2019

Avoiding Adverse Effects: New Ideas in Drug Discovery for Targeting PPARγ

Trey M. Patton

Avoiding Adverse Effects: New Ideas in Drug Discovery for Targeting PPARγ

By

Trey Michael-Douglas Patton

Bachelor of Arts in Chemistry, High Point University

High Point, North Carolina, 2017

Master's Thesis

Presented in partial fulfillment of the requirements for the degree of

Master of Sciences In

Pharmaceutical Sciences and Drug Design

The University of Montana

Missoula, MT

May 2019

Approved by:

Scott Whittenburg, Dean of The Graduate School

Graduate School

Dr. Travis Hughes, Co-advisor,

Department of Biomedical & Pharmaceutical Sciences

Dr. Philippe Diaz, Co-advisor,

Department of Biomedical & Pharmaceutical Sciences

Dr. Keith Parker,

Department of Biomedical & Pharmaceutical Sciences

Dr. Nigel Priestley, Chair

Department of Chemistry

Patton, Trey, M.Sc., Spring 2019

Pharmaceutical Sciences and Drug Design

Travis Hughes, Co-advisor

Philippe Diaz, Co-advisor

Acknowledgements

This work was accomplished with the guidance of Dr. Travis Hughes and Dr. Philippe Diaz. I would like to thank the Department of Biomedical and Pharmaceutical Sciences for giving me the opportunity to contribute to this work as well as supporting me while doing so. This department made suffering through the Montana weather worth it. I would also like to give special thanks to my advisors, committee, colleagues, family, and teachers. Without their guidance and support, I never would have accomplished as much as I was able to.

First, Dr. Travis Hughes and Dr. Philippe Diaz, thank you for the guidance, patience, and opportunity to contribute to your work. I came into this program hoping to learn more about computer-aided drug design and organic synthesis and I was met with great opportunities for both. Dr. Zahra Heidari and Dr. Ian Chrisman, thank you both for answering countless questions about simulations and PPARγ no matter how trivial or poorly worded they were. Michelle Nemetchek and Mariah Rayl, I cannot show enough gratitude for your kindness and willingness to make yourselves available on short notice. Dr. Desiree Mendes and Dr. Nicholas Wageling, thank you both for pushing me to think of answers to my questions regarding synthesis preparation, reaction conditions, mechanisms, etc. instead of giving me the answer. Ben Uhlenbruck, I came into the Diaz lab with little organic chemistry lab experience and I am forever grateful for your patience, willingness to help, and keeping an eye on my reactions. To my family, I would not be where I am today without your constant support. Finally, a thanks to the educators throughout my academic career that laid the foundation for my curiosity in the sciences.

Abstract

Peroxisome proliferator-activated receptor gamma (PPARγ) has been a drug target to treat type 2 diabetes for the last 20 years when rosiglitazone and pioglitazone were approved by the FDA in 1999. While effective at increasing insulin sensitivity, these drugs cause serious adverse effects due to their full agonist characteristics. For that reason, drug discovery efforts have attempted to reduce or prevent the amount of agonist character of new PPARγ targeting ligands. Unfortunately, there have been no new FDA approved drugs for the receptor. There is a need for new ideas to produce better quality pharmaceuticals that lessen the impact of adverse effects. This work aims to propose and expand on new ideas: biased agonism and interactions that anchor a consistent binding mode.

Table of Contents

List of Tables, Figures, and Schemes

Chapter 1 Introduction

1.1 Peroxisome proliferator-activated receptor gamma

Peroxisome proliferator-activated receptors (PPARs) are a class of ligand-activated transcription factors that regulate gene transcription and belong to the nuclear hormone receptor superfamily. The PPARs are essential in energy metabolism and consist of three, uniquely expressed subtypes: PPARα, PPARβ/δ, and PPARγ.[1] PPARγ has three isoforms, and the second isoform is the primary target of this work. PPARγ2 is expressed mainly in adipose tissue and is the master regulator of adipogenesis.[2] It is a drug target for the treatment of type 2 diabetes mellitus because the receptor is linked to increasing insulin sensitivity. Initially, researchers were unaware of the target these drugs were affecting, and the mechanism of action was unknown. Eventually, PPARγ was identified as the drug target. For simplicity, PPARγ2 will be referred to as PPARγ from this point forward and all residue counts will be in PPARγ2 numbering.

PPARγ is similar to other nuclear receptors in that it has a conserved domain structure consisting of domains A-F (Figure 1a).[3] Domain A/B is the activation function 1 (AF-1) region, C is the DNA binding domain, and D is the hinge. Domain E/F is comprised of the ligand binding domain (LBD), which is made up of the ligand binding pocket, the activation function 2 (AF-2) region, and the dimer interface.[4] The PPARγ LBD consists of 13 α-helices (H1-H12, H2'), an Ω-loop, and a β-sheet region (Figure 1b). A large (1200 $Å^2$) T-shaped ligand binding pocket is located at the center of the LBD and is divided into three branches, each with unique ligand binding characteristics that may contribute to different functional effects. Branch I and Branch

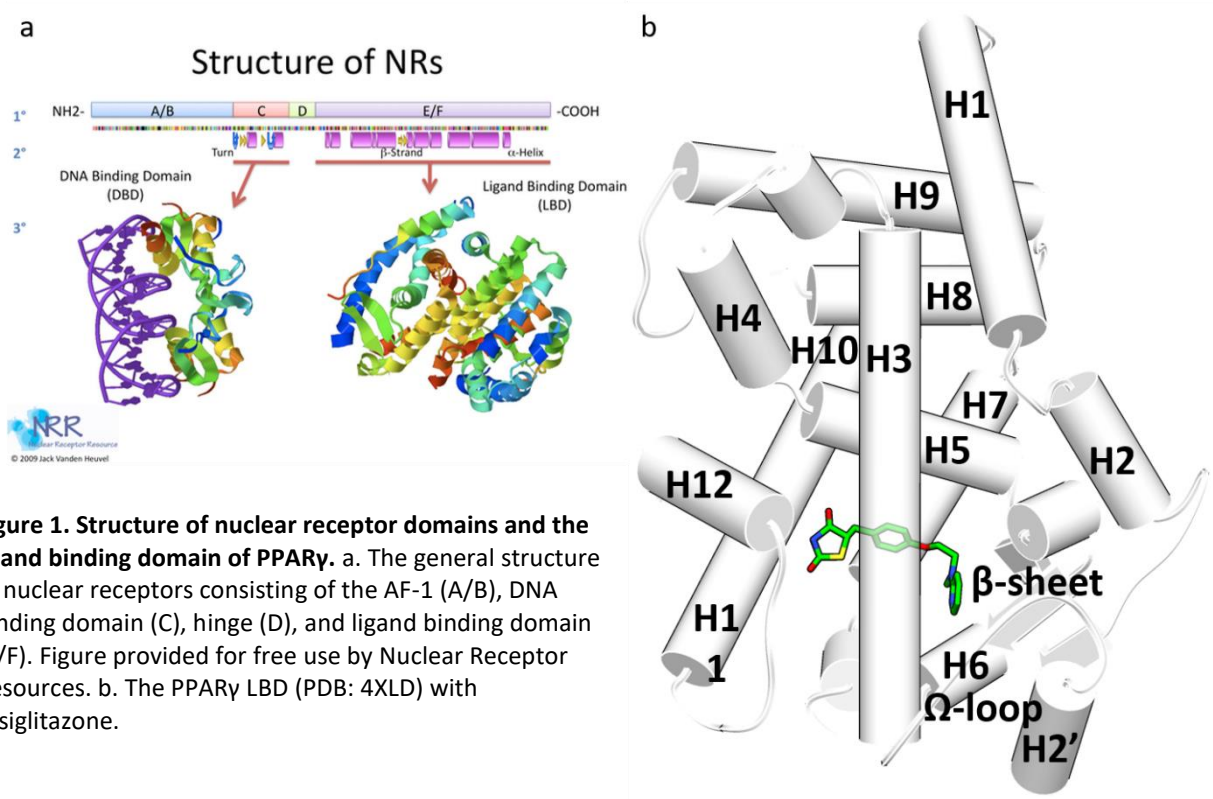


Figure 1. Structure of nuclear receptor domains and the ligand binding domain of PPARγ. a. The general structure of nuclear receptors consisting of the AF-1 (A/B), DNA binding domain (C), hinge (D), and ligand binding domain (E/F). Figure provided for free use by Nuclear Receptor Resources. b. The PPARγ LBD (PDB: 4XLD) with rosiglitazone.

II are hydrophobic cavities consisting of H3, H5, H11, and H12 and H2', H3, H6, H7 and the β-sheet region, respectively.[5] Branch III is both hydrophobic and hydrophilic in character, consisting of H2, H3, H5, and the β-sheet region.[5] The AF-2 region is made up of H3, H4, H5, and H12 and provides a site for coactivator binding.[5] PPARγ is an obligate heterodimer with retinoid X receptor α (RXRα), where the two make interactions at the LBD dimer interface.[6,7] The ligand bound state of the PPARγ:RXRα heterodimer will then direct the formation of a complex with coregulator proteins (corepressors and coactivators) that bind at the AF-2 region.[4]

In the absence of agonist binding, the PPARγ:RXRα heterodimer can recruit corepressor proteins to form a complex that represses PPAR activated transcription.[8,9] A small number of corepressors have been observed to interact with PPARγ, and this work will be looking specifically at nuclear receptor corepressor 1 (NCoR) and silencing mediator of retinoic acid and thyroid hormone receptor (SMRT).[10] In the presence of ligand binding, the PPARγ:RXRα

heterodimer will bind to DNA through a PPAR response element (PPRE) causing corepressors to dissociate and the heterodimer to recruit coactivators that induce transcriptional activity.[11] There have been a larger number of coactivators identified to interact with PPARγ, with examples including CREB-binding protein (CBP-1), mediator of RNA polymerase II transcription subunit 1 (MED1), and PPARγ coactivator 1 alpha (PGC1α).[10] Endogenous ligands of PPARγ are fatty acids and fatty acid metabolites, eicosanoids, and even serotonin has been observed to bind to PPARγ.[12–14] Independent of endogenous ligands, a diverse collection of synthetic ligands intended to increase insulin sensitivity in T2DM patients have been developed.

1.2 Synthetic ligands that target PPARγ

The diverse collection of synthetic ligands can be separated into four classes: full agonists, partial agonists, antagonists, and inverse agonists. Each class may produce unique conformational changes in PPARγ that lead to unique functional effects.

Full agonists fully activate PPARγ transcription through a well described mechanism. One class of full agonists, the thiazolidinediones (TZDs), have progressed into two FDA approved antidiabetic agents: rosiglitazone (Avandia) and pioglitazone (Actos). Crystal structures of rosiglitazone-bound PPARγ have shown putative hydrogen bond interactions between the thiazolidinedione functional group and Branch I residues His323 (H5), His449 (H11), and Tyr473 (H12) that stabilize H12.[15–17] H12 stabilization locks the AF-2 region into a favorable binding surface for coactivators, and this leads to full activation of PPARγ transcription.[4] While effective at increasing insulin sensitivity in patients with T2DM, rosiglitazone and pioglitazone have been under scrutiny after post-marketing surveillance showed an association with serious adverse effects that stem from PPARγ's role as the master

regulator of adipogenesis and full activation. These drugs upregulate adipogenesis leading to weight gain and increased risk of bone fractures due to increased adipocyte differentiation in bones.[18,19] These adverse effects have been attributed to the interactions described above regarding H12 stabilization and full activation of PPARγ. Drug discovery efforts have been conducted to identify new ligands that continue to increase insulin sensitivity while partially activating transcription by stabilizing other regions of PPARγ through H12 independent mechanisms.[20]

These ligands, called partial agonists, do not fully stabilize H12, preventing PPARγ from becoming locked in an activated conformation. This partial stabilization results in lower transcriptional activity compared to the TZDs. Currently, the maximum transcriptional output observed by H12 independent partial agonists is 80% that of rosiglitazone.[5] These ligands occupy the regions around Branch II and III and interact with H3 and the β-sheet region through hydrogen bond interactions with residues such as Arg288 and Ser342 and various hydrophobic interactions. The exact mechanism that leads to partial, indirect stabilization of H12 has yet to be identified; however, it is believed that stabilization of H3 and the β-sheet is responsible.[5] As antidiabetic agents, partial agonists improve insulin sensitivity while showing no indications of increasing fatty acid storage *in vitro*. Animal model experiments have shown similar, positive results.[21] Despite these improvements, no partial agonists have reached FDA approval as PPARγ targeting antidiabetic agents. This lack of success contributed to the development of ligands that aim to inhibit S273 phosphorylation which may correlate with insulin sensitization.

Preventing the activation of PPARγ involves the recruitment of corepressors to the PPARγ:RXRα heterodimer. This can be accomplished through the binding of antagonists and/or

inverse agonists. The mechanism that leads to heterodimer recruitment of corepressors is not well understood, and the activity of antagonism and inverse agonism are often conflated in the literature. There is, however, a clear distinction between these two ligand types that have been observed through structural differences and intensity of corepressor affinity in defined biochemical assays. In general, the proposed mechanism that leads to corepressor recruitment involves moving H12 away from the activated conformation into either a disordered, extended conformation or some other unknown conformation that provides a favorable binding site for corepressors. Antagonist-bound PPARγ has fluidity in the conformations PPARγ can adopt leading to both corepressors and coactivators to be recruited and producing basal level activity. The effects observed in both antagonists and inverse agonists likely arise from various mechanisms. One of these mechanisms may be that functional groups on the ligands prevent key intra-protein interactions between residues that stabilize H12 in Branch I (Figures 2c and 2d).[22]

To summarize, each class of ligand causes PPARγ to adopt unique conformational ensembles, as evidenced by ^{19}F NMR, that reflect the type of activity that is observed.[23] Full agonists and inverse agonists cause PPARγ to adopt the activated and deactivated conformations, respectively, with minimal variability. Partial agonists and antagonists have more conformational fluidity causing the receptor to adopt both the activated and deactivated conformations. Fluorine NMR has been used to support the flexibility, or lack thereof, of PPARγ bound to each class of ligand. Full agonists and inverse agonists have thin peaks in NMR spectra indicating less variability, while partial agonists and antagonists have wider peaks indicating multiple conformations (Figure 2).[23]

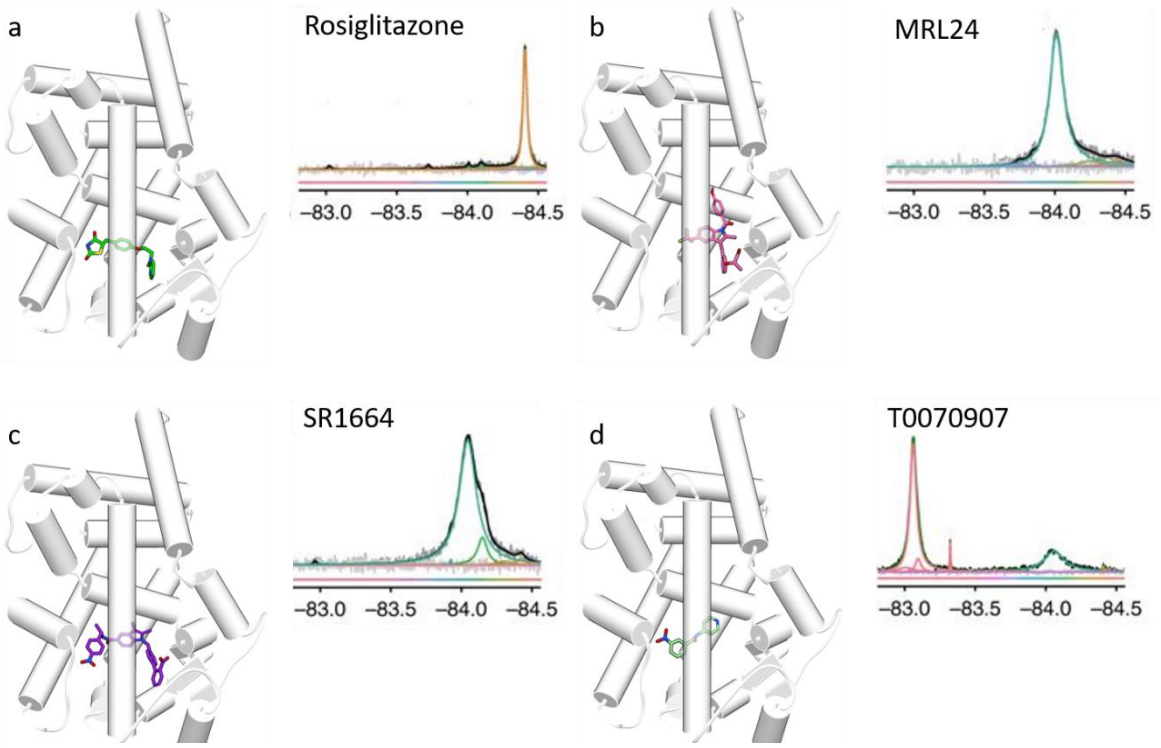


Figure 2. Full agonist, partial agonist, antagonist, and inverse agonist aligned to rosiglitazone-bound PPARγ and their respective protein NMR data. a. Full agonist rosiglitazone bound to PPARγ (PDB: 4XLD). Rosiglitazone (green) occupies Branches I, II, and III. The thin NMR peak suggests PPARγ adopts a stable conformation. b. Partial agonist MRL24 aligned in PPARγ (PDB: 2Q5P, Chain A). MRL24 (pink) occupies Branches II and III. The wide NMR peak suggests PPARγ adopts multiple conformations. c. Antagonist SR1664 aligned in PPARγ (PDB: 4R2U, Chain A). SR1664 (purple) occupies Branches I and III. The wide NMR peak suggests PPARγ adopts multiple conformations. d. Inverse agonist T0070907 aligned in PPARγ (PDB: 6C1I, Chain A). T0070907 (teal) occupies Branch I but can also occupy Branches II and III due to free rotation. The thin NMR peak suggests PPARγ adopts a stable conformation.
NMR data was modified from Figure 2 of the Nat. Comm. publication by *Chrisman et. al.* doi: 10.1038/s41467-018-04176-x

The development of partial agonists, antagonists, and inverse agonists have yielded a diverse series of ligands aimed at increasing insulin sensitivity while reducing or preventing the adverse effects associated with full agonists. Despite the time and effort spent designing, synthesizing, and evaluating such ligands, none have been able to progress through clinical trials and FDA approval. The mechanisms of action for the different ligand types are still being elucidated, and this work will aid in the development of a comprehensive drug design strategy that can be used to discover compounds that treat T2DM by targeting PPARγ without adverse effects. Currently, there is a need for new ideas and approaches to identify this drug design

strategy. Biased agonism in PPARγ is a new idea, and the work covered in Chapter 2 will explore this concept.

1.3 Biased agonism in PPARγ

Biased agonism, also called functional selectivity, is the ligand-dependent selectivity for a specific signal transduction pathway over other pathways at the same receptor.[24] Biased agonism is most extensively described in G-protein coupled receptors. Two papers describe observations of biased agonism exclusively in PPARγ. The first paper contributes the ability for the PPARγ ligand magnolol to induce the transcription of PPAR response element (PPRE) mediated by the RXRα:PPARγ heterodimer instead of RXR-response element (RXRE) mediated by the RXRα:RXRα homodimer to biased agonism.[25] The second paper describes bisphenol A diglycidyl ether (BADGE) as a biased ligand due to agonist effects in one cell type or tissue and an antagonist effect in another cell type or tissue.[26,27]

In this work, biased agonism involves the use of a ligand that biases the coregulator that is recruited to the PPARγ:RXRα heterodimer. Many nuclear receptor coregulators have been identified.[10] It is believed in the literature that each coregulator is associated with the upregulation or downregulation of a specific set of genes; however, these associations have yet to be fully described in PPARγ and nuclear receptors in general.[28] If a ligand can introduce a bias that causes the recruitment of a coregulator that does not upregulate genes involved in adipogenesis we could, in theory, prevent the adverse effects associated with current treatments even if the drug is a full agonist. For example, PGC1α has been associated with upregulating the transcription of genes that cause the browning of white adipose tissue, a process that ultimately reduces weight in mice.[29] A drug could be developed that biases the

recruitment of PGC1α and could be used as a treatment for obesity. This definition of biased agonism in PPARγ has not been researched yet, and the following work in Chapter 2 looks at the selective recruitment of corepressors to the PPARγ:RXRα heterodimer, specifically, the bias recruitment of NCoR over SMRT. This may lead to specialized effects based on the genes regulated by NCoR. The aim of this research is to modulate the bias effect observed in nTZDpa, a partial agonist of PPARγ. If biased agonism is a possibility in PPARγ, this may apply to other nuclear receptors due to their highly conserved structures, and strategies can be developed that exploit coregulator recruitment bias in all nuclear receptors.

1.4 Objectives and Methodologies

This thesis is a combination of two separate projects. The first project explores the structure-activity relationships between PPARγ and nTZDpa that cause selective recruitment of the corepressor NCoR. The second project is the characterization of a novel PPARγ partial agonist, NMP422, with a focus on the ligand's structural configuration that allows a dual binding mode. Overall, each project suggests new ideas that can be used in drug discovery to reduce or prevent adverse effects in PPARγ targeting drugs through biased agonism and consistent protein-ligand interactions.

In the first project, covered in Chapter 2, nTZDpa shows biased agonist characteristics and the cause of these characteristics are explored. *In silico* screening is used with the goal of providing support to the importance of the 5-chloro-2-indole carboxylic acid scaffold at directing a consistent binding mode across a series of 30 chemically diverse analogs. Modern synthetic routes are then used to create fragments of nTZDpa to be screened *in vitro* to pinpoint the functional groups responsible for biased agonism.

In the second project, covered in Chapter 3, the crystal structure of NMP422, a novel PPARγ partial agonist, is used to build and analyze molecular dynamics simulations meant to identify protein-ligand interactions that formed a testable hypothesis for a dual binding mode observed in the simulations. NMP422 adopts a conformation in accelerated molecular dynamics simulations that is reminiscent of PPARγ full agonists; however, it shows partial agonist characteristics. PPRE transactivation assays were used to determine the transcriptional activity compared to rosiglitazone. Finally, ^{19}F NMR was used to support the dual binding mode observed in molecular dynamics simulations.

