TOWARD BETTER DEVELOPMENT OF PROTEIN-PROTEIN INTERACTION INHIBITORS: APPLICATIONS TO THE β-CATENIN/T-CELL FACTOR PROTEIN-PROTEIN INTERACTION

by

Jonathan Leon Catrow

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Chemistry

The University of Utah

August 2017

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Jonathan Leon Catrow**

has been approved by the following supervisory committee members:

Haitao Ji	, Chair	**12/09/2016** Date Approved
Bethany Anne Koehntop	, Member	**12/09/2016** Date Approved
Jon D. Rainier	, Member	**12/09/2016** Date Approved
Ryan P. Steele	, Member	**12/09/2016** Date Approved
Markus Babst	, Member	**12/09/2016** Date Approved

and by **Cynthia Burrows** , Chair/Dean of

the Department/College/School of **Chemistry**

and by David B. Kieda, Dean of The Graduate School.

ABSTRACT

The dysregulation of protein–protein interaction (PPI) networks has been implicated in many diseases. Designing therapeutic small-molecule inhibitors of these interactions is a challenging field for medicinal chemistry. This work advances the techniques for discovering more potent PPI inhibitors through integration of computational and biochemical techniques.

High-throughput screening using fluorescence polarization and AlphaScreen assays identified an acyl hydrazone-containing inhibitor of the β-catenin/Tcf4 PPI, a key mediator of the canonical Wnt signaling pathway. By removing the undesirable acyl hydrazone moiety, a new compound, 4-(5H-[1,2,5]oxadiazolo[3',4':5,6]pyrazino[2,3-b]indol-5-yl)butanoic acid, was developed to selectively inhibit the β-catenin/Tcf4 interaction. The ethyl ester of this compound was tested in zebrafish embryos and shown to inhibit Wnt signaling in vivo at 2 and 10 μM concentrations.

Differences between the PPI interface and the active site of traditional targets add to the difficulty of discovering PPI inhibitors. Herein, the relationship between inhibitor potency and ligand burial—defined as the fraction of the solvent accessible surface areas of the bound over unbound ligand, θ_l—in the PPI surface was evaluated. A positive correlation between θ_l and inhibitor potency was discovered. However, this correlation was secondary to the strong nonbonding interactions. A study of five PPI targets with corresponding inhibitor-bound crystal structures also revealed that empirical scoring

functions were slightly better at identifying known inhibitors out of the putatively inactive test set, and the Lamarckian genetic algorithm was more successful at pose prediction.

Due to the nature of the PPI surface, directly targeting the binding site may be difficult. A novel combination of computational methods explored the druggability, selectivity, and potential allosteric regulation of PPIs. Solvent mapping confirmed that Tcf4, E-cadherin, APC and axin use the same binding site on β-catenin in different ways. Evolutionary trace analysis indicated that the region surrounding W504 of β-catenin might be a potentially allosteric site. Site-directed mutagenesis testing results for a W504I β-catenin mutant resulted in three-fold increased binding of Tcf4 to β-catenin over the wild-type. This new site is promising for the discovery of future allosteric inhibitors of the β-catenin/Tcf4 PPI. The combined results from these studies reveals ways to better design PPI inhibitors.

TABLE OF CONTENTS

LIST OF FIGURES

Figures

LIST OF TABLES

Tables

LIST OF ABBREVIATIONS

Abbreviations

APC: adenomatous polyposis coli

AR: androgen receptor

Arf1: adenosine diphosphate ribosylation factor

ARNO: cytohesin-2

BAR-1: β-catenin/armadillo-related protein 1

BCL9: B-cell lymphoma 9

BIR3: baculovirus inhibitor of apoptosis protein repeat 3

BRD4: bromodomain-containing protein 4

BRG1: Brahma-related gene 1

C-Abl: Abelson murine leukemia viral oncogene homolog 1 kinase

CBP: CREB-binding protein

cDNA: complimentary DNA

C_T: threshold cycle

DCC: dynamic combinatorial chemistry

DNA: deoxyribonucleic acid

DTT: dithiothreitol

ET: evolutionary trace

Fz: frizzled protein

FP: fluorescence polarization

GAPDH: glyceraldehyde 3-phosphate dehydrogenase

GPCR: G-protein coupled receptor

GSK3: glycogen synthase kinase 3

HPrK: histidine-containing phosphocarrier protein kinase/phosphatase

HPRT: hypoxanthine phosphoribosyltransferase 1 gene

HTVS: high-throughput virtual screening

HMP-2: humpback 2

IC_{50}: half maximal inhibitory concentration

IN-LEDGF: lens epithelium-derived growth factor

IFN: interferon

IFNAR: interferon alpha/beta receptor 1

INT1: integrator complex subunit 1 gene

JUP: plakoglobin

K_d: dissociation constant

K_i: inhibitory constant

LE: ligand efficiency

LEF1: lymphoid enhancer-binding factor 1

logp: octanol/water partition coefficient

LRP: lipoprotein receptor-related protein

Md2: lymphocyte antigen 96

MDM2: murine double minute 2 homolog protein

MEC: mammary epithelial cell

MM/GBSA: molecule mechanics/meneralized Borne solvent accessible

MS: mass spectroscopy

MLL: mixed lineage leukemia protein

MSA: multiple sequence alignment

NEF: negative regulatory factor protein

Ni-NTA: nickel-nitrilotriacetic acid

NMR: nuclear magnetic resonance spectroscopy

PAIN: Pan-assay interference

PPI: protein–protein interaction

Pygo: pygopus protein

q-AP-MS: quantitative affinity pulldown mass spectroscopy

Rac1: ras-related C3 botulinum toxin substrate 1

rmsd: root-mean-square deviation

RT-qPCR: quantative reverse transcriptase real-time polymerase chain reaction

rvET: real-value evolutionary trace

SAR: structure–activity relationship

SCA: statistical coupling analysis

SP: standard precision

SYS-1: symmetrical sister cell hermaphrodite gonad defect 1

Tiam1: T-lymphoma invasion and metastasis-inducing protein 1

Tcf4: T-cell factor 4

TLR4: toll-like receptor 4

Wnt: wingless/integration complex subunit 1

WT: wild-type

WDR5: tryptophan-aspartic acid repeat protein 5

XIAP: X-linked inhibitor of apoptosis protein

$|r_s|$: Spearman’s rho

ACKNOWLEDGEMENTS

I would like to acknowledge my advisor Dr. Haitao Ji for his continued support, and his wife Dr. Min Zhang for her efforts in the biological testing presented herein. Her hard work was essential for making the work presented in the second and fourth chapters a success. I would also like to acknowledge all members of the Ji group who have made themselves available to help when asked. My wife, too, deserves special thanks for loving me during my graduate studies, as well as aiding in the creation of the first and final images of this document.

CHAPTER 1

INTRODUCTION

1.1 Protein–protein interactions as drug targets

Protein–protein interactions (PPIs) form the cornerstone of many important cellular processes such as signaling pathways, cellular adhesion networks, and gene transcription.[1] The disruption or dysregulation of PPI networks is responsible for many diseases.[1] Therefore, PPIs make for desirable therapeutic targets. Despite composing a large portion of the human proteome,[2,3] disproportionately few PPI inhibitors have advanced through clinical trials.[1,3] Therefore, there is a great need to improve methods to discover PPI inhibitors.[1]

PPIs are distinct from enzymes and other traditional drug targets including G-protein coupled receptors (GPCRs), proteases, and kinases; these distinctions lead to challenges in inhibitor design.[1,4] PPIs occur between two proteins over a large contacting surface upwards of 3,000 $Å^2$,[1] whereas enzymes have active site surfaces of less than 1,000 $Å^2$.[1] Key interacting residues of PPIs are dispersed throughout the larger surface, instead of clustered together in an active site.[1] The enzyme active site has a native small-molecule substrate which can be a template for an inhibitor, but PPIs have no such small-molecule substrate.[1] The enzyme active site is often deeply buried in the protein surface, yet PPI surfaces are shallow and solvent exposed.[1] As a result, PPI inhibitors often have

low ligand efficiency (LE),[2,4,5] poor pharmacokinetic properties[6] and the presence of pan-assay interference (PAIN) substructures.[4,7]

The large surfaces of the PPI interface result in higher molecular weight inhibitors than inhibitors of traditional targets with similar potency.[1] Because much of the PPI interaction surface is hydrophobic, small-molecule inhibitors of PPIs tend to be more hydrophobic.[1] This hydrophobicity creates high octanol/water partition coefficient (logp) values.[1] These characteristics create challenges when designing inhibitors with desirable absorptions, distribution, metabolism, and excretion pharmacokinetic properties.[1] The presence of PAIN substructures lead to false positive hits, potential reactivity, poor bioavailability or toxicity.[7] PPIs also face challenges with inhibitor selectivity. A single PPI site may accommodate multiple binding partners. Such is the case with β-catenin, which uses the same binding surface to interact with five different proteins.[8-12] These challenges have led to specific strategies being developed to discover drug-like PPI inhibitors.

Many of the same techniques used for drug discovery in traditional targets may be used to discover PPI inhibitors.[1] However, specific accommodations must be made for differences between the deep enzyme active site and the large, shallow PPI surface. For example, many enzyme inhibitors have been designed from the native binding ligand of the enzyme. This is not possible with PPIs, but it is possible to create a small molecule that mimics a protein binding partner. Such molecules, known as peptidomimetic, include cyclic petides[13] and stapled α-helices.[14] However, many peptidomimetic inhibitors lack good drug-like properties.[6] Designing drug-like small-molecule PPI inhibitors is possible, but the molecule must overcome the daunting task of inhibiting the much larger protein

partner. One successful technique for inhibitor design is the use of hot spots and hot regions.[15–17]

Certain residues, known as hot spots, contribute significantly more binding energy ($\geq$ 1.5 kcal mol^{-1}) to the binding interaction than surrounding residues.[15] These residues are often found in deeper pockets on the protein surface.[15–17] The pockets are surrounded by hydrophobic residues, which decrease the enthalpic penalty of the binding partner.[18,19] Areas on the PPI interface containing multiple hot spots are dubbed hot regions.[18,19] The number of hot regions on a PPI surfaces varies from protein to protein. Hot spots and hot regions provide a means to effectively shrink the amount of surface that a small molecule needs to cover in order to inhibit a PPI. Therefore, effectively using hot regions is essential to designing successful PPI inhibitors.

Because of poor site druggability and overlapping binding partners, directly targeting the binding site, known as the orthosteric site, may not be the best option. In this case, it may be possible to use another site on the surface of the protein to modulate the PPI.[20] This other site is known as an allosteric site. An allosteric site is one where the binding of a protein or small molecule affects binding at a second site. The exact mechanism of allostery differs between different proteins.[21–24] Several computational programs are able to predict potential allostery with good success. These are Statistical Coupling Analysis (SCA)[25] and Evolutionary Trace analysis (ET).[26] SCA was developed by the Ranganathan lab to determine energetically linked residues within proteins. A recent example of allostery discovered by SCA is the discovery of potential allosteric regulation in the dihydrofolate reductase enzyme, where no previously known allostery existed.[27] SCA has also been able to identify the potential network through which the

allosteric signal is conveyed, such as in kinesin.[28] ET, developed by the Lichtarge lab, achieved similar success at discovering latent allostery. However, the method for determining potentially allosteric sites is based on evolutionary conservation.[26] These scores correlate to the potential functional importance of the residue. Regions of potentially important residues on the surface of a protein with no known function may indicate latent allosteric regulation. This was recently shown with the discovery of allostery in dopamine receptor 2 using ET.[29] Regardless if an inhibitor is of an orthosteric or allosteric nature, it still faces the problems associated with the PPI surface. The following chapters focus on a specific example, β-catenin, that exemplifies the challenges associated with PPI inhibitor development.

1.2. The β-catenin/T-cell factor 4 interaction and the Wnt signaling pathway

β-catenin is an excellent example of a protein involved in PPI networks associated with disease.[8,30] β-catenin has roles in the androgen receptor (AR) ligand-dependent transcription pathway,[31] adherin cell adhesion network,[32] and the canonical Wingless/Integration complex subunit 1 (Wnt) signaling pathway (Figure 1.1).[30] β-catenin interacts with AR to increase ligand-dependent AR activity in prostate cancer cells.[31] The interactions between β-catenin, E-cadherin, and α-catenin are critical to link the extracellular components of a tight junction to the intracellular actin matrix.[32] In the canonical Wnt signaling pathway, β-catenin interacts with axin, adenomatous polyposis coli (APC), and T-cell factor 4 (Tcf4) proteins.[30] The β-catenin/APC and β-catenin/axin PPIs are responsible for sequestering excess cytosolic β-catenin for destruction.[34] Once

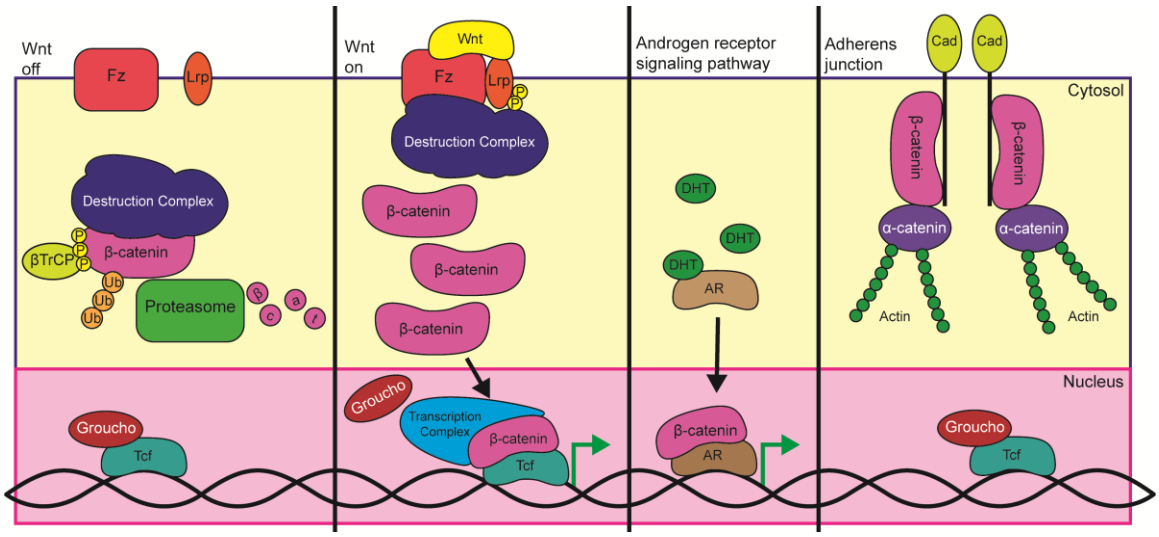


Figure 1.1. β-catenin is involved in multiple cellular pathways. In the canonical Wnt signaling pathway, β-catenin regulates transcription of Wnt target genes.[30] An excess of β-catenin can act as a co-activator of the androgen receptor in a ligand-dependent manner.[31] β-catenin is also a critical part of adherens junctions, acting as a link between cadherins and α-catenin.[32] Figure derived from work described by Clevers et al[31] and Yang et al.[32]

inside the nucleus, β-catenin displaces Groucho from Tcf4 to form a transcription complex.[33] Other proteins in the transcription complex that directly interact with β-catenin include B-cell lymphoma-9 (BCL9),[30,35] CREB-binding protein (CBP),[36] Brahma-related gene 1 (BRG1),[37] and pygopus (pygo).[38] This transcription complex is responsible for promoting the transcription of Wnt target genes such as *c-myc*,[39] *cyclin D1*,[40,41] and *survivin*.[42] The dysregulation of Wnt signaling has been implicated in a variety of cancers, fibroses, and neurological disorders.[30]

The canonical Wnt signaling pathway was first discovered as a cell polarity pathway in *Drosophila*,[43] and it was later determined to be involved in oncogenesis with the discovery of the highly conserved integration complex subunit 1 gene (*INT1*).[44] The Wnt signaling pathway begins when a Wnt effector protein binds the extra-cellular domains of Frizzled (Fz)[45] and lipoprotein-related receptor (LRP).[46] Axin is then sequestered away from the destruction complex by LRP, and the kinase ability of the destruction complex is ablated. This leaves excess cytosolic β-catenin free to be imported into the nucleus and form the transcription complex responsible for transcribing Wnt target genes.[30] In the absence of the Wnt effector protein, axin is phosphorylated and sequesters β-catenin in the destruction complex.[47] In the destruction complex, β-catenin is phosphorylated by glycogen synthase kinase 3 (GSK3), and is ultimately ubiquitinated and degraded by the proteasome.[30,48,49] Problems occur when the Wnt signaling pathway becomes activated without the presence of the Wnt effector protein.

The aberrant activation, or dysregulation, of the Wnt signaling pathway is often the result of genetic mutations affecting the various components of the pathway.[30] Multiple points exist where these mutations can occur. Mutations in axin[50,51] or APC[52,53]

are commonly found in colon cancers and disrupt function of the destruction complex. Mutations of β-catenin itself can also result in oncogenesis.[54] These mutations are found in the intrinsically disordered N-terminal tail where β-catenin is phosphorylated by GSK3.[48,49] The lack of phosphorylation prevents β-catenin from being degraded by the proteasome. Under these conditions, β-catenin accumulates in the cytosol and is imported into the nucleus. Regardless of the source, all of these mutations lead to the untimely formation of the β-catenin/Tcf4 transcription complex and transcription of Wnt target genes. Because the β-catenin/Tcf4 PPI is the most downstream component of the Wnt signaling pathway before transcription, targeting this PPI poses lower risk than disrupting the upstream Wnt signaling pathway. This makes the β-catenin/Tcf4 PPI a prime therapeutic target.

Unfortunately, the binding mode of APC, E-cadherin, and Tcf4 are highly overlapped and shown in Figure 1.2A.[52] This presents a significant challenge for the development of a selective inhibitor of the β-catenin/Tcf4 PPI. The β-catenin/Tcf4 PPI has a dissociation constant (K_d) of 7 nM.[55] This is lower than the β-catenin/APC and β-catenin/E-cadherin PPIs, which have K_d values of 0.6 μM and 41 nM, respectively.[55-57] Any benefit of disrupting the β-catenin/Tcf4 interaction would be rendered moot by the disruption of the β-catenin/APC interaction critical to the formation of the destruction complex.[58] The disruption of the β-catenin/E-cadherin interaction would interfere with cell adhesion and could result in tumor metastasis.[58] Therefore, any inhibitor designed to inhibit the β-catenin/Tcf4 PPI must be selective.

This may be accomplished by targeting specific β-catenin hot regions. The structure of β-catenin consists of three distinct domains. The N- and C-terminal domains

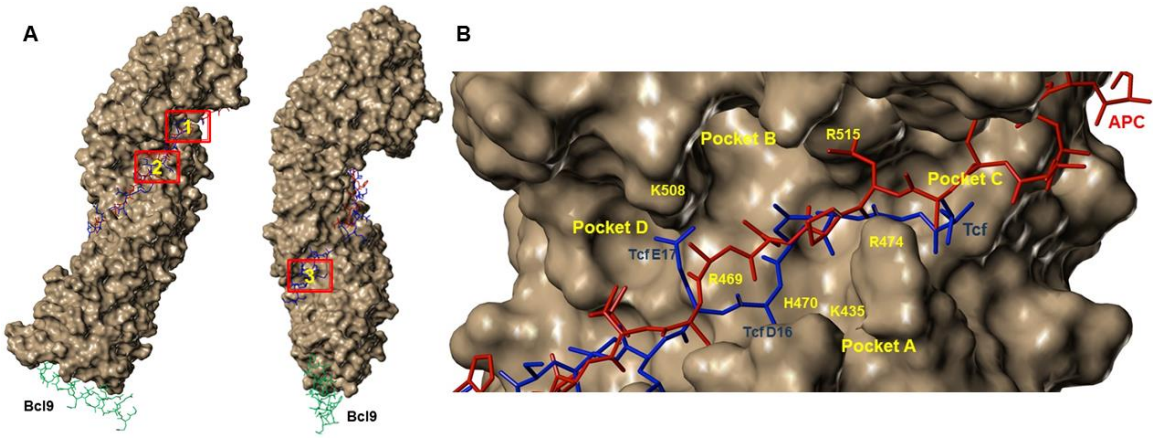


Figure 1.2. (A) Axin, APC, Tcf, and E-cadherin bind to a long groove on the armadillo repeat (ARM) domain of β-catenin.[8] The binding groove has three hot regions (boxed). Bcl9 binds a distinct site near the N-terminal end of the ARM domain. (B) Hot region 1, also known as the deep cleft, contains four pockets.[61] Binding partners APC (red, pdbID 1TH1),[11] Tcf (blue, pdbID 2GL7),[35] and E-cadherin (not shown) have very similar binding modes in this hot region. Image created from superimposition of backbone atoms of 1TH1 and 2GL7.

of β-catenin are intrinsically disordered. The middle domain consists of a repeating motif of two α-helices which form hairpin loops, known as an armadillo repeat (ARM). E-cadherin, APC, and Tcf4, though mutually exclusive in binding, share a long, positively-charged groove that spans the length of the ARM domain of β-catenin (Figure 1.2A).[8] Alanine scanning of the ARM domain revealed differences in the energetic contributions of residues to each interaction. K436A, R469A, and H470A mutants showed no transactivation of the Wnt-signaling pathway, while other nearby residue mutants, such as R457A, showed little-to-no reduction in Wnt signaling function.[60] The mutants that did not activate Wnt-signaling were found to be defective in the β-catenin/Tcf4 interaction.[60] The study also showed K435A, R469A, and H470A had a greater effect on Tcf4 binding than on APC binding.[60] Conversely, β-catenin mutants K345A and W383A were more disruptive to the β-catenin/APC interaction and less disruptive to the β-catenin/Tcf4 interaction.[60] The alanine scanning also revealed that the hot spot residues for binding Tcf4, APC, E-cadherin, and axin were clustered into three hot regions on the ARM domain of β-catenin. Hot region 1, also known as the deep cleft, is the binding site of the $G_{13}ANDE_{17}$ motif of Tcf4 (Figure 1.2B).[8,60] There are four main pockets in hot region 1, dubbed A, B, C, and D shown in Figure 1.2B.[61] Pocket A contains residues K435, a hot spot for the β-catenin/Tcf4, β-catenin/APC, and β-catenin/E-cadherin PPIs, and H470, a hot spot that is very selective for the β-catenin/Tcf4 PPI.[8] Pocket B is a deep hydrophobic pocket that is not utilized by any known binding partner. It is comprised of residues I507, V511, G512, L539, L535, I569, G572, and C573. Pocket C is a shallow, hydrophobic surface consisting of the arginine channel, R474 and R515, and the nearby hydrophobic surface. The final pocket, pocket D, contains hot spot K508, as well as E462, L506, I507,

A509, L539, and I569. The next hot region on the large, positively-charged groove of β-catenin is hot region 2.[8,60] The residues important for Tcf4 binding in hot region 2 include Y306, K312, K335, W338, R342, K345, Y354, and R376. The final hot region, hot region 3, includes key residues R212, R225, F253, H260, K270, R274, and R292.[8,61] Included in these hot regions are three residues, K292, K335, and R376 that interact with a pS/pT-motif on E-cadherin[10] and APC.[11,56] The phosphorylation of E-cadherin increases its affinity for β-catenin 1,000-fold.[10] Tcf4 does not undergo phosphorylation and relies more on hot region 1 than hot regions 2 or 3.[8] Because of the therapeutic importance of β-catenin, it has been the target of many drug design campaigns.[57] Some of these have been successful while others have been stifled by the problems common to PPIs.

1.3 Previously discovered Wnt inhibitors

Much work has gone into the discovery of inhibitors of the β-catenin/Tcf4 PPI (Figure 1.3). The first β-catenin/Tcf4 inhibitors were discovered by Lepourcelet et al.[62] by a high-throughput screen. These include compounds **1** (PKF115-584), **2** (CGP049090), and **3** (PKF118-380). These inhibitors obtained half maximal inhibitory concentration (IC_{50}) values of 3.2 μM, 0.8 μM, and 8.7 μM, respectively.[62] The binding mode of these inhibitors was not elucidated, and subsequent testing for selectivity of these compounds showed little difference in their ability to inhibit the β-catenin/Tcf4 interaction over the β-catenin/APC and β-catenin/E-cadherin interactions.[61] In 2006, **4** (PNU 74654) was discovered by Trosset, et al.[63] through high-throughput screening and obtained a K_d of 450 nM.[63] Three more inhibitors of the downstream canonical Wnt

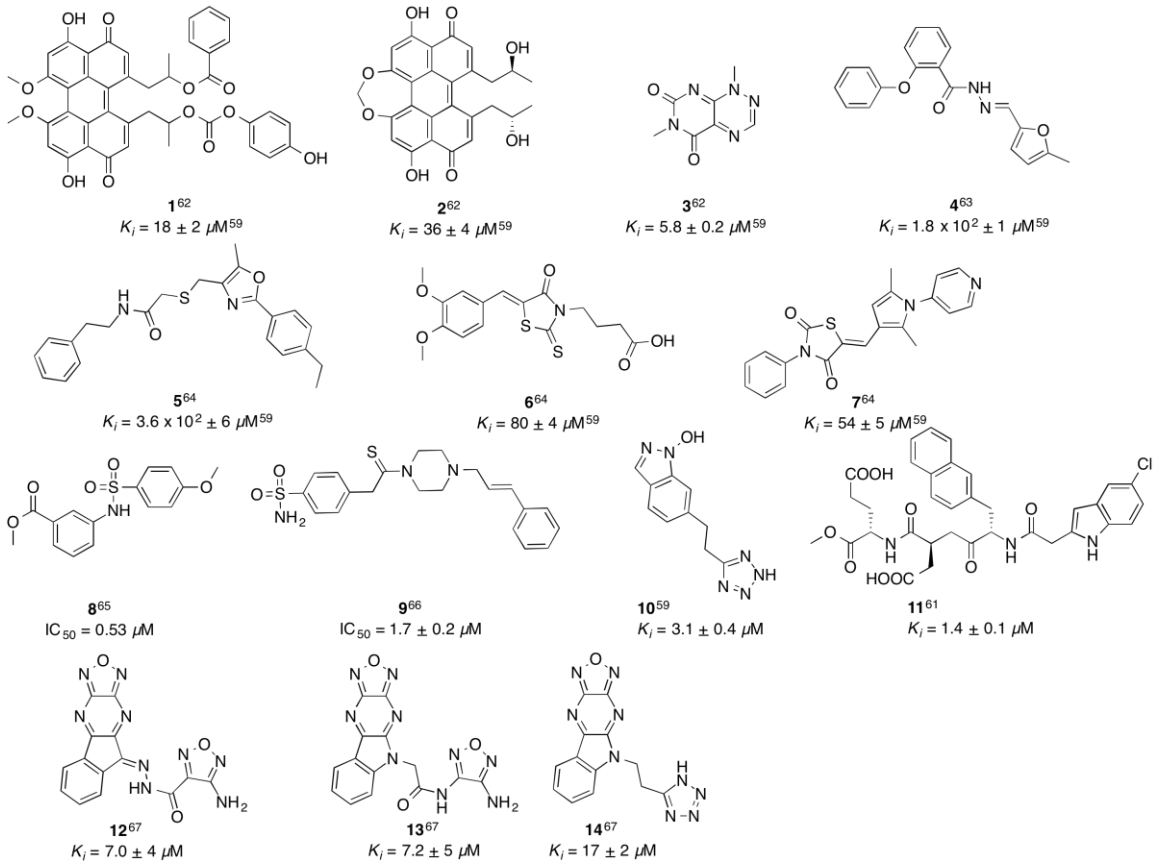


Figure 1.3. Many compounds have been developed to inhibit the *β*-catenin/Tcf4 PPI. In the case of PKF115-584, PKF115-584, CGP049090, PNU 74654, iCRT3, iCRT5, and iCRT14, the values indicate the results of a fluorescence polarization assay by Min Zhang to directly measure disruption of the *β*-catenin/Tcf interaction.[59] Adapted with permission from Catrow, J. L.; Zhang, Y.; Zhang, M.; Ji, H. Discovery of Selective Small-Molecule Inhibitors for the β-Catenin/T-Cell Factor Protein–Protein Interaction through the Optimization of the Acyl Hydrazone Moiety. *J. Med. Chem.* **2015**, *58* (11), 4678-4692. Copyright 2015 American Chemical Society.

signaling pathway, **5** (iCRT3), **6** (iCRT5), and **7** (iCRT14) were discovered by Gonsalves, et al.[64] These inhibitors were shown to disrupt the Wnt signaling pathway but not specifically the β-catenin/Tcf4 PPI. When assessed for the ability to inhibit the β-catenin/Tcf4 PPI, these inhibitors had weaker inhibitory constant (K_i) values of 364 μM, 80.3 μM, and 53.5 μM, respectively.[59] More recently, Hwang et al. identified **8** (methyl 3-{[(4-methyl-phenyl) sulfonyl]amino}benzoate) as a Wnt inhibitor with an IC_{50} of 530 nM that binds to the ARM domain of β-catenin.[65] Fang et al. developed **9** (LF3),[66] which inhibits the β-catenin/Tcf4 PPI with an IC_{50} of 1.7 μM and does not inhibit the β-catenin/E-cadherin PPI.

The Ji lab achieved success at developing inhibitors for the β-catenin/Tcf4 PPI.[59,61,67] The first potent inhibitor in the lab, **10**, was developed by postdoctoral fellow Dr. Binxun Yu.[59] Compound **10** was created through bioisostere replacement using human Tcf4 residues E17 and D16—these residues of Tcf4 form salt bridges with β-catenin hot spots K508 and K435, respectively.[59] Compound **10** achieved a K_i of 3.14 μM. Isothermal titration calorimetry studies using K435A, R469A, and K508A β-catenin mutants were used to validate the binding mode of **10**.[59] The results of this study indicated that **10** bound hot region 1. Next, Dr. Zheng Huang developed a peptidomimetic inhibitor **11** (UUT-02).[61] This compound was effective at inhibiting the β-catenin/Tcf4 PPI by mimicking the $G_{13}ANDE_{17}$ portion of Tcf4. Compound **11** marks a significant milestone in the development of β-catenin/Tcf4 PPI inhibitors. It is the first inhibitor to achieve selectivity for the β-catenin/Tcf4 PPI over the β-catenin/APC and β-catenin/E-cadherin PPIs.[61] The success of **11** puts hot region 1 in the spotlight as the means to selectively inhibit the β-catenin/Tcf4 interaction.

Herein, some of the commonly encountered problems of PPI inhibitor design are tackled. First, common challenges associated with small-molecule inhibitor design are addressed through the optimization of **12** (UZI/7116003).[67] Compound **12** was identified by Dr. Min Zhang as an inhibitor of the β-catenin/Tcf4 PPI, but this small molecule contained the acyl hydrazone PAIN substructure.[67] From this inhibitor, a new scaffold, **13**, was developed by Dr. Yongqiang Zhang and optimized to interact with β-catenin hot spot K435, resulting in **14**.[67] This new series of inhibitors were specific for the β-catenin/Tcf4 PPI over the β-catenin/E-cadherin and β-catenin/APC PPIs, and showed Wnt inhibitory activity both in vitro and in vivo.[67]

Furthermore, the abilities of common high-throughput virtual screening (HTVS) programs to identify potent PPI inhibitors are examined. Past discovery of PPI inhibitors by HTVS has proved challenging, and it has been noted that current software may be ill-suited for HTVS of PPIs.[2] Herein, the effects of different docking algorithms and scoring functions were assessed for PPIs, and it was found that neither program proved better at handling HTVS of PPI inhibitors. In order to enhance the HTVS results, a physics-based re-ranking based on molecular mechanics/generalized Borne solvent accessible (MM/GBSA) binding energy calculations. The effect of ligand burial[2] on inhibitor potency was also examined, and a weak, but significant, correlation was observed, indicating that a more buried inhibitor was more potent.

Lastly, ET analysis, solvent mapping and molecular dynamic (MD) simulations were applied to β-catenin to explore ligand selectivity and identify a novel allosteric site. Solvent mapping analysis of different β-catenin crystal structures with binding partners confirms that Tcf4 relies strongly on hot region 1 for binding, while E-cadherin, APC and

axin rely comparatively more on hot regions 2 and 3. ET sequence analysis was used to discover a novel, potentially allosteric region on the surface of β-catenin. This site does not interact with Tcf4, APC, or E-cadherin, and the disruption of this site increases β-catenin/Tcf4 binding. The discovery of this site will prove useful for small-molecule allosteric regulation of the β-catenin/Tcf4 PPI.

1.4 References

1. Wells, J. A.; McClendon, C. L. Reaching for High-Hanging Fruit in Drug Discovery at Protein–Protein Interfaces. *Nature* **2007,** *450* (7172), 1001–1009.

2. Gowthaman, R.; Deeds, E. J.; Karanicolas, J. Structural Properties of Non-Traditional Drug Targets Present New Challenges for Virtual Screening. *J. Chem. Inf. Model.* **2013,** *53* (8), 2073–2081.

3. Makley, L. N.; Gestwicki, J. E. Expanding the Number of 'Druggable' Targets: Non-Enzymes and Protein–Protein Interactions. *Chem. Biol. Drug. Des*. **2013**, *81* (1), 22–32.

4. Xu, D.; Jalal, S. I.; Sledge, G. W.; Meroueh, S. O. Small-Molecule Binding Sites to Explore Protein–Protein Interactions in the Cancer Proteome. *Mol. Biosyst.* **2016,** *12* (10), 3067–3087.

5. Kuntz, I. D.; Chen, K.; Sharp, K. A.; Kollman, P. A. The Maximal Affinity of Ligands. *Proc. Natl. Acad. Sci. U.S.A.* **1999,** *96* (18), 9997–10002.

6. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings. *Adv. Drug. Deliv. Rev.* **2001,** *46* (1–3), 3–26.

7. Baell, J. B.; Holloway, G. A. New Substructure Filters for Removal of Pan Assay Interference Compounds (PAINS) from Screening Libraries and for Their Exclusion in Bioassays. *J. Med. Chem.* **2010,** *53* (7), 2719–2740.

8. Xu, W.; Kimelman, D. Mechanistic Insights from Structural Studies of β-Catenin and its Binding Partners. *J. Cell. Sci.* **2007,** *120* (19), 3337–3344.

9. Poy, F.; Lepourcelet, M.; Shivdasani, R. A.; Eck, M. J. Structure of a Human Tcf4-β-Catenin Complex. *Nat. Struct. Biol.* **2001,** *8* (12), 1053–1057.

10. Huber, A. H.; Weis, W. I. The Structure of the β-Catenin/E-Cadherin Complex and the Molecular Basis of Diverse Ligand Recognition by β-Catenin. *Cell* **2001,** *105* (3), 391–402.

11. Xing, Y.; Clements, W. K.; Le Trong, I.; Hinds, T. R.; Stenkamp, R.; Kimelman, D.; Xu, W. Crystal Structure of a β-Catenin/APC Complex Reveals a Critical Role for APC Phosphorylation in APC Function. *Mol. Cell* **2004,** *15* (4), 523–533.

12. Xing, Y.; Clements, W. K.; Kimelman, D.; Xu, W. Crystal Structure of a β-Catenin/Axin Complex Suggests a Mechanism for the β-Catenin Destruction Complex. *Genes Dev.* **2003,** *17* (22), 2753–2764.

13. Russo, A.; Aiello, C.; Grieco, P.; Marasco, D. Targeting "Undruggable" Proteins: Design of Synthetic Cyclopeptides. *Curr. Med. Chem.* **2016,** *23* (8), 748–762.

14. Iyer, V. V. A Review of Stapled Peptides and Small Molecules to Inhibit Protein–Protein Interactions in Cancer. *Curr. Med. Chem.* **2016,** *23* (27), 3025–3043.

15. Clackson, T.; Wells, J. A. A Hot Spot of Binding Energy in a Hormone-Receptor Interface. *Science* **1995**, *267 (5196)*, 383–386.

16. Bogan, A. A.; Thorn, K. S. Anatomy of Hot Spots in Protein Interfaces. *J. Mol. Biol.* **1998**, *280 (1)*, 1–9.

17. Ma, B.; Elkayam, T.; Wolfson, H.; Nussinov, R. Protein–Protein Interactions: Structurally Conserved Residues Distinguish Between Binding Sites and Exposed Protein Surfaces. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100 (10)*, 5772–5777.

18. Keskin, O.; Ma, B.; Nussinov, R. Hot Regions in Protein–Protein Interactions: the Organization and Contribution of Structurally Conserved Hot Spot Residues. *J. Mol. Biol*. **2005**, *345 (5)*, 1281–1294.

19. Tuncbag, N.; Gursoy, A.; Keskin, O. Identification of Computational Hot Spots in Protein Interfaces: Combining Solvent Accessibility and Inter-Residue Potentials Improves the Accuracy. *Bioinformatics* **2009**, *25 (12)*, 1513–1520.

20. Cossins, B. P.; Lawson, A. D. Small Molecule Targeting of Protein–Protein Interactions Through Allosteric Modulation of Dynamics. *Molecules* **2015**, *20 (9)*, 16435–1645.

21. Schueler-Furman, O.; Wodak, S. J. Computational Approaches to Investigating Allostery. *Curr. Opin. Struct. Biol.* **2016**, *41*, 159–171.

22. Monod, J.; Wyman, J.; Changeux, J. P. On the Nature of Allosteric Transitions: a Plausible Model. *J. Mol. Biol.* **1965**, *12*, 88–118.

23. Koshland, D.E. Jr.; Némethy, G.; Filmer D. Comparison of Experimental Binding Data and Theoretical Models in Proteins Containing Subunits. *Biochemistry*. **1966**, *5* (1), 365–368

24. Jaffe, E. K. Morpheeins. A New Structural Paradigm for Allosteric Regulation. *Trends Biochem. Sci.* **2005**, *30 (9)*, 490–497.

25. Lockless, S. W.; Ranganathan, R. Evolutionarily Conserved Pathways of Energetic Connectivity in Protein Families. *Science* **1999,** *286* (5438), 295–299.

26. Lichtarge, O.; Bourne, H. R.; Cohen, F. E. An Evolutionary Trace Method Defines Binding Surfaces Common to Protein Families. *J. Mol. Biol.* **1996,** *257* (2), 342–358.

27. Reynolds, K. A.; McLaughlin, R. N.; Ranganathan, R. Hot Spots for Allosteric Regulation on Protein Surfaces. *Cell* **2011,** *147* (7), 1564–1575.

28. Richard, J.; Kim, E. D.; Nguyen, H.; Kim, C. D.; Kim, S. Allostery Wiring Map for Kinesin Energy Transduction and its Evolution. *J. Biol. Chem.* **2016,** *291* (40), 20932–20945.

29. Sung, Y. M.; Wilkins, A. D.; Rodriguez, G. J.; Wensel, T. G.; Lichtarge, O. Intramolecular Allosteric Communication in Dopamine D2 Receptor Revealed by Evolutionary Amino Acid Covariation. *Proc. Natl. Acad. Sci. U.S.A.* **2016,** *113* (13), 3539–3544.

30. Clevers, H.; Nusse, R. Wnt/β-Catenin Signaling and Disease. *Cell* **2012,** *149* (6), 1192–1205.

31. Yang, F.; Li, X.; Sharma, M.; Sasaki, C. Y.; Longo, D. L.; Lim, B.; Sun, Z. Linking β-Catenin to Androgen-Signaling Pathway. *J. Biol. Chem.* **2002,** *277* (13), 11336–11344.

32. Yap, A. S.; Brieher, W. M.; Gumbiner, B. M. Molecular and Functional Analysis of Cadherin-Based Adherens Junctions. *Annu. Rev. Cell. Dev. Biol.* **1997,** *13*, 119–146.

33. Li, V. S.; Ng, S. S.; Boersema, P. J.; Low, T. Y.; Karthaus, W. R.; Gerlach, J. P.; Mohammed, S.; Heck, A. J.; Maurice, M. M.; Mahmoudi, T.; Clevers, H. Wnt Signaling Through Inhibition of β-Catenin Degradation in an Intact Axin1 Complex. *Cell* **2012,** *149* (6), 1245–1256.

34. Daniels, D. L.; Weis, W. I. β-Catenin Directly Displaces Groucho/TLE Repressors from Tcf/Lef in Wnt-Mediated Transcription Activation. *Nat. Struct. Mol. Biol.* **2005,** *12* (4), 364–371.

35. Sampietro, J.; Dahlberg, C. L.; Cho, U. S.; Hinds, T. R.; Kimelman, D.; Xu, W. Crystal Structure of a β-Catenin/BCL9/Tcf4 Complex. *Mol. Cell.* **2006,** *24* (2), 293–300.

36. Takemaru, K. I.; Moon, R. T. The Transcriptional Coactivator CBP Interacts with β-Catenin to Activate Gene Expression. *J. Cell. Biol.* **2000,** *149* (2), 249–254.

37. Stadeli, R.; Hoffmans, R.; Basler, K. Transcription under the Control of Nuclear Arm/β-Catenin. *Curr. Biol.* **2006,** *16* (10), R378–385.

38. Parker, D. S.; Jemison, J.; Cadigan, K. M. Pygopus, a Nuclear PHD-Finger Protein Required for Wingless Signaling in Drosophila. *Development* **2002,** *129* (11), 2565–2576.

39. He, T. C.; Sparks, A. B.; Rago, C.; Hermeking, H.; Zawel, L.; da Costa, L. T.; Morin, P. J.; Vogelstein, B.; Kinzler, K. W. Identification of C-MYC as a Target of the APC Pathway. *Science* **1998,** *281* (5382), 1509–1512.

40. Tetsu, O.; McCormick, F. β-Catenin Regulates Expression of Cyclin D1 in Colon Carcinoma Cells. *Nature* **1999,** *398* (6726), 422–426.

41. Shtutman, M.; Zhurinsky, J.; Simcha, I.; Albanese, C.; D'Amico, M.; Pestell, R.; Ben-Ze'ev, A. The Cyclin D1 Gene Is a Target of the β-Catenin/LEF-1 Pathway. *Proc. Natl. Acad. Sci. U.S.A.* **1999,** *96* (10), 5522–5527.

42. Zhang, T.; Otevrel, T.; Gao, Z.; Gao, Z.; Ehrlich, S. M.; Fields, J. Z.; Boman, B. M. Evidence that APC Regulates Survivin Expression: A Possible Mechanism Contributing to the Stem Cell Origin of Colon Cancer. *Cancer Res.* **2001,** *61* (24), 8664–8667.

43. Nusslein-Volhard, C.; Wieschaus, E. Mutations Affecting Segment Number and Polarity in Drosophila. *Nature* **1980,** *287* (5785), 795–801.

44. Nusse, R.; Varmus, H. E. Many Tumors Induced by the Mouse Mammary Tumor Virus Contain a Provirus Integrated in the Same Region of the Host Genome. *Cell* **1982,** *31* (1), 99–109.

45. Bhanot, P.; Brink, M.; Samos, C. H.; Hsieh, J. C.; Wang, Y.; Macke, J. P.; Andrew, D.; Nathans, J.; Nusse, R. A New Member of the Frizzled Family from Drosophila Functions as a Wingless Receptor. *Nature* **1996,** *382* (6588), 225–230.

46. Wehrli, M.; Dougan, S. T.; Caldwell, K.; O'Keefe, L.; Schwartz, S.; Vaizel-Ohayon, D.; Schejter, E.; Tomlinson, A.; DiNardo, S. Arrow Encodes an LDL-Receptor-Related Protein Essential for Wingless Signaling. *Nature* **2000,** *407* (6803), 527–530.

47. Mao, J.; Wang, J.; Liu, B.; Pan, W.; Farr, G. H.; Flynn, C.; Yuan, H.; Takada, S.; Kimelman, D.; Li, L.; Wu, D. Low-Density Lipoprotein Receptor-Related Protein-5 Binds to Axin and Regulates the Canonical Wnt Signaling Pathway. *Mol. Cell* **2001,** *7* (4), 801–809.

48. Siegfried, E.; Chou, T. B.; Perrimon, N. Wingless Signaling Acts Through Zeste-White 3, the Drosophila Homolog of Glycogen Synthase Kinase-3, to Regulate Engrailed and Establish Cell Fate. *Cell* **1992,** *71* (7), 1167–1179.

49. Peifer, M.; Sweeton, D.; Casey, M.; Wieschaus, E. Wingless Signal and Zeste-White 3 Kinase Trigger Opposing Changes in the Intracellular Distribution of Armadillo. *Development* **1994,** *120* (2), 369–380.

50. Satoh, S.; Daigo, Y.; Furukawa, Y.; Kato, T.; Miwa, N.; Nishiwaki, T.; Kawasoe, T.; Ishiguro, H.; Fujita, M.; Tokino, T.; Sasaki, Y.; Imaoka, S.; Murata, M.; Shimano, T.; Yamaoka, Y.; Nakamura, Y. AXIN1 Mutations in Hepatocellular Carcinomas, and Growth Suppression in Cancer Cells by Virus-Mediated Transfer of AXIN1. *Nat. Genet.* **2000,** *24* (3), 245–250.

51. Lammi, L.; Arte, S.; Somer, M.; Jarvinen, H.; Lahermo, P.; Thesleff, I.; Pirinen, S.; Nieminen, P. Mutations in AXIN2 Cause Familial Tooth Agenesis and Predispose to Colorectal Cancer. *Am. J. Hum. Genet.* **2004,** *74* (5), 1043–1050.

52. Kinzler, K. W.; Nilbert, M. C.; Su, L. K.; Vogelstein, B.; Bryan, T. M.; Levy, D. B.; Smith, K. J.; Preisinger, A. C.; Hedge, P.; McKechnie, D. Identification of FAP Locus Genes from Chromosome 5q21. *Science* **1991,** *253* (5020), 661–665.

53. Nishisho, I.; Nakamura, Y.; Miyoshi, Y.; Miki, Y.; Ando, H.; Horii, A.; Koyama, K.; Utsunomiya, J.; Baba, S.; Hedge, P. Mutations of Chromosome 5q21 Genes in FAP and Colorectal Cancer Patients. *Science* **1991,** *253* (5020), 665–669.

54. Morin, P. J.; Sparks, A. B.; Korinek, V.; Barker, N.; Clevers, H.; Vogelstein, B.; Kinzler, K. W. Activation of β-Catenin-Tcf Signaling in Colon Cancer by Mutations in β-Catenin or APC. *Science* **1997,** *275* (5307), 1787–1790.

55. Sun, J.; Weis, W. I. Biochemical and Structural Characterization of β-Catenin Interactions with Nonphosphorylated and CK2-Phosphorylated Lef-1. *J. Mol. Biol.* **2011,** *405* (2), 519–530.

56. Ha, N. C.; Tonozuka, T.; Stamos, J. L.; Choi, H. J.; Weis, W. I. Mechanism of Phosphorylation-Dependent Binding of APC to β-Catenin and its Role in B-Catenin Degradation. *Mol. Cell* **2004,** *15* (4), 511–521.

57. Choi, H. J.; Huber, A. H.; Weis, W. I. Thermodynamics of β-Catenin-Ligand Interactions: the Roles of the N- And C-Terminal Tails in Modulating Binding Affinity. *J. Biol. Chem.* **2006,** *281* (2), 1027–1038.

58. Garber, K. Drugging the Wnt Pathway: Problems and Progress. *J. Natl. Cancer Inst.* **2009,** *101* (8), 548–550.

59. Yu, B.; Huang, Z.; Zhang, M.; Dillard, D. R.; Ji, H. Rational Design of Small-Molecule Inhibitors for β-Catenin/T-Cell Factor Protein–Protein Interactions by Bioisostere Replacement. *ACS Chem. Biol.* **2013,** *8* (3), 524–529.

60. von Kries, J. P.; Winbeck, G.; Asbrand, C.; Schwarz-Romond, T.; Sochnikova, N.; Dell'Oro, A.; Behrens, J.; Birchmeier, W. Hot Spots in β-Catenin for Interactions with LEF-1, Conductin and APC. *Nat. Struct. Biol.* **2000,** *7* (9), 800–807.

61. Huang, Z.; Zhang, M.; Burton, S. D.; Katsakhyan, L. N.; Ji, H. Targeting the Tcf4 G13ANDE17 Binding Site to Selectively Disrupt β-Catenin/T-Cell Factor Protein–Protein Interactions. *ACS Chem. Biol.* **2014,** *9* (1), 193–201.

62. Lepourcelet, M.; Chen, Y. N.; France, D. S.; Wang, H.; Crews, P.; Petersen, F.; Bruseo, C.; Wood, A. W.; Shivdasani, R. A. Small-Molecule Antagonists of the Oncogenic Tcf/β-Catenin Protein Complex. *Cancer Cell* **2004,** *5* (1), 91–102.

63. Trosset, J. Y.; Dalvit, C.; Knapp, S.; Fasolini, M.; Veronesi, M.; Mantegani, S.; Gianellini, L. M.; Catana, C.; Sundstrom, M.; Stouten, P. F.; Moll, J. K. Inhibition of Protein–Protein Interactions: The Discovery of Druglike β-Catenin Inhibitors by Combining Virtual and Biophysical Screening. *Proteins* **2006,** *64* (1), 60–67.

64. Gonsalves, F. C.; Klein, K.; Carson, B. B.; Katz, S.; Ekas, L. A.; Evans, S.; Nagourney, R.; Cardozo, T.; Brown, A. M.; DasGupta, R. An RNAi-Based Chemical Genetic Screen Identifies Three Small-Molecule Inhibitors of the Wnt/Wingless Signaling Pathway. *Proc. Natl. Acad. Sci U.S.A.* **2011,** *108* (15), 5954–5963.

65. Hwang, S. Y.; Deng, X.; Byun, S.; Lee, C.; Lee, S. J.; Suh, H.; Zhang, J.; Kang, Q.; Zhang, T.; Westover, K. D.; Mandinova, A.; Lee, S. W. Direct Targeting of β-Catenin by a Small Molecule Stimulates Proteasomal Degradation and Suppresses Oncogenic Wnt/β-Catenin Signaling. *Cell Rep.* **2016,** *16* (1), 28–36.

66. Fang, L.; Zhu, Q.; Neuenschwander, M.; Specker, E.; Wulf-Goldenberg, A.; Weis, W. I.; von Kries, J. P.; Birchmeier, W. A Small-Molecule Antagonist of the β-Catenin/TCF4 Interaction Blocks the Self-Renewal of Cancer Stem Cells and Suppresses Tumorigenesis. *Cancer Res.* **2016,** *76* (4), 891–901.

67. Catrow, J. L.; Zhang, Y.; Zhang, M.; Ji, H. Discovery of Selective Small-Molecule Inhibitors for the β-Catenin/T-Cell Factor Protein–Protein Interaction Through the Optimization of The Acyl Hydrazone Moiety. *J. Med. Chem.* **2015,** *58* (11), 4678–4692.

CHAPTER 2

DISCOVERY AND OPTIMIZATION OF SELECTIVE SMALL-MOLECULE INHIBITORS FOR THE β-CATENIN/T-CELL FACTOR 4 PROTEIN–PROTEIN INTERACTION

Adapted with permission from Catrow, J. L.; Zhang, Y.; Zhang, M.; Ji, H. Discovery of Selective Small-Molecule Inhibitors for the β-Catenin/T-Cell Factor Protein–Protein Interaction through the Optimization of the Acyl Hydrazone Moiety. *J. Med. Chem.* **2015**, *58* (11), 4678-4692.

Authors Note: This chapter contains material originally published from a collaboration of Jonathan Leon Catrow, Dr. Yongiang Zhang, and Dr. Min Zhang. The text has been rewritten and reorganized to include further work performed on this series of inhibitors. Figures 2.1–2.11 are adapted with permission from the original manuscript.

2.1 Introduction

The canonical Wnt signaling pathway is critical for regulating cell growth, differentiation, and migration.[1,2] The hyperactivation of the Wnt signaling pathway is

commonly found in cancer stem cells, which are resistant to traditional therapies.[3-5] Mutations in axin, APC, and β-catenin, which represent the key regulatory components of the Wnt signaling pathway, result in the untimely expression of Wnt target genes including *c-myc*,[6] *cyclin D1*,[7] and *survinin*.[8] The penultimate step of the Wnt signaling pathway is the formation of a transcription complex initiated by the β-catenin/Tcf4 PPI.[1] Therefore, the β-catenin/Tcf4 PPI is an appealing target to inhibit the canonical Wnt signaling pathway.[9]

The crystal structure of β-catenin in complex with human Tcf4 (PDB ID 2GL7)[10] shows that the interacting surface between β-catenin and Tcf4 is approximately 3,500 Å^2, while most PPI surfaces are between 1,200-2,500 Å^2.[11] The crystallographic binding mode of human Tcf4 reveals that most interactions with β-catenin occur in three hot regions on the β-catenin ARM domain.[10,12-18] Hot region 1 contains β-catenin hot spots K435 and K508, which form charge-charge interactions with human Tcf4 residues D16 and E17, respectively. Disruption of the Tcf4 D16/β-catenin K435 interaction reduced the β-catenin/Tcf4 interaction by 1000-fold.[14,18] Previous surface plasmon resonance (SPR) studies indicated that these two interactions in hot region 1 were more important than interactions in other hot regions.[16] In hot region 2, β-catenin K312 and K345 form charge-charge interactions with Tcf4 E24 and E29, respectively.[10-13] Hot region 3 contains the hydrophobic interactions of Tcf4 V44 and L48 with β-catenin F253, I256, F293, A295, and A296.[19] In addition to Tcfs, these hot regions also bind E-cadherin (human E-cadherin residues 819-876, PDB ID 1I7X)[20] and the third 20 amino acid repeat of APC (human APC residues 1477-1519, PDB ID 1TH1),[21] and axin (human axin residues 466-482, PDB ID 1QZ7).[19] Examination of the crystallographic binding modes

of these proteins indicated that Tcf4, E-cadherin, and APC interacted with all three hot regions, while axin only interacted with hot region 3.[22] The β-catenin/APC and β-catenin/axin interactions are crucial for regulating β-catenin degradation through the destruction complex.[19,21] The β-catenin/E-cadherin interaction is a component of cell-cell adhesion, and the disruption of this PPI could disrupt adhesion and result in tumor metastasis.[20] Therefore, it is critical that an inhibitor that binds β-catenin be selective for the β-catenin/Tcf4 PPI over the β-catenin/E-cadherin and β-catenin/APC PPIs. The β-catenin/Tcf4 interaction is the strongest of the three binding partners (K_d = 7 nM),[15,23,24] while the β-catenin/E-cadherin (K_d = 41 nM)[23,25,26] and β-catenin/APC (K_d = 0.6 μM)[21,23,25,27,28] are weaker. The tight interaction between β-catenin/Tcf4 complicates inhibitor design.

Several successful β-catenin/Tcf4 inhibitors have been reported, most of which were not found to be selective (Figure 2.1).[29] A high-throughput screening identified six natural products out of 52,000 compounds, **1**, **2**, **3**, **15** (PKF118-744), **16** (ZTM000990), and **17** (PKF222-815) as inhibitors of the β-catenin/Tcf4 PPI.[29] Despite achieving good potency, compounds **1, 3, 15,** and **16** contain quinine-like PAIN substructure,[30] while **2** contains the toxoflavin PAIN substructure.[30,31] These PAIN substructures are associated with a high rate of false-positive hits and toxicity.[32] Compound **4** was identified through virtual and biophysical screening and contained the acyl hydrazone PAIN substructure.[30] No further study or optimization was reported. The iCRT series of inhibitors, **5**, **6,** and **7**, disrupted axin knockdown-induced Wnt signaling, downregulated Wnt target genes, and inhibited the growth of colorectal cancer cells.[33] However, **5** was reported to undesirably interfere with androgen signaling,[34] and rhodanine-like PAIN substructures were present

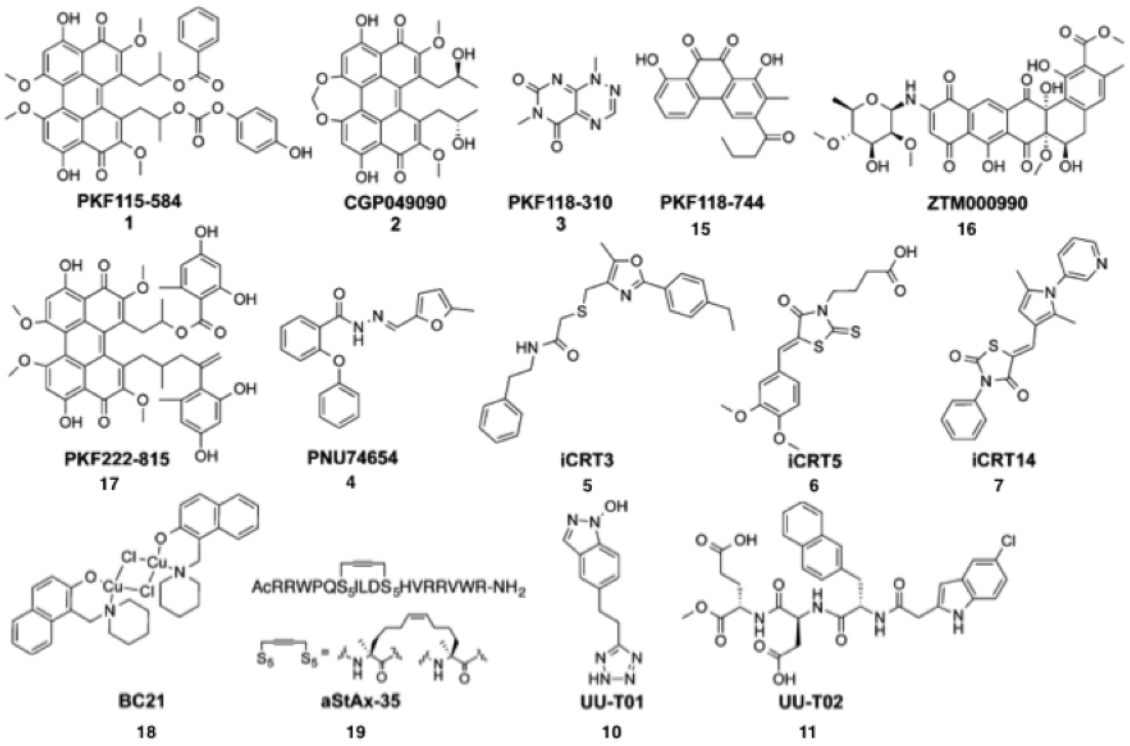


Figure 2.1. Previous β-catenin/Tcf4 PPI inhibitors. Compounds **1**, **2**, **15,** and **16** contain quinine-like PAIN substructure.[30] Compound **3** contains the toxoflavin PAIN substructure. Compound **4** contains the acyl hydrazone PAIN substructure. Compounds **6** and **7** contain a rhodanine-like PAIN substructure. Adapted with permission from Catrow, J. L.; Zhang, Y.; Zhang, M.; Ji, H. Discovery of Selective Small-Molecule Inhibitors for the β-Catenin/T-Cell Factor Protein–Protein Interaction through the Optimization of the Acyl Hydrazone Moiety. *J. Med. Chem.* **2015**, *58* (11), 4678-4692. Copyright 2015 American Chemical Society.

in **6** and **7**, which may lead to nonspecific binding and off-target effects.[30,35,36] A virtual screen of the Tcf4 $D_{16}ELISF_{21}$ binding site on β-catenin found that **18** (BC21) was a potent inhibitor of the β-catenin/Tcf4 interaction with cell-based activity,[37] but no further study of the selectivity of **18** has been reported. Compound **19** (aStAx-35) is a stapled *α*-helix peptidomimetic compound based on axin residues S471-V480 and was designed to inhibit the β-catenin/Tcf4 interaction by binding hot region 3.[38] However, it may also disrupt the β-catenin/axin PPI, which is responsible for regulating the degradation of β-catenin through the destruction complex.[19,39] Compound **10** (UUT-01) was developed by postdoctoral fellow Dr. Binxun Yu to mimic the binding mode of Tcf4 D16 and E17 to β-catenin K435 and K508, respectively.[16] Using the Tcf4 $G_{13}ANDE_{17}$ structure as a guide, Dr. Zheng Huang developed a peptidomimetic inhibitor, **11** (UUT-02), of the β-catenin/Tcf4 PPI.[39] This compound was not only a potent β-catenin/Tcf4 inhibitor, but it was selective for the β-catenin/Tcf4 PPI over the β-catenin/E-cadherin and β-catenin/APC PPIs.[39] As such, **11** served as a proof of concept that the $G_{13}ANDE_{17}$ binding site of hot region 1 could be used for selective inhibition of the β-catenin/Tcf4 PPI.[39]

This chapter presents an evaluation and optimization of **12**, a β-catenin/Tcf4 inhibitor shown to bind the hot region 1 of β-catenin. Compound **12**, which was discovered by high-throughput screening, inhibited the β-catenin/Tcf4 interaction, disrupted transactivation of the Wnt signaling pathway and inhibited cancer cell growth. However, this compound contained the acyl hydrazone PAIN substructure.[30] Despite problems reported with the acyl hydrazone structure, it is highly represented in reported small molecule inhibitors. A literature search for biological activity of the acyl hydrazone

substructure resulted in 11,490 publications between 1993 and 2014 (Figure 2.2). The acyl hydrazone moiety is an important component in dynamic combinatorial chemistry (DCC).[40-42] The formation of the acyl hydrazone substructure is important for DCC because it is reversible, yet it is stable under physiological pH.[42] DCC studies using the formation of the acyl hydrazone moiety resulted in inhibitors of acetylcholinesterase,[43] aspartic proteases,[44] β-tryptase,[45] and *Bacillus subtilis* histidine-containing phosphocarrier protein kinase/phosphatase (HPrK).[46] The acyl hydrazone substructure is also present in two FDA approved drugs, the antibiotic nitrofurantoin and the calcium sensitizer levosimendan. Therefore, a need exists to better understand the biological activities of acyl hydrazone-containing inhibitors as well as to develop efficient means of PAIN substructure removal and optimization of acyl hydrazone-containing compounds.

