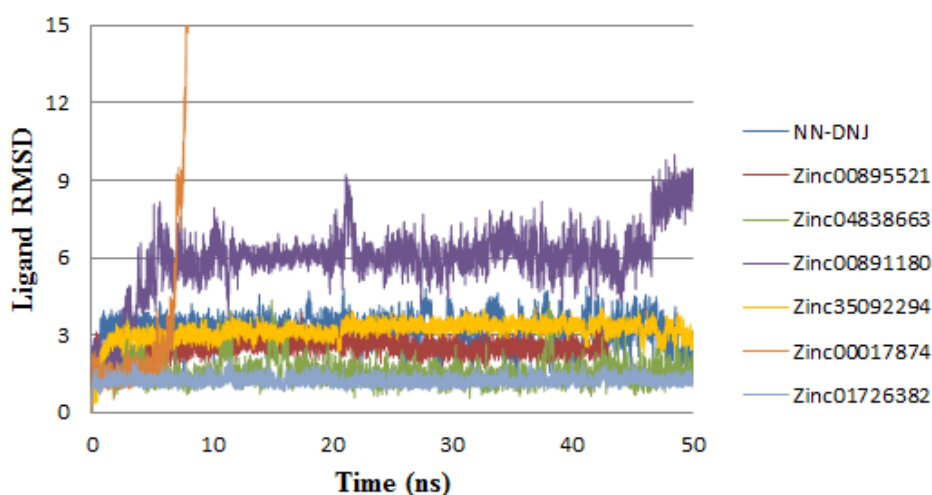


Figure 4.58. RMSD values of ligands bound to GBA active site through MD simulation.

Figure 4.58 summarized the RMSD values of selected compounds bound to GBA. Zinc04838663 and Zinc01726382 had the lowest RMSD values implying that their flexibilities were very low and they seemed to be stable. In contrast, the most mobile compound was Zinc00017874 and it left protein around 10 ns.

Average RMSF values of each protein-ligand complex were plotted as shown in Figure 4.59. One of the largest fluctuations in RMSF value occurred while GBA was bound to Zinc00017874 and that flexibility leads to break down of protein-ligand complex.

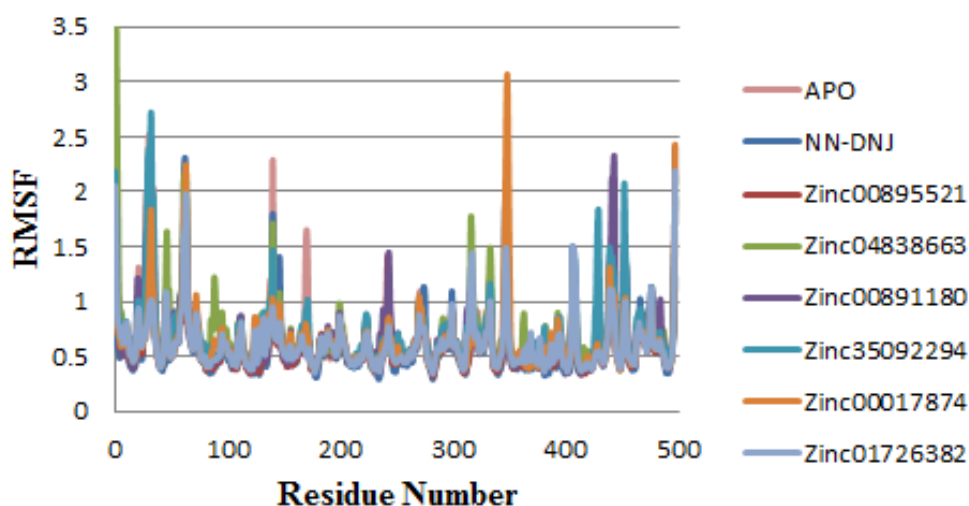


Figure 4.59. RMSF values of residues in the presence of different ligands

Each of compounds simulated with Desmond were compared in terms of their overall protein stabilities, RMSF values of important mutated residues such as Leu 444 and Asn 370, and RMSF values of glycosylated residue Asn 19. Overall protein RMSD and RMSF values of mutated residues were shown in Table 4.44. The table displayed that binding of compound 5 and compound 7 stabilized the GBA like binding of NN-DNJ, because its overall RMSD value is lower than those of Apo-GBA structure. In addition the most stable protein-ligand complex is Zinc01726382-GBA and Leu 444 was more flexible than other mutated residues while glycosylated residue, Asn 19, was also more stable than Leu 444. Leu 444 was more rigid in the presence of Compound 7 than in the presence of NN-DNJ. This result suggest that Compound 7 can be a good ligand for the treatment of Type 2 and Type 3 Gaucher disease.

Table 4.44. Protein RMSD and RMSF values of significant mutated residues and glycosylated Asn 19.

Compound	Overall Protein RMSD	RMSF (Å)					
		Leu 444	Asn 370	Phe 213	Asn 188	Gly 202	Asn 19
APO-GBA	1.17	0.63	0.44	0.42	0.60	0.65	0.66
NN-DNJ	1.00	0.73	0.38	0.41	0.56	0.65	0.51
Zinc 00895521	1.21	0.75	0.43	0.46	0.59	0.67	0.58
Zinc 04838663	1.45	0.88	0.47	0.44	0.64	0.78	0.74
Zinc 00891180	1.29	1.27	0.44	0.45	0.72	0.68	0.86
Zinc 00017874	1.09	0.72	0.41	0.46	0.69	0.72	0.63
Zinc 35092294	1.30	0.91	0.52	0.48	0.69	0.69	0.71
Zinc 01726382	1.09	0.69	0.46	0.43	0.65	0.70	0.69

5. CONCLUSIONS AND RECOMMENDATIONS

5.1. Conclusions

In this thesis, computational drug discovery methods were utilized in order to find novel pharmacological chaperones and novel scaffolds against Gaucher Disease. Integrated study of both structure and ligand based pharmacophore modeling, molecular docking followed by substructure search of ZINC database with the common scaffolds of different ligand sets and induced fit docking studies were performed. As a result of pharmacophore modeling, ADDHR and AAADR hypotheses were obtained for ligand-based modeling performed using experimentally known ligands while HRH hypothesis was found for structure-based modeling. In order to validate the ligand-based hypotheses, structure-activity relationship (QSAR) model was generated. As a result, both hypotheses were found statistically important (For ADDHR, 0.9091 R^2 , 0.2345 RMSE, 0.7659 Q^2 , for AAADR, 0.8785 R^2 , 0.2467 RMSE and 0.7468 Q^2) and they are sufficient to use in drug discovery studies for Beta-Glucocerebrosidase (GBA). ZINC Database consisting of 500000 molecules were filtered with ADDHR, AAADR and HRH hypotheses and docked to the active site of GBA in SP and XP mode after high-throughput virtual screening. Molecules were eliminated with respect to XP GScore cutoff (-6.5 kcal/mol) and strain energy cutoff (4 kcal/mol). Remained compounds for each hypothesis (96 ligands for HRH, 218 ligands for AAADR and 241 ligands for ADDHR) were clustered with respect to their structural similarities and substructure search was done for each cluster. Common scaffolds obtained by substructure search were utilized to retrieve new molecules from ZINC website and they are docked in the same way to the active site of GBA. After filtration of molecules with respect to their docking scores, ADME and drug properties, the top molecules were redocked to GBA active site flexibly (IFD). Molecules were compared with respect to their induced fit docking scores, blood-brain barrier coefficients, human oral absorption percentages. Consequently, best molecules (Zinc00895521, Zinc04838663, Zinc00551825, Zinc00891180, Zinc00017874, Zinc35092294 and Zinc01726382) were simulated for 50 ns at 1 atm-310 K in order to analyze their stability with GBA. All compounds except Zinc0017478 remained bound to GBA until the end of simulation. In

addition, Zinc00891180 has dioxino-quinolione scaffold which has been already used for kinase CHK1 inhibitor and allosteric modulator for melanocortin receptor. Zinc04838663, Zinc00551825 and Zinc20135476 share a common scaffold called tricyclic pyrido-thieno-pyrimidine used as anticancer agent. These scaffolds are expected to bind GBA active site and they can be used to generate new compounds having positive blood brain barrier coefficient enabling molecule to pass blood-brain barrier. As a result, they offer new insights for Type 2 and Type 3 Gaucher Disease.



Table 5.1. Properties of existing and proposed compounds for GBA.

Compound		MW (Da)	IFD XP GScore (kcal/mol)	QLogBB	LLE	# of Bondswith Active Site	# of BondswithHydrophobic Pocket	Experimental IC50 for GBA (nM)*	Predicted IC50 using IFD (nM)
ExistingCompounds	NB-DNJ	219.28	-10.877	-0.794	7.43	7	1	270000 (pH:5.5)	21.43
	NN-DNJ	289.41	-8.364	-1.354	5.77	7	1	660 (pH:5.5)	1267.34
	Isofagomine	147.17	-7.368	-0.524	5.62	3	1	60 (pH:n/a)	6426.19
	Ambroxol	378.11	-6.151	0.273	1.46	2	-	4100 (pH:n/a)	46045.7
	Aminocyclitol	319.44	-11.701	-1.974	6.85	7	1	1800 (pH:5.2)	5.62
	Fingolimod	307.47	-10.265	-0.733	4.15	4	-	-	57.88
	U18666A	387.60	-7.022	0.158	-3.25	1	-	-	11196.5
	Spisulosine	285.51	-7.471	-0.995	-0.17	3	1	-	5401.48
	GW4869	504.59	-6.223	-1.932	-0.81	-	1	-	40966.1
New Proposed Compounds	Zinc00895521	525.00	-10.174	-0.088	2.32	1	-	-	67.10
	Zinc04838663	327.44	-9.346	0.201	3.54	1	1	-	257.35
	Zinc00551825	325.43	-10.142	0.121	3.15	1	1	-	70.68
	Zinc20135476	331.38	-11.420	-0.591	4.82	4	1	-	8.87
	Zinc00891180	469.98	-9.905	0.13	1.27	-	2	-	103.85
	Zinc00017874	332.30	-10.072	-0.952	3.84	1	1	-	79.19
	Zinc35092294	250.25	-9.18	-1.002	6.56	3	1	-	103.84
	Zinc01726382	388.46	-12.36	-0.62	5.17	4	-	-	1.93

*Data taken from BindingDB.com

5.2. Recommendations for Future Works

In ligand-based modeling, different pharmacophore models can be created in order to observe the influence of different featured ligands on GBA.

New molecules generated from proposed scaffolds can be tested in vitro studies. Vendor of Zinc00551825 and Zinc20135476 is Specs, while Zinc00891180 and Zinc0895521 can be bought from Asinex. Zinc04838663 are sold by ChemDiv company. Nevertheless, other compounds do not have any vendor and they need to be synthesized.

The activity and stability tests of GBA can be performed by various techniques. Thermal stability of GBA can be determined at different concentrations of selected compounds using differential scanning fluorimetry. In one study, they determined the thermal stability of GBA in the absence and presence of isofagomine. They monitored the increase in emission of NanoOrange and undergo fluorescent enhancement when it exposures to the hydrophobic parts of the protein. The fluorescence intensity of NanoOrange raises as unfolded GBA concentration increases in solution[15] . Besides, SDS-PAGE and Western plotting analysis can be performed using different types of GBA enzyme such as wildtype or mutated forms of N370S, L444P, G202R, N188S. First, the proteins were subjected to SDS-PAGE, then transferred to Immobilon-P Transfer nitrocellulose membranes. They were incubated overnight at 4 °C with anti-GBA monoclonal antibody. Then the membranes were washed and filled again with a secondary antibody sheep anti-mouse IgG Peroxidase Conjugate for 2 hours at room temperature. After washing the membrane, detection reaction was carried out by incubating the membrane in a luminol solution consisting of 100 mM Tris-HCl pH 9, coumaric acid 0.2mM, luminol 1.25mM and 0.3% H₂O₂, and the membrane was exposed to Amersham Hyper filmTM ECL. In addition, inhibition assay can be done in order to establish the best concentration of the compounds to be used in the therapy. Crude extract of COS-cells and fibroblasts were utilized for the activity measurements, and 4-methylumbelliferly- β -D-glucoside substrate placed in 0.5 M citrate buffer and Triton X-100 and taurocholate were used for the assay. The cells were transfected for 6 days using different concentrations of iminosugars and aminocyclitols, and enzyme activities were determined[13].

REFERENCES

1. Boyd, R. E., G. Lee, P. Rybczynski, E. R. Benjamin, R. Khanna, B. a. Wustman, and K. J. Valenzano, "Pharmacological chaperones as therapeutics for lysosomal storage diseases," *Journal of Medicinal Chemistry*, Vol. 56, No. 7, pp. 2705–2725, 2013.
2. Díaz L., J. Bujons, A. Delgado, H. Gutiérrez-De-Terán, and J. Åqvist, "Computational prediction of structure-Activity Relationships for the Binding of Aminocyclitols to β -Glucocerebrosidase," *Journal of Chemical Information and Modeling*, Vol. 51, No. 3, pp. 601–611, 2011.
3. Rogers S. , S. Patnaik, F. Schoenen, W. Zheng, J. Choi, O. Motabar, N. Southall, W. Westbroek, E. Goldin, E. Sidransky, W. Leister, J. J. Marugan, and J. Aube, "Discovery, SAR, and Biological Evaluation of Non-inhibitory Chaperones of Glucocerebrosidase," *Probe Reports from NIH Molecular Libraries Program*, No. Md, 2010.
4. Gault, C. R., L. M. Obeid, and Y. A. Hannun, "An overview of sphingolipid metabolism: From synthesis to breakdown," *Advances in Experimental Medicine and Biology*, Vol. 688, pp. 1–23, 2010.
5. Pralhada Rao, R., N. Vaidyanathan, M. Rengasamy, A. Mammen Oommen, N. Somaiya, and M. R. Jagannath, "Sphingolipid metabolic pathway: an overview of major roles played in human diseases.," *Journal of Lipids*, Vol. 2013, p. 178910, 2013.
6. Alfonso, P., S. Pampin, J. Estrada, J. C. Rodriguez-Rey, P. Giraldo, J. Sancho, and M. Pocova, "Miglustat (NB-DNJ) works as a chaperone for mutated acid beta-glucosidase in cells transfected with several Gaucher disease mutations," *Blood Cells, Molecules and Diseases*, Vol. 35, No. 2, pp. 268–276, 2005.

7. Maegawa, G. H. B., M. B. Tropak, J. D. Buttner, B. A. Rigat, M. Fuller, D. Pandit, L. Tang, G. J. Kornhaber, Y. Hamuro, J. T. R. Clarke, and D. J. Mahuran, "Identification and characterization of ambroxol as an enzyme enhancement agent for Gaucher disease," *Journal of Biological Chemistry*, Vol. 284, No. 35, pp. 23502–23516, 2009.
8. Dvir, H., M. Harel, A. a McCarthy, L. Toker, I. Silman, A. H. Futerman, and J. L. Sussman, "X-ray structure of human acid-beta-glucosidase, the defective enzyme in Gaucher disease.," *The EMBO Journal*, Vol. 4, No. 7, pp. 704–709, 2003.
9. Lieberman, R.L., J.A. D'aquino, D.Ringe, G.A. Petsko, "Effects of pH and iminosugar pharmacological chaperones on lysosomal glucosidase structure and stability," *Biochemistry*, Vol. 48, No. 22, pp. 4816–4827, 2009.
10. Brumshtein, B., H. M. Greenblatt, T. D. Butters, Y. Shaaltiel, D. Aviezer, I. Silman, A. H. Futerman, and J. L. Sussman, "Crystal Structures of Complexes of *N* -Butyl- and *N* -Nonyl-Deoxynojirimycin Bound to Acid β -Glucosidase," *Journal of Biological Chemistry*, Vol. 282, No. 39, pp. 29052–29058, 2007.
11. Moyses, C., "Substrate reduction therapy: clinical evaluation in type 1 Gaucher disease," *Philosophical Transactions of the Royal Society B Biological Sciences*, Vol. 358, No. 1433, pp. 955–960, 2003.
12. Sawkar, A. R., W.-C. Cheng, E. Beutler, C.-H. Wong, W. E. Balch, and J. W. Kelly, "Chemical chaperones increase the cellular activity of N370S beta -glucosidase: a therapeutic strategy for Gaucher disease.," *Proceedings of the National Academy of Sciences U. S. A.*, Vol. 99, No. 24, pp. 15428–33, 2002.
13. Sanchez-Olle, G., J. Duque, M. Egido-Gabas, J. Casas, M. Lluch, A. Chabas, D. Grinberg, and L. Vilageliu, "Promising results of the chaperone effect caused by iminosugars and aminocyclitol derivatives on mutant glucocerebrosidases causing Gaucher disease," *Blood Cells, Molecules and Diseases*, Vol. 42, No. 2, pp. 159–

166, 2009.

