Title
Structure, Function, and Druggability of SLC Transporters and Kinases

Author
Stecula, Adrian

Publication Date
2016

Peer reviewed|Thesis/dissertation

Structure, Function, and Druggability of SLC Transporters and Kinases

by

Adrian Wladyslaw Stecula

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Pharmaceutical Sciences and Pharmacogenomics

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

"We have no need of other worlds. We need mirrors.
We don't know what to do with other worlds.
A single world, our own, suffices us;
but we can't accept it for what it is."

Stanisław Lem, Solaris

Acknowledgements

This work would not have been possible without the support from my advisors, Andrej Sali and Kathy Giacomini.

I would like to thank Andrej for taking a chance on a pharmaceutical science student with a background in chemical engineering. With his unyielding sense of reason, he has a way of introducing logic into the illogical. He taught me how to think critically about complex biological systems and how to manage both time and people. My time in his lab has been defined by two principles. 1) Instead of competing, collaborate. 2) Life is nothing but a modeling problem.

I would like to thank Kathy for taking me under her wing from the very beginning of my time at UCSF. Her unique mentorship style allowed me to be very independent. She enabled me to ask and pursue my own questions, and to develop new methods to be used by others in the future. As a pioneer in the field of transporters, she taught me how to pave my own path, and how to overcome the fear of asking the most basic questions.

In addition to my advisors, I was fortunate to interact with several other PIs and members of their labs. First, I would like to thank Greg Amidon, my academic advisor at the University of Michigan, for encouraging me to pursue a graduate degree. I want to thank Robert Stroud for his guidance on the intricacies of membrane transporter expression and purification. I also want to thank members of the Stroud lab, particularly Thomas and Yaneth, who made even the most discouraging purification trials very enjoyable. I want to thank Yifan Cheng for creating a state-of-the-art electron microscopy core, and Jean-Paul for staying late on countless occasions to collect data for me. I would also like to acknowledge Francis Brodsky. I would not have come to UCSF if it weren't for her persuasive lecture on clathrin and her truly charismatic personality. Although we didn't meet until recently, I would also like to thank James Fraser and Trever Bivona for bringing me onboard for a very ambitious project exploring the cancer proteome.

The completion of this PhD was made possible by the inclusive environment created by the members of both the Sali and Giacomini labs. In the Sali lab, I would like to thank my first and most important mentor, Avner. His work on modeling and virtual screening of the SLC transporters created a framework for my studies. I also want to thank Charles and Peter, the two pillars of the Sali lab, for scientific discussions, Taco Bell lunches, whiskey club, and inventing "mushroom" projects. It always amazed me just how productive our conversations were. I would like to thank Ilan, Patrick, Daniel, Barak, Ignacia, Hao and Kate for the everyday discussions and help in accessing the ever-changing lab resources. In the Giacomini lab, I would like to thank Sook Wah. She is the kindest, most resourceful person I have ever met. I also want to thank Ethan for providing me with guidance during my first few months in the lab. I would like to acknowledge Sirj, James, Ling, and Arik for the frequent discussions while troubleshooting the experiments.

I owe a huge debt of gratitude to my classmates Aishwarya, Megan, and Xiaomin. Although small, our cohort has made significant scientific progress in the fields of SLC/ABC transporters and HIV latency. I would like to thank them for the constant stream of support and frequent conversations about our failed experiments/dreams.

My journey in higher education began in a very different time and place. At every step along the educational ladder, I've been fortunate to meet incredible people. To start, I'm very thankful to have met my friend Neil on the first day of school in America. Whether it's running through the woods, falling from a plane, or playing poker, whatever life brings my way, he has always been there for me. My days at the University of Michigan are synonymous with Patryk and Michael. I still think about the late nights at "the Dude," eating Panda, or keeping each other awake in quantum chemistry. The college experience also would not be complete without Kuba and Wojtek, my "other" older brothers. It's incredible that there was a time when all of us attended UM Dearborn together.

Over the past five years, my personal life and my work life have become mutually informative. I would not have been able to get through this period, especially 2016, without the help of my friends here in San Francisco. First and foremost, I would

like to thank Joanna and Frank for the easy days and the wild nights, the Tahoe trips, the beach fires, the concerts, and the Yellow Submarines. They have provided me with a home (in every sense of the word) when I needed it most. I would not be where I am today if it wasn't for their moral and financial support. I would also like to thank my friend Jorge. From our first rendezvous by Mission Creek, through the frequent Lucky 13 outings, to his days at UCLA, he was always someone I could depend on.

Next, I want to thank Ryan, the one person to truly define my time at UCSF, for the countless hours discussing science, family, and the future. Whether it's orange chicken dinners, rollercoaster trips, Utah Fridays, or exploring niche music genres, I have grown to consider him as an integral part of my graduate career and a member of my family.

Lastly, I have to thank my family. I would like to thank Ania and Jarek for opening up their home and providing a foundation for a new life in America. The impact of their generosity over the past two decades cannot be overstated. I want to thank my sister-in-law Katie, her parents, Steve and Cindy, and her brothers, Steve and Nick. Katie, thank you for your advice during the most difficult of times. You have an extraordinary gift of laughter. I want to thank my grandparents, Józef and Maria, for doing everything in their power and more to help me succeed. Thank you for playing an active role in my daily life, no matter how many oceans separate us. Finally, I want to thank my mom, Dorota, and my brother, Dominik. We have come so far with so little. Mom, thank you for creating a home where education was valued. Thank you for driving us to school sometimes 8 times a day to make sure we made our practices, clubs, and societies. Thank you for enabling us to have dreams that were much bigger than our past or our present, completely defying what was possible for a first generation single-parent immigrant family. Dominik, thank you for always being by my side and for paving the way forward. I wouldn't be here without you or your bad jokes.

Structure, Function, and Druggability of SLC Transporters and Kinases

Adrian Stecula

Abstract

All functions of a protein involve its physical interactions, however transient, with specific other molecules (ligands), large or small. Protein structure and its dynamics, determined by sequence, in turn determine which molecules are able to bind to the protein. Druggable proteins are defined as those whose function can be modulated with a small molecule. The contents of this dissertation focus on leveraging the knowledge about protein structure and changes in protein structure to identify small molecule modulators of protein function, thus expanding the druggable proteome. The studied proteins include members of the kinase and SLC transporter superfamilies, both important drug targets. First, we separately examine the impact of mutations on the structures of two kinases, FLT3 and BCR-ABL, and the resistance of these mutants to drugs. We provide a structure-based rationale for the cause of resistance and offer treatment alternatives. Second, we address the modulation of function of human organic cation transporters (OCTs), either through phosphorylation or structure-guided screens to discover novel small molecule inhibitors of these transporters. We establish that by combining docking and *in vitro* high-throughput screens, competitive and non-competitive ligands of OCTs can be predicted accurately. Third, we examine the quaternary structure of the human concentrative nucleoside transporters (CNTs) to gain new insight into their functions and use structure-guided screens to discover novel ligands to modulate them. We show that human concentrative nucleoside transporter 3 forms homo-oligomers, thus encouraging efforts on finding allosteric inhibitors. Finally, we present a large-scale study of the impact of cancer mutations on protein structure, with the hope of expanding the druggable proteome through the discovery of mutant-specific binding pockets that would allow for selective, functional inhibition or activation.

Table of Contents

List of Tables

List of Figures

List of Figures (continued)

List of Figures (continued)

List of Figures (continued)

Introduction

Protein structure depends on its sequence

Proteins are molecular workhorses of the cell. A protein consists of a linear chain of amino acid residues, whose sequence determines its three-dimensional structure and dynamics, under given conditions.[1] The 20 amino acid residue types form proteins with properties that are far more complex than the sum of their parts.[2] Many proteins spontaneously fold into their unique native structures, while others require molecular chaperones,[3,4] but all are the embodiment of the transition between 1-dimensional gene sequences storing the heritable makeup of the cell and the 3-dimensional world of life.[5]

The field of structural biology is concerned with the identification and understanding of structure and function of macromolecules on the molecular level.[6,7] The history of structural biology begins with discoveries by Linus Pauling. Several years before the determination of the first high-resolution protein structure (of myoglobin) by John Kendrew,[8] Pauling working with Robert Corey and Herman Branson, laid the foundation by proposing regularly repeating structural elements, now called the α-helix and the β-sheet, to serve as the backbone of protein secondary structure.[9,10] It was a landmark study because it implied that the conformation of the polypeptide could be accurately predicted if the conformational preferences of its component residues were precisely known.

A major theme in structural biology is how structure determines function. Proteins are not rigid structures, but instead adopt a variety of conformations under native conditions.[11] The (free) energy landscape defines all possible protein conformations and their relative concentrations.[12] Transitions between conformations are often associated with common protein functions, such as signaling, transport, chemical reaction catalysis, regulation, cell movement, and cytoskeleton assembly.[6]

All functions of a protein involve its physical interactions, however transient, with specific other molecules (ligands), large or small. The formation of a protein-ligand complex depends on multiple bonds, including specific weak non-covalent interactions (e.g., electrostatic interactions, van der Waals interactions, π effect), the non-specific hydrophobic effect, and sometimes covalent interactions.[13-16] The binding thus requires shape complementarity between the target (the protein) and the ligand.[17,18] Structure and its dynamics, determined by sequence, therefore in turn determine which molecules are able to bind to the protein.[6,19]

Changes in protein structure lead to changes in protein function

Protein structure can be affected in numerous ways. Temperature, pH, post-translational modifications, mutations in amino acid sequence, and presence of small molecules such as urea[20] are all common factors that affect protein folding.[21] Any change in the protein structure can have a significant effect on its function.[6]

Cells have evolved regulatory mechanisms that utilize this principle. For example, phosphorylation, the most ubiquitous of over 200 identified post-translational modifications,[22] is a covalent, reversible, kinase-mediated transfer of a negatively charged phosphate group onto serine, threonine, and tyrosine amino acid residues.[23] Protein phosphorylation can result in either activation or inhibition of protein activity. Phosphorylation can activate protein activity through allosteric conformational changes, as seen in glycogen phosphorylase,[23] or create a recognition site for other proteins, such as SH2 domains of kinases such as ABL1.[24] Phosphorylation can inhibit protein activity either by acting as a steric block, as seen in isocitrate dehydrogenase, or by impeding substrate recognition, as seen in CDK2. Phosphorylation can also cause order-to-disorder transitions, as seen in K^+ channel inactivation domain, or disorder-to-order transitions, as seen in the CBP/CREB binding.[24]

Amino acid mutations are the most direct way of altering the protein structure and dynamics. In the post-genome era, we have been able to sequence genomes to discover even rare inter-individual mutations.[25-27] Mutations play a fundamental role in evolution, because they introduce variation into genomes, thus contributing to phenotypic variation in individual's characteristics, including risk of disease.[25] The analysis of this data has led to the identification of several mutation types, including those affecting the coding regions of the genome. While synonymous mutations (no change in the coding amino acid residue) can affect protein function through changes in transcription, splicing, mRNA transport or translation,[28] non-synonymous mutations alter the protein sequence, thus affecting structure and possibly function.

Non-synonymous mutations have been linked to the pathogenesis of diseases such as sickle cell anemia, which is caused by a hemoglobin β-chain E6V mutation.[29] The mutation of a polar, hydrophilic residue to a nonpolar, hydrophobic residue leads to an association of hydrophobic regions of hemoglobin molecules in the cytoplasm, resulting in hemoglobin polymerization, and ultimately generation of hemoglobin fibers that damage the membrane and cytoskeleton of red blood cells.[30] Non-synonymous mutations in proteins have also been shown to act as "driver mutations," mutations that confer a growth advantage, in cancer.[31] The most common mutations, occurring in almost every type of cancer at rates from 38% – 50%, involve the inactivation of p53, a tumor suppressor.[32] One such mutation is the Y220C mutation, which creates a surface cavity that destabilizes p53.[33] As seen in these examples, even a single amino acid residue substitution can have profound functional consequences.

Proteins as targets of small molecule modulators

Rational ligand design involves the design or search for small molecules that are complementary to a binding site on the surface of the target, therefore allowing for the binding of the molecule.[34-36] Drugs are a subset of protein ligands that elicit a desired pharmacological effect.[34,37]

Druggability is defined as "the likelihood of being able to modulate a target with a small molecule."[38] Protein druggability depends on the presence of a binding site for a small molecule to bind to with high affinity and selectivity. Since our methods for binding site identification and prediction change over time, so does druggability.[39] Thus, druggability is both a property of the protein and the current state of knowledge. Most successful drugs achieve their activity by competing for a binding site on a protein with an endogenous small molecule. Current estimates place the number of druggable proteins at approximately 3,000.[40]

Druggability is not the only factor determining whether or not a protein becomes a drug target. The other major consideration is of course the protein's link to disease. Current estimates propose that there are approximately 3,000 disease-modifying genes. The overlap between druggable and disease-modifying proteins is the current number of drug targets (approximately 1,500).[40] While it might be difficult to increase the number of disease-modifying genes without new biological insights, recent efforts have expanded the number of druggable proteins and protein families. Historically, only the "active" or "orthosteric" sites have been considered as targets of drug binding.[34] However, consideration of allosteric sites,[41] cryptic sites,[42] and covalent drugs that might bind to pockets with weak affinities[43,44] significantly increases the number of druggable proteins.

Protein druggability can be affected by changes in the protein structure either positively or negatively. In the negative case, mutations lead to structural changes that prevent drugs from binding and eliciting their pharmacological response. In the positive case, the structural changes create mutant-specific binding sites, which allow for mutant-specific modulation. Precision medicine is a term coined for the synthesis of established clinical protocols with molecular profiling to create patient-specific diagnostic, prognostic, and therapeutic strategies.[45,46] Thus, because changes in protein structure caused by mutations and/or inter-individual genetic differences have the potential to affect drug response, this area of research is critical to the advancement of precision medicine.

The overall goal of this dissertation research was to use structural information to find novel ways of modulating protein function and to ultimately expand the druggable human proteome. The following are abstracts of the chapters in this dissertation. Each chapter focuses on leveraging knowledge about protein structure and changes in protein structure to identify small molecule ligands for modulating protein function. The studied proteins include members of the kinase and SLC transporter superfamilies, both important drug targets.

Chapter abstracts

In Chapter 1, we separately examine the impact of mutations on the structure of two kinases and the resistance of these mutants to drugs.

First, we characterized acquired kinase domain (KD) point mutations in BCR-ABL which frequently cause resistance to ABL tyrosine kinase inhibitors (TKIs) in chronic myeloid leukemia (CML). BCR-ABL T315I is the most problematic mutation being highly resistant to all approved ABL TKIs except ponatinib. Ponatinib inhibits all single nucleotide substitution KD mutants but is vulnerable to "compound" mutations. We profiled the seven clinically reported BCR-ABL T315 mutations and determined that three mutations were pan-resistant to all approved ABL TKIs (T315F/L/V). T315L confers clinical and *in vitro* resistance to ponatinib but is sensitive *in vitro* to the approved TKI axitinib. To prospectively guide clinical decision-making for novel T315 mutants, we created all 19 substitutions at the T315 residue and found that nearly all confer moderate to high resistance to imatinib, dasatinib, nilotinib, and bosutinib. Importantly, seven T315 mutations conferred less relative resistance to axitinib than to ponatinib *in vitro*. Further, we provided a structure-based rationale for the impact of clinically observed mutations that allowed us to predict the impact of uncharacterized mutations.

Second, we focused on activating mutations in the FLT3 kinase, which occur in approximately 30% of adult acute myeloid leukemia instances, primarily consisting of internal tandem duplication (ITD) mutations (~25%) and point mutations in the tyrosine kinase domain (~5%), commonly at the activation loop residue D835. Secondary kinase domain mutations in FLT3-ITD, particularly at the D835 residue, are frequently associated with acquired clinical resistance to effective FLT3 tyrosine kinase inhibitors (TKIs). Molecular docking studies have suggested that D835 mutations primarily confer resistance by stabilizing an active Asp-Phe-Gly in ('DFG-in') kinase conformation

unfavorable to the binding of type II FLT3 TKIs, which target a 'DFG-out' inactive conformation. We profiled the activity of active type II FLT3 TKIs against D835 kinase domain mutants that have been clinically detected to date. We found that type II inhibitors (quizartinib, sorafenib, ponatinib, and PLX3397) retain activity against specific D835 substitutions. Modeling studies suggest that bulky hydrophobic substitutions (D835Y/ V/I/F) at this residue are particularly resistant, whereas mutations that preserve interactions between D835 and S838 are relatively sensitive (D835E/N).

In Chapter 2, we address the modulation of function of human organic cation transporters (OCTs), either through phosphorylation or structure-guided screens to discover novel small molecule inhibitors of the transporter.

First, we report a unique phosphorylation-dependent link between drug transporters and tyrosine kinase inhibitors (TKIs), which has uncovered widespread phosphotyrosine-mediated regulation of drug transporters. We initially found that OCTs, uptake carriers of metformin and oxaliplatin, were inhibited by several clinically used TKIs. Mechanistic studies showed that these TKIs inhibit the Src family kinase Yes1, which was found to be essential for OCT2 tyrosine phosphorylation and function. Yes1 inhibition *in vivo* diminished OCT2 activity, significantly mitigating oxaliplatin-induced acute sensory neuropathy. Along with OCT2, other SLC-family drug transporters are potentially part of an extensive 'transporter-phosphoproteome' with unique susceptibility to TKIs. On the basis of these findings, we proposed that TKIs, a rapidly expanding class of therapeutics, can functionally modulate pharmacologically important proteins by inhibiting protein kinases essential for their post-translational regulation.

Second, we identified competitive and non-competitive OCT1 interacting ligands in a library of 1,780 prescription drugs by combining *in silico* and *in vitro* methods. Ligands were predicted by docking against a comparative model based a eukaryotic homolog. In parallel, high-throughput screening (HTS) was conducted using the fluorescent probe substrate ASP^+ in cells overexpressing human OCT1. Thirty OCT1 competitive ligands, defined as ligands predicted *in silico* as well as found by HTS, were identified. Finally, virtual screening of 29,332 metabolites predicted 146 competitive OCT1 ligands, of which an endogenous neurotoxin, 1-benzyl-1,2,3,4-tetrahydroisoquinoline, was experimentally validated. We establish that by combining docking and *in vitro* HTS, competitive and non-competitive ligands of OCT1 can be predicted.

In Chapter 3, we examine the quaternary structure of the human concentrative nucleoside transporters (CNTs) to gain new insight into their function and use structure-guided screens to discover novel ligands to modulate their function.

First, through a series of cysteine mutants at highly conserved positions guided by homology models, we cross-linked human CNT3 protomers in a cell-based assay, thus verifying the existence of hCNT3 homo-trimers in human cells. The presence and absence of cross-links at specific locations along TM9 informed us about structural differences between vcCNT and CNT3. Analysis of the trimerization domain using both structural modeling and sequence co-evolution analysis indicated that oligomerization is critical for the stability and function of hCNT3.

Second, we describe our multi-year effort to express and purify the human CNT3 for the purpose of structure determination with either X-ray crystallography or single-particle electron cryo-microscopy. An atomic or near-atomic resolution structure of the human CNT3 homo-trimer could provide us with insight into the dimerization interface for allosteric modulation, structural features governing substrate specificity, and conformational changes associated with the transport cycle. Despite a significant effort in which we tested several constructs, lipid extracts, detergent and thermostability screens, we were largely unsuccessful in obtaining even a low-resolution structure of the human transporter.

Third, we virtually screened hundreds of thousands of compounds, encompassing a large, never-before-tested chemical space, in search of novel hCNT3 ligands, both substrates and inhibitors. Following a virtual screen, a subset of compounds was tested in a cell-based assay. Ticagrelor, an FDA-approved platelet aggregation inhibitor, was identified as a novel and potent (IC_{50} of 6.47 ± 1.27 μM) inhibitor of hCNT3-mediated uridine uptake. This result serves as an encouraging first step in structure-based screening for small molecule modulators of this important transporter family.

Finally, Chapter 4 presents a large-scale study of the impact of cancer mutations on protein structure, with the hope of expanding the druggable proteome through the discovery of mutant-specific binding pockets that would allow for selective, functional inhibition or activation. The underlying hypothesis is that mutations in cancer genes alter the thermodynamic stability of the mutated protein, creating new cryptic small molecule binding pockets in otherwise undruggable

proteins and allowing for selective targeting of the mutant but not native form of the cancer-driving protein. Our approach seeks to re-prioritize critical cancer mutations for therapeutic targeting in patients, uncover the biochemical, biological, and functional effects of these therapeutically unused cancer mutations, and provide mutant-selective activators and inhibitors that would unlock the therapeutic potential of currently undruggable tumor suppressor. PTEN, a phosphatase frequently mutated in cancer, is examined as a case study.

References

1. Winnacker E-L, Huber R. *Protein Structure and Protein Engineering*. (Winnacker E-L, Huber R, eds.). Berlin, Heidelberg: Springer Science & Business Media; 1988.
2. National Research Council (US) Committee on Research Opportunities in Biology. Opportunities in Biology. 1989. doi:10.17226/742.
3. Ellis RJ, van der Vies SM. Molecular chaperones. *Annu Rev Biochem*. 1991;60:321-347. doi:10.1146/annurev.bi.60.070191.001541.
4. Laskey RA, Honda BM, Mills AD, Finch JT. Nucleosomes are assembled by an acidic protein which binds histones and transfers them to DNA. *Nature*. 1978;275(5679):416-420.
5. Berg JM, Tymoczko JL, Stryer L. *Biochemistry*. W. H. Freeman; 2002.
6. Alberts B. *Molecular Biology of the Cell*. Courier Corporation; 1989.
7. Astbury WT. Molecular biology or ultrastructural biology? *Nature*. 1961;190:1124.
8. Kendrew JC, Bodo G, Dintzis HM, Parrish RG. A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature*. 1958.
9. Pauling L, Corey RB, Branson HR. The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chain. *Proceedings of the National Academy of Sciences*. 1951;37(4):205-211. doi:10.1073/pnas.37.4.205.
10. Pauling L, Corey RB. The pleated sheet, a new layer configuration of polypeptide chains. *Proceedings of the National Academy of Sciences*. 1951;37(5):251-256.
11. Ansari A, Berendzen J, Bowne SF, et al. Protein states and proteinquakes. *Proceedings of the National Academy of Sciences*. 1985;82(15):5000-5004.
12. Onuchic JN, Wolynes PG, Luthey-Schulten Z, Socci ND. Toward an outline of the topography of a realistic protein-folding funnel. *Proceedings of the National Academy of Sciences*. 1995;92(8):3626-3630.
13. Spolar RS, Ha JH, Record MT. Hydrophobic effect in protein folding and other noncovalent processes involving proteins. *Proceedings of the National Academy of Sciences*. 1989;86(21):8382-8385.
14. Pauling L. Nature of forces between large molecules of biological interest. *Nature*. 1948;161(4097):707-709.
15. Wells JA, McClendon CL. Reaching for high-hanging fruit in drug discovery at protein–protein interfaces. *Nature*. 2007;450(7172):1001-1009.
16. Kessel A, Ben-Tal N. *Introduction to Proteins*. CRC Press; 2010.

References (continued)

17. Koshland DE. Application of a Theory of Enzyme Specificity to Protein Synthesis. *Proceedings of the National Academy of Sciences*. 1958;44(2):98-104.
18. Koshland DE. Correlation of Structure and Function in Enzyme Action. *Science*. 1963;142(3599):1533-1541.
19. Tsai CJ, Ma B, Nussinov R. Folding and binding cascades: shifts in energy landscapes. *Proceedings of the National Academy of Sciences*. 1999;96(18):9970-9972.
20. Bennion BJ, Daggett V. The molecular basis for the chemical denaturation of proteins by urea. *Proceedings of the National Academy of Sciences*. 2003;100(9):5142-5147. doi:10.1073/pnas.0930122100.
21. Anfinsen CB. The formation and stabilization of protein structure. *Biochem J*. 1972;128(4):737-749.
22. Walsh C. *Posttranslational Modification of Proteins*. Roberts and Company Publishers; 2006.
23. Johnson LN, Barford D. The effects of phosphorylation on the structure and function of proteins. *Annu Rev Biophys Biomol Struct*. 1993;22(1):199-232. doi:10.1146/annurev.bb.22.060193.001215.
24. Johnson LN. The regulation of protein phosphorylation. *Biochem Soc Trans*. 2009;37(Pt 4):627-641. doi:10.1042/BST0370627.
25. Sachidanandam R, Weissman D, Schmidt SC, et al. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature*. 2001;409(6822):928-933. doi:10.1038/35057149.
26. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature*. 2001;409(6822):860-921. doi:10.1038/35057062.
27. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature*. 2001;409(6822):860-921. doi:10.1038/35057062.
28. Goymer P. Synonymous mutations break their silence. *Nat Rev Genet*. 2007;8(2):92-92.
29. Ingram VM. A specific chemical difference between the globins of normal human and sickle-cell anaemia haemoglobin. *Nature*. 1956;178(4537):792-794.
30. Barabino GA, Platt MO, Kaul DK. Sickle cell biomechanics. *Annu Rev Biomed Eng*. 2010;12(1):345-367. doi:10.1146/annurev-bioeng-070909-105339.
31. Stratton MR, Campbell PJ, Futreal PA. The cancer genome. *Nature*. 2009;458(7239):719-724. doi:10.1038/nature07943.

References (continued)

32. Olivier M, Hollstein M, Hainaut P. TP53 mutations in human cancers: origins, consequences, and clinical use. *Cold Spring Harb Perspect Biol*. 2010;2(1):a001008-a001008. doi:10.1101/cshperspect.a001008.
33. Boeckler FM, Joerger AC, Jaggi G, Rutherford TJ, Veprintsev DB, Fersht AR. Targeted rescue of a destabilized mutant of p53 by an in silico screened drug. *Proc Natl Acad Sci USA*. 2008;105(30):10360-10365. doi:10.1073/pnas.0805326105.
34. Baron R, ed. *Computational Drug Discovery and Design*. Vol 819. New York, NY: Springer New York; 2012. doi:10.1007/978-1-61779-465-0.
35. Teague SJ. Implications of protein flexibility for drug discovery. *Nat Rev Drug Discov*. 2003;2(7):527-541. doi:10.1038/nrd1129.
36. Robertson JG. Mechanistic Basis of Enzyme-Targeted Drugs. *Biochemistry*. 2005;44(15):5561-5571. doi:10.1021/bi050247e.
37. Tozer TN, Rowland M. *Introduction to Pharmacokinetics and Pharmacodynamics*. Lippincott Williams & Wilkins; 2006.
38. Owens J. Determining druggability. *Nat Rev Drug Discov*. 2007;6(3):187-187. doi:10.1038/nrd2275.
39. Kozakov D, Hall DR, Napoleon RL, Yueh C, Whitty A, Vajda S. New Frontiers in Druggability. *J Med Chem*. 2015;58(23):9063-9088. doi:10.1021/acs.jmedchem.5b00586.
40. Hopkins AL, Groom CR. The druggable genome. *Nat Rev Drug Discov*. 2002;1(9):727-730. doi:10.1038/nrd892.
41. Christopoulos A. Allosteric binding sites on cell-surface receptors: novel targets for drug discovery. *Nat Rev Drug Discov*. 2002;1(3):198-210. doi:10.1038/nrd746.
42. Cimermancic P, Weinkam P, Rettenmaier TJ, et al. CryptoSite: Expanding the Druggable Proteome by Characterization and Prediction of Cryptic Binding Sites. *J Mol Biol*. 2016;428(4):709-719. doi:10.1016/j.jmb.2016.01.029.
43. Singh J, Petter RC, Baillie TA, Whitty A. The resurgence of covalent drugs. *Nat Rev Drug Discov*. 2011;10(4):307-317. doi:10.1038/nrd3410.
44. London N, Miller RM, Krishnan S, et al. Covalent docking of large libraries for the discovery of chemical probes. *Nat Chem Biol*. 2014;10(12):1066-1072. doi:10.1038/nchembio.1666.
45. Mirnezami R, Nicholson J, Darzi A. Preparing for Precision Medicine. *N Engl J Med*. 2012;366(6):489-491. doi:10.1056/NEJMp1114866.

References (continued)

46. Giacomini KM, Yee SW, Ratain MJ, Weinshilboum RM, Kamatani N, Nakamura Y. Pharmacogenomics and patient care: one size does not fit all. *Science Translational Medicine*. 2012;4(153):153ps18-153ps18. doi:10.1126/scitranslmed.3003471.

Chapter 1: Structural Impact of Mutations on Resistance of Tyrosine Kinases to Inhibitors

There are approximately 500 kinases in the human genome.[1] By definition, each of them performs the catalysis of phosphorylation of a target protein, where the γ-phosphate from ATP is transferred onto the hydroxyl groups of serine, threonine, or tyrosine residue, thus regulating its biological activity.[2] The reversible phosphorylation of proteins is one of the most critical post-translational modifications within cells, as it is a universal, ubiquitous mechanism for cellular control and signal transduction.[2] Thus, kinases are involved in the vast majority of cellular processes, including metabolism, transcription, cell movement, apoptosis, cell cycle control, and differentiation.[1]

A recent survey of human disease due to germline mutations in kinases implicated 50 of approximately 500 kinases as the predominant cause for specific neurological disorders, skeletal disorders, haematological and vascular disorders, immunological disorders, and metabolic disorders.[3] In terms of germline and somatic mutations in cancer, protein kinases are the most frequently mutated family of genes to contribute to neoplastic malignancies.[3,4] Many kinases act as tumor suppressors in healthy cells, and mutations often lead to disregulation of pathways they control. Due to their central role in the pathogenesis of human cancers, they have emerged as one of the most important drug target families. Approximately 30 kinase inhibitors with activity against one or multiple kinases currently approved for clinical use.[5]

Tyrosine kinases are a large subfamily of the human kinome composed of 90 members[1] that specifically phosphorylates tyrosine residues. They have been implicated in a wide variety of human tumors, including chronic myelogenous leukemia (CML), breast cancer, lung cancer, and gastrointenstinal tumors. These tumours display large-scale activation of tyrosine kinase activity.[6] Inhibitors of tyrosine kinases have been shown to be safe and effective for specific tumor types. The most prototypical example is imatinib, marketed in the United States by Novartis under the name Gleevec.[7] It targets the constitutively active BCR-ABL fusion protein, a hallmark of CML, resulting from a translocation between chromosomes 9 and 22 that juxtaposes the sequences of the *bcr* and *c-abl* genes.[6] The elucidation of the structural mechanism of imatinib inhibition of BCR-ABL was instrumental in the development of other tyrosine kinase inhibitors (TKIs).[8]

The first crystal structure of a eukaryotic kinase, the catalytic subunit of protein kinase A, was determined in 1991.[9] The same protein kinase fold has since been seen in over 5,000 human kinase structures (PDB search on December 18, 2016), and is extremely conserved among all serine, threonine, and tyrosine kinases.[10] In general, kinases are made up of two domains, the N-terminal lobe composed of a 5-stranded β-sheet and an α-helix, and a C-terminal lobe composed primarily of α-helices (**Figure 1.1**). The ATP binding site is located between the two lobes. A central loop, known as the activation loop, once phosphorylated, stabilizes the kinase in the active conformation compatible with ATP binding. It can undergo large conformational changes between inactive and active states.[10] The conformation of the loop is closely coupled to the highly conserved DFG motif (aspartic acid – phenylalanine - glycine), residing at its base.

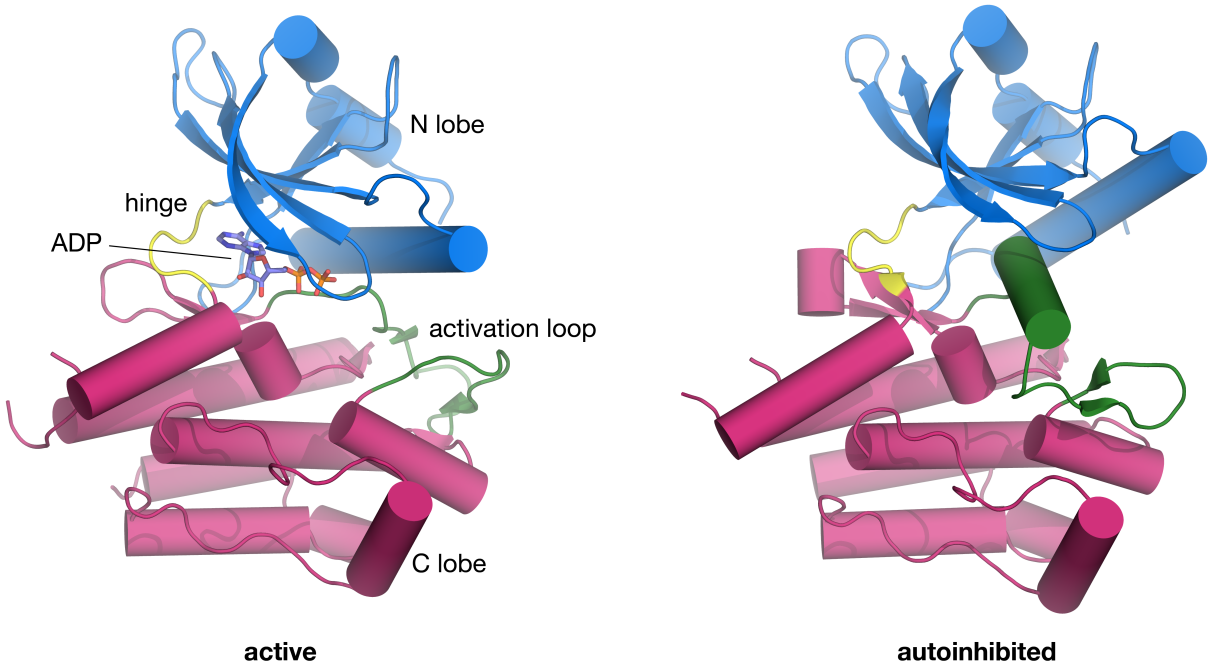


Figure 1.1 Cartoon representation of the active and autoinhibited states of c-Kit tyrosine kinase. Co-crystal structure of the active c-Kit (PDB ID 1PKG)(ref. 11) with ADP is shown on the left. The hinge region (yellow) connects the N (blue) and C (pink) lobes. The activation loop (green) is shown in the DFG-in conformation. Crystal structure of the autoinhibited c-Kit (PDB ID 1T45)(ref. 12) is shown on the right. The activation loop is shown in the DFG-out conformation.

The placement of the loop governs the accessibility of the ATP binding site to the inhibitors. At least two types of inhibitors have been identified. Type II inhibitors exhibit a binding mode compatible only with the inactive DFG-out activation loop conformation. In addition to the ATP binding site, these inhibitors access an allosteric pocket adjacent to it. Imatinib, is an example of a type II inhibitor. In contrast, Type I inhibitors are compatible with numerous activation loop conformations, and do not require a specific DFG-in or DFG-out conformation.[13] Mutations affecting the activation loop conformation are thus predicted to affect type II inhibitor binding, but not type I inhibition.

Mutations within ATP binding site frequently confer resistance to both type I and type II inhibitors. By directly affecting inhibitor binding, they represent cases that are difficult to treat with first or second generation inhibitors.[6] Many of the TKIs currently on the market rely on a conserved threonine (ABL residue T315) for binding specificity.[14] It has been termed the "gatekeeper" residue[15], because it controls access to the allosteric pocket adjacent to the ATP binding site. The T315I mutation itself accounts for approximately 20% of clinically reported resistance.[14,16] Next-generation inhibitors are being actively developed against specific mutations. Ponatinib, for example, has been designed to be compatible with the isoleucine side chain by circumventing it with a linear carbon-carbon triple bond.[17]

This chapter rationalizes the drug resistance in the current treatment of leukemia in terms of structures of kinases ABL1 and FLT3. The first section examines the resistance of all currently approved ABL TKIs against novel clinical mutations at the gatekeeper residue of BCR-ABL. It presents a VEGFR TKI axitinib as a valuable option for the treatment of the resistant cases, providing an example of repurposing an already approved drug. The second section examines mutations within the activation loop of FLT3, a receptor tyrosine kinase frequently targeted in acute myeloid leukemia. It proves that all activation loop mutations are not equally destabilizing to type II inhibitors and calls for the inclusion of patients with the sensitive mutations in clinical trials of new type II TKIs.

Novel TKI-resistant BCR-ABL gatekeeper residue mutations retain *in vitro* sensitivity to axitinib

The contents of this section have been published in an abbreviated form in the following article:

Novel TKI-resistant BCR-ABL1 gatekeeper residue mutations retain in vitro sensitivity to axitinib.

*Elisabeth A Lasater, Evan S Massi, *Adrian Stecula, Julieta Politi, Sophia K Tan, Catherine C Smith, Martha Gunthorpe, Jarrod P Holmes, Farid Chehab, Andrej Sali, and Neil P Shah*
Leukemia. 2016;30(6):1405.

*The author's contribution included comparative structure modeling of the BCR-ABL T315 mutants to rationalize ponatinib resistance and the development of a SOAP-Protein score-based model to predict the ponatinib IC_{50} values for uncharacterized mutants.

Abstract

Acquired kinase domain (KD) point mutations in BCR-ABL are the most frequent cause of resistance to ABL tyrosine kinase inhibitors (TKIs) in chronic myeloid leukemia (CML). BCR-ABL T315I is the most problematic mutation being highly resistant to all approved ABL TKIs except ponatinib. Ponatinib inhibits all single nucleotide substitution KD mutants but is vulnerable to "compound" mutations. We profiled the seven BCR-ABL T315 mutations that have been clinically reported and identified that three mutations are pan-resistant to all approved ABL TKIs (T315F/L/V). T315L confers clinical and *in vitro* resistance to ponatinib but is sensitive *in vitro* to the approved TKI axitinib. To prospectively guide clinical decision-making for novel T315 mutants, we created all 19 substitutions at the T315 residue and found that nearly all confer moderate to high resistance to imatinib, dasatinib, nilotinib and bosutinib. Importantly, seven T315 mutations confer less relative resistance to axitinib than to ponatinib *in vitro*. Further, we provide a structure-based rationale for the impact of clinically observed mutations that allowed us to predict the impact of uncharacterized mutations. Clinical trial evaluation of axitinib in the setting of select T315 mutants is warranted, and medicinal chemistry efforts are needed for the treatment of pan-resistant T315 mutants identified herein.

Introduction

Tyrosine kinase inhibitor (TKI) therapy has revolutionized the treatment of chronic myeloid leukemia (CML) and other cancers, but acquired TKI resistance due to secondary kinase domain (KD) mutations in BCR-ABL represents a commonly encountered clinical problem. The prototypic ABL TKI imatinib is vulnerable to a large number of resistance-conferring secondary KD mutations.[1] The clinical management of imatinib-resistant disease has been successfully guided by *in vitro* studies of mutant sensitivities to alternative TKIs, which are largely predictive of clinical responsiveness.[18-23]

Mutations of the "gatekeeper" residue (T315 in BCR-ABL), which restricts access to a deeper hydrophobic pocket within the kinase,[24] have been particularly problematic not only for ABL TKIs,[18-22,25] but for inhibitors of numerous other pathologically activated kinases including KIT,[26] epidermal growth factor receptor (EGFR),[27] anaplastic lymphoma kinase (ALK),[28] and fms-like tyrosine kinase-3 (FLT3).[29] The nature of the observed substitution is influenced by the native gatekeeper codon and in BCR-ABL, an isoleucine substitution (T315I) has been most commonly reported at this position. EGFR has a different codon for threonine at this position, and TKI-resistant methionine substitutions are commonly identified.[27] Gatekeeper substitutions can confer TKI resistance through one or more of several mechanisms, including steric hindrance,[30] increased ATP affinity,[31] altered kinase conformation dynamics,[14] and altered substrate preference.[32]

Resolution of the ABL-imatinib co-crystal structure demonstrated that imatinib binds to an inactive kinase confirmation[8] and that the threonine residue at position 315 makes a critical, stabilizing hydrogen bond with imatinib. The second-generation ABL TKIs dasatinib[33] and nilotinib[34] also make critical contact with the 315 residue, while bosutinib makes extensive van der Waals contacts with this residue,[35] thereby providing a structural rationale for the vulnerability of these TKIs to the T315I mutation. Alternative amino acid substitutions at this residue have been associated with acquired clinical resistance to particular ABL TKIs, including T315A[36] which confers resistance to dasatinib but not other TKIs,[37,38] and T315N which has been rarely described in patients with resistance to imatinib.[39]

The third generation ABL TKI ponatinib was rationally designed based upon the structure of ABL T315I. Ponatinib does not rely upon hydrogen bonding to the gatekeeper threonine, but utilizes a

triple bond to skirt around the bulky isoleucine substitution.[23] Ponatinib represents the only approved TKI option for BCR-ABL T315I mutant CML. While ponatinib can be highly active in this setting, it is unfortunately associated with a substantial risk of thrombotic events,[40] and alternative strategies may be required for safe long-term management of T315I-associated CML. Due to its invulnerability to amino acid substitutions resulting from a single nucleotide change, ponatinib has been referred to as a "pan-BCR-ABL" inhibitor.[23] However, select "compound" mutations, which can create two or more amino acid substitutions or alternatively a single amino acid substitution if both nucleotide changes occur within the same codon, have been shown to confer ponatinib resistance *in vitro* and clinically.[41] While compound mutations can be commonly detected following acquired resistance to sequential ABL TKI therapy,[36,42,43] they can also occasionally be detected following TKI monotherapy.[43,44]

A mutagenesis screen of BCR-ABL T315I identified compound mutations that could arise on the T315I backbone and confer acquired resistance to ponatinib; highly resistant compound mutants primarily at positions Y253 and E255 (i.e. T315I in addition to mutation at Y253 or E255) were described.[23] Moderately resistant compound T315 mutants were also identified *in vitro* and recently the BCR-ABL T315M mutation was identified in a patient that progressed while on ponatinib therapy.[41] With the increasing prevalence of CML, the extensive use of sequential ABL TKI therapy and the ability to more effectively suppress mutants that arise from single nucleotide change, it is expected that compound mutations will be increasingly encountered in the clinical management of CML. However, due to the nature of the mutagenesis screens that identify resistance to ABL TKIs, clinically problematic compound mutations are not easily predictable, and the resistance profile for these mutations when they arise is largely unknown.

Here we profile the seven BCR-ABL gatekeeper mutations that have been clinically described to date including a compound T315L mutation that we identified in a patient following dasatinib monotherapy and found to confer *in vitro* pan-resistance to all approved ABL TKIs and clinical resistance to ponatinib. Encouragingly, we demonstrate that this mutant is sensitive *in vitro* to the approved vascular endothelial growth factor receptor (VEGFR) TKI axitinib, which has been reported to have affinity for the ABL T315I mutant[45] and has shown the potential for clinical activity in a single case of BCR-ABL T315I mutant CML.[46] In order to prospectively identify other problematic compound gatekeeper mutations for ponatinib and to help guide clinical decision-making, we generated all possible amino acid substitutions at the T315 residue of BCR-ABL and

assessed the impact on sensitivity to all five approved ABL TKIs (imatinib, dasatinib, bosutinib, nilotinib and ponatinib) as well as axitinib. Further, we rationalized the impact of clinically observed mutations against ponatinib using comparative structural models, which enabled us to successfully predict resistance of uncharacterized mutations. In all, we identified a total of nine T315 substitutions that confer a high degree of resistance to ponatinib *in vitro*. Each of these mutants is highly resistant to all other approved ABL TKIs, but several retain sensitivity to axitinib. Therefore, clinical trial evaluation of axitinib in patients with select T315 mutants is warranted.

Materials and methods

Patients All patients gave informed consent for the collection of samples according to the Declaration of Helsinki. All research involving human subjects was approved by the Institutional Review Board at the University of California, San Francisco (UCSF).

Cell lines Ba/F3 cells were propagated in RPMI-1640 with 10% fetal bovine serum (FBS; Omega Scientific) supplemented with murine IL3 (Peprotech).

Tyrosine kinase inhibitors Axitinib, bosutinib and nilotinib were purchased from Selleckchem. Dasatinib and imatinib were purified at the University of California, San Francisco (UCSF). Ponatinib was a kind gift from ARIAD Pharmaceuticals.

BCR-ABL T315 mutant cell lines *Kpn1-BsrG1* fragments of ABL containing the appropriate T315 mutation were synthesized by Integrated DNA Technologies (IDT) and ligated into a *Kpn1(-)* Bluescript-BCR-ABL plasmid. The *SalI-EcoRI* fragment from Bluescript-BCR-ABL was isolated and ligated with the *EcoRI-ClaI* and *ClaI-SalI* fragments from MSCVpuro-BCR-ABL. Plasmid stocks were co-transfected with EcoPack into 293T cells grown in DMEM with 10% FBS using *Trans*IT-LT1 (Mirus) transfection reagent per manufacturer's protocol. Viral supernatant was collected at 48 hours and filtered through 0.45 μM PVDF filters (Millipore). Ba/F3 cells were resuspended in viral supernatant supplemented with murine IL3 and centrifuged for 2 hours at 1000 g at 32 °C. Additional media was added and cells were grown at 37 °C. After 48 hours, cells were collected and resuspended in fresh media plus murine IL3 and 4 μg/mL puromycin (Invitrogen). Following 72 hour selection, cells were washed 2 times with PBS and resuspended in fresh media with no murine IL3.

Cell viability assay Exponentially growing BCR-ABL mutant cells were plated at 5 x 10^3 cells/well in 0.1 mL in 96-well opaque plates with the appropriate concentration of inhibitor in triplicate. Cells were incubated for 48 hours at 37 °C and proliferation was assessed by CellTiter Glo reagent (Promega) according to manufacturer's recommendations using SpectraMax Pro Software (Molecular Devices). For each inhibitor, values were normalized to the median value of the untreated sample and a mean value was calculated. IC_{50} values were generated using nonlinear best-fit regression analysis with Prism 5 software (GraphPad).

Immunoblotting Exponentially growing BCR-ABL T315 mutant cells were treated for 3 hours with 75 nM axitinib and lysates were prepared as previously described.[47] Immunoblotting was performed using anti-phospho-ABL (Y412), anti-phosho-STAT5 (Y694) and anti-STAT5 (Cell Signaling Technology) and anti-cABL (Calbiochem).

Modeling of BCR-ABL mutants Comparative protein structure models of the BCR-ABL T315 mutants were created using MODELLER 9.14.[48] The crystal structure of the native ABL bound to ponatinib served as the template (PDB ID 3OXZ).[17] The ligand from the crystal structure was treated as the BLK residue type in the alignment and was essentially copied from the template structure into a model as a rigid body, retaining proper stereochemistry in the surrounding protein-binding site and protein-ligand interactions. One hundred models were created for each T315 mutant using the automodel class with default settings. The MODELLER objective function ensures that models have proper stereochemistry (*i.e.*, the bond lengths, bond angles, dihedral angles, and chirality are close to ideal values). The non-bonded atom-atom interactions in each model were evaluated using the protein orientation-dependent statistically optimized atomic potential (SOAP-Protein).[49] For each set of 100 models, the maximum SOAP-Protein score of the clinically observed mutations was plotted against experimentally determined ponatinib IC_{50} values on a semi-log plot and an exponential trend line model was fitted using Tableau 8.2.[50] The fit was then used to predict the ponatinib IC_{50} values of the remaining mutations.

Results

Detection of a compound BCR-ABL T315 mutation following ABL TKI therapy A male patient presented to the University of California, San Francisco two weeks after having initiated 140 mg dasatinib daily monotherapy for *de novo* T-lymphoid blast crisis CML. The patient clinically

responded to dasatinib, as evidenced by rapid reduction in splenomegaly and leukocytosis, but relapsed after approximately six weeks of treatment. Direct sequencing of the BCR-ABL KD revealed the presence of substitutions at two nucleotide positions within the T315 codon, resulting in the genesis of T315I and T315L mutations (**Figure 1.2A**, middle panel). While the T315I mutation, which represents a single nucleotide substitution (ACT → ATT) from native BCR-ABL, was readily detectable, the T315L mutation, which requires a two-nucleotide substitution at this residue (ACT → CTT), represented the predominant isoform. The patient was subsequently treated with 45 mg ponatinib daily but had no objective response. After four weeks of ponatinib treatment, re-evaluation of the BCR-ABL KD revealed persistence of only the T315L mutation (**Figure 1.2A**, right panel). The disappearance of the T315I mutation while on ponatinib and the enrichment of the T315L allele in the setting of clinical progression strongly suggest clinical resistance of T315L to ponatinib.

Several clinically identified T315 mutants retain sensitivity to axitinib Given the lack of response of the T315L mutation to ponatinib and the unknown resistance profile of this mutation to the other ABL TKIs, we sought to characterize all BCR-ABL gatekeeper mutations that have been identified clinically against all approved ABL TKIs (T315A[36]/F(Smith manuscript submitted)/I[51]/L[52]/M[41]/N[39]/V[53]). We transformed Ba/F3 cells to growth factor independence with each of these seven BCR-ABL T315 mutants and assessed their sensitivities to the approved ABL TKIs. Of the clinically described gatekeeper mutants, BCR-ABL T315A demonstrated the most relative sensitivity (<10-fold compared to native BCR-ABL) to imatinib, nilotinib and bosutinib (**Figure 1.2B**, **Table 1.1**). Ponatinib was the most active against these seven mutants, with T315A/I/N/V all retaining sensitivity (**Figure 1.2B**). However, the T315F/L/M mutants conferred a high degree of resistance to ponatinib and all other approved ABL TKIs (**Figure 1.2B**).

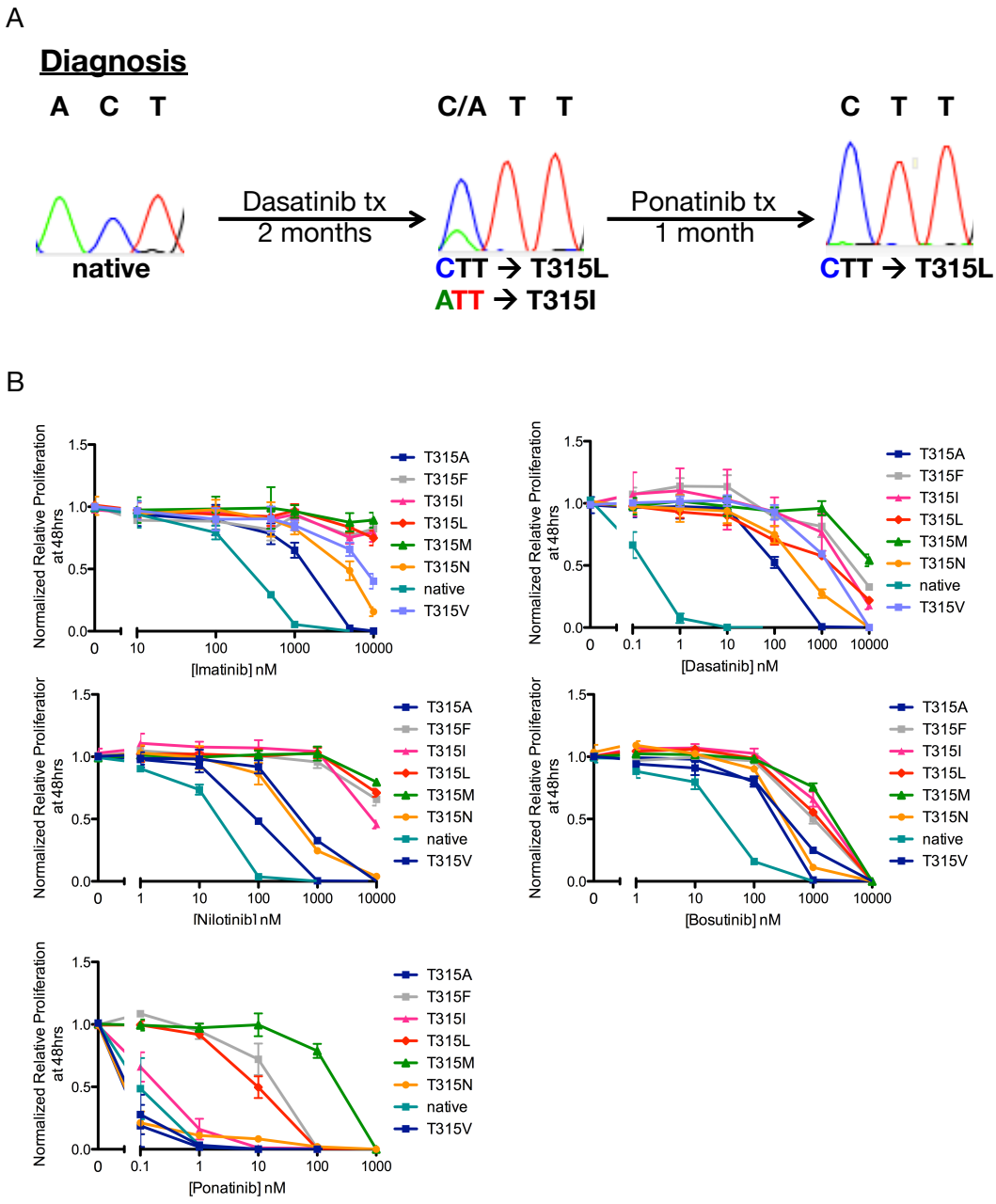


Figure 1.2. Acquisition of BCR-ABL T315L mutation in a patient following dasatinib monotherapy and characterization of clinical mutants against ABL TKIs. A) Sequencing traces of the T315 codon at diagnosis (left panel), following relapse on dasatinib treatment (middle panel) and following four weeks of subsequent treatment with ponatinib (right panel). ACT = T, ATT = I, CTT = L. B) Representative IC_{50} curves for the BCR-ABL T315 mutations described clinically (error bars represent s.d. of triplicates from the same experiment). Cells were treated with increasing concentrations of imatinib, dasatinib, nilotinib, bosutinib or ponatinib for 48 hours. Data is representative of 3 independent experiments.

Table 1.1 Relative fold increase in IC_{50} values compared to native BCR-ABL for ABL TKIs or BCR-ABL T315I for axitinib.

IC50 Fold Increase (Native = 1)					
	Nilotinib	Ponatinib	Bosutinib	Dasatinib	Imatinib
Mutations Observed in Patients					
T315A	4.52	<1	7.18	672.42	5.80
T315F	>500	185.43	33.52	24907.23	>47
T315I	414.58	2.27	47.32	15374.28	>47
T315L	>500	103.47	35.80	11381.96	>47
T315M	>500	2355.65	63.99	>66667	>47
T315N	19.57	<1	9.36	2254.00	17.50
native	1.00	1.00	1.00	1.00	1.00
T315V	27.18	<1	13.33	7683.94	40.09
All Other Substitutions					
T315C	7.43	<1	6.06	1481.13	6.62
T315D	6.31	<1	8.06	1480.49	8.20
T315E	>500	14567.25	37.56	19673.70	>47
T315G	2.91	<1	8.56	94.88	5.44
T315H	>500	4427.80	96.40	17741.52	>47
T315K	>500	5025.93	12.59	6948.18	>47
T315P	<1	<1	5.77	<1	1.43
T315Q	>500	9475.63	62.16	21113.24	>47
T315R	>500	13645.27	85.16	10000.00	>47
T315S	1.72	<1	9.79	2.10	4.67
T315W	28.03	<1	9.70	2566.22	15.15
T315Y	324.73	42.83	12.51	8432.50	>47

IC50 Fold Increase (T315I = 1)	
	Axitinib
Mutations Observed in Patients	
T315A	6.47
T315F	2.36
T315I	1.00
T315L	<1
T315M	4.97
T315N	<1
native	6.74
T315V	<1
All Other Substitutions	
T315C	2.82
T315D	3.57
T315E	12.60
T315G	3.06
T315H	35.23
T315K	9.95
T315P	13.97
T315Q	7.42
T315R	42.25
T315S	22.60
T315W	3.13
T315Y	<1

Sensitive	≤ 2
Moderately Resistant	2.01-4
Resistant	4.01-10
Highly Resistant	>10

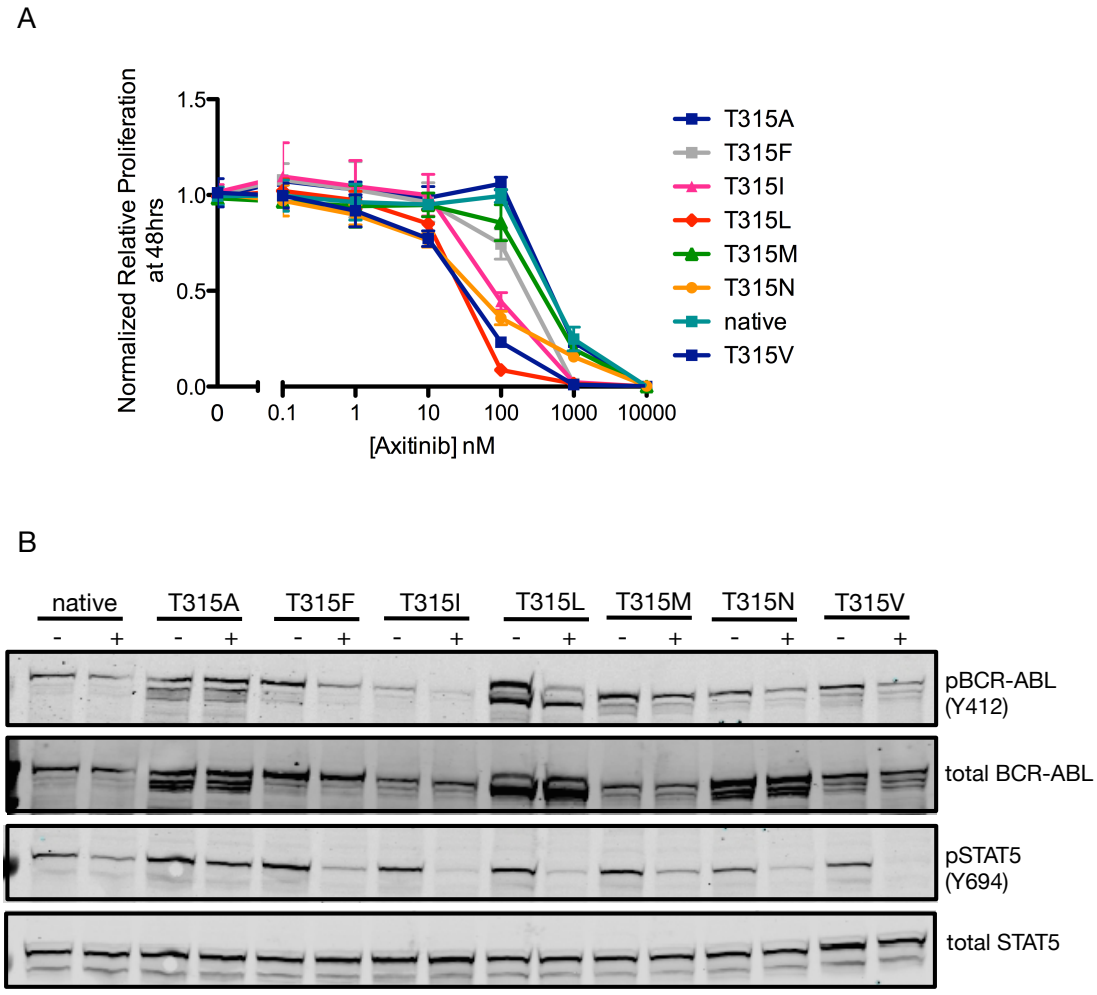


Figure 1.3 Assessment of clinically described BCR-ABL T315 mutant cell line sensitivity to axitinib. A) Representative IC_{50} curves for the BCR-ABL T315 mutations described clinically (error bars represent s.d. of triplicates from the same experiment). Cells were treated with increasing concentrations of axitinib for 48 hours. Data is representative of 3 independent experiments. B) Western blot analysis using anti-phospho-ABL, anti-ABL, anti-phospho-STAT5 and anti-STAT5 on lysates from BCR-ABL mutant cell lines treated for 3 hours with 75 nM axitinib.

We sought to determine if axitinib, which has been reported to be active against BCR-ABL T315I, retains activity against any of these seven clinically-isolated mutants. Encouragingly, BCR-ABL T315L/N/V were more sensitive to axitinib treatment than the T315I mutant (**Figure 1.3A**, **Table 1.1**). Of the remaining clinically described T315 substitutions, T315F and T315M both conferred a significant degree of cross-resistance to axitinib (**Figure 1.3A**).

To confirm that the activity of axitinib against select T315 mutants is a consequence of BCR-ABL kinase inhibition, we assessed the impact of a clinically achievable concentration of axitinib (75 nM) on BCR-ABL phosphorylation at the activation loop residue Y412 and STAT5 phosphorylation (Y694) a downstream effector of BCR-ABL activity.[54] Consistent with the IC_{50} data, axitinib treatment substantially decreased BCR-ABL and STAT5 phosphorylation in the T315F/I/L/N/V BCR-ABL mutant cell lines (**Figure 1.3B**). The T315F mutant showed mild resistance to axitinib based on the IC_{50} data; therefore the potential of this mutation to respond clinically to axitinib strongly depends upon the maximum achievable plasma concentration of the inhibitor. The BCR-ABL native, T315A and T315M mutant cell lines were resistant to axitinib treatment and only showed a slight reduction in either BCR-ABL or STAT5 phosphorylation (**Figure 1.3B**), again consistent with the IC_{50} profile. Taken together, these data suggest that axitinib directly inhibits kinase activity of select T315 mutants and might be clinically active in BCR-ABL T315L/N/V-associated cases.

Structural modeling rationalizes resistance of BCR-ABL T315 mutants to ponatinib To understand the structural basis of resistance toward ponatinib, we modeled the seven clinical mutants bound to ponatinib. A mutation in the binding site is less likely to result in resistance when the binding site is more complementary to the inhibitor, as estimated by the SOAP-Protein score (the proper stereochemistry of any complex model is ensured by construction).[49] Ponatinib is a type II inhibitor that was specifically designed to accommodate the T315I side chain with its ethynyl linker.[23] Ponatinib, unlike imatinib, does not form a hydrogen bond with the hydroxyl group of T315,[34] but it appears to be stabilized by a total of five hydrogen bonds with the protein. The T315 hydrogen bond drives both the potency and specificity of imatinib;[17] thus, its absence, resulting from mutations at the 315 position, leads to almost complete resistance to imatinib (**Figure 1.2B**). The carbon-carbon triple bond eliminates the hydrogen bond requirement for ponatinib, instead making a hydrophobic contact with the I side chain. In addition to minimizing the steric

incompatibility between the ligand and the gatekeeper residue, the triple bond also contributes to a broader specificity of ponatinib compared to imatinib (**Figure 1.4**).[23]

The impact of clinically observed mutations T315A/F/I/L/M/N/V can be discussed in three categories:

First, mutations T315A and T315V, which introduce side chains that are small and hydrophobic, are most sensitive to ponatinib. Like T315, these side chains are predicted not to sterically clash with ponatinib and to make favorable van der Waals interactions with the ethynyl linker and A ring groups. Because the alanine side chain is smaller than that of T, it creates a binding site that is larger and potentially more accessible for ponatinib binding.

Second, mutation T315I represents the mild resistance category of ponatinib sensitivity. As mentioned above, ponatinib makes a favorable hydrophobic interaction with the isoleucine side chain. However, the added methyl group of the isoleucine side chain compared to valine leads to a minor steric clash with the A ring of ponatinib within the model based on the co-crystal structure of ponatinib and native ABL1. This unfavorable interaction requires a subtle change in the binding conformation of ponatinib to alleviate the strain, as shown by the crystal structure of ponatinib bound to T315I (**Figure 1.4**).[17] The result is a slight displacement of the A ring from the hydrophobic pocket, which could account for a decrease in ponatinib potency.

Third, mutations T315F/L/M, which introduce larger, bulky hydrophobic side chains, represent the high resistance category of ponatinib sensitivity. The bulky side chains are sterically incompatible with the binding mode of ponatinib observed in the crystal structure (**Figure 1.4**).

The T315N mutant was excluded from the training set because structural changes necessary to accommodate the ligand are larger than those allowed by the conservative assumptions of homology modeling.

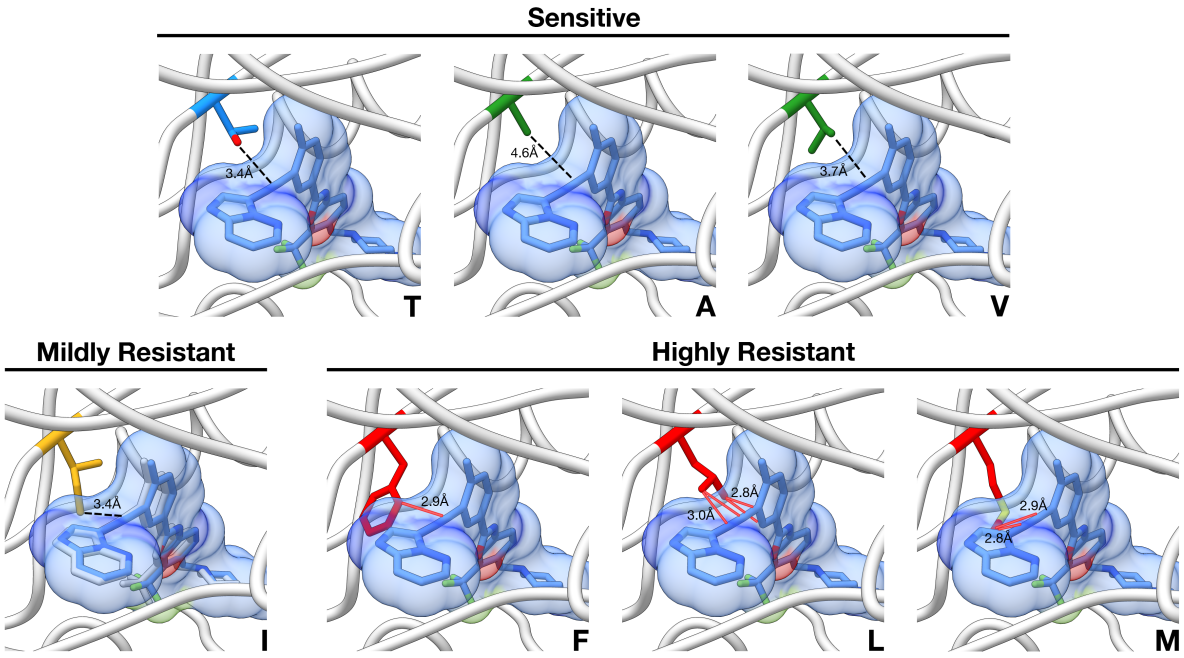
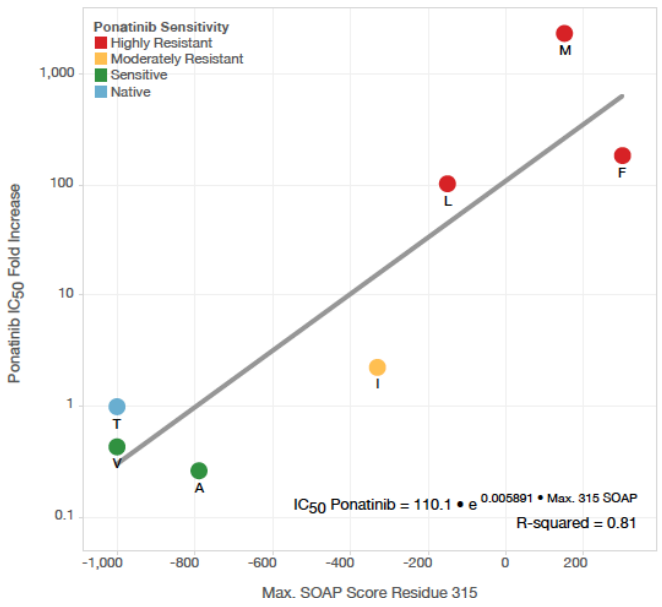


Figure 1.4 Crystal structures and comparative models of the ABL1 kinase domain bound to inhibitors. Cartoon representation of the crystal structure of ponatinib bound to native ABL1 (PDB ID 3OXZ), the model of ponatinib bound to native ABL1 (T), ABL1 T315A and ABL1 T315V (sensitive group), the crystal structure of ponatinib bound to ABL1 T315I (PDB ID 3IK3; mildly resistant group), the model of ponatinib bound to ABL1 T315F, ABL1 T315L, and ABL1 T315M (highly resistant group). Black lines indicate the closest position between the side chain and ponatinib. The red lines indicate steric clashes between the side-chain atoms and ponatinib. In the T315I panel, the pose of ponatinib from the native ABL1 crystal structure is overlaid in white, to indicate the slight change in the binding conformation.

A



B

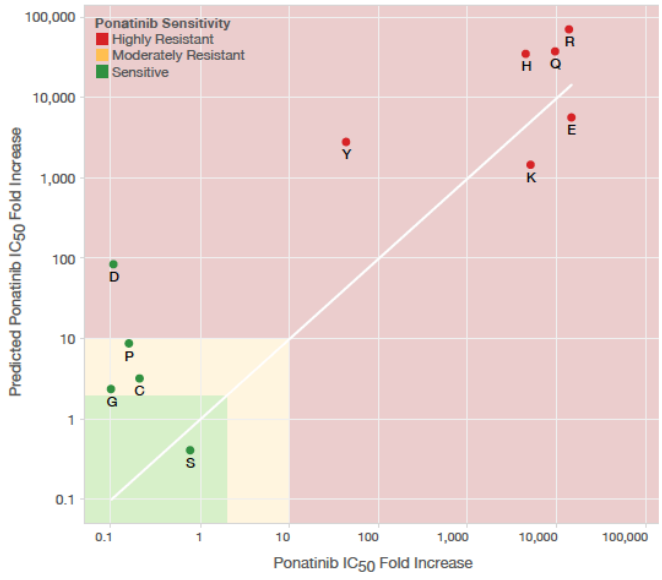


Figure 1.5 Correlations between mutant SOAP scores and ponatinib IC50 values. A) Maximum SOAP score of residue 315 for each of the clinically observed mutants extracted from the comparative structural models. The best fit exponential line is shown in grey, and its equation and coefficient of determination are shown in the bottom right corner. The points are colored according to the experimentally determined ponatinib sensitivity. B) The experimentally determined fold increase in ponatinib sensitivity relative to native BCR-ABL is plotted against the predicted ponatinib sensitivity. The regions of the plot corresponding to high resistance, moderate resistance, and sensitivity to ponatinib are colored in red, yellow, and green, respectively. The line y = x is shown in white.

The maximum SOAP scores for residue 315 for each of the clinically observed mutants correlated well with the ponatinib IC_{50} values, with an R-squared value of 0.81 (**Figure 1.5A**). This correlation indicates that SOAP-Protein score of a comparative model is a sensitive metric able to capture the physicochemical compatibility between the mutant and the ligand, and may serve as a predictive metric for other, as of yet unmapped mutations.[49]

In contrast to ponatinib, we have not performed an equivalent analysis for axitinib because of the relative uncertainty in the modeling of axitinib interactions with the native BCR-ABL and its mutants. The crystal structures of axitinib bound to the kinase domain of the native and T315I mutant versions of ABL1 were published during the preparation of this manuscript.[46] Native ABL1 binds axitinib in the inactive DFG-out conformation, while the T315I ABL1 mutant binds axitinib with the A-loop in the active DFG-in conformation. The differences between the two structures make it difficult to model other mutants with confidence, because it is unknown whether these mutations are activating and stabilizing towards the active kinase conformation, as is the case with T315I.[14]

Several BCR-ABL T315 mutants are highly resistant to ponatinib. To prospectively predict the sensitivity of other atypical gatekeeper substitutions that may be clinically encountered, we modeled the 12 remaining T315 amino acid substitutions. Based on the SOAP-Protein score for residue 315, we calculated the predicted ponatinib IC_{50} value for each mutant using the correlation derived from the clinical T315 mutant dataset (**Figure 1.5A**). Subsequently, we created Ba/F3 cell lines containing each of the remaining mutations and experimentally determined the IC_{50} values of ponatinib (**Table 1.1**). The structural models allowed us to successfully predict the category of ponatinib sensitivity in 7 out of 12 mutants (**Figure 1.5B**). In addition, three other mutants had predicted IC_{50} values that were less than an order of magnitude different from the experimentally determined value (**Figure 1.5B**). Mutations T315E/H/K/Q/R/Y were correctly predicted to be highly resistant to ponatinib. These mutations are characterized by large side chains, which are incompatible with the ponatinib binding mode. The models also correctly predicted the increased sensitivity of the T315S mutation towards ponatinib. Like the alanine and valine side chains, the smaller serine side chain leads to a larger binding pocket and reduces the potential for a steric clash with the compound. The models failed to predict the sensitivity of the T315W/D mutants (W not shown in **Figure 1.5B**). As with T315N, it is conceivable that the models of these mutants are

relatively inaccurate because the structural changes compared to the template structure are larger than allowed by template-based homology modeling by MODELLER.

IC_{50} values of these 12 T315-mutant transformed Ba/F3 cell lines were also determined for imatinib, dasatinib, bosutinib, nilotinib, and axitinib (**Figure 1.6**, **Table 1.1**). BCR-ABL T315P was more sensitive to TKI treatment with dasatinib, nilotinib or ponatinib than native BCR-ABL. Nearly all T315 mutants demonstrated a >10-fold increased relative resistance to dasatinib compared to native BCR-ABL (**Figure 1.6**, **Table 1.1**). With bosutinib and imatinib, all mutants were more resistant to TKI treatment than native BCR-ABL, although several were only moderately resistant. Of the five approved ABL TKIs, ponatinib maintained the broadest activity against T315 mutations, with 9/19 mutants having an IC_{50} value less than native BCR-ABL (**Figure 1.6**, **Table 1.1**). Axitinib, which has been reported to have greater activity against BCR-ABL T315I than native BCR-ABL, also demonstrated increased potency against T315L/N/V/Y mutations compared to T315I (**Figure 1.3**, **Figure 1.6**, and **Table 1.1**). Altogether, axitinib was relatively more active *in vitro* against six gatekeeper mutants (T315F/K/L/M/Q/Y) when compared with the approved ABL TKIs, including ponatinib.

In all, we identified nine BCR-ABL T315 mutations (E/F/H/K/L/M/Q/R/Y) that conferred >10-fold resistance to all ABL TKIs relative to native BCR-ABL (**Table 1.1**). Of these ABL TKI pan-resistant mutations, T315L and T315Y showed greater sensitivity to axitinib than T315I. The fact that multiple mutations are pan-resistant to all ABL inhibitors, including mutations (T315L/M) that have been described clinically, warrants the investigation of novel strategies, such as alternative TKIs or allosteric inhibitors, to treat these vulnerabilities.

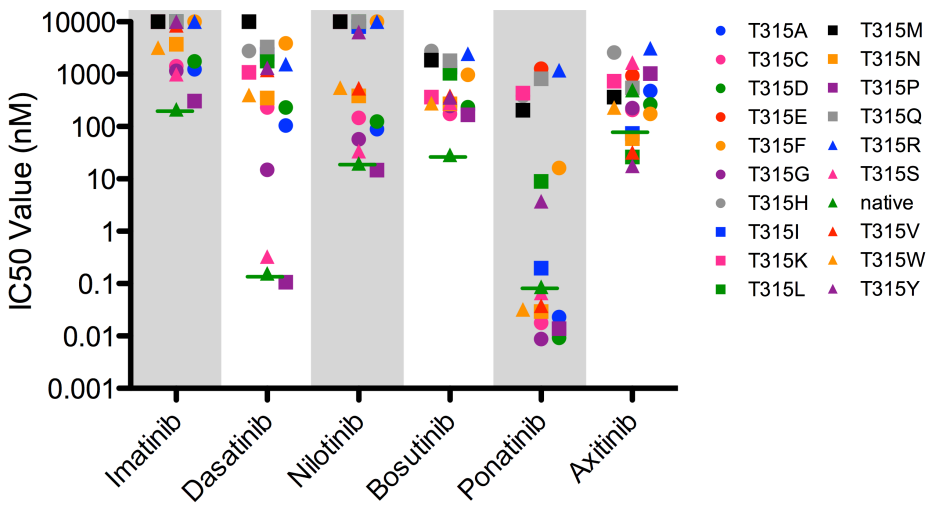


Figure 1.6 IC_{50} Values for BCR-ABL T315 mutant cell lines treated with ABL TKIs and axitinib. Average IC_{50} values for each BCR-ABL T315 mutant cell line. Cells were treated with increasing concentrations of indicated TKI for 48 hours. Data is representative of 3 independent experiments. Native BCR-ABL is indicated by a green line for each ABL TKI and T315I for axitinib.

Discussion

Through the development of second and third-generation ABL TKIs, treatment options are available for BCR-ABL-positive leukemia patients with all amino acid substitutions that arise from single nucleotide changes within the BCR-ABL KD. However, the acquisition of multiple nucleotide substitutions can result in novel amino acid mutations. Here we described the acquisition of a compound gatekeeper mutation (T315L) that arose in concert with a T315I mutation in a single patient on dasatinib monotherapy. Given that BCR-ABL T315I is highly resistant to dasatinib, it was unexpected that monotherapy would select for the additional compound T315L mutation. This finding reinforces the observation that even without obvious selective pressure, multiple nucleotide substitutions can occur. Traditional *in vitro* saturation mutagenesis screens do not typically have the capacity to generate multiple substitutions, and the potential repertoire of compound mutants that can conceivably confer clinical TKI resistance therefore remains unknown.

