2012

Identification Of Pan-Ligands For Peroxisome Proliferator-Activated Receptors (Ppar) Using Computational Virtual Screening With Molecular Docking

Hamid Dafalla Ismail
North Carolina Agricultural and Technical State University

IDENTIFICATION OF PAN-LIGANDS FOR PEROXISOME PROLIFERATOR-ACTIVATED RECEPTORS (PPAR) USING COMPUTATIONAL VIRTUAL SCREENING WITH MOLECULAR DOCKING

by

Hamid Dafalla Ismail

A thesis submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of
MASTER OF SCIENCE

Department: Computational Science & Engineering
Major: Computational Sciences
Major Professor: Dr. Mary A. Smith

North Carolina A&T State University
Greensboro, North Carolina
2012

School of Graduate Studies
North Carolina Agricultural and Technical State University

This is to certify that the Master's Thesis of

Hamid Dafalla Ismail

has met the thesis requirements of
North Carolina Agricultural and Technical State University

Greensboro, North Carolina
2012

Approved by:

_____	_____
Dr. Mary A. Smith Major Professor	Dr. Marwan U. Bikdash Committee Member
_____	_____
Dr. Gregory D. Goins Committee Member	Dr. Scott H. Harrison Committee Member
_____	_____
Dr. Marwan U. Bikdash Department Chairperson	Dr. Sanjiv Sarin Associate Vice Chancellor for Research and Graduate Dean

DEDICATION

I lovingly dedicate this thesis to my wife, Amani, and my daughter, Amna, who supported me each step of the way, and to my mother and the memory of my father.

BIOGRAPHICAL SKETCH

Hamid D. Ismail was born in 1967, in Sudan. He received the Bachelor of Science degree in Veterinary Sciences in 1994 and the Higher Diploma in Computer Programming and Statistics in 1999 from the University of Khartoum. He worked at the University of Khartoum as a computer center administrator from 1995 to 1999, and then moved to Qatar, where he occupied different senior administrative positions in a multi-national environment. In 2007, Hamid moved to USA and worked as a biotechnology laboratory research assistant at NC A&T State University. Alongside his position, Hamid attended graduate school in computational sciences program. He was recognized as a top ten graduate student with a 4.0 GPA in 2012 by the College of Engineering at NC A&T State University. After finishing his master degree, Hamid plans to pursue a Ph.D. degree in a discipline relevant to computational sciences, bioinformatics and/or biophysics.

ACKNOWLEDGMENTS

First, I would like to express my utmost gratitude to the distinguished faculty members who honored me by serving on my committee: Dr. Mary Smith (Chair of the Department of Biology), Dr. Marwan Bikdash (Chair of the Department of Computational Sciences & Engineering.), Dr. Gregory Goins, and Dr. Scott Harrison for their profound reviewing, insightful comments, and meaningful questions.

I am deeply grateful to my advisor, Dr. Smith. She provided me with guidance, ideas, and support throughout the course of my research. Despite her busy schedule, she was always willing to give me the time that I needed to discuss my topics even if she had to come to her office on the weekends and stay all day long. She also gave me the opportunity to attend two workshops at the Pittsburgh Supercomputing Center (PSC). These workshops enhanced my knowledge in bioinformatics and molecular dynamics, and educated me profoundly about the tools and techniques that I used in my research.

My thanks are due to Dr. Bikdash for his support, willingness, and readiness to help. Dr. Bikdash always welcomed me to his office and dealt with my requests with immediate and appropriate attention.

I am very grateful to Dr. Sharita Womack from the Department of Biology for the proof-reading and grammatical corrections.

I would like to extend my thanks to the faculty members in the Departments of Computational Science and Engineering and Biology: Dr. Kenneth Flurchick, Ram Mohan, and Vinaya Kelkar, from whom I developed my skills in programming, mathematics, and statistics.

My thanks also go to Pittsburgh Supercomputing Center (PSC), for the grant they awarded me to use their computer platforms.

I am grateful to Dr. Morris Clarke from Winston-Salem State University for his critique and valuable advice.

My thanks also go to Dr. Worku from the Department of Animal Sciences for her support and encouragement to pursue a degree in computational sciences.

My thanks are due to Dr. Hugh Nicholas from the Pittsburgh Supercomputing Center for assisting in framing the general idea about the project.

I would like to express my warm thanks to Dr. Troy Wymore from the PSC for the profound insight I gained from the lectures he delivered about the homology modeling of proteins and enzymatic reactions in two separate workshops.

My special thanks go to Dr. Martin Field, the developer of pDynamo software, for his rich lectures about molecular dynamics and quantum chemistry and for his answers to my questions that helped me in my project.

I would like to extend my thanks to the executive assistant of the Department of Computational Science & Engineering, Lydia Leak, and the staff of the Department of Biology, Ashley Tuck and Donna Robertson for their kindness and assistance in the processing of my paperwork.

I always thank the most influential teacher in my life, Obeid Farah, my elementary school teacher, from whom I learned discipline, responsibility, and honesty.

Finally I would like to thank my wife Amani and my daughter Amna for the unlimited moral support, encouragement, and inspiration they always provide me.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURES **PAGE**

LIST OF TABLES

FIGURES **PAGE**

LIST OF ABBREVIATION

3D	Three dimensions
ABNR	Adopted basis Newton-Raphson optimization method
ADT	AutoDock Tools
AF	Activation function
ASCII	American Standard Code for Information Interchange
CASTp	Computed Atlas of Surface Topography of proteins
CHARMM	Chemistry at Harvard for Molecular Dynamics
DBD	DNA binding domain
DLG	Docking log file
DNA	Deoxyribonucleic Acid
DPF	Docking Parameter File
GA	Genetic algorithm
GB	Gigabyte
GPF	Grid Parameter File
HAT	Histone acetyltransferase
HDL	High-density lipoprotein
IUPAC	International Union of Pure and Applied Chemistry
LBD	Ligand binding domain
LDL	Low-density lipoprotein
MOLE2	File format for molecule
MS	Metabolic Syndrome

NCoR	Nuclear receptor co-repressor
NR	Nuclear receptor
PCG-1	Peroxisome proliferator-activated receptor gamma coactivator 1
PDB	Protein Data Bank
PDBQT	Protein Data Bank, Partial Charge (Q), & Atom Type (T))
PP	Peroxisome Proliferators
PPAR	Peroxisome proliferator-activated receptors
PPARα	Peroxisome proliferator-activated receptors alpha
PPARδ	Peroxisome proliferator-activated receptors delta
PPARγ	Peroxisome proliferator-activated receptors gamma
PPRE	Peroxisome proliferator DNA response element
RAM	Random access memory
RE	Response Element
RMSD	Root Mean Square Deviation
RXR	Retinoid X receptor
SD	Steepest Descend optimization method
SMILES	Simplified molecular input line entry specification
TZD	Thiazolidinediones
VDW	Van Der Waals interactions
IR	Important Residues

ABSTRACT

Ismail, Hamid Dafalla. IDENTIFICATION OF PAN-LIGANDS FOR PEROXISOME PROLIFERATOR-ACTIVATED RECEPTORS (PPAR) USING COMPUTATIONAL VIRTUAL SCREENING WITH MOLECULAR DOCKING. **(Major Professor: Dr. Mary A. Smith)**, North Carolina Agricultural and Technical State University.

Peroxisome Proliferator-activated Receptors (PPARs) are ligand-activated nuclear receptors known for their major role in metabolic syndrome (MS). Abdominal obesity, high blood pressure, increased glucose levels and low concentrations of high-density lipoprotein characterize MS. Numerous studies proposed developing pan-agonists as potent drug candidates for the treatment and control of metabolic syndrome. The objective of this study was to use virtual screening with molecular docking to identify potential pan-PPAR ligands from the ZINC database. The 3D structural files of the receptor ligand binding domains (LBD), obtained from the Protein Data Bank (PDB), were energetically minimized and the binding pockets on each LBD were identified and measured. The screening was performed by docking each compound from the lead-like database to the LBD of the three receptors using the AutoDock software. The evaluation of the docking was based on the free energy of binding, position of the compound inside the binding pocket, and the protein residues that were involved in the binding. Twenty-seven out of approximately four million lead-like compounds were found to position themselves very well in the binding pockets of the three PPARs with minimal free energy of binding. These pan-PPAR ligands may be strong candidates as pan-PPAR agonists that should be investigated further.

CHAPTER 1

INTRODUCTION

Peroxisome proliferator-activated receptors (PPARs) are a group of globular proteins that belong to the nuclear receptor super-family. They are activated when they bind to ligands and play important roles as transcription factors and major regulators of the genes involved in lipid and carbohydrates metabolism, and storage of fatty acids [1].

PPARs are divided into three isotypes; PPARα, PPARδ, and PPARγ, that differ in locations, ligand specificity, and functions [2]. PPARα is produced in the liver, heart, and muscle. It can be activated endogenously by fatty acids to regulate the genes involved in the metabolism of lipids and lipoproteins in the liver and the oxidation of fatty acids in skeletal muscles. Exogenously it can bind to a diverse set of ligands including prostaglandins, plasticizers, and synthetic fibrate drugs. The PPARδ gene is expressed ubiquitously but most abundantly in brain, adipose tissue, skin, and kidney. It is activated by both saturated and unsaturated fatty acids, and it has been found to have a role in the regulation of fatty acid oxidation and in modulating the level of the high density lipoprotein (HDL) [3]. PPARγ gene is expressed in three isoforms (γ1, γ2, and γ3) in many organs including heart, muscle, large intestine, pancreas, spleen, adipose tissue and macrophages [4]. Endogenously, PPARγ has a low affinity for naturally occurring fatty acids, eicosanoids, prostaglandins and their metabolites. It exhibits a preference for essential polyunsaturated fatty acids. Thus, it may act as a fatty acid sensor along with the target genes associated with lipid and glucose homeostasis [5].

PPARγ agonists have been found to increase the action of insulin and to lower the serum glucose in patients with type 2 diabetes mellitus [6].

A recent study has shown that approximately 40 million individuals in the United States are afflicted with the metabolic syndrome (MS) and the prevalence increases with age in the western societies. MS, is a cluster of disorders characterized by abdominal obesity, high blood pressure, increased fasting glucose levels and low concentrations of HDL which can lead to insulin insensitivity and type 2 diabetes [7] Since PPARs mediates lipid and glucose metabolism, they have been intense pharmacological targets for the treatment and control of MS.

There are a number of drugs that have been developed to treat the different disorders of MS by targeting specific PPAR isotypes. Fibrates are synthetic ligands of PPARα that are used as lipid lowering drugs, while synthetic agonists of PPARγ, the TZDs, are used to increase insulin sensitivity in type 2 diabetic patients. Independently, these drugs are ineffective as therapeutic treatments for the cluster of disorders that constitute the MS. Therefore, researchers are searching for PPAR ligands with superior therapeutic and metabolic actions [8, 9]. Some researchers are trying to develop pan-PPAR agonists, which are synthetic ligands that are capable of activating the three PPAR isotypes simultaneously. Pan-PPAR agonists offer the potential to increase drug efficacy and reduce the risk factors associated with polypharmacy [10, 11]. The significant structural similarity of the ligand binding domain of the three isotypes makes it possible to find such alternative drugs. Bezafibrate was the first drug to be identified and clinically investigated [12] followed by Indeglitazar, which was identified by coupling

low-affinity biochemical screening with high-throughput co-crystallography [10]. Unfortunately, these pan-agonists did not pass phase 2 clinical trials. Thus, investigations are pursuing different directions and adopting alternative approaches in search of more efficient pan-agonists.

The primary goal of this study was to apply computational techniques to identify potential pan-PPAR ligands among the millions of lead-like compounds stored in digital format in the data repository managed by ZINC, a free online database of commercially-available compounds for ligand discovery. To achieve the goal of this study, three key factors had to be investigated and determined: the size of the major cavity where ligand binding takes place; the residues that are the most frequently involved in binding; and the best positions for the ligands in the binding cavity in the three isotypes along with corresponding receptor-ligand conformations. The significance of these three factors is obvious. The major cavity was developed throughout the evolutionary path to host the ligand. Studies have shown that the binding cavity is relatively large and formed by hydrophobic residues to force out any water molecules and to attract certain small amphipathic compounds [13]. Exploration of the cavities in the three isotypes was conducted by determining the cavities and voids on the proteins' surfaces computationally using geometric triangulation. The other two factors; the important residues in the contact and the positions of ligands and the corresponding conformations, were tested after conducting virtual screening with molecular docking. The residues, which are critical for hydrogen bonding, were determined statistically by their probability distributions, based on the observed frequencies of involvement in ligand binding. The

residues of high probability distributions were considered the most important in the contacts essential for ligand binding. The positions were tested visually by observing the locations of the docked ligands and determining whether they are well-positioned in the binding pocket. The quality of the positions of ligands and the complex conformations were evaluated by calculating the free energy of binding which is the measure of the binding fitness.

Virtual screening exploits the large libraries of chemical structures to identify those structures which are most likely to bind to a target. Moreover, it gives some information about the binding nature and whether it is good hit to be considered for further investigation.

The study was mainly focused on identifying pan-PPAR ligands from lead-like compounds, which are classified as such based on their molecular weight, amenability for chemical optimization and other properties [14]. This study does not cover toxicity and other physicochemical properties of the identified ligands, which require laboratory investigations. However, the identified ligands offer the opportunity for further research on drug discovery.

CHAPTER 2

LITERATURE REVIEW

2.1 Peroxisome Proliferator-Activated Receptors (PPARs)

The peroxisome proliferator-activated receptors (PPARs) were given their name because the first discovered member of these receptors, the mouse PPARα, was shown to be activated by a diverse group of compounds that causes the proliferation of peroxisomes. PPARs are members of the nuclear receptor (NR) superfamily that have evolved only in metazoan animals as transcriptional factors that mediate a variety of metabolic processes [15]. Three different isotypes of PPARs, known as PPARα, PPARγ, and PPARδ, have been identified. They show distinct tissue distributions, physiological roles, and distinct but overlapping ligand specificity [16]. PPARs are ligand-activated transcription factors that function as gene regulators by binding to some small molecules such as hormones or dietary components [17]. They regulate both target gene expression and repression when they bind to a ligand [18]. The most conserved domains of this protein subfamily are the DNA binding domain (DBD) and the ligand binding domain (LBD) that, in addition to binding to ligands, is required for dimerization and interaction with transcriptional co-factors [19, 20]. They play important roles in lipid and glucose homeostasis, regulation of cellular differentiation, and tumorigenesis [21]. Although PPARs can form homodimers, they must heterodimerize with retinoid X receptor (RXR) to carry out most of their functions [22]. Despite their structural similarities, each member of the PPAR family is localized to certain parts of the body. Location of receptors partially determines their function in the body and the different roles they can

play in medicine as drug targets. PPARα is localized in liver, kidney, heart, and muscles and is important for the uptake and oxidation of fatty acids and lipoprotein metabolism [23]. Therefore, it is the target for the lipid lowering fibrates [24]. PPARγ is found in adipose tissue, large intestine, and macrophages [25]. It plays an important role in adipocyte differentiation as well as a receptor for the well-known class of anti-diabetic insulin sensitizer drugs, the thiazolidinediones (TZD) [26]. PPARδ is found in most cell types and it can bind to a number of agonists that play important roles in dyslipidemia, cancer treatment, and cell differentiation in the central nervous system [27].

2.2 Structure of PPARs

The structure of PPARs consists of five regions that differ in the degrees of homology. These regions include, from N-terminal to C-terminal, the region (A/B), a DNA binding domain (C), a hinge region (domain D), and a ligand binding domain (E and F) (Figure 2.1).

The A/B domain is highly variable among the members of the PPAR family. It has a ligand-independent transcriptional activation function, which is referred to as AF-1. The AF-1 mediated gene activation often displays promoter-dependent and cell-type-dependent specificities, indicating that this domain may be responsible for interactions with cell-specific co-regulatory proteins.

The C- region is identified as the most conserved region. It contains the DNA binding domain (DBD) which is composed of two zinc finger modules. These two zinc modules are the characteristic feature of the DNA binding motif that distinguishes the

nuclear receptor from other DNA binding proteins. Each zinc finger contains a group of four Cysteine (Cys) residues which co-ordinate a single zinc atom [28].

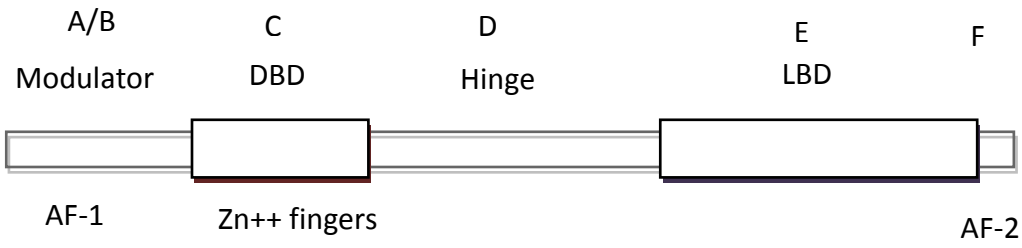


Figure 2.1. A typical nuclear receptor with its five characteristic domains

Domain D, the hinge region that connects domains C and E, is not well conserved and varies significantly in length among the nuclear receptors. This domain participates in DNA rotation and co-repressor interaction.

Domain E is the second most conserved region. It includes the ligand-binding domain (LBD), which is made up of 13 α-helices and four short β-sheets. The helices and the sheets are organized in three layers to form a hydrophobic pocket, where the ligand binds. The PPAR structures obtained with X-ray crystallography show an exceptionally spacious ligand-binding pocket compared to other nuclear receptors, which explains the promiscuity of the PPAR ligand binding site.

The C-terminal region is a multifunctional domain that mediates ligand binding, receptor dimerization, and transcriptional activation or repression [28].