2.2 Methods

2.2.1. Protein expression and purification. Wild-type (WT) or mutant β-catenin complimentary DNA (cDNA) was cloned into either a Novagen pET28b vector or a pEHISTEV vector (provided by Dr. Hanting Liu of St. Andrew University, UK) resulting in a C-terminal 6X-His tag or N-terminal 6X-His tag fusion, respectively. Plasmids were transformed into Novagen BL21 DE *E. coli* cells and cultured in LB media containing 30 μg/mL kanamycin at 37 °C until the OD_{600} reached approximately 0.8. At this time, β-catenin expression was induced by 400 μM of isopropyl β-D-1-thiogalactopyranoside (IPTG). Proteins were expressed at 20 °C overnight. To extract proteins, cells were sonically lysed, and the protein was purified using nickel-nitriloaceticacid (Ni-NTA) affinity chromatography (30210, Qiagen). Purified protein was dialyzed against a buffer

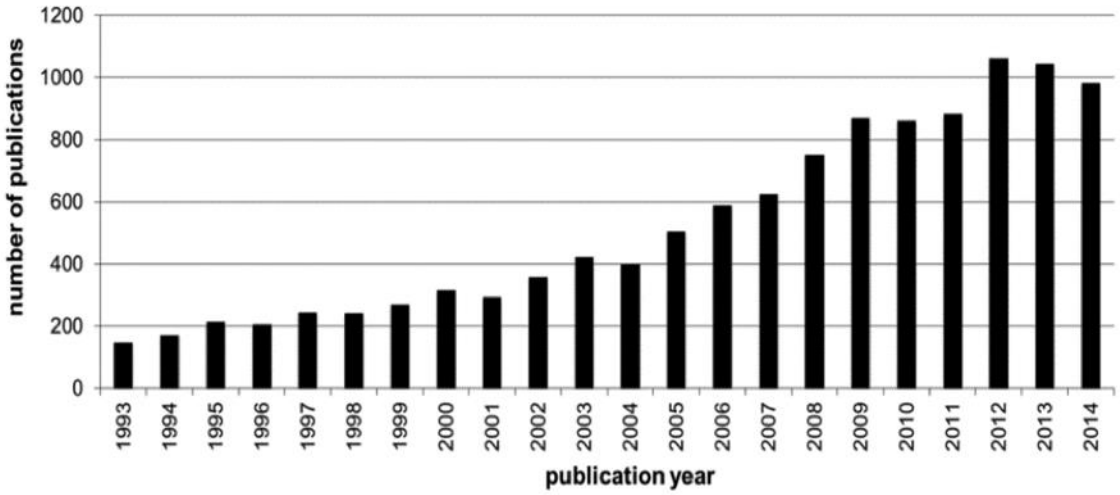


Figure 2.2. Publications involving biological activities of the acyl hydrazone moiety. Results of a SciFinder search result for the publications that include the biological activities of the acyl hydrazone sructure between 1993–2014. Adapted with permission from Catrow, J. L.; Zhang, Y.; Zhang, M.; Ji, H. Discovery of Selective Small-Molecule Inhibitors for the β-Catenin/T-Cell Factor Protein–Protein Interaction through the Optimization of the Acyl Hydrazone Moiety. *J. Med. Chem.* **2015**, *58* (11), 4678-4692. Copyright 2015 American Chemical Society.

of 20 mM Tris, 100 mM NaCl, 10% glycerol, and 3 mM dithiothreitol (DTT) at pH 8.8 overnight at 4 °C. β-catenin was isolated at 95% purity as determined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis. A thermal shift assay to evaluate protein stability, unfolding, and aggregation was performed using a BioRad iCycler iQ real-time detection system, Sypro Orange dye, and a temperature gradient of 1 °C/min. Circular dichroism (CD) spectroscopy was used to evaluate mutant protein folding. A Jasco J-85 spectropolarimeter was used to record CD spectra with a 1 mm path-length quartz cell at room temperature (rt). CD spectra scans were conducted from 90 nm to 260 nm in 1 nm increments. Three CD spectra scans were averaged together. Spectra were base-line corrected using a blank sample containing everything except protein. Samples of 1–5 μM protein were prepared in a buffer of 10 mM potassium phosphate and 100 mM potassium fluoride and adjusted to pH 7. Protein was stored in 100 μL aliquots at -80 °C. High performance liquid chromatography (HPLC) purified (> 95% purity) C-terminally biotinylated human Tcf4 (residues 7-51), C-terminally fluorescein-labeled human Tcf4 (residues 7-51), C-terminally fluoroscein-labeled human E-cadherin (residues 819-873) and C-terminally fluoroscein-labeled human APC (residues 1477-1519) were synthesized by InnoPep, Inc. The synthesized peptide structures were validated using liquid chromatography/mass spectrometry.

2.2.1 Fluorescence polarization (FP) and AlphaScreen assays. 10 mM stock solutions in DMSO were prepared for all screened compounds. For the initial screen, all compounds were tested at a 50 μM concentration in 1% (v/v) DMSO solution. Compounds achieving $\geq$ 50% reduction of signal in both the FP assay and AlphaScreen were examined by a counter screen to eliminate false positive hits. Compounds that

passed the counter screen were further evaluated through a second round of the FP and AlphaScreen assays to establish a dose-dependent response. For the FP assay, an assay buffer of 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4, 1% (v/v) DMSO, 100 μg/mL bovine γ-globin, and 0.01% (v/v) Trition-X100 was used to reduce aggregation. 10 nM β-catenin and 2.5 nM fluoroscein-labeled Tcf4 were incubated in the assay buffer for 15 min at 4 °C. A range of concentrations of the test compound were tested in the assay buffer in a 96-well plate at a final well volume of 100 μL. The reaction was allowed to equilibrate while gently shaking on an orbital mixer in the dark for 1.5 h at 4 °C before polarization values were recorded. IC_{50} values were calculated using a nonlinear least-squares analysis with GraphPad Prism 5.0 and K_i values were derived from the IC_{50}. All samples were tested in triplicate.

A buffer of 20 mM HEPES, 100 mM NaCl, 0.1% BSA, and 0.001% Triton-X100 was used for AlphaScreen testing. Five nM of C-terminally biotinylated Tcf4 and 20 nM of N-terminally 6X His-tagged β-catenin were incubated in buffer for 45 min. A range of concentrations of the test compound were tested in the assay buffer in a 384-well plate at a final volume of 100 μL. The reaction was allowed to equilibrate while gently shaking on an orbital mixer in the dark for 2 h at 4 °C before emission values were recorded. IC_{50} and K_i values were determined in the same manner as the FP assay. All samples were tested in triplicate.

2.2.3 Cell viability assay. Wnt-active colorectal cancer cell lines SW480, HT29, and HCT116, as well as HEK293 cells were seeded in 96-well plates at a density of 4000 cells/well and incubated overnight at 37 °C. A range of test compound concentrations were added and the cells were incubated for an additional 72 h. After 72 h, cells were

incubated for 3 h in 10 μL of a fresh solution of 1:19 phenazine methosulfate (PMS, Sigma) solution (0.92 mg/mL) and 3-4,5-dimethylthiazol-2-yl)-5-3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2*H*-terazolium (MTS, Promega). The effect of the tested compounds was measured using A_{490} and IC_{50} values were calculated. All compounds were tested in triplicate.

2.2.4 Cell transfection and TOPFlash luciferase reporter gene assay. FuGENE 6 (E2962, Promega) was used to transfect HEK293 and SW480 cells in a 96-well plate per manufacturer's instructions. HEK293 cells were cotransfected with 45 ng of either TOPFlash or FOPFlash reporter gene and 40 ng of pCMV-RL reporter for normalization and 135 ng of pcDNA3.1-β-catenin. SW480 cells were cotransfected with 60 ng of either TOPFlash or FOPFlash reporter gene and 40 ng of pCMV-RL reporter for normalization. Cells were cultured in Dulbecco's modified Eagle medium (DMEM) containing 10% of fetal bovine serum at 37 °C for 24 h. After 24 h, varying concentrations of inhibitors were added and the cells incubated for another 24 h. Inhibitor reporter activity was measured using a Dual-Glo system (E2940, Promega). Luciferase activity was normalized to pCMV-RL reporter activity and activity was compared to cells treated with only DMSO. The TOPFlash reporter gene was performed in triplicate.

2.2.5 Quantitative reverse transcriptase real time PCR analysis (RT-qPCR). SW480 cells at a density of 1,000,000 cells/mL were treated with a range of concentrations of compounds at 37 °C for 24 h. All mRNAs were extracted using TRIzol (155960269, Life Technologies), and cDNA was synthesized using the Superscript III first-strand kit (18080-051, Invitrogen). Quantitative real-time PCR was performed with the iQ SYBR green supermix kit (170-8880, BIO-RAD) using a Bio-RAD iQ^5 multicolor

real time PCR reaction system. Threshold cycle (C_T) values were normalized to a human *Glyceraldehyde 3-phosphate dehydrogenase* (*GAPDH)* internal reference gene using primer pairs forward 5'-GAAGGTGAAGGTCGGAGTC-3' and reverse 5'-GAAGATGGTGATGGGATTTC-3'. Human *hypoxanthine phosphoribosyltransferase 1* (*HPRT)*, a house-keeping reference gene, was used as a control, with primers forward 5'-GCTATAAATTCTTTGCTGACCTGCTG-3' and reverse 5'-AATTACTTTTATGTCCCCTGTTGACTGG-3'. *Axin2*, *cyclin D-1* and *c-myc* cDNA levels were monitored for Wnt target gene expression; *Axin2* forward 5'-AGTGTGAGGTCCACGGAAAC-3', and reverse 5'-CTTCACACTGCGATGCATTT-3'; *cyclin D-1* forward 5'-ACAAACAGATCATCCGCAAACAC-3', and reverse 5'-TGTTGGGGCTCCTCAGGTTC-3'; *c-myc* forward 5'-CTTCTCTCCGTCCTCGGATTCT-3', and reverse 5'-GAAGGTGATCCAGACTCTGACCTT-3'. Compounds were tested in triplicate.

2.2.6 Western blotting. SW480 colorectal cancer cells at a density of 1,000,000 cells/mL were treated with a range of concentrations of compounds and incubated at 37 °C for 24 h. Cells were chemically lysed in a buffer of 50 mM Tris, 150 mM NaCl, 1% Nonidet P40, 0.5% sodium deoxycholate, 1% SDS, and protease inhibitors at pH 7.4. The supernatant was collected after 20 min of centrifugation at 12,000 rpm and 4 °C and loaded onto an 8% SDS-PAGE for separation. Separated proteins were transferred onto a nitrocellulose membrane for immunoblot assays. Primary antibodies against c-myc (D84C12, Cell Signaling), cyclin D1 (sc-853, Santa Cruz Biotechnology, Inc), and β-tubulin, a control (sc-55529, Santa Cruz Biotechnology, Inc.), proteins were used. The secondary antibody consisted of IRDye 680LT goat antimouse IgG (827-11080, LiCOR)

or IRDye 800CW goat antirabbit IgG (827-08365, LiCOR). Western blot bands were detected using LiCOR Odyssey infrared imaging system and quantified using LiCOR Image Studio Lite 4.0 software. Compounds were tested in duplicate.

2.2.8 Coimmunoprecipitation assay. SW480 colorectal cancer cells at a density of 1,000,000 cells/mL were treated with a range of concentrations of **12** and incubated at 37 °C for 24 h. Cells were chemically lysed in a buffer of 50 mM Tris, 150 mM NaCl, 1% Nonidet P-40, 2 mM EDTA, and protease inhibitors at pH 7.4. The resulting lysates were preadsorbed onto A/G plus agarose (sc-2003, Santa Cruz Biotechnology, Inc) for 1 h at 4 °C. A/G plus agarose was then added and the mixture incubated for 3 h at 4 °C. The beads were washed 5 times with the same buffer used for cell lysis at 4 °C. Bound proteins were eluted by boiling the beads in SDS sample buffer. Eluted proteins were then loaded onto SDS-PAGE for separation. The separated proteins were transferred onto nitrocellulose membranes for Western blot analysis. Primary antibodies against β-catenin (610153, BD Biosciences) and human Tcf4 (05-511, Millipore) and secondary antibodies IRDye 680LT goat antimouse IgG (827-11080, LiCOR) were used. Images were captured using the LiCOR infrared imaging system. Compound **12** was tested in duplicate.

2.2.9 Ligand docking using AutoDock Vina. A blind docking of **12** was performed using AutoDock Vina.[47] Overlapping grids were generated to cover the entire β-catenin (PDB ID 2GL7)[10] surface. Exhaustiveness was increased to 12 (exhaustiveness = 12), and 18 ligand conformations were generated for each grid (num_modes = 18). All other parameters were left as default values. Compound **49** was docked into a grid centered on the Tcf4 G_{13}ANDE$_{17}$ binding region of β-catenin.

2.2.10 Ligand docking using Glide 5.8. All ligand structures were prepared by Schrodinger LigPrep[48] using Epik[49] to generate protonation and tautomeric states at pH 7.0. Ligand structures were minimized using the OPLS_2005 force field[50] in generalized Born solvent accessible (GBSA) implicit water solvent conditions. Atomic partial charges were assigned by the OPLS_2005 force field. A grid encompassing all residues of the human Tcf4 $G_{13}ANDE_{17}$ binding region of β-catenin was generated with default Glide parameters.[51] The ligand scaling factor was adjusted to 0.5 for atoms with partial charges lower than 0.15. 10,000 poses were generated during the initial phase of docking and the 1,000 best poses were kept for a 5,000-step maximum energy minimization. Up to 100 poses per ligand were minimized post-docking. A maximum of 100,000 ligand poses were generated per docking run with a maximum of 50 poses reported per ligand. All remaining settings were kept as default for Glide standard precision.[51,52]

2.2.11 Ligand docking using AutoDock 4.2. AutoDock Tools and AutoTors were used to generate united atom ligand coordinates utilizing Gasteiger-Marsili atomic partial charges and define torsion trees.[53] Grids covering all residues of the human Tcf4 $G_{13}ANDE_{17}$ binding region of β-catenin were generated with AutoGrid 4 with 0.375 Å grid spacing. Ligands were docked using 100 iterations of a Lamarckian genetic algorithm to generate 100 poses per ligand.[53]

2.2.12 Compound sources, general chemical methods, reagents, and materials for chemical synthesis. Compounds **20**-**23** and **27**-**30** were purchased from Sigma-Aldrich. Compounds **12** (UZI/7116003) and **26** (UZI/2587998) were purchased from the Zelinsky Institute, Inc. (Newark, Delaware, USA). Compound **12** was resynthesized by Dr. Yongqiang Zhang. Compound **24** (STOCK1S-52622) was purchased from

InterBioScreen (Moscow, Russia). Compounds **8** and **9** were resynthesized by the Ji lab (University of Utah, Salt Lake City, Utah, USA). Chemical reagents were purchased from Sigma-Aldrich, Acros Organics, or Ark Pharm, Inc. and used as is without further purification unless otherwise stated. ^{1}H and ^{13}C NMR spectra were recorded using a Varian VXR-500 (500 MHz, 125.7 MHz ^{13}C), Varian Inova-400 (400 MHz ^{1}H, 100 MHz ^{13}C) or Varian Unity-300 (300 MHz ^{1}H, 75 ^{13}C) spectrometers (Figures 2S.1 – 2S.20). Chemical shifts were reported as parts per million. ^{1}H reference peaks were set at 7.26 ppm, 2.5 ppm and 2.05 ppm for $CHCl_3$, $(CD_2H)_2SO$, and $(CD_2H)_2CO$. ^{13}C reference peaks were set at 77.23 ppm, 39.52 ppm, and 29.84 ppm for $CDCl_3$, DMSO-d_6, and acetone-d_6. Low- and high- resolution mass spectra were obtained using a Mmass Quattro II mass spectrometer with an ESI source. Thin-layer chromatography was conducted with E. Merck precoated silica gel 60 F254 plates with visualization accomplished with a UV-visible lamp. Column chromatography was performed using 234-400 mesh SilaFlash F60.

2.2.13 Chemical synthesis. All synthesized compounds except for **34**-**40** and **44**-**49** were synthesized by Dr. Yongqiang Zhang. Compounds **45, 47**, and **49** were synthesized by undergraduate ACCESS fellow Brooklyn Brakey.

9H-Indeno[1,2-b][1,2,5]oxadiazolo[3,4-e]pyrazin-9-one (**24**). 1,2,5-oxadiazole-3,4-diamine (0.18 g, 1.80 mmol) was added to a solution of ninhydrin (0.32 g, 1.80 mmol) in a solvent mixture of ethanol and glacial acetic acid (1:1, 10 mL). After stirring for 18 h at rt, the mixture was heated to a gentle reflux for another 6 h. It was then cooled to rt, and the resulting precipitate was filtered and washed with water to yield **24** as a pale yellow solid (0.29 g, 72% yield). ^{1}H NMR (300 MHz, DMSO-d_6) δ ppm 8.31 (d, J = 7.2

Hz, 1H), 8.18–8.02 (m, 2H), 7.97 (d, *J* = 7.2 Hz, 1H). ^{13}C NMR (125 MHz, DMSO-d_6) δ ppm 189.6, 161.7, 140.0, 140.0, 138.7, 136.5, 126.1. HRMS (ESI) calcd for $C_{11}H_4N_4O_2$ $(M + Na)^+$ 247.0232, found 247.0230.

(Z)-9-Hydrazono-9H-indeno[1,2-b][1,2,5]oxadiazolo[3,4-e]pyrazine (**25**). Hydrazine hydrate (0.33 g, 6.70 mmol) was added to a solution of **24** (0.15 g, 0.67 mmol) in a mixture of ethanol and glacial acid (1:1, 10 mL). The mixture was heated to a gentle reflux for 2 h and then cooled to rt. The resulting precipitate was filtered and washed with water to afford **24** as a red solid (0.14 g, 93% yield). ^{1}H NMR (300 MHz, DMSO-d_6) δ ppm 10.48 (d, *J* = 13.8 Hz, 1H), 10.40 (d, *J* = 13.8 Hz, 1H), 8.07 (d, *J* = 7.5 Hz, 1H), 7.73–7.66 (m, 2H), 7.46 (t, *J* = 7.5 Hz, 1H). ^{13}C NMR (75 MHz, DMSO-d_6) δ ppm 161.8, 155.0, 152.7 152.5, 144.8, 135.9, 129.8, 128.9, 128.5, 124.7, 119.7. HRMS (ESI) calcd for $C_{11}H_6N_6O$ $(M – H)^–$ 237.0525, found 237.0531.

4-Amino-1,2,5-oxadiazole-3-carbohydrazide (**43**). HCl gas was bubbled into a solution of 4-amino-1,2,5-oxadiazole-3-carboxylic acid (0.50 g, 3.87 mmol) in methanol (30 mL). The mixture was heated to a gentle reflux for 8 h. Methanol was then removed by vacuum to give the crude product as a pale yellow oil. Diethyl ether (30 mL) was added to this residue and stirred for 30 min. The resulting precipitate was filtered to give desired product **42** (0.50 g, 91% yield) as a white solid. It was used directly in the next step without further purification. Hydrazine hydrate (0.08g, 1.57 mmol) was added to a solution of **42** (0.15 g, 1.05 mmol) in methanol (10 mL), and the mixture was heated to a gentle reflux. After 15 h, the solvent was removed completely under vacuum to give **43** (0.09 g, 60% yield) as a white solid. ^{1}H NMR (500 MHz, acetone-d_6) δ ppm 9.50 (brs, 1H), 5.76 (s, 2H), 4.50 (brs, 2H). ^{13}C NMR (-d_6) δ ppm 157.0, 156.0, 140.0. HRMS (ESI)

calcd for $C_3H_5N_5O_2$ $(M + H)^+$ 144.0516, found 144.0519.

(Z)-4-Amino-N′-(9H-indeno[1,2-b][1,2,5]oxadiazolo[3,4-e]pyrazin-9-ylidene)-1,2,5-oxadiazole-3-carbohydrazide (**12**). Compound **43** (0.03 g, 0.22 mmol) and glacial acetic acid (1 mL) was added to a solution of **24** (0.05 g, 0.22 mmol) in ethanol (10 mL). The reaction mixture was heated to a gentle reflux overnight and then poured into ice water. The resulting precipitate was filtered and washed with water to give **12** as a yellow solid (0.06 g, 78% yield). ^{1}H NMR (500 MHz, DMSO-d_6) δ ppm 13.55 (brs, 1H), 8.26 (d, J = 7.5 Hz, 1H), 8.06 (d, J = 7.0 Hz, 1H), 7.92 (t, J = 7.5 Hz, 1H), 7.82 (t, J = 7.0 Hz, 1H), 6.01 (s, 2H). ^{13}C NMR (125 MHz, DMSO-d_6) δ ppm 162.8, 158.6, 157.1, 155.4, 153.2, 152.3, 152.0, 144.3, 140.4, 136.8, 134.6, 133.7, 125.2, 122.8. HRMS (ESI) calcd for $C_{14}H_7N_9O_3$ $(M – H)^–$ 348.0594, found 348.0604.

5H-[1,2,5]Oxadiazolo[3′,4′:5,6]pyrazino[2,3-b]indole (**31**). 1,2,5-oxadiazole-3,4-diamine (0.14 g, 1.36 mmol) was added to a solution of isatin (0.20 g, 1.36 mmol) in glacial acetic acid (5 mL). The resulting mixture was heated to a gentle reflux for 15 h and then poured into ice water. The resulting precipitate was filtered and washed with water to afford **31** as a red solid (0.05 g, 17% yield). ^{1}H NMR (300 MHz, DMSO-d_6) δ ppm 12.35 (brs, 1H), 8.19 (d, J = 7.8 Hz, 1H), 7.73 (t, J = 7.5 Hz, 1H), 7.39 (d, J = 8.1 Hz, 1H), 7.32 (t, J = 7.2 Hz, 1H). ^{13}C NMR (75 MHz, DMSO-d_6) δ ppm 154.0, 152.6, 152.5, 151.9, 148.8, 136.4, 125.3, 123.1, 118.9, 113.4. HRMS (ESI) calcd for $C_{10}H_5N_5O$ $(M – H)^–$ 210.0416, found 210.0427.

Ethyl-2-(5H-[1,2,5]oxadiazolo[3′,4′:5,6]pyrazino[2,3-b]indol-5-yl)acetate (**32**). K_2CO_3 (0.08 g, 0.57 mmol) was added to a solution of **31** (0.08 g, 0.38 mmol) and ethyl 2-bromoacetate (0.09 g, 0.57 mmol) in DMSO (10 mL). The resulting mixture was

heated to 80 °C overnight. After 15 h, the reaction mixture was cooled to rt and diluted with ethyl acetate (80 mL), and washed with brine (20 mL × 3). The organic layer was dried over Na_2SO_4, filtered and concentrated. The residue was purified by column chromatography (silica gel, hexanes:acetone = 5:1) to afford **32** (0.08 g, 73% yield) as a red solid. ^{1}H NMR (500 MHz, DMSO-d_6) δ ppm 8.28 (d, *J* = 7.5 Hz, 1H), 7.85 (t, *J* = 7.0 Hz, 1H), 7.69 (d, *J* = 8.5 Hz, 1H), 7.42 (t, *J* = 7.5 Hz, 1H), 5.22 (s 2H), 4.18 (q, *J* = 7.0 Hz, 2H), 1.21 (t, *J* = 7.0 Hz, 3H). ^{13}C NMR (125 MHz, DMSO-d_6) δ ppm 168.0, 153.2, 152.3, 152.1, 151.6, 148.9, 136.5, 125.3, 124.1, 118.5, 112.4, 62.3, 43.4, 14.7. HRMS (ESI) calcd for $C_{14}H_{11}N_5O_3$ $(M + Na)^+$ 320.0760, found 320.0759.

2-(5H-[1,2,5]Oxadiazolo[3′,4′:5,6]pyrazino[2,3-b]indol-5-yl)acetic Acid (**33**). LiOH (0.05 g, 2.15 mmol) was added to a solution of **32** (0.08 g, 0.27 mmol) in a solvent mixture (14 mL, THF:MeOH:H_2O = 4:2:1). The mixture was stirred for 8 h at rt. Then, the pH value was adjusted to 4–5 with HCl (1 M), diluted with water (50 mL) and extracted with ethyl acetate (20 mL × 3). The combined organic phase was dried over Na_2SO_4 and concentrated to give **33** (0.05g, 68%) as an orange solid. ^{1}H NMR (500 MHz, DMSO-d_6) δ ppm 13.36 (brs, 1H), 8.28 (d, *J* = 8.0 Hz, 1H), 7.84 (t, *J* = 8.0 Hz, 1H), 7.69 (d, *J* = 8.0 Hz, 1H), 7.41 (t, *J* = 8.0 Hz, 1H), 5.10 (s, 2H). ^{13}C NMR (125 MHz, DMSO-d_6) δ ppm 169.4, 153.2, 153.0, 152.1, 151.6, 149.1, 136.5, 125.3, 124.0, 118.4, 112.4, 43.4. HRMS (ESI) calcd for $C_{12}H_7N_5O_3$ $(M – H)^-$ 268.0471, found 268.0485.

2-(5H-[1,2,5]Oxadiazolo[3′,4′:5,6]pyrazino[2,3-b]indol-5-yl)-N-(4-amino-1,2,5-oxadiazol-3-yl)acetamide (**13**). Isobutyl chloroformate (0.09 g, 0.67 mmol) was added to a solution of **33** (0.18 g, 0.67 mmol) and 4-methylmorpholine (0.14 g, 0.34 mmol) in THF (15 mL) at −15 °C. The resulting mixture was stirred for 1 h at the same

temperature. Then 1,2,5-oxadiazole-3,4-diamine (0.10 g, 1.00 mmol) was slowly added. The temperature was allowed to rise to rt gradually and stirred for another 1 h. The mixture was diluted with ethyl acetate (80 mL), and washed with brine (20 mL × 3). The organic layer was dried over Na_2SO_4, filtrated, and concentrated. The product was purified by column chromatography (silica gel, hexanes:acetone = 2:1 to 1:1) to afford **13** (0.12 g, 51% yield) as an orange solid. ^{1}H NMR (500 MHz, DMSO-d_6) δ ppm 11.13 (brs, 1H), 8.32 (d, *J* = 7.5 Hz, 1H), 7.87 (t, *J* = 8.0 Hz, 1H), 7.73 (d, *J* = 8.0 Hz, 1H), 7.45 (t, *J* = 7.5 Hz, 1H), 6.04 (s, 2H), 5.29 (s, 2H). ^{13}C NMR (125 MHz, DMSO-d_6) δ ppm 166.7, 153.5, 152.9, 152.4, 152.1, 152.0, 149.3, 144.3, 136.5, 125.3, 124.0, 118.7, 112.6, 44.8. HRMS (ESI) calcd for $C_{14}H_9N_9O_3$ $(M - H)^-$ 350.0750, found 350.0755.

Ethyl 4-(5H-[1,2,5]Oxadiazolo[3′,4′:5,6]pyrazino[2,3-b]indol-5-yl)butanoate (**36**). NaH (0.0030 g, 0.12 mmol) was added to a solution of **31** (0.017 g, 0.08 mmol) in DMF (5 mL) at 0 °C. The resulting mixture was stirred for 0.5 h at 0 °C before adding ethyl 4-bromobutanoate (0.023 g, 0.12 mmol). The reaction solution was allowed to warm to rt gradually then stirred overnight. The mixture was diluted with ethyl acetate (50 mL), and washed with brine (20 mL × 3). The organic layer was dried over Na_2SO_4, filtrated, and concentrated. The product was purified through column chromatography (silica gel, hexanes:acetone = 3:1 to 1:1) to afford **36** as an orange solid (0.010 g, yield 58%). ^{1}H NMR (400 MHz, $CDCl_3$) δ ppm 8.20 (d, *J* = 7.2 Hz, 1H), 7.71 (t, *J* = 7.2 Hz, 1H), 7.42 (d, *J* = 8.0 Hz, 1H), 7.31 (t, *J* = 7.2 Hz, 1H), 4.32 (t, *J* = 6.8 Hz, 2H), 4.07 (q, *J* = 6.0 Hz, 2H), 2.43 (t, *J* = 6.4 Hz, 2H), 2.20–2.12 (m, 2H) 1.19 (t, *J* = 6.0 Hz, 3H). ^{13}C NMR (100 MHz, $CDCl_3$) δ ppm 172.5, 151.8, 151.3, 151.2, 150.4, 148.0, 135.6, 125.1, 123.1, 118.2, 110.6, 60.7, 40.9, 30.9, 22.7, 14.1. HRMS (ESI) calcd for $C_{16}H_{15}N_5O_3$ $(M + Na)^+$

348.1073, found 348.1081.

Methyl 3-(5H-[1,2,5]Oxadiazolo[3′,4′:5,6]pyrazino[2,3-b]indol-5-yl)propanoate (**34**) 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) (0.24 g, 1.57 mmol) was added to a solution of **32** (0.10 g, 0.47 mmol) in acetonitrile (15 mL). The resulting mixture was stirred for 0.5 h before adding methyl acrylate (0.14 g, 1.57 mmol). The reaction solution was heated to 50 °C and stirred for another 24 h. It was then diluted with ethyl acetate (50 mL), and washed with brine (20 mL × 3). The organic layer was dried over Na_2SO_4, filtrated, and concentrated. The product was purified by column chromatography (silica gel, hexanes:acetone = 3:1 to 1:1) to afford **34** as an orange solid (0.084 g, yield 60%). ^{1}H NMR (400 MHz, $CDCl_3$) δ ppm 8.33 (d, *J* = 7.6 Hz, 1H), 7.79 (t, *J* = 8.4 Hz, 1H), 7.51 (d, *J* = 8.4 Hz, 1H), 7.40 (t, *J* = 7.6 Hz, 1H), 4.60 (t, *J* = 6.8 Hz, 2H), 3.66 (s, 3H), 3.02 (t, *J* = 6.8 Hz, 2H). ^{13}C NMR (100 MHz, $CDCl_3$) δ ppm 171.3, 151.8, 151.4, 151.2, 150.4, 147.9, 135.5, 125.2, 123.3, 118.4, 110.8, 52.1, 37.9, 21.1. HRMS (ESI) calcd for $C_{14}H_{11}N_5O_3$ $(M + Na)^+$ 320.0760, found 320.0760.

3-(5H-[1,2,5]Oxadiazolo[3′,4′:5,6]pyrazino[2,3-b]indol-5-yl)propanoic Acid (**35**). Compound **35** was synthesized using the same procedure as **33**. Orange solid, yield 80%. ^{1}H NMR (400 MHz, DMSO-d_6) δ ppm 12.46 (brs, 1H), 8.26 (d, *J* = 7.6 Hz, 1H), 7.85 (t, *J* = 8.0 Hz, 1H), 7.76 (d, *J* = 8.0 Hz, 1H), 7.40 (t, *J* = 7.6 Hz, 1H), 4.48 (t, *J* = 6.8 Hz, 2H), 2.83 (t, *J* = 7.2 Hz, 2H). ^{13}C NMR (100 MHz, DMSO-d_6) δ ppm 172.7, 153.6, 152.1, 152.0, 151.5, 148.7, 136.9, 125.0, 123.3, 118.5, 112.3, 32.1, 31.1. HRMS (ESI) calcd for $C_{13}H_9N_5O_3$ $(M + Na)^+$ 306.0603, found 306.0612.

4-(5H-[1,2,5]Oxadiazolo[3′,4′:5,6]pyrazino[2,3-b]indol-5-yl)butanoic Acid (**37**). Compound **37** was synthesized using the same procedure as **33**. Orange solid, and yield

65%. ^{1}H NMR (400 MHz, DMSO-d_6) δ ppm 12.08 (brs, 1H), 8.27 (d, J = 7.6 Hz, 1H), 7.86 (t, J = 8.0 Hz, 1H), 7.70 (d, J = 8.0 Hz, 1H), 7.41 (t, J = 7.2 Hz, 1H), 4.30 (t, J = 6.8 Hz, 2H), 2.38 (t, J = 7.2 Hz, 2H), 2.07–2.00 (m, 2H). ^{13}C NMR (100 MHz, DMSO-d_6) δ ppm 174.4, 153.7, 152.2, 152.0, 148.8, 136.1, 125.1, 123.3, 118.6, 111.8, 110.0, 41.1, 31.2, 23.1. HRMS (ESI) calcd for $C_{14}H_{11}N_5O_3$ $(M + Na)^+$ 320.0760, found 320.0761.

2-(5H-[1,2,5]Oxadiazolo[3′,4′:5,6]pyrazino[2,3-b]indol-5-yl)acetonitrile (**38**). NaH (60%) (0.007 g, 0.17 mmol) was added to a solution of **31** (0.03 g, 0.14 mmol) in DMF (10 mL) at 0 °C. The resulting mixture was stirred for 0.5 h at the same temperature before 2-bromoacetonitrile (0.02 g, 0.17 mmol) was added into it. It was then allowed to warm to rt gradually and stirred for another 1 h. The reaction mixture was then quenched with water (30 mL) and diluted with ethyl acetate (60 mL). The organic phase was washed with brine (20 mL × 2), dried over Na_2SO_4, and concentrated. The residue was purified by column chromatography (silica gel, hexanes:acetone = 5:1) to afford **38** (0.025 g, 71% yield) as a yellow solid. ^{1}H NMR (500 MHz, acetone-d_6) δ ppm 8.37 (d, J = 8.0 Hz, 1H), 7.97 (t, J = 8.5 Hz, 1H), 7.85 (d, J = 8.5 Hz, 1H), 7.56 (t, J = 8.0 Hz, 1H), 5.51 (s 2H). ^{13}C NMR (125 MHz, acetone-d_6) δ ppm 152.7, 152.3, 151.7, 151.1, 147.17, 136.0, 125.0, 124.2, 119.2, 114.3, 111.3, 29.4. HRMS (ESI) calcd for $C_{12}H_6N_6O$ $(M + Na)^+$ 273.0501, found 273.0508.

4-(5H-[1,2,5]Oxadiazolo[3′,4′:5,6]pyrazino[2,3-b]indol-5-yl)butanenitrile (**40**). Compound **40** was synthesized using the same procedure as **36**. Orange solid, yield 50%. ^{1}H NMR (400 MHz, DMSO-d_6) δ ppm 7.61–7.64 (m, 2H), 7.16 (t, J = 7.6 Hz, 1H), 6.97 (d, J = 8.0 Hz, 1H), 3.87 (t, J = 7.2 Hz, 2H), 2.48 (t, J = 7.2 Hz, 2H), 2.11–2.07 (m, 2H). ^{13}C NMR (100 MHz, DMSO-d_6) δ ppm 158.4, 150.2, 138.6, 125.8, 124.7, 118.6, 117.7,

109.8, 38.9, 23.6, 15.0. HRMS (ESI) calcd for $C_{14}H_{10}N_6O$ $(M + Na)^+$ 301.0814, found 301.0808.

3-(5H-[1,2,5]Oxadiazolo[3′,4′:5,6]pyrazino[2,3-b]indol-5-yl)propanenitrile (**39**). DBU (0.072 g, 0.47 mmol) was added to a solution of **31** (0.20 g, 0.95 mmol) in acetonitrile (10 mL). The solution was stirred for 0.5 h. Then, acrylonitrile (0.075 g, 1.4 mmol) was added. The reaction solution was heated to 50 °C and stirred for 24 h. It was then diluted with ethyl acetate (50 mL), washed with brine (20 mL × 3). The organic layer was dried over Na_2SO_4, filtrated, and concentrated. The product was purified by column chromatography (silica gel, hexanes:acetone = 5:1 to 1:1) to afford **39** as orange solid (yield 62%). ^{1}H NMR (400 MHz, DMSO-d_6) δ ppm 8.28 (d, J = 7.6 Hz, 1H), 7.89–7.84 (m, 2H), 7.42 (t, J = 7.2 Hz, 1H), 4.61 (t, J = 6.8 Hz, 2H), 3.10 (t, J = 6.4 Hz, 2H). ^{13}C NMR (100 MHz, DMSO-d_6) δ ppm 153.4, 152.0, 151.9, 151.5, 148.3, 136.2, 125.1, 123.7, 119.1, 118.5, 112.2, 37.7, 16.4. HRMS (ESI) calcd for $C_{13}H_8N_6O$ $(M + Na)^+$ 287.0657, found 287.0661.

5-(2-(2H-Tetrazol-5-yl)ethyl)-5H-[1,2,5]oxadiazolo[3′,4′:5,6]pyrazino[2,3-b]indole (**14**). nBu$_3$SnN$_3$ (0.13 g, 0.38 mmol) was added to a solution of **39** (0.02 g, 0.076 mmol) in toluene (10 mL). The resulting mixture was heated to reflux for 48 h. It was then cooled to rt, and the pH value was adjusted to 4–5 with HCl (1 M) and diluted with ethyl acetate (60 mL). The organic phase was washed with brine (20 mL × 2), dried over Na_2SO_4, and concentrated. The residue was purified by column chromatography (silica gel, CH_2Cl_2:MeOH = 10:1 to 5:1) to afford **14** (15.00 mg, 65% yield) as red solid. ^{1}H NMR (500 MHz, DMSO-d_6) δ ppm 8.28 (d, J = 7.8 Hz, 1H), 7.81 (t, J = 8.1 Hz, 1H), 7.58 (d, J = 8.0 Hz, 1H), 7.39 (t, J = 7.8 Hz, 1H), 4.66 (t, J = 6.6 Hz, 2H), 3.46 (t, J = 6.6

Hz, 2H). ^{13}C NMR (125 MHz, DMSO-d_6) δ ppm 154.3, 153.8, 152.2, 152.2, 151.8, 148.7, 136.3, 125.3, 123.7, 118.7, 111.9, 40.4, 22.3. HRMS (ESI) calcd for $C_{13}H_9N_9O$ $(M - H)^-$ 306.0852, found 306.0871.

5-(3-(2H-Tetrazol-5-yl)propyl)-5H-[1,2,5]oxadiazolo[3′,4′:5,6]pyrazino[2,3-b]indole (**41**). Compound **41** was synthesized using the same procedure as **14**. Red solid, yield 57%. ^{1}H NMR (500 MHz, DMSO-d_6) δ ppm 8.26 (d, J = 8.0 Hz, 1H), 7.85 (t, J = 8.0 Hz, 1H), 7.72 (d, J = 8.5 Hz, 1H), 7.40 (t, J = 7.5 Hz, 1H), 4.40 (t, J = 6.5 Hz, 2H), 3.02 (t, J = 7.5 Hz, 2H), 2.32–2.26 (m, 2H). ^{13}C NMR (125 MHz, DMSO-d_6) δ ppm 153.8, 152.4, 152.2, 151.9, 148.9, 136.4, 125.3, 123.6, 118.8, 112.1, 41.3, 25.6, 21.1. HRMS (ESI) calcd for $C_{14}H_{11}N_9O$ $(M + Na)^+$ 344.0984, found 344.0990.

6H-indolo[2,3-b]quinoxaline (**44**)**.** *o*-phenylenediamine (0.25 g, 2.3 mmol) was added to a solution of isatin (1.0 g, 6.8 mmol) in glacial acetic acid (10 ml) at rt. The mixture was heated to gentle reflux for 24 h and then poured into ice water. The resulting precipate was filtered to yield **44** as a yellow solid (0.36 g, 72% yield) ^{1}H NMR (400 MHz, DMSO-d6) δ ppm 12.02 (s, 1H), 8.34 (d, J = 7.6 Hz, 1H), 8.24 (d, J = 8.3 Hz,1H), 8.06 (d, J = 8.3 Hz, 1H), 7.79 (t, J = 7.6 Hz, 1H), 7.70 (dd, J = 1.3, J = 7.2 Hz, 2H), 7.57 (d, J = 7.9 Hz, 1H), 7.36 (t, J = 7.5 Hz, 1H). ^{13}C NMR (100 MHz, DMSO-d6) δ ppm 146.3, 144.4, 140.6, 140.2, 139.0, 131.8, 129.5, 129.2, 127.9, 126.4, 122.7, 121.1, 119.4, 112.4.

7,9-dichloro-6H-indolo[2,3-b]quinoxaline (**45**). Compound **45** was synthesized using the same procedure as **44**. Yellow solid (yield 80%). Due to solubility issues, no NMR spectra for this compound were obtained.

3-(6H-indolo[2,3-b]quinoxalin-6-yl)propanenitrile (**46**). Compound **46** was

synthesized using the same procedure as **39**. Yellow solid (56 % yield). ^{1}H NMR (400 MHz, DMSO-d6) δ ppm 8.50 (d, *J* = 7.7 Hz, 1H), 8.32 (dd, *J* = 1.1 Hz, *J* = 8.3 Hz, 1H), 7.79 (ddd, *J* = 1.5 Hz, *J* = 6.9 Hz, *J* = 8.4 Hz, 2H), 7.76-7.70 (m, 2H), 7.56 (d, 1H, J=8.2Hz), 7.44 (t, 1H, J=7.6Hz), 4.81 (t, 2H, J=7.0Hz), 3.09 (t, 2H, J=7.0Hz). Due to solubility issues, no ^{13}C NMR spectrum was obtained.

3-(7,9-dichloro-6H-indolo[2,3-b]quinoxalin-6-yl)propanenitrile (**47**). Compoound **47** was synthesized using the same procedure as **39**. Yellow solid (55% yield). ^{1}H NMR (400 MHz, DMSO-d6) δ ppm 8.41 (d, *J* = 2.0 Hz, 1H), 8.32 (dd, *J* = 1.1 Hz, *J* = 8.3 Hz, 1H), (dd, *J* = 1.3 Hz, *J* = 8.4 Hz, 1H), 8.17 (dd, *J* = 1.2 Hz, *J* = 8.4 Hz, 1H), 7.84 (ddd, *J* = 1.5 Hz, *J* = 6.9 Hz, *J* = 8.4 Hz, 2H), 7.76 (ddd, *J* = 1.3 Hz, *J* = 7.0 Hz, *J* = 8.1Hz, 1H), 7.67 (d, *J* = 2.0 Hz, 1H), 5.26 (t, *J* = 7.3 Hz, 2H), 3.08 (t, *J* = 7.3 Hz, 2H). Due to solubility issues, no ^{13}C NMR spectrum was obtained.

6-(2-(2H-tetrazol-5-yl)ethyl)-6H-indolo[2,3-b]quinoxaline (**48**). nBu_3SnN_3 (0.37 g, 1.1 mmol) was added to a solution of **46** (0.060 g, 0.22 mmol) in toluene (20 mL) at rt. The resulting mixture was heated to reflux and stirred for 72 h. It was then cooled to rt and quenched with glacial acetic acid (10 ml). The resulting precipitate was filtered and washed with glacial acetic acid (10 ml) to afford **48** as a yellow solid. (0.10 g, 71% yield). ^{1}H NMR (400 MHz, DMSO-d6) δ ppm 8.34 (d, *J* = 7.7 Hz, 1H), 8.24 (dd, *J* = 1.3 Hz, *J* = 8.3 Hz, 1H), 8.04 (dd, *J* = 1.2 Hz, *J* = 8.4 Hz, 1H), 7.80 (ddd, *J* = 1.5 Hz, *J* = 6.9 Hz, *J* = 8.4 Hz, 1H), 7.74-7.68 (m, 2H), 7.58 (d, *J* = 8.2 Hz, 1H), 7.37 (t, *J* = 7.5 Hz, 1H), 4.84 (t, *J* = 6.8 Hz, 2H), 3.53 (t, *J* = 6.8 Hz, 2H). ^{13}C NMR (100 MHz, DMSO-d6) δ ppm 145.4, 144.3, 140.2, 140.1, 139.2, 131.8, 129.5, 129.6, 127.9, 126.6, 122.7, 121.6, 119.1, 110.4, 40.6, 22.9. HRMS for $C_{17}H_{13}N_7$: 315.1232, found: 315.1238.

6-(2-(2H-tetrazol-5-yl)ethyl)-7,9-dichloro-6H-indolo[2,3-b]quinoxaline (**49**). Compound **49** was synthesized using the same procedure as **48**. Yellow solid. Yield 69%. ^{1}H NMR (400 MHz, DMSO-d6) δ ppm 8.36 (d, J = 2, 1H), 8.314 (d, *J* = 8 Hz, 1H), 8.01 (d, *J* = 8.4 Hz, 1H), 7.89 (d, *J* = 2 Hz, 1H), 7.85 (t, *J* = 8 Hz, 1H), 7.76 (t, *J* = 8 Hz, 1H), 5.126 (t, *J* = 6.8 Hz, 2H), 3.521 (t, *J* = 6.8 Hz, 2H). ^{13}C NMR (100 MHz, DMSO-d6) δ ppm 165.5, 146.1, 140.8, 139.7, 138.2, 138.2, 132.2, 130.5, 129.6, 128.1, 127.6, 126.2, 123.7, 121.0, 117.2, 41.3, 24.5.

2.3 Results

2.3.1 High-throughput screening to discover drug-like, small-molecule β-catenin/Tcf4 PPI inhibitors. The high-throughput FP and AlphaScreen assays were previously optimized by Dr. Min Zhang for the discovery of β-catenin/Tcf4 PPI inhibitors.[54] The FP assay used C-terminally fluorescein-labeled human Tcf4 peptide consisting of residues 7-51 and unlabeled human β-catenin ARM domain (residues 138-686). Tcf4 was selected for labeling because the β-catenin/Tcf4 complex is over 10-fold larger than the fluorescein-labeled Tcf4. Therefore, it will exhibit the most depolarization when the large β-catenin/Tcf4 complex is disrupted. In the AlphaScreen assay, the streptavidin-coated donor bead is bound to a C-terminally biotinylated human Tcf4 (residues 7-51), and the nickel-chelate acceptor bead is bound to a C-terminally 6X His-tagged β-catenin ARM domain (residues 138-686). Laser excitation at 680 nm causes photosensitizers inside the donor bead to convert ambient oxygen into a higher energy singlet state. If the acceptor bead is within 2,000 Å of the donor bead, as is the case in the β-catenin/Tcf4 complex, the singlet oxygen will cause an emission of light at 570 nm.

Inhibition of the β-catenin/Tcf4 PPI will result in a larger distance between the donor and acceptor beads and a reduction in emission at 570 nM. Dr. Min Zhang screened 2,093 compounds from multiple libraries: 24 compounds from the Zinc is Not Commercialized (ZINC) library, 90 compounds from the synthetic LOPAC^{Pfizer} library, 117 natural products, 269 Sigma-Aldrich carboxylic acids and sulfonamides, and 1593 small molecules from the Diversity Set V from the Developmental Therapeutics Program of the National Cancer Institute/National Institute of Health. Compounds **12** and **20**-**22** inhibited the β-catenin/Tcf4 PPI in both FP and AlphaScreen assays (Figure 2.3). The most potent compound, **12**, achieved apparent K_i values of 5.8 $\pm$ 1.3 μM and 1.2 $\pm$ 0.4 μM in FP and AlphaScreen assays, respectively. Another potent dianion compound, **23**, was discovered in the FP screening, but was not further evaluated. Dr. Min Zhang evaluated **12** and **20**-**22** for cell-based inhibition of the Wnt signaling pathway through a TOPFlash luciferase reporter gene assay[55] and MTS cell viability assay.[56,57] Again, **12** was the most effective inhibitor tested. Compound **12** downregulated (IC_{50} = 0.71-0.86 μM) Wnt target gene transcription in a highly sensitive TOPFlash luciferase reporter gene assay (Figure 2.4A, and inhibited growth of colorectal cancer cells (IC_{50} = 0.85-1.1 μM) in a dose-dependent manner (Figure 2.4B). Of these four inhibitors, only **12** proved effective in all assays, and was therefore selected for further study and optimization.

*2.3.2 Effects of **12** on Wnt target gene expression and protein expression levels.* The effect of **12** on *Axin2*,[58] *cyclin D1*[7] *and c-myc*[6] Wnt target gene transcription was monitored by RT-qPCR by Dr. Min Zhang (Figure 2.5A). Compound **12** downregulated transcription in all tested Wnt target genes in SW480 cells. Transcription of the *HPRT* house-keeping gene was unaffected. Dr. Min Zhang evaluated protein expression of two

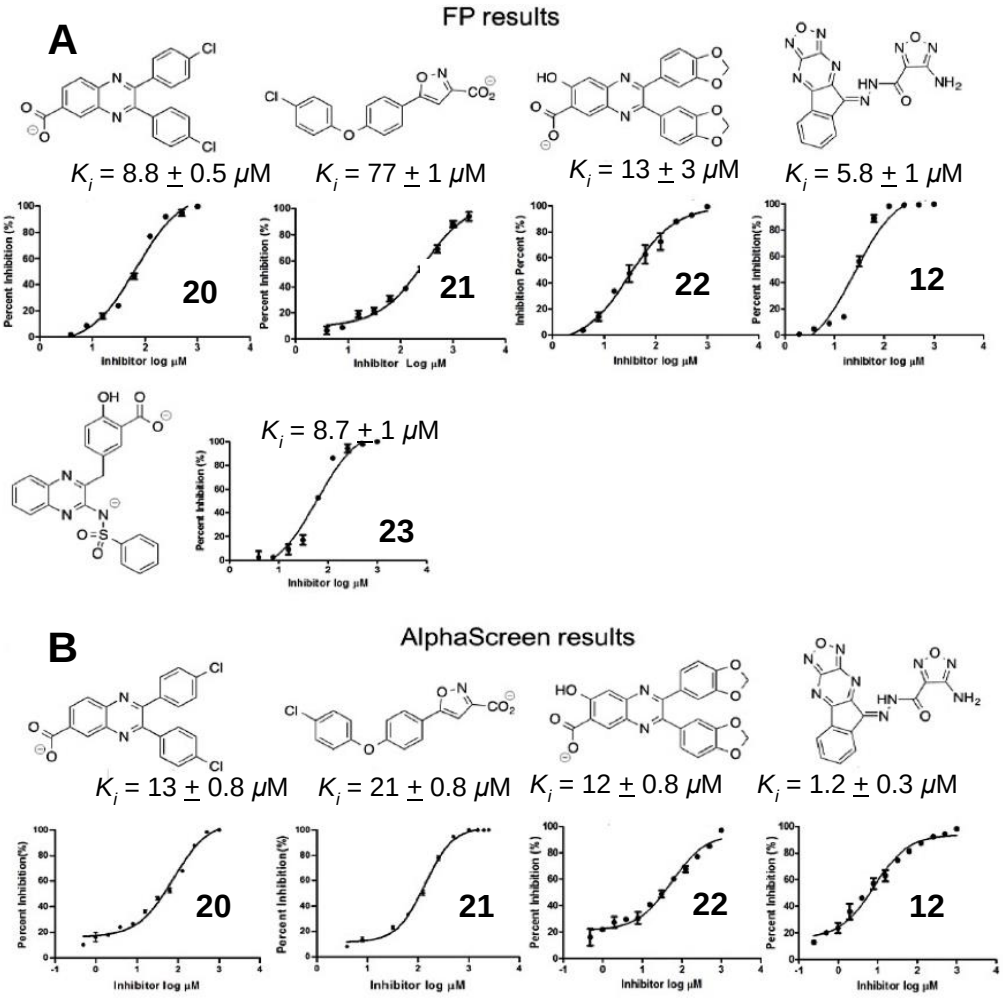
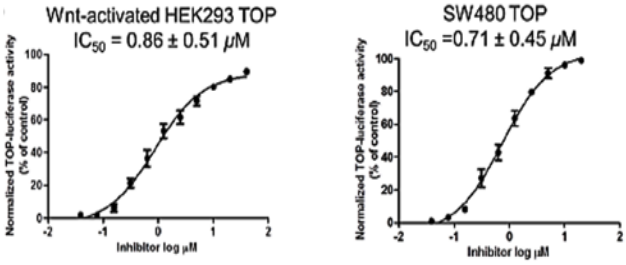


Figure 2.3. Top high-throughput screening compounds FP and AlphaScreen assay results. (A) FP assay results for top tested compounds. (B) AlphaScreen results of top compounds. Compound **23** was not tested in AlphaScreen because of the potential to interfere with the assay results. Adapted with permission from Catrow, J. L.; Zhang, Y.; Zhang, M.; Ji, H. Discovery of Selective Small-Molecule Inhibitors for the β-Catenin/T-Cell Factor Protein–Protein Interaction through the Optimization of the Acyl Hydrazone Moiety. *J. Med. Chem*. **2015**, *58* (11), 4678-4692. Copyright 2015 American Chemical Society.

A



B

TOPFlash Luciferase results IC_{50} ± S.D. (μM)		
Compound	Wnt-activated HEK293	SW480
12	0.86 ± 0.5	0.71 ± 0.5
20	> 1.0 x 10^2	n.d.
21	39 ± 6	n.d.
22	> 1.0 x 10^2	n.d.

C

MTS cell viability results IC_{50} ± S.D. (μM)			
Compound	SW480	HCT116	HT29
12	0.85 ± 0.1	0.99 ± 0.1	1.1 ± 0.1
20	38 ± 1	22 ± 6	41 ± 1
21	> 4.0 x 10^2	> 4.0 x 10^2	> 4.0 x 10^2
22	1.1 x 10^2 ± 20	2.5 x 10^2 ± 20	3.0 x 10^2 ± 20

D

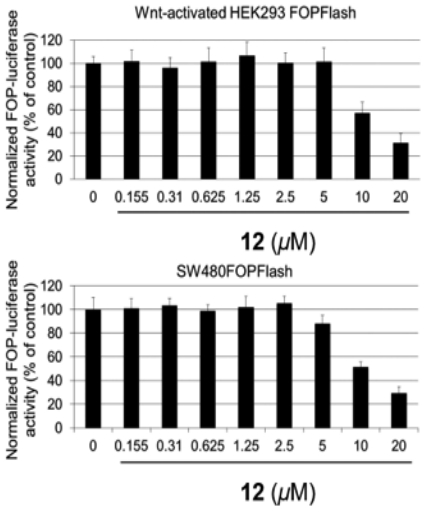


Figure 2.4. Cell-based assay results of top compounds. (A,B) Compound **12** exhibited the best inhibition of Wnt signaling in the TOPflash luciferase assay. (C) Compound **12** exhibited the best inhibition of cancer cell growth in all three tested cell lines. (D) Compound **12** inhibits FOPFlash luciferase activity. Adapted with permission from Catrow, J. L.; Zhang, Y.; Zhang, M.; Ji, H. Discovery of Selective Small-Molecule Inhibitors for the β-Catenin/T-Cell Factor Protein–Protein Interaction through the Optimization of the Acyl Hydrazone Moiety. *J. Med. Chem.* **2015**, *58* (11), 4678-4692. Copyright 2015 American Chemical Society.