14. Aguilar-Moncayo, M., M. I. García-Moreno, A. Trapero, M. Egido-Gabás, A. Llebaria, J. M. G. Fernández, and C. O. Mellet, "Bicyclic (galacto)nojirimycin analogues as glycosidase inhibitors: effect of structural modifications in their pharmacological chaperone potential towards β -glucocerebrosidase.," *Organic and Biomolecular Chemistry*, Vol. 9, pp. 3698–3713, 2011.
15. Kornhaber, G. J., M. B. Tropak, G. H. Maegawa, S. J. Tuske, S. J. Coales, D. J. Mahuran, and Y. Hamuro, "Isofagomine induced stabilization of glucocerebrosidase.," *Chembiochem*, Vol. 9, No. 16, pp. 2643–9, 2008.
16. Kato, A., S. Miyauchi, N. Kato, R. J. Nash, Y. Yoshimura, I. Nakagome, S. Hirono, H. Takahata, and I. Adachi, "Docking and SAR studies of d- and l-isofagomine isomers as human β -glucocerebrosidase inhibitors," *Bioorganic Medicinal Chemistry*, Vol. 19, No. 11, pp. 3558–3568, 2011.
17. Lin, H., Y. Sugimoto, Y. Ohsaki, H. Ninomiya, A. Oka, M. Taniguchi, H. Ida, Y. Eto, S. Ogawa, Y. Matsuzaki, M. Sawa, T. Inoue, K. Higaki, E. Nanba, K. Ohno, and Y. Suzuki, "N-Octyl- α -valienamine up-regulates activity of F213I mutant beta-glucosidase in cultured cells: A potential chemical chaperone therapy for Gaucher disease," *Biochimica Biophysica Acta - Molecular Basis of Diseases*, Vol. 1689, No. 3, pp. 219–228, 2004.
18. Wang, G.-N., G. Twigg, T. D. Butters, S. Zhang, L. Zhang, L.-H. Zhang, and X.-S. Ye, "Synthesis of N-substituted ϵ -hexanolactams as pharmacological chaperones for the treatment of N370S mutant Gaucher disease.," *Organic and Biomolecular Chemistry*, Vol. 10, No. 15, pp. 2923–7, 2012.
19. Li, Z., T. Li, S. Dai, X. Xie, X. Ma, W. Zhao, W. Zhang, J. Li, and P. G. Wang, "New Insights into the Pharmacological Chaperone Activity of C2-Substituted Glucoimidazoles for the Treatment of Gaucher Disease," *ChemBioChem*, Vol. 14,

No. 10, pp. 1239–1247, 2013.

20. Serra-Vinardell, J. , L. Díaz, H. Guitiérrez-De Terán, G. Sánchez-Ollé, J. Bujons, H. Michelakakis, I. Mavridou, J. M. F. G. Aerts, A. Delgado, D. Grinberg, L. Vilageliu, and J. Casas, “Selective chaperone effect of aminocyclitol derivatives on G202R and other mutant glucocerebrosidases causing Gaucher disease,” *The International Journal of Biochemistry and Cell Biology*, Vol. 54, pp. 245–254, 2014.
21. Schitter, G., A. J. Steiner, G. Pototschnig, E. Scheucher, M. Thonhofer, C. A. Tarling, S. G. Withers, K. Fantur, E. Paschke, D. J. Mahuran, B. A. Rigat, M. B. Tropak, C. Illaszewicz, R. Saf, A. E. Statz, and T. M. Wrodnigg, “Fluorous iminoalditols: A new family of glycosidase inhibitors and pharmacological chaperones,” *ChemBioChem*, Vol. 11, No. 14, pp. 2026–2033, 2010.
22. Zimran, A., G. Altarescu, and D. Elstein, “Pilot study using ambroxol as a pharmacological chaperone in type 1 Gaucher disease,” *Blood Cells, Molecules and Diseases*, Vol. 50, No. 2, pp. 134–137, 2013.
23. Rigat B., D. Mahuran, “Diltiazem, a L-type Ca²⁺ channel blocker, also acts as a pharmacological chaperone in Gaucher patient cells,” *Molecular Genetics and Metabolism*, Vol. 96, No. 4, pp. 225–232, 2009.
24. Kato, A., I. Nakagome, K. Sato, A. Yamamoto, I. Adachi, R. J. Nash, G. W. J. Fleet, Y. Natori, Y. Watanabe, T. Imahori, Y. Yoshimura, H. Takahata, and S. Hirono, “Docking study and biological evaluation of pyrrolidine-based iminosugars as pharmacological chaperones for Gaucher disease,” *Organic and Biomolecular Chemistry*, 2016.
25. Laigre, E., D. Hazelard, J. Casas, J. Serra-Vinardell, H. Michelakakis, I. Mavridou, J. M. F. G. Aerts, A. Delgado, and P. Compain, “Investigation of original multivalent iminosugars as pharmacological chaperones for the treatment of gaucher disease,” *Carbohydrate Research*, pp. 4–10, 2016.

26. LigPrep, Schrödinger, LLC, New York, NY, 2015.
27. ConfGen, Schrödinger, LLC, New York, NY, 2015.
28. QikProp, Schrödinger, LLC, New York, NY, 2015.
29. Madhavi Sastry, G., M. Adzhigirey, T. Day, R. Annabhimoju, and W. Sherman, "Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments," *Journal of Computational Aided Molecular Design*, Vol. 27, No. 3, pp. 221–234, 2013.
30. Schrödinger Suite 2015-4 Protein Preparation Wizard Epik version 3.4, Schrödinger, LLC, New York, NY, 2015; Impact version 6.9, Schrödinger, LLC, New York, NY, 2015; Prime version 4.2, Schrödinger, LLC, New York, NY, 2015.
31. Glide, Schrödinger, LLC, New York, NY, 2015.
32. Phase, Schrödinger, LLC, New York, NY, 2015.
33. Dixon, S. L., A. M. Smondyrev, E. H. Knoll, S. N. Rao, D. E. Shaw, and R. A. Friesner, "PHASE: A new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results," *Journal of Computational Aided Molecular Design*, Vol. 20, No. 10–11, pp. 647–671, 2006.
34. Dixon, S. L., A. M. Smondyrev, and S. N. Rao, "PHASE: A novel approach to pharmacophore modeling and 3D database searching," *Chemical Biology and Drug Design*, Vol. 67, No. 5, pp. 370–372, 2006.
35. Salam, N. K., R. Nuti, and W. Sherman, "Novel method for generating structure-based pharmacophores using energetic analysis," *Journal of Chemistry Information and Modeling*, Vol. 49, No. 10, pp. 2356–2368, 2009.

36. Loving, K., N. K. Salam, and W. Sherman, "Energetic analysis of fragment docking and application to structure-based pharmacophore hypothesis generation," *Journal of Computational Aided Molecular Design*, vol. 23, no. 8, pp. 541–554, 2009.
37. Bento, A. P., A. Gaulton, A. Hersey, L. J. Bellis, J. Chambers, M. Davies, F. A. Krüger, Y. Light, L. Mak, S. McGlinchey, M. Nowotka, G. Papadatos, R. Santos, and J. P. Overington, "The ChEMBL bioactivity database: An update," *Nucleic Acids Research*, Vol. 42, No. D1, 2014.
38. Schrödinger, *Phase-Score Hypotheses Step*, 2014, http://gohom.win/ManualHom/Schrodinger/Schrodinger_2015-2_docs/maestro/help_Maestro/phase/score_hypotheses.html, accessed at October 2016.
39. Schrödinger, *Glide: A complete solution for ligand-receptor docking*, 2010, <https://isp.ncifcrf.gov/files/isp/uploads/2010/07/glide.pdf>, accessed at October 2016.
40. Friesner R. A., R. B. Murphy, M. P. Repasky, L. L. Frye, J. R. Greenwood, T. a. Halgren, P. C. Sanschagrin, and D. T. Mainz, "Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein-Ligand Complexes," *Journal of Medicinal Chemistry*, No. 49, pp. 6177–6196, 2006.
41. Schrodinger Software Release 2015-2, "Induced Fit Docking," *Schrodinger Press*, pp. 1–80, 2015.
42. Induced Fit Docking, Schrödinger, LLC, New York, NY, 2015
43. Hopkins, A. L., G. M. Keserü, P. D. Leeson, D. C. Rees, and C. H. Reynolds, "The role of ligand efficiency metrics in drug discovery.," *Nature Reviews Drug Discovery*, Vol. 13, No. 2, pp. 105–21, 2014.
44. Abad-Zapatero, C., O. Perišić, J. Wass, A. P. Bento, J. Overington, B. Al-Lazikani,

- and M. E. Johnson, "Ligand efficiency indices for an effective mapping of chemico-biological space: The concept of an atlas-like representation," *Drug Discovery Today*, Vol. 15, No. 19–20, pp. 804–811, 2010.
45. Canvas, Schrödinger, LLC, New York, NY, 2015.
46. Sastry, M., J. F. Lowrie, S. L. Dixon, and W. Sherman, "Large-scale systematic analysis of 2D fingerprint methods and parameters to improve virtual screening enrichments," *Journal of Chemistry Information and Modeling*, Vol. 50, No. 5, pp. 771–784, 2010.
47. Duan, J., S. L. Dixon, J. F. Lowrie, and W. Sherman, "Analysis and comparison of 2D fingerprints: Insights into database screening performance using eight fingerprint methods," *Journal of Molecular Graphics and Modeling*, Vol. 29, No. 2, pp. 157–170, 2010.
48. Irwin, J. J., T. Sterling, M. M. Mysinger, E. S. Bolstad, and R. G. Coleman, "ZINC: A free tool to discover chemistry for biology," *Journal of Chemistry Information and Modeling*, Vol. 52, No. 7, pp. 1757–1768, 2012.
49. Prime, Schrödinger, LLC, New York, NY, 2015.
50. Uciechowska, U., J. Schemies, M. Scharfe, M. Lawson, K. Wichapong, M. Jung, and W. Sippl, "Binding free energy calculations and biological testing of novel thiobarbiturates as inhibitors of the human NAD⁺ dependent histone deacetylase Sirt2," *Medchemcomm*, Vol. 3, No. 2, p. 167, 2012.
51. Desmond Molecular Dynamics System, D. E. Shaw Research, New York, NY, 2015.
52. Guo, Z., U. Mohanty, J. Noehre, T. K. Sawyer, W. Sherman, and G. Krilov, "Probing the alpha-helical structural stability of stapled p53 peptides: Molecular dynamics simulations and analysis: Research article," *Chemical Biology and Drug Design*,

- Vol. 75, No. 4, pp. 348–359, 2010.
53. Shivakumar, D., J. Williams, Y. Wu, W. Damm, J. Shelley, and W. Sherman, “Prediction of absolute solvation free energies using molecular dynamics free energy perturbation and the opls force field,” *Journal of Chemical Theory and Computation*, Vol. 6, No. 5, pp. 1509–1519, 2010.
 54. Maestro, Schrödinger, LLC, New York, NY.
 55. Shargel, L., S. Wu-Pong, and A. Yu, “Applied Biopharmaceutics and Pharmacokinetics, 5th Edition,” *Annals of Internal Medicine*, Vol. 94, p. 826, 2004.
 56. LigPrep, Schrödinger, New York, NY, 2015.
 57. Vilar, S., M. Chakrabarti, and S. Costanzi, “Prediction of passive blood-brain partitioning: Straightforward and effective classification models based on in silico derived physicochemical descriptors,” *Journal of Molecular Graphics and Modeling*, Vol. 28, No. 8, pp. 899–903, 2010.
 58. Backman, T. W. H., Y. Cao, and T. Girke, “ChemMine tools: An online service for analyzing and clustering small molecules,” *Nucleic Acids Research*, Vol. 39, No. supplementary 2, pp. 486–491, 2011.
 59. Yap, T. L., J. M. Gruschus, A. Velayati, W. Westbroek, E. Goldin, N. Moaven, E. Sidransky, and J. C. Lee, “Alpha-Synuclein interacts with glucocerebrosidase providing a molecular link between Parkinson and gaucher diseases,” *Journal of Biological Chemistry*, Vol. 286, No. 32, pp. 28080–28088, 2011.
 60. Alfonso, P., S. Pampín, J. Estrada, J. C. Rodríguez-Rey, P. Giraldo, J. Sancho, and M. Pocoví, “Miglustat (NB-DNJ) works as a chaperone for mutated acid beta-glucosidase in cells transfected with several Gaucher disease mutations,” *Blood Cells, Molecules and Diseases*, Vol. 35, No. 2, pp. 268–276, 2005.

61. Kitchen, D., H. Decornez, J. Furr, and J. Bajorath, "Docking and scoring in virtual screening for drug discovery: methods and applications," *Nature Reviews Drug Discovery*, Vol. 3, No. 11, pp. 935–949, 2004.
62. Woods, R. J., *What is Binding Affinity*, 2012 http://glycam.org/legacy_courses/2012/MolecularModeling2012/lectureNotes/Lecture6_RW.pdf, accessed at October 2016.
63. Pol-Fachin, L., M. Siebert, H. Verli, and M. L. Saraiva-Pereira, "Glycosylation is crucial for a proper catalytic site organization in human glucocerebrosidase," *Glycoconjugate Journal*, pp. 237–244, 2016.
64. Savjani, K. T., A. K. Gajjar, and J. K. Savjani, "Drug solubility: importance and enhancement techniques.," *ISRN Pharmacology*, Vol. 2012, No. 100 mL, p. 195727, 2012.
65. Zindell, R., E. R. Walker, J. Scott, P. Amouzegh, L. Wu, M. Ermann, D. Thomson, M. B. Fisher, C. L. Fullenwider, H. Grbic, P. Kaplita, B. Linehan, M. Patel, M. Patel, S. Löbbe, S. Block, C. Albrecht, M. J. Gemkow, D. T. Shih, and D. Riether, "Aryl 1,4-diazepane compounds as potent and selective CB2 agonists: Optimization of drug-like properties and target independent parameters," *Bioorganic and Medicinal Chemistry Letters*, Vol. 21, No. 14, pp. 4276–4280, 2011.
66. Gillis, E. P., K. J. Eastman, M. D. Hill, D. J. Donnelly, and N. A. Meanwell, "Applications of Fluorine in Medicinal Chemistry," *Journal of Medicinal Chemistry*, Vol. 58, No. 21, pp. 8315–8359, 2015.
67. Liou, B., W. D. Haffey, K. D. Greis, and G. A. Grabowski, "The LIMP-2/SCARB2 binding motif on acid beta-Glucosidase: Basic and applied implications for gaucher disease and associated neurodegenerative diseases," *Journal of Biological Chemistry*, Vol. 289, No. 43, pp. 30063–30074, 2014.

68. Everse, J., K. E. Everse, M. B. Grisham, *Peroxidase in Chemistry and Biology*, 2nd Edition, CRC Press, Boston, 1990.



APPENDIX A: PHARMACOPHORE MODELING

Table A.1. Glide GScore and strain energy values of 96 ligands obtained by structure-based modeling

ZINC ID	Glide GScore	Strain Energy (kcal/mol)	ZINC ID	Glide GScore	Strain Energy (kcal/mol)
ZINC04878437	-9,552	0	ZINC12302834	-7,065	3,691
ZINC00169855	-9,407	0,566	ZINC12371101	-7,058	3,71
ZINC00169855	-9,347	1,158	ZINC13325320	-7,029	3,615
ZINC16525222	-9,234	0,891	ZINC12561953	-7,028	0,869
ZINC09912163	-8,936	3,057	ZINC00897322	-6,985	1,465
ZINC04878437	-8,909	0	ZINC19810261	-6,953	0,787
ZINC16525222	-8,831	2,116	ZINC01436893	-6,932	1,918
ZINC60380058	-8,492	0,631	ZINC00068436	-6,865	2,932
ZINC05798744	-8,315	2,865	ZINC00088895	-6,845	0,383
ZINC09914588	-8,189	2,649	ZINC00052503	-6,845	0,593
ZINC32690757	-8,189	1,517	ZINC04190444	-6,84	0,523
ZINC05042732	-8,124	3,509	ZINC00170038	-6,84	0,03
ZINC02775425	-8,114	1,555	ZINC04744809	-6,82	0,082
ZINC60380098	-8,075	1,917	ZINC00128435	-6,819	0,7
ZINC09910065	-7,999	0,025	ZINC12561953	-6,817	2,127
ZINC33417810	-7,965	3,453	ZINC16741444	-6,785	2,561
ZINC19908618	-7,932	3,203	ZINC19913216	-6,785	3,442
ZINC01287873	-7,926	1,977	ZINC40964152	-6,773	2,189
ZINC12559505	-7,884	3,936	ZINC04017167	-6,73	0,112
ZINC16525222	-7,877	2,643	ZINC05342844	-6,729	0,614
ZINC04878446	-7,809	0,591	ZINC04878437	-6,722	1,413
ZINC18055803	-7,737	2,487	ZINC00128435	-6,714	2,655
ZINC02755138	-7,701	0,944	ZINC00633905	-6,694	2,778
ZINC19878580	-7,692	1,7	ZINC40968669	-6,641	1,879
ZINC04385877	-7,674	0,276	ZINC08917037	-6,641	2,829
ZINC05351001	-7,646	0,957	ZINC13548420	-6,639	3,833
ZINC33417804	-7,607	0,824	ZINC13107497	-6,627	0,749
ZINC05602789	-7,595	2,788	ZINC13646199	-6,624	3,241
ZINC40984488	-7,582	1,951	ZINC24539367	-6,608	0,741
ZINC19691349	-7,574	2,384	ZINC09909119	-6,605	1,811
ZINC02770384	-7,553	1,232	ZINC19359372	-6,6	0

Table A.1. Glide GScore and strain energy values of 96 ligands obtained by structure-based modeling (cont.)