As shown by Uniprot database [29], the human PPARα consists of 468 amino acids (Figure 2.2) organized in 14 helices and 3 sheets (Tables 2.1 and 2.2). The PPARα DBD is localized between the residues 99 and 173. The DBD contains two zinc finger

motifs at the residual location 102 - 122 and 139 – 161. The LBD extends over 189 amino acids from residue 280 to residue 468.

```
        10         20         30         40         50         60
MVDTESPLCP LSPLEAGDLE SPLSEEFLQE MGNIQEISQS IGEDSSGSFG FTEYQYLGSC

        70         80         90        100        110        120
PGSDGSVITD TLSPASSPSS VTYPVVPGSV DESPSGALNI ECRICGDKAS GYHYGVHACE

       130        140        150        160        170        180
GCKGFFRRTI RLKLVYDKCD RSCKIQKKNR NKCQYCRFHK CLSVGMSHNA IRFGRMPRSE

       190        200        210        220        230        240
KAKLKAEILT CEHDIEDSET ADLKSLAKRI YEAYLKNFNM NKVKARVILS GKASNNPPFV

       250        260        270        280        290        300
IHDMETLCMA EKTLVAKLVA NGIQNKEAEV RIFHCCQCTS VETVTELTEF AKAIPGFANL

       310        320        330        340        350        360
DLNDQVTLLK YGVYEAIFAM LSSVMNKDGM LVAYGNGFIT REFLKSLRKP FCDIMEPKFD

       370        380        390        400        410        420
FAMKFNALEL DDSDISLFVA AIICCGDRPG LLNVGHIEKM QEGIVHVLRL HLQSNHPDDI

       430        440        450        460        470
FLFPKLLQKM ADLRQLVTEH AQLVQIIKKT ESDAALHPLL QEIYRDMY
```

Figure 2.2. The sequence of human PPARα (Uniprot ID: Q07869)

Table 2.1

PPARα regions

Region	Location	Number	Description
DNA binding	99 – 173	75	Nuclear receptor
Zinc finger	102 – 122	21	NR C4-type
Zinc finger	139 – 161	23	NR C4-type
Region	280 – 468	189	Ligand-binding
Region	304 – 433	130	Required for heterodimerization with RXRA

Table 2.2

PPARα secondary structures

Secondary structure	Location	Number of residues
Helix (H1)	201 – 217	17
Helix (H2)	222 – 228	7
Beta strand	239 – 241	3
Helix (H3)	244 – 251	8
Helix (H4)	268 – 292	25
Helix (H5)	302 – 321	20
Helix (H6)	322 – 324	3
Beta strand	329 – 332	4
Helix (H7)	333 – 335	3
Beta strand	337 – 340	4
Helix (H8)	341 – 346	6
Helix (H9)	351 – 353	3
Helix (H10)	356 – 366	11
Helix (H11)	372 – 383	12
Helix (H12)	394 – 415	22
Helix (H13)	422 – 450	29
Helix (H14)	458 – 467	10

The human PPARδ is composed of 441 amino acids that form 14 helices, 3 sheets, and a turn (Figure 2.3, Tables 2.3 and 2.4). The PPARδ DBD extends between residue 71 and residue 145. The zinc finger motifs are located at 74 – 94 and 111 – 133. The LBD extends over 189 residues between residue 280 and residue 468.

```
        10         20         30         40         50         60
MEQPQEEAPE VREEEEKEEV AEAEGAPELN GGPQHALPSS SYTDLSRSSS PPSLLDQLQM

        70         80         90        100        110        120
GCDGASCGSL NMECRVCGDK ASGFHYGVHA CEGCKGFFRR TIRMKLEYEK CERSCKIQKK

       130        140        150        160        170        180
NRNKCQYCRF QKCLALGMSH NAIRFGRMPE AEKRKLVAGL TANEGSQYNP QVADLKAFSK

       190        200        210        220        230        240
HIYNAYLKNF NMTKKKARSI LTGKASHTAP FVIHDIETLW QAEKGLVWKQ LVNGLPPYKE

       250        260        270        280        290        300
ISVHVFYRCQ CTTVETVREL TEFAKSIPSF SSLFLNDQVT LLKYGVHEAI FAMLASIVNK

       310        320        330        340        350        360
DGLLVANGSG FVTREFLRSL RKPFSDIIEP KFEFAVKFNA LELDDSDLAL FIAAIILCGD

       370        380        390        400        410        420
RPGLMNVPRV EAIQDTILRA LEFHLQANHP DAQYLFPKLL QKMADLRQLV TEHAQMMQRI

       430        440
KKTETETSLH PLLQEIYKDM Y
```

Figure 2.3. The sequence of human PPARδ (Uniprot ID: Q03181)

Table 2.3

PPARδ regions

Region	Location	Number	Description
DNA binding	71 – 145	75	Nuclear receptor
Zinc finger	74 – 94	21	NR C4-type
Zinc finger	111 – 133	23	NR C4-type
Region	254 – 441	188	Ligand-binding

Table 2.4

PPARδ secondary structures

Secondary structure	Location	Number of residues
Helix (H1)	170 – 173	4
Helix (H2)	175 – 189	15
Helix (H3)	194 – 200	7
Beta strand	211 – 213	3
Helix (H4)	216 – 223	8
Turn	224 – 226	3
Helix (H5)	241 – 264	24
Turn	268 – 272	5
Helix (H6)	275 – 294	20
Helix (H7)	295 – 297	3
Beta strand	302 – 305	4
Helix (H8)	306 – 308	3
Beta strand	310 – 313	4
Helix (H9)	314 – 318	5
Helix (H10)	322 – 339	18
Helix (H11)	345 – 356	12
Helix (H12)	367 – 388	22
Helix (H13)	395 – 423	29
Helix (H14)	431 – 437	7

The Human PPARγ is the longest isotype as it consists of 505 amino acids forming 14 helices and 4 sheets (Figure 2.4, Tables 2.5 and 2.6). The location of the DBD is between the residue 136 and residue 210 while the zinc finger motifs are located at 139 – 159 and 176 – 198. The PPARγ LBD consists of 189 residues from residue 254 to residue 441.

```
        10          20          30          40          50          60
MGETLGDSPI DPESDSFTDT LSANISQEMT MVDTEMPFWP TNFGISSVDL SVMEDHSHSF

        70          80          90         100         110         120
DIKPFTTVDF SSISTPHYED IPFTRTDPVV ADYKYDLKLQ EYQSAIKVEP ASPPYYSEKT

       130         140         150         160         170         180
QLYNKPHEEP SNSLMAIECR VCGDKASGFH YGVHACEGCK GFFRRTIRLK LIYDRCDLNC

       190         200         210         220         230         240
RIHKKSRNKC QYCRFQKCLA VGMSHNAIRF GRMPQAEKEK LLAEISSDID QLNPESADLR

       250         260         270         280         290         300
ALAKHLYDSY IKSFPLTKAK ARAILTGKTT DKSPFVIYDM NSLMMGEDKI KFKHITPLQE

       310         320         330         340         350         360
QSKEVAIRIF QGCQFRSVEA VQEITEYAKS IPGFVNLDLN DQVTLLKYGV HEIIYTMLAS

       370         380         390         400         410         420
LMNKDGVLIS EGQGFMTREF LKSLRKPFGD FMEPKFEFAV KFNALELDDS DLAIFIAVII

       430         440         450         460         470         480
LSGDRPGLLN VKPIEDIQDN LLQALELQLK LNHPESSQLF AKLLQKMTDL RQIVTEHVQL

       490         500
LQVIKKTETD MSLHPLLQEI YKDLY
```

Figure 2.4. The sequence of human PPARγ (Uniprot ID: P37231)

Table 2.5

PPARγ regions

Region	Location	Number	Description
Chain	1 – 505	505	PPARγ
DNA binding	136 – 210	75	Nuclear receptor
Zinc finger	139 – 159	21	NR C4-type
Zinc finger	176 – 198	23	NR C4-type
Region	205 – 280	76	Interaction with FAM120B By similarity
Region	317 – 505	189	Ligand-binding

Table 2.6

PPARγ secondary structures

Secondary structure	Location	Number of residues
Helix (H1)	236 – 253	18
Helix (H2)	258 – 266	9
Beta strand	275 – 277	3
Helix (H3)	280 – 287	8
Helix (H4)	289 – 292	4
Beta strand	301 – 303	3
Helix (H5)	305 – 329	25
Helix (H6)	339 – 358	20
Helix (H7)	359 – 361	3
Beta strand	364 – 369	6
Turn	370 – 373	4
Beta strand	374 – 377	4
Helix (H8)	378 – 383	6
Helix (H9)	388 – 390	3
Helix (H10)	393 – 403	11
Helix (H11)	409 – 420	12
Helix (H12)	431 – 452	22
Helix (H13)	459 – 487	29
Helix (H14)	495 – 501	7

2.3 Ligand Binding Domain of PPARs

The ligand effects on PPARs are mediated through the LBD, a region of 189 amino acid residues at the C-terminal end of the receptor. Besides ligand binding, the LBD also contains the transcriptional activation function 2 (AF-2), which consists of the

residues that play a key role in dimerization and trans-activation [28]. The structures of the three isotypes are very similar. The LBD is composed of 13 α-helices and four strands of ß-sheet. The ligand binding pocket is Y-shaped and it consists of a mouth opening and two arms. The PPAR's ligand binding pocket is about 1400$Å^3$, which allows the PPARs to interact with a number of different ligands. There are conserved polar resides, which are involved in a hydrogen bonding network that interacts with the ligands upon binding [30]. The conserved residues that form a hydrogen bonding network in one of the two arms assist in holding the AF2-helix in an active conformation to promote co-activator binding. The second arm is highly hydrophobic and is thus ideal for binding the hydrophobic tail of fatty acids via van der Waals interactions (VDW). It has been found that about 80% of the LBD residues are conserved across PPAR isotypes, while the remaining 20% of the residues create the ligand specificity of each isotype [31].

2.3.1 PPARα LBD. The ligand binding site in PPARα is situated in a large cavity and guarded by helices H3, H5, H7, H11, and H12 (Figure 2.5). The cavity spans the LBD between the AF2 helix and the 3-stranded β- sheet. The cavity splits upward and downward at the level of the β-sheet, along an axis parallel to helix H3, forming upper and lower distal cavities. The mouth opening to cavity is found between helix H3 and the 3-stranded anti-parallel β-sheet. PPARα does not bind to ligands with large carboxylate head groups because in place of TYR-314 , PPARγ has a smaller equivalent residue, HIS-351. The loop is highly flexible and partly covers the opening to the ligand-binding site. The mouth opening is further guarded by TYR-334 which forms a hydrogen bond with GLU-282 [32].

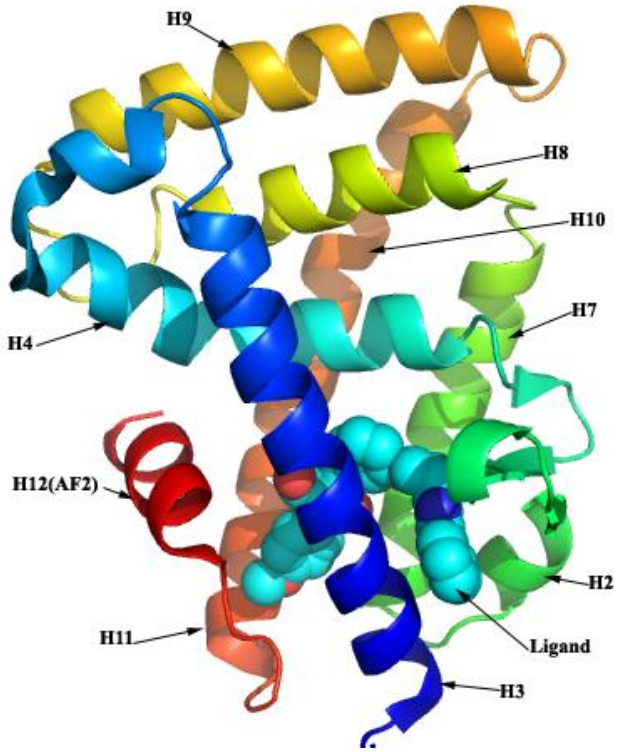


Figure 2.5. PPARα LBD in complex with a ligand

2.3.2 PPARδ LBD. The PPARδ LBD consists of 12 α-helices and a three-stranded β-sheet (Figure 2.6). The elements of the secondary structure create a large ligand-binding cavity of an approximate volume of 1300 $A^{\circ 3}$. The significantly narrower cavity adjacent to the AF-2, helix 12, prevents PPARδ from binding to the large headed TZDs [33].

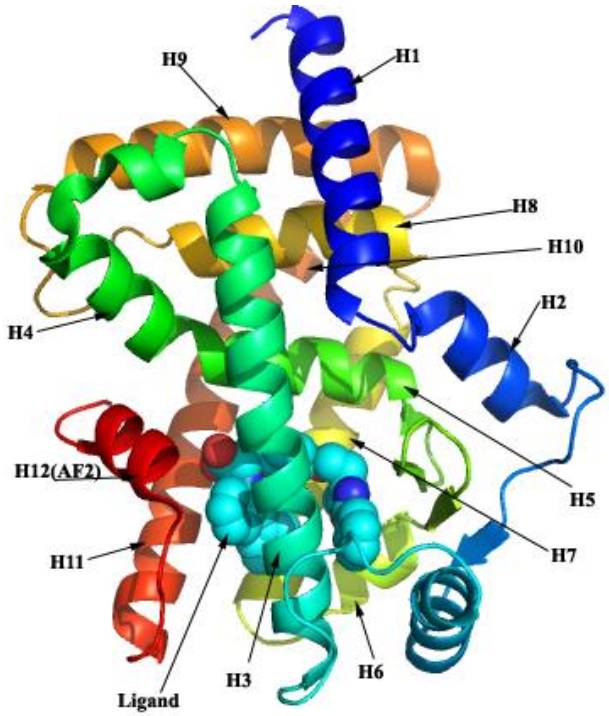


Figure 2.6. PPARδ LBD in complex with a ligand

2.3.3 PPARγ LBD. PPARγ LBD is comprised of three layers made up of 13 helices and a small four-stranded β-sheet (Figure 2.7). It is very similar to that of PPARα, with the exception of an extra helix, called H2`, located between the first β-strand and H3. The two outer layers of the LBD are formed by the three long helices (H3, H7, and H10/H11). The region between α-helix 1 and α-helix 3 contains an insertion of more than 20 amino acid residues forming two helices, 2a and 2b, separated by a β-strand. The middle layer is formed by helices (H4, H5, H8, and H9) and it occupies only the top half of the domain leaving very large cavity at the bottom half of the LBD (1400 $Å^3$). This LBD pocket is relatively large and has a three-arm Y shape, allowing ligands of multiple branches or a single branch to bind to the LBD in multiple conformations. PPARγ binds to ligands with large carboxylate head groups because of

the smaller HIS-351 that allows the ligand to enter the binding pocket [34]. On the lower half of the LBD, the right-hand side is closed by a two-stranded β-sheet and, on the left-hand side, by the short C-terminal α-helix (H12) of the receptor, which constitutes the receptor's ligand-dependent AF-2 region. Helix 12 covers the ligand-binding pocket and changes its conformation upon binding with ligands. This helix forms hydrophobic interactions with the rest of the receptor, a weak hydrogen bond between HIS-351 and TYR-501, and a salt bridge between LYS-347 and ASP-503. This salt bridge is critical for transcriptional activation and has a stabilizing effect on the active LBD [35].

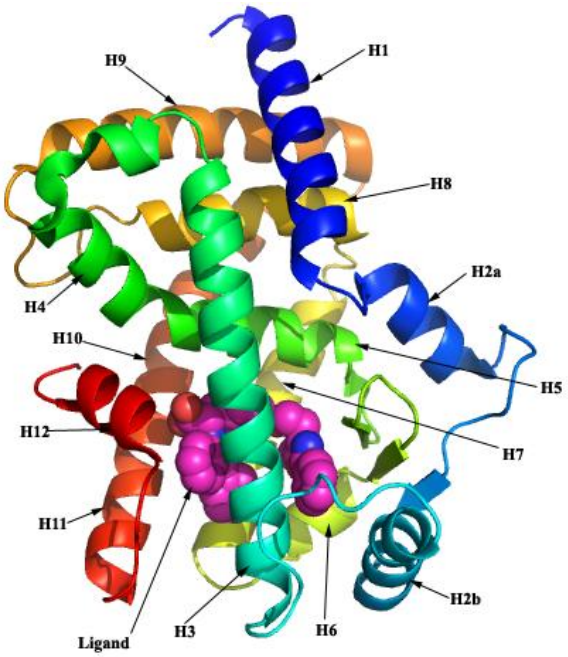


Figure 2.7. PPARγ LBD in complex with a ligand

2.3.4 AF-2 domain. The AF-2 domain is essential for ligand binding to the PPAR LBD and the function of the receptor (Figure 2.8). When a ligand binds to the LBD, helix H12 of AF-2 closes on the ligand-binding concavity, reducing conformational

flexibility of the LBD and forming a structure that is ideal for co-activator binding. It has been determined that residues GLU-352, ARG-425, ARG-471, and TYR-505 (in PPARγ) are involved in a hydrogen bonding network that stabilizes the AF-2 helix in an active conformation upon ligand binding [36, 28].

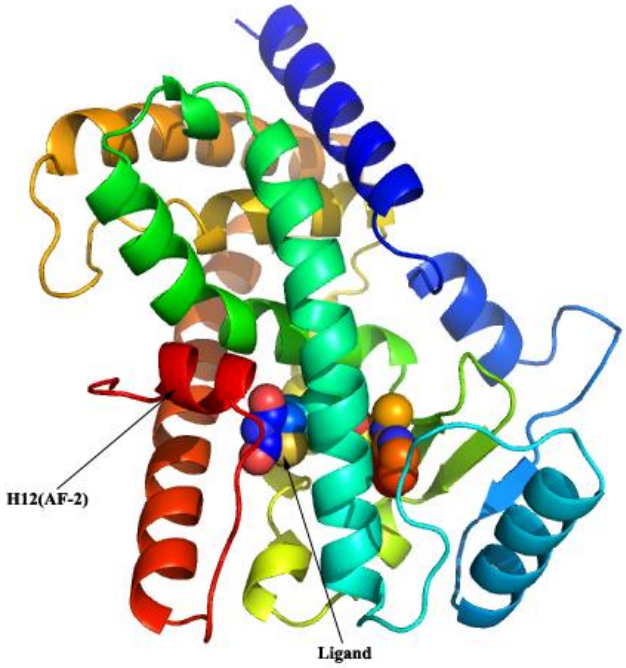


Figure 2.8. PPARγ LBD with AF-2 domain and a ligand

2.4 PPAR DNA Binding Domain

The DBD (Figure 2.9) is located in the center of the receptor and is comprised of highly conserved residues, two zinc-binding sites and sequence specific residues that direct the binding of PPAR to a consensus DNA sequence AGGTCA. This consensus sequence designates the peroxisome proliferator DNA response element (PPRE) [36, 28]. It has been shown that the DNA PPRE conformation contributes to the binding of PPAR

to DNA via a head-to-tail interaction between the PPAR DBD and RXR DBD using residues GLN-206 and ARG-209 on RXRα and ASN-188 on PPARγ [37, 38].

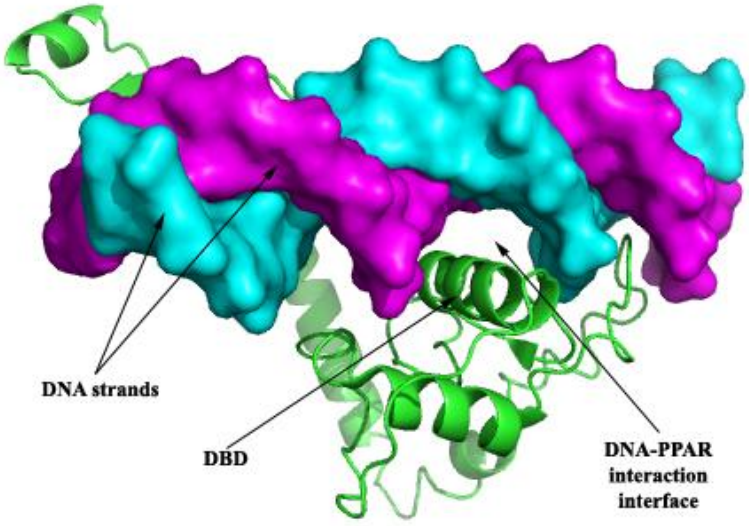


Figure 2.9. PPAR DBD binding to DNA

2.5 PPAR Ligands

PPARs are activated by the peroxisome proliferators (PPs), a group of chemicals such as the fibrate drugs, WY-14, 643, and plasticizers. In addition, PPARs are also activated by endogenous fatty acids and their metabolic products. PPs are similar in their chemical structures to endogenous and dietary fatty acids and their metabolites. Both PPs and fatty acids contain carboxylic functional groups and a hydrophobic tails. Many ligands of the PPARs have been identified; some of them are able to activate all three isotypes with varying efficacy and potency while others exhibit isotype specificity [39, 40].

A group of synthetic drugs such as troglitazone, pioglitazone, and rosiglitazone have been developed to treat and control metabolic diseases like diabetes. These drugs function similarly to the endogenous ligands that bind to the PPARγ LBD to activate the receptor [41]. It has been found that rosiglitazone creates hydrogen interactions with its TZD group and HIS-351 and HIS-477. The sulfur atom of the TZD occupies a hydrophobic pocket formed by PHE-391, GLN-314, PHE-310, LEU-358, ILE-354 and LEU-497, and the central benzene ring occupies a pocket formed by CYS-313 and MET-392 [42]. PPARγ regulates the genes responsible for lipid metabolism and homeostasis and fatty acid transport as well as the genes that function in insulin signaling and glucose transport [43]. PPARα has been found to play a significant role in the regulation of uptake and oxidation of fatty acids [44]. Therefore, PPARα became an important target for atherosclerosis drugs because it reduces LDL cholesterol and increases HDL cholesterol. The fibrates, which are a group of carboxylic acids, can bind to PPARα as agonists. So, they can be used as drugs to treat hypercholesterolemia and hyperlipidemia [45]. PPARδ is expressed all over the body and might play a role in a number of diseases and disorders including cancer. Synthetic PPARδ agonists are used in the treatment of some diseases of central nervous system [46].

Pan-PPAR agonists are a class of compounds that are capable of activating the three PPAR isotypes simultaneously. Recent studies suggest that pan-agonists offer the potential to improve the treatment of metabolic syndrome, a group of clinical conditions that include hypertension, obesity and type 2 diabetes mellitus. Bezafibrate is a clinically tested pan-PPAR ligand that acts as a lipid-lowering drug. It has been shown to increase

high density lipoprotein, lower triglycerides, enhance insulin sensitivity, and reduce blood glucose level. Therefore, bezafibrate significantly reduces the risks of cardiovascular conditions and metabolic syndrome but also it has been reported that it has a number of side effects including muscle problems [47].