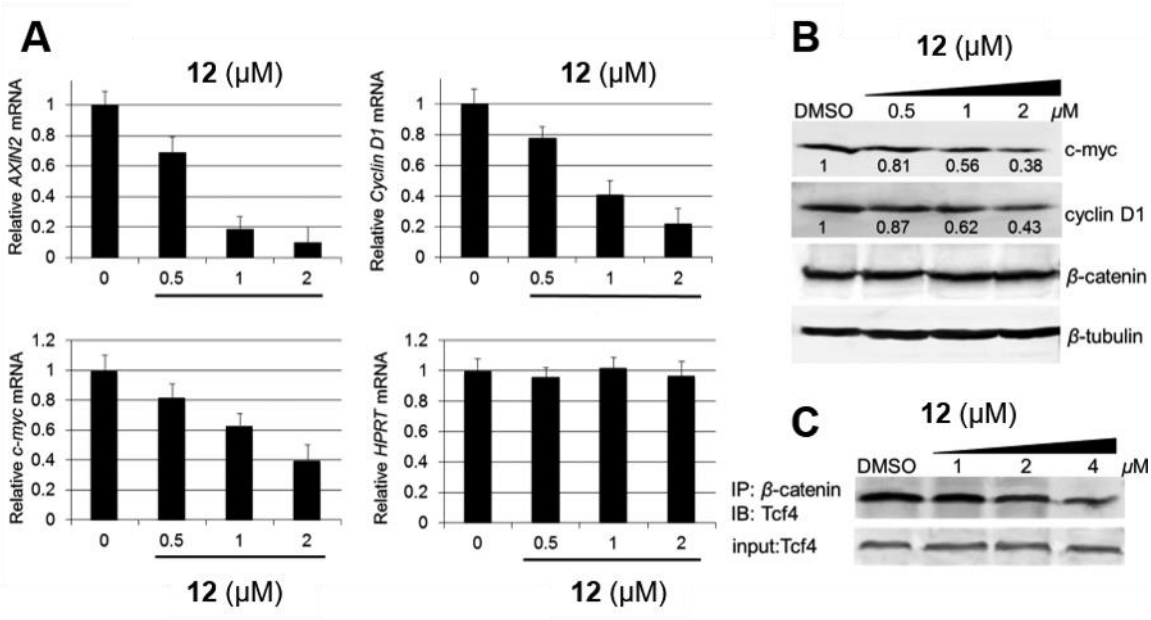


Figure 2.5. Further testing of **12**. (A) RT-qPCR results of **12**. Compound **12** inhibited Wnt target gene transcription of all three tested genes (B) Western blot analysis of c-myc and cyclin D1 protein concentrations in the presence of **12**. (C) Coimmunoprecipitation analysis of the ability of **12** to inhibit the β-catenin/Tcf PPI. Adapted with permission from Catrow, J. L.; Zhang, Y.; Zhang, M.; Ji, H. Discovery of Selective Small-Molecule Inhibitors for the β-Catenin/T-Cell Factor Protein–Protein Interaction through the Optimization of the Acyl Hydrazone Moiety. *J. Med. Chem.* **2015**, *58* (11), 4678-4692. Copyright 2015 American Chemical Society.

Wnt target genes, *cyclin D1* and *c-myc*, and β-catenin protein levels by coimmunoprecipitation and western blot analysis (Figure 2.5B). Compound **12** reduced Wnt target gene protein levels in a dose-dependent manner, but did not reduce protein levels of β-catenin or the β-tubulin control. These results indicated that **12** inhibited cellular Wnt signaling, but did not disrupt the highly important β-catenin degradation pathway. To test the ability of **12** to inhibit the cellular formation of the β-catenin/Tcf4 complex, Dr. Min Zhang performed a second coimmunoprecipitation and western blot analysis. The results of this experiment indicated that **12** inhibited the formation of the β-catenin/Tcf4 complex in SW480 cells (Figure 2.5C).

*2.3.3 Removal of PAIN substructures and optimization of **12**.* The structure of **12** is not without problems. The first is the oxadiazolopyrazine structure which is similar to the benzafurazan PAIN substructure.[30,59-61] The benzene ring of the benzafurazan moiety is potentially reactive with singlet oxygens[62] and nucleophiles.[63] However, it is unlikely that the oxadiazolopyrazine will suffer from these problems because of the lack of a benzene ring. The second PAIN substructure is the acyl hydrazone moiety.[30] The acyl hydrazine moiety has been reported to be potentially reactive and cause frequent false positives in biochemical assays.[64-66] The N-N single bond of the acyl hydrazone may be reactive with nucleophiles.[64-66] However, **12** contains an intramolecular hydrogen bond between a nitrogen atom of the oxadiazolopyrazine and the hydrogen atom of the acyl hydrazone that may stabilize the structure. Lastly, DNA intercalation is a concern with compounds containing large, planar heterocylic rings, but the hydrophobicity of the oxadiazolopyrazine and low polarizability of the tetracyclic ring make **12** an unlikely DNA intercalator.[67,68]

Dr. Yongqiang Zhang performed a fragmentation study on compound **12** in order to better understand the manner in which **12** was able to inhibit the β-catenin/Tcf4 PPI. Compounds **24** and **25** were synthesized (Figure 2.6A), and **26**-**30** were purchased commercially. A condensation reaction between ninhydrin and 3,4-diaminofurazan resulted in fragment **23**. Fragment **24** was generated by a condensation reaction between **24** and hydrazine hydrate in glacial acetic acid. Compounds **24-30** were tested by the FP assay by Dr. Min Zhang (Figure 2.7). While most fragments exhibited some inhibitory activity, all were significantly less potent than **12**.

The optimization of **12** began with the resynthesis of **12** by Dr. Yongqiang Zhang (Figure 2.6B). To resynthesize **12**, first 4-amino-1,2,5-oxadiazole-3-carboxylic acid was refluxed in methanol under acidic conditions to generate the methyl ester **42**. Hydrazinolysis of **42** generated **43**. A condensation reaction between **24** and **43** afforded compound **12** with moderate yield. The resynthesized **12** achieved K_i values of 7 $\pm$ 4 μM and 1.7 $\pm$ 3 μM in the FP and AlphaScreen assays, respectively. A blind docking of **12** was conducted on the surface of β-catenin using AutoDock Vina[47] to determine the possible binding sites. AutoDock Vina scored the top poses for hot regions 1, 2, and 3, at -8.4, -4.6, and -3.7 kcal mol^{-1}, respectively. Therefore, it was predicted that **12** preferentially bound hot region 1. Compound **12** was docked into hot region 1 using AutoDock Vina, AutoDock4,[51] and Schrödinger Glide[53] to examine the binding mode. One binding mode recurred in all instances (Figure 2.8A). Most importantly, the 4-amino-1,2,5-oxadiazole ring was predicted to act as a hydrogen bond acceptor for the side-chain of hot spot K435 and the backbone NH group of N430. The oxadiazolopyrazine ring nitrogen atoms formed hydrogen bond interactions with hot spot

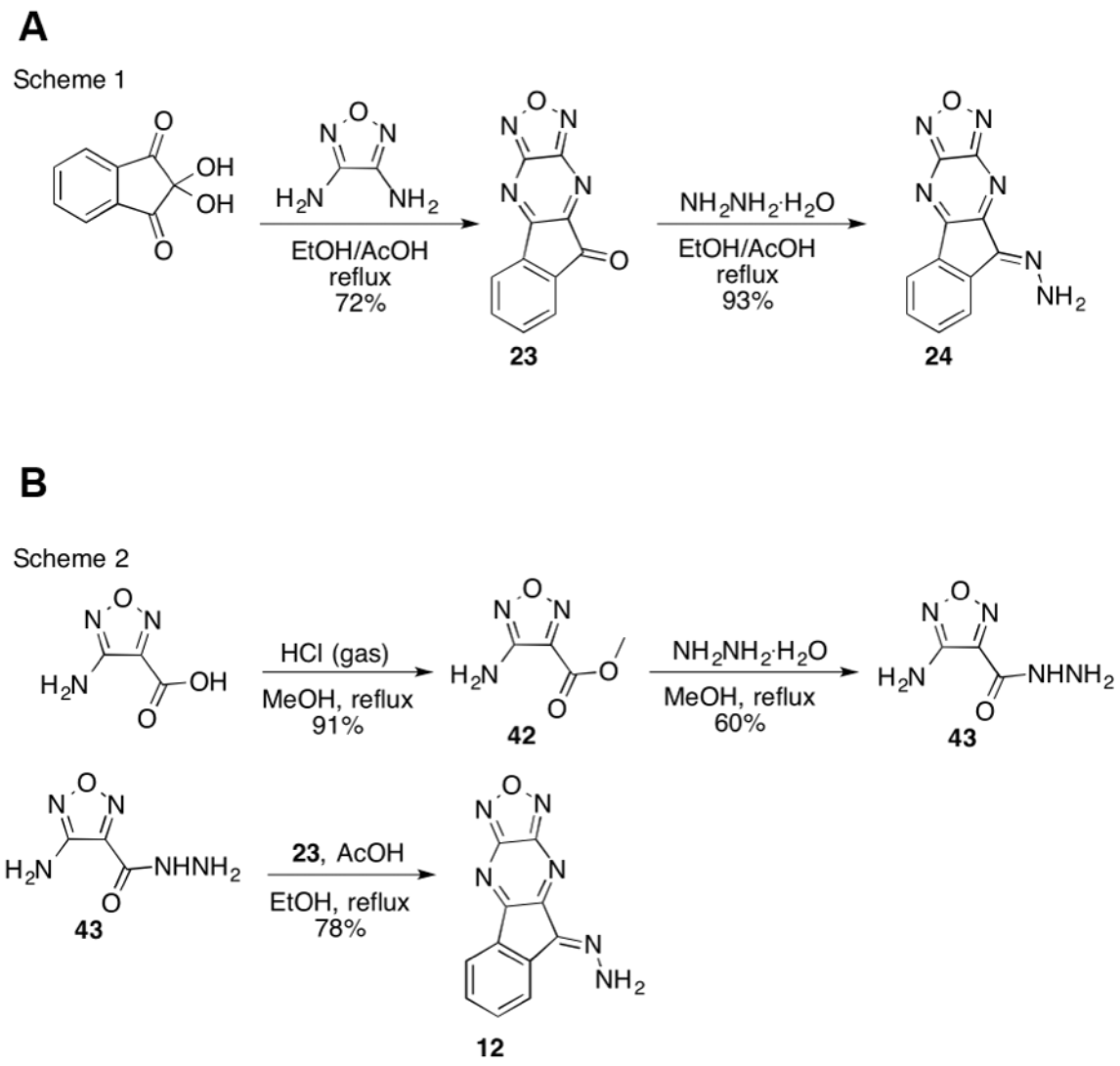


Figure 2.6. Synthetic schemes 1 and 2. Adapted with permission from Catrow, J. L.; Zhang, Y.; Zhang, M.; Ji, H. Discovery of Selective Small-Molecule Inhibitors for the β-Catenin/T-Cell Factor Protein–Protein Interaction through the Optimization of the Acyl Hydrazone Moiety. *J. Med. Chem.* **2015**, *58* (11), 4678-4692. Copyright 2015 American Chemical Society.

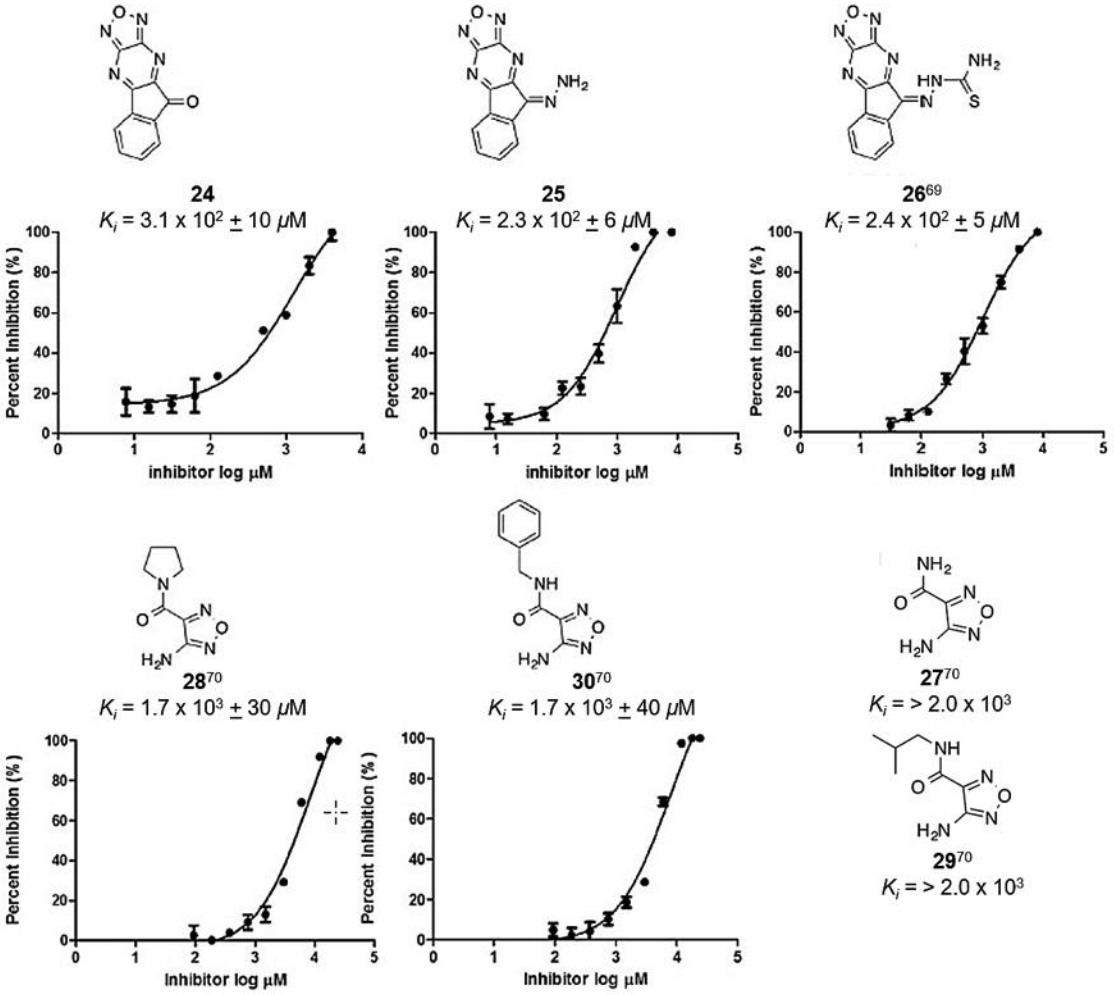
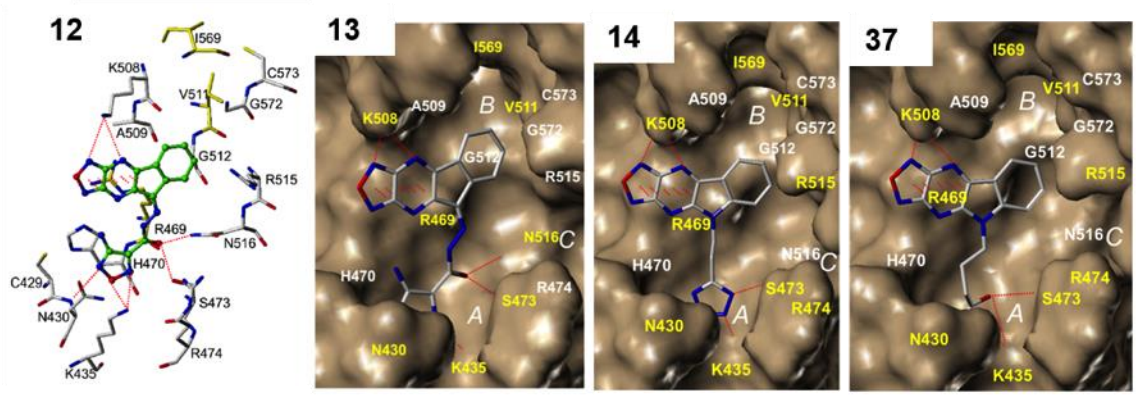


Figure 2.7. FP assay results of fragmentation study. Adapted with permission from Catrow, J. L.; Zhang, Y.; Zhang, M.; Ji, H. Discovery of Selective Small-Molecule Inhibitors for the β-Catenin/T-Cell Factor Protein–Protein Interaction through the Optimization of the Acyl Hydrazone Moiety. *J. Med. Chem.* **2015**, *58* (11), 4678-4692. Copyright 2015 American Chemical Society.

A



B

Site-directed mutagenesis FP assay K_i ± S.D. (μM)				
	12	**13**	**14**	**37**
WT	7.0 ± 4	7.2 ± 5	17 ± 6	3.4 ± 3
V511S	20 ± 2	31 ± 4	53 ± 4	17 ± 2
V511S/I569S	36 ± 3	88 ± 9	1.8 X 10^2 ±6	59 ± 4
R469A	46 ± 4	42 ± 6	52 ± 8	42 ± 4

Figure 2.8. Docking and site-directed mutagenesis studies of inhibitors. (A) Results of Glide docking of inhibitors **12-14** and **37** representing the progression of inhibitor development (B) FP results of mutant β-catenins with compounds **12-14** and **37**. Adapted with permission from Catrow, J. L.; Zhang, Y.; Zhang, M.; Ji, H. Discovery of Selective Small-Molecule Inhibitors for the β-Catenin/T-Cell Factor Protein–Protein Interaction through the Optimization of the Acyl Hydrazone Moiety. *J. Med. Chem.* **2015**, *58* (11), 4678-4692. Copyright 2015 American Chemical Society.

K508. Other predicted interactions included a hydrogen bond between the amide carbonyl group of **12** and the side chains of N516 and S473. The large, heterocyclic ring structure was predicted to from a cation-π interaction with the side-chain of R469. The benzene ring was predicted to project into hydrophobic pocket B. Pocket B is not utilized by the β-catenin/Tcf4 interaction. Therefore, pocket B represents an excellent region to perform site-directed mutagenesis studies to evaluate the binding mode. Dr. Min Zhang created pocket B B-catenin mutants R469A, V511S, and V511/I569S to test the binding mode of **12** (Figure 2.8B).[16,39] FP binding assays using the mutant β-catenin and fluorescein-labeled Tcf4 indicated no change in the apparent K_d values of the mutant proteins. Resynthesized **12** achieved K_i values of 46 $\pm$ 4, 20 $\pm$ 2 μM, and 36 $\pm$ 3 μM against R469A, V511S, and V511S/I569S mutant β-catenin/Tcf4 PPIs, respectively. These results indicate that **12** had some interaction with hot region 1. However, the acyl hydrazone PAIN substructure needed to be removed before further optimization could occur. Compound **13** was synthesized by Dr. Yongqiang Zhang to generate a small molecule similar to **12** but without the acyl hydrazone PAIN substructure. Isatin and 3,4-diaminofurazan were refluxed in glacial acetic acid to yield **31** by a condensation reaction. Compound **32** was synthesized by reacting **31** with ethyl bromoacetate, and the ethyl ester was hydrolyzed to yield **36**. Compound **13** was generated by peptide coupling of 3,4-diaminofurazan and **33**. The potentially reactive imine carbon at the dibenzylic position of **12** was replaced with a nitrogen atom. This new compound, **13**, is made more stable by completing an aromatic system within the heterocyclic ring yet is structurally similar to **12**. The atoms responsible for hydrogen bonds with K435, K508 and the cation-π interaction with R469 were retained. Compound **13** had similar K_i values to **12**

when tested by Dr. Min Zhang in the FP and AlphaScreen assays, 7 ± 5 μM and 2.7 ± 0.5 μM, respectively. FP assay analysis of **13** with the previously described mutant β-catenin yielded similar results as **12** (Figure 2.8B), indicating that **13** bound to hot region 1.

Once the acyl hydrazone PAIN substructure was removed, a series of inhibitors were designed to optimize the interaction between 4-amino-1,2,5-oxadiazole portion of **13** and K435 (Figure 2.9). Because of the positively-charged amino group of K435, this series focused on creating compounds with carboxylic acids and tetrazoles, a carboxylic acid bioisostere, with a 1-4 carbon alkyl linker. An aza-Michael addition to **31** generated **34** and **39**, and nucleophilic substitution to **31** generated **34**, **36**, **38**, and **40**. Carboxylic acids **35** and **37** were generated by hydrolysis of **34** and **36**, respectively. Tetrazoles **14** and **41** were generated by a [2 + 3] cycloaddition reaction with **39** and **40**, respectively. The ester compounds **32**, **34**, and **36** and nitrile compounds **38-40** showed poor ability to inhibit the β-catenin/Tcf4 PPI in the FP assay.

Compounds **14**, **33**, **35**, **37**, and **41** inhibited the β-catenin/Tcf4 PPI in biochemical assays (Table 2.1). Of the new inhibitors, carboxylic acid **36** proved the most potent inhibitor in the FP and AlphaScreen assays, with K_i values of 3 ± 1 μM and 1.0 ± 0.1 μM, respectively. Compound **14** was the most potent tetrazole compound tested with FP (17 ± 2 μM) and AlphaScreen (18 ± 2 μM).

Compounds **14** and **37** were docked with AutoDock Vina, Schrödinger Glide, and AutoDock4 to evaluate the binding modes. The Glide docking results are shown in Figure 2.10A. The docking modes of the new compounds are predicted to be similar to **12**. The hydrogen bonds between the oxadiazolopyrazine ring and K508 were retained, as was the cation-π interaction with R469. The carboxylic acid moiety of **37** and tetrazole moiety of

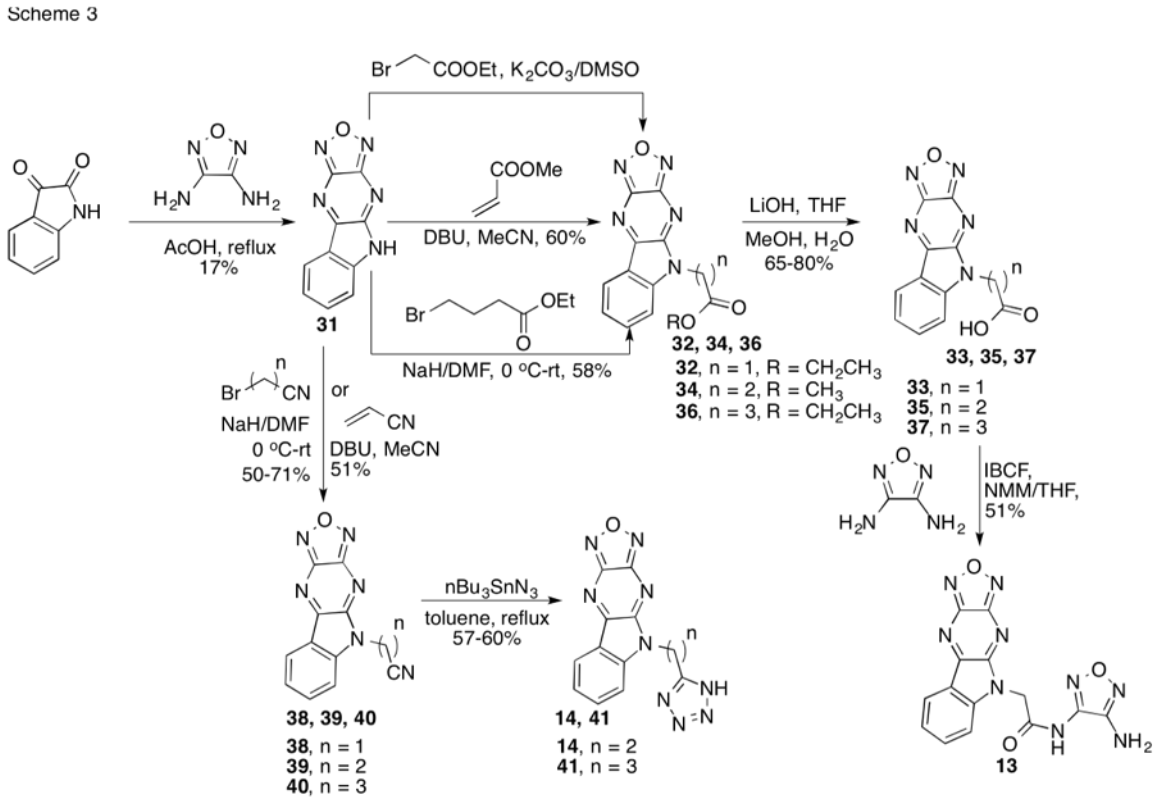


Figure 2.9. Synthetic scheme 3. Adapted with permission from Catrow, J. L.; Zhang, Y.; Zhang, M.; Ji, H. Discovery of Selective Small-Molecule Inhibitors for the β-Catenin/T-Cell Factor Protein–Protein Interaction through the Optimization of the Acyl Hydrazone Moiety. *J. Med. Chem.* **2015**, *58* (11), 4678-4692. Copyright 2015 American Chemical Society.

Table 2.1 FP and AlphaScreen assay results of new compounds (Ki ± S.D. [μM]).

Compound	FP result	AlphaScreen result
12	7.0 ± 4	1.7 ± 0.3
13	7.2 ± 5	2.7 ± 5
14	17 ± 2	18 ± 2
31	2.8×10^2 ± 10	1.2×10^2 ± 9
32	1.5×10^2 ± 8	63 ± 5
33	5.8 ± 2	4.1 ± 0.5
34	1.0×10^2 ± 9	1.0×10^2 ± 3
35	11 ± 4	3.30 ± 0.6
36	1.2×10^2 ± 4	52 ± 2
37	3.4 ± 1	1.0 ± 0.1
38	2.4×10^2 ± 10	$> 13 \times 10^2$
39	47 ± 5	39 ± 2
40	1.6×10^2 ± 9	1.0×10^2 ± 4
41	51 ± 13	74 ± 2

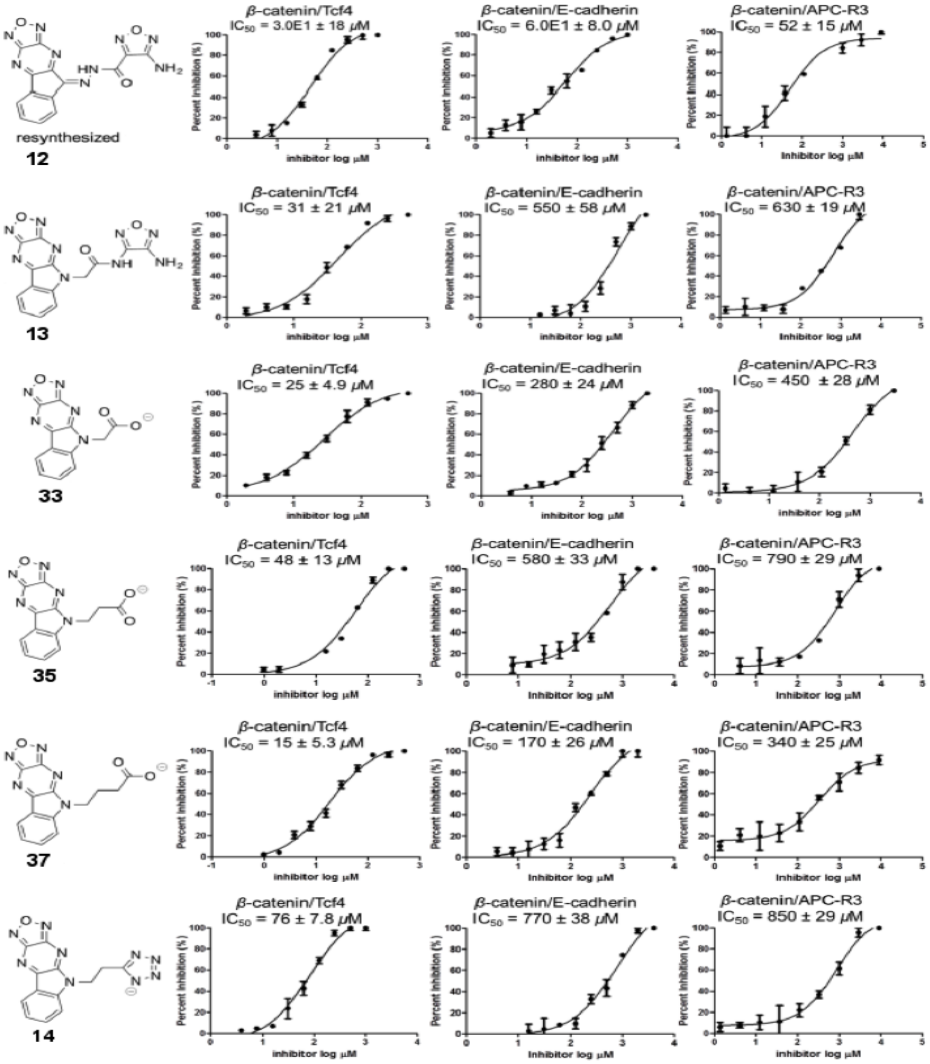


Figure 2.10**.** FP assay results for β-catenin/Tcf4, β-catenin/E-cadherin and β-catenin/APC PPIs. All derivatives of **12** were selective for the β-catenin/Tcf4 PPI. Adapted with permission from Catrow, J. L.; Zhang, Y.; Zhang, M.; Ji, H. Discovery of Selective Small-Molecule Inhibitors for the β-Catenin/T-Cell Factor Protein–Protein Interaction through the Optimization of the Acyl Hydrazone Moiety. *J. Med. Chem.* **2015**, *58* (11*)*, 4678-4692. Copyright 2015 American Chemical Society.

14 were predicted to interact strongly with hot spot K435. Hydrogen bonds with the backbone amide hydrogen of H470 and side chain hydroxyl of S473, and the hydrophobic interaction between the benzene and pocket B were also maintained. FP assays using the previously described mutant β-catenin were used to validate the binding modes of **14** and **37**. Both **14** and **37** were responsive to the site-directed mutagenesis β-catenin mutants in the FP assay (Figure 2.8B).

Compounds **12**, **13**, **14**, **33**, **35**, and **37** were evaluated for selectivity for the β-catenin/Tcf4 PPI over the β-catenin/E-cadherin and β-catenin/APC PPIs. Selectivity was evaluated through the FP assay utilizing C-terminally fluorescein-labeled human Tcf4 (residues 7-51), C-terminally fluoroscein-labeled human E-cadherin (residues 819-873), and C-terminally fluoroscein-labeled human APC (residues 1477-1519). With the exception of **12**, all compounds tested showed good selectivity for the β-catenin/Tcf4 PPI (Figure 2.10). Compound **12** was not as selective as the other compounds, because the acyl hydrazone PAIN substructure caused nonspecific interactions with other hot regions. While these results are promising for the selective inhibition of the β-catenin/Tcf4 PPI, further study is required to better understand the selective nature of the inhibitors.

*2.3.4 Biological characterization of compound **12** derivatives.* Dr. Min Zhang evaluated **13**, **14**, **34**, **36**, and **37** with a TOPFlash Wnt reporter gene assay using pcDNA3.1-β-catenin transfected HEK293 cells. Only **14** downregulated the transcription of Wnt target genes in a dose-dependent manner (Figure 2.11A). The Wnt-independent FOPFlash luciferase activity remained unaffected even at high concentrations of **1**(Figure 2.10B). Similar results were observed when Dr. Min Zhang tested compounds **13**, **14**, **34**, **36,** and **37** for inhibition of colorectal cancer cell growth through the MTS cell viability

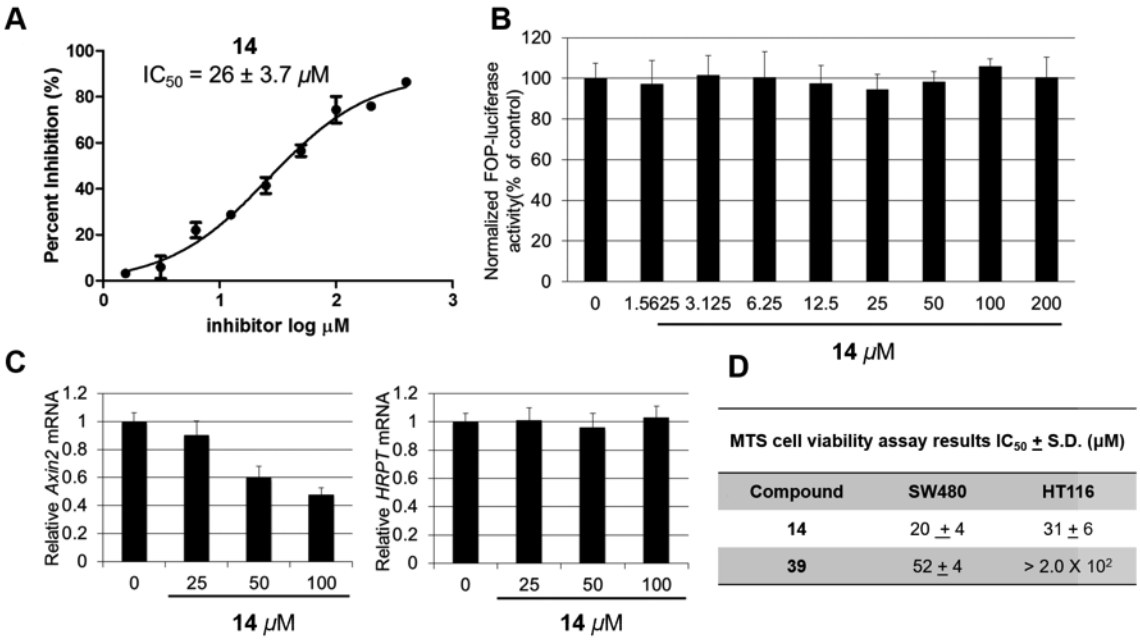


Figure 2.11. Cell-based assay testing of **14**. (A) TOPFlash luciferase results for **14**. (B) FOP-luciferase control. Compound **14** does not show off-target activity even at high doses. (C) RT-qPCR results for **14**. Compound **14** inhibits Wnt target gene transcription, while the HRPT transcription remains unaffected. (D) MTS cell viability assay for **14** and **39**. Compound **14** inhibited the growth of both SW480 and HT115 colorectal cancer cells. Adapted with permission from Catrow, J. L.; Zhang, Y.; Zhang, M.; Ji, H. Discovery of Selective Small-Molecule Inhibitors for the β-Catenin/T-Cell Factor Protein–Protein Interaction through the Optimization of the Acyl Hydrazone Moiety. *J. Med. Chem.* **2015**, *58* (11), 4678-4692. Copyright 2015 American Chemical Society.

assay. Again, only **14** inhibited growth of SW480 and HCT116 cell lines with an IC_{50} of 20 $\pm$ 4 μM and 31 $\pm$ 6 μM, respectively (Figure 2.11D). Compound **39** inhibited the growth of SW480 cells but not HCT116 cells (Figure 2.11D). Additionally, **14** downregulated *Axin2* gene transcription when examined by RT-qPCR (Figure 2.11C). The inactivity of the other compounds may be due to low cell permeability in the tested cell lines.

To evaluate the ability of these compounds to inhibit the Wnt signaling pathway in vivo, Dr. David Hutcheson and Janelle Tardif tested **36** in a transgenic Wnt reporter zebrafish model (Figure 2.12).[71,72] The *headless* mutation in zebrafish has been linked to the canonical Wnt signaling pathway.[73] Dorsky et al.[72] developed a transgenic zebrafish line (TOPdGFP) which expresses destabilized green fluorescent protein (dGFP) when canonical Wnt signaling is active. Therefore, a positive result for the disruption of Wnt signaling is reduced expression of dGFP in the head and eye region. Zebrafish embryos were exposed to 2 μM and 10 μM of **36**. The zebrafish embryos exposed to 10 μM of **36** showed markedly less development in the eye and brain. Transgenic embryos exposed to 2 μM and 10 μM of **36** showed a dose-dependent response. While these results are quite preliminary, they highlight the ability of **36** to disrupt the Wnt signaling pathway in vivo.

2.3.5 Further development of biologically active B-*catenin/Tcf4 PPI inhibitors.* One of the largest problems facing **31** is the low synthetic yield, 15-18%. This was hypothesized to be due to the electron-poor nature of 3,4-diaminofurazan. To remedy this, the starting material 3,4-diaminofurzan was replaced with ortho-phenylenediamine (Figure 2.13).[74] The resulting reaction had a dramatically increased yield of 80%. Because the tetrazole-containing compound **14** was the only compound active in cancer

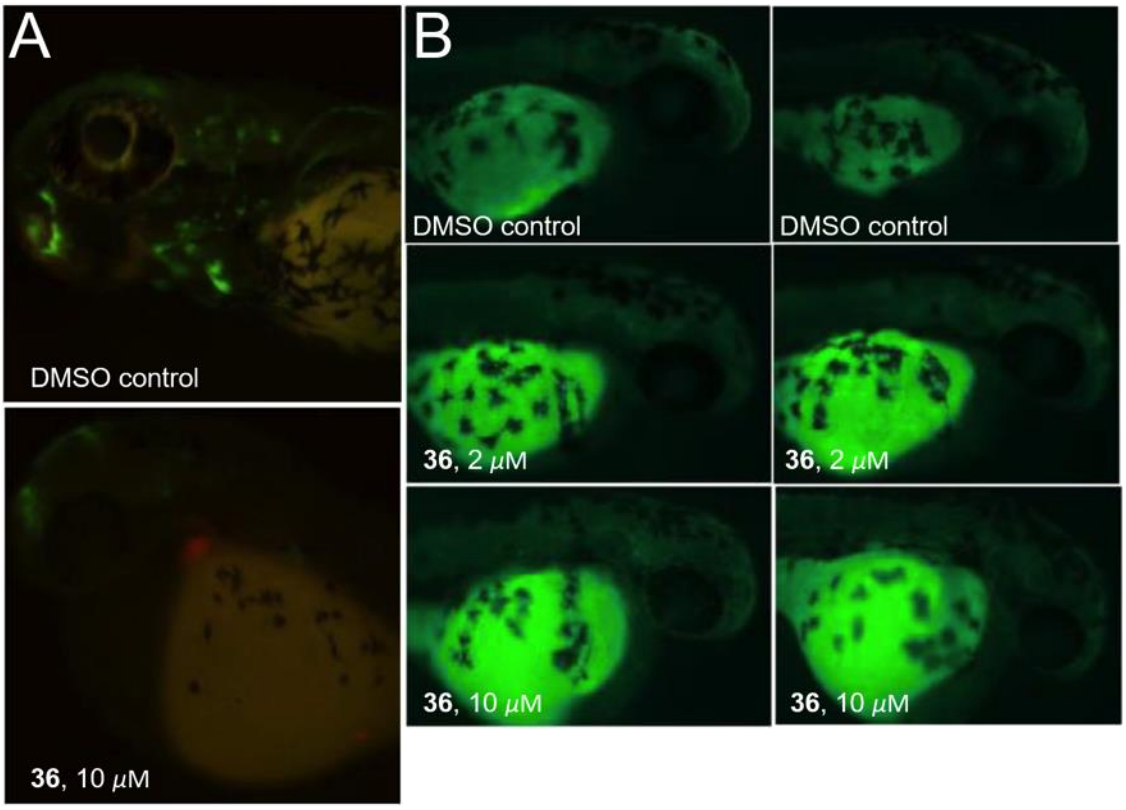


Figure 2.12. Results of compound **36** on transgenic zebrafish model. (A) Developmental effects on the head and eye region as a result of treatment with 10 μM **36**. (B) Expression of dGFP in TOPdGFP zebrafish embryos exposed to varying concentrations of **36**. Images taken by the Dorsky lab.

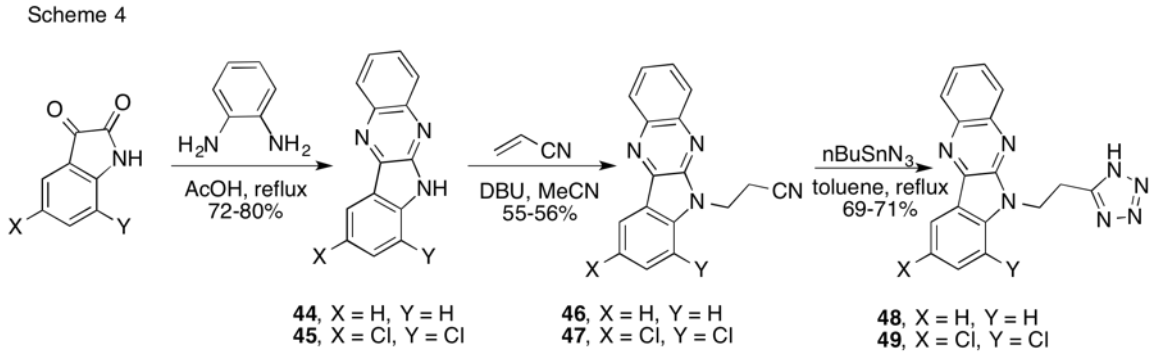


Figure 2.13. Synthetic scheme 4.

cells, compounds **48** and **49** were created. Compounds **48** and **49** were synthesized using the same synthetic route as **14**.

A ligand-based structure activity relationship study was planned for this series of inhibitors to explore adding additional hydrophobic interactions to pocket B. Undergraduate ACCESS fellow Brooklyn Brakey synthesized **49** for this study. Compounds **48** and **49** were tested by Dr. Min Zhang in the FP assay and obtained K_i values of 66 $\pm$ 7 μM and 12 $\pm$ 1 μM, respectively (Figure 2.14A). Compound **48** had reduced activity compared to **14**. This reduction is likely due to the loss of a hydrogen bond from the furazan nitrogen with K508. The addition of the chlorine at the 5-position of the isatin ring is predicted by AutoDock Vina to interact with pocket B, resulting in increased activity (Figure 2.14B). These final two compounds were submitted to Dr. Katrin Guillen of the Welm lab for testing on mammary epithelial cells (MECs)[75] (Figure 2.15). MECS were embedded in matrigel (4 μL base, 8 μL dome, 200 agg/dom) and treated with compounds **14**, **37**, **48,** and **49** 48 h post embedding and again at 96 h post-embedding. The effects of these compounds on the cells were scored 144 h post-embedding. Compound **48** produced the best result with all of the cells in either a cyst or solid state. Compound **49** induced the solid extruding cell type, indicating mild cytotoxicity. Compounds **14** and **37** had less activity. Despite the promising results in the TOPdGFP zebrafish and primary cell MEC 3D screen, there is still much to do in order to optimize this series of inhibitors.

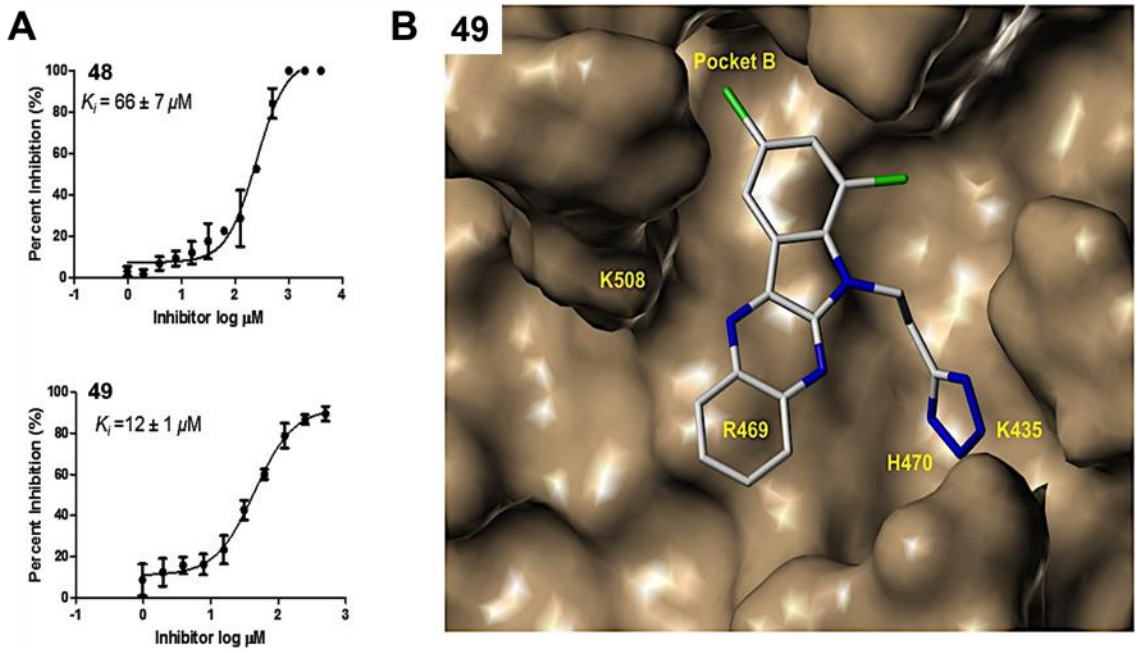


Figure 2.14. Results of **48** and **49**. (A) Compounds **48** and **49** show an inhibitory effect on the β-catenin/Tcf PPI. (B) Binding mode proposed for **49**.

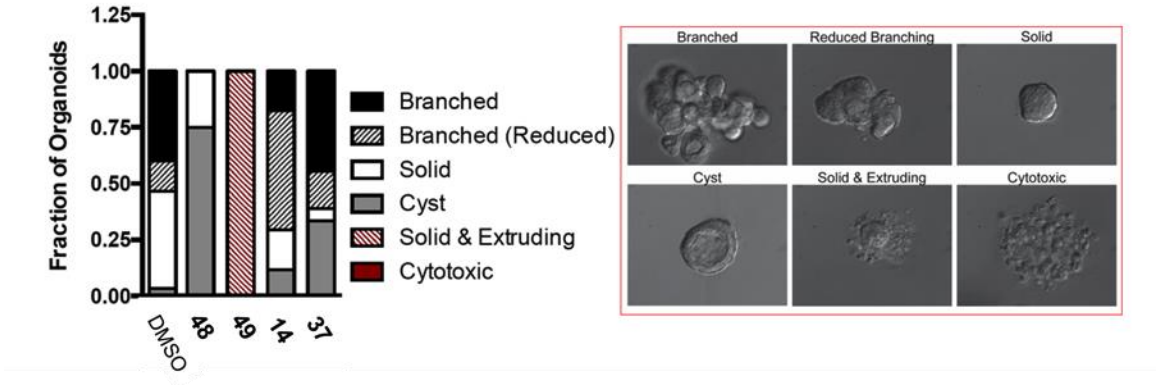


Figure 2.15. Results of testing compounds on MECs. Compound **45** proved most capable of inducing the cyst state in MECs. Images taken by Dr. Katrin Guillen.

2.4 Discussion

Because the aberrant activation of the Wnt signaling pathway is responsible for many diseases, there is an urgent need to develop selective, small-molecule Wnt inhibitors.[9] The β-catenin/Tcf4 PPI is the penultimate step of the Wnt signaling pathway and represents an appealing target for small-molecule Wnt inhibitors. Because E-cadherin and APC overlap the Tcf4 binding site on β-catenin, a small molecule inhibitor of the β-catenin/Tcf4 PPI must be selective for inhibition of the β-catenin/Tcf4 over the β-catenin/E-cadherin and β-catenin/APC PPIs. This selectivity is important because a small molecule that disrupts the β-catenin/E-cadherin or β-catenin/APC PPIs risks disrupting cell adhesion networks or the β-catenin degradation pathway.

Finding effective, selective, PAIN substructure-free inhibitors of this PPI has proven difficult. In this study, FP and AlphaScreen identified **12** as an inhibitor of the β-catenin/Tcf PPI. Compound **12** was effective at inhibiting Wnt target gene activation and cancer cell growth. However, **12** contained the acyl hydrazone PAIN substructure. The acyl hydrazone moiety has been implicated as causing false positives and is possibly reactive with nucleophiles.[30] The formation of this substructure is an important component of DCC because it is reversible and controllable by pH change.[42] The acyl hydrazone moiety can also act as both a hydrogen bond donor and acceptor, thereby increasing the potency of the discovered molecule.[42] Though many biologically active acyl hydrazone-containing molecules have been reported, this substructure is only found in two clinically accepted small-molecule drugs. Therefore, a careful evaluation of any acyl hydrazone containing molecule must be made before proceeding with optimization.

The optimization of **12** presented a path for removing the acyl hydrazone moiety

while maintaining inhibitor potency. By replacing the acyl hydrazone with a peptide bond, the potentially unstable N-N single bond was removed. Docking studies of **12** and **13** ensured that a similar predicted binding mode was maintained even after the acyl hydrazone was removed. These compounds interacted with hot region 1, which has been previously established as a site to selectively inhibit the β-catenin/Tcf4 PPI.

The further optimization of **13** was performed to generate compounds **14** and **37**. Compounds **14** and **37** were predicted to form charge-charge interactions with the β-catenin hot spot K435 and increase inhibitor potency and cell-based activity. This hot spot, which interacts with D16 of human Tcf4, was shown to be crucial for the β-catenin/Tcf4 PPI. Compounds **13**, **14,** and **37** achieved increased selectivity for the β-catenin/Tcf4 PPI compared to **12**. This increased selectivity is likely a result of less promiscuous binding due to the removal of the acyl hydrazone PAIN substructure. Compound **14** also downregulated Wnt target gene transcription in a TOPFlash luciferase reporter gene assay and inhibited the growth of Wnt-active colorectal cancer cells.

Compound **44** was created to overcome the low yield of the first step in the synthesis of **31**. Compound **49** was synthesized from **44** to increase interaction with hydrophobic pocket B of hot region 1. Compound **49** showed increased inhibition of the β-catenin/Tcf4 PPI compared to **14** in biochemical testing. Additionally, **48** and **49** were effective at inhibiting the growth of MECs.

2.5 Conclusion

Small-molecule inhibition of PPIs is a challenging field of medicinal chemistry. The β-catenin/Tcf4 PPI is one particularly difficult drug target because of the very large

contacting surfaces involved.[11] Selective inhibition of the β-catenin/Tcf4 PPI is complicated by binding site overlap between Tcf4, E-cadherin and APC. These three proteins share three hot regions on β-catenin.[22] However, hot region 1, containing key lysine residues K435 and K508, is the most important hot region for Tcf4 binding. Previous studies with UUT-02 have shown that inhibitors binding this hot region are capable of selective inhibition of the β-catenin/Tcf4 PPI.[39]

Biochemical and cell-based assays identified **12** as a selective, drug-like, small-molecule inhibitor of the β-catenin/Tcf4 PPI. Site-directed mutagenesis assays confirmed that **12** bound to hot region 1 of β-catenin. The acyl hydrazone PAIN substructure was identified in **12**. Compound **12** was optimized to generate **13**, **14**, **37**, **48,** and **49** which lack the acyl hydrazone substructure. Compounds **13**, **14,** and **37** achieved greater selectivity for the β-catenin/Tcf4 PPI than **12**. Site-directed mutagenesis confirmed that these inhibitors utilized a similar binding mode in hot region 1 as **12**. This further confirms the use of hot region 1 of β-catenin as a small molecule binding site to selectively inhibit the β-catenin/Tcf4 PPI. This series of inhibitors has also exhibited the ability to inhibit Wnt signaling *in vivo*, as **35** was shown to inhibit Wnt signaling in zebrafish, and **48** inhibited growth of MEC cells.

Even though the selectivity is still lower than what could be desired, it should be noted that in the biochemical assays, Tcf4 is the tightest binding partner.[15,23,24] Therefore, the selectivity of these molecules is significant. Further optimization could increase binding to hot region 1 and increase selectivity. Compound **49** was used to explore this by utilizing Cl atoms to occupy pocket B. In addition to further optimizing the pocket B interactions, the use of pockets C and D could yield even more potent inhibitors.

2.6. References

1. Clevers, H.; Nusse, R. Wnt/β-Catenin Signaling and Disease. *Cell* **2012**, *149*, 1192–1205.

2. Klaus, A.; Birchmeier, W. Wnt Signaling and its Impact on Development and Cancer. *Nature Rev. Cancer* **2008**, *8*, 387–398.

3. Malanchi, I.; Peinado, H.; Kassen, D.; Hussenet, T.; Metzger, D.; Chambon, P.; Huber, M.; Hohl, D.; Cano, A.; Birchmeier, W.; Huelsken, J. Cutaneous Cancer Stem Cell Maintenance is Dependent on β-Catenin Signaling. *Nature* **2008**, *452*, 650–653.

4. Barker, N.; Ridgway, R. A.; van Es, J. H.; van de Wetering, M.; Begthel, H.; van den Born, M.; Danenberg, E.; Clarke, A. R.; Sansom, O. J.; Clevers, H. Crypt Stem Cells as the Cells-Of-Origin of Intestinal Cancer. *Nature* **2009**, *457*, 608–611.

5. Yeung, J.; Esposito, M. T.; Gandillet, A.; Zeisig, B. B.; Griessinger, E.; Bonnet, D.; So, C. W. E. β-Catenin Mediates the Establishment and Drug Resistance of MLL Leukemic Stem Cells. *Cancer Cell* **2010**, *18*, 606–618.

6. He, T.-C.; Sparks, A. B.; Rago, C.; Hermeking, H.; Zawel, L.; da Costa, L. T.; Morin, P. J.; Vogelstein, B.; Kinzler, K. W. Identification of C-MYC as a Target of the APC Pathway. *Science* **1998**, *281*, 1509–1512.

7. Tetsu, O.; McCormick, F. β-Catenin Regulates Expression of Cyclin D1 in Colon Carcinoma Cells. *Nature* **1999**, *398*, 422–426.

8. Kim, P. J.; Plescia, J.; Clevers, H.; Fearon, E. R.; Altieri, D. C. Survivin and Molecular Pathogenesis of Colorectal Cancer. *Lancet* **2003**, *362*, 205–209.

9. Kim, J.-S.; Crooks, H.; Foxworth, A.; Waldman, T. Proof-of-Principle: Oncogenic β-Catenin is a Valid Molecular Target for the Development of Pharmacological Inhibitors. *Mol. Cancer Ther.* **2002**, *1*, 1355–1359.

10. Sampietro, J.; Dahlberg, C. L.; Cho, U. S.; Hinds, T. R.; Kimelman, D.; Xu, W. Crystal Structure of a β-Catenin/BCL9/Tcf4 Complex. *Mol. Cell* **2006**, *24*, 293–300.

11. Wells, J. A.; McClendon, C. L. Reaching for High-Hanging Fruit in Drug Discovery at Protein–Protein Interfaces. *Nature* **2007,** *450* (7172), 1001–1009.

12. Graham, T. A.; Ferkey, D. M.; Mao, F.; Kimelman, D.; Xu, W. Tcf4 Can Specifically Recognize β-Catenin Using Alternative Conformations. *Nature Struct. Biol.* 2001, *8*, 1048–1052.

13. Poy, F.; Lepourcelet, M.; Shivdasani, R. A.; Eck, M. J. Structure of a Human Tcf4−β-Catenin Complex. *Nature Struct. Biol.* **2001**, *8*, 1053–1057.

14. Graham, T. A.; Weaver, C.; Mao, F.; Kimelman, D.; Xu, W. Crystal Structure of a β-Catenin/Tcf Complex. *Cell* **2000**, *103*, 885–896.

15. Knapp, S.; Zamai, M.; Volpi, D.; Nardese, V.; Avanzi, N.; Breton, J.; Plyte, S.; Flocco, M.; Marconi, M.; Isacchi, A.; Caiolfa, V. R. Thermodynamics of the High-Affinity Interaction of TCF4 with β-Catenin. *J. Mol. Biol.* **2001**, *306*, 1179–1189.

16. Yu, B.; Huang, Z.; Zhang, M.; Dillard, D. R.; Ji, H. Rational Design of Small-Molecule Inhibitors for β-Catenin/T-Cell Factor Protein–Protein Interactions by Bioisostere Replacement. *ACS Chem. Biol*. **2013**, *8*, 524–529.

17. von Kries, J. P.; Winbeck, G.; Asbrand, C.; Schwarz-Romond, T.; Sochnikova, N.; Dell'Oro, A.; Behrens, J.; Birchmeier, W. Hot Spots in β-Catenin for Interactions with LEF-1, Conductin and APC. *Nature Struct. Biol*. **2000**, *7*, 800–807.

18. Gail, R.; Frank, R.; Wittinghofer, A. Systematic Peptide Array-Based Delineation of the Differential β-Catenin Interaction with Tcf4, E-Cadherin, and Adenomatous Polyposis Coli. *J. Biol. Chem.* **2005**, *280*, 7107–7117.

19. Xing, Y.; Clements, W. K.; Kimelman, D.; Xu, W. Crystal Structure of a β-Catenin/Axin Complex Suggests a Mechanism for the β-Catenin Destruction Complex. *Genes Dev.* **2003**, *17*, 2753–2764.

20. Huber, A. H.; Weis, W. I. The Structure of the β-Catenin/E-Cadherin Complex and the Molecular Basis of Diverse Ligand Recognition by β-Catenin. *Cell* **2001**, *105*, 391–402.

21. Xing, Y.; Clements, W. K.; Le Trong, I.; Hinds, T. R.; Stenkamp, R.; Kimelman, D.; Xu, W. Crystal Structure of a β-Catenin/APC Complex Reveals a Critical Role for APC Phosphorylation in APC Function. *Mol. Cell* **2004**, *15*, 523–533.

22. Xu, W.; Kimelman, D. Mechanistic Insights from Structural Studies of β-Catenin and its Binding Partners. *J. Cell Sci.* **2007,** *120* (19), 3337-3344.

23. Choi, H.-J.; Huber, A. H.; Weis, W. I. Thermodynamics of β-Catenin–Ligand Interactions: the Roles of the N- And C-Terminal Tails in Modulating Binding Affinity. *J. Biol. Chem*. **2006**, *281*, 1027–1038.

24. Sun, J.; Weis, W. I. Biochemical and Structural Characterization of β-Catenin Interactions with Nonphosphorylated and CK2-Phosphorylated Lef-1. *J. Mol. Biol.* **2011**, *405*, 519–530.

25. Ha, N.-C.; Tonozuka, T.; Stamos, J. L.; Choi, H.-J.; Weis, W. I. Mechanism of Phosphorylation-Dependent Binding of APC to β-Catenin and its Role in β-Catenin Degradation. *Mol. Cell* **2004**, *15*, 511–521.

26. Choi, H.-J.; Gross, J. C.; Pokutta, S.; Weis, W. I. Interactions of Plakoglobin and β-Catenin with Desmosomal Cadherins: Basis of Selective Exclusion of α- and β-Catenin from Desmosomes. *J. Biol. Chem.* **2009**, *284*, 31776–31788

27. Tickenbrock, L.; Köβmeier, K.; Rehmann, H.; Herrmann, C.; Müller, O. Differences Between the Interaction of β-Catenin with Non-Phosphorylated and Single-Mimicked Phosphorylated 20-Amino Acid Residue Repeats of the APC Protein. *J. Mol. Biol*. **2003**, *327*, 359–367.

28. Liu, J.; Xing, Y.; Hinds, T. R.; Zheng, J.; Xu, W. The Third 20 Amino Acid Repeat is the Tightest Binding Site of APC for β-Catenin. *J. Mol. Biol.* **2006**, *360*, 133–144.

29. Lepourcelet, M.; Chen, Y.-N. P.; France, D. S.; Wang, H.; Crews, P.; Petersen, F.; Bruseo, C.; Wood, A. W.; Shivdasani, R. A. Small-Molecule Antagonists of the Oncogenic Tcf/β-Catenin Protein Complex. *Cancer Cell* **2004**, *5*, 91–102.

30. Baell, J. B.; Holloway, G. A. New Substructure Filters for Removal of Pan Assay Interference Compounds (PAINS) from Screening Libraries and for their Exclusion in Bioassays. *J. Med. Chem.* **2010**, *53*, 2719–2740.