ZINC ID	Glide GScore	Strain Energy (kcal/mol)	ZINC ID	Glide GScore	Strain Energy (kcal/mol)
ZINC04878446	-7,529	0,332	ZINC05800807	-6,589	0,375
ZINC12560091	-7,527	3,288	ZINC12379115	-6,585	1,085
ZINC00301795	-7,526	1,757	ZINC02271264	-6,581	0,851
ZINC09360332	-7,489	0,625	ZINC00939219	-6,577	0,55
ZINC13644821	-7,461	0,531	ZINC04178909	-6,574	0,69
ZINC12560424	-7,451	1,509	ZINC05343749	-6,57	1,469
ZINC02229507	-7,426	3,051	ZINC01509727	-6,564	2,298
ZINC19878922	-7,413	0	ZINC04878437	-6,55	2,045
ZINC09912526	-7,404	3,996	ZINC16580763	-6,544	0,86
ZINC19794477	-7,328	1,404	ZINC06721730	-6,542	1,708
ZINC20357596	-7,303	2,545	ZINC20149721	-6,535	2,099
ZINC00053231	-7,276	3,131	ZINC04512202	-6,534	1,109
ZINC60380080	-7,231	3,54	ZINC19760123	-6,522	1,437
ZINC00129734	-7,219	0,011	ZINC00626336	-6,518	1,05
ZINC00273506	-7,205	0,782	ZINC40985612	-6,518	1,851
ZINC58657943	-7,135	2,939	ZINC06743284	-6,515	3,022

Table A.2. ADME properties of survived 96 compounds of HRH hypothesis

ZINC ID	Glide GScore	SA SA	FO SA	FI SA	PI SA	WP SA	QPlo gBB	QPlog Po/w	QPI ogS	%H OA	Strain Energy (kcal/mol)
ZINC04878437	-9,552	514,03	346,021	159,441	8,567	0	-0,913	2,098	-3,835	83,687	0
ZINC00169855	-9,407	544,629	261,08	193,581	73,8	16,169	-1,438	1,056	-3,681	71,792	0,566
ZINC00169855	-9,347	477,937	195,947	193,091	73,268	15,632	-1,156	0,752	-2,681	70,096	1,158
ZINC16525222	-9,234	650,496	49,099	163,224	380,691	57,482	-0,699	2,064	-3,587	72,054	0,891
ZINC09912163	-8,936	783,418	264,96	241,462	274,868	2,128	-2,631	2,635	-5,842	72,912	3,057
ZINC04878437	-8,909	514,198	345,979	159,396	8,823	0	-0,913	2,1	-3,838	83,708	0
ZINC16525222	-8,831	654,884	84,401	168,26	346,457	55,765	-1,186	2,628	-4,975	85,291	2,116
ZINC60380058	-8,492	505,872	202,663	68,782	160,385	74,042	-0,157	2,72	-3,474	100	0,631

Table A.2. ADME properties of survived 96 compounds of HRH hypothesis (cont.)

ZINC ID	Glide GScore	SA SA	FO SA	FI SA	PI SA	WP SA	QPlo gBB	QPlog Po/w	QPI ogS	%H OA	Strain Energy (kcal/mol)
ZINC057 98744	-8,315	567,8 39	99,14 6	194, 641	241, 663	32,38 9	-1,567	0,637	-1,678	60,302	2,865
ZINC099 14588	-8,189	713,2 49	333,6 76	230, 939	115, 918	32,71 5	-2,125	1,346	-4,071	63,734	2,649
ZINC326 90757	-8,189	538,2 36	46,06 7	183, 232	218, 127	90,81	-0,972	2,687	-4,053	83,096	1,517
ZINC050 42732	-8,124	560,8 89	324,7 39	88,1 13	138, 484	9,553	-0,544	3,144	-4,265	100	3,509
ZINC027 75425	-8,114	546,0 57	114,9 91	165, 71	245, 492	19,86 4	-1,129	2,492	-3,416	84,93	1,555
ZINC603 80098	-8,075	568,3 2	102,1 72	127, 267	174, 655	164,2 27	-0,572	3,811	-4,933	100	1,917
ZINC099 10065	-7,999	587,9 34	281,4 42	106, 251	149, 565	50,67 5	-0,513	3,432	-4,67	100	0,025
ZINC334 17810	-7,965	651,0 56	141,4 68	127, 438	178, 903	203,2 47	-0,636	4,354	-5,966	100	3,453
ZINC199 08618	-7,932	596,8 41	333,1 48	202, 122	61,5 72	0	-1,846	2,654	-4,132	69,024	3,203
ZINC012 87873	-7,926	683,4 58	206,0 85	172, 59	279, 628	25,15 5	-1,701	3,062	-5,03	87,1	1,977
ZINC125 59505	-7,884	706,2 76	247,0 55	87,3 04	353, 462	18,45 5	-0,589	4,047	-5,056	100	3,936
ZINC165 25222	-7,877	651,1 64	82,6	169, 443	346, 195	52,92 5	-1,191	2,595	-4,905	84,902	2,643
ZINC048 78446	-7,809	436,3	265,6 04	159, 646	11,0 5	0	-0,818	1,152	-2,558	78,114	0,591
ZINC180 55803	-7,737	579,5 88	162,7 8	174, 32	242, 488	0	-1,469	1,372	-4,123	76,91	2,487
ZINC027 55138	-7,701	647,9 21	338,6 59	167, 144	107, 598	34,52 1	-1,395	3,209	-5,196	78,211	0,944
ZINC198 78580	-7,692	529,6 93	230,8 29	37,3 55	247, 133	14,37 6	0,03	2,765	-3,255	100	1,7
ZINC043 85877	-7,674	604,0 16	359,6 8	144, 707	93,6 09	6,019	-1,124	2,125	-4,118	86,347	0,276
ZINC053 51001	-7,646	623,9 87	203,3 24	153, 857	168, 495	98,31	-1,015	2,387	-4,438	86,329	0,957
ZINC334 17804	-7,607	605,9 46	135,7 14	126, 435	226, 423	117,3 75	-0,76	3,683	-4,817	100	0,824
ZINC056 02789	-7,595	484,7 35	200,0 86	141, 737	120, 63	22,28 1	-0,792	1,371	-2,703	82,436	2,788
ZINC409 84488	-7,582	834,8 05	652,3 31	57,9 6	124, 514	0	0,15	2,532	-2,924	78,559	1,951

Table A.2. ADME properties of survived 96 compounds of HRH hypothesis (cont.)

ZINC ID	Glide GScore	SA SA	FO SA	FI SA	PI SA	WP SA	QPlo gBB	QPlog Po/w	QPI ogS	%H OA	Strain Energy (kcal/mol)
ZINC196 91349	-7,574	744,4 35	385,5 78	94,0 82	215, 641	49,13 4	0,15	3,406	-3,704	80,85	2,384
ZINC027 70384	-7,553	628,5 58	183,2 72	164, 77	116, 168	164,3 48	-0,998	3,791	-6,273	82,024	1,232
ZINC196 90917	-7,541	531,0 75	358,6 17	110, 791	61,6 67	0	-0,099	1,801	-2,581	79,413	1,938
ZINC048 78446	-7,529	448,1 85	257,4 94	178, 539	12,1 52	0	-1,026	1,056	-2,749	74,347	0,332
ZINC125 60091	-7,527	761,6 96	305,7 88	110, 531	254, 299	91,07 9	-0,826	4,03	-5,844	100	3,288
ZINC003 01795	-7,526	508,1 45	86,95 5	180, 602	224, 266	16,32 1	-1,217	1,9	-3,009	78,935	1,757
ZINC093 60332	-7,489	651,2 92	371,8 24	66,7 55	212, 712	0	-0,383	4,194	-5,308	100	0,625
ZINC136 44821	-7,461	641,6 47	351,7 79	68,8 31	221, 037	0	-0,39	4,137	-5,138	100	0,531
ZINC125 60424	-7,451	614,3 62	62,77 9	96,5 71	293, 787	161,2 25	-0,241	5,109	-6,083	100	1,509
ZINC022 29507	-7,426	642,6 8	90,13 7	37,6 29	403, 017	111,8 98	0,048	4,306	-4,858	100	3,051
ZINC198 78922	-7,413	543,5 24	288,4 28	67,4 66	187, 629	0	-0,184	2,686	-3,825	100	0
ZINC099 12526	-7,404	663,0 67	404,7 74	119, 398	110, 561	28,33 4	-1,048	3,382	-4,778	100	3,996
ZINC197 94477	-7,328	649,9 06	201,8 15	120, 249	247, 719	80,12 3	-0,381	3,166	-2,989	85,797	1,404
ZINC203 57596	-7,303	598,7 73	240,6 42	58,5 83	261, 478	38,06 9	0,212	3,167	-3,244	96,272	2,545
ZINC000 53231	-7,276	596,0 48	170,8 68	48,5 81	365, 625	10,97 4	-0,034	4,293	-4,716	100	3,131
ZINC603 80080	-7,231	489,1 48	173,4 74	144, 471	171, 203	0	-1,027	1,743	-2,512	84,148	3,54
ZINC001 29734	-7,219	519,2 26	202,9 22	71,9 01	186, 349	58,05 4	0,063	3,759	-5,08	100	0,011
ZINC002 73506	-7,205	568,6 3	146,3 27	150, 388	250, 019	21,89 5	-1,078	2,057	-3,777	84,985	0,782
ZINC586 57943	-7,135	679,1 95	278,8 06	90,7 14	231, 376	78,29 8	-0,581	3,956	-5,835	100	2,939
ZINC123 02834	-7,065	649,5 77	340,2 92	90,3 02	217, 78	1,204	-0,386	1,636	-1,193	81,922	3,691
ZINC123 71101	-7,058	726,1 96	135,3 68	126, 616	358, 112	106,1	-0,369	3,287	-6,441	74,752	3,71

Table A.2. ADME properties of survived 96 compounds of HRH hypothesis (cont.)

ZINC ID	Glide GScore	SA SA	FO SA	FI SA	PI SA	WP SA	QPlo gBB	QPlog Po/w	QPI ogS	%H OA	Strain Energy (kcal/mol)
ZINC133 25320	-7,029	543,63	191,672	205,398	136,788	9,773	-1,592	1,233	-3,312	70,821	3,615
ZINC125 61953	-7,028	630,111	260,37	81,872	184,557	103,313	-0,208	3,613	-5,209	100	0,869
ZINC008 97322	-6,985	451,263	305,294	86,704	59,266	0	-0,333	1,82	-2,552	94,405	1,465
ZINC198 10261	-6,953	723,508	461,859	79,556	181,684	0,409	-0,146	1,888	-2,056	81,494	0,787
ZINC014 36893	-6,932	614,185	390,511	92,384	50,989	80,301	-0,325	4,565	-5,591	100	1,918
ZINC000 68436	-6,865	556,621	277,175	131,101	70,766	77,579	-0,629	3,77	-5,324	100	2,932
ZINC000 88895	-6,845	539,291	202,944	71,791	176,174	88,381	0,125	4,097	-5,575	100	0,383
ZINC000 52503	-6,845	604,004	213,249	109,439	281,316	0	-0,702	3,426	-4,394	100	0,593
ZINC041 90444	-6,84	656,04	128,722	131,226	328,013	68,079	-0,881	3,86	-5,677	100	0,523
ZINC001 70038	-6,84	543,257	375,896	52,932	114,428	0	-0,311	3,019	-3,12	100	0,03
ZINC047 44809	-6,82	580,536	165,625	127,637	173,456	113,818	-0,735	2,644	-4,052	92,284	0,082
ZINC001 28435	-6,819	536,052	152,697	235,197	102,813	45,346	-1,468	1,266	-4,574	65,96	0,7
ZINC125 61953	-6,817	595,639	254,56	71,857	183,406	85,816	-0,103	3,471	-4,564	100	2,127
ZINC167 41444	-6,785	678,982	325,643	101,04	252,299	0	-0,457	3,925	-5,387	100	2,561
ZINC199 13216	-6,785	718,961	378,479	80,239	188,209	72,034	0,022	2,073	-2,282	82,464	3,442
ZINC409 64152	-6,773	739,599	385,851	107,095	246,652	0	-0,466	2,207	-2,963	82,414	2,189
ZINC040 17167	-6,73	533,025	246,601	145,774	94,292	46,358	-0,806	3,878	-4,822	96,431	0,112
ZINC053 42844	-6,729	593,777	238,633	165,903	183,217	6,024	-1,126	2,901	-4,485	87,292	0,614
ZINC048 78437	-6,722	523,11	351,91	161,953	9,247	0	-0,975	2,098	-3,988	83,263	1,413
ZINC001 28435	-6,714	547,328	160,805	234,463	106,324	45,737	-1,497	1,371	-4,744	66,695	2,655
ZINC006 33905	-6,694	820,722	264,07	103,811	342,983	109,858	-0,747	4,597	-6,684	100	2,778

Table A.2. ADME properties of survived 96 compounds of HRH hypothesis (cont.)

ZINC ID	Glide GScore	SA SA	FO SA	FI SA	PI SA	WP SA	QPlo gBB	QPlog Po/w	QPI ogS	%H OA	Strain Energy (kcal/mol)
ZINC409 68669	-6,641	709,3 76	268,4 02	110, 982	252, 62	77,37 2	-0,206	2,599	-3,283	84,053	1,879
ZINC089 17037	-6,641	697,0 85	37,44 2	206, 266	344, 474	108,9 03	-1,77	2,391	-5,08	77,457	2,829
ZINC135 48420	-6,639	635,5 05	360,9 4	125, 022	80,4 69	69,07 5	-0,191	2,321	-3,351	80,041	3,833
ZINC131 07497	-6,627	632,8 53	470,4	109, 206	53,2 46	0	-0,732	3,95	-5,166	100	0,749
ZINC136 46199	-6,624	690,6 54	244,1 91	105, 96	340, 503	0	-0,812	4,19	-5,275	100	3,241
ZINC245 39367	-6,608	715,5 22	452,0 6	80,1 1	182, 844	0,508	-0,138	1,848	-1,906	81,169	0,741
ZINC099 09119	-6,605	569,7 97	343,6 06	147, 556	77,7 94	0,842	-1,153	1,758	-3,203	83,713	1,811
ZINC193 59372	-6,6	674,3 18	504,0 76	51,0 24	119, 219	0	0,738	2,013	-1,419	75,068	0
ZINC135 84585	-6,598	684,4	237,1 24	73,4 46	373, 83	0	-0,393	4,868	-6,057	100	2,678
ZINC058 00807	-6,589	520,1 85	72,47 6	17,2 19	308, 022	122,4 68	0,715	4,476	-5,176	100	0,375
ZINC123 79115	-6,585	663,7 34	165,2 41	182, 078	174, 247	142,1 68	-1,472	2,147	-3,806	74,698	1,085
ZINC022 71264	-6,581	593,3 76	155,0 02	52,7 4	354, 528	31,10 7	0,056	4,11	-4,917	100	0,851
ZINC009 39219	-6,577	591,0 83	17,92 2	63,0 54	443, 674	66,43 4	-0,144	4,234	-4,709	100	0,55
ZINC041 78909	-6,574	610,8 59	131,1 5	78,9 15	370, 599	30,19 5	-0,413	4,26	-4,964	100	0,69
ZINC053 43749	-6,57	595,9 15	147,8 56	178, 002	203, 011	67,04 6	-1,117	2,467	-4,443	82,697	1,469
ZINC015 09727	-6,564	623,4 27	118,3 12	96,3 52	380, 864	27,89 9	-0,324	4,033	-5,59	100	2,298
ZINC048 78437	-6,55	511,4 92	348,2 87	157, 092	6,11 3	0	-0,872	2,14	-3,792	84,331	2,045
ZINC165 80763	-6,544	604,2 51	264,5 59	75,2 44	264, 448	0	-0,534	2,875	-3,374	100	0,86
ZINC067 21730	-6,542	651,0 95	143,0 96	51,2 12	448, 462	8,325	-0,045	4,767	-6,107	100	1,708
ZINC201 49721	-6,535	680,3 27	329,5 82	87,9 11	215, 811	47,02 3	-0,101	4,037	-4,268	96,389	2,099
ZINC045 12202	-6,534	510,2 99	291,1 53	157, 595	61,5 51	0	-0,988	0,647	-2,685	75,503	1,109

Table A.2. ADME properties of survived 96 compounds of HRH hypothesis (cont.)