Chapter 2: Identifying the region responsible for corepressor recruitment bias caused by the PPARγ partial agonist nTZDpa

T. M. Patton, Ian M. Chrisman, Michelle D. Nemetchek, Philippe Diaz, and Travis S. Hughes

2.1 Abstract

Peroxisome proliferator-activated receptor gamma (PPARγ) is a drug target for the treatment of type 2 diabetes mellitus due to its role in regulating insulin sensitivity. The current FDA approved drugs Avandia (rosiglitazone) and Actos (pioglitazone) are effective at increasing insulin sensitivity but cause unwanted adverse effects attributed to their full agonist characteristics. Drug discovery efforts have attempted to reduce or prevent these adverse effects by creating ligands that decrease transcriptional activity to no avail. There is a need for new ideas to discover novel and effective PPARγ antidiabetic agents, and one such idea is using biased agonists. Herein, we explore this concept of biased agonism using fragments of the PPARγ partial agonist nTZDpa, a ligand that can selectively recruit the corepressor NCoR.

A structure-activity relationship was determined and the regions of nTZDpa responsible for the bias effect was identified.

2.2 Introduction

PPARγ is a ligand-activated transcription factor targeted for the treatment of type 2 diabetes mellitus (T2DM). Current FDA approved antidiabetic agents, Avandia (rosiglitazone) and Actos (pioglitazone), are full agonists of PPARγ that fully activate transcriptional activity. This full activation is associated with adverse effects such as weight gain and increased risk of bone fractures, sparking drug discovery efforts to prevent them. A variety of PPARγ ligands aiming to partially activate or deactivate the transcriptional output produced by the receptor have been developed and evaluated. Despite these efforts, there have been no new PPARγ targeting antidiabetic drugs to successfully reach FDA approval. A new approach is needed to revitalize PPARγ drug discovery, and one such approach is exploring the concept of biased agonism.

Activation and deactivation of PPARγ involves the binding of coactivators and corepressors, respectively, to the PPARγ:RXRα heterodimer. There are several unique coactivators and corepressors that make up each class of coregulator proteins and each one upregulates or downregulates the transcription of a specific set of genes.[30] One example of this has been observed in PPARγ coactivator 1 alpha (PGC1α), a coactivator that upregulates the transcription of genes that cause white adipose tissue to brown, a process that leads to weight loss in mice. Biased agonism, also called functional selectivity or selective nuclear receptor modulation in the context of nuclear receptors, is a bias effect where a ligand can preferentially activate one signaling pathway over other pathways. In the case of PPARγ, selective

recruitment of coregulators results in the upregulation or downregulation of specific gene sets. nTZDpa is a PPARγ partial agonist that showed promising activity *in vitro* but fell short of useful *in vivo* activity; however, a recently resolved crystal structure implies that nTZDpa may have inverse agonist character due to observed destabilization of H12. Further coregulator recruitment experiments supported the partial inverse agonist characteristics and identified nTZDpa as a biased agonist with selective recruitment of the corepressor NCoR over the corepressor SMRT.

Herein, we look to identify the region of nTZDpa responsible for this bias effect through *in silico* screening and determination of structure-activity relationships between nTZDpa fragments and PPARγ. We wanted to conserve the 5-chloro-2-indole carboxylic acid scaffold to maintain a consistent binding mode, so a series of analogs were designed with a focus on modifying the thiophenyl (R^1), benzyl chloride (R^2), and thio group (R^3) of nTZDpa. These analogs introduced bulkier functional groups, hydrogen bond donor and acceptor groups, and conversion of the thio group to a sulfoxide, sulfone, and sulfonamide (Table 1) at these regions. These modifications were chosen due to the region each group occupies in the nTZDpa crystal structure. Specifically, the thiophenyl group occupies a pocket that is made up of both hydrophobic and hydrophilic residues in the Ω-loop, lower H3, and β-sheet regions. The benzyl chloride group occupies a hydrophobic pocket in the upper H3 and H4/5 region.

Each analog was screened *in silico* to support a consistent binding mode regardless of functional group present. From there, fragments of nTZDpa were synthesized and screened *in vitro* to identify the region responsible for the bias effect. Upon identifying this region,

additional analogs were synthesized to establish a structure-activity relationship for the biased agonism of nTZDpa.

2.3 Results

2.3.1 Identifying nTZDpa as a biased agonist

A series of inverse agonists, antagonists, apo state, and rosiglitazone as a negative control were screened *in vitro* for affinity of the corepressors NCoR and SMRT using fluorescence anisotropy experiments under four conditions: PPARγ ligand binding domain (LBD), full-length protein (FL), full-length with heterodimer partner retinoid X receptor (FL+RXR), and the heterodimer bound to DNA (Figure 3). In each condition, nTZDpa showed a higher affinity for NCoR than SMRT compared to apo PPARγ, identifying the ligand as biased toward NCoR recruitment.

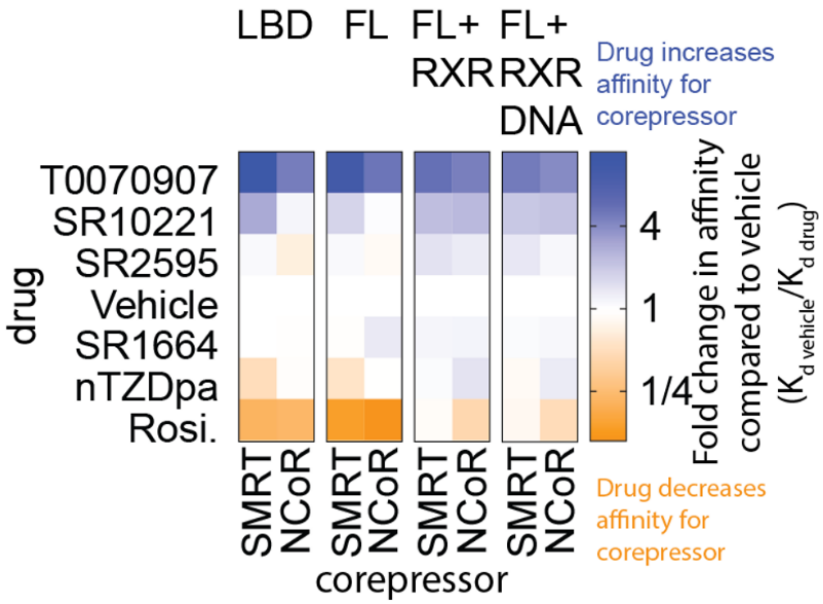


Figure 3. Fluorescence anisotropy experiments measuring the affinity for corepressors NCoR and SMRT. Full agonist rosiglitazone, antagonist SR1664, and inverse agonists T0070907, SR10221, SR2595, and nTZDpa bound to PPARγ.
Abbreviations: ligand binding domain (LBD), full-length (FL), full-length with heterodimer RXR (FL+RXR), and full-length with RXR and DNA (FL+RXR+DNA).

2.3.2 First series analog development

The first series of nTZDpa analogs aimed to preserve the binding mode observed in crystal structures using the 5-chloro-2-indole carboxylic acid as a scaffold for all analogs. Modifications focused on introducing bulkiness and hydrogen bond acceptors and/or donors to three regions of nTZDpa: the thiophenyl (R^1), the benzyl chloride (R^2), and the thio group (R^3). The aim of these modifications was to probe the size of the binding pocket and to probe for new hydrogen bond interactions. Table 1 is a complete list of each analog.

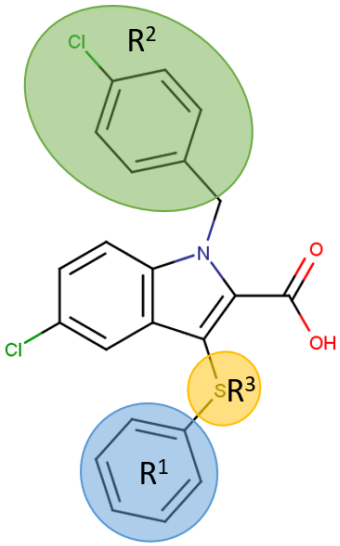


Table 1. First series analogs

Compound	R^1	R^2	R^3
nTZDpa	**Ph**	**Ph-4-Cl**	**S**
1	Ph	Ph-4-N-acetamide	S
2	Ph	Ph-4-propan-2-one	S
3	Ph	Ph-4-ethenone	S
4	Ph	Ph-4-OH	S
5	Ph	naphthyl	S
6	Ph	phenanthrenyl	S
7	Ph	cyclohexyl	S
8	Ph	cyclopentyl	S
9	Ph	cyclobutyl	S
10	Ph	cyclopropyl	S
11	Ph	adamantyl	S
12	naphthyl	Ph-4-Cl	S
13	phenanthrenyl	Ph-4-Cl	S
14	biphenyl	Ph-4-Cl	S
15	terphenyl	Ph-4-Cl	S
16	Ph-4-OH	Ph-4-Cl	S
17	Ph-4-aniline	Ph-4-Cl	S
18	Ph-4-SH	Ph-4-Cl	S
19	Ph-4-$NHCH_3$	Ph-4-Cl	S
20	Ph-4-CH_2NH_2	Ph-4-Cl	S
21	Ph-4-CH_2OH	Ph-4-Cl	S
22	Ph-4-CH_2SH	Ph-4-Cl	S
23		Ph-4-Cl	pentyl
24	Ph	Ph-4-Cl	S(O)
25	Ph	Ph-4-Cl	$S(O)_2$
26	NH_2	Ph-4-Cl	$S(O)_2$
27	NH-Ph	Ph-4-Cl	$S(O)_2$
28	pyridinium	Ph-4-Cl	$S(O)_2$
29	pyrrolidinyl	Ph-4-Cl	$S(O)_2$
30	Ph	Ph-4-acetamide	S

Abbreviations: Phenyl (Ph), Cl (chloro), alcohol (OH), thiol (SH), amine (NH), sulfoxide (S(O)), sulfone ($S(O)_2$).

2.3.3 Docking

In crystal structures, nTZDpa has significant interactions with R288 on H3 and the backbone of S342 in the β-sheet region. Each analog was built *in silico* and docked into the nTZDpa-bound crystal structure of PPARγ (PDB: 2Q5S) using nTZDpa as a template. The docked final structure was chosen based on the relative location of the carboxylic acid and the indole nitrogen compared to nTZDpa in its crystal structure (Figure 4).

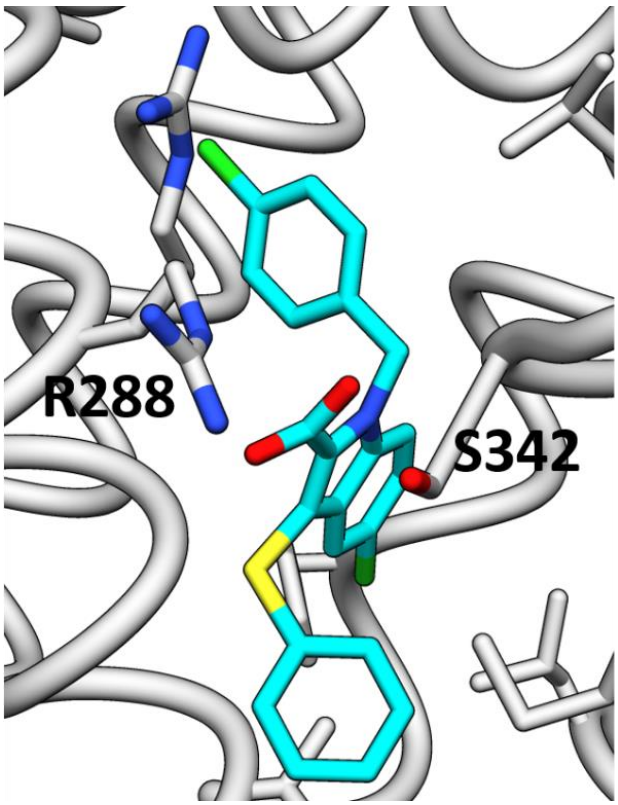


Figure 4. Binding mode of nTZDpa. The 5-chloro-2-indole carboxylic acid scaffold is believed to direct the binding mode to this region of PPARγ through interactions with R288 of H3 and the backbone of S342 of the β-sheet. The R^1 group of nTZDpa occupies the lower portion of helix 3 around the Ω-loop. The R^2 group occupies the upper portion of H3 around H5. PDB: 2Q5S

2.3.4 In silico screening

Each analog and nTZDpa were prepared for molecular dynamics simulations (MD) using the AMBER force field where 500 ns of conventional MD were used to calculate boost parameters and 1.5 μs of accelerated MD were completed. RMSD free energy maps for the accelerated MD were generated and used to select frames corresponding to the lowest energy well (**Figures S1-S4**). Each simulation was analyzed, and the frequency of

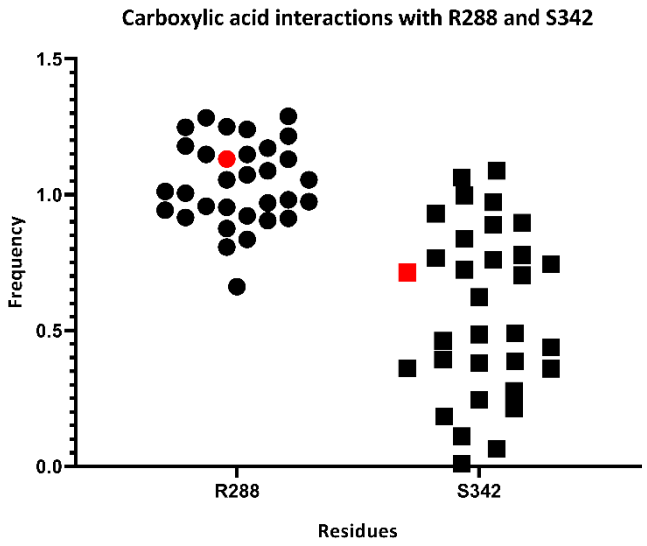
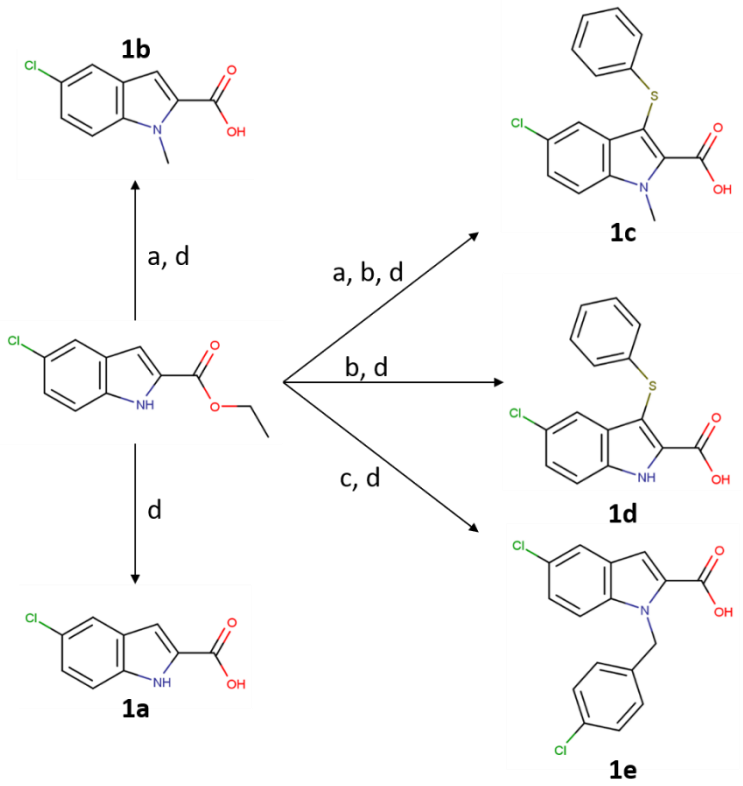


Figure 5. Carboxylic acid interactions with R288 and S342 for each analog. The red data labels are the frequency of interactions between nTZDpa and R288 and S342.

hydrogen bond interactions are shown in Figure 5. Overall, each analog retained interactions with either R288, S342, or both residues simultaneously at high frequency providing support that the desired binding mode is maintained. Interactions between the analogs and R288 are more consistent than with S342 (Figure 5).

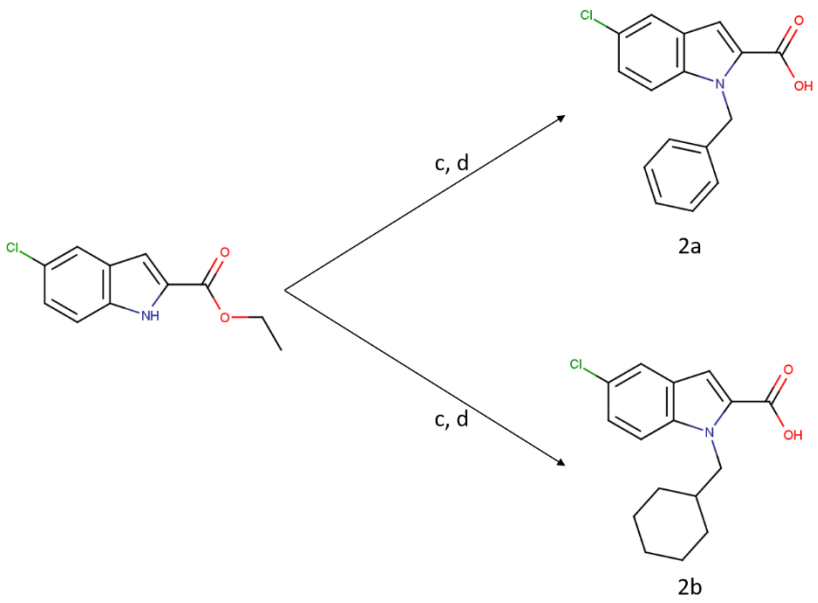
2.3.5 Synthesis and *in* vitro screening of nTZDpa fragments and 1e analogs

The synthetic route for each fragment (**1a-1e**) is shown in Scheme 1 and analogs of fragment **1e** (**2a** and **2b**) are shown in Scheme 2. Commercially available ethyl 5-chloro-2-indole carboxylate was N-methylated, sulfenylated, coupled with 4-chlorobenzyl chloride, benzyl bromide, or (bromomethyl)cyclohexane, and deprotected to yield each compound. The compounds were then screened *in vitro* for PPARγ binding affinity and coregulator binding affinities using fluormone competitive binding assays and fluorescence anisotropy assays, respectively (Table 2 and Figure 6).



Scheme 1. Synthesis of nTZDpa fragments. a. Iodomethane, K_2CO_3, acetone, 2 hr at rt. b. Thiophenol, N-chlorosuccinimide, CH_2Cl_2, MeCN, -78°C to 0°C, 1hr at 0°C. c. 4-chlorobenzyl chloride, NaH, TBAI, DMF, 0°C to 16 hr at rt. d. NaOH, 1:1 THF:EtOH, overnight at rt.

The fluormone competitive binding assays showed that fragments **1a** and **1b**, consisting mainly of the 5-chloro-2-indole carboxylic acid scaffold, resulted in micromolar affinities.[23] Fragments **1c** and **1d**, lacking the benzyl chloride group, had affinities in the hundreds of nanomolar. Finally, fragment **1e** and compounds **2a** and **2b**, consisting of only benzyl chloride, benzyl, and cyclohexylmethyl groups, respectively, had the highest affinities for PPARγ compared to the other fragments. The results of these assays indicate several important features. Either R^1 or R^2 is required for binding, but a hydrophobic group at R^2 has more impact on binding affinity than a hydrophobic group at R^1. Despite the impact inclusion of a hydrophobic group at R^2 has on binding, none of the fragments had the same affinity as nTZDpa, so both R^1 and R^2 make contributions to the strong binding affinity of nTZDpa.



Scheme 2. Synthesis of 1e analogs. c. Benzyl bromide or (Bromomethyl)cyclohexane, NaH, TBAI, DMF, 30 min. at 0°C to 16 hr at rt. d. NaOH, 1:1 THF:EtOH, overnight at rt.

The fluorescence anisotropy coregulator recruitment assays produced a variety of results (Figure 6). Because fragments **1a** and **1b** had high micromolar affinities, we assumed the fragments had no appreciable effect on the binding affinities of SMRT, NCoR, or MED1. Interestingly, fragment **1c** had a higher affinity for the coactivator MED1 indicating an increased agonist character than nTZDpa. Fragment **1d** had similar affinities for each peptide, each being less than the vehicle. Fragment **1e** had a slightly higher affinity for SMRT than nTZDpa, but both ligands caused SMRT and MED1 to dissociate from the PPARγ indicated by a lower affinity for both peptides. Most importantly, fragment **1e** had a similar affinity for NCoR as nTZDpa. This indicates that the R^2 region, consisting of a benzyl chloride moiety, is the region of nTZDpa responsible for increased NCoR affinity.

Both **2a** and **2b** had agonist character due to their increased affinity for MED1 and reduced affinity for NCoR. This gives us some insight into the structure-activity relationship. The

Table 2. K_i values of nTZDpa fragments and 1e analogs

Compound	Ki (nM)
Rosiglitazone	**5.33**
nTZDpa	**0.0083**
1a	41399
1b	13066
1c	539
1d	299
1e	79
2a	141
2b	54

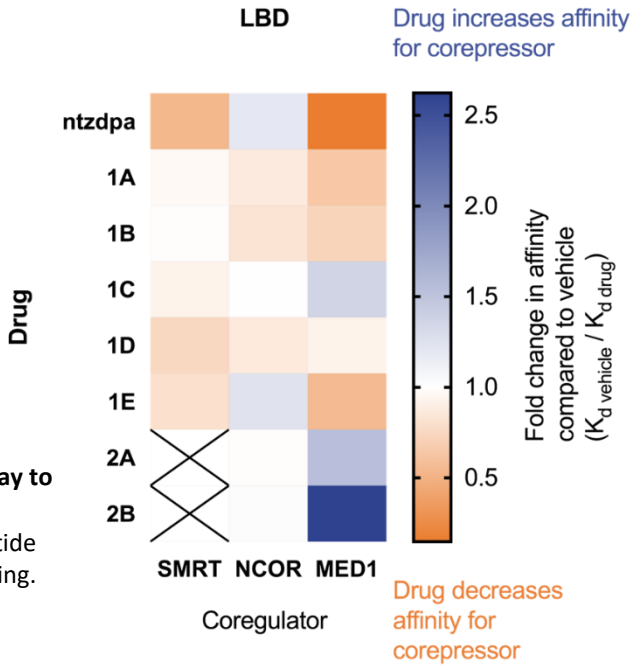


Figure 6. Fluorescence anisotropy assay to screen for biased agonism in nTZDpa fragments and 1e analogs. SMRT peptide was unavailable at the time of publishing.

agonist character suggests that compounds **2a** and **2b** may bind more similarly to fragment **1c** than fragment **1e**. Further experiments will need to be completed to validate the binding modes of fragments **1c** and **1e** and compounds **2a** and **2b**. Regardless, based on these data, we have support that the chloro group of the benzyl chloride is necessary for the binding mode that leads to the bias effect.