The tables 2.7 - 2.12 [48] contain the best known endogenous and exogenous ligands for the three PPAR isotypes. These ligands differ in their binding affinities and they range from very strong ligands to weak or partial ligands.

Table 2.7

PPARα exogenous ligands

Ligand description
• Fibrate and nonfibrate hypolipidemic drugs (bezafibrate, clofibrate, ciprofibrate, fenofibrate, gemfibrozil, nafenopin, Wy-14643) • Phytanic acid • Indomethacin • Dehydroepiandrosterone (DHEA) • Phthalates (DEHP, MEHP) • MK-886a • Valproic acid (VPA) • Telmisartan • Phytol • Pefluorinated fatty acids (PFOA, PFDA), perfluorosulfonic acid (PFOS) • Phenobarbital (PB) • Oxirane compounds • ETYA • Epoxyisoprostane • DRF-2519 • Bm 17.0744 • Benz[a]anthracene Di- and trichloroacetic acid (DCA, TCA)

Table 2.8

PPARα Endogenous ligands

Ligand description
• LPL-treated VLDL • VLDL • 5,6-epoxyeicosatrienoic acids (EET); 8,9-EET; 11,12-EET; 14,14-EET; 20-hydroperxyeicosatetraenoic acid (20-HETE), 20,14,15-hydroxyepoxyeicosatrienoic acids (20,14,15-HEET) • 2-arachidonylglycerol (2-AG); 15-S-hydroxyeicosatetraenoic-glycerol ester (15-S-HETE-G) • Long-chain alkylamines • Saturated and unsaturated fatty acids • PGD2, PGD1 • Leukotriene B4 (LTB4)

Table 2.9

PPARδ exogenous ligands

Ligand description
• Tetradecylthioacetic (TTA) • Wy-14643 • VPA • Bezafibrate • Sulindac sulfidea • Benz[a]anthracene • Treprostinil sodium • GW-501516

Table 2.10

PPARδ endogenous ligands

Ligand description
• LPL-treated VLDL • VLDL • Mono- and polyunsaturated fatty acids from triglycerides • Saturated and unsaturated fatty acids • Prostaglandin A1 (PGA1), PGD2, PGD1 • oxVLDL, 13-S-hydroxyoctadecadienoic acid (13-S-HODE), 4-hydroxynonenol (4-HNE)

Table 2.11

PPARγ exogenous ligands

Ligand description
• Thiazolidinediones (ciglitazone, pioglitazone, rosiglitazone, troglitazone, N-(2-[4-[2,4-dioxo (1,3-thiazolidin-5-yl)methyl]phenoxy]ethyl)-5-(1,2-dithiolan-3-yl)-N-methylpentanamide) • GW9662a • Indomethacin • Phthalate esters (MEHP, DEHP) • Bisphenol A diglycidyl ether (BADGE)a • Wy-14643 • Glimepiride • Diclofenac • Anandamide • Sulindac • JTP-426467 • Pemoline • Phenylacetate • Nimesulide • Curcumin • 2-bromopalmitate • Tolbutamide, chlorpropamide, gliclazide, Glibenclamide

Table 2.12

PPARγ endogenous ligands

Ligand description
• OxLDL, 9-S-hydroxyoctadecadienoic acid (9-S-HODE), 13-S-HODE • LPL-treated VLDL • 15-S-HETE • Lysophosphatidic acid (LPA) • Hexadecylazelaicphosphatidylcholine (AzPC) • 13-S-HODE, 15-S-HETE, 5-S-HETE, 12-S-HETE • Polyunsaturated acids including linoleic acid, linolenic acid, arachidonic acid, and eicosapentaenoic acid (EPA) • PGD1, PGD2, PGA1 • Nitroalkene derivatives of linoleic acid (LNO2)

2.6 Heterodimerization

Once a PPAR is activated, it binds to DNA as a heterodimeric complex with retinoic X receptor (RXR). The PPAR-RXR complex regulates the gene expression by interacting with the PPREs, which are located close to the target genes. There are several genes that contain these PPRE motifs including acyl-CoA oxidase, peroxisomal bifunctional enzyme, liver fatty acid-binding protein, microsomal CYP4A, and others. Members of the RXR-interacting subgroup of NRs typically bind to DNA elements containing two copies of a direct repeat array spaced by 1–6 nucleotides (DR1–DR6). The consensus binding site (AGGTCA) is similar for all isotypes. The interaction of PPAR and RXR directs the complex to bind to DR1 motifs. However, if PPAR interacts with another transcription factor, this complex may no longer associate with DNA and the interacting protein may be removed from its normal site of action [49].

2.7 PPAR Cofactors and Corepressors

Transcriptional coactivators and corepressors are a group of protein that interacts with nuclear receptors to repress or enhance their transcriptional activities [50]. In the absence of ligand, PPARs may bind to corepressors, which decrease the activity of the receptor [51]. Ligand binding induces a conformational change in the nuclear receptor that favors binding to coactivators. The receptor-coactivator complex can then activate gene transcription by means of recruiting chromatin-modifying enzymes (acetyltransferases) or by forming a 'bridge' with the pre-initiation complex at the hormone-regulated promoter [52, 53]. The PPAR coactivators such as PCG-1 contain a nuclear receptor interaction domain (LXXLL) and possess enzymatic activity (histone acetyltransferase, HAT) [54].

2.8 The Mechanism of PPAR Transcriptional Regulation

The PPAR responds to ligands in processes that lead to the expression of a number of target genes. The ligand binding to PPAR initiates a signaling pathway that leads to dimerization of PPAR with retinoid X receptor (RXR). The target genes are activated through direct binding of the PPAR-RXR heterodimer to the PPREs. A PPRE region may be found as multiple copies close to the promoter region of a target gene [55]. Once a ligand binds in the pocket of PPAR LBD, favorable interaction forces between the ligand and helix 12 firmly place helix 12 over the binding pocket, closing the pocket off from interaction with water molecules. In this way, the helix 12 acts like a trap door, it closes once the ligand enters the binding pocket. This conformational change leads to

the activation of PPAR targets by increasing the affinity of nuclear receptors for coactivators [56].

2.9 Computational Ligand-Protein Docking

Docking involves finding the most favorable binding modes of a ligand to the target of interest. The binding mode of a ligand with respect to the receptor can be uniquely defined by its state variables. These consist of its position, orientation, and conformation [57, 58]. Computational docking is used to predict the binding modes of two or more molecules. It uses molecular dynamics force fields to calculate the free energy of binding of the ligand-protein complex and a search method such as genetic algorithm to identify all possible conformations. Docking is defined as a multi-step process in which each step introduces one or more additional degrees of complexity. The docking begins with the application of docking algorithms that position small molecules or probes in the active site. Molecules may contain many conformations [58]. The sampling of these conformations can be performed to determine the identification of the best conformation. Algorithms use scoring functions that are designed to predict the biological activity through the evaluation of interactions between compounds and potential targets [57]. The docking scoring function depends on estimation of electrostatic and van der Waals interactions. The techniques used in molecular docking can be classified based on the three types of protein representations; atomic, surface, or grid. AutoDock uses GRID to calculate interaction energy and to perform the docking process [57, 60].

$$L + R \overset{\Delta G_{Binding}}{\longleftrightarrow} LR$$

Binding of a ligand (L) to a receptor (R) to form a ligand-receptor complex (LR) only takes place if the generated energy of interaction ($\Delta G_{binding}$) overcomes the repulsive van der Waals forces. Molecular interaction fields are used to investigate the energetic conditions that arise between molecules approaching each other. The molecular fields describe the variation of interaction energy between a target molecule and a probe moved in 3D grid, which has been set around the target. The probes reflect the chemical characteristics of a binding partner. GRID, a program used by AutoDock, can calculate the molecular interaction fields from Cartesian coordinates. The interaction energy is calculated on a grid of points surrounding the target molecule. At each grid point the interaction energy between the probe and the target is calculated using the following empirical energy function [61]:

$$\Delta G_{binding} = E_{vdw} + E_{el} + E_{hb} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

where $\Delta G_{binding}$ is the total interaction energy, E_{vdw} is the van der Waals interaction energy, E_{el} is the electrostatic energy, and E_{hb} is the interaction energy due to hydrogen bond formation.

The repulsive force of the van der Waals energy can be calculated by an empirical energy function. For non-polar molecules the balance between the attractive forces and the short-range repulsive forces of van der Waals energy can be estimated with the Lennard-Jones potential [62].

$$E_{vdw} = \sum_{j=1}^{N} \sum_{i=1}^{N} 4\varepsilon \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] \ldots\ldots\ldots (2)$$

where ε is the well depth, σ is the collision diameter of the respective atoms *i* and *j*, and *r* is the distance between two atoms.

The electrostatic term of the energy function is calculated by Coulomb equation[62].

$$E_{coul}(r) = \sum_{i=1}^{N_A} \sum_{j=1}^{N_B} \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \quad \text{..... (3)}$$

where *N* is the number of atoms in molecules A and B, respectively, and *q* the charge on each atom, ε_0 is the permittivity of free space, which is an experimentally determined quantity with the approximate value 8.854×10^{-12} C^2 N^{-1} m^{-2}, and *r* is the distance between two atoms.

A hydrogen bond is an interaction between a positively charged hydrogen atom and an electronegative acceptor atom. The resulting distance between acceptor and donor atom is less than the sum of their van der Waals radii. In contrast to other non-covalent forces, the hydrogen bonding interaction is directional as it depends on the orientation of the acceptor atom. The GRID method uses an explicit energy term in experimental crystallographic data for hydrogen bonds**.** The functional form of this term has been developed to fit experimental data [62].

In the grid, parameters for each type of atom in the molecule ar defined. The parameters describe the strength of the non-bonded interactions namely; van der Waals, the electrostatic, and the hydrogen bond interactions [63].

The free energy of binding for the formation of a protein-ligand complex can also be expressed with the following equation [64]:

$$\Delta G_{bind} = \Delta H - T\,\Delta S \quad \text{..... (4)}$$

where *ΔH* represents the enthalpy, *T* represents the temperature in Kelvin, and *ΔS* represents the entropy. Under equilibrium conditions

$$L + R \overset{K_{eq}}{\Longleftrightarrow} LR$$

The free energy of binding can be expressed as follows:

$$\Delta G° = \Delta H° - T\,\Delta S° = - R\,T \text{ Log } (K\text{eq}) \quad \ldots. (5)$$

where R is constant and Keq is the kinetic equilibrium constant.

2.10 Drug-like Compounds and Lead-like Compounds

In drug discovery, a number of properties have been outlined for orally deliverable drug-like molecules. These properties include appropriate molecular weight, ionization constants, lipophilicity, polar surface area, and number of hydrogen donors or acceptors. The drug-like molecule must be able to cross the cell membrane, while retaining the ability to be transported in plasma [65]. Lipinski formulated the so-called "rule of 5" as the general guidelines for boundary definition for orally administered drug-like compounds. These guidelines include the following criteria: 1) the molecular weight of the orally administered drug must not be more than 500; 2) the maximum number of hydrogen bond donors in the molecule must not exceed 5; 3) the maximum number of hydrogen bond acceptors in the molecule must not exceed 10; 4) the logarithm of partition coefficient between octanol and water (LogP) should not be more than 5; and finally 5) the above-mentioned four rules only apply to passive transport. These criteria recognize that penetration through cell membranes is accomplished only rarely by molecules of high molecular weight, unless there is an active transport mechanism [66].

High throughput screening of the chemical compounds can lead to the discovery of lead compounds with higher MW, higher lipophilicity, and lowered solubility which can be optimized and enhanced as s clinical candidates [67]. Thus criteria for lead-

likeness must include validated biological activity in screens, that is patentable, and have a good initial Drug Metabolism and PharmacoKinetics profile (DMPK). Property ranges for lead-like compounds can be defined as: 1-5 rings, 2-15 rotatable bonds, MW less than 400, up to 8 acceptors, up to 2 donors, and a logP range of 0.0 to 3.0. The average differences in comparisons between drugs and leads include 2 less rotatable bonds, MW 100 lower, and a reduction in logP of 0.5 to 1.0 log units. Thus, one of the key objectives in the identification of lead-like compounds for screening is the need for smaller, less lipophilic compounds that, upon optimization, will yield compounds that still have drug-like properties [68].

CHAPTER 3

MATERIALS AND METHODS

3.1 PPAR LBD Structure Files

Protein Data Bank files (PDB) for PPARs were obtained from Protein Data Bank repository. The PDB files provide standard representation for macromolecular structure data derived from X-ray diffraction. A PDB file representing the structure of the LBD was selected for each of the three PPAR isotypes. The selected PDB files were 1I7G (PPARα) [69], 2XYW (PPARβ/δ) [70], and 2ZK0 (PPARγ) [71].

3.2 Surface Topography of PPARs

The ligand-binding on a receptor requires cavities on the receptor surface as well as specific amino acid positioning within it that create the physicochemical properties needed for a LBD to perform its function. Therefore, exploring the cavities on the surfaces of the PPARs is important in the virtual screening to determine whether the docking of the ligands makes sense in terms of location and chemical bonding. To explore the surface topography of the PPARs, CASTp (Computed Atlas of Surface Topography of proteins), was used to locate and measure pockets and voids on the PPAR LBDs. CASTp is based on recent theoretical and algorithmic results of computational geometry to identify the cavities analytically [72, 73, 74]. CASTp is capable of identifying and measuring pocket and pocket mouth openings using a series of computational geometry methods, including Delaunay triangulation, alpha shape, discrete flow, and Voronoi diagram [74, 75]. The program starts the identification of the cavities by finding the polygon of the convex hull of the molecule that encloses all atoms. The

polygon was triangulated by the Delaunay method. Then the Voronoi diagram was formed from the Voronoi cells. The alpha shape was formed by omitting the edges and vertices outside the molecule (Figure 3.1). Once the alpha shape was formed, the area, volume, and cavities were measured. The cavities represent the empty triangles [75].

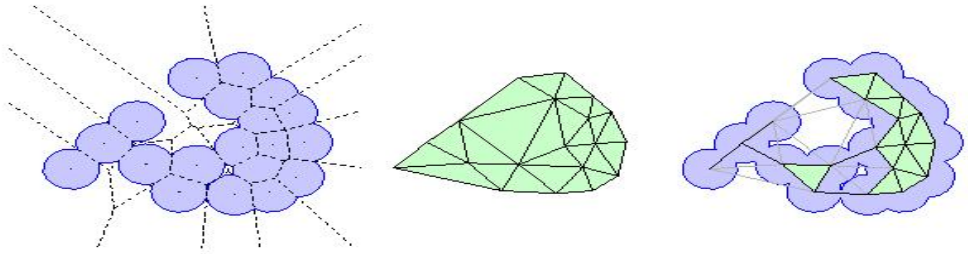


Figure 3.1. A Voronoi diagram and Delaunay triangulation of a molecule

The PDB file for each PPAR LBD was examined with CASTp. The cavities of the LBDs were obtained and visualized with PyMol CASTp plug-in [76] as shown in Figure 3.2.

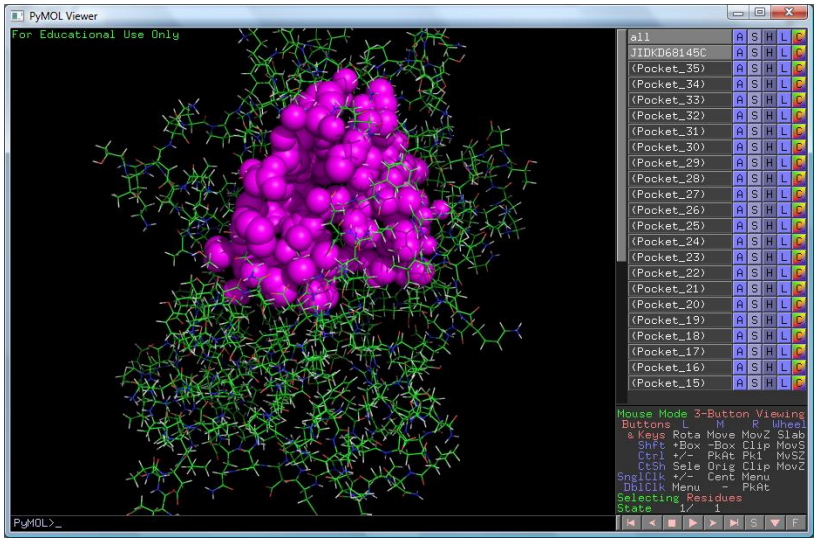


Figure 3.2. PyMol with CASTp plug-in showing a cavity

3.3 Virtual Screening with Molecular Docking

Virtual screening with molecular docking was used to search for pan-PPAR ligands from more than four million compounds. Screening was performed through different processes that include ligand library filtering, ligand preparation, PPAR preparations, grid parameter file (GPF) preparation, docking parameter file (PDF) preparation, and generation of python docking script that perform calculation of grid interaction fields and implementation of a search algorithm to search for the top hits with the minimal binding energy. The entire process was performed under the linux platform to access the open source programs such as python and AutoDock. The computational cost of the virtual screening conducted in this study was inexpensive and performed in two-week run-time using a multi-core personal computer of 4 GB random access memory (RAM).

The virtual screening encompassed 4 separate procedures: the PPAR preparation, compound preparation, docking, and docking results evaluation. These four procedures were further split into different steps. The flow chart in Figure 3.3 summarizes the different procedures and steps that were followed during the virtual screening processes in this study.

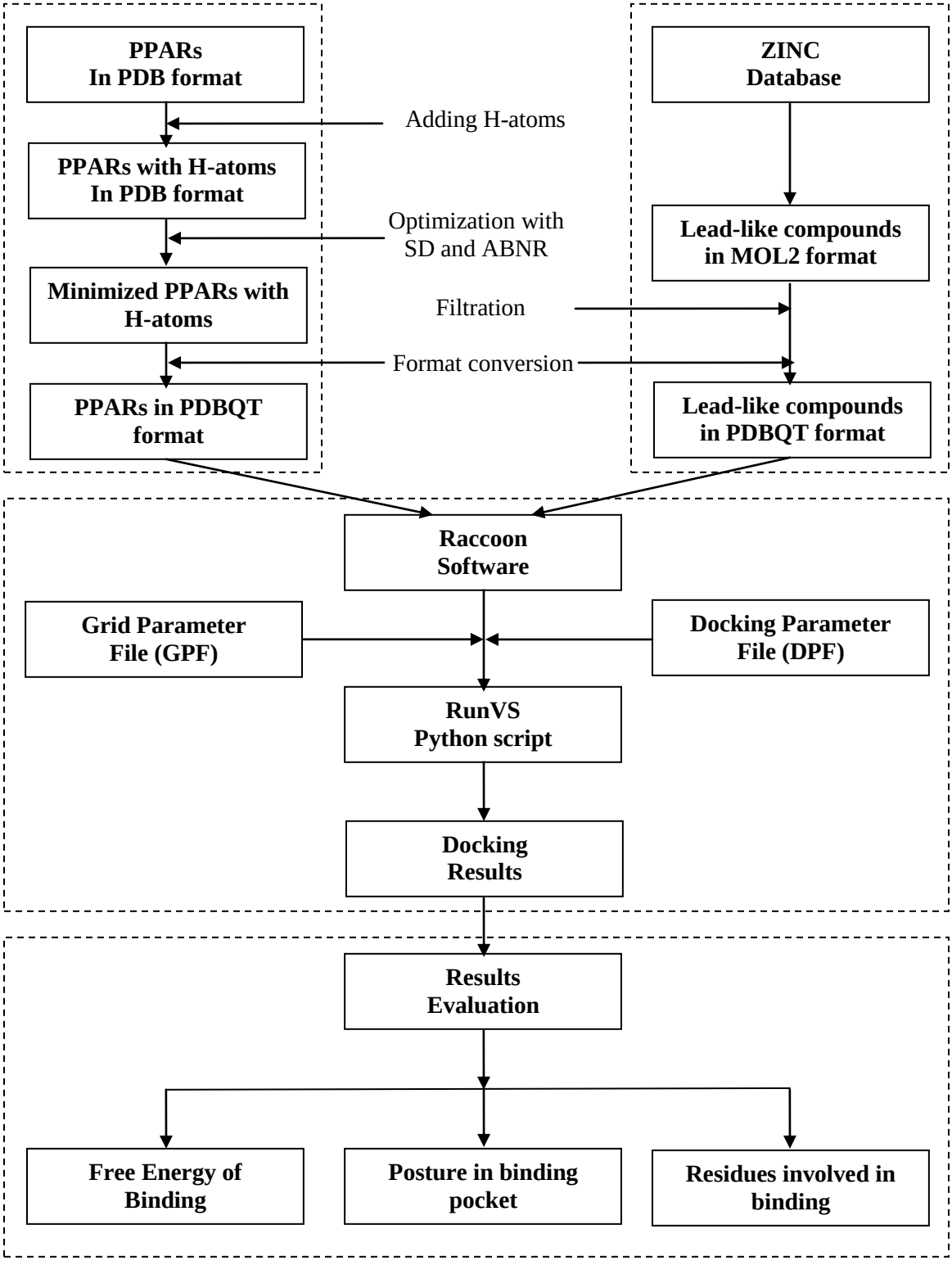


Figure 3.3 A flowchart of the key steps of the molecular docking protocol

3.3.1 Ligand libraries. The compounds for screening were obtained from the ZINC database, a free database of commercially available compounds. ZINC is provided by the Shoichet Laboratory in the Department of Pharmaceutical Chemistry at the University of California, San Francisco (UCSF) [79]. Subsets of more than four million lead-like compounds collected in libraries were downloaded in MOL2 format. Those compounds were filtered based on a primary automated virtual screening so that only the compounds, the most likely to dock with PPARs, were chosen for the final docking with AutoDock. The filtering was performed with iGEMDOCK, an award winning suite of automated screening tools developed by the Department of Biological Science and Technology & Institute of Bioinformatics National Chiao Tung University [81, 82, 83, 84]. iGEMDOCK uses a simple empirical fitness scoring function to predict the compounds that are likely to dock to the target proteins:

Fitness = van der Waals energy + H-bond energy + Electrostatic energy

3.3.2 PPARs preparation. A number of modifications were performed to the LBD PDB files. These modifications included the addition of the missing hydrogen atoms, removing water molecules, and removing extra chains and alternate locations. A python script was used to remove the unneeded records of the PDB file while the addition of the missing hydrogen atoms was performed with CHARMM (Chemistry at HARvard Molecular Mechanics), a widely-used program for macromolecular simulations (Appendix A). Before docking, the PPAR molecules were minimized with CHARMM to identify a set of coordinates representing a stable molecular conformation and minimal potential energy. The algorithm used to search for the global minimum energy was the

Steepest Descent (SD) and Adopted Basis Newton-Raphson (ABNR) methods (see Appendix A). After modifications and minimization, OpenBabel was used to convert the PDB files to PDBQT, a file format used by AutoDock.