ZINC ID	Glide GScore	SA SA	FO SA	FI SA	PI SA	WP SA	QPIo gBB	QPlog Po/w	QPI ogS	%H OA	Strain Energy (kcal/mol)
ZINC19760123	-6,522	746,605	475,969	69,924	200,713	0	-0,244	3,927	-4,746	100	1,437
ZINC00626336	-6,518	622,053	239,317	139,843	152,948	89,945	-0,911	2,918	-4,478	91,814	1,05
ZINC40985612	-6,518	748,638	489,618	72,603	97,762	88,655	-0,195	3,85	-4,421	100	1,851
ZINC06743284	-6,515	593,79	210,657	185,852	181,708	15,573	-1,252	2,357	-4,185	80,724	3,022

Table A.3 Clusters involving less than 4 ligands

Cluster	ZINC ID
2	Zinc196909173
3	Zinc19359372
5	Zinc19760123 Zinc12302834
7	Zinc32690757
9	Zinc19908618

Table A.4. ZINC ID, GScore and strain energy of compounds retrieved by using the scaffold of cluster 6

ZINC ID	Glide GScore	Strain Energy (kcal/mol)	ZINC ID	Glide GScore	Strain Energy (kcal/mol)
ZINC00629284	-7,94	2,027	ZINC00666373	-7,611	1,622
ZINC00036815	-7,876	0,554	ZINC00036961	-7,58	3,03
ZINC00127024	-7,837	0,488	ZINC00196582	-7,554	1,774
ZINC00069184	-7,835	3,401	ZINC00256328	-7,554	1,594
ZINC00069201	-7,835	3,295	ZINC00115929	-7,536	3,507
ZINC00706076	-7,832	2,412	ZINC00706097	-7,505	2,019
ZINC00036440	-7,821	0,544	ZINC00122581	-7,487	0,214
ZINC00085452	-7,472	3,082	ZINC00706452	-7,294	3,368
ZINC00116097	-7,454	2,104	ZINC00132100	-7,154	1,17
ZINC00178860	-7,343	0,953	ZINC00494426	-7,021	2,176
ZINC00195247	-7,332	0,239	ZINC00055546	-6,974	0,105

Table A.5. ZINC ID, GScore and strain energy of compounds retrieved by scaffold of cluster 8

ZINC ID	Glide GScore	Strain Energy (kcal/mol)	ZINC ID	Glide GScore	Strain Energy (kcal/mol)
ZINC90243750	-8,928	3,157	ZINC21236699	-6,838	2,561
ZINC89599321	-8,628	1,187	ZINC04252138	-6,808	2,639
ZINC24711411	-8,436	3,627	ZINC08568916	-6,704	2,41
ZINC12520499	-7,971	2,206	ZINC08558671	-6,696	3,263
ZINC21831587	-7,858	3,306	ZINC09355188	-6,621	2,404
ZINC84750167	-7,793	1,004	ZINC12529365	-6,599	0,658
ZINC13346182	-7,337	3,54	ZINC16097394	-6,596	2,551
ZINC89599319	-7,323	1,509	ZINC02650032	-6,546	2,168
ZINC13345994	-7,317	3,597	ZINC13545028	-6,546	2,168
ZINC23514761	-7,311	3,477	ZINC31737285	-6,535	1,459
ZINC09477322	-7,049	3,022			

Table A.6. ZINC ID, GScore and strain energy of compounds retrieved by scaffold of cluster 10

ZINC ID	Glide GScore	Strain Energy (kcal/mol)	ZINC ID	Glide GScore	Strain Energy (kcal/mol)
ZINC00128435	-7.466	0.301	ZINC05634055	-6.735	0.232
ZINC00128435	-7.466	0.301	ZINC05634055	-6.735	0.232
ZINC00128435	-6.766	0.874	ZINC05634057	-7.084	1.735
ZINC00274360	-7.066	0	ZINC05817474	-6.79	1.21
ZINC01654585	-8.008	0.329	ZINC05817475	-6.824	2.636
ZINC01703450	-6.533	0.956	ZINC06733513	-6.94	1.281
ZINC04149791	-6.763	1.092	ZINC06733513	-6.94	1.281
ZINC04301469	-7.434	1.151	ZINC06733513	-6.94	1.283
ZINC04452966	-7.657	0.22	ZINC08567006	-6.374	2.737
ZINC04640113	-6.399	0.9	ZINC08568085	-6.758	0.94
ZINC04741517	-6.363	0.037	ZINC08568086	-6.736	0.051
ZINC04878429	-6.935	2.278	ZINC08568090	-6.722	0.104
ZINC04878433	-6.721	0	ZINC08568093	-7.183	1.15
ZINC04878436	-8.353	1.943	ZINC08568099	-7.566	3.946
ZINC04878437	-6.74	0.717	ZINC08740857	-6.401	0.842
ZINC04878446	-8.034	0.24	ZINC09009096	-6.77	0.727
ZINC04880297	-6.766	1.934	ZINC09009097	-6.566	3.06
ZINC05019639	-6.425	0.754	ZINC11374053	-7.804	0.099

Table A.6. ZINC ID, GScore and strain energy of compounds retrieved by scaffold of cluster 10 (cont.)

ZINC ID	Glide GScore	Strain Energy (kcal/mol)	ZINC ID	Glide GScore	Strain Energy (kcal/mol)
ZINC05095393	-6.656	0.247	ZINC11936431	-6.915	2.755
ZINC05095394	-6.798	0.284	ZINC14809693	-7.759	2.472
ZINC05130057	-6.507	0	ZINC14944921	-7.641	2.557
ZINC05130058	-7.259	0.845	ZINC14944951	-7.788	2.23
ZINC05148262	-6.428	0.165	ZINC16676237	-6.462	0.678
ZINC05162489	-6.682	0.197	ZINC19288139	-8.617	2.287
ZINC05488483	-6.767	0.563	ZINC19661977	-6.543	2.174
ZINC05488489	-7.432	2.172	ZINC19700804	-8.11	0.777
ZINC05545513	-6.792	0.004	ZINC20868605	-8.079	2.576
ZINC05610008	-6.716	0.777	ZINC21962498	-6.67	0.476
ZINC05610025	-6.692	0	ZINC21967740	-6.648	0
ZINC05610025	-6.649	0.934	ZINC22935067	-7.443	1
ZINC05610079	-6.716	0.015	ZINC31897406	-8.008	0.329
ZINC23399033	-8.854	0.114	ZINC32509038	-6.855	1.479
ZINC23428850	-9.428	3.07	ZINC25009275	-7.493	2.585

Table A.7. Interactions of ligands in cluster 1 of structure-based modeling

ZINC ID	Trp 179	Asp 127	Trp 381	Ser 345	Glu 235	Asn 396	Glu 340	Phe 347	Tyr 244	Tyr 313	Lys 346	Phe 246	Asn 234	Water
ZINC00 213230		H (s)	H (s)				H (s)						H (s)	
ZINC00 214263		H (s)	H (s)				H (s)						H (s)	
ZINC00 279749		H (s)	H (s)				H (s)			Pi-Pi S.			H (s)	
ZINC03 683456	H (s)	H (s)					H (s) (2)							
ZINC05 343749		H (s)	H (s)	H (b)	H (s)									
ZINC06 753198		H (s)	H (s)	H (b)	H (s)					Pi-Pi S.				
ZINC08 996357		H (s)	Pi-Pi S.			H (s)			Pi-Pi S.					
ZINC09 010544	H (s)					H (s) (2)							H (s)	

Table A.7. Interactions of ligands in cluster 1 of structure-based modeling (cont.)

ZINC ID	Trp 179	Asp 127	Trp 381	Ser 345	Glu 235	Asn 396	Glu 340	Phe 347	Tyr 244	Tyr 313	Lys 346	Phe 246	Asn 234	Water
ZINC09089201						H (s) (2)	H (s)							
ZINC09241405	H (s)			H (s)					Pi-Pi S.		H (s)			H (s)
ZINC09353855			Pi-Pi S.								H (s)	Pi-Pi S.		
ZINC09447905	H (s)				H (s)					Pi-Pi S.			H (s)	
ZINC12463555	H (s)	H (s)		H (b)		H (s)		Pi-Pi S.						
ZINC12873329	H (s)	H (s)			H (s)		H (s)							
ZINC13728463			Pi-Pi S.			H (s) (2)		Pi-Pi S.						
ZINC14879274	H (s)	H (s)							Pi-Pi S.	Pi-Pi S.				
ZINC19397245		H (s)												
ZINC19789949	H (s)					H (s) (2)							H (s)	
ZINC19841201				H (s)			H (s)				H (s)			
ZINC20538875	H (s)					H (s)								

Table A.8. Interactions of ligands in cluster 4

ZINC ID	Trp 179	Asp 127	Trp 381	Ser 345	Glu 235	Asn 396	Glu 340	Phe 347	Tyr 244	Tyr 313	Lys 346	Phe 246	Asn 234	Water
ZINC01726382	H (s)	H (s) (2)	H (s)											
ZINC03818225						H (s)	H (s)						H (s)	
ZINC04687140			H (s)			H (s)	H (s)							
ZINC05323884				H (b)	H (s)								H (s)	
ZINC04728750	H (s)	H (s) (2)	H (s)	H (b)		H (s)	H (s)		H (s)				H (s)	
ZINC04033947	H (s)		H (s)			H (s)		Pi-Pi S.	Pi-Pi S.				H (s)	

Table A.8. Interactions of ligands in cluster 4 (cont.)

ZINC ID	Trp 179	Asp 127	Trp 381	Ser 345	Glu 235	Asn 396	Glu 340	Phe 347	Tyr 244	Tyr 313	Lys 346	Phe 246	Asn 234	Water
ZINC04 650245	H (s)	H (s) (2)	H (s)	H (b)		H (s)	H (s)		H (s)				H (s)	
ZINC04 581688	H (s)				H (s)					Pi-Pi S.			H (s)	H (s)
ZINC04 581736	H (s)									Pi-Pi S.			H (s)	H (s)
ZINC04 316723						H (s)		Pi-Pi S.					H (s)	
ZINC04 743962	H (s)						H (s)							H (s)
ZINC05 258075		H (s)		H (s)		H (s)		Pi-Pi S.			H (s)			
ZINC05 154911													H (s)	
ZINC04 687139	H (s)												H (s)	H (s)
ZINC04 348040	H (s)					H (s)		Pi-Pi S.					H (s)	
ZINC04 743969	H (s)				H (s)								H (s)	H (s)
ZINC00 924612			H (s)					Pi-Pi S.	Pi-Pi S.	Pi-Pi S.			H (s)	
ZINC01 022581					H (s)	H (s)	H (s) (2)	Pi-Pi S.	Pi-Pi S.					
ZINC01 717584						H (s) (2)	H (s)							
ZINC04 488805				H (s)			H (s)							
ZINC04 524123	H (s)	H (s)	Pi-Pi S.				H (s)	Pi-Pi S.					H (s)	
ZINC01 298527	H (s)	H (s)	H (s)											
ZINC00 886837			Pi-Pi S.							Pi-Pi S.				H (s)
ZINC00 784562				H (b)		H (s)		Pi-Pi S.	Pi-Pi S.					
ZINC01 070395				H (s)							H (s)			
ZINC00 124158				H (s)							H (s)			
ZINC04 838248				H (s)							H (s)			

Table A.8. Interactions of ligands in cluster 4 (cont.)

ZINC ID	Trp 179	Asp 127	Trp 381	Ser 345	Glu 235	Asn 396	Glu 340	Phe 347	Tyr 244	Tyr 313	Lys 346	Phe 246	Asn 234	Water
ZINC03 434247	H (s)	H (s)		H (s)		H (s)					H (s)			
ZINC00 195055				H (s)							H (s)			
ZINC00 705368			Pi-Pi S.		H (s)					Pi-Pi S.				H (s)
ZINC00 243679	H (s)									Pi-Pi S.				H (s)
ZINC04 912920				H (s)							H (s)			
ZINC00 421743	H (s)	H (s)				H (s)								H (s)
ZINC04 913783				H (s)							H (s)			
ZINC01 004334	H (s)												H (s)	H (s)
ZINC00 195125				H (s)							H (s)			
ZINC01 726383						H (s) (2)		Pi-Pi S.	Pi-Pi S.					
ZINC04 903360	H (s)		H (s)					Pi-Pi S.					H (s)	
ZINC01 049890	H (s)		Pi-Pi S.	H (b)									H (s)	
ZINC02 427877				H (s)				Pi-Pi S.	Pi-Pi S.	Pi-Pi S.	H (s)			H (s)
ZINC01 049873				H (b)		H (s) (2)								
ZINC00 787014		H (s)	H (s)							Pi-Pi S.				H (s)
ZINC04 815668	H (s)		Pi-Pi S.&H (s)										H (s)	
ZINC03 505496				H (s)							H (s)			
ZINC00 895379	H (s)									Pi-Pi S.			H (s)	H (s)
ZINC04 676095	H (s)	H (s)(2)				H (s)		Pi-Pi S.						
ZINC00 886718		H (s)	H (s)					Pi-Pi S.		Pi-Pi S.				H (s)

Table A.8. Interactions of ligands in cluster 4 (cont.)

ZINC ID	Trp 179	Asp 127	Trp 381	Ser 345	Glu 235	Asn 396	Glu 340	Phe 347	Tyr 244	Tyr 313	Lys 346	Phe 246	Asn 234	Water
ZINC04 249559			Pi-Pi S.&H (s)	H (s)							H (s)	Pi-Pi S.	H (s)	
ZINC01 172808			Pi-Pi S.		H (s)					Pi-Pi S.				H (s)
ZINC01 948887	H (s)			H (s)							H (s)		H (s) (2)	H (s)
ZINC04 940487	H (s)	H (s)	Pi-Pi S.											H (s)
ZINC01 225992			Pi-Pi S.	H (s)		H (s)		Pi-Pi S.			H (s)			
ZINC01 379166	Pi-Pi S.		Pi-Pi S.		H (s)	H (s)		Pi-Pi S.						
ZINC04 676317		H (s)								Pi-Pi S.				H (s)
ZINC05 160559	H (s)	H (s)	H (s)	H (s)						Pi-Pi S.	H (s)			
ZINC04 868325	H (s)											Pi-Pi S.		
ZINC00 886720		H (s)	H (s)							Pi-Pi S. (2)				

Table A.9. Interactions between GBA and ligands in cluster 6

ZINC ID	Trp 179	Asp 127	Trp 381	Ser 345	Glu 235	Asn 396	Glu 340	Phe 347	Tyr 244	Tyr 313	Lys 346	Phe 246	Asn 234	Water
ZINC00 629284				H (s)							H (s)	Pi-Pi S.		
ZINC00 036815				H (s)		H (s)(2)					H (s)			
ZINC00 127024	H (s)			H (s)		H (s)					H (s)		H (s)	
ZINC00 069184				H (s)			H (s)				H (s)			
ZINC00 069201				H (s)			H (s)				H (s)			
ZINC00 706076				H (s)&(b)	H (s)			Pi-Pi S.						
ZINC00 036440				H (s)			H (s)				H (s)			

Table A.9. Interactions between GBA and ligands in cluster 6 (cont.)

ZINC ID	Trp 179	Asp 127	Trp 381	Ser 345	Glu 235	Asn 396	Glu 340	Phe 347	Tyr 244	Tyr 313	Lys 346	Phe 246	Asn 234	Water
ZINC00666373				H (s)			H (s)				H (s)			
ZINC00036961		H (s)		H (s)		H (s)			H (s)		H (s)			
ZINC00196582	H (s)			H (s)		H (s)					H (s)		H (s)	
ZINC00256328	H (s)			H (s)&(b)						Pi-Pi S.	H (s)		H (s)	
ZINC00115929				H (s)&(b)		H (s)					H (s)			
ZINC00706097				H (s)&(b)	H (s)			Pi-Pi S.						
ZINC00122581	H (s)			H (s)		H (s)					H (s)		H (s)	
ZINC00085452				H (s)			H (s)				H (s)			
ZINC00116097				H (s)			H (s)				H (s)			
ZINC00178860		H (s)		H (s)		H (s)					H (s)			
ZINC00195247	H (s)			H (s)		H (s)					H (s)		H (s)	
ZINC00706452				H (s)		H (s)					H (s)			
ZINC00132100				H (s)&(b)							H (s)			H (s)
ZINC00494426				H (s)			H (s)				H (s)			
ZINC00055546				H (s)&(b)	H (s)			Pi-Pi S.						
ZINC00365829		H (s)	H (s)	H (s)							H (s)			

Table A.10. Interactions between GBA and ligands in cluster 8

ZINC ID	Trp 179	Asp 127	Trp 381	Ser 345	Glu 235	Asn 396	Glu 340	Phe 347	Tyr 244	Tyr 313	Lys 346	Phe 246	Asn 234
ZINC902 43750	H (s)	H (s)			H (s)	H (s)							
ZINC895 99321				H (b)									H (s)
ZINC247 11411				H (b)&H (s)		H (s)		Pi-Pi S..	Pi-Pi S..		H (s)		
ZINC125 20499					H (s)			Pi-Pi S..		Pi-Pi S..			H (s)
ZINC218 31587						H (s)						Pi-Pi S..	
ZINC847 50167													H (s)
ZINC133 46182	H (s)					H (s)		Pi-Pi S..	Pi-Pi S..				H (s)
ZINC895 99319													H (s)
ZINC133 45994		H (s)		H (b)		H (s)							H (s)
ZINC235 14761			Pi-Pi S.			H (s)							
ZINC094 77322			Pi-Pi S.	H (s)							H (s)	Pi-Pi S.	
ZINC212 36699		H (s)		H (s)		H (s)					H (s)		
ZINC042 52138				H (b)		H (s)				Pi-Pi S..			
ZINC085 68916				H (s)		H (s)			Pi-Pi S..		H (s)	Pi-Pi S.	
ZINC085 58671				H (s)		H (s)			Pi-Pi S..		H (s)		
ZINC093 55188				H (s)		H (s)			Pi-Pi S..		H (s)		
ZINC125 29365				H (b)		H (s)(2)				Pi-Pi S.. (2)			
ZINC160 97394				H (b)		H (s)		Pi-Pi S..		Pi-Pi S..			
ZINC026 50032				H (s)		H (s)			Pi-Pi S..		H (s)	Pi-Pi S..	
ZINC135 45028				H (s)		H (s)			Pi-Pi S..		H (s)	Pi-Pi S..	
ZINC317 37285				H (b)		H (s)(2)				Pi-Pi S.. (2)			

Table A.11. Interactions between GBA and ligands in cluster 10

ZINC ID	Trp 179	Asp 127	Trp 381	Ser 345	Glu 235	Asn 396	Glu 340	Phe 347	Tyr 244	Tyr 313	Lys 346	Phe 246	Asn 234	Water
ZINC23 428850		H (s)		H (s)		H (s)				Pi-Pi S.				
ZINC23 399033				H (b)				Pi-Pi S.		Pi-Pi S.			H (s)	
ZINC19 288139				H (b)		H (s)		Pi-Pi S.						
ZINC04 878433				H (s) &H (b)	H (s)	H (s)					H (s)			
ZINC04 878436	H (s)	H (s)	H (s)	H (b)									H (s)	
ZINC20 868605	H (s)			H (b)			H (s)							
ZINC04 878446	H (s)	H (s)	H (s)			H (s)							H (s)	
ZINC01 654585	H (s)	H (s)							H (s)					
ZINC31 897406	H (s)	H (s)							H (s)					
ZINC08 568099				H (s) &H (b)							H (s)			
ZINC11 374053	H (s)					H (s)							H (s)	
ZINC14 944951				H (b)										
ZINC14 809693				H (s)						Pi-Pi S.	H (s)		H (s)	
ZINC04 452966				H (b)		H (s)								
ZINC14 944921				H (b)										
ZINC25 009275						H (s) (2)		Pi-Pi S.				Pi-Pi S.		
ZINC17 188137	H (s)	H (s)	H (s)			H (s)							H (s)	
ZINC00 128435				H (s) &H (b)	H (s)						H (s)			
ZINC22 935067				H (s)			H (s)				H (s)			
ZINC04 301469	H (s) & Pi-Pi S.	H (s)				H (s)	H (s)					Pi-Pi S.	H (s)	

Table A.11. Interactions between GBA and ligands in cluster 10 (cont.)