31. Baell, J.; Walters, M. A. Chemistry: Chemical Con Artists Foil Drug Discovery. *Nature* **2014**, *513*, 481–483.

32. Trosset, J.-Y.; Dalvit, C.; Knapp, S.; Fasolini, M.; Veronesi, M.; Mantegani, S.; Gianellini, L. M.; Catana, C.; Sundström, M.; Stouten, P. F. W.; Moll, J. K. Inhibition of Protein–Protein Interactions: the Discovery of Druglike β-Catenin Inhibitors by Combining Virtual and Biophysical Screening. *Proteins* **2006**, *64*, 60–67.

33. Gonsalves, F. C.; Klein, K.; Carson, B. B.; Katz, S.; Ekas, L. A.; Evans, S.; Nagourney, R.; Cardozo, T.; Brown, A. M.; DasGupta, R. An RNAi-Based Chemical Genetic Screen Identifies Three Small-Molecule Inhibitors of the Wnt/Wingless Signaling Pathway. *Proc. Natl. Acad. Sci. U.S.A.* ***2011***, *108*, 5954–5963.

34. Lee, E.; Mada, A.; David, G.; Garabedian, M. J.; DasGupta, R.; Logan, S. K. Inhibition of Androgen Receptor and β-Catenin Activity in Prostate Cancer. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 15710–15715.

35. Baell, J. B. Observations on Screening-Based Research and Some Concerning Trends in the Literature. *Future Med. Chem.* **2010**, *2*, 1529–1546.

36. Tomašić, T.; Mašič, L. P. Rhodanine as a Privileged Scaffold in Drug Discovery. *Curr. Med. Chem.* **2009**, *16*, 1596–1629.

37. Tian, W.; Han, X.; Yan, M.; Xu, Y.; Duggineni, S.; Lin, N.; Luo, G.; Li, Y. M.; Han, X.; Huang, Z.; An, J. Structure-Based Discovery of a Novel Inhibitor Targeting the β-Catenin/Tcf4 Interaction. *Biochemistry* **2012**, *51*, 724–731.

38. Grossmann, T. N.; Yeh, J. T.-H.; Bowman, B. R.; Chu, Q.; Moellering, R. E.; Verdine, G. L. Inhibition of Oncogenic Wnt Signaling Through Direct Targeting of β-Catenin. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 17942–19747.

39. Huang, Z.; Zhang, M.; Burton, S. D.; Katsakhyan, L. N.; Ji, H. Targeting the Tcf4 G13ANDE17 Binding Site to Selectivity Disrupt β-Catenin/T-Cell Factor Protein–Protein Interactions. *ACS Chem. Biol.* **2014**, *9*, 193–201.

40. Cousins, G. R. L.; Poulsen, S.-A.; Sanders, J. K. M. Dynamic Combinatorial Libraries of Pseudo-Peptide Hydrazone. *Macrocycles Chem. Commun*. **1999**, 1575–1576.

41. Poulsen, S.-A. Direct Screening of a Dynamic Combinatorial Library Using Mass Spectrometry. *J. Am. Soc. Mass Spectrom*. **2006**, *17*, 1074–1080.

42. Bhat, V. T.; Caniard, A. M.; Luksch, T.; Brenk, R.; Campopiano, D. J.; Greaney, M. F. Nucleophilic Catalysis of Acylhydrazone Equilibration for Protein-Directed Dynamic Covalent Chemistry. *Nature Chem*. **2010**, *2*, 490–497.

43. Bunyapaiboonsri, T.; Ramström, O.; Lohmann, S.; Lehn, J.-M.; Peng, L.; Goeldner, M. Dynamic Deconvolution of a Pre-Equilibrated Dynamic Combinatorial Library of Acetylcholinesterase Inhibitors. *ChemBioChem* **2001**, *2*, 438–444.

44. Mondal, M.; Radeva, N.; Köster, H.; Park, A.; Potamitis, C.; Zervou, M.; Klebe, G.; Hirsch, A. K. H. Structure-Based Design of Inhibitors of the Aspartic Protease Endothiapepsin by Exploiting Dynamic Combinatorial Chemistry. *Angew. Chem., Int. Ed*. **2014**, *53*, 3259–3263.

45. Jiang, Q.-Q.; Sicking, W.; Ehlers, M.; Schmuck, C. Discovery of Potent Inhibitors of Human β-Tryptase from Pre-Equilibrated Dynamic Combinatorial Libraries. *Chem. Sci*. 2015, *6*, 1792–1800.

46. Bunyapaiboonsri, T.; Ramström, H.; Ramström, O.; Haiech, J.; Lehn, J.-M. Generation of Bis-Cationic Heterocyclic Inhibitors of Bacillus Subtilis Hpr Kinase/Phosphatase from a Ditopic Dynamic Combinatorial Library. *J. Med. Chem*. **2003**, *46*, 5803–5811.

47. Trott, O.; Olson, A. J. Autodock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **2010**, *31* (2), 455–461.

48. Sastry, G. M.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. Protein and Ligand Preparation: Parameters, Protocols, and Influence on Virtual Screening Enrichments. *J. Comput.-Aided Mol. Des.* **2013**, *27* (3), 221–234.

49. *Schrödinger Suite 2015-1 Protein Preparation Wizard;* Schrödinger, LLC, New York, NY, 2015.

50. Harder, E.; Damm, W.; Maple, J.; Wu, C.; Reboul, M.; Xiang, J.Y.; Wang, L.; Lupyan, D.; Dahlgren, M. K.; Knight, J. L.; Kaus, J. W.; Cerutti, D. S.; Krilov, G.; Jorgensen, W. L.; Abel, R.; Friesner, R. A. OPLS3: a Froce Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. *J. Chem. Theory Comput.*, ***2015,*** *12* (1), 281-296.

51. Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: a New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **2004**, *47* (7), 1739–1749.

52. Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: a New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *J. Med. Chem.* **2004**, *47* (7), 1750–1759.

53. Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. Autodock4 and Autodocktools4: Automated Docking with Selective Receptor Flexibility. *J. Comput. Chem*. **2009**, *30*, 2785–2791.

54. Zhang, M.; Huang, Z.; Yu, B.; Ji, H. New Homogeneous High-Throughput Assays for Inhibitors of β-Catenin/Tcf Protein–Protein Interactions. *Anal. Biochem*. **2012**, *424*, 57–63.

55. Korinek, V.; Barker, N.; Morin, P. J.; van Wichen, D.; de Weger, R.; Kinzler, K. W.; Vogelstein, B.; Clevers, H. Constitutive Transcriptional Activation by a β-Catenin–Tcf Complex in APC–/– Colon Carcinoma. *Science* **1997**, *275*, 1784–1787.

56. Ilyas, M.; Tomlinson, I. P. M.; Rowan, A.; Pignatelli, M.; Bodmer, W. F. β-Catenin Mutations in Cell Lines Established from Human Colorectal Cancers. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 10330–10334.

57. Sparks, A. B.; Morin, P. J.; Vogelstein, B.; Kinzler, K. W. Mutational Analysis of the APC/β-Catenin/Tcf Pathway in Colorectal Cancer. *Cancer Res.* **1998**, *58*, 1130–1134.

58. Leung, J. Y.; Kolligs, F. T.; Wu, R.; Zhai, Y.; Kuick, R.; Hanash, S.; Cho, K. R.;

Fearon, E. R. Activation of AXIN2 Expression by β-Catenin–T Cell Factor. A Feedback Repressor Pathway Regulating Wnt Signaling. *J. Biol. Chem.* **2002**, *277*, 21657–21665.

59. Metz, J. T.; Huth, J. R.; Hajduk, P. J. Enhancement of Chemical Rules for Predicting Compound Reactivity Towards Protein Thiol Groups. *J. Comput.-Aided Mol. Des.* **2007**, *21*, 139–144.

60. Huth, J. R.; Song, D.; Mendoza, R. R.; Black-Schaefer, C. L.; Mack, J. C.; Dorwin, S. A.; Ladror, U. S.; Severin, J. M.; Walter, K. A.; Bartley, D. M.; Hajduk, P. J. Toxicological Evaluation of Thiol-Reactive Compounds Identified Using a LA Assay to Detect Reactive Molecules by Nuclear Magnetic Resonance. *Chem. Res. Toxicol.* **2007**, *20*, 1752–1759.

61. Dahlin, J. L.; Nissink, J. W. M.; Strasser, J. M.; Francis, S.; Higgins, L.; Zhou, H.; Zhang, Z.; Walters, M. A. PAINS in the Assay: Chemical Mechanisms of Assay Interference and Promiscuous Enzymatic Inhibition Observed During a Sulfhydryl-Scavenging HTS. *J. Med. Chem.* **2015**, *58*, 2091–2113.

62. Takabatake, T.; Miyazawa, T.; Hasegawa, M.; Foote, C. S. Reaction of 4,7-Dimethylbenzofurazan with Singlet Oxygen. *Tetrahedron Lett.* **2001**, *42*, 987–989.

63. Okabe, K.; Wada, R.; Ohno, K.-i.; Uchiyama, S.; Santa, T.; Imai, K. Development of Hydrophilic Fluorogenic Derivatization Reagents for Thiols: 4-(N-Acetylaminosulfonyl)-7-Fluoro-2,1,3-Benzoxadiazole and 4-(N-Trichloroacetylaminosulfonyl)-7-Fluoro-2,1,3-Benzoxadiazole. *J. Chromatogr., A* **2002**, *982*, 111–118.

64. Jadhav, A.; Ferreira, R. S.; Klumpp, C.; Mott, B. T.; Austin, C. P.; Inglese, J.; Thomas, C. J.; Maloney, D. J.; Shoichet, B. K.; Simeonov, A. Quantitative Analyses of Aggregation, Autofluorescence, and Reactivity Artifacts in a Screen for Inhibitors of a Thiol Protease. *J. Med. Chem.* **2010**, *53*, 37–51.

65. Che, J.; King, F. J.; Zhou, B.; Zhou, Y. Chemical and Biological Properties of Frequent Screening Hits. *J. Chem. Inf. Model.* **2012**, *52*, 913–926.

66. Yu, B.; Reynisson, J. Bond Stability of the "Undesirable" Heteroatom–Heteroatom Molecular Moieties for High-Throughput Screening Libraries. *Eur. J. Med. Chem.* **2011**, *46*, 5833–5837.

67. Řeha, D.; Kabeláč, M.; Ryjáček, F.; Šponer, J.; Šponer, J. E.; Elstner, M.; Suhai, S.; Hobza, P. Intercalators. 1. Nature of Stacking Interactions between Intercalators (Ethidium, Daunomycin, Ellipticine, And 4′,6-Diaminide-2-Phenylindole) and DNA Base Pairs. Ab Initio Quantum Chemical, Density Functional Theory, and Empirical Potential Study. *J. Am. Chem. Soc.* **2002**, *124*, 3366–3376.

68. Reynisson, J.; Schuster, G. B.; Howerton, S. B.; Williams, L. D.; Barnett, R. N.; Cleveland, C. L.; Landman, U.; Harrit, N.; Chaires, J. B. Intercalation of Trioxatriangulenium Ion in DNA: Binding, Electron Transfer, X-Ray Crystallography, and Electronic Structure. *J. Am. Chem. Soc.* **2003**, *125*, 2072–2083.

69. Guedat, P.; Jacq, X.; Colland, F.; Daviet, L.; Formstecher, E.; Rain, J.-C.; Colombo, M. Preparation of Tetracyclic Compounds Ascysteine Proteases Inhibitors. U.S. Patent 7,875,613, January 25, 2011.

70. Ichikawa, T.; Kato, T.; Takenishi, T. New Synthesis of Adenine and 4-aminoimidazole-5-carboxamide. *J. Heterocycl. Chem*. **1965**, 2253−255.

71. Shinya, M.; Eschbach, C.; Clark, M.; Lehrach, H.; Furutani-Seiki, M., Zebrafish Dkk1, Induced by the Pre-MBT Wnt Signaling, is Secreted from the Prechordal Plate and Patterns the Anterior Neural Plate. *Mech. Dev.* **2000,** *98* (1-2), 3-17.

72. Dorsky, R. I.; Sheldahl, L. C.; Moon, R. T. A Transgenic Lef1/B-Catenin-Dependent Reporter Is Expressed in Spatially Restricted Domains Throughout Zebrafish Development. *Dev. Biol.* **2002,** *241* (2), 229-37.

73. Kim, C. H.; Oda, T.; Itoh, M.; Jiang, D.; Artinger, K. B.; Chandrasekharappa, S. C.; Driever, W.; Chitnis, A. B. Repressor Activity of Headless/Tcf3 is Essential for Vertebrate Head Formation. *Nature* **2000,** *407* (6806), 913-6.

74. Avula, S.; Komsani, J. R.; Koppireddi, S.; Yadla, R. Microwave-Assisted Synthesis of 6-{(5-Aryl-1,3,4-Oxadiazol-2-Yl)Methyl}-6H-Indolo[2,3-B]Quinoxalines. *J. Hetero. Chem.* **2015,** *52* (6), 1737-1742.

75. Basham, K. J.; Kieffer, C.; Shelton, D. N.; Leonard, C. J.; Bhonde, V. R.; Vankayalapati, H.; Milash, B.; Bearss, D. J.; Looper, R. E.; Welm, B. E. Chemical Genetic Screen Reveals a Role for Desmosomal Adhesion in Mammary Branching Morphogenesis. *J. Biol. Chem.* **2013,** *288* (4), 2261-70.

2.7 Supplementary materials of NMR spectra

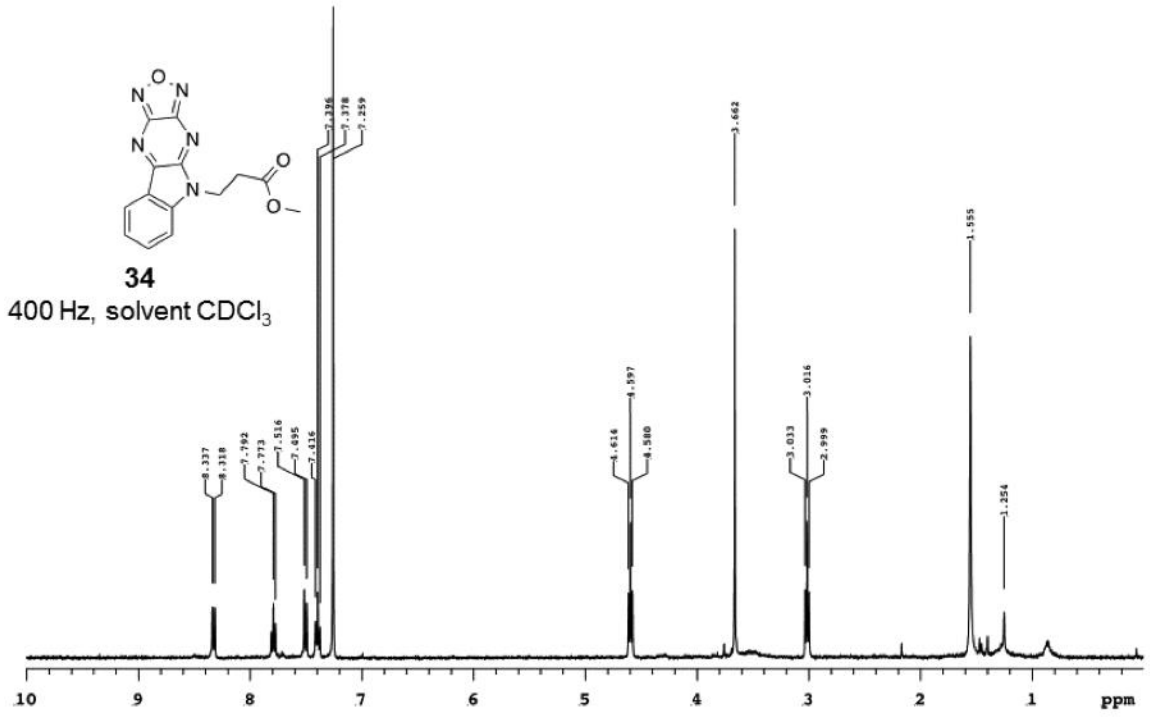


Figure 2S.1. Compound **34** ^{1}H NMR.

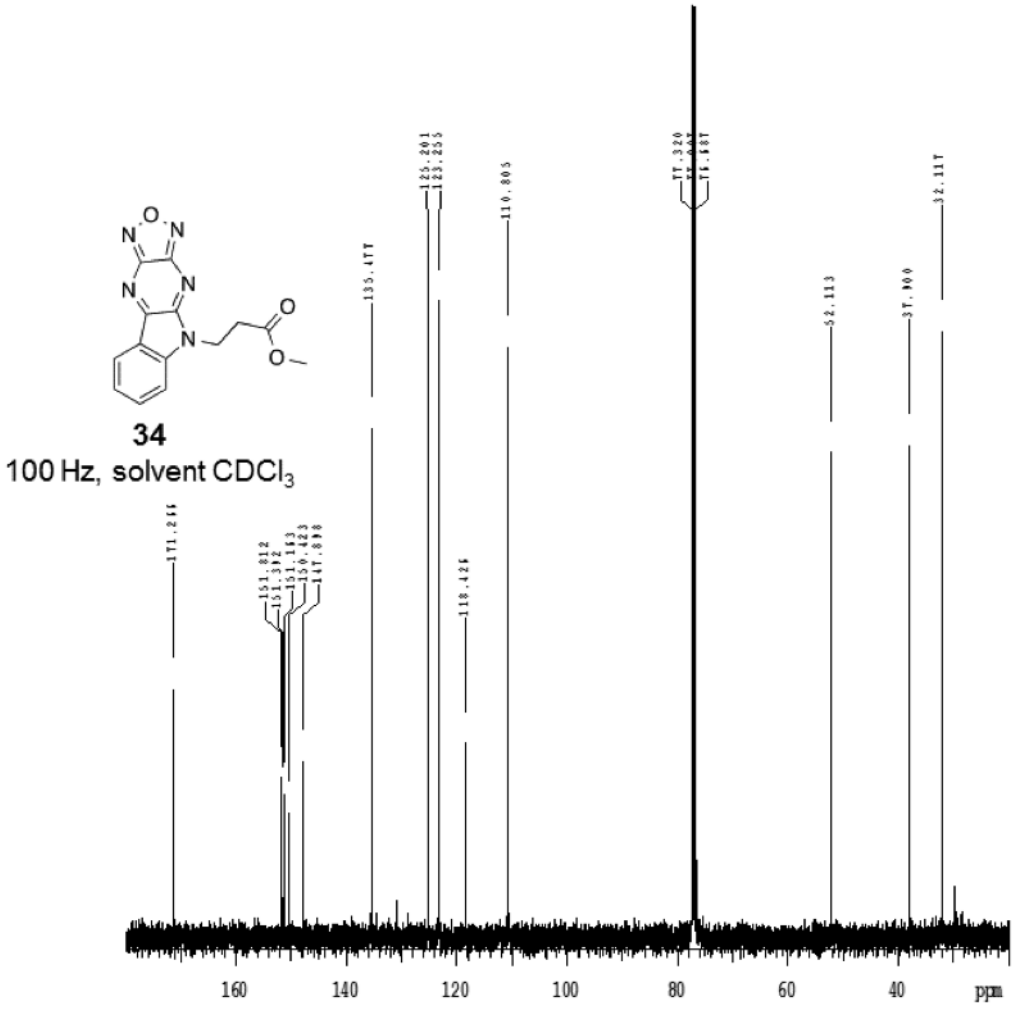


Figure 2S.2. Compound **34** ^{13}C NMR.

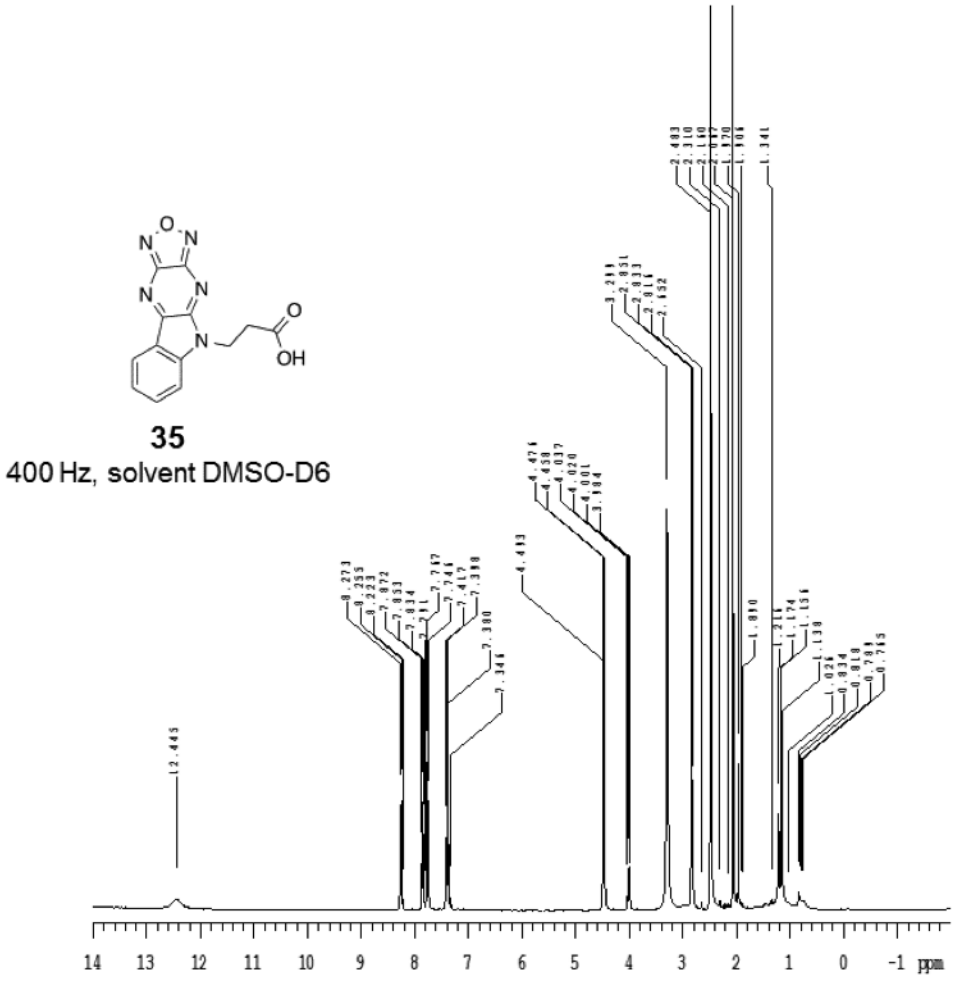


Figure 2S.3. Compound **35** ^{1}H NMR.

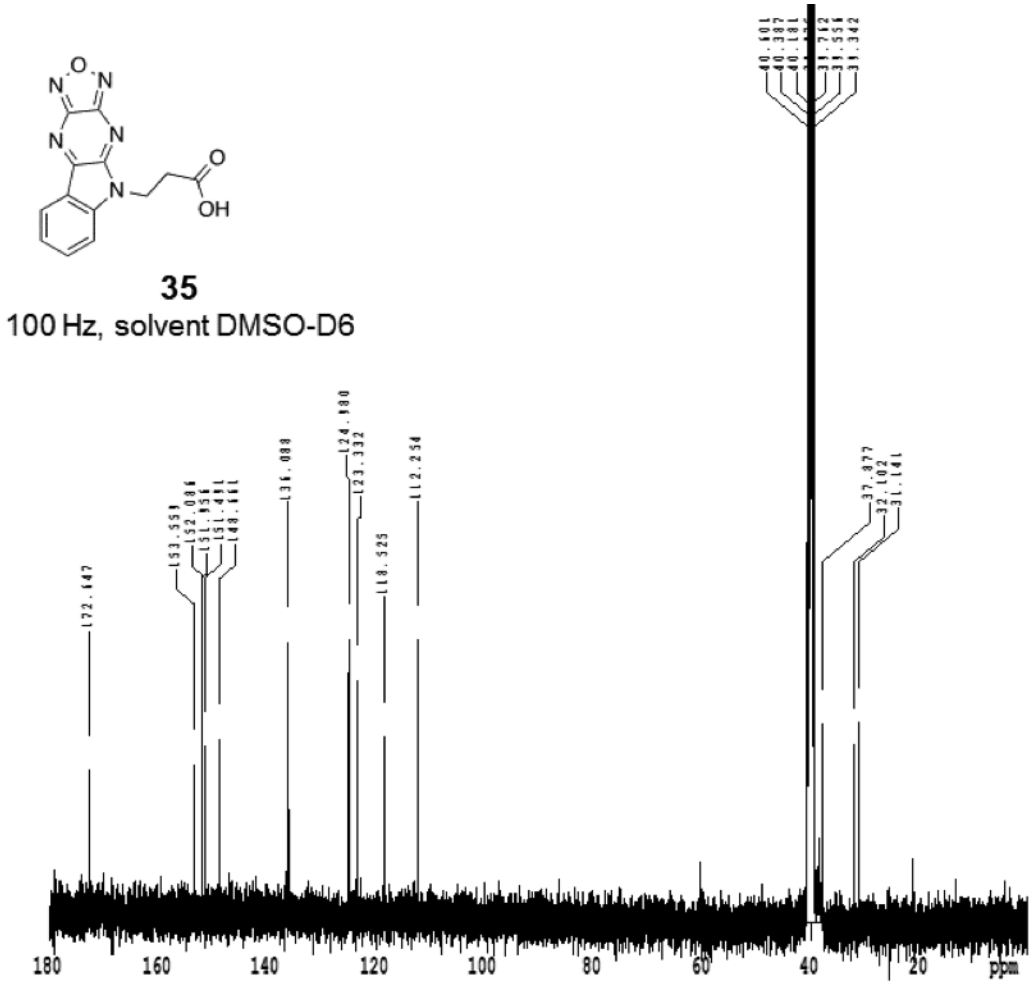


Figure 2S.4. Compound **35** ^{13}C NMR.

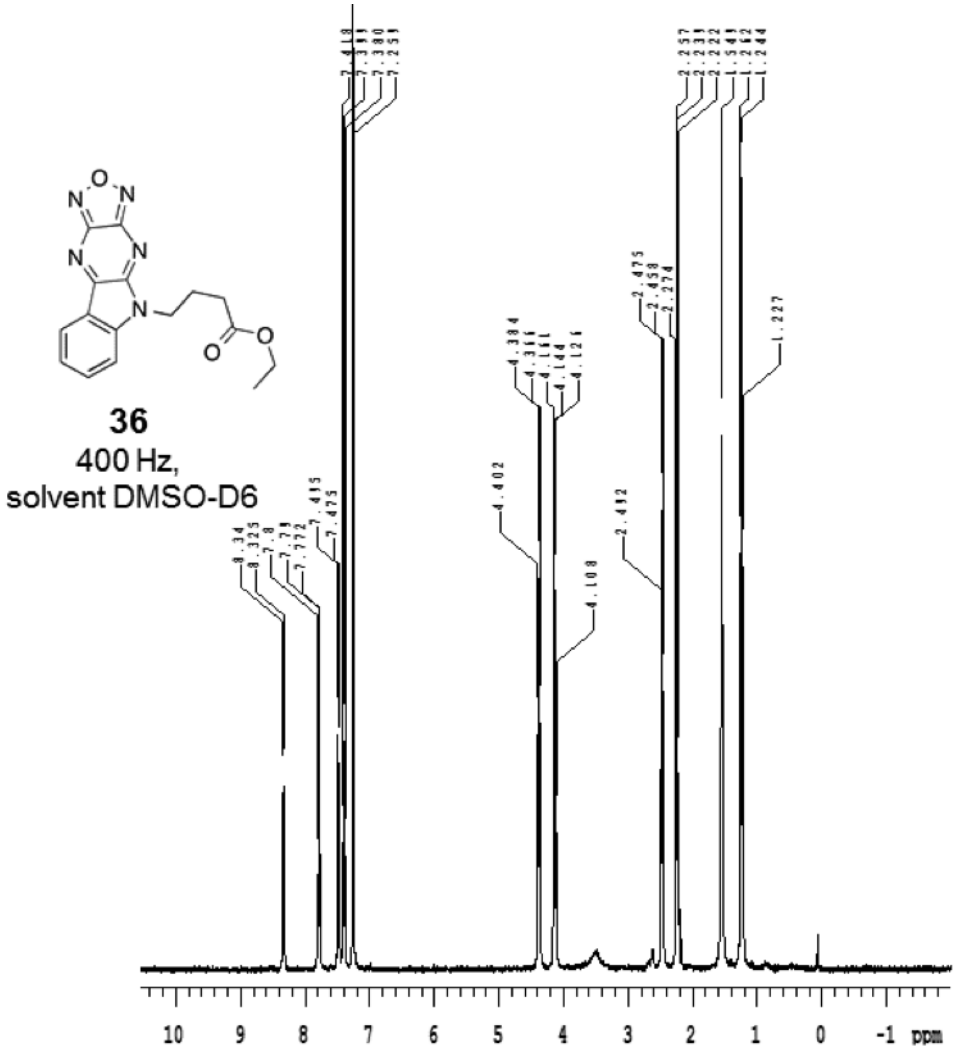


Figure 2S.5. Compound **36** ^{1}H NMR.

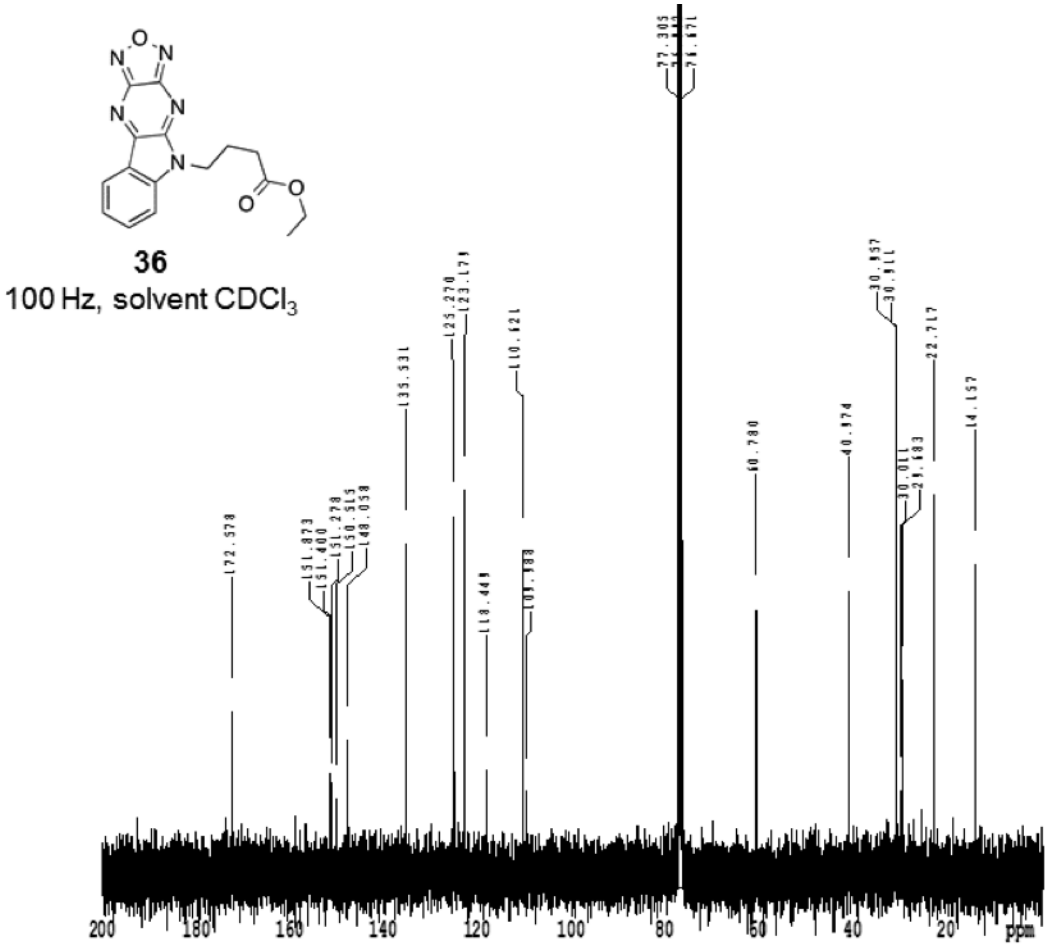


Figure 2S.6. Compound **36** ^{13}C NMR.

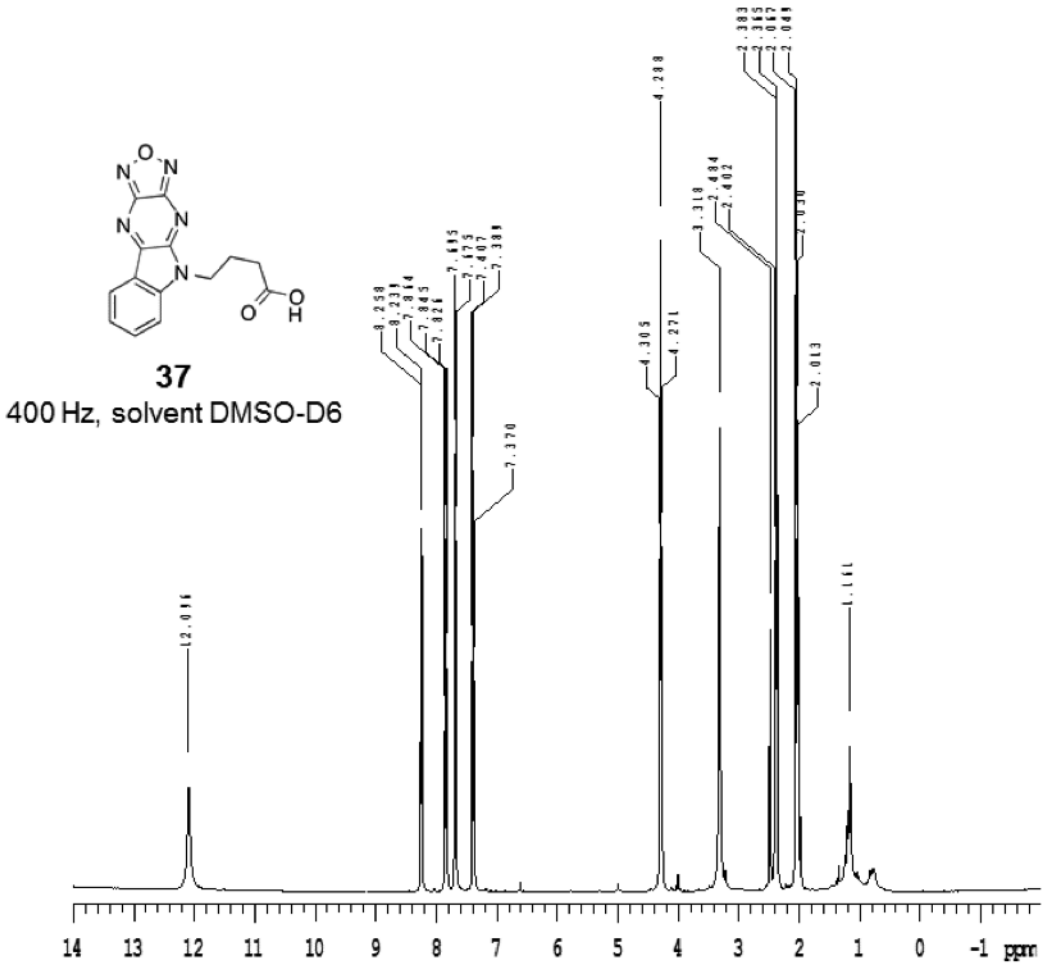


Figure 2S.7. Compound **37** ^{1}H NMR.

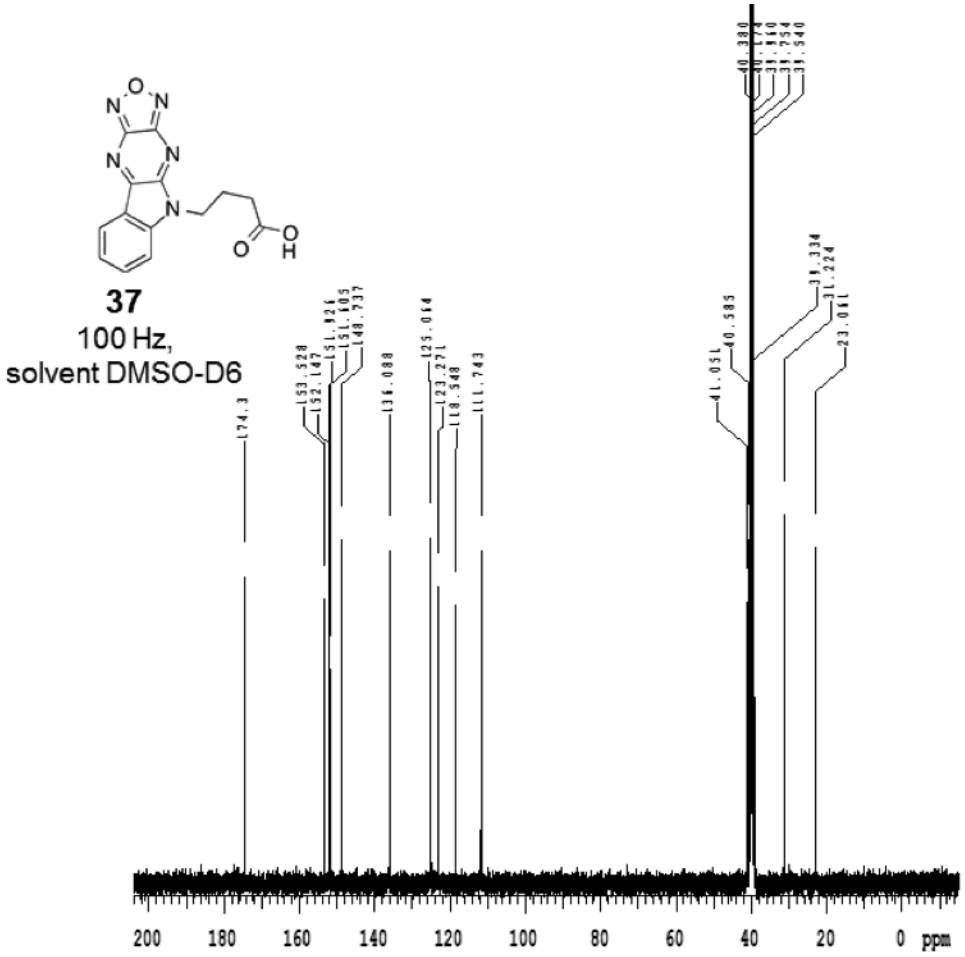


Figure 2S.8. Compound **37** ^{13}C NMR.

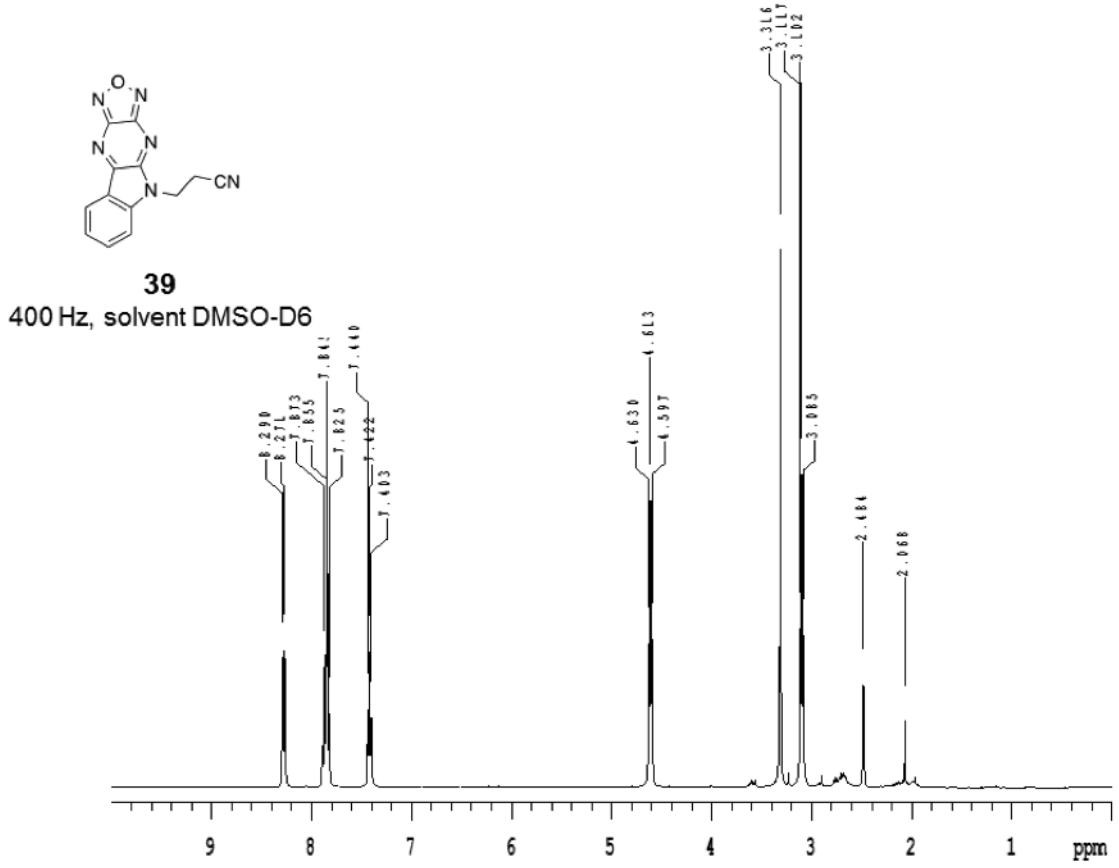


Figure 2S.9. Compound **39** ^{1}H NMR.

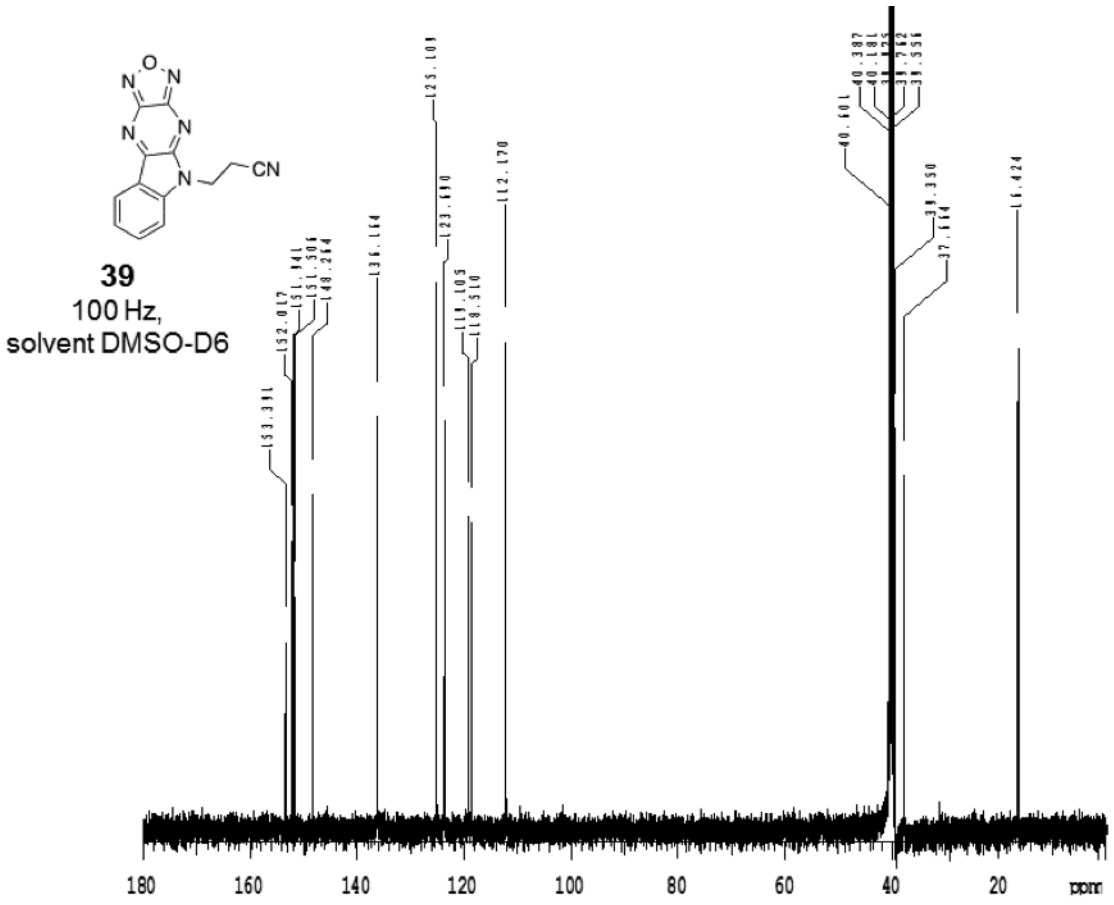


Figure 2S.10. Compound **39** ^{13}C NMR.

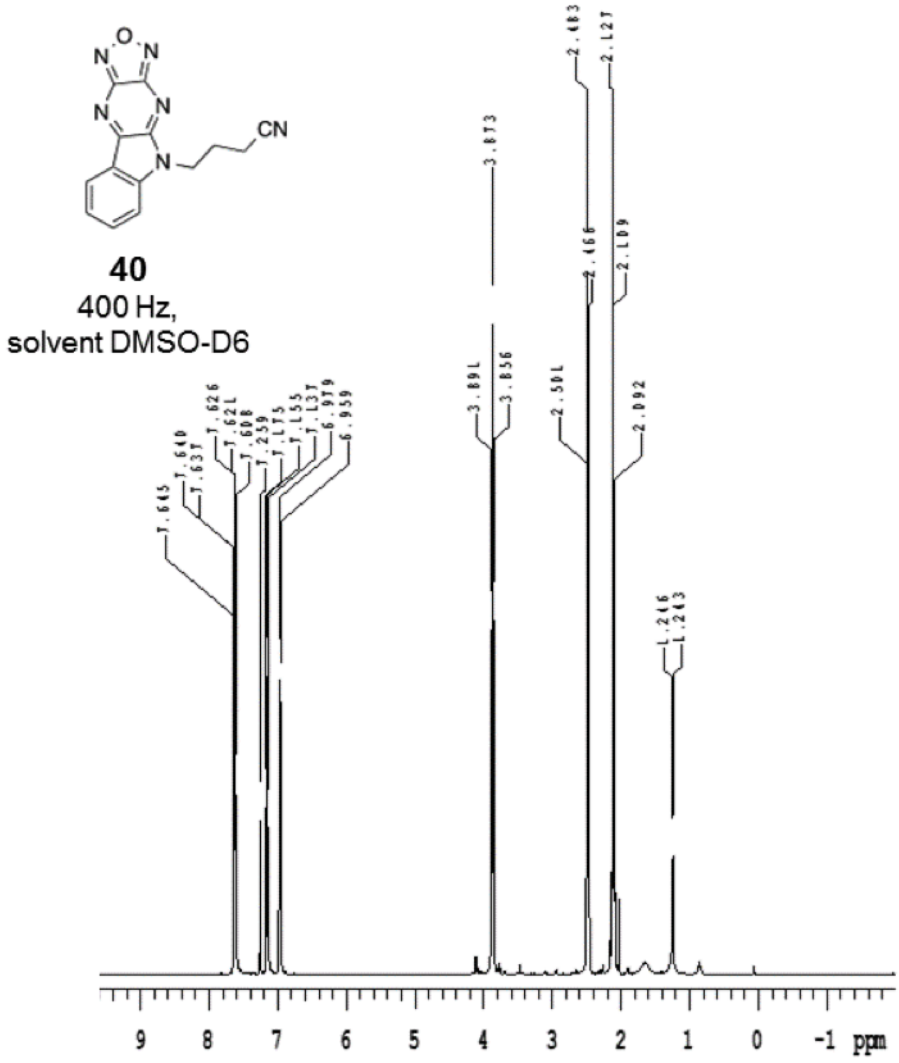


Figure 2S.11. Compound **40** ^{1}H NMR.

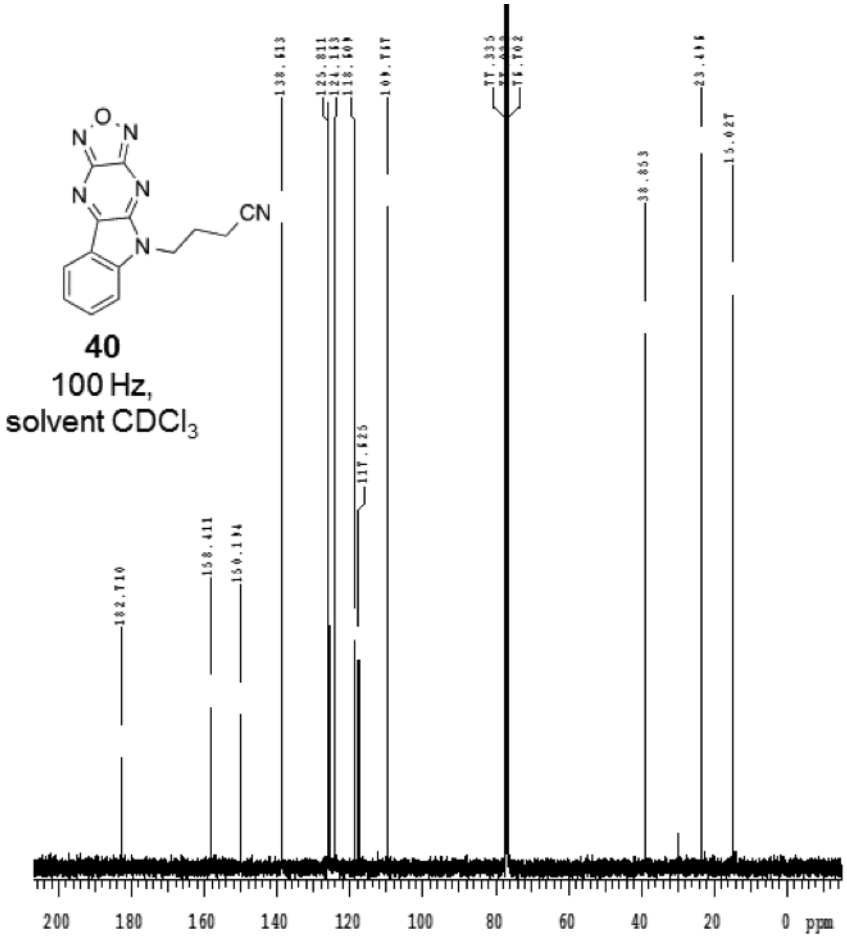


Figure 2S.12. Compound **40** ^{13}C NMR.

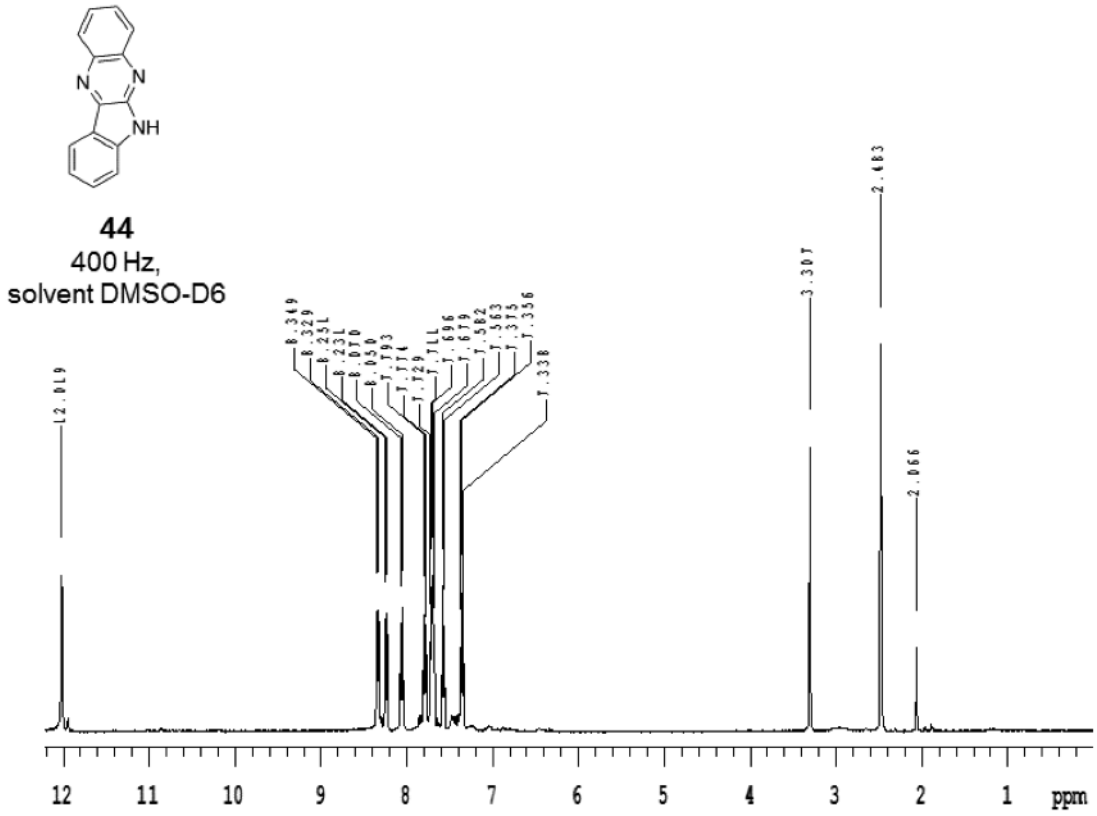


Figure 2S.13. Compound **44** ^{1}H NMR.

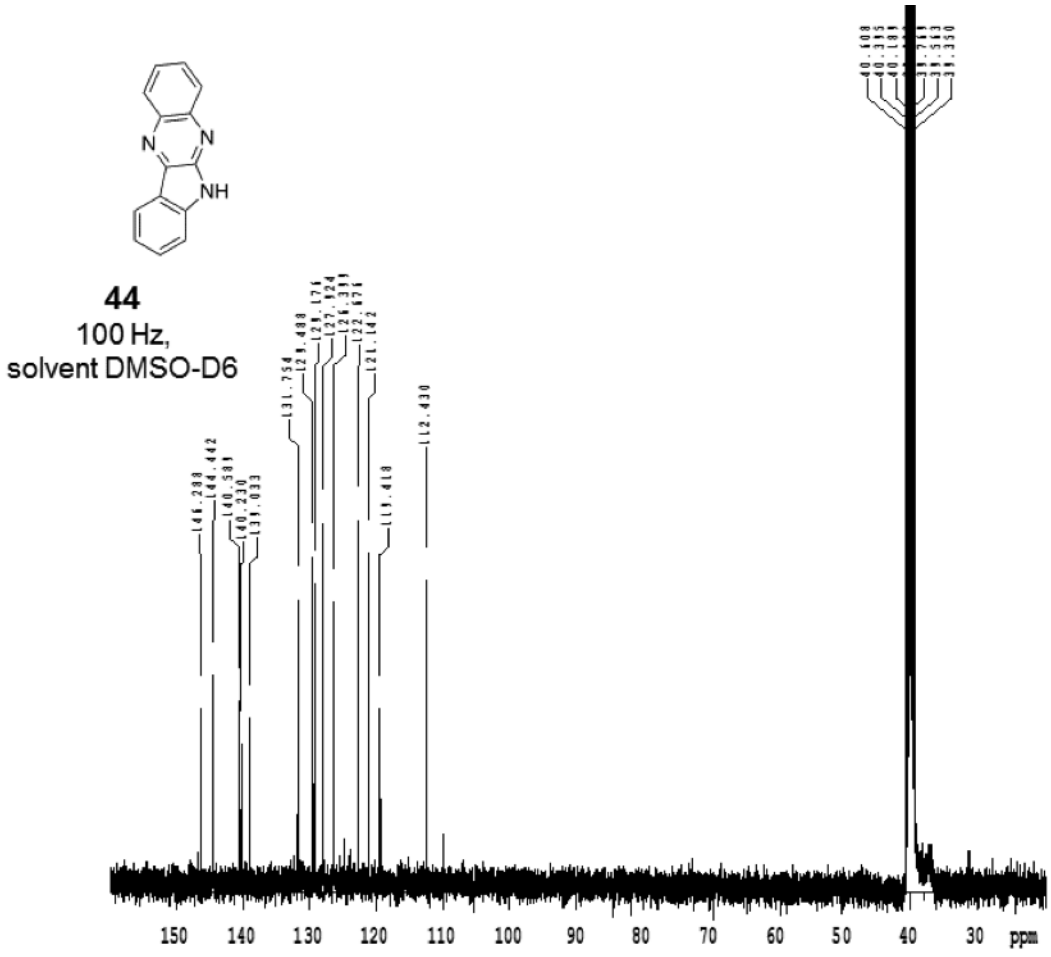


Figure 2S.14. Compound **44** ^{13}C NMR.

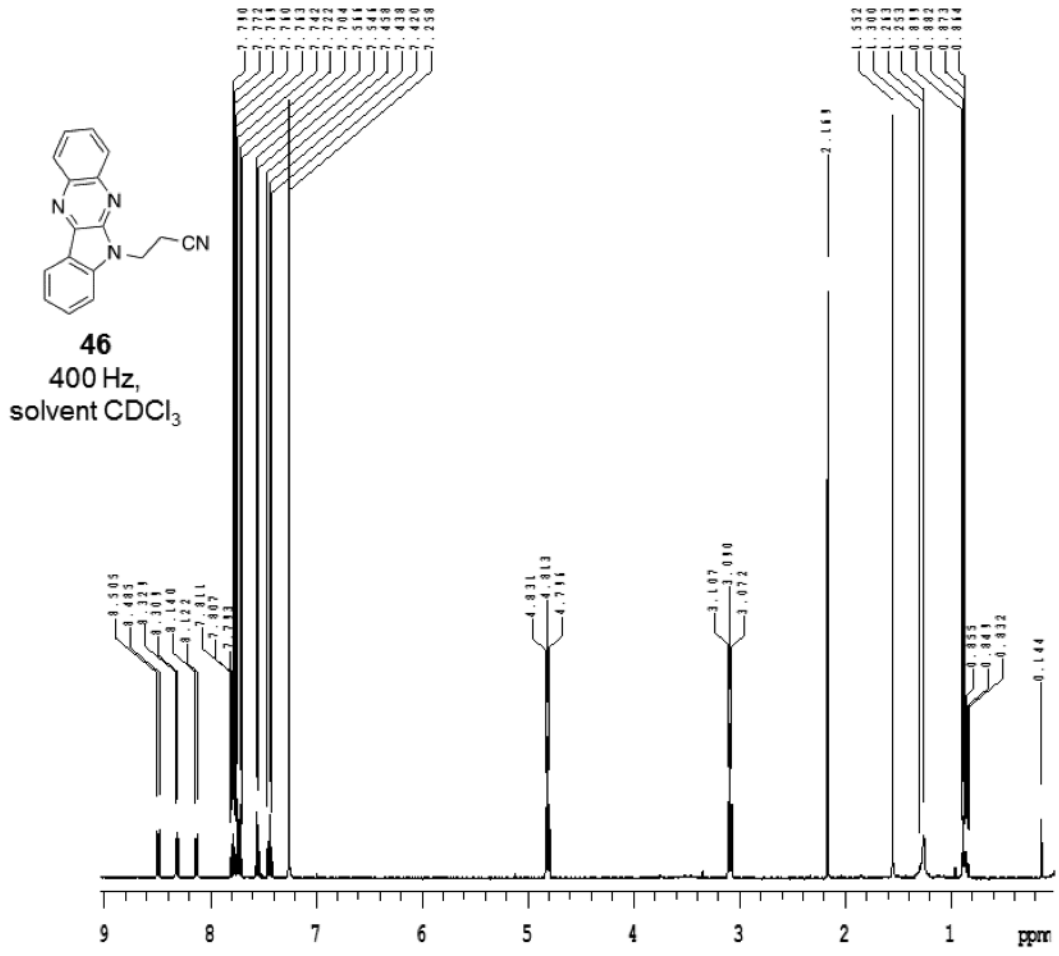


Figure 2S.15. Compound **46** ^{1}H NMR.