3.3.3 Ligand preparation. OpenBabel was also used to convert the filtered lead-like compounds from MOL2 file to PDBQT file format. The hydrogen atoms were added by AutoDockTools to the compounds which were missing the hydrogen atoms and the partial charge of each compound was calculated as well and added to the compound.

3.3.4 Preparation of grid parameter files. The grid parameter file contained the information about the parameters and methods for the grid computing. The parameters were about grid-based potential energies files known as grid maps. A grid map was calculated for each atom type existing in the compound to be docked. A grid map consisted of a three dimensional lattice of regularly spaced points, surrounding the entire LBD (Figure 3.4). Each point within the grid map was the sum of the pair-wise potential interaction energy of a probe atom of a particular type with each of the atoms in the LBD.

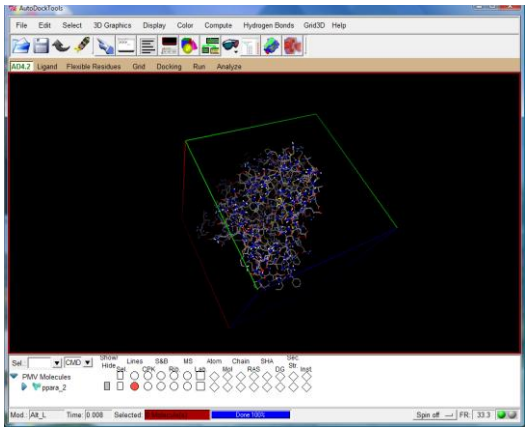


Figure 3.4. AutoDockTools

3.3.5 Docking of the compounds. Raccoon, developed by AutoDock, was used to automate docking. The software needed the compound files, receptor files, grid parameter template file (GPF), and docking parameter script file (DPF) as inputs. All required input files created in the previous steps except for the docking script file, which was generated by Raccoon. The docking script file contained the parameters needed to compute the pair-wise energy and the search method and search parameters needed to search for the global optimum. Genetic algorithm (GA) was used as a search method.

When the docking scrip file was run, the compounds were tested one by one against the three PPAR isotypes. The output of the docking process was saved in the docking log files (DLG), which contains all docking results such as energies, hydrogen bonds, electrostatic and van der Waals interactions. Each compound tested for docking had a separate docking log file. The docking files were analyzed one by one to evaluate the docking results of the compounds.

3.3.6 Evaluation of the docking results. The docking log files of the identified pan-PPAR ligands were evaluated with AutoDockTools. The DLG file for each ligand was opened with AutoDock and analyzed to identify the best conformations in terms of free energy of binding and docking position (Figure 3.5 and 3.6). The chosen conformations of each ligand-PPAR complex were then built-up and saved as PDB files, which were visualized with PyMol.

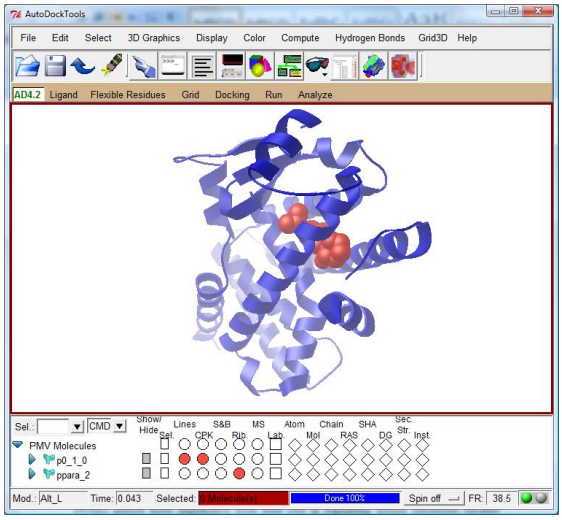


Figure 3.5. Visualizing a ligand docked to a PPAR with AutoDockTools

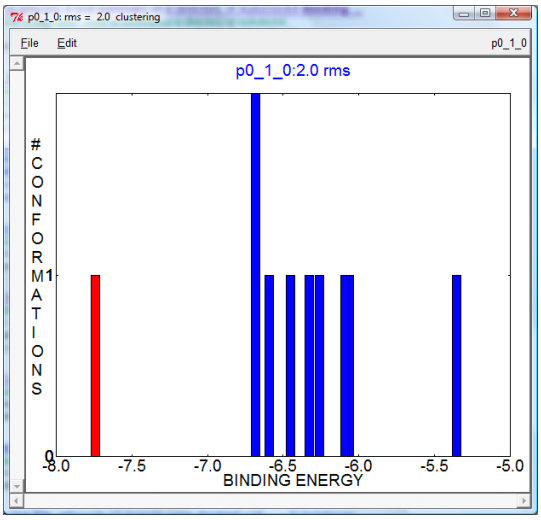


Figure 3.6. Binding energy histogram

3.3.7 Docking control PPAR ligands. The technique used in virtual screening was validated with four known PPAR ligands: unique alpha PPARα ligand, PPARδ ligand, PPARγ ligand, and pan PPAR ligand. These control ligands are known for their binding affinity and furthermore there are high resolution resolved 3D PPAR structures in complex with these ligands. The control ligands were obtained as SDF file format from

Protein Data Bank and converted to AutoDock file format. Then they were prepared, docked, and the results were analyzed with the same technique used for the compounds under investigation. Finally the results of the docked control ligands were compared to the already known results, namely the resolved 3D structures, locations of the binding site, and residues involved in binding. Table 3.1 shows the PDB files that were used in the comparison.

Table 3.1

The PPAR structures resolved by X-ray diffraction in complex with the control ligands

PDB ID	PPAR	Ligand
2P54	PPARα	GW735
2ZNP	PPARδ	TIPP-204
3K8S	PPARγ	T2384
3ET1, 3ET2, and 3ET3	PPARα, PPARδ, and PPARγ	Indeglitazar

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Selection of Compounds for Ligand Screening

The ZINC database was selected as the source for the test compounds used in this study because it is the largest database of commercially-available lead-like compounds in ready-to-dock 3D format. Only lead-like compounds were used for screening because of their potential to be used as drugs. These compounds are capable of passing into the blood via the digestive track when they are used orally. They meet the Lipinski's rule of five for orally-administered drugs, the main criteria to be selected as a PPAR ligand candidate in this study.

4.2 Characterization of PPAR Cavities

CASTp was used to identify the cavities and voids on each PPAR's LBD. The results showed that PPARs have pockets and voids of different shapes and sizes. Some of these cavities have one or more mouth openings while others do not have any. The finding confirmed that the binding site is located in the largest cavity. We found that PPARγ has the cavity with the largest molecular surface area (1496.05 A^2) followed by PPARα (1307.46 A^2), and PPARδ (1274.94 A^2) (Table 4.1). The size of the cavity may determine the size of the ligand that binds to it. The isotype with more spacious cavity might be able to bind to diverse sets of ligands varying in size and chemical nature. The relatively small size of PPARδ may explain why there are less known common ligands between this isotype and the others. On the other hand, the relative close cavity area of

PPARγ and PPARα indicates that dual ligands for these two isotypes are more likely to be found.

Table 4.1

The major pockets in the PPAR isotypes

Isotype	Pocket Area (A^2)		# of mouths
	SA	MS	
PPARα	648.009	1307.46	3
PPARβ/δ	755.686	1274.94	3
PPARγ	719.46	1496.05	2

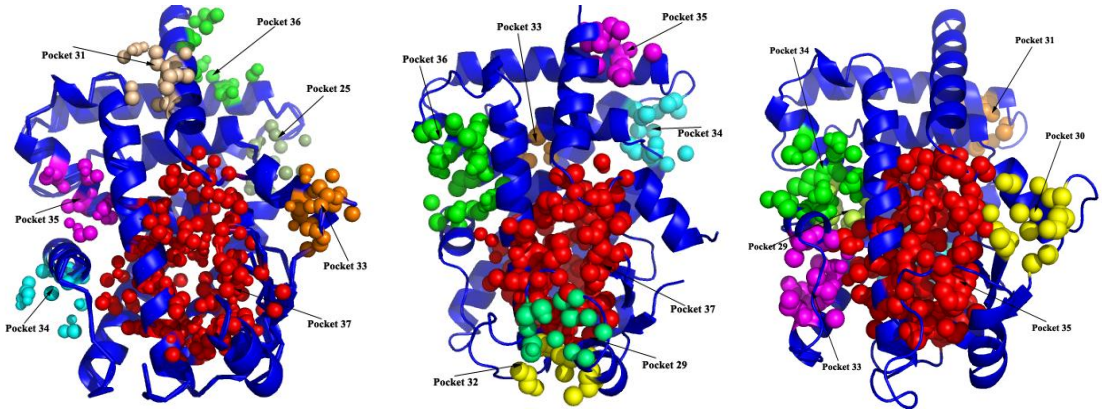
In addition to the major cavities, CASTp calculation predicted several smaller cavities and concavities in the ligand binding domains of the PPARS. The cavities of each isotype along with their areas, volumes, and number of mouths, are listed in Appendix B. PPARα and PPARδ have 37 cavities numbered from 1 to 37, while PPARγ has 35. The major cavities found in PPARα, PPARδ, and PPARγ were 37, 37, and 35 respectively. The total areas of the largest cavities in each isotype (Table 4.2) indicate the space of their potential binding sites. Our finding supports previous studies that showed that the PPARγ has a more spacious ligand binding site than the other two PPAR isotypes. The spacious binding pocket suggests that the PPARγ binding site is more promiscuous and consequently more likely to accept diverse ligands in terms of size and chemical composition.

Table 4.2

The total area and volume of the pockets of each PPAR isotype

Isotype	Pocket Area (A^2)	
	SA	MS
PPARα	1013.362	2955.85
PPARδ	1182.285	3139.59
PPARγ	1294.879	3547.23

In visualizing the top cavities with PyMOL, we observed that the largest cavity on each isotype was located at the same binding site positions, reported by others in previous studies (Figure 4.1). Based on this evidence we used the largest cavity to test our compounds for docking to the PPAR isotypes.



(A) PPARα (1I7G) (B) PPARδ (2XYW) (C) PPARγ (2ZK0)

Figure 4.1. The top seven cavities on PPAR LBDs

Given the importance of the largest cavity in each PPAR isotype for ligand binding, we identified the residues that form these cavities. Figures 4.2- 4.4 highlight the residues that form the binding cavities in the three isotypes, for pocket 37 on PPARα and PPARδ and for pocket 35 on PPARγ. The figures also show a similar sequence pattern,

shaded in green, across the three isotypes. Further investigation reveals that the amino acids forming the cavities are mostly hydrophobic. In addition, the first and last amino acids in the cavities of the three isotype are tyrosine residues. This co-incidence might suggest a particular importance of tyrosine for ligand binding.

Figure 4.2. The residues that form the largest cavity (pocket 37) on PPARα LBD

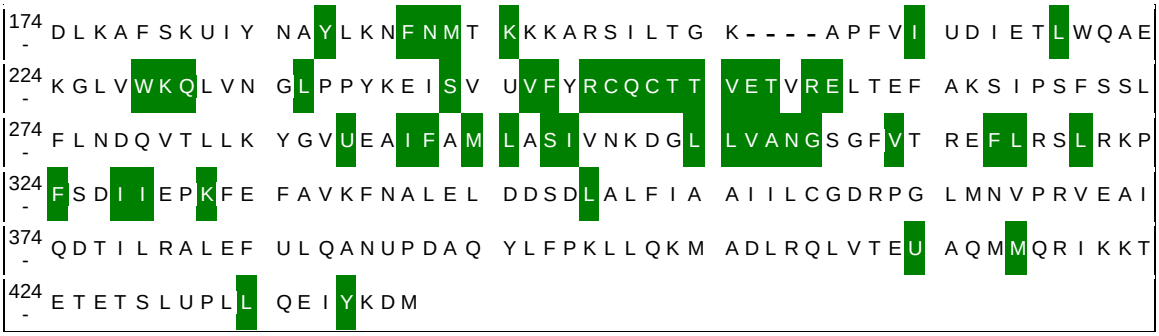


Figure 4.3. The residues that form the largest cavity (pocket 37) on PPARδ LBD

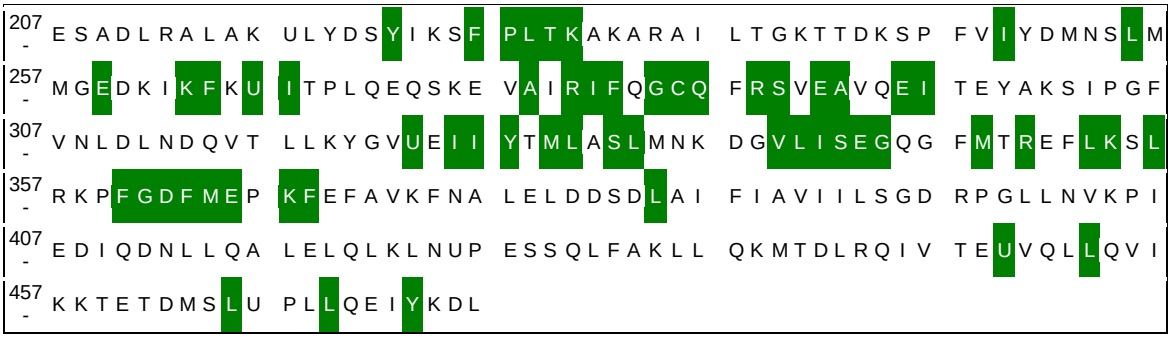


Figure 4.4. The residues that form the largest cavity (pocket 35) on PPARγ LBD

These findings are consistent with the need for hydrophobicity of some residues, which is essential to push out the solvent and furnish a suitable environment for ligand binding. Since ligands are hydrophobic or amphipathic, hydrophobic forces play a key role in forcing the ligands inside the ligand binding cavity. Variations in hydrophobic force contribution may account for the variation in the affinity for certain ligands across the different isotypes.

4.3 Identification of Pan-Ligands by AutoDock

Twenty seven out of approximately 4 million lead-like compounds met the criteria for consideration as pan-PPAR ligands. They were selected based on their position in the binding cavity and the free energy of binding after docking. The identified ligands were found to position themselves very well in the binding cavity with minimal free energy of binding. A cutoff of -6.0 kcal/mol of free energy of binding was set as threshold. It was shown that all ligands tended to bind to the binding cavity and position themselves behind Helix3. The number of hydrogen bonds formed between the ligands and the receptor residues never exceeded 5, complying with the Lipinski's rule of five for orally administered drugs. These results suggest that these ligands are strong candidate as potential drugs.

The Figures 4.5 – 4.7 show one of the identified pan-PPAR ligands binding to each PPAR isotype. The first image from left shows a wire structure model that depicts the hydrogen bond between the ligand (ID # p0.1-0) and the residue SER-280 of PPARα LBD. The image in the middle shows a surface protein structure model for the binding

pocket and the ligand inside. The image on the right is a ribbon model that shows the position of the ligand, which is close to Helix-3.

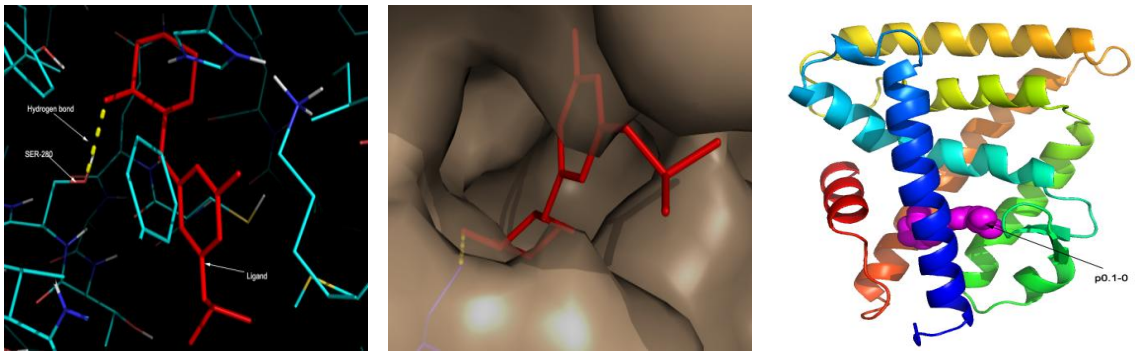


Figure 4.5. Ligand p0.1-0 docked to the PPARα LBD

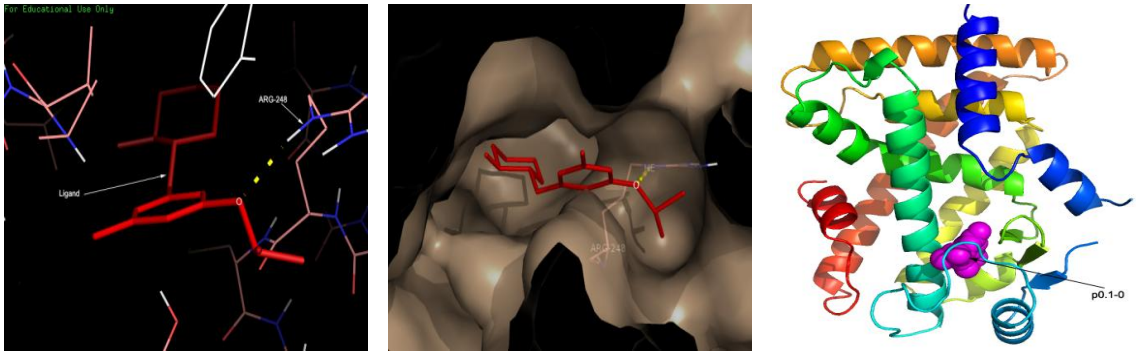


Figure 4.6. Ligand p0.1-0 docked to PPARδ LBD

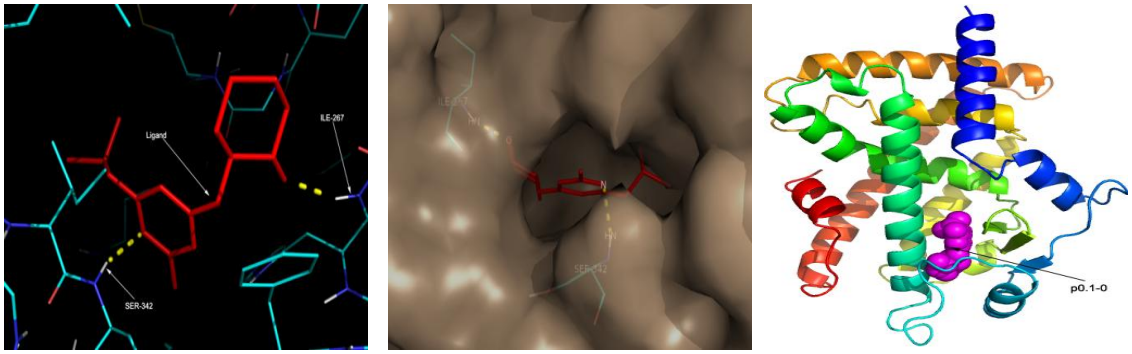


Figure 4.7. Ligand p0.1-0 docked to PPARγ LBD

The complete list of the identified pan-PPAR ligands with their chemical structures and names can be found in Appendix C, while their docking information such as free energy of binding and the hydrogen bonds can be found in Appendix D.