ZINC ID	Trp 179	Asp 127	Trp 381	Ser 345	Glu 235	Asn 396	Glu 340	Phe 347	Tyr 244	Tyr 313	Lys 346	Phe 246	Asn 234	Water
ZINC05 488489				H (s) &H (b)	H (s)					Pi-Pi S.	H (s)			H (s)
ZINC05 130058				H (s) &H (b)	H (s)			Pi-Pi S.						
ZINC08 568093				H (s) &H (b)							H (s)			
ZINC05 634057				H (s) &H (b)	H (s)			Pi-Pi S.						
ZINC00 274360	Pi-Pi S.&H (s)	H (s)				H (s)						Pi-Pi S.	H (s)	H (s)
ZINC06 733513				H (s) &H (b)	H (s)									
ZINC04 878429	H (s)	H (s)	H (s)		H (s)									
ZINC11 936431		H (s)		H (s)				Pi-Pi S.			H (s)			
ZINC32 509038	Pi-Pi S.											Pi-Pi S.		
ZINC05 817475				H (s) &H (b)	H (s)			Pi-Pi S.						
ZINC05 095394				H (s) &H (b)	H (s)			Pi-Pi S.			H (s)			
ZINC05 545513				H (s) &H (b)	H (s)	H (s)					H (s)	Pi-Pi S.		
ZINC05 817474				H (s) &H (b)	H (s)			Pi-Pi S.						
ZINC09 009096				H (s)		H (s)			H (s)		H (s)			
ZINC05 488483				H (s) &H (b)	H (s)						H (s)			
ZINC04 880297	H (s)	H (s)	H (s)	H (b)									H (s)	
ZINC04 149791				H (s) &H (b)	H (s)			Pi-Pi S.						
ZINC08 568085				H (s) &H (b)	H (s)						H (s)			H (s)
ZINC04 878437	H (s)	H (s)	H (s)			H (s)							H (s)	
ZINC08 568086				H (s) &H (b)	H (s)									
ZINC05 634055				H (s) &H (b)	H (s)			Pi-Pi S.						

Table A.11. Interactions between GBA and ligands in cluster 10 (cont.)

ZINC ID	Trp 179	Asp 127	Trp 381	Ser 345	Glu 235	Asn 396	Glu 340	Phe 347	Tyr 244	Tyr 313	Lys 346	Phe 246	Asn 234	Water
ZINC08568090				H (s) &H (b)	H (s)									
ZINC05610008				H (s) &H (b)	H (s)			Pi-Pi S.						
ZINC05610079				H (s) &H (b)	H (s)			Pi-Pi S.						
ZINC05610025				H (s) &H (b)	H (s)									
ZINC05162489				H (s) &H (b)	H (s)			Pi-Pi S.						
ZINC21962498				H (s)				Pi-Pi S.			H (s)			
ZINC05095393				H (s) &H (b)	H (s)						H (s)			
ZINC21967740			Pi-Pi S.	H (s)						Pi-Pi S.	H (s)			
ZINC09009097				H (s) &H (b)	H (s)						H (s)			
ZINC19661977	H (s)	H (s)						Pi-Pi S.						
ZINC01703450			H (s)			H (s) (2)								
ZINC05130057				H (b)										

Table A.12. QLogBB, QLogPo/w and XP GScore of 182 ligands obtained from ChEMBL

CHEMBL ID	XP GScore	QLogPo/w	QLogBB	Percent Human Oral Absorption
CHEMBL589341	-11,602	-0,865	-2,353	40,472
CHEMBL1766488	-10,83	-1,995	-3,16	11,095
CHEMBL1766477	-10,823	-1,612	-2,57	0
CHEMBL1766489	-10,64	-1,132	-3,661	15,747
CHEMBL1766469	-10,443	-2,296	-2,19	0
CHEMBL1766350	-10,431	1,324	-2,495	17,089
CHEMBL1766479	-10,415	-0,39	-2,87	28,486
CHEMBL1766476	-10,391	-1,198	-2,223	11,697
CHEMBL1766471	-10,201	-0,767	-1,903	28,404
CHEMBL1766478	-10,158	-2,049	-2,879	0
CHEMBL1082596	-10,123	-0,061	-3,075	29,741
CHEMBL1766484	-9,983	-0,844	-2,739	16,937
CHEMBL1766487	-9,966	-0,5	-2,271	21,482

Table A.12. QLogBB, QLogPo/w and XP GScore of 182 ligands obtained from ChEMBL
(cont.)

CHEMBL ID	XP GScore	QLogPo/w	QLogBB	Percent Human Oral Absorption
CHEMBL1766470	-9,963	0,465	-2,493	35,506
CHEMBL1085392	-9,938	-0,839	-2,574	26,9
CHEMBL2219816	-9,92	-0,535	-1,909	36,619
CHEMBL590070	-9,886	0,533	-2,707	49,71
CHEMBL1085391	-9,868	-1,98	-2,167	18,624
CHEMBL1766468	-9,799	-1,672	-2,912	2,071
CHEMBL1085393	-9,77	-0,474	-2,819	28,352
CHEMBL1766481	-9,768	-2,435	-2,312	12,864
CHEMBL1651632	-9,738	1,072	-1,42	70,621
CHEMBL1085130	-9,729	-0,295	-2,613	29,895
CHEMBL1766472	-9,7	-0,439	-2,698	10,376
CHEMBL1082491	-9,688	-1,183	-2,108	25,286
CHEMBL1766475	-9,687	-1,203	-1,713	26,559
CHEMBL1766483	-9,649	-2,037	-2,672	6,376
CHEMBL1766485	-9,595	-1,522	-1,9	24,766
CHEMBL1086532	-9,585	1,18	-3,234	39,884
CHEMBL1086533	-9,574	-1,073	-2,286	27,717
CHEMBL1651630	-9,544	0,351	-1,652	62,128
CHEMBL1651637	-9,438	0,756	-1,742	64,538
CHEMBL1085859	-9,427	-0,639	-2,651	28,906
CHEMBL1766482	-9,418	0,118	-2,72	34,596
CHEMBL1766482	-9,418	0,118	-2,72	34,596
CHEMBL1083794	-9,414	0,397	-3,247	32,313
CHEMBL3220377	-9,386	-1,636	-2,086	18,818
CHEMBL2219956	-9,378	-0,5	-1,898	37,234
CHEMBL1682495	-9,325	1,399	-1,209	71,003
CHEMBL3220376	-9,234	-1,444	-2,167	23,419
CHEMBL1085631	-9,197	-0,452	-2,679	29,101
CHEMBL1651633	-9,09	1,189	-1,961	65,725
CHEMBL574645	-9,062	1,297	-1,496	68,169
CHEMBL1083795	-9,052	-0,227	-1,974	38,831
CHEMBL1651635	-8,96	0,323	-1,686	61,609
CHEMBL1766473	-8,953	-0,814	-2,2	26,092
CHEMBL1651639	-8,912	1,297	-1,986	66,208
CHEMBL1085632	-8,887	-0,208	-2,441	32,3
CHEMBL1651631	-8,85	0,559	-1,589	62,392
CHEMBL1084109	-8,827	0,428	-2,637	38,081
CHEMBL1085633	-8,785	0,558	-3,123	36,016
CHEMBL1766486	-8,76	-1,768	-2,374	11,11

Table A.12. QLogBB, QLogPo/w and XP GScore of 182 ligands obtained from ChEMBL
(cont.)

CHEMBL ID	XP GScore	QLogPo/w	QLogBB	Percent Human Oral Absorption
CHEMBL1651628	-8,745	-0,09	-1,553	58,629
CHEMBL1029	-8,709	-1,044	-0,794	54,853
CHEMBL1651638	-8,667	0,721	-1,582	63,957
CHEMBL1651550	-8,64	-0,516	-0,983	57,994
CHEMBL1651636	-8,57	0,324	-1,454	61,493
CHEMBL1651549	-8,555	-0,941	-0,882	54,592
CHEMBL2164031	-8,509	3,074	-3,243	81,715
CHEMBL1651551	-8,491	-0,172	-1,073	60,389
CHEMBL1561	-8,402	-2,391	-1,288	36,504
CHEMBL408500	-8,392	0,384	-1,354	62,565
CHEMBL1090539	-8,37	-0,619	-1,174	55,126
CHEMBL1086996	-8,366	0,937	-1,475	64,402
CHEMBL1682494	-8,336	-0,12	-0,769	61,986
CHEMBL492080	-8,3	0,657	-1,818	70,31
CHEMBL88413	-8,288	-2,61	-1,242	19,336
CHEMBL1651626	-8,254	0,027	-1,34	59,137
CHEMBL1651552	-8,196	0,058	-1,295	60,693
CHEMBL1682493	-8,07	-1,496	-0,642	47,773
CHEMBL1089184	-8,064	-1,404	-0,841	50,855
CHEMBL413645	-7,989	0,316	-1,304	58,065
CHEMBL307429	-7,891	-2,114	-0,751	44,092
CHEMBL1651627	-7,871	0,325	-1,567	60,063
CHEMBL236484	-7,793	3,956	-1,1	100
CHEMBL1088892	-7,79	-1,148	-0,967	52,096
CHEMBL375376	-7,725	0,98	-1,522	62,175
CHEMBL236273	-7,654	2,596	-0,877	95,651
CHEMBL87120	-7,629	-2,513	-1,01	22,221
CHEMBL1089194	-7,576	-0,482	-1,123	57,821
CHEMBL392251	-7,567	2,758	-0,794	95,293
CHEMBL261397	-7,56	-1,656	-0,869	58,231
CHEMBL314403	-7,486	-0,041	-1,418	55,845
CHEMBL314403	-7,486	-0,041	-1,418	55,845
CHEMBL236275	-7,39	2,773	-0,719	96,305
CHEMBL3233945	-7,33	-1,788	-0,684	45,394
CHEMBL237562	-7,29	2,352	-0,898	93,654
CHEMBL526330	-7,247	-1,792	-0,699	44,379
CHEMBL526330	-7,247	-1,792	-0,699	44,379
CHEMBL429079	-7,196	2,936	-0,805	100
CHEMBL441134	-7,192	2,714	-0,72	100

Table A.12. QLogBB, QLogPo/w and XP GScore of 182 ligands obtained from ChEMBL
(cont.)

CHEMBL ID	XP GScore	QLogPo/w	QLogBB	Percent Human Oral Absorption
CHEMBL236482	-7,16	2,293	-0,934	92,787
CHEMBL1089185	-7,127	-0,997	-0,39	58,646
CHEMBL110830	-7,05	-1,805	-0,456	49,046
CHEMBL235416	-7,004	2,274	-1,446	89,157
CHEMBL235843	-6,984	3,113	-0,711	100
CHEMBL394672	-6,948	2,588	-1,079	95,143
CHEMBL449940	-6,869	0,494	-1,798	67,94
CHEMBL259717	-6,859	1,871	-0,851	90,509
CHEMBL396565	-6,843	2,501	-0,97	94,978
CHEMBL237561	-6,828	3,578	-0,548	100
CHEMBL486065	-6,807	0,651	-1,803	70,544
CHEMBL1766480	-6,725	-0,782	-1,444	36,287
CHEMBL406884	-6,69	2,036	-0,748	90,593
CHEMBL258515	-6,689	5,37	-0,084	100
CHEMBL88408	-6,688	-0,236	-1,58	64,414
CHEMBL444504	-6,687	0,388	-1,75	68,659
CHEMBL219813	-6,643	-0,554	-1,778	44,384
CHEMBL103769	-6,586	6,257	-0,433	94,701
CHEMBL491718	-6,552	-0,002	-1,5	68,164
CHEMBL236926	-6,525	3,076	-0,8	100
CHEMBL86772	-6,485	-0,574	-1,594	61,231
CHEMBL235415	-6,445	2,763	-0,439	100
CHEMBL488314	-6,425	-1,037	-1,148	50,57
CHEMBL236925	-6,422	3,425	-0,502	100
CHEMBL237130	-6,419	3,047	-0,473	100
CHEMBL1090540	-6,415	-0,155	-0,781	62,433
CHEMBL237348	-6,392	3,873	-1,006	100
CHEMBL206468	-6,376	-1,562	-0,483	51,614
CHEMBL237560	-6,336	3,234	-0,895	100
CHEMBL264147	-6,22	3,195	-0,818	94,706
CHEMBL490715	-6,2	-0,132	-1,837	65,359
CHEMBL86865	-5,867	-0,021	-1,041	72,49
CHEMBL403714	-5,849	3,931	0,039	100
CHEMBL3233942	-5,655	-1,157	-0,364	54,774
CHEMBL235197	-5,646	4,685	-0,193	100
CHEMBL392038	-5,628	2,665	-0,845	96,863
CHEMBL313375	-5,602	0,183	-1,18	72,847
CHEMBL356023	-5,573	0,495	-0,048	75,787
CHEMBL235635	-5,566	4,533	-0,858	100

Table A.12. QLogBB, QLogPo/w and XP GScore of 182 ligands obtained from ChEMBL
(cont.)