2.4 Methods

2.4.1 Protein purification

A pET-46 plasmid carrying the genes for ampicillin resistance and N terminally 6xHis-tagged PPARγ containing a tobacco etch virus (TEV) nuclear inclusion protease recognition site between the His tag and protein of interest was transformed into chemically competent E. coli BL21(DE3) Gold cells (Invitrogen). Cells were grown in either ZYP-5052 autoinduction media or terrific broth (TB). Cells grown in TB at 37 °C were induced at an OD600 of approximately 0.8 by the addition of 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and the temperature lowered to 22 °C. Induction proceeded for 16 h prior to harvesting. Harvested cells were

homogenized into 50 mM phosphate (pH 8.0), 300 mM KCl, 1 mM tris(2-carboxyethyl)phosphine (TCEP), and lysed using a C-5 Emulsiflex high-pressure homogenizer (Avestin). Lysates were then clarified and passed through two Histrap FF 5 ml columns in series (GE Healthcare). Protein was eluted using a gradient from 15 to 500 µM imidazole. Fast protein liquid chromatography was performed on either an NGC Scout system (Bio-Rad) or an ÄKTA Start (GE Healthcare). Eight milligrams of recombinant 6xHis-tagged TEV was added to eluted protein followed by dialysis into 50 mM Tris (pH 8.0), 200 mM NaCl, 1 mM TCEP, and 4 mM EDTA. The protein was again passed through HisTrap FF columns in order to separate cleaved protein from TEV as well as the cleaved 6xHis tag. The cleavage step was only performed on protein which would be used for NMR or FP, but the protein used for Fluormone competitive binding assay did not have the 6xHis tag removed. The protein was then further purified by gel filtration using a HiLoad 16/600 Superdex 200 PG (GE Healthcare). Size exclusion was performed in 25 mM MOPS (pH 8.0), 300 mM KCl, 1 mM TCEP, and 1 mM EDTA buffer. Protein was then dialyzed into 25 mM 3-(N-morpholino)propanesulfonic acid (MOPS) (pH 7.4), 25 mM KCl, and 1 mM EDTA buffer. Protein purity in excess of 95% was determined by gradient 4–20% sodium dodecyl sulfate-polyacrylamide gel electrophoresis analysis (NuSep). Protein concentration was determined using ε280 = 12,045 M^{-1} cm^{-1}.

2.4.2 Delipidation of PPARγ

To delipidate PPARγ LBD, purified protein was diluted to 0.8 mg ml−1 and batched with Lipidex 1000 (Perkin-Elmer) at an equal volume. This mixture was batched for 1 h at 37 °C and 100 rpm. Immediately following this treatment, protein was pulled through a gravity column by syringe. To increase yield, it was found that the speed of elution was important; protein could

not remain on the resin at room temperature in excess of 3 min. Two more column volumes of pre-warmed 25 mM MOPS, 25 mM KCl, and 1 mM EDTA were also pulled through in the same manner. Quality of delipidation was then estimated by 19F NMR, and loss of lipid can be most easily detected by a reduction in the peak at −84.1 ppm. All protein in this work has been delipidated.

2.4.3 Fluormone competitive binding assay

PPARγ ligand inhibition constants (K_i) were measured using a protocol adapted from LanthaScreen TR-FRET PPARγ competitive binding assay (Invitrogen, catalog number PV4894). Assay was performed by plating a mixture of 8 nM 6xHis-PPARγ-LBD, 2.5 nM LanthaScreen Elite Tb-anti-His antibody, 5 nM LanthaScreen Fluormone Pan-PPAR Green (Invitrogen, catalog number PV4896), and 12-point serial dilutions of PPARγ ligands from 50 µM to 140 fM. This mixture was added to wells of low-volume 384-well black plates (Grenier Bio-one) to a final volume of 16 µL. All dilutions were made in 25 mM MOPS (pH 7.4), 25 mM KCl, 1 mM EDTA, 0.01% fatty-acid free BSA (EMD Millipore), 0.01% Tween, and 5 mM TCEP. Assay titrations were performed in duplicate. Plates were incubated in the dark for 2 h at room temperature before being read on a Synergy H1 microplate reader (BioTek). TR-FRET was measured by excitation at 330 nm/80 nm and emission at 495 nm/10 nm for terbium and 520 nm/25 nm for Fluormone. Change in TR-FRET was calculated by 520 nm/495 nm ratio. Nonlinear curve fitting was performed using Prism 7.0b (Graphpad Software Inc.) as described above for the TR-FRET data, including manual exclusion of highest three concentrations for nTZDpa due to instrinsic fluorescence of the ligand.

Fluormone competitive binding assays were completed using two technical replicates

and the experiment was repeated independently once. The number chosen for technical replicates and independent experiment replicates were based on previous experience of limited variability inherent in these biochemical assays.

2.4.4 Ki calculation

The inhibition constant for each PPARγ ligand was calculated by applying a corrected Cheng-Prusoff:[31]

$$\mathrm{Ki} = (\mathrm{Lb})(\mathrm{IC50})(\mathrm{Kd})(\mathrm{Lo})(\mathrm{Ro}) + \mathrm{Lb}(\mathrm{Ro} - \mathrm{Lo} + \mathrm{Lb} - \mathrm{Kd})$$

where IC50 is the concentration of the ligand that produces 50% displacement of the Fluormone tracer, Lo is the concentration of Fluormone in the assay (5 nM), and Kd is the binding constant of Fluormone to wt or the two BTFA-labeled mutants, Ro is the total receptor concentration, and Lb is the concentration of bound Fluormone in the assay with no addition of test ligand. The affinity of Fluormone for the two BTFA-labeled mutant proteins was determined via TR-FRET by titration of Fluormone into each mutant bound to Elite Tb-anti-His antibody. Dissociation constant of Fluormone for wt was measured as 7.9 ± 0.2 for PPARγ LBD.

2.4.5 Fluorescence anisotropy peptide binding assays

Fluorescence anisotropy peptide binding assays were performed by plating a mixture of 50 nM peptide with an N-terminal FITC tag, 12-point serial dilutions of PPARγ-LBD wt, and PPARγ ligands from 50 µM to 24 nM. PPARγ-LBD and PPARγ ligands were added at a 1:1 ratio. This mixture was added to wells of low-volume 384-well black plates (Grenier Bio-one, catalog number 784076) to a final volume of 16 µL. Peptides were synthesized by Lifetein LLC (Somerset, NJ, USA) for the for MED1 peptide, sequence: NTKNHPMLMNLLKDNPAQD; and the

NCoR peptide, sequence: GHSFADPASNLGLEDIIRKALMG (2251–2273). Other peptides were purchased from ThermoFisher (Waltham, MA, USA) for MED1 peptide, sequence: NTKNHPMLMNLLKDNPAQD (catalog number PV4549) and SMRT, sequence: HASTNMGLEAIIRKALMGKYDQW (catalog number PV4424). All dilutions were made in 25 mM MOPS (pH 7.4), 25 mM KCl, 1 mM EDTA, 0.01% fatty-acid-free bovine serum albumin (BSA) (EMD Millipore, catalog number 126575), 0.01% Tween, and 5 mM TCEP. Assay titrations were performed in duplicate. Plates were incubated in the dark at room temperature for 2 h before being read on a Synergy H1 microplate reader (BioTek). Anisotropy was measured by excitation at 485 nm/20 nm and emission at 528 nm/20 nm for FITC. Data were fit using nonlinear regression: agonist vs. response – variable slope 4 parameters, in Prism 7.0b.

Anisotropy assays were completed using two technical replicates and the experiment was repeated independently once. The number chosen for technical replicates and independent experiment replicates were based on previous experience of limited variability inherent in these biochemical assays.

2.4.6 *In silico* screening of nTZDpa analogs

Each analog was built computationally using Discovery Studio and saved as a mol2 file. The crystal structure of nTZDpa-bound PPARγ was retrieved from the protein databank (PDB: 2Q5S). Unresolved residues from the crystal structure were modeled into the PDB file using the Modeller extension within UCSF Chimera.[32,33] This refined structure was then used in the GOLD molecular docking software to dock each analog into the PDB file using nTZDpa as a template. The resulting docked analog files were compared to the nTZDpa crystal structure and the most nTZDpa-like binding mode, based on carboxylic acid and indole nitrogen positioning, was

chosen. The analog-docked structure file was formatted to use the Amber 16 molecular modeling package and AMBER ff14SB force field.[34]

Conventional molecular dynamics simulations were run at a 500 ns timescale to calculate boost parameters to be used by accelerated molecular dynamics simulations for each analog and nTZDpa. All ligands were subjected to the same preparation steps, simulation conditions, and data analysis unless stated otherwise. The formatted ligand-docked structure files were submitted to the H++ server to determine the protonation states of titratable residues at pH 7.4 (http://biophysics.cs.vt.edu/H++).[35] The resulting pdb file was modified using pdb4amber in AmberTools14 for use with tleap.[36] Each analog was extracted and submitted to RED Server for RESP charge derivation and geometry optimization.[37,38] Structures were immersed in a TIP3P octahedron water box and water molecules were extended to 10 Å from protein atoms.[39] Na+ atoms were added until the system was neutralized and K+ and Cl- ions were added to concentrate the system to 50 mM.[40] The resulting system was equilibrated using nine-steps of minimization and restrained simulations protocol. In the first step, 2000 steps of a 5 kcal mol^{-1} $Å^{-2}$ force constant was applied on the protein heavy atoms. Then, an MD simulation was performed for 15 ps with shake under constant volume periodic boundary conditions (NVT). This was followed by two rounds of steepest descent minimization with a 2 and 0.1 kcal mol^{-1} $Å^{-2}$ spring constant for 2000 steps each round. The system underwent another simulation, this time with no restraints followed by three rounds of simulations using 1, 0.5, and 0.5 kcal mol^{-1} $Å^{-2}$ force constants on heavy atoms for 5 ps, 10 ps, and 10 ps, respectively. Finally, a simulation without restraints was performed for 200 ps under NPT conditions. Hydrogen mass repartitioning and SHAKE algorithm were used to allow an integration time step of 4 fs.

Production runs of constant pressure replicates were performed from randomized initial velocities. The pressure was controlled by a Monte Carlo barostat with a pressure relaxation time (taup) of 2 ps. A temperature of 310 K was kept constant using Langevin dynamics with a collision frequency (gamma_ln) of 3 ps^{-1}. The Ewald41 particle mesh with an 8.0 Å cutoff was used to treat electrostatic interactions. Accelerated molecular dynamics simulations were completed at 1.5 µs timescales using the with a timestep of 3 fs. Production runs used a dual boosting approach in which two separate boost potentials are applied to the torsional and total potential terms. Boost parameters were calculated using the average dihedral and total potential energies obtained from 500 ns of conventional molecular dynamics runs. Accelerated molecular dynamics production runs continued from the final frame of the conventional molecular dynamics runs.

All production simulations were performed using pmemd.cuda or pmemd.cuda.MPI. Hydrogen bond analysis were completed using CPPTRAJ in the AmberTools14 Toolbox.[41] A toolkit of Python scripts of "PyReweighting" was used to reweight the biased accelerated molecular dynamics frames and to calculate free-energy profiles (Figures S1, S2, S3, S4, S5).[42]

2.4.7 Synthesis of nTZDpa fragments and 1e analogs

All starting materials and reagents were available commercially unless stated otherwise. Acetone was dried using nitrogen and magnesium sulfate ($MgSO_4$) prior to use. Melting points were determined using a Stuart Scientific melting point apparatus SMP3. Proton and ^{13}C spectrums were determined using either a Bruker 400 MHz or Bruker 700 MHz NMRs. High resolution liquid chromatography mass spectrometry was completed using using an Agilent 6520 Accurate Mass Q-TOF LC/MS with Agilent 1200 binary LC.

General procedure a: N-methylation of indole.[43] The ethyl ester indole (1eq) was dissolved in acetone (60 mM) and K_2CO_3 (4 eq) and iodomethane (2 eq) were added to the reaction vessel. Stir this mixture at room temperature overnight. The following day white solids were visible in the reaction vessel. The reaction was poured into water and extracted three times with ethyl acetate. The combined organic layers were washed with brine, dried over $MgSO_4$, filtered, and concentrated in vacuo.

General procedure b: sulfenylation of indole.[43] Thiophenol (1.2 eq) was added to a solution of N-chlorosuccinimide (1.2 eq) in CH_2Cl_2 (0.2 M) at -78°C. The mixture was warmed to 0°C, where the reaction changed from clear to bright yellow. After 15 minutes at this temperature, the ethyl ester indole (1 eq) was dissolved in a 1:1 solution of CH_2Cl_2 and MeCN (0.2 M) and was added to the reaction vessel and stirred at 0°C for one hour. The reaction was quenched with water and extracted three times with CH_2Cl_2. The combined organic layers were washed with brine, dried over $MgSO_4$, filtered, and concentrated in vacuo.

General procedure c: alkylation of indole.[43] The ethyl ester indole (1 eq) was dissolved in DMF (0.1 M) and was added to a suspension of sodium hydride (60 % in mineral oil, 1.2 eq) in DMF (0.1 M) at 0°C. The reaction was warmed to room temperature and stirred for 30 minutes. 4-chlorobenzyl chloride or benzyl bromide or (bromomethyl)cyclohexane (1.5 eq) and TBAI (1 eq) were added, and the reaction was stirred at room temperature overnight. The reaction was quenched with water and extracted three times with ethyl acetate. The combined organic layers were washed with brine, dried over $MgSO_4$, filtered, and concentrated in vacuo.

General procedure d: ethyl ester deprotection.[43] The ethyl ester indole (1 eq) was dissolved in 1:1 THF:EtOH (0.1 M), and 1M NaOH (5 eq) was added to the reaction vessel. The reaction was stirred at room temperature overnight. Upon completion, the mixture was transferred to a round bottom flask and concentrated in vacuo to remove the reaction solvent, acidified with 1M HCl, and extracted three times with ethyl acetate. The combined organic layers were washed with brine, dried over $MgSO_4$, filtered, and concentrated in vacuo.

5-chloro-1H-indole-2 carboxylic acid (1a). Following general procedure d, the ethyl ester indole (80 mg, 0.358 mmol) was purified by Flash Chromatography (DCM:MeOH 0% to 15% gradient) yielding the title compound as a brown solid (64.1 mg, 91.6% yield). Decomp., 230°C. **H NMR** (400 MHz, MeOD) δ 7.62 (s, 1H), 7.41 (d, 1H), 7.21 (d, 1H), 7.09 (s, 1H); **C NMR** (700 MHz, DMSO) δ 163.05, 135.99, 130.78, 128.37, 124.89, 124.77, 121.43, 114.54, 107.13; **HRMS** ESI (m/z): $[M-H]^-$ calcd for $C_9H_5ClNO_2$ 194.001, found 193.978.

5**-chloro-1-methyl-indole-2 carboxylic acid (1b).** Following general procedure a, the ethyl ester indole (180 mg, 0.814 mmol) was purified by Flash Chromatography (Heptane:Ethyl acetate 0% to 30% gradient) yielding the N-methylated intermediate compound as a white solid (157.3 mg, 73% yield). **H NMR** (400 MHz, $CDCl_3$) δ 7.63 (s, 1H), 7.30 (s, 2H), 7.22 (s, 1H), 4.38 (q, 2H), 4.06 (s, 3H), 1.41 (t, 3H).

The ethyl ester indole intermediate (51.6 mg, 0.217 mmol) underwent general procedure d, and was purified by Flash Chromatography (DCM:MeOH 0% to 15% gradient) yielding the title compound as a white solid (41.7mg, 92% yield). Decomp., 230°C. **H NMR** (400 MHz, MeOD) δ 7.62 (s, 1H), 7.45 (d, 1H), 7.27 (d, 1H), 7.20 (s, 1H), 4.05 (s, 3H); **C NMR** (700 MHz, DMSO) δ

163.30, 137.99, 131.32, 126.78, 125.13, 124.74, 121.45, 113.06, 108.83, 32.17; **HRMS** ESI (m/z): $[M-H]^-$ calcd for $C_{10}H_7ClNO_2$ 208.017, found 207.972.

5-chloro-1-methyl-3-(phenylthiol)-indole-2 carboxylic acid (1c). Following general procedure b, the ethyl ester indole (100 mg, 0.447 mmol) was purified by Flash Chromatography (Heptane:Ethyl acetate 0% to 30% gradient) yielding the sulfenylated intermediate compound as a white solid (77.4 mg, 52% yield). **H NMR** (400 MHz, $CDCl_3$) δ 9.22 (s, 1H), 7.60 (s, 1H), 7.38 (d, 1H), 7.31 (d, 1H), 7.22-7.08 (m, 5H), 4.38 (q, 2H), 1.30 (t, 3H).

The sulfenylated intermediate (42 mg, 0.127 mmol) underwent general procedure a, and was purified by Flash Chromatography (Heptane:Ethyl acetate 0% to 20% gradient) yielding the N-methylated intermediate compound as a white solid (38.5 mg, 87% yield). **H NMR** (400 MHz, $CDCl_3$) δ 7.63 (s, 1H), 7.37-7.29 (br, 2H), 7.20-7.14 (br, 2H), 7.12-7.04 (br, 3H) 4.34 (q, 2H), 4.04 (s, 3H), 1.25 (t, 3H).

Finally, the N-methylated intermediate (35.2 mg, 0.102 mmol) underwent general procedure d, and was not purified further (29.1 mg, 90% yield). Melting point, 190°C. **H NMR** (400 MHz, MeOD) δ 7.42 (d, 1H), 7.30 (s, 1H), 7.20 (d, 1s), 7.10-7.03 (m, 2H), 6.98 (dd, 4H) 3.94 (s, 3H); **C NMR** (700 MHz, DMSO) δ 162.53, 137.96, 136.87, 135.43 129.50, 129.31, 126.84, 126.49, 125.86, 125.57, 119.42, 114.05, 105.85, 40.36, 33.12, ; **HRMS** ESI (m/z): $[M-H]^-$ calcd for $C_{16}H_{11}ClNO_2S$ 316.020, found 315.965.

5-chloro-3-(phenylthiol)-1H-indole-2 carboxylic acid (1d).

The sulfenylated intermediate (36.2 mg, 0.109 mmol) underwent general procedure d, and was purified by Flash Chromatography (DCM:MeOH 0% to 15% gradient) yielding the title

compound as a white solid (17.1 mg, 51% yield). Decomp., 200°C. **H NMR** (400 MHz, MeOD) δ 7.47 (d, 1H), 7.38 (s, 1H), 7.25 (d, 1H), 7.21-7.15 (m, 2H), 7.12-7.05 (m, 3H); **C NMR** (700 MHz, DMSO) δ 162.27, 138.12, 134.84, 130.75, 129.46, 126.82, 125.96, 125.73, 125.50, 119.36, 115.41, 105.62, 40.71, 29.39; **HRMS** ESI (m/z): $[M-H]^-$ calcd for $C_{15}H_9ClNO_2S$ 302.005, found 301.936.

5-chloro-1-(4-chlorobenzyl)-indole-2 carboxylic acid (1e). Following general procedure c, the ethyl ester indole (100 mg, 0.447 mmol) was purified by Flash Chromatography (Heptane:Ethyl acetate 0% to 30% gradient) yielding the alkylated intermediate compound as a white solid (126.2 mg, 81% yield). **H NMR** (400 MHz, $CDCl_3$) δ 7.70 (s, 1H), 7.34 (s, 1H), 7.31-7.22 (m, 4H), 6.97 (d, 2H), 5.80 (s, 2H), 4.36 (q, 2H), 1.39 (t, 3H).

The alkylated intermediate (40 mg, 0.115 mmol) underwent general procedure d, and was purified by Flash Chromatography (DCM:MeOH 0% to 15% gradient) yielding the title compound as a white solid (36.5mg, 99% yield). Decomp., 217°C. **H NMR** (400 MHz, $CDCl_3$) δ 7.61 (s, 1H), 7.34 (s, 1H), 7.19 (s, 1H), 7.13 (d, 3H), 6.86 (d, 2H), 5.68 (s, 2H); **C NMR** (700 MHz, DMSO) δ 163.09, 137.85, 137.63, 132.14, 130.32, 128.95, 128.57, 127.10, 125.70, 125.40, 121.87, 113.39, 110.28, 46.97, 40.32; **HRMS** ESI (m/z): $[M-H]^-$ calcd for $C_{16}H_{10}Cl_2NO_2$ 318.009, found 317.963.

5-chloro-1-(benzyl)-indole-2 carboxylic acid (2a). Following general procedure c, the ethyl ester indole (111.83 mg, 0.5 mmol) was purified by Flash Chromatography (Heptane:Ethyl acetate 0% to 30% gradient) yielding the alkylated intermediate compound as a yellow solid (139.9 mg,

89.2% yield). **H NMR** (400 MHz, $CDCl_3$) δ 7.64 (s, 1H), 7.29 (s, 1H), 7.25-7.14 (br, 5H), 6.99 (d, 2H), 5.78 (s, 2H), 4.31 (q, 2H), 1.33 (t, 3H).