4.4 Identification of Important Residues for Hydrogen Bonding

After docking, the PPAR residues that formed hydrogen bonds, and most likely to participate in receptor/ligand contacts were determined, in addition to the number of H-bonds an isotype makes with each pan-ligand. The identification of the important residues complemented the identification of the pan-ligands. Future studies of the pan-ligands can use these residues to predict where contacts for binding may occur. Determining the number of H-bonds between the receptor and the test compounds was important for confirming that the pan-ligands passed the Lipinski's rule of five.

We found that each ligand forms a number of non-covalent H-bonds ranging between 2 to 3 with residues in the binding sites of each isotype. However, the residues involved in binding differed from one isotype to another for the same ligand. By evaluating the probability distribution of the residues involved in hydrogen bonding with the ligands across the identified ligand conformations, we noticed that some residues had a high probability of involvement in hydrogen bonding. Tables 4.3- 4.5 show the residues that were engaged in the ligand-protein contact in the three isotypes along with their frequencies and relative frequencies. These residues were weighted by their frequencies which indicate how many times a residue was found to be involved in a hydrogen bond. The relative frequency, the frequency of a residue divided by the total frequencies; indicate the probability distribution for each residue. These probability

distributions quantify the likelihood that these residues might be involved in hydrogen bonding with any other ligand that binds to these LBDs. These residues can be qualified as Important Residues (IR) for their role in hydrogen bonding. In PPARα, TYR-334, ASN-219, ALA-333, THR-279, MET-220, and GLU-286 were found to be the most frequent hydrogen bond formers (Table 4.3). Figure 4.8 shows the probability distributions of the most hydrogen bond forming residues in PPARα..

Table 4.3

The important residues of PPARα LBD and their probability distributions

Residue	Frequency	Relative Frequency
TYR334	47	0.30
ASN219	25	0.16
ALA333	22	0.14
THR279	19	0.12
MET220	13	0.08
GLU286	6	0.04
GLU282	5	0.03
SER280	3	0.02
SER323	3	0.02
CYS275	2	0.01
CYS278	2	0.01
HSD440	2	0.01
LEU331	2	0.01
GLY335	1	0.01
LYS364	1	0.01
THR283	1	0.01
TYR33	1	0.01
Total	155	1.0

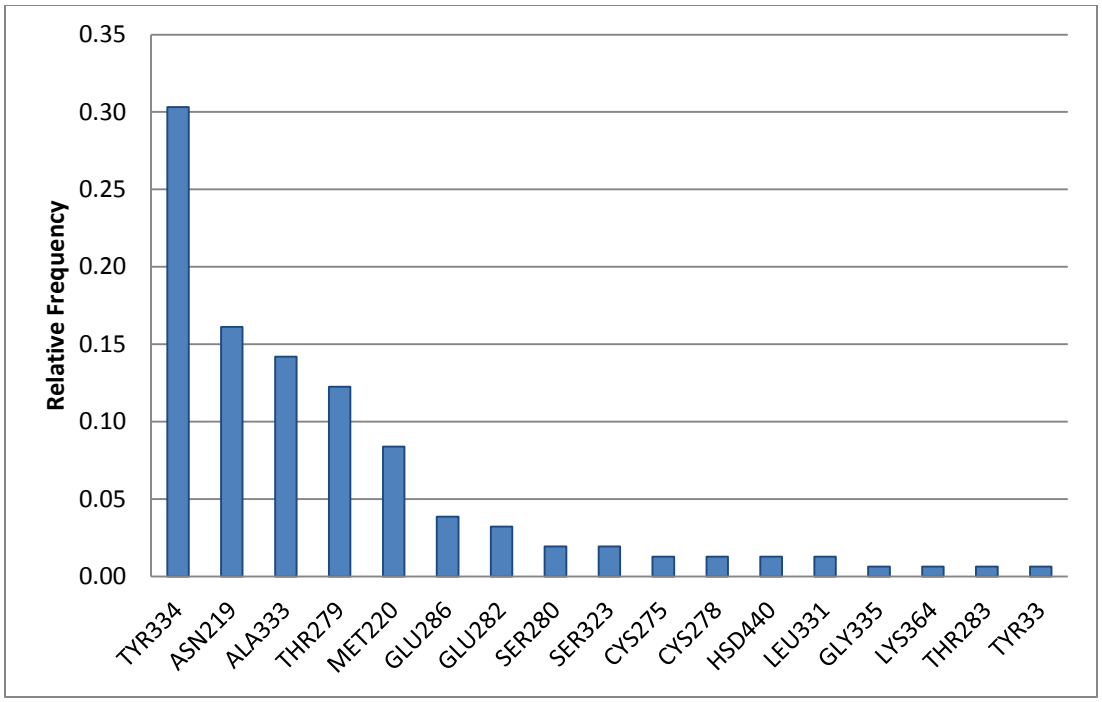


Figure 4.8. The important residues of PPARα LBD and their probabilities

In PPARδ, the most frequent hydrogen bond forming residues are ASN-307, ALA-306, THR252, MET192, THR-256, LYS-229, and GLN-230 (Table 4.4). Figure 4.9 shows the probability distributions of the most frequent hydrogen bond forming residues in PPARδ.

Table 4.4

The important residues of PPARδ LBD and their probability distributions

Residue	Frequency	Relative Frequency
ASN307	32	0.26
ALA306	19	0.15
THR252	16	0.13
MET192	13	0.10
THR256	10	0.08
LYS229	7	0.06
GLN230	5	0.04

Table 4.4 (Cont).

CYS249	4	0.03
HSD287	3	0.02
LEU304	3	0.02
ARG248	2	0.02
MET293	2	0.02
THR253	2	0.02
TYR437	2	0.02
GLY308	1	0.01
HSD413	1	0.01
ILE290	1	0.01
SER296	1	0.01
Total1	124	1.0

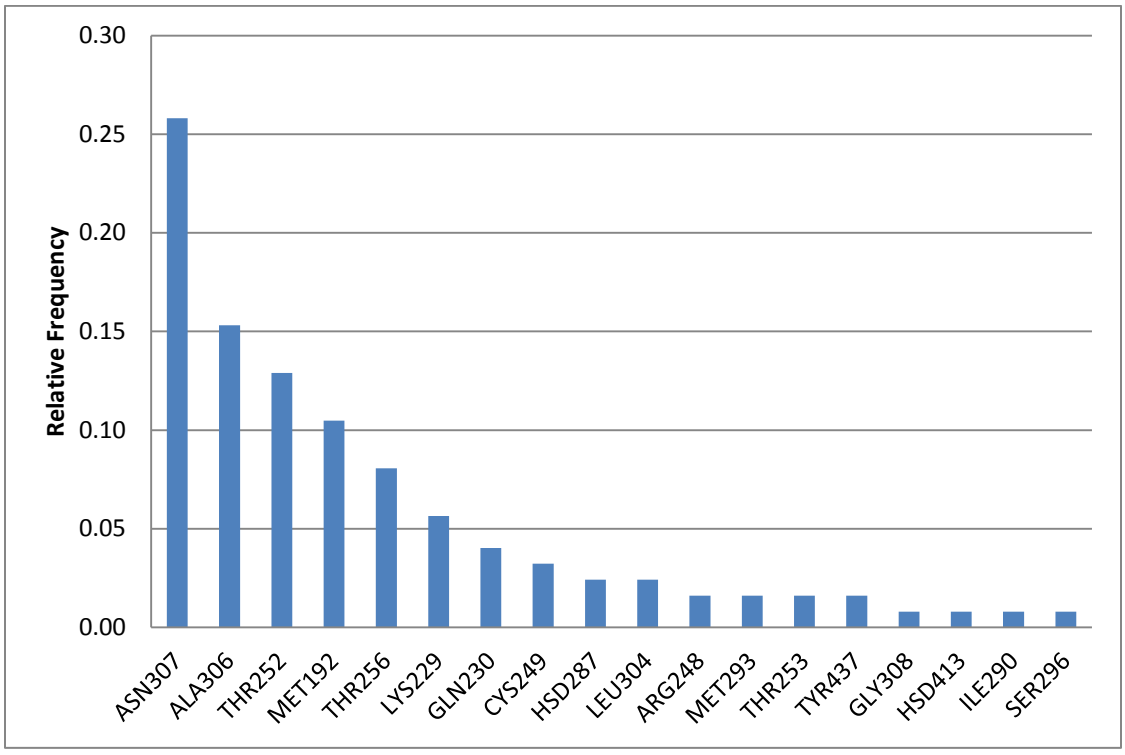


Figure 4.9. The important residues of PPARδ LBD and their probabilities

The most frequent hydrogen bond forming residues in PPARγ were SER-342, ILE-267, GLU-343, HIS-266, ARG-288, GLU-291, and LYS-244. Table 4.5 lists the frequencies and relative frequencies (probability distributions) for the hydrogen bond forming residues in PPARγ while Figure 4.10 shows a graphical view of the probability distributions.

Table 4.5

The important residues of PPARγ LBD and their probability distributions

Residue	Frequency	Relative Frequency
SER342	46	0.36
ILE267	24	0.19
GLU343	12	0.09
HSD266	11	0.09
ARG288	10	0.08
GLU291	8	0.06
LYS244	5	0.04
CYS285	2	0.02
GLN345	2	0.02
GLU369	2	0.02
LYS474	2	0.02
HSD449	1	0.01
SER245	1	0.01
SER289	1	0.01
Total	127	1.00

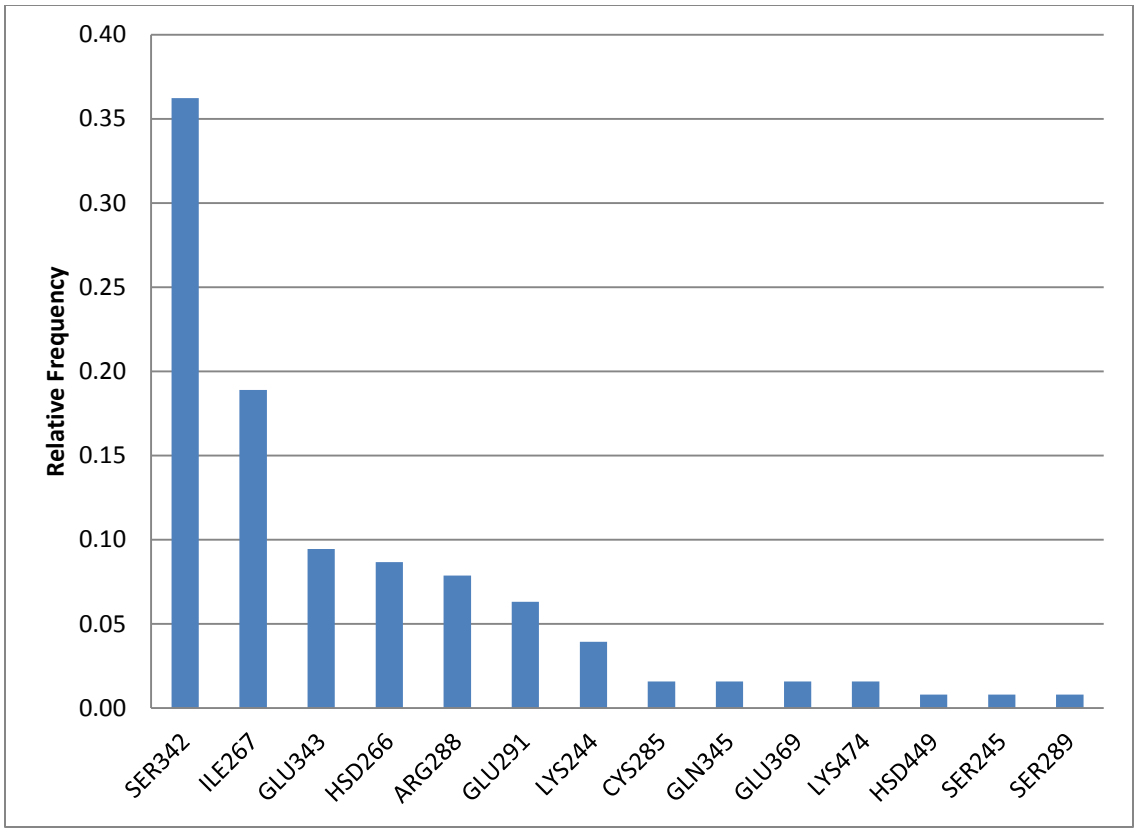


Figure 4.10. The important residues of PPARγ LBD and their probabilities

The contacts for the LBDs of the three isotypes with the ligands are listed in Appendix D. The LBD residues contributing to the contacts together with the atoms that play a role as electron donors or acceptors in the residue are shown. The number of hydrogen bonds agree with Lipinski's Rule of Five, which states that the number of hydrogen bond should not be more than five. A cutoff of 2.8 Å was set as a threshold to exclude any interaction above that distance.

By studying the locations of these important residues, we noticed that these residues belong to the residual groups that form the main binding cavities in the three PPAR isotypes respectively (Figures 4.11). Such findings strongly support the results of

computational prediction of the binding cavities as shown in the beginning of this chapter. The involvement of these residues in hydrogen bonding suggests that the mutation of such residues may affect the affinity of the LBDs of the receptors to ligands.

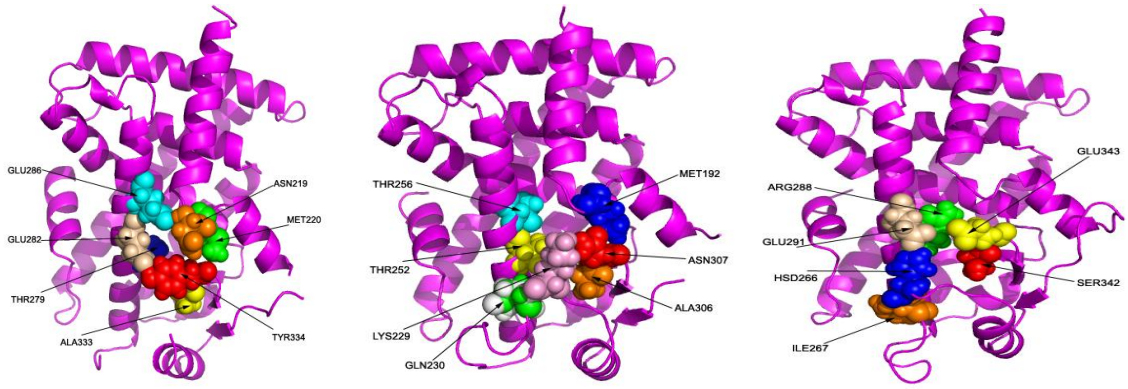


Figure 4.11. The important residues in (a) PPARα (b) PPARδ (c) PPARγ

4.5 Chemical Structures of the Identified Pan-ligands

Once the pan-ligands were identified and selected, the chemical structure and International Union of Pure and Applied Chemistry (IUPAC) names for each compound was resolved with Accelrys Draw. SMILES strings for each ligand was obtained by OpenBabel. The chemical structures and SMILES strings for the 27 pan-ligands are found in Appendix C.

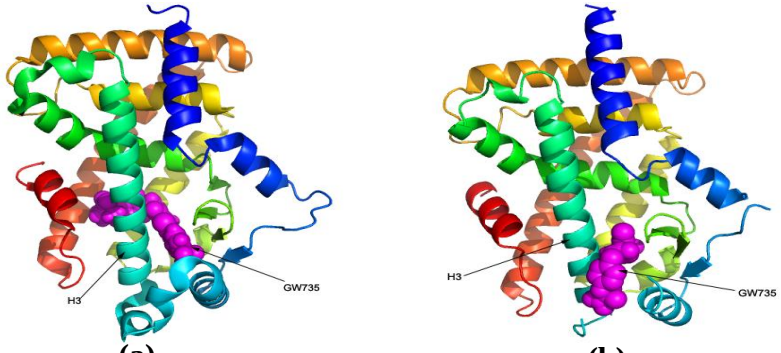
4.6 Validation of the Computational Techniques Using Control Ligands

The known PPAR agonists listed in Table 4.2 were used as controls to validate the docking technique used in this study. These controls were prepared and docked in the same way as the lead test compounds were done. The control results were quite convincing in showing that the control ligands docked and positioned themselves in the pockets of their respective PPAR LBD and did not dock or docked very poorly to

unexpected PPAR targets, complying with the specificity rules imposed on the LBDs of the three isotypes.

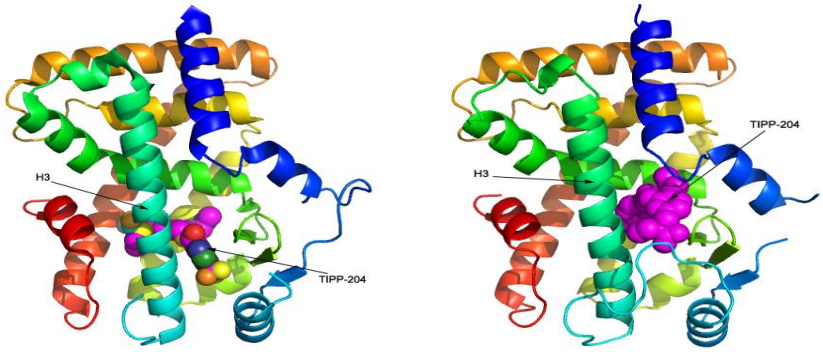
The control ligands did not dock exactly as in the structures resolved with X-ray diffraction but a small margin of error was expected. Figures 4.12 and 4.13 compare the actual X-ray structures with the structures obtained with the molecular docking. Tables 4.7 and 4.8 indicate the free energy of binding and hydrogen bonds for each conformation.

We noticed that the ligands positioned themselves well on the binding site close enough to Helix3 so that some residues from this helix may be involved in the contacts with the ligands. In general, the slight deviations that were observed do not necessarily invalidate the techniques used to identify the pan-ligands. The results show enough degree of certainty to support the computational techniques used in this study.