CHEMBL ID	XP GScore	QLogPo/w	QLogBB	Percent Human Oral Absorption
CHEMBL235636	-5,544	2,365	-0,69	95,441
CHEMBL1818433	-5,5	-1,508	-0,423	53,187
CHEMBL501385	-5,497	-1,592	-0,524	50,555
CHEMBL501385	-5,497	-1,592	-0,524	50,555
CHEMBL237132	-5,463	3,377	-0,784	100
CHEMBL236276	-5,397	2,432	-1,015	93,759
CHEMBL219190	-5,335	0,611	-1,176	64,114
CHEMBL491744	-5,328	1,249	-1,487	77,145
CHEMBL259848	-5,308	3,227	-1,224	94,695
CHEMBL427852	-5,288	3,288	-1,62	91,519
CHEMBL428551	-5,184	2,275	-1,129	90,965
CHEMBL1818321	-5,135	-1,508	-0,416	53,184
CHEMBL235844	-5,12	2,879	-1,324	93,842
CHEMBL1088891	-5,118	-1,126	-0,844	53,879
CHEMBL153479	-5,115	2,282	0,273	86,868
CHEMBL237131	-5,081	2,457	-1,054	93,761
CHEMBL1361379	-5,017	5,205	-0,361	100
CHEMBL237779	-4,877	3,92	-0,64	100
CHEMBL2012656	-4,811	4,176	-2,492	96,139
CHEMBL402477	-4,777	2,663	-1,26	91,58
CHEMBL259853	-4,712	4,428	-0,092	100
CHEMBL406121	-4,695	2,146	-1,58	81,321
CHEMBL237350	-4,632	2,013	-1,326	82,441
CHEMBL264146	-4,615	2,928	-1,044	94,259
CHEMBL235845	-4,602	2,81	-0,796	95,381
CHEMBL236274	-4,458	2,663	-0,803	94,973
CHEMBL236483	-4,392	1,62	-1,496	79,683
CHEMBL260291	-4,229	3,162	-0,067	100
CHEMBL259330	-4,185	5,715	-0,267	100
CHEMBL394426	-4,095	2,853	-0,669	100
CHEMBL260290	-4,094	5,433	-0,038	100
CHEMBL405616	-3,664	2,179	-1,426	83,927
CHEMBL259854	-3,636	5,055	-0,257	100
CHEMBL262347	-3,434	2,426	-1,46	84,792
CHEMBL406704	-3,398	5,888	0,01	100
CHEMBL236065	-3,156	2,559	-0,99	93,467

Table A.13. 218 survived compounds from hypothesis AAADR.305

ZINC ID	Glide GScore	Strain Energy (kcal/mol)	ZINC ID	Glide GScore	Strain Energy (kcal/mol)
ZINC14825969	-10,288	1	ZINC00409966	-8,435	1,267
ZINC04420512	-9,88	1	ZINC02556389	-8,427	1
ZINC17027347	-9,839	1	ZINC08535052	-8,425	1,528
ZINC16926741	-9,811	1	ZINC23125610	-8,424	1,976
ZINC04962317	-9,535	0,181	ZINC23141951	-8,421	2,202
ZINC12190458	-9,423	3,073	ZINC49984693	-8,402	1,637
ZINC04510621	-9,419	1,229	ZINC07419419	-8,34	2,121
ZINC16545897	-9,384	2,377	ZINC03954199	-8,328	1
ZINC18996501	-9,34	1,931	ZINC16525119	-8,315	0,298
ZINC04301176	-9,326	1,479	ZINC04179506	-8,289	1
ZINC36500912	-9,249	1	ZINC08694197	-8,272	0,183
ZINC00390702	-9,219	2,84	ZINC12716430	-8,255	1,848
ZINC23141942	-9,013	3,884	ZINC09927237	-8,252	0,02
ZINC04720912	-9,002	2,688	ZINC03785451	-8,245	1
ZINC04900336	-8,978	1	ZINC13810826	-8,245	1,243
ZINC00034199	-8,943	3,808	ZINC03417650	-8,219	3,857
ZINC65426053	-8,924	0,299	ZINC04370910	-8,198	3,856
ZINC00411061	-8,912	1	ZINC04460533	-8,181	1
ZINC17027345	-8,806	1	ZINC06796347	-8,181	2,438
ZINC01081243	-8,798	1	ZINC65394819	-8,169	0,222
ZINC00063885	-8,749	2,618	ZINC06813601	-8,15	2,457
ZINC27101140	-8,721	0,843	ZINC05461303	-8,145	3,361
ZINC23141954	-8,558	2,819	ZINC16545543	-8,144	0,518
ZINC00053683	-8,553	1	ZINC03118617	-8,134	1,306
ZINC00864152	-8,549	3,167	ZINC01438243	-8,103	1
ZINC00975848	-8,523	1,132	ZINC09963047	-8,101	0,118
ZINC23125576	-8,521	1,461	ZINC04460543	-8,098	1
ZINC17021209	-8,481	0,26	ZINC11694209	-8,094	3,245
ZINC00625962	-8,447	3,8	ZINC13131030	-8,091	1
ZINC17782397	-8,441	3,295	ZINC23142040	-8,001	1,09
ZINC09981778	-7,966	2,1	ZINC04333900	-7,996	1,085
ZINC08395503	-7,937	2,704	ZINC08012860	-7,983	1,107
ZINC07421430	-7,926	3,069	ZINC17972105	-7,976	2,734
ZINC00864182	-7,9	1,657	ZINC00895827	-7,436	3,629
ZINC08580285	-7,899	1,179	ZINC08395500	-7,434	2,846
ZINC00036720	-7,892	0,035	ZINC27101123	-7,433	1,415
ZINC04962259	-7,88	2,864	ZINC03954290	-7,407	0,326

Table A.13. 218 survived compounds from hypothesis AAADR.305 (cont.)

ZINC ID	Glide GScore	Strain Energy (kcal/mol)	ZINC ID	Glide GScore	Strain Energy (kcal/mol)
ZINC00626106	-7,862	1,92	ZINC00056408	-7,406	2,054
ZINC04927547	-7,851	3,171	ZINC17013666	-7,38	1
ZINC00706076	-7,832	2,412	ZINC08566730	-7,379	1,617
ZINC04223000	-7,815	2,203	ZINC36503694	-7,358	2,824
ZINC23141948	-7,807	0,558	ZINC20345022	-7,354	1,487
ZINC01241436	-7,806	2,743	ZINC12445447	-7,351	2,013
ZINC06721507	-7,777	1	ZINC04385869	-7,35	0,105
ZINC04385868	-7,766	0,098	ZINC04305927	-7,336	2,263
ZINC00390620	-7,765	1,856	ZINC06188436	-7,322	1,429
ZINC00062379	-7,762	1	ZINC06158223	-7,311	1,352
ZINC00705595	-7,76	2,51	ZINC06476748	-7,308	1
ZINC17020646	-7,735	0,688	ZINC16545160	-7,303	0,473
ZINC16526083	-7,717	3,056	ZINC23126481	-7,297	0,923
ZINC02977329	-7,709	0,718	ZINC09963309	-7,295	1,532
ZINC08996140	-7,701	2,554	ZINC00136975	-7,293	3,779
ZINC23126480	-7,688	1,5	ZINC01015229	-7,276	0,208
ZINC04922815	-7,672	1,073	ZINC13138949	-7,272	2,647
ZINC05636997	-7,661	3,082	ZINC16309772	-7,267	0,83
ZINC16916067	-7,654	1,155	ZINC04173589	-7,264	0,017
ZINC23142029	-7,63	0,662	ZINC01663005	-7,259	1,675
ZINC36500696	-7,63	2,574	ZINC05718805	-7,254	0,271
ZINC00633233	-7,629	1,368	ZINC00115028	-7,252	1,402
ZINC17020456	-7,619	1	ZINC09915050	-7,202	1,053
ZINC18063911	-7,612	0,069	ZINC01705122	-7,197	0,396
ZINC16545157	-7,61	0,997	ZINC46580563	-7,196	0,613
ZINC07418730	-7,604	1,979	ZINC04963545	-7,174	2,41
ZINC09909370	-7,599	2,265	ZINC20919734	-7,172	0,859
ZINC00895891	-7,596	3,317	ZINC19782479	-7,144	2,707
ZINC08548244	-7,578	0,975	ZINC15227631	-7,143	0
ZINC04534031	-7,57	1	ZINC10435769	-7,141	1,261
ZINC18172834	-7,565	0,015	ZINC08395098	-7,126	2,591
ZINC05606175	-7,491	0,051	ZINC06076241	-7,125	1
ZINC08973356	-7,477	3,027	ZINC12315359	-7,12	2,247
ZINC05181796	-7,465	2,178	ZINC20731248	-7,114	3,638
ZINC01592678	-7,451	1,141	ZINC17013362	-7,089	1,406
ZINC02698931	-7,44	0,788	ZINC04926460	-7,071	2,596
ZINC01189398	-7,057	1	ZINC00062908	-6,724	2,907

Table A.13. 218 survived compounds from hypothesis AAADR.305 (cont.)

ZINC ID	Glide GScore	Strain Energy (kcal/mol)	ZINC ID	Glide GScore	Strain Energy (kcal/mol)
ZINC16524877	-7,049	0	ZINC02411301	-6,71	1,645
ZINC02712938	-7,032	0,941	ZINC01054624	-6,703	0,081
ZINC01320041	-7,016	1	ZINC09962923	-6,691	2,026
ZINC02339328	-7,011	1,356	ZINC04025419	-6,688	1,028
ZINC00705599	-7,01	2,532	ZINC12929423	-6,686	2,833
ZINC19808023	-7,001	3,161	ZINC00000221	-6,685	0,08
ZINC17957794	-6,997	1,14	ZINC09962944	-6,678	1,961
ZINC09427294	-6,994	3,596	ZINC04384926	-6,675	1,925
ZINC12444060	-6,994	1,965	ZINC05819122	-6,661	3,871
ZINC16928937	-6,992	1	ZINC04935438	-6,656	1
ZINC04963361	-6,98	1,904	ZINC23139485	-6,654	2,214
ZINC00078701	-6,968	0,212	ZINC16491019	-6,65	0,3
ZINC08405313	-6,954	2,582	ZINC19075503	-6,641	2,164
ZINC10423060	-6,946	1,318	ZINC09214236	-6,627	1
ZINC09909329	-6,937	0,986	ZINC11976899	-6,615	3,116
ZINC18276478	-6,919	1,954	ZINC22190195	-6,608	1,689
ZINC17020453	-6,916	1	ZINC20070821	-6,605	1
ZINC12371147	-6,907	2,792	ZINC04995654	-6,603	0,019
ZINC12384360	-6,904	1,292	ZINC00859213	-6,594	1,281
ZINC18222655	-6,895	2,702	ZINC12418930	-6,593	2,156
ZINC18128341	-6,888	0,011	ZINC00708281	-6,587	2,562
ZINC04946582	-6,885	1	ZINC01635835	-6,572	0,066
ZINC02937316	-6,884	3,588	ZINC16932732	-6,544	1
ZINC12382148	-6,876	1	ZINC04386210	-6,537	2,817
ZINC09846035	-6,871	2,285	ZINC23139291	-6,531	1
ZINC17021040	-6,869	1	ZINC19909525	-6,52	2,197
ZINC00864212	-6,863	2,624	ZINC08395691	-6,513	0,021
ZINC16523720	-6,849	0	ZINC01250606	-6,509	1,12
ZINC06585101	-6,846	0,001	ZINC23126241	-6,802	2,345
ZINC02698919	-6,84	0,917	ZINC17004909	-6,773	1
ZINC06215994	-6,831	0,387	ZINC05955681	-6,768	3,165
ZINC17005625	-6,823	1	ZINC08996085	-6,752	3,42
ZINC01302040	-6,818	0,57	ZINC17021036	-6,748	1
ZINC04181427	-6,807	0,884	ZINC11028238	-6,747	2,203
ZINC00062588	-6,725	0,573	ZINC08763098	-6,746	3,694

Table A.14. 241 survived compounds from hypothesis ADDHR.125

ZINC ID	Glide GScore	Strain Energy (kcal/mol)	ZINC ID	Glide GScore	Strain Energy (kcal/mol)
ZINC16931850	-10,761	0,697	ZINC18127668	-8,716	0,343
ZINC00626542	-10,399	2,66	ZINC06379786	-8,57	2,319
ZINC17042784	-10,102	3,635	ZINC23138223	-8,55	2,908
ZINC03644152	-10,039	2,059	ZINC12243080	-8,545	0,652
ZINC08437921	-9,417	2,589	ZINC19758216	-8,531	2,819
ZINC05313299	-9,22	2,167	ZINC04172287	-8,529	0,107
ZINC20919734	-9,205	1,077	ZINC12301354	-8,518	2,802
ZINC05424268	-9,193	2,81	ZINC03855454	-8,467	3,473
ZINC16030071	-9,193	3,489	ZINC21872977	-8,46	1,706
ZINC64074652	-9,164	1,526	ZINC00036720	-8,397	0,339
ZINC65394603	-9,139	2,34	ZINC07419419	-8,396	1,974
ZINC12562308	-9,133	3,964	ZINC00839381	-8,365	3,154
ZINC04720912	-9,037	1,158	ZINC00211165	-8,362	3,47
ZINC08391923	-8,996	3,323	ZINC17020433	-8,33	2,784
ZINC15598837	-8,906	1,062	ZINC08694197	-8,32	0,088
ZINC12445381	-8,887	2,391	ZINC02311670	-8,311	2,459
ZINC12299401	-8,886	0,171	ZINC23141948	-8,26	0,423
ZINC02474545	-8,862	0,609	ZINC08391922	-8,223	1,004
ZINC01113312	-8,86	2,652	ZINC26914931	-8,221	3,982
ZINC05807493	-8,819	0,07	ZINC04386115	-8,206	2,457
ZINC23138226	-8,772	1,404	ZINC09312212	-8,19	0,795
ZINC00056474	-8,771	1,284	ZINC04384926	-8,159	1,894
ZINC00633243	-8,723	3,396	ZINC00046785	-7,886	2,1
ZINC09912526	-8,149	2,582	ZINC04385877	-7,877	0,443
ZINC04934725	-8,136	0,645	ZINC02346097	-7,864	0,013
ZINC01251044	-8,12	0	ZINC00662900	-7,861	1,229
ZINC04830978	-8,116	2,65	ZINC03727897	-7,858	0,214
ZINC19691296	-8,103	0,082	ZINC04473711	-7,819	1,551
ZINC12656254	-8,094	0,261	ZINC12323927	-7,815	0,236
ZINC06413597	-8,093	3,003	ZINC00212950	-7,808	2,651
ZINC04512144	-8,067	1,249	ZINC01097322	-7,804	0,464
ZINC08828766	-8,062	1,883	ZINC15565205	-7,803	2,142
ZINC00267760	-8,06	1,219	ZINC08996140	-7,763	2,039
ZINC04512143	-8,058	1,023	ZINC09242875	-7,75	0,328

Table A.14. 241 survived compounds from hypothesis ADDHR.125 (cont.)

ZINC ID	Glide GScore	Strain Energy (kcal/mol)	ZINC ID	Glide GScore	Strain Energy (kcal/mol)
ZINC23141942	-8,057	1,841	ZINC09969361	-7,749	2,111
ZINC13121470	-8,019	2,637	ZINC19691259	-7,743	0,33
ZINC00062908	-8,009	1,052	ZINC00085454	-7,73	0,993
ZINC13521711	-7,972	0,94	ZINC00034911	-7,705	3,014
ZINC00666370	-7,959	0,805	ZINC01177336	-7,697	2,519
ZINC19800002	-7,956	3,306	ZINC04094552	-7,693	0,239
ZINC04962263	-7,942	1,901	ZINC60382051	-7,692	0,281
ZINC00069174	-7,94	2,652	ZINC17972158	-7,68	0,376
ZINC08337082	-7,925	3,018	ZINC00034786	-7,652	3,977
ZINC20309467	-7,904	1,579	ZINC02311187	-7,645	0,81
ZINC01249445	-7,896	0,596	ZINC30822743	-7,63	1,275
ZINC23141945	-7,614	0,875	ZINC09601202	-7,344	2,315
ZINC15678668	-7,6	0,608	ZINC04619528	-7,336	2,575
ZINC09914633	-7,598	2,924	ZINC12561965	-7,327	3,668
ZINC09908995	-7,592	0,904	ZINC12468193	-7,317	3,892
ZINC08579862	-7,59	0,38	ZINC65426162	-7,316	0,945
ZINC00494426	-7,559	3,234	ZINC00070837	-7,314	1,127
ZINC65394729	-7,556	2,27	ZINC04491809	-7,278	1,869
ZINC00625908	-7,54	2,155	ZINC09908804	-7,273	3,81
ZINC18210026	-7,523	3,159	ZINC00124149	-7,265	1,034
ZINC04163045	-7,521	1,51	ZINC19800004	-7,259	1,938
ZINC23126477	-7,519	2,377	ZINC12559741	-7,243	2,509
ZINC00044093	-7,516	0,858	ZINC01436936	-7,23	2,641
ZINC17013859	-7,516	2,024	ZINC23139331	-7,22	1,854
ZINC05116168	-7,503	1,885	ZINC09316431	-7,215	1,511
ZINC09908803	-7,501	0,607	ZINC01915942	-7,208	1,05
ZINC00662894	-7,497	0,407	ZINC04558835	-7,205	0,01
ZINC06753265	-7,493	0,088	ZINC12298516	-7,205	0,218
ZINC60381995	-7,489	1,551	ZINC19078014	-7,157	0,897
ZINC12812424	-7,484	2,505	ZINC10435769	-7,15	1,195
ZINC04126334	-7,476	1,42	ZINC05274696	-7,135	0,268
ZINC00301795	-7,457	2,666	ZINC09243252	-7,134	3,023
ZINC17013713	-7,455	0	ZINC12300512	-7,12	2,996
ZINC00049154	-7,452	1,274	ZINC08845714	-7,116	3,312

Table A.14. 241 survived compounds from hypothesis ADDHR.125 (cont.)

ZINC ID	Glide GScore	Strain Energy (kcal/mol)	ZINC ID	Glide GScore	Strain Energy (kcal/mol)
ZINC09243769	-7,452	0,28	ZINC07423101	-7,114	3,32
ZINC17837522	-7,452	2,844	ZINC02488838	-7,107	0,022
ZINC00937984	-7,45	3,949	ZINC00803232	-7,104	3,166
ZINC08672730	-7,449	2,322	ZINC01261638	-7,101	1,908
ZINC04949816	-7,445	2,56	ZINC04113588	-7,091	3,168
ZINC08741845	-7,445	3,169	ZINC01659824	-7,087	3,11
ZINC23138193	-7,425	2,381	ZINC14921296	-7,081	3,537
ZINC22024613	-7,413	0,379	ZINC12914279	-7,075	0,971
ZINC23138055	-7,402	0,674	ZINC36505981	-7,073	3,331
ZINC18082434	-7,399	3,244	ZINC02366446	-7,063	2,152
ZINC08973276	-7,397	3,135	ZINC23139614	-7,055	3,577
ZINC03528696	-7,396	0,741	ZINC09358395	-7,038	0,001
ZINC05043294	-7,393	3,811	ZINC23139297	-7,027	0,997
ZINC05426055	-7,393	0,518	ZINC00169855	-7,022	0,672
ZINC01044296	-7,387	1,939	ZINC11935634	-7,008	1,793
ZINC18995720	-7,366	0,111	ZINC04180136	-7,007	3,421
ZINC03417650	-7,363	2,313	ZINC12301680	-7,003	0,762
ZINC07977273	-7,352	0,058	ZINC13284954	-7	1,811
ZINC11417099	-7,349	2,515	ZINC17042699	-6,995	0,787
ZINC06040039	-7,348	0,158	ZINC17020163	-6,991	1,305
ZINC00042035	-6,984	1,219	ZINC13131878	-6,687	1,72
ZINC23141994	-6,963	0,577	ZINC02260513	-6,681	3,761
ZINC22020302	-6,954	3,435	ZINC16661600	-6,674	2,393
ZINC03903085	-6,945	0,66	ZINC00407682	-6,667	3,434
ZINC00812842	-6,939	0,259	ZINC05343215	-6,662	1,265
ZINC17957794	-6,938	0,965	ZINC18222651	-6,655	0,275
ZINC16995010	-6,933	2,47	ZINC26960715	-6,65	2,397
ZINC06700796	-6,926	1,259	ZINC19691251	-6,648	2,978
ZINC06999195	-6,926	0,053	ZINC10272098	-6,645	1,486
ZINC12562318	-6,922	0,244	ZINC32771231	-6,642	0,32
ZINC20731248	-6,921	2,869	ZINC04172270	-6,616	0,13
ZINC04946348	-6,917	0,007	ZINC09242983	-6,588	3,061
ZINC06714635	-6,906	0,001	ZINC04313342	-6,574	0,966
ZINC23138187	-6,89	0,168	ZINC23139300	-6,563	0,898

Table A.14. 241 survived compounds from hypothesis ADDHR.125 (cont.)