The alkylated intermediate (139.9 mg, 0.446 mmol) underwent general procedure d, and was purified by Flash Chromatography (DCM:MeOH 0% to 10% gradient) yielding the title compound as a light brown solid (57.6mg, 45.3% yield). Decomp., 170°C. **H NMR** (400 MHz, $CDCl_3$) δ 7.67 (s, 1H), 7.39 (s, 1H), 7.20 (dd, 5H), 6.90 (s, 2H), 5.49 (s, 2H); **C NMR** (400 MHz, MeOD) δ 163.38, 138.32, 137.73, 129.64, 128.12, 127.08, 126.77, 125.97, 124.83, 121.12, 112.09, 109.81, 47.28; **HRMS** ESI (m/z): $[M-H]^-$ calcd for $C_{16}H_{11}ClNO_2$ 284.048, found 283.982.

5-chloro-1-(cyclohexylmethyl)-indole-2 carboxylic acid (2b). Following general procedure c, the ethyl ester indole (111.83 mg, 0.5 mmol) was purified by Flash Chromatography (Heptane:Ethyl acetate 5% to 25% gradient) yielding the alkylated intermediate compound as a yellow solid (87.0 mg, 54.4% yield). **H NMR** (400 MHz, $CDCl_3$) δ 7.62 (s, 1H), 7.42-7.27 (br, 2H), 7.23 (s, 1H), 4.41-4.33 (m, 4H), 1.89-1.77 (br, 1H), 1.71-1.59 (br, 3H), 1.51 (d, 2H), 1.40 (t, 3H), 1.19-0.97 (br, 5H).

The alkylated intermediate (87.0 mg, 0.272 mmol) underwent general procedure d, and was purified by Flash Chromatography (DCM:MeOH 0% to 10% gradient) yielding the title compound as a light yellow solid (42.7 mg, 53.7% yield). Decomp., 177°C. **H NMR** (400 MHz, $CDCl_3$) δ 7.67 (s, 1H), 7.39 (s, 1H), 7.37-7.27 (m, 2H), 4.41 (d, 2H), 1.91-1.79 (br, 1H), 1.73-1.62 (br, 3H), 1.52 (d, 2H), 1.19-1.02 (br, 5H); **C NMR** (400 MHz, MeOD) δ 163.43, 137.96, 129.53, 126.66, 125.56, 124.46, 120.95, 112.24, 109.50, 50.02, 39.41, 30.44, 29.38, 29.35, 26.06, 25.57; **HRMS** ESI (m/z): $[M-H]^-$ calcd for $C_{16}H_{17}ClNO_2$ 290.095, found 290.027.

2.5 Discussion

The data resulting from the fluorescence anisotropy coregulator affinity experiments indicate several important characteristics that can be applied to future work, specifically, the impact the R^1 and R^2 regions have on coregulator recruitment. Based on the crystal structure binding mode seen in Figure 4, the R^1 region of nTZDpa occupies Branch II of the ligand binding pocket and the R^2 region of nTZDpa occupies Branch III of the ligand binding pocket. The agonist character of fragment **1c** is assumed to be the consequence of Branch II binding, and this suggests Branch II is involved in MED1 recruitment. In contrast, the biased character of fragment **1e** suggests Branch III is involved in NCoR recruitment.

Analogs of fragment **1e**, compounds **2a** and **2b**, provided support to the necessity of the chloro group in the R^2 region for the bias effect observed with nTZDpa. These compounds both lacked the chloro group and had a higher affinity for MED1 compared to NCoR. This suggests they bind more similarly to fragment **1c**, occupying the Branch II region of the ligand binding pocket and giving them agonist characteristics.

2.6 Conclusions

The first series of analogs screened *in silico* maintained interactions between the carboxylic acid of the scaffold and residues R288 and S342 suggesting that modifications to the thiophenyl, benzyl chloride, and thio group will retain a consistent binding mode. Fluormone binding affinity assays showed that the hydrophobic pocket occupied by the benzyl chloride group has a larger influence on binding affinity than the thiophenyl group; however, both groups are needed to reach picomolar binding affinity like nTZDpa. It is important to note that the focus of this research was not to modify the affinity of nTZDpa, but instead, to modify the

affinity for NCoR. In this case, the data suggests the Branch II region is involved in MED1 recruitment and the Branch III region is involved in NCoR recruitment. Future work will focus on modifying the chloro group of fragment **1e** to other hydrophobic functional groups to further probe the effects Branch III has on NCoR recruitment.

2.7 Future directions

The future of this project should be quite linear. The work completed thus far has created a great starting point to further explore biased agonism in PPARγ. The specific area that needs to be addressed in the immediate future is validation of the binding mode of fragment **1e** through crystallography and molecular dynamics simulations. While the crystal structure is being resolved, new **1e** analogs that vary the position of the chloro group and replace the chloro group with other hydrophobic alkyl chains can be synthesized. The intent of these analogs is to probe the size of the hydrophobic pocket and determine whether this pocket is involved in the bias effect. Examples of chloro group replacements are methyl, isopropyl, and tert-butyl groups. Furthermore, the agonist character of fragment **1c** and compounds **2a** and **2b** should be explored further to determine whether the Branch II region influences H12 independent agonism. The first step in this work would be to validate the binding mode through crystallography and molecular dynamics simulations. From there, planning of analogs of fragment **1c** can begin.

Chapter 3: Helix 3 and β-sheet interactions anchor multiple-binding modes in novel PPARγ partial agonist NMP422

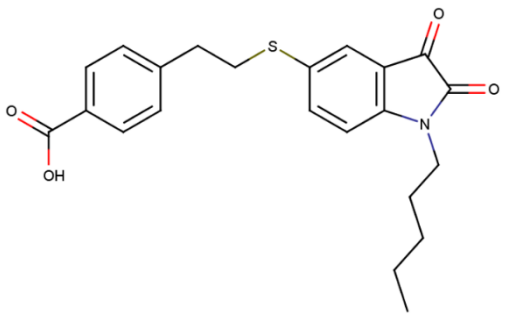
T. M. Patton, Ian M. Chrisman, Mariah L. Rayl, Michelle D. Nemetchek, Tung-Chung Mou, Desiree E. Mendes, Philippe Diaz, and Travis S. Hughes

3.1 Abstract

A class of FDA approved peroxisome proliferator-activated receptor gamma (PPARγ) targeting antidiabetic agents called the glitazones, Avandia (rosiglitazone) and Actos (pioglitazone), are associated with unwanted adverse effects due to their full agonist characteristics. To avoid these adverse effects, drug discovery efforts have focused on avoiding full agonist like binding modes to reduce transcriptional activity through helix 12 independent mechanisms, but there have not been any FDA approved drugs resulting from these efforts. NMP422 is a novel PPARγ partial agonist that adopts two unique binding modes anchored by helix 3 and β-sheet interactions. In molecular dynamics simulations, these interactions allow the major ligand binding mode to be like the glitazones, but flipped, making the ligand independent of helix 12 interactions that would lead to full agonism. Identifying the interactions that lead to this flipped binding mode could lead to new strategies in developing chemically diverse PPARγ ligands with a steady binding mode.

3.2 Introduction

Drug discovery efforts to reduce the unwanted adverse effects of full agonist PPARγ drugs led to the development of partial agonists as described in Section 1.2. The TZDs, and other full agonists, generally consist of four components: an acidic head, an aromatic ring, a linker, and an aromatic or heterocyclic tail that binds to the receptor in a horseshoe shape around the backside of helix 3 (Figure 1b).[44] The acidic head is situated near the junction of helix 4/5, helix 11, and helix 12 where it makes putative interactions described in Section 1.2. The aromatic ring is placed directly behind helix 3 forming hydrophobic interactions with Cys285 and Met364, while the linker and heterocyclic tail are placed into the β-sheet side of helix 3 where the groups interact with Val339, Leu330, and Ile341.[16]



	PPARγ	PPARδ	PPARα
Ki (nM)	578	> 50,000	6712

Figure 7. Structure of NMP422 and selectivity for PPARγ. The selectivity of NMP422 was determined by fluormone competitive binding affinity assays.

NMP422, a novel partial agonist, is an intermediate for a promiscuous PPAR modulator that showed unexpected selectivity and modest potency for PPARγ (Figure 7).[45] Using the resolved crystal structure described in molecular dynamics simulations, and with support from fluorine NMR, we found that NMP422 is able to adopt two ligand binding modes. One mode is TZD-like, forming a horseshoe conformation around helix 3, but the carboxylic acid forms bonds to residues in the helix 3-β-sheet region instead of the putative interactions that stabilize helix 12. The other binding mode maintains the interactions in the helix 3-β-sheet region, but the horseshoe conformation is not formed around helix 3.

3.3 Results

3.3.1 *In vitro* assays

The binding affinity assays were completed using all three isoforms of PPAR in fluormone competitive binding assays. Of those, NMP422 was selective for PPARγ with an affinity of 578 nM (Figure 7).

NMP422 was evaluated for the *in vitro* binding of hPPARγ protein in HEK293T cells to

determine its level of transactivation compared to the transactivation of rosiglitazone. NMP422 transactivation was found to be 53.6% of rosiglitazone transactivation and was found to be significantly different from rosiglitazone and pioglitazone by one-way ANOVA (Figure 8).

3.3.2 Overall crystal structure

The crystal structure of NMP422-bound PPARγ is a homodimer complex consisting of a chain A and chain B. Chain A did not show large variations when compared to the crystal structure of rosiglitazone-bound PPARγ (Figure 9) indicating the protein is in an active conformation where helix 12 is stabilized. Similar to many of the ~200 other PPARγ crystal structures, chain B showed PPARγ in a conformation where helix 12 was in a position parallel to

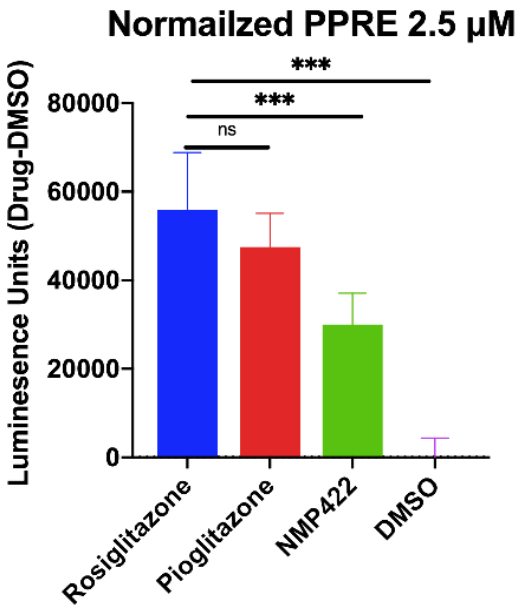


Figure 8. PPRE assays. a. Peptide affinity binding assay is a placeholder b. HEK293T cells were saturated to a ligand concentration of 2.5 µM and incubated for 18 hr.
n.s. not significant. (***) $p < 0.0001$.

	PPAR
Data collection	
Beamline	SSRL 14-1
Wavelength (Å)	1.19
Ligand	NMP-422
Resolution range (Å)*	30.05 - 2.60 (2.64 - 2.60)
Space group	C2
Unit cell dimensions **a, b, c (Å)** **α, β, γ (◦)**	92.9 62.2 120.1 90.0 102.5 90.0
Unique reflections*	20553 (1893)
Redundancy*	5.3 (4.3)
Completeness (%)*	98.5 (91.6)
Mean I/σ (I)*	15.7 (2.2)
Wilson B-factor	50.3
R_{sym} †,*	0.09 (0.65)
Refinement	
R_{work} §,*	0.1771 (0.2920)
R_{free} §,*	0.2510 (0.3995)
Number of total atoms	
protein	4239
NMP-422	59
Solvent	34
Total protein residues	530
RMS deviations **Bonds lengths (Å)** **Bond angles (◦)**	0.009 1.33
Ramachandran favored (%)††	95.2
Ramachandran outlier (%)††	0.6
Average B-factor	
Macromolecules	65.08
PEG	72.26
water	48.84
Number of TLS groups	1

Table 6. Crystallography data collection and refinement statistics. † $R_{sym}=\sum_{hkl}\sum_i |I_i(hkl)-\langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the ith observation of the intensity of the reflection hkl. . § $R_{work}=\sum_{hkl} ||F_{obs}|-|F_{calc}||/\sum_{hkl}|F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure-factor amplitudes for each reflection hkl. R_{free} was calculated with 5% of the diffraction data that were selected randomly and excluded from refinement. * Data for highest resolution shell are given in brackets. ††

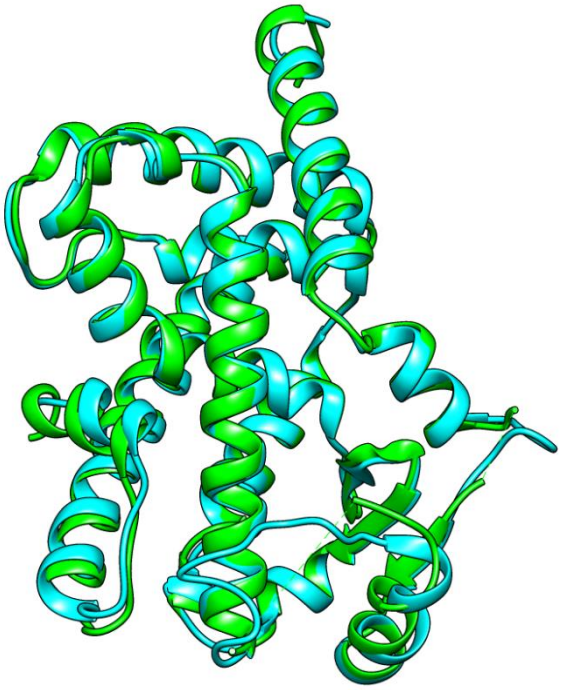


Figure 9. Active conformation of agonist-bound PPARγ. Crystal structure of Chain A of NMP422-bound (cyan) and rosiglitazone-bound (green) PPARγ (PDB: 4XLD).

helix 3 indicating an inactive conformation that coactivators will not be able to bind to. Both

conformations of helix 12 involve unit cell crystal contacts, indicating that the structures may not accurately represent the lowest energy structures in solution.[23] Molecular dynamics (MD) simulations showed the same active conformation with no large variations in conformation. This supports that NMP422 favors the chain A "active" helix 12 conformation in solution.

3.3.3 Ligand binding mode

The ligand binding mode of NMP422 showed consistent hydrogen bond interactions between the carboxylic acid of NMP422 and Arg288 and Ser342 in the helix 3-β-sheet region of the crystal structure, conventional molecular dynamics (cMD) simulations, and accelerated molecular dynamics (aMD) simulations. The rest of the ligand showed variability in ligand binding mode and interactions in all three instances.

The crystal structure showed NMP422 occupying the helix 3-β-sheet region and formed putative interactions with residues Ile281, Gly284, Cys285, Arg288, Ser389, and Ala292 of helix 3, residues Ile326, Tyr327, Met329, Leu330, and Leu333 of helix 5, residues Val339, Ile341, Ser342, Glu343, and Met348 of the β-sheet, residue Leu353 of helix 6, and residues Phe363, Met364, and Lys367 of helix 7 (Figure 10a). NMP422 was in a compact binding mode with the pentyl group participating in interactions with the residues of the upper part of helix 3 and with helix 5. The isatin was positioned to interact with the residues of helix 3 and helix 7, and the thioether linker interacted with the residues of the β-sheet region.

cMD simulations consistently showed NMP422 adopt a pi-stacking conformation that occupied the helix 3 and β-sheet region in representative structures generated by k means

clustering (Figure S18). Overall, the ligand occupied similar regions as the crystal structure. The aMD simulations had more variability in occupied regions indicating two ligand binding modes.

The first ligand binding mode is TZD-like, adopting the horseshoe conformation around helix 3 and interacting with residues on helix 3, helix 4/5, helix 11, and helix 12 (Figure 10b). An important difference between the binding of NMP422 and the TZDs was observed, mainly that the horseshoe of NMP422 is flipped. The TZD's thiazolidinedione group occupies the area around helix 4/5, helix 11, and helix 12 allowing for the putative interactions that stabilize helix 12. NMP422, however, occupies this same region with its pentyl group, and the direct hydrogen bond interactions that stabilize helix 12 do not form. Instead, the carboxylic acid of NMP422 forms hydrogen bond interactions with residues of helix 3 and the β-sheet, stabilizing the region that indirectly stabilizes helix 12. In five aMD simulations, 55% of clustered frames (285,000 of 516,000 total frames) adopted the TZD-like conformation. The second ligand binding mode is completely independent of the helix 12 region, exclusively occupying the helix 3-β-sheet region (Figure 10c).

3.3.4 Fluorine NMR

At the time of publishing this thesis, analysis and replicates of ^{19}F NMR are still being completed and analyzed by members of the Hughes lab. The standalone publication of this project will contain these data. Initial ^{19}F NMR shows NMP422 aligning with PPARγ full agonists rosiglitazone and pioglitazone indicating a similar binding mode, but NMP422 also has an additional minor peak downstream of the major peak (Figure 11). The minor peak indicates a potential second binding mode, consistent with aMD simulations.

3.4 Methods

3.4.1 Protein purification

The methods used in protein purification were the same as those described in section 2.4.1.

3.4.2 Delipidation of PPARγ

The methods used for delipidation of PPARγ were the same as those described in section 2.4.2.

3.4.3 Fluormone competitive binding assay

The methods used in fluormone competitive binding assays were the same as those described in section 2.4.3.

3.4.4 Ki calculation

The methods used in Ki calculations were the same as those described in section 2.4.4.

3.4.5 PPRE transactivation assay

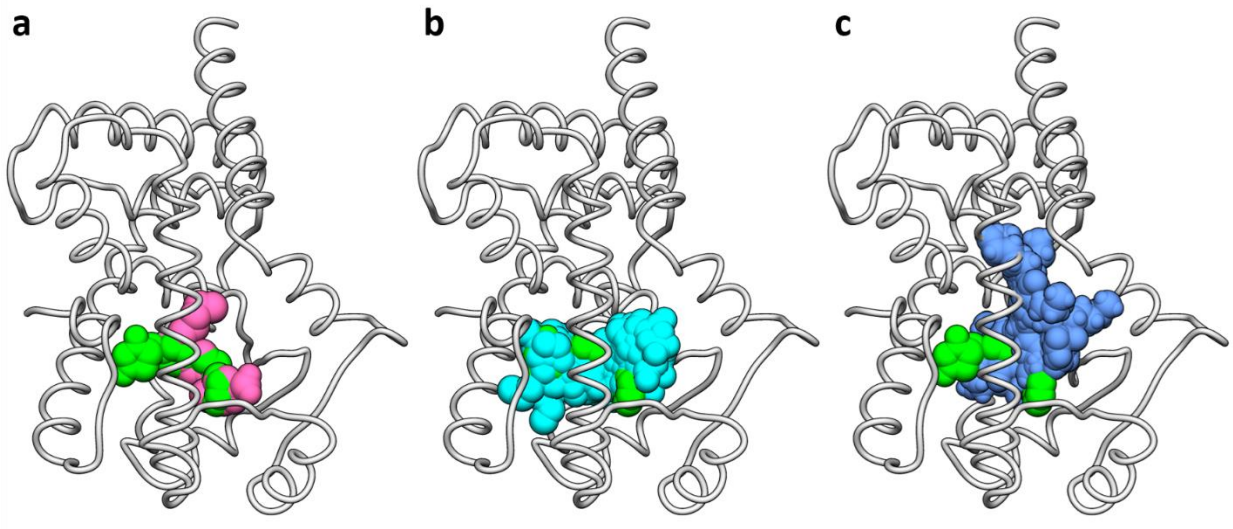


Figure 10. Comparison of NMP422 and rosiglitazone ligand binding modes for crystal structure and aMD simulation clusters. Rosiglitazone (green; PDB: 4XLD) in ligand binding pocket of NMP422-bound PPARγ (white) crystal structure compared to: a. crystal structure (pink), b. TZD-like clusters (cyan), and c. non-TZD-like clusters.

PPARγ expression plasmid and a luciferase reporter plasmid containing PPARγ response element sequences were obtained from the Kojetin lab and used based on protocols from previous studies.[22,46–48] HEK293 cells were thawed and seeded in T75 flasks. The cells underwent two passages at greater than 90% confluency post-thaw and then were plated at 3.82 million cells per well in a T75 flask. The cells were then transfected with a pCMV6 full-length PPARγ expression plasmid and the luciferase reporter plasmid using X-tremegene 9 (Roche) in accordance with manufacturer's protocols. 18 hours after beginning transfection, the medium was harvested, and the cells were removed from the T75 flask by gentle pipetting. The cells were resuspended in the harvested medium and plated into white 384-well plates (Thermo Fisher Scientific) at 10,000 cells per well in 20 µL harvested medium per well. The cells were placed in the incubator for 4 hours and then 20 µL of fresh medium with ligand (5 µM) or vehicle control was added to the cells to achieve a final concentration of 2.5 µM of ligand in total medium volume of 40 µL per well. 18 hours after addition of the ligand, the luciferase

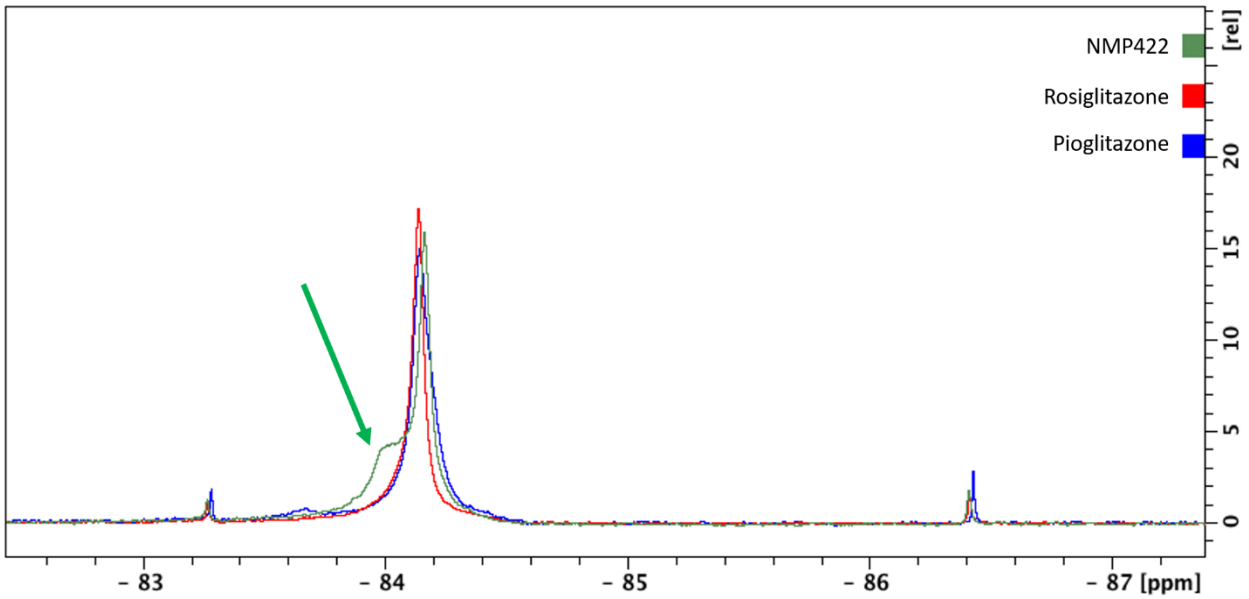


Figure 11. NMR of NMP422, rosiglitazone, and pioglitazone. The green arrow indicates the secondary binding mode in the helix 3-β-sheet region adopted by NMP422. All ligands are at 1.25x the concentration of PPARγ (150 µM). Mutant PPARγ (Q322C, C313A) was used with BFTA tag located at residue 322.

activity was measured using 20 µL of BriteLite Plus (Perkin Elmer) and read on Synergy H1 Hybrid Multi-Mode Reader (BioTek). Statistical analysis was completed using a One-Way ANOVA.