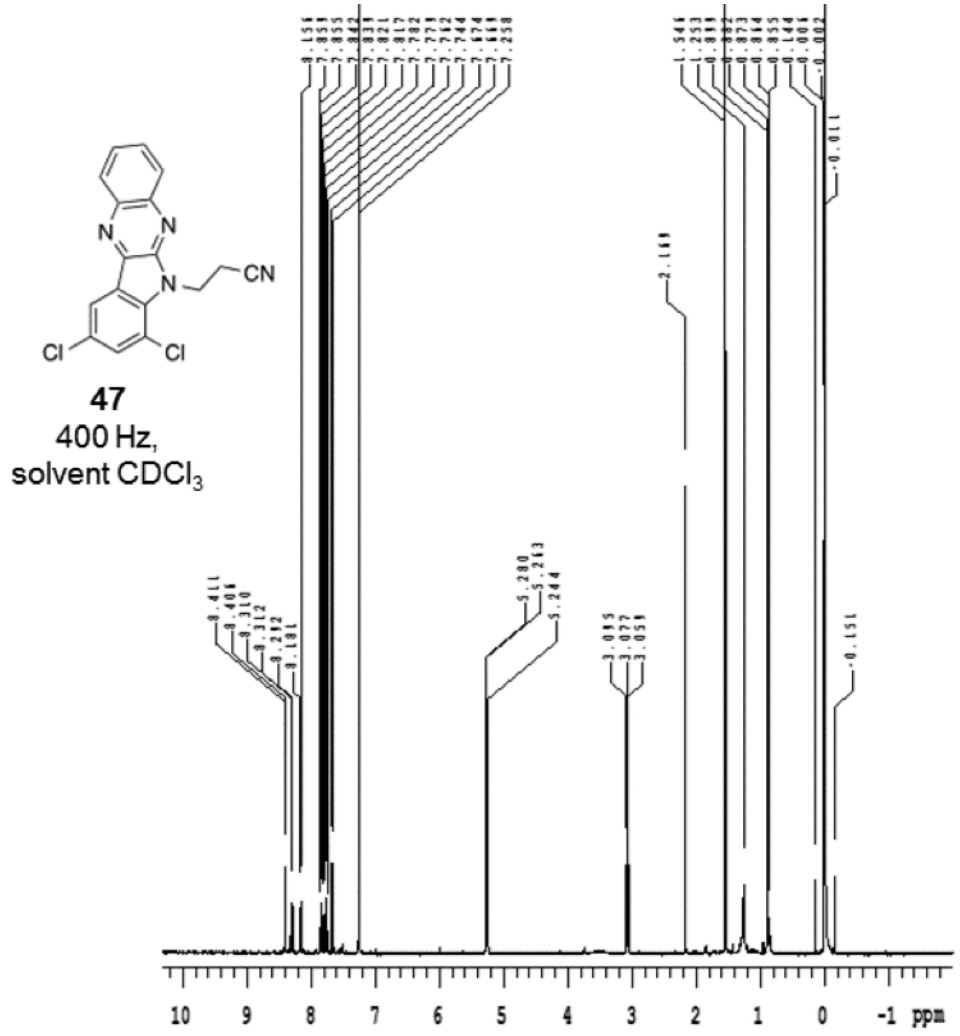


Figure 2S.16. Compound **47** ^{1}H NMR.

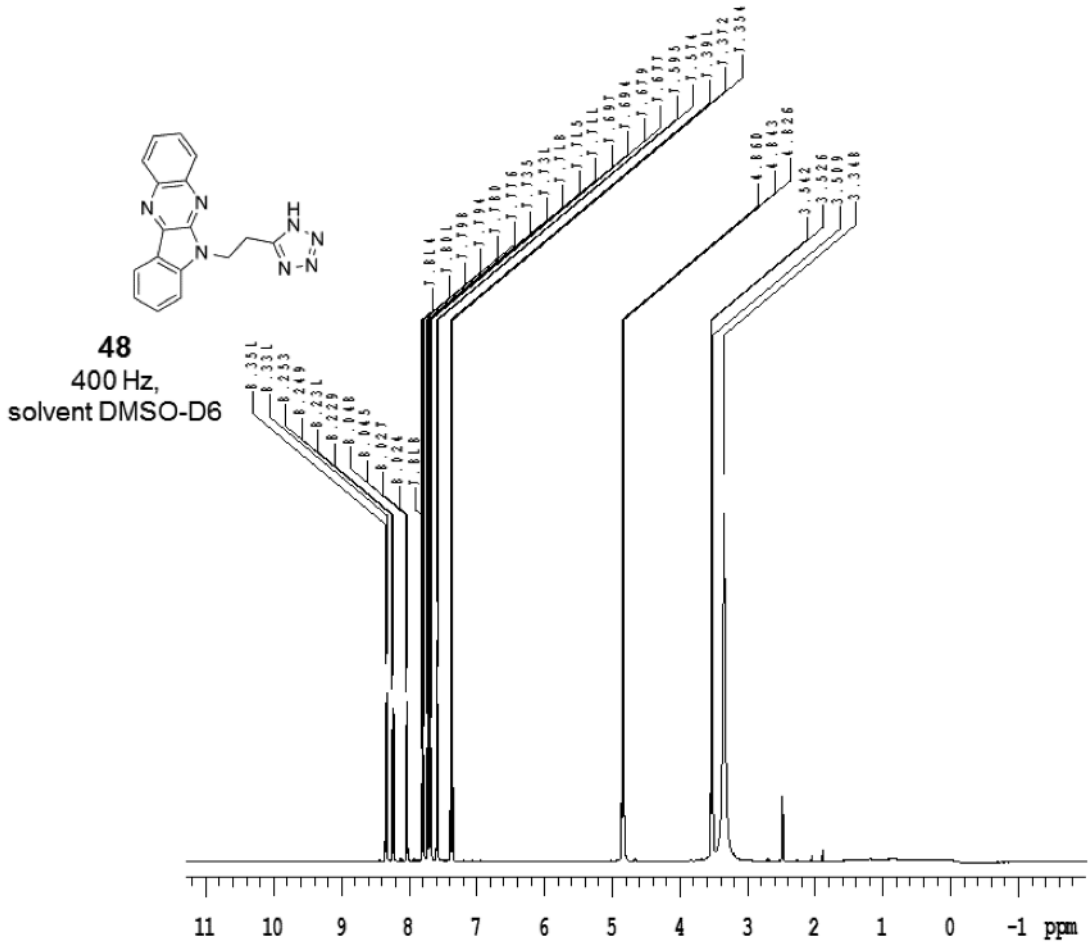


Figure 2S.17. Compound **48** ^{1}H NMR.

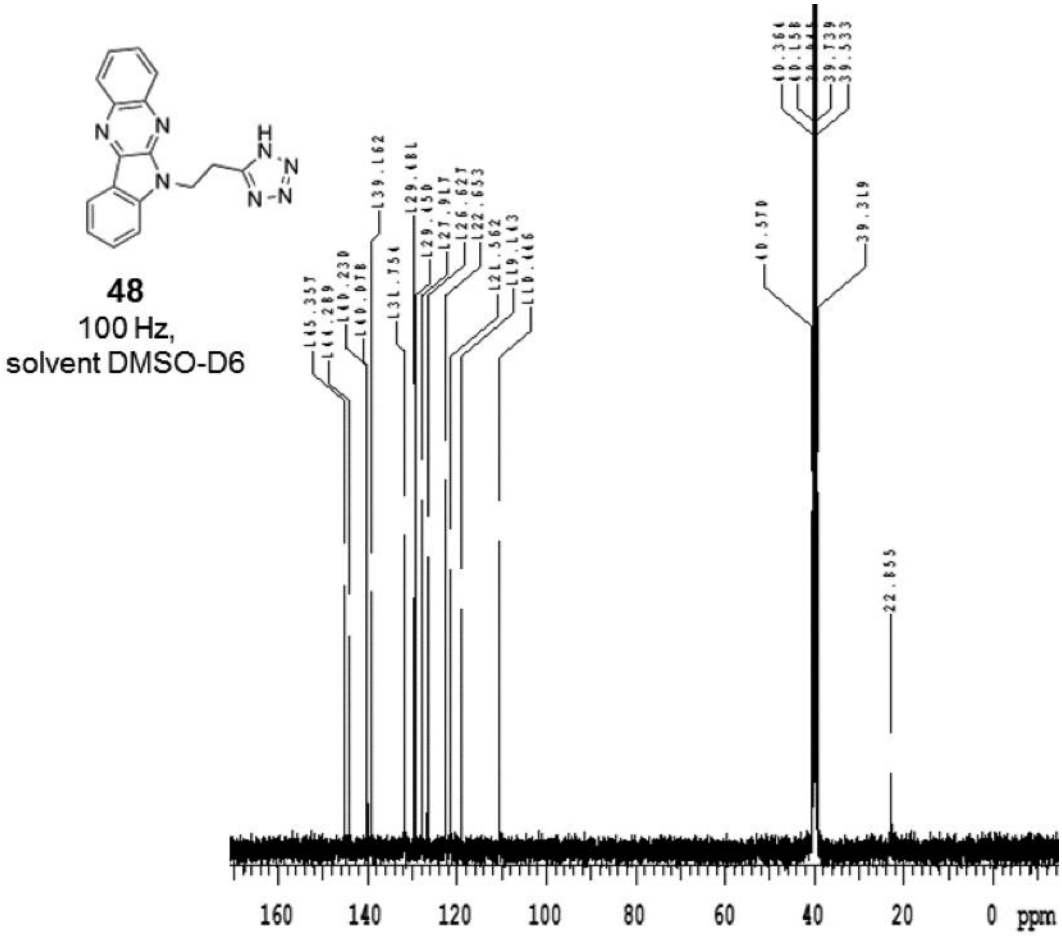


Figure 2S.18. Compound **48** ^{13}C NMR.

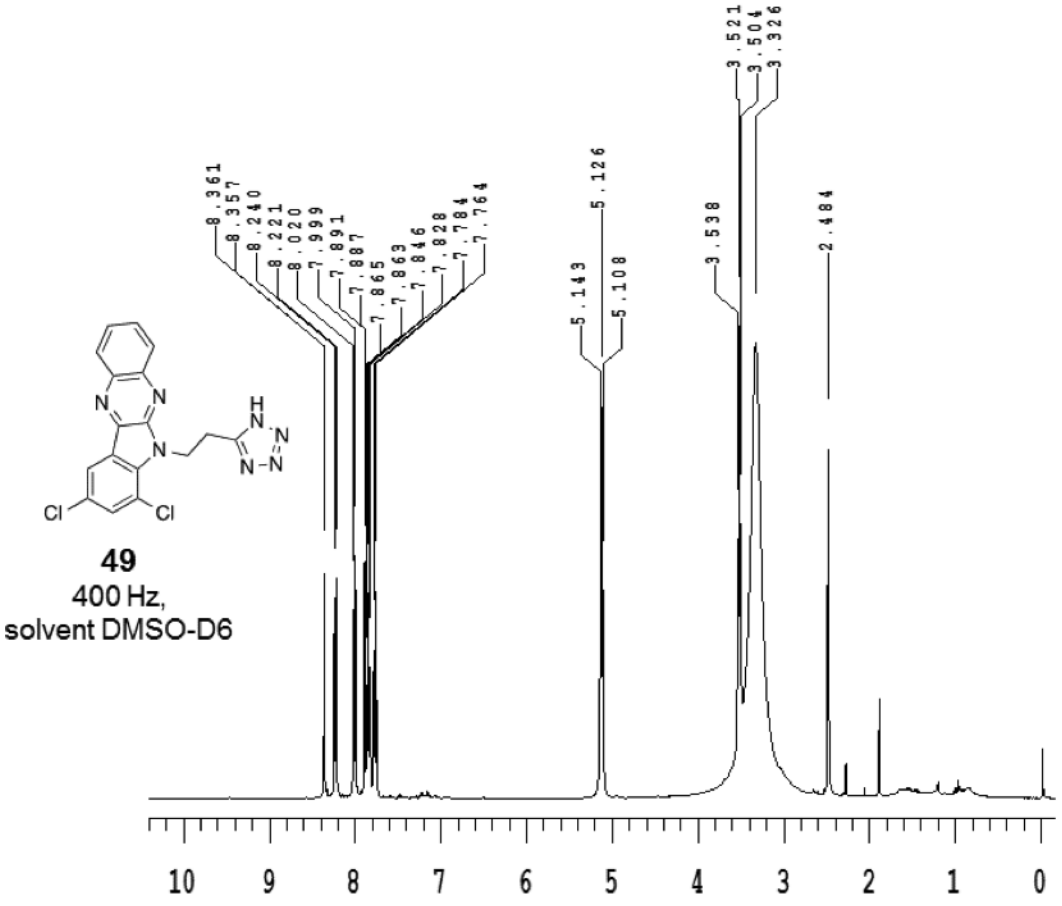


Figure 2S.19. Compound **49** ^{1}H NMR.

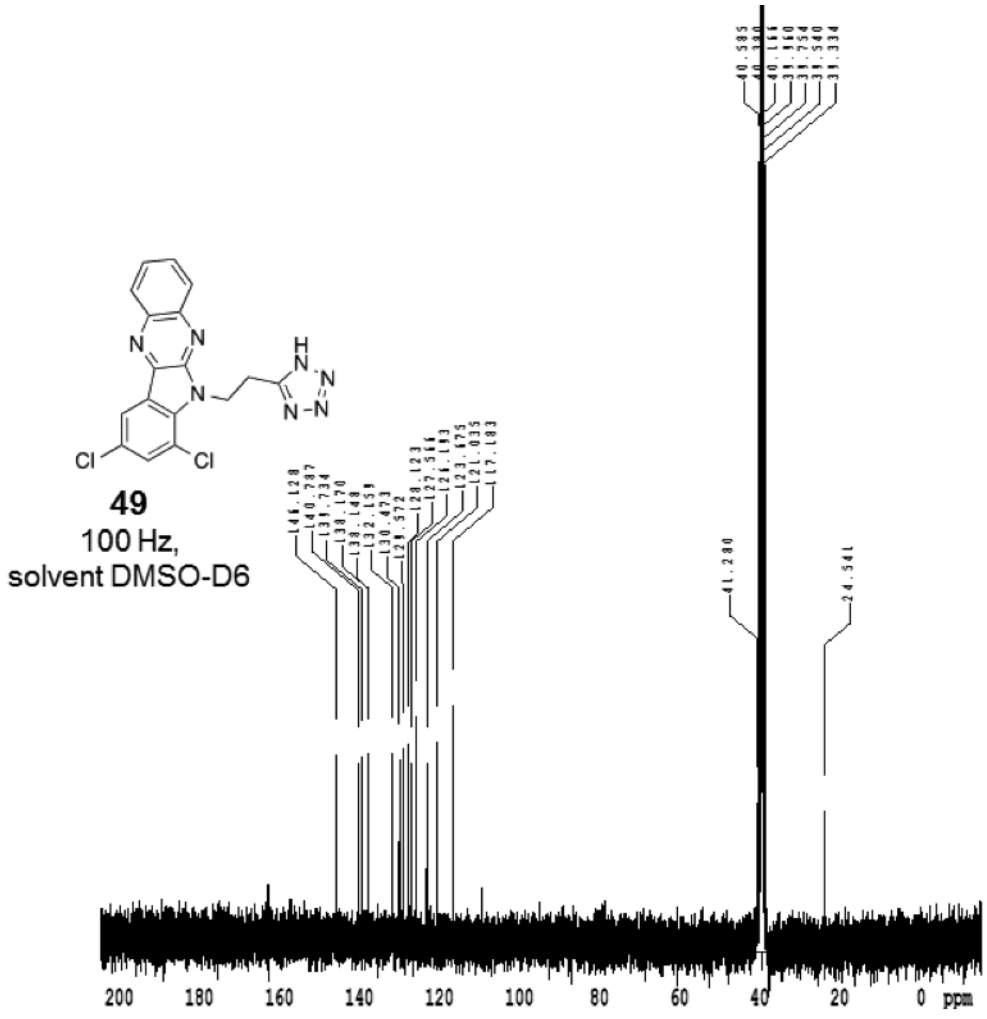


Figure 2S.20. Compound **49** ^{13}C NMR.

CHAPTER 3

HIGH-THROUGHPUT VIRTUAL SCREENING CHALLENGES AND CONSIDERATIONS FOR PROTEIN–PROTEIN INTERACTION INHIBITORS

3.1 Introduction

PPIs play a crucial role in regulating biological pathways. The dysregulation of these pathways results in diseases. PPIs are ubiquitous, and more than 40,000 PPIs have been experimentally identified with even more predicted.[1–3] The abundance of PPIs and their high relevance in disease makes them desirable drug targets, and targeting these interactions will greatly expand the druggable genome. Despite their therapeutic appeal, the discovery of small molecules to disrupt these interactions has been met with difficulty. Notably, few PPI inhibitors have advanced to clinical trials.[4]

Nontraditional PPIs differ from traditional drug targets such as GPCRs, kinases, proteases, and other enzymes. These differences are predominantly related to the size, shape, and chemical characteristics of PPI interfaces.[5] PPIs have large contacting surface areas, upwards of 3,000 $Å^2$, in contrast to traditional drug targets that have contacting surface areas of less than 1,000 $Å^2$. Traditional targets often display an active site with deep, well-defined pockets containing residues that readily bind a substrate or catalyze a reaction. PPI binding sites are often flat, shallow, and devoid of catalytic residues. When

taken as a whole, these compounding factors result in several challenges for the discovery of small-molecule inhibitors.[6,7] Regardless of the problems associated with the PPI interface, this group of nontraditional drug targets comprise a large portion of the druggable genome but constitute a disproportionately small number of therapeutic targets.

One strategy for overcoming the challenges presented by PPI interfaces is the use of hot spots.[8–10] These are the regions at the PPI interface containing residues that exhibit a $\Delta\Delta G$ of $\geq$ 1.5 kcal mol^{-1} when examined by alanine scanning.[8] Hot spot residues on the target protein are often preorganized into pockets that reside on the PPI interface.[11,12] The surrounding residues help occlude solvent from the hot spot pockets, thereby reducing entropic and enthalpic penalties for binding.[12] Hot spots function similarly to "anchor residues" proposed by Rajamani et al.[13] By use of anchor residues, large solvent accessible surface areas (SASAs) of one protein binding partner are buried through interactions with the anchor residues of the other protein partner. Hot spots present a site where small molecules may more efficiently bind and disrupt PPIs.[14] Hot spots have been successfully used for structure-based design, high-throughput screening and HTVS of PPI inhibitors.[15-20] Although high-throughput screening, rational design, and peptidomimetic strategies have achieved varying levels of success, HTVS provides an inexpensive starting strategy for obtaining hits from large virtual libraries of millions of compounds.

Despite the advantages, HTVS has had limited success for discovering PPI inhibitors. Table 3.1 surveys PPI inhibitors identified through HTVS as well as the docking programs and the supplementary techniques used to discover them. Docking

Table 3.1. HTVS success in discovering small-molecule inhibitors for PPIs.

PPI Target	Docking Program	Supplementary Techniques	Reported Activity	Reference
Annexin/S100A10	GOLD		IC_{50} = 3 μM	79
Arno/Arf	FlexX	X-score	K_{iapp} = 47.9 μM	80
Arno/Arf	SurFleX	Hot-spots, Fragment-based design	K_{iapp} = 540 μM	16
Bcl-2/BH3-domain	DOCK		IC_{50} = 4 μM	81
Bcl-6/corepressor	DOCK		K_d = 138 μM	82
BRD/Kac	Glide		IC_{50} = 4.7 μM	83
C-abk0/14-3-3	GOLD	Pharmacophore matching	LD_{50} = 1.04	29
Frataxin/Ubiquitin	Glide, Autodock	Consensus scoring	IC_{50} = 45 μM	84
LEDGF/P75	Glide	Pharmacophore matching, Scaffold hopping	IC_{50} = 0.4 μM	28
LEDGF/P75	GOLD	Q-SAR, Pharmacophore	IC_{50} = 1.37 μM	27
IFN/IFNAR	GOLD	pharmacophore	K_d = 4 μM	30
Keap1/Nrf2	Glide		K_{d2} = 2.9 μM	85
MDM2/p53	GOLD	Pharmacophore, PoliMorph, BEDROC	30 μM	26
S100B/p53	DOCK	NMR-based screening and docking	K_d = 1 μM	86
S100B/p53	GOLD	Pharmacophore, fragment-based design	K_d = 100–150 μM	32
TLR-4/MD2	MM/PBSA	MM/PBSA	K_F < 10 μM	24
uPAR/uPA	AutoDock	Multiple scoring functions and Ensemble docking	IC_{50} = 8–10 μM	87
uPAR/uPA	Glide		K_i = 0.7 μM	88

algorithms and scoring functions for most docking software, including Glide and AutoDock were validated using a test set of traditional drug targets.[20–23] This has necessitated the use of alternative methods to enhance HTVS results for PPI targets.[20] As shown in Table 3.1, the use of supplementary techniques that go beyond scoring and ranking potential hits has been common and highlighted the need to improve HTVS for PPIs. The supplementary techniques include molecular mechanics based calculation of theoretical binding energies,[24,25] the use of hot spot knowledge,[15–19] and pharmacophore modeling.[26-32] Pharmacophore modeling approaches have been used to discover small-molecule PPI inhibitors for murine double minute 2 homolog (MDM2)/p53, Abelson murine leukemia viral oncogene homolog 1 kinase (c-Abl)/14-3-3, Lens epithelium-derived growth factor (IN-LEDGF)/p75, interferon (IFN)/Interferon alpha/beta receptor 1 (IFNAR), and S100B/p53 PPIs.[26–32] The other methods have been developed to aid HTVS for discovering PPI inhibitors. Geppert et al. combined the methodologies for molecular descriptors, pharmacophore modeling, and HTVS to search inhibitors for IFN/IFNAR and trypsin-inhibitor/trypsin PPIs.[30,31] Molecular mechanics/Poisson-Boltzmann Surface Area (MM/PBSA) physics-based calculations were performed to aid in re-ranking HTVS results for the Toll-like receptor 4 (TLR4)/Lymphocyte antigen 96 (Md2) PPI.[26] Hot spot pocket knowledge has been used to enhance HTVS results for cytohesin-2 (Arno)/ adenosine diphosphate Ribosylation Factor 1 (Arf1) and negative regulatory factor (NEF)/SH3-domain PPIs[15,33] as well as the discovery of submicromolar inhibitors for G$\beta\gamma$ subunit-mediated PPIs.[16] The protruding hot spots of the ligand protein at the PPI interface have also been used as the template for HTVS to discover Ras-related C3 botulinum toxin substrate 1 (Rac1)/ T-lymphoma invasion and metastasis-inducing

protein 1 (Tiam1) inhibitors.[34] Despite the use of these complementary techniques, most small molecules discovered through HTVS for PPIs achieve only micromolar activities with low success rates.

On the other hand, it is possible to discover or design potent inhibitors for PPIs. Techniques such as structure–activity relationship (SAR) by NMR,[35,36] high-throughput screening by fluorescence polarization[37] or luminescent reporter assays,[38] and structure-based design[39] all have produced small-molecule inhibitors with submicromolar potency. Inhibitors with nanomolar to subnanomolar potency have been developed for X-linked inhibitor of apoptosis protein (XIAP)/baculovirus inhibitor of apoptosis protein repeat 3 (BIR3),[40] menin/mixed lineage leukemia protein (MLL),[41] MDM2/p53,[42] bromodomain-containing protein 4 (BRD4)/histone,[43,44] calpain/calpastatin,[45] Bcl-2/BH3-domain,[46] Bcl-X_L/BH3-domain,[36,47] and WD repeat protein 5 (WDR5) /MLL PPIs.[48,49] These examples prompted us to investigate the factors that might be able to increase the HTVS hit rate of small-molecule PPI inhibitors.

One approach to enhancing the rate for the discovery of more potent PPI inhibitors through HTVS is to identify the compounds that are more likely to bind hot spots.[15,16] A higher degree of ligand burial will be more likely to have higher ligand efficiency due to more contacts with the binding site. Mysinger et al. pointed out the problems of current scoring functions and the need to select small molecules that engage all of their functional groups with the target protein in their HTVS to discover small-molecule inhibitors for the chemokine receptor (CXCR4)/chemokine (CXCL12) PPI.[50] More importantly, the ligand will be more likely to interact with the hot spot pocket because it may induce the pocket formation.[6,7] In a recent paper comparing traditional

drug targets with PPIs, Gowthaman et al. described a property of inhibitors relating to their binding of shallow PPI surfaces.[24,25] This property, called θ_I, is the fraction of the bound over unbound SASAs of the inhibitor, as shown in Figure 3.1.[24,25] Evaluation of the SASA has proven to be useful tool to better understand PPIs. Indeed, the amount of buried surface area between two protein-protein interacting partners correlates with the K_d.[51] The more buried the two PPI partners are in each other, the lower the K_d. Examining the SASA has also provided the ability to computationally detect PPI hot spots.[13] Furthermore, SASA has proven a useful tool in determining the location where a small molecule likely binds.[52,53] Therefore, the basis for using SASA in evaluation of inhibitor burial is well supported. While θ_I is likely relevant for the inhibitors of both traditional and nontraditional targets, it is particularly important for PPI inhibitors because hot spot residues are often more buried in the PPI interface.[5,6] As the buried surface area was correlated with the PPI K_d, it was hypothesized that θ_I may be the equivalent measure of this relationship between small-molecule inhibitors and PPI targets.

Gowthaman, et al. concluded that traditional virtual screening software was ill adapted for PPIs.[24,25] They found that PPI inhibitors were more solvent exposed and have lower ligand efficiency than their traditional target counterparts. The authors stated that θ_I did not correlate with inhibitor potency between traditional and PPI targets or between different PPI targets. Herein, the relationship between θ_I and inhibitor potency for a single PPI target and the relationship between θ_I and the other HTVS associated properties are explored. The goals were to evaluate the feasibility of using θ_I to aid in identifying potent PPI hits and identify the determinants that affect the use of HTVS tools for the discovery of small-molecule PPI inhibitors.

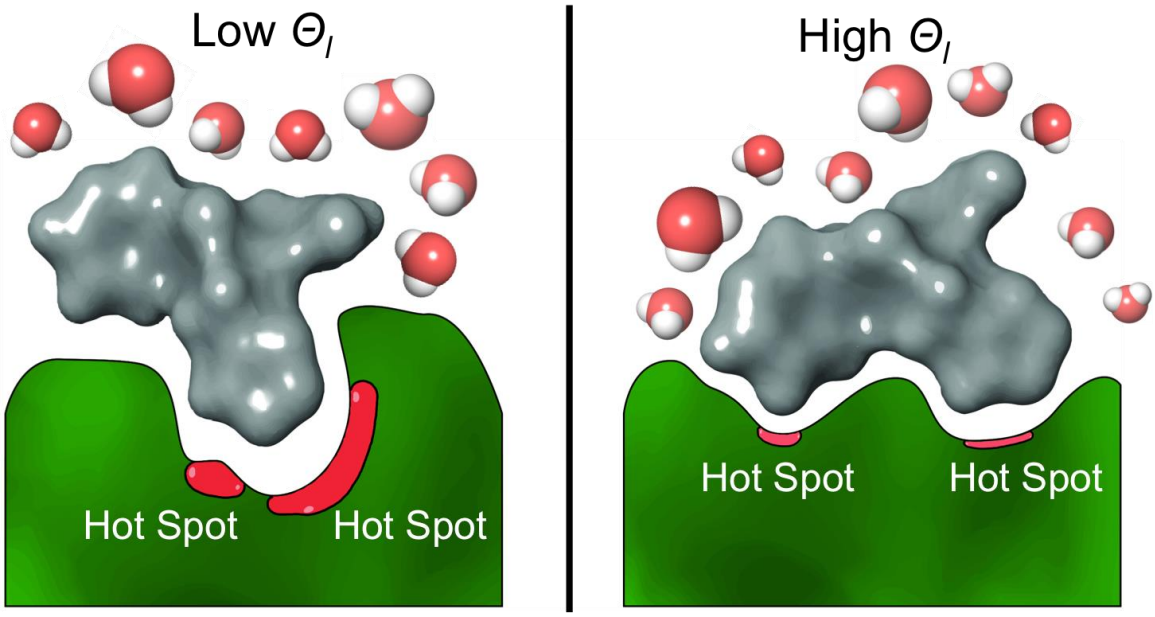


Figure 3.1. Illustrated example showing the differences in ligand burial. The inhibitor for the PPI site has a higher θ_I value because more of the ligand is solvent exposed compared to the lower θ_I of the enzyme inhibitor.

3.2 Methods

3.2.1 MDM2 crystal structure inhibitor study. Crystal structures of human or humanized MDM2 in complexes with potent small-molecule inhibitors for the MDM2/p53 PPI (3VZV,[55] 3W69,[56] 4OBA,[41] 4ERE,[41] 4ERF,[41] 4OQ3,[57] 4ODE,[58] 4OGN,[60] 4OGT,[60] 4OGV,[60] 4OCC,[60] 4ZYF,[59] 4ZYI,[61] 4LWU,[60] 4JVE,[61] 4JVR,[63] 4QO4,[62] 4IPF,[63] 4QOC[64], 4OAS[45]) were retrieved from the RCSB server. The structures and biological activities of these potent MDM2-bound inhibitors ($IC_{50} \leq 100$ nM) are shown in Figure 3.2. Ligands were prepared by adding appropriate bonds and hydrogen atoms using Schrödinger Maestro.[65] Proteins were prepared using the Schrödinger protein preparation workflow.[66,67] SASA calculations were carried out using the Schrödinger script *atomic_sasa.py*. The SASA was measured for the bound and unbound ligand structures. The fraction of SASA exposed for an inhibitor-protein complex, θ_I, was calculated by dividing the SASA of the bound ligand to that of the unbound ligand. The same 3D structure was used for the SASA calculations of the unbound ligand. Statistical analyses were performed using Microsoft Excel and Schrödinger Strike (Figure 3.3a).[68]

3.2.2 Bcl-X_L study. Compounds **1a**, **1b**, **8a-e**, **8h**, **8i**, **8k**, **8n**, **10f**, **10g**, **10i**, **10j**, **14a**, **14b**, **18**, **21**, **22**, **23a**, **23b**, **23e**, **23f**, and **23p** from Bruncko and colleagues' study of the structure-guided design and synthesis of Bcl-X_L and Bcl-2 dual inhibitors[69] were prepared using Schrödinger Ligprep and were protonated for physiological pH[70]. The structures and biological activities of these Bcl-X_L inhibitors are shown in Figure 3.4. These inhibitors represent the SAR study in which their K_i values were definitive, above the limit of detection, and more selective for Bcl-X_L over Bcl-2. The NMR structures of Bcl-X_L (PDB IDs 2O1Y, 2O2M, and 2O2N) were prepared using the Schrödinger

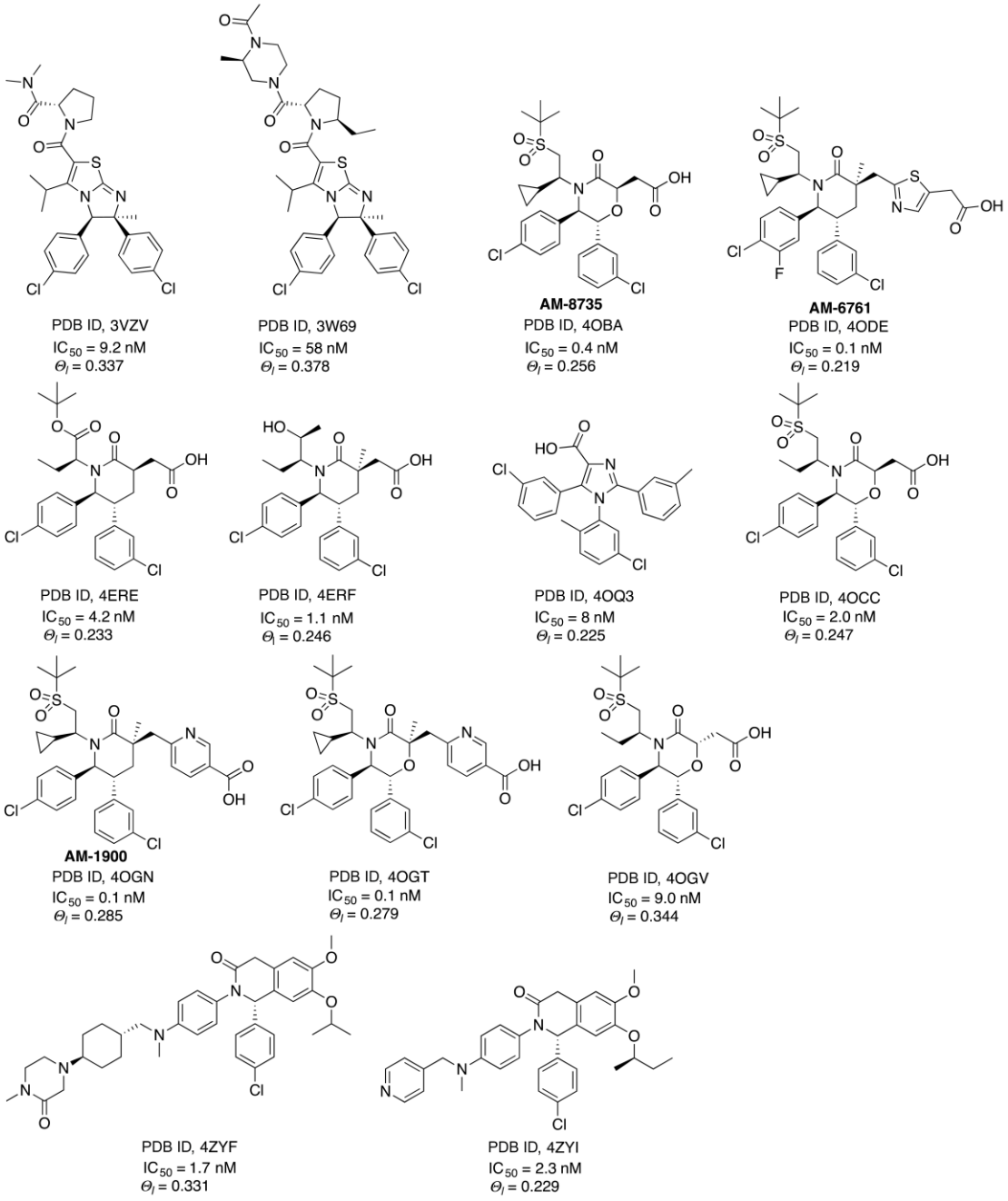


Figure 3.2. Structures and biological activities of MDM2/p53 PPI inhibitors.

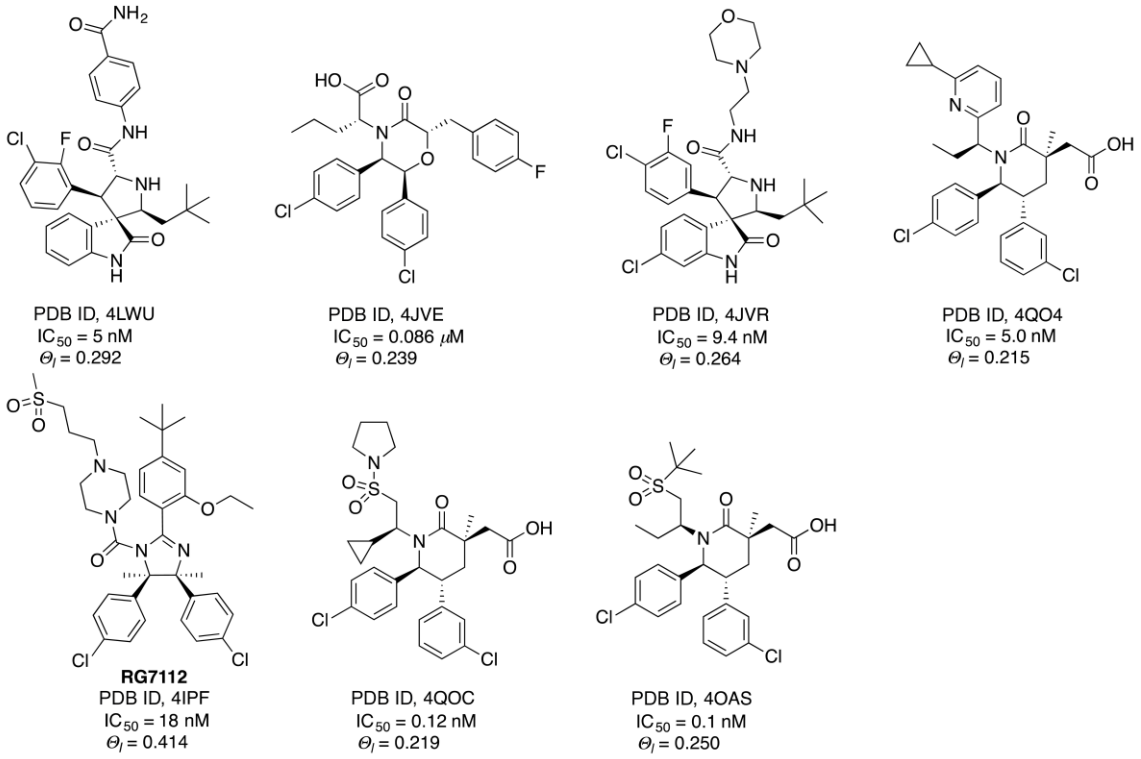


Figure 3.2. Continued.

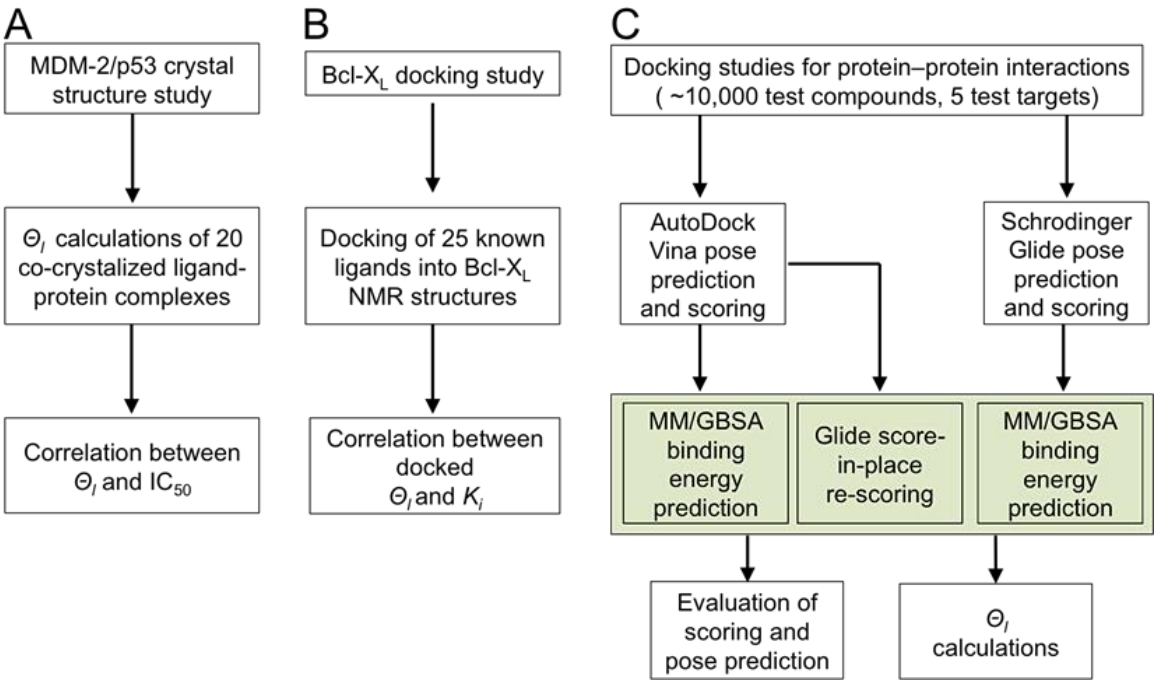


Figure 3.3. Workflows to evaluate the key factors for HTVS to discover inhibitors for PPIs. (A) Evaluation of the relationship between the burial and the potency of MDM2/p53 inhibitors that are co-crystalized with MDM2. (B) Evaluation of the relationship between the burial and the potency of a series of Bcl-X_L inhibitors. (C) The docking studies of approximately 10,000 test compounds using AutoDock Vina and Schrodinger Glide for five PPI targets. The flowchart shows the algorithms for pose prediction and the steps for evaluation of the scoring functions.

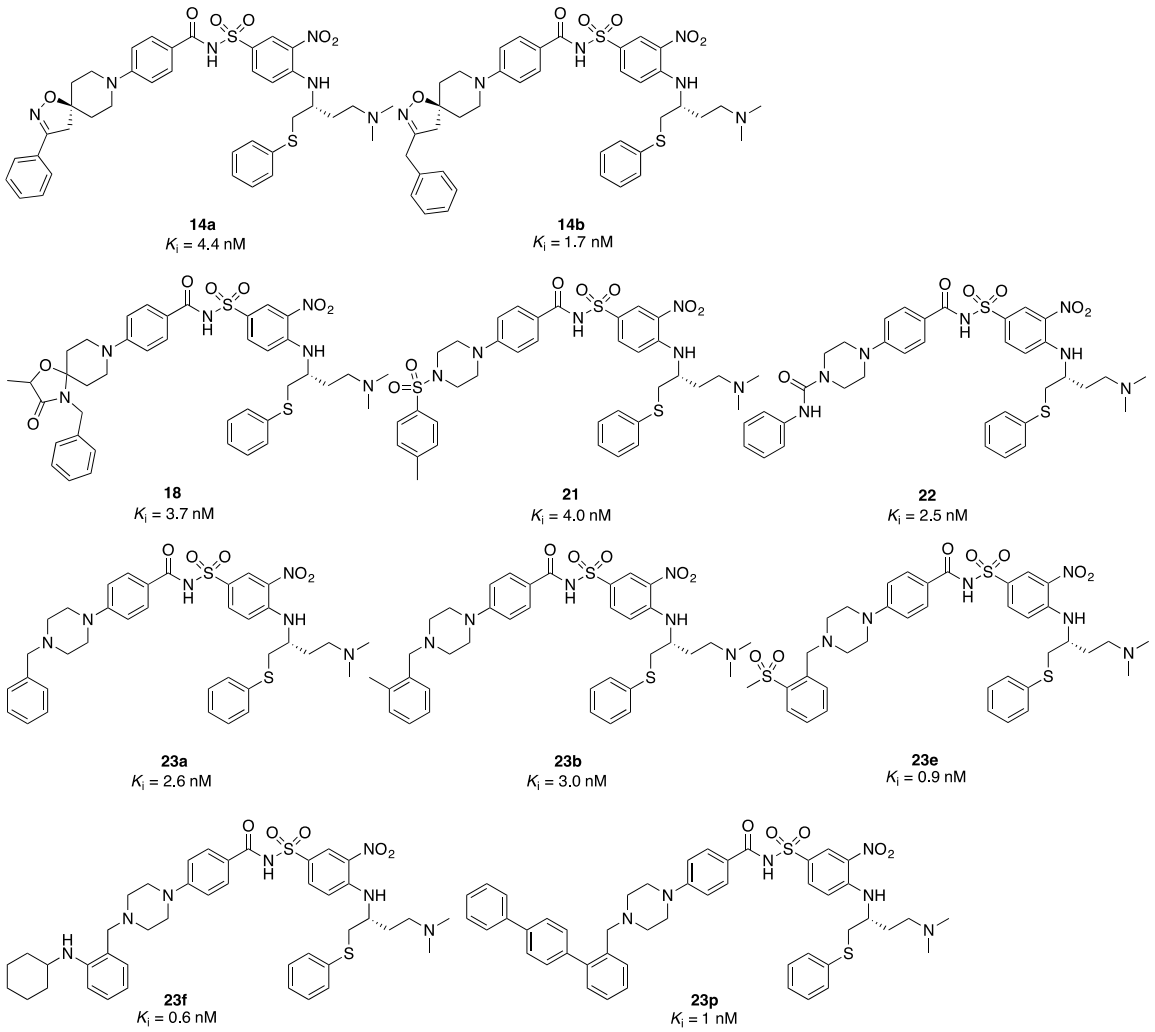


Figure 3.4. Structures and biological activities of Bcl-X_L/BH3-domain PPI inhibitors. Asterisk indicated mean ± standard error for three or more experiments run in duplicate for **8a**.[69]

protein preparation workflow[69]. Ligands were docked using AutoDock Vina with a 30 Å grid centered on the structure of the bound ligand encompassing all relevant residues. The exhaustiveness parameter was set to 12. Nine docked poses were generated for each compound. SASA calculations were carried out using the Schrödinger *atomic_sasa.py* script. The inhibitor poses were visually analyzed for features such as the sizes and shapes of the inhibitors. Statistical calculations were carried out using the data of the complex chosen as the best fit for each ligand based on visual shape similarity (Figure 3.3B). Statistical tests were calculated using Microsoft Excel and Schrödinger Strike.[70]

3.2.3 Docking study. Structures representing five potent PPI inhibitors in complexes with their respective targets were chosen to evaluate the performance of HTVS. These five structures are the XIAP-BIR3 domain (PDB ID, 1TFT)[71], Bcl-X_L (PDB ID, 1YSI)[36], BRD4 (PDB ID, 2YEL)[40], WDR5 (PDB ID, 3UR4)[72], and MDM2 (PDB ID, 4ERF)[41] and were downloaded from the RCSB web server. The receptor was prepared by removing all crystallographic waters, solvents, and ligands. Receptor structures were processed using the Schrödinger protein preparation workflow.[69] Missing side-chain atoms were added using Schrödinger Prime.[73,74] A subset of 10,000 structures from the ZINC library for HTVS with properties similar to that of the inhibitors used in Gowthaman and colleagues' study[24,25] was downloaded and prepared through Schrödinger Ligprep. These inhibitors have molecular weights between 200 and 750 Da, net charge between –1 and +2, and xlogP between –4.9 and +6.5. The protonation state was generated at pH = 7.0 using Epik[69]. Ligand structures encountering an error on any part of the preparation were discarded resulting in 9,300 structures. The native ligands from the crystal structures were extracted and prepared identically with that of the test

set. A cubic grid was generated for each structure centered on the bound ligand with a size of 30 $Å^3$. Ligands were docked using the Glide standard precision (SP)[27,58] default settings and the enhanced planarity for conjugated π systems. Ser, Thr, and Tyr side chain hydroxyl groups were allowed to rotate. A single output pose was generated for each docking job. Ligands were docked using Autodock Vina[76] with grid parameters similar to the Glide grid with exhaustiveness of 12. Nine poses were generated for each docked small molecule. The output poses from AutoDock Vina were rescored in place using Glide SP. Further, the output poses of both AutoDock Vina and Glide SP docking were rescored using Schrödinger Prime MM/GBSA calculations[77] with a minimization method. MM/GBSA enrichment was defined as the scoring rank divided by the MM/GBSA rank. SASA calculations were carried out using the Schrödinger *atomic_sasa.py* script. Statistical calculations were performed using Microsoft Excel.[70] To examine the correlation between docking scores and θ_I values, the known inhibitors were docked to its respective PPI target with AutoDock Vina. The settings were nummodes = 20, exhaustiveness = 12, energy_range = 5. Twenty poses were generated for 1TFT, 1YSI, 3UR4, and 4ERF. The docking of the ligand for 2YEL generated 19 poses. SASA and θ_I calculations and statistical analyses were performed using Microsoft Excel and Schrödinger Strike.[70]

3.3 Results

3.3.1 MDM2 crystal structure inhibitor study. A correlation was observed between IC_{50} and θ_I for the examined MDM2/p53 PPI inhibitors with Spearman's rho ($|r_s|$) = 0.327 (Figure 3.5a). The observed correlation was statistically significant at 90%

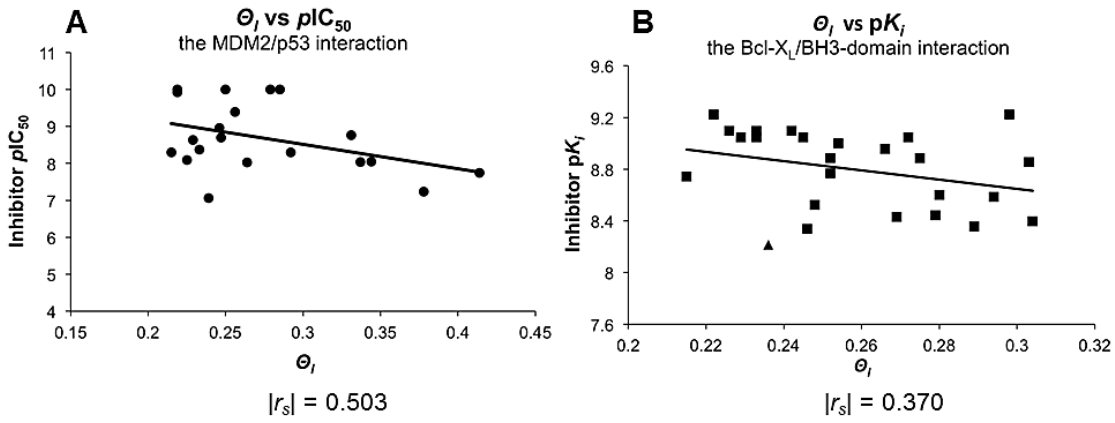


Figure 3.5. Correlation between inhibitor burial and ligand potency. (A) A moderate correlation was observed between inhibitor potency (pIC_{50}) and burial (θ_l) for 20 MDM2/p53 inhibitors. (B) Correlation between θ_l and K_i for a series of inhibitors for the Bcl-X_L/BH3-domain PPI. Compound **8a** was marked with a triangle due to the larger error in the measurement of its K_i value.

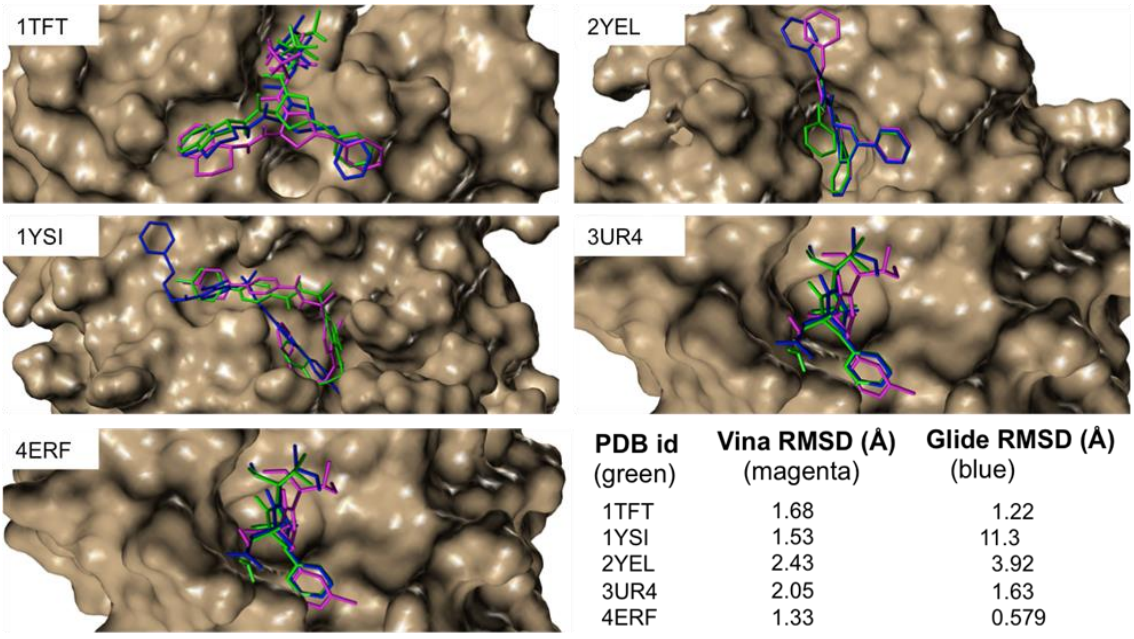
confidence (α = 0.1) for the size of the data set (n = 20). The inhibitors from structures 4ODE, 4OGN and 4OGT appear to be more potent than the other inhibitors that are equally buried. These inhibitors gain much of their extra potency from the strong ionic interaction with hot spot residue H96.[60] The other outlier of the group, 4JVE, is relatively not as potent despite its high degree of burial. The analysis of the binding mode of the inhibitors in the crystal structures showed that most of the potency was derived from hydrophobic contacts and most structures lack the ionic and hydrogen bonding interactions that were observed in 4ODE, 4OGN, and 4OGT. When these structures were removed, a stronger relationship was observed between IC_{50} and θ_I with $|r_s|$ = 0.503 (n = 16). This correlation was statistically significant at the 90% confidence interval. These data led to the conclusion that ligand burial was important, but secondary to the strong non-bonding interactions with the hot spot residues.

3.3.2 Bcl-X_L inhibitor study. A correlation, shown in Figure 3.5B, was observed between θ_I and K_i in a series of Bcl-X_L inhibitors ($|r_s|$ = 0.370, α = 0.05, n = 25). Compound **8a** was observed to have a higher K_i than the rest of the data set. Compound **8a** likely represents a case where the effect of θ_I is eclipsed by the other factors. If the data point for **8a** is removed because it has a relatively large reported standard error value[71] (standard error = 1.2) and the relatively higher K_i value, the $|r_s|$ value rises to 0.450 and shows a stronger correlation (n = 24).

3.3.3 Docking study. Table 3.2 and Figure 3.6 show a summary of the docking and scoring results. The success of the docking study was defined by two criteria: (1) the top-scored docking pose of the bioactive compound was similar to the crystallographic binding conformation; (2) the potent inhibitor observed in the crystal or NMR structure

Table 3.2. Results of the docking study for PPI inhibitors.

Structure (PDB id)	Vina success (Rank ≤ 1,000)	Glide success (Rank ≤ 1,000)	Vina docking–Glide re-scoring (Rank ≤ 1,000)	Vina docking–MM/GBSA (Enrichment)	Glide docking–MM/GBSA (Enrichment)
1TFT	successful (rmsd = 1.68 Å)	successful (rmsd = 1.22 Å)	failure	1.8×10^2	0.16
1YSI	scoring failure (rmsd = 1.53 Å)	pose failure (rmsd = 11.3 Å)	failure	0.83	0.81
2YEL	pose failure (rmsd = 2.43 Å)	pose failure (rmsd = 3.99 Å)	failure	0.011	0.2
3UR4	pose failure (rmsd = 2.05 Å)	scoring failure (rmsd = 1.63 Å)	successful	0.55	0.22
4ERF	scoring failure (rmsd = 1.33 Å)	successful (rmsd = 0.579 Å)	successful	7.2×10^2	96



PDB id (green)	Vina RMSD (Å) (magenta)	Glide RMSD (Å) (blue)
1TFT	1.68	1.22
1YSI	1.53	11.3
2YEL	2.43	3.92
3UR4	2.05	1.63
4ERF	1.33	0.579

Figure 3.6. Comparison of the docked poses generated by Schrodinger Glide and AutoDock Vina. The crystallographic or NMR poses are shown green. The poses that were predicted by AutoDock Vina and Schrodinger Glide SP were colored magenta and blue, respectively. AutoDock Vina was more successful in pose prediction, but Glide proved to be more accurate for three instances — 1TFT, 3UR4, and 4ERF.

as scored within the top 1,000 out of all the compounds in the test set. Based on the two criteria, AutoDock Vina successfully identified a single complex, 1TFT. The inhibitors of structures 1TFT, 1YSI, 2YEL, 3UR4, and 4ERF achieved the docking poses with root-mean-square deviations (rmsds) of 1.68 Å, 1.53 Å, 2.43 Å, 2.05 Å, and 1.33 Å, respectively. The docking results for 2YEL and 3UR4 were pose failures because their rmsds from the crystallographic binding conformations were > 2 Å, as shown in Table 3.1. The 2YEL pose failure was the result of an incorrect amide bond dihedral angle. The rmsd of the docking pose for 3UR4 was just out of the boundary of our definition of a good pose (rmsd = 2 Å). This pose is still acceptable and can be useful for guiding inhibitor optimization. It should be noted that the fourth pose ranked by the AutoDock Vina scoring function for 3UR4 had an rmsd of 1.19 Å, but this pose was not counted as a success because only the first pose would be considered in virtual screening. Glide SP had two successes, 1TFT and 4ERF, with rmsds of 1.22 Å and 0.579 Å, respectively. 3UR4 was a scoring failure, but achieved an rmsd of 1.68 Å. The Glide results for 1YSI and 2YEL were pose failures with rmsds of 11.3 Å and 3.92 Å, respectively. The re-scoring of the AutoDock Vina poses using Glide score-in-place resulted in two successes for the top picks (3UR4 and 4ERF). The sole original success in the AutoDock Vina study, 1TFT, was scored worse by Glide score-in-pace. The result for 2YEL was not improved by Glide score-in-place likely due to the initial AutoDock Vina pose failure. An enrichment value was calculated when the poses were re-evaluated using MM/GBSA. Enrichment was defined as rank of docking score divided by the rank of MM/GBSA energy. A positive enrichment (enrichment > 1) meant that the ranking was improved by the MM/GBSA energy calculation, while a negative enrichment (enrichment < 1) refers

to when the inhibitor was scored worse by MM/GBSA than the scoring function. As shown in Table 3.2, two inhibitors from the AutoDock Vina study, 1TFT and 4ERF, were significantly positively enriched. One inhibitor from the Glide study, 4ERF, was also positively enriched. All remaining structures from both studies were not enriched in the MM/GBSA ranking.