ZINC ID	Glide GScore	Strain Energy (kcal/mol)	ZINC ID	Glide GScore	Strain Energy (kcal/mol)
ZINC09293097	-6,888	2,502	ZINC23139200	-6,546	1,846
ZINC00708978	-6,876	1,211	ZINC00244317	-6,545	1,211
ZINC00381248	-6,871	0,488	ZINC05340498	-6,545	1,673
ZINC36505027	-6,841	3,656	ZINC00639817	-6,541	1,209
ZINC09446022	-6,836	1,139	ZINC01669236	-6,539	0,483
ZINC00223310	-6,834	1,056	ZINC09334797	-6,538	1,579
ZINC23139337	-6,828	0,793	ZINC27577746	-6,538	0,643
ZINC01187428	-6,821	3,378	ZINC01320041	-6,513	0,126
ZINC00053738	-6,806	0,398	ZINC01717601	-6,723	2,767
ZINC23138181	-6,797	3,141	ZINC00051123	-6,709	0,067
ZINC00338808	-6,794	0,251	ZINC00487709	-6,707	1,665
ZINC03289817	-6,77	3,314	ZINC07422686	-6,702	3,262
ZINC04482171	-6,761	1,113	ZINC04140192	-6,701	2,785
ZINC05296209	-6,759	0,893	ZINC06142862	-6,7	1,917
ZINC05295710	-6,757	0,425	ZINC22911740	-6,697	1,812
ZINC20567513	-6,757	0,404	ZINC12063456	-6,695	0,586
ZINC01705122	-6,749	0,958	ZINC11784871	-6,693	3,158
ZINC05342038	-6,74	0,973	ZINC00082370	-6,724	0
ZINC09446022	-6,726	2,389			

Table A.15. ZINC ID, XP GScore and strain values of ligands for cluster 11 for ADDHR.125

ZINC ID	Glide GScore	Strain Energy (kcal/mol)
ZINC08845748	-8.728	1.497
ZINC08845659	-7.172	3.471
ZINC08845492	-6.921	0.643
ZINC19897786	-6.921	0.645
ZINC05343657	-6.905	1.841
ZINC09124654	-6.87	1.924
ZINC39911539	-6.779	2.213
ZINC05343380	-6.778	1.298
ZINC27583739	-6.775	3.272
ZINC08574721	-6.767	2.805
ZINC05343839	-6.695	1.36
ZINC06743272	-6.666	3.968

Table A.15. ZINC ID, XP GScore and strain values of ligands for cluster 11 for
ADDHR.125 (cont.)

ZINC ID	Glide GScore	Strain Energy (kcal/mol)
ZINC08845678	-6.601	2.732
ZINC05343749	-6.574	1.465
ZINC05343749	-6.572	1.456
ZINC05342038	-6.561	1.381
ZINC05342038	-6.561	1.374
ZINC09242983	-6.55	3.263
ZINC06753198	-6.539	1.494
ZINC05342038	-6.505	1.395

Table A.16. ZINC ID, XP GScore and strain values of ligands for cluster 6 for
AAADR.305

ZINC ID	XP GScore	Strain Energy (kcal/mol)
ZINC04715033	-10,885	1,019
ZINC03953883	-9,859	1,485
ZINC04533421	-9,793	3,617
ZINC35092286	-9,065	1,807
ZINC35092294	-8,73	3,137
ZINC04715023	-8,513	0,718
ZINC43063572	-7,953	0,376
ZINC05606432	-7,928	0,624
ZINC43063560	-6,927	1,905
ZINC05606432	-6,891	0,432

Table A.17. ZINC ID, XP GScore and strain values of ligands for cluster 7 for
AAADR.305

ZINC ID	Glide GScore	Strain Energy (kcal/mol)	ZINC ID	Glide GScore	Strain Energy (kcal/mol)
ZINC04096954	-11.138	3.203	ZINC03875122	-9.004	2.199
ZINC03778859	-10.558	3.871	ZINC01638930	-8.654	0.625
ZINC01081233	-10.281	3.371	ZINC01733348	-8.62	3.015
ZINC01728740	-10.173	3.642	ZINC03802196	-8.588	3.857
ZINC04808155	-10.161	3.84	ZINC04018120	-8.546	0.111
ZINC01531100	-10.02	1.778	ZINC00525876	-7.89	0.17

Table A.17. ZINC ID, XP GScore and strain values of ligands for cluster 7 for
AAADR.305 (cont.)

ZINC ID	Glide GScore	Strain Energy (kcal/mol)	ZINC ID	Glide GScore	Strain Energy (kcal/mol)
ZINC05437232	-10.017	0.826	ZINC02583634	-7.741	3.13
ZINC01081238	-9.997	3.784	ZINC04556966	-7.42	2.734
ZINC04095646	-9.995	0.349	ZINC04425887	-7.235	0.859
ZINC03782862	-9.796	1.456	ZINC01609441	-7.208	2.905
ZINC05437235	-9.742	2.457	ZINC03779784	-6.581	1.453
ZINC01078624	-9.495	3.232	ZINC05437233	-9.108	0.325
ZINC02381519	-9.2	0.19	ZINC03964331	-9.039	2.886

Table A.18. ZINC ID, XP GScore and strain values of ligands for cluster 16 for
AAADR.305

ZINC ID	Glide GScore	Strain Energy (kcal/mol)	ZINC ID	Glide GScore	Strain Energy (kcal/mol)
ZINC15150733	-9.675	1.395935	ZINC01335891	-8.027	1.228692
ZINC00895521	-9.444	3.132142	ZINC01335987	-8.023	2.466678
ZINC01335983	-9.077	2.192696	ZINC01336016	-8.02	2.123585
ZINC01329311	-8.867	1.875198	ZINC01335719	-8.003	2.885993
ZINC00891187	-8.841	1.511775	ZINC01336015	-7.989	1.084039
ZINC00891180	-8.805	3.109452	ZINC01335936	-7.984	3.963245
ZINC01335991	-8.673	2.316026	ZINC00880256	-7.959	2.221141
ZINC01335753	-8.673	2.166678	ZINC00588112	-7.933	3.183165
ZINC00891392	-8.654	0.950712	ZINC01336014	-7.882	1.003907
ZINC01335704	-8.64	3.665825	ZINC00891155	-7.874	2.881532
ZINC00896096	-8.616	2.733222	ZINC01335854	-7.848	2.031937
ZINC01335993	-8.552	2.375883	ZINC01336012	-7.841	3.925958
ZINC00895413	-8.53	3.367912	ZINC01378050	-7.839	2.467342
ZINC01335816	-8.517	0.893489	ZINC01378050	-7.839	2.467342
ZINC01335803	-8.473	2.171277	ZINC00895222	-7.833	3.508911
ZINC01335989	-8.468	2.194434	ZINC01335876	-7.828	2.365605
ZINC00891491	-8.409	3.181933	ZINC01335893	-7.824	1.30882
ZINC00891077	-8.408	3.694926	ZINC00891153	-7.805	1.030723
ZINC01335759	-8.404	2.907663	ZINC00901972	-7.804	1.557521
ZINC01335836	-8.404	0.76404	ZINC00891195	-7.803	3.772153
ZINC01335683	-8.401	3.714738	ZINC00891256	-7.759	1.76033
ZINC01335925	-8.4	1.910568	ZINC01335894	-7.706	2.047153
ZINC01335766	-8.391	2.204919	ZINC00919456	-7.699	2.038081

Table A.18. ZINC ID, XP GScore and strain values of ligands for cluster 16 for
AAADR.305 (cont.)

ZINC ID	Glide GScore	Strain Energy (kcal/mol)	ZINC ID	Glide GScore	Strain Energy (kcal/mol)
ZINC40395475	-8.333	0.843148	ZINC01335932	-7.68	2.164505
ZINC00891051	-8.328	2.700447	ZINC00891199	-7.659	1.187626
ZINC00895815	-8.318	2.515092	ZINC00895972	-7.651	3.365434
ZINC01335903	-8.316	2.421359	ZINC00891199	-7.647	0.368576
ZINC01335843	-8.291	2.11404	ZINC01335907	-7.639	3.780619
ZINC00425081	-8.279	0.54338	ZINC00891470	-7.616	2.537695
ZINC01225000	-8.259	2.681347	ZINC00891200	-7.603	1.493261
ZINC00880416	-8.187	2.707164	ZINC01335667	-7.552	2.429745
ZINC01335767	-8.185	2.405188	ZINC00891157	-7.516	3.172545
ZINC01378109	-8.163	2.069092	ZINC00891073	-7.502	3.700681
ZINC01335937	-8.15	2.207223	ZINC00880225	-7.485	3.872971
ZINC01335924	-8.149	3.708039	ZINC01335939	-7.469	1.355918
ZINC01335924	-8.149	3.708039	ZINC01335944	-7.396	3.489854
ZINC00895767	-8.143	2.943474	ZINC00891048	-7.396	3.346712
ZINC01335773	-8.124	2.711086	ZINC00880314	-7.569	2.476606
ZINC00880400	-8.105	2.061745	ZINC00891058	-7.392	2.383234
ZINC01335985	-8.067	2.195564	ZINC01335973	-7.398	3.704835
ZINC00895068	-8.052	2.165234	ZINC01335979	-7.366	3.345663
ZINC00920280	-7.237	1.296203	ZINC00895616	-7.325	3.808489
ZINC01335923	-7.172	2.028343	ZINC00009064	-7.466	1.387921
ZINC00890581	-7.172	3.796983	ZINC00918722	-7.279	3.095249
ZINC00891443	-7.124	3.128079	ZINC00919539	-7.247	1.708083
ZINC01335764	-7.111	3.579468	ZINC00890780	-6.837	0.707514
ZINC00890843	-7.091	3.915644	ZINC01335845	-6.836	0.124605
ZINC01335797	-7.061	1.477459	ZINC00491033	-6.749	0.621862
ZINC01377551	-6.983	3.349107	ZINC00919368	-6.713	3.427099
ZINC01335853	-6.933	2.000634	ZINC01335928	-6.698	2.257943
ZINC01335707	-6.841	3.158131	ZINC00891081	-6.666	3.253176
ZINC00425079	-6.516	0.146527	ZINC00891170	-6.631	2.564489
ZINC00891262	-6.504	0.413046	ZINC00895810	-6.561	2.668146
ZINC00887077	-6.503	0.827294			

Table A.19. Significant interactions of ligands in cluster 7 of ADDHR.125 with GBA

ZINC ID	Trp 17 9	As p 127	Trp 38 1	Ser 345	Gl u 23 5	As n 396	Gl u 34 0	Ph e 34 7	Tyr 244	Ty r 31 3	Ly s 34 6	Ph e 24 6	As n 234
ZINC1246941 6				H(b)			H(s)						
ZINC0055182 5				H(b)			H(s)	π - π					
ZINC1832476 3				H(b)			H(s)	π - π					
ZINC0248883 8				H(b)			H(s)	π - π					
ZINC2013547 6	H(s)	H(s)				H(s)	H(s)						H(s)
ZINC3340349 8				H(b)			H(s)	π - π					
ZINC2013547 6	H(s)	H(s)				H(s)	H(s)						H(s)
ZINC2010088 1				H(b)		H(s)	H(s)	π - π					
ZINC0483866 3				H(b)			H(s)	π - π					
ZINC2013547 2	H(s)	H(s)	π - π				H(s)			π - π			H(s)
ZINC1832476 5				H(b)		H(s)				π - π			
ZINC0417230 1				H(b)			H(s)	π - π					
ZINC2006446 1				H(b)			H(s)	π - π					
ZINC0489118 9				H(b)&H(s))						π - π			
ZINC1832476 5				H(b)		H(s)				π - π			

Table A.20. Significant interactions of ligands in cluster 9 of ADDHR.125 with GBA

ZINC ID	Trp 179	Asp 127	Trp 381	Ser 345	Glu 235	Asn 396	Glu 340	Phe 347	Tyr 244	Tyr 313	Lys 346	Phe 246	Asn 234
ZINC00728611				H(s)					H(s)		H(s)	π - π	
ZINC00812896				H(s)			H(s)	π - π			H(s)		
ZINC02057831				H(s)							H(s)	π - π	
ZINC02336211				H(s)			H(s)	π - π			H(s)		
ZINC00817961	H(s)			H(s)					π - π		H(s)		H(s)
ZINC00466889				H(s)			H(s)					H(s)	
ZINC01384119			π - π	H(s)					π - π		H(s)		
ZINC01423544				H(s)			H(s)					H(s)	
ZINC02059291	H(s)			H(s)		H(s)						H(s)	H(s)
ZINC01384109			π - π	H(s)					π - π		H(s)		
ZINC01422341				H(s)			H(s)	π - π			H(s)		
ZINC00666373				H(s)			H(s)				H(s)		
ZINC01422345				H(s)			H(s)				H(s)		
ZINC00036961		H(s)		H(s)		H(s)			H(s)		H(s)		
ZINC00531136				H(s)		H(s)		π - π	H(s)		H(s)		
ZINC00071726						H(s)							H(s)
ZINC00115929				H(s)		H(s)					H(s)		
ZINC00844377	H(s)			H(s)		H(s)					H(s)		H(s)
ZINC00844343	H(s)			H(s)		H(s)					H(s)		H(s)
ZINC02057819	H(s)			H(s)		H(s)					H(s)		H(s)

Table A.21. Significant interactions of ligands in cluster 11 of ADDHR.125 with GBA

ZINC ID	Trp 179	Asp 127	Trp 381	Ser 345	Glu 235	Asn 396	Glu 340	Phe 347	Tyr 244	Tyr 313	Lys 346	Phe 246	Asn 234
ZINC08845748			π - π	H(b)		H(s)							
ZINC08845659		H(s)	H(s)	H(b)	H(s)			π - π		π - π			
ZINC08845492		H(s)									H(s)		
ZINC19897786		H(s)									H(s)		
ZINC05343657		H(s)	H(s)	H(b)	H(s)			π - π		π - π			
ZINC09124654		H(s)	H(s)	H(b)	H(s)			π - π		π - π			
ZINC39911539					H(s)					π - π	H(s)		
ZINC05343380				H(b)				π - π		π - π			
ZINC27583739		H(s)	H(s)	H(b)	H(s)								
ZINC08574721		H(s)	H(s)	H(b)	H(s)					π - π			
ZINC05343839		H(s)	H(s)	H(b)	H(s)					π - π			
ZINC06743272		H(s)	H(s)	H(b)	H(s)					π - π			
ZINC08845678		H(s)	H(s)	H(b)	H(s)					π - π			

Table A.21. Significant interactions of ligands in cluster 11 of ADDHR.125 with GBA
(cont.)

ZINC ID	Trp 179	Asp 127	Trp 381	Ser 345	Glu 235	Asn 396	Glu 340	Phe 347	Tyr 244	Tyr 313	Lys 346	Phe 246	Asn 234
ZINC05343749		H(s)	H(s)	H(b)	H(s)								
ZINC05342038		H(s)	H(s)	H(b)	H(s)					π - π			
ZINC09242983		H(s)	H(s)	H(b)									
ZINC06753198		H(s)	H(s)	H(b)	H(s)					π - π			

Table A.22. Significant interactions of ligands in cluster 13 of ADDHR.125 with GBA

ZINC ID	Trp 179	Asp 127	Trp 381	Ser 345	Glu 235	Asn 396	Glu 340	Phe 347	Tyr 244	Tyr 313	Lys 346	Phe 246	Asn 234
Zinc00634793		H(s)		H(b)		H(s)	H(s)						
Zinc00640019	H(s)			H(b)		H(s)		π - π	π - π	π - π			
Zinc00485734	H(s)		π - π	H(b)(2)		H(s)							H(s)
Zinc00863220	H(s)	H(s)	π - π	H(b)		H(s)				π - π			
Zinc01108051	H(s)							π - π	π - π	π - π			
Zinc00924613			H(s)					π - π		π - π			H(s)
Zinc01269582	H(s)	H(s)		H(b)		H(s)				π - π			
Zinc00824792				H(b)		H(s)		π - π			π - π		
Zinc00407170			H(s)	H(s)						π - π	H(s)		H(s)
Zinc00518118	H(s)			H(b)	H(s)	H(s)							H(s)
Zinc00420705			H(s)							π - π			H(s)
Zinc00287483	H(s)	H(s)				H(s)							
Zinc00407171			H(s)	H(s)						π - π	H(s)		H(s)
Zinc00092990	H(s)							π - π	π - π	π - π			
Zinc00531403				H(b)		H(s)							
Zinc00662961	H(s)							π - π	π - π	π - π			
Zinc00287536				H(s)		H(s)(2)	H(s)			π - π	H(s)		
Zinc00266365	H(s)		π - π		H(s)								H(s)
Zinc01069211	H(s)		π - π H(s)			H(s)							H(s)
Zinc00049541			H(s)							π - π			H(s)
Zinc01021695	H(s)	H(s)				H(s)		π - π		π - π			

Table A. 23. Significant interactions of ligands in cluster 6 of AAADR.305 with GBA

ZINC ID	Trp 179	Asp 127	Trp 381	Ser 345	Glu 235	Asn 396	Glu 340	Phe 347	Tyr 244	Tyr 313	Lys 346	Phe 246	Asn 234
ZINC04715033	H (s)	H (s)(2)	H (s)	H (b) (2)		H (s) (2)	H(s)	π - π (2)					

Table A.23. Significant interactions of ligands in cluster 6 of AAADR.305 with GBA
(cont.)