3.4.6 Preparation of NMR samples

BTFA labeling was done during lysis before the purification of protein by adding 10mM BTFA to the protein lysate and purifying as stated previously. NMR samples were prepared to a final concentration of 150 µM protein in 470 µL volume containing 10% D2O. Addition of ligand was done in two separate injections of compound to reduce precipitation. Injections were spaced 30–60 min apart to allow time for binding. All ligands were dissolved in D6-dimethylsulfoxide (DMSO). Deuterated solvents were obtained from Cambridge Isotope Laboratories Inc. and were at least 99% isotopically pure. Final concentrations of ligand were 1.25x ligand to protein (187.5 µM). Following this, buffered D2O was added.

3.4.7 NMR spectroscopy

Acquisition of spectra was performed using a Bruker 700 MHz NMR system equipped with a QCI-F cryoprobe. Chemical shifts were calibrated using an internal separated KF reference in 20 mM KPO4 (pH 7.4) and 50 mM KCl contained in a coaxial tube inserted into the NMR sample tube. KF was set to be −119.522 ppm, which is the shift of the KF signal with respect to the 19F basic transmitter frequency for the instrument (658.8462650 MHz) at 298.2 K, the temperature at which samples were run. Routine 1D fluorine spectra were acquired utilizing the zgfhigqn.2 pulse program (Bruker Topspin 3.5), which consists of a 90° pulse followed by acquisition with proton decoupling (acquisition = 0.7 s). Settings were D1 = 1.2 s, AQ = 0.82 s. Approximately 500 to 4000 transients were collected. The fitting

algorithm32 assumes Lorentzian lineshapes of similar phase. Intermediate exchange effects and field inhomogeneity are likely present in some of these spectra, which will result in inaccuracies in the fitted models; however, notwithstanding these limitations, the deconvolution method provides an objective view of the possible underlying spectral structure and populations. Select NMR spectra were replicated in two different ways: (1) Some NMR samples were measured via NMR initially and then days to weeks later to determine if certain parts of the spectrum changed. Any changes would indicate that non-reversible processes contribute to that part of the signal, such as unfolding or degradation of the protein. (2) Some spectra were run twice utilizing protein from the same batch as utilized for the first spectra or from an entirely different protein preparation.

3.4.8 Crystallography data collection and structure determination

The PPARγ LDB was co-crystallized with a 1.05x molar excess of NMP422. Diffracting crystals grew overnight by hanging drop at 20°C (small rock crystals). Condition yielding the best crystals were 1.0 M NaCit at pH 7.0 with 13 mg mL^{-1} PPARγ LBD. Crystals were harvested and put in a cryoprotectant containing 15 % v/v glycerol. The structure was solved to a resolution of 2.6 Å. Crystal structure data was calculated using MolProbity.[49]

3.4.9 Molecular dynamics simulations

All molecular dynamics simulations were completed using the same methods described in Section 2.4.6. Any variations to these methods are recorded below.

Three independent cMD simulations, production runs a, b, and c, were completed at timescales of 15.26 µs, 15.37 µs, 15.23 µs, respectively, using Chain A of the crystal structure. Five independent aMD simulations, production runs a, b, c, d, and e, were completed using the

dual boosting approach described in Section 2.4.2. Boost parameters for aMDs were calculated from each respective cMD run (e.g. boost parameters for aMD production run a were calculated from cMD production run a, aMD production run b calculated from cMD production b, etc). aMD runs a-c were continued from the end of their respective cMD runs (e.g. aMD run a was continued from cMD run a, etc.). Boost parameters for aMD production runs d and e were calculated from the collective average dihedral and total potential energies from cMD production runs a, b, and c. The production runs for d and e started from the final step of the nine-step minimization, described in Section 2.4.2.

Hydrogen bond analysis and K means clustering were completed using CPPTRAJ and hydrophobic contact analysis was completed using CPPTRAJ in the AmberTools18 Toolbox.[50]

3.5 Discussion

The PPRE transactivation assay determines that NMP422 acts as a partial agonist of PPARγ with lower transcriptional activity compared to the full agonists, rosiglitazone and pioglitazone. The crystal structure and each simulation showed high frequencies of interactions between the carboxylic acid of the ligand and Arg288 and Ser342, anchoring the functional group to the helix 3-β-sheet region. This, in combination with the flexibility of NMP422 and large ligand binding pocket, allows for the dual binding mode observed as well as the partial agonist characteristic of the ligand.

In simulations, we observed flexibility from NMP422 in the ligand binding pocket indicating multiple binding modes, with aMD simulations identifying occupation of regions like the TZDs and unlike the TZDs as major binding modes. Analysis of the cMD simulations showed NMP422 consistently adopting a pi-stacking conformation and occupying the same relative

region around the helix 3-β-sheet region (Figure S18). This binding mode seemed to be further supported by aMD simulations as 45% of frames showed the ligand occupied the same region. Most frames; however, indicate that NMP422 adopts the flipped TZD-like binding mode. The dual binding modes are further supported in fluorine NMR where we observed NMP422 produce a peak that overlaps with rosiglitazone and pioglitazone, but also an additional peak that indicates a second binding mode (Figure 11).

The TZDs and other PPARγ full agonists generally consist of four components: an acidic head, an aromatic ring, a linker, and an aromatic or heterocyclic tail that wrap around helix 3.[44] The acidic head is situated near the junction of helix 4/5, helix 11, and helix 12 where it participates in the interactions that stabilize helix 12 and makes hydrophobic contacts with Phe363, Gln286, Phe282, and Leu469.[15–17] The aromatic ring is placed directly behind helix 3 forming hydrophobic interactions with Cys285 and Met364. The linker and heterocyclic tail are located in the helix 3-β-sheet region where the groups interact with Val339 and Leu330 and Ile341 and Arg288, respectively.[15] NMP422 has a similar functional group configuration as the TZDs and the same interactions were observed at high frequencies in aMD simulations; however, the acidic head of NMP422 is located in the same region as the pyridyl group of rosiglitazone and the pentyl group is in the same region as the thiazolidinedione group, giving NMP422 the flipped binding mode (Figure 10b). The cause of the flip is unknown, and further work will need to be done to determine the cause. One reason may have to do with the location of the carboxylic acid.

There are many PPARγ partial agonists in the Protein Data Bank that contain a carboxylic acid in the helix 3-β-sheet region that forms hydrogen bonds with Arg288 and Ser342.[51] Two

specific examples can be seen with the PPARγ antagonist SR1664 (PDB: 4R2U) and the PPARγ partial agonist nTZDpa (PDB: 2Q5S). The carboxylic acids that form these interactions are bonded directly to an aromatic ring, acting as hydrogen bond acceptors and electron-withdrawing groups. In contrast, many PPARγ full agonists have at least one carbon separating acceptor groups from aromatic rings or have acceptor groups that are part of alicyclic systems. Based on these observations, we believe that functional groups that act as hydrogen bond acceptors and electron-withdrawing groups bound directly to aromatic rings cause preferential binding in the helix 3-β-sheet region instead of the helix 12 region.

3.6 Conclusions

The search for compounds that improve insulin sensitivity comparable to rosiglitazone and pioglitazone with lower transactivation is still ongoing. NMP422 is configured similarly to the TZDs but with a partial agonist profile that showed unique dual binding mode characteristics. One mode resembled the TZDs, but with a flipped conformation that causes the acidic head of the ligand to form interactions with residues Arg288 and Ser342 in the helix 3-β-sheet region. The direct cause of this flip needs to be investigated further; however, we believe aromatic rings bonded directly to hydrogen bond accepting groups that also withdraw electron-density from the ring system play a pivotal role. By identifying the exact cause of the flipped conformation, drug discovery efforts can be focused on creating more chemical diversity in PPARγ ligand libraries while maintaining a consistent binding mode independent of helix 12.

3.7 Future directions

As previously stated, the preferential binding in the helix 3-β-sheet region directed by benzoic acid like moieties needs to be tested further. There are two strategies that can be used

to aid in supporting this hypothesis. The first strategy is to make modifications only to the NMP422 structure. By adding one or two carbons between the aromatic ring and carboxylic acid in NMP422, we can identify whether the direct connection between the two functional groups is essential for the NMP422 binding mode. Ultimately, the intensity of electron-withdrawing should be reduced or eliminated while retaining the hydrogen bond acceptor capacity of the carboxylic acid resulting in binding in the helix 12 region.

The binding mode of each ligand can be validated using a combination of methods. Because helix 12 stabilization is caused by interactions with residues on helix 4/5, helix 11, and helix 12, we can use transactivation assays to determine agonist properties. Analogs that output transcriptional activity on par with full agonists would allow us to assume that the acceptor groups are binding in the helix 12 region. For more definitive clarification, we can use x-ray crystallography, NMR, and molecular dynamics simulations to identify the binding mode of the analogs.

References

1. Berger, J. & Moller, D. E. The mechanisms of action of PPARs. *Annu. Rev. Med.* **53,** 409–435 (2002).

2. Rosen, E. D., Sarraf, P., Troy, A. E. PPARγ is required for the differentiation of adipose tissue in vivo and in vitro. *Mol. Cell* **4,** 611–617 (1999).

3. Chandra, V. *et al.* Structure of the intact PPAR-γ-RXR-α nuclear receptor complex on DNA. *Nature* **456,** 350–356 (2008).

4. Helsen, C. & Claessens, F. Looking at nuclear receptors from a new angle. *Molecular and Cellular Endocrinology* **382,** 97–106 (2014).

5. Kroker, A. J. & Bruning, J. B. Review of the structural and dynamic mechanisms of PPAR γ partial agonism. *PPAR Res.* **2015,** (2015).

6. Kliewer, S. A., Umensono, K. Noonan, D. J., Heyman, R. A., and Evans, R. M. Convergence of 9-cis retinoic acid and peroxisome proliferator signalling pathways through heterodimer formation of their receptors. *Nature* **358,** 771–774 (1992).

7. Keller, H. *et al.* Fatty acids and retinoids control lipid metabolism through activation of peroxisome proliferator-activated receptor-retinoid X receptor heterodimers. *Proc. Natl. Acad. Sci. U. S. A.* **90,** 2160–2164 (1993).

8. Andreas J. Hörlein, Anders M. Näär, Thorsten Heinzel, Joseph Torchia, Bernd Gloss, Riki Kurokawa, Aimee Ryan, Yasutomi Kamei, Mats Söderström, C. K. G. & M. G. R. Ligand-independent repression by the thyroid hormone receptor mediated by a nuclear

receptor co-repressor. *Nature* **377,** 397–404 (1995).

9. Evans, J. D. C. and R. M. A transcriptional co-repressor that interacts with nuclear hormone receptors. *Nature* **377,** 454–457 (1995).

10. Koppen, Arjen, Houtman, Rene, Pijnenburg, Dirk, Jeninga, Ellen H., Ruijtenbeek, Rob, and Kalkhoven, E. Nuclear Receptor-Coregulator Interaction Profiling Identifies TRIP3 as a Novel Peroxisome Proliferator-activated Receptor Cofactor□.pdf. *Mol. Cell. Proteomics* **8,** 2212–2226 (2009).

11. Zoete, V., Grosdidier, A. & Michielin, O. Peroxisome proliferator-activated receptor structures: Ligand specificity, molecular switch and interactions with regulators. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids* **1771,** 915–925 (2007).

12. Schupp, M. & Lazar, M. A. Endogenous ligands for nuclear receptors: Digging deeper. *J. Biol. Chem.* **285,** 40409–40415 (2010).

13. Forman BM *et al.* 15-Deoxy-delta 12, 14-prostaglandin J2 is a ligand for the adipocyte determination factor PPAR gamma. *Cell* **85,** 803–12 (1995).

14. Waku, T. *et al.* The nuclear receptor PPARγ individually responds to serotonin-and fatty acid-metabolites. *EMBO J.* **29,** 3395–3407 (2010).

15. Milburn, M. V *et al.* Ligand binding and co-activator assembly of the peroxisome proliferator-activated receptor-γ. *Nature* **395,** 137–143 (1998).

16. Souza, P. C. T. *et al.* Medium Chain Fatty Acids Are Selective Peroxisome Proliferator

Activated Receptor (PPAR) γ Activators and Pan-PPAR Partial Agonists. *PLoS One* **7,** e36297 (2012).

17. Gelin, M., Delfosse, V., Allemand, F., Hoh, F., Sallaz-Damaz, Y., Pirocchi, M., Bourguet, W., Ferrer, J.lL., Labesse, G., A. Combining 'dry' co-crystallization and in situ diffraction to facilitate ligand screening by X-ray crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **D71,** 1777–1787 (2015).

18. Yki-Järvinen, H. Thiazolidinediones. *N. Engl. J. Med.* **351,** 1106–1118 (2004).

19. Aubert, R. E., Herrera, V., Chen, W., Haffner, S. M. & Pendergrass, M. Rosiglitazone and pioglitazone increase fracture risk inwomen andmen with type 2 diabetes. *Diabetes, Obes. Metab.* **12,** 716–721 (2010).

20. Nettles, K. W. *et al.* Partial Agonists Activate PPARγ Using a Helix 12 Independent Mechanism. *Structure* **15,** 1258–1271 (2007).

21. Zhang, F., Lavan, B. E. & Gregoire, F. M. Selective modulators of PPAR-γ activity: Molecular aspects related to obesity and side-effects. *PPAR Res.* **2007,** (2007).

22. Marciano, D. P. *et al.* Pharmacological repression of PPARγ promotes osteogenesis. *Nat. Commun.* **6,** 1–7 (2015).

23. Chrisman, I. M. *et al.* Defining a conformational ensemble that directs activation of PPARγ. *Nat. Commun.* **9,** 1–16 (2018).

24. Smith, J. S., Lefkowitz, R. J. & Rajagopal, S. Biased signalling: From simple switches to allosteric microprocessors. *Nat. Rev. Drug Discov.* **17,** 243–260 (2018).

25. Zhang, H. *et al.* Molecular determinants of magnolol targeting both RXRα and PPARγ. *PLoS One* **6,** (2011).

26. Bishop-bailey, D., Hla, T. & Warner, T. D. Bisphenol A diglycidyl ether (BADGE) is a PPAR g agonist in an. 651–654 (2000). doi:10.1038/sj.bjp.0703628

27. Pastelín-Hernández, G. *et al.* Rosiglitazone, a Ligand to PPAR γ , Improves Blood Pressure and Vascular Function through Renin-Angiotensin System Regulation . *PPAR Res.* **2019,** 1–12 (2019).

28. Glass, C. K. & Rosenfeld, M. G. The coregulator exchange in transcriptional functions of nuclear receptors. *Genes Dev.* **14,** 121–141 (2000).

29. Spiegelman, B. M. *et al.* β-Aminoisobutyric Acid Induces Browning of White Fat and Hepatic β-Oxidation and Is Inversely Correlated with Cardiometabolic Risk Factors. *Cell Metab.* **19,** 96–108 (2014).

30. Meyer, M.B., & Pike, J. W. Corepressors (NCoR and SMRT) as well as coactivators are recruited to positively regulated 1α,25-dihydroxyvitamin D3-responsive genes. *J. Steroid Biochem. Mol. Biol.* **136,** 120–124 (2013).

31. Kenakin, T. P. *Pharmacologic Analysis of Drug/Receptor Interaction 2nd Edition*. (1993).

32. Eswar, N., et al. Comparative protein structure modeling using Modeller. *Curr. Protoc. Bioinforma.* **Chapter 5,** (2006).

33. Pettersen, E. F. *et al.* UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25,** 1605–1612 (2004).

34. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and testing of a general Amber force field. *J. Comput. Chem.* **25,** 1157–1174 (2004).

35. Gordon, J. C. *et al.* H++: A server for estimating pKas and adding missing hydrogens to macromolecules. *Nucleic Acids Res.* **33,** 368–371 (2005).

36. D.A. Case, V. Babin, J.T. Berryman, R.M. Betz, Q. Cai, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R. E., Duke, H. Gohlke, A.W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. K., T.S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, W. and P. A. AMBER 14. (2014).

37. Vanquelef, E. *et al.* R.E.D. Server: A web service for deriving RESP and ESP charges and building force field libraries for new molecules and molecular fragments. *Nucleic Acids Res.* **39,** 511–517 (2011).

38. Cornell, W. D., Cieplak, P., Bayly, C. I. & Kollman, P. A. Application of RESP Charges To Calculate Conformational Energies, Hydrogen Bond Energies, and Free Energies of Solvation. *J. Am. Chem. Soc.* **115,** 9620–9631 (1993).

39. Jorgensen, W.L., Chandrasekhar, J., Madura, J. D. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79,** 926–935 (1983).

40. Joung, I. S. & Cheatham III, T. E. Molecular Dynamics Simulations of the Dynamic and Energetic Properties of Alkali and Halide Ions Using Water-Model-Specific Ion Parameters. *J. Phys. Chem. B* **113,** 13279–13290 (2009).

41. Roe, D. R. & Cheatham, T. E. PTRAJ and CPPTRAJ: Software for processing and analysis of

molecular dynamics trajectory data. *J. Chem. Theory Comput.* **9,** 3084–3095 (2013).

42. Miao, Y. *et al.* Improved Reweighting of Accelerated Molecular Dynamics Simulations for Free Energy Calculation. *J. Chem. Theory Comput.* **10,** 2677–2689 (2014).

43. Kim, W. *et al.* Discovery and Optimization of nTZDpa as an Antibiotic Effective Against Bacterial Persisters. *ACS Infect. Dis.* **4,** 1540–1545 (2018).

44. Chou, Y.-C. *et al.* Novel Indole-Based Peroxisome Proliferator-Activated Receptor Agonists: Design, SAR, Structural Biology, and Biological Activities. *J. Med. Chem.* **48,** 8194–8208 (2005).

45. Diaz, P., Barbuis, C. M. Novel biaromatic compounds that modulate PPAR type receptors and cosmetic/pharmaceutical compositions comprised thereof. (2008).

46. Choi, J. H. *et al.* Antidiabetic actions of a non-agonist PPARγ ligand blocking Cdk5-mediated phosphorylation. *Nature* **477,** 477–481 (2011).

47. Kuruvilla, D. S. *et al.* Ligand and Receptor Dynamics Contribute to the Mechanism of Graded PPARγ Agonism. *Structure* **20,** 139–150 (2012).

48. Kamenecka, T. M. *et al.* An alternate binding site for PPARγ ligands. *Nat. Commun.* **5,** (2014).

49. Chen, V. B. *et al.* MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66,** 12–21 (2010).

50. D.A. Case, I.Y. Ben-Shalom, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, D. Ghoreishi, M.K. Gilson, H. Gohlke, A.W. Goetz, D. Greene, R

Harris, N. Homeyer, S. Izadi, A. Kovalenko, T. Kurtzman, T.S. Lee, S. LeGra, D. M. Y. and P. A. K. AMBER 2018. (2018).

51. Guasch, L. *et al.* Structural insights for the design of new PPARgamma partial agonists with high binding affinity and low transactivation activity. *J. Comput. Aided. Mol. Des.* **25,** 717–728 (2011).