The θ_l results did not correlate with the docking scores or the MM/GBSA $\Delta\Delta G_{binding}$ values between different ligands. The θ_l values provided no enrichment in ranking the known potent inhibitors that were complexed in the crystal or NMR structures. However, when examining the relationship between the AutoDock Vina score and θ_l, we found a moderate-to-strong correlation ($|r_s|$ = 0.318-0.769) between these two variables within the same ligand, as shown in Figure 3.7. The correlation was significant (α = 0.1) for the docking poses in all of the PPI targets.

3.4 Discussion

The MDM2/p53 PPI is a high priority target due to its involvement in many aggressive and therapy-resistant cancers.[77] As a result, many classes of inhibitors have been developed for this PPI.[78] This has made the MDM2/p53 PPI the preferred target from which to begin the study of ligand burial in PPI inhibitors. Gowthaman et al.[20] did not observe a statistically significant relationship between θ_l and inhibitor potency when comparing traditional targets with PPIs or between different PPI complexes. This is likely because the property θ_l is relative to both inhibitor and protein structures, so such a comparison between different protein structures might be inaccurate.[24,25] In the Gowthama. study, the value of θ_l indeed varies significantly across different

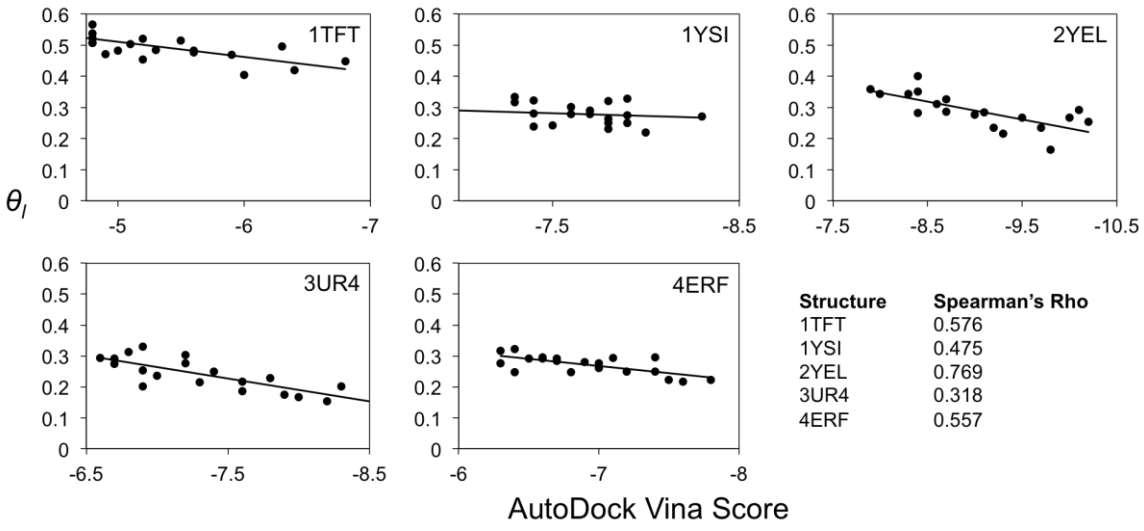


Figure 3.7. Correlation between docking score and ligand burial. Varying levels of correlations between θ_I and the AutoDock Vina score were observed for the PPI targets. This indicates that AutoDock Vina tends to score more buried poses higher than less buried poses. This is likely because the more buried the ligand is, the more favorable interactions it may have with the receptor.

inhibitor–protein complexes. The work presented in this study was focused on a single target, the MDM2/p53 PPI. The studied crystal structures were complexed with potent, nonfragment, small-molecule inhibitors (IC_{50} < 100 nM). In this data set a moderate but significant correlation between inhibitor potency (IC_{50}) and the fraction of the ligand that remains solvent exposed (θ_l) was found. This relationship was not more strongly correlated because ligand burial was secondary to strong nonbonding interactions. The examination of the inhibitor binding mode in MDM2 crystal structures, 4ODE, 4OGT, and 4OGN, indicated that the inhibitors gained more potency through ionic interactions with an additional MDM2 hot spot, H96. On the other end of the spectrum was 4JVE, in which the inhibitor was as buried as 4ERE and 4ERF but exhibited much weaker potency due to the absence of strong ionic and/or hydrogen bonds. The introduction of strong noncovalent bonding such as ionic and H-bonding interactions between the potential small-molecule inhibitors and the PPI targets becomes an exciting and important research direction for the discovery of potent PPI inhibitors, in particular considering many PPI interfaces are shallow, hydrophobic, and require low desolvation penalty. The current drug design strategies for introducing ionic functional groups and/or H-bond donors and acceptors cannot overcome the low desolvation penalty requirement to bind shallow PPI interfaces.

To further explore the relationship between inhibitor potency and burial, we examined the Bcl-X_L/BH3-domain PPI, another important therapeutic target. In contrast to the MDM2 crystal structure inhibitor study, a specific series of Bcl-X_L inhibitors with the variations on the designated side chains was chosen for this study.[71] As shown in Figure 3.6, the property θ_l was found to be correlated with the K_i values of the Bcl-

X_L/BH3-domain inhibitors. Furthermore, Bcl-X_L is known to have a flexible interaction surface, a structural feature of many PPI targets. Therefore, the result in Figure 3.6 indicated that θ_I could be used to assess the potency of PPI inhibitor derivatives. This result also echoes a previous study that found most nonfragment inhibitors which are structurally similar bind in similar fashion.[79]

Next the effect of ligand burial on HTVS was evaluated, confirming similar observations from Gowthaman et al.[24,25] that current virtual screening programs were ill-suited for docking PPI inhibitors. The ability of the common docking programs to determine a proper pose and to identify potent PPI inhibitors from a series of putatively inactive compounds was evaluated. The inhibitor-bound protein complexes were chosen for this study because they are all examples of highly potent inhibitors bound to a PPI interface. It was found that AutoDock Vina was more accurate in delivering a valid pose with lower rmsds from the biologically active conformations in the crystal or NMR structures. However, AutoDock Vina was less effective at scoring the known inhibitors relative to the putatively inactive compounds. In contrast, Glide more often identified the known inhibitors, but generated less than optimal predicted poses. One possible explanation for these results could be the differences of the docking algorithms and the scoring functions used by the two programs. AutoDock Vina uses a Lamarckian genetic algorithm to dock the ligand and a knowledge-based scoring function to rank the poses. This knowledge-based scoring function was originally proven using traditional targets, not PPI targets, when it was developed.[78] Glide fared better at scoring known inhibitors, but produced inferior poses. It was unclear if increasing the number of poses produced by Glide SP would have aided in pose prediction, but only the top picks are usually assessed

in virtual screening. AutoDock Vina uses a machine learning scoring function, described by Trott and Olson,[78] as opposed to the empirical scoring function used by Glide and Gold. The empirical scoring function may be more appropriate for PPIs because it is more physics-based. Success in HTVS is to not only identify a potential hit but also to produce a valid binding pose that could be used to guide further inhibitor optimization. Therefore, the AutoDock Vina poses were rescored using Glide SP score-in-place, with the belief that the empirical scoring function would be better at identifying known inhibitors. As expected, the poses predicted by AutoDock Vina and rescored by Glide gave better results for two docked ligands, 3UR4 and 4ERF. As observed by Trott and Olson,[78] the Lamarckian genetic algorithm for docking was able to more accurately predict the poses for the known ligands. However, the one success of AutoDock Vina, 1TFT, was ranked worse with Glide rescoring. The MM/GBSA reranking was overall no more effective than Glide rescoring of the AutoDock Vina poses. In some cases, MM/GBSA reranking was worse. The results of the HTVS study suggested that an empirical scoring method for ranking potential hits might improve results, but neither rescoring using different scoring functions nor the computational estimates of binding energy by MM/GBSA were sufficient to identify potent PPI hits.

Lastly, the relationship of θ_I to the different docked poses of the same ligand was examined using AutoDock Vina. A moderate-to-strong correlation was observed for each of the five PPIs, indicating that the docking program better scored the poses that were more deeply buried (Figure 3.7). The property θ_I alone cannot rank HTVS hits because it is secondary to the strong noncovalent interactions. Regardless, given the evidence from Gowthaman et al.[20] and the studies presented here, it would be wise to consider θ_I as an

adjunct property when selecting hits from an HTVS. A more buried ligand will provide a better starting point for inhibitor development because it can more efficiently use the limited space available in PPI pockets, some of which contain hot spots.

3. 5 Conclusion

PPIs represent a class of appealing drug targets where conventional methods for drug discovery have been met with difficulty. This is due to the nature of the PPI interface which contrasts in size and shape with traditional drug targets. One way to overcome these problems is through the use of hot spots,[8-10] or the regions in the PPI interface that contribute more to binding energy. HTVS has had moderate success at finding hits for PPIs; however, most of these hits were typically discovered with the aid of other methods and exhibited low potency. Two critical components need to be assessed in HTVS: correctly ranking hits with good potency, and predicting a valid pose which can be optimized further. This evaluated the ability of two common molecular docking programs with distinct docking and scoring algorithms to identify known, highly potent, small-molecule inhibitors for PPIs. It was found that the choice of scoring function is important when ranking HTVS results, and the use of MM/GBSA calculated $\Delta\Delta G_{binding}$ to rerank inhibitors was met with mixed results. The study evaluated an additional property, θ_l, which is relevant to all small molecule–protein interactions, but addresses specific challenges in inhibiting PPIs, such as low ligand efficiency due to shallow binding surfaces. This property relates with the efficiency of ligand burial and, therefore, with hot spots. θ_l shows a positive correlation with IC_{50} or K_i when examining known inhibitors for the MDM2/p53 and Bcl-X_L/BH3-domain PPIs. While no effect of θ_l

on reranking docking results was observed, this could be due to the strong non-covalent interactions. Therefore, θ_l is an important property to consider when evaluating HTVS results.

3.6 References

1. Peri, S.; Navarro, J. D.; Amanchy, R.; Kristiansen, T. Z.; Jonnalagadda, C. K.; Surendranath, V.; Niranjan, V.; Muthusamy, B.; Gandhi, T. K. B.; Gronborg, M.; Ibarrola, N.; Deshpande, N.; Shanker, K.; Shivashankar, H. N.; Rashmi, B. P.; Ramya, M. A.; Zhao, Z.; Chandrika, K. N.; Padma, N.; Harsha, H. C.; Yatish, A. J.; Kavitha, M. P.; Menezes, M.; Choudhury, D. R.; Suresh, S.; Ghosh, N.; Saravana, R.; Chandran, S.; Krishna, S.; Joy, M.; Anand, S. K.; Madavan, V.; Joseph, A.; Wong, G. W.; Schiemann, W. P.; Constantinescu, S. N.; Huang, L.; Khosravi-Far, R.; Steen, H.; Tewari, M.; Ghaffari, S.; Blobe, G. C.; Dang, C. V.; Garcia, J. G. N.; Pevsner, J.; Jensen, O. N.; Roepstorff, P.; Deshpande, K. S.; Chinnaiyan, A. M.; Hamosh, A.; Chakravarti, A.; Pandey, A. Development of Human Protein Reference Database as an Initial Platform for Approaching Systems Biology in Humans. *Genome Res.* **2003**, *13* (10), 2363–2371.

2. Mishra, G. R.; Suresh, M.; Kumaran, K.; Kannabiran, N.; Suresh, S.; Bala, P.; Shivakumar, K.; Anuradha, N.; Reddy, R.; Raghavan, T. M.; Menon, S.; Hanumanthu, G.; Gupta, M.; Upendran, S.; Gupta, S.; Mahesh, M.; Jacob, B.; Mathew, P.; Chatterjee, P.; Arun, K. S.; Sharma, S.; Chandrika, K. N.; Deshpande, N.; Palvankar, K.; Raghavnath, R.; Krishnakanth, R.; Karathia, H.; Rekha, B.; Nayak, R.; Vishnupriya, G.; Kumar, H. G. M.; Nagini, M.; Kumar, G. S. S.; Jose, R.; Deepthi, P.; Mohan, S. S.; Gandhi, T. K. B.; Harsha, H. C.; Deshpande, K. S.; Sarker, M.; Prasad, T. S. K.; Pandey, A. Human Protein Reference Database–2006 Update. *Nucleic Acids Res.* **2006**, *34*, D411–D414.

3. Keshava Prasad, T. S.; Goel, R.; Kandasamy, K.; Keerthikumar, S.; Kumar, S.; Mathivanan, S.; Telikicherla, D.; Raju, R.; Shafreen, B.; Venugopal, A.; Balakrishnan, L.; Marimuthu, A.; Banerjee, S.; Somanathan, D. S.; Sebastian, A.; Rani, S.; Ray, S.; Harrys Kishore, C. J.; Kanth, S.; Ahmed, M.; Kashyap, M. K.; Mohmood, R.; Ramachandra, Y. L.; Krishna, V.; Rahiman, B. A.; Mohan, S.; Ranganathan, P.; Ramabadran, S.; Chaerkady, R.; Pandey, A. Human Protein Reference Database–2009 Update. *Nucleic Acids Res.* **2009**, *37*, D767–D772.

4. Arkin, M. R.; Tang, Y.; Wells, J. A. Small-Molecule Inhibitors of Protein-Protein Interactions: Progressing Toward the Reality. *Chem. Biol.* **2014**, *21* (9), 1102–1114.

5. Wells, J. A.; Mcclendon, C. L. Reaching for High-Hanging Fruit in Drug Discovery at

Protein–Protein Interfaces. *Nature* **2007**, *450* (7172), 1001–1009.

6. Laraia, L.; Mckenzie, G.; Spring, D. R.; Venkitaraman, A. R.; Huggins, D. J. Overcoming Chemical, Biological, and Computational Challenges in the Development of Inhibitors Targeting Protein–Protein Interactions. *Chem. Biol.* **2015**, *22* (6), 689–703.

7. Sperandio, O.; Reynès, C. H.; Camproux, A.-C.; Villoutreix, B. O. Rationalizing the Chemical Space of Protein–Protein Interaction Inhibitors. *Drug Discov. Today* **2010**, *15* (5–6), 220–229.

8. Clackson, T.; Wells, J. A. A Hot Spot of Binding Energy in a Hormone-Receptor Interface. *Science* **1995**, *267* (5196), 383–386.

9. Bogan, A. A.; Thorn, K. S. Anatomy of Hot Spots in Protein Interfaces. *J. Mol. Biol.* **1998**, *280* (1), 1–9.

10. Ma, B.; Elkayam, T.; Wolfson, H.; Nussinov, R. Protein–Protein Interactions: Structurally Conserved Residues Distinguish Between Binding Sites and Exposed Protein Surfaces. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100* (10), 5772–5777.

11. Keskin, O.; Ma, B.; Nussinov, R. Hot Regions in Protein–Protein Interactions: the Organization and Contribution of Structurally Conserved Hot Spot Residues. *J. Mol. Biol.* **2005**, *345* (5), 1281–1294.

12. Tuncbag, N.; Gursoy, A.; Keskin, O. Identification of Computational Hot Spots in Protein Interfaces: Combining Solvent Accessibility and Inter-Residue Potentials Improves the Accuracy. *Bioinformatics* **2009**, *25* (12), 1513–1520.

13. Rajamani, D.; Thiel, S.; Vajda, S.; Camacho, C. J. Anchor Residues in Protein–Protein Interactions. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101* (31), 11287–11292.

14. Tesmer, J. J. G. Pharmacology. Hitting the Hot Spots of Cell Signaling Cascades. *Science* **2006**, *312* (5772), 377–378.

15. Bonacci, T. M.; Mathews, J. L.; Yuan, C.; Lehmann DM, Malik S, Wu D, Font JL, Bidlack JM, Smrcka AV. Differential Targeting of Gβγ-Subunit Signaling with Small Molecules. *Science* **2006**, *312* (5772), 443–446.

16. Rouhana, J.; Hoh, F.; Estaran, S.; Henriquet, C.; Boublik, Y.; Kerkour, A.; Trouillard, R.; Martinez, J.; Pugnière, M.; Padilla, A.; Chavanieu, A. Fragment-Based Identification of a Locus in The Sec7 Domain of Arno for the Design of Protein–Protein Interaction Inhibitors. *J. Med. Chem.* **2013**, *56* (21), 8497–8511.

17. Yu, B.; Huang, Z.; Zhang, M.; Dillard, D. R.; Ji, H. Rational Design of Small-Molecule Inhibitors for *β*-Catenin/T-Cell Factor Protein–Protein Interactions by

Bioisostere Replacement. *ACS Chem. Biol.* **2013**, *8* (3), 524–529.

18. Catrow, J. L.; Zhang, Y.; Zhang, M.; Ji, H. Discovery of Selective Small-Molecule Inhibitors for the *β*-Catenin/T-Cell Factor Protein–Protein Interaction Through the Optimization of the Acyl Hydrazone Moiety. *J. Med. Chem.* **2015**, *58* (11), 4678–4692.

19. Hoggard, L. R.; Zhang, Y.; Zhang, M.; Panic, V.; Wisniewski, J. A.; Ji, H. Rational Design of Selective Small-Molecule Inhibitors for *β*-Catenin/B-Cell Lymphoma 9 Protein–Protein Interactions. *J. Am. Chem. Soc.* **2015**, *137* (38), 12249–12260.

20. Gowthaman, R.; Deeds, E. J.; Karanicolas, J. Structural Properties of Non-Traditional Drug Targets Present New Challenges for Virtual Screening. *J. Chem. Inf. Model.* **2013**, *53* (8), 2073–2081.

21. Huey, R.; Morris, G. M.; Olson, A. J.; Goodsell, D. S. A Semiempirical Free Energy Force Field with Charge-Based Desolvation. *J. Comput. Chem.* **2007**, *28* (6), 1145–1152.

22. Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: a New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **2004**, *47* (7), 1739–1749.

23. Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: a New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *J. Med. Chem.* **2004**, *47* (7), 1750–1759.

24. Joce, C.; Stahl, J. A.; Shridhar, M.; Hutchinson, M. R.; Watkins, L. R.; Fedichev, P. O.; Yin, H. Application of a Novel In Silico High-Throughput Screen to Identify Selective Inhibitors for Protein–Protein Interactions. *Bioorg. Med. Chem. Lett.* **2010**, *20* (18), 5411–5413.

25. Zhang, X.; Wong, S. E.; Lightstone, F. C. Toward Fully Automated High Performance Computing Drug Discovery: a Massively Parallel Virtual Screening Pipeline for Docking and Molecular Mechanics/Generalized Born Surface Area Rescoring to Improve Enrichment. *J. Chem. Inf. Model.* **2014**, *54* (1), 324–337.

26. Lu, Y.; Nikolovska-Coleska, Z.; Fang, X.; Gao, W.; Shangary, S.; Qiu, S.; Qin, D.; Wang, S. Discovery of a Nanomolar Inhibitor of the Human Murine Double Minute 2 (MDM2)-P53 Interaction Through an Integrated, Virtual Database Screening Strategy. *J. Med. Chem.* **2006**, *49* (13), 3759–3762.

27. Christ, F.; Voet, A.; Marchand, A.; Nicolet, S.; Desimmie, B. A.; Marchand, D.;

Bardiot, D.; Van Der Veken, N. J.; Van Remoortel, B.; Strelkov, S. V.; De Maeyer, M.; Chaltin, P.; Debyser, Z. Rational Design of Small-Molecule Inhibitors of the LEDGF/P75-Integrase Interaction and HIV Replication. *Nat. Chem. Biol.* **2010**, *6* (6), 442–448.

28. Serrao, E.; Debnath, B.; Otake, H.; Kuang, Y.; Christ, F.; Debyser, Z.; Neamati, N. Fragment-Based Discovery of 8-Hydroxyquinoline Inhibitors of the HIV-1 Integrase-Lens Epithelium-Derived Growth Factor/P75 (IN-LEDGF/P75) Interaction. *J. Med. Chem.* **2013**, *56* (6), 2311–2322.

29. Mancini, M.; Corradi, V.; Petta, S.; Barbieri, E.; Manetti, F.; Botta, M.; Santucci, M. A. A New Nonpeptidic Inhibitor of 14-3-3 Induces Apoptotic Cell Death in Chronic Myeloid Leukemia Sensitive or Resistant to Imatinib. *J. Pharmacol. Exp. Ther.* **2011**, *336* (3), 596–604.

30. Geppert, T.; Bauer, S.; Hiss, J. A.; Conrad, E.; Reutlinger, M.; Schneider, P.; Weisel, M.; Pfeiffer, B.; Altmann, K.-H.; Waibler, Z.; Schneider, G. Immunosuppressive Small Molecule Discovered by Structure-Based Virtual Screening for Inhibitors of Protein–Protein Interactions. *Angew. Chem. Int. Ed.* **2012**, *51* (1), 258–261.

31. Geppert, T.; Reisen, F.; Pillong, M.; Hähnke, V.; Tanrikulu, Y.; Koch, C. P.; Perna, A. M.; Perez, T. B.; Schneider, P.; Schneider, G. Virtual Screening for Compounds that Mimic Protein–Protein Interface Epitopes. *J. Comput. Chem.* **2012**, *33* (5), 573–579.

32. Agamennone, M.; Cesari, L.; Lalli, D.; Turlizzi, E.; Del Conte, R.; Turano, P.; Mangani, S.; Padova, A. Fragmenting the S100B-P53 Interaction: Combined Virtual/Biophysical Screening Approaches to Identify Ligands. *Chemmedchem* **2010**, *5* (3), 428–435.

33. Betzi, S.; Restouin, A.; Opi, S.; Arold, S. T.; Parrot, I.; Guerlesquin, F.; Morelli, X.; Collette, Y. Protein–Protein Interaction Inhibition (2P2I) Combining High Throughput and Virtual Screening: Application to the HIV-1 Nef Protein. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104* (49), 19256–19261.

34. Gao, Y.; Dickerson, J. B.; Guo, F.; Zheng, J.; Zheng, Y. Rational Design and Characterization of a Rac Gtpase-Specific Small Molecule Inhibitor. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101* (20), 7618–7623.

35. Petros, A. M.; Huth, J. R.; Oost, T.; Park, C.-M.; Ding, H.; Wang, X.; Zhang, H.; Nimmer, P.; Mendoza, R.; Sun, C.; Mack, J.; Walter, K.; Dorwin, S.; Gramling, E.; Ladror, U.; Rosenberg, S. H.; Elmore, S. W.; Fesik, S. W.; Hajduk, P. J. Discovery of a Potent and Selective Bcl-2 Inhibitor Using SAR by NMR. *Bioorg. Med. Chem. Lett.* **2010**, *20* (22), 6587–6591.

36. Oltersdorf, T.; Elmore, S. W.; Shoemaker, A. R.; Armstrong, R. C.; Augeri, D. J.;

Belli, B. A.; Bruncko, M.; Deckwerth, T. L.; Dinges, J.; Hajduk, P. J.; Joseph, M. K.; Kitada, S.; Korsmeyer, S. J.; Kunzer, A. R.; Letai, A.; Li, C.; Mitten, M. J.; Nettesheim, D. G.; Ng, S.; Nimmer, P. M.; O'Connor, J. M.; Oleksijew, A.; Petros, A. M.; Reed, J. C.; Shen, W.; Tahir, S. K.; Thompson, C. B.; Tomaselli, K. J.; Wang, B.; Wendt, M. D.; Zhang, H.; Fesik, S. W.; Rosenberg, S. H. An Inhibitor of Bcl-2 Family Proteins Induces Regression of Solid Tumours. *Nature* **2005**, *435* (7042), 677–681.

37. Nikolovska-Coleska, Z.; Wang, R.; Fang, X.; Pan, H.; Tomita, Y.; Li, P.; Roller, P. P, Krajewski, K.; Saito, N. G.; Stuckey, J. A.; Wang, S. Development and Optimization of a Binding Assay for the XIAP BIR3 Domain Using Fluorescence Polarization. *Anal. Biochem.* **2004**, *332* (2), 261–273.

38. Chung, C.-W.; Coste, H.; White, J. H.; Mirguet, O.; Wilde, J.; Gosmini, R. L.; Delves, C.; Magny, S. M.; Woodward, R.; Hughes, S. A.; Boursier, E. V.; Flynn, H.; Bouillot, A. M.; Bamborough, P.; Brusq, J.-M. G.; Gellibert, F. J.; Jones, E. J.; Riou, A. M.; Homes, P.; Martin, S. L.; Uings, I. J.; Toum, J.; Clément, C. A.; Boullay, A.-B.; Grimley, R. L.; Blandel, F. M.; Prinjha, R. K.; Lee, K.; Kirilovsky, J.; Nicodeme, E. Discovery and Characterization of Small Molecule Inhibitors of the BET Family Bromodomains. *J. Med. Chem.* **2011**, *54* (11), 3827–3838.

39. Rew, Y.; Sun, D.; Gonzalez-Lopez De Turiso, F.; Bartberger, M. D.; Beck, H. P.; Canon, J.; Chen, A.; Chow, D.; Deignan, J.; Fox, B. M.; Gustin, D.; Huang, X.; Jiang, M.; Jiao, X.; Jin, L.; Kayser, F.; Kopecky, D. J.; Li, Y.; Lo, M.-C.; Long, A. M.; Michelsen, K.; Oliner, J. D.; Osgood, T.; Ragains, M.; Saiki, A. Y.; Schneider, S.; Toteva, M.; Yakowec, P.; Yan, X.; Ye, Q.; Yu, D.; Zhao, X.; Zhou, J.; Medina, J. C.; Olson, S. H. Structure-Based Design of Novel Inhibitors of the MDM2-P53 Interaction. *J. Med. Chem.* **2012**, *55* (11), 4936–4954.

40. Cai, Q.; Sun, H.; Peng, Y.; Lu, J.; Nikolovska-Coleska, Z.; Mceachern, D.; Liu, L.; Qiu, S.; Yang, C.-Y.; Miller, R.; Yi, H.; Zhang, T.; Sun, D.; Kang, S.; Guo, M.; Leopold, L.; Yang, D.; Wang, S. A Potent and Orally Active Antagonist (SM-406/AT-406) of Multiple Inhibitor of Apoptosis Proteins (Iaps) in Clinical Development for Cancer Treatment. *J. Med. Chem.* **2011**, *54* (8), 2714–2726.

41. Borkin, D.; Pollock, J.; Kempinska, K.; Purohit, T.; Li, X.; Wen, B.; Zhao, T.; Miao, H.; Shukla, S.; He, M.; Sun, D.; Cierpicki, T.; Grembecka, J. Property Focused Structure-Based Optimization of Small Molecule Inhibitors of the Protein-Protein Interaction Between Menin and Mixed Lineage Leukemia (MLL). *J. Med. Chem.* **2016**, *59* (3), 892–913.

42. Sun, D.; Li, Z.; Rew, Y.; Gribble, M.; Bartberger, M. D.; Beck, H. P.; Canon, J.; Chen, A.; Chen, X.; Chow, D.; Deignan, J.; Duquette, J.; Eksterowicz, J.; Fisher, B.; Fox, B. M.; Fu, J.; Gonzalez, A. Z.; Gonzalez-Lopez De Turiso, F.; Houze, J. B.; Huang, X.; Jiang, M.; Jin, L.; Kayser, F.; Liu, J. J.; Lo, M.-C.; Long, A. M.;

Lucas, B.; Mcgee, L. R.; Mcintosh, J.; Mihalic, J.; Oliner, J. D.; Osgood, T.; Peterson, M. L.; Roveto, P.; Saiki, A. Y.; Shaffer, P.; Toteva, M.; Wang, Y.; Wang, Y. C.; Wortman, S.; Yakowec, P.; Yan, X.; Ye, Q.; Yu, D.; Yu, M.; Zhao, X.; Zhou, J.; Zhu, J.; Olson, S. H.; Medina, J. C. Discovery Of AMG 232, A Potent, Selective, and Orally Bioavailable MDM2-P53 Inhibitor in Clinical Development. *J. Med. Chem.* **2014**, *57* (4), 1454–1472.

43. Filippakopoulos, P.; Qi, J.; Picaud, S.; Shen, Y.; Smith, W. B.; Fedorov, O.; Morse, E. M.; Keates, T.; Hickman, T. T.; Felletar, I.; Philpott, M.; Munro, S.; Mckeown, M. R.; Wang, Y.; Christie, A. L.; West, N.; Cameron, M. J.; Schwartz, B.; Heightman, T. D.; La Thangue, N.; French, C. A.; Wiest, O.; Kung, A. L.; Knapp, S.; Bradner, J. E. Selective Inhibition of BET Bromodomains. *Nature* **2010**, *468* (7327), 1067–1073.

44. Gosmini, R.; Nguyen, V. L.; Toum, J.; Simon, C.; Brusq, J.-M. G.; Krysa, G.; Mirguet, O.; Riou-Eymard, A. M.; Boursier, E. V.; Trottet, L.; Bamborough, P.; Clark, H.; Chung, C.-W.; Cutler, L.; Demont, E. H.; Kaur, R.; Lewis, A. J.; Schilling, M. B.; Soden, P. E.; Taylor, S.; Walker, A. L.; Walker, M. D.; Prinjha, R. K.; Nicodème, E. The Discovery Of I-BET726 (GSK1324726A), A Potent Tetrahydroquinoline Apoa1 Up-Regulator and Selective BET Bromodomain Inhibitor. *J. Med. Chem.* **2014**, *57* (19), 8111–8131.

45. Low, K. E.; Ler, S.; Chen, K. J.; Campbell, R. L.; Hickey, J. L.; Tan, J.; Scully, C. C. G.; Davies, P. L.; Yudin, A. K.; Zaretsky, S. Rational Design of Calpain Inhibitors Based on Calpastatin Peptidomimetics. *J. Med. Chem.* **2016**, *59* (11), 5403–5415.

46. Souers, A. J.; Leverson, J. D.; Boghaert, E. R.; Ackler, S. L.; Catron, N. D.; Chen, J.; Dayton, B. D.; Ding, H.; Enschede, S. H.; Fairbrother, W. J.; Huang, D. C. S.; Hymowitz, S. G.; Jin, S.; Khaw, S. L.; Kovar, P. J.; Lam, L. T.; Lee, J.; Maecker, H. L.; Marsh, K. C.; Mason, K. D.; Mitten, M. J.; Nimmer, P. M.; Oleksijew, A.; Park, C. H.; Park, C.-M.; Phillips, D. C.; Roberts, A. W.; Sampath, D.; Seymour, J. F.; Smith, M. L.; Sullivan, G. M.; Tahir, S. K.; Tse, C.; Wendt, M. D.; Xiao, Y.; Xue, J. C.; Zhang, H.; Humerickhouse, R. A.; Rosenberg, S. H.; Elmore, S. W. ABT-199, a Potent and Selective BCL-2 Inhibitor, Achieves Antitumor Activity While Sparing Platelets. *Nat. Med.* **2013**, *19* (2), 202–208.

47. Tse, C.; Shoemaker, A. R.; Adickes, J.; Anderson, M. G.; Chen, J.; Jin, S.; Johnson, E. F.; Marsh, K. C.; Mitten, M. J.; Nimmer, P.; Roberts, L.; Tahir, S. K.; Xiao, Y.; Yang, X.; Zhang, H.; Fesik, S.; Rosenberg, S. H.; Elmore, S. W. ABT-263: A Potent and Orally Bioavailable Bcl-2 Family Inhibitor. *Cancer Res.* **2008**, *68* (9), 3421–3428.

48. Karatas, H.; Townsend, E. C.; Cao, F.; Chen, Y.; Bernard, D.; Liu, L.; Lei, M.; Dou, Y.; Wang, S. High-Affinity, Small-Molecule Peptidomimetic Inhibitors of MLL1/WDR5 Protein–Protein Interaction. *J. Am. Chem. Soc.* **2013**, *135* (2), 669–

682.

49. Grebien, F.; Vedadi, M.; Getlik, M.; Giambruno, R.; Grover, A.; Avellino, R.; Skucha, A.; Vittori, S.; Kuznetsova, E.; Smil, D.; Barsyte-Lovejoy, D.; Li, F.; Poda, G.; Schapira, M.; Wu, H.; Dong, A.; Senisterra, G.; Stukalov, A.; Huber, K. V. M.; Schönegger, A.; Marcellus, R.; Bilban, M.; Bock, C.; Brown, P. J.; Zuber, J.; Bennett, K. L.; Al-Awar, R.; Delwel, R.; Nerlov, C.; Arrowsmith, C. H.; Superti-Furga, G. Pharmacological Targeting of the Wdr5-MLL Interaction in C/Ebpα N-Terminal Leukemia. *Nat. Chem. Biol.* **2015**, *11* (8), 571–578.

50. Mysinger, M. M.; Weiss, D. R.; Ziarek, J. J.; Gravel, S.; Doak, A. K.; Karpiak, J.; Heveker, N.; Shoichet, B. K.; Volkman, B. F. Structure-Based Ligand Discovery for the Protein–Protein Interface of Chemokine Receptor CXCR4. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109* (14), 5517–5522.

51. Chen, J.; Sawyer, N.; Regan, L. Protein–Protein Interactions: General Trends in the Relationship Between Binding Affinity and Interfacial Buried Surface Area. *Protein Sci.* **2013**, *22* (4), 510–515.

52. Munteanu, C. R.; Pimenta, A. C.; Fernandez-Lozano, C.; Melo, A.; Cordeiro, M. N. D. S.; Moreira, I. S. Solvent Accessible Surface Area-Based Hot-Spot Detection Methods for Protein-Protein and Protein-Nucleic Acid Interfaces. *J. Chem. Inf. Model.* **2015**, *55* (5), 1077–1086.

53 Li, Z.; Wong, L.; Li, J. DBAC: A Simple Prediction Method for Protein Binding Hot Spots Based on Burial Levels and Deeply Buried Atomic Contacts. *BMC Syst. Biol.* **2011**, 5 Suppl 1, S5.

54. Miyazaki, M.; Naito, H.; Sugimoto, Y.; Kawato, H.; Okayama, T.; Shimizu, H.; Miyazaki, M.; Kitagawa, M.; Seki, T.; Fukutake, S.; Aonuma, M.; Soga, T. Lead Optimization of Novel P53-MDM2 Interaction Inhibitors Possessing Dihydroimidazothiazole Scaffold. *Bioorg. Med. Chem. Lett.* **2013**, *23* (3), 728–732.

55. Miyazaki, M.; Naito, H.; Sugimoto, Y.; Yoshida, K.; Kawato, H.; Okayama, T.; Shimizu, H.; Miyazaki, M.; Kitagawa, M.; Seki, T.; Fukutake, S.; Shiose, Y.; Aonuma, M.; Soga, T. Synthesis and Evaluation of Novel Orally Active P53-MDM2 Interaction Inhibitors. *Bioorg. Med. Chem.* **2013**, *21* (14), 4319–4331.

56. Gonzalez, A. Z.; Eksterowicz, J.; Bartberger, M. D.; Beck, H. P.; Canon, J.; Chen, A.; Chow, D.; Duquette, J.; Fox, B. M.; Fu, J.; Huang, X.; Houze, J. B.; Jin, L.; Li, Y.; Li, Z.; Ling, Y.; Lo, M.-C.; Long, A. M.; Mcgee, L. R.; Mcintosh, J.; Mcminn, D. L.; Oliner, J. D.; Osgood, T.; Rew, Y.; Saiki, A. Y.; Shaffer, P.; Wortman, S.; Yakowec, P.; Yan, X.; Ye, Q.; Yu, D.; Zhao, X.; Zhou, J.; Olson, S. H.; Medina, J. C.; Sun, D. Selective and Potent Morpholinone Inhibitors of the MDM2-P53 Protein–Protein Interaction. *J. Med. Chem.* **2014**, *57* (6), 2472–2488.

57. Vaupel, A.; Bold, G.; De Pover, A.; Stachyra-Valat, T.; Lisztwan, J. H.; Kallen, J.; Masuya , K.; Furet, P. Tetra-Substituted Imidazoles as a New Class of Inhibitors of the P53-MDM2 Interaction. *Bioorg. Med. Chem. Lett.* **2014**, *24* (9), 2110–2114.

58. Gonzalez, A. Z.; Li, Z.; Beck, H. P.; Canon, J.; Chen, A.; Chow, D.; Duquette, J.; Eksterowicz, J.; Fox, B. M.; Fu, J.; Huang, X.; Houze, J.; Jin, L.; Li, Y.; Ling, Y.; Lo, M.-C.; Long, A. M.; Mcgee, L. R.; Mcintosh, J.; Oliner, J. D.; Osgood, T.; Rew, Y.; Saiki, A. Y.; Shaffer, P.; Wortman, S.; Yakowec, P.; Yan, X.; Ye, Q.; Yu, D.; Zhao, X.; Zhou, J.; Olson, S. H.; Sun, D.; Medina, J. C. Novel Inhibitors of the MDM2-P53 Interaction Featuring Hydrogen Bond Acceptors as Carboxylic Acid Isosteres. *J. Med. Chem.* **2014**, *57* (7), 2963–2988.

59. Holzer, P.; Masuya, K.; Furet, P.; Kallen, J.; Valat-Stachyra, T.; Ferretti, S.; Berghausen, J.; Bouisset-Leonard, M.; Buschmann, N.; Pissot-Soldermann, C.; Rynn, C.; Ruetz, S.; Stutz, S.; Chène, P.; Jeay, S.; Gessier, F. Discovery of a Dihydroisoquinolinone Derivative (NVP-CGM097): a Highly Potent and Selective MDM2 Inhibitor Undergoing Phase 1 Clinical Trials in P53wt Tumors. *J. Med. Chem.* **2015**, *58* (16), 6348–6358.

60. Zhang, Z.; Ding, Q.; Liu, J.-J.; Zhang, J.; Jiang, N.; Chu, X.-J.; Bartkovitz, D.; Luk, K.-C.; Janson, C.; Tovar, C.; Filipovic, Z. M.; Higgins, B.; Glenn, K.; Packman, K.; Vassilev, L. T.; Graves, B. Discovery of Potent and Selective Spiroindolinone MDM2 Inhibitor, RO8994, for Cancer Therapy. *Bioorg. Med. Chem.* **2014**, *22* (15), 4001–4009.

61. Gonzalez-Lopez De Turiso, F.; Sun, D.; Rew, Y.; Bartberger, M. D.; Beck, H. P.; Canon, J.; Chen, A.; Chow, D.; Correll, T. L.; Huang, X.; Julian, L. D.; Kayser, F.; Lo, M.-C.; Long, A. M.; Mcminn, D.; Oliner, J. D.; Osgood, T.; Powers, J. P.; Saiki, A. Y.; Schneider, S.; Shaffer, P.; Xiao, S.-H.; Yakowec, P.; Yan, X.; Ye, Q.; Yu, D.; Zhao, X.; Zhou, J.; Medina, J. C.; Olson, S. H. Rational Design and Binding Mode Duality of MDM2-P53 Inhibitors. *J. Med. Chem.* **2013**, *56* (10), 4053–4070.

62. Yu, M.; Wang, Y.; Zhu, J.; Bartberger, M. D.; Canon, J.; Chen, A.; Chow, D.; Eksterowicz, J.; Fox, B.; Fu, J.; Gribble, M.; Huang, X.; Li, Z.; Liu, J.; Lo, M.-C.; Mcminn, D.; Oliner, J. D.; Osgood, T.; Rew, Y.; Saiki, A. Y.; Shaffer, P.; Yan, X.; Ye, Q.; Yu, D.; Zhao, X.; Zhou, J.; Olson, S. H.; Medina, J. C.; Sun, D. Discovery of Potent and Simplified Piperidinone-Based Inhibitors of the MDM2-P53 Interaction. *ACS Med. Chem. Lett.* **2014**, *5* (8), 894–899.

63. Tovar, C.; Graves, B.; Packman, K.; Filipovic, Z.; Higgins, B.; Xia, M.; Tardell, C.; Garrido, R.; Lee, E.; Kolinsky, K.; To, K.-H.; Linn, M.; Podlaski, F.; Wovkulich, P.; Vu, B.; Vassilev, L. T. MDM2 Small-Molecule Antagonist RG7112 Activates P53 Signaling and Regresses Human Tumors in Preclinical Cancer Models. *Cancer Res.* **2013**, *73* (8), 2587–2597.

64. Wang, Y.; Zhu, J.; Liu, J.; Chen, X.; Mihalic, J.; Deignan, J.; Yu, M.; Sun, D.; Kayser, F.; Mcgee, L. R.; Lo, M.-C.; Chen, A.; Zhou, J.; Ye, Q.; Huang, X.; Long, A. M.; Yakowec, P.; Oliner, J. D.; Olson, S. H.; Medina, J. C. Optimization Beyond AMG 232: Discovery and SAR of Sulfonamides on a Piperidinone Scaffold as Potent Inhibitors of the MDM2-P53 Protein-Protein Interaction. *Bioorg. Med. Chem. Lett.* **2014**, *24* (16), 3782–3785.

65. *Schrödinger Release 2015-1: Maestro*, Version 10.1, Schrödinger, LLC, New York, NY, 2015.

66. *Schrödinger Release 2015-1: Schrödinger Suite 2015-1 Protein Preparation Wizard; Epik*, Version 3.1, Schrödinger, LLC, New York, NY, 2015; Impact Version 6.6, Schrödinger, LLC, New York, NY, 2015; Prime Version 3.9, Schrödinger, LLC, New York, NY, 2015.

67. Sastry, G. M.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. Protein and Ligand Preparation: Parameters, Protocols, and Influence on Virtual Screening Enrichments. *J. Comput.-Aided Mol. Des.* **2013**, *27* (3), 221–234.

68. Zaiontz C. (2015) Real Statistics Using Excel. http://www.real-statistics.com. (accessed October 31, 2016)

69. Bruncko, M.; Oost, T. K.; Belli, B. A.; Ding, H.; Joseph, M. K.; Kunzer, A.; Martineau, D.; Mcclellan, W. J.; Mitten, M.; Ng, S.-C.; Nimmer, P. M.; Oltersdorf, T.; Park, C.-M.; Petros, A. M.; Shoemaker, A. R.; Song, X.; Wang, X.; Wendt, M. D.; Zhang, H.; Fesik, S. W.; Rosenberg, S. H.; Elmore, S. W. Studies Leading to Potent, Dual Inhibitors of Bcl-2 and Bcl-Xl. *J. Med. Chem.* **2007**, *50* (4), 641–662.

70. *Schrödinger Release 2015-1: Ligprep*, Version 3.3, Schrödinger, LLC, New York, NY, 2015.

71. Oost, T. K.; Sun, C.; Armstrong, R. C.; Al-Assaad, A.-S.; Betz, S. F.; Deckwerth, T. L.; Ding, H.; Elmore, S. W.; Meadows, R. P.; Olejniczak, E. T.; Oleksijew, A.; Oltersdorf, T.; Rosenberg, S. H.; Shoemaker, A. R.; Tomaselli, K. J.; Zou, H.; Fesik, S. W. Discovery of Potent Antagonists of the Antiapoptotic Protein XIAP for the Treatment of Cancer. *J. Med. Chem.* **2004**, *47* (18), 4417–4426.

72. Senisterra, G.; Wu, H.; Allali-Hassani, A.; Wasney, G. A.; Barsyte-Lovejoy, D.; Dombrovski, L.; Dong, A.; Nguyen, K. T.; Smil, D.; Bolshan, Y.; Hajian, T.; He, H.; Seitova, A.; Chau, I.; Li, F.; Poda, G.; Couture, J.-F.; Brown, P. J.; Al-Awar, R.; Schapira, M.; Arrowsmith, C. H.; Vedadi, M. Small-Molecule Inhibition of MLL Activity by Disruption of its Interaction with WDR5. *Biochem. J.* **2013**, *449* (1), 151–159.

73. Jacobson, M. P.; Pincus, D. L.; Rapp, C. S.; Day, T. J. F.; Honig, B.; Shaw, D. E.;

Friesner, R. A. A Hierarchical Approach to All-Atom Protein Loop Prediction. *Proteins* **2004**, *55* (2), 351–367.

74. Jacobson, M. P.; Friesner, R. A.; Xiang, Z.; Honig, B. On the Role of the Crystal Environment in Determining Protein Side-Chain Conformations. *J. Mol. Biol.* **2002**, *320* (3), 597–608.

75. *Schrödinger Release 2015-1: Prime*, Version 3.9, Schrödinger, LLC, New York, NY, 2015.

76. Trott, O.; Olson, A. J. Autodock Vina: Improving The Speed And Accuracy Of Docking With A New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **2010**, *31* (2), 455–461.

77. Carson, D. A.; Lois, A. Cancer Progression and P53. Lancet. 1995, 346 (8981), 1009–1011.

78. Estrada-Ortiz, N.; Neochoritis, C. G.; Dömling, A. How to Design a Successful P53-MDM2/X Interaction Inhibitor: a Thorough Overview Based on Crystal Structures. *Chemmedchem* **2016**, *11* (8), 757–772.

79. Boström, J.; Hogner, A.; Schmitt S. Do Structurally Similar Ligands Bind in a Similar Fashion? *J. Med. Chem.* **2006**, *49* (23), 6716–6725.

79. Reddy, T. R. K.; Li, C.; Guo, X.; Myrvang, H. K.; Fischer, P. M.; Dekker, L. V. Design, Synthesis, and Structure-Activity Relationship Exploration of 1-Substituted 4-Aroyl-3-Hydroxy-5-Phenyl-1H-Pyrrol-2(5H)-One Analogues as Inhibitors of the Annexin A2-S100A10 Protein Interaction. *J. Med. Chem.* **2011**, *54* (7), 2080–2094.

80. Viaud, J.; Zeghouf, M.; Barelli, H.; Zeeh, J.-C.; Padilla, A.; Guibert, B.; Chardin, P.; Royer, C. A.; Cherfils, J.; Chavanieu, A. Structure-Based Discovery of an Inhibitor of Arf Activation by Sec7 Domains Through Targeting of Protein-Protein Complexes. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104* (25), 10370–10375.

81. Enyedy, I. J.; Ling, Y.; Nacro, K.; Tomita, Y.; Wu, X.; Cao, Y.; Guo, R.; Li, B.; Zhu, X.; Huang, Y.; Long, Y.-Q.; Roller, P. P.; Yang, D.; Wang, S. Discovery of Small-Molecule Inhibitors of Bcl-2 Through Structure-Based Computer Screening. *J. Med. Chem.* **2001**, *44* (25), 4313–4324.

82. Cerchietti, L. C.; Ghetu, A. F.; Zhu, X.; Da Silva, G. F.; Zhong, S.; Matthews, M.; Bunting, K. L.; Polo, J. M.; Farès, C.; Arrowsmith, C. H.; Yang, S. N.; Garcia, M.; Coop, A.; Mackerell, A. D. Jr.; Privé, G. G.; Melnick, A. A Small-Molecule Inhibitor of BCL6 Kills DLBCL Cells In Vitro and In Vivo. *Cancer Cell* **2010**, *17* (4), 400–411.

83. Vidler, L. R.; Filippakopoulos, P.; Fedorov, O.; Picaud, S.; Martin, S.; Tomsett, M.; Woodward, H.; Brown, N.; Knapp, S.; Hoelder, S. Discovery of Novel Small-Molecule Inhibitors of BRD4 Using Structure-Based Virtual Screening. *J. Med. Chem.* **2013**, *56* (20), 8073–8088.

84. Lavecchia, A.; Di Giovanni, C.; Cerchia, C.; Russo, A.; Russo, G.; Novellino, E. Discovery of a Novel Small Molecule Inhibitor Targeting the Frataxin/Ubiquitin Interaction via Structure-Based Virtual Screening and Bioassays. *J. Med. Chem.* **2013**, *56* (7), 2861–2873.

85. Zhuang, C.; Narayanapillai, S.; Zhang, W.; Sham, Y. Y.; Xing, C. Rapid Identification of Keap1-Nrf2 Small-Molecule Inhibitors Through Structure-Based Virtual Screening and Hit-Based Substructure Search. *J. Med. Chem.* **2014**, *57* (3), 1121–1126.

86. Markowitz, J.; Chen, I.; Gitti, R.; Baldisseri, D. M.; Pan, Y.; Udan, R.; Carrier, F.; Mackerell, A. D. Jr.; Weber, D. J. Identification and Characterization of Small Molecule Inhibitors of the Calcium-Dependent S100B-P53 Tumor Suppressor Interaction. *J. Med. Chem.* **2004**, *47* (21), 5085–5093.

87. Khanna, M.; Wang, F.; Jo, I.; Knabe, W. E.; Wilson, S. M.; Li, L.; Bum-Erdene, K.; Li, J.; Sledge, G. W. Jr.; Khanna, R.; Meroueh, S. O. Targeting Multiple Conformations Leads to Small Molecule Inhibitors of the Upar·Upa Protein–Protein Interaction that Block Cancer Cell Invasion. *ACS Chem. Biol*. **2011**, *6* (11), 1232–1243.

88. Liu, D.; Zhou, D.; Wang, B.; Knabe, W. E.; Meroueh, S. O. A New Class of Orthosteric Upar·Upa Small-Molecule Antagonists are Allosteric Inhibitors of the Upar·Vitronectin interaction. *ACS Chem. Biol.* **2015**, *10* (6), 1521–1534.

CHAPTER 4

EVOLUTIONARY TRACE AND SEQUENCE ANALYSIS REVEAL INSIGHTS INTO CATENIN-LIKE PROTEINS, IMPLICATIONS FOR DRUG DESIGN, SELECTIVITY, AND POTENTIAL FORALLOSTERIC REGULATION

4.1 Introduction

The PPI between the ARM domain protein *β*-catenin and Tcf4 is a promising, yet elusive, drug target.[1] This PPI is the penultimate step in the Wnt signaling pathway where it mediates the formation of a nuclear transcription complex.[2] The canonical Wnt signaling pathway is responsible for the expression of Wnt target genes that are important for stem cell differentiation and survival.[3] Mutations in proteins associated with canonical Wnt signaling[4,5] result in the dysregulation of the Wnt signaling pathway, accumulation of *β*-catenin, and the untimely overexpression of Wnt target genes. Mutations in APC, a protein responsible for sequestering excess cytosolic *β*-catenin for degradation, are present in most cases of sporadic colorectal cancers.[6] Other mutations in Wnt-associated proteins have been identified in cancerous skin stem cells,[7,8] oral cancer,[9,10] and numerous others.[5,10] Inhibiting the formation of the *β*-catenin/Tcf4 downstream nuclear complex represents a way to disrupt the aberrant Wnt target gene expression and halt cancer stem cell renewal.[1] Unfortunately, *β*-catenin is a promiscuous protein with over

one hundred direct binding partners.[11,12,13] Several proteins, E-cadherin,[14] APC,[15] axin[16] and Tcf[17], are known to bind to a long, positively-charged groove that runs the length of the ARM domain of β-catenin. β-catenin functions as a structural protein by binding to E-cadherin at adherens junctions;[18] axin and APC mediate cytosolic β-catenin levels by sequestering excess β-catenin for degradation in the destruction complex.[19,20] Crystal structures for β-catenin/Tcf (PDB IDs 1JPW,[21] 1JDH,[22] 2GL7),[17] β-catenin/E-cadherin (PDB IDs 1I7X,[14] 1I7W)[14] and β-catenin/APC (PDB IDs 1TH1,[15] 1JPP,[23] 1T08)[24] reveal three shared hot regions on the ARM domain of β-catenin.[25,26] These hot regions are surface areas where residues contribute significantly greater to binding energy than the surrounding residues. Hot region 1, with K435, forms the binding site of the $Dx\theta\theta x\Phi x_{2\text{-}7}E$ motif of E-cadherin, APC and Tcf4 (Figure 4.1).[21,22,25,27]This hot region is the most important of the three hot regions for binding Tcf4,[22] even though E-cadherin and APC both share a similar binding mode. Hot region 2 contains K312, K345, and Y306. Hot region 3 contains residue R212 and is important for binding axin, in addition to Tcf4, E-cadherin, and APC.[26] In addition to the three hot regions, β-catenin contains a separate binding site near the N-terminal region of the ARM domain for Bcl9, another component of the nuclear transcription complex.[17]

E-cadherin and APC also utilize phosphorylation to enhance their binding to β-catenin.[14,15] Residues T1487, S1504, S1505, S1507, and S1510 of APC are capable of being phosphorylated, as are residues S684, S686, S690, and S692 in E-cadherin. These phosphorylated residues interact with K292, K335, and R376 of β-catenin.[26] Tcf4 does not undergo phosphorylation; therefore, it relies less on these residues. While β-catenin fulfills multiple roles, similar proteins in the same family have developed more specific

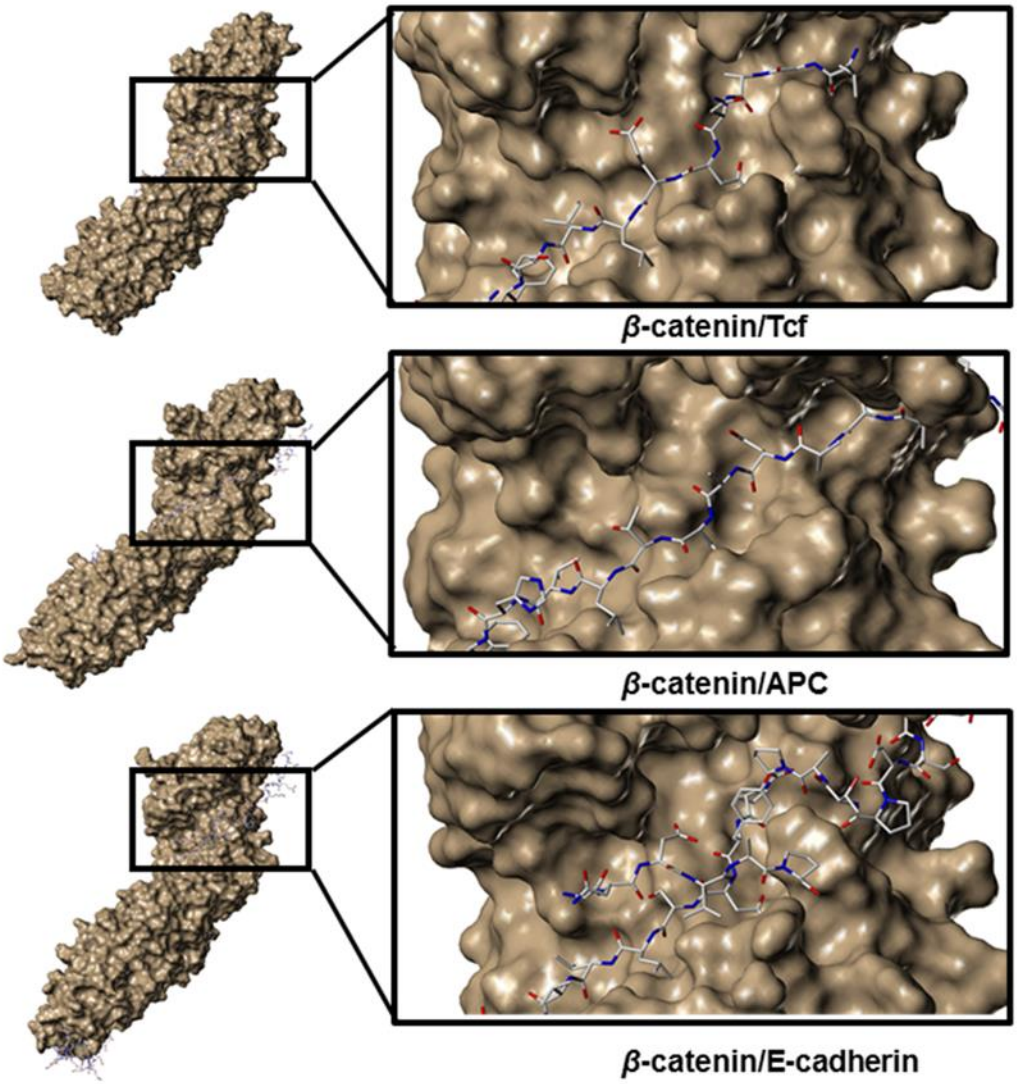


Figure 4.1. Binding mode of Tcf4, APC, and E-cadherin in *β*-catenin hot region 1. Image created from 2GL7,[17] 1TH1,[15] and 1I7X.[14]

function and binding.[28,29] *β*-catenin is one member of the large family of ARM domain proteins. The armadillo repeat motif that defines this family consists of two α-helices forming a hairpin-like structure[30,31] that is repeated to form the ARM domain. These proteins compose a functionally diverse, yet structurally similar, family of proteins responsible for cellular signaling, adhesion, and structural networks.[31,32] The ARM family of proteins includes such notable members as *β*-catenin, plakoglobin (JUP, *γ*-catenin), APC, and importin.[33] A subset of these ARM proteins has functional overlap with *β*-catenin. These catenin-like proteins include JUP in humans, and symmetrical sister cell hermaphrodite gonad defect 1 (SYS-1), humpback 2 (HMP-2), armadillo repeat domain containing WRM-1 and *β*-catenin/armadillo-related protein 1 (BAR-1) in *C. elegans.*[28,29] Sequence similarity between catenin-like proteins ranges from very high, between plakoglobin and *β*-catenin,[34] or very low, between SYS-1 and *β*-catenin.[30] Consequently, these proteins are identified by armadillo repeat structural motif and function rather than primary sequence identity.[29,35]

The most closely related ARM family member to *β*-catenin is JUP, which shares 64% sequence identity overall and 79% sequence identity in the ARM domain.[34] All the important lysine residues are conserved, as are the residues that interact with the $Dx\theta\theta x\Phi x_{2\text{-}7}E$ motif of Tcf4, APC and E-cadherin,[36] and the residues responsible for binding the pS/pT-motif of APC and E-cadherin.[29] Despite this similarity, JUP has a more specialized function. JUP is a key structural protein of the desmosome, linking desmosomal cadherin proteins to desmoplakin.[34] Unlike *β*-catenin, which is found in both the cytosol and nucleus, JUP is predominantly localized to the desmosome and has a higher affinity for desmosomal cadherins than the E-cadherin found at adherens

junctions.[36] Although JUP binds to Tcf4 in biochemical assays, the interaction with DNA is considerably weaker than that of the *β*-catenin/Tcf4 complex.[36]

The four other types of catenin-like proteins, SYS-1, HMP-2, WRM-1, and BAR-1, are found almost exclusively in the genus *Caenorhabditis*, but not in other organisms.[28,29,37] These proteins are specialized in function unlike the promiscuous *β*-catenin even though they share similar structures and key residues. These specialized divergent catenins have only minimal sequence identity with *β*-catenin (full sequence/ARM domain): BAR-1 (22%/24%), HMP-2 (27%/31%), WRM-1 (16%/18%), and SYS-1 (12%/14%).