More distinct TKI-resistant mutations have been detected at the T315 residue (n = 7) than at any other position in the BCR-ABL KD, underscoring the importance of this residue as an escape mechanism to TKIs. In order to prospectively characterize the sensitivity profiles of variant gatekeeper mutations in BCR-ABL, we generated all possible mutations at the T315 residue and assessed them against all five approved ABL TKIs (imatinib, dasatinib, nilotinib, bosutinib and ponatinib) as well as the multi-kinase inhibitor axitinib, which has been described to bind to ABL T315I more avidly than to native ABL.[45] All substitutions at this codon were capable of transforming Ba/F3 cells to growth factor independence and displayed a range of resistance to ABL TKIs and axitinib. For gatekeeper mutations that have been previously described, our model recapitulated published reports and clinical observation.[25,37,39,41] Of all clinically detected mutations at the T315 residue that have been described (T315A/F/I/L/M/N/V), the T315F/L/M mutations are highly resistant to all ABL TKIs tested, including ponatinib, suggesting that patients who acquire these mutations will require novel treatment strategies. Given the described activity of axitinib against BCR-ABL T315I,[46] we hypothesized that activity may be retained against a leucine substitution, and indeed, BCR-ABL T315L was more sensitive to axitinib than BCR-ABL T315I. BCR-ABL T315F displayed minor resistance to axitinib while BCR-ABL T315M demonstrated 5-fold resistance compared to T315I. Based on the reported maximum plasma concentration of axitinib (~72 nM),[55] it is unclear if adequate plasma concentrations can be safely achieved to

effectively inhibit these mutants. Like ponatinb, axitinib is also a multi-kinase inhibitor and therefore the benefit that arises from use of this compound will have to be balanced against the clinical risks.

In all, nine T315 mutations (E/F/H/K/L/M/Q/R/Y), each of which requires changes of at least a two nucleotides, demonstrated a high degree of resistance to all five approved ABL TKIs. The *in vitro* activity of axitinib against T315F/L/Y warrants clinical investigation. On the other hand, the T315E/H/K/M/Q/R substitutions, which are highly pan-resistant to all approved ABL TKIs, confer substantial *in vitro* resistance to axitinib, and are therefore not predicted to respond clinically. For patients who evolve these substitutions prompt consideration of allogeneic stem cell transplantation seems prudent. We noted that substitutions with any of the three basic amino acids (K/H/R) conferred resistance to all compounds tested. However, no other group of amino acids displayed uniform response illustrating the difficulty in predicting drug sensitivity based on amino acid substitution.

Overall, ponatinib maintained activity against the largest number of gatekeeper substitutions with 9/19 mutants demonstrating more sensitivity to TKI treatment than native BCR-ABL. Strikingly, the relative sensitivity of gatekeeper mutants appeared bimodal for all approved ABL TKIs, including ponatinib, with very few "moderately resistant" mutations observed. In contrast, few mutants were "highly resistant" to axitinib, confirming that axitinib binds to BCR-ABL in a manner that is fundamentally different from the approved ABL TKIs. The recently published structures of ABL complexed with axitinib will allow for investigation of the basis of this observation and help direct rational modifications of axitinib to further increase its activity and selectivity toward native BCR-ABL and gatekeeper mutants. On this front, we utilized structural comparative models to successfully rationalize and predict ponatinib resistance for the majority of T315 mutants. This provides an expandable platform for predicting the resistance profile of additional ABL TKIs for the treatment of novel BCR-ABL T315 mutations. Furthermore, the work we described herein promises to provide clinicians with useful guidance toward selection of appropriate approved or investigational therapy, and the pan-resistant mutants we identified represent the next frontier in the management of BCR-ABL-positive leukemias.

FLT3 D835 mutations confer differential resistance to type II FLT3 inhibitors

The contents of this section have been published in the following article:

FLT3 D835 mutations confer differential resistance to type II FLT3 inhibitors.

*Catherine C Smith, Kimberly Lin, *Adrian Stecula, Andrej Sali, and Neil P Shah*
Leukemia. 2015;29(12):2390.

*The author's contribution included comparative structure modeling of the FLT3 D835 mutants to rationalize type II inhibitor resistance.

Introduction

Activating mutations in FLT3 occur commonly in acute myeloid leukemia (AML), including internal tandem duplication (ITD) and point mutations in the tyrosine kinase domain, typically at the activation loop (AL) residue D835. Recent studies of potent FLT3 inhibitors in FLT3 mutant patients has demonstrated clinical activity for sorafenib,[56] quizartinib (AC220),[57] ponatinib,[58] and crenolanib.[59] However, duration of response to FLT3 inhibitors is limited by resistance-conferring secondary kinase domain (KD) mutations. Highly resistant FLT3-ITD AL D835V/Y/F and gatekeeper F691L mutations confer acquired clinical resistance to quizartinib[60] and acquired D835Y/H and F691L mutations have been detected at the time of resistance to sorafenib.[56,61]

The most common residue implicated in clinical resistance to FLT3 TKI therapy is D835.[56,60-62] Molecular docking analysis suggests that D835 mutants induce an active "DFG-in" kinase conformation unfavorable for binding by type II inhibitors such as sorafenib, quizartinib, ponatinib and PLX3397.[60,62] Type I inhibitors (e.g. crenolanib) bind a "DFG-in" conformation and retain activity against D835 mutants.[63] Despite the fact that D835 mutations have been commonly associated with *in vitro* and clinical resistance to type II FLT3 inhibitors, differences in the spectrum of D835 mutations identified at the time of clinical resistance to FLT3 TKIs (e.g. D835H mutations observed with sorafenib but not quizartinib resistance) suggest that relative resistance of D835 substitutions to type II FLT3 TKIs is not uniform, though the number of cases analyzed to date is small. *In vitro* mutagenesis screens have identified different resistant D835 substitutions for individual FLT3 TKIs.[60] Nevertheless, clinical trials of type II FLT3 inhibitors commonly exclude

patients with any FLT3 D835 mutation due to a prevailing assumption that all FLT3 D835 substitutions uniformly confer resistance to type II inhibitors. We sought to experimentally determine the degree of resistance conferred by individual D835 mutations and to further characterize molecular mechanisms underlying this resistance with the goal of informing clinical trial design and molecular testing.

Materials and methods

Ba/F3 cell lines were created and proliferation assays performed as previously described.[29] Comparative protein structure models of FLT3 mutants were created with MODELLER 9.14,[48] using the crystal structures of the auto-inhibited FLT3 (PDB ID 1RJB)[64] and the co-crystal structure of FLT3 with quizartinib (PDB ID 4RT7),[62] as templates. For each D835 mutant, we generated 100 models using the automodel class with default settings, separately for each template. The models had acceptable protein orientation-dependent statistically optimized atomic potential (SOAP-Protein) scores.[49] They were clustered visually into up to 5 classes based on the conformation of the mutated side chain.

Results and discussion

We profiled all D835 substitutions previously reported to cause FLT3 TKI resistance in patients,[56,60,65] as well as D835 mutations occurring in patients as cataloged in the Sanger COSMIC database or the Cancer Genome Atlas. IC_{50} for proliferation of Ba/F3 cells expressing FLT3-ITD D835 mutants profiled for the clinically active FLT3 inhibitors quizartinib,[57] sorafenib,[56] ponatinib,[58] PLX3397,[62] and crenolanib[59] is shown in **Table 1.2** and are in general keeping with previously reported values.[47,60,61,63,66] Relative resistance compared to FLT3-ITD is shown in **Table 1.3**. Surprisingly, individual D835 substitutions conferred a wide range of resistance to all tested type II inhibitors. As previously reported, FLT3-ITD D835V/Y/F mutations cause a high degree of resistance to all type II inhibitors.[47,60] Deletion of the D835 residue or substitution with the bulky residue isoleucine also resulted in a high degree of resistance. The basic substitution D835H caused intermediate resistance, which may explain why this residue has been observed in clinical resistance to sorafenib,[56] but not to the more potent inhibitor quizartinib.[60] Overall, D835A/E/G/N mutations conferred the least degree of resistance to the type II inhibitors. Consistent with our experimental observations, we identified only highly resistant D835 mutations (D835V/Y/F/I) in

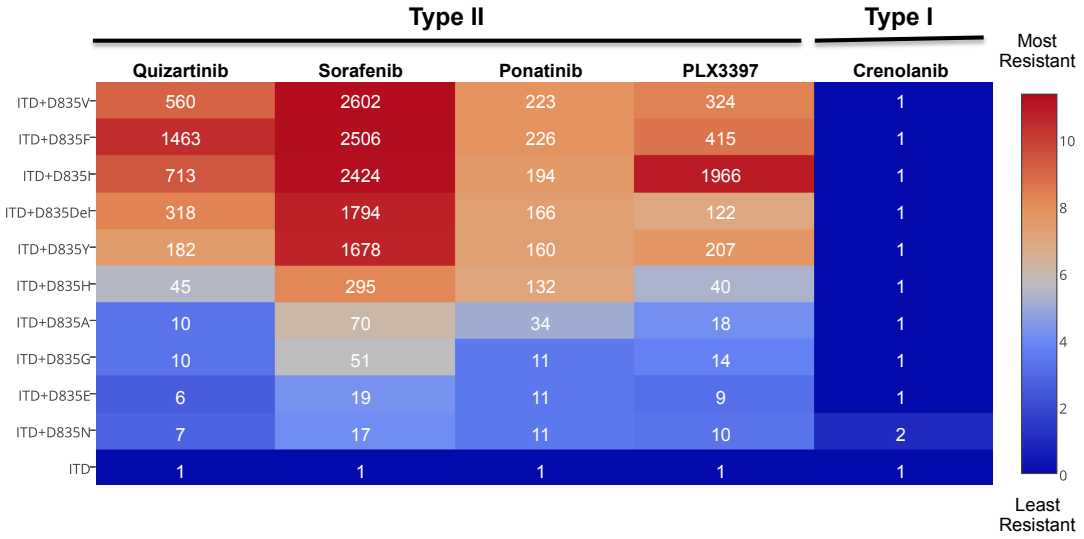
patients who relapsed after responding to quizartinib.[60] As expected, D835 mutations retained sensitivity to the type I inhibitor crenolanib and consistent with previous reports, it is expected that other type I inhibitors such as sunitinib, would also retain activity against these mutations.[61]

Type II inhibitors bind to the conformation coupled to the DFG-out position of the kinase AL (residues 829-856 in FLT3).[13] As previously noted, D835 is predicted to play a critical role in the stabilization of the DFG-out conformation by serving as an amino-terminal capping residue for the short, one-turn α-helix.[12,60,64] Alpha helices have a macrodipole, with a positive pole near the N-terminus and a negative pole near the C-terminus.[67,68] Short helices in particular may be stabilized with single residues forming favorable interactions with the helix dipole near the ends of the helix. The presumed negative charge of the D835 side chain at the N-terminus of the short α-helix spanning residues 835-839 is an example of such an interaction. D835 may also stabilize the helix by forming a hydrogen bond with either the main-chain amide or the side chain hydroxyl of S838, as seen in the crystal structures of the auto-inhibited and quizartinib bound FLT3 structures, respectively. Similar interactions are observed for the equivalent aspartic acid residue in the KIT structure (PDB ID 1T45).[12]

Table 1.2 IC_{50} for proliferation of FLT3 kinase domain mutations in the presence of FLT3 inhibitors.

	Quizartinib		Sorafenib		Ponatinib		Crenolanib		Sunitinib		Midostaurin	
Mutation	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
ITD	0.31	0.26	2.06	0.98	1.75	0.88	15.38	4.86	4.20	2.27	4.55	0.71
ITD+F691L	102.36	20.77	2063.33	149.70	26.37	0.45	122.27	2.71	23.18	0.66	7.07	0.94
ITD+A848P	5.11	2.05	208.90	32.77	63.75	10.18	27.02	2.78	20.08	2.04	2.85	0.18
ITD+N676K	2.79	0.40	19.93	3.18	3.11	0.29	101.04	7.21	17.94	0.94	28.02	4.53
ITD+D835Y	56.75	5.25	4717.33	936.94	279.97	16.93	22.06	2.51	13.94	1.47	6.78	1.15
ITD+D835V	174.80	10.24	7315.67	276.50	390.03	19.35	19.08	0.62	15.81	1.19	4.10	0.61
ITD+D835F	456.97	9.77	7047.33	108.93	396.00	58.78	18.33	3.02	15.97	1.09	5.28	1.54
ITD+D835H	13.95	4.88	830.63	56.67	230.97	40.69	20.63	1.14	15.56	1.82	3.95	0.43
ITD+D835A	3.15	0.24	196.87	10.04	59.44	9.07	17.50	2.04	7.54	1.57	3.47	0.40
ITD+D835E	1.83	0.44	52.66	7.88	19.18	0.56	17.35	1.26	10.15	0.48	1.93	0.33
ITD+D835G	2.99	0.43	144.53	6.91	19.45	0.25	18.17	3.45	13.32	2.00	2.88	0.50
ITD+D835I	222.73	43.29	6815.33	348.01	340.00	59.41	17.31	5.99	12.58	0.53	3.18	0.48
ITD+D835N	2.13	0.64	47.46	4.48	19.21	2.25	26.07	2.99	12.28	1.35	3.56	0.37
ITD+D835Del	99.34	10.06	5043.67	326.71	290.23	31.42	18.35	2.00	8.19	0.74	3.61	0.53

Table 1.3 Relative resistance of FLT3 inhibitors to FLT3-ITD kinase domain mutations compared with ITD alone. Blue indicates most sensitive; red indicates most resistant. Number indicates fold-resistance compared with ITD alone for each inhibitor.



It is not straightforward to rationalize the effect of mutations near or in the short α-helix on the distant drug-binding active site. Nevertheless, it has been suggested that the short α-helix, which is part of the AL, is coupled to the drug-binding site over a distance.[60] Therefore, we focus here on describing the impact of D835 mutations on the short α-helix. We modeled each of the mutants using the crystal structures of FLT3 in the auto-inhibited state and bound to quizartinib. It is possible to correlate three categories of resistance to type II inhibitors with predicted structural features of each mutant (**Figure 1.7**).

The most sensitive mutants (D835E/N) are characterized by the predicted preservation of the hydrogen bond between D835 and S838 based on models utilizing both the apo and holo structures as templates. Side chains of the D835E/N mutations are predicted to form hydrogen bonds with the side chain hydroxyl group of S838, the main chain amide group of S838, or the main chain amide group of M837, depending on the modeled conformations of the neighboring side chains. As a result, the short α-helix and thus the coupled DFG-out conformation are conserved, retaining type II inhibitor binding. In contrast, the short or lacking side chains of the residues in the more resistant mutants (D835G/A) cannot form any hydrogen bonds. Thus, the short α-helix may not form, shifting the equilibrium from the DFG-out to the DFG-in conformation, rationalizing the observed resistance.

The most highly resistant mutants (D835Y/V/I/F) are large and bulky hydrophobic amino acid residues. In addition to an inability to hydrogen bond with S838, these large side chains are predicted to be sterically incompatible with the short α-helix in models based on both the apo and holo structures, thus further shifting the equilibrium towards the active DFG-in conformation.

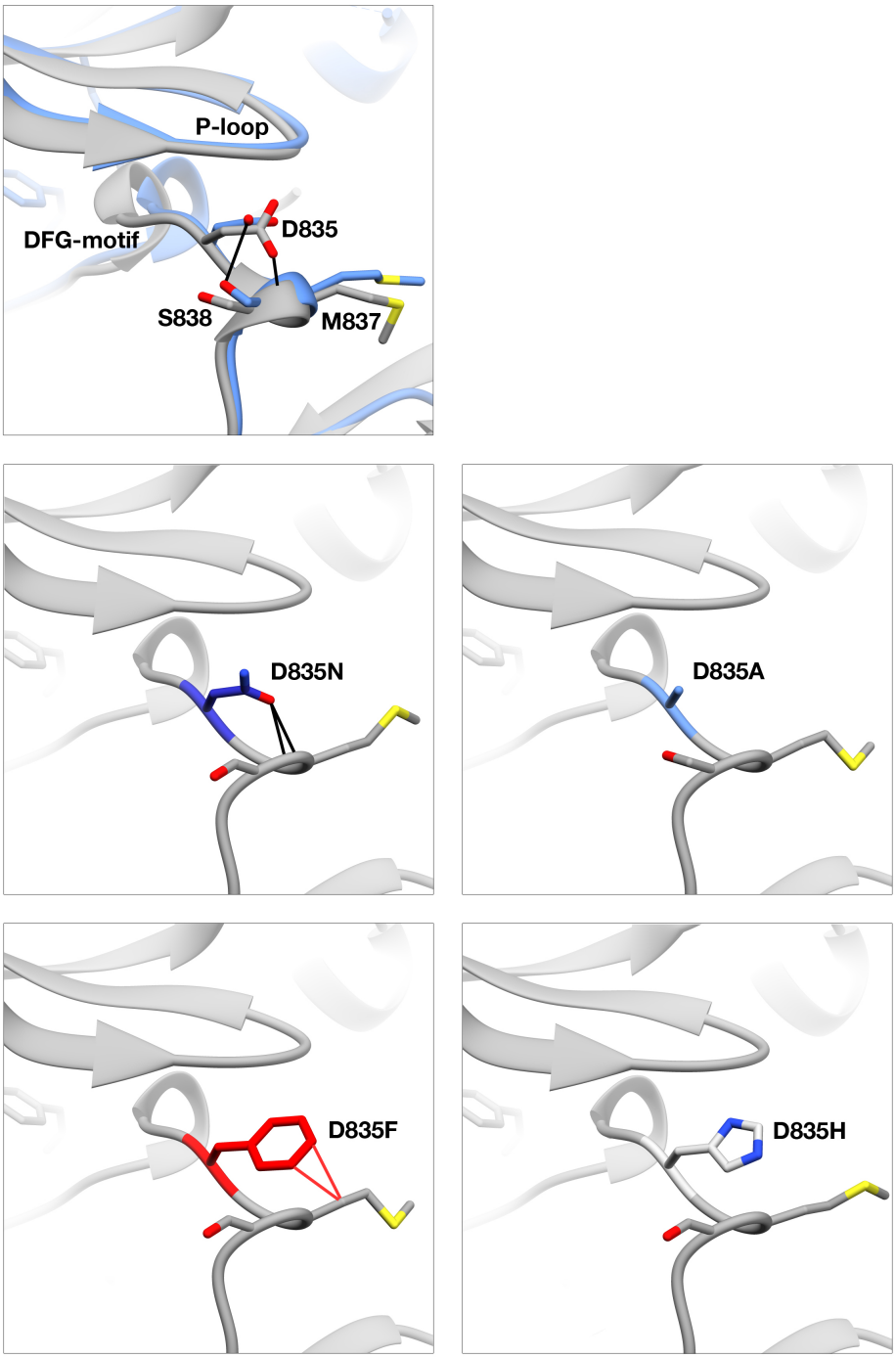


Figure 1.7 Comparative structure models of FLT3 mutants. Cartoon representation of the activation loop, where residue 835 is highlighted in the center of each panel. Predicted hydrogen bonds and steric clashes are indicated by solid black and red lines, respectively. The crystal structures of FLT3 in the autoinhibited state (PDB ID 1RJB)[64] (gray) and bound to quizartinib (PDB ID 4RT7)[62] (blue) are shown in the leftmost panel. Predicted orientations of mutant side chains from the three categories of resistance to type II inhibitors and histidine are shown in the other panels. All models shown are based on the FLT3 structure bound to quizartinib.

Unlike the mutations discussed above, the models of the moderately resistant D835H mutation based each of the two template structures differ from each other. We assume the model based on the quizartinib template is more accurate because holo states are generally more accurately modeled based on holo than apo templates. In addition, to accommodate the binding mode of quizartinib, the DFG motif is predicted to be displaced out of the DFG pocket (**Figure 1.7**), thus creating an essential edge-to-face aromatic interaction between F830 and the middle phenyl ring of quizartinib.[62] As a result, the predicted side chain conformations of the D835H mutant are unable to form any hydrogen bonds. The intermediary type II inhibitor resistance of the D835H mutation is therefore predicted to be a consequence of its inability to both accommodate the binding mode of the type II inhibitors and preserve the hydrogen bonds that stabilize the short α-helix in the auto-inhibited conformation. Our conclusions based on the holo state model are in agreement with those based on an apo state model.[61]

Our data suggest that some clinically relevant D835 mutants retain sensitivity to type II inhibitors at clinically achievable drug concentrations and propose a molecular mechanism for differences in sensitivity for individual D835 mutants to type II FLT3 TKIs. While it has been long recognized that the D835 residue plays an important role in stabilizing the inactive conformation of FLT3,[5] our molecular modeling studies implicate critical hydrogen bonding interactions that mediate the formation of a short α-helix critical to the DFG-out conformation, with the side chain hydroxyl group of S838, the main chain amide group of S838, and the main chain amide group of M837, as necessary for type II inhibitor binding. Mutations that preserve this short α-helix do not appreciably bias the kinase active conformation and therefore cause nominal resistance. Notably, current clinical assays report only the presence or absence of D835 mutations. From a practical perspective, our findings argue that the common practice of uniformly excluding patients with any D835 mutation from participation in type II FLT3 TKI clinical trials is misguided, as less resistant D835 substitutions such as D835N/E/G/A may retain sensitivity to TKI therapy. The development of clinical assays that report the exact nature of amino acid substitution(s) at D835 is therefore indicated. Further studies to elucidate molecular mechanisms of resistance mediated by FLT3 KD mutants to type II TKIs will require co-crystal structural analyses, particularly with type I inhibitors bound to the active conformation of FLT3.

References

1. Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S. The protein kinase complement of the human genome. *Science*. 2002;298(5600):1912-1934. doi:10.1126/science.1075762.
2. Adams JA. Kinetic and catalytic mechanisms of protein kinases. *Chem Rev*. 2001;101(8):2271-2290. doi:10.1021/cr000230w.
3. Lahiry P, Torkamani A, Schork NJ, Hegele RA. Kinase mutations in human disease: interpreting genotype|[ndash]|phenotype relationships. *Nat Rev Genet*. 2010;11(1):60-74. doi:10.1038/nrg2707.
4. Futreal PA, Coin L, Marshall M, et al. A census of human cancer genes. *Nat Rev Cancer*. 2004;4(3):177-183. doi:10.1038/nrc1299.
5. Knapp S, Sundström M. Recently targeted kinases and their inhibitors-the path to clinical trials. *Current Opinion in Pharmacology*. 2014;17:58-63. doi:10.1016/j.coph.2014.07.015.
6. Baselga J. Targeting Tyrosine Kinases in Cancer: The Second Wave. *Science*. 2006;312(5777):1175-1178. doi:10.1126/science.1125951.
7. Buchdunger E, Zimmermann J, Mett H, Meyer T. Inhibition of the Abl protein-tyrosine kinase in vitro and in vivo by a 2-phenylaminopyrimidine derivative. *Cancer Res*. 1996.
8. Schindler T, Bornmann W, Pellicena P, Miller WT, Clarkson B, Kuriyan J. Structural Mechanism for STI-571 Inhibition of Abelson Tyrosine Kinase. *Science*. 2000;289(5486):1938-1942. doi:10.1126/science.289.5486.1938.
9. Knighton DR, Zheng JH, Eyck Ten LF, et al. Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science*. 1991;253(5018):407-414.
10. Huse M, Kuriyan J. The Conformational Plasticity of Protein Kinases. *Cell*. 2002;109(3):275-282. doi:10.1016/S0092-8674(02)00741-9.
11. Mol CD, Lim KB, Sridhar V, et al. Structure of a c-kit product complex reveals the basis for kinase transactivation. *J Biol Chem*. 2003;278(34):31461-31464. doi:10.1074/jbc.C300186200.
12. Mol CD, Dougan DR, Schneider TR, et al. Structural basis for the autoinhibition and STI-571 inhibition of c-Kit tyrosine kinase. *J Biol Chem*. 2004;279(30):31655-31663. doi:10.1074/jbc.M403319200.
13. Wodicka LM, Ciceri P, Davis MI, et al. Activation State-Dependent Binding of Small Molecule Kinase Inhibitors: Structural Insights from Biochemistry. *Chemistry & Biology*. 2010;17(11):1241-1249. doi:10.1016/j.chembiol.2010.09.010.
14. Azam M, Seeliger MA, Gray NS, Kuriyan J, Daley GQ. Activation of tyrosine kinases by

References (continued)

mutation of the gatekeeper threonine. *Nat Struct Mol Biol*. 2008;15(10):1109-1118. doi:10.1038/nsmb.1486.

15. Liu Y, Shah K, Yang F, Witucki L, Shokat KM. A molecular gate which controls unnatural ATP analogue recognition by the tyrosine kinase v-Src. *Bioorg Med Chem*. 1998;6(8):1219-1226.
16. Shah NP, Nicoll JM, Nagar B, et al. Multiple BCR-ABL kinase domain mutations confer polyclonal resistance to the tyrosine kinase inhibitor imatinib (STI571) in chronic phase and blast crisis chronic myeloid leukemia. *Cancer Cell*. 2002;2(2):117-125.
17. Zhou T, Commodore L, Huang W-S, et al. Structural Mechanism of the Pan-BCR-ABL Inhibitor Ponatinib (AP24534): Lessons for Overcoming Kinase Inhibitor Resistance. *Chemical Biology & Drug Design*. 2010;77(1):1-11. doi:10.1111/j.1747-0285.2010.01054.x.
18. Soverini S, Colarossi S, Gnani A, et al. Contribution of ABL Kinase Domain Mutations to Imatinib Resistance in Different Subsets of Philadelphia-Positive Patients: By the GIMEMA Working Party on Chronic Myeloid Leukemia. *Clin Cancer Res*. 2006;12(24):7374-7379. doi:10.1158/1078-0432.CCR-06-1516.
19. Shah NP, Tran C, Lee FY, Chen P, Norris D, Sawyers CL. Overriding imatinib resistance with a novel ABL kinase inhibitor. *Science*. 2004;305(5682):399-401. doi:10.1126/science.1099480.
20. O'Hare T, Walters DK, Stoffregen EP, et al. In vitro Activity of Bcr-Abl Inhibitors AMN107 and BMS-354825 against Clinically Relevant Imatinib-Resistant Abl Kinase Domain Mutants. *Cancer Res*. 2005;65(11):4500-4505. doi:10.1158/0008-5472.CAN-05-0259.
21. Puttini M, Coluccia AML, Boschelli F, et al. In vitro and In vivo Activity of SKI-606, a Novel Src-Abl Inhibitor, against Imatinib-Resistant Bcr-Abl+ Neoplastic Cells. *Cancer Res*. 2006;66(23):11314-11322. doi:10.1158/0008-5472.CAN-06-1199.
22. Redaelli S, Piazza R, Rostagno R, et al. Activity of Bosutinib, Dasatinib, and Nilotinib Against 18 Imatinib-Resistant BCR/ABL Mutants. *Journal of Clinical Oncology*. September 2016. doi:10.1200/JCO.2008.17.27.issue-3;wgroup:string:Publication.
23. O'Hare T, Shakespeare WC, Zhu X, et al. AP24534, a Pan-BCR-ABL Inhibitor for Chronic Myeloid Leukemia, Potently Inhibits the T315I Mutant and Overcomes Mutation-Based Resistance. *Cancer Cell*. 2009;16(5):401-412. doi:10.1016/j.ccr.2009.09.028.
24. Zuccotto F, Ardini E, Casale E, Angiolini M. Through the "gatekeeper door": exploiting the active kinase conformation. *J Med Chem*. 2010;53(7):2681-2694. doi:10.1021/jm901443h.
25. Weisberg E, Manley PW, Cowan-Jacob SW, Hochhaus A, Griffin JD. Second generation

References (continued)

inhibitors of BCR-ABL for the treatment of imatinib-resistant chronic myeloid leukaemia. *Nat Rev Cancer*. 2007;7(5):345-356. doi:10.1038/nrc2126.

26. Tamborini E, Bonadiman L, Greco A, et al. A new mutation in the KIT ATP pocket causes acquired resistance to imatinib in a gastrointestinal stromal tumor patient. *Gastroenterology*. 2004;127(1):294-299. doi:10.1053/j.gastro.2004.02.021.
27. Kobayashi S, Boggon TJ, Dayaram T, et al. EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. *New Engl J Med*. 2005;352(8):786-792. doi:10.1056/NEJMoa044238.
28. Katayama R, Shaw AT, Khan TM, et al. Mechanisms of Acquired Crizotinib Resistance in ALK-Rearranged Lung Cancers. *Science Translational Medicine*. 2012;4(120):120ra17-120ra17. doi:10.1126/scitranslmed.3003316.
29. Smith CC, Wang Q, Chin C-S, et al. Validation of ITD mutations in FLT3 as a therapeutic target in human acute myeloid leukaemia. *Nature*. 2012;485(7397):260-263. doi:10.1038/nature11016.
30. Daub H, Specht K, Ullrich A. Strategies to overcome resistance to targeted protein kinase inhibitors. *Nat Rev Drug Discov*. 2004;3(12):1001-1010. doi:10.1038/nrd1579.
31. Yun C-H, Mengwasser KE, Toms AV, et al. The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP. *Proc Natl Acad Sci USA*. 2008;105(6):2070-2075. doi:10.1073/pnas.0709662105.
32. Skaggs BJ, Gorre ME, Ryvkin A, et al. Phosphorylation of the ATP-binding loop directs oncogenicity of drug-resistant BCR-ABL mutants. *Proceedings of the National Academy of Sciences*. 2006;103(51):19466-19471. doi:10.1073/pnas.0609239103.
33. Tokarski JS, Newitt JA, Chang CYJ, et al. The structure of Dasatinib (BMS-354825) bound to activated ABL kinase domain elucidates its inhibitory activity against imatinib-resistant ABL mutants. *Cancer Res*. 2006;66(11):5790-5797. doi:10.1158/0008-5472.CAN-05-4187.
34. Nagar B, Bornmann WG, Pellicena P, et al. Crystal structures of the kinase domain of c-Abl in complex with the small molecule inhibitors PD173955 and imatinib (STI-571). *Cancer Res*. 2002;62(15):4236-4243. doi:10.1016/S0962-8924(99)01549-4.
35. Levinson NM, Boxer SG. Structural and Spectroscopic Analysis of the Kinase Inhibitor Bosutinib and an Isomer of Bosutinib Binding to the Abl Tyrosine Kinase Domain. Ramchandran R, ed. *PLoS ONE*. 2012;7(4):e29828. doi:10.1371/journal.pone.0029828.
36. Shah NP, Skaggs BJ, Branford S, et al. Sequential ABL kinase inhibitor therapy selects for

References (continued)

compound drug-resistant BCR-ABL mutations with altered oncogenic potency. *J Clin Invest*. 2007;117(9):2562-2569. doi:10.1172/JCI30890.

37. Burgess MR, Skaggs BJ, Shah NP, Lee FY, Sawyers CL. Comparative analysis of two clinically active BCR-ABL kinase inhibitors reveals the role of conformation-specific binding in resistance. *Proc Natl Acad Sci USA*. 2005;102(9):3395-3400. doi:10.1073/pnas.0409770102.
38. Soverini S, Colarossi S, Gnani A, et al. Resistance to dasatinib in Philadelphia-positive leukemia patients and the presence or the selection of mutations at residues 315 and 317 in the BCR-ABL kinase domain. *Haematologica*. 2007;92(3):401-404. doi:10.3324/haematol.10822.
39. Al-Ali H-K, Heinrich MC, Lange T, et al. High incidence of BCR-ABL kinase domain mutations and absence of mutations of the PDGFR and KIT activation loops in CML patients with secondary resistance to imatinib. *Hematol J*. 2004;5(1):55-60. doi:10.1038/sj.thj.6200319.
40. Cortes JE, Kim DW, Pinilla-Ibarz J, et al. A phase 2 trial of ponatinib in Philadelphia chromosome-positive leukemias. *New Engl J Med*. 2013;369(19):1783-1796. doi:10.1056/NEJMoa1306494.
41. Zabriskie MS, Eide CA, Tantravahi SK, et al. BCR-ABL1 Compound Mutations Combining Key Kinase Domain Positions Confer Clinical Resistance to Ponatinib in Ph Chromosome-Positive Leukemia. *Cancer Cell*. 2014;26(3):428-442. doi:10.1016/j.ccr.2014.07.006.
42. Khorashad JS, Kelley TW, Szankasi P, et al. BCR-ABL1 compound mutations in tyrosine kinase inhibitor–resistant CML: frequency and clonal relationships. *Blood*. 2013;121(3):489-498. doi:10.1182/blood-2012-05-431379.
43. Gibbons DL, Pricl S, Posocco P, et al. Molecular dynamics reveal BCR-ABL1 polymutants as a unique mechanism of resistance to PAN-BCR-ABL1 kinase inhibitor therapy. *Proc Natl Acad Sci USA*. 2014;111(9):3550-3555. doi:10.1073/pnas.1321173111.
44. Soverini S, De Benedittis C, Polakova KM, et al. Unraveling the complexity of tyrosine kinase inhibitor-resistant populations by ultra-deep sequencing of the BCR-ABL kinase domain. *Blood*. 2013;122(9):blood–2013–03–487728–1648. doi:10.1182/blood-2013-03-487728.
45. Davis MI, Hunt JP, Herrgard S, et al. Comprehensive analysis of kinase inhibitor selectivity. *Nature Biotechnology*. 2011;29(11):1046-1051. doi:10.1038/nbt.1990.
46. Pemovska T, Johnson E, Kontro M, et al. Axitinib effectively inhibits BCR-ABL1(T315I) with a distinct binding conformation. *Nature*. 2015;519(7541):102-105. doi:10.1038/nature14119.
47. Smith CC, Lasater EA, Zhu X, et al. Activity of ponatinib against clinically-relevant AC220-

References (continued)

resistant kinase domain mutants of FLT3-ITD. *Blood*. 2013;121(16):3165-3171. doi:10.1182/blood-2012-07-442871.

48. Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol*. 1993;234(3):779-815. doi:10.1006/jmbi.1993.1626.
49. Dong GQ, Fan H, Schneidman-Duhovny D, Webb B, Sali A. Optimized atomic statistical potentials: assessment of protein interfaces and loops. *Bioinformatics*. 2013;29(24):3158-3166. doi:10.1093/bioinformatics/btt560.
50. Mackinlay J, Hanrahan P, Stolte C. Show me: automatic presentation for visual analysis. *IEEE Trans Vis Comput Graph*. 2007;13(6):1137-1144. doi:10.1109/TVCG.2007.70594.
51. Gorre ME, Mohammed M, Ellwood K, et al. Clinical Resistance to STI-571 Cancer Therapy Caused by BCR-ABL Gene Mutation or Amplification. *Science*. 2001;293(5531):876-880. doi:10.1126/science.1062538.
52. How GF, Müller MC, Lim LC. Acquisition of the novel ABL kinase domain mutation T315L in a relapsed Philadelphia-positive acute lymphoblastic leukemia patient. *Leuk Res*. 2012;36(1):e20-e21. doi:10.1016/j.leukres.2011.08.011.
53. Redaelli S, Mologni L, Rostagno R, et al. Three novel patient-derived BCR/ABL mutants show different sensitivity to second and third generation tyrosine kinase inhibitors. *Am J Hematol*. 2012;87(11):E125-E128. doi:10.1002/ajh.23338.
54. Ilaria RL, Van Etten RA. P210 and P190(BCR/ABL) induce the tyrosine phosphorylation and DNA binding activity of multiple specific STAT family members. *J Biol Chem*. 1996;271(49):31704-31710.
55. Chen Y, Tortorici MA, Garrett M, Hee B, Klamerus KJ, Pithavala YK. Clinical pharmacology of axitinib. *Clin Pharmacokinet*. 2013;52(9):713-725. doi:10.1007/s40262-013-0068-3.
56. Man CH, Fung TK, Ho C, et al. Sorafenib treatment of FLT3-ITD(+) acute myeloid leukemia: favorable initial outcome and mechanisms of subsequent nonresponsiveness associated with the emergence of a D835 mutation. *Blood*. 2012;119(22):5133-5143. doi:10.1182/blood-2011-06-363960.
57. Cortes JE, Tallman MS, Schiller G, et al. Results Of a Phase 2 Randomized, Open-Label, Study Of Lower Doses Of Quizartinib (AC220; ASP2689) In Subjects With FLT3-ITD Positive Relapsed Or Refractory Acute Myeloid Leukemia (AML). *Blood*. 2013;122(21):494-494.
58. Shah NP, Talpaz M, Deininger MWN, et al. Ponatinib in patients with refractory acute myeloid

References (continued)

leukaemia: findings from a phase 1 study. *Br J Haematol*. 2013;162(4):548-552. doi:10.1111/bjh.12382.

59. Randhawa JK, Kantarjian HM, Borthakur G, et al. Results of a Phase II Study of Crenolanib in Relapsed/Refractory Acute Myeloid Leukemia Patients (Pts) with Activating FLT3 Mutations. *Blood*. 2014;124(21):389-389.
60. Smith CC, Wang Q, Chin C-S, et al. Validation of ITD mutations in FLT3 as a therapeutic target in human acute myeloid leukaemia. *Nature*. 2012;485(7397):260-263. doi:10.1038/nature11016.
61. Baker SD, Zimmerman EI, Wang Y-D, et al. Emergence of Polyclonal FLT3 Tyrosine Kinase Domain Mutations during Sequential Therapy with Sorafenib and Sunitinib in FLT3-ITD–Positive Acute Myeloid Leukemia. *Clin Cancer Res*. 2013;19(20):5758-5768. doi:10.1158/1078-0432.CCR-13-1323.
62. Smith CC, Zhang C, Lin KC, et al. Characterizing and Overriding the Structural Mechanism of the Quizartinib-Resistant FLT3 "Gatekeeper" F691L Mutation with PLX3397. *Cancer Discov*. 2015;5(6):668-679. doi:10.1158/2159-8290.CD-15-0060.
63. Smith CC, Lasater EA, Lin KC, et al. Crenolanib is a selective type I pan-FLT3 inhibitor. *Proc Natl Acad Sci USA*. 2014;111(14):5319-5324. doi:10.1073/pnas.1320661111.
64. Griffith J, Black J, Faerman C, et al. The structural basis for autoinhibition of FLT3 by the juxtamembrane domain. *Mol Cell*. 2004;13(2):169-178.
65. Baker SD, Zimmerman EI, Wang Y-D, et al. Emergence of polyclonal FLT3 tyrosine kinase domain mutations during sequential therapy with sorafenib and sunitinib in FLT3-ITD-positive acute myeloid leukemia. *Clin Cancer Res*. 2013;19(20):5758-5768. doi:10.1158/1078-0432.CCR-13-1323.
66. Galanis A, Ma H, Rajkhowa T, et al. Crenolanib is a potent inhibitor of FLT3 with activity against resistance-conferring point mutants. *Blood*. 2014;123(1):94-100. doi:10.1182/blood-2013-10-529313.
67. Fairman R, Shoemaker KR, York EJ, Stewart JM, Baldwin RL. Further studies of the helix dipole model: effects of a free alpha-NH3+ or alpha-COO- group on helix stability. *Proteins*. 1989;5(1):1-7. doi:10.1002/prot.340050102.
68. Hol WG, van Duijnen PT, Berendsen HJ. The alpha-helix dipole and the properties of proteins. *Nature*. 1978;273(5662):443-446.

Chapter 2: Activation and Inhibition of Organic Cation Transporters

The first organic cation transporter (OCT) was cloned in 1994 from rats,[1] followed closely by the human OCT1 in 1997[2,3] and the other members of the human OCT family, OCT2 and OCT3.[4] As facilitative diffusion uniporters, they are important in the transport of monovalent and divalent cations, weak bases, and some neutral molecules.[2,5] OCT1 is primarily expressed on the basolateral membrane of hepatocytes,[6] OCT2 on the basolateral membrane of renal proximal tubules,[7] while OCT3 is more ubiquitously expressed in the heart, skeletal muscle, brain, small intestine, liver, lung, kidneys, and other tissues.[8,9]

OCTs have been primarily studied from the standpoint of drug absorption and excretion because they are present in pharmacologically most important tissues - intestine, liver, and kidney. OCT1 and OCT2 transport the blockbuster antidiabetic drug metformin,[9-11] nucleoside antivirals zidovudine, dalcitabine, lamivudine, and tenofovir,[12] anticancer drug daunorubicin,[13], and platinum-containing drugs oxaliplatin,[14] cisplatin,[15] and picoplatin,[16] to name only a few.[5] Their polyspecificity allows them to transport a wide variety of structurally dissimilar endogenous molecules, including the recently discovered thiamine,[17,18] metabolites choline and creatinine, and neurotransmitters acetylcholine, dopamine, epinephrine, and serotonin.[8]

In addition to its role in the distribution of platinum-containing antineoplastic compounds, OCT1 has recently been identified as a mediator of bendamustine uptake and cytotoxicity in chronic lymphocytic leukemia (CLL) cells.[19] Bendamustine was approved by the FDA in 2008 and is sold under the name Treanda for the treatment of several hematological malignancies.[20] Also, OCT1 is an important clinical determinant of imatinib treatment response in CML patients. Patients with high pre-treatment expression of OCT1 mRNA had a higher probability of achieving a cytogenetic response and a superior progression-free and overall survival.[21] In a separate study, the GG genotype of SLC22A1 rs683369, corresponding to a missense mutation L160F, correlated with high rate of loss of response or treatment failure to imatinib therapy.[22] It has since been shown that imatinib is not a substrate of OCTs,[23] and is in fact an inhibitor of all human OCTs.[24] One proposed explanation for these inconsistencies is that OCT1 expression is a surrogate for expression of other transporters relevant to imatinib distribution, including OATP1A2, which does transport imatinib.[23]

Although no crystal structures have been determined for the human OCTs or any of their close homologs, OCTs are members of the major facilitator superfamily (MFS) and are predicted to have the same MFS protein fold composed of 12 transmembrane helices, a large extracellular loop between TM 1/2, and a large intracellular loop between TM seen in the crystal structures of lactose permease LacY[25,26] and phosphate transporter PiPT (**Figure 2.1**).[27] Studies on residues lining the translocation pathway, explorations in substrate specificity, and identification of the more than one ligand binding site with different affinities have all been achieved through mutagenesis studies guided by comparative structure modeling of human OCTs.[28-32] Further, the extracellular loop has been shown to be required for plasma membrane targeting and homo-oligomerization between monomers, for both OCT1 and OCT2.[33,34] In line with other MFS transporters, their mechanism of

transport is hypothesized to follow the "rocker-switch" mechanism in which the pseudo-symmetrical N- and C-terminal halves of the transporter rock back and forth against each other, along the axis defined by the domain interface.[35]

This chapter addresses two longstanding questions in the field of human organic compound transport. How are human OCTs activated? And what inhibits them? The first section addresses these questions by examining the relationship between the Src family kinase Yes1 and OCT2. It shows that phosphorylation of OCT2 by Yes1 *in vitro* is critical to its function, while Yes1 inhibition *in vivo* diminished OCT2 activity. This study adds to the current understanding of the interplay between kinases and SLC transporters in the phosphoproteome.[36] In fact, the presence of several potential protein kinase phosphorylation sites in the intracellular loops is a common feature among OCTs, suggesting that their activity can be subjected to kinase regulation.[37] Previous work has shown that rat OCT1 is activated by protein kinase C-dependent phosphorylation.[38] In an equivalent study with human OCT2, PKA stimulation resulted in inhibition, while inhibition of calmodulin reduced the apparent affinity of OCT2 for TEA, suggesting a direct effect of this pathway on the molecular structure of the binding site of hOCT2.[37] TKIs might therefore have additional functions as modulations of pharmacologically important SLC transporters, outside of their direct inhibition through competitive binding.[24]

The second section focuses solely on the question of what inhibits them, through the lens of OCT1 with a combination of high-throughput screening, *in silico* molecular docking against a comparative structure model, structure-activity relationship (SAR) modeling, and cell based validation of hits. By using a comprehensive, multifaceted approach, the study illustrates potential new clinical drug interactions and identifies the physicochemical properties characterizing OCT1 inhibitors.

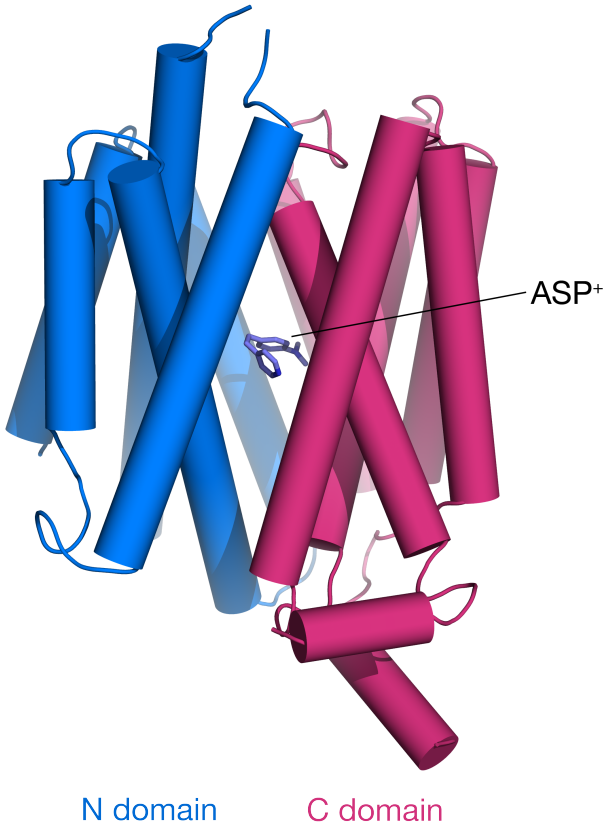


Figure 2.1 Comparative structure model of human OCT1 with a docked pose of ASP+. Model uses the co-crystal structure of PiPT (PDB ID 4J05)(ref. 27) with a phosphate ion in the inward-facing occluded state as the template.

A phosphotyrosine switch regulates organic cation transporters

The contents of this section have been published in the following article:

A phosphotyrosine switch regulates organic cation transporters.

*Jason A. Sprowl, Su Sien Ong, Alice A. Gibson, Shuiying Hu, Guoqing Du, Wenwei Lin, Lie Li, Shashank Bharill, Rachel A. Ness, *Adrian Stecula, Steven M. Offer, Robert B. Diasio, Anne T. Nies, Matthias Schwab, Guido Cavaletti, Eberhard Schlatter, Giuliano Ciarimboli, Jan H. M. Schellens, Ehud Y. Isacoff, Andrej Sali, Taosheng Chen, Sharyn D. Baker, Alex Sparreboom, and Navjotsingh Pabla*
Nature Communications. 2016;7:10880.

*The author's contribution included comparative structure modeling of the human OCT2 to identify the location of the phosphorylation site and guide mutagenesis studies.

Introduction

An enclosed but selectively permeable cellular membrane is a ubiquitous feature of all life forms.[39,40] Membrane transporters are the evolutionally conserved gatekeepers that govern this selective cellular permeability.[41] They regulate the uptake and efflux of essential molecules such as amino acids, nucleosides, sugars, inorganic ions as well as therapeutic compounds.[41,42] Therefore, transporters have wide ranging influences on normal human physiology and pathophysiology and are key determinants of therapeutic response to drugs.

The human genome is thought to encode more than 400 membrane transporter genes belonging to two major superfamilies: ATP-binding cassette (ABC) transporters and solute carriers (SLC), which are involved in almost every essential biological process.[42-45] Among these, about 20 'multispecific' transporters belonging to either superfamily have been extensively implicated in drug transport.[44,46,47] Drug transporters are highly expressed in the intestine, kidney, liver and endothelial barriers, where they regulate absorption, distribution, metabolism and excretion of drugs.[44,46] At the cellular level, transporter-mediated uptake or efflux can impart drug-sensitive or -resistant phenotypes in target cells, thereby affecting therapeutic efficacy.[48] Likewise, transporter-mediated uptake in non-target tissues can contribute to drug toxicities.[47,49] As a result, along with drug-metabolizing enzymes, transporters have emerged as critical determinants of drug disposition, therapeutic efficacy and adverse drug reactions.

Due to their predominant role in determining clinical response to therapeutics, multiple regulatory aspects of drug transporters have been widely studied. Genetic polymorphisms,[50] epigenetic mechanisms,[51] dietary ingredients,[52] and drug–drug interactions[44] that functionally modulate drug transporters can profoundly affect therapeutic outcomes. However, surprisingly, phosphotyrosine-mediated regulation of drug transporters has not been extensively studied. Here we report phosphotyrosine-mediated regulation of several pharmacologically important SLC-family transporters, including organic cation transporters (OCTs),[53] multidrug and toxin-extrusion transporters (MATEs),[54] and organic anion transporting polypeptides (OATPs).[55] Notably, we propose that clinically used tyrosine kinase inhibitors (TKIs)[56] can inhibit protein kinases required for tyrosine phosphorylation of drug transporters, thereby influencing transporter function.

Results

Small-molecule screen identifies potent OCT2 inhibitors The SLC-family member OCT2 is expressed in the renal tubular cells,[53] dorsal root ganglia (DRG),[57] and brain,[58] where it regulates uptake of endogenous organic cations like creatinine[59] and catecholamines.[60] OCT2 is also involved in the uptake of widely used therapeutics including the anti-diabetic drug metformin[11] and platinum-based chemotherapeutics.[57,61] The three major debilitating side effects of platinum-based anti-cancer drugs, namely peripheral neurotoxicity,[57] nephrotoxicity,[61,62] and ototoxicity,[63] are dependent on OCT2-mediated uptake of cisplatin or oxaliplatin in DRGs, renal tubular cells, and hair cells of the cochlea, respectively. Identification of potent OCT2 inhibitors, which could be combined with chemotherapy, has the potential to provide clinical benefits by reducing these toxicities. Thus, with the aim to identify OCT2 inhibitors, we carried out a small-molecule screen using the St Jude Children's Research Hospital bioactive compound library (8086 compounds). For this screen, we utilized OCT2 expressing HEK293 cells and uptake of fluorescent substrate 4-(4-(dimethylamino)styryl)-N-methylpyridinium (ASP^+) was used as an indicator of OCT2 function (**Figure 2.2A**). Imipramine, a known OCT2 inhibitor, was used as a positive control (**Figure 2.2A**). We identified 433 compounds that exhibited >90% inhibition of OCT2 function (**Figure 2.2B**). About half of these active compounds were neurological drugs, which was not surprising because they are structurally similar to endogenous OCT2 substrates like norepinephrine and serotonin.[58] Due to their undesirable neurological effects these agents were not explored further.

Multiple TKIs also inhibited OCT2 function, and since they are widely used as anti-cancer agents[56] we pursued them further (**Figure 2.2B**). FDA-approved TKIs were tested for their ability to inhibit OCT2 in secondary screens using tetraethylammonium (TEA) as OCT2 substrate (**Figure 2.2C**). Dose–response experiments in OCT2 expressing Hela (Hela-OCT2) and HEK293 (HEK293-OCT2) cells showed that multiple TKIs can inhibit OCT2 function at sub-micromolar concentrations (**Figure S2.1A**). Creatinine is an OCT2 substrate and interestingly, the TKIs (bosutinib, dasatinib, nilotinib, pazoponib, sunitinib and vandetinib) that inhibited OCT2 function *in vitro* can cause increase in serum creatinine levels (**Figure S2.1B**) in patients, which could be a result of reduced creatinine excretion as a result of renal OCT2 inhibition. Among these TKIs, dasatinib,[56,64] an oral BCR-ABL and Src family kinase inhibitor approved for the treatment of leukemia, was found to be the most potent OCT2 inhibitor (**Figure 2.2C**), with OCT2 inhibition seen at nanomolar concentrations (**Figure 2.2D**).

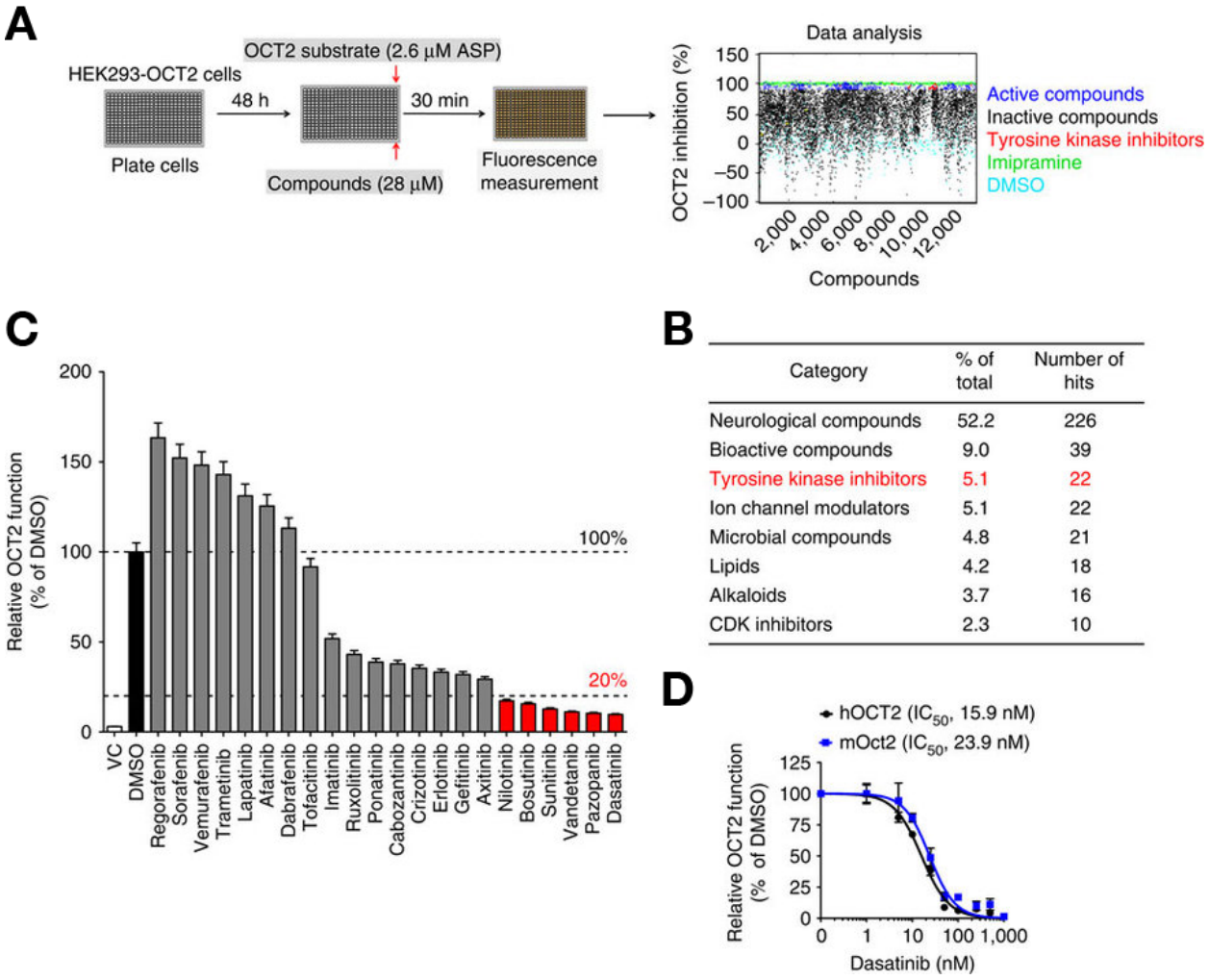


Figure 2.2 Small molecule HTS of OCT2 inhibitors. A) Scheme depicting the assay conditions used in the primary screen for OCT2 inhibitors. HEK293-OCT2 cells were plated in 384-well plates, followed by incubation with small molecules and uptake assay using ASP+ as OCT2 substrate. Imipramine and DMSO were used as positive and negative controls respectively. Compounds that inhibited OCT2 activity by ≥90% were considered as active compounds. B) Active compounds (433 out of 8,086) were categorized into distinct groups and the categories with more than 2% hits are shown here. C) Secondary screens were carried out using Hela-OCT2 cells. These cells were preincubated with 10μM concentration of TKIs for 15 min, followed by incubation with OCT2 substrate [14C]-TEA. Data are presented as percentage OCT2 activity (TEA uptake) as compared to DMSO group. D) HEK293-hOCT2 and HEK293-mOct2 cells were preincubated with varying concentrations of dasatinib for 15min, followed by incubation with 2μM [14C]-oxaliplatin for 15min. Data are presented as percentage OCT2 activity (oxaliplatin uptake) as compared with DMSO group. All experimental values are presented as mean±s.e. The height of error bar = 1 s.e.

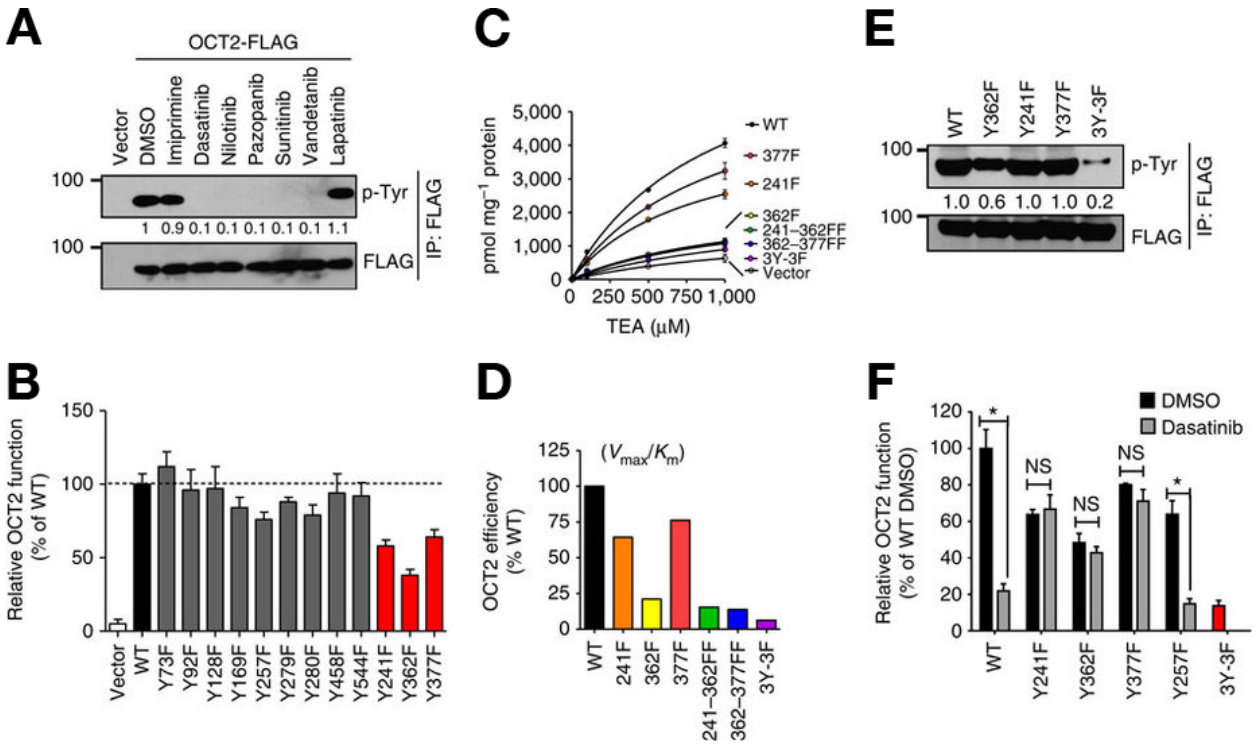


Figure 2.3 Functional regulation of OCT2 by tyrosine phosphorylation. A) Hela-Vector or Hela-OCT2 cells were treated with either DMSO, TKIs (1 mM) or 500 mM imipramine for 30 min. Cell lysates were then used for immunoprecipitation of FLAG-OCT2 with mouse anti-FLAG antibodies, followed by western blot analysis with rabbit phospho-tyrosine and FLAG antibodies. B) Plasmids for OCT2 mutants were transiently transfected into Hela cells and 24 h later, uptake assays (15 min) were performed using [14C]-TEA (2 mM). TEA uptake levels were normalized to protein levels in each group. The graph represents relative OCT2 function (TEA uptake) as compared with wild-type OCT2 transfected group. C,D) Plasmids for indicated OCT2 mutants were transiently transfected into Hela cells and 24 h later, uptake assays were performed using varying concentration of [14C]-TEA. TEA uptake levels were normalized to protein levels in each group and graphs represent Michaelis–Menten kinetics and calculated ratio of K_m/V_{max} values, which are indicative of transporter activity. E) Hela cells were transiently transfected with indicated FLAG-tagged OCT2 constructs, followed by immunopreicipation with anti-FLAG bead conjugated antibodies and western blot analysis by FLAG and phosphotyrosine antibodies. F) Hela cells were transiently transfected with indicated plasmids, followed by TEA uptake assays in the presence of DMSO or 1 mM dasatinib. * indicates statistically significant as compared to respective DMSO treated group (Student's t-test). All experimental values are presented as mean±s.e. The height of error bar = 1 s.e.

We next considered if TKIs inhibited OCT2 through a competitive mechanism. We found that OCT2 inhibiting TKIs, including dasatinib were not OCT2 substrates (**Figure S2.1C**) and dasatinib-mediated OCT2 inhibition was found to be non-competitive (**Figure S2.1D**). In addition, dasatinib-mediated OCT2 inhibition was found to be reversible (**Figure S2.1E**). A previous study[24] also reported dasatinib-mediated OCT2 inhibition albeit with significantly less potency. In that study,[24] the substrates and inhibitors were co-incubated, while we preincubated cells with dasatinib. Indeed, preincubation with TKIs was sufficient and, in fact, more effective in inhibiting OCT2 function than co-incubation (**Figure S2.1F**).

Tyrosine phosphorylation is essential for OCT2 function To gain further mechanistic insights, we performed surface biotinylation assays and found that TKIs did not affect membrane OCT2 expression (**Figure S2.2A**). Although tyrosine phosphorylation of OCTs has not been studied, we considered the possibility that TKIs might inhibit OCT2 function by modulating its tyrosine phosphorylation. Indeed, immunoprecipitation (IP) of FLAG-OCT2 from Hela-OCT2 cells showed that OCT2 was tyrosine phosphorylated (**Figure 2.3A**). Importantly, OCT2-inhibiting TKIs (dasatinib, nilotinib, pazoponib, sunitinib and vandetinib) also inhibited its tyrosine phosphorylation. This was not observed for lapatinib, a TKI that did not inhibit OCT2 as well as for imipramine, a competitive OCT2 inhibitor (**Figure 2.3A**). Tyrosine phosphorylation of OCT2 was further corroborated by reverse-IP experiments, where we immunoprecipitated total tyrosine phosphorylated proteins using a phosphotyrosine antibody and could pull-down FLAG-OCT2 (**Figure S2.2B**). Moreover, tyrosine phosphorylation of OCT2 was inhibited by dasatinib in a time-dependent manner, which indicated that OCT2 tyrosine phosphorylation could be a dynamic process (**Figure S2.2C**).

To identify the phosphotyrosine sites in OCT2 we utilized computational prediction tools[65] and PhosphoSite,[66,67] a database of post-translational modifications derived from discovery-mode tandem mass spectrometry (MS) experiments. Subsequently, functional assays of several tyrosine-to-phenylalanine OCT2 mutants showed that mutations at three sites (241, 362 and 377) substantially reduced OCT2 function, without affecting membrane OCT2 expression (**Figure 2.3B** and **Figure S2.2D,E**). These three sites were originally selected on the basis of MS data in PhosphoSite,[66,67] which showed that the Y241 site in OCT2 is tyrosine phosphorylated while sites corresponding to OCT2 Y362 and Y377 are tyrosine phosphorylated in the closely related transporters OCT1 and OCTN2. Interestingly, tyrosine-to-phenylalanine mutations at

computationally predicted tyrosine sites had no effect on OCT2 function, while all the three sites for which MS data exist in PhosphoSite were found to be important for OCT2 function. Further kinetic experiments showed that the three mutants had reduced OCT2 function, with Y362F mutation having the most dramatic effect (**Figure 2.3C,D**). Immunoprecipitation experiments showed that the Y362F mutant had significantly reduced tyrosine phosphorylation, which was further diminished in the triple-mutant (Y241F-Y362F-Y377F, denoted as 3Y–3F; **Figure 2.3E**). These three sites are evolutionarily conserved (**Figure S2.2F**). While these OCT2 mutants had lower baseline function, they were resistant to dasatinib-mediated OCT2 inhibition (**Figure 2.3F**).

To understand how post-translational modifications affect OCT2 function, we considered the possibility that tyrosine phosphorylation may regulate OCT2 oligomerization. However, single-molecule subunit analysis[68,69] experiments showed that OCT2 most likely exists as a monomer (**Figure S2.3A-C**).

Although the crystal structure of OCT2 is not known, previous OCT1 structural models[27] and new OCT2 modeling data (**Figure S2.3D**) suggest that the Y362 site is localized close to the substrate-binding domain. This raises the possibility that the negative charge provided by a phosphotyrosine may enhance substrate binding of positively charged organic cations. However, we cannot rule out the possibility that the Y362F mutant has reduced function due to diminished substrate binding, but the presence of MS data[66] showing tyrosine phosphorylation at this site in OCT1 combined with our results that Y362F OCT2 mutant has reduced tyrosine phosphorylation and is resistant to dasatinib-mediated OCT2 inhibition suggests that the reduced function of Y362F mutant could be due to decreased tyrosine phosphorylation.

siRNA screen identifies Yes1 as OCT2 phosphorylating kinase To identify upstream kinases involved in OCT2 phosphorylation, we performed a primary siRNA screen using a human protein kinase siRNA library (779 genes; Dharmacon) to knockdown protein kinases in HEK293-OCT2 cells followed by functional assays (**Figure 2.4A**). The siRNAs (tyrosine kinases) that reduced OCT2 function to ≤75% without affecting cell viability were selected for further consideration (**Figure 2.4B**). Next, we performed a deconvolution secondary screen with ASP^+ as OCT2 substrate (**Figure 2.4C**) and an additional secondary screen utilizing pooled siRNAs (Sigma) with TEA as OCT2 substrate (**Figure 2.4C** and **Figure S2.4**). On the basis of the positive hits from both these secondary screens, we narrowed down the putative targets to KDR, LYN and Yes1 (**Figure 2.4C**). To

unambiguously identify the TKI-sensitive protein kinase that phosphorylates OCT2, we carried out tertiary screens utilizing a chemical genetics approach.[70] The TKI resistance of gatekeeper mutants of KDR, LYN and Yes1 was initially confirmed (**Figure S2.5A**). We then transfected Hela-OCT2 cells with either the wild-type or TKI-resistant (gatekeeper) mutants[70] of KDR, LYN and Yes1, followed by dasatinib treatment and OCT2 uptake assays. As shown in **Figure 2.4D**, only the TKI-resistant Yes1 mutant was able to rescue OCT2 inhibition by dasatinib. The TKI-resistant Yes1 mutant was also able to significantly rescue OCT2 inhibition by other TKIs (**Figure S2.5B**). Moreover, Yes1 knockdown significantly reduced OCT2 tyrosine phosphorylation (**Figure 2.4E**). Indeed, previous analysis of kinase inhibition selectivity[71] has shown that all the TKIs that inhibited OCT2 function can inhibit Yes1 (**Figure S2.6A**). Moreover, other non-TKI Yes1 inhibitors like dorsomorphin[72] were also found to be potent OCT2 inhibitors (**Figure S2.6B**). Importantly, OCT2 phosphotyrosine sites have sequences similar to known Src family kinase substrates (**Figure S2.7A**). *In vitro* kinase assays with purified Yes1 and OCT2 proteins showed that Yes1 can phosphorylate OCT2 and this tyrosine phosphorylation is significantly reduced in the Y362F and 3Y–3F mutants (**Figure S2.7B**). These data suggest that Yes1 is the TKI-sensitive kinase that can directly phosphorylate OCT2. On the basis of these studies, we propose that these three tyrosines may be phosphorylated by Yes1, but Y362 might be the major phosphotyrosine site that has greater functional relevance. On the basis of these studies, Yes1 inhibition is likely a key mechanism of dasatinib-mediated OCT2 inhibition, however, the role of other kinases as well as other non-kinase mechanisms cannot be excluded for other TKIs.

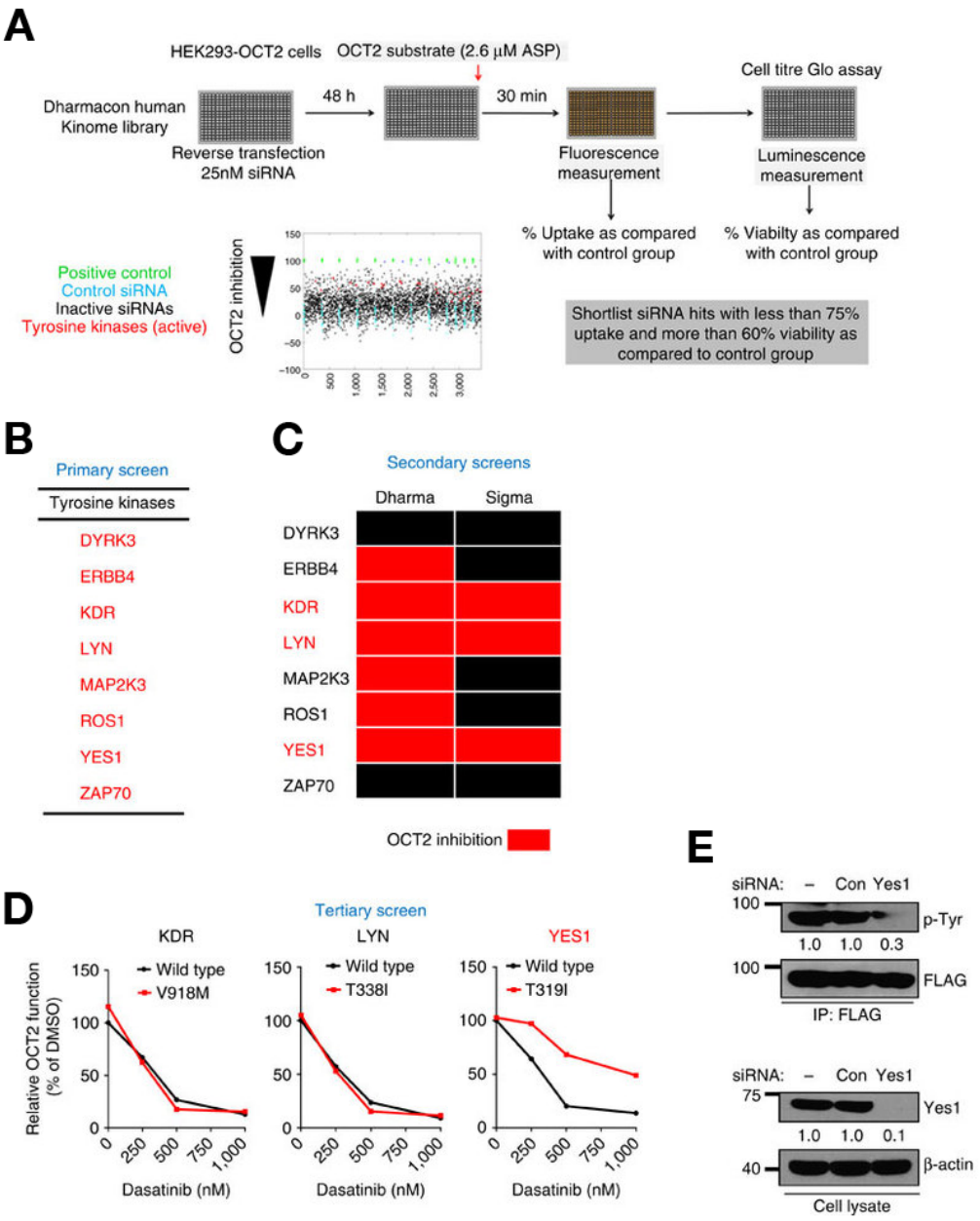


Figure 2.4 Functional regulation of OCT2 by tyrosine phosphorylation. A) Scheme depicting the assay conditions used in the primary siRNA kinome screen to identify an OCT2 phosphorylating kinase. HEK293-OCT2 cells were reverse transfected with the siRNA library and plated in 384-well plates, followed by incubation functional uptake and viability assays. B) Positive hits from the primary screen. Only the tyrosine kinase hits (indicated in red) were used for secondary screens. C) Schematic representation of secondary screens: deconvoluted screen using Dharmacon siRNA and Sigma siRNA screen. The siRNA that inhibited OCT2 function to at least 75% are indicated in red. D) Hela-OCT2 cells were transfected with either wild-type or dasatinib resistant KDR, LYN or Yes1 plasmids and 24 h later, [14C]-TEA uptake assays were performed in the presence or absence of varying concentrations of dasatinib. The graph represents relative OCT2 function ([14C]-TEA uptake) as compared with DMSO group for each plasmid. E) Hela-OCT2 cells were transfected with Sigma siRNA (scrambled control or Yes1) and 48 h later, OCT2 was immunoprecipitated to determine its tyrosine phosphorylation. Whole-cell lysate was used to confirm Yes1 knockdown.

Tyrosine phosphorylation is conserved in OCTs While further exploring the OCT2 regulation by Yes1, we noticed that OCT2 has a proline-rich (PXXPR) sequence, which is known to bind Src Homology 3 (SH3)[73,74] domain present in Yes1 (**Figure 2.5A**). Indeed, co-immunoprecipitation experiments showed that Yes1 can physically associate with OCT2 (**Figure 2.5B**), and mutations in the proline-rich SH3-binding domain reduced OCT2 function (**Figure 2.5C**) and tyrosine phosphorylation (**Figure 2.5D**). Although Yes1 is a non-receptor tyrosine kinase, it is both myristoylated and palmitoylated, which allows it to tether to membranes both in the endoplasmic reticulum (ER) and the plasma membrane.[75] Tyrosine phosphorylated OCT2 was present in both the ER and plasma-membrane fractions (**Figure S2.7C**) and it is likely that Yes1 might phosphorylate OCT2 in the ER and/or at the plasma membrane (**Figure 2.5E**). The OCT2 phosphotyrosine sites are conserved in the related transporters OCT1 and OCT3, which are also inhibited by dasatinib (**Figure S2.8A,B**). The proline-rich SH3-binding sequence is also conserved in OCT1, OCT3, OCTN1 and OCTN2 (**Figure S2.8A**). In humans, a naturally occurring single nucleotide variant in the OCT1 gene, causing a P283L change, is known to reduce OCT1 function.[76] Interestingly, this site is located in the proline-rich SH3 binding sequence of OCT1. To examine whether the phosphorylation-mediated regulation of OCT2 is conserved in OCT1 and OCT3, we carried out functional assays after mutagenesis of relevant sites. As shown in **Figure 2.5F**, OCT1 and OCT3 mutants that lacked either the putative phosphorylation sites or the proline-rich motif had significantly reduced function.

The existence of a conserved regulatory mechanism in OCT1, OCT2, and OCT3 led us to question if other drug transporters are also regulated by tyrosine phosphorylation. Interestingly, MS data from global phosphoproteome studies[66,67] suggest that a substantial fraction of clinically relevant transporters, drug-metabolizing enzymes, and ion channels have sites that are tyrosine phosphorylated (**Figure S2.9**). To test whether these phosphotyrosine modifications have a functional role, we generated tyrosine-to-phenylalanine mutants for the SLC-family drug-transporters MATE1[54] and OATP1B1.[55] Mutations at these conserved sites significantly reduced transporter activity, indicating functional relevance (**Figure S2.10A-D**). Moreover, distinct TKIs can inhibit these SLC-family drug transporters (**Figure S2.11**). These studies suggest that tyrosine phosphorylation may be a potentially widespread mechanism of drug transporter regulation that might be a target of deregulation by clinically used TKIs.

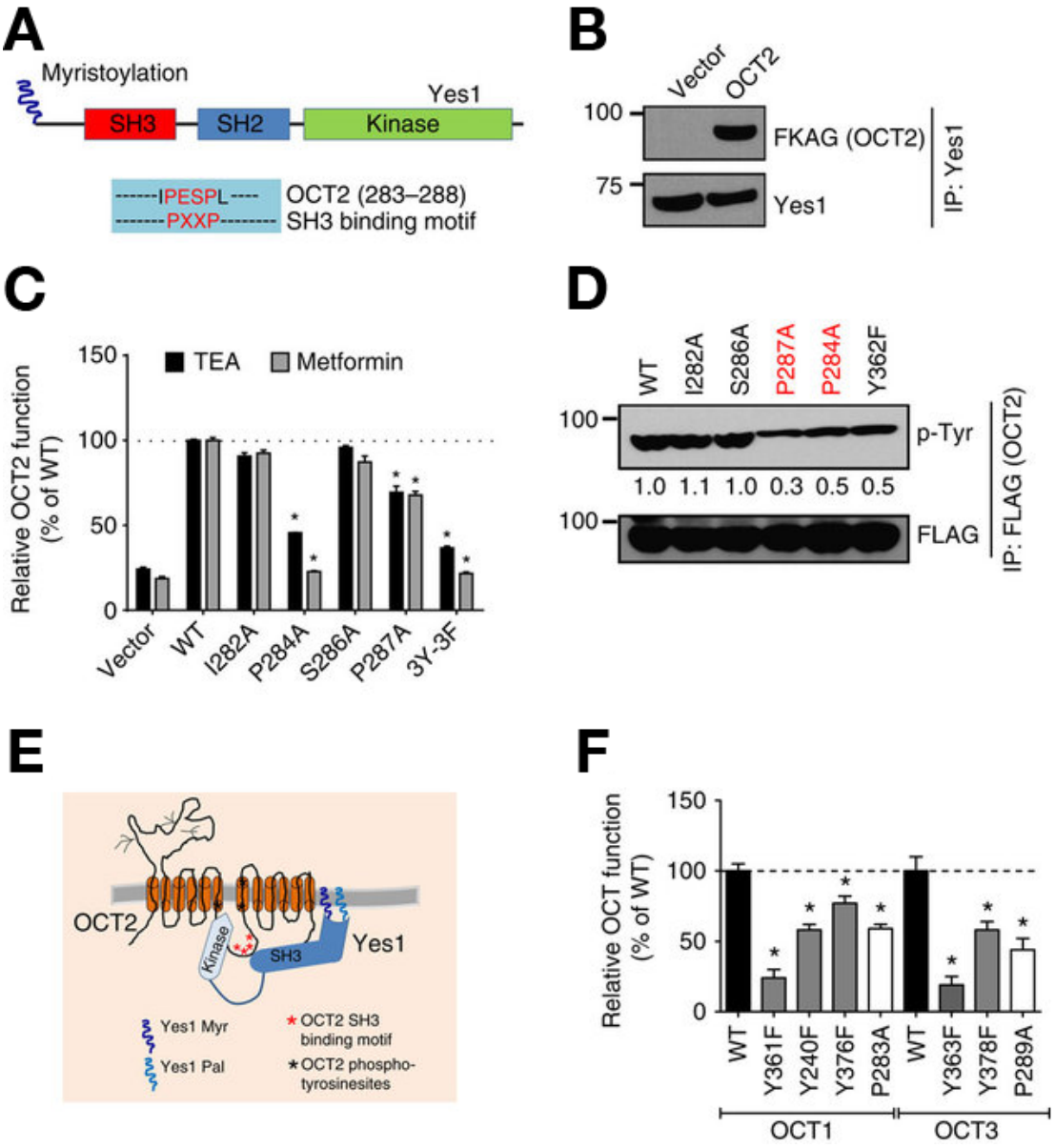


Figure 2.5 Yes1-mediated regulation of OCT2 tyrosine phosphorylation. A) Schematic representation of Yes1 protein (upper panel) showing the SH3 domain. The lower panel shows the putative proline-rich SH3 binding sequence in OCT2. B) Endogenous Yes1 was immunoprecipitated from Hela-Vector and Hela-OCT2 cell lysates using a mouse anti-Yes1 antibody, followed by western blot analysis with rabbit anti-FLAG and Yes1 antibodies. C) Plasmids for OCT2 mutants were transiently transfected into Hela cells and 24 h later, uptake assays (15 min) were performed using [14C]-TEA (2 mM) or [14C]-metformin (50 mM). The uptake levels were normalized to protein concentration in each group. The graph represents relative OCT2 function (TEA or metformin uptake) as compared to wild-type OCT2 transfected group. * indicates statistically significant as compared with wild-type group ($P<0.05$, Student's t-test). D) Hela cells were transiently transfected with indicated FLAG-tagged OCT2 constructs, followed by immunopreicipation with anti-FLAG antibodies and western blot analysis by FLAG and phosphotyrosine antibodies. E) Proposed model of Yes1 and OCT2 interaction. F) Plasmids for OCT1 and OCT3 mutants were transiently transfected in Hela cells and 24 h later, uptake assays (15 min) were performed using [14C]-TEA (2 mM). The uptake levels were normalized to protein concentration and the graph represents relative OCT2 function (TEA uptake) as compared with respective wild-type group. * indicates statistically significant as compared to wild-type group ($P<0.05$, Student's t-test). All experimental values are presented as mean±s.e. The height of error bar = 1 s.e.