ZINC ID	Trp 179	Asp 127	Trp 381	Ser 345	Glu 235	Asn 396	Glu 340	Phe 347	Tyr 244	Tyr 313	Lys 346	Phe 246	Asn 234
ZINC03953883	H (s)	H (s)(2)	H(s)			H(s)							
ZINC04533421		H(s)	H (s)	H (b)			H(s)(2)						H(s)
ZINC35092286		H(s)	H(s)		H (s)	H(s)		π - π				π - π	
ZINC35092294	H (s)	H(s) (2)	H (s)			H(s)						π - π	
ZINC04715023	H(s)	H(s)			H(s)	H(s)	H(s)					π - π	
ZINC43063572		H(s)				H(s)(2)	H(s)					π - π	
ZINC05606432					H(s)	H(s)	H(s)(2)						H(s)
ZINC43063560		H(s)	H(s)		H(s)		H(s)			π - π			H(s)

Table A.24. Significant interactions of ligands in cluster 7 of AAADR.305 with GBA

ZINC ID	Trp 179	Asp 127	Trp 381	Ser 345	Glu 235	Asn 396	Glu 340	Phe 347	Tyr 244	Tyr 313	Lys 346	Phe 246	Asn 234	Water	Ser 237
Zinc04096954	H(s)	H(s)		H(s)							H(s)		H(s)		H(b)
Zinc03778859	H(s)	H(s)(2)	H(s)	H(b)		H(s)(2)							H(s)		
Zinc01081233	H(s)	H(s)(2)	H(s)	H(b)		H(s)(2)							H(s)		
Zinc01728740	H(s)	H(s)(2)	H(s)	H(b)		H(s)(2)							H(s)		
Zinc04808155				H(b)		H(s)							H(s)		
Zinc01531100		H(s)									H(s)		H(s)	H(s)	
Zinc05437232	H(s)					H(s)(2)							H(s)	H(s)	
Zinc01081238	H(s)	H(s)(2)	H(s)	H(b)		H(s)(2)							H(s)		
Zinc04095646	H(s)			H(s)		H(s)					H(s)				
Zinc03782862	H(s)	H(s)	H(s)	H(s)		H(s)							H(s)		
Zinc05437235	H(s)					H(s)(2)							H(s)	H(s)	

Table A.24. Significant interactions of ligands in cluster 7 of AAADR.305 with GBA
(cont.)

ZINC ID	Trp 179	Asp 127	Trp 381	Ser 345	Glu 235	Asn 396	Glu 340	Phe 347	Tyr 244	Tyr 313	Lys 346	Phe 246	Asn 234	Water	Ser 237
Zinc01078624	H(s)	H(s)(2)	H(s)	H(b)		H(s)(2)							H(s)		
Zinc02381519	H(s)	H(s)		H(b)		H(s)									
Zinc05437233	H(s)					H(s)(2)							H(s)		
Zinc03964331						H(s)	H(s)	$\pi-\pi$	$\pi-\pi$				H(s)		
Zinc03875122	H(s)			H(s)&H(b)	H(s)	H(s)					H(s)		H(s)		
Zinc01638930	H(s)			H(b)		H(s)					H(s)				
Zinc01733348	H(s)(2)	H(s)(2)	H(s)	H(b)		H(s)									
Zinc03802196	H(s)	H(s)	H(s)			H(s)				$\pi-\pi$			H(s)		
Zinc04018120		H(s)	H(s)	H(b)						$\pi-\pi$			H(s)	H(s)	
Zinc00525876	H(s)	H(s)		H(b)		H(s)					H(s)				
Zinc02583634	H(s)	H(s)	H(s)	H(b)		H(s)(2)							H(s)		
Zinc04556966	H(s)	H(s)		H(s)			H(s)(2)			$\pi-\pi$	H(s)		H(s)		
Zinc04425887	H(s)	H(s)				H(s)					H(s)				
Zinc01609441	H(s)	H(s)		H(b)		H(s)	H(s)						H(s)		
Zinc03779784	H(s)	H(s)(2)	H(s)	H(b)											

Table A.25. Significant interactions of ligands in cluster 12 of AAADR.305 with GBA

ZINC ID	Trp 179	Asp 127	Trp 381	Ser 345	Glu 235	Asn 396	Glu 340	Phe 347	Tyr 244	Tyr 313	Lys 346	Phe 246	Asn 234
Zinc00032841		H(s)				H(s)		$\pi-\pi$					H(s)
Zinc00106233	H(s)					H(s)							
Zinc00145814								$\pi-\pi$					H(s)

[illegible]

Table A.25. Significant interactions of ligands in cluster 12 of AAADR.305 with GBA
(cont.)

ZINC ID	Trp 179	Asp 127	Trp 381	Ser 345	Glu 235	Asn 396	Glu 340	Phe 347	Tyr 244	Tyr 313	Lys 346	Phe 246	Asn 234
Zinc01255543		H(s)	π - π & H(s)					π - π					
Zinc01269540	H(s)	H(s)	π - π				H(s)						H(s)
Zinc01303070	H(s)	H(s)	π - π			H(s)							
Zinc01307568				H(s)				π - π			H(s)	π - π	
Zinc01394651	H(s)												H(s)
Zinc01394661				H(s)							H(s)		H(s)
Zinc01394918						H(s)							
Zinc01407282						H(s)		π - π					H(s)

Table A.26. Significant interactions of ligands in cluster 16 of AAADR.305 with GBA

ZINC ID	Trp 179	Asp 127	Trp 381	Ser 345	Glu 235	Asn 396	Glu 340	Phe 347	Tyr 244	Tyr 313	Lys 346	Phe 246	Asn 234
15150733	H(s)	H(s)(2)	π - π	H(s)		H(s)	H(s)	π - π		π - π	H(s)		H(s)
00895521							H(s)	π - π					
01335983				H(s)						π - π	H(s)		
01329311				H(s)						π - π	H(s)		
00891187				H(s)				π - π		π - π	H(s)		
01335753				H(s)						π - π	H(s)		
00891180				H(s)						π - π	H(s)		
01335991				H(s)						π - π	H(s)		
00891392				H(s)				π - π		π - π	H(s)		
01335704				H(s)						π - π	H(s)		
00896096				H(s)				π - π		π - π	H(s)		
01335993				H(s)				π - π		π - π	H(s)		
00895413				H(s)						π - π	H(s)		
01335816				H(s)						π - π	H(s)		
01335803				H(s)				π - π		π - π	H(s)		
01335989				H(s)				π - π	π - π	π - π	H(s)		
00891491				H(s)						π - π	H(s)		
00891077				H(s)						π - π	H(s)		
01335799				H(s)						π - π	H(s)		
01335836				H(s)						π - π	H(s)		
01335683				H(s)						π - π	H(s)		
01335925				H(s)				π - π		π - π	H(s)		
01335766				H(s)						π - π	H(s)		
40395475	H(s)	H(s)				H(s)	H(s)(2)						H(s)
00891051				H(s)						π - π	H(s)		
00895815				H(s)						π - π	H(s)		

Table A.26. Significant interactions of ligands in cluster 16 of AAADR.305 with GBA
(cont.)

ZINC ID	Trp 179	Asp 127	Trp 381	Ser 345	Glu 235	Asn 396	Glu 340	Phe 347	Tyr 244	Tyr 313	Lys 346	Phe 246	Asn 234
01335903				H(s)						π - π	H(s)		
01335843				H(s)				π - π		π - π	H(s)		
00425081						H(s)		π - π	π - π				
01225000				H(s)				π - π		π - π	H(s)		
00880416				H(s)						π - π	H(s)		
01335767				H(s)						π - π	H(s)		
01378109				H(s)						π - π	H(s)		
01378109				H(s)						π - π	H(s)		
01335937				H(s)				π - π		π - π	H(s)		
01335924				H(s)						π - π	H(s)		
00895767				H(s)						π - π	H(s)		
01335773				H(s)						π - π	H(s)		
00880400				H(s)						π - π	H(s)		
01335985				H(s)						π - π	H(s)		
00895068				H(s)				π - π		π - π	H(s)		
01335891				H(s)						π - π	H(s)		
01335987				H(s)				π - π		π - π	H(s)		
01336016				H(s)						π - π	H(s)		
01335719				H(s)						π - π	H(s)		
01336015				H(s)						π - π	H(s)		
01335936				H(s)						π - π	H(s)		
00588112													
01336014				H(s)						π - π	H(s)		
00891155				H(s)						π - π	H(s)		
01335854				H(s)						π - π	H(s)		
01336012				H(s)						π - π	H(s)		
01378050				H(s)						π - π	H(s)		
00895222													
01335876				H(s)						π - π	H(s)		
01335893				H(s)						π - π	H(s)		
00891153				H(s)				π - π		π - π	H(s)		
00901972													
00891195				H(s)						π - π	H(s)		
00891256				H(s)							H(s)		
01335894				H(s)						π - π	H(s)		
00919456				H(s)						π - π	H(s)		
01335932				H(s)				π - π		π - π	H(s)		
00891199				H(s)						π - π	H(s)		
00895872				H(s)						π - π	H(s)		
01335907				H(s)						π - π	H(s)		
00891200				H(s)						π - π	H(s)		

[illegible]

Table A.27. Significant interactions of ligands in cluster 17 of AAADR.305 with GBA

ZINC ID	Trp 179	Asp 127	Trp 381	Ser 345	Glu 235	Asn 396	Glu 340	Phe 347	Tyr 244	Tyr 313	Lys 346	Phe 246	Asn 234
Zinc00003944				H(s)& H(b)		H(s)	H(s)						
Zinc00017874	H(s)			H(s)				π - π					
Zinc00021604				H(s)		H(s)	H(s)			π - π	H(s)		
Zinc00022009							H(s) (2)		H(s)				
Zinc00022010	H(s)	H(s)			H(s)		H(s)					π - π	
Zinc00026427						H(s)	H(s)						
Zinc00056389	H(s)									π - π	H(s)		
Zinc00109315	H(s)									π - π	H(s)		
Zinc00109327	H(s)			H(s)& H(b)						π - π	H(s)		H(s)
Zinc00109388	H(s)									π - π	H(s)		
Zinc00116819	H(s)									π - π	H(s)		
Zinc00116833	H(s)									π - π	H(s)		

Table A.28. RMSD values of hit 8 ligands

Zinc ID			RMSD (Å)
Zinc00895521	Pose 1	Pose 2	7.1479
		Pose 3	2.7998
		Pose 4	6.5790
		Pose 5	6.8003
Zinc00891180	Pose 1	Pose 2	3.4850
		Pose 3	4.9019
		Pose 4	1.5970
		Pose 5	3.2083
Zinc00017874	Pose 1	Pose 2	0.5064
		Pose 3	0.6532
		Pose 4	3.8670
		Pose 5	4.0029
Zinc04838663	Pose 1	Pose 2	0.3483
		Pose 3	4.6262
		Pose 4	4.3457
		Pose 5	4.4218
Zinc00551825	Pose 1	Pose 2	0.4864
		Pose 3	5.3012
		Pose 4	4.8218
		Pose 5	4.8512

Table A.28. RMSD values of hit 8 ligands (cont.)

Zinc ID			RMSD (Å)
Zinc20135476	Pose 1	Pose 2	0.5327
		Pose 3	3.1843
		Pose 4	3.6493
		Pose 5	0.5787
Zinc35092294	Pose 1	Pose 2	2.1250
		Pose 3	2.1072
		Pose 4	2.4272
		Pose 5	1.3773
Zinc01726382	Pose 1	Pose 2	0.8817
		Pose 3	0.6376
		Pose 4	0.7045
		Pose 5	4.31

Figure A.1. Second and third set of ligands used for QSAR studies.

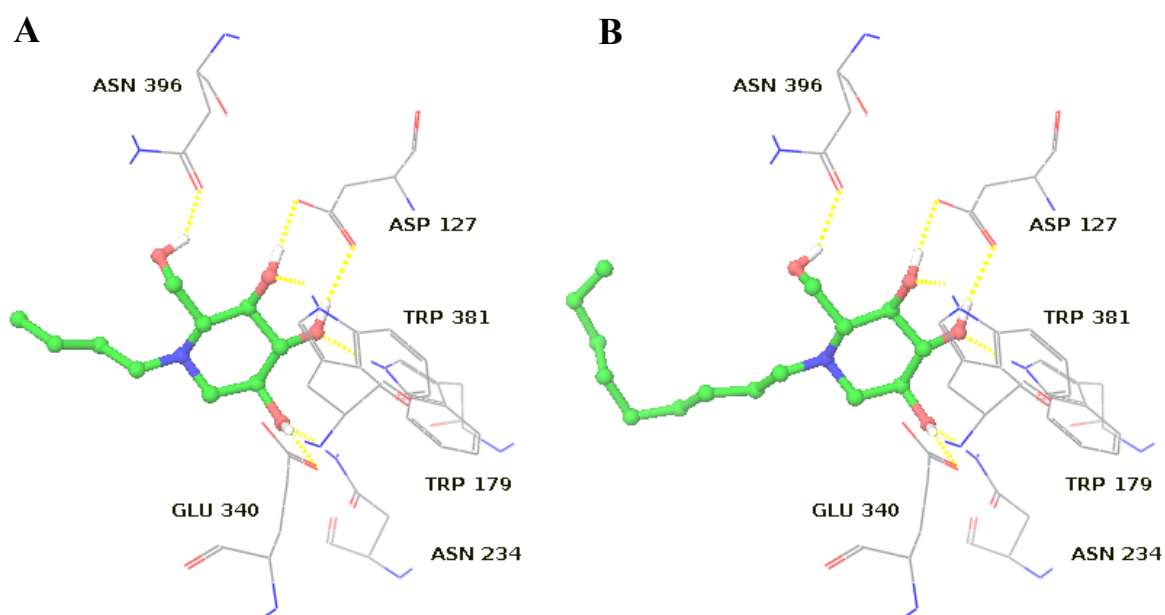


Figure A.2. Interaction maps of NB-DNJ (Miglustat) (A) and NN-DNJ (B) to GBA active site

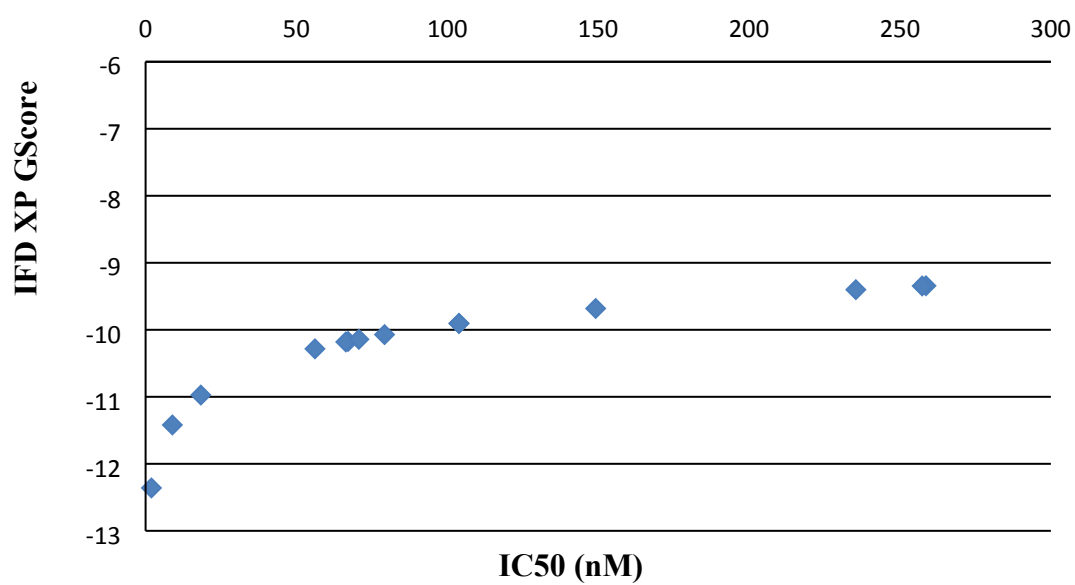


Figure A.3. IFD XP GScore versus predicted IC₅₀ of compounds obtained by ligand-based model