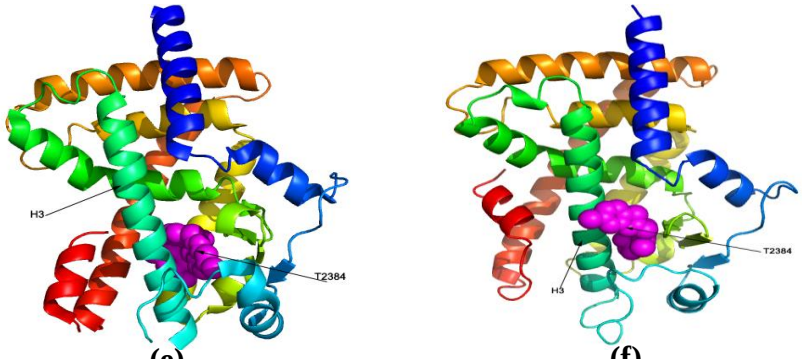
(a) **(b)**

PPARα LBD in complex with GW735 (a) resolved by X-ray diffraction (2P54) (b) obtained with molecular docking



(c) **(d)**

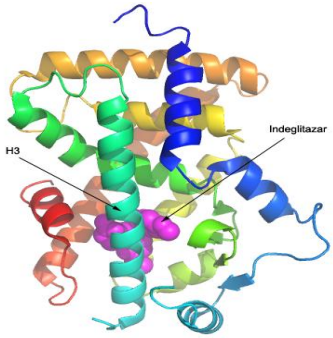
PPARδ LBD in complex with TIPP-204 (c) resolved by X-ray diffraction (2ZNP) (d) obtained with molecular docking



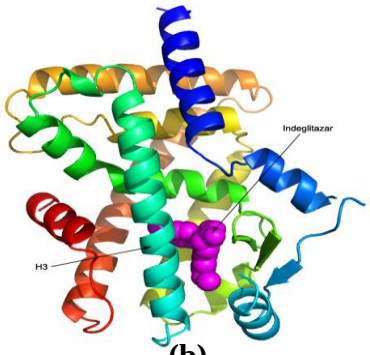
(e) **(f)**

PPARγ LBD in complex with T2384 (e) resolved by X-ray diffraction (3K8S) (f) obtained with molecular docking

Figure 4.12. PPAR LBDs in complex with the control ligands

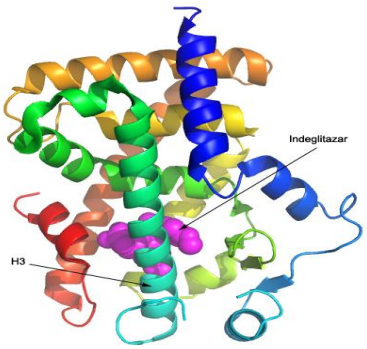


(a)



(b)

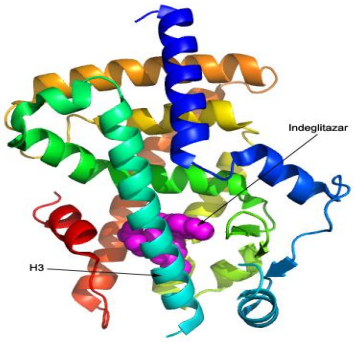
PPARα LBD in complex with Indeglitazar (a) resolved by X-ray diffraction (3ET1) (b) obtained with molecular docking



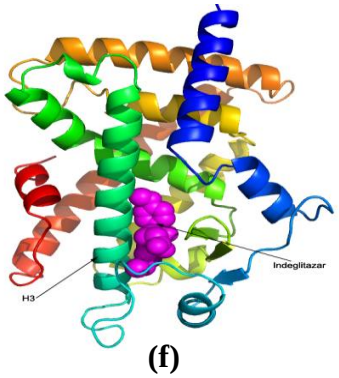
(c)

(d)

PPARδ LBD in complex with Indeglitazar (c) resolved by X-ray diffraction (3ET2) (d) obtained with molecular docking



(e)



(f)

PPARγ LBD in complex with Indeglitazar (e) resolved by X-ray diffraction (3ET3) (f) obtained with molecular docking

Figure 4.13. PPAR LBDs in complex with the control pan-ligand

Table 4.6

The residues of the three PPARs that form H-bond with the control ligands

PPAR	Control Ligand	Binding Energy	H-bonds
PPARα	GW735	-7.57 -6.43	GLN264:HN1 GLN264:HN3
PPARδ	TIPP-204	-6.8	LYS229:HN
PPARγ	T2384	-7.37 -6.44	HSD266:HD1 THR297:HG1 LYS319:HZ3

Table 4.7

The residues of the three PPARs that form H-bond with the control pan-ligand

PPAR	Control Ligand	Binding Energy	H-bonds
PPARα	Indeglitazar	-7.88 -7.43 -7.31 -7.19 -7.17 -6.8	GLN445:HE21 LYS448:HZ1 LYS204:HZ2 LYS208:HZ2 HSD411:HD1 LYS204:HZ2 LYS208:HZ2 HSD411:HD1 TYR334:HN GLN264:HN1 CYS275:HG1
PPARδ	Indeglitazar	-7.75 -7.65 -7.37 -7.37	ASP174:HN3 LEU175:HN LYS176:HZ2 SER179:HG1 LYS229:HN THR252:HG1 HSD389:HD1 LYS398:HZ2 LYS402:HZ1 MET192:HN ASN307:HN
PPARγ	Indeglitazar	-7.33 -6.93 -6.44	HSD266:HD1 ARG288:HE LYS230:HZ1 THR238:HG1 LYS240:HZ2

CHAPTER 5

CONCLUSION

In this study, twenty seven out of the four million lead-like compounds, stored in the ZINC database, were identified as PPAR pan-ligands based on results obtained by computational analysis techniques. As pan-ligands the compounds comply with Lipinski's rule of five and dock very well in the binding cavity of the three isotypes with a minimal free energy of binding. All the twenty seven successfully met these criteria when evaluated on the basis of their docking position in the binding sites, free energy of binding, and the number of hydrogen bonds they form. In addition to meeting the criteria, these pan-ligands were found to be chemically similar to known PPAR's ligands that have been used as drugs. They have the hydrophobic carboxylic and hydrophilic properties required for them to pass through the gates of the binding cavities of the three isotypes and form hydrogen bonds with the polar residues within the binding sites. The entrance of the ligands into the cavities to be in contact with the binding site is the most critical part of the PPAR activation process that mediates the allosteric change in the AF2-helix that leads to the interaction of the receptors with the DNA to regulate the target genes. Thus, the identified pan-ligands are strong potential pan-PPAR agonists.

To validate the computational techniques used in this study, known PPAR ligands were tested against the three isotypes as positive and negative controls. The control ligands behaved similarly and met the study criteria, asserting that the computational techniques were capable of discriminating for the pan-PPAR ligands.

We found that the computational virtual screening with molecular docking was efficient, fast, straightforward and economical, but confirmation of the ligands as pan-agonists will require laboratory investigation. Each ligand represents the potential for discovering a potent pan-PPAR agonist in the laboratory. Drug optimization may be needed to ensure that these compounds can activate the PPARs simultaneously to control multiple genes involved in the metabolic syndrome with minimal side effects.

REFERENCES

[1] Beatrice Artice Desvergne and Walter Wahli. (1999). “Peroxisome Proliferator Activated Receptors: Nuclear Control of Metabolism”. Institute of Animal Biology, University of Lausanne, CH-1015 Lausanne, Switzerland.

[2] Salman Azhar, (2010). “Peroxisome proliferator-activated receptors, metabolic syndrome and cardiovascular disease”. Future Cardiol.; 6(5): 657–691. doi:10.2217/fca.10.86.

[3] V. A. Javiya, J. A. Patel. (2006). “The role peroxisome proliferator-activated receptors in human diseases”. Indian J Pharmacol. Vol 38, issue 4. 243-253.

[4] Berger J, Moller DE (2002). "The mechanisms of action of PPARs". Annu. Rev. Med. 53: 409-35. doi:10.1146/annurev.med.53.082901.104018. PMID 11818483.

[5] Grigorios Krey, Olivier Braissant, Fabienne L’Horset, Eric Kalkhoven, Mai Perroud, Malcolm G. Parker, and Walter Wahli. (1997). "Fatty Acids, Eicosanoids, and Hypolipidemic Agents Identified as Ligands of Peroxisome Proliferator-Activated Receptors by Coactivator-Dependent Receptor Ligand Assay". Molecular Endocrinology Laboratory (F.L., E.K., M.G.P.) Imperial Cancer Research Fund 44 Lincoln’s Inn Fields London WC2A 3PX, U.K.

[6] Baichun Yang, Kathleen K Brown, Lihong Chen, Kevin M Carrick, Lisa G Clifton, Judi A McNulty, Deborah A Winegar, Jay C Strum, Stephen A Stimpson1 and Gregory L Pahel. (2004). "Serum adiponectin as a biomarker for in vivo PPARgamma activation and PPARgamma agonist-induced efficacy on insulin sensitization/lipid lowering in rats". BMC Pharmacology 2004, 4:23 doi:10.1186/1471-2210-4-23.

[7] Marc-Andre Cornier, Dana Dabelea, Teri L. Hernandez, Rachel C. Lindstrom, Amy J. Steig, Nicole R. Stob, Rachael E. Van Pelt, Hong Wang, and Robert H. Ecke. (2008). "The Metabolic Syndrome". Endocrine Reviews 29(7):777–822. doi: 10.1210/er.2008-0024.

[8] Taro E. Akiyama, Peter T. Meinke, and Joel P. Berger. (2005). "PPAR Ligands: Potential Therapies for Metabolic Syndrome". Current Science Inc. ISSN 1534–4827.

[9] Salman Azhar. (2010). "Peroxisome proliferator-activated receptors, metabolic syndrome and cardiovascular disease". Future Cardiol. 2010 September ; 6(5): 657–691. doi:10.2217/fca.10.86.

[10] Dean R. Artisa, Jack J. Lina, Chao Zhanga, Weiru Wanga, Upasana Mehraa, Mylene Perreaultb, David Erbeb, Heike I. Krupkaa, Bruce P. Englanda, James

Arnolda, Alexander N. Plotnikova, Adhirai Marimuthua, Hoa Nguyena, Sarah Willb, Maxime Signaevskyc, John Kralc, John Cantwella, Calvin Settachatgulla, Douglas S. Yana, Daniel Fonga, Angela Oha, Shenghua Shia, Patrick Womacka, Benjamin Powella, Gaston Habetsa, Brian L. Westa, Kam Y. J. Zhanga, Michael V. Milburna, George P. Vlasukb, K. Peter Hirtha, Keith Nolopa, Gideon Bollaga, Prabha N. Ibrahima, and James F. Tobinb. (2008). "Scaffold-based discovery of indeglitazar, a PPAR pan-active anti-diabetic agent". The National Academy of Sciences of the USA. vol. 106, No. 1. 262–267.

[11] Takuji Oyama Kenji Toyota Tsuyoshi Waku Yuko Hirakawa Naoko Nagasawa Jun-ichi Kasuga Yuichi Hashimoto Hiroyuki Miyachib and Kosuke Morikawaa. (2009). "Adaptability and selectivity of human peroxisome proliferator-activated receptor (PPAR) pan agonists revealed from crystal structures". Biological Crystallography ISSN 0907-4449.

[12] M. Khazaei E. Salehi and B. Rashidi. (2012). "Pan-PPAR Agonist, Bezafibrate, Restores Angiogenesis in Hindlimb Ischemia in Normal and Diabetic Rats". Hindawi Publishing Corporation International Journal of Peptides Volume 2012, Article ID 637212, 5 pages doi:10.1155/2012/637212.

[13] Alberts B, Johnson A, Lewis J, et al. Molecular Biology of the Cell. 4th edition. New York: Garland Science; 2002. Protein Function. Available from: http://www.ncbi.nlm.nih.gov/books/NBK26911/

[14] Tudor I. Oprea, Andrew M. Davis, Simon J. Teague, and Paul D. Leeson. (2001). "Is There a Difference between Leads and Drugs? A Historical Perspective". AstraZeneca R&D Mo¨lndal, EST Lead Informatics, S 431 83 Mo¨lndal, Sweden, and AstraZeneca R&D Charnwood, Medicinal Chemistry, Loughborough, Leics., LE11 5RH, U.K.J. Chem. Inf. Comput. Sci. 2001, 41, 1308-1315.

[15] Tamas Varga, Zsolt Czimmerer, and Laszlo Nagy. (2011). “PPARs are a unique set of fatty acid regulated transcription factors controlling both lipid metabolism and inflammation”. Biochim Biophys Acta. 2011 August; 1812(8): 1007–1022. doi: 10.1016/j.bbadis.2011.02.014

[16] Liliane Michalik, Beatrice Desvergne, Christine Dreyer, Mathilde Gavillet, Ricardo N. Laurini and Walter Wahli. (2002). “PPAR expression and function during vertebrate development”. Int. J. Dev. Biol. 46: 105-114.

[17] Frances M. Sladek. (2010). “What are nuclear receptor ligands?”. Molecular and Cellular Endocrinology. 2010.

[18] Markus Rieck, Wolfgang Meissner, Simone Ries, Sabine Muller-Brusselbach, and Rolf Muller. (2008). “Ligand-Mediated Regulation of Peroxisome Proliferator-Activated Receptor (PPAR) β/δ: A Comparative Analysis of PPAR-Selective

Agonists and All-trans Retinoic Acid". Mol Pharmacol 74:1269–1277, Vol. 74, No. 5.

[19] Ningwu Huang, Elmar vom Baur, Jean-Marie Garnier, Thierry Lerouge, Jean-Luc Vonesch, Yves Lutz, P.Chambon and Regine Losson. (1998). "Two distinct nuclear receptor interaction domains in NSD1, a novel SET protein that exhibits characteristics of both corepressors and coactivators". The EMBO Journal Vol.17 No.12 pp.3398–3412.

[20] Pierre Germain, Bart Staels, Catherine Dacquet, Michael Spedding, and Vincent Laudet. (2006). "Overview of Nomenclature of Nuclear Receptors". doi: 10.1124/pr.58.4.2 Pharmacological Reviews vol. 58 no. 4 685-704.

[21] Rolf Muller, Martin Komhoff, Jeffrey M. Peters, and Sabine Muller-Brusselbach. (2008). "A Role for PPARβ/δ in Tumor Stroma and Tumorigenesis". PPAR Research Volume 2008, Article ID 534294, 5 pages doi:10.1155/2008/534294.

[22] Lap Shu Alan Chan and Richard A.Wells. (2009). " Cross-Talk between PPARs and the Partners of RXR: A Molecular Perspective". PPAR Research Volume 2009, Article ID 925309, 9 pages doi:10.1155/2009/925309

[23] A J Vidal-Puig, R V Considine, M Jimenez-Liñan, A Werman, W J Pories, J F Caro, J S Flier. (1997). "Peroxisome Proliferator–activated Receptor Gene Expression in Human Tissues". J Clin Invest.; 99(10):2416–2422 doi:10.1172/JCI119424.\

[24] Christopher D. Kane, Kimberly A. Stevens, James E. Fischer, Mehrdad Haghpassand, Lori J. Royer, Charles Aldinger, Katherine T. Landschulz1, Panayiotis Zagouras, Scott W. Bagley, William Hada, Robert Dullea, Cheryl M. Hayward and Omar L. Francone. (2008). "Molecular Characterization of Novel and Selective Peroxisome Proliferator-Activated Receptor α Agonists with Robust Hypolipidemic Activity in Vivo". doi: 10.1124/mol.108.051656.

[25] Yuji Naito, Tomohisa Takagi, and Toshikazu Yoshikawa.(2010). " Gastrointestinal Cytoprotection by PPARγ Ligands". Hindawi Publishing Corporation, PPAR Research, Volume 2010, Article ID 108632, 8 pages, doi:10.1155/2010/108632.

[26] Tatjana Albrektsen, Klaus Stensgaard Frederiksen, William E. Holmes, Esper Boel, Karen Taylor, and Jan Fleckner. (2002). "Novel Genes Regulated by the Insulin Sensitizer Rosiglitazone During Adipocyte Differentiation". doi: 10.2337/diabetes.51.4.1042 Diabetes vol. 51 no. 4 1042-1051.

[27] Claire Bastie, Dorte Holst, Danielle Gaillard, Chantal Jehl-Pietri and Paul A. Grimaldi. (1999). "Expression of Peroxisome Proliferator-activated Receptor PPARδ Promotes Induction of PPARγ and Adipocyte Differentiation in 3T3C2

Fibroblasts". doi: 10.1074/jbc.274.31.21920. The Journal of Biological Chemistry, 274, 21920-21925.

[28] Ana Aranda and Angel Pascual. (2001). "Nuclear Hormone Receptors and Gene Expression". Physiological Review Vol. 81, No. 3.

[29] Magrane M. and the UniProt consortium. UniProt Knowledgebase: a hub of integrated protein data. Database, 2011: bar009 (2011).

[30] Vincent Zoete, Aurelien Grosdidier, Olivier Michielin. (2007). "Peroxisome proliferator-activated receptor structures: Ligand specificity, molecular switch and interactions with regulators". Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids. Volume 1771, Issue 8, Pages 915–925.

[31] Zoete V, Grosdidier A, Michielin O. (2007). "Peroxisome proliferator-activated receptor structures: ligand specificity, molecular switch and interactions with regulators". Biochim Biophys Acta.;1771(8):915-25. Epub 2007 Jan 18. PMID:17317294 doi:10.1016/j.bbalip.2007.01.007.

[32] Philippe Cronet, Jens F.W. Petersen, Rutger Folmer, Niklas Blomberg, Kristina Sjöblom, Ulla Karlsson, Eva-Lotte Lindstedt, Krister Bamberg. (2001). "Structure of the PPARα and –γ Ligand Binding Domain in Complex with AZ 242; Ligand Selectivity and Agonist Activation in the PPAR Family". Structure Volume 9, Issue 8, Pages 699–706.

[33] Liliane Michalik, Johan Auwerx, Joel P. Berger, V. Krishna Chatterjee, Christopher K. Glass, Frank J. Gonzalez, Paul A. Grimaldi, Takashi Kadowaki, Mitchell A. Lazar, Stephen O'Rahilly, Colin N. A. Palmer, Jorge Plutzky, Janardan K. Reddy, Bruce M. Spiegelman, Bart Staels and Walter Wahli. (2006). " Peroxisome Proliferator-Activated Receptors". doi: 10.1124/pr.58.4.5 Pharmacological Reviews vol. 58 no. 4 726-741.

[34] Shu-Hsien Sheu, Taner Kaya, David J. Waxman, and Sandor Vajda. (2005). "Exploring the Binding Site Structure of the PPARγ Ligand-Binding Domain by Computational Solvent" Mapping. Biochemistry, 44 (4), pp 1193–1209 DOI: 10.1021/bi048032c.

[35] Ferdinand Molnar, Merja Matilainen, and Carsten Carlberg. (2005). "Structural Determinants of the Agonist-independent Association of Human Peroxisome Proliferator-activated Receptors with Coactivators". JBC Papers in Press, DOI 10.1074/jbc.M502463200.

[36] Ferdinand Molnar, Merja Matilainen, and Carsten Carlberg. (2005). "Structural Determinants of the Agonist-independent Association of Human Peroxisome Proliferator-activated Receptors with Coactivators". THE JOURNAL OF

BIOLOGICAL CHEMISTRY Vol. 280, No. 28, Issue of July 15, pp. 26543–26556.

[37] Chandra V, Huang P, Hamuro Y, Raghuram S, Wang Y, Burris TP, Rastinejad F. (2008). “Structure of the intact PPAR-gamma-RXR- nuclear receptor complex on DNA. Nature”. 456(7220):350-6. PMID:19043829 doi:10.1038/nature07413

[38] W Jurkowski, K Roomp, I Crespo, J G Schneider, and A del Sol. (2011). “PPARγ population shift produces disease-related changes in molecular networks associated with metabolic syndrome”. Cell Death Dis 2(8): e192. Published online 2011 August 11. doi: 10.1038/cddis.2011.74

[39] Jeffrey M. Peters . Connie Cheung . Frank J. Gonzalez. (2005). “Peroxisome proliferator-activated receptor-! and liver cancer: where do we stand?”. Springer-Verlag. J Mol Med (2005) 83: 774–785 DOI 10.1007/s00109-005-0678-9.

[40] J. Christopher Corton, Steven P. Anderson and Anja Stauber. (2000). “Central Role of Peroxisome Proliferator–Activated Receptors in the Actions of Peroxisome Proliferators”. Annual Review of Pharmacology and Toxicology Vol. 40: 491-518 DOI: 10.1146/annurev.pharmtox.40.1.491

[41] Christopher Weidner, Jens C. de Groot, Aman Prasad, Anja Freiwald, Claudia Quedenau, Magdalen Kliem, Annabell Witzke, Vitam Kodelja, Chung-Ting Han, Sasch Giegold, Matthias Baumann, Bert Klebl, Karsten Siems, Lutz Müller-Kuhrt, Annette Schürmann, Rita Schüler, Andreas F. H. Pfeiffer, Frank C. Schroeder, Konra Büssow, and Sascha Sauer. (2012). "Amorfrutins are potent antidiabetic dietary natural products". doi: 10.1073/pnas.1116971109 PNAS April 16, 2012.