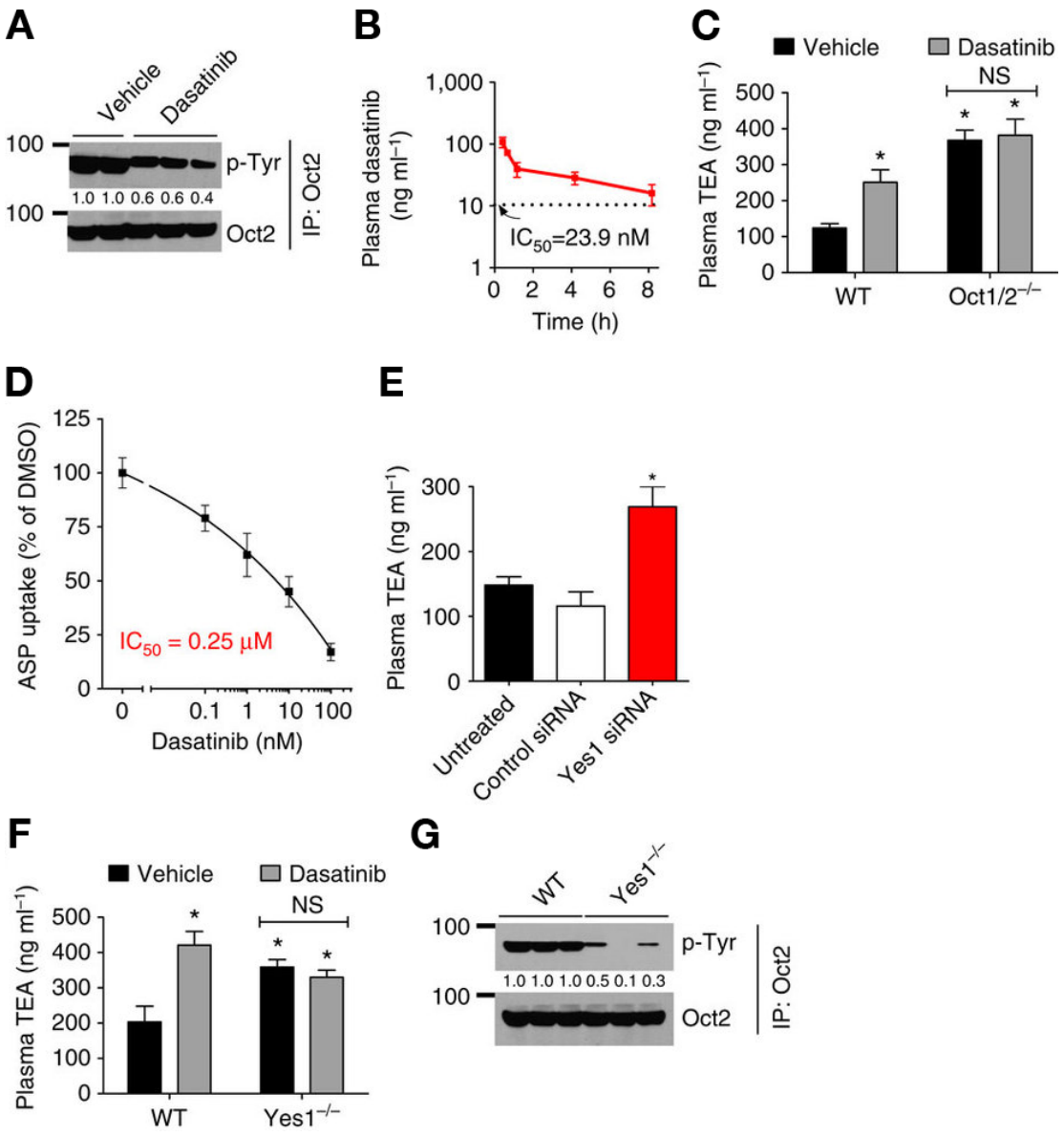


Figure 2.6 OCT2 tyrosine phosphorylation and functional regulation in vivo. A) Wild-type FVB mice were injected with either vehicle or dasatinib (15 mg/kg, p.o.) and 30 min later, the mice were euthanized and the kidneys were collected. Kidney tissue lysates were then used to immunoprecipitate endogenous Oct2 followed by western blot analysis by phosphotyrosine and Oct2 antibodies. B) Male FVB mice were injected with dasatinib (15 mg/kg) followed by pharmacokinetic analysis of dasatinib levels in the plasma. C) Wild-type and Oct1/2 -/- mice were injected with either vehicle or dasatinib (15 mg/kg, p.o.) and 30 min later they were injected with 0.2 mg/kg [14C]-TEA (i.v.), followed by plasma collection at 5 min. The graph represents plasma TEA levels from n = 5 mice per group. * indicates statistically significant as compared with wild-type vehicle group ($P<0.05$, Student'st-test). D) Isolated renal tubules were co-incubated with dasatinib in the presence of the OCT2 substrate ASP+ (30 min), and relative uptake was measured compared with control group. E) Wild-type FVB mice were injected with either control or Yes1 siRNA by hydrodynamic tail-vein injection (25 mg in 0.5 ml of PBS). Three days later, the mice were injected i.v. with a 0.2 mg/kg dose of [14C]-TEA, and plasma levels of TEA were measured at 5 min (n = 5 mice per group). Kidneys were also collected to determine Yes1 knockdown. F) Wild-type and Yes1 -/- mice (n = 5) were injected with 0.2 mg/kg [14C]-TEA (i.v.) and plasma levels of TEA were measured at 5 min. G) Kidney tissue lysates from wild-type and Yes1 -/- were used to immunoprecipitate endogenous Oct2 followed by western blot analysis by phosphotyrosine and Oct2 antibodies. All experimental values are presented as mean±s.e. The height of error bar = 1 s.e.

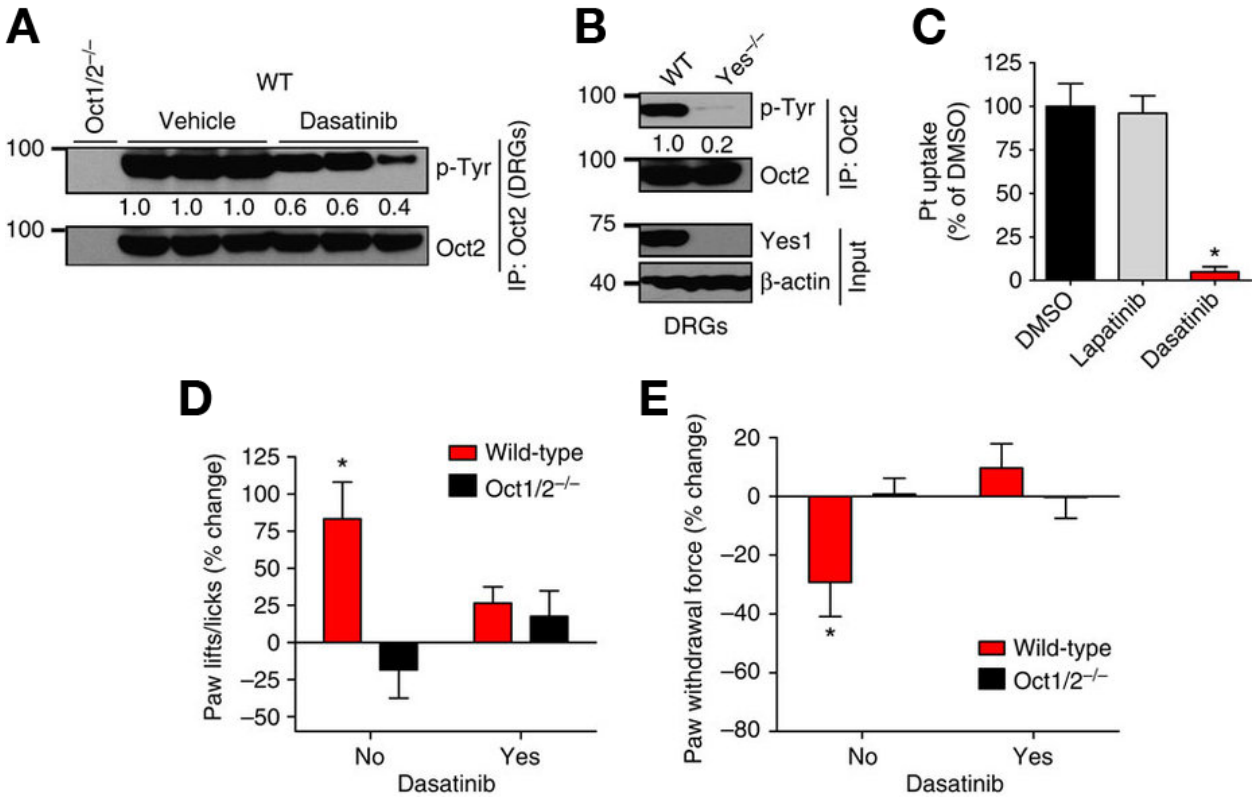


Figure 2.7 Yes1 inhibition mitigates Oct2-dependent oxaliplatin neurotoxicity. A) Wild-type FVB mice were injected with vehicle and the Oct1/2 -/- mice were injected with either vehicle or dasatinib (15 mg/kg, p.o.) and 30 min later, DRGs were collected. DRG lysates were then used to immunoprecipitate endogenous Oct2 followed by western blot analysis by phospho-tyrosine and Oct2 antibodies. B) DRGs were collected from Wild-type and Yes1 -/- mice. The upper panel shows representative blots from experiments where DRG lysates were used to immunoprecipitate endogenous Oct2 followed by western blot analysis by phosphotyrosine and Oct2 antibodies. The lower panel shows western blot results from total DRG lysates showing that Yes1 is expressed in DRGs in the wild-type mice. C) DRGs were collected from wild-type FVB mice, followed by satellite cell isolation and culture. The primary satellite cells were then plated in six-well plates followed by oxaliplatin uptake assays in the presence of DMSO, lapatinib or dasatinib (30 min). The graph represents relative oxaliplatin uptake as compared to DMSO group. * indicates a statistically significant difference compared with the DMSO group. D) Sensitivity to cold associated with a single dose of oxaliplatin (40 mg/kg) in wild-type mice pretreated with vehicle or dasatinib (15 mg/kg, p.o.) as determined by a cold-plate test. The number of paw lifts or licks at baseline and following exposure to a temperature of 4 °C for 5 min at 24 h after drug administration was determined (n = 5). The graph represents relative percentage change in paw lifts/licks as compared with baseline values. E) Mechanical allodynia associated with a single dose of oxaliplatin (40 mg/kg) in wild-type mice pretreated with vehicle or dasatinib (15 mg/kg, p.o.), as determined by a Von Frey Hairs test. The force required to induce paw withdrawal in grams at baseline was measured following 24 h after drug administration (n = 5). The graph represents relative percentage change in paw withdrawal force as compared to baseline values. * indicates a statistically significant difference as compared with the baseline (untreated) values. All experimental values are presented as mean±s.e. The height of error bar = 1 s.e.

Yes1-mediated regulation of OCT2 function in the kidney We next examined whether OCT2 is tyrosine phosphorylated *in vivo*, initially in the kidney, a major site of OCT2 expression.[77] Endogenous OCT2 in murine kidneys was tyrosine phosphorylated, and dasatinib administration significantly inhibited OCT2 tyrosine phosphorylation (**Figure 2.6A**). Dasatinib administration followed by pharmacokinetic analysis showed that plasma dasatinib levels were significantly higher than the concentrations required for inhibiting OCT2 function *in vitro* (**Figure 2.6B**). Next, we utilized a previously established *in vivo* functional assay[61] where the OCT2 substrate TEA is injected in mice and OCT2 inhibition under these conditions leads to increased plasma TEA levels. In the wild-type mice, dasatinib treatment led to increased plasma TEA levels (**Figure 2.6C**) and reduced urinary TEA excretion (**Figure S2.12A**), indicating OCT2 inhibition. Importantly, in the Oct1/2 deficient mice, dasatinib treatment did not alter the TEA levels in the plasma or urine (**Figure 2.6C** and **Figure S2.12A**). Dasatinib-mediated OCT2 inhibition was further confirmed in *ex vivo* renal tubule uptake assays (**Figure 2.6D**).[78] Next, we used hydrodynamic siRNA injection[79] to knockdown Yes1 in the kidneys, which resulted in reduced OCT2 phosphorylation (**Figure S2.12B**) and diminished OCT2 function (**Figure 2.6E** and **Figure S2.12C**). Yes1 knockout mice also had reduced OCT2 function (**Figure 2.6F** and **Figure S2.12D**) and lower OCT2 tyrosine phosphorylation (**Figure 2.6G**). The levels of dasatinib in the plasma of WT and Yes1-deficient mice was not significantly different (**Figure S2.12E**). These studies using multiple experimental approaches established that Yes1-mediated OCT2 phosphorylation is critical for OCT2 function *in vivo*.

Yes1 inhibition ameliorates oxaliplatin-induced acute sensory neuropathy Finally, we examined whether inhibition of Yes1 can diminish OCT2 function in DRGs and reduce oxaliplatin-induced acute sensory neuropathy.[57] Indeed, dasatinib treatment reduced OCT2 tyrosine phosphorylation in DRGs, the primary site of oxaliplatin uptake[57] and peripheral neurotoxicity (**Figure 2.7A**). Moreover, DRGs isolated from Yes1-deficient mice had significantly reduced OCT2 tyrosine phosphorylation (**Figure 2.7B**). Since OCT2 is expressed in satellite cells in DRGs,[57] we cultured primary satellite cells from murine DRGs. Uptake assays showed that dasatinib but not lapatinib inhibited oxaliplatin uptake in these cells (**Figure 2.7C**). To assess the therapeutic potential of dasatinib in preventing oxaliplatin-induced acute sensory neuropathy, we performed experiments in previously established mouse models.[57] Dasatinib treatment significantly mitigated both mechanical allodynia and cold sensitivity, markers of oxaliplatin-induced acute sensory neuropathy (**Figure 2.7D,E**), without affecting the systemic disposition of oxaliplatin (**Figure S2.13A,B**) or the metabolism and plasma levels of fluoropyrimidines such as fluorouracil,[80] which are commonly

given concurrently with oxaliplatin (**Figure S2.13C,D**). Furthermore, Yes1-deficient mice were also protected from oxaliplatin-induced acute sensory neuropathy (**Figure S2.14**). While we observed that Yes1 inhibition resulted in significant mitigation of oxaliplatin-induced acute sensory neuropathy, the effect of dasatinib on chronic neuropathy is currently unknown. However, recent studies have shown that chronic neuropathy correlates with the severity of acute sensory neuropathy,[81] raising the possibility that reduction in acute symptoms could also provide protection from chronic neuropathy.

Discussion

Tyrosine phosphorylation is a key regulatory mechanism of intra- and intercellular signaling in metazoans.[82] Recent global phosphoproteomic analysis using high-throughput, high-sensitivity MS[66,67,83,84] have revealed an unexpectedly complex repertoire of proteins that can be tyrosine phosphorylated. Extensive tyrosine phosphorylation in secreted or extracellular proteins was one such unexpected observation, which later led to the identification of VLK (vertebrate lonesome kinase) as a secreted tyrosine kinase that acts in the extracellular environment.[85] Analysis of these global phosphoproteome data[66,67] have now revealed that membrane transporters in general, and 'multispecific' drug transporters in particular, are tyrosine phosphorylated. Functional studies showed that many important SLC-family drug transporters (OCT1, OCT2, OCT3, MATE1 and OATP1B1)[46] might be regulated through tyrosine phosphorylation. This regulatory mechanism was explored in depth for OCT2, which led to the identification of the Src family kinase Yes1 as the kinase responsible for OCT2 tyrosine phosphorylation both *in vitro* and *in vivo*.

Identification of tyrosine phosphorylation as a crucial mechanism of drug transporter regulation has potentially wide ranging implications. Drug–drug interactions or genetic polymorphisms that alter transporter function are known to greatly affect therapeutic outcomes, including efficacy as well as toxicities.[47] Our study suggests that deregulation of post-translational modifications of drug transporters might have similar pharmacological consequences. We show that TKIs can be a potential cause of such deregulation. Tyrosine kinases are major targets of not only currently used drugs, but also of ongoing drug discovery programs.[56] Interestingly, several clinically used TKIs are known to modulate drug transporter functions.[24,86,87] Some of these interactions could be competitive, while, in case of ABC transporters, TKIs could occupy ATP-binding sites. However, as we show here for OCT2, TKIs can also inhibit drug transporters by targeting tyrosine kinases

essential for their function. Multiple TKIs that target Yes1 and other potent Yes1 inhibitors like dorsomorphin inhibited OCT2 function. Importantly, TKI-resistant Yes1 was able to rescue OCT2 inhibition, providing compelling evidence in support of our hypothesis.

The Yes1 kinase is a member of Src family of non-receptor tyrosine kinases (SFK) involved in diverse cellular processes.[88] Among SFKs, Src, Fyn and Yes1 are expressed in a variety of cell types, while the rest are restricted to haematopoietic cells.[89] Yes1 is highly expressed in the epithelial cells of the kidney, intestine, liver and lung.[90] While Yes1-deficient mice do not show any overt phenotype, they have significantly reduced transcytosis of polymeric immunoglobulin A across epithelial cells.[91] Our study suggests that through OCT2 phosphorylation, Yes1 may also regulate the transport of organic cations across epithelial cells. Interestingly, Yes1 has been implicated in the phosphorylation of occluding,[92] which is essential for epithelial tight junction maintenance. Higher Yes1 expression in epithelial cells, its role in transcytosis of macromolecules, and tight junction maintenance, combined with our identification of its role in OCT2 regulation raises the possibility that among the SFKs, Yes1 might have functionally diverged to regulate transport processes across epithelia.

The clinically used Yes1 inhibitor dasatinib[64] reduced OCT2 function in both cell culture and *ex vivo* models. Importantly, we carried out proof-of-principle experiments to determine the *in vivo* effect of pharmacological or genetic inhibition of Yes1 on OCT2 function at two major sites the kidneys and DRGs. These studies showed that renal OCT2 phosphorylation and function was significantly reduced by Yes1 inhibition *in vivo*. Similarly, dasatinib-mediated Yes1 inhibition or genetic Yes1 knockdown reduced OCT2 phosphorylation and function in the DRGs. These studies utilizing two models, which provide an *in vivo* readout of OCT2 function, suggest that Yes1 is essential for OCT2 phosphorylation and function *in vivo*. These studies also suggest that Yes1 could be pharmacologically targeted to reduce OCT2 function. OCT2 inhibition has the potential to mitigate acute sensory neuropathy,[57] a debilitating and common side effect of chemotherapy,[93] which occurs in part, due to OCT2-mediated oxaliplatin uptake in the DRGs. However, the clinical translation of these findings would require in-depth studies to determine the effect of combining oxaliplatin and dasatinib on both the anti-cancer efficacy and acute as well as chronic neuropathy.

Collectively, in the current study, we have uncovered the underappreciated role of post-translational regulation of drug transporters. These drug transporters are highly expressed in

intestine, liver, and kidney, major tissues that regulate absorption, distribution, metabolism, and excretion of drugs, including TKIs. Due to TKI accumulation in these tissues, drug transporters are uniquely vulnerable to phosphorylation-mediated interaction with TKIs. We propose that widely used therapeutics that target protein kinases, either due to inhibition of 'on-target' or 'off-target' protein kinases, can alter post-translational modifications of drug transporters and possibly drug-metabolizing enzymes and ion channels, which has widespread pharmacological implications.

Methods

Cell culture and reagents Parental Hela, HEK293 cells were obtained from American Type Culture Collection (ATCC). Hela, HEK293, Hela-OCT2 and HEK293-OCT2 cells were cultured in DMEM supplemented with 10% FBS and grown at 37 °C in a humidified incubator containing 5% CO_2 as described previously21. Lipofectamine 2000 or LTX (Life Technologies) reagent was used for transient transfections, followed by uptake assays, 24 h later. Tyrosine kinase inhibitors were obtained from Sigma-Aldrich or Selleckchem. Radiolabelled compounds were obtained from American Radiochemicals or Moravek Biochemicals.

Site-directed mutagenesis The OCT1, OCT2, OCT3, MATE1, OATP1B1, YES1, LYN and KDR plasmids with pCMV6-Entry (C-terminal FLAG tagged) backbone were obtained from Origene. The QuikChange II XL Site-Directed Mutagenesis Kit (Agilent) was utilized to generate mutants, according to suggested methods. The mutagenesis primers were designed using the QuikChange Primer Design program and synthesized by Integrated DNA Technologies. These plasmids were sequenced to confirm successful mutagenesis and then used for transient transfection experiments.

Protein analysis Whole-cell lysates from cultured cells and tissues were made in modified RIPA buffer (20 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1 mM Na_2EDTA, 1 mM EGTA, 1% NP-40, 2.5 mM sodium pyrophosphate, 1 mM beta-glycerophosphate, protease and phosphatase inhibitors) supplemented with 1% SDS. Membrane extracts from cultured cells were prepared using Cell surface protein isolation kit from Pierce (89881), while the Endoplasmic reticulum isolation kit was obtained from Sigma (ER0100). Lysates for immunoprecipitation were made in modified RIPA buffer supplemented with 0.1% SDS. Lysates for co-immunoprecipitation experiments were made in modified RIPA buffer supplemented with 0.2% β-maltoside. Immunoprecipitation was carried out using anti-FLAG (EZview Red ANTI-FLAG M2 Affinity Gel, Sigma) and anti-phosphotyrosine

(P-Tyr-1000 Rabbit mAb Sepharose) beads. Invitrogen Bis-tris gradient mini or midi-gels were used for western blot analysis, followed by detection by ECL reagent (Cell Signaling). Primary antibodies used were from cell signaling: FLAG (14793) and Phospho-Tyrosine (8954), Santa Cruz Biotech: Yes1 (8403), Phospho-Yes1 (130182), Na^+/K^+ ATPase (sc28800) and β-actin (47778), Alpha diagnostic international: Oct2 (OCT21-A) and Abcam: Transferrin receptor (ab84036) and PDI (ab5484). All primary antibodies were used at 1:1,000 dilution. Secondary antibodies were from Jackson Immunoresearch and used at 1:4,000 dilutions. The enzyme activity of recombinant DPD was measured as previously described.[94]

Cellular accumulation studies Uptake experiments were performed with tetraethylammonium, oxaliplatin, metformin and ASP^+ using standard methods,[57,61] in the presence or absence of inhibitors with results normalized to uptake values in cells transfected with an empty vector or DMSO-treated groups. In typical uptake experiments, medium was removed, cells were rinsed with PBS, followed by preincubation with either DMSO or inhibitors for 15 min, followed by addition of substrate and uptake measurement after 10–15 min incubation. Fluorescence measurements were done for ASP^+, while uptake for other substrates was measured by scintillation counter as described previously.[95]

Small-molecule HTS for OCT2 inhibitors HEK293 cells stably overexpressing OCT2 were used for optimizing the assay conditions for HTS of OCT2 inhibitors. Briefly, for the primary screen, cells were plated at 3,500 cells per well in 25 ml medium per well (384-well black clear bottom plates, tissue culture treated) and on day 3, compounds were transferred (final test compound at 28 mM concentration and positive control was 140 mM imipramine) along with OCT2 substrate ASP^+ at 5.6 mM and incubated for 10 min, followed by three wash cycles and fluorescence measurement (excitation 492 nm/emission 590 nm). Data analysis and positive hit selection was done according to standard methods. Compounds that inhibited OCT2 activity by 490% were considered as positive hits. TKIs, which were selected for further investigation, were then validated in secondary screens using standard uptake assays.

siRNA kinome screening HEK293-OCT2 cells were used for the siRNA kinome screening using methods similar to a previous study.[96] Briefly, the Dharmacon human siRNA library targeting protein kinases and related genes (total 779 genes) and containing four pooled siRNAs for each gene was utilized for this purpose. Briefly, the HEK293-OCT2 cells were plated in 384-well plates and

reverse transfected with 25 nM siRNA using Lipofectamine RNAiMAX reagent (Life Technologies). At 48 h post-transfection, cells were incubated with using 10 mM ASP^+ 4-(4-(dimethylamino)styryl)-N-methylpyridinium iodide followed by fluorescence readout of OCT2 function. This was followed by CellTiter-Glo Luminescent Cell Viability Assay (Promega). The siRNAs that reduced OCT2 function to ≤75% of control siRNA without affecting cell viability were selected for secondary screen. Deconvoluted secondary screen were performed by methods similar to the primary screen. In a separate secondary screen, Hela-OCT2 cells were reverse transfected with pooled siRNA (Sigma) and plated in 24-well plates, followed by uptake assays with 2 mM TEA after 48 h.

Single-molecule imaging Single-molecule imaging on *X. laevis* oocytes was performed after 1 day of expression at 18 degrees using Total Internal Reflection Fluorescence Microscopy (TIRFM) set-up as described elsewhere.[68,69,97] Briefly, oocytes were manually devitellinized and placed on high refractive index coverglass (n = 1.78) and imaged using Olympus 100X, numerical aperture 1.65 oil immersion objective at room temperature. mEGFP-tagged OCT2 (at C terminus) was excited using a phoxX 488 (60 mW) laser. Six hundred frames at the rate of 20 Hz were acquired for subunit counting. Only single, immobile and diffraction-limited spots were analysed. The number of bleaching steps was determined manually for each single spot included in the analysis. The error bars in subunit-counting data show statistical uncertainty in counting and are given by On, where n = number of counts. In addition, SiMPull was performed on whole oocyte or only plasma membrane as described previously. Briefly, channels/flow chambers were prepared on coverslips passivated with monofunctional and biotinylated polyethylene glycol. Biotinylated anti-EGFP antibody (Abcam) was then immobilized by incubating 40 nM of antibody on Neutravidin (Thermofisher) coated channels. Sample lysate was flown through the channels and TIRFM imaging and single-molecule counting analysis was done as described above.

OCT2 structural modeling Human OCT2 (SLC22A2; accession number O15244) was modeled based on the 2.9 Å resolution X-ray structure of a high-affinity phosphate transporter PiPT (PDB ID 4J05).[27] The template was selected because of the shared MFS fold assignment structure quality, and sequence similarity to OCT2. The sequence alignment was obtained by manually refining the combined output from PROMALS3D[98] and MUSCLE[99] servers. One hundred models were generated using the automodel class of MODELLER 9.13[100] and subsequently assessed by the normalized discrete optimized protein energy (zDOPE)[101] potential and the protein orientation-

dependent statistically optimized atomic potential (SOAP-Protein-OD).[102] The top scoring models were used for mapping possible phosphorylation sites. The ligand was copied from the PiPT structure as a rigid body for illustrative purposes only.

Yes1 kinase assay Yes1 recombinant human protein was obtained from Life technologies (A15557). To purify OCT2 proteins, FLAG-tagged wild-type or mutant OCT2 constructs (described in the site-directed mutagenesis section) were subcloned into pT7CFE1-CHis plasmid (Thermo Fischer). These constructs were then used for *in vitro* translation using a HeLa cell lysate-based kit (1-Step Human Coupled IVT Kit – DNA; 88881, Life Technologies). The *in vitro* translated proteins were then purified using His Pur cobalt spin columns (Thermo Scientific). For *in vitro* kinase assays, recombinant Yes1 and purified OCT2 proteins were incubated in a kinase buffer (Cell Signaling, 9802) supplemented with cold ATP (Cell signaling, 9804) at 30 °C for 30 min. After the incubation period, the reaction was terminated and OCT2 proteins were immunoprecipitated by FLAG tagged beads (described in the Protein Analysis section) followed by western blot analysis to determine OCT2 tyrosine phosphorylation.

Animal experiments Yes1-deficient mice (129S7 background, stock no. 002280) were obtained from Jackson laboratories and Oct1/2-deficient mice (FVB background, model no. 6622) were obtained from Taconic and heterozygous mice were bred in-house to obtain wild-type and knock-out littermates. For some experiments, wild-type FVB mice were obtained from Taconic. For all experiments, age matched (8–12 week) male mice were used. Animals were housed in a temperature-controlled environment with a 12 hour light cycle and given a standard diet and water *ad libitum*. All animals were housed and handled in accordance with the Institutional Animal Care and Use Committee of St Jude Children's Research Hospital or approved by a governmental committee overseeing animal welfare at the University of Münster and performed in accordance with national animal protection laws. For TEA experiments, [14C]-TEA (0.2 mg/kg) was injected intravenously followed by plasma collection 5 min later. Urinary bladder, renal and liver tissues were also collected and tissue or plasma TEA levels were measured by scintillation counter. Similar experiments were carried out after dasatinib (15 mg/kg oral gavage) administration and TEA injection (intravenous), 30 min later. For hydrodynamic injection, siRNAs from Ambion (25 mg in 0.5 ml of PBS) or 0.5 ml of PBS was rapidly injected into the tail vein. After 48 h, [14C]-TEA (0.2 mg/kg) was injected followed by plasma and tissue collection. Plasma TEA levels were used to determine OCT2 function, while renal tissues were processed for western blot analysis to

determine Yes1 knockdown and OCT2 tyrosine phosphorylation. The *ex vivo* renal tubule uptake assays were performed as described previously.[78,103] For pharmacokinetic analyses, male wild-type or Oct1/2 mice were administered oxaliplatin (i.p. 10–40 mg/kg) and/or dasatinib (oral gavage 15 mg/kg). Plasma was collected at various time intervals. Urine was collected from animals housed in metabolic cages for 72 h after oxaliplatin administration. Oxaliplatin levels were analysed by flameless atomic absorption spectrometry[57] and dasatinib levels were measured by LC-MS/MS as described previously.[104]

Mice DRG isolation L4 DRGs were extracted from male 8- to 12-week-old mice as reported previously19. For whole-cell lysate preparation and immunoprecipitation experiments, DRGs from two mice were pooled, lysed in RIPA buffer, followed by standard immunoprecipitation and western blot analysis. For primary satellite cell culture, the extracted DRGs were further processed according to a modified protocol based on a previous study.[105] Briefly, DRGs from two mice were collected in a 1.5-ml tube containing 500 ml of ice-cold PBS without Ca^{2+} and Mg^{2+}, supplemented with antibiotics and D-glucose (6 mg/ml). DRGs were then digested in 500 ml of 5 mg/ml collagenase for 20-60 min at 37°C; followed by addition of 100 ml of 0.25% trypsin, in the last 10 min of incubation. At the end of the incubation, 600 ml DMEM (10% FBS and antibiotics) media was added and DRGs were mechanically dissociated using a 1-ml tip and transferred to a 25-cm^2 flask (with 10 ml medium), and incubated at 37°C for 2–3 h. At the end of this pre-plating time, neurons floating in the incubation media were removed by discarding the media, while satellite cells remained attached to the flask and were supplemented with fresh medium. After 4–7 days, when the cells reached confluence, they were detached with trypsin-EDTA and resuspended in fresh medium, and plated at a density of 100,000 cells per ml in six-well plates for further experiments.

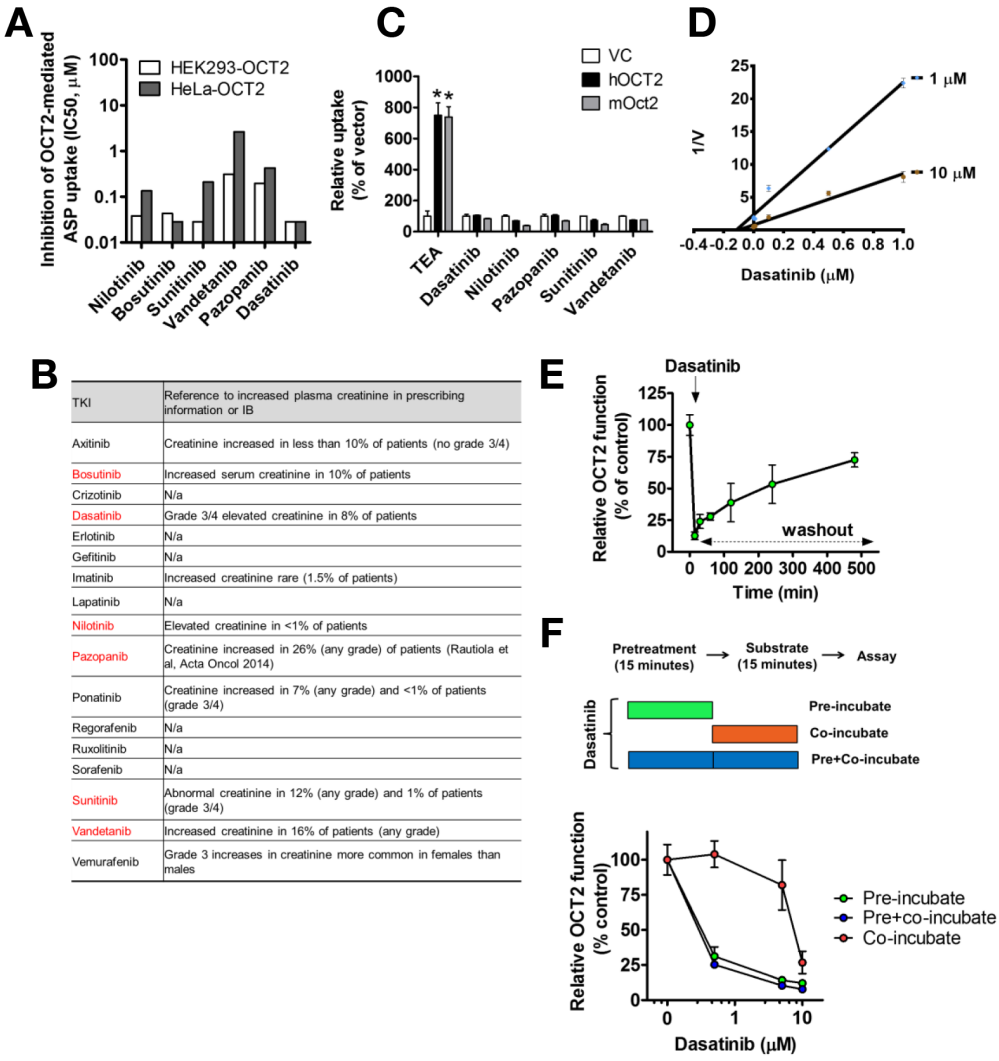
Mouse models of oxaliplatin neuropathy The mouse models of oxaliplatin neuropathy have been previously reported19. Thermal sensitivity associated with a single i.p. oxaliplatin dose (10 mg/kg) in male wild-type and Oct1/2 -/- mice was assessed by a cold-plate test. The number of paw lifts and licks when exposed to a temperature of 4 °C for 5 min was obtained for each mouse at 24 h before receiving oxaliplatin to determine the baseline levels. Data were recorded as the percentage change in the number of paw lifts or paw licks compared with baseline values when the animals were exposed to the same temperature 24 h. post-oxaliplatin administration. Mechanical allodynia was determined by a Von Frey Hairs test as described previously. Paw withdrawal was assessed in

triplicate on each hind paw with 5-min intervals. Data were recorded as the percent change of force (in g) necessary to promote paw withdrawal before and after oxaliplatin administration. Investigators conducting the experiments were blinded to the mouse genotypes. To determine the effect of dasatinib treatment on oxaliplatin neuropathy, male wild-type FVB mice were injected with dasatinib (oral gavage), followed by oxaliplatin injection intra-peritoneal 30 min later. Cold-plate assay and Von Frey Hairs test were then performed as described above.

Statistical considerations Data are presented as mean with s.e., unless stated otherwise. Statistical calculations (Student's t-test or analysis of variance) were done using a Graph-pad Prism. $P<0.05$ was considered statistically significant.

Supplemental information

Supplemental figures and tables are included below.



B

TKI	Reference to increased plasma creatinine in prescribing information or IB
Axitinib	Creatinine increased in less than 10% of patients (no grade 3/4)
Bosutinib	Increased serum creatinine in 10% of patients
Crizotinib	N/a
Dasatinib	Grade 3/4 elevated creatinine in 8% of patients
Erlotinib	N/a
Gefitinib	N/a
Imatinib	Increased creatinine rare (1.5% of patients)
Lapatinib	N/a
Nilotinib	Elevated creatinine in <1% of patients
Pazopanib	Creatinine increased in 26% (any grade) of patients (Rautiola et al, Acta Oncol 2014)
Ponatinib	Creatinine increased in 7% (any grade) and <1% of patients (grade 3/4)
Regorafenib	N/a
Ruxolitinib	N/a
Sorafenib	N/a
Sunitinib	Abnormal creatinine in 12% (any grade) and 1% of patients (grade 3/4)
Vandetanib	Increased creatinine in 16% of patients (any grade)
Vemurafenib	Grade 3 increases in creatinine more common in females than males

Figure S2.1 TKIs as OCT2 inhibitors. A) IC_{50} values of OCT2 inhibition in two different cell lines overexpressing OCT2. B) List of TKIs that are known to cause increased serum creatinine levels, which could be a result of OCT2 inhibition. TKIs that have potent OCT2 inhibition activity are highlighted in red. C) HEK293 cells stably expressing vector, hOCT2 or mOct2 were incubated with indicated radiolabeled compounds (2 μM) and cellular accumulation (15 min) was measured by scintillation counter. The graph represents relative uptake values as compared to individual vector group, after normalization of protein levels. * indicates statistically significant as compared to vector (VC) group. D) HEK293-OCT2 cells were used for uptake assays ([14C]-TEA for 15 minutes) at two indicated TEA concentrations and in the presence of 0-1μM dasatinib. The graph depicts a Dixon plot. E) HEK293-OCT2 cells were treated with 10 μM dasatinib for 15 min. Then the drug was washed out, and TEA uptake assays were carried out at different time points between 0 and 8 hrs. Baseline OCT2 activity was determined in untreated cells at each time point. F) HEK293-OCT2 cells were either pretreated with dasatinib for 15 min, followed by TEA uptake assay in the absence of dasatinib (Pre-incubate group) or pretreated with DMSO, followed by TEA uptake assay in the presence of dasatinib (Co-incubate group) or pretreated with dasatinib, followed by TEA uptake assay in the presence of dasatinib (Pre+Co-incubate group).

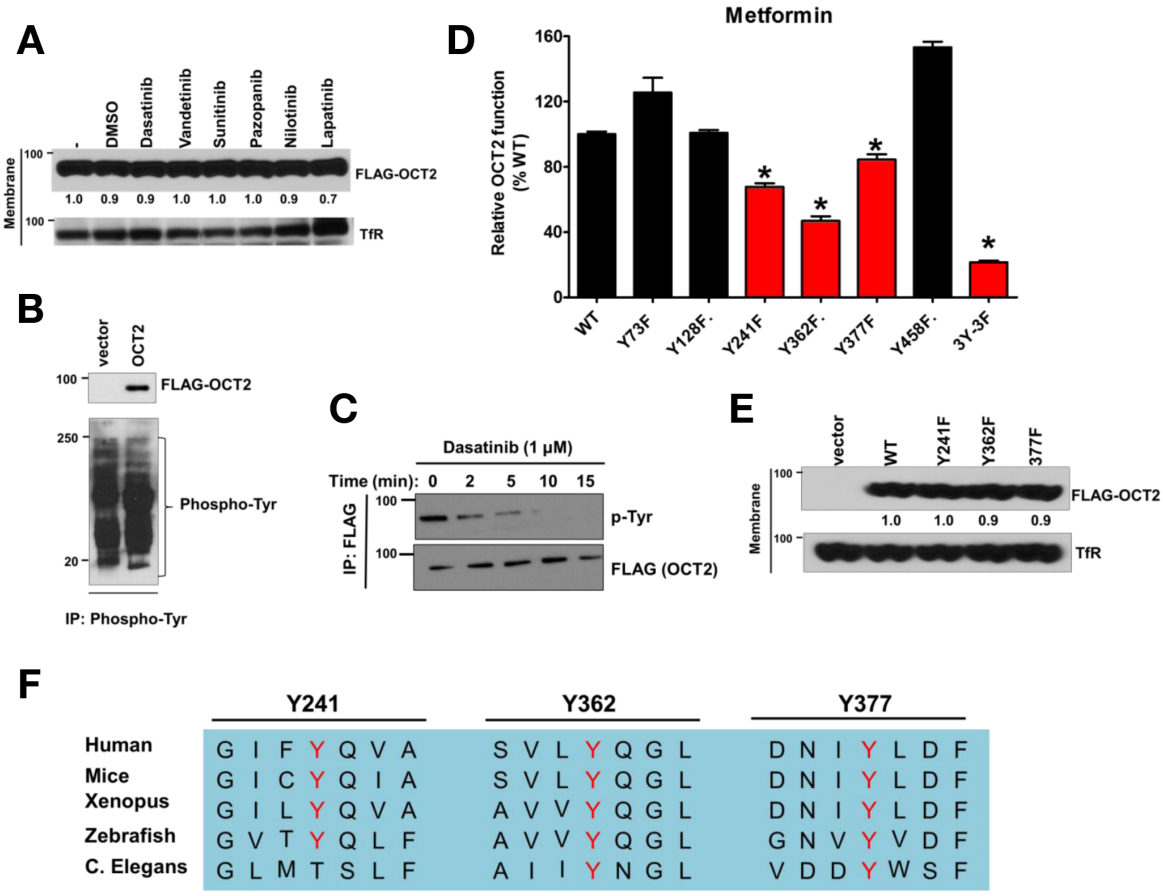


Figure S2.2 Regulation of OCT2 tyrosine phosphorylation. A) Hela-OCT2 cells were treated with either DMSO or indicated TKIs (10 μM) for 30 min followed by membrane extraction and western blot analysis. Transferrin receptor (TfR) served as loading control. B) Hela-vector or Hela-OCT2 cells were lysed and a phosphotyrosine antibody was used for immunoprecipitation followed by western blot analysis. C) Hela-OCT2 cells were treated with 1μM dasatinib for indicated time points, followed by immunoprecipitation of FLAG-OCT2 and western blot analysis of tyrosine phosphorylated OCT2. D) Hela cells were transfected with indicated wild-type or OCT2 mutants and 24 hours later 50 μM metformin was used for uptake assays (15 min). The graph represents relative uptake as compared to WT group. * indicates statistically significant as compared to WT group. E) Hela cells were transfected with vector, WT or mutant OCT2, followed by membrane preparation 24 hours after transfection. Transferrin receptor (TfR) served as loading control. F) The OCT2 protein sequence from indicated organisms was aligned by a multiple sequence alignment program (MAFFT).

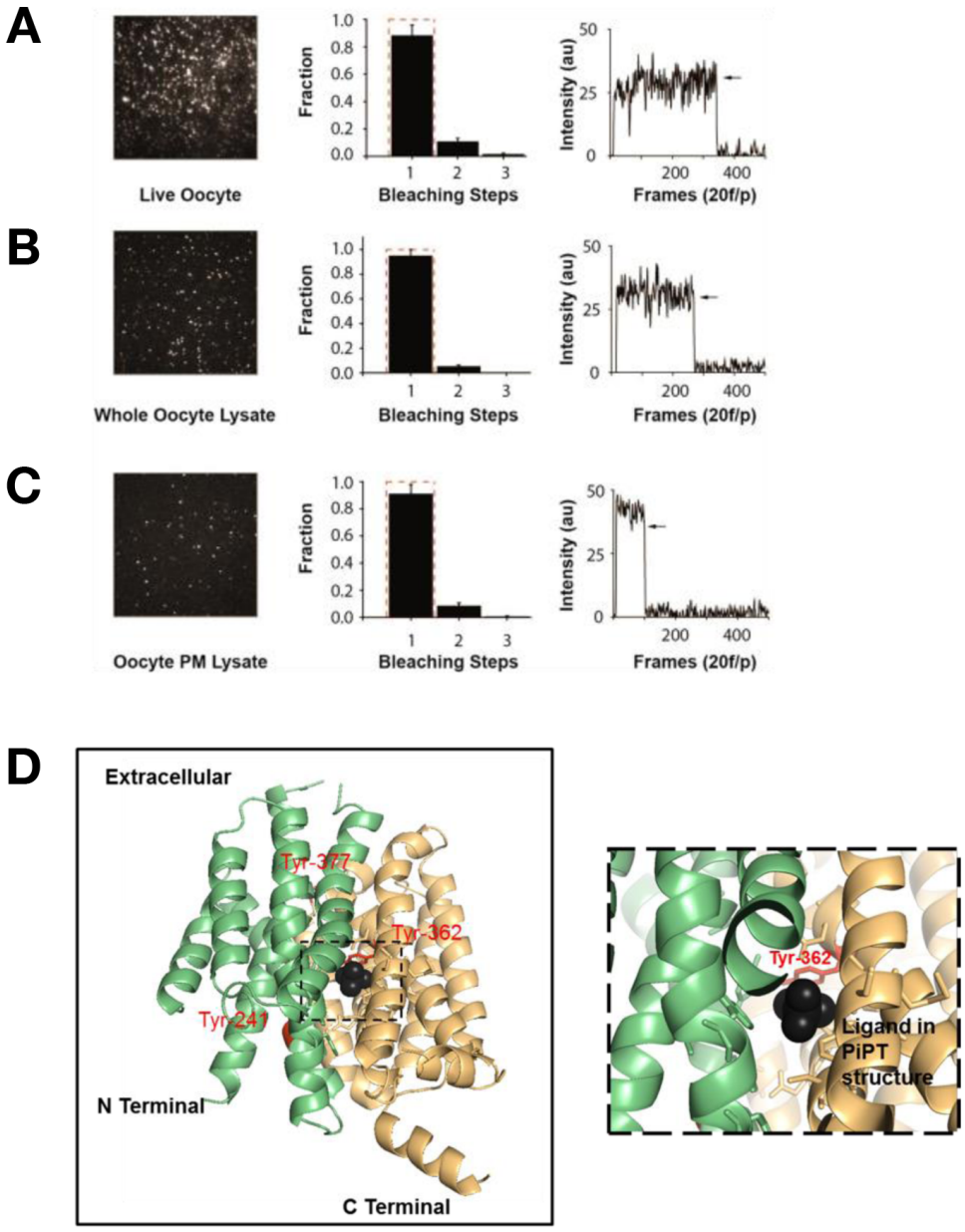


Figure S2.3 Relevance of OCT2 tyrosine phosphorylation. A-C) Single-molecule subunit counting of organic cation transporter 2 Single-molecule irreversible photobleaching to count the number of mEGFPs per fluorescent spot (i.e., number of subunits per transporter) of OCT2 in live Xenopus oocyte plasma membrane A), Using whole oocyte lysate in SiMPull B) and Using oocyte plasma membrane lysate only in SiMPull C). (Left) Images show first frame of the movie to indicate density of spots. (Middle) Average frequency distributions of number of bleaching steps (black bars) with error bars indicating SEM. Dashed red line indicates theoretical binominal distribution for monomer. (Right) Fluorescence traces from single spots showing single step of photobleaching. D) OCT2 structural model showing the three proposed phosphotyrosine sites and the predicted substrate binding site.

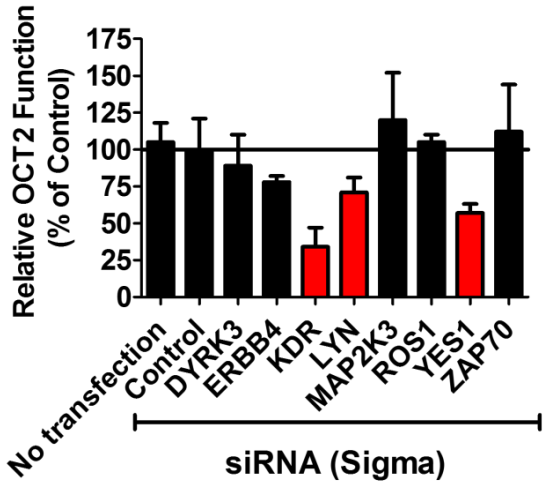


Figure S2.4 Secondary kinase screen to identify OCT2 phoshorylating protein kinases. Hela-OCT2 cells were reverse transfected with 25 nM siRNA (Sigma, pooled) for indicated protein kinases and 72 hours later, OCT2 functional assays were performed using TEA as a substrate (2 μM, 15 min). The graph represents relative uptake as compared to control siRNA transfected group. An arbitrary cut-off of ≤75% activity as compared to Control group was used to further short list candidate kinases (KDR, LYN and Yes1) that significantly affected OCT2 function.

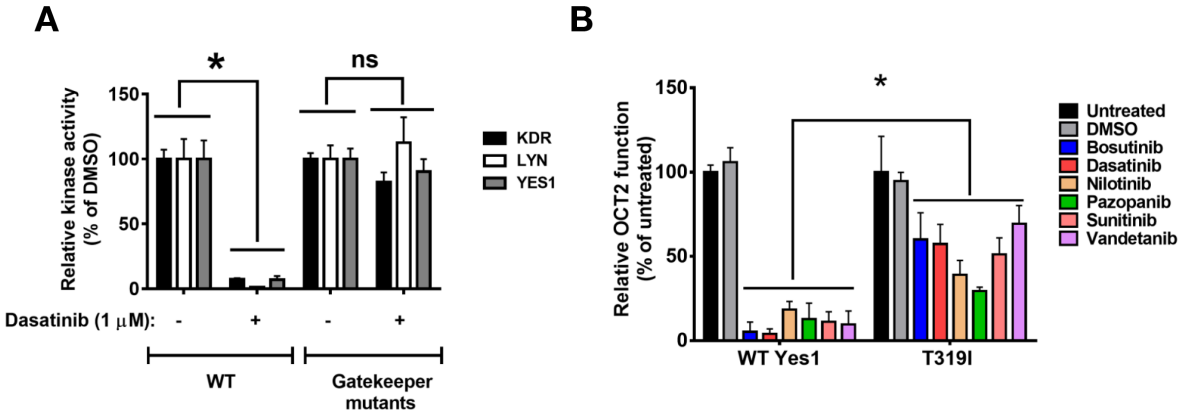


Figure S2.5 Effect of selected TKI resistant Kinase mutants on OCT2 function. A) Hela cells were transfected with WT and gatekeeper mutants of KDR, LYN and YES1. These cells were then treated with either DMSO (-) or 1μM Dasatinib for 15 min, followed by FLAG mediated immunoprecipitation of WT or gatekeeper kinases. The immunoprecipitates were then used for an *in vitro* kinase assay using Myelin Basic Protein (MBP) as a substrate (30 min, 30 °C, 32P ATP). SDS-PAGE gels were run and relative kinase activity was calculated based on the densitometric analysis. * indicates statistically significant as compared to DMSO group B) Hela-OCT2 cells were transfected with either wild type or T319I gatekeeper mutant, followed by pre-treatment with indicated TKIs (1μM and 15 minutes). The pre-treatment was followed by uptake assays (15 minutes) using [14C]-TEA (2μM). * indicates statistically significant as compared to WT Yes1 treated groups.

A

TKI	Yes1 K_d (nM)	OCT2 Inhibition
Bosutinib	4	Yes
Dasatinib	0.3	Yes
Erlotinib	2200	Yes
Lapatinib	>10,000	No
Nilotinib	1100	Yes
Pazopanib	5000	Yes
Sunitinib	120	Yes
Sorafenib	>10,000	No
Vandetinib	120	Yes

B

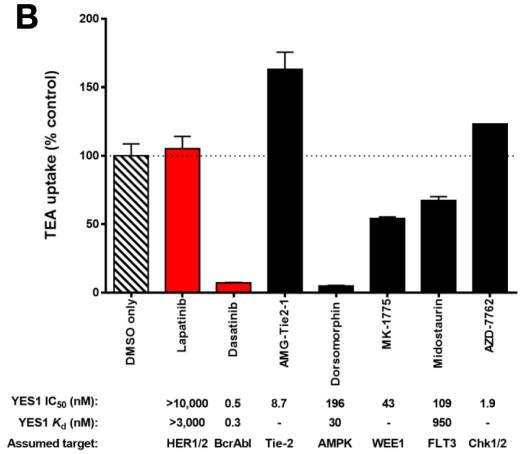
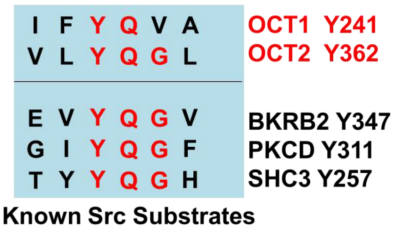
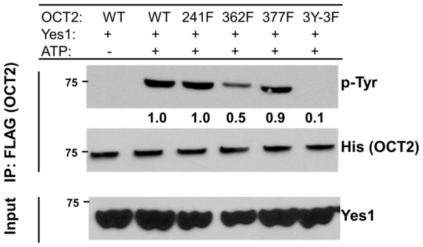


Figure S2.6 Small molecule Yes1 inhibitors reduce OCT2 function. A) TKIs that inhibit OCT2 are also Yes1 inhibitors (ref. 71). B) Hela-OCT2 cells were pre-treated (1 μM) with indicated compounds including Yes1 inhibitors followed by OCT2 uptake assays using TEA as substrate (2 μM, 15 min). The data is presented as relative OCT2 activity as compared to DMSO group. Hyphen (-) in the lower panel indicates that Kd values are unknown.

A



B



C

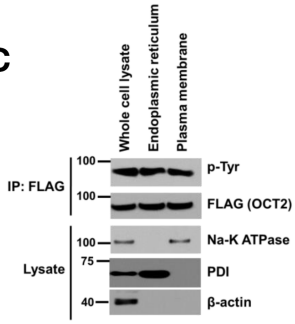


Figure S2.7 Yes1 can phosphorylate OCT2. A) The OCT2 protein sequence surrounding the Y241 and Y362 sites was aligned with other known Src family kinase substrates (PhosphoSite). Multiple Src family substrates had sites similar to sequence surrounding Y241 and Y362 sites in OCT2. B) Purified FLAG-OCT2 wild-type or Y/F mutants and Yes1 were used in a kinase assay in the presence or absence of ATP, followed by His-OCT2-IP and western blot analysis of OCT2 tyrosine phosphorylation and determination of total OCT2 and Yes1 protein levels. C) Indicated subcellular fraction were isolated from Hela-OCT2 cells, followed by FLAG-IP of OCT2 and western blot analysis. The lower panel shows the blots from the indicated lysates and Na+/K+ ATPase (plasma membrane), Protein disulfide isomerase (PDI) (Endoplasmic reticulum) and β-actin (cytosol) were used to check the purity of the isolated fractions.

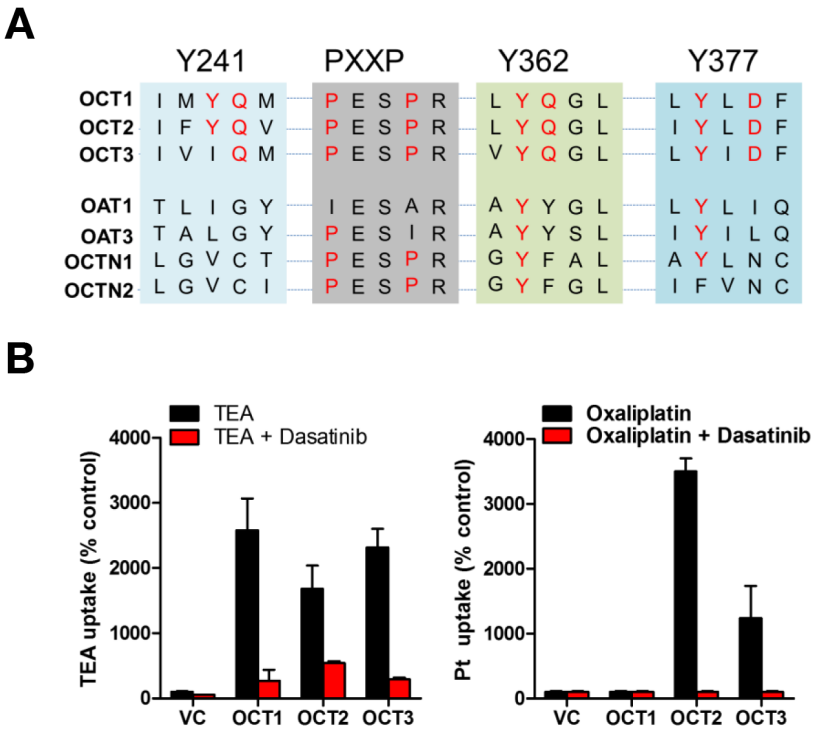


Figure S2.8 Tyrosine phosphorylation sites are conserved in OCTs. A) The protein sequence of indicated human organic cation and anion transporters of SLC22 family proteins was aligned by a multiple sequence alignment program (MAFFT). B) HEK293 stably expressing OCT1, OCT2 and OCT3 were used for uptake assays (SLC22A - TEA and Oxaliplatin) in the presence of 10 μM dasatinib. The graphs represent relative transporter activity for indicated substrates as compared to vector (VC) group.

A

Drug Transporter	Phospho-Tyrosine site(s)
OCT1	361
OCT2	241
MATE1	299
OAT2	252
OCTN2	486
OATP1B1	367,422,425,625,645
OATP4C1	473, 543
PGP	187
MRP1	277,490,920,1508,1522
MRP2	616,885
MRP3	229,231,720
MRP4	45,1255,1259
MRP5	10,17,1166,1202,1423
BCRP	464

B

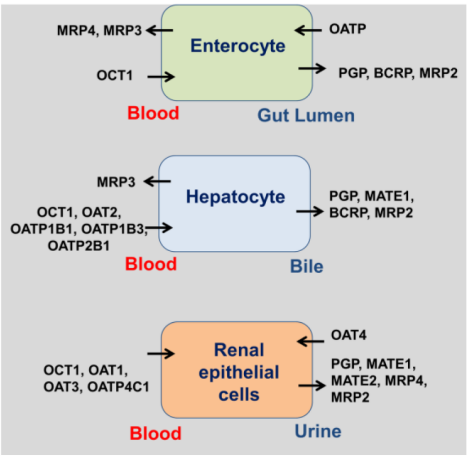


Figure S2.9 Drug-transporters and tyrosine phosphorylation. A) Analysis of high-throughput global phosphorylation data (PhosphoSite) suggests that multiple drug-transporters have phosphotyrosine sites. B) Representative figure showing the localization of important drug-transporters.

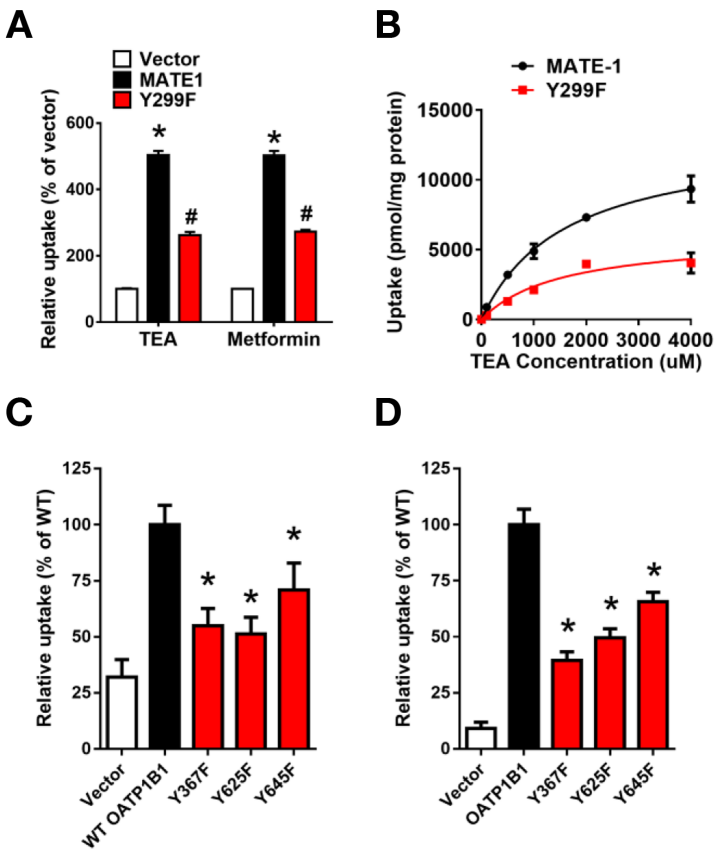


Figure S2.10 MATE1 and OATP1B1 mutants that lack phospho-tyrosine sites have reduced activity. A,B) HEK293 were transfected with indicated MATE1 constructs and uptake assays (15 min incubation) were performed using 2 μM TEA and 50 μM metformin as substrates C,D) HEK293 were transfected with indicated OATP1B1 constructs and uptake assays were performed using C) estradiol-17β-d-glucuronide (E2G) (0.1 μM; 5 min incubation) and D) 8-fluorescein-cAMP (8-FcA) (5 μM; 10 min incubation) as substrates. Graphs represent relative uptake activity as compared to WT group. * indicates statistically significant as compared to WT group.

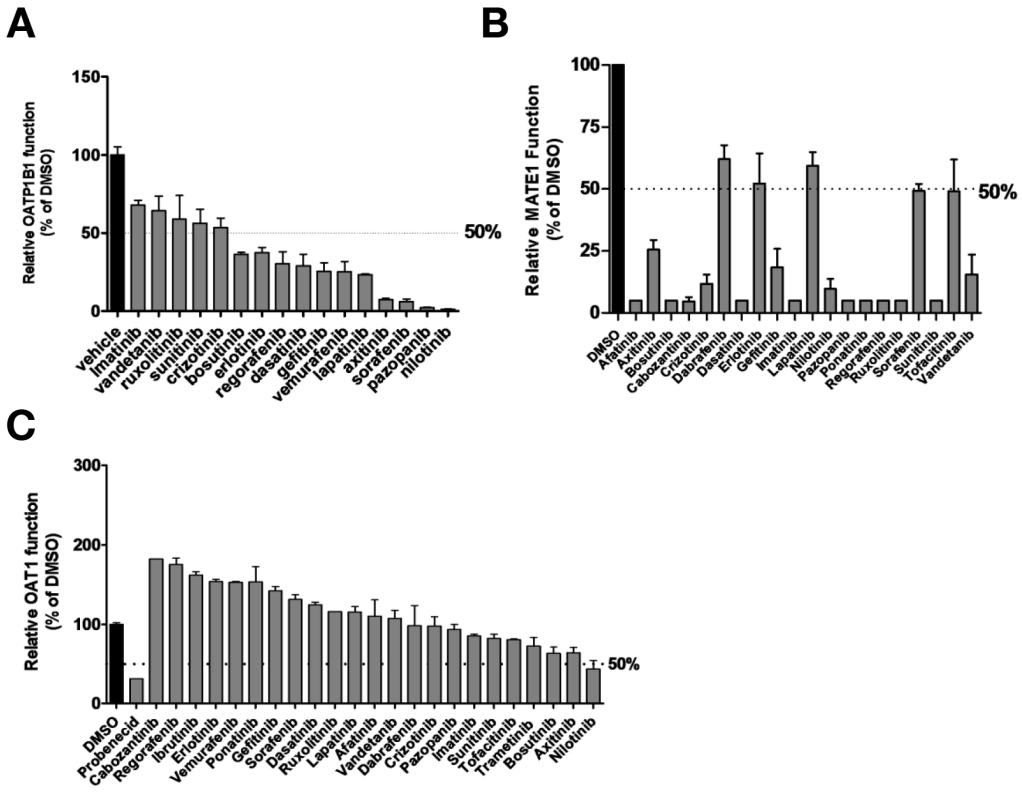


Figure S2.11 Distinct TKI inhibit multiple SLC transporters. HEK293 cells expressing indicated transporters were pre-incubated for 15 min with 10 μM TKIs, followed by uptake assays using indicated substrates for A) OATP1B1 (0.1 μM estradiol-17β-d-glucuronide (E2G) for 5 min) B) MATE1 (2 μM TEA for 15 min) and C) OAT1 (5μM para-aminohippurate (PAH) for 15 min).

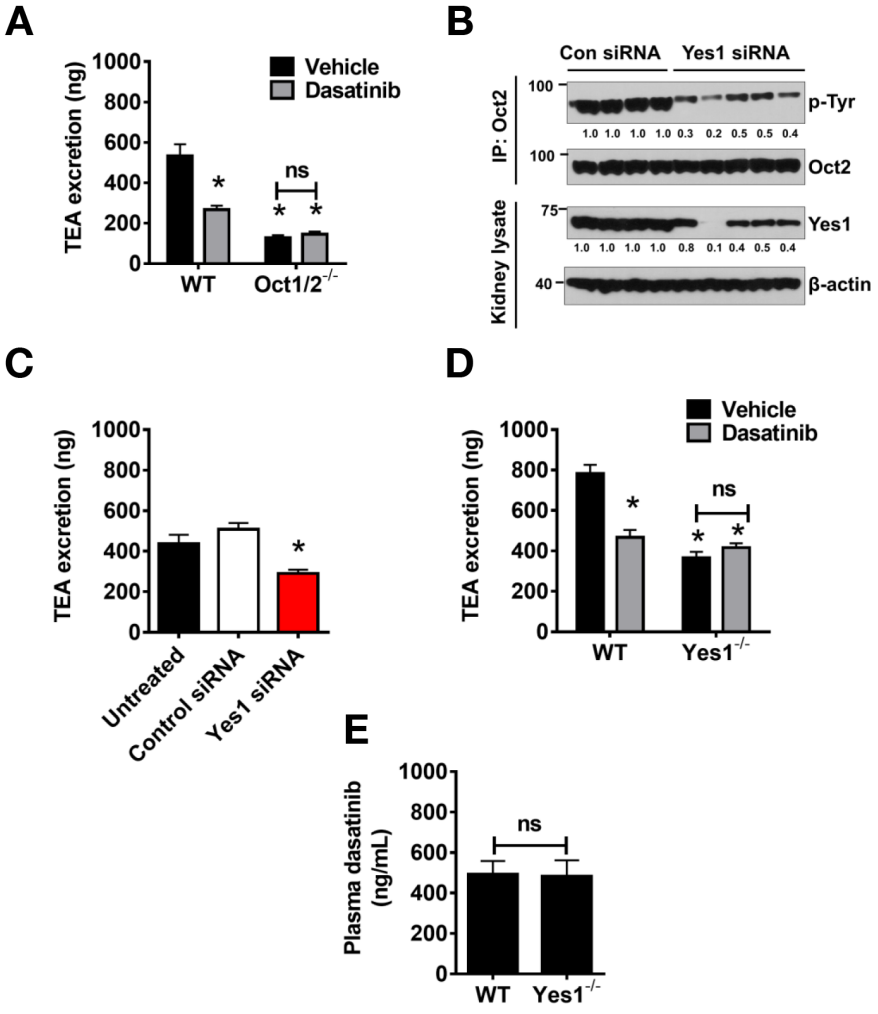


Figure S2.12 Effect of genetic or pharmacological Yes1 inhibition on OCT2 function and phosphorylation in vivo. A) Wild type and Oct1/2 -/- mice were injected with either vehicle or dasatinib (15 mg/kg, p.o.) and 30 min later they were injected with 0.2 mg/kg C14-TEA (i.v.), followed by urine collection at 5 min. The graph represents urine TEA levels from n = 5 mice per group. * indicates statistically significant as compared to WT vehicle group. B) FVB mice were injected with control or Yes1 siRNA and 3 days later, kidneys were collected followed by OCT2 immunoprecipitation and OCT2 phosphotyrosine levels were detected by western blot analysis. Yes1 and actin levels were determined in total cellular lysates. C) Wild type FVB mice were injected with either control or Yes1 siRNA by hydrodynamic tail-vein injection (25 µg in 0.5 ml of PBS). Three days later, the mice were injected i.v. with a 0.2 mg/kg dose of [14C]-TEA, and urine levels of TEA were measured at 5 min. * indicates statistically significant as compared to untreated group D) Wild type and Yes1 -/- mice were injected with 0.2 mg/kg [14C]-TEA (i.v.) and urine levels of TEA were measured at 5 minutes. * indicates statistically significant as compared to WT vehicle. E) Wild type and Yes1 -/- mice were injected with dasatinib (15 mg/kg, p.o.) and 30 min, plasma was collected for dasatinib measurement.

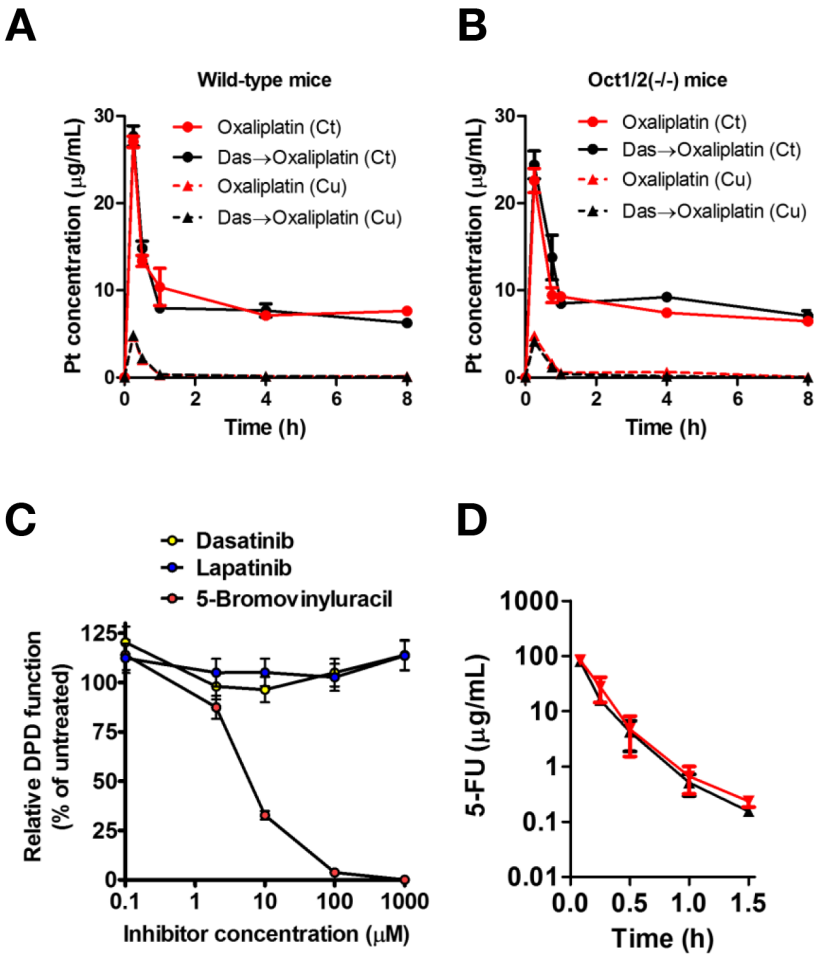


Figure S2.13 Pharmacokinetic analysis of dasatinib, oxaliplatin and fluorouracil (5-FU). A) Plasma concentration-time profiles of total platinum (Ct) and unbound platinum (Cu) in wild type mice following oxaliplatin with and without dasatinib (10 mg/kg, p.o.). B) Plasma concentration-time profiles of total platinum (Ct) and unbound platinum (Cu) in Oct1/2 -/- mice following oxaliplatin with and without dasatinib (10 mg/kg, p.o.). C) DPD enzyme activity was measured in the presence of the indicated concentrations of dasatinib, lapatinib and 5-bromovinyluracil, a known DPD inhibitor. D) Plasma-concentration time profiles of 5-FU (75 mg/kg, i.v.) in wild type mice in the presence (red) or absence (black) of dasatinib (mg/kg, p.o.; 15 min before 5-FU).

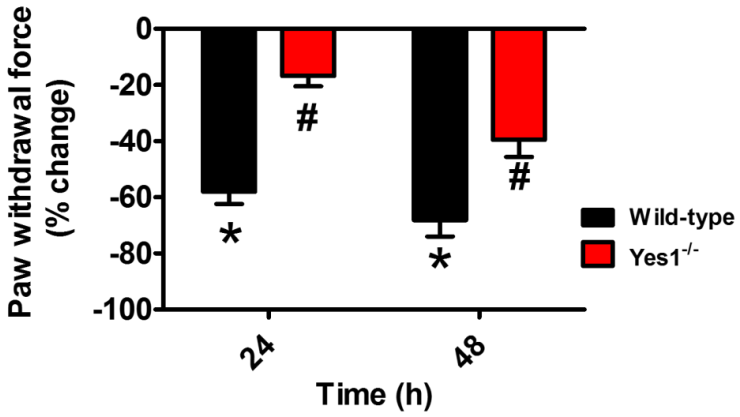


Figure S2.14 Yes1-deficient mice are protected from oxaliplatin neurotoxicity. Mechanical allodynia associated with a single dose of oxaliplatin (40 mg/kg) in wild type and Yes1 deficient mice as determined by a Von Frey Hairs test. The force required to induce paw withdrawal in grams at baseline was measured following 24 h after drug administration. The graph represents relative percentage change in paw withdrawal force as compared to baseline values. * indicates a statistically significant difference as compared with the baseline (untreated) values.

Discovery of competitive and non-competitive ligands of the organic cation transporter 1 (OCT1)

The contents of this section have been submitted for publication in the following article:

Discovery of competitive and non-competitive ligands of the organic cation transporter 1 (OCT1).

*Eugene C Chen, Natalia Khuri, Xiaomin Liang, *Adrian Stecula, Huan-Chieh Chien, Sook Wah Yee, Yong Huang, Andrej Sali, and Kathleen M Giacomini*
Journal of Medicinal Chemistry. Submitted 2016.

*The author's contribution included comparative structure modeling of human OCT1 and docking of a prescription drug and endogenous metabolite libraries into the homology model, to supplement the high-throughput screening results.

Introduction

Organic cation transporter 1 (OCT1, SLC22A1), a polyspecific membrane transporter is among the most abundantly expressed transporters in human liver. Localized to the sinusoidal membrane of hepatocytes, OCT1 mediates the hepatic uptake of a diverse array of small positively charged hydrophilic compounds, including many endogenous bioactive amines (e.g., dopamine, histamine, and serotonin).[106] Recently, we identified OCT1 as a high capacity transporter of thiamine in the liver, and showed that the transporter plays a key role in modulating hepatic energy status and lipid content.[107]

Although the transporter clearly has important endogenous functions,[107,108] OCT1 has been characterized primarily as a drug transporter, capable of transporting a wide variety of prescription drugs, including the anti-diabetic drug, metformin, and the opioid analgesic, morphine. Genetic variants of OCT1 with reduced function have been associated with decreased response to metformin,[109] as well as high systemic plasma levels of morphine and the active metabolite of the opioidergic drug, tramadol.[110] Further, administration of the calcium channel blocker, verapamil (a potent inhibitor of OCT1), has been shown to reduce response to metformin, presumably through reducing hepatic drug levels.[111] In recognition of its critical role in drug disposition and response, OCT1 was included in a group of transporters of clinical importance by the International Transporter Consortium.[46] In 2012, the European Medicines Agency (EMA) recommended *in vitro*

inhibition studies against OCT1 for investigational drugs in its Guidance on the Investigation of Drug Interactions.[112]

To date, more than 50 inhibitors of OCT1 transport have been identified by *in vitro* inhibition studies using radioactive or fluorescent probe substrates.[113,114] However, these studies have not identified the mechanisms by which inhibitors modulate OCT1 transport. Growing evidence suggests that OCT1-mediated transport can be inhibited in a "substrate-dependent" manner due to the presence of multiple, possibly overlapping binding sites on the protein.[115] Identification and characterization of OCT1 ligands could be facilitated by the availability of an atomic structure. However, the three-dimensional structure of human OCT1 or its mammalian orthologs has not yet been determined. Although several residues important for substrate binding have been reported and rationalized with OCT1 comparative models built using atomic structures of bacterial homologs,[8,28,29,31,116,117] accurate prediction of the binding site(s) remains challenging because of the low sequence identity between bacterial proteins and human OCT1. Recently, a structure of a high-affinity phosphate transporter (PiPT) from the fungus *Piriformospora indica* was determined by X-ray crystallography.[27] The transporter shares approximately 20% sequence identity with human solute carrier (SLC) transporters, especially within the SLC22 family, thus providing a new opportunity for comparative modeling of OCT1 and virtual screening.[27]

We used a combination of *in silico* and HTS methods to identify prescription drugs and endogenous metabolites that are ligands of OCT1, with the goals of predicting clinical drug interactions and understanding their interaction with OCT1 protein. To this end, we screened a prescription drug library *in silico* for compounds that interact with a predicted binding site on OCT1, using comparative structure modeling and virtual docking. In parallel, we conducted HTS for inhibitors of OCT1-mediated transport of the fluorescent ligand, ASP^+, against the same prescription drug library. We identified 167 ligands in the screened library and predicted 30 competitive ligands. Moreover, we showed that by combining structure-guided ligand discovery with structure-activity relationship models from HTS data, competitive ligands of OCT1 could be predicted from endogenous and exogenous metabolites.