Supplemental Information

Chapter 2 Supplemental Information

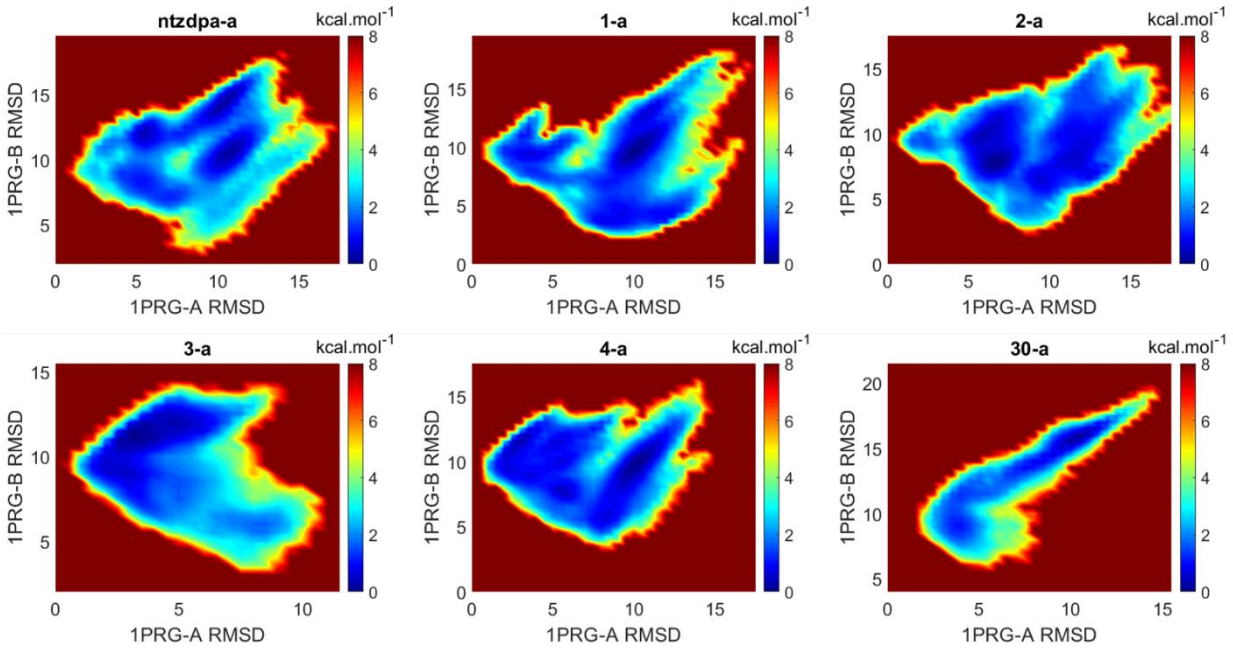


Figure S1. Free-energy topology maps of nTZDpa and R^2 hydrogen bond donor and acceptor groups changes.

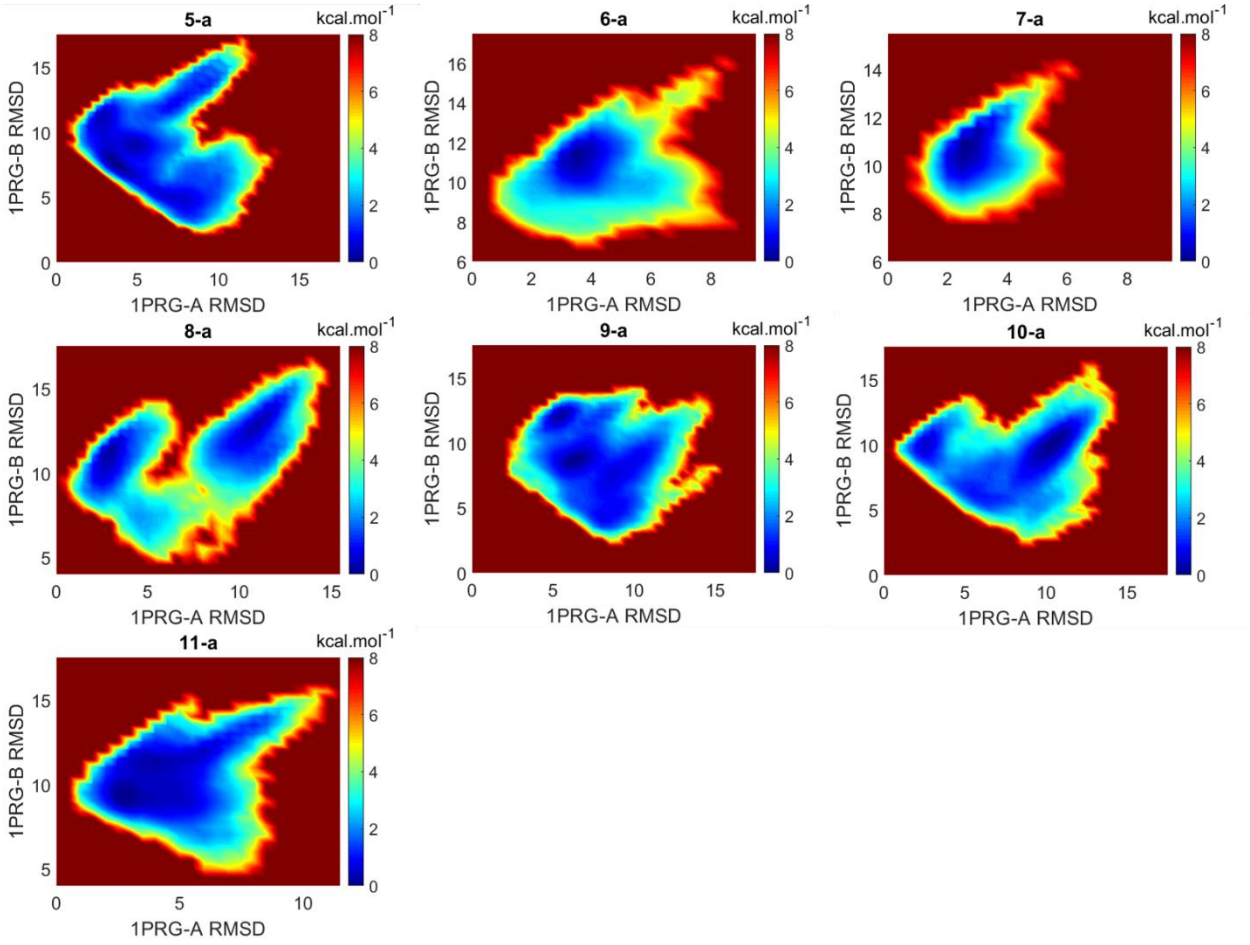


Figure S2. Free-energy topology maps of R^2 bulkiness changes.

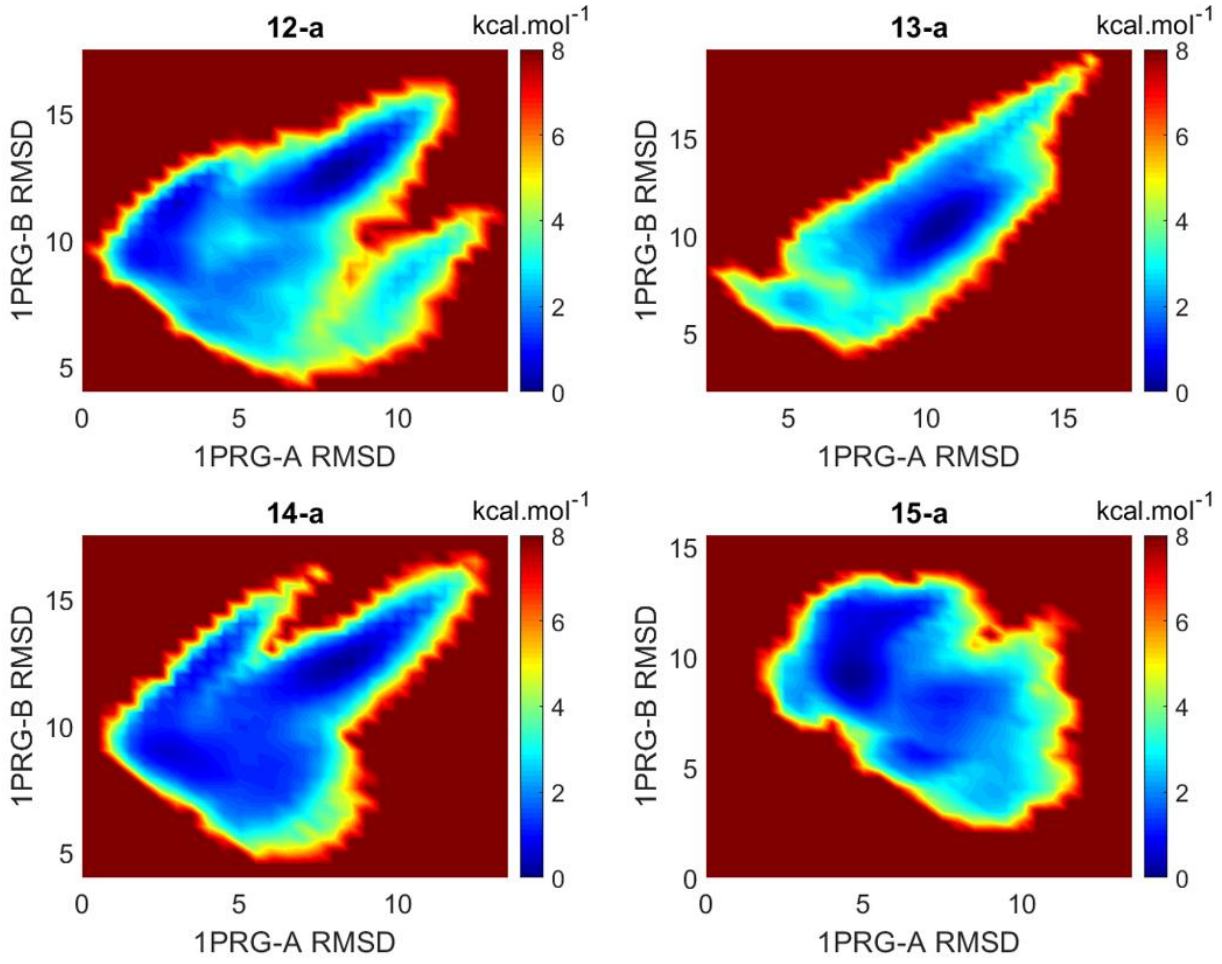


Figure S3. Free-energy topology maps of R^1 bulkiness changes.

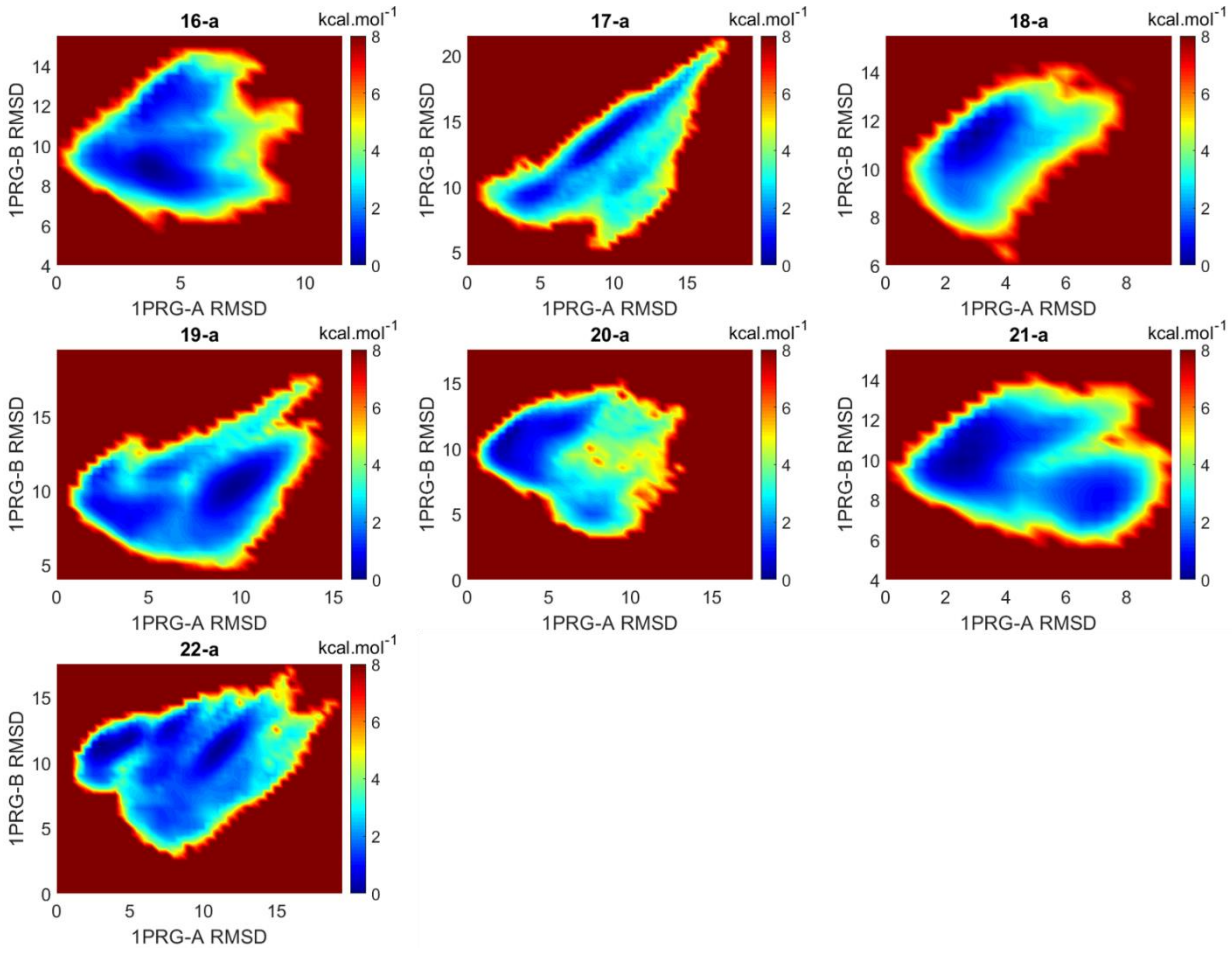


Figure S4. Free-energy topology maps of R^1 hydrogen bond donor and acceptor groups changes.

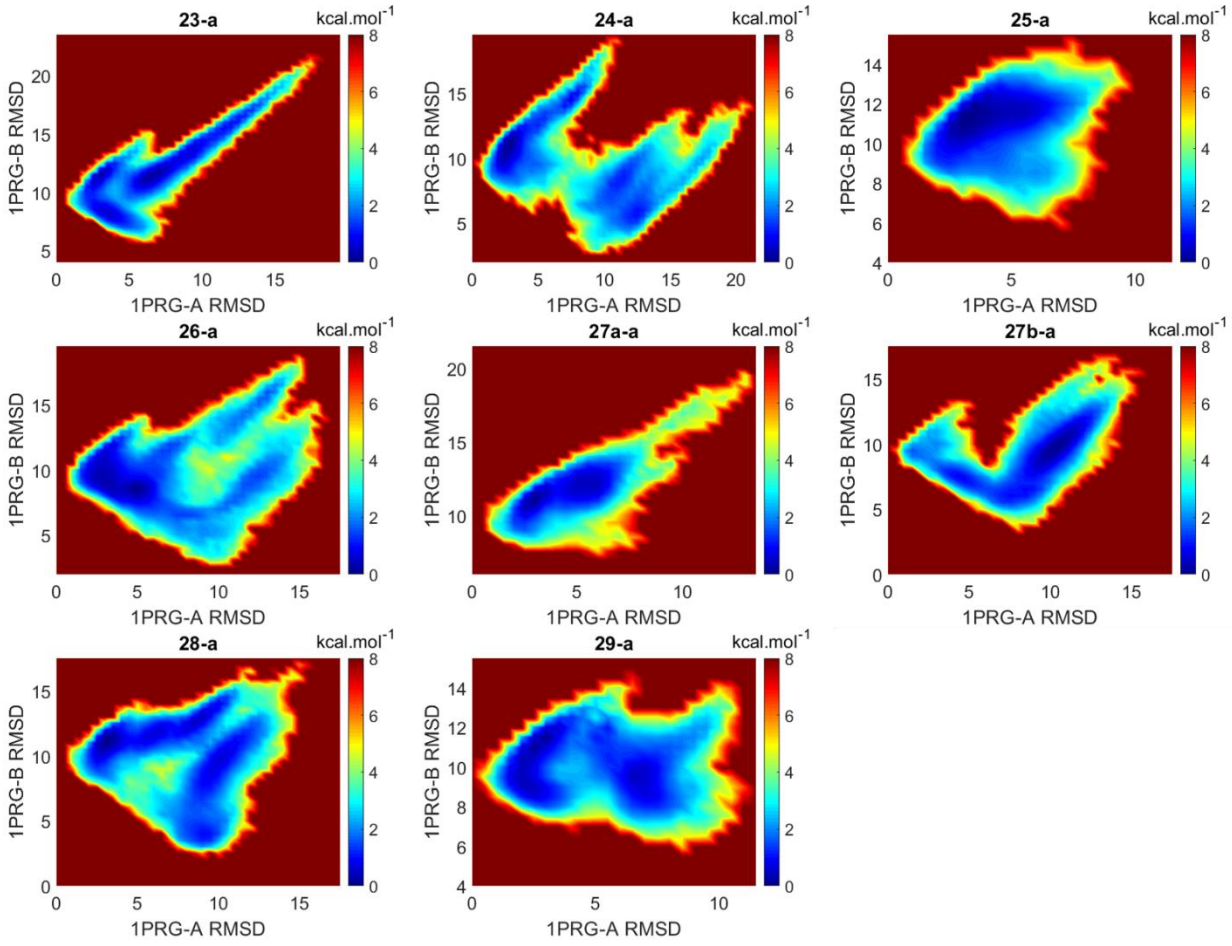


Figure S5. Free-energy topology maps of nTZDpa and R^3 thio-group changes.

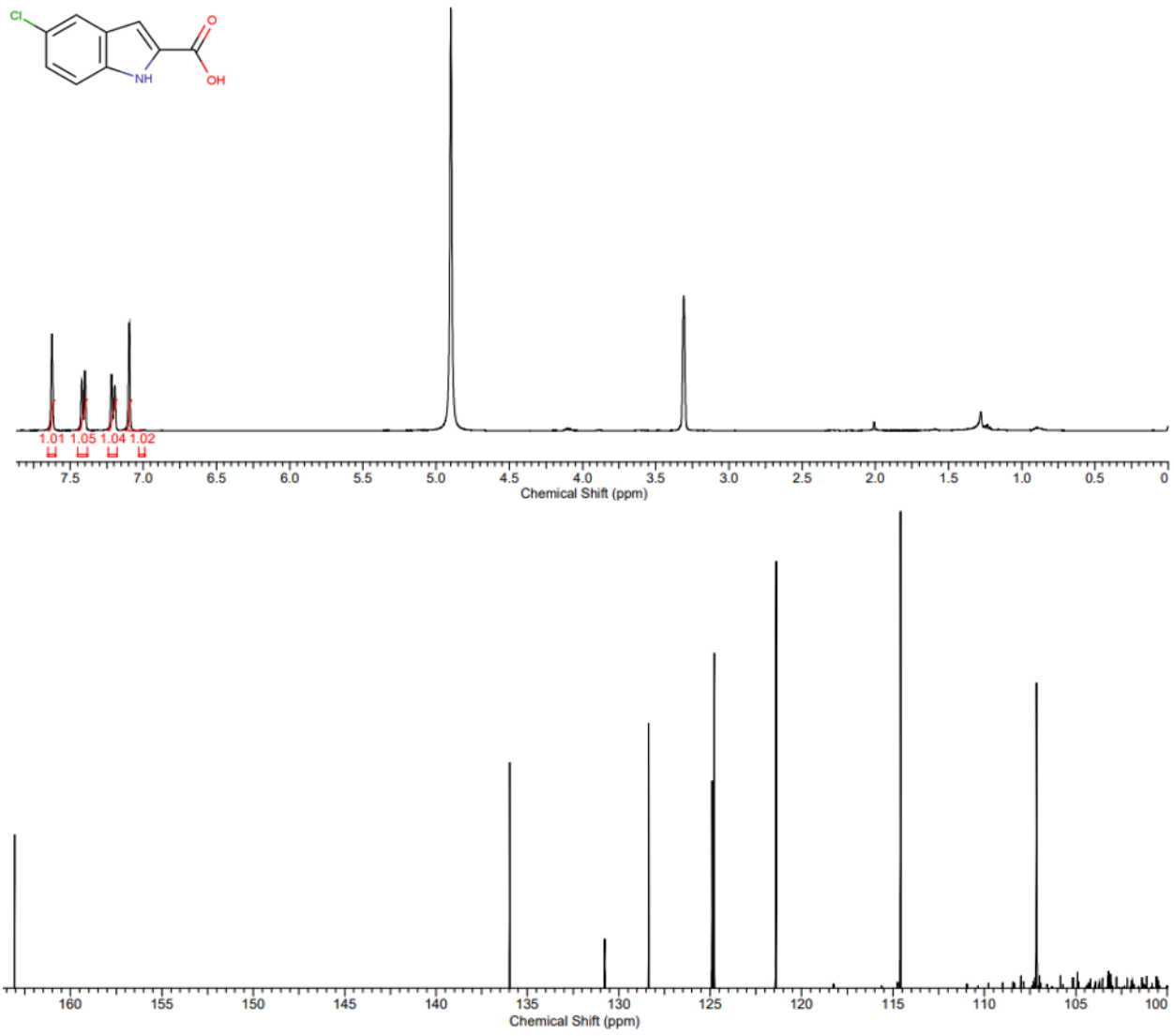


Figure S6. H and C NMR for compound 1a.

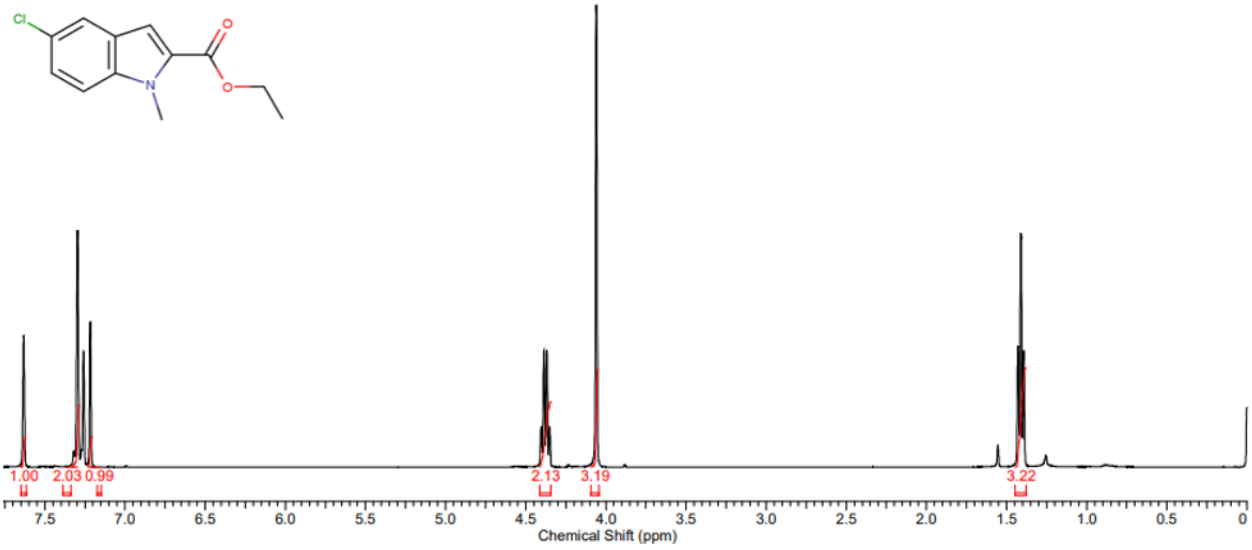


Figure S7. H NMR for intermediate of 1b.