[42] Milton Hamblin, Lin Chang, Yanbo Fan, Jifeng Zhang, and Y. Eugene Chen. (2009). “PPARs and the Cardiovascular System”. Antioxidants & Redox Signaling. Volume 11, Number 6, 200 Mary Ann Liebert, Inc. DOI: 10.1089=ars.2008.2280

[43] Eddy Karnieli and Michal Armoni. (2008). “Transcriptional regulation of the insulin-responsive glucose transporter GLUT4 gene: from physiology to pathology”. Am J Physiol Endocrinol Metab 295: E38–E45,. First published May 20, 2008; doi:10.1152/ajpendo.90306.2008.

[44] Huiyun Liang and Walter F. Ward. (2006). “PGC-1α: a key regulator of energy metabolism”. doi: 10.1152/advan.00052.2006 Adv Physiol Educ vol. 30 no. 4 145-151.

[45] Jordi Camps,AnabelGarcıa-Heredia,AnnaRull,CarlosAlonso-Villaverde, GerardAragones, Raul Beltran-Debon, Esther Rodrıguez-Gallego, and Jorge Joven. (2012). “PPARs in Regulation of Paraoxonases: Control of Oxidative Stress and

Inflammation Pathways". PPAR Research Volume 2012, Article ID 616371, 10 pages doi:10.1155/2012/616371.

[46] Berger J, Moller DE. (2002). "The mechanisms of action of PPARs". Annu Rev Med.;53:409-35. PMID:11818483 doi:10.1146/annurev.med.53.082901.104018.

[47] Alexander Tenenbaum, Michael Motro and Enrique Z Fisman. (2005). "Dual and pan-peroxisome proliferator-activated receptors (PPAR) co-agonism: the bezafibrate lessons". Cardiovascular Diabetology , 4:14 doi:10.1186/1475-2840-4-14.

[48] Joshua P. Gray. PPAR ligands. Retrieved from http://www.joshuapgray.com/.

[49] Fraydoon Rastinejad, Trixie Wagner, Qiang Zhao and Sepideh Khorasanizadeh. (2000). "Structure of the RXR–RAR DNA-binding complex on the retinoic acid response element DR1". The EMBO Journal (2000) 19, 1045 – 1054 doi:10.1093/emboj/19.5.1045.

[50] Christopher K. Glass1 and Michael G. Rosenfeld. (2000). "The coregulator exchange in transcriptional functions of nuclear receptors". doi: 10.1101/gad.14.2.121 Genes & Dev. 2000. 14: 121-141.

[51] Christine Yu, Kathleen Markan, Karla A. Temple, Dianne Deplewski, Matthew J. Brady, and Ronald N. Cohen. (2005). "The Nuclear Receptor Corepressors NCoR and SMRT Decrease Peroxisome Proliferator-activated Receptor γ Transcriptional Activity and Repress 3T3-L1 Adipogenesis". Published, JBC Papers in Press, February 3, 2005, DOI 10.1074/jbc.M409468200.

[52] Mary C. Thomas and Cheng-Ming Chiang. (2006). "The General Transcription Machinery and General Cofactors". Critical Reviews in Biochemistry and Molecular Biology, 41:105–178, 2006.

[53] Judit Osz, Yann Brélivet, Carole Peluso-Iltis, Vincent Cura, Sylvia Eiler, Marc Ruff, William Bourguet, Natacha Rochel, and Dino Moras. (2012). "Structural basis for a molecular allosteric control mechanism of cofactor binding to nuclear receptors". Published online before print February 21, 2012, doi: 10.1073/pnas.1118192109 PNAS March 6, 2012 vol. 109 no. 10 E588-E594.

[54] Rick B. Vega, Janice M. Huss, and Daniel P. Kelly. (2000). "The Coactivator PGC-1 Cooperates with Peroxisome Proliferator-Activated Receptor a in Transcriptional Control of Nuclear Genes Encoding Mitochondrial Fatty Acid Oxidation Enzymes". Molecular and Cellular Biology, 0270-7306/00/$04.0010 Mar. 2000, p. 1868–1876 Vol. 20, No. 5

[55] Christophe Blanquart, Olivier Barbier, Jean-Charles Fruchart, Bart Staels, and Corine Glineur. (2002). “Peroxisome Proliferator-activated Receptor α (PPAR α) Turnover by the Ubiquitin-Proteasome System Controls the Ligand-induced Expression Level of Its Target Genes”. Published, JBC Papers in Press, July 12, 2002, DOI 10.1074/jbc.M110598200.

[56] Albert J. Courey. (2008). “Mechanisms in Transcriptional Regulation”. Blackwell Publishing.

[57] Douglas B. Kitchen,Hélène Decornez, John R. Furr and Jürgen Bajorath. (2004). “Docking and Scoring in Virtual Screening for Drug Dscovery: Method and Applications”. doi:10.1038/nrd1549. Nature Review, volume 3, 935-949.

[58] Romano T. Kroemer. (2007). “Structure-Based Drug Design: Docking and Scoring”. Bentham Science Publishers Ltd; Current Protein and Peptide Science, 2007, 8, 312-328.

[59] P. Therese Lang, Demetri Moustakas.(2012). “DOCK 6.5 Users Manual”. University of California.

[60] Daniel Seeliger · Bert L. de Groot. (2010). “Ligand docking and binding site analysis with PyMOL and Autodock/Vina”. J Comput Aided Mol Des (2010) 24:417–422 DOI 10.1007/s10822-010-9352-6.

[61] Julien Michel. (2006). “The Use of Free Energy Simulations as Scoring Functions”. (Doctoral dissertation) . University of Southampton.

[62] H.-D. Holtje and G. Folkers. (1997). “Molecular Modeling Basic Principles and Applications”. VCH Verlagsgesellschaft mbH, Germany.

[63] Garrett M. Morris, David S. Goodsell, Ruth Huey, William E. Hart, Scott Halliday, Rik Belew, and Arthur J. Olson . (1999). “Autodock User Manual Version 3.03”. The Scripps Research Institute.

[64] Remo Perozzo, Gerd Folkers, and Leonardo Scapozza. (2004). “Thermodynamics of Protein–Ligand Interactions: History, Presence, and Future Aspects”. JOURNAL OF RECEPTORS AND SIGNAL TRANSDUCTION Vol. 24, Nos. 1 & 2, pp. 1–52.

[65] Daniel F. Veber, Stephen R. Johnson, Hung-Yuan Cheng, Brian R. Smith, Keith W. Ward and Kenneth D. Kopple. (2002). “Molecular Properties That Influence the Oral Bioavailability of Drug Candidates”. J. Med. Chem. 2002, 45, 2615-2623.

[66] Christopher A. Lipinski. (2004). “Lead- and drug-likecompounds: the rule-of-five revolution”. http://dx.doi.org/10.1016/j.ddtec.2004.11.007.

[67] Hongyu Zhao. (2007). “Scaffold selection and scaffold hopping in lead generation: a medicinal chemistry perspective”. Drug Discovery Today, Volume 12, Numbers 3/4.

[68] Mike M Hann1 and Tudor I Oprea (2004). “Pursuing the leadlikeness concept in pharmaceutical research”. Current Opinion in Chemical Biology 2004, 8:255–263.

[69] PDB ID: 1I7G. Cronet, P., Petersen, J.F., Folmer, R., Blomberg, N., Sjoblom, K., Karlsson, U., Lindstedt, E.L., Bamberg, K. (2001). “Structure of the PPARalpha and -gamma ligand binding domain in complex with AZ 242; ligand selectivity and agonist activation in the PPAR family”. Journal: (2001) Structure 9: 699-706.

[70] PDB ID: 2XYW. Keil, S., Matter, H., Schonafinger, K., Glien, M., Mathieu, M., Marquette, J.-P., Michot, N., Haag-Diergarten, S., Urmann, M., Wendler, W.(2011). “Sulfonylthiadiazoles with an unusual binding mode as partial dual peroxisome proliferator-activated receptor (PPAR) agonists with high potency and in vivo efficacy”. Journal: Chem Med. Chem 6: 633-653.

[71] PDB ID: 2ZK0. Waku, T., Shiraki, T., Oyama, T., Fujimoto, Y., Maebara, K., Kamiya, N., Jingami, H., Morikawa, K. (2009). “Structural insight into PPARgamma activation through covalent modification with endogenous fatty acids”. Journal: J. Mol Biol. 385: 188-199.

[72] J. Liang, H. Edelsbrunner, and C. Woodward. (1998) "Anatomy of protein pockets and cavities: Measurement of binding site geometry and implications for ligand design". Prot Sci, 7, 1884-1897.

[73] J. Liang, H. Edelsbrunner, P. Fu, P.V. Sudhakar and S. Subramaniam. (1998) "Analytical shape computing of macromolecules II: identification and computation of inaccessible cavities inside proteins". Proteins. 33, 18-29.

[74] J. Liang, H. Edelsbrunner, P. Fu, P.V. Sudhakar and S. Subramaniam. (1998) "Analytical shape computing of macromolecules I: molecular area and volume through alpha shape". Proteins. 33, 1-17.

[75] H. Edelsbrunner. (1994). “Three Dimensional Alpha Shapes”. ACM Transactions in Graphics, vol. 13, no. 1, page 43-72.

[76] The PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC.

[77] AutoDock (Version 4.2). [Computer software] . Molecular Graphics Laboratory, Department of Molecular Biology, The Scripps Research Institute.

[78] CHARMM (Chemistry at Harvard Molecular Mechanics) (c352b). Harvard University Cambridge, Massachusetts 02138, U. S. A.

[79] Irwin, Sterling, Mysinger, Bolstad and Coleman, J. Chem. Inf. Model. 2012, accepted for publication DOI: 10.1021/ci3001277.

[80] V. Lounnas and G. Vriend. AsteriX: A Web Server To Automatically Extract Ligand Coordinates from Figures in PDF Articles. J. Chem. Inf. Model., 2012, 52 (2), pp 568–576 DOI: 10.1021/ci2004303.

[81] J.-M. Yang and C.-C. Chen, "GEMDOCK: A generic evolutionary method for molecular docking", Proteins: Structure, Function and Bioinformatics, vol. 55, pp. 288-304, 2004.

[82] J.-M. Yang, Y.-F. Chen, T.-W. Shen, B. S. Kristal, and D. F. Hsu, "Consensus Scoring Criteria for Improving Enrichment in Virtual Screening ", Journal of Chemical Information and Modeling 2005.

[83] J.-M. Yang and C.-Y. Kao, "Flexible ligand docking using a robust evolutionary algorithm", Journal of Computational Chemistry ,vol. 21, pp. 988-998, 2000.

[84] J.-M. Yang, "Development and evaluation of a generic evolutionary method for protein-ligand docking", Journal of Computational Chemistry, vol. 25, pp. 843-857, 2004.

[85] Accelrys Draw (Version 4). Accelrys, Inc. USA.

APPENDIX A
CHARMM SCRIPTS

```
! protein topology and parameter
open read card unit 10 name toppar/top_all22_prot.rtf
read rtf card unit 10
open read card unit 20 name toppar/par_all22_prot.prm
read para card unit 20 flex
! Read pdb file and extract chain A and give it a name (PROA)
open read card unit 10 name 1i7g.pdb
read sequence pdb unit 10
generate PROA setup warn first NTER last CTER
open read card unit 10 name 1i7g.pdb
read coor pdb  unit 10 resid
!Add h-atoms
ic param
ic build
prnlev 0
hbuild
prnlev 5
ENERGY
open write card unit 10 name li7g_PPARA.pdb
write coor pdb  unit 10
open write card unit 10 name li7g_PPARA.crd
write coor unit 10 card
open write unit 10 card name li7g_PPARA.psf
write psf  unit 10 card
stop
```

```
! protein topology and parameter
open read card unit 10 name toppar/top_all22_prot.rtf
read  rtf card unit 10
open read card unit 20 name toppar/par_all22_prot.prm
read para card unit 20 flex
! Read PSF and Coordinates
open read unit 10 card name li7g_PPARA.psf
read psf  unit 10 card
open read unit 10 card name li7g_PPARA.crd
read coor unit 10 card
mini SD   nstep 1000 nprint 100
mini ABNR nstep 1000 nprint 100
open write unit 10 card name li7g_PPARA_min.pdb
write coor unit 10 pdb
open write unit 10 card name li7g_PPARA_min.crd
write coor unit 10 card
close unit 10
stop
```

CHARMM scripts: The first one adds H-atoms to the PDB file. The second one minimizes the structure.

APPENDIX B
PPAR CAVITIES

Cavities and voids of PPARα LBD

ID	No. of mouth	Cavity Area	
		Solvent Accessible	Molecular
8	1	0.106	17.58
9	1	1.392	16.79
10	0	0.456	34.04
11	1	4.785	23.99
12	0	0.002	25.20
13	1	1.668	23.68
14	1	5.680	20.4
15	0	0.281	31.87
16	1	3.895	27.61
17	0	0.875	37.59
18	1	2.216	34.68
19	1	4.276	40.33
20	1	1.086	5.62
21	1	1.627	33.25
22	2	8.325	45.79
23	1	26.759	36.87
24	1	6.783	104.99
25	2	29.290	70.05
26	1	12.071	36.78
27	1	9.662	68.14
28	1	11.486	71.24
29	0	15.545	119.53
30	1	20.759	64.72
31	1	28.657	64.05
32	0	13.144	79.02
33	3	25.118	119.84
34	1	29.032	71.09
35	1	38.292	79.57
36	1	61.626	110.56
37	3	648.009	1307.46

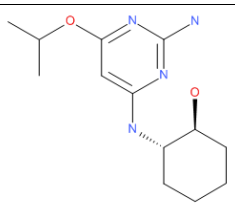
Cavities and voids of PPARδ LBD

ID	No. of mouths	Cavity Area	
		Solvent Accessible	Molecular
3	1	0.206	19.43
4	0	0.027	27.40
5	1	0.428	21.30
6	1	0.297	19.10
7	0	0.507	42.44
8	0	1.178	41.85
9	0	1.208	54.39
10	1	3.034	24.38
11	0	1.181	42.33
12	0	1.023	39.30
13	0	2.191	50.08
14	0	0.838	44.27
15	1	2.964	27.27
16	1	5.242	32.08
17	2	4.883	21.07
18	0	3.073	74.17
19	0	3.400	52.64
20	1	6.831	50.39
21	1	10.322	49.47
22	0	1.071	40.01
23	1	20.477	42.13
24	1	11.038	37.63
25	0	3.664	58.63
26	1	12.523	54.69
27	1	21.350	45.26
28	1	25.586	54.72
29	1	26.945	73.11
30	0	8.546	87.22
31	1	19.265	82.95
32	1	34.013	68.65
33	1	24.598	79.65
34	1	49.931	83.28
35	1	30.088	65.15
36	2	88.528	214.26
37	3	755.686	1274.94

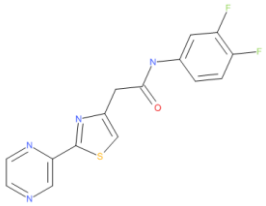
Cavities and voids of PPARγ LBD

ID	No. of mouths	Cavity Area	
		Solvent Accessible	Molecular
3	1	1.515	1.82
4	1	2.014	2.09
5	0	0.014	27.44
6	0	0.021	26.53
7	0	0.179	31.09
8	1	2.348	11.87
9	1	0.149	16.64
10	1	3.279	38.74
11	1	2.839	27.68
12	1	6.005	21.48
13	1	1.702	28.63
14	0	3.107	61.78
15	1	12.729	39.90
16	0	5.702	61.54
17	1	18.197	53.57
18	1	6.637	42.32
19	0	3.920	55.26
20	0	4.681	59.45
21	1	9.083	50.56
22	1	20.961	45.16
23	1	15.702	67.32
24	1	17.030	49.34
25	1	25.948	70.15
26	1	14.059	55.63
27	1	21.181	82.00
28	0	17.818	109.30
29	0	35.510	149.18
30	1	52.065	101.12
31	2	51.016	143.22
32	2	68.229	144.00
33	1	73.232	196.95
34	1	78.542	153.67
35	2	719.460	1496.05

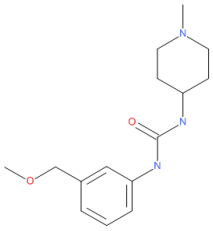
APPENDIX C
THE IDENTIFIED PAN-PPAR LIGANDS



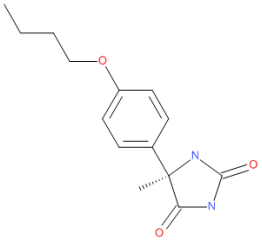
(1S,2S)-2-[(2-amino-6-isopropoxy-pyrimidin-4-yl)amino]cyclohexanol
c1(cc(nc(n1)N)N[C@@H]1[C@H](CCCC1)O)OC(C)C
p0.1-0



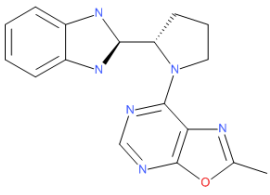
N-(3,4-difluorophenyl)-2-(2-pyrazin-2-ylthiazol-4-yl)acetamide
C(C(=O)Nc1ccc(c(c1)F)F)c1nc(sc1)c1nccnc1
p0.2-0



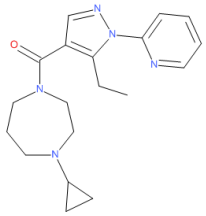
1-[3-(methoxymethyl)phenyl]-3-(1-methyl-4-piperidyl)urea
N(C(=O)Nc1cc(ccc1)COC)C1CCN(C)CC1
p0.3-1



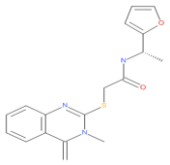
(5R)-5-(4-butoxyphenyl)-5-methyl-imidazolidine-2,4-dione
c1(ccc(cc1)[C@]1(NC(=O)NC1=O)C)OCCCC
p0.4-0



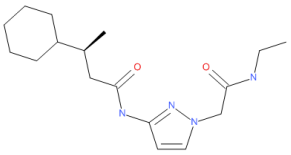
7-[(2S)-2-(2,3-dihydro-1H-benzimidazol-2-yl)pyrrolidin-1-yl]-2-methyl-oxazolo[5,4-d]pyrimidine
N1(CCC[C@H]1[C@H]1Nc2c(N1)cccc2)c1c2nc(C)oc2ncn1
p0.5-1



(4-cyclopropyl-1,4-diazepan-1-yl)-[5-ethyl-1-(2-pyridyl)pyrazol-4-yl]methanone
c1(c(cnn1c1ncccc1)C(=O)N1CCN(CCC1)C1CC1)CC
p0.6-0



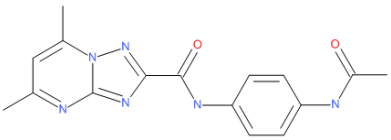
N-[(1S)-1-(2-furyl)ethyl]-2-(3-methyl-4-methylene-quinazolin-2-yl)sulfanyl-acetamide
c1(nc2ccccc2c(=C)n1C)SCC(=O)N[C@@H](C)c1occc1
p0.7-1



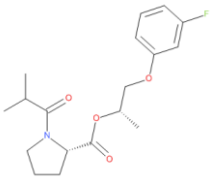
(3S)-3-cyclohexyl-N-[1-[2-(ethylamino)-2-oxo-ethyl]pyrazol-3-yl]butanamide

N(C(=O)C[C@H](C)C1CCCCC1)c1ccn(n1)CC(=O)NCC

p0.8-1

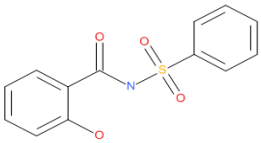


N-(4-acetamidophenyl)-5,7-dimethyl-[1,2,4]triazolo[1,5-a]pyrimidine-2-carboxamide
C(=O)(Nc1ccc(cc1)NC(=O)C)c1nc2n(c(cc(n2)C)C)n1
p0.9-1



[(1S)-2-(3-fluorophenoxy)-1-methyl-ethyl] (2S)-1-(2-methylpropanoyl)pyrrolidine-2-carboxylate

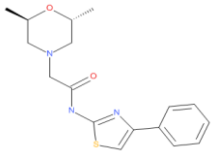
O([C@@H](C)COc1cc(ccc1)F)C(=O)[C@H]1N(CCC1)C(=O)C(C)C
p0.11-1