Results

Comparative model of OCT1 and its validation by docking of known substrates We modeled human OCT1 based on the 2.9 Å structure of a high-affinity phosphate transporter PiPT, from *Piriformospora indica*, crystallized in an inward-facing occluded state with a bound substrate (PDB ID 4J05).[27] The best scoring 3-dimensional model was selected using the normalized discrete optimized protein energy (zDOPE) potential.[101] The zDOPE score of -0.22 suggests that 60% of its C-alpha atoms are within 3.5 Å of their correct positions.[118] The model includes all 12 transmembrane helices (TMHs), organized into the N- and C-terminal domains, and putative binding sites were identified in the translocation cavity between the two domains.[116] The following regions were not modeled: a large extracellular loop between TMHs 1 and 2, an intracellular loop between TMHs 6 and 7, and the intracellular N- and C-termini. While these regions may play a role in OCT1 homo-oligomerization, studies in rat showed that homo-oligomerization does not affect substrate affinity and transport function.[33] Because ASP^+ was used as a probe substrate, we first docked ASP^+ into two predicted binding sites and selected the site with the best score (-6.08) for all subsequent docking. Next, we validated the accuracy of the comparative model of the OCT1 transporter by confirming that (1) known OCT1 substrates docked favorably against the predicted binding site and (2) residues implicated in OCT1 transport[8,28,29,31] were localized in the predicted binding site, as follows. First, we docked 15 known OCT1 endogenous and drug substrates against the predicted binding site (**Figure 2.8A**). Twelve out of fifteen substrates (80%) had favorable (negative) docking scores, ranging from -24.44 for acyclovir to -2.88 for oxaliplatin (**Figure 2.8B** and **Table S2.1**). Positive scores for the three compounds (e.g., prostaglandins and pentamidine) resulted from steric clashes between ligand and transporter atoms, indicating that either the predicted binding site is too small to accommodate larger OCT1 ligands or that these compounds bind at a different site in the translocation pore. We also performed an enrichment analysis of docked substrates and decoys and computed a logAUC metric of 22.94, which suggests docking accuracy comparable to previously reported virtual screening experiments for human SLC transporters.[119,120] Second, we analyzed favorable docking poses to determine the frequency of predicted hydrogen bonds between binding site residues and substrate molecules. Thirteen residues in TMH 4, 10, and 11 formed hydrogen bonds with substrate molecules (**Figure 2.8A**). Previous mutagenesis studies and homology modeling efforts suggested that these residues are important for substrate binding and OCT1-mediated transport in other species.[8,28,29,31,116,117] In particular, negatively charged D474 is important for interactions with positively charged OCT1 substrates. Additionally, in rat, mutations

of Y221 and D474 resulted in the reduced uptake of tetraethylammonium (TEA).[28] Finally, docked substrates formed non-covalent interactions with W217, T225, I449, and Q447 in TMH 4, 10, and 11. The equivalent residues in rat have also been implicated in ligand-transporter interactions.[29]

Prediction of new ligands by virtual screening and validation by HTS We predicted new ligands of OCT1 by docking each one of the 1,780 compounds in the Pharmakon drug library (MicroSource Discovery Systems) against the predicted binding site on the OCT1 model. From the 1,780 compounds, 471 putative OCT1 ligands were predicted (normalized docking scores less than -1). These predictions were then tested by HTS of the entire Pharmakon library. The HTS assay relied on the uptake of the fluorescent substrate ASP^+ by OCT1 overexpressing HEK cells to assess the activity of the transporter.[114,121] The uptake of ASP^+ was linear for the first 5 minutes, and the K_m was determined to be 21.2 μM (**Figure 2.9A,B**). Thus, a substrate concentration of 2 μM was used to minimize the effect of substrate concentration on the IC_{50} values and an incubation time of 2 minutes was used to measure the initial rate of uptake. The average Z-prime score of the HTS was 0.80, indicating an reproducible assay.[122] Of the 1,780 Pharmakon compounds, 167 compounds (9%) were determined to be OCT1 inhibitors (**Figure 2.9C**), defined as compounds that inhibited 50% or more of ASP^+ uptake at 20 μM. Drugs known to inhibit OCT1 activity at 20 μM were generally confirmed by the screen. Of the 167 compounds, 30 were also predicted as inhibitors by virtual screening. The overall accuracy of virtual screening was 70%. The sensitivity and specificity of predictions were 77% and 12%. The low specificity is not surprising because the docking pipeline was executed in a fully automated fashion; in contrast, typical structure-based virtual screening involves manual post-docking selection of ligand poses.[119,120,123]

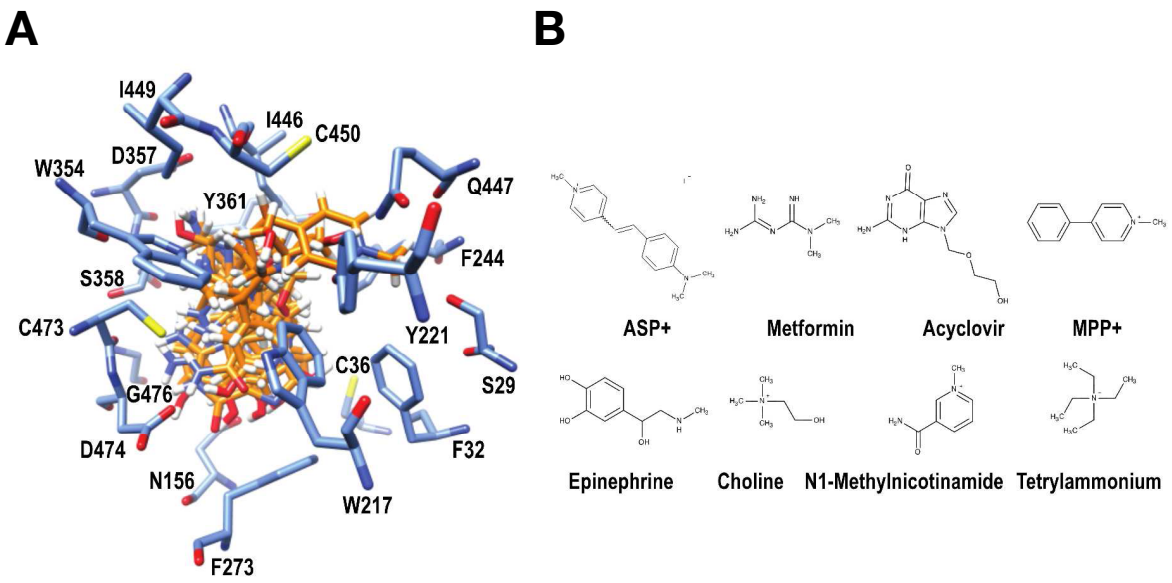


Figure 2.8 Predicted binding site of OCT1 and representative substrates. A) Thirteen residues (S29, F32, C36, N156, Y221, F273, W354, Y361, I446, S358, C450, C473, and D474) formed hydrogen bonds with docked substrate molecules. Six residues in the predicted binding site, W217, F244, I449, D357, Q447, and G476, participated in noncovalent and/or polar substrate-transporter interactions. OCT1 residues are shown as cornflower blue sticks. Structures of 12 favorably docked known OCT1 substrates are shown as orange sticks. Oxygen, nitrogen, sulfur, and hydrogen atoms are depicted in red, dark blue, yellow, and white, respectively. B) 2D structures of representative OCT1 substrates are drawn using MarvinView 14.7.7.0 (Chemaxon).

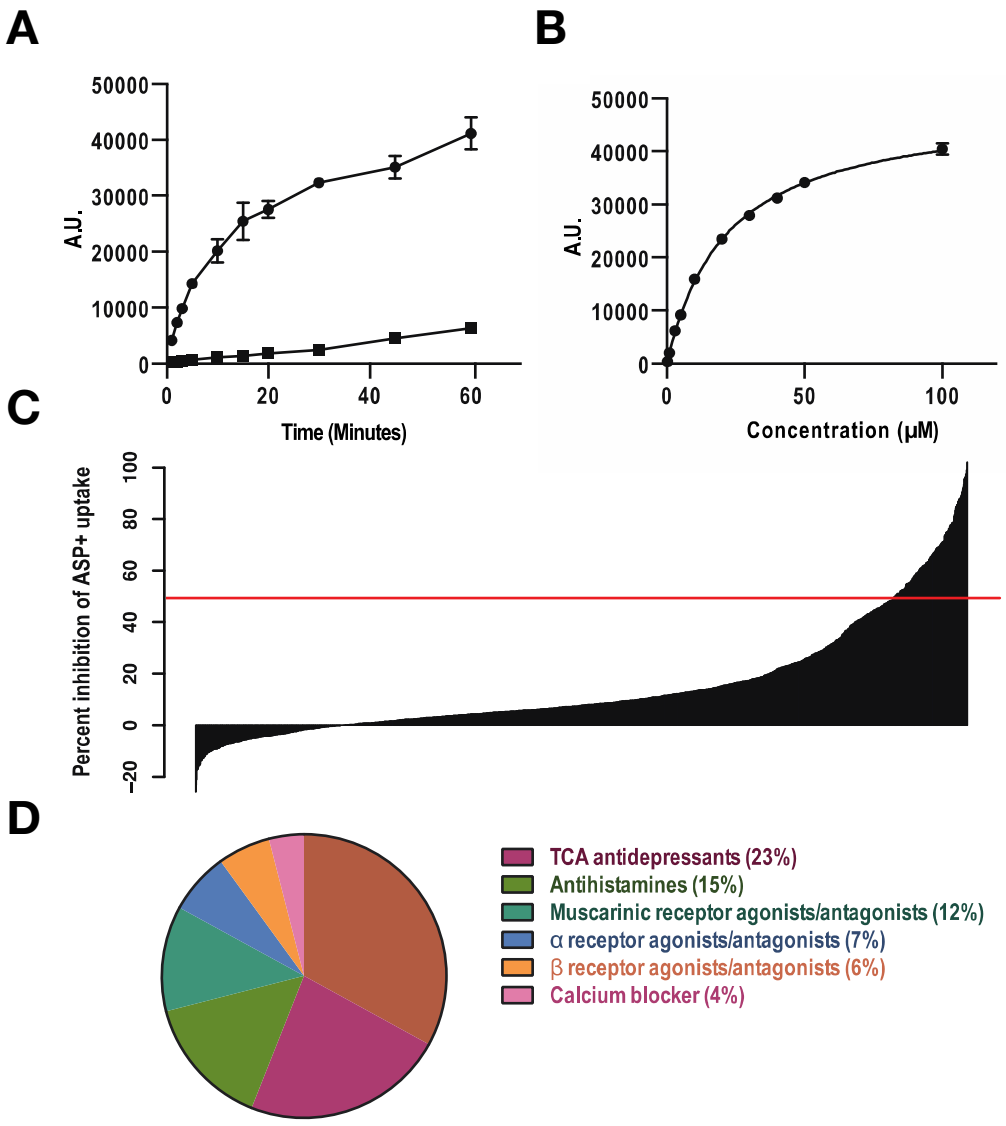


Figure 2.9 Uptake of ASP+ and HTS screening data. A) Time-dependent ASP+ uptake in HEK cells overexpressing OCT1 (●) or empty vector (■). B) Overexpressing OCT1 increases ASP+ uptake in HEK cells. ASP+ uptake studies were conducted in HEK cells overexpressing OCT1 or empty vector. Cells were incubated with increasing concentrations of ASP+ for 2 minutes. The uptake kinetic parameters were calculated using the difference in ASP+ accumulation between cells overexpressing OCT1 and empty vector cells. Data represent mean ± s.d., n = 6 per data point. C) Distribution of inhibition values from HTS of 1,780 compounds. A total of 167 Inhibitors were identified among the 1,780 Pharmakon compounds. D) The distribution of the 167 inhibitors in various pharmacological classes.

We examined the compounds predicted to be OCT1 ligands by the comparative model, but not identified as inhibitors by HTS. In view of the published literature, we identified 13 previously reported substrates and 5 inhibitors of OCT1. Among these predicted ligands, cimetidine (rank 21), metformin (rank 84), and thiamine (rank 200) inhibited OCT1-mediated uptake of ASP^+ by 4.13%, 11.1%, and -1.5%, respectively. The docking poses of all three compounds predicted favorable interaction with the comparative model; in particular, metformin and cimetidine formed hydrogen bonds with D474 (**Figure 2.10A,B**). The inability of HTS to identify some of the previously reported ligands can be explained by their OCT1 affinity, which is much weaker than that of ASP^+. For example, the reported IC_{50} values of cimetidine and metformin for inhibition of OCT1-mediated transport of YM155 and MPP^+ were 149 μM and 1,230 μM, respectively[124,125], and the IC_{50} of thiamine was determined to be 4.1 mM (**Figure S2.15**). Because our HTS assay measured inhibition of OCT1 activity by test compounds at 20 μM, it was unable to identify ligands such as cimetidine, metformin, and thiamine.

HTS identified tricyclic antidepressants, antihistamines, and α-adrenergic receptor agonists, agreeing with previously published results.[114] HTS also identified a high proportion of ligands from other drug classes, including calcium channel blockers, β-adrenergic receptor agonists/antagonists, and muscarinic acetylcholine receptor agonists/antagonists (**Figure 2.9D**). Selected hits from different classes of drugs were validated *in vitro* (**Table 2.1** and **Figure S2.16**). Finally, HTS identified drugs that were less studied to interact with OCT1, including carvedilol[126] (an anti-hypertensive medication) and ethopropazine (an anti-Parkinsonian agent). Selected hits not known previously to interact with OCT1 were validated by determining their IC_{50} values *in vitro* (**Figure 2.11**). Analysis of the physicochemical properties showed that OCT1 ligands tend to have fewer hydrogen bond donors and acceptors and are less polar, but more lipophilic than non-ligands (**Figure 2.12**). As expected, ligands were more likely to be positively charged (**Figure 2.12**).

Table 2.1 Summary of IC_{50} for selected inhibition studies.

Name	IC_{50} μM	IC_{50}, 95% Confidence Intervals
Ketoconazole	2.60	2.49 to 2.78
Closantel	3.00	2.83 to 3.18
Dobutamine	4.17	3.73 to 4.67
Alfuzosin	14.87	11.92 to 18.54
Erlotinib	16.24	11.34 to 23.26
Carbetapentane	1.55	1.19 to 2.02
Clotrimazole	11.97	9.97 to 14.38
Bithionol	2.23	1.98 to 2.52
Carvediolol	3.43	3.02 to 3.90
Clonidine	18.98	15.82 to 22.77
Trimethoprim	50.68	41.47 to 61.93
Guanabenz	4.85	4.17 to 5.64
Pyrimethamine	13.57	10.74 to 17.14
Dichlorophene	8.41	6.612 to 10.69
Imipramine	7.95	5.912 to 10.68
Cloperastine	14.89	14.00 to 15.83
Dextromethorphan	10.45	8.66 to 12.61
Propafenone	15.54	14.04 to 17.20
Tacrine	21.72	18.77 to 25.14
Ethopropazine	20.46	18.31 to 22.86
Nitroprusside	43.84	38.08 to 50.46
Sunitinib	6.10	5.47 to 6.79
Desipramine	9.18	8.08 to 10.41
Doxepin	11.19	9.89 to 12.66
Camylofine	9.12	8.03 to 10.36
Thiamine	4354	3704 to 5119

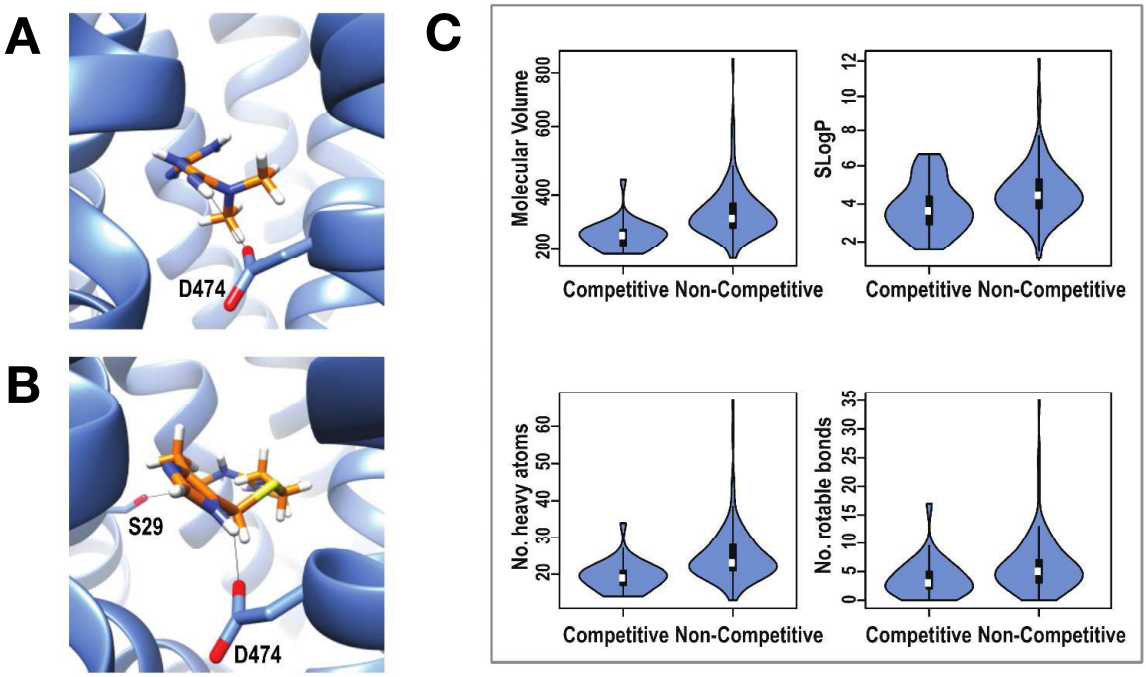


Figure 2.10 Docking results and physiochemical properties of OCT1 ligands in the Pharmakon library. Compounds are shown in orange sticks and hydrogen bonds as black dotted lines. A) The predicted pose of metformin and its hydrogen bond with aspartic acid residue 474. B) Predicted pose of cimetidine and its hydrogen bonds with D474 and S29. C) Differences in distribution of physicochemical properties for predicted competitive (n=30) and non-competitive (n=137) inhibitors of OCT1. Only significantly different distributions are shown (Student t-test p-value < 0.05).

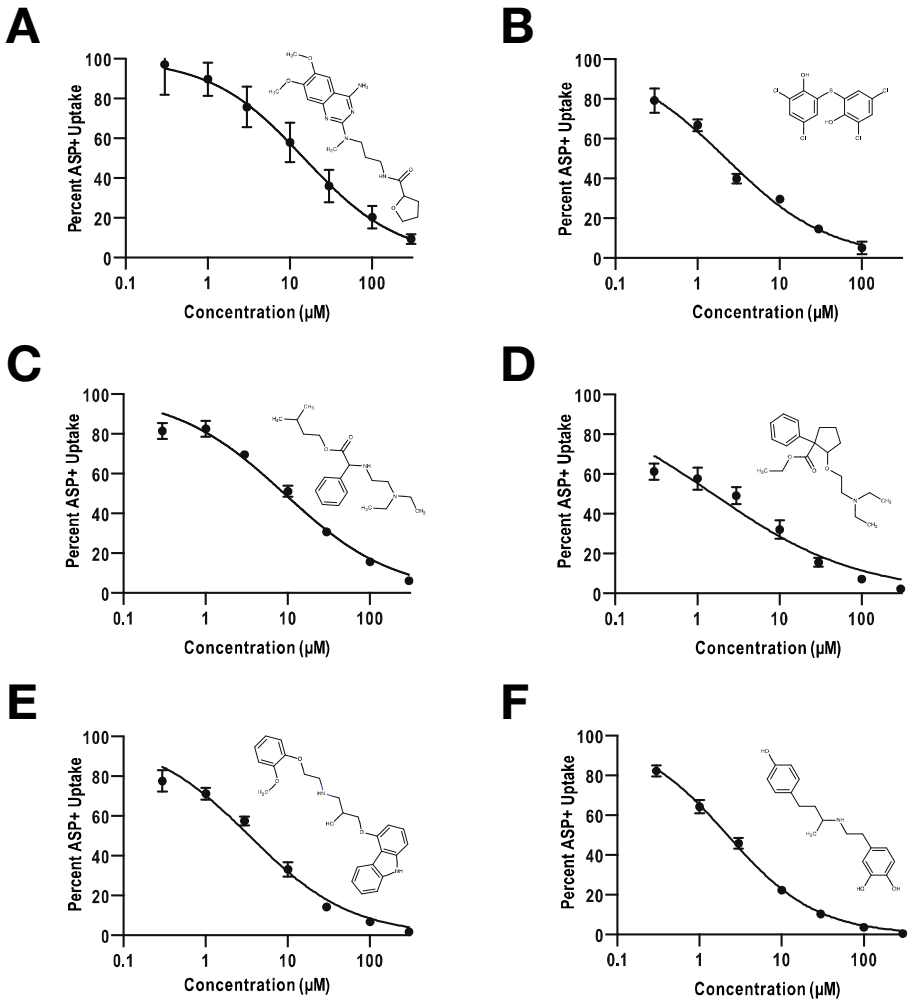


Figure 2.11 Selected inhibition studies of previously unknown OCT1 ligands. Estimated IC_{50} with 95% confidence limits in parenthesis. A) Alfuzosin, an α_1 adrenergic receptor antagonist IC_{50} = 14.9 (11.9; 18.5) μM. B) Bithionol, an anthelmintic IC_{50} = 2.2 (2.0; 2.5) μM. C) Camylofine, an antimuscarinic IC_{50} = 9.1 (8.0; 2.5) μM. D) Carbetapentane, an antitussive IC_{50} = 1.6 (1.2; 2.0) μM. E) Carvedilol, a nonselective β/α_1 adrenergic receptor antagonist IC_{50} = 3.4 (3.0; 3.9) μM. F) Dobutamine, a Sympathomimetic IC_{50} = 2.2 (2.1; 2.3) μM. Data represent mean ± s.d., n = 6 per data point.

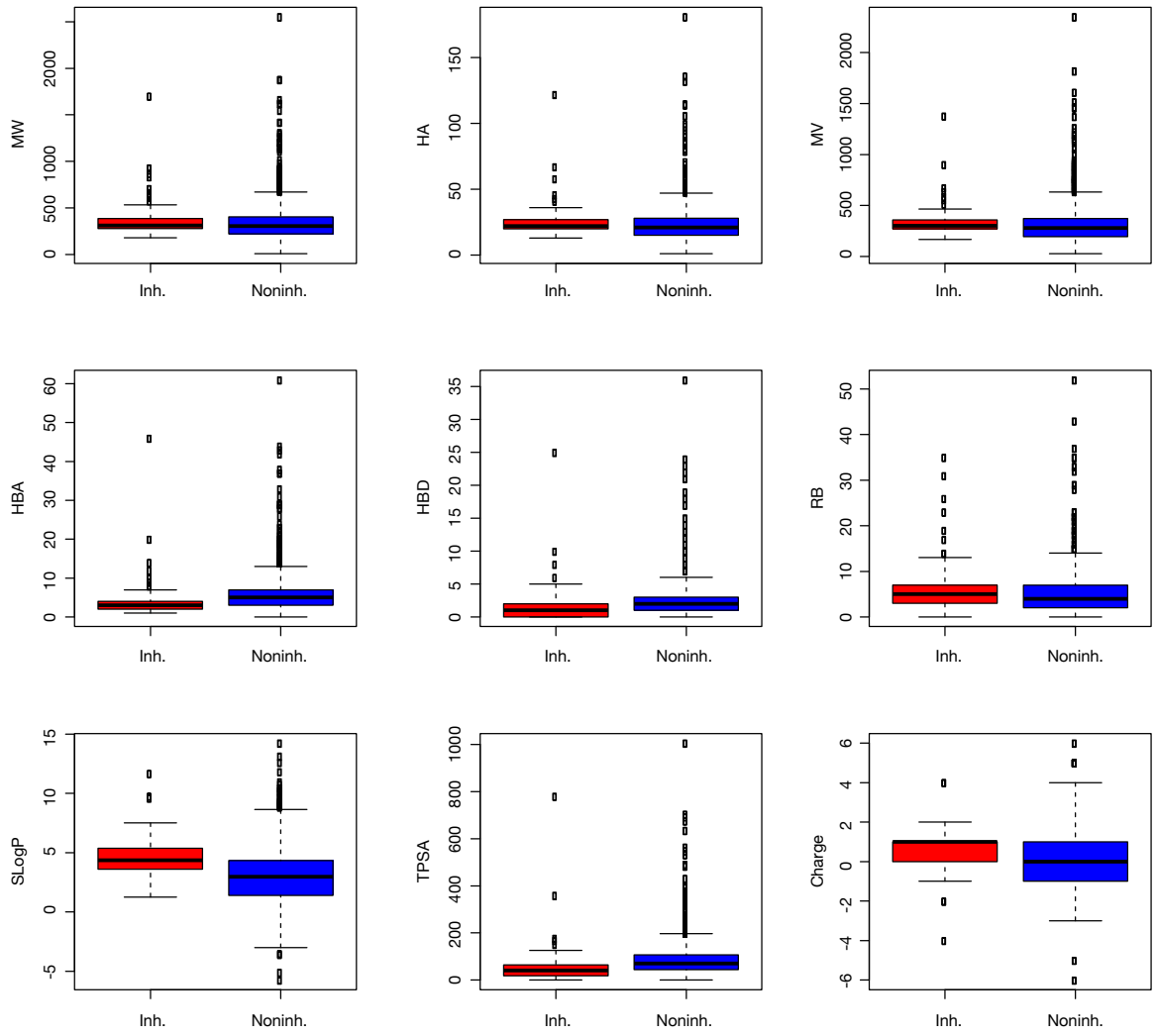


Figure 2.12 Differences in physicochemical properties of 167 inhibitors and 1,613 nonhibitors. Boxplots of MW: molecular weight, HA: number of heavy atoms, MV: molecular volume, HBD: number of hydrogen bond donors, HBA: number of hydrogen bond acceptors, RB: number of rotatable bonds, SLogP, TPSA: total polar surface area, and Charge at pH 7.4. Statistically significant differences were estimated using Student's t-test. Distributions of HBD, HBA, SLogP, TPSA, and Charge were significantly different between inhibitors and noninhibitors (p-value < 0.0001).

Prediction of competitive ligands Thirty of the 1,780 compounds in the library were identified as ligands by both virtual screening and HTS. Because docking was performed against the predicted binding site, we hypothesize that it can only identify competitive inhibitors or substrates. In contrast, HTS can identify both competitive and non-competitive inhibitors or substrates. Thus, the 30 compounds that were identified with both methods are likely to be competitive inhibitors or even substrates. If docking is accurate, the remaining 137 compounds are likely to bind non-competitively. To validate these predictions, Lineweaver-Burk plots were constructed for selected compounds in each class (**Figure 2.13**). Tacrine and ethopropazine, neither previously known to inhibit OCT1, were confirmed to be competitive inhibitors of ASP^+ and metformin uptake by OCT1 (**Figure 2.13A,B** and **Figure S2.17**). Furthermore, we assessed the relative competitiveness of these inhibitors by calculating K_{is} (the dissociation constant for the transporter-inhibitor complex) and K_{ii} (the dissociation constant for transporter-substrate-inhibitor complex) using previously published methods.[127] The ratios of K_{ii} and K_{is} for 14 compounds are summarized in **Table S2.2**. A larger K_{ii}/K_{is} value indicates a mode of inhibition that is relatively more competitive, whereas a lower K_{ii}/K_{is} value reflects less competitiveness. Our K_{ii}/K_{is} values were from 1.26 to 95.62 (**Table S2.2**). Thiamine had the highest value of K_{ii}/K_{is}, which indicated that it competitively inhibits ASP^+ uptake. In addition, thiamine has been validated as an OCT1 substrate in our previous publication.[107] Among the 14 compounds, compounds that were predicted to be competitive inhibitors resulted in relatively higher K_{ii}/K_{is}, whereas those that were predicted to be non-competitive inhibitors had lower K_{ii}/K_{is} values.

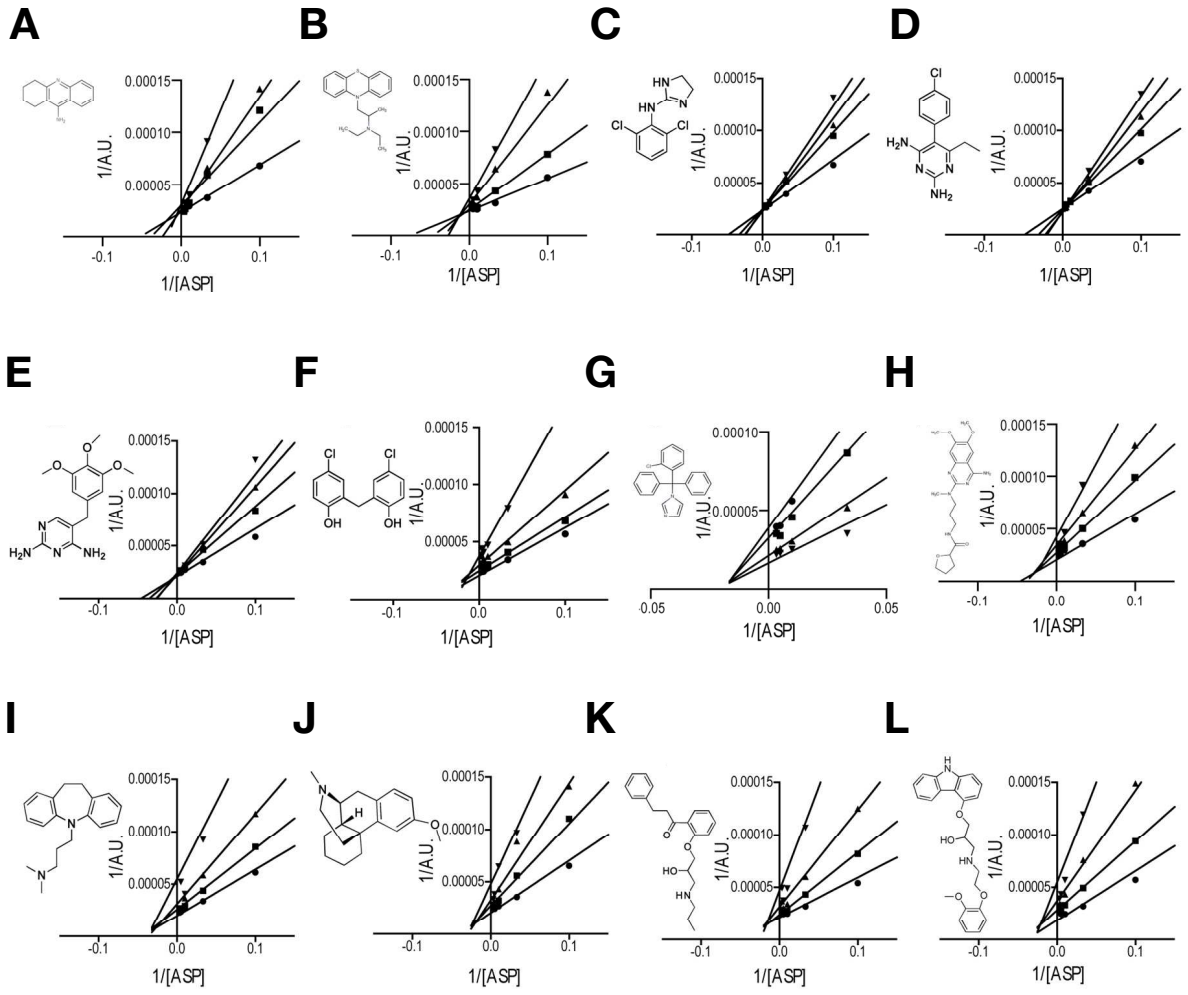


Figure 2.13 Lineweaver-Burk plots for discriminating between competitive and non-competitive inhibitors of OCT1. The inhibitory effects of selected inhibitors at various concentrations (●<■<▲<▼) were measured with increasing concentration of ASP+. The inhibitory effect of A) tacrine, B) ethopropazine, C) clonidine, D) pyrimethamine, E) trimethoprim, F) dichlorophene, G) alfuzosin, H) clotrimazole, I) imipramine, J) dextromethorphan, K) propafenone, L) carvedilol on ASP+ uptake by OCT1. Data represent mean values, n = 3 per data point. Error bars were not plotted in the other panels for the sake of clarity.

Physicochemical properties of putative competitive and non-competitive inhibitors We analyzed physicochemical properties of 30 putative competitive and 137 non-competitive ligands. Our analysis revealed that non-competitive ligands were significantly larger and more hydrophobic than competitive ligands (**Figure 2.10C**), predicting that the two types of ligands might bind to different sites on the transporter.

In liver, several solute carrier (SLC) transporters participate in the uptake of drugs across the sinusoidal membrane into hepatocytes. A comparison of our HTS result with those for two other liver uptake transporters, OATP1B1 (SLCO1B1) and OATP1B3 (SLCO1B3),[128,129] allowed us to identify OCT1-selective versus pan-inhibitors (i.e., compounds that inhibited transport of all three liver transporters). Fifty compounds inhibited the three liver transporters, whereas 112 inhibited OCT1 only. As expected, differences in charge were significant (Student t-test p-values < 0.05; **Figure S2.18**). Additionally, we found that OCT1 inhibitors were significantly smaller and less hydrophobic (Student t-test p-values < 0.05). This result showed that multi-way comparison of HTS data of several transporters can help in identifying compounds that are selective for a specific transporter and underscores the need for HTS of additional SLC transporters.

We next determined the fraction of inhibitors with different predicted inhibitory mechanisms among the OCT1-selective inhibitors and pan-inhibitors. Eight compounds inhibited OCT1 and OATP1B1, but not OATP1B3. Of these 8 compounds, only ethacridine lactate was predicted to competitively inhibit OCT1. Likewise, only 2 out of 10 OCT1/OATP1B3 inhibitors were predicted to inhibit OCT1 competitively and only 2 predicted competitive inhibitors were identified among 30 pan-inhibitors. Interestingly, 22 out of 30 predicted competitive ligands of OCT1 (73%) were found among the OCT1-selective inhibitors. That is, these 22 were not inhibitors of OATP1B1 or OATP1B3. In contrast, only 5 out 30 predicted competitive ligands (16%) were also inhibitors of OATP1B1 and OATP1B3 (the remaining 3 compounds were not screened in OATP1B1/B3 HTS). This finding not only supports the accuracy of our comparative model, but also highlights the importance of combining structure-guided ligand discovery with HTS. By combining experimental HTS with docking against comparative model of OCT1, we can efficiently predict competitive OCT1-selective inhibitors with 73% accuracy. Therefore, a structure-guided approach greatly accelerates the identification of selective inhibitors and provides valuable information for drug-drug interaction studies, compared to the conventional trial-and error approach. Pan-inhibitors,

OCT1/OATP1B1 and OCT1/OATP1B3 inhibitors were enriched for non-competitive inhibitors compared to OCT1-selective inhibitors.

Structure-activity modeling and validation Data from our OCT1 HTS experiment was used as a training set to construct a binary structure-activity relationship (SAR) model correlating molecular features of 1,780 compounds from Pharmacon library with their inhibitory activities, discretized into two classes: inhibitors and non-inhibitors. The random forest (RF) algorithm[130] was employed to build an ensemble classifier (SAR-I). We evaluated the accuracy of the SAR-I model (i.e., the area under the Receiver Operating Characteristic curve – ROC AUC), by 100 repeated cross-validation runs (**Figure 2.14A**). The average ROC AUC of RF classifiers in this retrospective validation was 0.89±0.3. This accuracy is comparable to the accuracies of retrospective validation of SAR models for other transporters.[121,131] In addition, we estimated the accuracy of RF-based SAR models in prospective validation as follows. First, we used molecular features and inhibitory outcomes of 183 compounds from a small previously published OCT1 inhibition screen[114] to develop a new SAR model (SAR-II). Next, SAR-II was utilized to predict the class of the 1,780 Pharmakon compounds. The ROC AUC for this prospective validation was 0.84 (**Figure 2.14B**). Thus, the decrease in accuracy measures between retrospective and prospective validation was only 5%, strongly suggesting that our OCT1 SAR-I model is highly accurate (approximately 84%). We also used SAR-II model to predict the sensitivity and specificity of the classifier at different cutoff values (**Figure 2.14C**). The sensitivity and specificity of SAR-II model at a cutoff value of 0.6 were respectively 92% and 43%; at a cutoff value of 0.4, they were 82% and 65%, respectively. Finally, the observed and SAR-II predicted classification scores of the 1,780 Pharmakon compounds were moderately correlated (Pearson correlation coefficient of 0.50; **Figure 2.14D**). These results suggest that SAR models can accurately predict OCT1 ligands by virtual screening.

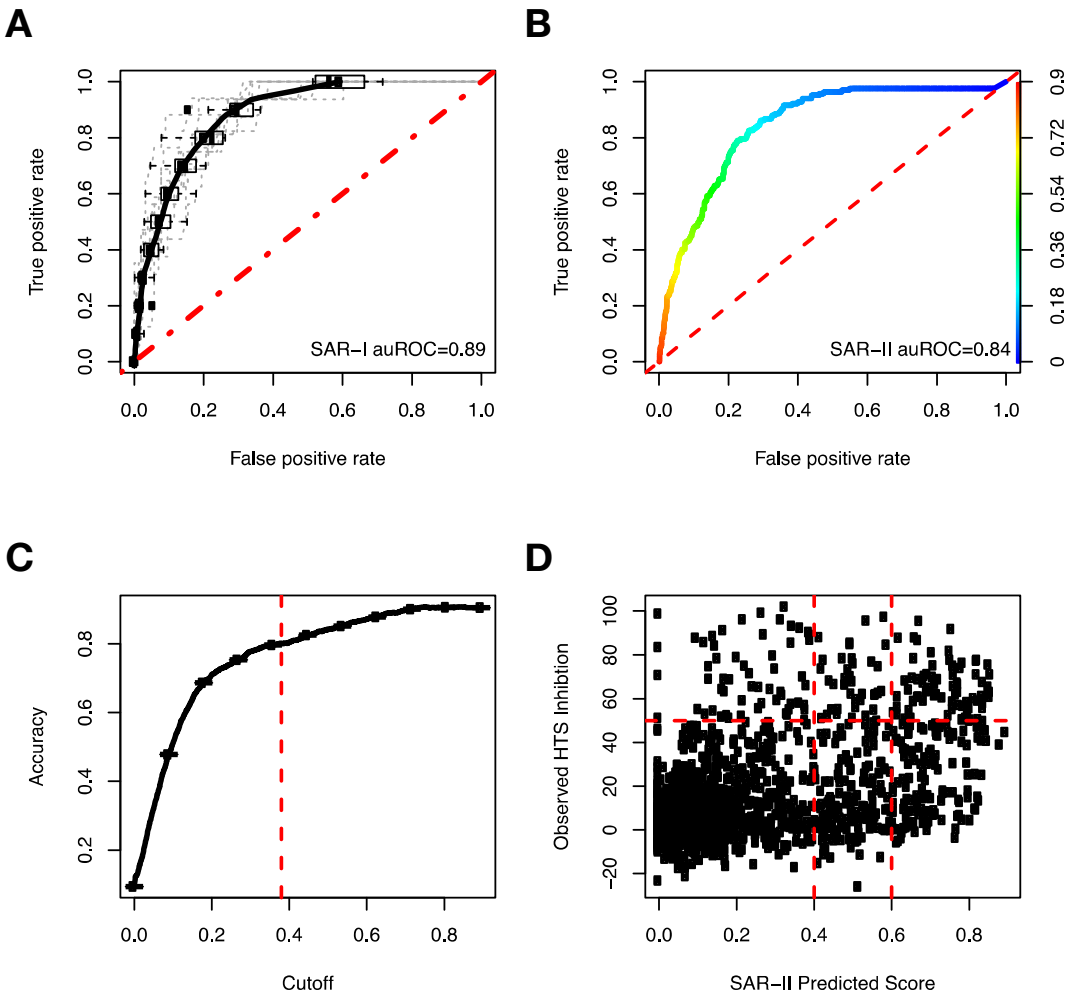


Figure 2.14 Results of structure-activity relationship modeling. A) ROC curves for 100 retrospective cross-validation runs. Average ROC curve is shown in black. ROC curve of a random classifier is shown as a read dotted line. B) ROC curve for the SAR model tested in prospective validation of 1,780 predicted inhibition values. Performance at different cutoff values (shown on the right y-axis) are indicated by rainbow colors. C) Accuracy of classification of 1,780 compounds using SAR-II model as a function of cutoff. A cutoff of 0.38 is indicated by the red dotted line. D) Observed vs. predicted inhibition values for 1,780 compounds. Two classification cutoffs are drawn vertically and HTS classification cutoff is drawn horizontally.

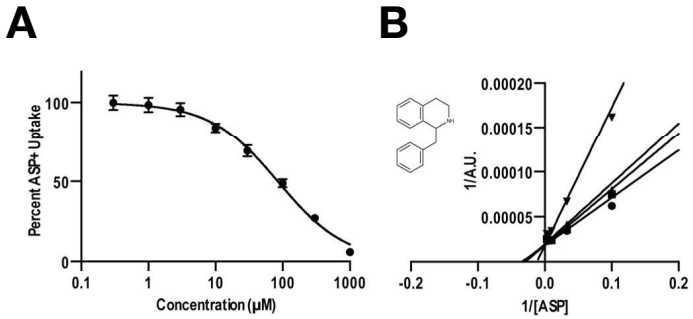


Figure 2.15 The inhibitory effects of 1BnTIQ on OCT1 transport. A) 1BnTIQ inhibited ASP+ uptake by OCT1 and the IC_{50} was determined to be 82.1 μM. B) The inhibitory effect of 1BnTIQ on ASP+ uptake by OCT1 indicated competitive inhibition. Data represent mean ± s.d., n = 3 per data point.

Virtual screening of endogenous and drug metabolites In addition to searching for OCT1 ligands among the 1,780 prescription drugs, we applied our structure-based and SAR model methods to predict OCT1 ligands among a larger set of 29,332 endogenous and drug metabolites in the Human Metabolome Database (HMDB).[132] 864 out of 29,332 (3%) of compounds docked favorably. We then computed ligand-OCT1 interaction values for the 864 compounds using the SAR-I model, allowing us to predict 146 competitive ligands. Among these 146 compounds, 1-Benzyl-1,2,3,4-tetrahydroisoquinoline (1BnTIQ), an endogenous amine present at high level in the cerebrospinal fluid of Parkinson's disease patients,[133,134] docked favorably against OCT1 (rank 14) and had an SAR-I inhibition score of 0.48. 1BnTIQ inhibited OCT1 at 82.1 μM, and Lineweaver-Burk plot confirmed that 1BnTIQ inhibited OCT1 competitively (**Figure 2.15**). This validation suggests that combined docking and SAR-I virtual screening can accurately predict OCT1 metabolite ligands, in addition to prescription drug ligands.

Discussion and conclusions

OCT1, a protein of great pharmacologic interest, transports a wide array of drugs into and out of the liver, and thus serves as a major determinant of drug metabolism and action. Because of its clinical importance, OCT1 has become a focus of many pharmacogenomics and drug interaction studies, which have prompted the EMA to recommend that all new drugs undergo *in vitro* testing to assess their liability to interact with OCT1.[112] The goals of the current study were (i) to develop robust computational models to predict the interaction of new molecular entities with OCT1, and (ii) to use a combination of experimental and computational approaches to identify prescription drug and metabolite ligands of OCT1. By combining *in silico* and *in vitro* approaches, we sought to gain information about whether a ligand binds competitively or non-competitively on OCT1.

Three major findings emerged from the current studies. First, a comparative structure model of OCT1 successfully discriminated ligands from non-ligands of the transporter. Second, by combining molecular docking and HTS approaches, we were indeed successful in determining whether a ligand interacts competitively or non-competitively with the transporter. Third, we identified 30 and 137 prescription drugs as competitive and non-competitive ligands of OCT1, respectively, including drugs not known to interact with the transporter. We now discuss each one of these findings in turn.

OCT1 ligands can be identified accurately by virtual screening against an OCT1 comparative model Accurate prediction of inhibitors and substrates of OCT1 is challenging. First, although several atomic structures have been resolved for bacterial members of the Major Facilitator Superfamily of transporters, low sequence identity between bacterial homologs and human OCT1 casts doubt on the accuracy of comparative models built using these structures as templates.[28] Second, OCT1 is considered a polyspecific transporter that transports compounds of different sizes and molecular features. For example, OCT1 mediates the uptake of compounds ranging from small cations such as tetraethylammonium, monoamines (metformin), chemotherapy drugs (oxaliplatin), to hormone-like lipid compounds (prostaglandin E1).[14,109,135,136] This broad specificity may result in inaccurate docking and SAR models. For example, ligand-based models may under-predict ligands, especially in chemical spaces not sampled by the training set. Similarly, docking generally does not consider multiple binding sites, which are characteristic of polyspecific transporters. Finally, docking also depends on the accuracy of the target structure and therefore, a combination of several computational and experimental validation experiments should be performed to ascertain the applicability of a homology model for docking.

Here, we built a comparative model of human OCT1 in an inward-facing occluded conformation, using a recently determined structure of its eukaryotic homolog, a phosphate transporter from *Piriformospora indica* (PipT).[27] We confirmed that 80% of known structurally diverse substrates can be docked favorably against the predicted binding site (**Table S2.1**). Three compounds that had unfavorable docking scores against this binding site had steric clashes with side chains of the binding cavity, underscoring the need for modeling transporters in alternative conformations. Unfortunately, there are no template structures for alternative conformations of OCT1 and molecular dynamics simulations still lack computing power to model conformational changes of transporters. The average pairwise Tanimoto coefficient of substrates used to validate the predicted binding site was 0.33, indicating that the compounds were structurally unrelated (**Figure S2.19**). In an unbiased validation of the OCT1 comparative model, we docked 1,780 compounds from the Pharmakon library and validated 70% of predicted binders and non-binders *in vitro*. Previous SAR and pharmacophore models of OCT1-ligand interactions identified hydrophobicity and charge as the main physicochemicalproperties required for inhibition.[114,137,138] Our screening results confirmed that charge and hydrophobicity positively correlated with inhibitory activity of compounds (**Figure 2.12**). In addition, ligands had fewer hydrogen bond donors and acceptors and were less polar than non-ligands.

A combination of virtual screening and HTS can determine whether a ligand binds competitively or non-competitively It has been frequently assumed that inhibition of SLC-mediated transport is primarily competitive, although some recent studies have challenged this assumption by showing that competitive, non-competitive, and mixed type inhibition can also occur.[115,139-141] The International Transporter Consortium recently pointed out that the lack of understanding of inhibition mechanisms remains a limiting factor in transporter studies in drug development.[142] We showed that inhibitors of OCT1-mediated ASP^+ transport can be divided into two groups based on their docking scores against the OCT1 model. Predicted competitive inhibitors of ASP^+ transport are compounds that are identified by experimental screening as well as predicted by virtual screening against the predicted substrate binding site; our assay and calculation do not distinguish between competitive inhibitors and substrates. In contrast, predicted non-competitive inhibitors are compounds that are identified by experimental screening as well as predicted not to bind by virtual screening. Competitive ligands were significantly smaller and less hydrophobic than the non-competitive ligands (**Figure 2.10C**). Because we modeled OCT1 in an inward-facing occluded conformation, we predict competitive ligands from only those compounds that fit into the compact translocation cavity. In addition, a broad range of K_{ii}/K_{is} (**Table S2.2**) suggested various degrees of competitiveness among these inhibitors. Some of the inhibitors may bind to another binding cavity, or to alternative OCT1 conformations, and these will not be predicted as competitive ligands by our docking approach. Furthermore, interactions with OCT1 have been shown to be ligand dependent.[115] We used ASP^+ as our primary probe substrate in this study, but the mechanism of ligand-dependent interaction will be needed to be further investigated.

Additionally, we showed that 87% of predicted competitive ligands were selective for OCT1. In contrast, only 13% of predicted competitive ligands were identified among inhibitors of three hepatic uptake transporters OCT1, OATP1B1, and OATP1B3. This result provides further confidence in the accuracy of the comparative model and its ability to predict competitive ligands.

To predict competitive ligands in large virtual library of potential ligands, we combined the SAR model (built using our HTS data) and docking against a comparative OCT1 model. Indeed, 146 putative competitive ligands among endogenous and exogenous metabolites were predicted in the HMDB library; one of them, 1BnTIQ, was experimentally tested and validated (**Figure 2.15**). 1BnTIQ is an endogenous amine detected in human cerebrospinal fluid that accumulates in patients with Parkinson's disease[133] and is able to induce parkinsonism in both mice and monkeys.[143,144] 1BnTIQ

inhibits complex I in the mitochondria and induces dopaminergic death in the same manner as MPP^+, a neurotoxin also known to induce parkinsonism. Structurally similar to MPP^+, 1BnTIQ is hydrophilic and requires an uptake mechanism to enter cells. A likely uptake mechanism is suggested by our identification of 1BnTIQ as an OCT1 ligand.

Novel OCT1 inhibitors can be identified by HTS Because of the critical role OCT1 plays in drug disposition and response, efforts have been made to identify and characterize OCT1 inhibitors. For example, 20 pharmacologically diverse antidepressants and 14 antipsychotics were screened using an OCT1 mediated radiolabeled MPP^+ uptake assay, identifying drugs that could potentially inhibit 50% or more OCT1 activity in the brain.[113] In another study, 191 drugs from various sources were compiled, followed by a medium-throughput identification of 62 inhibitors.[114]

In this study, we conducted an extensive HTS of 1,780 drugs that have reached at least clinical trials in the United States or are marketed in Europe and/or Asia. We were able to confirm most of the previously known inhibitors. We also estimate that we identified at least 100 compounds previously unknown to interact with the transporter. Moreover, we grouped the identified OCT1 ligands into therapeutic classes, including tricyclic antidepressants, antihistamines, steroids, and α-adrenergic receptor agonists, all of which were previously reported to be more likely to interact with OCT1.[114] Our HTS also identified additional drug classes that were enriched in OCT1 ligands, including β-adrenergic receptor agonists/antagonists, calcium channel blockers, and muscarinic acetylcholine receptors agonists/antagonists. Among the inhibitors identified by HTS, select compounds were validated by determining their IC_{50} values (**Figure 2.11** and **Table 2.1**). Five of them (carbetapentane, carvedilol, erlotinib, griseofulvin, and ketoconazole) had IC_{50} values that were at least 10% of their maximum plasma concentrations achieved after therapeutic doses of the drug (**Table 2.2**). These estimates suggest the possibility of clinical drug-drug interactions with OCT1 substrates. Further, OCT1 inhibitors may potentially have beneficial effects on hepatic steatosis.[107] With the exception of erlotinib and ketoconazole, the drugs noted here were newly identified OCT1 ligands. We also estimated their IC_{50} for mouse OCT1 and other two human SLC uptake transporters (OCT2 and MATE1) (**Table S2.3**).

Table 2.2 Identified OCT1 inhibitors that could cause drug-drug interactions.

Name	IC_{50}*, μM	C_{MAX} †, μM	$C_{Portal\ Vein}$ ‡	C_{MAX} / IC_{50}
Carbetapentane	1.6	0.2	N.A.	0.12
Carvedilol	3.4	0.4	0.7	0.12
Erlotinib	16.2	4.8	7.0	0.30
Griseofulvin	7.3	4.5	4.6	0.62
Ketoconazole	2.6	6.6	10.0	2.54

* IC_{50} is the estimated half maximal inhibitory concentration (Materials and Methods).

† C_{MAX} values were obtained from http://www.micromedexsolutions.com/

‡ C_{Portal} Vein values were calculated based on equation previously described[60]

In conclusion, we developed a comparative structural model of OCT1 that discriminates ligands from non-ligands, and used the model together with an *in vitro* HTS assay. By combining the two approaches, we were able to predict whether a ligand binds competitively or non-competitively. The structure-guided approach also accurately predicted inhibitors specific to OCT1 rather than two other hepatic drug transporters, OATP1B1 and OATP1B3. Finally, we conducted a virtual screen against a metabolite library using both comparative and SAR models built from HTS data, and accurately identified and validated the parkinsonism-producing neurotoxin, 1BnTIQ, as a competitive inhibitor of OCT1.

Experimental procedures

Chemicals The MicroSource Pharmakon compound library (Gaylordsville, CT) was obtained through the Small Molecular Discovery Center at University of California, San Francisco (San Francisco, CA). 4-(4-(dimethylamino)styryl)-N-methylpyridinium (ASP^+) iodide was purchased from Molecular Probes (Grand Island, NY). All other chemicals were purchased from Sigma-Aldrich (St. Louis, MO). All chemicals used in the studies are purchased from commercial vendors and with 98% or higher purity. All cell culture media and supplements were purchased from Life Technologies (Carlsbad, CA) except fetal bovine serum, which was purchased from GE Healthcare Life Sciences (South Logan, UT).

Cell Culture Human embryonic kidney (HEK-293) cell line stably overexpressing OCT1 was established previously in our laboratory.[107] The cells were maintained in Dulbecco's Modified Eagle's Medium (DMEM H-21) supplemented with 75 μg/ml of hygromycin B, penicillin (100 U/ml), streptomycin (100 mg/ml), and 10% fetal bovine serum in a humidified atmosphere with 5% CO_2 at 37 °C.

***In vitro* uptake studies** HEK-293 cells overexpressing OCT1 were seeded in black, clear bottom poly-D-lysine coated 96-well plates (Greiner Bio-One, Monroe, NC) and allowed to grow for 48 hours until approximately 90% confluency. For uptake kinetics study, cells were incubated with HBSS containing serial dilution of ASP^+ for 2 minutes at 37 °C. At the end of experiments, the media were aspirated and the cells were washed twice with ice-cold HBSS containing 50 μM spironolactone. The K_m and V_{max} were calculated by fitting the data to Michaelis-Menten equations. For time course study, cells were incubated with HBSS containing 2 μM ASP^+ at 37 °C. At various

time points, the experiment was stopped as previously described. For IC_{50} determination, cells were incubated with HBSS containing 2 μM ASP^+ and serial dilution of inhibitors for 2 minutes at 37°C. IC_{50} was determined using appropriate curve fitting. For Lineweaver-Burk plots, cells were incubated with HBSS containing serial dilution of ASP^+ and the inhibitor of interest at 4 different fixed concentration for 2 minutes at 37 °C. The reciprocal value of ASP^+ uptake at each inhibitor concentration was fitted with linear regression. The signal of ASP^+ was measured using an Analyst AD plate reader (Molecular Devices, Sunnyvale, CA) with excitation and emission filters tuned at 485 and 585 nm wavelength, respectively. All statistical analysis and curve fitting were done using GraphPad Prism software (La Jolla, CA).

High-throughput screening The high-throughput screen was performed at the Small Molecule Discovery Center at the University of California, San Francisco. HEK-293 cells overexpressing OCT1 were seeded in black, clear bottom poly-D-lysine coated 96-well plates (Greiner Bio-One, Monroe, NC) and allowed to grow for 48 hours until approximately 90% confluency using methods established previously.[121] Cells were incubated with HBSS containing 2 μM ASP^+ and 20 μM of test compounds at ambient temperature for approximately 2 minutes. At the end of the experiment, media were aspirated and cells were washed twice with HBSS containing 50 μM spironolactone. Nonspecific transport was determined in wells on each assay plate using 100 μM spironolactone as OCT1 inhibitor. The screen was carried out with a Biomek FXp liquid handler (Beckman Coulter, Brea, CA). Fluorescence was measured as previously described.

OCT1 structure modeling and docking Human OCT1 was modeled based on the 2.9 Å structure of a high-affinity phosphate transporter PiPT, from *Piriformospora indica*, crystallized in an inward-facing occluded state with bound phosphate.[116] The template was selected based on the shared MFS fold assignment,[143] structure quality, sequence similarity to OCT1, and the ligand-bound conformation. The sequence alignment was obtained by a manual refinement of gaps in the output from the PROMALS3D[98] and MUSCLE[99] servers. One hundred models were generated using the automodel class of MODELLER 9.14,[100] and the normalized discrete optimized protein energy (zDOPE) potential.[101] The top scoring model was used to predict putative binding sites with the FTMap web server.[145] Two of the predicted bindings sites were identified in the translocation cavity between the two domains.[116] ASP^+ probe substrate was docked against the two binding sites with UCSF DOCK 3.6.[146] The size of the docking box was 38 x 40 x 38 Å. The pose with the best docking score was used as the template for a subsequent round of comparative modeling by MODELLER

9.14, generating 100 new models of OCT1. Each of the models was then evaluated for ligand enrichment from a set of challenging decoys based on enrichment curves and corresponding logAUC values.[146,147] Sixty decoys were generated using the Database of Useful Decoys (DUD)[148] for each of the selected experimentally validated substrates of OCT1. The best scoring model was used for subsequent virtual screening. Compounds in the Pharmakon library were downloaded from the ZINC database and docked against the predicted binding site on the comparative model, using UCSF DOCK 3.6. A negative DOCK score predicts a favorable interaction with the transporter, while a positive score predicts an unfavorable (or unlikely) intermolecular interaction. Normalized docking scores were computed by subtracting the average docking score of all compounds (including the non-binders) from the docking score of an individual compound and dividing by the standard deviation of all docking scores. Online web server PAINS-Remover was used to check Pharmkon compounds for the likelihood of interference in screening.[149] None of the 1,780 compounds was reported as a pan-assay interference compound.

Structure-activity relationship modeling of OCT1 inhibition Two-dimensional structure files of the 1,780 compounds from the Pharmakon library were also downloaded from the ZINC database. 2,900 molecular descriptors and charge (pH 7.4) for each compound were computed using PaDEL software[150] and ChemAxon cxcalc program (http://www.chemaxon.com), respectively. Non-informative descriptors (i.e., molecular descriptors with near zero variance and redundant descriptors defined by a correlation higher than 0.95 to an accepted descriptor) Next, information content and correlation between descriptor values and percent inhibition of 1,780 compounds were computed with the cfs filtering algorithm in the Fselector package[151] in R and 21 most informative descriptors were retained for further modeling (Supplemental Data). Percent inhibition values determined by HTS were discretized into two outcomes. Values of at least 50% were mapped to '1' (i.e., ligand) and values of less than 50%, including negative values, were mapped to '0' (i.e., nonligand). Binary SAR models were built with the RF algorithm.[130] Their accuracy was estimated using a double loop five-fold cross-validation[121] protocol in the caret package in R. To evaluate the accuracy of models, the average area under the Receiver Operating Characteristic curve (ROC AUC) was computed from 100 repeated double-loop cross-validation runs. For prospective validation, 21 molecular descriptors for 183 compounds[114] were computed (Supplemental Data) and SAR models were optimized by repeated 5-fold cross-validation. The model was employed to predict inhibition values for the 1,780 compounds in the Pharmakon library.

Virtual Screening against HMDB library For virtual screening, molecular structure files of the HMDB compounds were downloaded from the ZINC database[152] and 21 molecular features (Supplemental Data) were computed for each compound. The ligand score ranging between 0 and 1 was computed for each compound by SAR-I. Compounds with the ligand score higher than 0.4 were labeled as ligands. In parallel, each compound was docked against predicted binding sites on the OCT1 transporter using UCSF DOCK 3.6. Normalized docking scores were computed by subtracting the average docking score of all docked compounds from the docking score of a compound and dividing by the standard deviation of all docking scores. Compounds with normalized docking scores less than -1 and ligand scores greater than 0.4 were labeled as competitive inhibitors.

Pairwise compound similarity computation and clustering The MayaChemTools (http://www.mayachemtools.org) package was used to compute two-dimensional Extended Connectivity Fingerprints (ECFP) and pairwise Tanimoto coefficients. The hclust function from the stats library[153] was used to perform hierarchical compound clustering.

Supplemental information

Supplemental figures and tables are included below.

Table S2.1 Summary of 15 substrates used to validate OCT1 docking site.

CID	Name	Smiles	DOCK score
2022	Acyclovir	C1=NC2=C(N1COCCO)NC(=NC2=O)N	-24.44
4091	Metformin	CN(C)C(=N)N=C(N)N	-21.64
39484	MPP+	C[N+]1=CC=C(C=C1)C2=CC=CC=C2	-20.29
3454	Ganciclovir	C1=NC2=C(N1COC(CO)CO)NC(=NC2=O)N	-18.71
457	N(1)-methylnicotinamide	C[N+]1=CC=CC(=C1)C(=O)N	-14.24
5816	Epinephrine	CNCC(C1=CC(=C(C=C1)O)O)O	-11.49
305	Choline	C[N+](C)(C)CCO	-11
5413	TEA	CC[N+](CC)(CC)CC	-10.47
681	Dopamine	C1=CC(=C(C=C1CCN)O)O	-6.40
5333955	ASP+	C[N+]1=CC=C(C=C1)C=CC2=CC=C(C=C2)N(C)C.[I-]	-6.08
439260	Norepinephrine	C1=CC(=C(C=C1C(CN)O)O)O	-5.69
6857599	Oxaliplatin	C1CCC(C(C1)N)N.C(=O)(C(=O)O)O.[Pt+2]	-2.88
4735	Pentamidine	C1=CC(=CC=C1C(=N)N)OCCCCCOC2=CC=C(C=C2)C(=N)N	103.78
5280363	Dinoprost	CCCCCC(C=CC1C(CC(C1CC=CCCCC(=O)O)O)O)O	128.26
5280360	Dinoprostone	CCCCCC(C=CC1C(CC(=O)C1CC=CCCCC(=O)O)O)O	171.01

Table S2.2 The ratios of Kii and Kis for selective inhibitors.

Name	K_{ii}/ K_{is}
Imipramine	1.26
Clotrimazole	1.27
Detromethorphan	1.44
Carvedilol	1.82
Propafenone	2.19
Alfuzosin	2.27
Dichlorophene	3.48
Ethopropazine	5.88
Trimethoprim	7.29
Pyrimethamine	8.04
Tacrine	10.05
1BnTIQ	14.60
Clonidine	58.40
Thiamine	95.62

Table S2.3 Comparison of IC50 of OCT1 inhibitors in 4 different cell lines.

Name	HEK-hOCT1	HEK-hOCT2	HEK-hMATE1	HEK-mOCT1
Carbetapentane	1.6	2.7	45.8	0.7
Carvedilol	3.4	12.1	57.7	2.6
Erlotinib	16.2	3.3	650.5	6.5
Griseofulvin	7.3	136.3	105.6	24.0
Ketoconazole	2.6	2.1	2.7	0.7

* IC_{50} (μM) is the estimated half maximal inhibitory concentration (Materials and Methods).

** IC_{50} is the average of two independent experiments

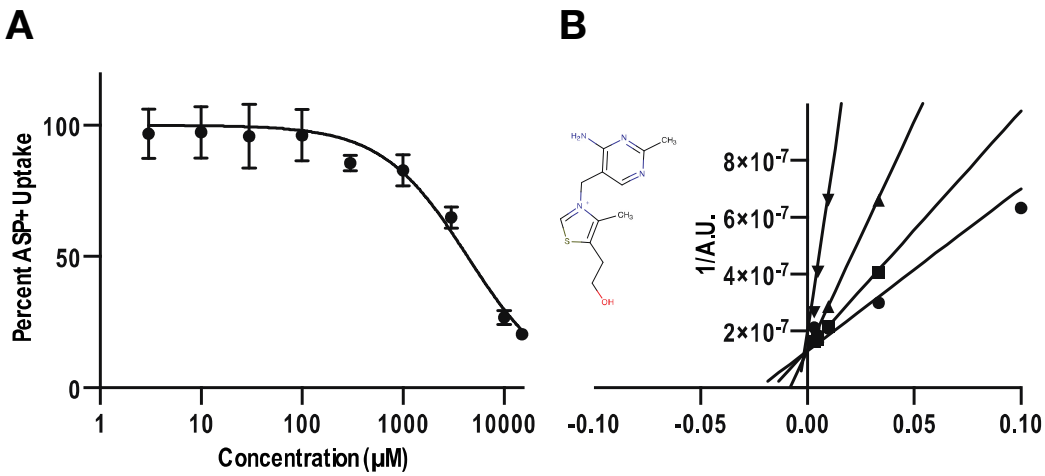


Figure S2.15 The effect of thiamine on the uptake of ASP+ by OCT1. (A) Thiamine inhibited ASP+ uptake by OCT1 and the IC_{50} was determined at 4.1 mM. (B) The inhibitory effect of thiamine at at 0 μM (●), 20 μM (■), 1000 μM (▲), 5 mM (▼) against increasing concentrations of ASP+ showed that thiamine inhibited OCT1 competitively. Data represent mean ± s.d., n = 3 per data point.

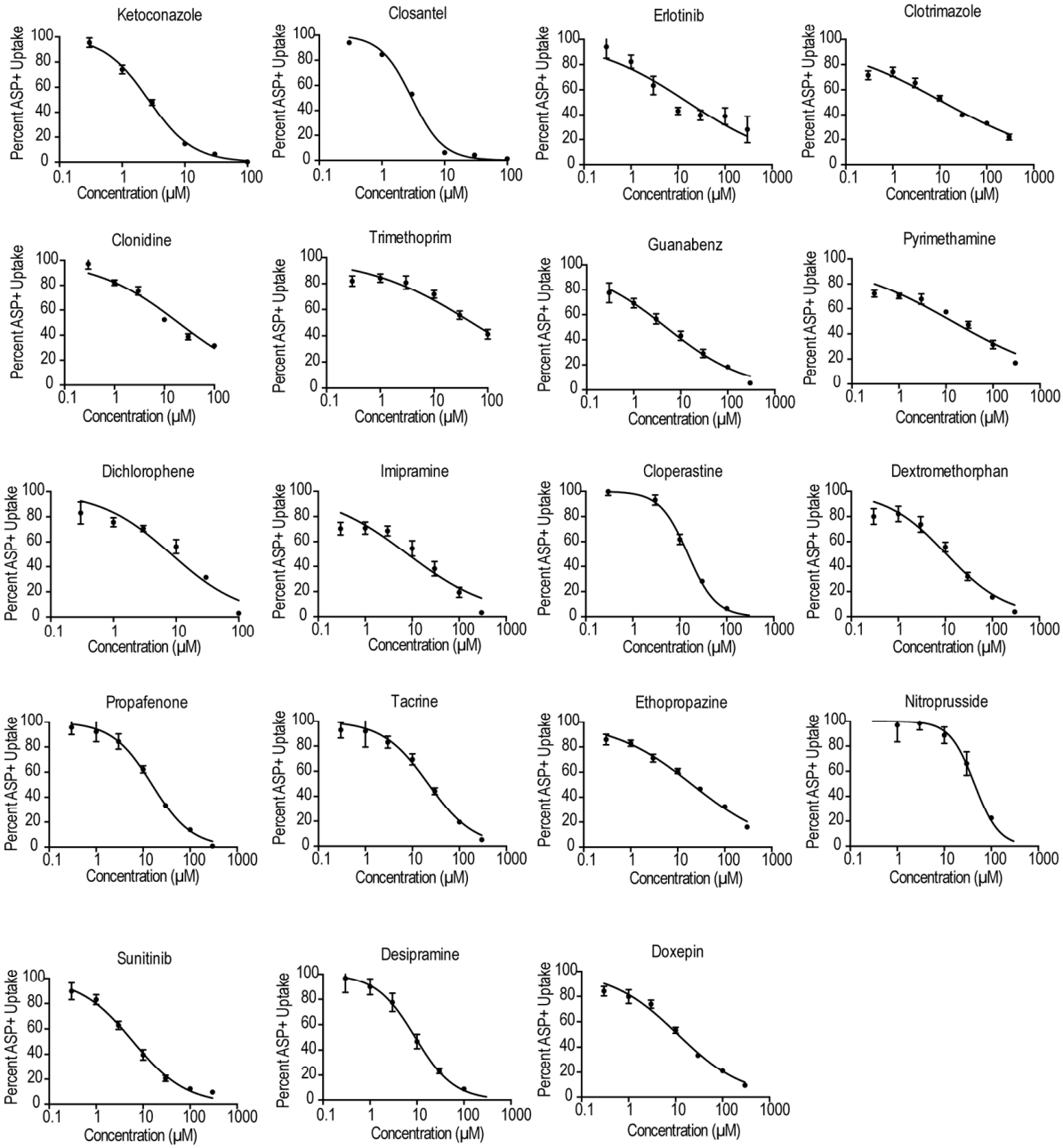


Figure S2.16 Selected inhibition studies of OCT1 ligands and their estimated IC_{50}. Data represent mean ± s.d., n = 6 per data point.

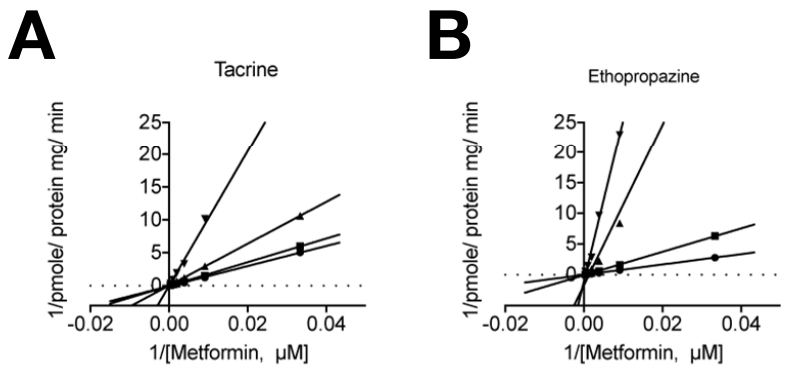


Figure S2.17 Lineweaver-Burk plots for discriminating between competitive and non-competitive inhibitors of OCT1 mediated metformin uptake. The inhibitory effects of selected inhibitors at various concentrations (●<■<▲<▼) were measured with increasing concentration of metformin. Data represent mean values, n = 3 per data point.

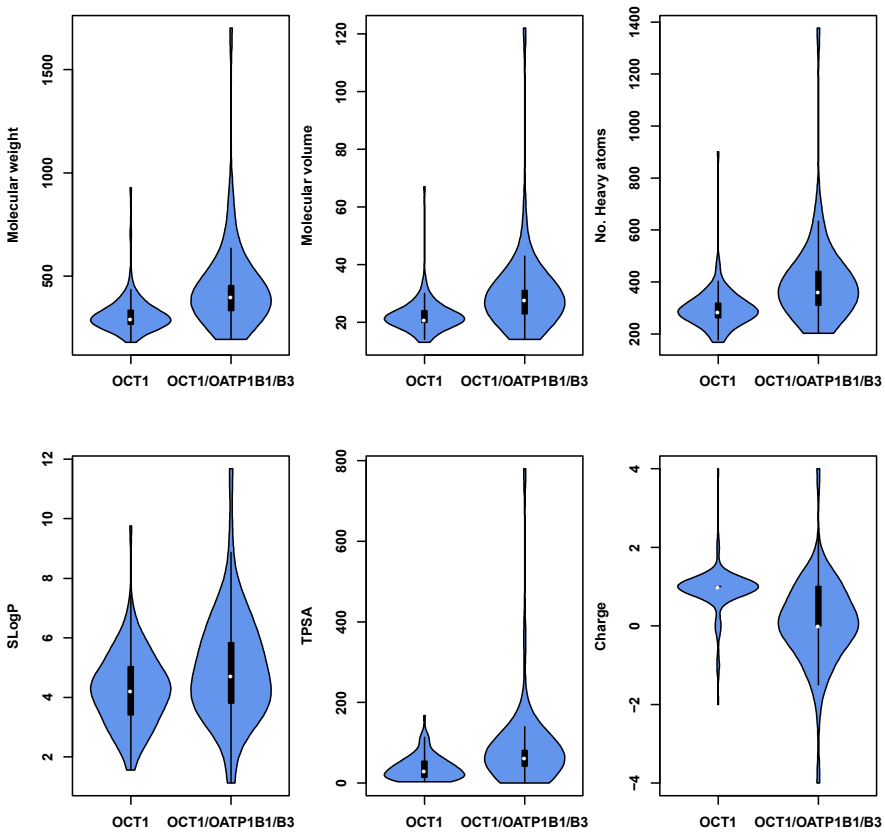


Figure S2.18 Differences in physicochemical properties of OCT1-selective inhibitors and pan-inhibitors. Only the distributions of statistically significant properties are shown (Student t-test p-value < 0.05).

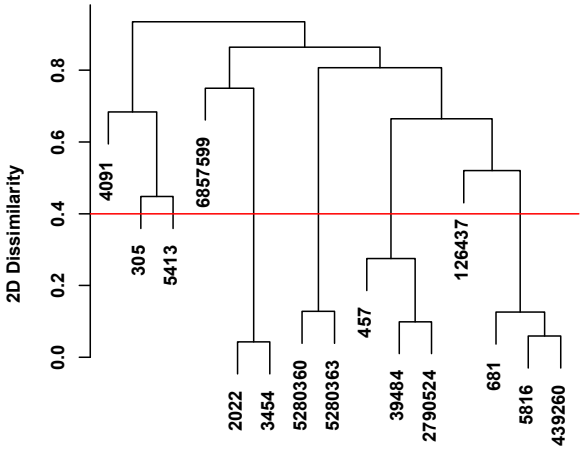


Figure S2.19 2D dissimilarity clustering of 15 known substrates of OCT1. Dissimilarity (i.e., 1 – pairwise Tanimoto similarity) between known OCT1 substrates (Table S1) is shown on the y-axis. The labels in the dendrogram denote PubChem CIDs of 15 known OCT1 substrates, which were used to validate the binding site on comparative OCT1 model. Compounds with dissimilarity score less than 0.4 (shown by the red line) are considered similar.

References

1. Gründemann D, Gorboulev V, Gambaryan S, Veyhl M, Koepsell H. Drug excretion mediated by a new prototype of polyspecific transporter. *Nature*. 1994;372(6506):549-552. doi:10.1038/372549a0.
2. Zhang L, Dresser MJ, Gray AT, Yost SC, Terashita S, Giacomini KM. Cloning and functional expression of a human liver organic cation transporter. *Mol Pharmacol*. 1997;51(6):913-921.
3. Zhang L, Schaner ME, Giacomini KM. Functional Characterization of an Organic Cation Transporter (hOCT1) in a Transiently Transfected Human Cell Line (HeLa). *J Pharmacol Exp Ther*. 1998;286(1):354-361. doi:10.1074/jbc.271.51.32599.
4. Koepsell H, Schmitt BM, Gorboulev V. Organic cation transporters. *Rev Physiol Biochem Pharmacol*. 2003;150(Chapter 2):36-90. doi:10.1007/s10254-003-0017-x.
5. Koepsell H. The SLC22 family with transporters of organic cations, anions and zwitterions. *Molecular Aspects of Medicine*. 2013;34(2-3):413-435. doi:10.1016/j.mam.2012.10.010.
6. Nies AT, Koepsell H, Winter S, et al. Expression of organic cation transporters OCT1 (SLC22A1) and OCT3 (SLC22A3) is affected by genetic factors and cholestasis in human liver. *Hepatology*. 2009;50(4):1227-1240. doi:10.1002/hep.23103.
7. Motohashi H, Sakurai Y, Saito H, et al. Gene expression levels and immunolocalization of organic ion transporters in the human kidney. *J Am Soc Nephrol*. 2002;13(4):866-874.
8. Koepsell H, Lips K, Volk C. Polyspecific organic cation transporters: structure, function, physiological roles, and biopharmaceutical implications. *Pharm Res*. 2007;24(7):1227-1251. doi:10.1007/s11095-007-9254-z.
9. Dresser MJ, Leabman MK, Giacomini KM. Transporters involved in the elimination of drugs in the kidney: organic anion transporters and organic cation transporters. *J Pharm Sci*. 2001;90(4):397-421.
10. Wang D-S, Jonker JW, Kato Y, Kusuhara H, Schinkel AH, Sugiyama Y. Involvement of organic cation transporter 1 in hepatic and intestinal distribution of metformin. *J Pharmacol Exp Ther*. 2002;302(2):510-515. doi:10.1124/jpet.102.034140.
11. Kimura N, Masuda S, Tanihara Y, et al. Metformin is a superior substrate for renal organic cation transporter OCT2 rather than hepatic OCT1. *DMPK*. 2005;20(5):379-386.
12. Errasti-Murugarren E, Pastor-Anglada M. Drug transporter pharmacogenetics in nucleoside-based therapies. *Pharmacogenomics*. 2010;11(6):809-841. doi:10.2217/pgs.10.70.
13. Andreev E, Brosseau N, Carmona E, Mes-Masson A-M, Ramotar D. The human organic cation

References (continued)

transporter OCT1 mediates high affinity uptake of the anticancer drug daunorubicin. *Sci Rep*. 2016;6:20508. doi:10.1038/srep20508.