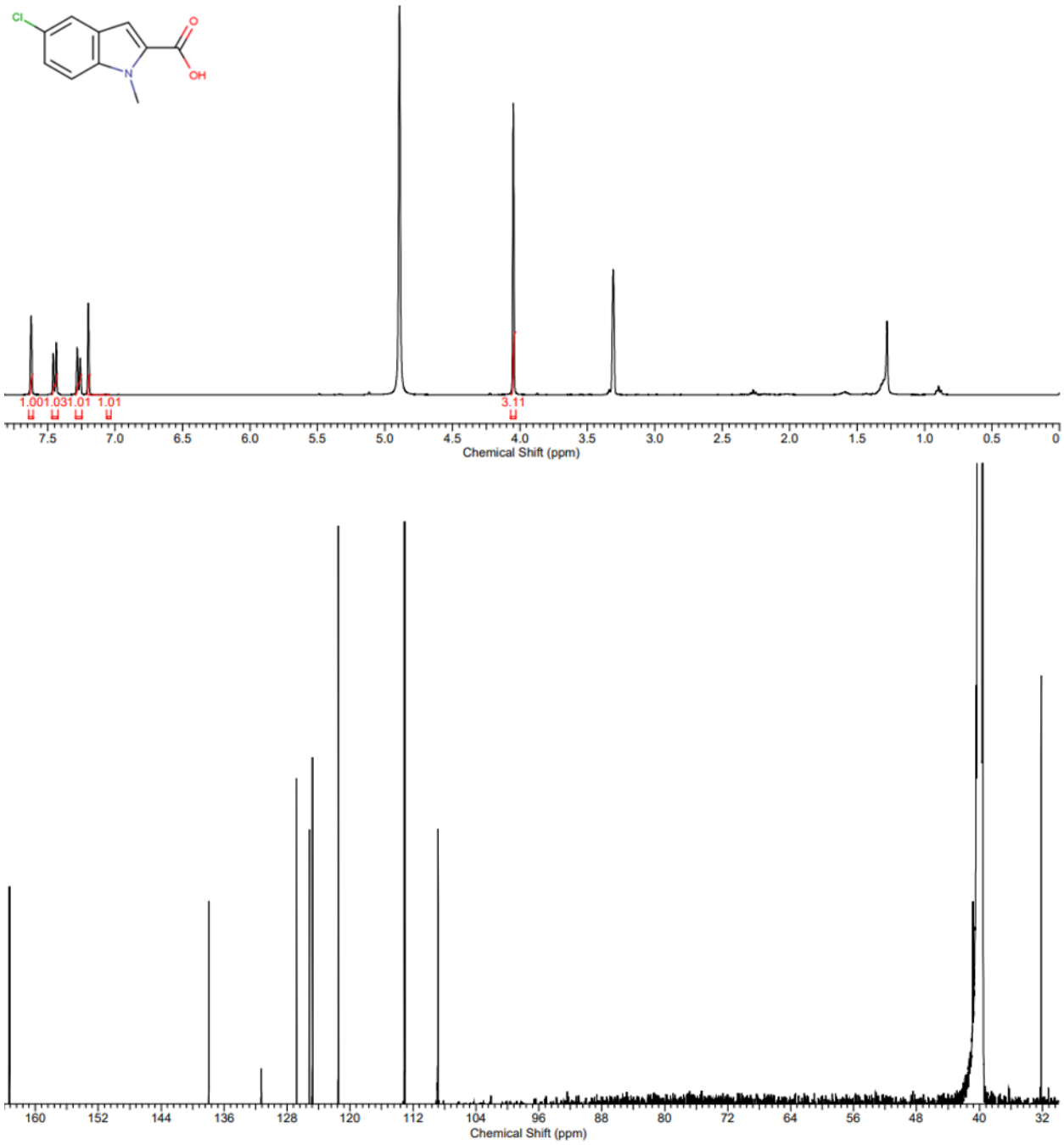


Figure S8. H and C NMR for compound 1b.

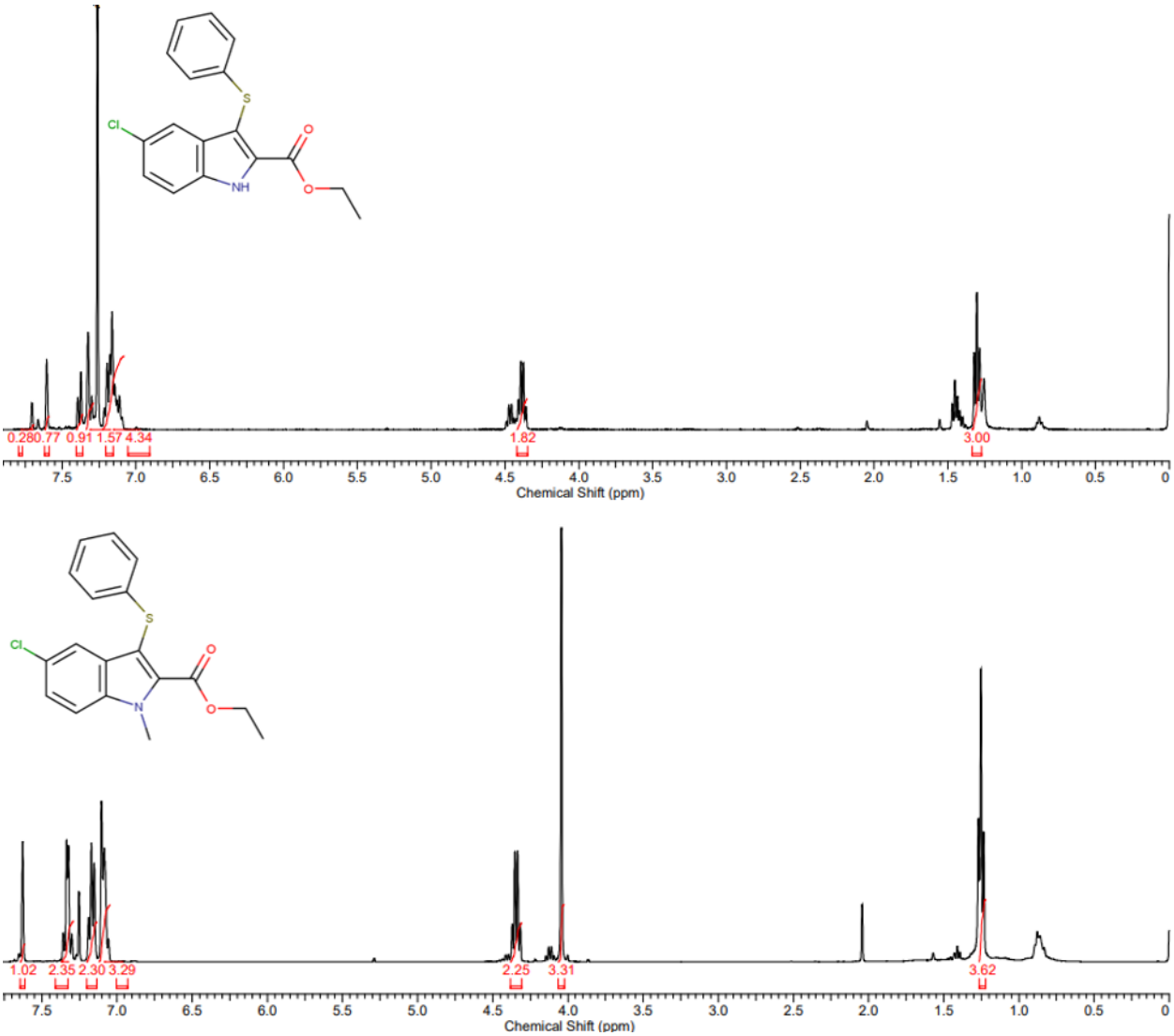


Figure S9. H NMR for intermediates of 1c.

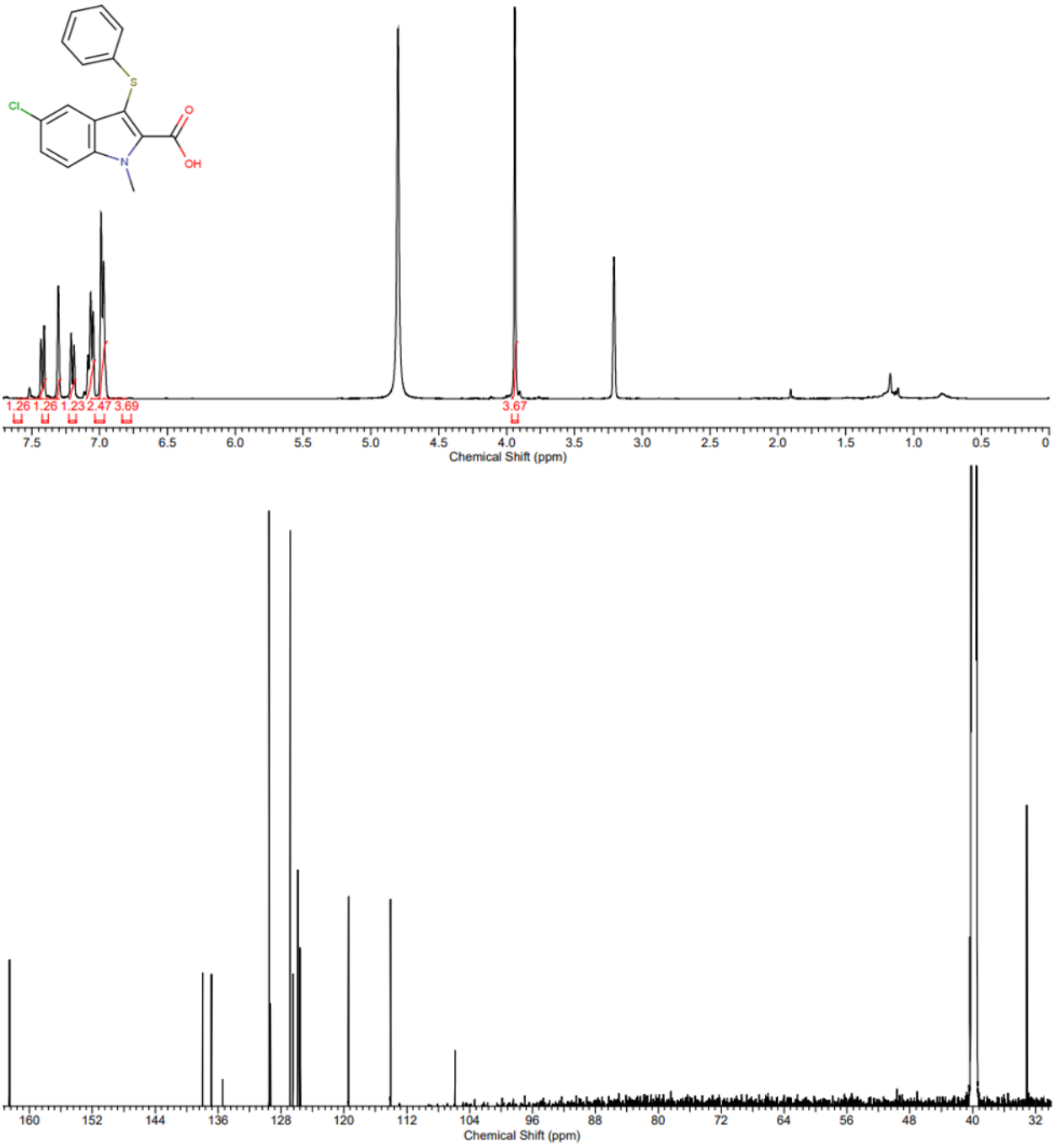


Figure S10. H and C NMR of compound 1c.

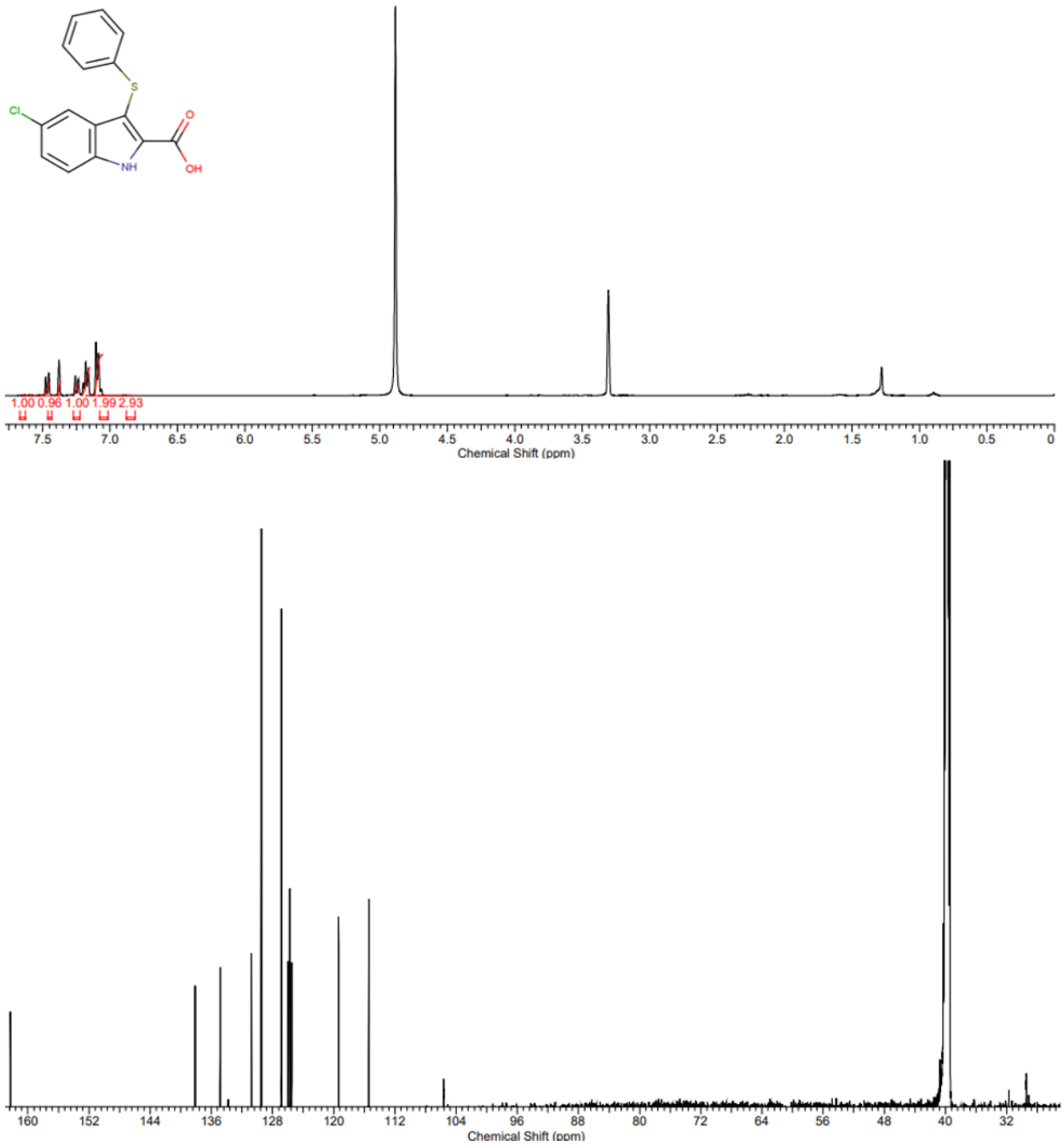


Figure S11. H and C NMR of compound 1d.

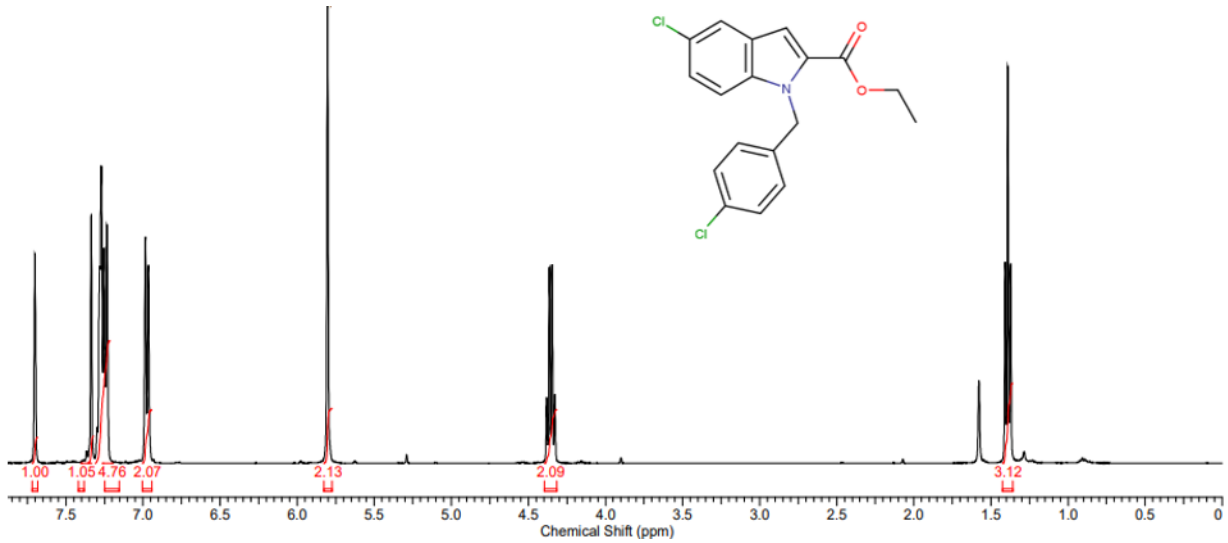


Figure S12. H NMR for intermediate of 1e.

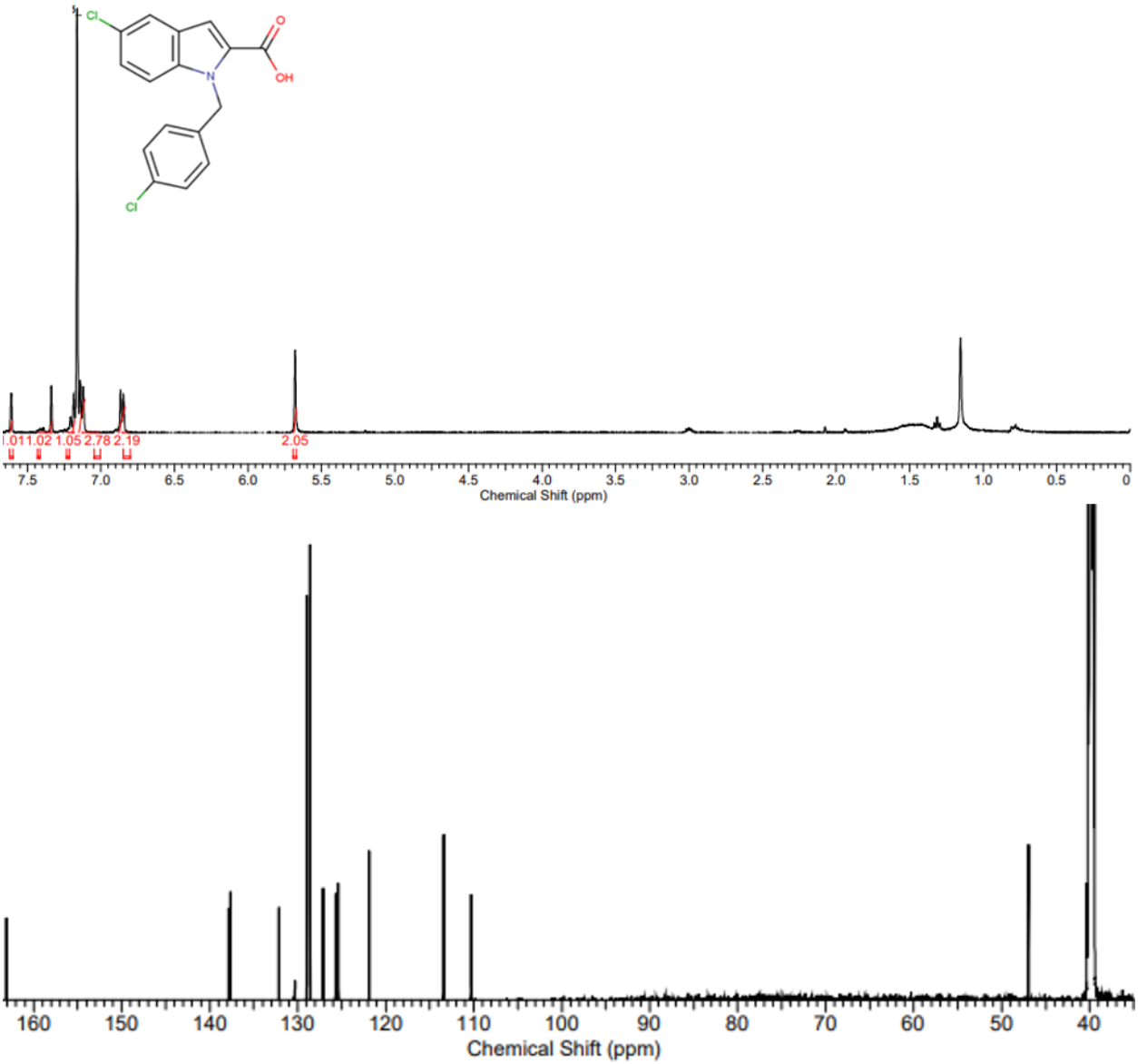


Figure S13. H and C NMR for compound 1e.