BAR-1 is the closest homologue in function, if not in sequence, to human *β*-catenin found in *C. elegans*.[28] BAR-1 fills *β*-catenin's role in canonical Wnt signaling. Homology modeling of BAR-1 indicates that it binds the *C. elegans* Tcf4 homologue POP-1 in a similar manner as *β*-catenin binds Tcf4.[28] BAR-1 has less conservation in hot regions 2 and 3 and does not interact with HMR-1, the *C. elegans* homologue of E-cadherin. HMP-2, by contrast, does not interact with POP-1. Instead, HMP-2 interacts with hammerhead cadherin-like protein HMR-1 in a manner similar to that of *β*-catenin and E-cadherin.[28,29] HMP-2 contains the key residues responsible for binding POP-1, but a nearby mutation may compromise the function of the K365 (*β*-catenin K435) hot spot.[28,29] WRM-1 interacts with neither POP-1 nor HMR-1. WRM-1 binds to Serine/threonine nemo-like kinase LIT-1 to inhibit POP-1 transcription.[28,29] The hot spot K497 (human *β*-catenin K435) is conserved, but the bulky side-chain of WRM-1 L491 blocks POP-1 E9 (human Tcf E17) access to this residue. SYS-1 is the most recently identified catenin-like protein.[29] The sequence similarity between *β*-catenin and SYS-1 is

considered insignificant, and it was not until a crystal structure of SYS-1[29] was produced that it was verified as a member of the catenin-like protein family. SYS-1 interacts with the *Caenorhabditis* Tcf4 homologue POP-1 to regulate gene transcription specific to producing asymmetrical daughter cells in *C. elegans*, but is not involved in canonical Wnt signaling.[29,35] The salt bridge between K539 of SYS-1 and D8 of POP-1 (K435 and D15 in human *β*-catenin and Tcf4) is conserved, but the other salt bridges found between *β*-catenin and Tcf4 are not present in the SYS-1/POP-1 crystal structure (PDB ID 3C2G).[30] Regardless, SYS-1 was able to rescue Wnt signaling in BAR-1 null mutants when attached to a *BAR-1* promotor region.[29,35]

While much study has been conducted on *β*-catenin and its binding partners, much remains to be understood about the differences between catenin-like proteins, binding-partner selectivity, and druggability. This is important because one of the biggest hurdles for disrupting the *β*-catenin/Tcf4 PPI is the risk of disrupting the *β*-catenin/E-cadherin and *β*-catenin/APC PPIs.[1] Clues about designing selective small-molecule inhibitors can be garnered from examining the differences between the different *C. elegans* catenin-like proteins. One lysine hot spot in hot region 1 (K435 in human *β*-catenin) is universally conserved between human *β*-catenin and JUP and the *C. elegans* catein-like proteins.[29] It is also the most important residue for interaction with Tcf4 and its *C. elegans* homologue POP-1, yet not all catenin-like proteins possess the ability to bind Tcf4/POP-1.[27-29] There is evidence that E-cadherin and APC, while binding similarly to K435 region of *β*-catenin, rely more on K292, K335 and R376.[27] These three residues are conserved in HMP-2 (K220, K264 and R306, respectively) whichs binds the *C. elegans* cadherin homologue HMR-1, but are not conserved in SYS-1 nor WRM-1

which do not bind HMR-1.[28,29] Furthermore, residues K292, K335 and R376, which are responsible for binding the pS/pT-motif of E-cadherin and APC,[27] are conserved in HMP-2. By exploiting these differences, the Ji lab succeeded in creating potent, selective inhibitors of the *β*-catenin/Tcf4 interaction.[38,39] However, much about the selectivity and differences in catenin-like proteins remains to be explored.

In this study, ET analysis is used to predict and rank functionally important residues. It has also been used to identify previously unknown allosteric sites. ET was used to identify a novel allosteric site in the D2 dopamine GPCR.[40] It is hypothesized that the inclusion of JUP and the four specialized *C. elegans* catenin-like proteins in ET analysis will yield important information as to the varying importance of the hot regions. ET may also be useful in discovering latent allosteric sites which can be used to selectively inhibit the *β*-catenin/Tcf4 interaction.

ET can help identify functionally important residues and potentially allosteric regions, but lacks data about the druggability of said residues. In order to assess the druggability of the differing hot regions, FTMap, a solvent mapping program developed by Kozakov et al.[41-44] is used. Because *β*-catenin has multiple binding partners binding to the same hot regions, solvent mapping different crystal structures may give insight into surface adaptability and the origin of selectivity. *β*-catenin was reported to be a flexible protein,[45] so an MD simulation was utilized to further explore the druggability of *β*-catenin. Combining FTMap with the MD simulations provided a way to analyze druggability changes in the three hot regions throughout the simulation. This unique combination of computational approaches has yielded new insights into the similarities and differences of catenin-like proteins and lead to the discovery of a new potential

allosteric site for regulating the β-catenin/Tcf4 PPI.

4.2 Methods

4.2.1 FTMap analysis. FTMap solvent mapping analysis was conducted on crystal structures 1TH1,[15] 1QZ7,[16] 1I7X,[17] and 2GL7[18] with ligands removed and MD trajectory snapshots of cluster centroids. Calculations were carried out using the FTMap web server with parameters specific for PPIs.[43] To facilitate comparison of probe interactions on a per residue level, the total number of probe interactions was normalized to a percentage. Because other nonbonding (van der Waals) interactions always outnumbered hydrogen bond interactions, they were also normalized in a similar manner. Results were analyzed and visualized using Sybyl,[46] R software[47] (*GGPlot2*,[48] *Reshape2*)[49] and UCSF Chimera.[50]

4.2.2 Evolutionary trace analysis. Primary sequences of β-catenin, plakoglobin (including γ-catenin), WRM-1, SYS-1, HMP-2, and BAR-1 were obtained from UniProt.[51] Accessions with greater than 90% sequence identity to human β-catenin (UniProt accession P35222) were discarded. Forty-nine sequences were included for the multiple sequence alignment (MSA) and subsequent ET analysis. The sequences were aligned using the ClustalOmega web server.[52,53] The sequences were trimmed so that only residues aligning to the ARM domain of human β-catenin were included in the ET analysis. Real-value ET (rvET) scores were calculated by the web server provided by the Lichtarge computational biology lab at Baylor College of Medicine.[54] rvET output data was analyzed with R software[48] and Schrödinger Maestro.[55] Unless otherwise noted, the residues discussed are all in reference to human β-catenin (PDB ID 2GL7)[24].

4.2.3 Molecular dynamics simulations. A *β*-catenin construct consisting of the main ARM domain of human *β*-catenin (PDB ID 2GL7,[17] residues 147-549 and 560-663), and the missing loop and C-terminal helix 12 (PDB ID 2Z6H,[56] residues 550-559 and 663-686) was created with Sybyl Biopolymer.[46] The construct was minimized by a 5000-step method in Schrödinger Macromodel.[57] The topologies for the simulation were generated with Gromacs[58] using the AMBER03 force field.[59] Simulations were conducted in a dodecahedron solvated with TIP3P water molecules, and ions were added to neutralize the system. The system was energy minimized by a 5000-step steepest descent method until the maximum force in the system (F_{max}) is less than 1000 kJ mol^{-1} nm^{-1}. Next, the system was subjected to NVT ensemble (constant number of particles, volume, and temperature) and NPT ensemble (constant number of particles, pressure, and temperature) equilibrations of 100 ps each while all nonsolvent heavy atoms were restrained. MD simulations were conducted using Gromacs for 20 ns at 310K. The trajectory was clustered by RMSD using the GROMOS[60] algorithm. Four groups were used for RMSD calculations, all protein heavy atoms, and hot regions 1-3. The following residues were used as definitions of the hot regions: hot region 1 contains residues G422, S425, N426, L428, C429, N430, K435, E462, P463, I465, C466, R469, H470, S473, R474, Q482, L506, I507, K508, A509, V511, G512, R515, N516, L519, L536, L539, R565, E568, I569, E571, G572, C573, G575, A576, H578, I579, R582, N609, R612, V613, E620, Y654; hot region 2 contains residues Q302, I303, A305, Y306, G307, G307, N308, Q309, S311, K312, I315, T332, Y333, E334, K335, L336, L337, W338, T339, S341, R342, V343, K345, V346, L347, S348, V349, C350, S374, Q375, R376, L377, V378, N380, T384; hot region 3 contains residues V173, V175, H176, Q177, S179,

A211, R212, C213, A215, G216, T217, H219, N220, S222, H223, H224, K242, D249, S250, F253, Y254, A256, T257, T258, L259, H260, N261, L264, H265, T289, N290, V291, K292, FA293, L294, A295, I296, T297, D299.

4.2.4 Fluorescence polarization binding assay of β-catenin mutants. Three *β*-catenin point mutants, W504A, W504L, and W504I, were created using overlapping PCR. Proteins were expressed using transformed BL21 *E. coli* cells and purified by Ni-NTA column. The proteins were dialyzed overnight at pH 8.8. Protein purity was verified using SDS-PAGE. Protein concentration was measured through a colorimetric Bradford assay. The effect of the point mutations on *β*-catenin/Tcf binding was determined through fluorescence polarization using a Tcf4 fluorescence tracer (excitation wavelength = 485 nm, emission wavelength = 535 nm). Wild-type *β*-catenin and Tcf4 tracer was used as a negative control, and Tcf4 tracer alone provided a positive control. K_d values were calculated using GraphPad Prism 5 software using a nonlinear least squares analysis.

4.3 Results

4.3.1 FTMap of crystal structures. Solvent mapping of the *β*-catenin/Tcf4/Bcl9 crystal structure (PDB ID 2GL7)[17] places five of the nine total probe cross-clusters in hot region 1, and one cross-cluster in hot region 2. When probe interactions were examined on a per residue basis, no single residue in hot regions 2 or 3 had more than 2% of the total solvent probe interactions in a single category (hydrogen bonds or other nonbonding interactions). In contrast, hot region 1 contained residues that had upwards of 10% of the total solvent interactions in either category. *β*-catenin structures with E-cadherin bound (1I7X)[14], APC bound (1TH1)[15] and axin bound (1QZ7)[16] were also evaluated for surface

changes affecting druggability (Figure 4.2). Hot region 1 maintained dominance in all the tested crystal structures, but increases in the number of probes at hot regions 2 and 3 were observed. The crystal structure with E-cadherin bound showed an increase of 10% in hot region 2 and 3% in hot region 3, while a decrease of 5% in hot region 1 compared to 2GL7. The crystal structure of axin bound *β*-catenin showed an increase of 4% and 9%, and a decrease of 13%, respectively. The phosphorylated APC bound crystal structure showed an increase of 2% and 18%, and a decrease of 22%, respectively. This is evidence of surface flexibility and surface adaptability of hot regions 2 and 3 in the presence of different binding partners. In order to further explore the importance of the *β*-catenin hot regions, ET analysis was performed on *β*-catenin and catenin-like proteins.

4.3.2 Multiple sequence alignment. The multiple sequence alignment of 49 *β*-catenin, JUP, SYS-1, HMP-2, WRM-1, and BAR-1 amino acid sequences from UniProt[51] indicate no residues are entirely conserved between species. Most hot regions have a high degree of conservation, but differences occur between different proteins and different species. This was surprising because even K435 was not entirely conserved in these sequences as previously thought. These differences in sequences between catenin-like proteins likely correspond to differences in function as observed in the *Caenorhabditis β*-catenin homologues.

4.3.3 ET analysis. ET analysis reports results as real-value ET scores (rvET); a lower score is indicative of a potentially more functionally important residue[54]. A full list of rvET scores may be found in the appendix of this chapter. Key residues in all three hot regions along the groove of *β*-catenin have low rvET scores. This indicates that ET has excellent power at predicting functional residues from our MSA. Evolutionary Trace

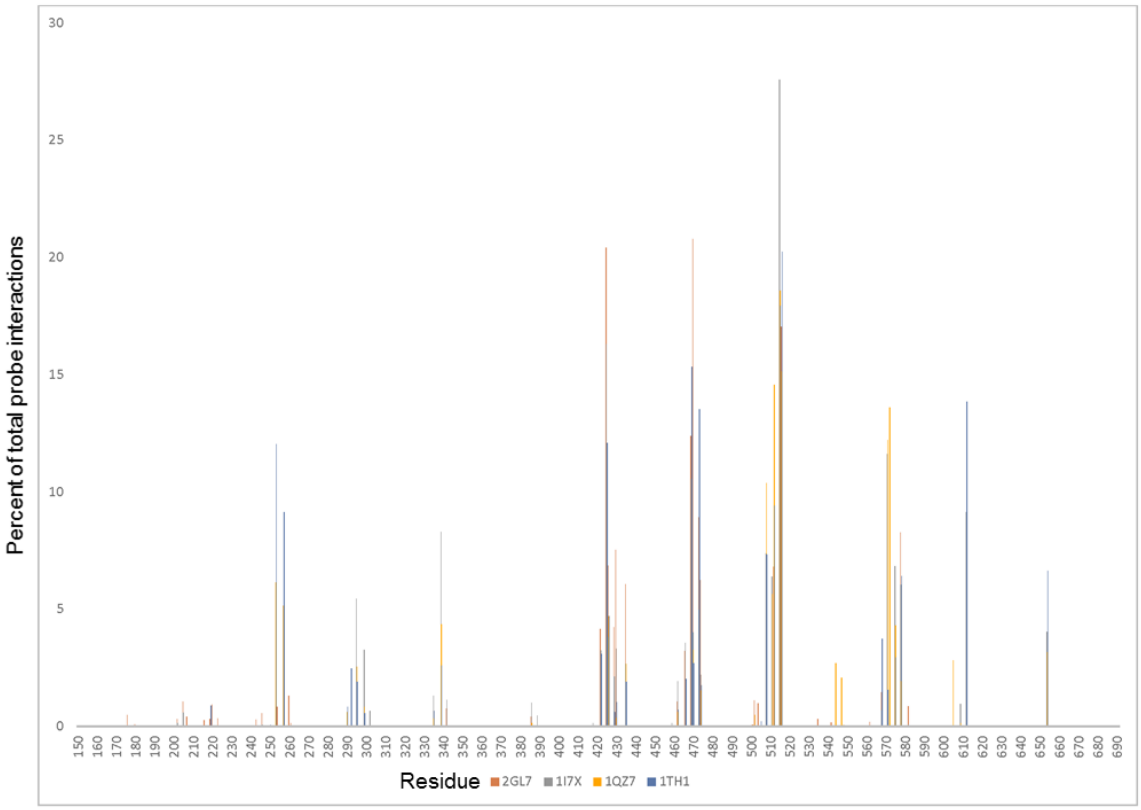
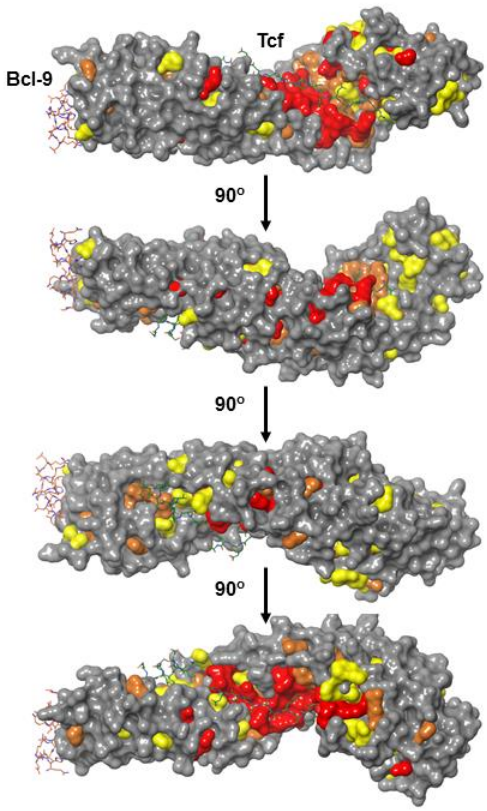


Figure 4.2. Probe interactions from solvent mapping the different crystal structures of *β*-catenin. Interactions in crystal structure 2GL7 are almost entirely located in hot region 1. The other hot regions have more probe interactions when other binding partners were present in the crystal structure. Interestingly, hot region 1 maintains a large amount of interactions even in 1QZ7, in which nothing was bound to hot region 1. This may indicate that while hot regions 2 and 3 are only accessible in the presence of a binding partner, hot region 1 is accessible when nothing is bound.

analysis confirms the importance of known hot-spot residues (Figure 4.3). Of the key lysines in the main groove of *β*-catenin, K312, 345, K435, and K508, K435 is ranked as the most evolutionarily important (rvET = 2.64), and K312 was the least important (rvET = 7.43) of the three key lysine residues. This is reflection of the conservation between the different variants of catenin-like proteins. Of the three known hot regions, hot region 1 showed the most evolutionary importance.

Few residues outside of the long positively-charged groove of the ARM domain had low rvET scores. Surprisingly, this extended to the N-terminal Bcl9 binding region. Reasons for this may be that it is not conserved in the worm *β*-catenin variants or that Bcl9 is dispensable in Wnt signaling.[61] One exception, residues which have very low ET scores but are not part of the surface of *β*-catenin, were the leucine residues scattered throughout the interior of the protein. This is expected due to the similarity of the armadillo repeat to the leucine-rich repeat motif in other alpha-solenoid structures.[62,63] Another exception was a region centered on W504 of human *β*-catenin (Figure 4.4). W504 itself had a low rvET score of 4.80, which was lower than K312 and K508. Surrounding it are residues H503 (7.01), P505 (4.29), R542 (5.80), and E562 (5.06). Disappointingly, FTMap did not indicate that this would be a druggable site in any of the tested crystal structures. An attempt was made to find alternative conformations of this pocket, which may be druggable, by MD simulation.

PPI surfaces are well known to be flexible, and *β*-catenin was reported to be a flexible protein in previous MD studies.[45] The *β*-catenin construct from the crystal structures 2GL7 and 26ZH was created so that the most complete structure of the ARM domain could be utilized in simulation. A 20 ns MD simulation was performed using the



Residue	rvET	Residue	rvET
R212	8.07	C429	3.83
K270	9.02	N430	3.45
R274	4.65	N434	1.32
K292	7.98	K435	2.64
I296	6.49	R457	9.94
K335	7.51	P463	5.61
W338	4.72	C466	5.83
R342	4.24	A467	9.59
K345	4.36	R469	7.15
Y354	3.85	H470	5.45
R376	6.13	S473	9.53
W383	5.83	R474	4.91
R386	3.42	K508	5.49
N387	4.24	R515	5.71
D390	4.61	Y654*	6.78
K394	8.83	S675*	9.58
N426	1.57	W504	4.8

Figure 4.3. The ET results are visualized on the β-catenin crystal structure 2GL7[17] (left). Residues with the lowest rvET scores are colored red. Those higher are orange and yellow. The W504 region can be seen on the second structure from the top.

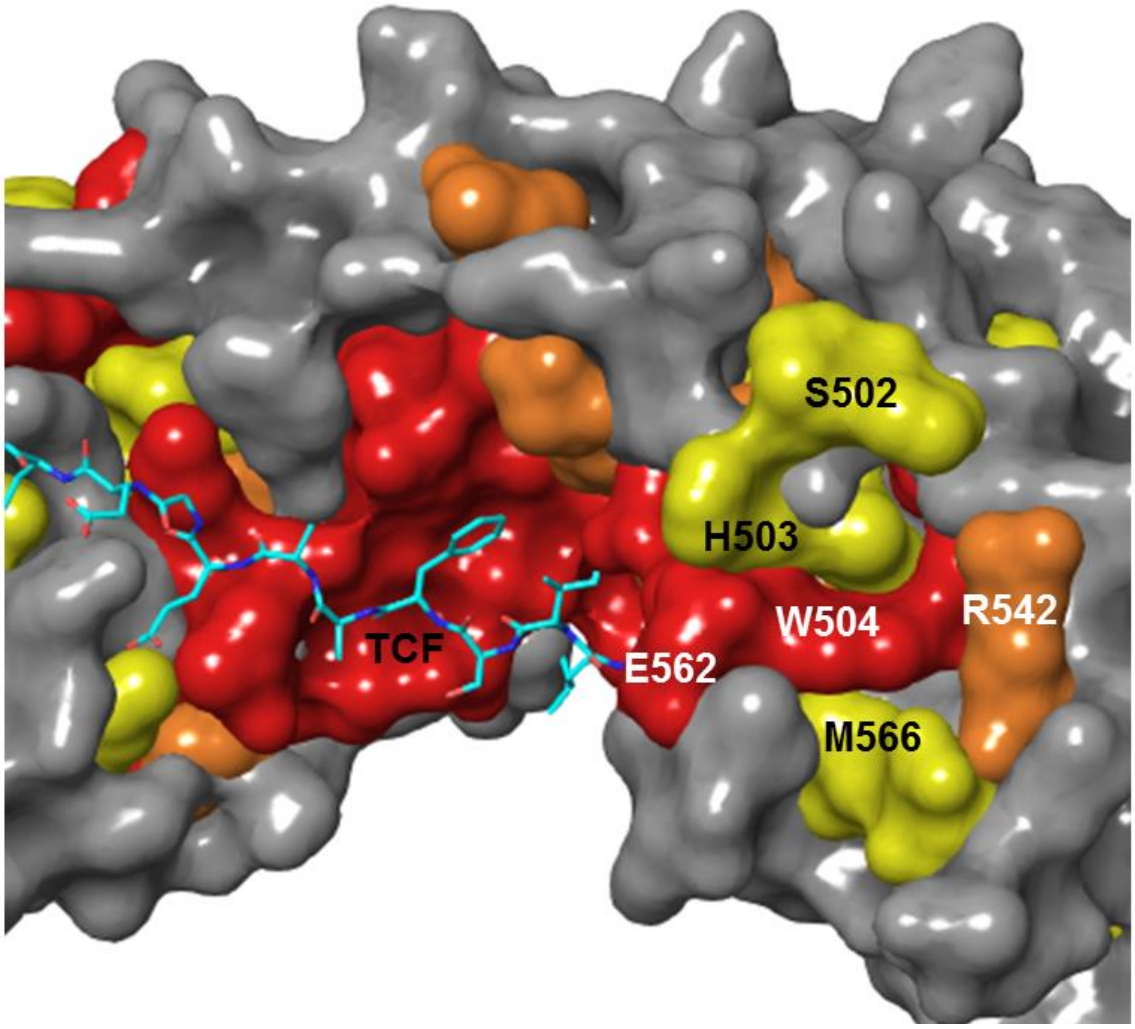


Figure 4.4. The region surrounding W504 has residues which are potentially functionally important. Tcf4 (teal) does not interact with this region. The number of residues in this area that could be functionally important leads us to the conclusion that this site may be a basis for potential allostery to the nearby hot region 1.

ARM domain of β-catenin. RMSD clustering of the MD simulations on protein heavy atoms with a 2Å cutoff found 53 clusters. The top 20 centroids of these clusters, which represent 90% of the simulation snapshots, and the initial 2GL7/2Z6H construct were analyzed with FTMap. The number of probe interactions was normalized between interaction types, because nonbonding interactions always outnumbered hydrogen bonding interactions. Then it was normalized between different protein structures. The changes in probe interactions per residue compared to the starting structure (Δprobes) were examined to detect changes in druggability throughout the simulation (Figure 4.5). Changes in Δprobes were classified in three categories: those which occurred within the main binding groove of β-catenin, those which occurred outside of the main binding groove of β-catenin with high rvET scores (rvET > 7), and those which occurred outside of known binding surfaces that had low rvET scores (rvET < 7). The analysis of the MD simulation showed several areas of the structure with increased druggability in the simulation as compared to the crystal structure in the first category. These were certain areas of the hot region 1 (Figure 4.6) which showed more probe interaction in simulation structures than the crystal structures. These subtle changes in side-chain orientation revealed alternative conformations of this hot region that could be used for drug discovery. Similar results were seen in hot region 2, hot region 3, and the W504 pocket (Figures 4.7-4.9).

In the third category was the region near W504, a region previously identified as potentially functionally important in ET analysis. This residue showed no probe interactions in the initial crystal structure FTMap analysis. However, FTMap cross-clusters were found in in this pocket in multiple snapshots from the MD simulation.

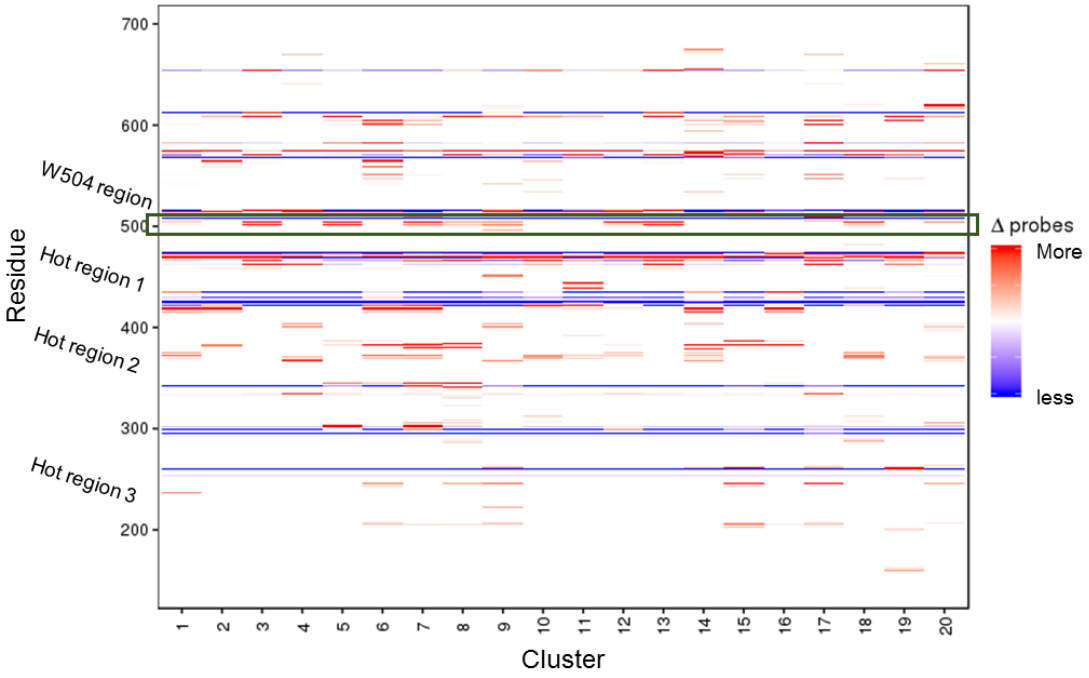


Figure 4.5. Comparison of probe changes from clustering all β-catenin heavy atoms. Red regions represent an increase in probe interactions; blue regions represent a decrease in probe interactions. The green box represents probe interactions that are occurring in the new potentially allosteric region of W504.

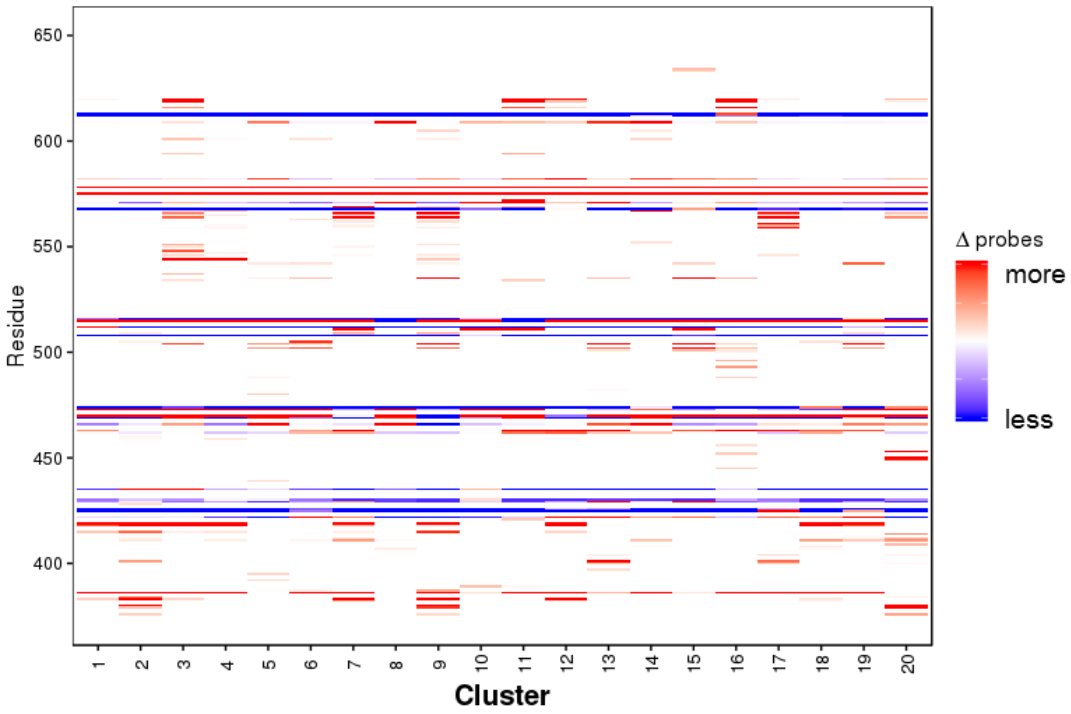


Figure 4.6. Comparison of probe changes from clustering hot region 1 heavy atoms.

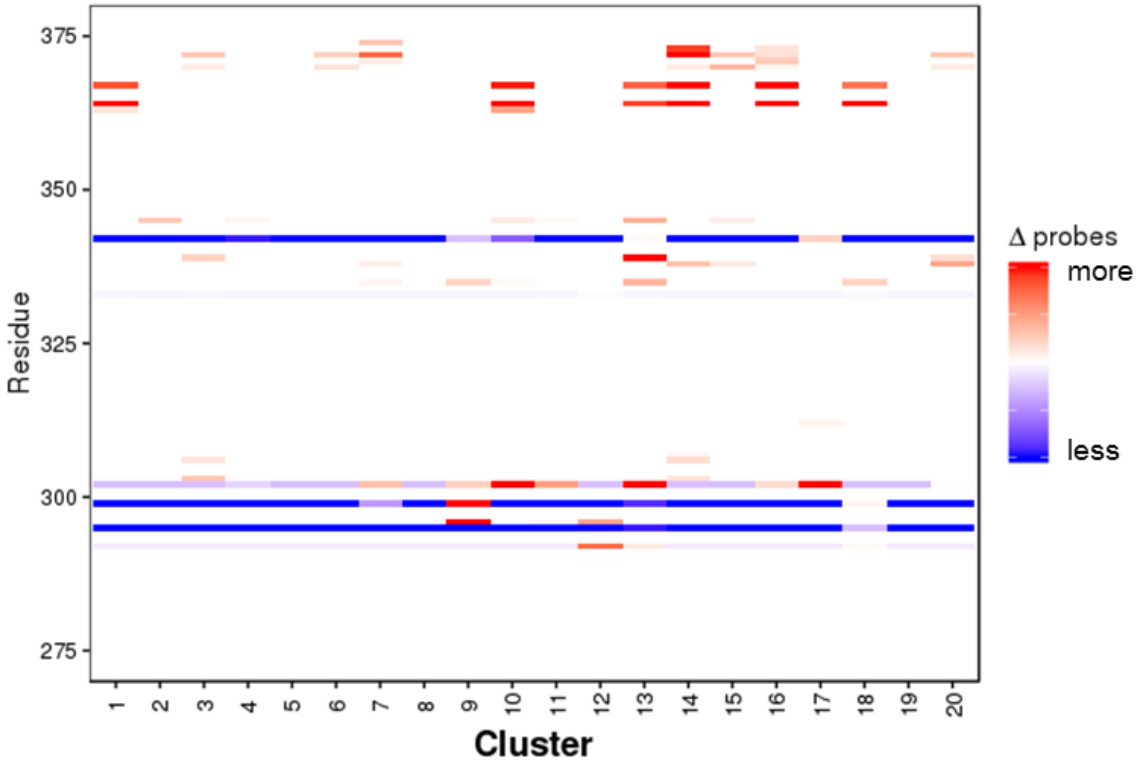


Figure 4.7 Comparison of probe changes from clustering all hot region 2 heavy atoms.

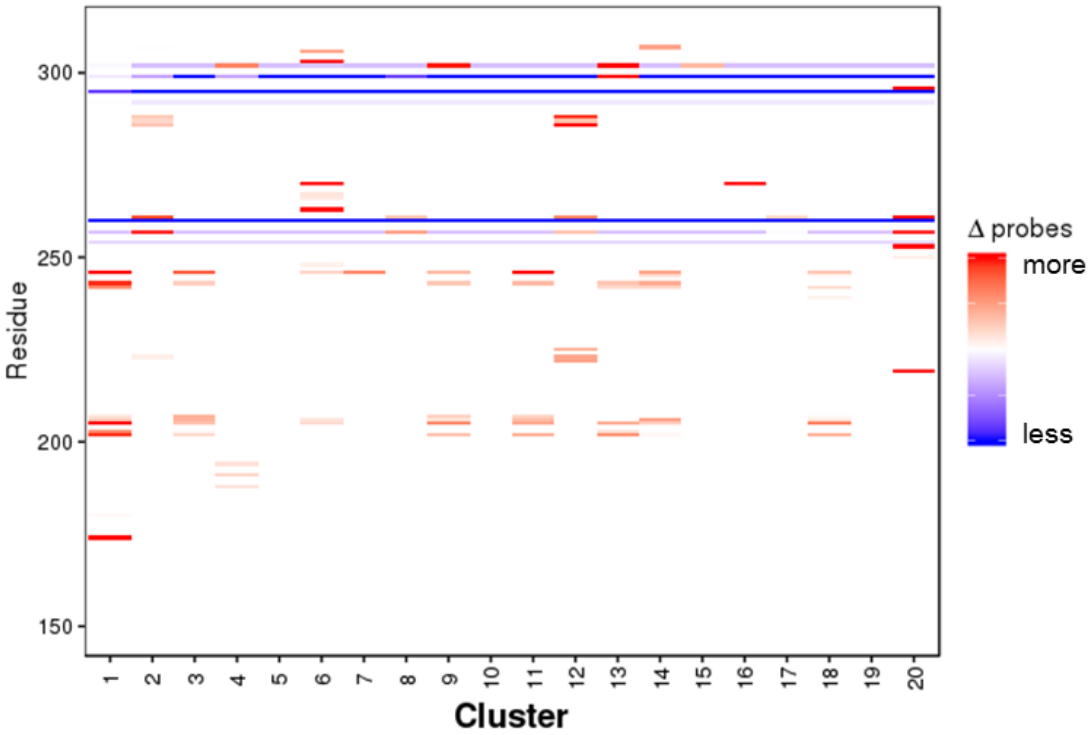


Figure 4.8. Comparison of probe changes from clustering hot region 3 heavy atoms.

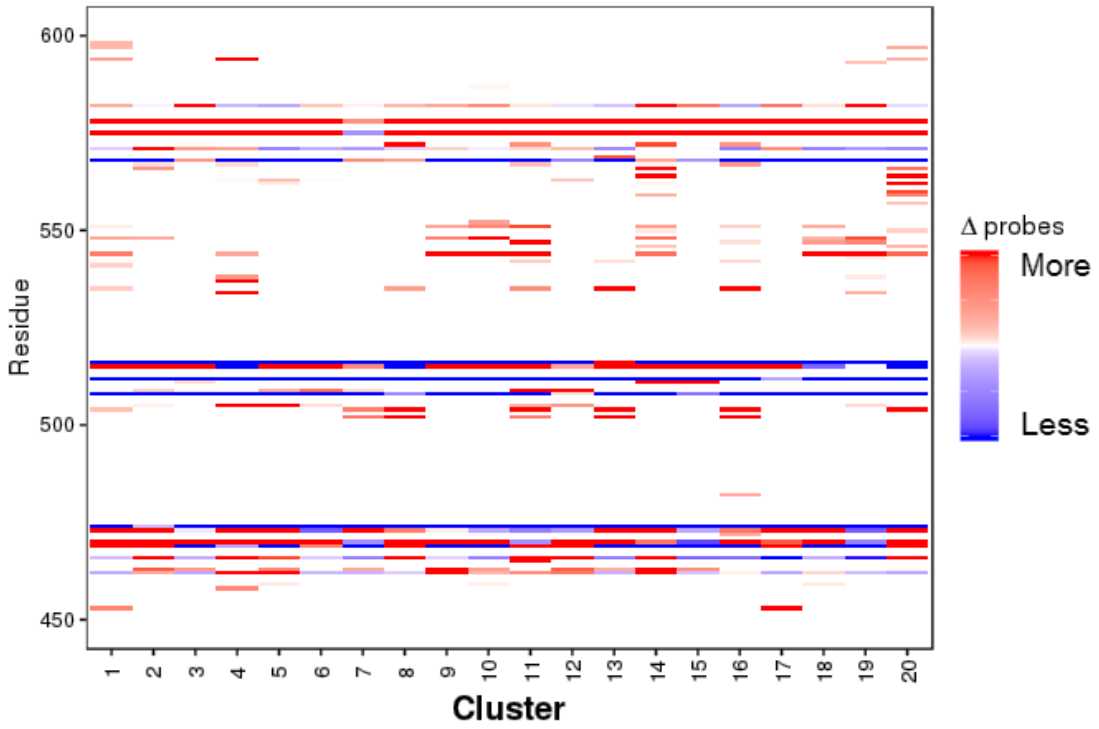


Figure 4.9. Comparison of probe changes from clustering the W504 site heavy atoms.

When taken together with the ET results, this pointed to the W504 site as a potential allosteric site. In order to probe for allostery between the W504 site and nearby hot region 1, a traditional alanine mutant and two mutants where the W504 side-chain was replaced with a bulky leucine or isoleucine side-chain were created. The intent with the latter two mutants was to mimic small-molecule binding. FP assays on β-catenin mutants were performed with labeled Tcf4 to test for changes in binding affinity. The results showed a modest increase in the ability of the W504I and W504L mutants to bind Tcf4, but no change in the W504A mutant (Figure 4.10). The W504I mutant decreased the K_d of the β-catenin/Tcf4 interaction from 2.3 nM to 0.72 nM. The W504L mutant decreased the K_d of the β-catenin/Tcf4 interaction to 1.3 nM. The W504A mutant remained unchanged at 2.1 nM. It was concluded that if allosteric modulation were achievable at this site, it would be by interaction with a bulky binding partner(s).

4.4 Discussion

The results from a combination of ET analysis, FTMap solvent mapping and MD simulations revealed important clues about the origins of catenin-like protein binding partner selectivity. Previous work on β-catenin inhibitors showed that hot region 1 is most important for the interaction involving Tcf4,[27] and the results from this study confirmed it. Hot regions 2 and 3 had residues with higher rvET scores, which is in line with research that indicated other binding partners, E-cadherin, APC and axin through phosphorylation.[23,24] This supported the notion that by altering the three hot regions, *Caenorhabditis* catenin-like proteins achieved specificity.

The solvent mapping results told a similar story. Evidence was seen of the varying

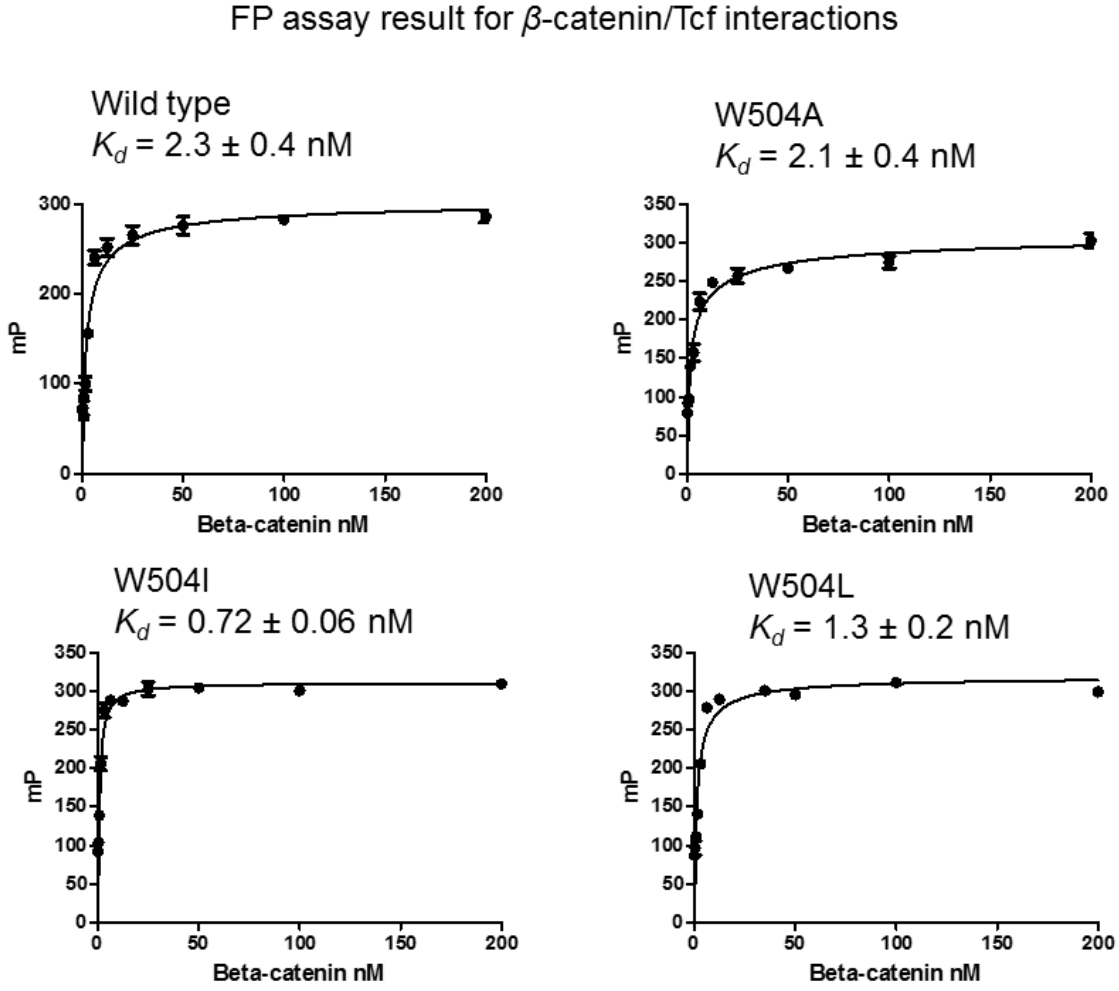


Figure 4.10. FP assay results of *β*-catenin point mutants. The W504A mutation showed no difference in Tcf binding. The W504I and W504L mutants bound Tcf4 more tightly.

importance and interactions of the three hot regions as well as hot region surface flexibility between binding partners when the different crystal structures were solvent mapped. In the Tcf4 bound crystal structure (PDB ID 2GL7), most of the solvent interactions were found in hot region 1, the hot region most important for binding Tcf4. The 1TH1, 1QZ7, and 1I7X structures were crystalized with binding partners that rely more on hot regions 2 and 3.[27] As expected, the number of probe interactions in these three structures in hot regions 2 and 3 increased, giving further evidence that these hot regions are more important for these binding partners than Tcf4.

MD simulations were used to reveal surface residue conformations that were not seen in other crystal structures. The combined use of FTMap and MD simulations were used in the past to reveal novel computational hot spots for drug design.[43] This was taken one step further by comparing the FTMap data to Evolutionary Trace scores in an attempt to identify novel allostery at potentially drugabble sites. A region near hot region 1 centered on residue W504 was identified as a potential novel allosteric site, because it did not contact Tcf4, E-cadherin, nor APC. It was first identified with ET analysis; however, it was not apparent that this region was suitable for binding a small molecule. By clustering the MD simulations based on the heavy atoms of the ARM domain and solvent mapping the representative structure, it was found that the conformation of this pocket may shift to one that is druggable and more accessible to FTMap probes. Because no known binding partners used this site, bulky amino acid side chains isoleucine and leucine were used to mimic a binding partner. The observed changes in Tcf4 binding were small, but the mutation was limited to a single residue with no known contact to Tcf4. Therefore, it was concluded that the potential for allostery between this site and the

nearby hot region 1 may be possible.

4.5 Conclusion

The β-catenin family of proteins, β-catenin, JUP, SYS-1, HMP-2, WRM-1, and BAR-1, fill important roles for cell development, differentiation, and adhesion.[1,32] The β-catenin interaction with Tcf4 is an important drug target because of its role in the canonical Wnt signaling pathway; the dysregulation of which is involved in many cancers and fibroses.[32] The downstream disruption of the canonical Wnt signaling pathway by inhibiting the β-catenin/Tcf4 PPI provides a means of disrupting cancer stem cell growth and division. This target is challenging due to the shared binding surface on β-catenin with Tcf4, APC, and E-cadherin. This study provided insights into achieving selectivity for the β-catenin/Tcf4 PPI, over the β-catenin/APC, and β-catenin/E-cadherin interactions. Solvent mapping by FTMap confirmed that Tcf4 relies heavily on hot region 1, while E-cadherin and APC utilize hot regions 2 and 3 in different surface conformations. Evolutionary Trace analysis of β-catenin, JUP, BAR-1, HMP-2, WRM-1, and SYS-1 revealed the origins of β-catenin binding partner selectivity and was able to discover a novel allosteric site near hot region 1. This novel allosteric site, centered on W504, did not have any interaction with probes from FTMap. However, when molecular dynamics simulations were conducted to reveal new FTMap hot spots not present in the analyzed crystal structures, several cluster centroids revealed probe interaction with W504 and the surrounding residues. W504 mutant β-catenin was tested for its ability to bind Tcf4 in order to confirm the presence of allosteric regulation between W504 and hot region 1. W504I/L mutant β-catenin proteins showed a mild positive allosteric effect on

Tcf4 binding in an FP assay. This unique combination of computational approaches has yielded valuable data from which better selective inhibitors of the *β*-catenin interaction may be designed.

4.6 References

1. Polakis, P. Drugging Wnt Signalling in Cancer. *EMBO J.* **2012,** *31*, 2737–2746.

2. Morin, P. J. β-Catenin Signaling and Cancer. *Bioessays.* **1999,** *21*, 1021–1030.

3. Nusse, R.; Varmus, H. E. Wnt Genes. *Cell.* **1992,** *69*, 1073 β 1087.

4. Rubinfeld, B.; Robbins, P.; El-Gamil, M.; Albert, I.; Porfiri, E.; Polakis, P. Stabilization of β-Catenin by Genetic Defects in Melanoma Cell Lines. *Science.* **1997,** *275*, 1790–1792.

5. Kinzler, K. W.; Vogelstein, B. Lessons from Hereditary Colorectal Cancer. *Cell.* **1996,** *87*, 159–170.

6. Barker, N.; Ridgway, R. A.; Van Es, J. H.; Van De Wetering, M.; Begthel, H.; Van Den Born, M.; Danenberg, E.; Clarke, A. R.; Sansom, O. J.; Clevers, H. Crypt Stem Cells as The Cells-of-Origin of Intestinal Cancer. *Nature* **2009,** *457*, 608–611.

7. Malanchi, I.; Peinado, H.; Kassen, D.; Hussenet, T.; Metzger, D.; Chambon, P.; Huber, M.; Hohl, D.; Cano, A.; Birchmeier, W.; Huelsken, J. Cutaneous Cancer Stem Cell Maintenance Is Dependent on β-Catenin Signaling. *Nature.* **2008,** *452*, 650–653.

8. Varelas, X.; Bouchie, M. P.; Kukuruzinska, M. A. Protein N-Glycosylation in Oral Cancer: Dysregulated Cellular Networks Among DPAGT1, E-Cadherin Adhesion and Canonical Wnt Signaling. *Glycobiology.* **2014,** *24*, 579–591.

9. Angelova, M.; Ferris, M.; Swan, K. F.; Mcferrin, H. E.; Pridjian, G.; Morris, C. A.; Sullivan, D. E. Kaposi's Sarcoma-Associated Herpesvirus G-Protein Coupled Receptor Activates the Canonical Wnt/B-Catenin Signaling Pathway. *Virology Journal* **2014,** *11*, 218.

10. Roose, J.; Clevers, H. TCF Transcription Factors: Molecular Switches in Carcinogenesis. *BBA Rev. Cancer.* **1999,** *1424*, M23–M37.

11. Rasad, T. S. K.; Goel, R.; Kandasamy, K.; Keerthikumar, S.; Kumar, S., Mathivanan,

S.; Telikicherla, D.; Raju, R.; Shafreen, B.; Venugopal, A.; Balakrishnan, L.; Marimuthu, A.; Banerjee, S.; Somanathan, D. S.; Sebastian, A.; Rani, S.; Ray, S.; Kishore, C. J. H.; Kanth, S.; Ahmed, M.; Kashyap, M.; Mohmood, R.; Ramachandra, Y. L.; Krishna, V.; Rahiman, A. B.; Mohan, S.; Ranganathan, P.; Ramabadran, S.; Chaerkady, R.; Pandey, A. Human Protein Reference Database - 2009 Update. *Nucleic Acids Res.* **2009,** *37*, D767–D772.

12. Mishra, G.; Suresh, M.; Kumaran, K.; Kannabiran, N.; Suresh, S.; Bala, P.; Shivkumar, K.; Anuradha, N.; Reddy, R.; Raghavan, T.M.; Menon, S.; Hanumanthu, G.; Gupta, M.; Upendran, S.; Gupta, S.; Mahesh, M.; Jacob, B.; Matthew, P.; Chatterjee, P.; Arun, K. S.; Sharma, S.; Chandrika, K. N.; Deshpande, N.; Palvankar, K.; Raghavnath, R.; Krishnakanth, K.; Karathia, H.; Rekha, B.; Rashmi, N. S.; Vishnupriya, G.; Kumar, H. G. M.; Nagini, M.; Kumar, G. S. S.; Jose, R.; Deepthi, P., Mohan, S. S., Gandhi, T. K. B., Harsha, H. C., Deshpande, K. S., Sarker, M., Prasad, T. S. K.; Pandey, A. Human Protein Reference Database - 2006 Update. *Nucleic Acids Res.* **2006,** *34*, D411–D414.

13. Peri, S.; Navarro, J. D.; Amanchy, R.; Kristiansen, T. Z.; Jonnalagadda, C. K.; Surendranath, V.; Niranjan, V.; Muthusamy, B.; Gandhi, T. K.; Gronborg, M.; Ibarrola, N.; Deshpande, N.; Shanker, K.; Shivashankar, H. N.; Rashmi, B. P.; Ramya, M. A.; Zhao, Z.; Chandrika, K. N.; Padma, N.; Harsha, H. C.; Yatish, A. J.; Kavitha, M. P.; Menezes, M.; Choudhury, D. R.; Suresh, S.; Ghosh, N.; Saravana, R.; Chandran, S.; Krishna, S.; Joy, M.; Anand, S. K.; Madavan, V.; Joseph, A.; Wong, G. W.; Schiemann, W. P.; Constantinescu, S. N.; Huang, L.; Khosravi-Far, R.; Steen, H.; Tewari, M.; Ghaffari, S.; Blobe, G. C.; Dang, C. V.; Garcia, J. G.; Pevsner, J.; Jensen, O. N; Roepstorff, P.;Deshpande, K. S.; Chinnaiyan, A. M.; Hamosh, A.; Chakravarti, A.; Pandey, A. Development Of Human Protein Reference Database as an Initial Platform for Approaching Systems Biology in Humans. *Genome Res.* **2003,** *13*, 2363–2371.

14. Xing, Y.; Clements, W. K.; Kimelman, D.; Xu, W. Crystal Structure of a β -Catenin/Axin Complex Suggests a Mechanism for the β -Catenin Destruction Complex. *Genes Dev.* **2003,** *17*, 2753–2764.

15. Xing, Y.; Clements, W. K.; Le Trong, I.; Hinds, T. R.; Stenkamp, R.; Kimelman, D.; Xu, W., Crystal Structure of a β-Catenin/APC Complex Reveals a Critical Role for APC Phosphorylation in APC Function. *Mol. Cell* **2004,** *15*, 523–533.

16. Xing, Y.; Clements, W. K.; Kimelman, D.; Xu, W. Crystal Structure of a Beta-Catenin/Axin Complex Suggests a Mechanism for the β-Catenin Destruction Complex. *Genes Dev.* **2003,** *17*, 2753–2764.

17 Sampietro, J.; Dahlberg, C. L.; Cho, U. S.; Hinds, T. R.; Kimelman, D.; Xu, W. Crystal Structure of a β-catenin/BCL9/Tcf4 Complex. *Mol. Cell* **2006,** *24*, 293–300.

18. Peifer, M.; McCrea, P. D.; Green, K. J.; Wieschaus, E.; Gumbiner, B. M. The Vertebrate Adhesive Junction Proteins β-Catenin and Plakoglobin and the Drosophila Segment Polarity Gene Armadillo Form a Multigene Family with Similar Properties. *J. Cell Biol.* **1992,** *118* (3), 681–691.

19. Su, L. K.; Vogelstein, B.; Kinzler, K. W. Association of the APC Tumor Suppressor Protein with Catenins. *Science* **1993,** *262* (5140), 1734–1737.

20. Rubinfeld, B.; Souza, B.; Albert, I.; Muller, O.; Chamberlain, S. H.; Masiarz, F. R.; Munemitsu, S.; Polakis, P. Association of the APC Gene Product with β-Catenin. *Science* **1993,** *262* (5140), 1731–1734.

21. Poy, F.; Lepourcelet, M.; Shivdasani, R. A.; Eck, M. J. Structure of a Human Tcf4-β-Catenin Complex. *Nat. Struct. Biol.* **2001,** *8*, 1053–1057.

22. Graham, T. A.; Ferkey, D. M.; Mao, F.; Kimelman, D.; Xu, W. Tcf4 can Specifically Recognize β-Catenin Using Alternative Conformations. *Nat. Struct. Biol.* **2001,** *8* (12), 1048–1052.

23. Eklof Spink, K.; Fridman, S. G.; Weis, W. I. Molecular Mechanisms of β-Catenin Recognition by Adenomatous Polyposis Coli Revealed by the Structure of an Apc-Beta-Catenin Complex. *EMBO J.* **2001,** *20* (22), 6203–6212.

24. Ha, N. C.; Tonozuka, T.; Stamos, J. L.; Choi, H. J.; Weis, W. I. Mechanism of Phosphorylation-Dependent Binding of APC to β-Catenin and its Role in Beta-Catenin Degradation. *Mol. Cell.* **2004,** *15* (4), 511–521.

25. Graham, T. A.; Weaver, C.; Mao, F.; Kimelman, D.; Xu, W. Crystal Structure of a β-Catenin/Tcf Complex. *Cell* **2000,** *103* (6), 885–896.

26. von Kries, J. P.; Winbeck, G.; Asbrand, C.; Schwarz-Romond, T.; Sochnikova, N.; Dell'Oro, A.; Behrens, J.; Birchmeier, W. Hot Spots in β-Catenin for Interactions with Lef-1, Conductin and APC. *Nat. Struct. Biol.* **2000,** *7*, 800–807.

27. Xu, W.; Kimelman, D. Mechanistic Insights from Structural Studies of β-Catenin and its Binding Partners. *J. Cell Sci.* **2007,** *120*, 3337–3344.

28. Natarajan, L.; Witwer, N. E.; Eisenmann, D. M. The Divergent Caenorhabditis Elegans β-Catenin Proteins Bar-1, Wrm-1 and Hmp-2 Make Distinct Protein Interactions but Retain Functional Redundancy In Vivo. *Genetics* **2001,** *159* (1), 159–172.

29. Liu, J.; Phillips, B. T.; Amaya, M. F.; Kimble, J.; Xu, W. The C. Elegans Sys-1 Protein is a Bona Fide β-Catenin. *Dev. Cell* **2008,** *14*, 751–761.

30. Huber, A. H.; Weis, W. I. The Structure of the β-Catenin/E-Cadherin Complex and

the Molecular Basis of Diverse Ligand Recognition by β-Catenin. *Cell* **2001,** *105*, 391–402.

31. Peifer, M.; Berg, S.; Reynolds, A. B. A Repeating Amino Acid Motif Shared by Proteins with Diverse Cellular Roles. *Cell* **1994,** *76* (5), 789–791.

32. Clevers, H.; Nusse, R. Wnt/ β-Catenin Signaling and Disease. *Cell* **2012,** *149*, 1192–1205.

33. Goujon, M.; McWilliam, H.; Li, W.; Valentin, F.; Squizzato, S.; Paern, J.; Lopez, R. A New Bioinformatics Analysis Tools Framework at Embl-Ebi. *Nucleic Acids Res.* **2010,** *38* (web server issue), W695-699.

34. Choi, H.-J.; Gross, J. C.; Pokutta, S.; Weis, W. I. Interactions of Plakoglobin and β-Catenin with Desmosomal Cadherins: Basis of Selective Exclusion of α- and β-Catenin from Desmosomes. *J. Biol. Chem.* **2009,** *284*, 31776–31788.

35. Kidd, A. R.; Miskowski, J. A.; Siegfried, K. R.; Sawa, H.; Kimble, J. A β-Catenin Identified by Functional Rather than Sequence Criteria and its Role in Wnt/Mapk Signaling. *Cell* **2005,** *121*, 761–772.

36. Shimizu, M.; Fukunaga, Y.; Ikenouchi, J.; Nagafuchi, A. Defining the Roles of β-Catenin and Plakoglobin in Lef/T-Cell Factor-Dependent Transcription Using β-Catenin/Plakoglobin-Null F9 Cells. *Mol. Cell. Biol.* **2008,** *28*, 825–835.

37. Korswagen, H. C.; Herman, M. A.; Clevers, H. C. Distinct β-Catenins Mediate Adhesion and Signaling Functions in C. Elegans. *Nature* **2000,** *406*, 527–532.

38. Catrow, J. L.; Zhang, Y.; Zhang, M.; Ji, H. Discovery of Selective Small-Molecule Inhibitors for the β-Catenin/T-Cell Factor Protein–Protein Interaction Through the Optimization of the Acyl Hydrazone Moiety. *J. Med. Chem.* **2015,** *58*, 4678–4692.

39. Huang, Z.; Zhang, M.; Burton, S. D.; Katsakhyan, L. N.; Ji, H. Targeting the Tcf4 G13Ande17 Binding Site to Selectively Disrupt β-Catenin/T-Cell Factor Protein–Protein Interactions. *ACS Chem. Biol.* **2014,** *9*, 193–201.

40. Sung, Y. M.; Wilkins, A. D.; Rodriguez, G. J.; Wensel, T. G.; Lichtarge, O. Intramolecular Allosteric Communication in Dopamine D2 Receptor Revealed by Evolutionary Amino Acid Covariation. *Proc Natl Acad Sci U S A* **2016,** *113* (13), 3539–3544.

41. Kozakov, D.; Grove, L. E.; Hall, D. R.; Bohnuud, T.; Mottarella, S. E.; Luo, L.; Xia, B.; Beglov, D.; Vajda, S. The FTMap Family of Web Servers for Determining and Characterizing Ligand-Binding Hot Spots of Proteins. *Nat. Protocols*. **2015,** *10* (5), 733–755.