14. Zhang S, Lovejoy KS, Shima JE, et al. Organic cation transporters are determinants of oxaliplatin cytotoxicity. *Cancer Res*. 2006;66(17):8847-8857. doi:10.1158/0008-5472.CAN-06-0769.
15. Yonezawa A, Masuda S, Yokoo S, Katsura T, Inui K-I. Cisplatin and oxaliplatin, but not carboplatin and nedaplatin, are substrates for human organic cation transporters (SLC22A1-3 and multidrug and toxin extrusion family). *J Pharmacol Exp Ther*. 2006;319(2):879-886. doi:10.1124/jpet.106.110346.
16. More SS, Li S, Yee SW, et al. Organic cation transporters modulate the uptake and cytotoxicity of picoplatin, a third-generation platinum analogue. *Mol Cancer Ther*. 2010;9(4):1058-1069. doi:10.1158/1535-7163.MCT-09-1084.
17. Chen L, Shu Y, Liang X, et al. OCT1 is a high-capacity thiamine transporter that regulates hepatic steatosis and is a target of metformin. *Proc Natl Acad Sci USA*. 2014;111(27):9983-9988. doi:10.1073/pnas.1314939111.
18. Chen L, Yee SW, Giacomini KM. OCT1 in hepatic steatosis and thiamine disposition. *Cell Cycle*. 2015;14(3):283-284. doi:10.1080/15384101.2015.1006532.
19. Arimany Nardi C, Montraveta A, Lee-Vergés E, et al. Human organic cation transporter 1 (hOCT1) as a mediator of bendamustine uptake and cytotoxicity in chronic lymphocytic leukemia (CLL) cells. - PubMed - NCBI. *Pharmacogenomics J*. 2015;15(4):363-371. doi:10.1038/tpj.2014.77.
20. Munakata W, Tobinai K. The discovery and the development of bendamustine for the treatment of non-Hodgkin lymphoma. *Expert Opin Drug Discov*. 2016;11(11):1123-1130. doi:10.1080/17460441.2016.1233174.
21. Wang L, Giannoudis A, Lane S, Williamson P, Pirmohamed M, Clark RE. Expression of the uptake drug transporter hOCT1 is an important clinical determinant of the response to imatinib in chronic myeloid leukemia. *Clinical Pharmacology & Therapeutics*. 2008;83(2):258-264. doi:10.1038/sj.clpt.6100268.
22. Kim DHD, Sriharsha L, Xu W, et al. Clinical relevance of a pharmacogenetic approach using multiple candidate genes to predict response and resistance to imatinib therapy in chronic myeloid leukemia. *Clin Cancer Res*. 2009;15(14):4750-4758. doi:10.1158/1078-0432.CCR-09-

References (continued)

0145.

23. Hu S, Franke RM, Filipski KK, et al. Interaction of imatinib with human organic ion carriers. *Clin Cancer Res*. 2008;14(10):3141-3148. doi:10.1158/1078-0432.CCR-07-4913.
24. Minematsu T, Giacomini KM. Interactions of tyrosine kinase inhibitors with organic cation transporters and multidrug and toxic compound extrusion proteins. *Mol Cancer Ther*. 2011;10(3):531-539. doi:10.1158/1535-7163.MCT-10-0731.
25. Abramson J, Smirnova I, Kasho V, Verner G, Kaback HR, Iwata S. Structure and mechanism of the lactose permease of Escherichia coli. *Science*. 2003;301(5633):610-615. doi:10.1126/science.1088196.
26. Radestock S, Forrest LR. The alternating-access mechanism of MFS transporters arises from inverted-topology repeats. *J Mol Biol*. 2011;407(5):698-715. doi:10.1016/j.jmb.2011.02.008.
27. Pedersen BP, Kumar H, Waight AB, Risenmay AJ. Crystal structure of a eukaryotic phosphate transporter. *Nature*. 2013. doi:doi:10.1038/nature12042.
28. Popp C, Gorboulev V, Müller TD, Gorbunov D, Shatskaya N, Koepsell H. Amino acids critical for substrate affinity of rat organic cation transporter 1 line the substrate binding region in a model derived from the tertiary structure of lactose permease. *Mol Pharmacol*. 2005;67(5):1600-1611. doi:10.1124/mol.104.008839.
29. Volk C, Gorboulev V, Kotzsch A, Müller TD, Koepsell H. Five Amino Acids in the Innermost Cavity of the Substrate Binding Cleft of Organic Cation Transporter 1 Interact with Extracellular and Intracellular Corticosterone. *Mol Pharmacol*. 2009;76(2):275-289. doi:10.1124/mol.109.054783.
30. Gorbunov D, Gorboulev V, Shatskaya N, et al. High-affinity cation binding to organic cation transporter 1 induces movement of helix 11 and blocks transport after mutations in a modeled interaction domain between two helices. *Mol Pharmacol*. 2008;73(1):50-61. doi:10.1124/mol.107.040170.
31. Koepsell H. Substrate recognition and translocation by polyspecific organic cation transporters. *Biol Chem*. 2011;392(1-2):95-101. doi:10.1515/BC.2011.009.
32. Egenberger B, Gorboulev V, Keller T, et al. A substrate binding hinge domain is critical for transport-related structural changes of organic cation transporter 1. *J Biol Chem*. 2012;287(37):31561-31573. doi:10.1074/jbc.M112.388793.
33. Keller T, Egenberger B, Gorboulev V, et al. The large extracellular loop of organic cation

References (continued)

transporter 1 influences substrate affinity and is pivotal for oligomerization. *J Biol Chem*. 2011;286(43):37874-37886. doi:10.1074/jbc.M111.289330.

34. Brast S, Grabner A, Sucic S, et al. The cysteines of the extracellular loop are crucial for trafficking of human organic cation transporter 2 to the plasma membrane and are involved in oligomerization. *The FASEB Journal*. 2012;26(3):976-986. doi:10.1096/fj.11-180679.
35. Forrest LR, Krämer R, Ziegler C. The structural basis of secondary active transport mechanisms. *Biochim Biophys Acta*. 2011;1807(2):167-188. doi:10.1016/j.bbabio.2010.10.014.
36. Ciarimboli G, Schlatter E. Regulation of organic cation transport. *Pflugers Arch*. 2004;449(5):423-441. doi:10.1007/s00424-004-1355-5.
37. Çetinkaya I, Ciarimboli G, Yalçinkaya G, et al. Regulation of human organic cation transporter hOCT2 by PKA, PI3K, and calmodulin-dependent kinases. *Am J Physiol Renal Physiol*. 2003;284(2):F293-F302. doi:10.1152/ajprenal.00251.2002.
38. Mehrens T, Lelleck S, Cetinkaya I, et al. The affinity of the organic cation transporter rOCT1 is increased by protein kinase C-dependent phosphorylation. *J Am Soc Nephrol*. 2000;11(7):1216-1224.
39. Singer SJ, Nicolson GL. The fluid mosaic model of the structure of cell membranes. *Science*. 1972;175(4023):720-731.
40. Grecco HE, Schmick M, Bastiaens PIH. Signaling from the living plasma membrane. *Cell*. 2011;144(6):897-909. doi:10.1016/j.cell.2011.01.029.
41. Kaback HR, Sahin-Tóth M, Weinglass AB. The kamikaze approach to membrane transport. *Nat Rev Mol Cell Biol*. 2001;2(8):610-620. doi:10.1038/35085077.
42. Borst P, Elferink RO. Mammalian ABC transporters in health and disease. *Annu Rev Biochem*. 2002;71(1):537-592. doi:10.1146/annurev.biochem.71.102301.093055.
43. He L, Vasiliou K, Nebert DW. Analysis and update of the human solute carrier (SLC) gene superfamily. *Hum Genomics*. 2009;3(2):195-206. doi:10.1186/1479-7364-3-2-195.
44. Nigam SK. What do drug transporters really do? *Nat Rev Drug Discov*. 2015;14(1):29-44. doi:10.1038/nrd4461.
45. Hediger MA, Clémençon B, Burrier RE, Bruford EA. The ABCs of membrane transporters in health and disease (SLC series): introduction. *Molecular Aspects of Medicine*. 2013;34(2-3):95-107. doi:10.1016/j.mam.2012.12.009.
46. International Transporter Consortium, Giacomini KM, Huang S-M, et al. Membrane

References (continued)

transporters in drug development. *Nat Rev Drug Discov*. 2010;9(3):215-236. doi:10.1038/nrd3028.

47. DeGorter MK, Xia CQ, Yang JJ, Kim RB. Drug transporters in drug efficacy and toxicity. *Annu Rev Pharmacol Toxicol*. 2012;52(1):249-273. doi:10.1146/annurev-pharmtox-010611-134529.
48. Borst P, Evers R, Kool M, Wijnholds J. A family of drug transporters: the multidrug resistance-associated proteins. *J Natl Cancer Inst*. 2000;92(16):1295-1302.
49. Sprowl JA, Sparreboom A. Uptake carriers and oncology drug safety. *Drug Metab Dispos*. 2014;42(4):611-622. doi:10.1124/dmd.113.055806.
50. Giacomini KM, Yee SW, Ratain MJ, Weinshilboum RM, Kamatani N, Nakamura Y. Pharmacogenomics and patient care: one size does not fit all. *Science Translational Medicine*. 2012;4(153):153ps18-153ps18. doi:10.1126/scitranslmed.3003471.
51. Gomez A, Ingelman-Sundberg M. Pharmacoepigenetics: its role in interindividual differences in drug response. *Clinical Pharmacology & Therapeutics*. 2009;85(4):426-430. doi:10.1038/clpt.2009.2.
52. Sparreboom A, Cox MC, Acharya MR, Figg WD. Herbal remedies in the United States: potential adverse interactions with anticancer agents. *J Clin Oncol*. 2004;22(12):2489-2503. doi:10.1200/JCO.2004.08.182.
53. Wright SH, Dantzler WH. Molecular and cellular physiology of renal organic cation and anion transport. *Physiological Reviews*. 2004;84(3):987-1049. doi:10.1152/physrev.00040.2003.
54. Tanaka Y, Hipolito CJ, Maturana AD, et al. Structural basis for the drug extrusion mechanism by a MATE multidrug transporter. *Nature*. 2013;496(7444):247-251. doi:10.1038/nature12014.
55. van de Steeg E, Wagenaar E, van der Kruijssen CMM, et al. Organic anion transporting polypeptide 1a/1b-knockout mice provide insights into hepatic handling of bilirubin, bile acids, and drugs. *J Clin Invest*. 2010;120(8):2942-2952. doi:10.1172/JCI42168.
56. Levitzki A. Tyrosine kinase inhibitors: views of selectivity, sensitivity, and clinical performance. *Annu Rev Pharmacol Toxicol*. 2013;53(1):161-185. doi:10.1146/annurev-pharmtox-011112-140341.
57. Sprowl JA, Ciarimboli G, Lancaster CS, et al. Oxaliplatin-induced neurotoxicity is dependent on the organic cation transporter OCT2. *Proc Natl Acad Sci USA*. 2013;110(27):11199-11204. doi:10.1073/pnas.1305321110.
58. Bacq A, Balasse L, Biala G, et al. Organic cation transporter 2 controls brain norepinephrine

References (continued)

and serotonin clearance and antidepressant response. *Mol Psychiatry*. 2012;17(9):926-939. doi:10.1038/mp.2011.87.

59. Ciarimboli G, Lancaster CS, Schlatter E, et al. Proximal tubular secretion of creatinine by organic cation transporter OCT2 in cancer patients. *Clin Cancer Res*. 2012;18(4):1101-1108. doi:10.1158/1078-0432.CCR-11-2503.
60. Gründemann D, Köster S, Kiefer N, et al. Transport of monoamine transmitters by the organic cation transporter type 2, OCT2. *J Biol Chem*. 1998;273(47):30915-30920.
61. Pabla N, Gibson AA, Buege M, et al. Mitigation of acute kidney injury by cell-cycle inhibitors that suppress both CDK4/6 and OCT2 functions. *Proc Natl Acad Sci USA*. 2015;112(16):5231-5236. doi:10.1073/pnas.1424313112.
62. Filipski KK, Mathijssen RH, Mikkelsen TS, Schinkel AH, Sparreboom A. Contribution of organic cation transporter 2 (OCT2) to cisplatin-induced nephrotoxicity. *Clinical Pharmacology & Therapeutics*. 2009;86(4):396-402. doi:10.1038/clpt.2009.139.
63. Ciarimboli G, Deuster D, Knief A, et al. Organic cation transporter 2 mediates cisplatin-induced oto- and nephrotoxicity and is a target for protective interventions. *The American Journal of Pathology*. 2010;176(3):1169-1180. doi:10.2353/ajpath.2010.090610.
64. Kantarjian H, Jabbour E, Grimley J, Kirkpatrick P. Dasatinib. *Nature reviews. Drug discovery*. September 2006:717-718.
65. Blom N, Gammeltoft S, Brunak S. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *J Mol Biol*. 1999;294(5):1351-1362. doi:10.1006/jmbi.1999.3310.
66. Hornbeck PV, Zhang B, Murray B, Kornhauser JM, Latham V, Skrzypek E. PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Research*. 2015;43(Database issue):D512-D520. doi:10.1093/nar/gku1267.
67. Hornbeck PV, Kornhauser JM, Tkachev S, et al. PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Research*. 2012;40(Database issue):D261-D270. doi:10.1093/nar/gkr1122.
68. Ulbrich MH, Isacoff EY. Subunit counting in membrane-bound proteins. *Nat Methods*. 2007;4(4):319-321. doi:10.1038/nmeth1024.
69. Jain A, Liu R, Ramani B, et al. Probing cellular protein complexes using single-molecule pull-down. *Nature*. 2011;473(7348):484-488. doi:10.1038/nature10016.

References (continued)

70. Li J, Rix U, Fang B, et al. A chemical and phosphoproteomic characterization of dasatinib action in lung cancer. *Nat Chem Biol*. 2010;6(4):291-299. doi:10.1038/nchembio.332.
71. Davis MI, Hunt JP, Herrgard S, et al. Comprehensive analysis of kinase inhibitor selectivity. *Nature Biotechnology*. 2011;29(11):1046-1051. doi:10.1038/nbt.1990.
72. Patel PR, Sun H, Li SQ, et al. Identification of potent Yes1 kinase inhibitors using a library screening approach. *Bioorg Med Chem Lett*. 2013;23(15):4398-4403. doi:10.1016/j.bmcl.2013.05.072.
73. Boggon TJ, Eck MJ. Structure and regulation of Src family kinases. *Oncogene*. 2004;23(48):7918-7927. doi:10.1038/sj.onc.1208081.
74. Kay BK, Williamson MP, Sudol M. The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *FASEB J*. 2000;14(2):231-241.
75. Sato I, Obata Y, Kasahara K, et al. Differential trafficking of Src, Lyn, Yes and Fyn is specified by the state of palmitoylation in the SH4 domain. *J Cell Sci*. 2009;122(Pt 7):965-975. doi:10.1242/jcs.034843.
76. Sakata T, Anzai N, Shin HJ, et al. Novel single nucleotide polymorphisms of organic cation transporter 1 (SLC22A1) affecting transport functions. *Biochem Biophys Res Commun*. 2004;313(3):789-793.
77. Koepsell H. Organic cation transporters in intestine, kidney, liver, and brain. *Annu Rev Physiol*. 1998;60(1):243-266. doi:10.1146/annurev.physiol.60.1.243.
78. Guckel D, Ciarimboli G, Pavenstädt H, Schlatter E. Regulation of organic cation transport in isolated mouse proximal tubules involves complex changes in protein trafficking and substrate affinity. *Cell Physiol Biochem*. 2012;30(1):269-281. doi:10.1159/000339063.
79. Hamar P, Song E, Kökény G, Chen A, Ouyang N, Lieberman J. Small interfering RNA targeting Fas protects mice against renal ischemia-reperfusion injury. *Proceedings of the National Academy of Sciences*. 2004;101(41):14883-14888. doi:10.1073/pnas.0406421101.
80. Diasio RB, Harris BE. Clinical pharmacology of 5-fluorouracil. *Clin Pharmacokinet*. 1989;16(4):215-237. doi:10.2165/00003088-198916040-00002.
81. Pachman DR, Qin R, Seisler DK, et al. Clinical Course of Oxaliplatin-Induced Neuropathy: Results From the Randomized Phase III Trial N08CB (Alliance). *J Clin Oncol*. 2015;33(30):3416-3422. doi:10.1200/JCO.2014.58.8533.
82. Hunter T. The genesis of tyrosine phosphorylation. *Cold Spring Harb Perspect Biol*.

References (continued)

2014;6(5):a020644-a020644. doi:10.1101/cshperspect.a020644.

83. Rikova K, Guo A, Zeng Q, et al. Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. *Cell*. 2007;131(6):1190-1203. doi:10.1016/j.cell.2007.11.025.
84. Rush J, Moritz A, Lee KA, et al. Immunoaffinity profiling of tyrosine phosphorylation in cancer cells. *Nature Biotechnology*. 2005;23(1):94-101. doi:10.1038/nbt1046.
85. Bordoli MR, Yum J, Breitkopf SB, et al. A secreted tyrosine kinase acts in the extracellular environment. *Cell*. 2014;158(5):1033-1044. doi:10.1016/j.cell.2014.06.048.
86. Hu S, Mathijssen RHJ, de Bruijn P, Baker SD, Sparreboom A. Inhibition of OATP1B1 by tyrosine kinase inhibitors: in vitro-in vivo correlations. *Br J Cancer*. 2014;110(4):894-898. doi:10.1038/bjc.2013.811.
87. Deng J, Shao J, Markowitz JS, An G. ABC transporters in multi-drug resistance and ADME-Tox of small molecule tyrosine kinase inhibitors. *Pharm Res*. 2014;31(9):2237-2255. doi:10.1007/s11095-014-1389-0.
88. Engen JR, Wales TE, Hochrein JM, et al. Structure and dynamic regulation of Src-family kinases. *Cell Mol Life Sci*. 2008;65(19):3058-3073. doi:10.1007/s00018-008-8122-2.
89. Stein PL, Vogel H, Soriano P. Combined deficiencies of Src, Fyn, and Yes tyrosine kinases in mutant mice. *Genes Dev*. 1994;8(17):1999-2007.
90. Zhao YH, Krueger JG, Sudol M. Expression of cellular-yes protein in mammalian tissues. *Oncogene*. 1990;5(11):1629-1635.
91. Luton F, Vergés M, Vaerman JP, Sudol M, Mostov KE. The SRC family protein tyrosine kinase p62yes controls polymeric IgA transcytosis in vivo. *Mol Cell*. 1999;4(4):627-632.
92. Chen Y-H, Lu Q, Goodenough DA, Jeansonne B. Nonreceptor tyrosine kinase c-Yes interacts with occludin during tight junction formation in canine kidney epithelial cells. *Mol Biol Cell*. 2002;13(4):1227-1237. doi:10.1091/mbc.01-08-0423.
93. Gamelin E, Gamelin L, Bossi L, Quasthoff S. Clinical aspects and molecular basis of oxaliplatin neurotoxicity: current management and development of preventive measures. *Semin Oncol*. 2002;29(5 Suppl 15):21-33. doi:10.1053/sonc.2002.35525.
94. Offer SM, Fossum CC, Wegner NJ, Stuflesser AJ, Butterfield GL, Diasio RB. Comparative functional analysis of DPYD variants of potential clinical relevance to dihydropyrimidine dehydrogenase activity. *Cancer Res*. 2014;74(9):2545-2554. doi:10.1158/0008-5472.CAN-13-2482.

References (continued)

95. Filipski KK, Loos WJ, Verweij J, Sparreboom A. Interaction of Cisplatin with the human organic cation transporter 2. *Clin Cancer Res*. 2008;14(12):3875-3880. doi:10.1158/1078-0432.CCR-07-4793.
96. Ong SS, Goktug AN, Elias A, Wu J, Saunders D, Chen T. Stability of the human pregnane X receptor is regulated by E3 ligase UBR5 and serine/threonine kinase DYRK2. *Biochem J*. 2014;459(1):193-203. doi:10.1042/BJ20130558.
97. Fogel H, Frere S, Segev O, et al. APP homodimers transduce an amyloid-β-mediated increase in release probability at excitatory synapses. *Cell Rep*. 2014;7(5):1560-1576. doi:10.1016/j.celrep.2014.04.024.
98. Pei J, Kim B-H, Grishin NV. PROMALS3D: a tool for multiple protein sequence and structure alignments. *Nucleic Acids Research*. 2008;36(7):2295-2300. doi:10.1093/nar/gkn072.
99. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*. 2004;32(5):1792-1797. doi:10.1093/nar/gkh340.
100. Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol*. 1993;234(3):779-815. doi:10.1006/jmbi.1993.1626.
101. Shen M-Y, Sali A. Statistical potential for assessment and prediction of protein structures. *Protein Sci*. 2006;15(11):2507-2524. doi:10.1110/ps.062416606.
102. Dong GQ, Fan H, Schneidman-Duhovny D, Webb B, Sali A. Optimized atomic statistical potentials: assessment of protein interfaces and loops. *Bioinformatics*. 2013;29(24):3158-3166. doi:10.1093/bioinformatics/btt560.
103. Sprowl JA, Lancaster CS, Pabla N, et al. Cisplatin-induced renal injury is independently mediated by OCT2 and p53. *Clin Cancer Res*. 2014;20(15):4026-4035. doi:10.1158/1078-0432.CCR-14-0319.
104. Furmanski BD, Hu S, Fujita K-I, et al. Contribution of ABCC4-mediated gastric transport to the absorption and efficacy of dasatinib. *Clin Cancer Res*. 2013;19(16):4359-4370. doi:10.1158/1078-0432.CCR-13-0980.
105. Capuano A, De Corato A, Lisi L, Tringali G, Navarra P, Russo Dello C. Proinflammatory-activated trigeminal satellite cells promote neuronal sensitization: relevance for migraine pathology. *Mol Pain*. 2009;5(1):43. doi:10.1186/1744-8069-5-43.
106. Boxberger KH, Hagenbuch B, Lampe JN. *Common Drugs Inhibit Human Organic Cation Transporter 1 (OCT1)-Mediated Neurotransmitter Uptake*. Drug Metabolism and Disposition;

References (continued)

2014. doi:10.1111/bph.12748/full.

107. Chen L, Shu Y, Liang X, et al. OCT1 is a high-capacity thiamine transporter that regulates hepatic steatosis and is a target of metformin. *Proc Natl Acad Sci USA*. 2014;111(27):9983-9988. doi:10.1073/pnas.1314939111.
108. Kell DB. Implications of endogenous roles of transporters for drug discovery: hitchhiking and metabolite-likeness. *Nat Rev Drug Discov*. 2016;15(2):143-143. doi:10.1038/nrd.2015.44.
109. Shu Y, Sheardown SA, Brown C, et al. Effect of genetic variation in the organic cation transporter 1 (OCT1) on metformin action. *J Clin Invest*. 2007;117(5):1422-1431. doi:10.1172/JCI30558.
110. Tzvetkov MV, Saadatmand AR, Lötsch J, Tegeder I, Stingl JC, Brockmöller J. Genetically polymorphic OCT1: another piece in the puzzle of the variable pharmacokinetics and pharmacodynamics of the opioidergic drug tramadol. *Clinical Pharmacology & Therapeutics*. 2011;90(1):143-150. doi:10.1038/clpt.2011.56.
111. Cho SK, Kim CO, Park ES, Chung JY. Verapamil decreases the glucose-lowering effect of metformin in healthy volunteers. *British Journal of Clinical Pharmacology*. 2014;78(6):1426-1432. doi:10.1111/bcp.12476.
112. European Medicines Agency. *Guideline on the Investigation of Drug Interactions*. Committee for Human Medicinal Products (CHMP); 2012.
113. Haenisch B, Drescher E, Thiemer L, et al. Interaction of antidepressant and antipsychotic drugs with the human organic cation transporters hOCT1, hOCT2 and hOCT3. *Naunyn-Schmiedeberg's Arch Pharmacol*. 2012;385(10):1017-1023. doi:10.1007/s00210-012-0781-8.
114. Ahlin G, Karlsson J, Pedersen JM, et al. Structural Requirements for Drug Inhibition of the Liver Specific Human Organic Cation Transport Protein 1. *J Med Chem*. 2008;51(19):5932-5942. doi:10.1021/jm8003152.
115. Belzer M, Morales M, Jagadish B, Mash EA, Wright SH. Substrate-Dependent Ligand Inhibition of the Human Organic Cation Transporter OCT2. *J Pharmacol Exp Ther*. 2013;346(2):300-310. doi:10.1124/jpet.113.203257.
116. Zhang X, Shirahatti NV, Mahadevan D, Wright SH. A conserved glutamate residue in transmembrane helix 10 influences substrate specificity of rabbit OCT2 (SLC22A2). *J Biol Chem*. 2005;280(41):34813-34822. doi:10.1074/jbc.M506342200.
117. Gorboulev V, Shatskaya N, Volk C, Koepsell H. Subtype-specific affinity for corticosterone of

References (continued)

rat organic cation transporters rOCT1 and rOCT2 depends on three amino acids within the substrate binding region. *Mol Pharmacol*. 2005;67(5):1612-1619. doi:10.1124/mol.104.008821.

118. Eramian D, Eswar N, Shen M-Y, Sali A. How well can the accuracy of comparative protein structure models be predicted? *Protein Science*. 2008;17(11):1881-1893. doi:10.1110/ps.036061.108.
119. Schlessinger A, Wittwer MB, Dahlin A, et al. High Selectivity of the γ-Aminobutyric Acid Transporter 2 (GAT-2, SLC6A13) Revealed by Structure-based Approach. *J Biol Chem*. 2012;287(45):37745-37756. doi:10.1074/jbc.M112.388157.
120. Schlessinger A, Geier E, Fan H, et al. Structure-based discovery of prescription drugs that interact with the norepinephrine transporter, NET. *Proc Natl Acad Sci USA*. 2011;108(38):15810-15815. doi:10.1073/pnas.1106030108.
121. Kido Y, Matsson P, Giacomini KM. Profiling of a Prescription Drug Library for Potential Renal Drug–Drug Interactions Mediated by the Organic Cation Transporter 2. *J Med Chem*. 2011;54(13):4548-4558. doi:10.1021/jm2001629.
122. Zhang J-H, Chung TDY, Oldenburg KR. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *J Biomol Screen*. 1999;4(2):67-73. doi:10.1177/108705719900400206.
123. Carlsson J, Coleman RG, Setola V, et al. Ligand discovery from a dopamine D3 receptor homology model and crystal structure. *Nat Chem Biol*. 2011;7(11):769-778. doi:10.1038/nchembio.662.
124. Minematsu T, Iwai M, Umehara K-I, Usui T, Kamimura H. Characterization of Human Organic Cation Transporter 1 (OCT1/SLC22A1)- and OCT2 (SLC22A2)-Mediated Transport of 1-(2-Methoxyethyl)-2-methyl-4,9-dioxo-3-(pyrazin-2-ylmethyl)- 4,9-dihydro-1H-naphtho[2,3-d]imidazolium Bromide (YM155 Monobromide), a Novel Small Molecule Survivin Suppressant. *Drug Metabolism and Disposition*. 2010;38(1):1-4. doi:10.1124/dmd.109.028142.
125. Bourdet DL, Pritchard JB, Thakker DR. Differential Substrate and Inhibitory Activities of Ranitidine and Famotidine toward Human Organic Cation Transporter 1 (hOCT1; SLC22A1), hOCT2 (SLC22A2), and hOCT3 (SLC22A3). *J Pharmacol Exp Ther*. 2005;315(3):1288-1297. doi:10.1124/jpet.105.091223.
126. Bachmakov I, Glaeser H, König J, Fromm MF. Einfluss von β-Blockern auf den Transport von Metformin durch den hepatischen Aufnahmetransporter für organische Kationen OCT1.

References (continued)

Diabetologie und Stoffwechsel. 2009;4(S 01):P_274. doi:10.1055/s-0029-1222078.

127. Barr JT, Jones JP. Inhibition of Human Liver Aldehyde Oxidase: Implications for Potential Drug-Drug Interactions. *Drug Metabolism and Disposition*. 2011;39(12):2381-2386. doi:10.1124/dmd.111.041806.
128. Gui C, Obaidat A, Chaguturu R, Hagenbuch B. Development of a cell-based high-throughput assay to screen for inhibitors of organic anion transporting polypeptides 1B1 and 1B3. *Curr Chem Genomics*. 2010;4(1):1-8. doi:10.2174/1875397301004010001.
129. De Bruyn T, van Westen GJP, Ijzerman AP, et al. Structure-based identification of OATP1B1/3 inhibitors. *Mol Pharmacol*. 2013;83(6):1257-1267. doi:10.1124/mol.112.084152.
130. Vladimir Svetnik, Andy Liaw, Christopher Tong, J Christopher Culberson, Robert P Sheridan A, Feuston BP. Random Forest: A Classification and Regression Tool for Compound Classification and QSAR Modeling. *J Chem Inf Comput Sci*. 2003;43(6):1947-1958. doi:10.1021/ci034160g.
131. Wittwer MB, Zur AA, Khuri N, et al. Discovery of Potent, Selective Multidrug and Toxin Extrusion Transporter 1 (MATE1, SLC47A1) Inhibitors Through Prescription Drug Profiling and Computational Modeling. *J Med Chem*. 2013;56(3):781-795. doi:10.1021/jm301302s.
132. Wishart DS, Jewison T, Guo AC, et al. HMDB 3.0—The Human Metabolome Database in 2013. *Nucleic Acids Research*. 2012;41(D1):gks1065–D807. doi:10.1093/nar/gks1065.
133. Kotake Y. Tetrahydroisoquinoline derivatives as possible Parkinson's disease-inducing substances. *Yakugaku Zasshi*. 2002;122(11):975-982. doi:10.1248/yakushi.122.975.
134. Kotake Y, Tasaki Y, Makino Y, Ohta S, Hirobe M. 1‐Benzyl‐1,2,3,4‐Tetrahydroisoquinoline as a Parkinsonism‐Inducing Agent: A Novel Endogenous Amine in Mouse Brain and Parkinsonian CSF. *Journal of Neurochemistry*. 1995;65(6):2633-2638. doi:10.1046/j.1471-4159.1995.65062633.x.
135. Zhang L, Schaner ME, Giacomini KM. Functional characterization of an organic cation transporter (hOCT1) in a transiently transfected human cell line (HeLa). *J Pharmacol Exp Ther*. 1998;286(1):354-361.
136. Kimura H, Takeda M, Narikawa S, Enomoto A, Ichida K, Endou H. Human organic anion transporters and human organic cation transporters mediate renal transport of prostaglandins. *J Pharmacol Exp Ther*. 2002;301(1):293-298.
137. Bednarczyk D, Ekins S, Wikel JH, Wright SH. Influence of Molecular Structure on Substrate

References (continued)

Binding to the Human Organic Cation Transporter, hOCT1. *Mol Pharmacol*. 2003;63(3):489-498. doi:10.1124/mol.63.3.489.

138. Moaddel R, Ravichandran S, Bighi F, Yamaguchi R, Wainer IW. Pharmacophore modelling of stereoselective binding to the human organic cation transporter (hOCT1). *British Journal of Pharmacology*. 2007;151(8):1305-1314. doi:10.1038/sj.bjp.0707341.
139. Harper JN, Wright SH. Multiple mechanisms of ligand interaction with the human organic cation transporter, OCT2. *Am J Physiol Renal Physiol*. 2013;304(1):F56-F67. doi:10.1152/ajprenal.00486.2012.
140. Ekins S, Polli JE, Swaan PW, Wright SH. Computational Modeling to Accelerate the Identification of Substrates and Inhibitors for Transporters That Affect Drug Disposition. *Clinical Pharmacology & Therapeutics*. 2012;92(5):661-665. doi:10.1038/clpt.2012.164.
141. Martínez-Guerrero LJ, Wright SH. Substrate-dependent inhibition of human MATE1 by cationic ionic liquids. *J Pharmacol Exp Ther*. 2013;346(3):495-503. doi:10.1124/jpet.113.204206.
142. Tweedie D, Polli JW, Berglund EG, et al. Transporter Studies in Drug Development: Experience to Date and Follow‐Up on Decision Trees From the International Transporter Consortium. *Clinical Pharmacology & Therapeutics*. 2013;94(1):113-125. doi:10.1038/clpt.2013.77.
143. Abe K, Taguchi K, Wasai T, et al. Biochemical and pathological study of endogenous 1-benzyl-1,2,3,4-tetrahydroisoquinoline-induced parkinsonism in the mouse. *Brain Res*. 2001;907(1-2):134-138.
144. Kotake Y, Yoshida M, Ogawa M, Tasaki Y, Hirobe M, Ohta S. Chronic administration of 1-benzyl-1,2,3,4-tetrahydroisoquinoline, an endogenous amine in the brain, induces parkinsonism in a primate. *Neuroscience Letters*. 1996;217(1):69-71. doi:10.1016/0304-3940(96)13065-2.
145. Brenke R, Kozakov D, Chuang G-Y, et al. Fragment-based identification of druggable “hot spots” of proteins using Fourier domain correlation techniques. *Bioinformatics*. 2009;25(5):621-627. doi:10.1093/bioinformatics/btp036.
146. Fan H, Irwin JJ, Webb BM, Klebe G, Shoichet BK, Sali A. Molecular docking screens using comparative models of proteins. *J Chem Inf Model*. 2009;49(11):2512-2527. doi:10.1021/ci9003706.

References (continued)

147. Fan H, Irwin JJ, Sali A. Virtual ligand screening against comparative protein structure models. *Methods Mol Biol*. 2012;819:105-126. doi:10.1007/978-1-61779-465-0_8.
148. Mysinger MM, Carchia M, Irwin JJ, Shoichet BK. Directory of Useful Decoys, Enhanced (DUD-E): Better Ligands and Decoys for Better Benchmarking. *J Med Chem*. 2012;55(14):6582-6594. doi:10.1021/jm300687e.
149. Baell JB, Holloway GA. 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. doi:10.1021/jm901137j.
150. Yap CW. PaDEL-descriptor: An open source software to calculate molecular descriptors and fingerprints. *Journal of Computational Chemistry*. 2011;32(7):1466-1474. doi:10.1002/jcc.21707.
151. Cheng T, Wang Y, Bryant SH. FSelector: a Ruby gem for feature selection. *Bioinformatics*. 2012;28(21):2851-2852. doi:10.1093/bioinformatics/bts528.
152. Irwin JJ, Sterling T, Mysinger MM, Bolstad ES, Coleman RG. ZINC: A Free Tool to Discover Chemistry for Biology. *J Chem Inf Model*. 2012;52(7):1757-1768. doi:10.1021/ci3001277.
153. Team RC. R: A language and environment for statistical computing. 2013.

Chapter 3: Structural Characterization of the Human Concentrative Nucleoside Transporter 3

As the name implies, human concentrative nucleoside transporters (CNTs, SLC28 family) have evolved to transport hydrophilic nucleosides across the cell membrane. First cloned in the mid-to-late 1990s,[1-4] they are characterized by a high affinity transport of nucleosides coupled to the gradient of sodium ions. Unlike the SLC29 family of nucleoside transporters, the equilibrative nucleoside transporters (ENTs), CNTs have been shown to only transport nucleosides (containing both ribose and deoxyribose) and their analogues. ENTs are characterized by a low affinity nucleoside transport and are additionally capable of transporting nucleobases[5] and monoamines.[6] Among the human CNTs, hCNT1 exhibits specificity towards pyrimidines, hCNT2 for purines, while hCNT3 transports both purines and pyrimidines (**Figure 3.1**).[7-9]

Nucleosides are metabolic precursors in nucleic acid synthesis with a wide variety of critical roles in cell homeostasis.[10,11] While the endogenous role of human CNTs is primarily the mediation of nucleoside salvage, from the standpoint of pharmacology, human CNTs are significant due to the importance of their drug substrates, the nucleoside analogs.[12] Nucleoside analogs are molecules built upon the templates of naturally occurring nucleosides. Being hydrophilic, nucleoside analogs require transport across the cell membrane by membrane transporters, including human CNTs. Once in the cell, they leverage their similarity to nucleosides and either become incorporated into newly synthesized DNA/RNA, which results in chain termination, or directly inhibit processing enzymes such as ribonucleotide reductase, both of which lead to apoptosis.[13] Nucleoside analogues are frequently used as first-line treatments of cancers, viral infections, and autoimmune diseases (**Figure 3.2**).[12] Two examples include gemcitabine, a drug frequently used for leukemia and solid tumors, and zidovudine, the first antiretroviral used in the treatment for HIV.[12]

Although the knowledge about human CNTs is growing, the current research is limited by the lack of knowing their structures. Human CNTs are predicted to contain 11 transmembrane helices, with the N-terminus facing the cytoplasm, and a glycosylated extracellular C-terminus.[14] Knowledge of the mechanism of transport for these transporters is necessary to gain insight into the nucleoside metabolism and signaling pathways, as well as to facilitate drug design and nucleoside drug delivery. In light of the recent crystal structure of a hCNT3 homolog from *Vibrio cholerae*, vcCNT,[14] it is now possible to model the human SLC28 family members, and use the resulting structural information to form hypotheses and guide functional studies. The X-ray crystal structure showed that the prokaryotic homolog vcCNT forms a homotrimer. This surprising finding could lead to a significant advancement because multimerization of membrane proteins often provides a mechanism for the regulation and stability of membrane structures.[15]

This chapter is focused on the structure and function of the human CNT3. The first section explores the quaternary structure of hCNT3. It establishes that human concentrative nucleoside transporter 3 forms homo-trimers and predicts differences between the structures of hCNT3 and vcCNT. Sequence co-evolution and conservation analysis of the trimerization interface indicates that oligomerization is critical for the stability and function of the transporter. The second section summarizes a multi-year effort to express and purify hCNT3 for the purpose of structure determination. Although ultimately unsuccessful in obtaining a good sample for X-ray crystallography or electron cryo-microscopy (cryo-EM), it should serve as a starting point for

future efforts. The last section presents the results of a virtual screen with a focus on finding substrates and modulators of hCNT3 function. A novel approach for finding CNT ligands was used, significantly expanding on the previously explored chemical space.

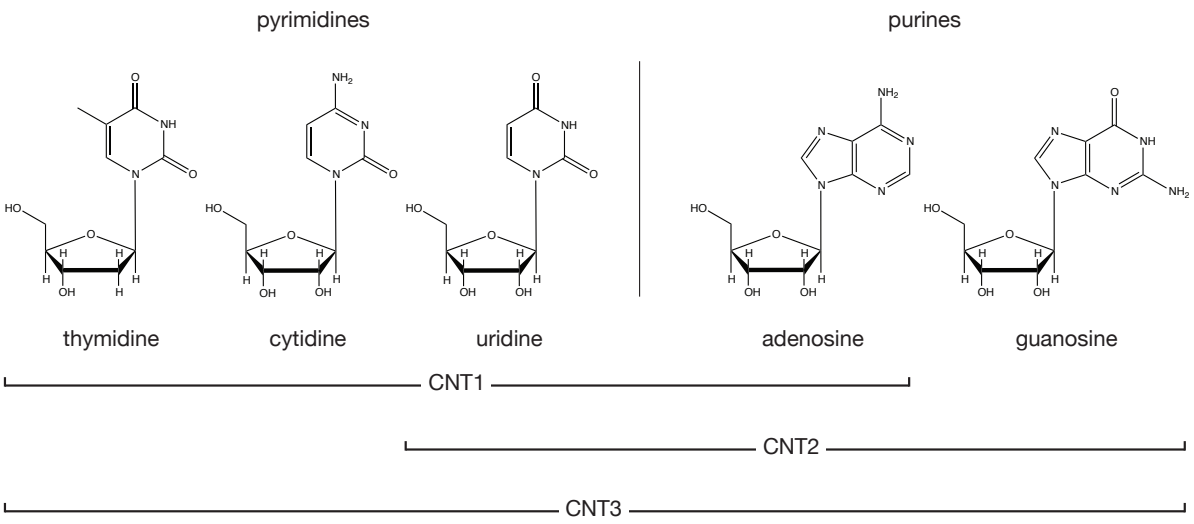


Figure 3.1 Schematic of the specificity of the human concentrative nucleoside transporters.

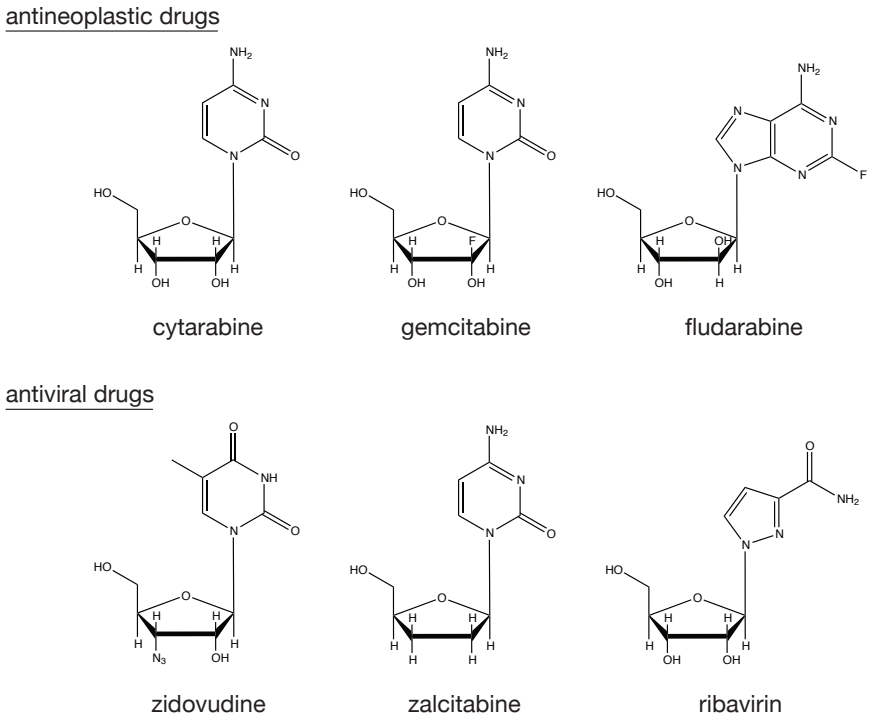


Figure 3.2 Molecular structures of nucleoside analogues commonly used in the clinic.

Cysteine cross-linking shows that human concentrative nucleoside transporter 3 forms a cyclic homo-trimer

The contents of this section have been submitted for publication in the following article:

Cysteine cross-linking shows that human concentrative nucleoside transporter 3 forms a cyclic homo-trimer

**Adrian Stecula, Avner Schlessinger, Andrej Sali, and Kathleen M. Giacomini*

*The author's contribution included experimental design, comparative structure modeling, execution of experiments, and writing of the manuscript.

Introduction

Nucleosides are metabolic precursors in nucleic acid synthesis with a wide variety of critical roles in cell homeostasis.[10,11] For example, when converted to nucleotides (e.g., ATP and GTP), they represent units of cellular currency, while as components of NAD and FAD, they act as coenzymes critical in energy metabolism. Some nucleosides can act alone as signaling molecules; for example, adenosine can act as an autocrine and paracrine hormone by serving as a ligand for P1 receptors.[16] Although most cell types are capable of *de novo* synthesis of nucleosides, others, like cells of the brain, muscle, erythrocytes, leukocytes, and bone marrow cells, lack these biosynthetic pathways.[17] In addition, *de novo* synthesis is energetically costly.[11] Nucleoside salvage (i.e., the recycling of extracellular nucleosides to replenish the intracellular levels) via nucleoside transporters is therefore an attractive strategy to achieve nucleoside homeostasis. Being largely hydrophilic, nucleosides require membrane transporters to cross the plasma membrane. To date, six human nucleoside transporters have been identified from two SLC transporter families, the concentrative nucleoside transporters (CNT) (SLC28 family) and the equilibrative nucleoside transporters (ENT) (SLC29 family).[18]

Of the three human members of the SLC28 family, hCNT3 is characterized by the broadest tissue distribution and substrate specificity. High levels of hCNT3 have been found in the pancreas, bone marrow, and mammary gland; lower levels have been found in the intestine, lung, prostrate, testis, and liver.[4] hCNT3 is a symporter that couples the transport of one nucleoside to the symport of two Na^+ ions or one proton.[17] Its pharmacological importance stems from its ability to transport a wide

variety of nucleoside-derived drugs,[12] including first-line treatments for viral infections, such as valacyclovir, or solid tumors, such as gemcitabine. As such, it is an important mediator of drug response and resistance to anti-cancer nucleoside analogs. Clinical examples include chronic lymphocytic leukemia (CLL) patients with elevated hCNT3 expression, who experienced a lower complete response rate to fludarabine therapy.[19] In addition, in pancreatic cancer, some uncharacterized hCNT3 polymorphisms, such as the nonsynonymous A25G mutation, have been shown to associate with gemcitabine toxicity.[20]

Although the regulation and function of hCNT3 have been subject to extensive characterization through mutagenesis and uptake studies over the past two decades, there have been only a few insights into the structure of these transporters.[2,21-25] The crystal structure of vcCNT provided the first breakthrough, by establishing a previously unknown protein-fold.[14] The most striking feature, however, proved to be the unexpected quaternary structure. No previous study has hypothesized CNTs to form oligomeric structures. Subsequently, it was established that the functional form of vcCNT is indeed a trimer, with clearly defined ligand and ion binding sites on each of the protomers.[14]

With the increasing number of crystal structures of membrane transporters, their oligomerization and symmetry have recently become topics of intense study due to a variety of possible functional roles.[15,26] For example, as shown by the studies on the BetP trimer, it is possible for stable BetP monomers to exist in the membrane on their own, while still retaining the ability to accumulate betaine.[27,28] However, the formation of the trimer is necessary for hyperosmotic stress stimulus sensing through the osmosensing C-terminal domain, which increases the transport activity based on the cytoplasmic K^+ concentration. The coupling between adjacent protomers therefore affects both the catalytic and regulatory aspects of the transport cycle. Each protomer in the trimer has been shown to adopt a distinct conformation along the cycle, further supporting inter-protomer coupling.[28,29]

As a trimer, BetP does not represent an outlier. Oligomerization is remarkably common among membrane proteins. A survey of the Protein Data Bank of Transmembrane Proteins shows that approximately 65% of membrane proteins form oligomers.[26] Amongst more than 50 SLC transporter families, several, including SLC1, SLC13, and SLC42, are predicted or have been shown to have members that form oligomeric structures.[18]

Here, we report a series of cross-linking studies to examine two main questions. 1) Does hCNT3 also form homo-trimers? 2) Is there a functional role for the formation of this quaternary structure? By using a homology model of hCNT3 based on the vcCNT structure (39% sequence identity) followed by cysteine mutagenesis, we determined that the trimer is preserved in hCNT3, using *in vitro* and cell-based assays. The cross-linking results can be parsimoniously rationalized through localized structural differences between hCNT3 and vcCNT in the trimerization domain. The model of the hCNT3 trimer also indicates how the translocation distance is significantly reduced in the trimer compared to that in the monomeric membrane transporters.

Experimental procedures

Expression and purification of hCNT3 in High Five insect cells hCNT3 was expressed in High Five insect cells using the Bac-to-Bac Baculovirus Expression System (Thermo Scientific). The gene encoding the full-length hCNT3 was cloned into the pFastBac1 vector, including a HRV 3C cleavable His_{10}-tag at the N-terminus (**Figure 3.3A**). The protein was purified as described previously.[30] Cells were spun down and solubilized in the lysis buffer (50 mM Tris pH 8.0, 300 mM NaCl, cOmplete Mini Protease Inhibitor Cocktail (Roche 11836153001)) before passing the solution through 5 cycles inside of Emulsiflex at 5,000-10,000 psi. The lysate was cleared with a 15 min spin at 15,000 x g at 4°C before pelleting the membranes with an 80 minute ultracentrifuge spin at 40,000 x g at 4°C. Membranes were resuspended in the resuspension buffer (50mM HEPES pH 7, 300 mM NaCl, 10% glycerol, 2 mM imidazole pH 7, 1% n-dodecyl-β-D-maltopyranoside (DDM) (Inalco 1758-1350), 1mg/mL iodoacetamide, cOmplete Mini Protease Inhibitor Cocktail (Roche 11836153001), 0.05% cholesteryl hemisuccinate, 0.1 mg/mL *E. coli* polar lipid extract) and stirred for 1 hour at 4°C. Excess lipids were pelleted with a 30 min spin at 35,000 rpm at 4°C. Supernatants were applied to TALON/Co^{2+} affinity resin and stirred for 2 hours. Protein was eluted off the column with 300 mM imidazole, following a series of washes. The protein was concentrated with 100kD cutoff Amicon concentrators (Sigma Z648043) and applied to a Superdex 200 size-exclusion column along with the running buffer (50 mM HEPES pH 7, 300 mM NaCl, 10% glycerol, 0.1% DDM). Fractions were collected every 500 μL to be further analysis. For Coomassie gel stains, fractions were loaded onto 4-15% Criterion Tris-HCl gels (Bio-Rad 3450027) under reducing and denaturing conditions and stained using SimplyBlue SafeStain (Invitrogen LC6060) following manufacturers' protocols. Protein size was estimated with Amersham ECL Full-Range Rainbow pre-stained protein ladder (GE RPN800E).

Glutaraldehyde cross-linking of purified CNT3 Glutaraldehyde cross-linking was performed following an adapted protocol from Perez, et al.[27] A volume of 100 µL of purified protein in running buffer was supplemented with 5 µL of 2.3% glutaraldehyde and incubated at 37°C for 5 min. The reaction was terminated by the addition of 10 µL of 1 M Tris-HCl (pH 8.0). Samples were loaded onto 4-15% Criterion Tris-HCl gels (Bio-Rad 3450027). Protein size was estimated with Precision Plus Protein™ All Blue pre-stained protein ladder (1610373). Polyclonal antibody Sigma HPA023311 against CNT3 was used for detection, with SuperSignal West Femto Chemiluminescent substrate (Thermo Scientific 34095) on the FluorChem E development system (proteinsimple).

Homology modeling of human CNTs Comparative protein structure models of wild type and mutant hCNT3 were created with MODELLER 9.10 (http://salilab.org/modeller),[31] using the 2.4 Å co-crystal structure of a concentrative nucleoside transporter from *Vibrio cholerae,* vcCNT, in complex with uridine (PDB ID 3TIJ)[14] as a template. Sequence alignment was created by a manual refinement of gaps in the alignment from PROMALS3D (**Figure S3.2**).[32] We generated 100 models using the automodel class with default settings. The models had acceptable normalized discrete optimized protein energy scores (zDOPE) in the range of -0.41 to -0.63.[33] The top scoring models were selected for analysis.

Cell culture of PK15NTD cells Porcine kidney tubular epithelium nucleoside transporter deficient cells (PK15NTD)[34] were donated by Dr. Chung-Ming Tse (The Johns Hopkins University School of Medicine, Baltimore, MD, USA). The cells were cultured in Eagle's minimal essential medium with Earle's balanced salt solution with 1 mM sodium pyruvate, 0.1 mM non-essential amino acids, 10% FBS, 100 I.U./mL penicillin, 100 µg/mL streptomycin, and 200 µg/mL hygromycin B at 37°C and 5% CO_2, as reported previously.[35,36]

Cloning and site-directed mutagenesis of human CNTs Full-length human CNT1, CNT2, and CNT3 cDNAs were cloned individually into the pcDNA5/FRT mammalian expression vector (Invitrogen V601020). SLC28A3 (UniProt ID Q9HAS3) cDNA were obtained from GE Dharmacon MGC cDNAs collection (MHS6278-202857241). Locations of cysteine mutants were selected based on the comparative structure models. Standard protocols for QuikChange II site-directed mutagenesis were followed (Agilent 200523), using KOD Xtreme Hot Start DNA polymerase (Novagen 71975-3) instead of *PfuUltra* High-Fidelity DNA polymerase and MAX Efficiency DH5α competent cells (Invitrogen 18258-012) instead of XL1-Blue supercompetent cells.

Cross-linking of human CNTs containing cysteine mutants in PK15NTD cells The protocol was adapted from Hastrup et al.[37] PK15NTD cells were seeded on non-treated polystyrene 60 mm dishes (Corning). Upon reaching 80-90% confluency, they were transfected with the appropriate construct and Lipofectamine LTX per the manufacturer's protocol (Invitrogen 15338-030). For co-expressing studies, half of the DNA material added came from each of the single mutant constructs.

Following a 48-hour incubation, the cells were washed with PBS (1.54 mM KH_2PO_4, 155.17 mM NaCl, 2.71 mM Na_2HPO_4, pH 7.4) and subsequently incubated in 100 μM $CuSO_4$ and 400 μM 1,10-phenanthroline for 10 minutes at room temperature. The cells were then washed twice with PBS and incubated in 10 mM *N*-ethylmaleimide (NEM) for 20 minutes at room temperature, before being scraped into PBS/protease inhibitor (PI) buffer (PBS, 10 mM NEM, cOmplete Mini Protease Inhibitor Cocktail (Roche 11836153001)). The suspension was pelleted at 800 x g for 5 min at 4°C. The pellet was resuspended in 0.2% digitonin in PBS/PI buffer for 20 min at 4°C. The suspension was pelleted at 2000 x g for 10 min. The pellet was resuspended in 100 μL of 1% Triton X-100 in PBS/PI at 4°C for 1 hour. The suspension was centrifuged at 14,000 x g for 30 min at 4°C.

The extract was then assayed for protein concentration with the BCA Protein Assay Kit (Pierce 23225) and 2 μg of protein was deglycosylated overnight with 1.5 μL of PNGase F (NEB P0704L) in a total volume of 20 μL, following the nonreducing manufacturer's protocol. Western blots were performed under denaturing but nonreducing conditions by mixing 15 μL of the sample with 4X NuPAGE LDS Sample Buffer (Invitrogen NP0008) and loading it onto NuPAGE 3-8% Tris-Acetate mini gel (Invitrogen EA0375). Protein size was estimated with HiMark pre-stained protein ladder (Invitrogen LC5699). Polyclonal antibody Sigma HPA023311 antibody against CNT3 was used for detection, with SuperSignal West Femto Chemiluminescent substrate (Thermo Scientific 34095) on the FluorChem E development system (proteinsimple).

[3H]uridine uptake Uptake studies were performed as described previously.[38] Briefly, PK15NTD cells were seeded on non-coated polystyrene 24-well plates (Corning). Upon reaching 70-80% confluency, they were transfected with the appropriate gene containing construct or an empty vector, and Lipofectamine LTX per the manufacturer's protocol (Invitrogen 15338-030). Following a 48-hour incubation, the cells were incubated for 10 minutes in a 37°C sodium-free buffer (5 mM HEPES, 10 mM glucose, 1 mM $CaCl_2$, 140 mM N-methyl-D-glucamine, 5 mM KH_2PO_4, 1 mM $MgCl_2$, pH 7.4). The uptake was initiated by the addition of 33.3 nM [5,6-^{3}H]-uridine (Moravek MT 799) in

a sodium-containing buffer (5 mM HEPES, 10 mM glucose, 1 mM $CaCl_2$, 5 mM KCl, 135 mM NaCl, 1 mM $MgCl_2$, 0.8 mM Na_2HPO_4, 3.3 mM NaH_2PO_4, pH 7.4) and terminated by washing the cells twice with 4°C sodium-free buffer. Cells were lysed by the addition of lysis buffer (0.1% SDS vol/vol, 0.1 N NaOH). Intracellular radioactivity was measured by scintillation counting and normalized per well of protein content using the BCA Protein Assay Kit (Pierce 23225).

Results

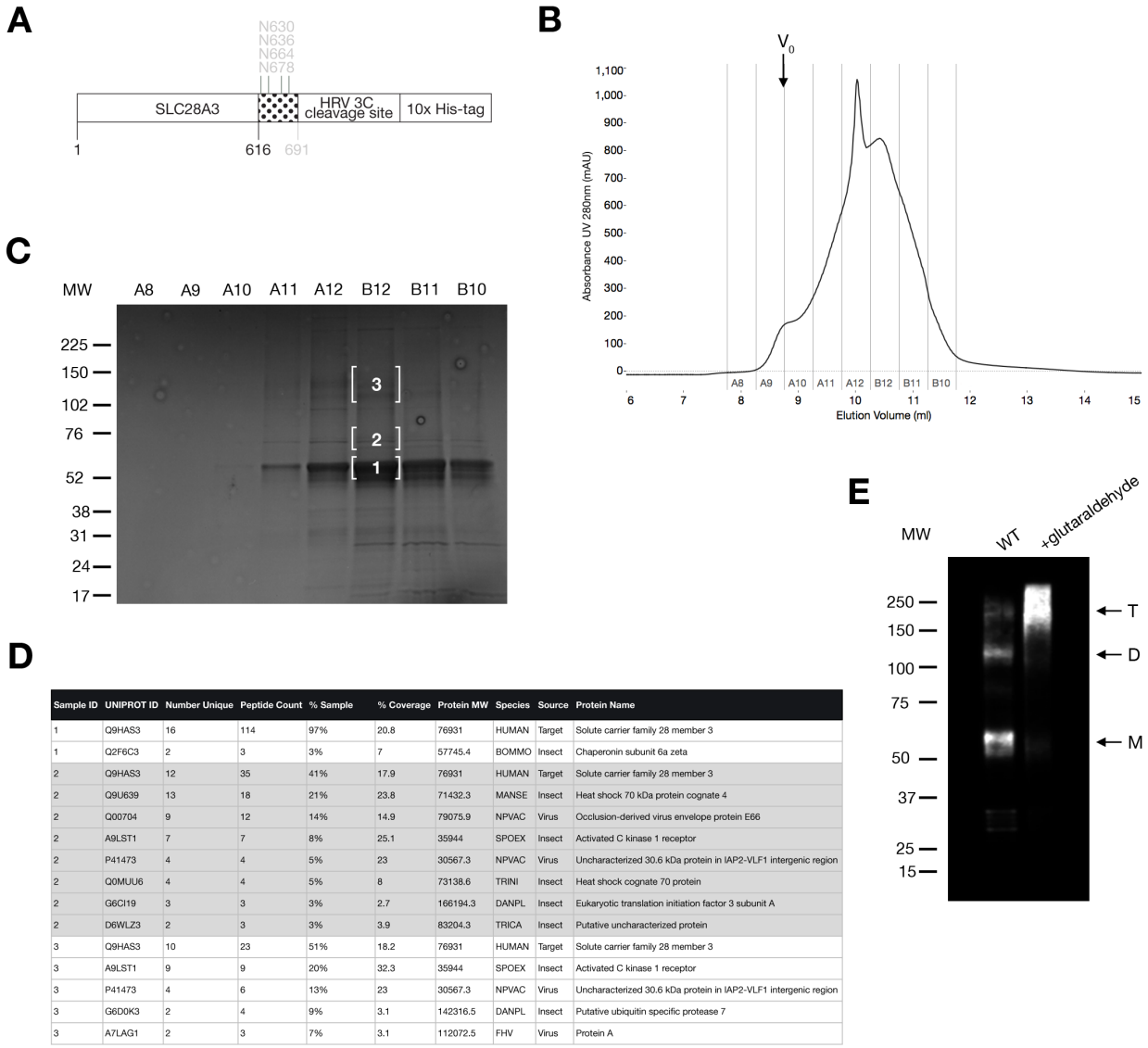
Purification of wild type hCNT3 To determine whether or not hCNT3 forms a homo-oligomer, we first expressed and purified the wild type protein. The created construct was a truncated variant of the full-length hCNT3, missing the 75 C-terminal residues predicted to be extracellular and to lack both secondary structure and glycosylation sites (**Figure 3.3A**). Earlier work performed in our laboratory has indicated that glycosylation most likely does not affect trafficking of CNTs.[39]

The size-exclusion chromatography profile of the sample has shown that using a mild, nonionic detergent like DDM preserves CNT3 as an oligomer. The elution peak corresponded to the micelle size of approximately 300 kDa, as expected for a trimer/detergent micelle complex (**Figure 3.3B**). Further analysis by mass spectrometry validated the purity of the fractions as mostly containing hCNT3 (**Figure 3.3C,D**). A Western blot of the sample indicated three potential species whose molecular weight was in agreement with the size of a CNT3 monomer, dimer, and trimer (**Figure 3.3E**). Nonspecific cross-linking of the detergent-solubilized sample using glutaraldehyde stabilized the highest molecular weight species (without cross-linking, any potential complex could dissociate during electrophoresis). This result is the first indication that hCNT3 forms a homo-trimer.

Homology modeling and selection of sites for cysteine mutagenesis To further examine the oligomerization of hCNT3, we turned to cysteine cross-linking. By introducing cysteines at key residue positions, we aimed to stabilize the complexes suggested in **Figure 3.3E** for further analysis. Cysteine cross-linking has been successfully used in previous studies of transporter oligomerization.[14,37,40,41] First, we created a homology model of the hCNT3 spanning the predicted TM4-TM11 helices, using the structure of vcCNT as a template. Cysteine mutagenesis sites were then chosen based on the proximity of interface residues and consideration of disulfide geometry.[42] Three separate sites along the trimerization interface, representing 4 separate pairs of cysteine

mutants, were selected to determine whether or not the hCNT3 forms a trimer similar to that of vcCNT (**Figure 3.4B**).

The first selected site is in the extracellular "crown" region of hCNT3, at the junction of the predicted IH1, EH, and TM6 helices. Both V274 and A285 are flanked in sequence by phenylalanine residues, which participate in the inter- and intra-protomer π-π stacking interactions in the vcCNT crystal structure (**Figure 3.4C**). The second site is also extracellular, positioned at the predicted interface of IH1 and the C-terminal end of TM9 (**Figure 3.4D**). The third site residing in the predicted transmembrane region was probed with two independent pairs of cysteines (**Figure 3.4E**). The first pair, A440 and A452, is equivalent to the pair shown to cross-link in vcCNT (A253 and A269, respectively).[14] Due to the symmetry of the trimer, it represents a double mutation on the predicted TM9, where the N-terminal A440 of one protomer intersects the centrally positioned A456 of another protomer. The second pair probes the junction of A456 with T297 located on TM6.



D

Sample ID	UNIPROT ID	Number Unique	Peptide Count	% Sample	% Coverage	Protein MW	Species	Source	Protein Name
1	Q9HAS3	16	114	97%	20.8	76931	HUMAN	Target	Solute carrier family 28 member 3
1	Q2F6C3	2	3	3%	7	57745.4	BOMMO	Insect	Chaperonin subunit 6a zeta
2	Q9HAS3	12	35	41%	17.9	76931	HUMAN	Target	Solute carrier family 28 member 3
2	Q9U639	13	18	21%	23.8	71432.3	MANSE	Insect	Heat shock 70 kDa protein cognate 4
2	Q00704	9	12	14%	14.9	79075.9	NPVAC	Virus	Occlusion-derived virus envelope protein E66
2	A9LST1	7	7	8%	25.1	35944	SPOEX	Insect	Activated C kinase 1 receptor
2	P41473	4	4	5%	23	30567.3	NPVAC	Virus	Uncharacterized 30.6 kDa protein in IAP2-VLF1 intergenic region
2	Q0MUU6	4	4	5%	8	73138.6	TRINI	Insect	Heat shock cognate 70 protein
2	G6CI19	3	3	3%	2.7	166194.3	DANPL	Insect	Eukaryotic translation initiation factor 3 subunit A
2	D6WLZ3	2	3	3%	3.9	83204.3	TRICA	Insect	Putative uncharacterized protein
3	Q9HAS3	10	23	51%	18.2	76931	HUMAN	Target	Solute carrier family 28 member 3
3	A9LST1	9	9	20%	32.3	35944	SPOEX	Insect	Activated C kinase 1 receptor
3	P41473	4	6	13%	23	30567.3	NPVAC	Virus	Uncharacterized 30.6 kDa protein in IAP2-VLF1 intergenic region
3	G6D0K3	2	4	9%	3.1	142316.5	DANPL	Insect	Putative ubiquitin specific protease 7
3	A7LAG1	2	3	7%	3.1	112072.5	FHV	Virus	Protein A

Figure 3.3 Expression and purification of hCNT3 from High Five insect cells. A) Schematic of the expression construct. B) SEC elution profile where the void peak is indicated with V_0 and the collected fractions are represented by grey vertical lines. C) Coomassie stained gel of the selected fractions, indicating the areas that were cut out for mass spectrometry analysis. D) Table containing the identity of peptides identified in each of the samples by mass spectrometry. E) Western blot of the WT sample before and after glutaraldehyde cross-linking.

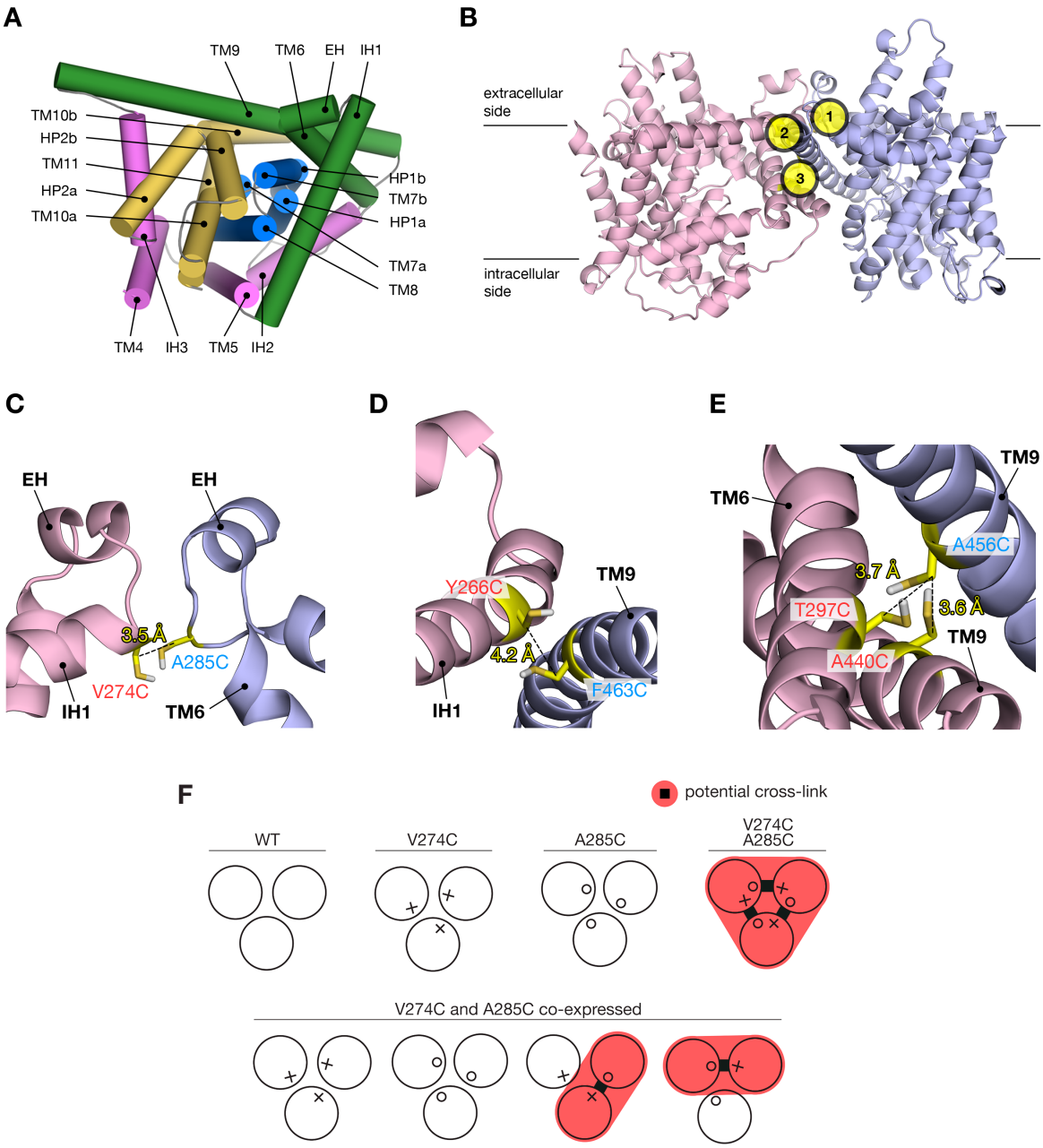


Figure 3.4 Comparative structure model of hCNT3 and the location of the mutated sites. A) A cartoon representation of the comparative structure model of hCNT3 based on the structure of vcCNT (PDB ID 3TIJ)(ref. 14). The trimerization domain is shown in green, scaffold domain in pink, and the inverted repeats performing the transport catalysis function in blue and yellow. B) Cartoon representation of the putative hCNT3 homo-trimer shown in the plane of the membrane. Two separate protomers are shown in blue and pink. Three mutagenesis sites are numbered and shown in yellow. C) Cartoon representation of the first mutagenesis site, where the side chains of V274C and A285C are shown as sticks. D) Cartoon representation of the second mutagenesis site, where the side chains of Y266C and F463C are shown as sticks. E) Cartoon representation of the third mutagenesis site, where the side chains of two separate pairs of mutants, T297C/A456C and A440C/A456C are shown as sticks. F) Schematic of the experimental design. Mutations are shown as X and O. Sites of potential cross-linking are indicated with a red highlight and a thick black line.

All residues selected for mutagenesis are conserved between the three human members of the SLC28 family, except for T297, which is a cysteine in both hCNT1 and hCNT2. Endogenous cysteines of hCNT3 are present only on the predicted helices TM1, IH3, TM10a, and TM11, and thus are located away from the trimerization regions and expected not to interfere with the cross-linking. A series of constructs was created, containing either single or double cysteine mutants. As shown in **Figure 3.4F** using the 274/285 pair as an example, due to the homo-trimer symmetry, it is possible to cross-link the vcCNT-like trimer only by expressing the hCNT3 construct containing a double mutation or co-expressing two single mutant constructs in a single cell.

Validation of hCNT3 homo-trimer formation with cysteine cross-linking To test whether or not the mutants fold properly and still traffic to the cell membrane, we transiently expressed the mutant constructs in the PK15NTD nucleoside transporter deficient porcine kidney cell line. This cell line was specifically chosen to ensure the lack of background concentrative nucleoside transporter expression. All mutants retained uridine transport function (**Figure 3.5A**), and are thus suitable for cell-based cross-linking tests.

To preserve the potential hCNT3 homo-trimer, we performed the cross-linking assay in PK15NTD cells transiently transfected with the mutant constructs. As hypothesized, no hCNT3 complexes survive the solubilization with Triton X-100 and electrophoresis, as demonstrated by a single band corresponding to the monomeric molecular weight of the wild type hCNT3. For the 274/285 site, when co-expressing both of the single mutants, we see a band corresponding to the molecular weight of the dimer, in line with a prediction that two protomers in a trimer are able to cross-link (**Figure 3.5B**). The sharp band that appears for the double mutant at the molecular weight of three monomers thus indicates that hCNT3 forms homo-trimers in cells.

Equivalent cross-linking studies for the second site, 266/463, are in agreement with the first site (**Figure 3.5C**). Co-expressing the single mutant constructs led to a sharp band corresponding to a dimeric species. Interestingly, there were two bands near the molecular weight of a homo-trimer in the double mutant. It is conceivable that there are two populations of homo-trimers, one with three cross-links and the other with two cross-links, each with a slightly different electrophoretic mobility.

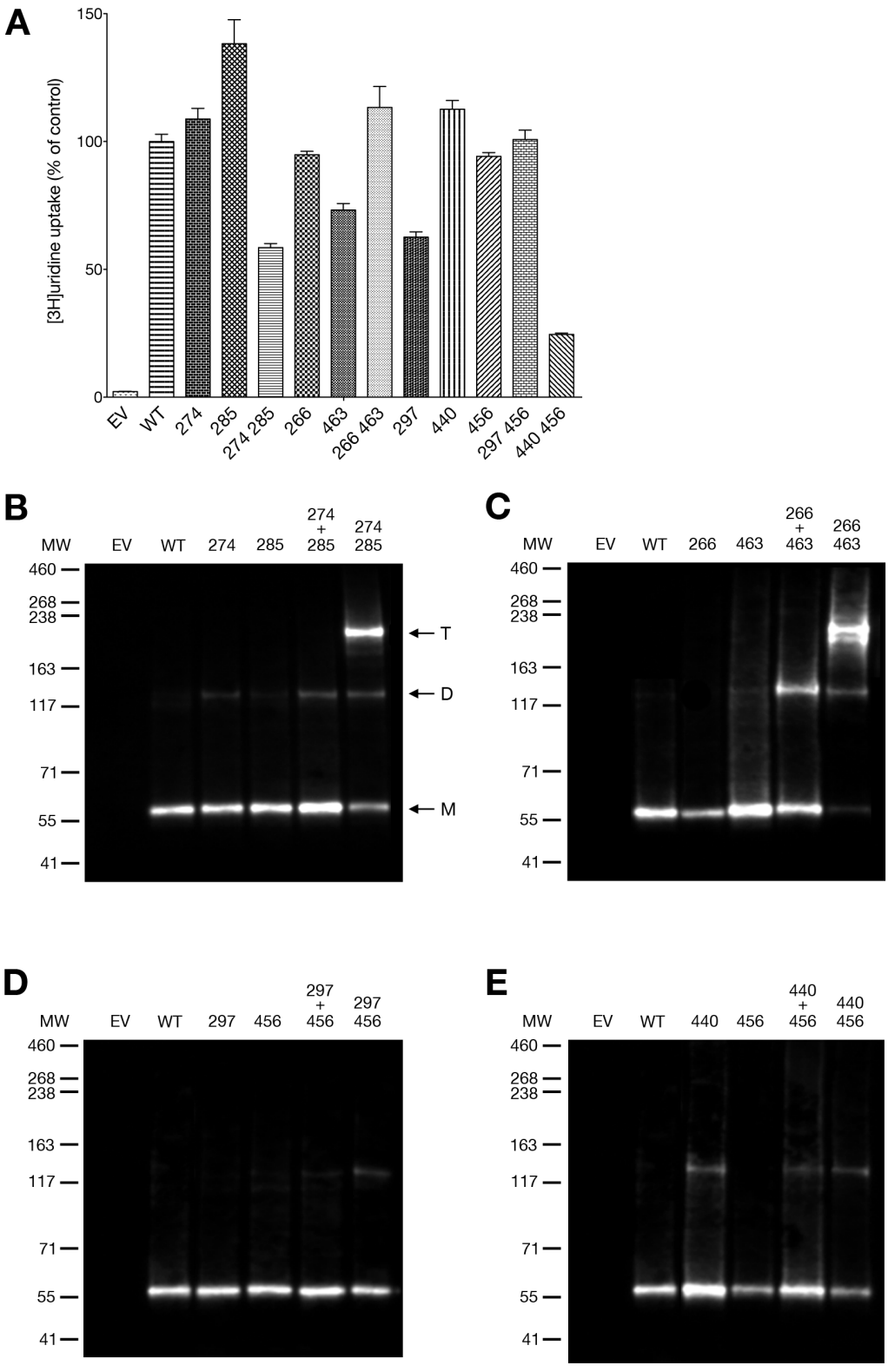


Figure 3.5 Functional validation and cross-linking of hCNT3 cysteine mutants. A) Uridine uptake inhibition studies in PK15NTD cells transiently transfected with hCNT3 constructs. B) Western blot of the cysteine cross-linking experiment involving V274C/A285C mutants, C) Y266C/F463C mutants, D) T297C/A456C mutants, and E) A440C/A456C mutants.

The third site, containing two separate cysteine pairs, surprisingly did not lead to the cross-linking of the trimer (**Figure 3.5D,E**). Because the equivalent residues in vcCNT do cross-link,[14] the absence of a cross-link in hCNT3 indicates a structural difference in the position and/or conformation of TM9 in vcCNT and hCNT3. In addition, the band corresponding to a dimer in the 440 single mutant suggests that the 440 residues are sufficiently close for disulfide bond formation. We repeated the experiment several times, including the replacement of the O_2/1,10-phenanthroline with molecular iodine, which is often a more effective oxidation agent for cross-linking transmembrane regions (**Figure S3.1**),[43] but with no change in results.

Discussion and conclusion

The main goal of this study was to determine whether or not hCNT3, like its bacterial homologue vcCNT, forms homo-trimers. By constructing a homology model, we leveraged structural information to introduce cysteine residues at key points along the hypothesized protomer-protomer interface. We have established that hCNT3 does trimerize, based on *in vitro* and cell-based experiments (**Figure 3.3** and **Figure 3.5**). We now discuss two major points in turn. First, we compare the crystallographic structure of vcCNT with a structural model of hCNT3 based on the cross-linking data. Second, we discuss the functional implications of trimer formation in hCNT3.