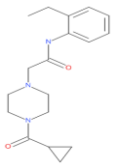
N-(benzenesulfonyl)-2-hydroxy-benzamide

N(C(=O)c1c(cccc1)O)S(=O)(=O)c1ccccc1

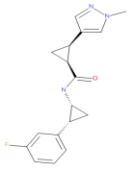
p0.13-0



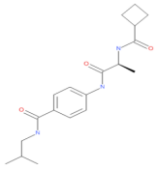
2-[(2R,6R)-2,6-dimethylmorpholin-4-yl]-N-(4-phenylthiazol-2-yl)acetamide
C(=O)(Nc1scc(n1)c1ccccc1)CN1C[C@@H](C)O[C@@H](C1)C
p0.14-0



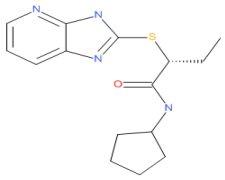
2-[4-(cyclopropanecarbonyl)piperazin-1-yl]-N-(2-ethylphenyl)acetamide
C(C(=O)Nc1c(cccc1)CC)N1CCN(CC1)C(=O)C1CC1
p0.15-0



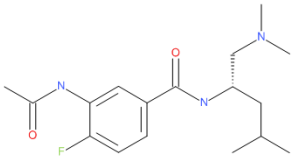
(1S,2R)-N-[(1R,2R)-2-(3-fluorophenyl)cyclopropyl]-2-(1-methylpyrazol-4-yl)cyclopropanecarboxamide
C(=O)(N[C@H]1[C@H](C1)c1cc(ccc1)F)[C@@H]1[C@@H](C1)c1cn(C)nc1
p0.17-0



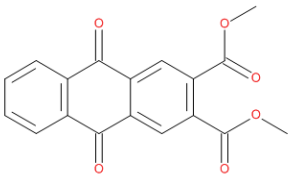
4-[[(2S)-2-(cyclobutanecarbonylamino)propanoyl]amino]-N-isobutyl-benzamide
c1(ccc(cc1)C(=O)NCC(C)C)NC(=O)[C@H](C)NC(=O)C1CCC1
p0.18-0



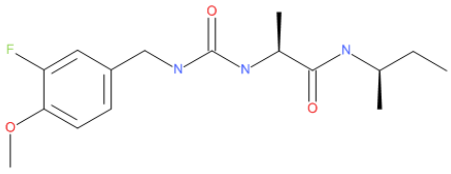
(2R)-N-cyclopentyl-2-(3H-imidazo[4,5-b]pyridin-2-ylsulfanyl)butanamide
[C@@H](CC)(C(=O)NC1CCCC1)Sc1[nH]c2c(n1)cccn2
p0.22-0



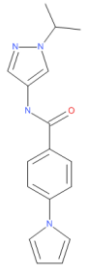
3-acetamido-N-[(1S)-1-(dimethylaminomethyl)-3-methyl-butyl]-4-fluoro-benzamide
N(C(=O)c1cc(c(cc1)F)NC(=O)C)[C@@H](CC(C)C)CN(C)C
p0.23-0



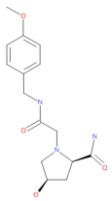
dimethyl 9,10-dioxoanthracene-2,3-dicarboxylate
c1(cc2c(cc1C(=O)OC)C(=O)c1ccccc1C2=O)C(=O)OC
p0.24-1



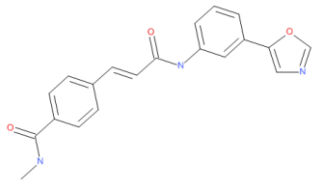
(2S)-2-[(3-fluoro-4-methoxy-phenyl)methylcarbamoylamino]-N-[(1R)-1-methylpropyl]propanamide
N(C(=O)NCc1cc(c(cc1)OC)F)[C@@H](C)C(=O)N[C@H](C)CC
p0.25-0



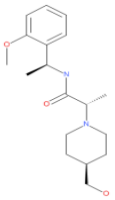
N-(1-isopropylpyrazol-4-yl)-4-pyrrol-1-yl-benzamide
c1(ccc(cc1)n1cccc1)C(=O)Nc1cn(nc1)C(C)C
p0.26-1



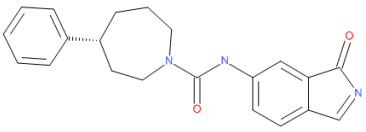
(2R,4R)-4-hydroxy-1-[2-[(4-methoxyphenyl)methylamino]-2-oxo-ethyl]pyrrolidine-2-carboxamide
N(C(=O)CN1[C@H](C[C@H](C1)O)C(=O)N)Cc1ccc(cc1)OC
p0.27-0



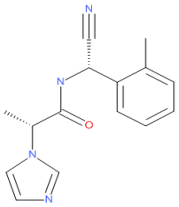
N-methyl-4-[(E)-3-(3-oxazol-5-ylanilino)-3-oxo-prop-1-enyl]benzamide
C(=O)(Nc1cc(ccc1)c1ocnc1)/C=C/c1ccc(cc1)C(=O)NC
p0.29-0



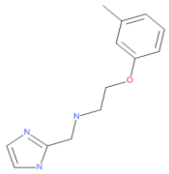
(2S)-2-[4-(hydroxymethyl)-1-piperidyl]-N-[(1S)-1-(2-methoxyphenyl)ethyl]propanamide
N(C(=O)[C@H](C)N1CC[C@@H](CC1)CO)[C@@H](C)c1c(cccc1)OC
p0.30-1



(4R)-N-(3-oxoisoindol-5-yl)-4-phenyl-azepane-1-carboxamide
[C@H]1(CCCN(CC1)C(=O)Nc1cc2c(cc1)C=NC2=O)c1ccccc1
p0.31-1

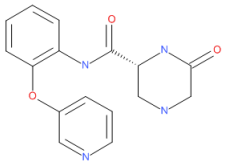


(2R)-N-[(S)-cyano(o-tolyl)methyl]-2-imidazol-1-yl-propanamide
[C@@H](C#N)(c1c(C)cccc1)NC(=O)[C@@H](C)n1ccnc1
p1.1-0



N-(1H-imidazol-2-ylmethyl)-2-(3-methylphenoxy)ethanamine
C(Oc1cccc(C)c1)CNCc1[nH]ccn1

p1.2-1



(2R)-6-oxo-N-[2-(3-pyridyloxy)phenyl]piperazine-2-carboxamide
c1ccc(c(c1)NC(=O)[C@@H]1NC(=O)CNC1)Oc1ncccc1
p1.3-1

APPENDIX D
H-BOND FORMING PPAR RESIDUES

Ligand ID#	PPARα		PPARγ		PPARβ/δ	
	Binging Energy	H-bond	Binging Energy	H-bond	Binging Energy	H-bond
p0.1-0	-7.74	SER280:HG1	-7.34	ILE267:HN	-7.94	ARG248:HE
	-6.68	TYR334:HN		SER342:HN	-7.73	THR253:OG1
	-6.65	TYR334:HN	-7.34	SER342:HN		CYS249:HG1
	-6.59	THR279:HG1	-6.99		-7.31	HSD287:HE2
		TYR334:HN	-6.77	SER342:HN		HSD413:HE2
	-6.45	TYR334:HN	-6.76	GLU343:HN		TYR437:OH
	-6.33	THR279:HG1	-6.72	SER342:HN	-7.13	THR252:HG1
		ALA333:HN	-6.38	LYS244:HN		ALA306:HN
	-6.26	TYR334:HN		LYS244:HZ3	-6.98	MET192:HN
		ALA333:HN		GLN345:HE22		ASN307:HN
		THR279:OG1	-6.27	LYS244:Hz3	-6.35	
	-6.09	TYR334:HN		GLN345:HE22		
		ALA333:HN				
		GLU282:OE1				
p0.2-1	-7.45	ALA333:HN	-7.38		-7.63	
		TYR334:HN	-7.37	ILE267:HN	-7.55	CYS249:HG1
	-7.29	MET220:O	-6.9	SER342:HN	-6.99	MET192:HN
		ASN219:HD22		GLU343:HN		THR252:HG1
	-7.0	LEU331:O	-6.79	SER342:HN	-6.91	GLN230:HN
		THR279:HG1	-6.71	CYS285:HG1	-6.8	
		TYR334:HN	-6.7	SER289:HG1	-6.66	
	-6.83	ALA333:HN			-6.38	
		TYR334:HN			-6.32	LYS229:HN
	-6.62					ALA306:HN
p0.3-0	-7.06	THR279:HG1	-7.55	SER342:HN	-6.31	
		TYR334:HN	-7.52	SER342:HN		
	-7.03	TYR334:HN	-7.38			
		ASN219:HD22	-7.37	SER342:HN		
		THR279:OG1	-7.34	HSD266:HD1		
	-7.0	TYR334:HN		GLU291:OE1		
	-6.86	ASN219:HD22	-7.22			
		TYR334:HN	-7.13			
	-6.59	MET220:HN	-7.06	ARG288:HE		
	-6.24	TYR334:HN	-6.63	SER342:HN		
		ALA333:HN				
		CYS278:HG1				

p0.4-1	-6.7 -6.68 -6.48 -6.03	ASN219:HD22 ASN219:HD22 TYR334:HN SER323:HG1 ASN219:HD22	-6.88 -6.81 -6.79 -6.71 -6.45 -6.36 -6.32 -6.06	HSD266:HD1 ILE267:HN SER342:HN ILE267:HN SER342:HN ILE267:HN SER342:HN ILE267:HN ILE267:O ILE267:O ILE267:HN ILE267:O ILE267:HN	-6.69 -6.67 -6.66 -6.5 -6.27 -6.03	GLY308:HN ALA306:HN ASN307:HN MET192:HN LYS229:HN MET192:O SER296:HG1 THR252:HG1
p0.5-0	-8.08 -8.06 -7.92 -7.67 -7.6	TYR334:HN TYR334:HN TYR334:HN ALA333:HN TYR334:HN SER280:HG1	-8.56 -8.48 -8.44 -8.37 -8.19 -8.07 -6.41	SER342:HN GLU343:HN SER342:HN SER342:HN SER342:HN ILE267:O	-8.26 -8.24 -8.19 -7.84 -7.34 -7.32 -7.32 -6.75	THR252:HG1 ASN307:HN LEU304:O ASN307:HN THR252:HG1 ASN307:HN LEU304:O ALA306:HN ASN307:HN ARG248:HE ASN307:HN
p0.7-0	-7.11 -6.93		-8.65 -7.94 -7.79 -7.67 -6.19	SER342:HN GLU343:HN ARG288:HE SER342:HN LYS244:HN SER245:O	-8.49 -7.96 -7.68 -7.18 -7.16 -6.88 -6.17	 ALA306:HN ASN307:HN GLN230:HN ASN307:HN ALA306:HN
p0.8-1	-7.83 -7.29 -7.03 -6.76 -6.6	TYR334:HN ALA333:HN ASN219:HD22 TYR334:HN	-8.27 -7.95 -7.41 -7.36 -7.32 -7.02	ARG288:HE SER342:HN ARG288:HE SER342:HN SER342:HN ARG288:HE	-8.03 -7.31 -7.18 -7.04 -6.94 -6.05	 ASN307:HN THR256:HG1 THR256:HG1

p0.9-0	-8.09 -7.7 -7.5 -7.49 -7.37	GLU282:OE1 TYR334:HN ASN219:HD22 THR279:HG1 ALA333:HN MET220:HN GLU282:OE1 CYS278:HG1 TYR334:HN	-9.73 -9.27 -9.26 -9.1 -8.43 -8.41 -8.18 -8.15	ILE267:HN SER342:HN GLU343:HN GLU291:OE1 GLU343:HN GLU343:HN GLU291:OE1 GLU343:HN GLU291:OE1 SER342:HN GLU291:OE1 HSD266:HD1	-8.01 -7.48 -7.19	THR252:HG1
p0.11-0	-6.83 -6.75 -6.25	 TYR334:HN ASN219:HD22	-6.76 -6.71 -7.17	 SER342:HN GLU291:OE1 HSD266:HD1	-7.09 -6.94 -6.69 -6.38 -6.06	ALA306:HN ASN307:HN ALA306:HN ASN307:HN LYS229:HN ASN307:HN THR256:HN
p0.13-0	-7.59 -7.45 -7.37 -7.33 -7.19 -7.02 -6.98 -6.94 -6.91	GLU286:OE2 MET220:HN ASN219:HD22 HSD440:HE2 MET220:O MET220:HN GLU286:OE2 MET220:HN ASN219:HD22 ASN219:HD22 ASN219:HD22 GLU286:OE2	-8.38 -7.89 -7.85 -7.84 -7.75 -7.64 -7.47 -7.43 -7.32	HSD266:HD1 ILE267:HN ILE267:HN ARG288:HE SER342:HN ILE267:HN SER342:HN ILE267:HN HSD266:HD1 SER342:HN SER342:HN SER342:HN	-8.45 -7.9 -7.84 -7.8 -7.77 -7.72 -7.6 -7.42 -7.14 -6.89	GLN230:HN ALA306:HN THR252:HG1 HSD287:HE2 LYS229:HN ASN307:HN LYS229:HN THR252:HG1 ALA306:HN LYS229:HN THR252:HG1 TYR437:OH HSD287:HE2 ILE290:O ASN307:HN ALA306:HN
p0.14-0	-8.0 -7.98 -7.86 -7.77 -7.66 -7.41 -6.28	THR279:HG1 TYR334:HN LYS364:HZ3	-8.57 -8.37 -7.97 -7.77 -7.63 -7.39 -7.36 -6.43	SER342:HN ILE267:O HSD266:HD1 HSD449:HE2 SER342:O SER342:HN SER342:HN GLU369:OE1	-7.96 -7.86 -7.64 -7.23 -7.11 -6.8	THR252:HG1 ALA306:HN ALA306:HN CYS249:O ASN307:HN

p0.15-0	-7.6 -6.92 -6.53 -6.17	 THR279:OG1 ASN219:HD22 TYR334:HN GLU286:OE2 TYR334:HN	-7.1 -6.92 -6.88	 GLU369:OE1	-6.29	
p0.17-0	-7.28 -7.03 -6.6 -6.23	TYR334:HN TYR334:HN CYS275:HG1	-6.77 -6.77 -6.58	 ARG288:HE	-7.09 -6.67 -6.65 -6.63 -6.61 -6.26 -6.18	MET192:HN ASN307:HN ALA306:HN ASN307:HN MET192:HN
p0.18-1	-7.51 -7.39 -7.38	THR279:OG1 ALA333:HN THR279:OG1 ASN219:HD22 ALA333:HN	-7.73 -7.67 -7.53 -7.36 -7.11	HSD266:HD1 ARG288:HE SER342:HN	-7.82 -7.26 -7.23 -6.49	THR252:HG1 ALA306:HN
p0.22-0	-7.48 -7.48 -7.34 -7.33 -7.28 -7.26 -7.17 -6.65 -6.49	MET220:HN TYR334:HN ALA333:HN TYR334:HN THR279:HG1 ALA333:HN TYR334:HN ALA333:HN TYR334:HN ALA333:HN TYR334:HN MET220:O MET220:HN ALA333:HN	-7.71 -7.59 -7.42 -7.34 -7.26 -7.14 -7.11 -6.81 -6.8	 SER342:HN CYS285 ILE267:O HSD266:HD1 HSD266:HD1	-7.4 -7.11 -6.99 -6.92 -6.84 -6.67	CYS249:HG1 MET192:HN ASN307:HN LYS229:HN ASN307:HN ASN307:HN THR256:HG1
p0.23-0	POOR		-6.3		-5.77	THR252:HG1 ALA306:HN
p0.24-0	-8.12 -7.95 -7.61 -7.1 -6.82	ASN219:HD22 MET220:HN TYR334:HN GLU286:OE2 ASN219:HD22 TYR334:HN CYS275:HG1	-8.44 -7.68 -7.5 -7.44 -7.29 -7.28 -6.59 -6.35	SER342:HN HSD266:HD1 LYS244:HZ3	-7.69 -7.2 -7.16	ALA306:HN ASN307:HN MET192:HN ASN307:HN

p0.25-1	-7.48 -6.73 -6.6 -6.47 -6.27	ASN219:HD22 MET220:HN TYR334:HN ALA333:HN THR279:OG1 ASN219:HD22 THR283:HG1 THR279:OG1 ALA333:HN	-6.83 -6.81 -6.45 -6.15	SER342:HN SER342:HN ILE267:O SER342:HN	-6.67 -6.63 -6.62 -6.51 -6.33 -6.2 -6.05	ASN307:HN ASN307:HN THR253:HG1 THR256:HG1 THR252:HG1
p0.26-1	-7.13 -7.1 -7.09 -6.88 -6.82 -6.64	ASN219:HD22 TYR334:HN ASN219:HD22 TYR334:HN TYR33:HN	-7.05 -07.02 -7.01 -6.96 -6.95		-7.49 -7.46 -7.41 -7.32 -7.3 -7.1 -7.07 -6.77	THR252:HG1 THR252:HG1 THR252:HG1 THR256:HG1
p0.27-0	-6.61 -6.49 -6.46 -6.44 -6.35 -6.21	THR279:HG1 THR279:OG1 TYR334:HN ALA333:HN ALA333:HN THR279:OG1 TYR334:HN ALA333:HN TYR334:HN ASN219:HD22 GLU286:OE1 GLU282:O LEU331:O HSD440:HE2	-7.13 -7.02 -6.39 -6.33 -6.28	LYS367:HZ3 ILE267:HN	-6.28 -6.08	MET192:HN ALA306:HN MET192:HN
p0.29-0	-9.16 -8.97 -7.87 -7.82 -7.78	SER280:HG1 SER323:HG1 SER323:HG1	HSD26 6:HD1 ARG2 88:HE		-8.28 -8.22 -8.15 -7.89 -7.68 -7.67 -7.33	MET192:O MET293:O MET192:O MET293:O ASN307:HN

p0.31-0	-9.33 -8.64	GLY335:HN	-9.43 -9.06 -9.02 -8.96 -8.92 -8.68 -8.65 -7.49 -7.02 -6.87	ARG288:HE LYS474:HZ2 LYS474:HZ2 	-9.54 -9.38 -9.28 -9.08 -8.92 -8.84 -8.44 	GLN230:HN THR256:HG1 THR256:HG1 ASN307:HN ASN307:HN
p1.1-0	-6.74	ASN219:HD22	-6.62 -6.57 -6.54 -6.48 -6.4 -6.35 -6.27	SER342:HN GLU343:HN ARG288:HE 	-7.25 -7.1 -7.09 -6.36 -6.23 -6.15	ASN307:HN ALA306:HN ASN307:HN THR256:HG1
p1.3-0	-8.1 -8.08 -7.76 -7.45 -7.06 -6.97	MET220:HN TYR334:HN ASN219:HD22 TYR334:HN THR279:HG1 TYR334:HN TYR334:HN ALA333:HN TYR334:HN GLU282:OE1 THR279:OG1 ASN219:HD22	-8.4 -8.18 -8.18 -7.73 -7.72 -7.65	GLU343:HN GLU291:OE1 SER342:HN GLU343:HN GLU291:OE1 SER342:HN GLU343:HN SER342:HN ILE267:HN SER342:HN ILE267:HN SER342:HN	-7.35 -6.35 -6.25 	GLN230:HN ASN307:HN ASN307:HN MET192:HN LEU304:O THR256:HG1

APPENDIX E
AMINO ACID LIST

Nonpolar (hydrophobic) amino acids

Amino acids	Three letter code	Single letter code
Glycine	Gly	G
Alanine	Ala	A
Valine	Val	V
Leucine	Leu	L
Isoleucine	Ile	I
Methionine	Met	M
Phenylalanine	Phe	F
Tryptophan	Trp	W
Proline	Pro	P

Polar (hydrophilic) amino acids

Amino acids	Three letter code	Single letter code
Serine	Ser	S
Threonine	Thr	T
Cysteine	Cys	C
Tyrosine	Tyr	Y
Asparagine	Asn	N
Glutamine	Gln	Q

Electrically charged (negative and hydrophilic) amino acids

Amino acids	Three letter code	Single letter code
Aspartic acid	Asp	D
Glutamic acid	Glu	E

Electrically charged (positive and hydrophilic) amino acids

Amino acids	Three letter code	Single letter code
Lysine	Lys	K
Arginine	Arg	R
Histidine	His/Hsd	H