42. Brenke, R.; Kozakov, D.; Chuang, G. Y.; Beglov, D.; Hall, D.; Landon, M. R.; Mattos, C.; Vajda, S. Fragment-Based Identification of Druggable 'Hot Spots' of Proteins Using Fourier Domain Correlation Techniques. *Bioinformatics*. **2009,** *25* (5), 621–627.

43. Kozakov, D.; Hall, D. R.; Chuang, G-Y.; Cencic, R.; Brenke, R.; Grove, L. E.; Beglov, D.; Pelletier, J.; Whitty, A.; Vajda, S. Structural Conservation of Druggable Hot Spots in Protein–Protein Interfaces. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108* (33), 13528–13533.

44. Bohnuud, T.; Beglov, D.; Ngan, C. H.; Zerbe, B. S.; Hall, D. R.; Brenke, R.; Vajda S.; Frank-Kamenetskii M. D.; Kozakov D. Computational Mapping Reveals Dramatic Effect of Hoogsteen Breathing on Duplex DNA Reactivity with Formaldehyde. *Nucleic Acids Res*. **2012**, *40* (16), 7644–7652.

45. Ritco-Vonsovici, M.; Ababou, A.; Horton, M. Molecular Plasticity of β-Catenin: New Insights from Single-Molecule Measurements and MD Simulation. *Protein Sci.* **2007,** *16* (9), 1984–1998.

46. *Sybyl Suite*, version 2.0; Certara USA, Inc., Princeton, NJ, 2016.

47. R Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-Project.org/. **2016**

48. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. Ucsf Chimera–a Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004,** *25* (13), 1605–1612.

49. Wickham, H. *Ggplot2: Elegant Graphics for Data Analysis*. Springer–Verlag New York. **2009.**

50. Wickham, H. Reshaping Data with the {Reshape} Package. *J. Stat. Soft.*, **2007,** *21* (12) , 1–20.

51. Uniprot: A Hub for Protein Information. *Nucleic Acids Res*. **2015**, 43, D204–212.

52. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Soding, J.; Thompson, J. D.; Higgins, D. G. Fast, Scalable Generation of High-Quality Protein Multiple Sequence Alignments Using Clustal Omega. *Mol. Syst. Biol.* **2011,** *7*, 539.

53. Goujon, M.; McWilliam, H.; Li, W.; Valentin, F.; Squizzato, S.; Paern, J.; Lopez, R. A New Bioinformatics Analysis Tools Framework at Embl-Ebi. *Nucleic Acids Res.* **2010,** *38* (Web Server issue), W695–699.

54. Wilkins, A. D.; Lua, R.; Erdin, S.; Ward, R. M.; Lichtarge, O. Sequence and Structure Continuity of Evolutionary Importance Improves Protein Functional Site Discovery and Annotation. *Protein Sci* **2010,** *19* (7), 1296–1311.

55. *Schrödinger Release 2016-3: MS Jaguar*, Schrödinger, LLC, New York, NY, **2016**.

56. Xing, Y.; Takemaru, K.; Liu, J.; Berndt, J. D.; Zheng, J. J.; Moon, R. T.; Xu, W. Crystal Structure of a Full-Length β-Catenin. *Structure* **2008,** *16* (3), 478–487.

57. *Schrödinger Release 2016-3: Macromodel*, Schrödinger, LLC, New York, NY, 2016.

58. Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. Gromacs: High Performance Molecular Simulations Through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015,** *1–2*, 19–25.

59. Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations. *J. Comput. Chem.* **2003,** *24*, 1999–2012.

60. Daura, X.; Gademann, K.; Jaun, B.; Seebach, W.; Gunsteren, W, F.; Mark, A, E. Peptide Folding: when Experiment Meets Simulation. *Angew. Chem. Int. Ed.* **1999,** *38*, 1433–7851.

61. Beaulieu, J. F. Tuning Wnt-β-Catenin Signaling via Bcl9 Proteins for Targeting Colorectal Cancer Cells. *EBioMedicine* **2015,** *2* (12), 1846–1847.

62. Enkhbayar, P.; Kamiya, M.; Osaki, M.; Matsumoto, T.; Matsushima, N. Structural Principles of Leucine-Rich Repeat (Lrr) Proteins. *Proteins* **2004,** *54* (3), 394–403.

63. Kobe, B.; Deisenhofer, J. The Leucine-Rich Repeat: A Versatile Binding Motif. *Trends Biochem. Sci.* **1994,** *19* (10), 415–421.

4.7 Supplementary materials of real-value evolutionary trace scores

Table 4S.1. Full list of ET scores

Residue	Variations	rvET scores	Residue	Variations	rvET scores
D144	IVRE.GDCSN	8.26	C429	AIHLCS	3.83
D145	RKN.QSED	8.08	N430	HSGNQ	3.45
A146	QRPSAH.YNTK	12.06	N431	KMPNH	3.81
E147	DEVC.NIA	7.83	Y432	QLESKGRVTPHAIMY	16.85
L148	FLSTM.EAVI	11.12	K433	HP.LTMARNEYKQWS	11.75
A149	AVNQ.TGS	9.22	N434	VN	1.32
T150	QNTALMR.VKEG	12.93	K435	KVTC	2.64
R151	VIDGR.SNYEQTKFL	11.85	M436	DEATFQKLSIYVM	17.83
A152	NQETAS.GV	8.93	M437	ILKRAEFHSTVMQ	17.02
I153	MLYNHI.VAS	10.53	V438	AFVMLT	6.2
P154	PWSQ.E	5.95	C439	IQYSVACFHT	8.85
E155	QRHDE.KNL	9.46	Q440	RAGENQSKF	10.91
L156	ILV.FA	6.67	V441	SKNYLARMFCHV	16.83
T157	LQKRIC.VTASP	12.34	G442	NGKRHMSA	12.69
K158	NLKQD.MHRESG	11.32	G443	AGVR	3.92
L159	ILAFSNQ.H	6.53	I444	ILHVTM	8.64
L160	IVLM.T	5.91	E445	FPGESDVMTQ	10.08
N161	YHKRATS.NIGCQ	15.43	A446	LTQIHNARSGV	9.85
D162	DNESA.GHQ	7.53	L447	LAVI	3.61
E163	GQETDSANK.VPH	9.89	V448	HCVILSMY	9.6
D164	LMTDEGKN.S	7.81	R449	TDKRGQHASC	12.27
Q165	PAVTY.ENQSL	11.56	T450	IVTLQAMC	12.48
V166	QTVSRAED.NML	11.6	V451	ILVC	6.29
V167	PQVICTE.	9.84	L452	.FNQKSRCLGEAITMYV	13.96
V168	VILHQTS.C	9.9	R453	.TVSDEQLNIRKAH	12.02
N169	QTAEYHVD.NKSGLM	15.56	A454	.TNYARHSQPICGFL	13.53
K170	LTFSVNKRQ.EH	12.02	G455	.NDGQLEPVIYATS	11.47
A171	SACT.V	7.82	D456	.QSIVRDEGAFTN	12.43
A172	YASTIV.	10.4	R457	.PCRIMSVKQL	9.94
V173	YMLRAITKQ.SVH	11.05	E458	.EPDLQFAG	8.19
M174	.FMRCGEYLHVIT	13.37	D459	.ERSD	5.87
V175	.QDLMVAI	8.18	I460	SMVTAI.Y	10.59
H176	APFSEQYH.N	10.42	T461	KRLVY.QTAFIS	9.55
Q177	QNSPHAM.GDEKRL	9.17	E462	YSVAE.	4.01

Table 4S.1. Continued.

Residue	Variations	rvET scores	Residue	Variations	rvET scores
L178	LMPIV.	6.02	P463	PTN.CSR	5.61
S179	WYMASC.T	9.77	A464	RPVSTC.LAIG	8.75
K180	YCQVTKR.NH	9.25	I465	LVAC.IM	7.24
K181	NYEDKM.TSR	10.64	C466	DSGAC.	5.82
E182	L.FWKGDEQVN	8.97	A467	EDAGTVIPC	9.59
A183	LTISACGP.K	8.14	L468	LFR	2.96
S184	DCSLEA.FN	6.34	R469	TANGPLKR	7.15
R185	IQKENHPRTM.LCFV	9.57	H470	DEINKH	5.45
H186	FSQLVAIDE.HTMKRYN	15.78	L471	AGFLMVHSIC	8.5
A187	LYAQSTRP.GN	7.91	T472	PSYATCN	5.64
I188	PRIYML.VTA	9.45	S473	KQGTLSVAHYC	9.53
M189	ITRFGAQ.VLNMP	10.05	R474	RK.MGN	4.91
R190	ITDESCVRANK.GQ	13.36	H475	NK.SADH	6.13
S191	SIAVNLCGDTY.H	11.67	Q476	RL.QPEASGHDVN	14.33
P192	PGKQASTIFRM.NH	12.96	E477	VA.NQSYMHLEDFG	12.75
Q193	YNGLQERKD.AVS	14.82	A478	C.DIASQYEV	6.27
M194	.ILVMAF	8.29	E479	EGIDQPVSNA	9.86
V195	ILVA.	7.36	M480	IPESQKDATMRVLHNF	12.41
S196	HEPDRTAFSKGLQN	12.87	A481	IVAPGETC	7.7
A197	QLVPATHEGC	10.86	Q482	CSNILVRQE	10.55
I198	VILM	6.18	N483	NKRHTSAYD	8.45
V199	VIFLRKAMC	9.69	A484	CSAMHFEITDG	13.58
R200	QNKEVDRHSGA	13.76	V485	LIVACMF	8.56
T201	MLCVKTA	10.03	R486	RLQDAIH	6.82
M202	FLV.IMTA	7.9	L487	TAKSGMDVQHELNIF	11.98
Q203	P.SRDTQVINGEA	12.99	H488	LMCAHKIG.EDQSFNT	15.3
N204	RKDSTLNHMAQ	13.26	Y489	NGS.FPQAYH	9.59
T205	EGTSVCAINFL	11.4	G490	NGAVES.TL	5.82
N206	NDEPSVGRHKTAQ	16.34	L491	FIVNLSPA.	7.51
D207	GQEDHSN	7.98	P492	LKPQTSRAE.	8.56
V208	PQVEDINKASLMG	16.44	V493	MSPADLF.TNIV	10.01
E209	QYHKDNEASR	10.99	V494	MILSDF.V	8.27
T210	DECVILTAS	9.76	V495	WVLAT.SIMC	7.02
A211	FKQILVETMAR	10.85	K496	IDNEHAR.SVK	11.42
R212	RIKHNAS	8.07	L497	PLVISRK.FQ	7.79
R212	RIKHNAS	8.07	L497	PLVISRK.FQ	7.79
C213	EDGNKFVPCYSA	13.73	L498	TMLV.I	5.02
T214	LSTACIV	8.48	H499	PQTKLSY.REHNDAC	13.21

Table 4S.1. Continued.

Residue	Variations	rvET scores	Residue	Variations	rvET scores
A215	ILVFCSATQG	12.54	P500	NTQ.PM	6.64
G216	CKEGLSA	5.41	P501	GN.PTEDAVQLS	6.97
T217	NIATSLV	7.6	S502	EQT.LMSKNVICY	10.54
L218	LAVIM	7.22	H503	.HRGSMNKQY	7.01
H219	IMRKFAHSQY	9.14	W504	.YTPWLA	4.8
N220	SQKENHYL	9.73	P505	.FPALSEM	4.29
L221	LPISVFM	9.4	L506	.NSVLI	5.95
S222	NTESVLCAY	8.87	I507	.VILA	7.59
H223	K.NHRQLTPEDYCS	13.23	K508	TRAQK.	5.49
H224	D.NSQTHRGLKMY	11.61	A509	KMLVSQH.A	5.27
R225	P.EVTYRQKSA	11.49	T510	TPELQA.VIM	12.67
E226	H.QDENVPTASG	11.24	V511	AIMVSL.C	10.99
G227	M.ISAG	5.55	G512	GNKTALQ.S	4.47
L228	K.SRVILAP	8.36	L513	PEQVTSIL.	8.02
L229	NKRLAIEDMTSV	11	I514	NYVTSMQLIF	12.55
A230	QAVFDLSNHIG	10.77	R515	EKRQLVN	5.71
I231	VTIML	8.18	N516	KQACREN	5.61
F232	FVAHRKLTWQY	7.68	L517	QVLTAICS	6.08
K233	QDHERTNKA	10.68	A518	QGAIPSVE	5.68
S234	IMLAP.SEVYC	10.03	L519	VAQELPSRM	6.98
G235	.GDEVIFNRQST	10.17	C520	CRAVEDLQGSNTY	11.03
G236	.GRNSFVC	5.42	P521	KRTECPDNGALISQ	10.24
I237	FLIV	8.77	A522	EDSRCVAQKHGT	12.6
P238	NEQPDTA	7.73	N523	IEDNASFLC	7.81
A239	CNSTPKIHDEVFA	13.25	H524	D.MRASLQIFHYV	8.88
L240	ILMV	6.46	A525	IVFASEPHRKYTGQ	15.61
V241	ILVYSC	6.28	P526	LIEDPGSVRFA	8.71
K242	MKQEAHRDN	13.17	L527	KILFHM	5.96
M243	GCYILSQFHM	10.57	R528	KRDQLSNIYG	12.48
L244	VLSTMI	5.56	E529	DQKSNGPE	10.21
G245	.NDRKAEGSTY	12.34	Q530	TVLAYIDKCQNHE	11.91
S246	.NSDMTFHYVI	9.82	G531	PVADGTSN	11.61
P247	.PDEVSAFTNQR	10.85	A532	GPNVFTILCA	10.54
V248	KVDFLISPGM	10.91	I533	LVIETGM	8.72
D249	NYKPALDFGEM	10.29	P534	LMRQEPFNHYWTAS	12.85
S250	EDVITSAKNP	8.3	R535	ELPNQRGSTHK	10.58
V251	KLIHVSA	3.98	L536	LVASNI.MT	9.36
L252	LKQYIVFR	6.44	V537	RWVIQ.TCFM	8

Table 4S.1. Continued.

Residue	Variations	rvET scores	Residue	Variations	rvET scores
F253	HSNACKYRTEF	10.56	Q538	SNMLIH.EQVTRD	10.29
Y254	LMICYTF	5.82	L539	TLMIV.AY	4.5
A255	KASTV	4.24	L540	ILS.FC	2.76
I256	FYALPMVI	8.48	V541	LKHQTR.YNISMVFAG	14.44
T257	EMKLYTSN	6.36	R542	RNSTV.QAK	5.8
T258	KRTALS	5.62	A543	SFPQYGA.TI	9.52
L259	LMVIF	6.69	H544	FSNHYPC.TQVIG	12.67
H260	VIFAQDRKHN	9.15	Q545	IFLAPKDEHT.QNY	15.61
N261	GQNRHST	7.55	D546	LDESYQ.TGHMA	6.76
L262	TAVILM	6.11	T547	LVSF.AIYNT	10.96
L263	LATSFIM	4.19	Q548	LSQREIPN.TKD	13.16
L264	SLRKTQAWYIM	11.6	R549	RSAHEQPV.DGNK	11.68
H265	EGRFDQVHY	11.08	R550	PAVKQE.STR	9.47
Q266	FL.VRSQIPHM	8.96	T551	FSACQTVM.LGR	14.74
E267	NGDRKSEPAVQ	13.54	S552	VMIWLTAG.QCSHPN	15.87
G268	PQRDESLGY	10.51	M553	PAITL.SRKVMN	12.93
A269	QTANPDSG	10.42	G554	KALIR.TNSDG	13.79
K270	VILRATQK	9.02	G555	DECSAT.INMPG	13.03
M271	HPRISVDANELMQTG	14.85	T556	GVSPLAT.DIQN	16.19
A272	PLTAQVMEIDS	11.35	Q557	VFLCRDI.TNEAGPQSM	15.97
V273	GSPVFLRIAT	10.8	Q558	LFRKIP.GMNSAQTDE	16.97
R274	MQLRC	4.65	Q559	NGRS.EPLTAKQI	15.67
L275	ILQDEHGRANKSM	14.55	F560	VLF.IHKWSNMRYA	14.43
A276	NYHELQKSRAC	11.09	V561	ITVFLDSHR.QAN	12.7
G277	PGQETHDSNL	9.07	E562	DGEKRQ.N	5.06
G278	ATGVLIC	7.68	G563	EDGAKRLNS.	8.28
L279	IVFMLTH	7.81	V564	NFLMRVIK.	6.92
Q280	FLKESTNIQPV	9.71	R565	RKEDA.LSYQP	11.64
K281	IVLKFRPHAGCN	9.36	M566	KRILHTV.MS	6.46
M282	IVLASM	9.3	E567	EKRVQLA.DSNMW	10.34
V283	FLVMITAYS	10.64	E568	NSED.G	6.56
A284	RKGDQHEPSATNY	13.52	I569	LVI.MF	5.76
L285	FLINSVWQKPH	8.66	V570	IVL.M	7.44
L286	ILVSM	4.66	E571	GANDKEVHQ.W	5.26
N287	IDSANKREQVH	13.05	G572	HSARLGEVC	5.49
K288	D.EAPRIYKH	10.08	C573	IVNSGTCA	11.23
T289	RVQCNHSTKDPG	15.03	T574	CALVMFIT	9.72
N290	LIV.KNSAD	7.1	G575	AISGRQVKCN	8.43

Table 4S.1. Continued.

Residue	Variations	rvET scores	Residue	Variations	rvET scores
V291	KPTMERVIY	8	A576	AVLCITE	4.85
K292	DNIEKRQ	7.98	L577	YLFIK	5.18
F293	YCGLVWNF	11.36	H578	SCAYTHRQE	6.61
L294	FLMASNKYIV	8.3	I579	WKNSLAMIVQ	7.99
A295	ISADGPVT	6.3	L580	VAILTM	5.72
I296	WLTFIVMQ	6.49	A581	FDYATCGS	8.1
T297	NGTSHALICVD	13.49	R582	RGKSCN	8.95
T298	NPKAVFLIT	11.74	D583	QDHKE	8.11
D299	NDEMS	8.2	V584	PKQELGAR.NHFISV	18.26
C300	PRQAICLTSG	11.17	H585	NTEASKRDH.LGYFVMIQ	15.77
L301	ILAV	5.72	N586	NSPVLI.AT	7.21
Q302	IMLQKHRY	7.79	R587	TELSR.KA	6.51
I303	KLAVNIMF	10.49	I588	QKADNET.SLMGI	16.49
L304	LVTI	2.81	V589	SDVILAEQ.YTF	16.78
A305	VICLSAT	9.2	I590	TLIV.CM	6.99
Y306	MLVSRYEDFHAI	11.24	R591	KRTSEIVA.DNYFQCL	13.63
G307	IMLKSETRGDNA	11.27	G592	QN.DHYFCRSLGAKE	16.38
N308	GRSDNH	8.75	L593	Q.GVLHFTEPMD	8.79
Q309	SCTKIAPEHNVYQG	13.78	N594	L.FVANIKMDQSRH	10.69
E310	YEHKRADSNPQ	12.32	T595	VI.YAHLQRGSTC	14.35
S311	RHQNSITEAGK	9.36	I596	IFVANL.T	8.88
K312	LMQKFR	7.43	P597	SPRTG.QHFAD	6.45
L313	ILMGEQYRNK	13.15	L598	LMPAIT.SV	10.57
I314	.RALMKTQEIVS	10.68	F599	LGN.FTVI	7.86
I315	.NIMAFLV	7.87	V600	LVIAQTNSC.MHY	10.5
L316	.SRHKQLVA	7.83	Q601	VRKPHL.QSE	8.78
A317	.QKEDRPLGACS	12.77	L602	LIF.EQYMC	6.88
S318	.IAGTMFLCSHN	12.05	L603	MITL.QKV	8.15
G319	.AGDNHSQTL	11.56	Y604	Q.PYARINFHSG	11.21
G320	ANSGVIL	7.82	S605	G.NDQTIESVMFR	11.59
P321	GFVILPAHS	7.72	P606	ADETSMPQ.NHVR	16.03
Q322	QPHTKNRASELVGMI	19.01	I607	ETVLISCD.NH	11.47
A323	NDLRIFKHEAQGS	12.32	E608	KVED.QIG	9.3
L324	PQLFV	3.7	N609	EQDSNP.RGHVL	10.61
V325	EMGAVLFI	10.15	I610	VRKLYIMT.	9.16
N326	TSRKEADLHQN	16.34	Q611	ALSCMTK.QHVE	8.64
I327	LMVFI	8.03	R612	QHRKPYGS.L	7.67
M328	PNLIMV	9.88	V613	YSGRQEP.AVF	9

Table 4S.1. Continued.

Residue	Variations	rvET scores	Residue	Variations	rvET scores
R329	QR.KETD	10.33	A614	SLHPEWAT.CVI	9.02
T330	N.EDRHTQSAIM	12.51	A615	YLVK.AST	5.68
Y331	.FKETCRLHYS	8.95	G616	STGAVE.NK	9.97
T332	EDP.SNGRTQIH	15.99	V617	IHRSTF.LVGA	8.25
Y333	LMHN.PDYQV	8.11	L618	DSCQIL.V	5.92
E334	ASKE.GPR	7.32	C619	CSLIQY.AV	8.51
K335	HMRN.EDKGY	7.51	E620	PINVASQME.LK	9.06
L336	LM.AFV	4.88	L621	LWFSIP.MV	5.65
L337	IALVT	6.94	A622	NSGCIMA.VE	10.32
W338	QITWMYF	4.72	Q623	LMSNTCQH.EGVAK	16.49
T339	VPTRNA	3.04	D624	L.ETHPSYDR	8.17
T340	IVNTGA	6.37	K625	NK.TLEPARD	10.34
S341	ITVLMSAC	7.48	E626	VMI.EDQHGANL	10.57
R342	RGSPQK	4.24	A627	K.NLVMGCFRYAS	10.16
V343	TSAEDLCV	8.69	A628	PNISCGLA.MV	7.18
L344	FLIVM	4.09	E629	TQLREKSDA.VGH	12.53
K345	DQKTRV	4.36	A630	FATKLIV.QMS	15.93
V346	LESMTAVQ	8.07	I631	ISANYCL.V	7.86
L347	LCNFMI	4.62	E632	NGSFLATREKD	9.62
S348	GNACS.	4.42	A633	VDRSGLIWTQNAEM	15.54
V349	LSVA.T	3.87	E634	LAEKTQDNIPCY	10.26
C350	DEKLASC	5.36	G635	VRGASQMWTPN	8.05
S351	SEPCAN.TLHQV	12.92	A636	VASLYFRIGP	9.05
S352	DNSIEAFH.KQM	11.48	T637	CSVDRNYMEFTGPA	14.17
N353	AYNKSDL	4.02	A638	DAVTILHQKSPENG	13.4
K354	IRKGLPA	3.85	P639	KPGTNVMAERLYI	9.87
P355	DILVSNPQHTKA	9.69	L640	ILERQTVF	6.15
A356	GAREVSNLKIC	7.79	T641	LVEDYQFTSGNMIA	8.97
I357	FILV	4.41	E642	EQRKGDFATNH	10.97
V358	VQRHSMIL	6.22	L643	HYQKLFTVAMS	7.94
E359	RKNTEVSAQD	13.12	L644	CSLGQRA.IV	10.4
A360	SLAEQDCFKTI	11.14	H645	PKNGTCR.HQI	7.87
G361	DGNEKSA	5.15	S646	TMLSPRA.F	8.01
G362	GMVASC	6.02	R647	RLKQPSN.HTA	13.39
M363	VIEALPRYM	7.82	N648	ADERNQYH	7.06
Q364	GQKESFDALVPI	13.02	E649	DNKETARSQMP	12.06
A365	ALMDVFI	8.96	G650	IGSANDRQYK	11.2
L366	IL	1.19	V651	WVTMGKCISFLE	10.16

Table 4S.1. Continued.

Residue	Variations	rvET scores	Residue	Variations	rvET scores
G367	TYCVIHASGL	13.79	A652	TSAHIN.VLQ	10.19
L368	TKEHGRNPSLQAVMD	13.79	T653	ITQRKPE.GAS	9.11
H369	VINLCRQHYF	8.83	Y654	DYSGR.HMT	6.78
L370	VILFKMC	5.83	A655	RSTDVI.PA	7.37
T371	QNSKVERDYHGATP	16.34	A656	.ACDRPQNSG	7.36
D372	YSVHCATGDN	9.72	A657	.MRTPNGYAIVL	8.44
P373	PQDAEGKTSRNHMV	15.6	V658	PSGATKYLNVI	11.78
S374	NSDE	4.92	L659	MTLPSIRNV	6.95
Q375	NSTDVALPQEHG	17.53	F660	LVYRKDMNPHIFS	12.8
R376	DNESRF	6.13	R661	ERQKSVCHN	10.12
L377	LVTS	5.18	M662	GMLKIFAT	10.28
V378	IVAQKLRTM.	10.07	S663	LFSKEI.TNAQ	11.92
Q379	RMTHYALIEQ.F	9.22	E664	TKEGDNHY.LARF	11.26
N380	ASHENCT.	7.37	D665	NDE.KPSVG	9.23
C381	GCTALISV.	8.07	K666	.KRHGQED	6.45
L382	CFAL.	4.81	P667	.LSPVNRK	9.89
W383	KSQGWEAYFD.VC	5.83	Q668	.QAPEDKGVSMHT	13.84
T384	LACTS.	6.16	D669	.PLIDERNAVM	10.92
L385	LIVM.C	6.52	Y670	.WFGYSHVLT	9.68
R386	LFDGR.	3.42	K671	H.KNCDARSL	11.92
N387	QKAVN.	4.24	K672	RS.MPQKDEGFNT	13.45
L388	VCMLI.	3.77	R673	N.MPQRSTHGAE	7.67
S389	SA.	2.31	L674	SPDTQLYMHFVIA	13.87
D390	DQEFN.	4.61	S675	DNEQFLRSPATI	9.58
A391	ASDEHNLVYMC.	11.53	V676	INVSLCPTHGMYA	12.56
A392	KRLPAVSGH.	9.01	E677	AECSNQRHP.D	10.45
T393	ASQELVTYKRN.	10.33	L678	KRQILNHY.VM	10.34
K394	LTSIDPKNHGR.	8.83	T679	ATQVRDP.FGNSE	10.22
Q395	KLRNYGHDEIMQVT.	14.26	S680	AVMLKQRN.EPTGSH	15.42
E396	TIDQSKLVAN.EP	13.13	S681	SGEVIQN.AP	9.25
G397	PNLSDVEGHATQ	11.42	L682	LKAEMQ.TPVIY	9.62
M398	LVFIMA.HTQ	10.28	F683	LGEF.YPAHCD	11.16
E399	EKRGTDSAQ	13.59	R684	SARVM.LPDGHQKFY	13.12
G400	NEQDHPIMVTKGSA	19.19	T685	RKVT.EPQINSDHGA	16.36
L401	IPTYLWSV	8.56	E686	FLKQE.SNDVMA	13.26
L402	LIV	4.33	P687	EDHKP.STNGAQ	17.2
G403	PQDHLSKTIMRVG	13.75	M688	NSAIML.TGVPEHQ	19.68
T404	FPASNQKRHVITMG	15.89	A689	DAQMVS.TLEPNI	18.56

Table 4S.1. Continued.

Residue	Variations	rvET scores	Residue	Variations	rvET scores
L405	LIV	6.19	W690	LMGDT.RYFWEI	8.66
V406	LMIVT	9.19	N691	LDSEA.GTFNPQM	15.34
Q407	RQASKMGCEDHNTL	16.89	E692	AEDQRT.GHSLPMN	20.16
L408	LRIVFSQM	9.34	T693	.ATYSNPVGQELD	11.66
L409	ILCVF	9.36	A694	.RQKPNEDASTH	11.3
G410	ERGSQNKAMDT	15.99	D695	.LFDPQNGEV	10.02
S411	HNATDSCPERGLVQI	16.55	L696	.LVPSGYMNIF	10.16
D412	PGENDSRVTAKQ	16.53	G697	.QSATRLGHCFDE	12.05
D413	DENSQ	6.35	L698	.EDMHQLANV	20.16
I414	DKLGVPAREITSM	14.03	D699	.AKHTNSQDEF	9.9
N415	EDQVTIAYNHSG	14.28	I700	.ASMQPLIV	20.16
V416	VILTMYS	11.92	G701	.AGDHSPLNT	20.16
V417	IVLKSRATN	10.81	A702	ILNMKA.DVCSYGT	15.74
T418	YAEQLMIT	5.12	Q703	FLIEGQMP.VHSTA	14.17
C419	SARCYF	2.5	G704	SHQRKP.TEDGLA	15.31
A420	GAICSV	5.66	E705	NTLRQP.EYDSIMHGA	16.07
A421	TLVGAMSC	10.41	P706	RSH.PQIADTNEM	15.82
G422	GQLD	3.42	L707	FT.PSHGVRADLQE	18.75
I423	FAVSICT	8	G708	.EDWLSHGAPVT	16.16
L424	LIM	3.54	Y709	.GLHSERYNT	11.74
S425	SWARGC	3.43	R710	.AHPQGMILRSYF	14.83
N426	NK	1.57	Q711	.LQVMSNYGAPDT	14.24
L427	VCLIM	3.08	D712	.ASNKVTQDMEGP	14.24
T428	VSATGIL	4.69	D713	.NEQPTSVDG	16.23

CHAPTER 5

FUTURE DIRECTIONS AND CONCLUSION

5.1 Future directions of novel β-catenin/Tcf inhibitors

Compounds **14** and **37** showed the ability of the 5*H*-[1,2,5]oxadiazolo-[3',4':5,6]pyrazino[2,3-*b*]indole (**31**) scaffold to inhibit the *β*-catenin/Tcf PPI.[1] Compound **37** achieved a 3.12 μM K_i in the FP assay and disrupted the canonical Wnt signaling pathway in zebrafish at a 10 μM concentration.[2,3] Compound **14** achieved a 17 μM K_i in the FP assay and a 20 μM IC_{50} against *β*-catenin-dependent colorectal cancer cell growth.[1] Despite these promising results, the removal of the acyl hydrazone PAIN structure was only the first step in the optimization of the compound. A 6*H*-indolo[2,3-*b*]quinoxaline (**44**) scaffold was developed to increase the synthetic yield of the first reaction, resulting in final compounds **48** and **49**. Compound **48** achieved a 67 μM K_i in the FP assay, and induced cell cycle arrest in MECs.[4] Compound **48** lost some activity compared to **14**; this is most likely due to the loss of a hydrogen bond between the furazan moiety and K508 of *β*-catenin. The SAR study is aimed at expanding the inhibitor into the hydrophobic pocket B and increasing potency. Compound **49** illustrates the potential for increased potency from this expansion. Finally, the molecule has low solubility in most solvents and the large flat scaffold may cause DNA intercalation. Therefore, further modifications are needed to increase efficacy and further develop drug-

like properties of the molecule: reinstatement the second hydrogen bond with K508; completion of the SAR of both the isatin and phenyldiamine moieties; and addition of more sp^3-hybridized carbons to the scaffold. Figure 5.1A illustrates potential substitutions which will be used and synthesized in a similar manner as scheme 4. With the inclusion of o-phenylenediamine, substitutions on the 3-, 4-, 5-, and 6-positions of the starting material are possible. Figure 5.1B shows potential o-phenylenediamine substitutions. Schemes 5-8 (Figure 5.2) present three synthetic routes to increase the number of sp^3-hybridized carbons to the scaffold or reinstate the hydrogen bond with K508. Figure 5.3 shows the predicted binding mode of the new compounds which differ from the original predicted binding mode of **14**. Once these studies are completed, a more complete battery of *in vivo* testing will be appropriate.

5.2 Practical applications of θ_I to the discovery of PPI inhibitors

Preliminary examination of the solvent exposed surface area of PPI inhibitors revealed a small, but significant, correlation between ligand burial (θ_I) and inhibitor potency in a limited test set. The two test proteins, Bcl-X_L and MDM2 were selected because of the large quantity of structural data available for protein-inhibitor complexes. It was hypothesized that if more protein-inhibitor complexes were examined, the relationship between θ_I and inhibitor potency would become clearer. θ_I alone was unable to identify a single correct inhibitor out of a putatively inactive test set. This result is not unexpected because the docking programs showed only mediocre performance in identifying PPI inhibitors. These results do not preclude that θ_I could function as an

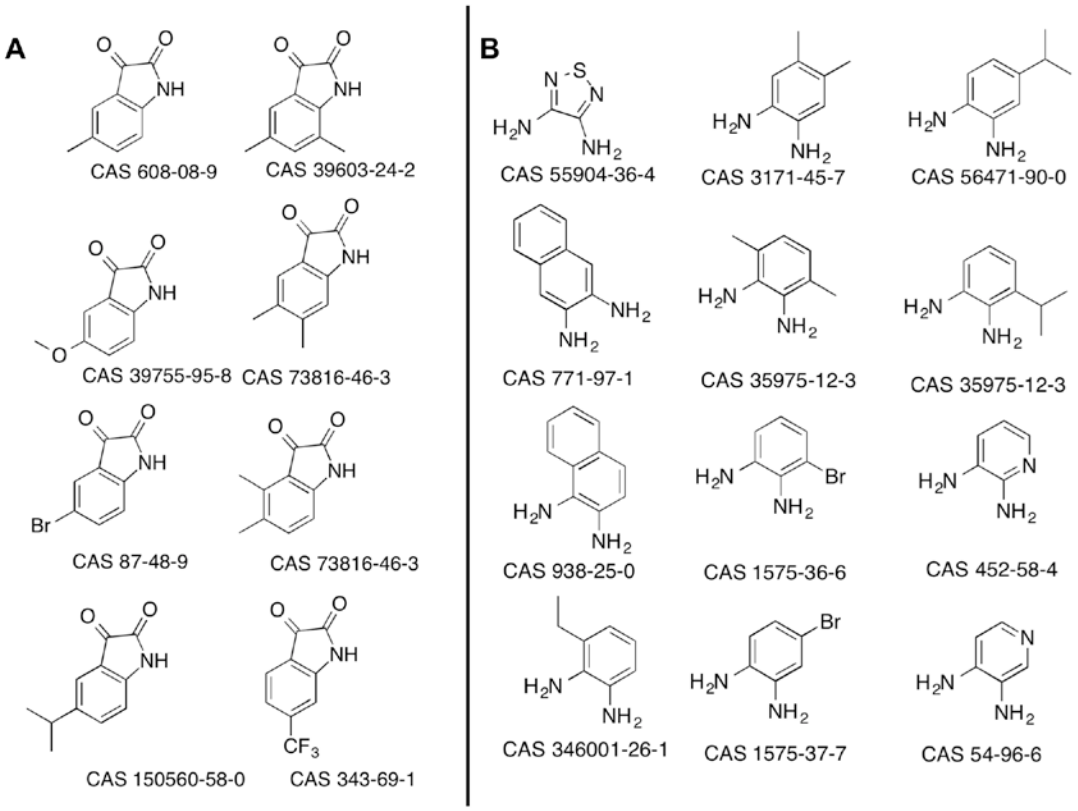


Figure 5.1. Potential substitutions based on available isatin (A) and o-phenylenediamine (B) starting materials.

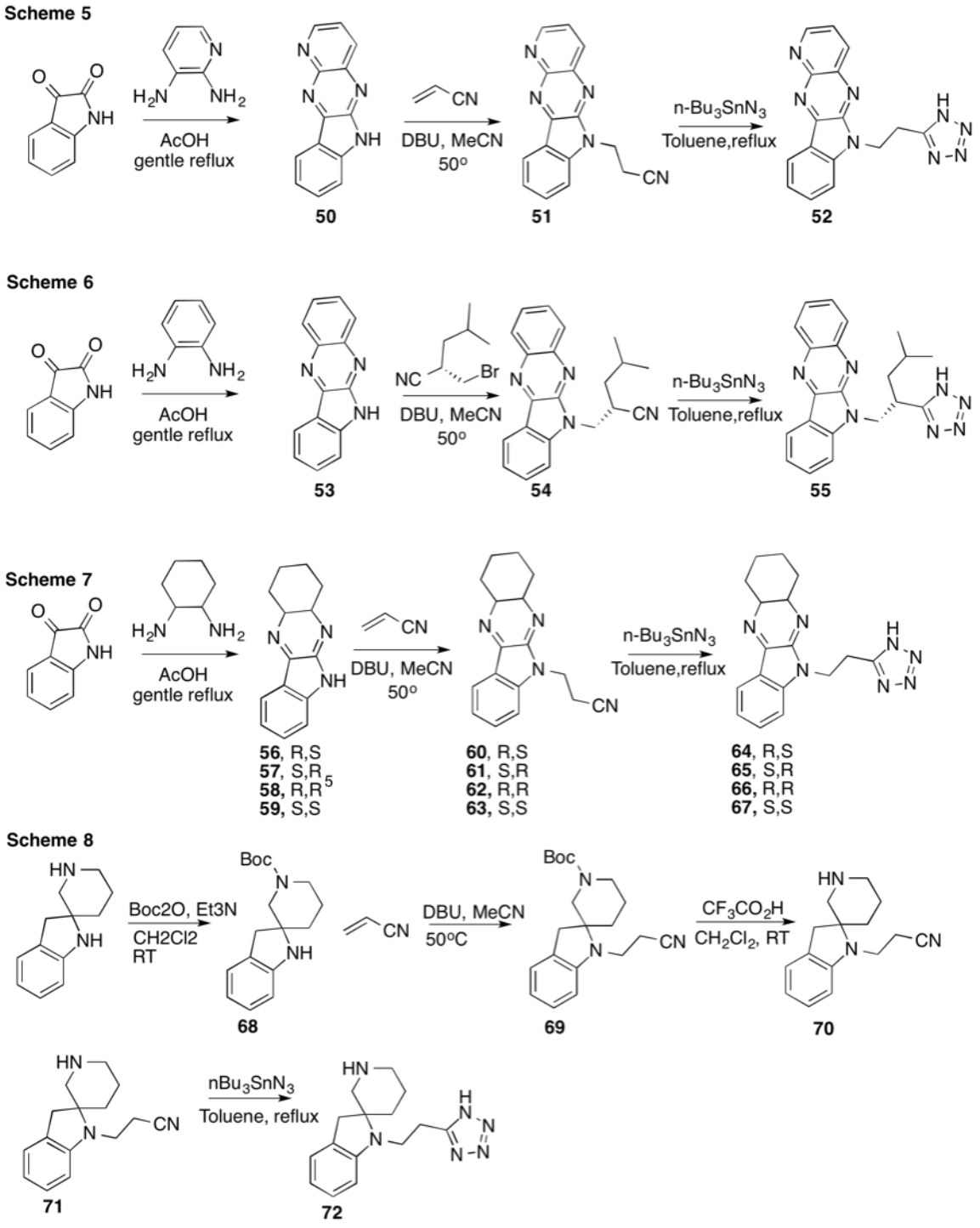


Figure 5.2. Schemes 5–8.

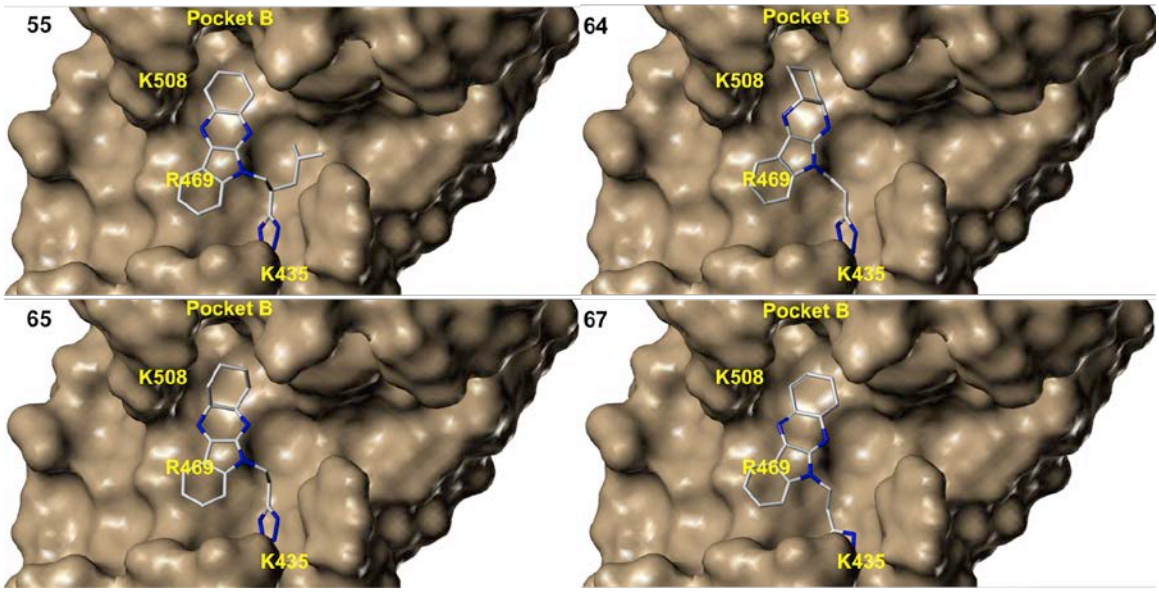


Figure 5.3. Predicted binding modes of proposed compounds. Because of the substitution of the benzene ring for a hexane ring, the binding mode of the inhibitor is reversed so that the aromatic ring interacts with R469.

ancillary property when selecting HTVS results to test. However, docking score, pharmacophore fit, and hot spot interactions in particular should remain the primary criteria for inhibitor selection. Therefore, to validate θ_l as an important property for HTVS of PPIs, there is a need to expand the limited samples that were used to establish its importance and apply θ_l to a novel HTVS.

The first objective is to review more protein-inhibitor complexes to strengthen the relationship between θ_l and potency. The first step would be to evaluate other MDM2 inhibitors, of which there are many with no crystal or NMR structures. These would need to be treated like the Bcl-X_L inhibitors, which were first docked, then evaluated. The second objective is to look for other PPI inhibitor crystal structures with large data sets. Examples include calmodulin, a messenger protein,[6] and BRD4, a bromodomain protein.[7] By examining inhibitors of these proteins, the relationship between θ_l and potency could be strengthened.

The second direction of the study would be developing a practical use for θ_l for improvement of HTVS of PPIs. This would be explored by incorporating it into a novel HTVS of a PPI. This would be accomplished by using the Bcl9 binding site on β-catenin. The Bcl9 binding site is shallow and lacks many deep pockets, which makes it a typical PPI binding surface. Drug-like, PAINs-free[8] ligands from the ZINC library[9] would be docked into the Bcl9 binding site using AutoDock Vina[10] and Schrödinger Glide.[11-13] A pharmacophore model would be created from the inhibitor 3',4'-difluoro-*N*-(4'-fluoro-2-(((*R*)-pyrrolidin-3-yl)oxy)-[1,1'-biphenyl]-4-yl)-6-(((*S*)-pyrrolidin-3-yl)oxy)-[1,1'-biphenyl]-3-carboxamide[14] developed by Logan Hoggard et al.[14] This inhibitor was designed to mimic the *i, i+3* and *i+7* interactions of the projecting hot spots of the Bcl9

helix-2.[14] This compound had a K_i of 2.1 $\pm$ 0.4 μM and was selective for the β-catenin/Bcl9 interaction over the β-catenin/E-cadherin PPI.[14] Two test sets of highly scored docked ligands that would be a good match to the pharmacophore model would have their solvent-exposed surface areas analyzed. These values would aid in the selection of molecules to test. Two separate test sets of significant size would be evaluated. One set would be picked with the aid of θ_l and one set would be picked without the aid of θ_l. It is hypothesized that there will be more effective inhibitors in the former set and fewer effective inhibitors in the latter set. The results of this would help further the understanding of the practical applications and limits of using solvent-exposed surface areas as a means to evaluate potential PPI inhibitors, as well as yield a new inhibitor of the β-catenin/Bcl9 PPI.

5.3 Further validation of potential allostery of the W504 pocket

ET analysis[15] identified the region surrounding β-catenin W504 as a functionally important residue. This is unusual because it had no contacts with any known β-catenin binding partner. Analysis of the β-catenin MD simulation with FTMap[16] revealed potential druggability of this site. Mutations of W504 produced an increased binding of β-catenin to Tcf, indicating potential allosteric communication between the W504 site and nearby hot region 1.

While evidence supported that the W504 pocket may be able to allosterically communicate with hot region 1, the evidence is still preliminary. Therefore, the next step would be to carefully evaluate that the site is indeed allosterically connected to hot region 1. The discovery of a small molecule capable of binding to the W504 binding site would

greatly enhance the significance of the W504 site. The small molecule may have a stronger effect on Tcf binding than the single point mutants of W504. This site would be an ideal target for an ensemble docking-based HTVS because many of the druggable conformations were discovered by an MD simulation. The MD simulation snapshots would be used generate the ensemble of structures for docking. The inclusion of receptor flexibility would allow for a conformation that can maximize interactions,[17] and would reduce the entropic penalty of binding by "softening" the protein surface.[17] Also, the MD simulation would allow for receptor conformations that exist only briefly to be sampled for docking.[17] The surface characteristics of the W504 pocket are aligned with those of a typical PPI surface. Therefore, this study would present an opportunity to apply the lessons learned in Chapter 3. Any inhibitor found would need to be well scrutinized to ensure binding to the W504 pocket and no other regions of β-catenin. Once a potent allosteric inhibitor is discovered, further study must be made into the mechanism of allosteric communication between the W504 region and hot region 1.

The mechanism of allosteric communication between the W504 pocket and hot region 1 is unknown. One possibility would be that W504 mutations affect hot region 1 through helix 8 of β-catenin. W504 is positioned at the end of this helix, and the disruption of these residues might be conveyed through the helix to key residue K508. The physical effects of the W504 mutation on Tcf binding would be examined through NMR[18] and MD simulations.[19] Understanding the mechanism of allostery would aid in designing small-molecule allosteric modulators that can inhibit, instead of enhance, Tcf binding.

β-catenin has over 100 direct interacting partners[20]. Therefore, quantitative

affinity pull-down mass-spectrometry (q-AP-MS) or “shotgun” proteomics would be a good method to characterize the potential allosteric or orthosteric effect of the W504 mutants at the same time (Figure 5.4).[21,22] An F9 β-catenin null cell line will be used to generate comparative cell lysates of both WT *β*-catenin and W504A mutants. The cell lysates will be digested and purified using HPLC and analyzed through mass spectrometry. The mass spectrometry data can be analyzed computationally to give quantitative information about changes in bound *β*-catenin populations. This experiment will yield data for any protein suspected of binding *β*-catenin that is affected by the W504 mutation.

5.4 Conclusion

In summary, the research presented here will help further the development of PPI inhibitors. These PPI inhibitors could act as small-molecule therapeutics or as probes which could aid in further discoveries about the natures of PPIs. The PPI interface is large and featureless which is problematic for small-molecule inhibitors[23]. As a result, multiple methods have been developed to aid in drugging the undruggable. Chief among these is the use of hot spots.[24]

One example of a PPI is the β-catenin/Tcf interaction. The β-catenin/Tcf PPI is a key mediator of the canonical Wnt signaling pathway.[25] This research presents a new series of small-molecule inhibitors that selectively inhibit the β-catenin/Tcf PPI. These compounds utilized the hot spots K435 and K508 to selectively disrupt the β-catenin/Tcf PPI. These compounds showed good results in vitro and in vivo. Further optimization of this compound requires consideration of the drug-like properties of this molecule.

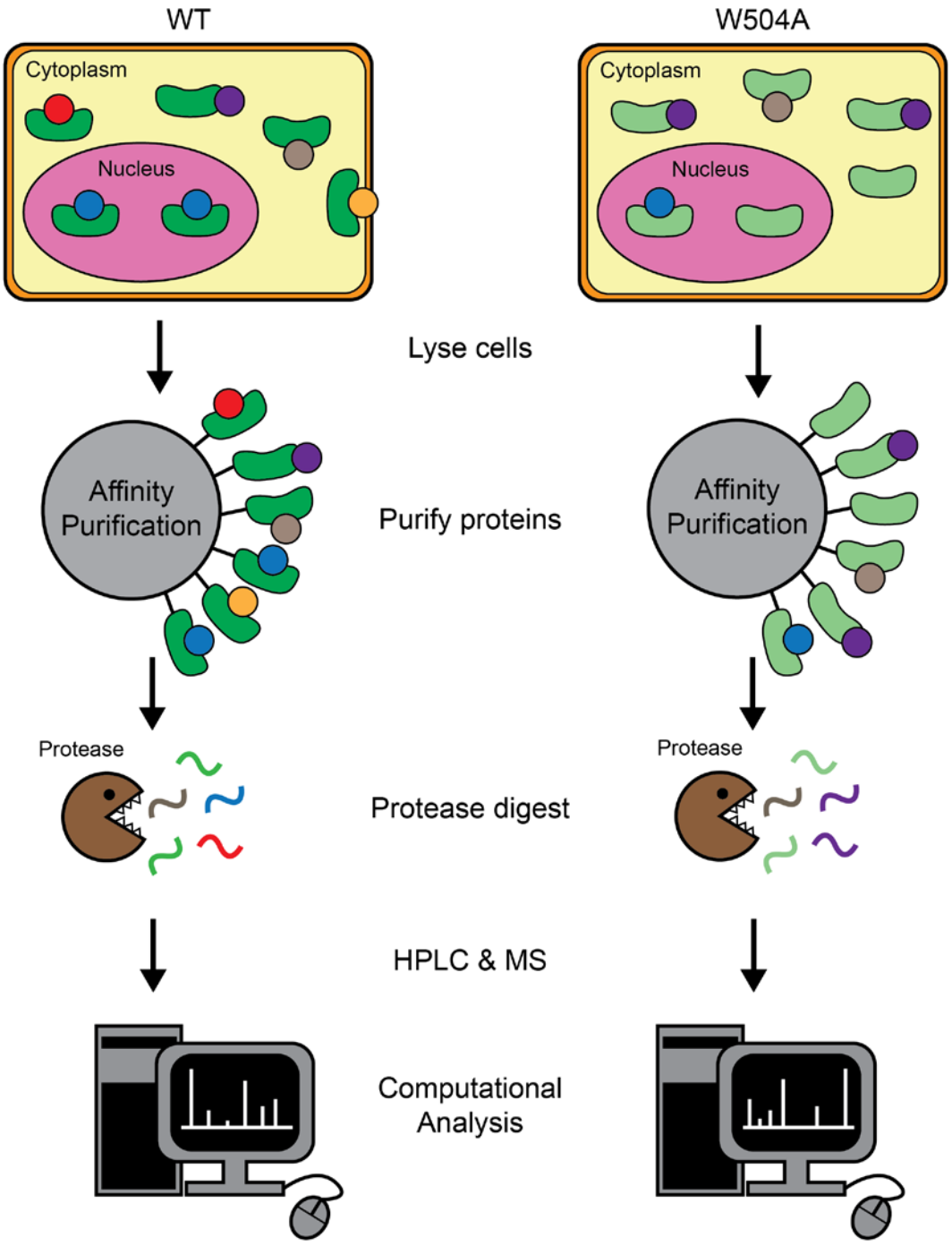


Figure 5.4. Diagram of q-AP-MS proteomics for β-catenin . Figure derived from methods described by Dunham et al.[22]

The research presented here examined the relationship of SASA and suitability of common docking algorithms for PPIs. The property, θ_I, was first discussed by Gowthaman et al.[26] This property was further explored to find a small, yet significant, correlation between θ_I and inhibitor potency. Because hot spots are often buried in the PPI interface, it was hypothesized a more buried ligand should have a better chance at interacting with hot spots. This was indeed the case, but θ_I was still overshadowed by other, stronger, nonbonding interactions.

Lastly, the origins of β-catenin binding partner selectivity and a potential for latent allostery in the ARM domain of β-catenin was explored through ET analysis. This analysis revealed the region around W504 as a potential allosteric site. This potential for allosteric regulation needs to be carefully validated before it may prove of use.

5.5 References

1. Catrow, J. L.; Zhang, Y.; Zhang, M.; Ji, H. Discovery of Selective Small-Molecule Inhibitors for the β-Catenin/T-Cell Factor Protein–Protein Interaction Through the Optimization of the Acyl Hydrazone Moiety. *J. Med. Chem.* **2015**, *58* (11), 4678–4692.

2. Shinya, M.; Eschbach, C.; Clark, M.; Lehrach, H.; Furutani-Seiki, M. Zebrafish Dkk1, Induced by the Pre-MBT Wnt Signaling, is Secreted from the Prechordal Plate and Patterns the Anterior Neural Plate. *Mech. Dev.* **2000,** *98* (1-2), 3–17.

3. Dorsky, R. I.; Sheldahl, L. C.; Moon, R. T. A Transgenic Lef1/β-Catenin-Dependent Reporter is Expressed in Spatially Restricted Domains Throughout Zebrafish Development. *Dev. Biol.* **2002,** *241* (2), 229–237.

4. Basham, K. J.; Kieffer, C.; Shelton, D. N.; Leonard, C. J.; Bhonde, V. R.; Vankayalapati, H.; Milash, B.; Bearss, D. J.; Looper, R. E.; Welm, B. E. Chemical Genetic Screen Reveals a Role for Desmosomal Adhesion in Mammary Branching Morphogenesis. *J. Biol. Chem.* **2013,** *288* (4), 2261–70.

5. Nami, N.; Hosseinzadeh, M.; Nami, N.; Haghdadi, M. Synthesis of Substituted pyrazino[5,6-b]pyrimidine and Some Indole Derivatives. *Phosp., Sulf. Sili. Rel.*

Elem. **2009,** 184 (11), 2846–55.

6. Stevens, F. C. Calmodulin: an Introduction. *Can. J. Biochem. Cell. Biol.* **1983,** *61* (8), 906-10.

7. Dey, A.; Ellenberg, J.; Farina, A.; Coleman, A. E.; Maruyama, T.; Sciortino, S.; Lippincott-Schwartz, J.; Ozato, K. A Bromodomain Protein, MCAP, Associates with Mitotic Chromosomes and Affects G(2)-to-M Transition. *Mol. Cell. Biol.* **2000,** *20* (17), 6537–49.

8. Baell, J. B.; Holloway, G. A. New Substructure Filters for Removal of Pan Assay Interference Compounds (PAINS) from Screening Libraries and for their Exclusion in Bioassays. *J. Med. Chem.* **2010,** *53* (7), 2719–40.

9. Irwin, J. J.; Shoichet, B. K. ZINC – A Free Database of Commercially Available Compounds for Virtual Screening. *J. Chem. Inf. Model.* **2005,** *45* (1), 177–82.

10. Trott, O.; Olson, A. J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization and Multithreading. *J. Comput. Chem.* **2010,** *31* (2), 455–61.

11. Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren,T. A.; Sanschagrin, P. C.; Mainz, D. T. Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein-Ligand Complexes. *J. Med. Chem.* **2006**, *49*, 6177–6196

12. Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening, *J. Med. Chem.* **2004**, *47*, 1750–1759

13. Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shaw, D. E.; Shelley, M.; Perry, J. K.; Francis, P.; Shenkin, P. S. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy, *J. Med. Chem.* **2004**, *47*, 1739–1749

14. Hoggard, L. R.; Zhang, Y.; Zhang, M.; Panic, V.; Wisniewski, J. A.; Ji, H. Rational Design of Selective Small-Molecule Inhibitors for β-catenin/B-cell Lymphoma 9 Protein–Protein Interactions. *J. Am. Chem. Soc.* **2015,** *137* (38), 12249–60.

15. Wilkins, A. D.; Lua, R.; Erdin, S.; Ward, R. M.; and Lichtarge, O. Sequence and Structure Continuity of Evolutionary Importance Improves Protein Functional Site Discovery And Annotation. *Protein Sci*. **2010**, *19*, 1296–1311.

16. Kozakov, D.; Grove, L. E.; Hall, D. R.; Bohnuud, T.; Mottarella, S. E.; Luo, L.; Xia,

B.; Beglov, D.; Vajda, S. The FTMap Family of Web Servers for Determining and Characterizing Ligand-binding Hot Spots of Proteins. *Nat. Protoc.* **2015,** *10* (5), 733–755.

17. Lexa, K. W.; Carlson, H. A. Protein Flexibility in Docking and Surface Mapping. *Q. Rev. Biophys.* **2012,** *45* (3), 301–343.

18. Manley, G.; Loria, J. P. NMR Insights into Protein Allostery. *Arch. Biochem. Biophys.* **2012,** *519* (2), 223–231.

19. Hertig, S.; Latorraca, N. R.; Dror, R. O. Revealing Atomic-Level Mechanisms of Protein Allostery with Molecular Dynamics Simulations. *PLoS Comput. Biol.* **2016,** *12* (6), e1004746.

20. Peri S, Navarro JD, Amanchy R, Kristiansen TZ, Jonnalagadda CK,Surendranath V, Niranjan V, Muthusamy B, Gandhi TK, Gronborg M, Ibarrola N, Deshpande N, Shanker K, Shivashankar HN, Rashmi BP, Ramya MA, Zhao Z, Chandrika KN, Padma N, Harsha HC, Yatish AJ, Kavitha MP, Menezes M, Choudhury DR, Suresh S, Ghosh N, Saravana R, Chandran S, Krishna S, Joy M, Anand SK, Madavan V, Joseph A, Wong GW, Schiemann WP, Constantinescu SN, Huang L, Khosravi-Far R, Steen H, Tewari M, Ghaffari S, Blobe GC, Dang CV, Garcia JG, Pevsner J, Jensen ON, Roepstorff P, Deshpande KS, Chinnaiyan AM, Hamosh A, Chakravarti A, Pandey A. Development of Human Protein Reference Database as an Initial Platform for Approaching Systems Biology in Humans. *Genome Research*. **2003**. *13*, 2363–2371.

21. Armean, I. M.; Lilley K. S.; Trotter M. W. Popular Computational Methods to Assess Multiprotein Complexes Derived from Label-free Affinity Purification and Mass Spectrometry (AP-MS) Experiments. *Mol. Cell Proteomics*. **2013**. *12* (1), 1–13.

22. Dunham W. H.,; Mullin M.; Gingras A.C. Affinity-Purification Coupled to Mass Spectrometry: Basic Principles and Strategies. *Proteomics*. **2012**. *12* (10), 1576–90.

23. Wells, J. A.; McClendon, C. L. Reaching for High-Hanging Fruit in Drug Discovery at Protein-protein Interfaces. *Nature* **2007,** *450* (7172), 1001–1009.

24. Clackson, T.; Wells, J. A. A Hot Spot of Binding Energy in a Hormone-Receptor Interface. *Science* **1995,** *267* (5196), 383–386.

25. Clevers, H.; Nusse, R. Wnt/beta-catenin Signaling and Disease. *Cell* **2012,** *149* (6), 1192–1205.

26. Gowthaman, R.; Deeds, E. J.; Karanicolas, J. Structural Properties of Non-Traditional Drug Targets Present New Challenges for Virtual Screening. *J. Chem. Inf. Model.* **2013,** *53* (8), 2073–2081.