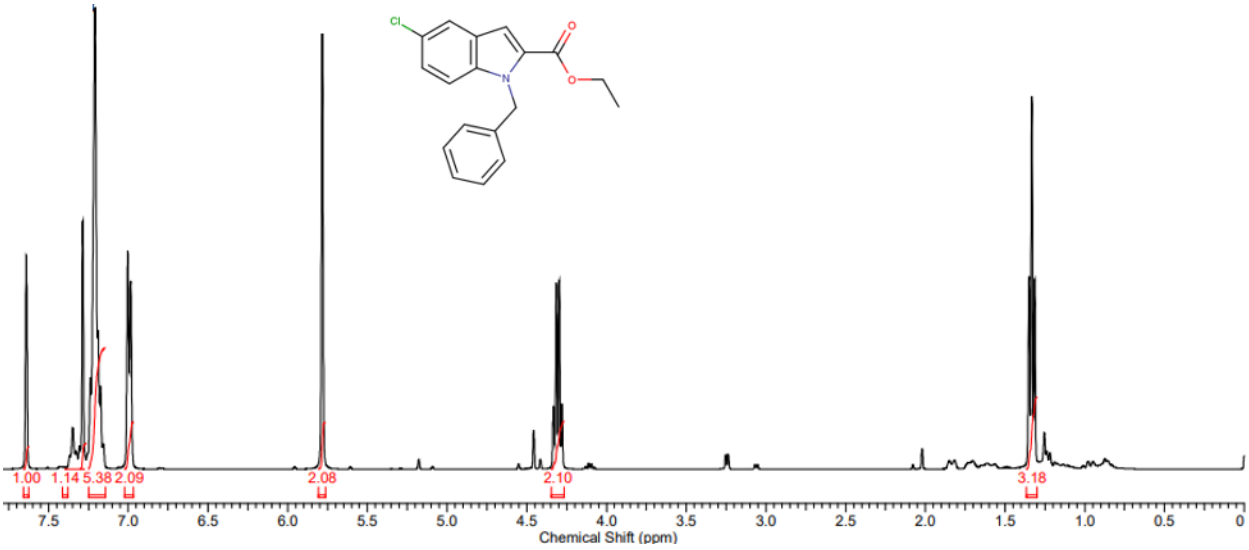


Figure S14. H NMR for intermediate of 2a.

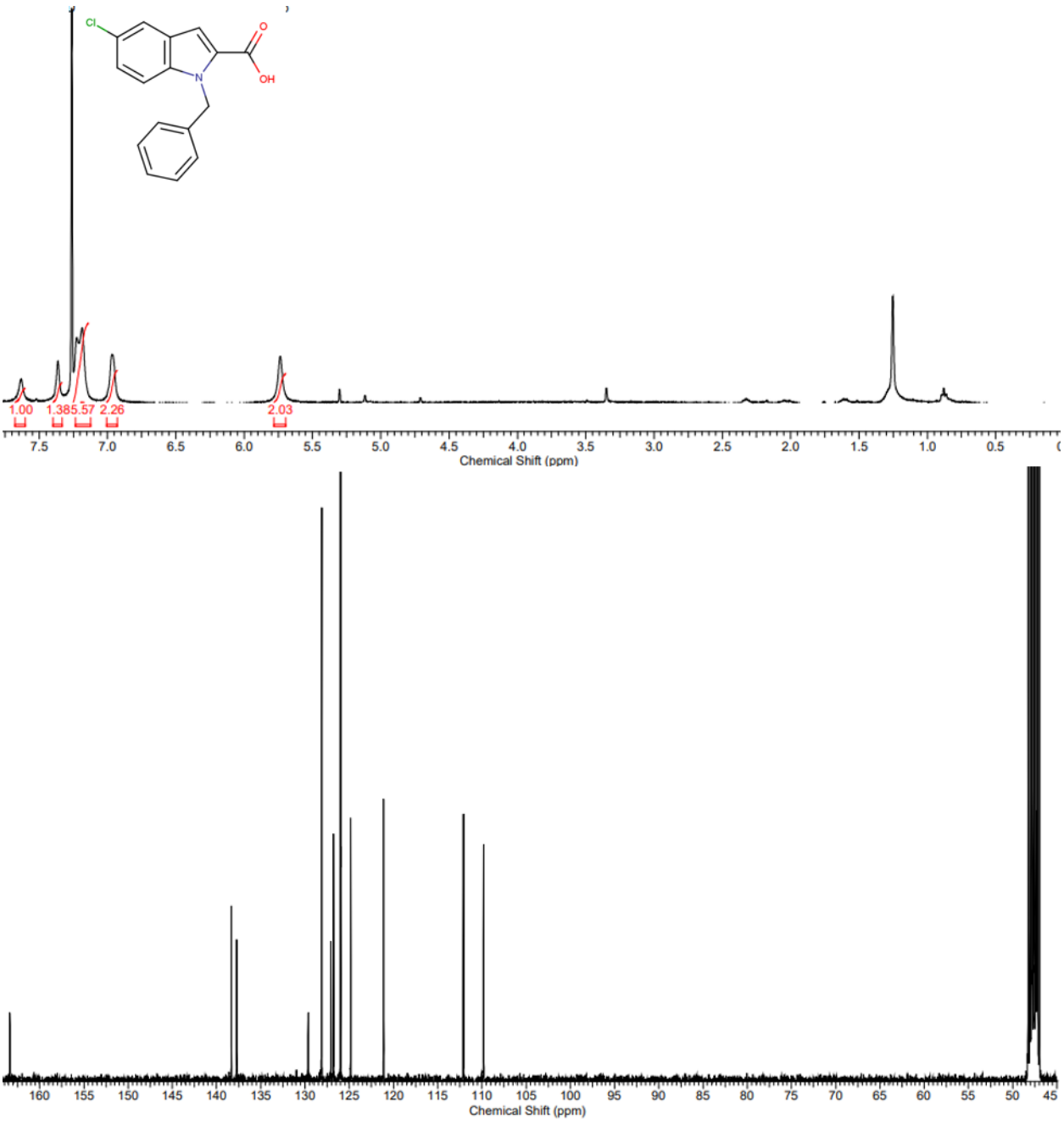


Figure S15. H and C NMR for compound 2a.

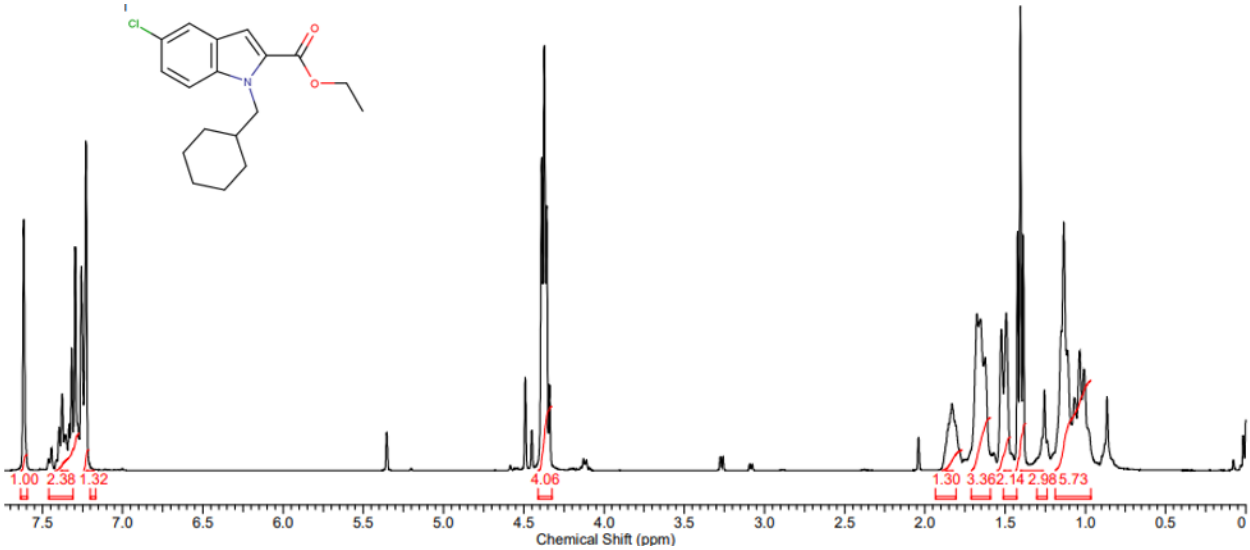


Figure S16. H NMR for intermediate of 2b.

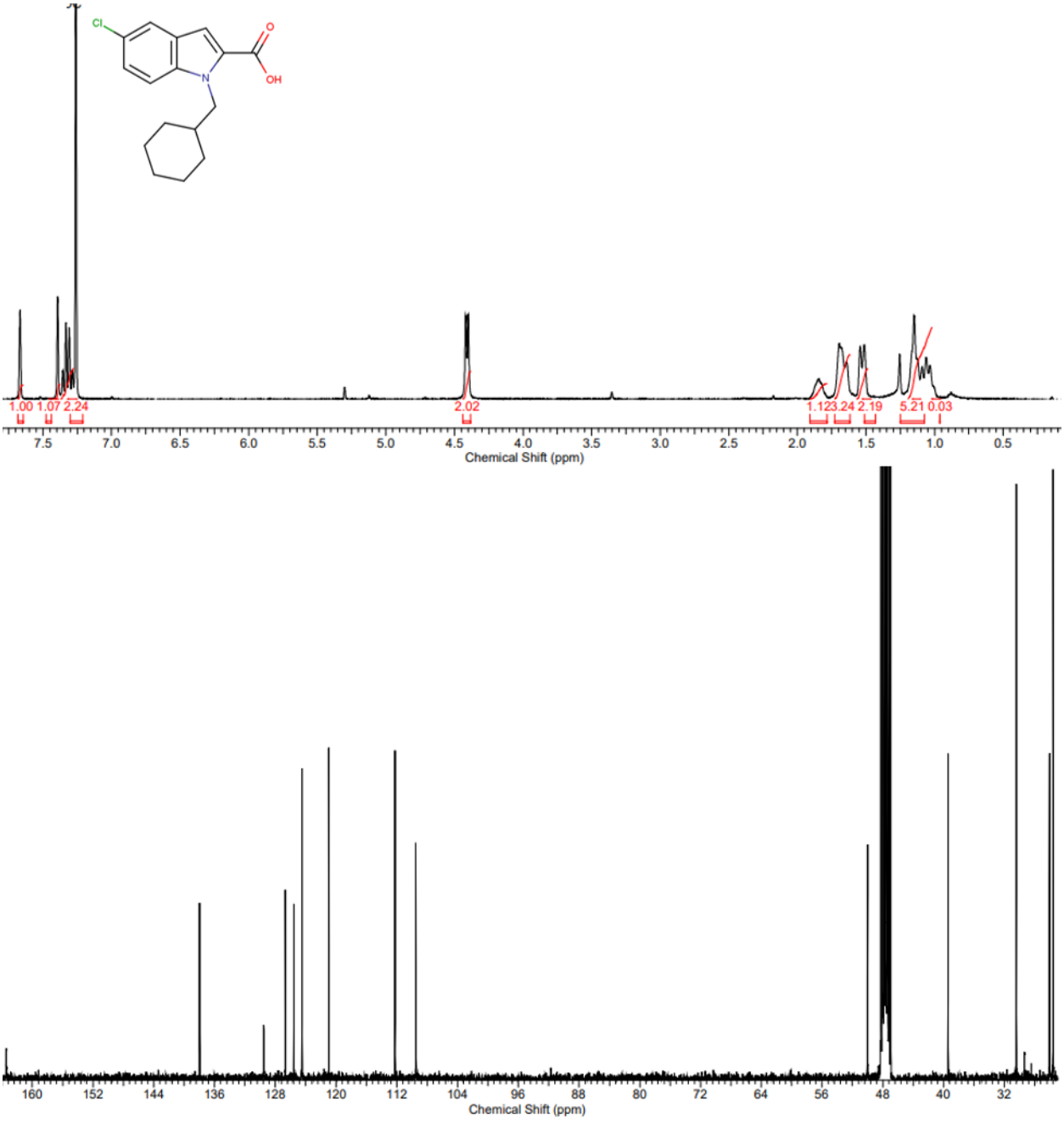


Figure S17. H and C NMR for compound 2b.

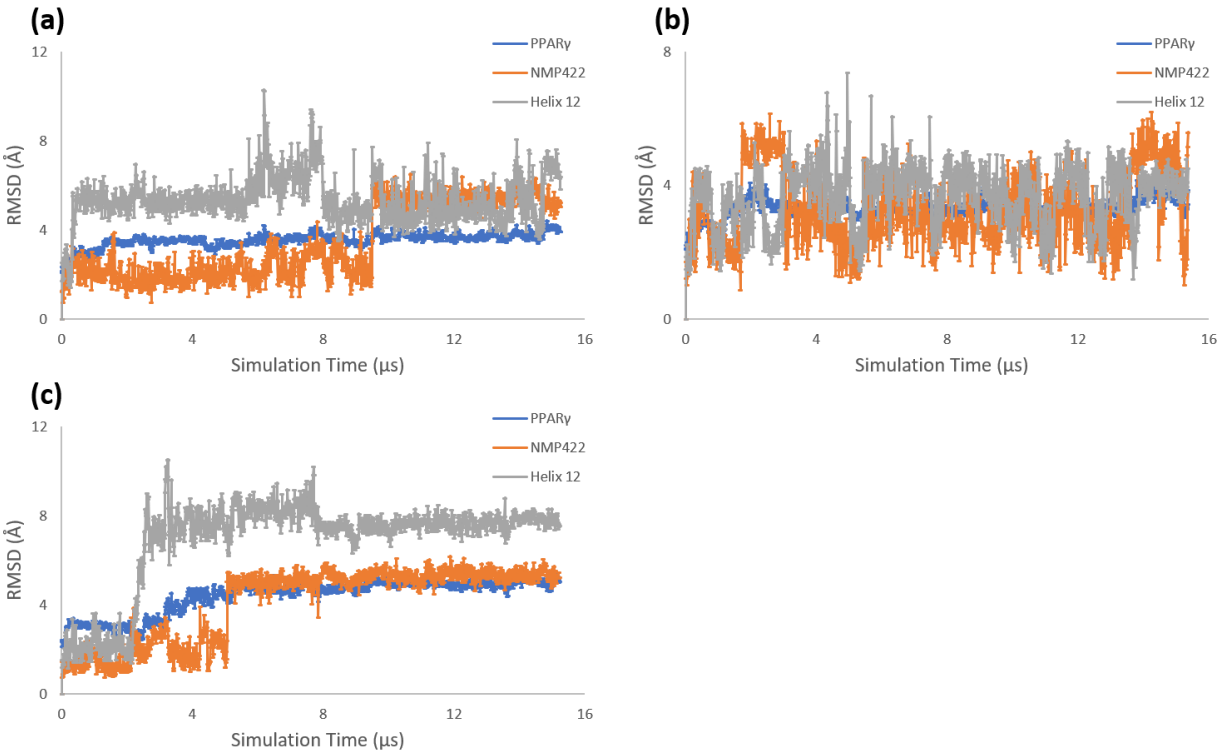


Figure S18. Root-mean square deviation versus time for NMP422 conventional molecular dynamics simulations. a. Production run a, 15.26 μs. b. Production run b, 15.37 μs. c. Production run c, 15.23 μs.

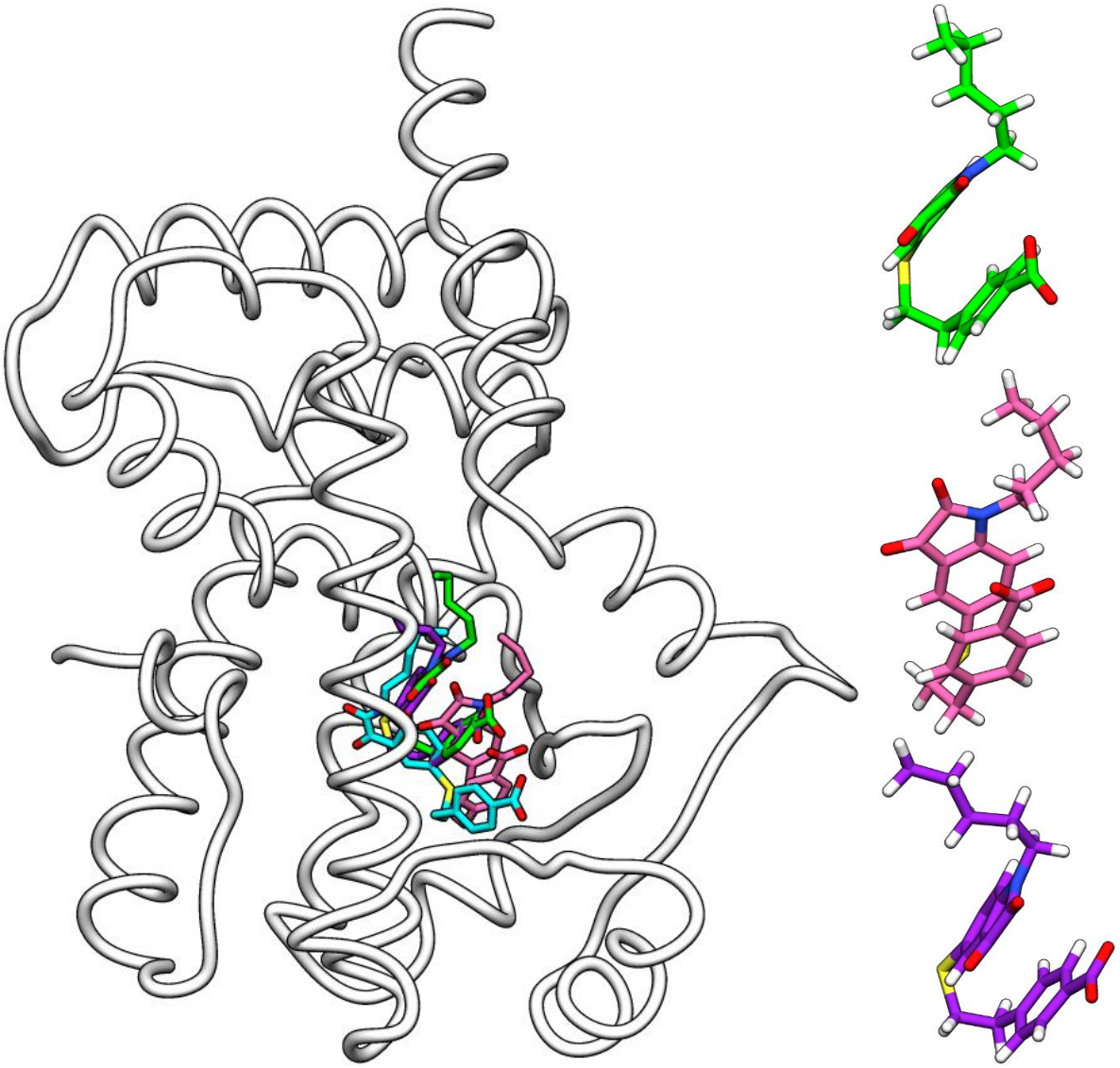


Figure S19. NMP422-bound PPARγ (white) in alignment with major k means clusters resulting from three conventional molecular dynamics simulations. NMP422 adopts a pi-stacking binding mode in run a (green), run b (pink), and run c (purple) while the binding mode in the crystal structure (cyan) does not.

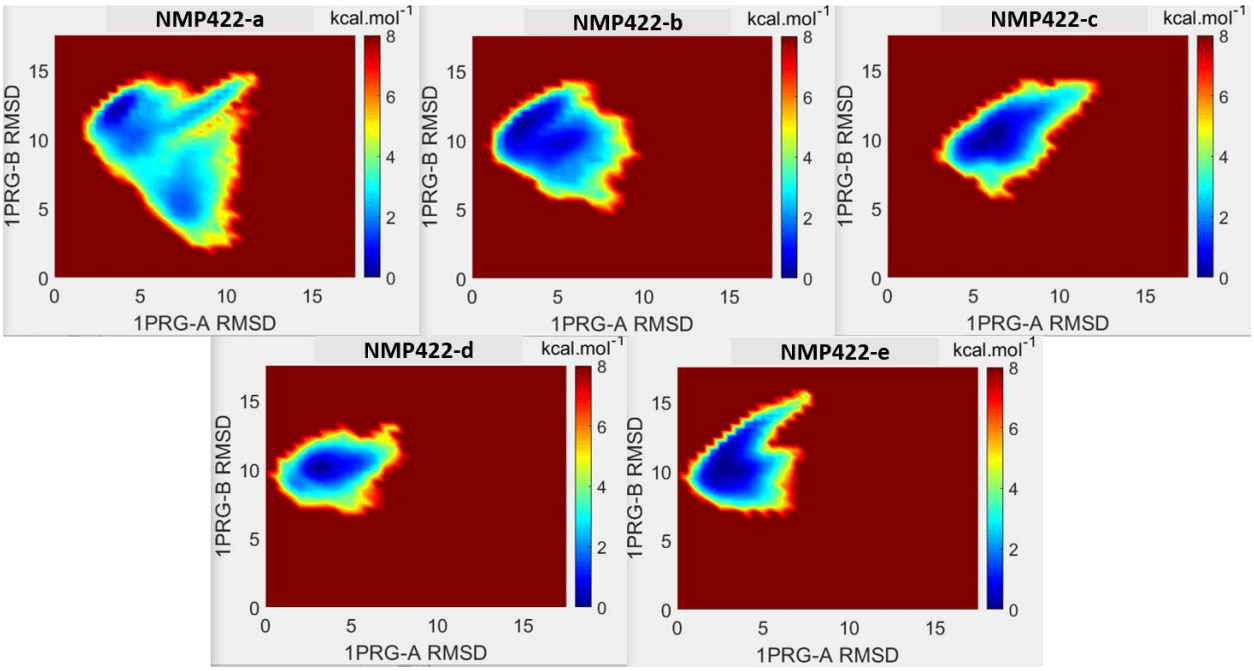


Figure S20. Free-energy topology maps of NMP422 accelerated molecular dynamics simulations.