Comparison between structural model of hCNT3 and vcCNT Unlike chemical cross-linking using agents such as DSS or BMB, the formation of a disulfide bond between two residues on separate protomers requires close proximity and specific geometry.[44] Thus, the method of cysteine cross-linking limits false-positives resulting from transient protein-protein interactions and provides residue-level insight into the configuration of the protomers in the trimer. Both selected extracellular mutated sites along IH1 led to a cross-linked trimer, validating the relative placement of helices IH1, EH, and TM6 in the hCNT3 homology model. In addition, these cross-links validate the fold of hCNT3, because interactions between IH1 and TMs 5-8 are some of its defining features.[14]

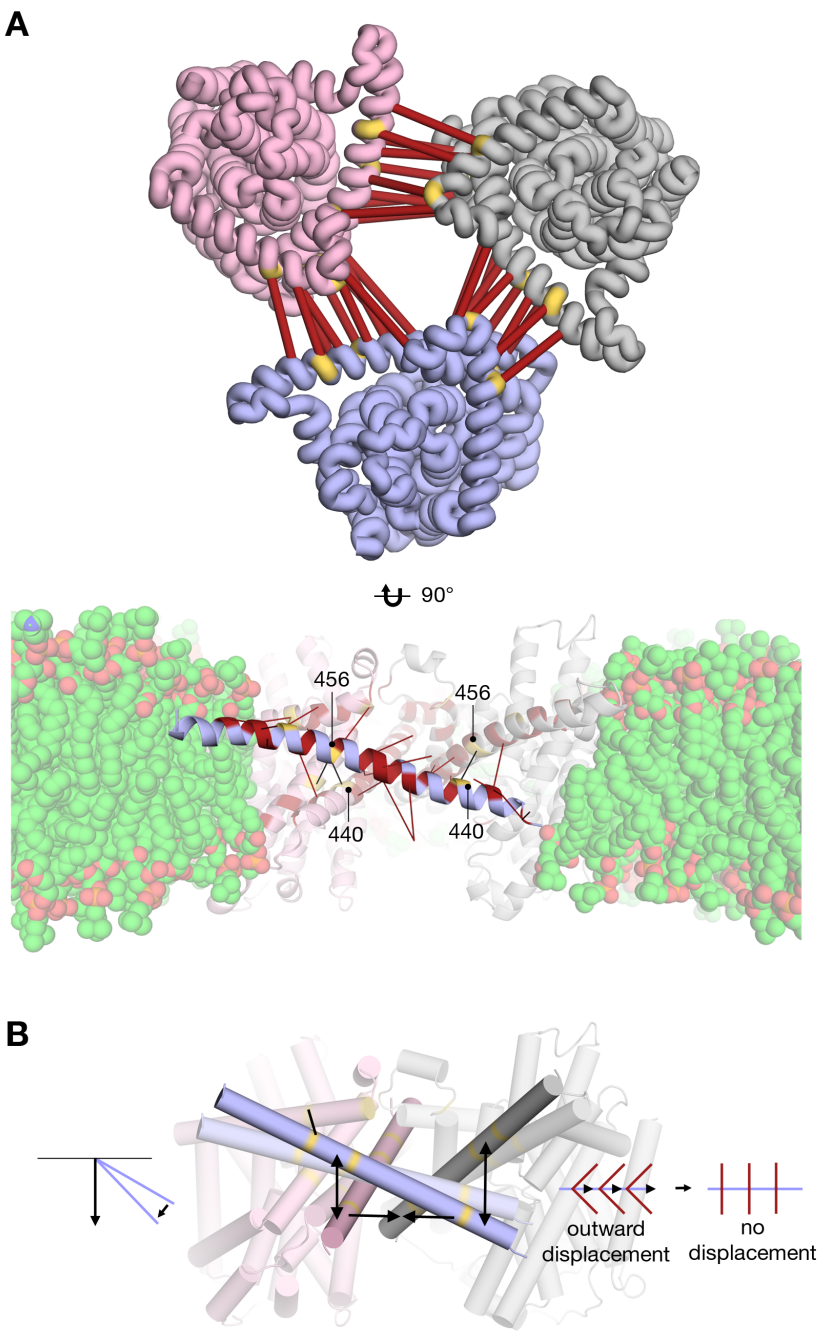


Figure 3.6 Residue covariation between different protomers and a schematic of the proposed differences between hCNT3 and vcCNT. A top) Cartoon representation of the hCNT3 comparative structure models is shown from the extracellular side. Three protomers are shown in different colors. Residue pairs with a high GREMLIN covariation score across the dimerization interface are shown in thick red bars. Location of mutated cysteines is shown in yellow. A bottom) Cartoon representation of the hCNT3 comparative structure model is shown with a simulated membrane aligned from the MemProtMD entry 3TIJ (ref. 45). Residue pairs involving TM9 residues with a high GREMLIN covariation score are shown in red. Red bars indicate the location of the other intra- or inter-protomer residue in a covariation pair. Location of muted cysteines is shown in yellow. Black lines are drawn between pairs of mutated cysteines. B) Cartoon representation of the comparative structure model of hCNT3 homo-trimer shown in the plane of the membrane, where TM9 on each of the protomers is highlighted. We propose a model where TM9 in the human structure is rotated in the plane of the membrane compared to the equivalent TM6 in vcCNT.

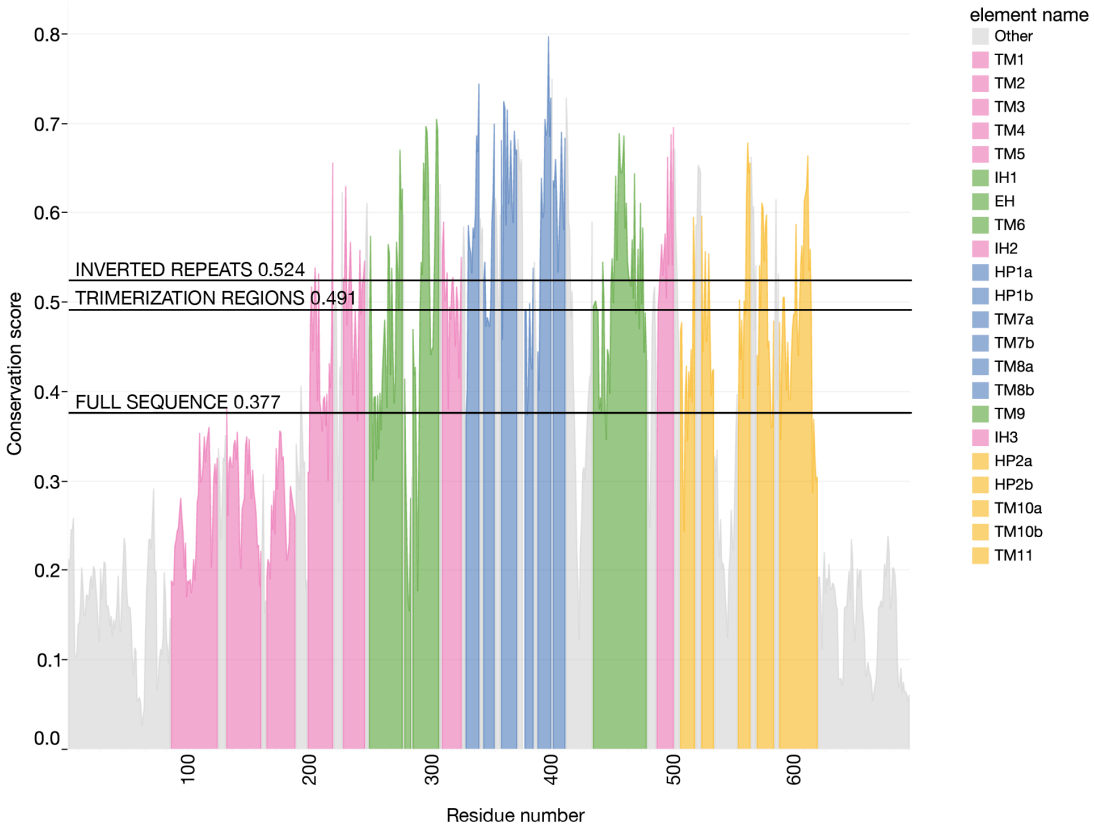


Figure 3.7 Graph of the sequence conservation score plotted for each residue. Colors follow the domains outlined in **Figure 3.4A**. Averages represent the full sequence, trimerization domain (IH1, EH, TM6, TM9), and inverted repeats (HP1a, HP1b, TM7a, TM7b, TM8a, TM8b, HP2a, HP2b, TM10a, TM10b, TM11).

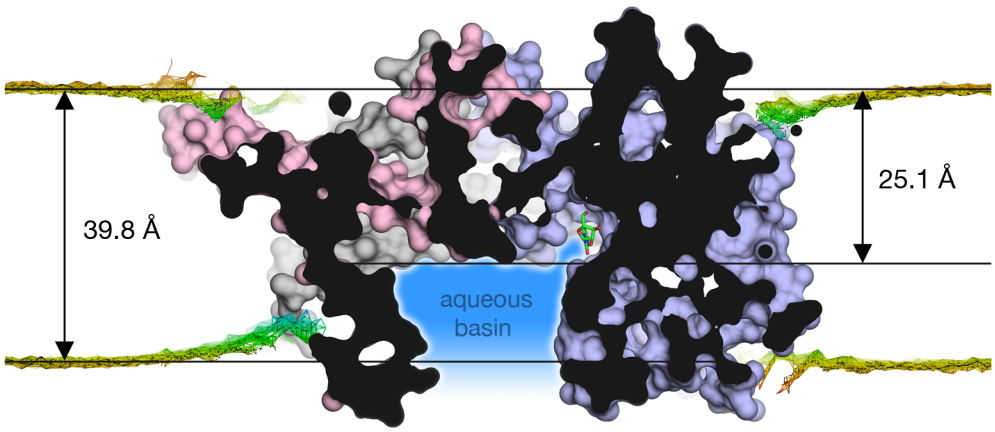


Figure 3.8 Cross-section of the hCNT3 comparative structure model. Three protomers are shown in different colors. Membrane is shown in yellow-green, colored according to local membrane deformations (MemProtMD)(ref. 45). Aqueous basin created in the center of the homo-trimer is shown in blue.

The differences between the known structure of vcCNT and the unknown structure of hCNT3 likely include the placement and/or conformation of TM9, the helix contributing to the trimerization domain (defined in **Figure 3.4A**). Based on the bacterial structure, TM9 is the longest helix in the structure, 63 Å long and spanning almost the entire thickness of the plasma membrane. None of the cross-links probing the cytoplasm-facing N-terminus of TM9 (440/456 and 297/456) led to the stabilization of the trimer. Instead, a dimer band surprisingly appeared for the 440 single mutant. To explain these results, we hypothesize that the N-terminal residues 440 in the unknown structure of hCNT3 are closer to each other than in the current hCNT3 model (Cβ-Cβ distance of 22 Å), while residue 456 is farther from both 297 and 440 than in the current model (4 Å). This hypothesis is consistent with an elevator-like movement[14,46] of the inverted repeats, in which the trimerization domain forms a stable, largely immobile scaffold, maintaining the overall architecture of the trimer. This elevator-like mechanism was also applied to the inward-occluded state to compute a model of the outward-open state.[47] In the elevator-like mechanism, the motions necessary for the transport cycle states are confined within the monomer, largely to the helices making up the transport domain, thus leaving the trimerization domain static throughout the cycle.

The GREMLIN score, based on an alignment 1629 sequences, shows that 8 and 14 of the 43 TM9 residues involved in respectively intra- and inter-protomer interactions co-vary (**Figure 3.6A**).[48] Because co-varying pairs of residues tend to be close in space, it is often possible to refine a structural model by minimizing distances between co-varying pairs of residues.[48] Applying this general consideration to the hCNT3 model, a translation of the helix deeper into the membrane is indicated by visual inspection (**Figure 3.6A**).

Combining the C-terminal TM9 266/463 and the TM9 440 cross-links, the co-evolutionary data, and the lack of cross-links for sites involving the N-terminal TM9 456 residue, has led us to propose that TM9 is rotated in the plane of the membrane in hCNT3 as compared to TM6 in vcCNT (**Figure 3.6B**). This parsimonious model satisfies the proximity requirements needed for the experimentally observed cross-link between A440C of adjacent protomers and the lack of a cross-link between the 440/456 and 297/456 residue pairs. The implications of these potential local structural differences are significant, because TM9 is predicted to interact directly with the nucleoside substrates. By creating a portion of the "thin" intracellular gate in the inward occluded conformation,[46,49] even slight positional changes in TM9 could affect the binding site volume, the specificity, and kinetics of transport.

Functional implications of trimerization Now that we have established that hCNT3 exists as a trimer, we address the question of its functional significance. The alignment of 883 sequences homologous to hCNT3 shows that the trimerization domain is as conserved in evolution as the inverted repeat domains, which perform the actual transport catalysis, indicating that the trimerization domain is indeed functionally important (**Figure 3.7**).[50]

As predicted by the structure of vcCNT, the translocation pathway of hCNT3 is confined to the monomer (i.e., each monomer contains an independent nucleoside and sodium binding sites). In addition, the helices in the trimerization and scaffold domains in the three protomers (**Figure 3.4A**) form one large aqueous basin within the membrane. This basin decreases the translocation distance for the ligand from approximately 40 Å to 25 Å (**Figure 3.8**).[45] An aqueous basin of similar depth has been reported within the trimeric structure of Glt_{Ph}, a eukaryotic glutamate transporter homologue from *Pyrococcus horikoshii*,[51] that is not related in sequence or fold to CNTs. The quaternary structure of the trimer ensures that the translocation pathway is not exposed to the membrane while also minimizing the size of the hCNT3 monomer, compared to translocation in a monomeric transporter. On its own, without the other two protomers, and thus without the aqueous basin, each protomer is likely unstable in the membrane due to the exposure of hydrophilic regions (e.g., parts of TM9) to the lipid membrane. Thus, we hypothesize that hCNT3 is an obligate trimer.

Symmetry is a common feature among protein complexes found in the PDB (approximately 85%).[52] Potential benefits of symmetrical homo-complexes include increased coding efficiency, reduced aggregation, enhanced error control, and enhancement to sensitivity of selection; the reason is that symmetrical interfaces contain duplicates of pairwise contacts, and thus any favorable/unfavorable mutations are amplified.[26,53-55] Although the sum of the three protomer-protomer interfaces is large (3 x 1250 $Å^2$)[56], it has been shown that only a small subset of residues, known as hotspots, contributes most of the free energy of binding.[57] Because oligomerization has the potential to create new binding sites,[26] it might be possible to disrupt the formation of the hCNT3 homo-trimer by targeting the dimeric interface. Such a binding site could be exploited for selectively modulating hCNT3, via allostery. Selective inhibition in particular might be appealing because hCNT3 substrate specificity overlaps with that of several other transporter families (e.g., ENTs).[12] Potential applications involve cases such as chronic lymphocytic leukemia patients characterized by up-regulated cytoplasmic hCNT3 levels, who experience a lower response rate to fludarabine therapy.

Targeting the cytoplasmic hCNT3 with allosteric inhibitors would leave the required ENT-mediated drug uptake on the plasma membrane unaffected.[19]

In conclusion, we have shown that hCNT3 forms an obligate homo-trimer in cells. By creating a homology model, we leveraged structural data to introduce cysteines at various points along the dimerization interface. Cross-linking results indicate local structural differences between hCNT3 and vcCNT in TM9 of the trimerization domain. The quaternary structure creates an aqueous basin that significantly shortens the substrate translocation distance.

Supplemental information

Supplemental figures are included below.

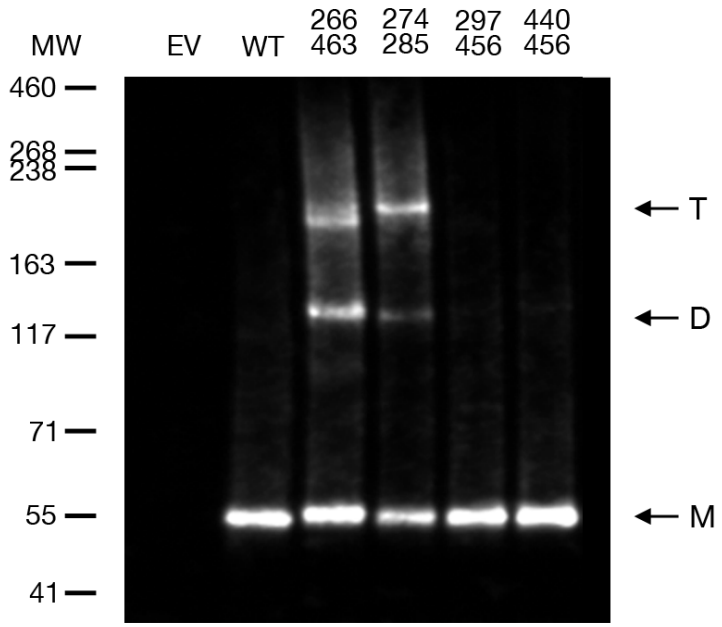


Figure S3.1 Western blot of the cysteine I_2 cross-linking experiment involving V274C/A285C, Y266C/F463C, T297C/A456C, and A440C/A456C double mutants.

Figure S3.2 Sequence alignment between vcCNT (3TIJ), hCNT1, hCNT2, and hCNT3. Residue conservation is indicated in the Clustal X color scheme.

hCNT3 expression and purification

Despite many advancements in the methods of structure determination, eukaryotic membrane proteins remain to be difficult targets. While they account for 25-35% of the human genome,[26,58] only 2,685 of the approximately 116,000 entries in the Protein Data Bank (PDB search on 21 December 2016) correspond to α-helical transmembrane protein structures.[59] One of the main reasons is what explicitly defines them – the presence of a hydrophobic transmembrane region. Methods used to study the structures of these proteins require their extraction from the cell membrane and rely on the introduction of detergents to create a detergent-lipid-protein micelle.[60] It is unrealistic, however, to expect such a simple extraction process to yield stable systems, when cells have evolved to use ~5% of their genes to synthesize thousands of lipid types, which stabilize the protein in the endogenous environment.[61,62] Unlike prokaryotic homologs, eukaryotic membrane proteins require extensive maturation and targeting apparatuses involving glycosylation and association with the proper lipid types, which occurs in the endoplasmic reticulum and the Golgi apparatus.[63-65]

New technologies have been developed to improve the efficiency, stability, and quality of purified eukaryotic membrane proteins. Milligram quantities of mammalian membrane proteins expressed in baculovirus-infected insect cells[63] and in stably and transiently transfected human embryonic kidney 293S cells[66,67] have recently been reported. Eukaryotic expression systems provide a near-native lipid environment and proper protein folding. Further, endogenous membrane lipid and cholesterol contents have increasingly become recognized as important to the stability of the solubilized proteins. Inclusion of lipids POPC, POPE, and POPG in the 3:1:1 ratio helped crystallize the rat Kv1.2 K1 channel,[68] while the addition of cholesteryl hemisuccinate (CHS) in the CHS:DDM ratio of 0.1:1 provided maximal stabilization for human G protein-coupled receptors.[69] New amphiphiles, including the maltose–neopentyl glycol (MNG) amphiphile family, show advantages over conventional detergents, leading to enhanced structural stability and successful crystallization.[70] Synthetic amphipathic polymers, amphipols, have been developed specifically to have a high affinity for the transmembrane surface of membrane proteins,[71] and have been used for structural determination of previously unreachable targets, such as the TRPV1 ion channel.[72]

Finally, several advancements within the field of cryo-electron microscopy (cryo-EM) have achieved the determination of near-atomic resolution structures by direct fitting of atomic models

into experimental high-resolution density maps.[73] In addition to allowing for the study of targets that are difficult to crystallize, cryo-EM has the capacity to capture proteins in different functional states, thus providing unprecedented biological insight.[74] Additional developments in fragment antigen binding (Fab), which can be readily generated using phage display, have allowed for the study of even small proteins and protein complexes.[75]

Here, we present a multi-year effort to express and purify human CNT3 for the purpose of structure determination with either X-ray crystallography or single-particle cryo-EM. A near-atomic resolution structure of the human CNT3 homo-trimer could provide us with insight into the dimerization interface, structural features governing substrate specificity, and conformational changes associated with the transport cycle. Despite a combinatorial approach with several constructs, lipid extracts, detergent and thermostability screens, we were largely unsuccessful in obtaining even a low-resolution structure of the human transporter. This section describes the expression and purification conditions that were attempted and the complications that were identified in the process.

Methods

Expression in insect cells hCNT3 was expressed in either High Five, Sf9, or Sf21 insect cells using the Bac-to-Bac Baculovirus Expression System (invitrogen). The gene encoding the full-length or truncated variant of hCNT3 was cloned into the pFastBac1 vector. The bacmid and the recombinant baculovirus were generated per the manufacturer's protocol (Invitrogen). Sf9 or Sf21 cells were used to amplify the viral stock. P2 or P3 viral stocks were used for expression studies using an MOI of 1. Cells were harvested 48 hours post infection.

Purification of hCNT3 The protocol for protein purification was described previously.[30] Cells were spun down and solubilized in the lysis buffer (50 mM Tris pH 8.0, 300 mM NaCl, cOmplete Mini Protease Inhibitor Cocktail (Roche)) before passing the solution through 5 cycles inside of Emulsiflex at 5,000-10,000 psi. The lysate was cleared with a 15 min spin at 15,000 x g at 4°C before pelleting the membranes with an 80 minute ultracentrifuge spin at 40,000 x g at 4°C. Membranes were then resuspended in the resuspension buffer (varies) and stirred for 1 hour at 4°C. Excess lipids were pelleted with a 30 min spin at 35,000 rpm at 4°C. Supernatants were applied to either the TALON (Takara Bio), Ni-NTA (Thermo Scientific), or amylose (NEB) affinity resin and stirred

for 2 hours. Immobilized protein was applied to a gravity flow column, washed with washing buffer, and either eluted off the column with 300 mM imidazole, or digested overnight with either HRV 3C (Thermo Scientific) or thrombin (Agard lab) proteases. The protein was concentrated with 100 kDa cutoff Amicon concentrators (Sigma) and applied to a Superdex 200 size-exclusion column along with the running buffer (varies). Fractions were collected every 500 µL for further analysis. For Coomassie gel stains, fractions were run on 4-15% Criterion Tris-HCl gels (Bio-Rad) under reducing and denaturing conditions and stained using SimplyBlue SafeStain (invitrogen) following manufacturer's protocols. For Western blots, polyclonal antibody against CNT3 was used for detection (Sigma HPA023311), with SuperSignal West Femto Chemiluminescent substrate (Thermo Scientific 34095) on the FluorChem E development system (proteinsimple).

Thermostability assay The protocol for the thermostability assay was described previously.[76] Briefly, the Protein Thermal Shift (PTS) dye was mixed with the protein purification buffer 1:20 (v/v). 10 µl of 1:20 PTS dye mixture was added to 120 µl of concentrated protein (>0.1 mg/mL) diluted in test buffer to attain a total volume of 130 µl. The software for a thermal assay was set to a temperature ramp beginning at 20° and ending at 90°C over a period of 1 hour. Excitation (450 nm) and emission wavelengths (500 nM) were set. The melting curves were analyzed to identify the midpoint of the transition.

Results

Construct 1 Expression and purification trials began with construct 1 (**Figure 3.9**). hCNT3 has been shown to be previously glycosylated,[39,77,78] therefore the 4 putative glycosylation sites were left intact. Since it was a preliminary expression trial, it was unknown whether the tag placement on N- versus C-terminus affects protein stability, so both were tried (N-terminal tag on construct 2). The baculovirus expression system was used because it has been shown to be quite robust and successful with the expression of other transmembrane proteins from higher eukaryotes.[64,79-81] The standard conditions used previously by the Stroud lab in the purification of ABC transporters were taken as the starting point of optimization.[30] DDM was used because it is a mild, nonionic detergent frequently used for solubilization of membrane proteins.[82] Its critical micelle concentration (CMC) is 0.12 mM (0.006%) in 0.2M NaCl (Thermo Fisher), thus we have ensured that at 0.1% we were above it. To determine where to expect the homo-trimer to elute, we injected the gel filtration standard (Bio-Rad). Based on the peaks in the SEC profile (**Figure 3.10**), we estimated the micelle

containing the hCNT3 homo-trimer to elute around 11 mL.[83] The SEC profile of construct 1 (run 1, **Figure 3.11**) showed that although there was a large amount of material, most of it was largely either aggregated or unfolded, and we did not follow up on this construct further.

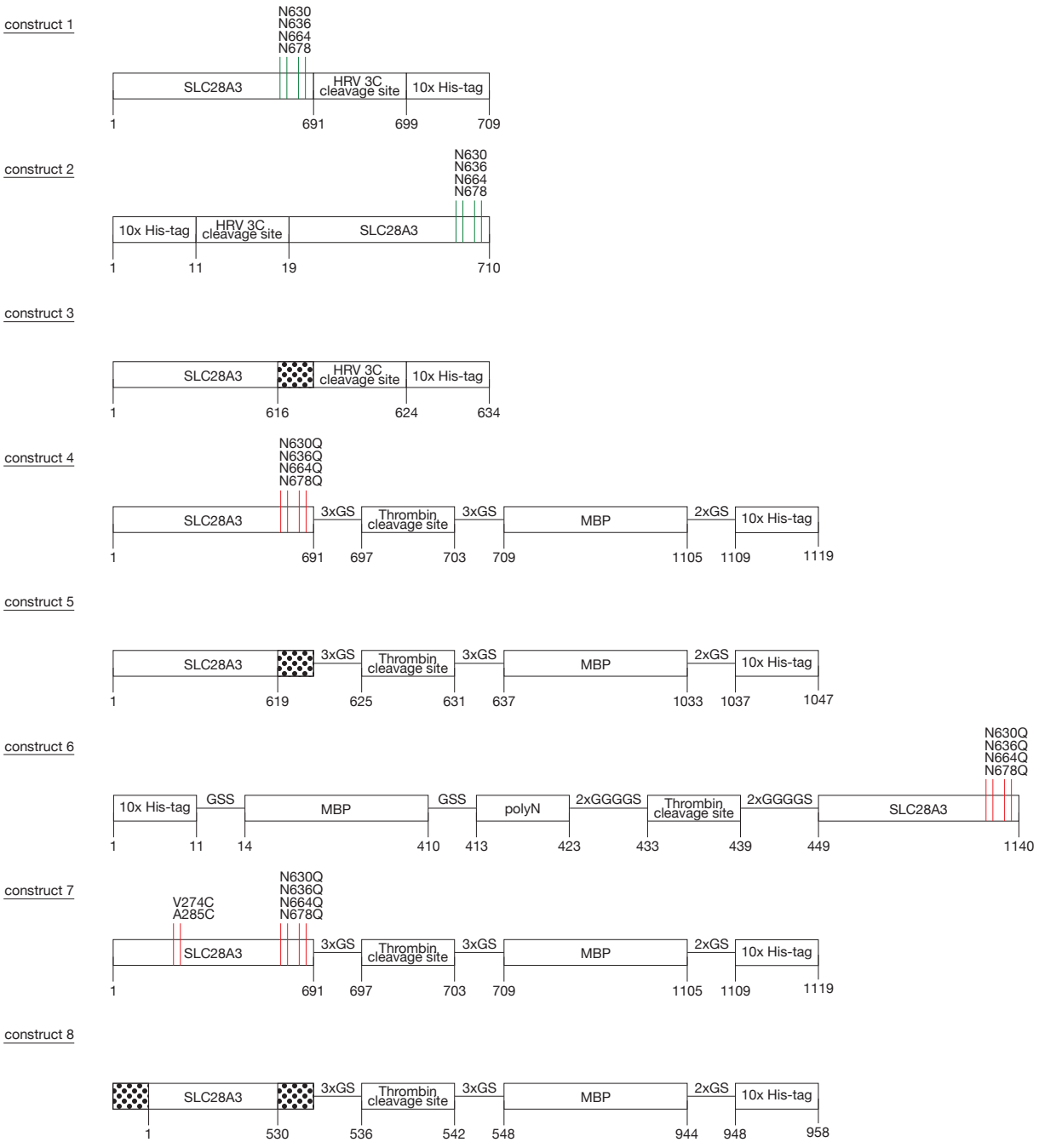


Figure 3.9 Constructs created for hCNT3 expression in insect cells. Glycosylation sites are indicated with green lines. Mutation sites are indicated with red lines. Length of segments not to scale.

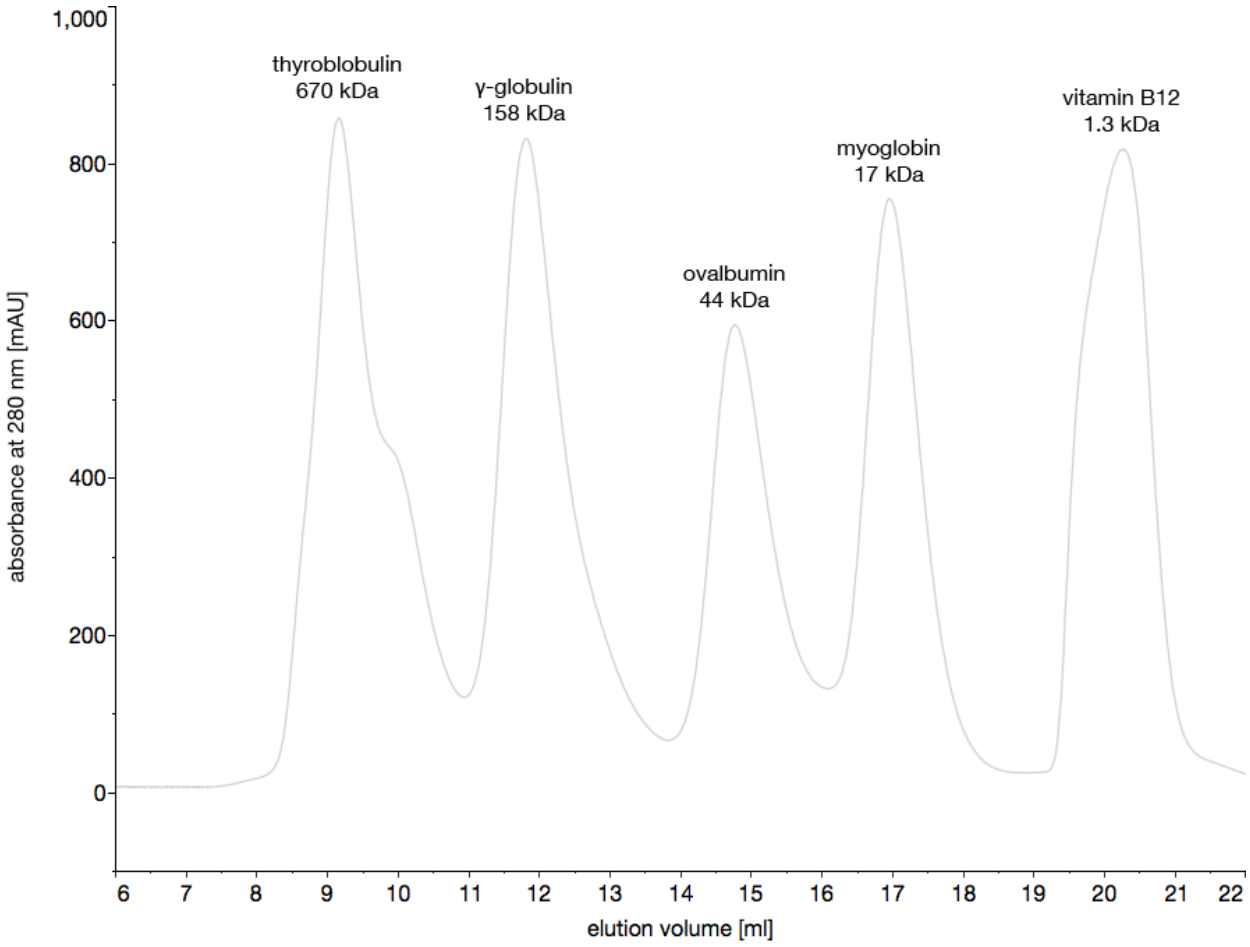


Figure 3.10 SEC profile of the gel filtration standards.

Construct:	1	Beads:	Ni-NTA
Insect cell line:	Sf21	Protease:	none
Amount of membrane (g):	1.4		

Resuspension buffer	Wash I buffer	Wash II buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	none	50 mM HEPES pH 7	50 mM HEPES pH 7
300 mM NaCl	300 mM NaCl		300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol		10% glycerol	10% glycerol
4 mM imidazole pH 7	2 mM imidazole pH 7		300 mM imidazole pH 7	0.1% DDM
1% DDM	0.1% DDM		0.1% DDM	
1 mg/mL iodoacetamide				
protease inhibitor				

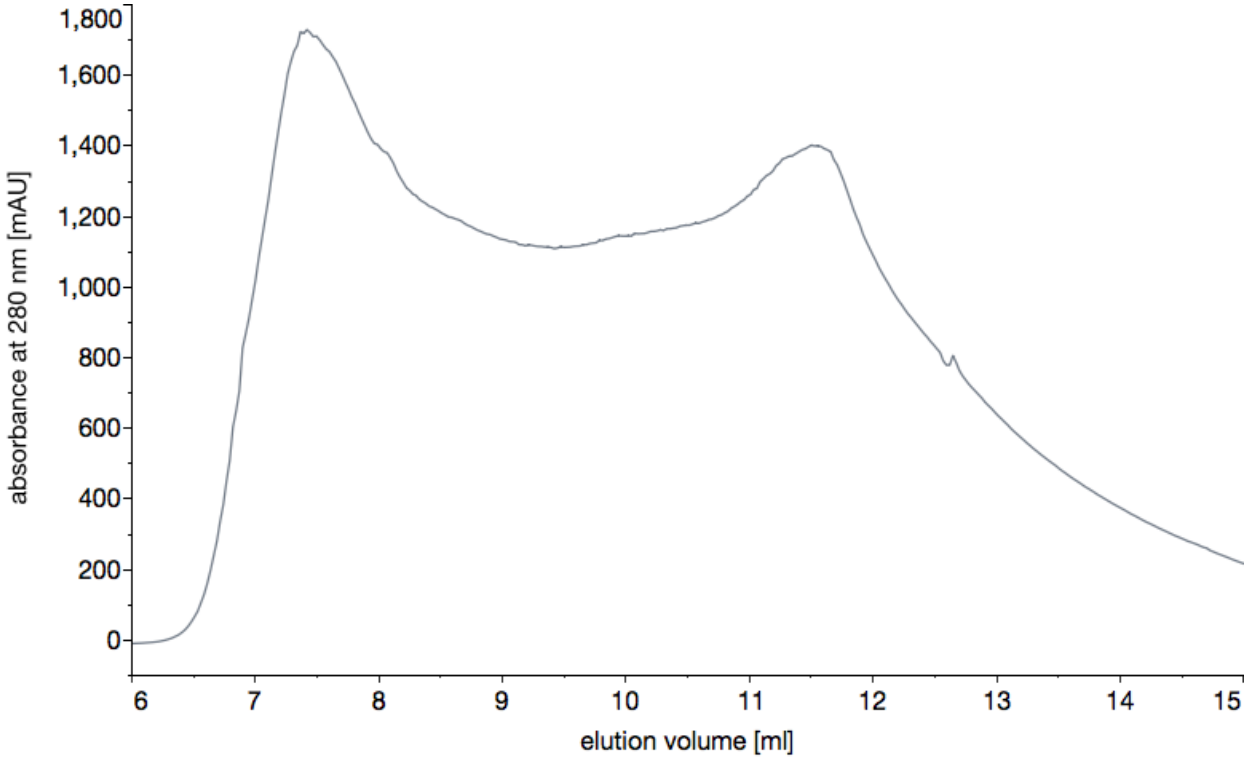


Figure 3.11 Purification conditions and elution profile of run 1.

Construct 3 Computational analysis of the hCNT3 sequence with PSIPRED[84] and DISOPRED[85] predicts the N-terminal residues 1-84 to be cytoplasmic and largely disordered and the C-terminal residues 614-691 to be extracellular and lacking secondary structure. In line with the size of vcCNT, we decided to create a truncated construct that eliminates the C-terminal region following the stop codon of the equivalent vcCNT position, which also removes the glycosylation sites, thus potentially reducing sample heterogeneity. Previous work has shown that N-glycosylation does not affect localization of rat CNT2,[39] which supports the hypothesis that the elimination of the extracellular C-terminus should not affect expression and folding.

The SEC profile of the same starting conditions as construct 1 (run 2, **Figure 3.12**) provided a slightly better profile. Gel analysis (**Figure 3.13**) indicated that the vast majority of the sample was hCNT3 but improvements could be made on the purity front by adding more stringent washing steps. In addition to altering the washing protocol, we tried including uridine, a high affinity substrate of hCNT3 for stabilization. Although, this combination did not result in a better profile (run 3, **Figure 3.14**), the addition of an extra washing step did lead to a significantly purer sample (**Figure 3.15**). Next, we tested different pH values, pH 8 (run 4, **Figure 3.16**), physiological pH 7.4 (run 5, **Figure 3.17**), and pH 6 (run 6, **Figure 3.18**), but none improved the sample stability over the previously tested pH 7. Unlike uridine, phloridzin is an inhibitor of hCNT3. We tried including it in the elution and SEC buffers with the hope of stabilizing the homo-trimer in a particular conformation. It also did not lead to an improvement in the sample stability (run7, **Figure 3.19**).

To further reconstitute the endogenous cell membrane environment, we tested adding *E. coli* polar lipid extract (Avanti Polar Lipids) (run 8, **Figure 3.20**), cholesteryl hemisuccinate (CHS) (run 9, **Figure 3.21**), or both (run 10, **Figure 3.22**). The sole addition of CHS led to problems with its dissolution. It was determined afterwards that it requires extensive sonication in order to go into solution. The *E. coli* polar lipid extract significantly improved the sample stability, and the combination with CHS completely eliminated the void peak. A mass spectrometry analysis of the Coomassie gel indicated that some viral and insect proteins still remain in the sample (**Figure 3.23**). A thermostability assay was performed to quickly scan different thermostabilizing buffer conditions (**Figure 3.24**). The three most stabilizing buffers were buffer 7 (50 mM MES pH 6.5), buffer 20 (200 mM NaCl, 50 mM MES pH 5.5), and buffer 22 (200 mM NaCl, 50 mM MES pH 6.5). The combination of CHS and *E. coli* polar lipid extract with DDM became the baseline of comparison for future trials.

Before these thermostabilizing conditions could be examined, we tested adding phospholipids POPC, POPE, and POPG in a 3:1:1 ratio (run 11, **Figure 3.25**), liver polar lipids extract (Avanti Polar lipids) (run 12, **Figure 3.26**), and a high 1M NaCl concentration (run 13, **Figure 3.27**), all with no general improvements. Even combining CHS and *E. coli* polar lipid extract with the low pH (run 14, **Figure 3.28**) and low NaCl salt concentration (run 15, **Figure 3.29**), the two trends from the thermostability assay, did not lead to an absence of the void peak.

Next, we tested a new class of Anatrace neopentyl glycol amphiphiles, octyl glucose neopentyl glycol (OGNG) (run 16, **Figure 3.30**) and lauryl maltose neopentyl glycol (LMNG) (run 17, **Figure 3.31**). DDM was used to extract the sample from the membrane, and subsequently exchanged for the amphiphiles during the IMAC step. Neither of them showed any promise in further stabilizing the sample even with CHS and *E. coli* polar lipid extract. To build on the baseline result and further decrease the homogeneity of the sample, we tried to cut the sample off the column with the HRV 3C protease (run 18, **Figure 3.32**). It again did not improve the SEC profile. Neither did attempting the most stabilizing buffers from the thermostability assay with construct 2 - substituting $NaSO_4$ for NaCl (run 19, **Figure 3.33**) or doubling the glycerol concentration (run 20, **Figure 3.34**). Since a replicate of the CHS and *E. coli* polar lipid extract, the best conditions, did not result in a same, consistent profile (run 21, **Figure 3.35**), construct 3 was abandoned.

Construct:	3	Beads:	TALON
Insect cell line:	High Five	Protease:	none
Amount of membrane (g):	34		

Resuspension buffer	Wash I buffer	Wash II buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	none	50 mM HEPES pH 7	50 mM HEPES pH 7
300 mM NaCl	300 mM NaCl		300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol		10% glycerol	10% glycerol
4 mM imidazole pH 7	2 mM imidazole pH 7		300 mM imidazole pH 7	0.1% DDM
1% DDM	0.1% DDM		0.1% DDM	
1 mg/mL iodoacetamide				
protease inhibitor				

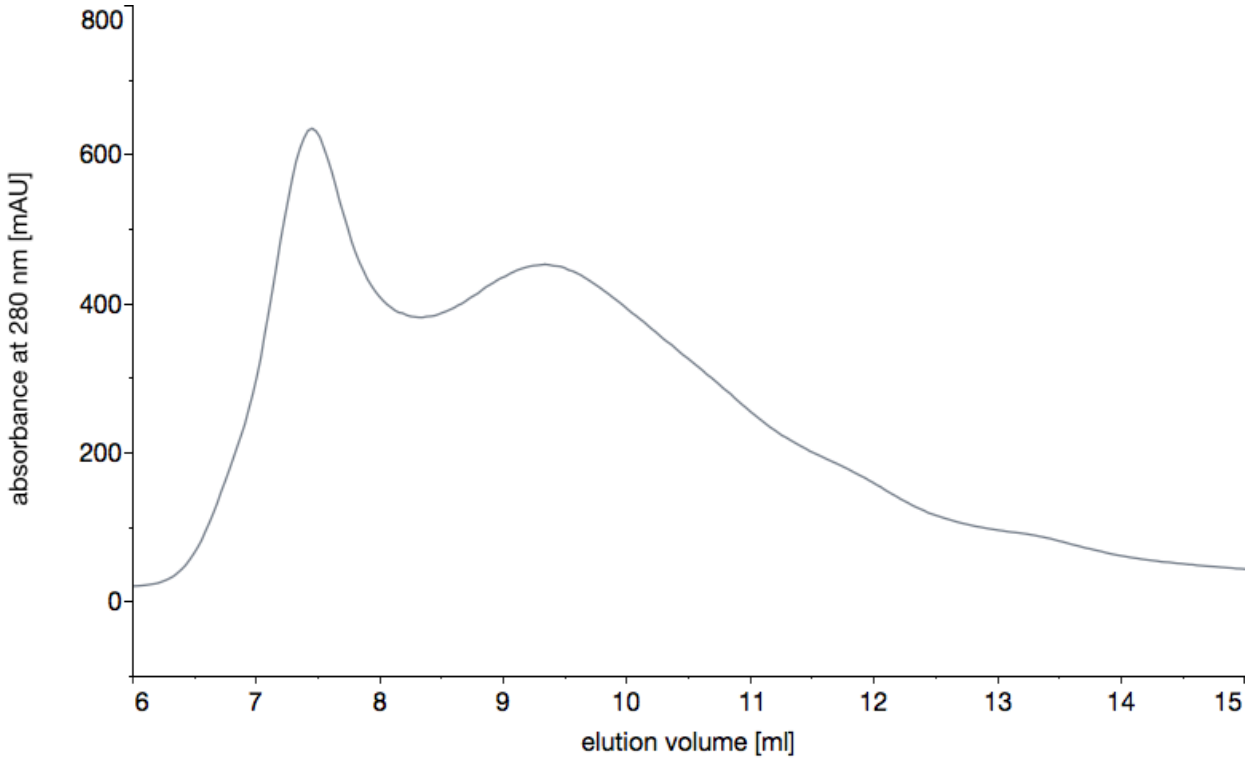


Figure 3.12 Purification conditions and elution profile of run 2.

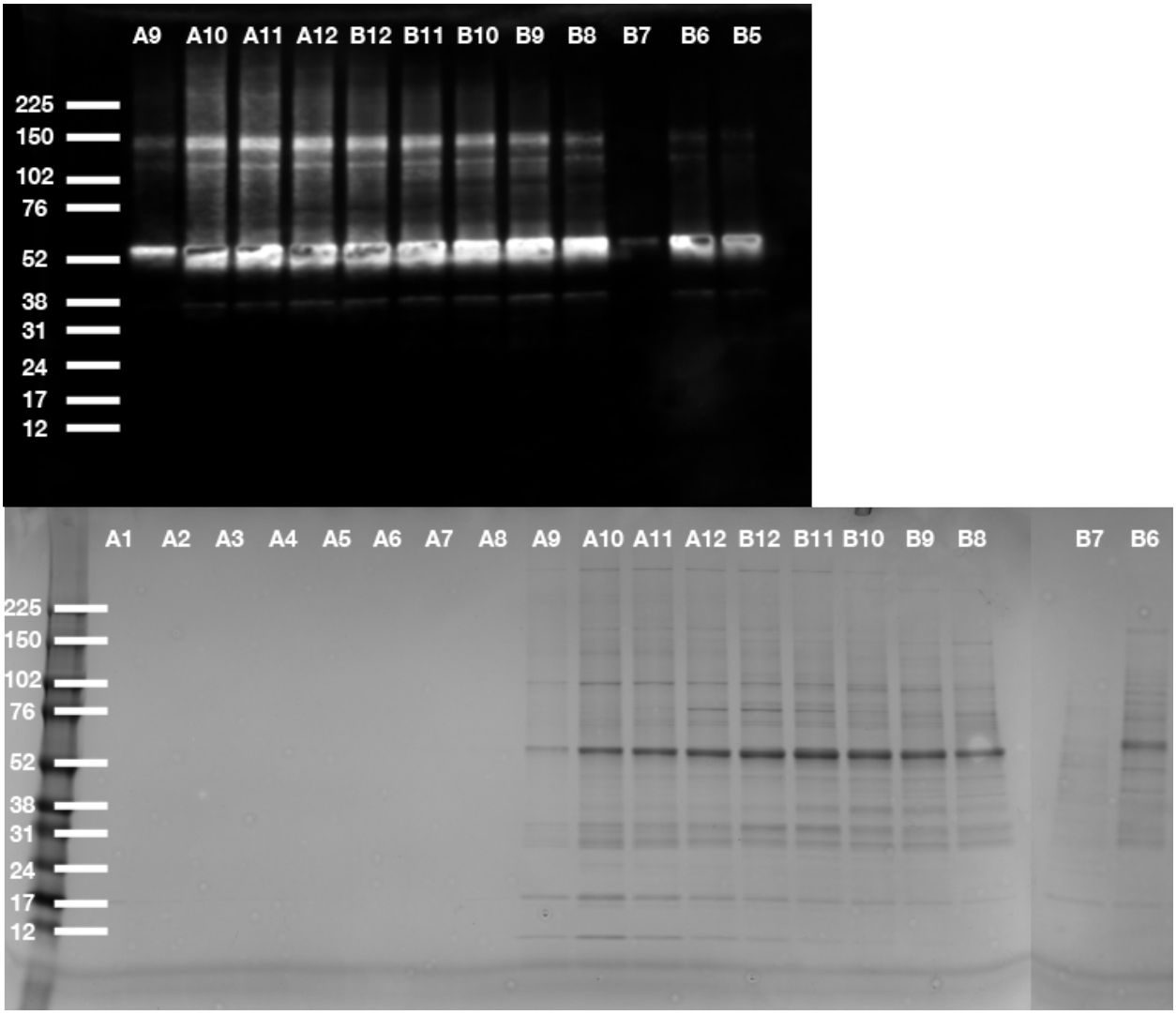


Figure 3.13 Western blot (top) and Coomassie gel stain (bottom) of run 2 fractions.

Construct:	3	Beads:	TALON
Insect cell line:	High Five	Protease:	none
Amount of membrane (g):	7.5		

Resuspension buffer	Wash I buffer	Wash II buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
4 mM imidazole pH 7	2 mM imidazole pH 7	10 mM imidazole pH 7	300 mM imidazole pH 7	0.1% DDM
1% DDM	0.1% DDM	0.1% DDM	0.1% DDM	2.2 mM uridine
1 mg/mL iodoacetamide			2.2 mM uridine	
protease inhibitor				

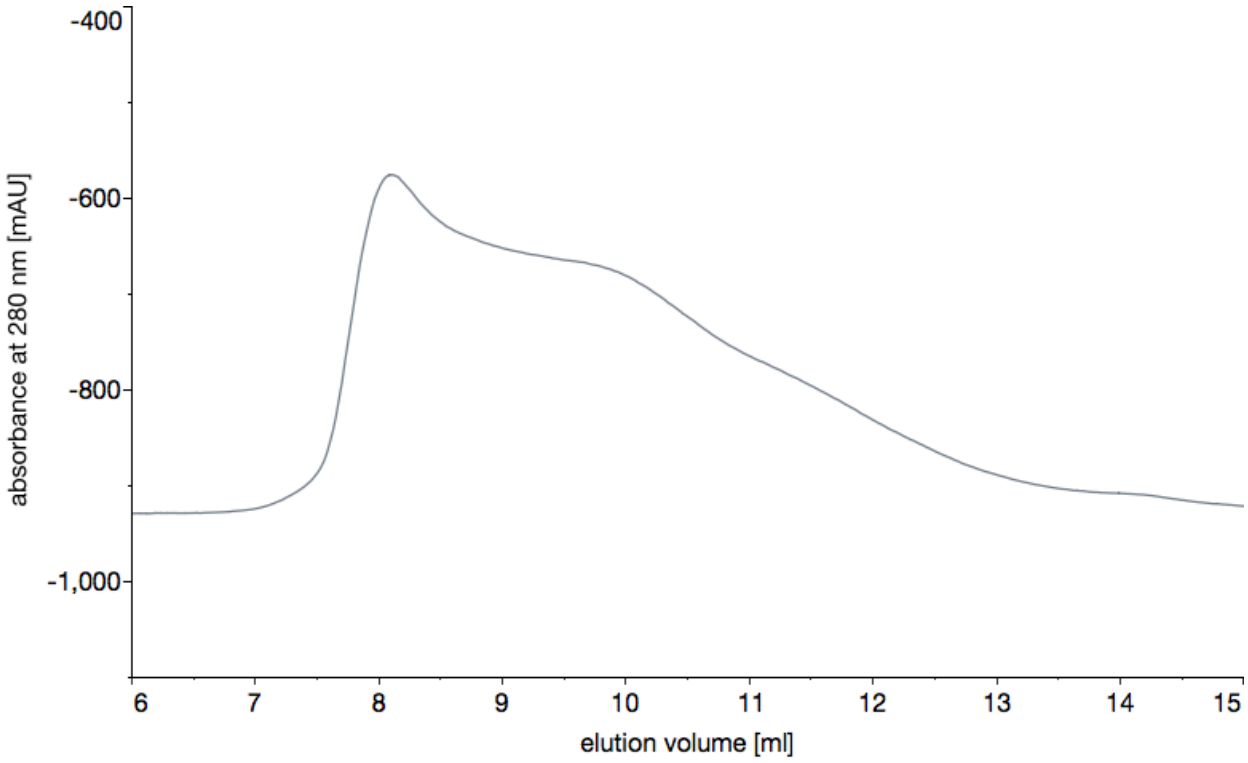


Figure 3.14 Purification conditions and elution profile of run 3.

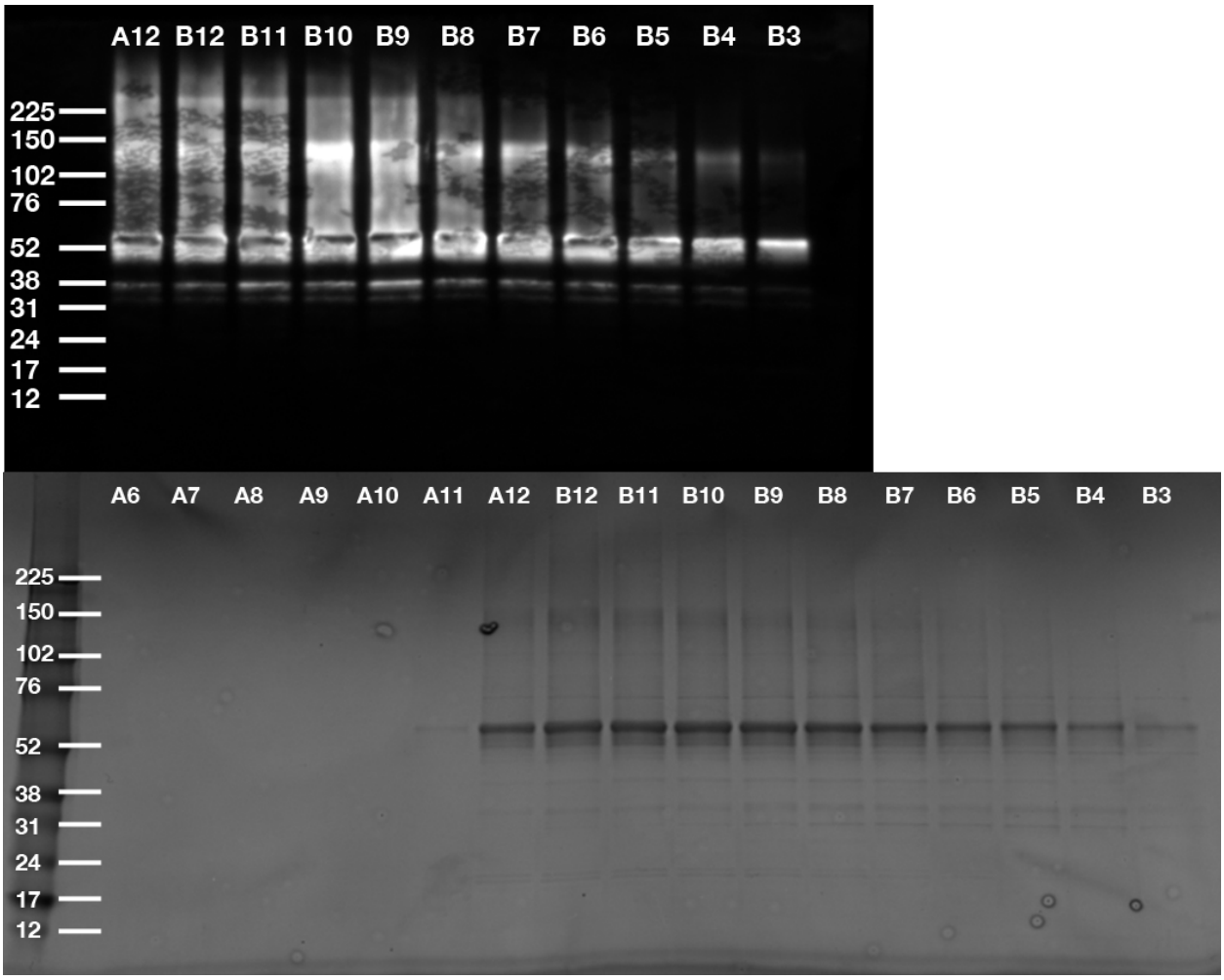


Figure 3.15 Western blot (top) and Coomassie gel stain (bottom) of run 3 fractions.

Construct:	3	Beads:	TALON
Insect cell line:	High Five	Protease:	none
Amount of membrane (g):	1.08		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 8	50 mM HEPES pH 8	50 mM HEPES pH 8	50 mM HEPES pH 8	50 mM HEPES pH 8
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
4 mM imidazole pH 8	2 mM imidazole pH 8	10/20 mM imidazole pH 8	300 mM imidazole pH 8	0.1% DDM
1% DDM	0.1% DDM	0.1% DDM	0.1% DDM	
1 mg/mL iodoacetamide				
protease inhibitor				

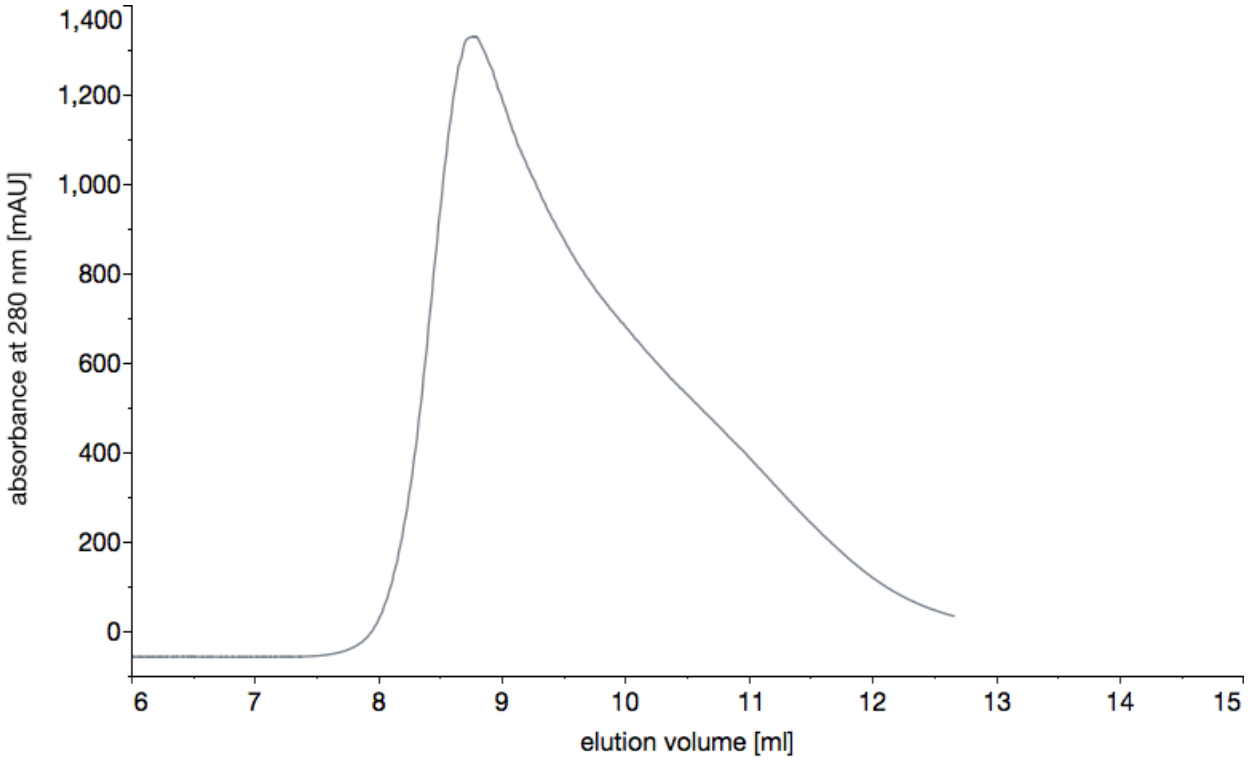


Figure 3.16 Purification conditions and elution profile of run 4.

Construct:	3	Beads:	TALON
Insect cell line:	High Five	Protease:	none
Amount of membrane (g):	1.08		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7.4	50 mM HEPES pH 7.4	50 mM HEPES pH 7.4	50 mM HEPES pH 7.4	50 mM HEPES pH 7.4
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
4 mM imidazole pH 7.4	2 mM imidazole pH 7.4	10/20 mM imidazole pH 7.4	300 mM imidazole pH 7.4	0.1% DDM
1% DDM	0.1% DDM	0.1% DDM	0.1% DDM	
1 mg/mL iodoacetamide protease inhibitor				

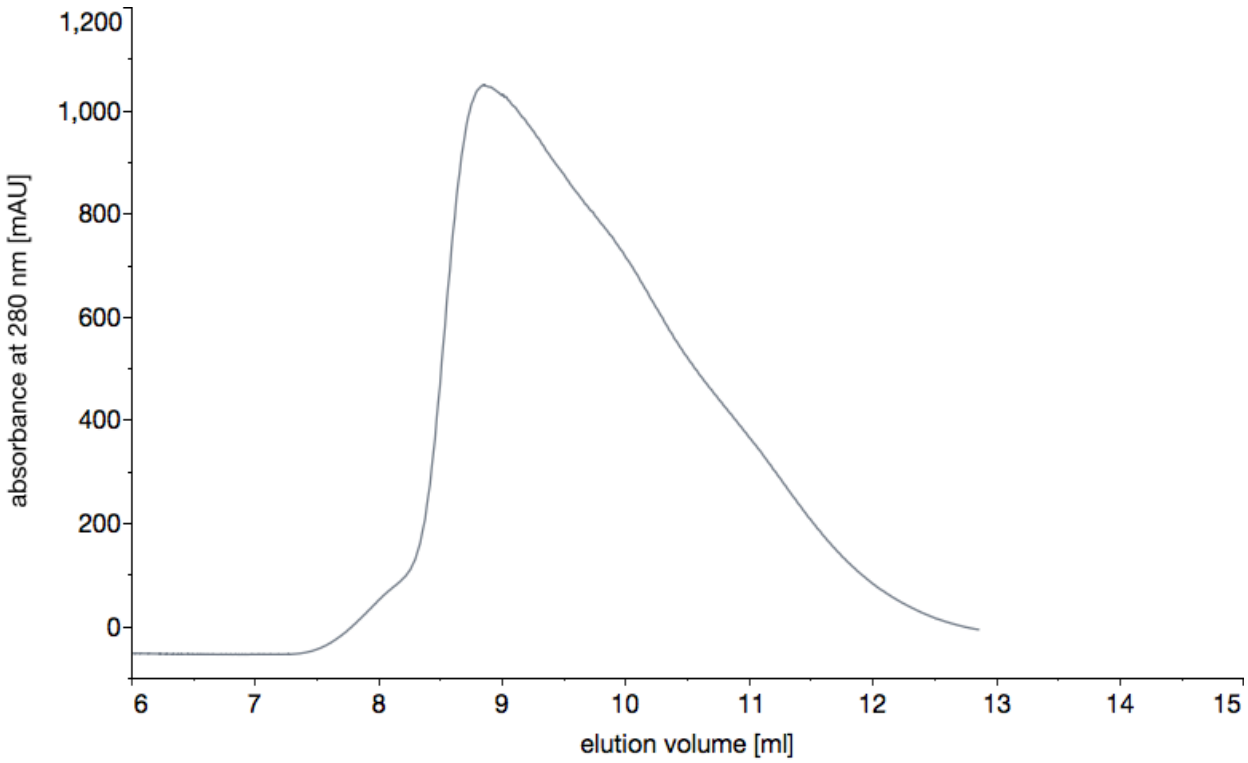


Figure 3.17 Purification conditions and elution profile of run 5.

Construct:	3	Beads:	TALON
Insect cell line:	High Five	Protease:	none
Amount of membrane (g):	1.08		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 6	50 mM HEPES pH 6	50 mM HEPES pH 6	50 mM HEPES pH 6	50 mM HEPES pH 6
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
4 mM imidazole pH 6	2 mM imidazole pH 6	10/20 mM imidazole pH 6	300 mM imidazole pH 6	0.1% DDM
1% DDM	0.1% DDM	0.1% DDM	0.1% DDM	
1 mg/mL iodoacetamide				
protease inhibitor				

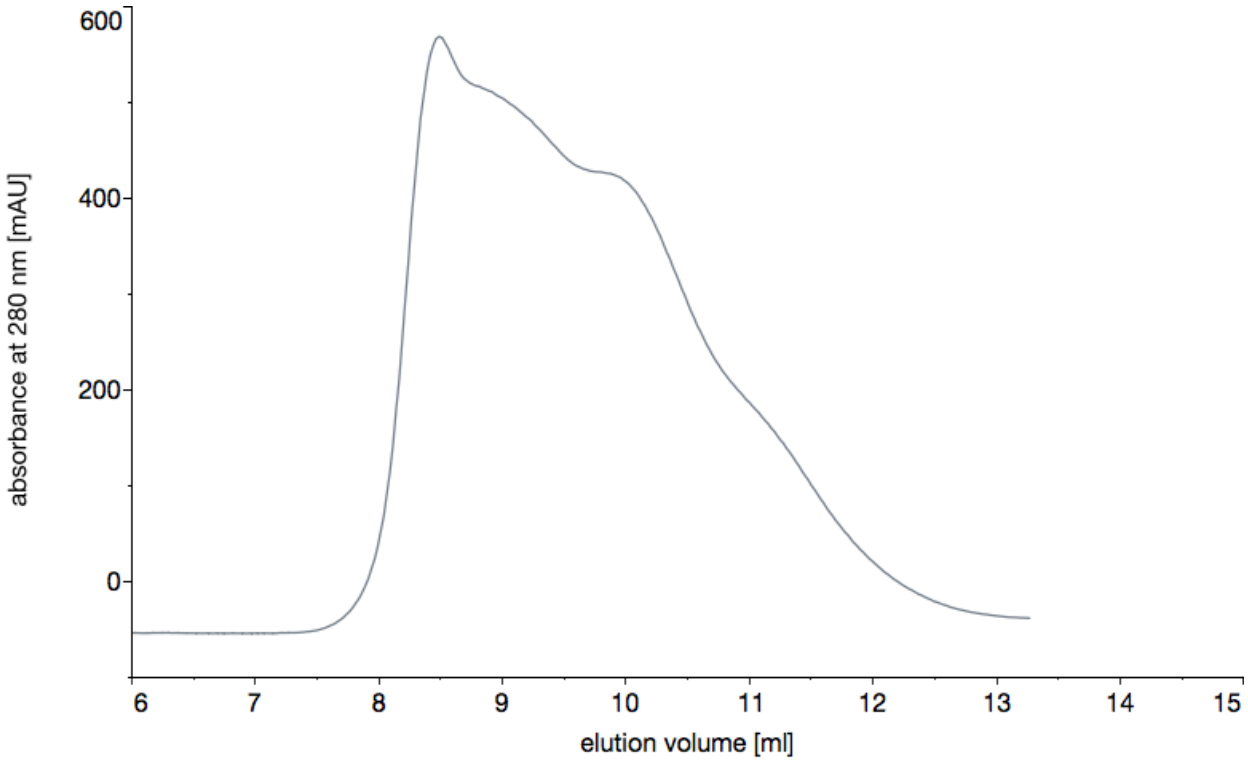


Figure 3.18 Purification conditions and elution profile of run 6.

Construct:	3	Beads:	TALON
Insect cell line:	High Five	Protease:	none
Amount of membrane (g):	1.08		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
4 mM imidazole pH 7	2 mM imidazole pH 7	10/20 mM imidazole pH 7	300 mM imidazole pH 7	0.1% DDM
1% DDM	0.1% DDM	0.1% DDM	0.1% DDM	1.6 mM phloridzin
1 mg/mL iodoacetamide			1.6 mM phloridzin	
protease inhibitor				

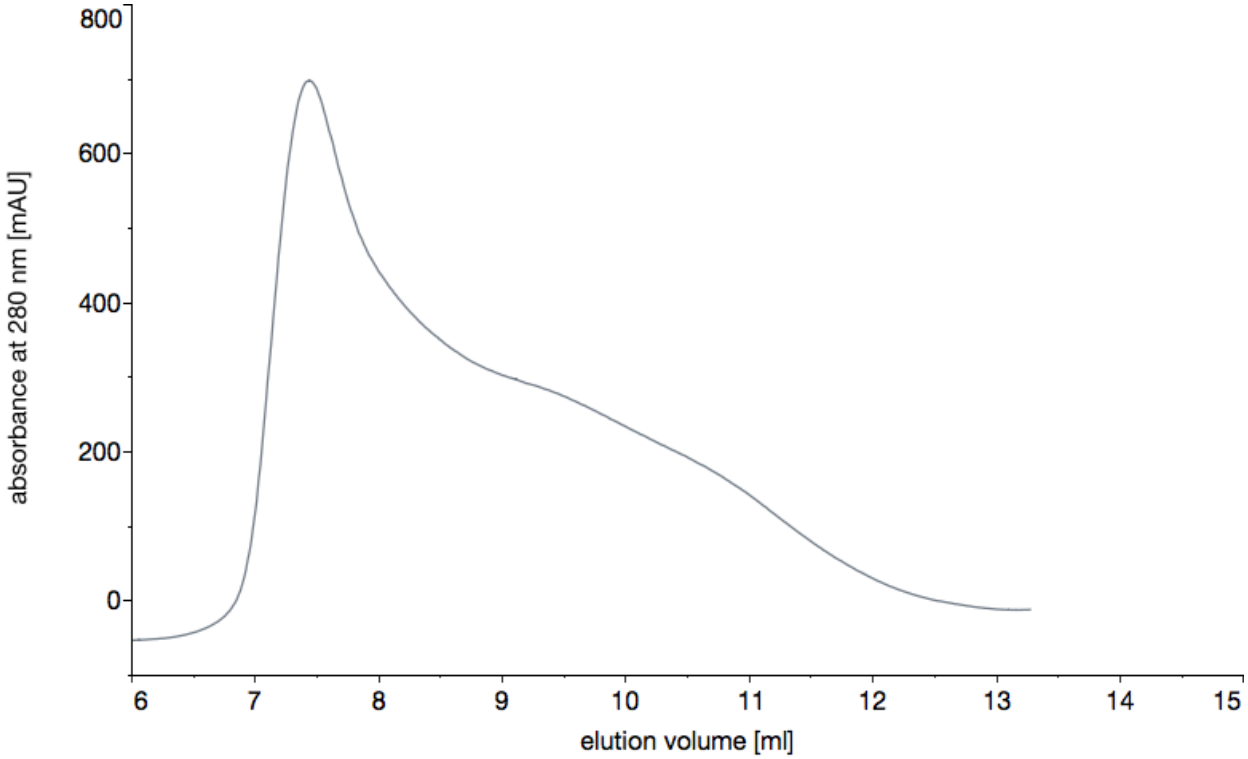


Figure 3.19 Purification conditions and elution profile of run 7.

Construct:	3	Beads:	TALON
Insect cell line:	High Five	Protease:	none
Amount of membrane (g):	0.97		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
2 mM imidazole pH 7	2 mM imidazole pH 7	10/20 mM imidazole pH 7	300 mM imidazole pH 7	0.1% DDM
1% DDM	0.1% DDM	0.1% DDM	0.1% DDM	0.1 mg/mL E.coli lipid extract
1 mg/mL iodoacetamide	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E.coli lipid extract	
protease inhibitor				
0.1 mg/mL E. coli lipid extract				

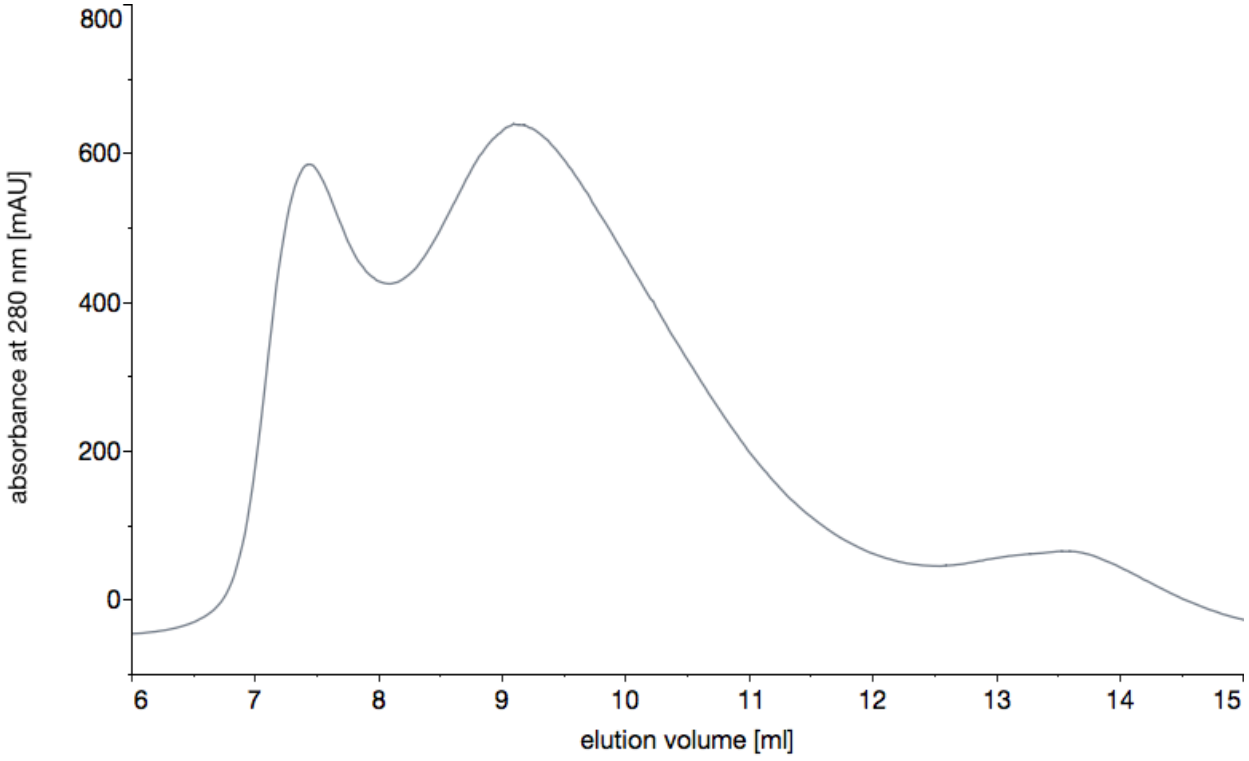


Figure 3.20 Purification conditions and elution profile of run 8.

Construct:	3	Beads:	TALON
Insect cell line:	High Five	Protease:	none
Amount of membrane (g):	1.1		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
2 mM imidazole pH 7	2 mM imidazole pH 7	10/20 mM imidazole pH 7	300 mM imidazole pH 7	0.1% DDM
1% DDM	0.1% DDM	0.1% DDM	0.1% DDM	0.20% CHS
1 mg/mL iodoacetamide	0.20% CHS	0.20% CHS	0.20% CHS	
protease inhibitor				
0.20% CHS				

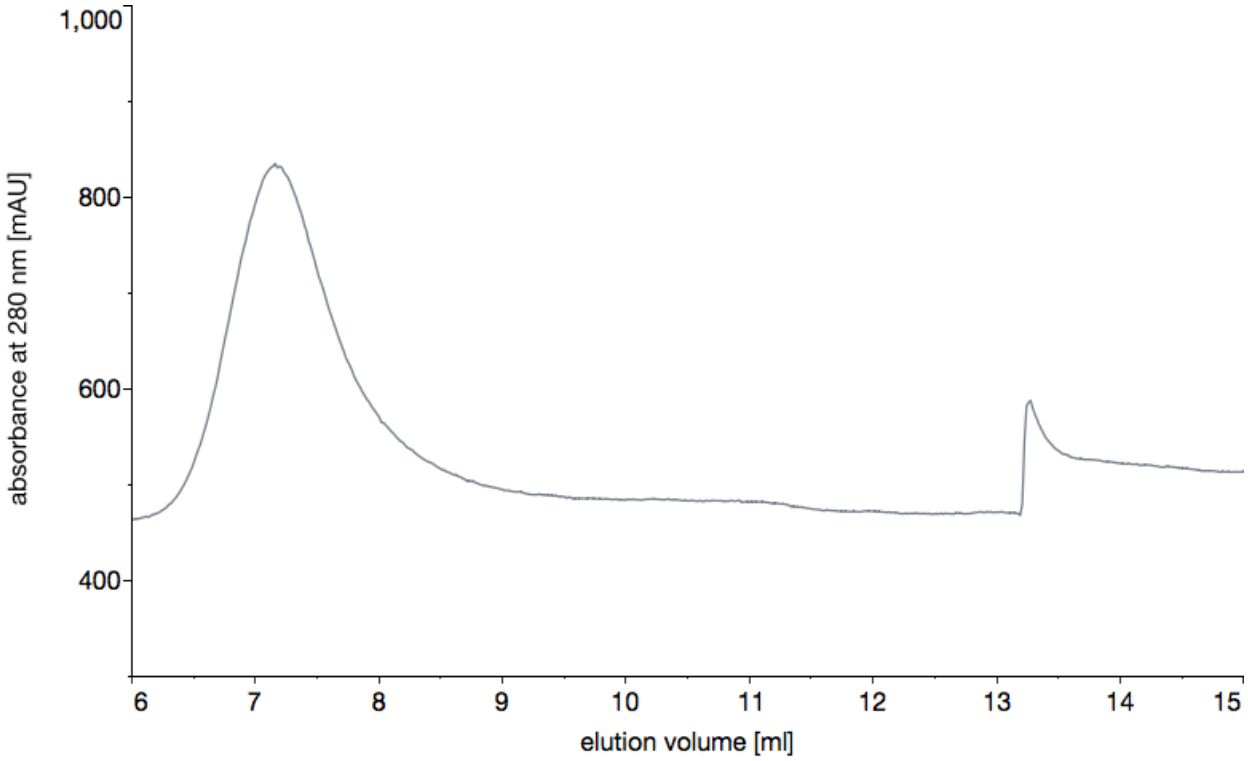


Figure 3.21 Purification conditions and elution profile of run 9.

Construct:	3	Beads:	TALON
Insect cell line:	High Five	Protease:	none
Amount of membrane (g):	0.85		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
2 mM imidazole pH 7	2 mM imidazole pH 7	10/20 mM imidazole pH 7	300 mM imidazole pH 7	0.1% DDM
1% DDM	0.1% DDM	0.1% DDM	0.1% DDM	
1 mg/mL iodoacetamide	0.05% CHS	0.05% CHS	0.05% CHS	
protease inhibitor	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	
0.05% CHS				
0.1 mg/mL E. coli lipid extract				

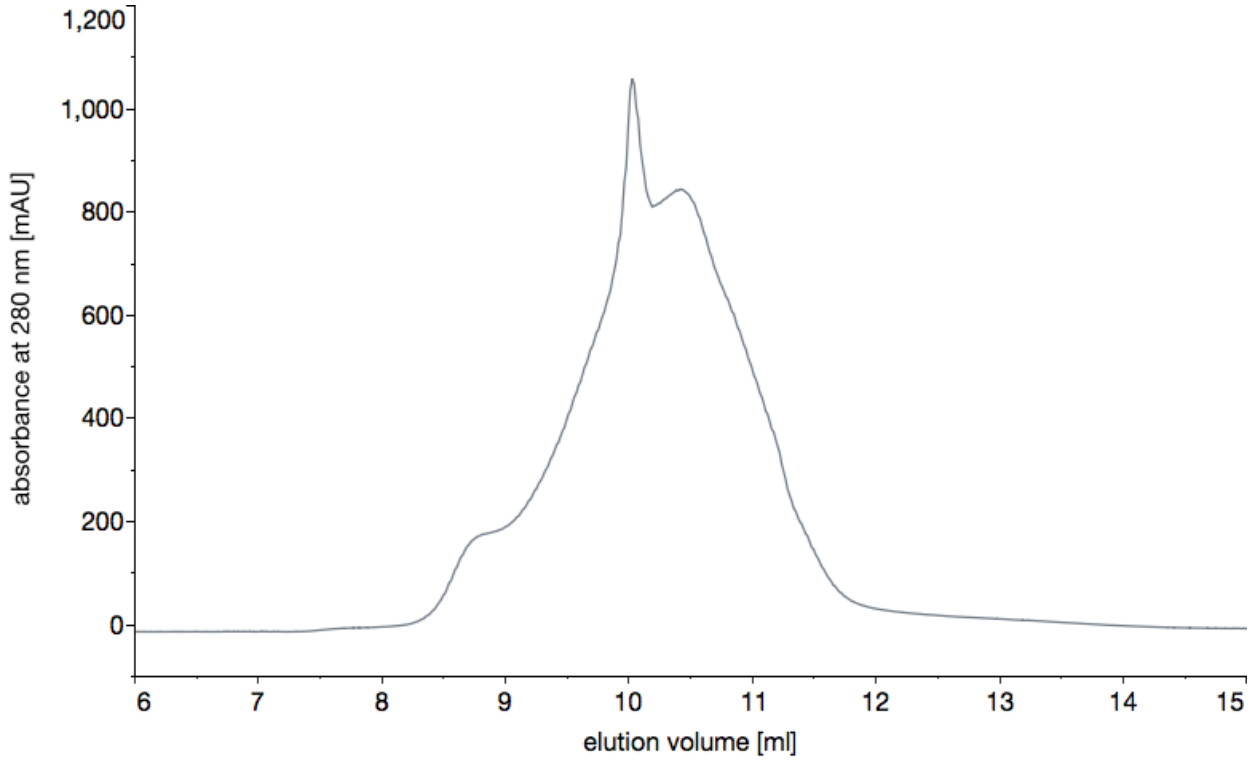
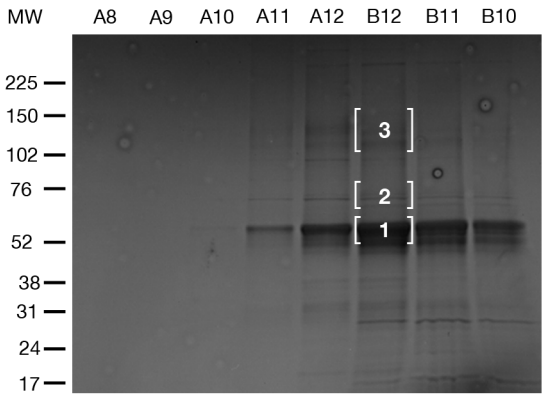


Figure 3.22 Purification conditions and elution profile of run 10.



Sample ID	UNIPROT ID	Number Unique	Peptide Count	% Sample	% Coverage	Protein MW	Species	Source	Protein Name
1	Q9HAS3	16	114	97%	20.8	76931	HUMAN	Target	Solute carrier family 28 member 3
1	Q2F6C3	2	3	3%	7	57745.4	BOMMO	Insect	Chaperonin subunit 6a zeta
2	Q9HAS3	12	35	41%	17.9	76931	HUMAN	Target	Solute carrier family 28 member 3
2	Q9U639	13	18	21%	23.8	71432.3	MANSE	Insect	Heat shock 70 kDa protein cognate 4
2	Q00704	9	12	14%	14.9	79075.9	NPVAC	Virus	Occlusion-derived virus envelope protein E66
2	A9LST1	7	7	8%	25.1	35944	SPOEX	Insect	Activated C kinase 1 receptor
2	P41473	4	4	5%	23	30567.3	NPVAC	Virus	Uncharacterized 30.6 kDa protein in IAP2-VLF1 intergenic region
2	Q0MUU6	4	4	5%	8	73138.6	TRINI	Insect	Heat shock cognate 70 protein
2	G6CI19	3	3	3%	2.7	166194.3	DANPL	Insect	Eukaryotic translation initiation factor 3 subunit A
2	D6WLZ3	2	3	3%	3.9	83204.3	TRICA	Insect	Putative uncharacterized protein
3	Q9HAS3	10	23	51%	18.2	76931	HUMAN	Target	Solute carrier family 28 member 3
3	A9LST1	9	9	20%	32.3	35944	SPOEX	Insect	Activated C kinase 1 receptor
3	P41473	4	6	13%	23	30567.3	NPVAC	Virus	Uncharacterized 30.6 kDa protein in IAP2-VLF1 intergenic region
3	G6D0K3	2	4	9%	3.1	142316.5	DANPL	Insect	Putative ubiquitin specific protease 7
3	A7LAG1	2	3	7%	3.1	112072.5	FHV	Virus	Protein A

Figure 3.23 Coomassie gel stain of run 10 fractions (top) and mass spectrometry output (bottom). Areas cut out of the gel are indicated on top of the gel. The peptides identified by mass spectrometry are organized by % abundance in each sample.

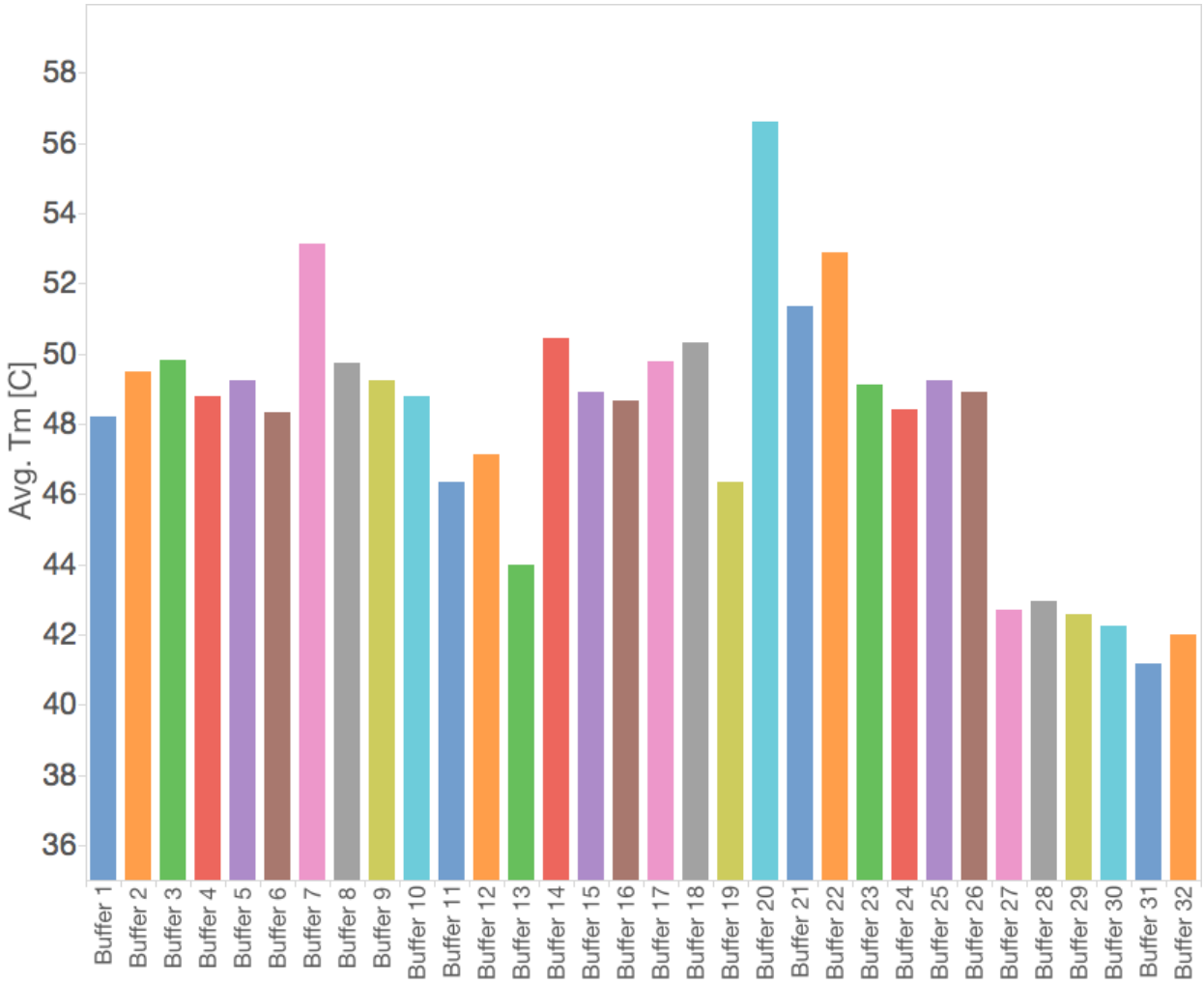


Figure 3.24 Summary of the thermostability assay results for run 10.

Construct:	3	Beads:	TALON
Insect cell line:	High Five	Protease:	none
Amount of membrane (g):	0.85		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
2 mM imidazole pH 7	2 mM imidazole pH 7	10/20 mM imidazole pH 7	300 mM imidazole pH 7	0.1% DDM
1% DDM	0.1% DDM	0.1% DDM	0.1% DDM	
1 mg/mL iodoacetamide	0.06 mg/mL POPC (3)	0.06 mg/mL POPC (3)	0.06 mg/mL POPC (3)	
protease inhibitor	0.02 mg/mL POPE (1)	0.02 mg/mL POPE (1)	0.02 mg/mL POPE (1)	
0.06 mg/mL POPC (3)	0.02 mg/mL POPG (1)	0.02 mg/mL POPG (1)	0.02 mg/mL POPG (1)	
0.02 mg/mL POPE (1)				
0.02 mg/mL POPG (1)				

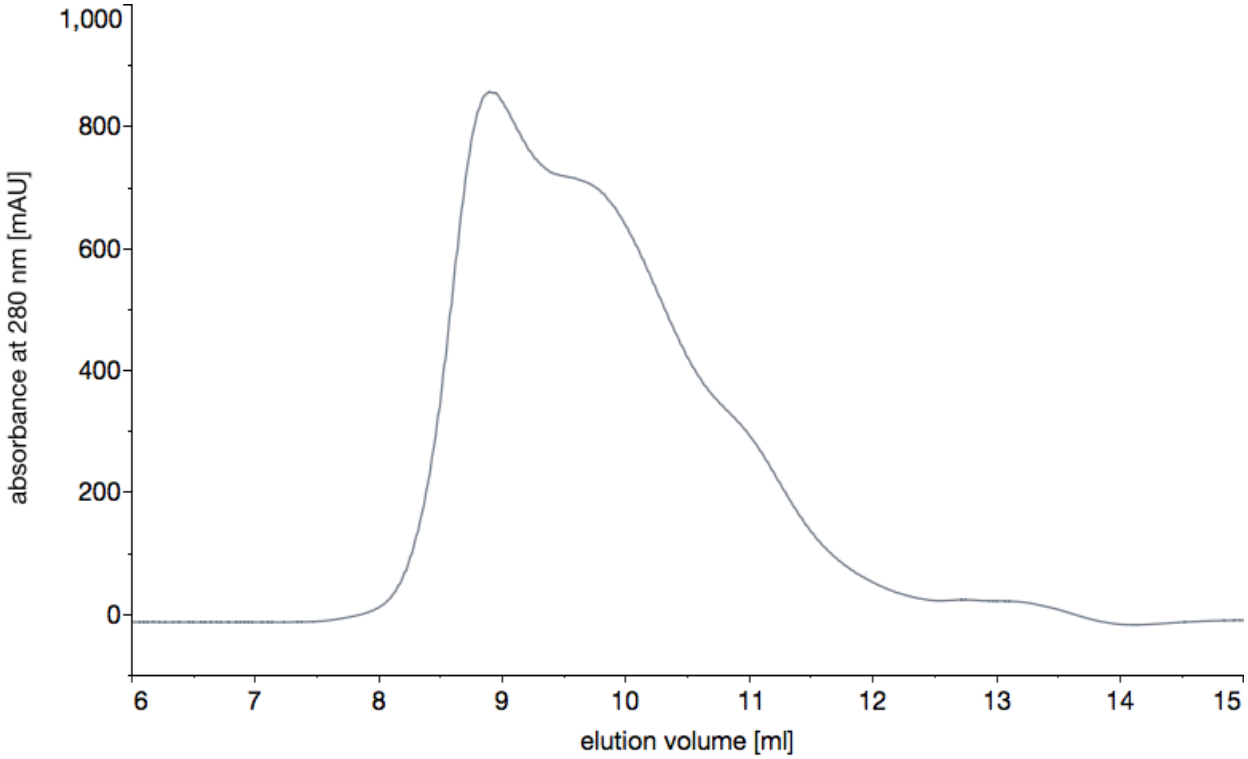


Figure 3.25 Purification conditions and elution profile of run 11.

Construct:	3	Beads:	TALON
Insect cell line:	High Five	Protease:	none
Amount of membrane (g):	0.85		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
2 mM imidazole pH 7	2 mM imidazole pH 7	10/20 mM imidazole pH 7	300 mM imidazole pH 7	0.1% DDM
1% DDM	0.1% DDM	0.1% DDM	0.1% DDM	0.1 mg/mL liver lipids extract
1 mg/mL iodoacetamide	0.1 mg/mL liver lipids extract	0.1 mg/mL liver lipids extract	0.1 mg/mL liver lipids extract	
protease inhibitor				
0.1 mg/mL liver lipids extract				

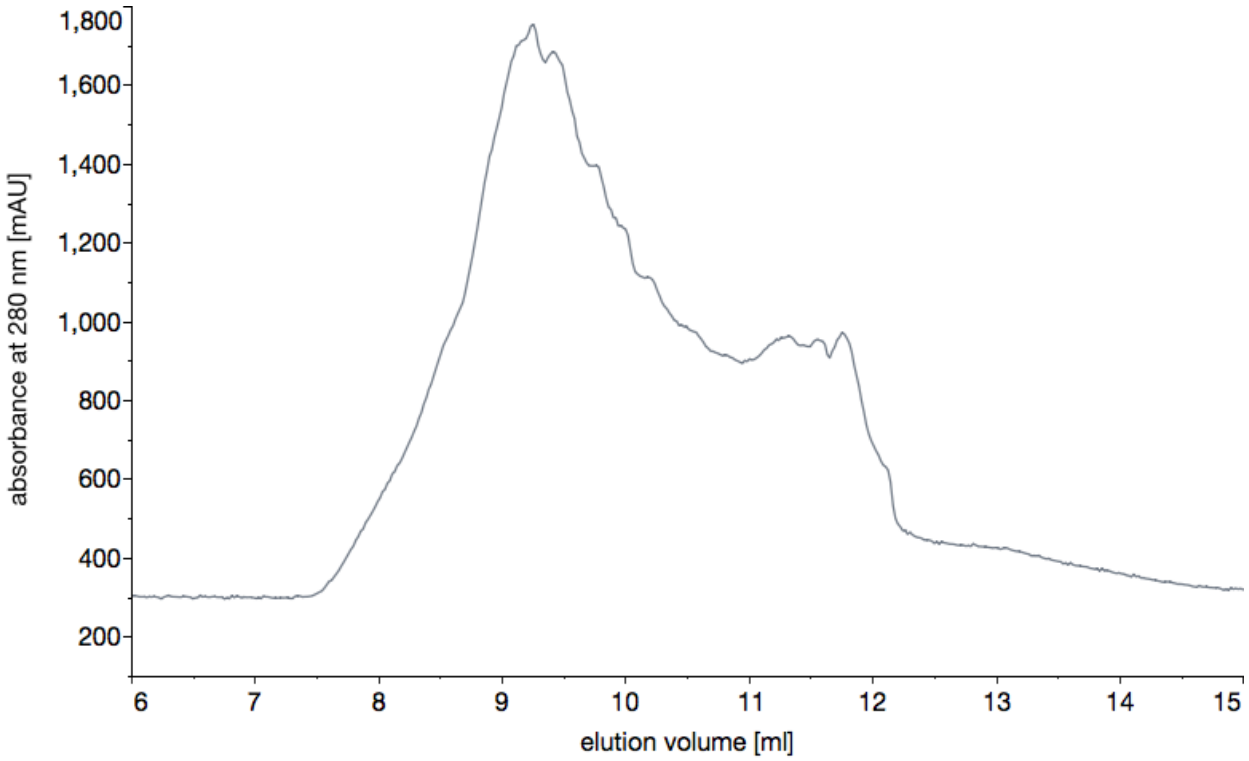


Figure 3.26 Purification conditions and elution profile of run 12.

Construct:	3	Beads:	TALON
Insect cell line:	High Five	Protease:	none
Amount of membrane (g):	1		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7
1 M NaCl	1 M NaCl	1 M NaCl	1 M NaCl	1 M NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
2 mM imidazole pH 7	2 mM imidazole pH 7	10/20 mM imidazole pH 7	300 mM imidazole pH 7	0.1% DDM
1% DDM	0.1% DDM	0.1% DDM	0.1% DDM	
1 mg/mL iodoacetamide	protease inhibitor	protease inhibitor	protease inhibitor	
protease inhibitor				

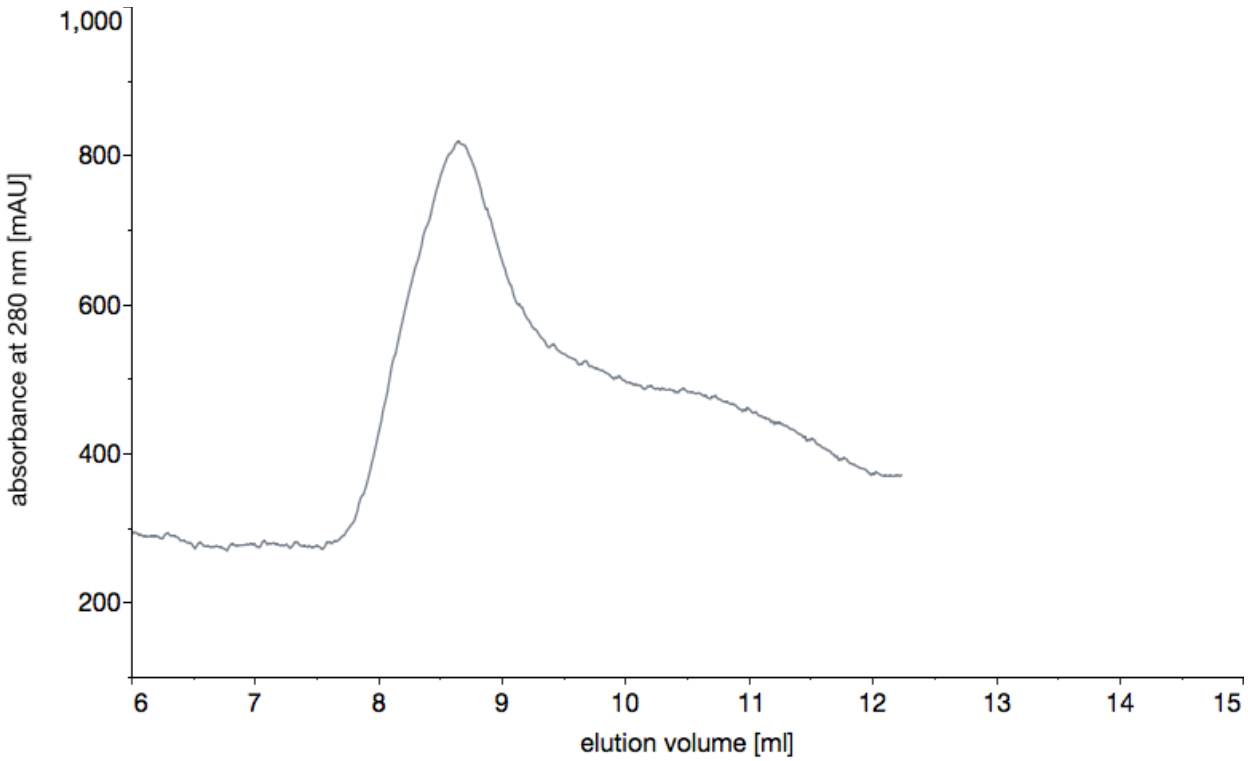


Figure 3.27 Purification conditions and elution profile of run 13.

Construct:	3	Beads:	TALON
Insect cell line:	High Five	Protease:	none
Amount of membrane (g):	0.85		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM MES pH 6	50 mM MES pH 6	50 mM MES pH 6	50 mM MES pH 6	50 mM MES pH 6
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
2 mM imidazole pH 6	2 mM imidazole pH 6	10/20 mM imidazole pH 6	300 mM imidazole pH 6	0.1% DDM
1% DDM	0.1% DDM	0.1% DDM	0.1% DDM	
1 mg/mL iodoacetamide	0.05% CHS	0.05% CHS	0.05% CHS	
protease inhibitor	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	
0.05% CHS				
0.1 mg/mL E. coli lipid extract				

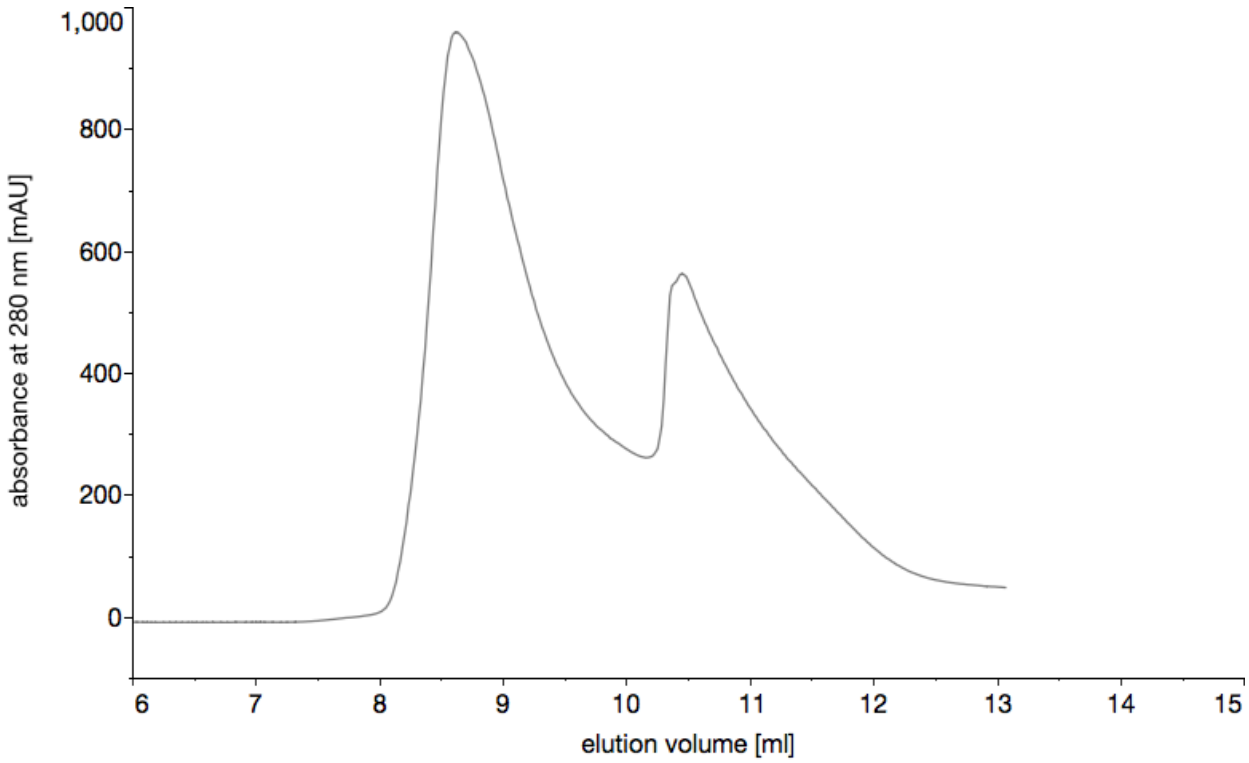


Figure 3.28 Purification conditions and elution profile of run 14.

Construct:	3	Beads:	TALON
Insect cell line:	High Five	Protease:	none
Amount of membrane (g):	1		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7
150 mM NaCl	150 mM NaCl	150 mM NaCl	150 mM NaCl	150 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
2 mM imidazole pH 7	2 mM imidazole pH 7	10/20 mM imidazole pH 7	300 mM imidazole pH 7	0.1% DDM
1% DDM	0.1% DDM	0.1% DDM	0.1% DDM	
1 mg/mL iodoacetamide	0.05% CHS	0.05% CHS	0.05% CHS	
protease inhibitor	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	
0.05% CHS				
0.1 mg/mL E. coli lipid extract				

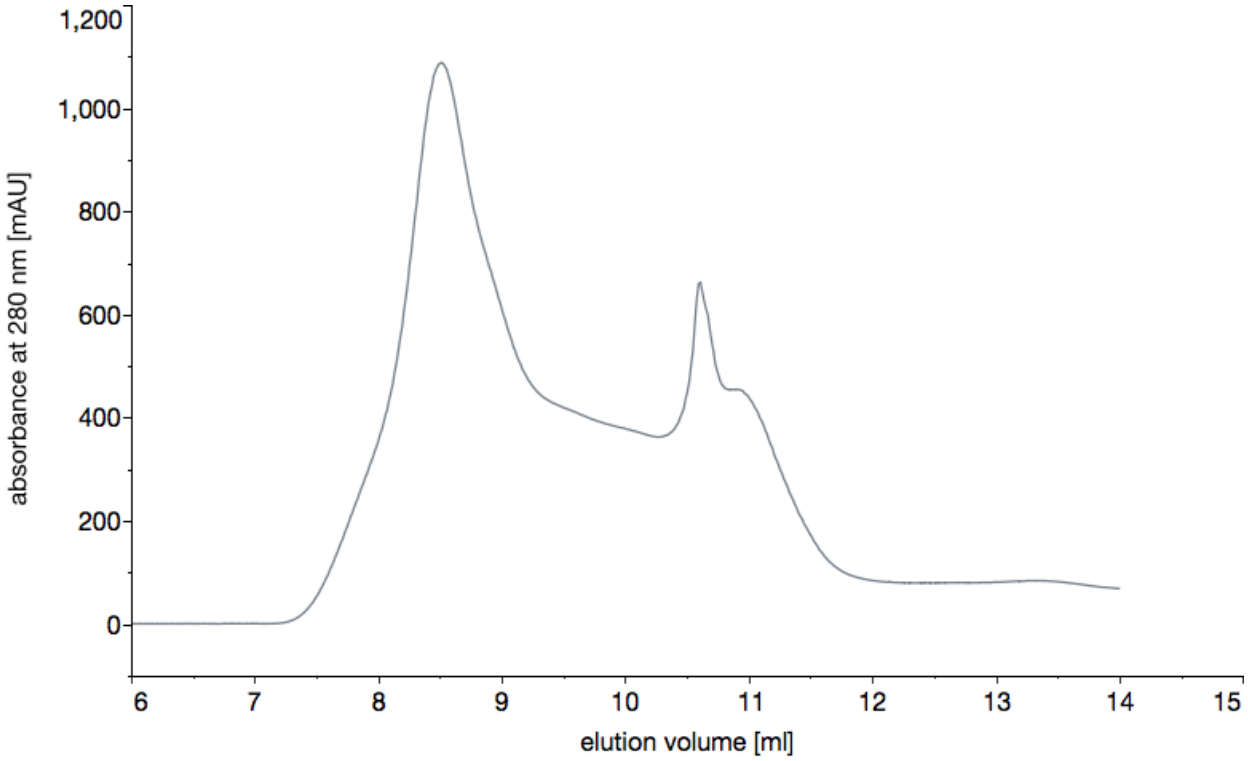


Figure 3.29 Purification conditions and elution profile of run 15.

Construct:	3	Beads:	TALON
Insect cell line:	High Five	Protease:	none
Amount of membrane (g):	1		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
2 mM imidazole pH 7	2 mM imidazole pH 7	10/20 mM imidazole pH 7	300 mM imidazole pH 7	0.15% OGNG
1% DDM	0.15% OGNG	0.15% OGNG	0.15% OGNG	0.05% CHS
1 mg/mL iodoacetamide	0.05% CHS	0.05% CHS	0.05% CHS	0.1 mg/mL E. coli lipid extract
protease inhibitor	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	
0.05% CHS				
0.1 mg/mL E. coli lipid extract				

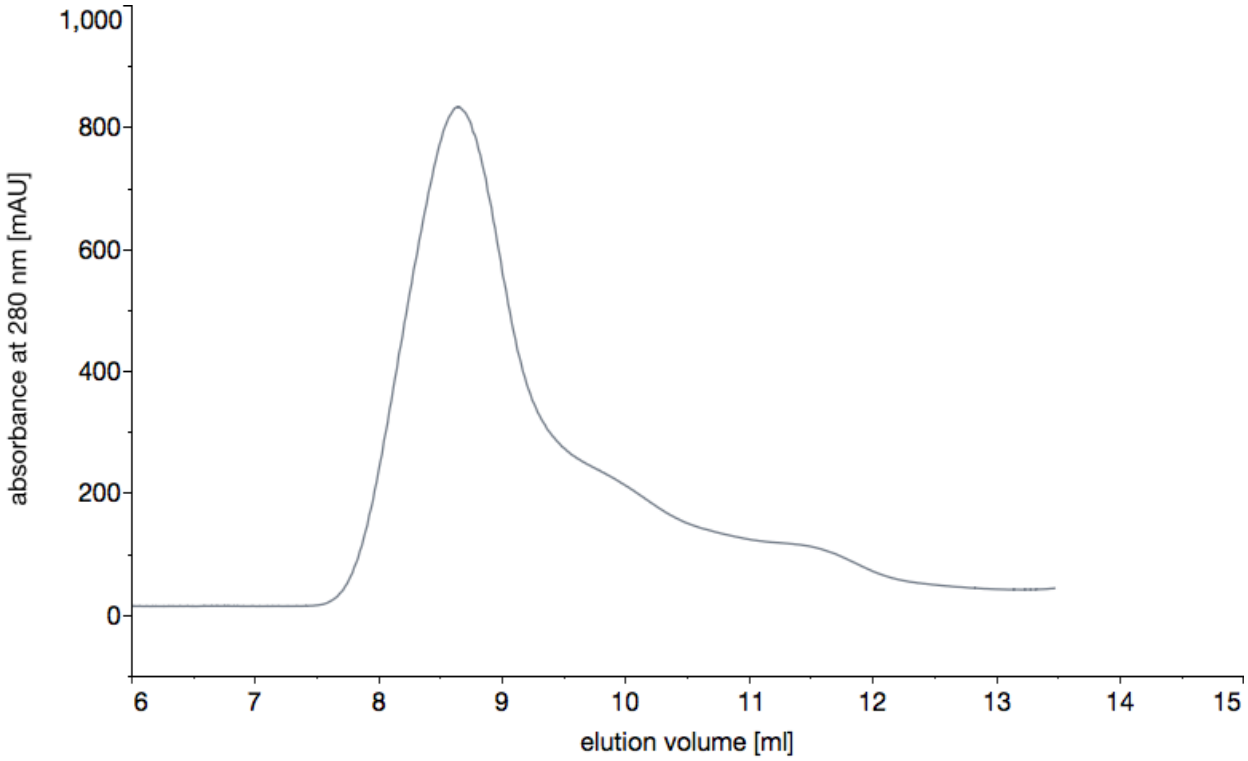


Figure 3.30 Purification conditions and elution profile of run 16.

Construct:	3	Beads:	TALON
Insect cell line:	High Five	Protease:	none
Amount of membrane (g):	1		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
2 mM imidazole pH 7	2 mM imidazole pH 7	10/20 mM imidazole pH 7	300 mM imidazole pH 7	0.05% LMNG
1% DDM	0.05% LMNG	0.05% LMNG	0.05% LMNG	0.05% CHS
1 mg/mL iodoacetamide	0.05% CHS	0.05% CHS	0.05% CHS	0.1 mg/mL E. coli lipid extract
protease inhibitor	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	
0.05% CHS				
0.1 mg/mL E. coli lipid extract				

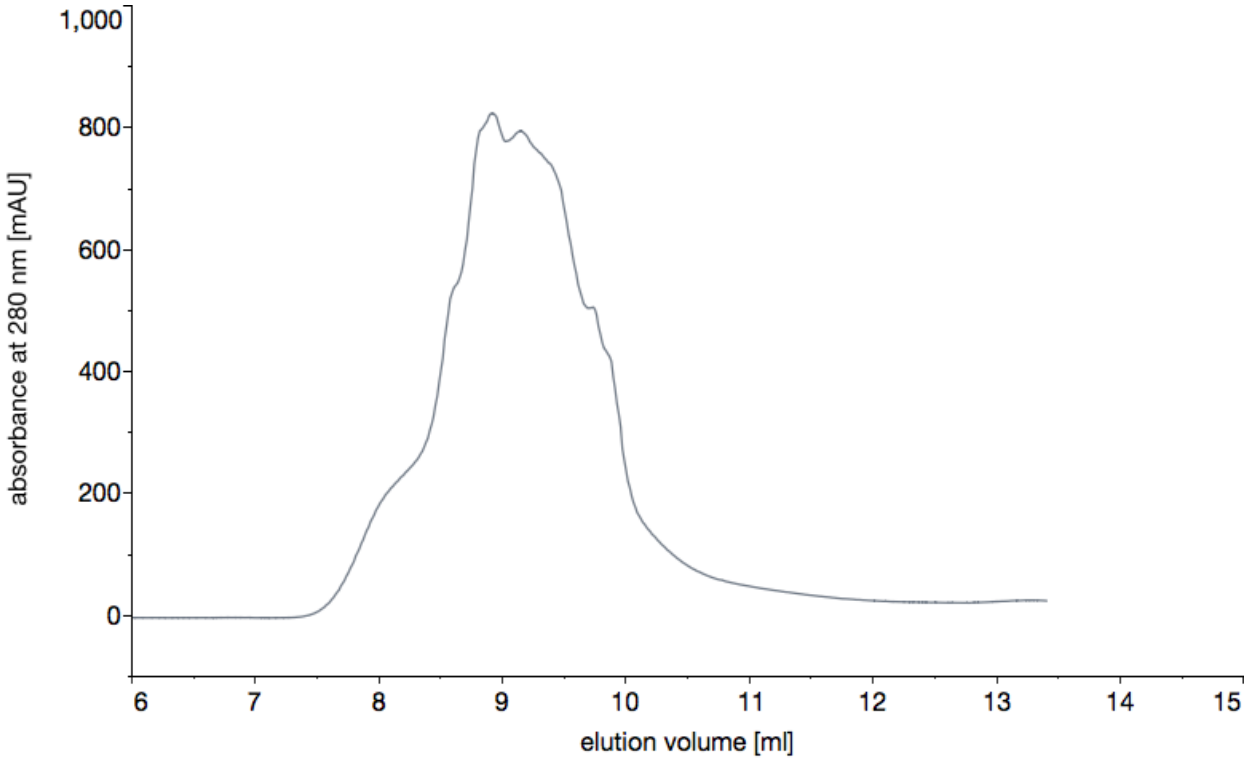


Figure 3.31 Purification conditions and elution profile of run 17.

Construct:	3	Beads:	TALON
Insect cell line:	High Five	Protease:	HRV 3C
Amount of membrane (g):	1.7		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
2 mM imidazole pH 7	2 mM imidazole pH 7	10/20 mM imidazole pH 7	300 mM imidazole pH 7	0.1% DDM
1% DDM	0.1% DDM	0.1% DDM	0.1% DDM	
1 mg/mL iodoacetamide	0.05% CHS	0.05% CHS	0.05% CHS	
protease inhibitor	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	
0.05% CHS				
0.1 mg/mL E. coli lipid extract				

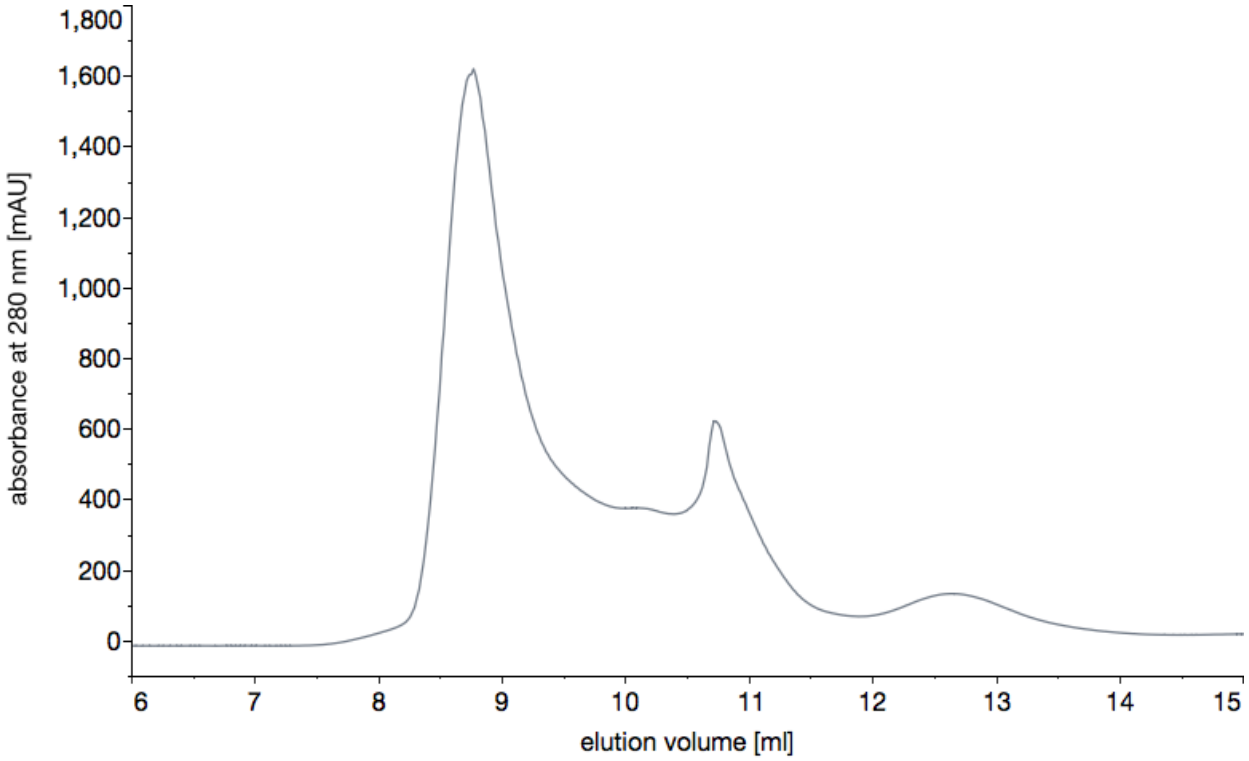


Figure 3.32 Purification conditions and elution profile of run 18.

Construct:	3	Beads:	TALON
Insect cell line:	High Five	Protease:	HRV 3C
Amount of membrane (g):	1		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7
200 mM NaSO4	200 mM NaSO4	200 mM NaSO4	200 mM NaSO4	200 mM NaSO4
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
2 mM imidazole pH 7	2 mM imidazole pH 7	10/20 mM imidazole pH 7	300 mM imidazole pH 7	0.1% DDM
1% DDM	0.1% DDM	0.1% DDM	0.1% DDM	0.1 mg/mL E. coli lipid extract
1 mg/mL iodoacetamide	0.05% CHS	0.05% CHS	0.05% CHS	
protease inhibitor	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	
0.05% CHS				
0.1 mg/mL E. coli lipid extract				

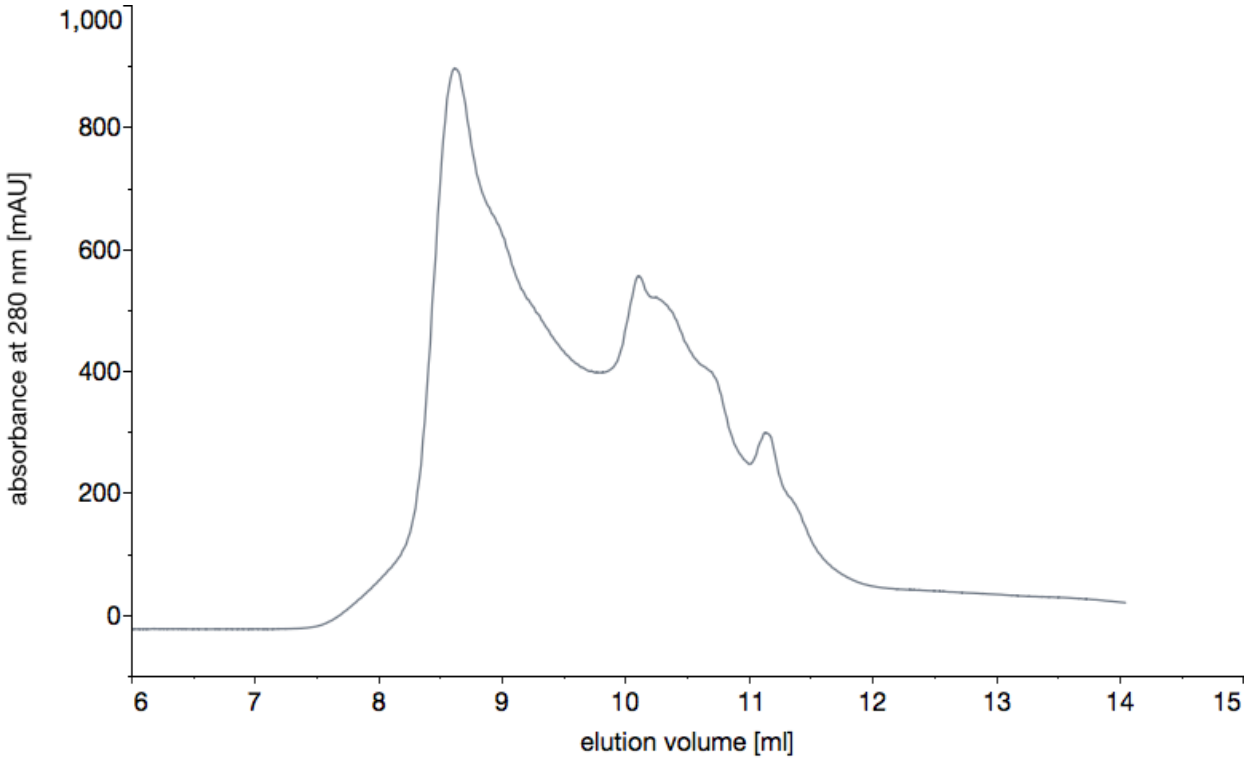


Figure 3.33 Purification conditions and elution profile of run 19.

Construct:	3	Beads:	TALON
Insect cell line:	High Five	Protease:	HRV 3C
Amount of membrane (g):	1		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM MES pH 6	50 mM MES pH 6	50 mM MES pH 6	50 mM MES pH 6	50 mM MES pH 6
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
20% glycerol	20% glycerol	20% glycerol	20% glycerol	20% glycerol
2 mM imidazole pH 6	2 mM imidazole pH 6	10/20 mM imidazole pH 6	300 mM imidazole pH 6	0.1% DDM
1% DDM	0.1% DDM	0.1% DDM	0.1% DDM	0.1 mg/mL E. coli lipid extract
1 mg/mL iodoacetamide	0.05% CHS	0.05% CHS	0.05% CHS	
protease inhibitor	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	
0.05% CHS				
0.1 mg/mL E. coli lipid extract				

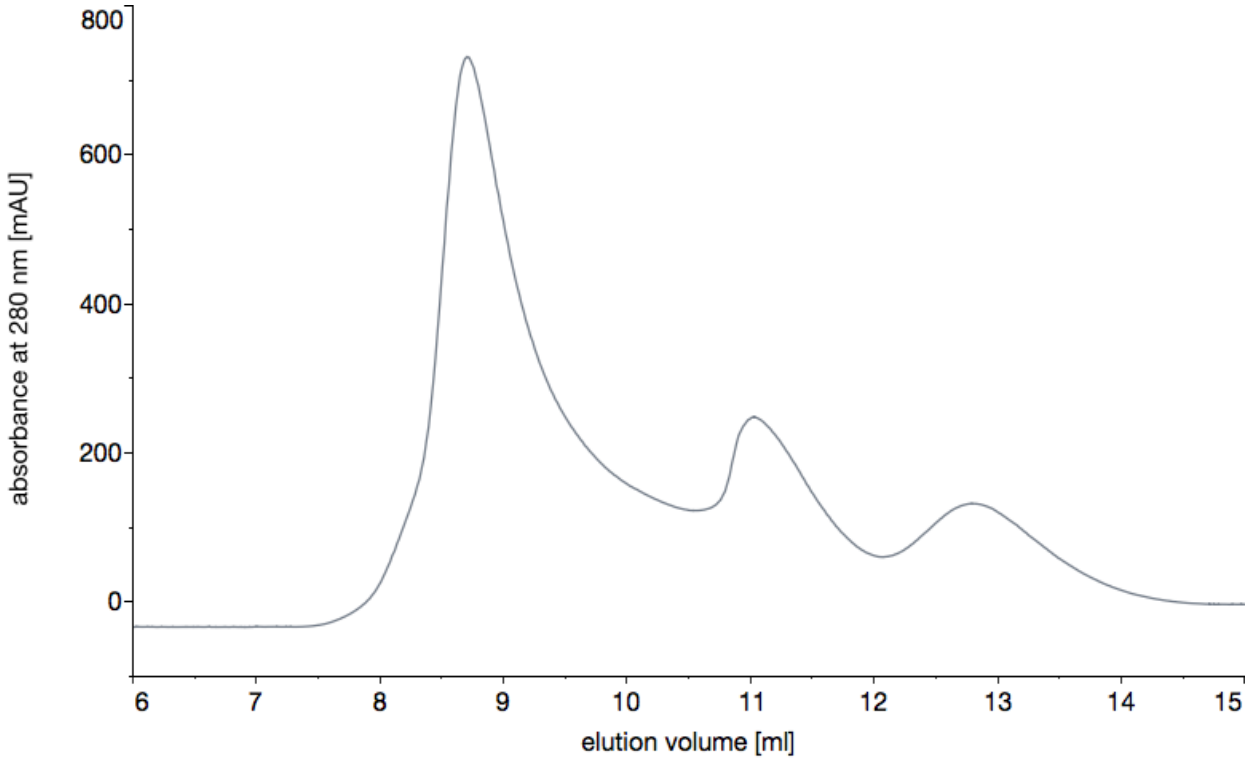


Figure 3.34 Purification conditions and elution profile of run 20.

Construct:	3	Beads:	TALON
Insect cell line:	High Five	Protease:	none
Amount of membrane (g):	1		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
2 mM imidazole pH 7	2 mM imidazole pH 7	10/20 mM imidazole pH 7	300 mM imidazole pH 7	0.1% DDM
1% DDM	0.1% DDM	0.1% DDM	0.1% DDM	
1 mg/mL iodoacetamide	0.05% CHS	0.05% CHS	0.05% CHS	
protease inhibitor	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	
0.05% CHS				
0.1 mg/mL E. coli lipid extract				

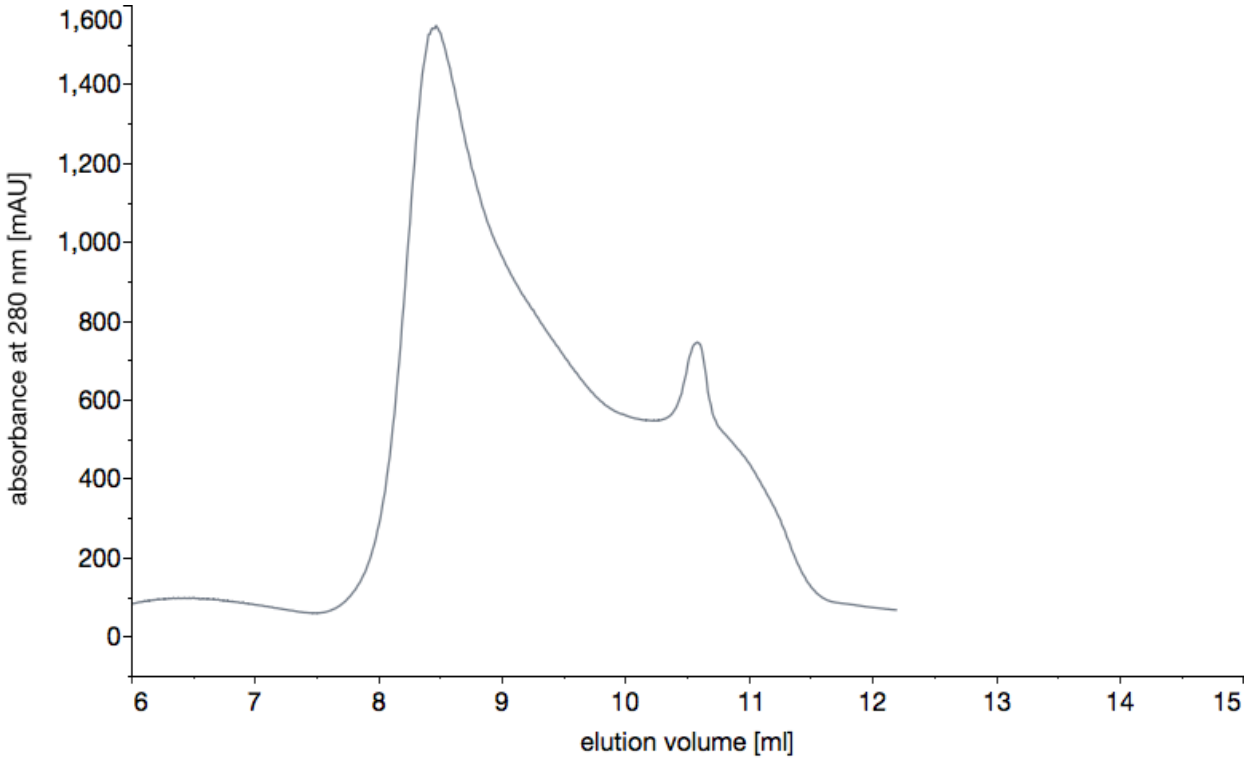


Figure 3.35 Purification conditions and elution profile of run 21.

Construct 2 To start, construct 2 trials were first attempted with CHS and *E. coli* polar lipid extract, the best conditions from construct 3 (run 22, **Figure 3.36**). Although noisy, the SEC profile showed promise, exhibiting a large peak around the correct elution volume. In our attempt to get rid of the void peak, we tried to cut the sample off the column with HRV 3C protease, but with limited success (run 23, **Figure 3.37**). A second thermostability screen was performed on the pooled fractions. The three most stabilizing buffers were buffer 4 (20% glycerol), buffer 15 (200 mM $NaSO_4$), and buffer 22 (50 mM MES pH 6). Two of those conditions, 20% glycerol with 50 mM MES pH 6, were incorporated into the next trial (run 24, **Figure 3.39**). Surprisingly it had the opposite effect, and actually led to a worse SEC profile. Since we were unable to replicate the initial result with just CHS and *E. coli* polar lipid extract (run 25, **Figure 3.40**), we moved on to the next construct.

Construct:	2	Beads:	TALON
Insect cell line:	High Five	Protease:	none
Amount of membrane (g):	1		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
2 mM imidazole pH 7	2 mM imidazole pH 7	10/20 mM imidazole pH 7	300 mM imidazole pH 7	0.1% DDM
1% DDM	0.1% DDM	0.1% DDM	0.1% DDM	
1 mg/mL iodoacetamide	0.05% CHS	0.05% CHS	0.05% CHS	
protease inhibitor	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	
0.05% CHS				
0.1 mg/mL E. coli lipid extract				

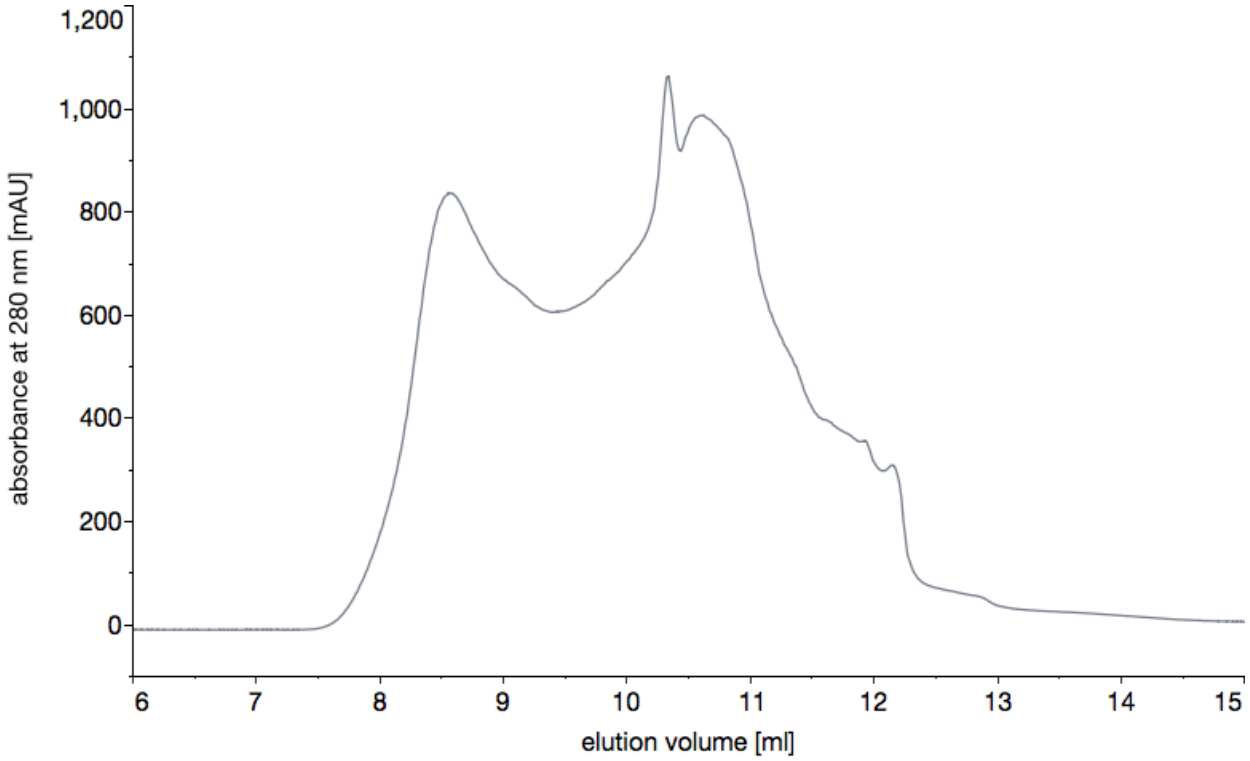


Figure 3.36 Purification conditions and elution profile of run 22.

Construct:	2	Beads:	TALON
Insect cell line:	High Five	Protease:	HRV 3C
Amount of membrane (g):	2.13		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
2 mM imidazole pH 7	2 mM imidazole pH 7	10/20 mM imidazole pH 7	300 mM imidazole pH 7	0.1% DDM
1% DDM	0.1% DDM	0.1% DDM	0.1% DDM	
1 mg/mL iodoacetamide	0.05% CHS	0.05% CHS	0.05% CHS	
protease inhibitor	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	
0.05% CHS				
0.1 mg/mL E. coli lipid extract				

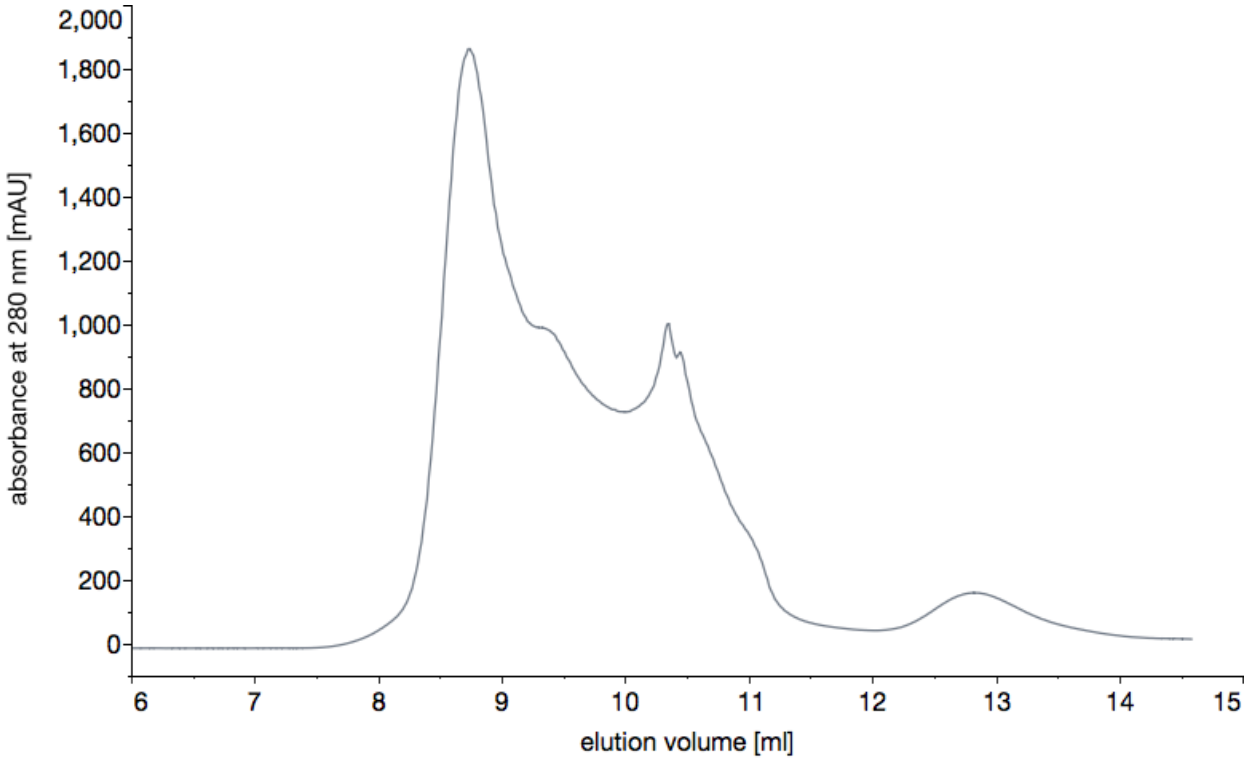


Figure 3.37 Purification conditions and elution profile of run 23.

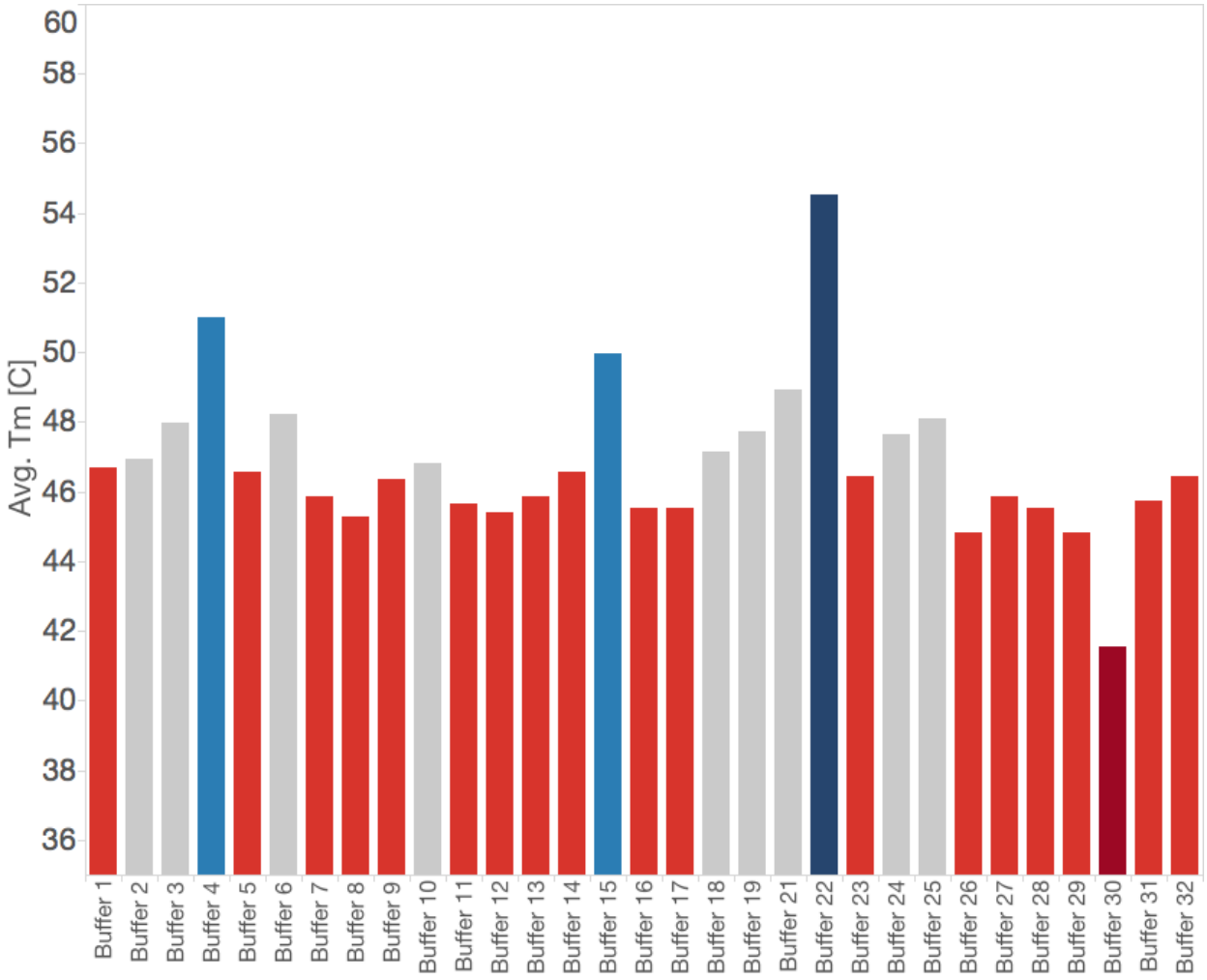


Figure 3.38 Summary of the thermostability assay results for run 23.

Construct:	2	Beads:	TALON
Insect cell line:	High Five	Protease:	HRV 3C
Amount of membrane (g):	1		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM MES pH 6	50 mM MES pH 6	50 mM MES pH 6	50 mM MES pH 6	50 mM MES pH 6
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
20% glycerol	20% glycerol	20% glycerol	20% glycerol	20% glycerol
2 mM imidazole pH 6	2 mM imidazole pH 6	10/20 mM imidazole pH 6	300 mM imidazole pH 6	0.1% DDM
1% DDM	0.1% DDM	0.1% DDM	0.1% DDM	0.1 mg/mL E. coli lipid extract
1 mg/mL iodoacetamide	0.05% CHS	0.05% CHS	0.05% CHS	
protease inhibitor	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	
0.05% CHS				
0.1 mg/mL E. coli lipid extract				

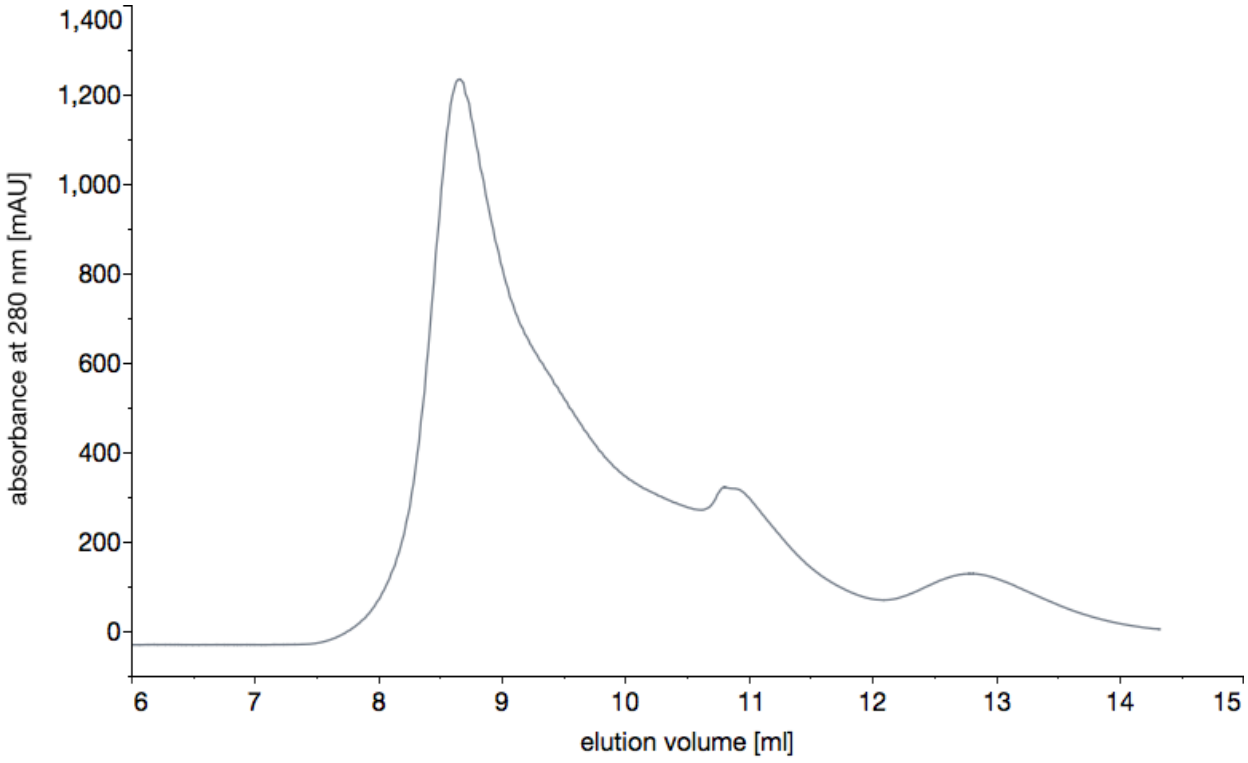


Figure 3.39 Purification conditions and elution profile of run 24.

Construct:	2	Beads:	TALON
Insect cell line:	High Five	Protease:	none
Amount of membrane (g):	1		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
2 mM imidazole pH 7	2 mM imidazole pH 7	10/20 mM imidazole pH 7	300 mM imidazole pH 7	0.1% DDM
1% DDM	0.1% DDM	0.1% DDM	0.1% DDM	
1 mg/mL iodoacetamide	0.05% CHS	0.05% CHS	0.05% CHS	
protease inhibitor	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	
0.05% CHS				
0.1 mg/mL E. coli lipid extract				

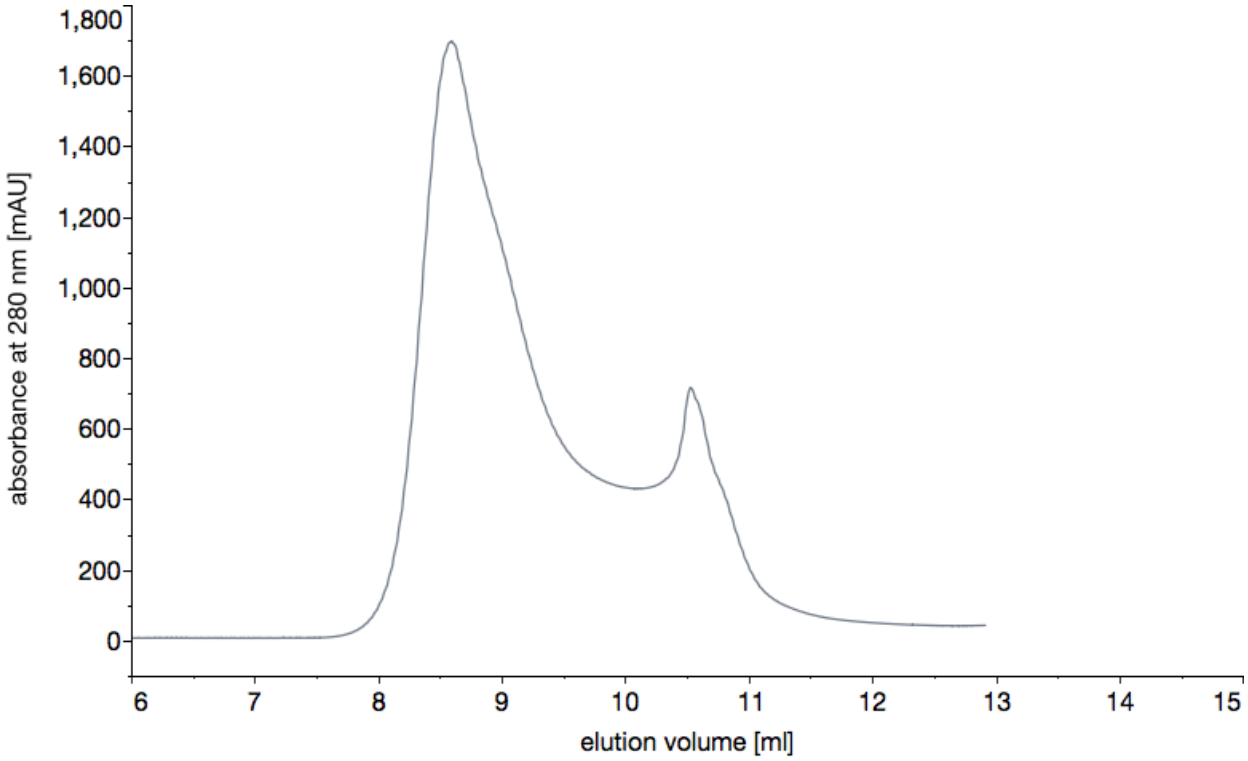


Figure 3.40 Purification conditions and elution profile of run 25.

Construct 4 To eliminate the potential heterogeneity stemming from glycosylation, in construct 4 we mutated the four putative glycosylated asparagine residues, N630, N636, N664, N678, to glutamine residues. It was also the first construct containing the maltose binding protein (MBP) as a fusion tag, which has been shown to increase protein solubility and stability.[86,87] As before, we treated the conditions containing CHS and *E. coli* polar lipid extract as the baseline. First, we tried a purification that did not contain CHS nor *E. coli* polar lipid extract (run 26, **Figure 3.41**). We eluted the full protein product off the TALON resin. The void peak appeared to shift to the right, but the vast majority of the sample still aggregated. Gels were in agreement with the predicted size estimates of approximately 65 kDa hCNT3 monomer and 42 kDa MBP, for a 107 kDa combined weight (**Figure 3.42**). Proetase digestion overnight efficiently cleaved the protein product, and encouraged us to cleave the next trial off the TALON resin, in addition to adding CHS and *E. coli* polar lipid extract (run 27, **Figure 3.43**). A second peak emerged, but it was unsymmetrical.

For the next round, we tried using amylose resin for the first time. Instead of relying on affinity towards TALON, the MBP tag was used for protein immobilization. Introduction of CHAPS, a zwitterionic detergent, did not lead to any improvement over DDM (run 28, **Figure 3.44**) Although a fresh membrane prep of construct 4 in Sf9 insect cells produced a significant second peak at the proper elution volume, the purification efficiency has significantly decreased when using the amylose resin (run 29, **Figure 3.45**). The same fresh membrane prep from Sf9 cells with TALON resin resulted in a clean SEC profile, although the void peak was still present (run 30, **Figure 3.46**). The gel indicated that the sample was quite pure (**Figure 3.47**). Brij-35, another nonionic detergent, was tried with negative results (run 31, **Figure 3.48**), before we attempted to repeat run 30 with more stringent washes (run 32, **Figure 3.49**). Since the SEC profile did not exhibit the same characteristics as before, we abandoned this construct.

Construct:	4	Beads:	TALON
Insect cell line:	High Five	Protease:	none
Amount of membrane (g):	1		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
2 mM imidazole pH 7	2 mM imidazole pH 7	10/20 mM imidazole pH 7	300 mM imidazole pH 7	0.1% DDM
1% DDM	0.1% DDM	0.1% DDM	0.1% DDM	
1 mg/mL iodoacetamide				
protease inhibitor				

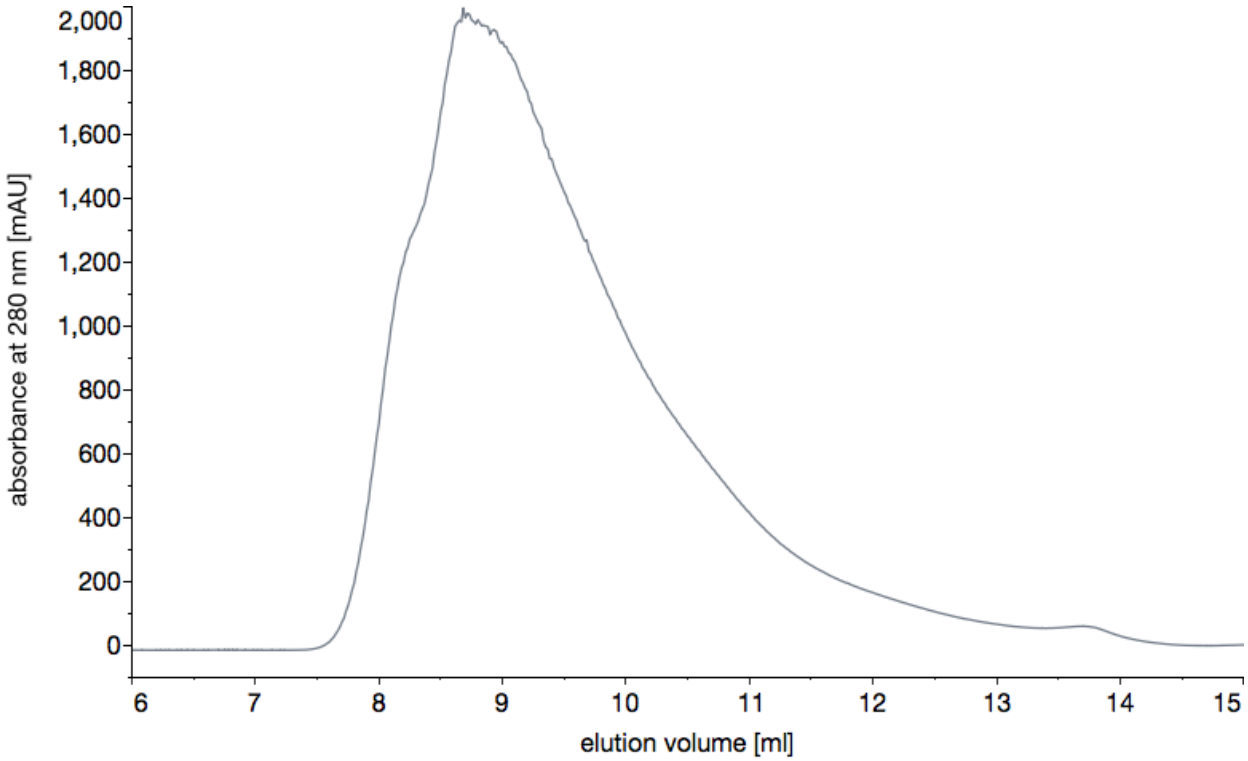


Figure 3.41 Purification conditions and elution profile of run 26.

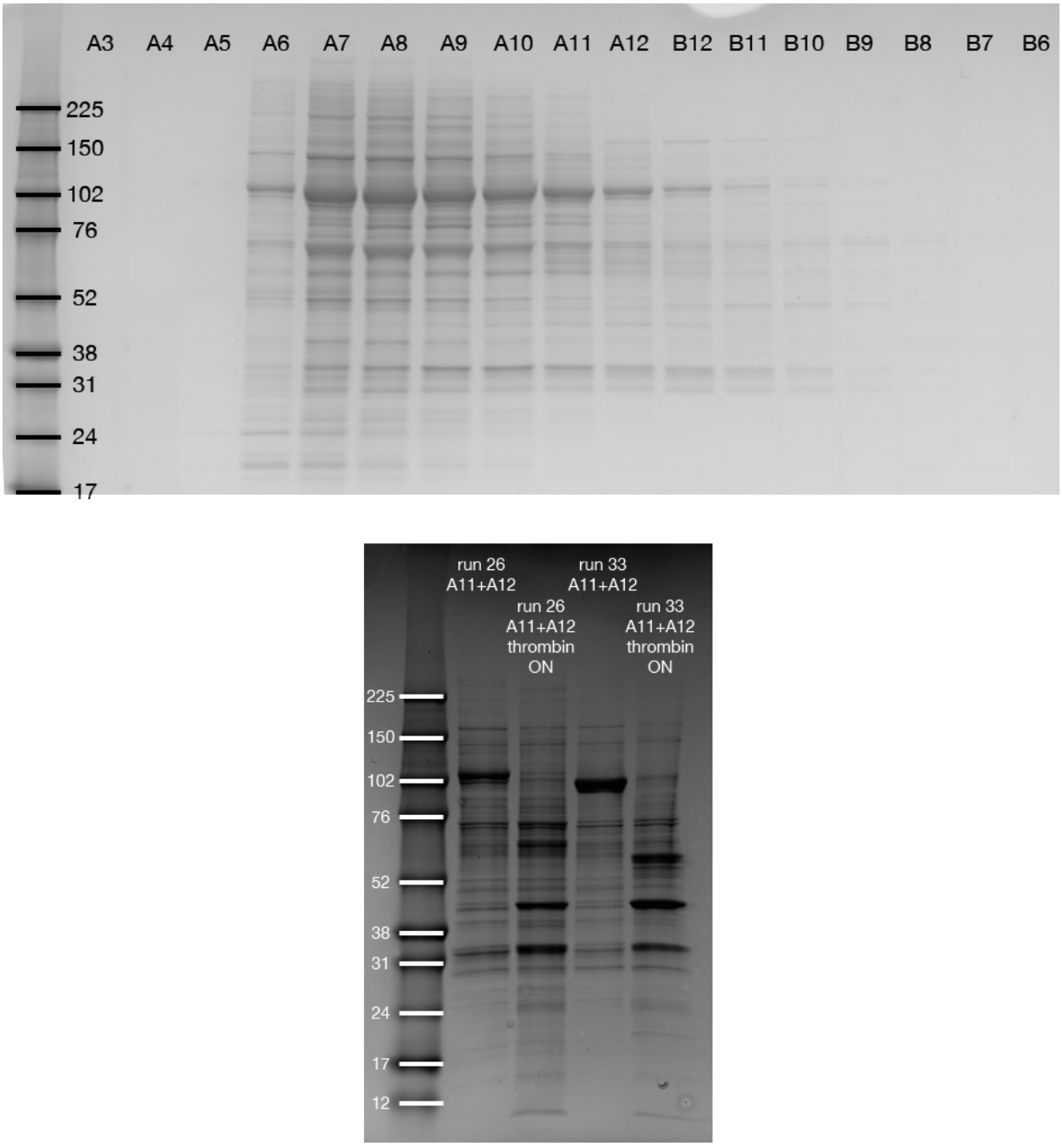


Figure 3.42 Coomassie gel stain of fractions from run 26 (top) and the effect of thrombin protease on run 26 and run 33 fractions (bottom).

Construct:	4	Beads:	TALON
Insect cell line:	High Five	Protease:	thrombin
Amount of membrane (g):	1		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
2 mM imidazole pH 7	2 mM imidazole pH 7	10/20 mM imidazole pH 7	300 mM imidazole pH 7	0.1% DDM
1% DDM	0.1% DDM	0.1% DDM	0.1% DDM	
1 mg/mL iodoacetamide	0.05% CHS	0.05% CHS	0.05% CHS	
protease inhibitor	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	
0.05% CHS				
0.1 mg/mL E. coli lipid extract				

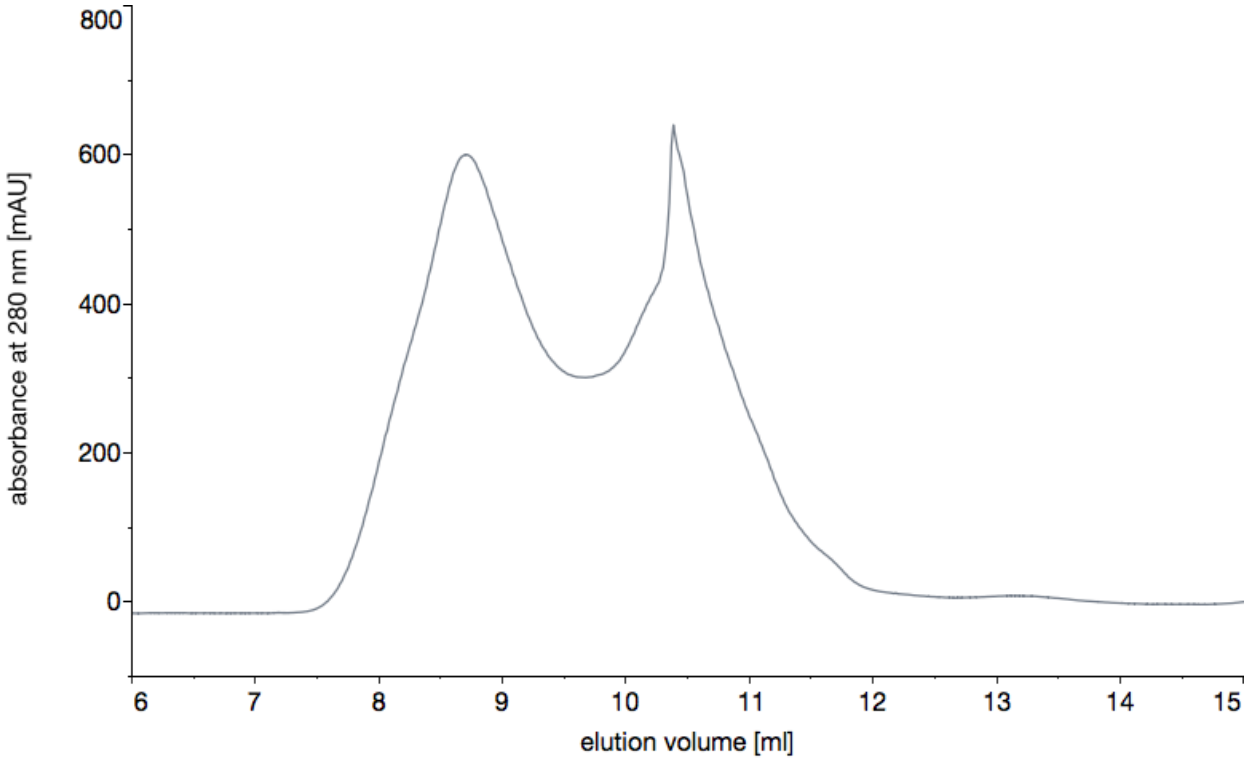


Figure 3.43 Purification conditions and elution profile of run 27.

Construct:	4	Beads:	amylose
Insect cell line:	High Five	Protease:	thrombin
Amount of membrane (g):	1.2		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
2 mM imidazole pH 7	2 mM imidazole pH 7	10/20 mM imidazole pH 7	300 mM imidazole pH 7	1.0% CHAPS
1% DDM	1.0% CHAPS	1.0% CHAPS	1.0% CHAPS	
1 mg/mL iodoacetamide	0.05% CHS	0.05% CHS	0.05% CHS	
protease inhibitor	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	
0.05% CHS				
0.1 mg/mL E. coli lipid extract				

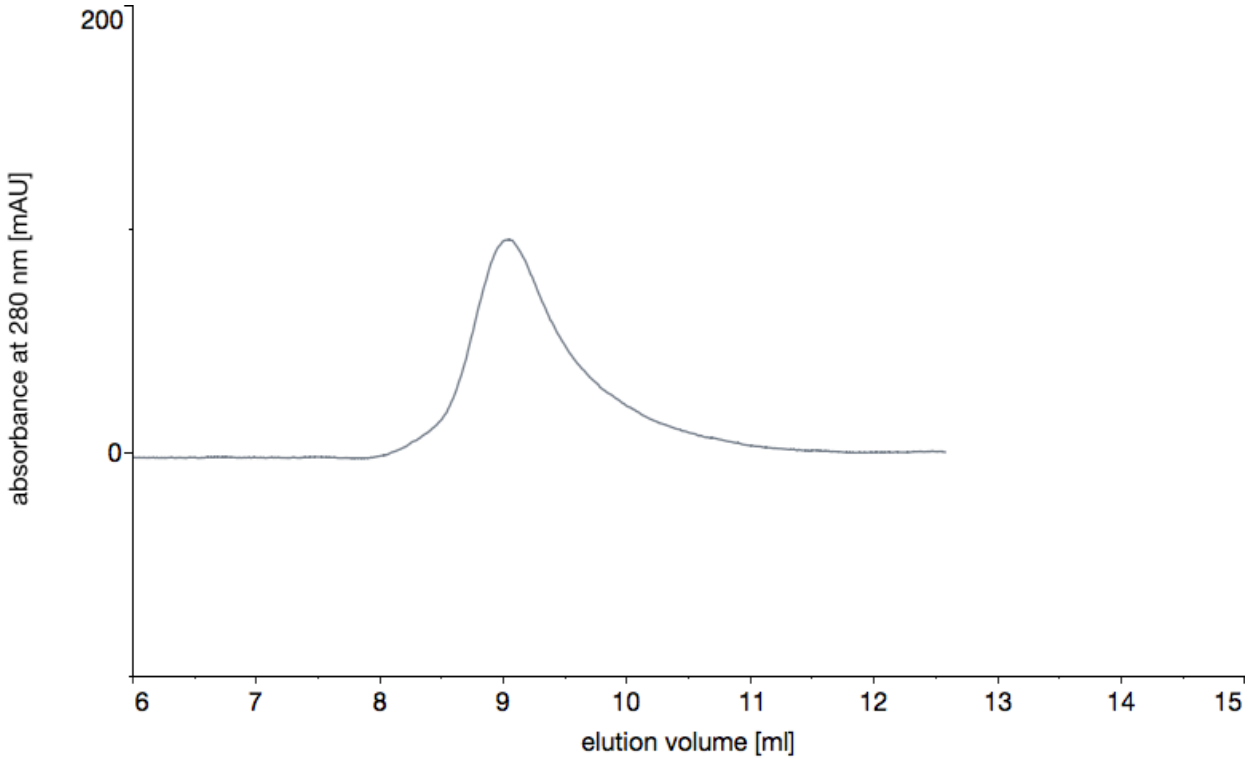


Figure 3.44 Purification conditions and elution profile of run 28.

Construct:	4	Beads:	amylose
Insect cell line:	Sf9	Protease:	thrombin
Amount of membrane (g):	1.4		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
2 mM imidazole pH 7	2 mM imidazole pH 7	10/20 mM imidazole pH 7	300 mM imidazole pH 7	0.1% DDM
1% DDM	0.1% DDM	0.1% DDM	0.1% DDM	
1 mg/mL iodoacetamide	0.05% CHS	0.05% CHS	0.05% CHS	
protease inhibitor	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	
0.05% CHS				
0.1 mg/mL E. coli lipid extract				

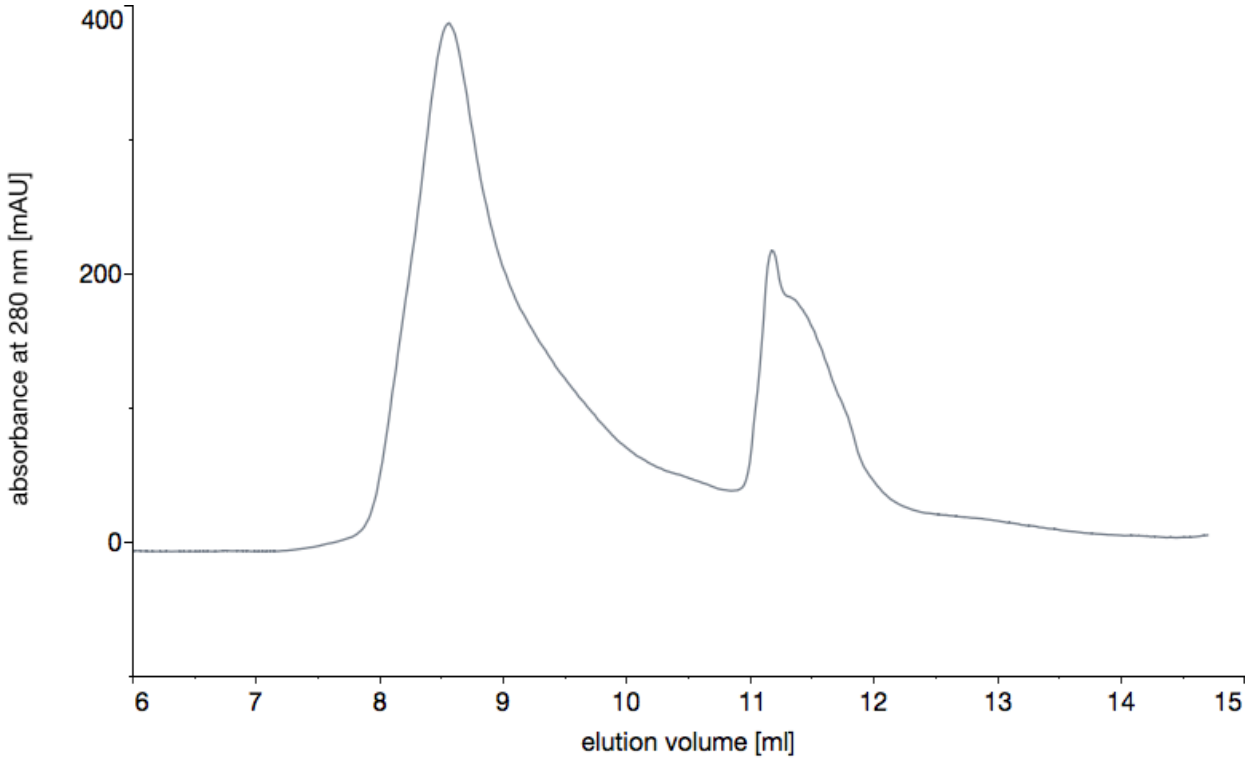


Figure 3.45 Purification conditions and elution profile of run 29.

Construct:	4	Beads:	TALON
Insect cell line:	Sf9	Protease:	thrombin
Amount of membrane (g):	1.4		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
2 mM imidazole pH 7	2 mM imidazole pH 7	10/20 mM imidazole pH 7	300 mM imidazole pH 7	0.1% DDM
1% DDM	0.1% DDM	0.1% DDM	0.1% DDM	
1 mg/mL iodoacetamide	0.05% CHS	0.05% CHS	0.05% CHS	
protease inhibitor	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	
0.05% CHS				
0.1 mg/mL E. coli lipid extract				

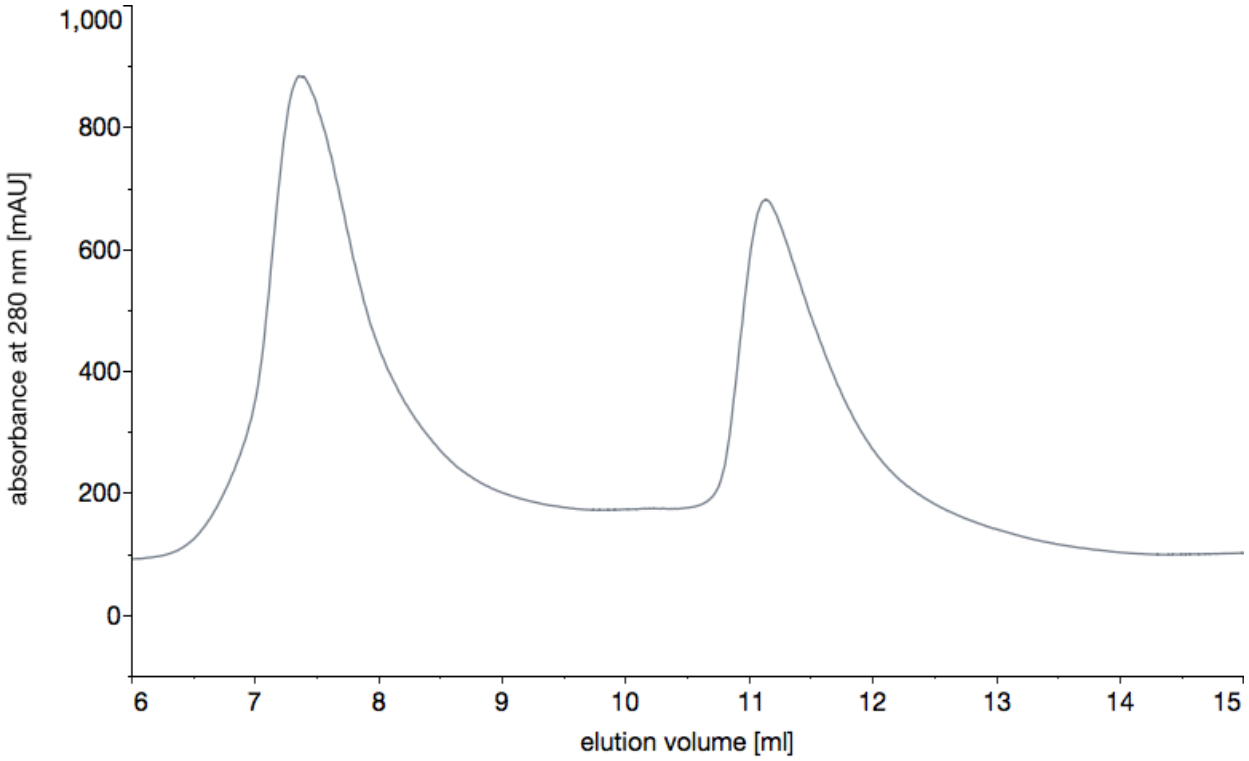


Figure 3.46 Purification conditions and elution profile of run 30.

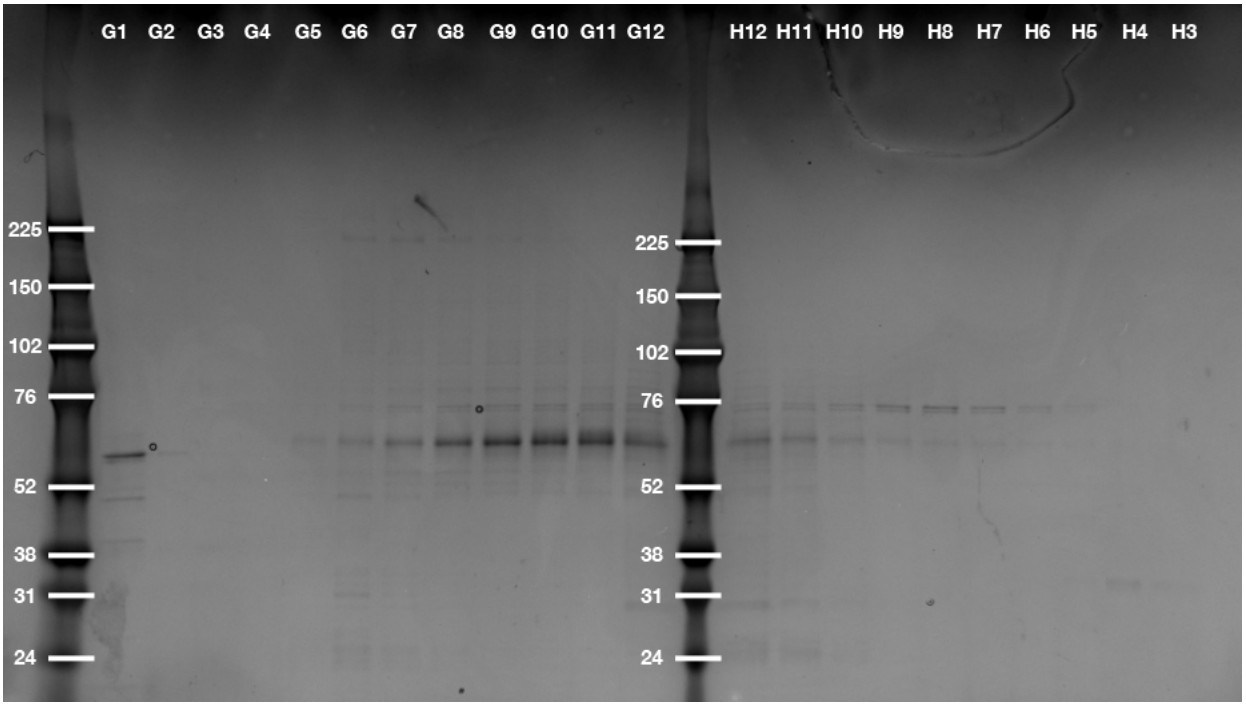


Figure 3.47 Coomassie gel stain of run 30.

Construct:	4	Beads:	TALON
Insect cell line:	High Five	Protease:	thrombin
Amount of membrane (g):	3		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
8 mM imidazole pH 7	8 mM imidazole pH 7	10/20 mM imidazole pH 7	300 mM imidazole pH 7	1.0% BRIJ
1.0% BRIJ	1.0% BRIJ	1.0% BRIJ	1.0% BRIJ	
1 mg/mL iodoacetamide	0.05% CHS	0.05% CHS	0.05% CHS	
protease inhibitor	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	
0.05% CHS				
0.1 mg/mL E. coli lipid extract				

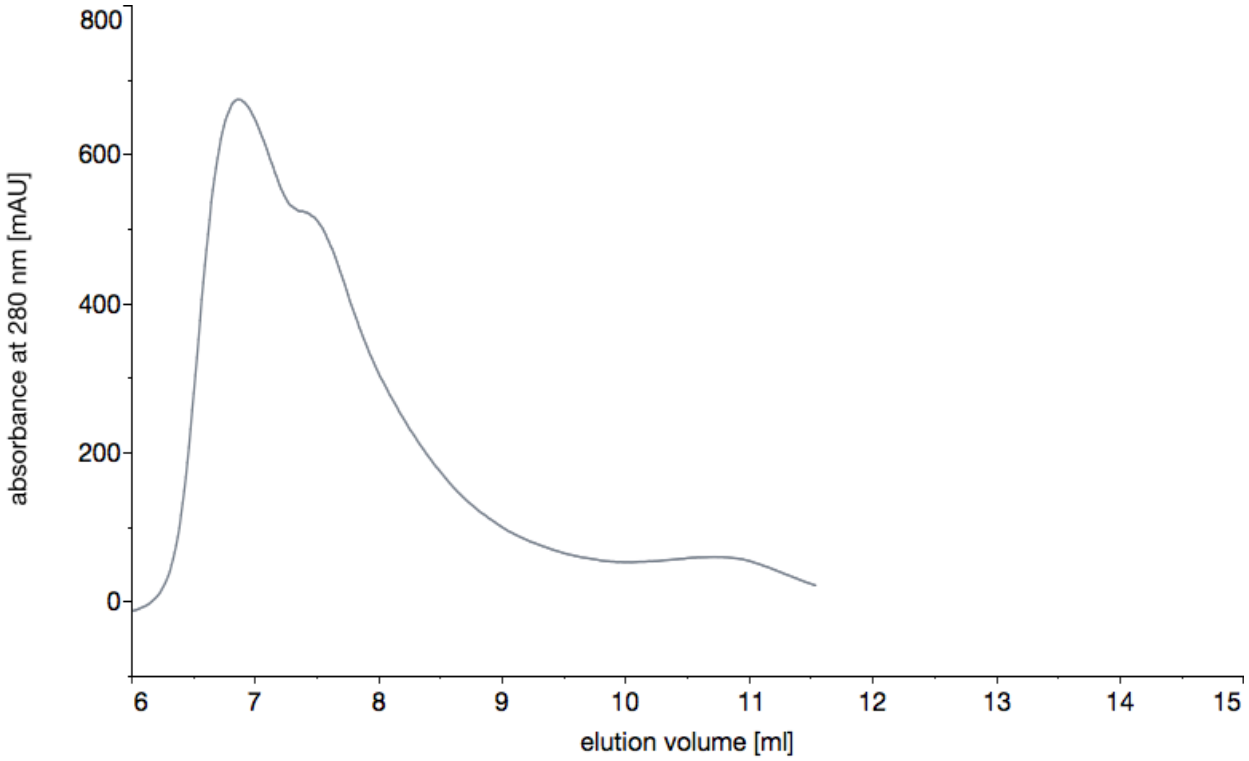


Figure 3.48 Purification conditions and elution profile of run 31.

Construct:	4	Beads:	TALON
Insect cell line:	Sf9	Protease:	thrombin
Amount of membrane (g):	3		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
8 mM imidazole pH 7	10 mM imidazole pH 7	20/30 mM imidazole pH 7	300 mM imidazole pH 7	0.1% DDM
1% DDM	0.1% DDM	0.1% DDM	0.1% DDM	
1 mg/mL iodoacetamide	0.05% CHS	0.05% CHS	0.05% CHS	
protease inhibitor	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	
0.05% CHS				
0.1 mg/mL E. coli lipid extract				

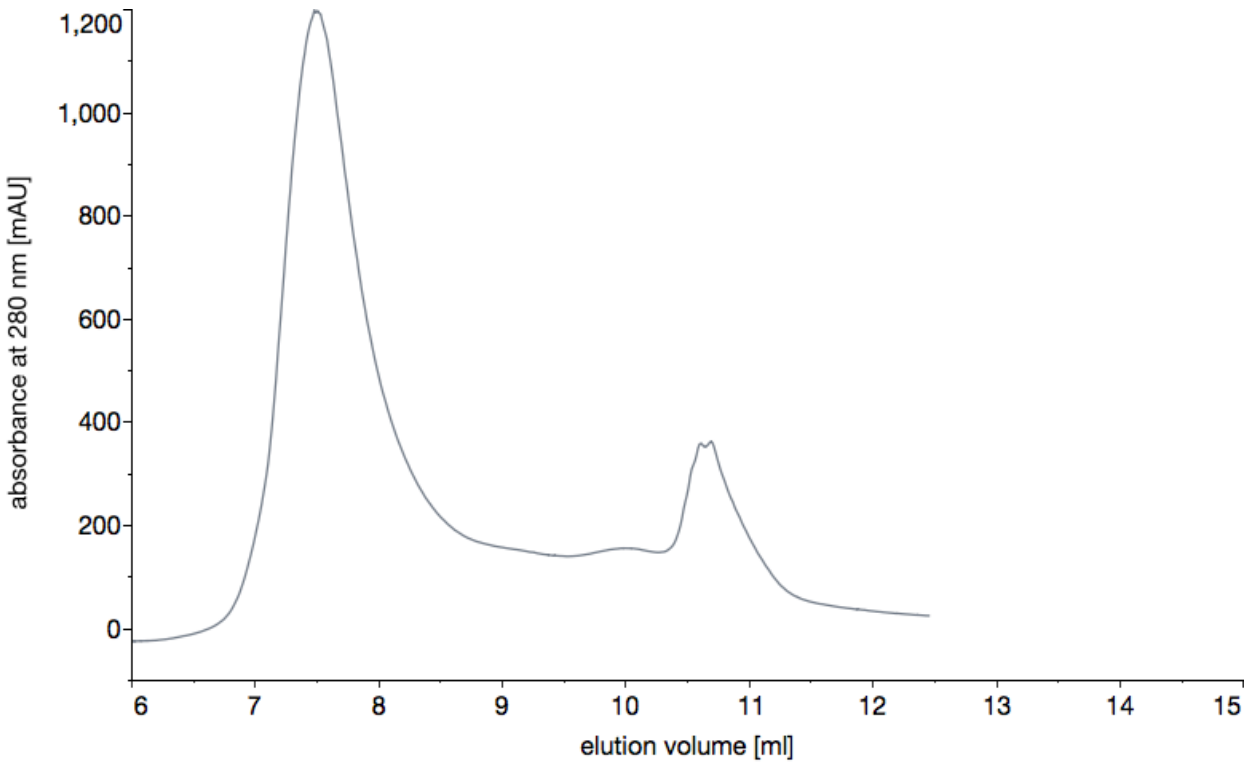


Figure 3.49 Purification conditions and elution profile of run 32.

Construct 5 To test whether MBP could enhance the stability of the most stable construct to date, construct 3, we created construct 5. Again, as the baseline, we assumed CHS and *E. coli* polar lipid extract with DDM. For the first trial, we omitted CHS and *E. coli* polar lipids, and eluted the entire protein fusion product off the TALON resin. Although the SEC profile did not show promise (run 33, **Figure 3.50**), subsequent cleavage with thrombin was efficient (**Figure 3.42**), and was incorporated into the ensuing runs. Next, we added CHS and *E. coli* polar lipid extract, and tried cutting the protein directly off the TALON resin (run 34, **Figure 3.51**). The void peak was much smaller than in previous runs, but the profile was quite noisy and irregular. To finish the screen of the library of common detergents in the Stroud lab, we tested $C_{12}E_8$, while using amylose resin for protein immobilization (run 35, **Figure 3.52**). The SEC profile had a sharp second peak, but the yield was significantly lower than with TALON. To test whether the type of resin was an issue, we then tested the same buffer conditions with TALON resin (run 36, **Figure 3.53**). It led to an even lower sample yield, which implied that $C_{12}E_8$ was simply inefficient at extracting hCNT3 from the membrane. In addition, we also tested DDM with CHS and *E. coli* polar lipid extract and cutting the protein directly off the amylose resin (run 37, **Figure 3.54**). Again, the final yield was low and a repeat with a 5-fold increase in membrane amount (run 38, **Figure 3.55**), even after a dialysis of DDM, which might interfere with amylose binding, did not improve sample stability. No further trials containing the construct were performed.

Construct:	5	Beads:	TALON
Insect cell line:	High Five	Protease:	none
Amount of membrane (g):	4.15		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
2 mM imidazole pH 7	2 mM imidazole pH 7	10/20 mM imidazole pH 7	300 mM imidazole pH 7	0.1% DDM
1% DDM	0.1% DDM	0.1% DDM	0.1% DDM	
1 mg/mL iodoacetamide				
protease inhibitor				

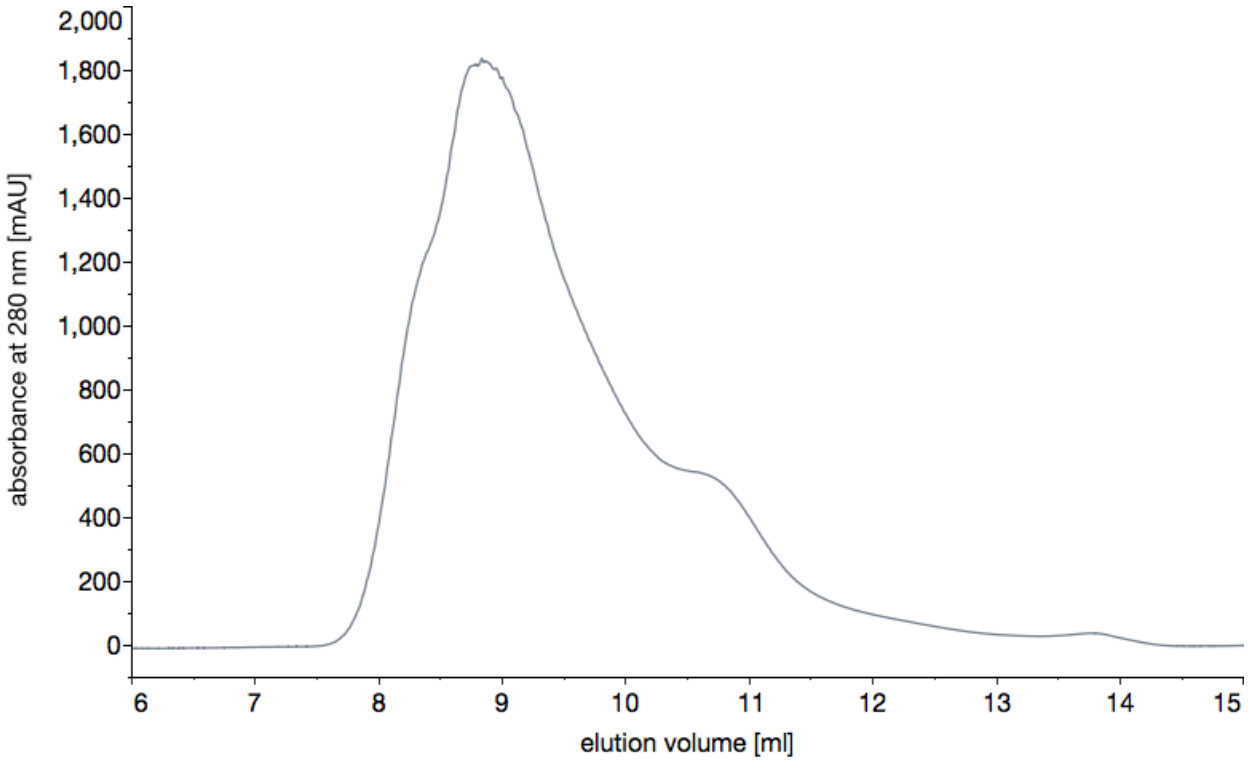


Figure 3.50 Purification conditions and elution profile of run 33.

Construct:	5	Beads:	TALON
Insect cell line:	High Five	Protease:	thrombin
Amount of membrane (g):	1		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
2 mM imidazole pH 7	2 mM imidazole pH 7	10/20 mM imidazole pH 7	300 mM imidazole pH 7	0.1% DDM
1% DDM	0.1% DDM	0.1% DDM	0.1% DDM	
1 mg/mL iodoacetamide	0.05% CHS	0.05% CHS	0.05% CHS	
protease inhibitor	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	
0.05% CHS				
0.1 mg/mL E. coli lipid extract				

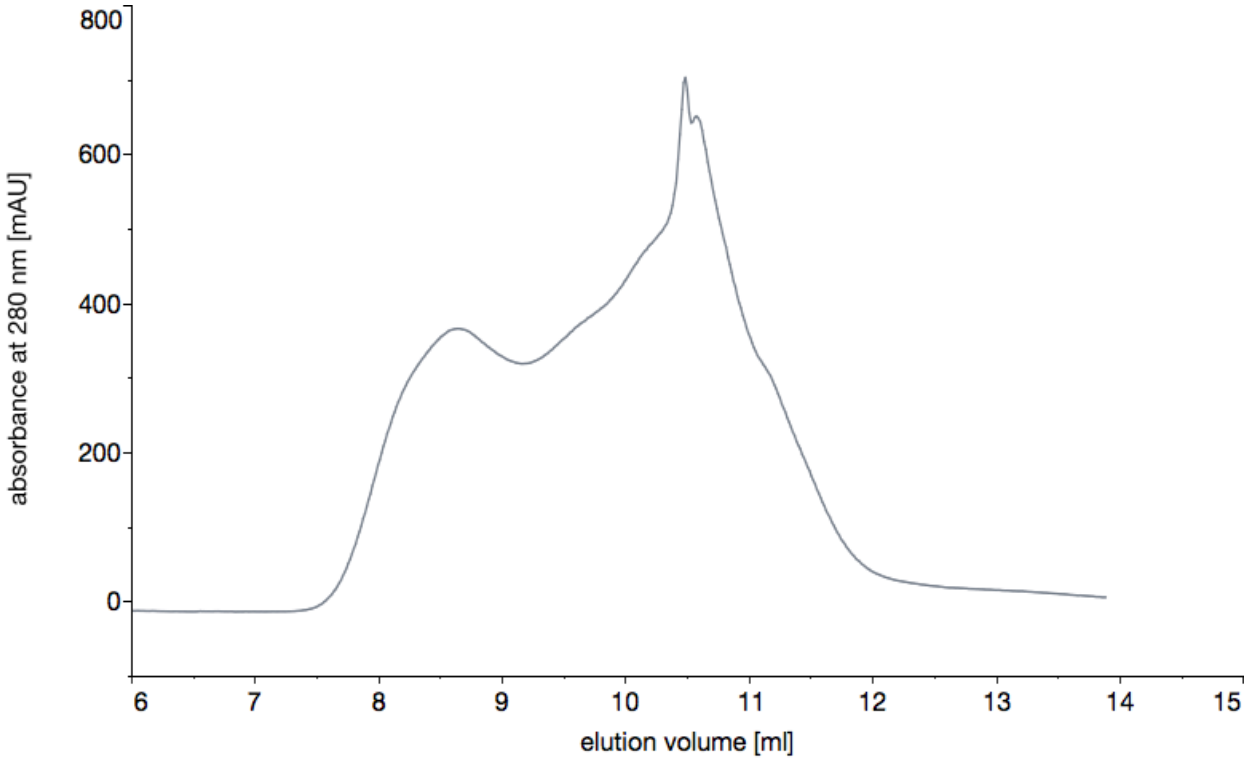


Figure 3.51 Purification conditions and elution profile of run 34.

Construct:	5	Beads:	amylose
Insect cell line:	High Five	Protease:	thrombin
Amount of membrane (g):	1		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
2 mM imidazole pH 7	2 mM imidazole pH 7	10/20 mM imidazole pH 7	300 mM imidazole pH 7	1.0% C12E8
1.0% C12E8	1.0% C12E8	1.0% C12E8	1.0% C12E8	
1 mg/mL iodoacetamide	0.05% CHS	0.05% CHS	0.05% CHS	
protease inhibitor	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	
0.05% CHS				
0.1 mg/mL E. coli lipid extract				

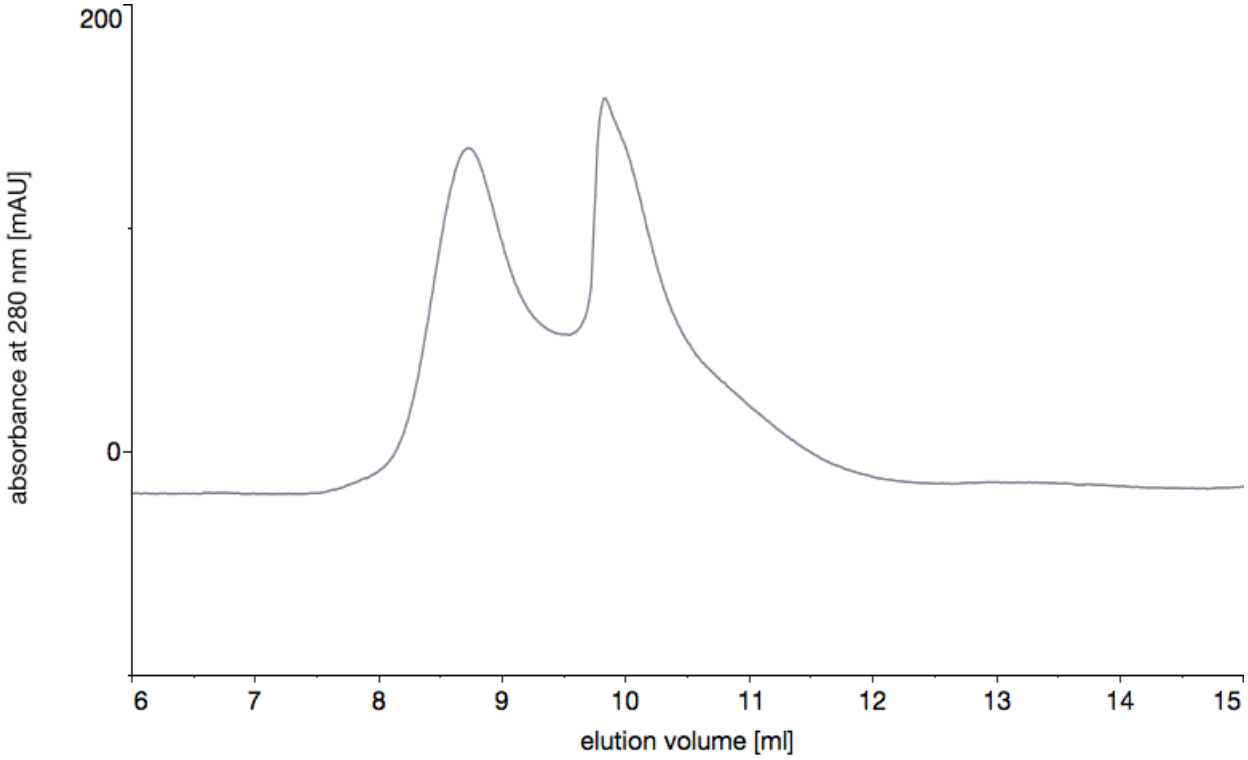


Figure 3.52 Purification conditions and elution profile of run 35.

Construct:	5	Beads:	TALON
Insect cell line:	High Five	Protease:	thrombin
Amount of membrane (g):	1.08		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
8 mM imidazole pH 7	8 mM imidazole pH 7	20/50 mM imidazole pH 7	300 mM imidazole pH 7	1.0% C12E8
1.0% C12E8	1.0% C12E8	1.0% C12E8	1.0% C12E8	
1 mg/mL iodoacetamide	0.05% CHS	0.05% CHS	0.05% CHS	
protease inhibitor	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	
0.05% CHS				
0.1 mg/mL E. coli lipid extract				

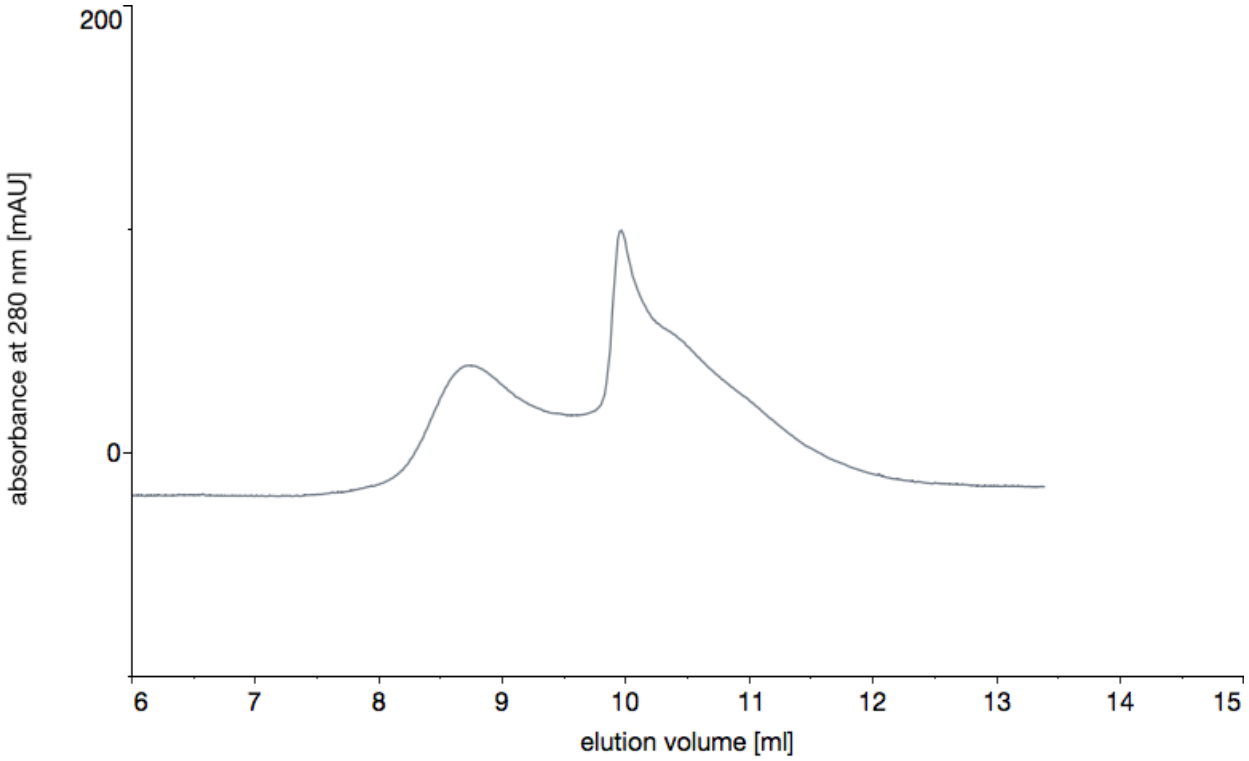


Figure 3.53 Purification conditions and elution profile of run 36.

Construct:	5	Beads:	amylose
Insect cell line:	High Five	Protease:	thrombin
Amount of membrane (g):	0.83		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
8 mM imidazole pH 7	8 mM imidazole pH 7	20/50 mM imidazole pH 7	300 mM imidazole pH 7	0.1% DDM
1.0% DDM	0.1% DDM	0.1% DDM	0.1% DDM	
1 mg/mL iodoacetamide	0.05% CHS	0.05% CHS	0.05% CHS	
protease inhibitor	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	
0.05% CHS				
0.1 mg/mL E. coli lipid extract				

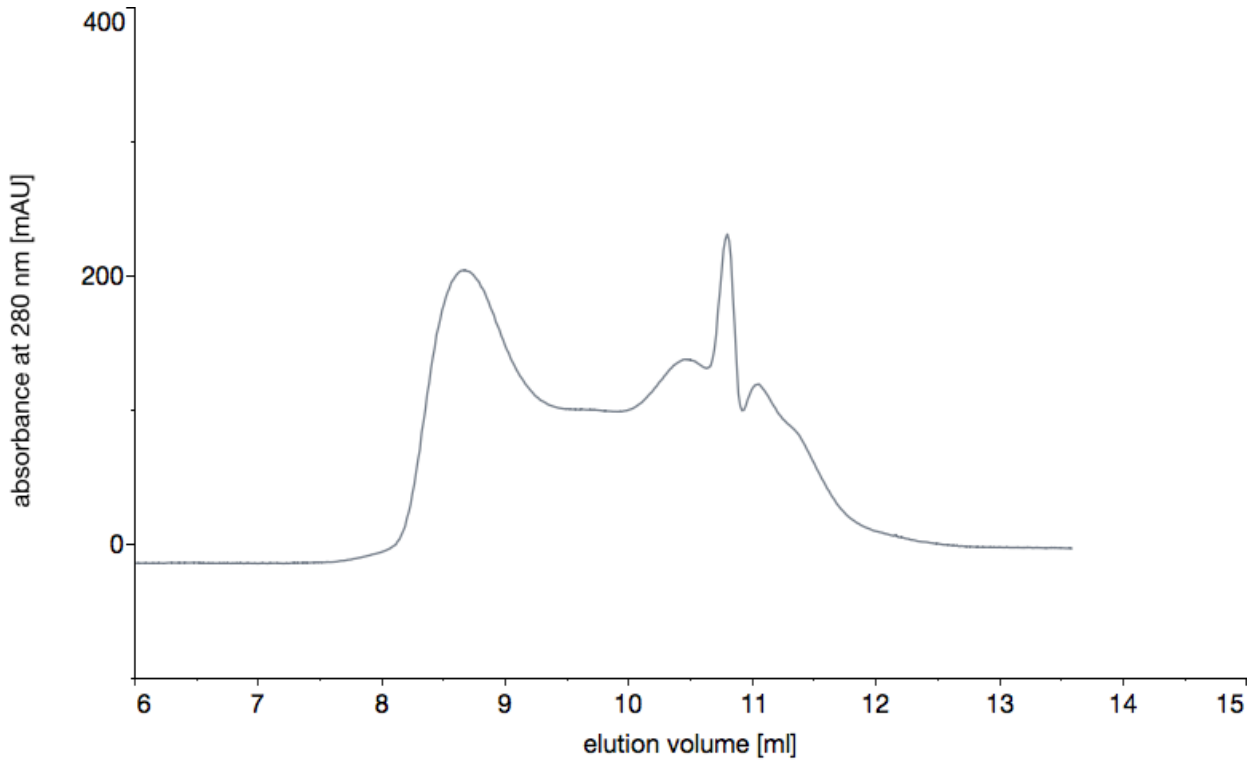


Figure 3.54 Purification conditions and elution profile of run 37.

Construct:	5	Beads:	amylose
Insect cell line:	High Five	Protease:	thrombin
Amount of membrane (g):	4.57		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
8 mM imidazole pH 7	8 mM imidazole pH 7	20/50 mM imidazole pH 7	300 mM imidazole pH 7	0.1% DDM
1.0% DDM	0.1% DDM	0.1% DDM	0.1% DDM	
1 mg/mL iodoacetamide	0.05% CHS	0.05% CHS	0.05% CHS	
protease inhibitor	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	
0.05% CHS				
0.1 mg/mL E. coli lipid extract				

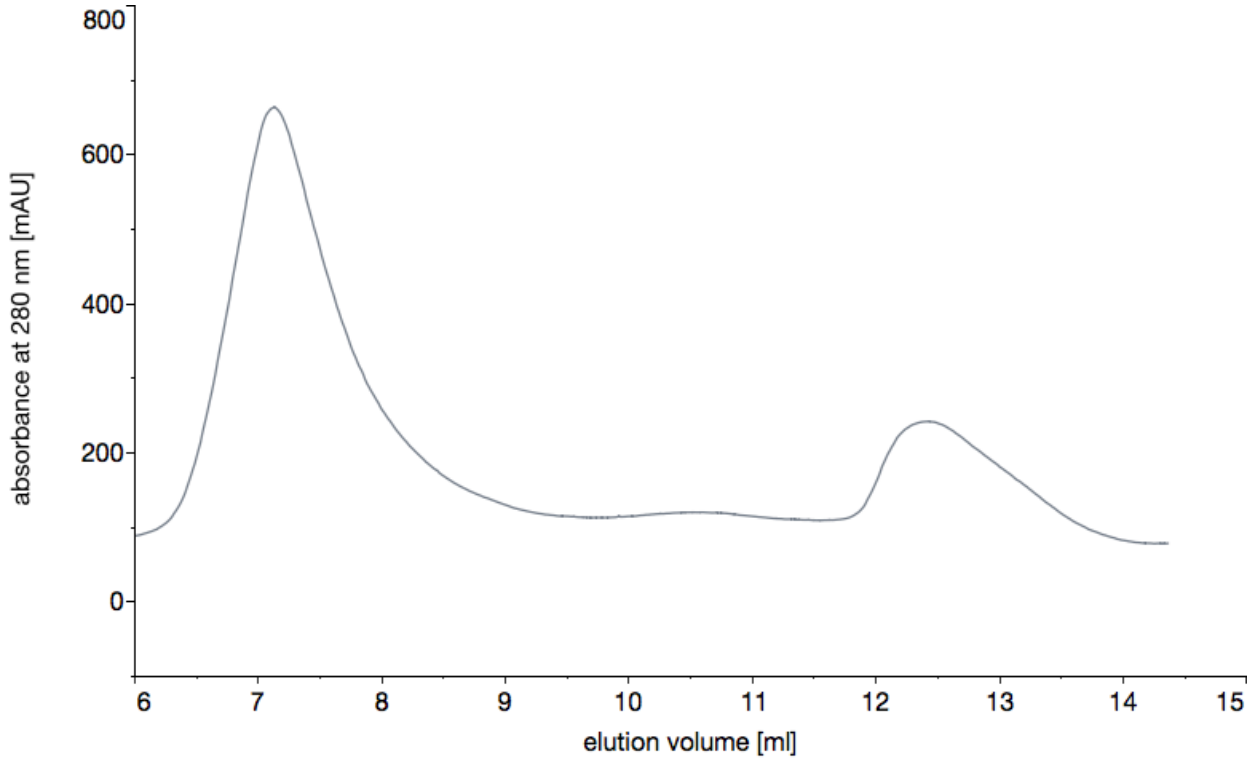


Figure 3.55 Purification conditions and elution profile of run 38.

Constructs 7 and 8 Constructs 7 and 8 are more refined versions of constructs 4 and 5. Construct 7 has an additional two mutations V274C and A285C, which were selected to cross-link the trimer together through the purification process. The sites were shown to cross-link in Chapter 3.1. Construct 8 is missing the first 90 amino acid residues which are predicted to be largely disordered, and is the first construct to lack parts of both the N- and C-termini. Despite using the most stabilizing conditions for construct 3, DDM with CHS and *E. coli* polar lipid extract, and cutting the protein off the column with thrombin, both of their SEC profiles show a largely aggregated void peak (run 39, **Figure 3.56** and run 40, **Figure 3.57**). The constructs were not examined further.

Construct:	7	Beads:	TALON
Insect cell line:	High Five	Protease:	thrombin
Amount of membrane (g):	4.68		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
8 mM imidazole pH 7	10 mM imidazole pH 7	20/40 mM imidazole pH 7	300 mM imidazole pH 7	0.1% DDM
1.0% DDM	0.1% DDM	0.1% DDM	0.1% DDM	
1 mg/mL iodoacetamide	0.05% CHS	0.05% CHS	0.05% CHS	
protease inhibitor	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	
0.05% CHS				
0.1 mg/mL E. coli lipid extract				

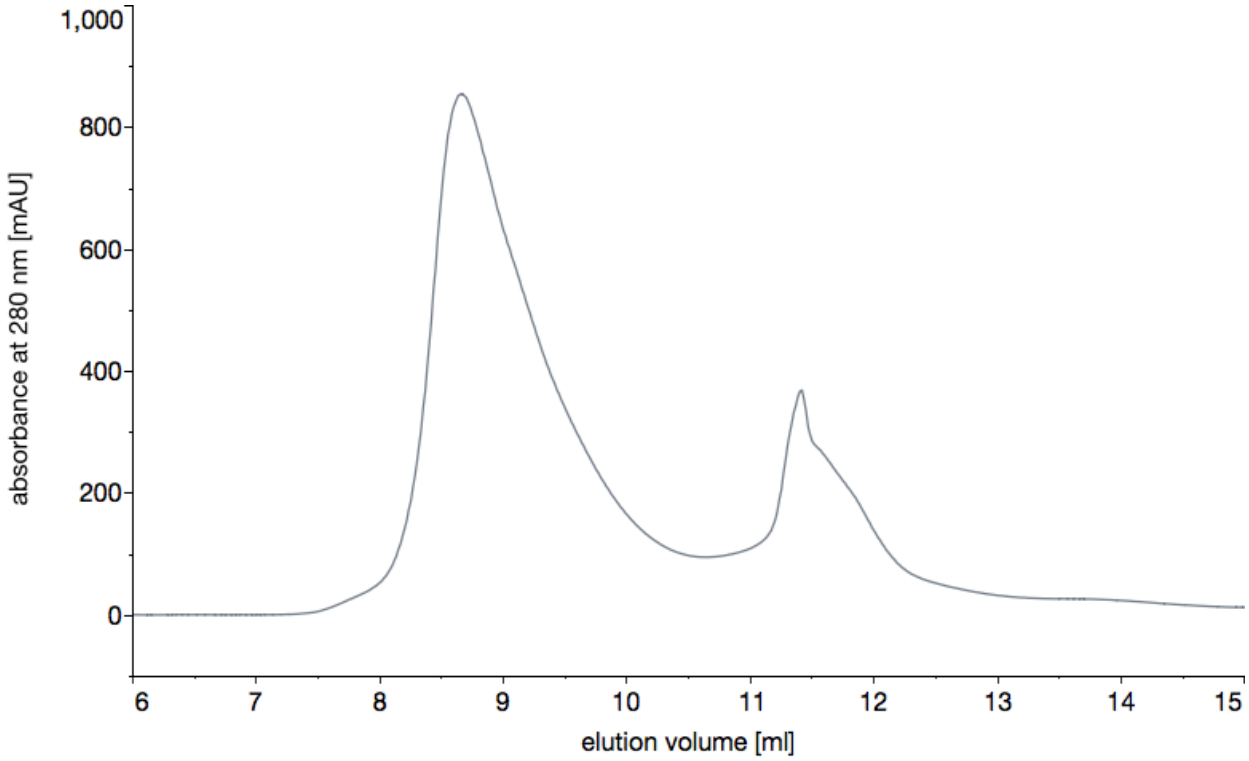


Figure 3.56 Purification conditions and elution profile of run 39.

Construct:	8	Beads:	TALON
Insect cell line:	High Five	Protease:	thrombin
Amount of membrane (g):	2.53		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
8 mM imidazole pH 7	10 mM imidazole pH 7	20/40 mM imidazole pH 7	300 mM imidazole pH 7	0.1% DDM
1.0% DDM	0.1% DDM	0.1% DDM	0.1% DDM	
1 mg/mL iodoacetamide	0.05% CHS	0.05% CHS	0.05% CHS	
protease inhibitor	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	
0.05% CHS				
0.1 mg/mL E. coli lipid extract				

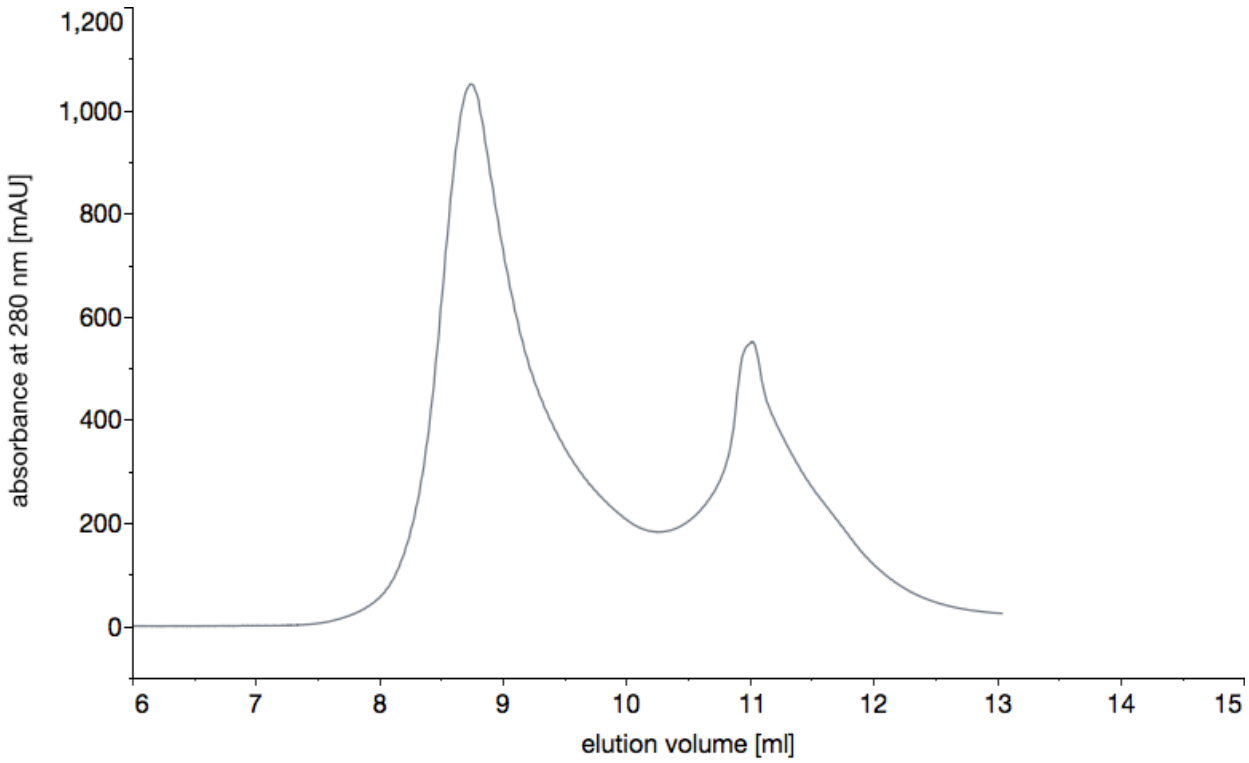


Figure 3.57 Purification conditions and elution profile of run 40.

Construct 6 The last construct examined required the most design. It contains an N-terminal MBP tag, a poly-asparagine sequence shown to enhance amylose resin binding,[86,87] long linker regions to ensure protease accessibility, as well as mutations of the 4 putative glycosylated asparagine residues. The first run consisted of using DDM with CHS and *E. coli* polar lipid extract, and cutting the protein off the TALON resin (run 41, **Figure 3.58**). Although the peak corresponding to the trimer was small, it was at the right elution volume. A fresh membrane prep from Sf9 insect cells was prepared, but it did not improve the results (run 42, **Figure 3.59**). The last trial consisted of reducing the glycerol concentration down to 0.5%, increasing the DDM concentration, and including CHS and *E. coli* polar lipid extract. The full-length protein product was eluted off TALON resin with imidazole and not cleaved by thrombin. This too did not improve the general stability of the protein sample.

Conclusion

A total of 8 different constructs were designed for expression and purification trials. Despite attempting to stabilize the hCNT3 homo-trimer with a variety of commonly employed methods, including detergent screens, lipid extracts, pH changes, glycerol and salt concentrations, IMAC resin types, and protein fusion tags, the purification always resulted in a mostly aggregated protein product with a large void peak on the SEC elution profile. These results should serve as the foundation for future work, as a catalog of conditions that have already been attempted, and which showed promise. The addition of MBP enhanced sample quality, as did CHS and *E. coli* polar lipid extract. Further work in the determination of the structure of CNT3 should focus on the screening of close homologs from higher eukaryotes, to identify those that fold properly and express in high quantities. While the focus of this project was on the human CNT3, a closely related homolog might be easier to purify and crystallize. Changing the expression system to HEK 293S would also be encouraged. The 43 trials described here represent the norm rather than an outlier when working with human SLC transporters. The importance of this transporter family should serve as the motivation for future efforts.

Construct:	6	Beads:	TALON
Insect cell line:	High Five	Protease:	thrombin
Amount of membrane (g):	6.85		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
8 mM imidazole pH 7	10 mM imidazole pH 7	20/30 mM imidazole pH 7	300 mM imidazole pH 7	0.1% DDM
1.0% DDM	0.1% DDM	0.1% DDM	0.1% DDM	
1 mg/mL iodoacetamide	0.05% CHS	0.05% CHS	0.05% CHS	
protease inhibitor	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	
0.05% CHS				
0.1 mg/mL E. coli lipid extract				

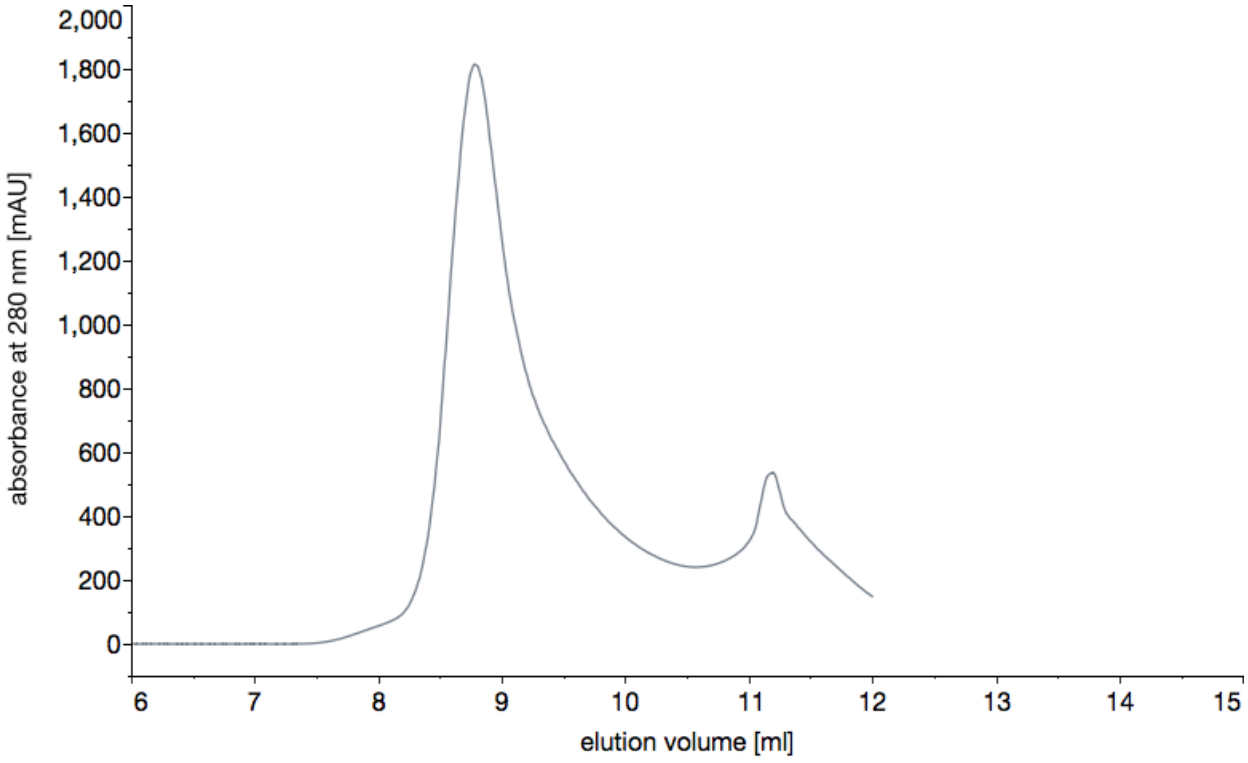


Figure 3.58 Purification conditions and elution profile of run 41.

Construct:	6	Beads:	TALON
Insect cell line:	Sf9	Protease:	none
Amount of membrane (g):	2.5		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
5% glycerol	0.5% glycerol	0.5% glycerol	0.5% glycerol	0.5% glycerol
8 mM imidazole pH 7	8 mM imidazole pH 7	20/40 mM imidazole pH 7	300 mM imidazole pH 7	0.25% DDM
1.5% DDM	0.25% DDM	0.25% DDM	0.25% DDM	
1 mg/mL iodoacetamide	0.05% CHS	0.05% CHS	0.05% CHS	
protease inhibitor	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	
0.05% CHS				
0.1 mg/mL E. coli lipid extract				

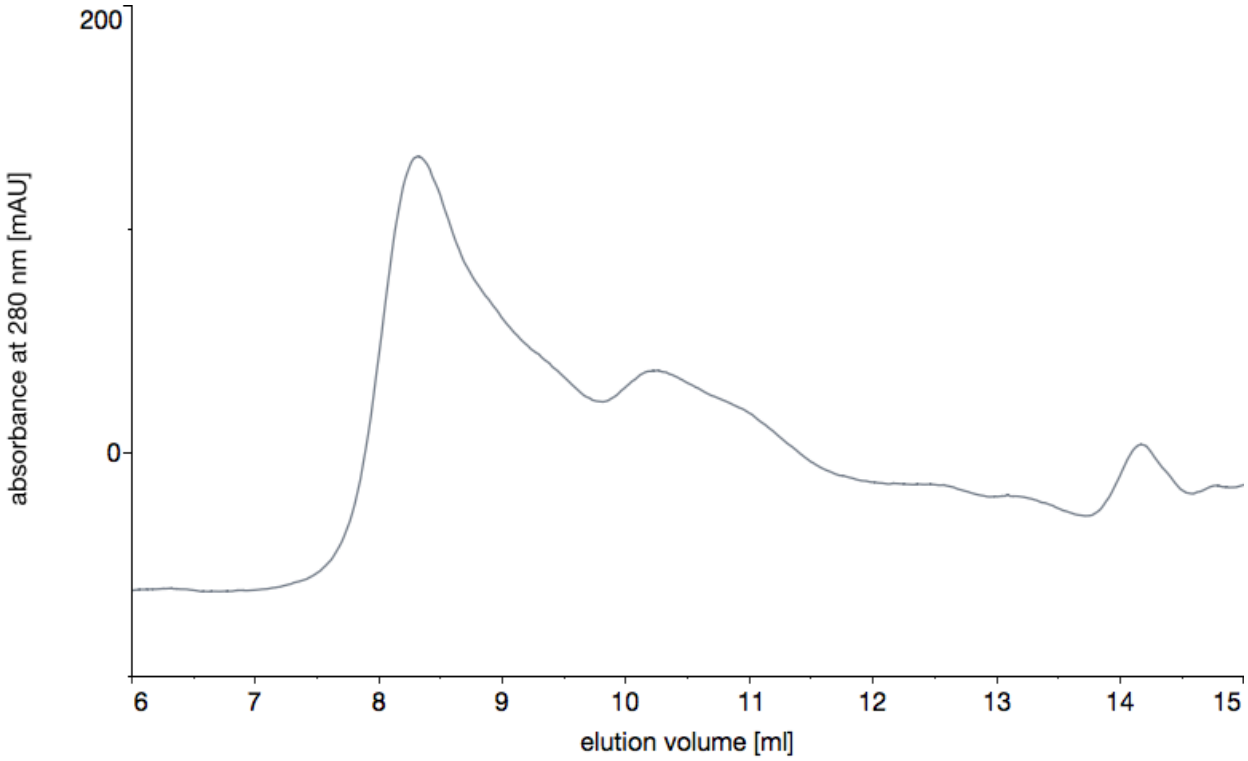


Figure 3.59 Purification conditions and elution profile of run 42.

Construct:	6	Beads:	TALON
Insect cell line:	Sf9	Protease:	thrombin
Amount of membrane (g):	2.5		

Resuspension buffer	Wash I buffer	Wash II/III buffer	Elution buffer	SEC buffer
50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7	50 mM HEPES pH 7
300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
10% glycerol	10% glycerol	10% glycerol	10% glycerol	10% glycerol
2 mM imidazole pH 7	2 mM imidazole pH 7	20/40 mM imidazole pH 7	300 mM imidazole pH 7	0.1% DDM
1.0% DDM	0.1% DDM	0.1% DDM	0.1% DDM	
1 mg/mL iodoacetamide	0.05% CHS	0.05% CHS	0.05% CHS	
protease inhibitor	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	0.1 mg/mL E. coli lipid extract	
0.05% CHS				
0.1 mg/mL E. coli lipid extract				

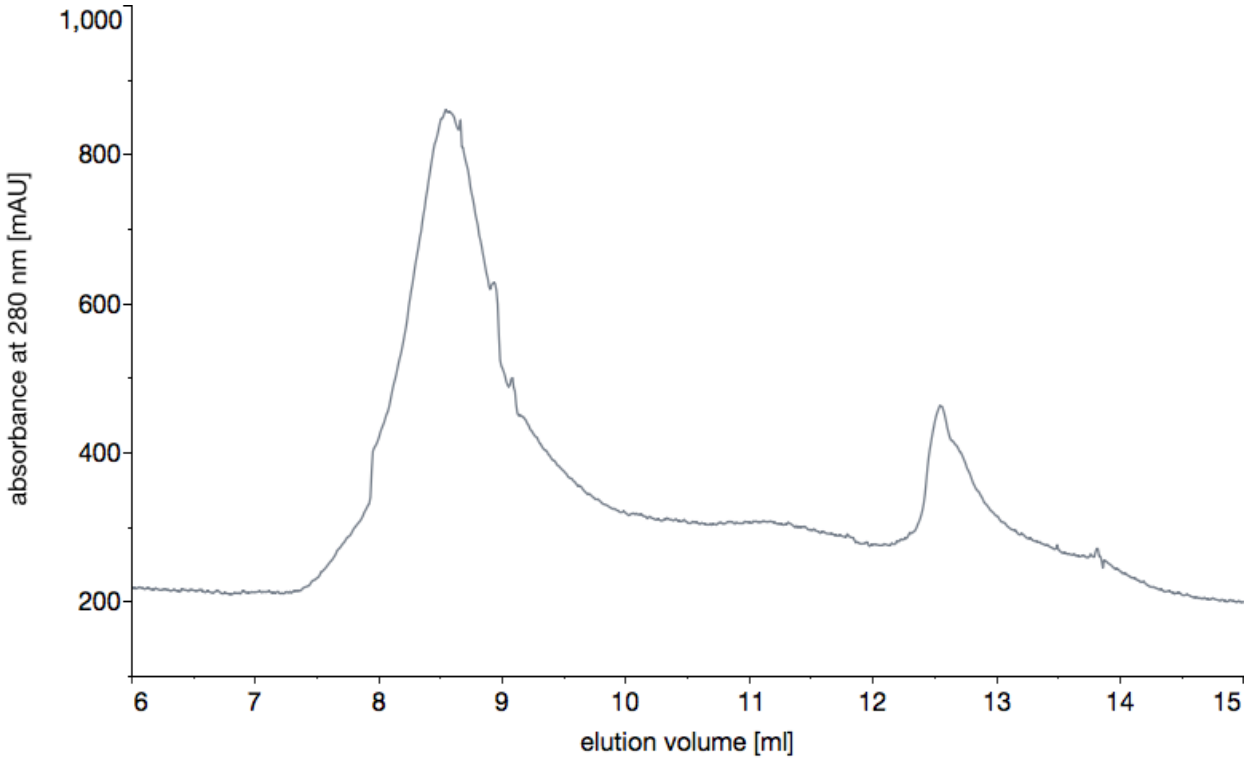


Figure 3.60 Purification conditions and elution profile of run 43.

Screening for novel ligands of hCNT3

The human genome encodes for 3 concentrative nucleoside transporters: CNT1 (SLC28A1) that is predominantly pyrimidine-selective, CNT2 (SLC28A2) that is purine-selective, and CNT3 (SLC28A3) that exhibits broad specificity.[88] The residues governing the ligand specificity of each of the CNT transporters have been determined through extensive chimeric studies.[2] By creating a chimera containing residues 1–300 of rat CNT1, 297–358 of rat CNT2, and 363–648 of rat CNT1, the specificity of rat CNT1 was altered such that the transporter became purine-selective. Further examination implicated hCNT1 residue S318 (hCNT3 residue G340) as the main determinant of ligand selectivity.[89] When mutated to the corresponding hCNT2 residue, glycine, the mutant was capable of both purine and pyrimidine uptake. Additional mutation of S352/L353 (hCNT3 residues S374/V375) of this chimera into the corresponding hCNT2 residues T347/V348 abolished pyrimidine transport, creating a purine-specific hCNT2-like tranporter.[77] Other mutations implicated over the years to be critical for substrate binding include hCNT1 residues E321 (hCNT3 residue E343) and E497 (hCNT3 residue E519).[90]

The details of nucleoside binding to CNTs were recently further clarified by the co-crystal structures of vcCNT with uridine and a series of other nucleosides and nucleoside analogues (zebularine, adenosine, ribavirin, cytidine, purrolo-cytidine, gemcitabine, 5-fluorouridine).[46] While the 2012 structure of vcCNT[14] allowed us to model the human CNTs for the first time, the coupling of equilibrium-binding and extensive crystallographic studies has allowed for the elucidation of the importance of specific residue/nucleoside interactions. Nucleoside interactions with the binding pocket can thus be divided into two groups: ribose interactions and nucleobase interactions.[46]

Amino acid residues that participate in the ribose interactions (hCNT3 residues E519, N565, S568) (**Figure 3.61A**) are conserved between vcCNT and all human CNTs. The same is not true for the nucleobase interactions. Because each human CNT has a different specificity, the exact nucleobase interactions vary across transporters. For hCNT3, those include Q341, T342, E343 (direct hydrogen bonds or through a water molecule), and V375 (van der Waals interaction) (**Figure 3.61A**). Since many nucleoside analogues contain a modified ribose sugar, they most likely sacrifice binding affinity stemming from these interactions. However, the total binding affinity can be additively compensated for by the nucleobase interactions, as shown by pyrrolo-gemcitabine.[46] Specificity also arises from the pocket volume. hCNT3 and hCNT2 can accommodate purines due to the larger

binding pockets created by a smaller side chain of valine (hCNT3 residue V375) near the nucleobase, instead of a leucine present in the pyrimidine-selective hCNT1.

As mentioned previously, the focus on CNTs stems not only from their endogenous role in nucleoside salvage, but also from the importance of their pharmacological substrates, the nucleoside analogs, which frequently find use in the treatment of viral infections and cancers.[12] To maximize the efficacy of these highly toxic agents, one strategy is to limit the distribution of these compounds into noncancerous tissues. This goal might require inhibition of transporters that transport nucleoside analogs into the cytoplasm. While highly selective inhibitors of ENTs, such as NBMPR, have been identified several decades ago,[91] the same is not true for human CNTs. Phloridzin, a naturally occurring flavonoid,[92] has long been considered as the standard hCNT inhibitor (K_i = 16 μM) (**Figure 3.62**). However, it is neither hCNT specific nor particularly potent. While recent studies have attempted to identify more potent inhibitors, they represent incremental improvements through the synthesis of phloridzin analogs,[93] benzopyranone derivatives,[36] or fused-pyrimidine nucleoside analogs.[94]

Here, we describe screening hundreds of thousands of compounds, encompassing a large, never-before-tested chemical space, in search of novel hCNT3 ligands, both substrates and inhibitors. Following a virtual screen, a subset of compounds was tested in a cell-based assay. Ticagrelor, an FDA-approved platelet aggregation inhibitor, was identified as a novel, potent (IC_{50} of 6.47 ± 1.27 μM) inhibitor of hCNT3-mediated uridine uptake. This result serves as an encouraging first step in the screening of significantly larger and varied libraries.

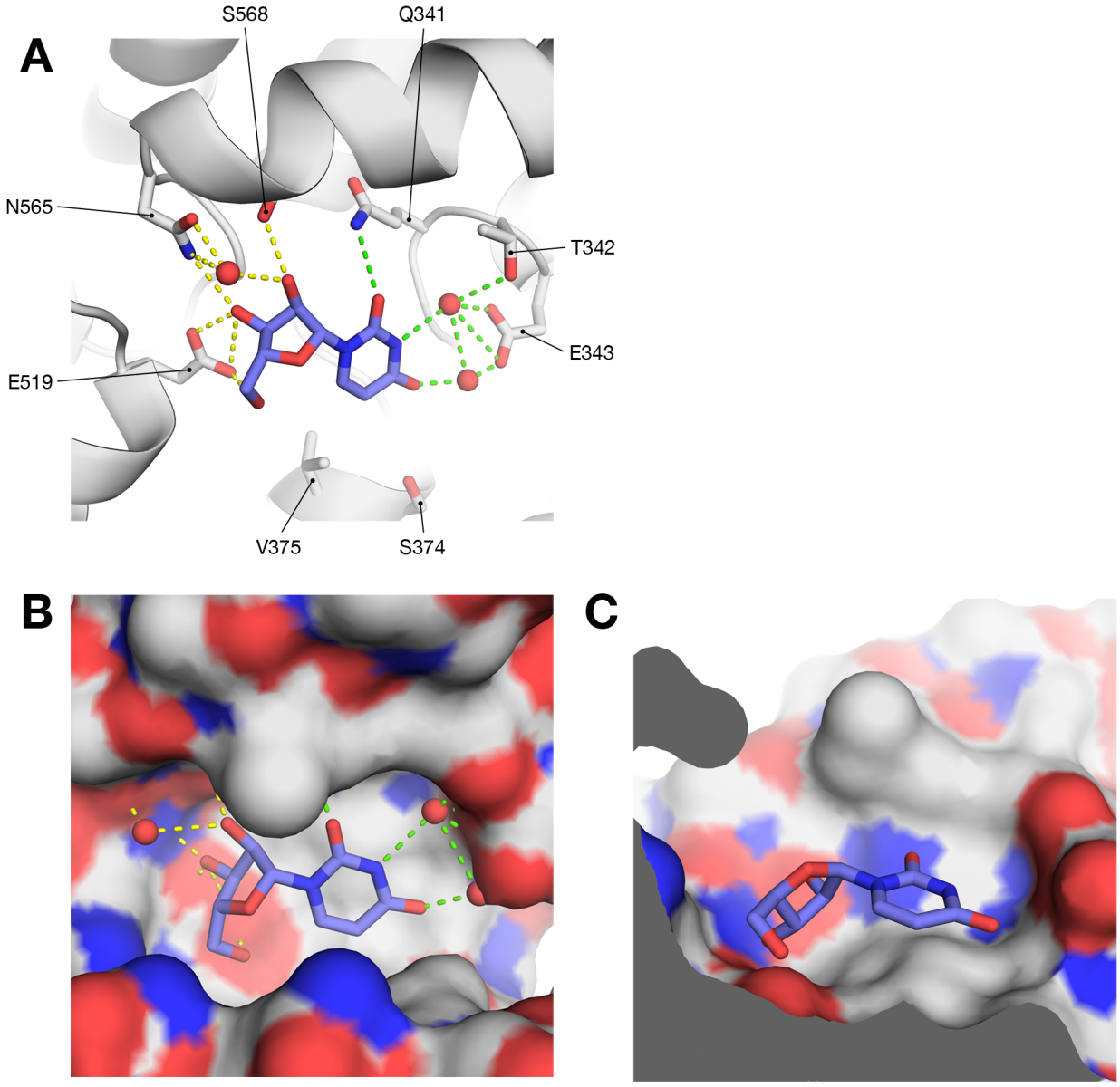


Figure 3.61. Cartoon and surface representations of the hCNT3 active site with uridine. Comparative structure model of hCNT3 was created using the co-crystal structure of vcCNT with uridine (PDB ID 3TIJ)(ref. 14) as a template. A) Residues participating in the stabilization of the ligand within the active site are shown as sticks. Crystallographic waters (PDB ID 4PD6) are shown as red spheres. Hydrogen bonds between uridine (marine) and the protein are shown as yellow (ribose interactions) and lime (nucleobase intractions) dashed lines. B) Surface representation of the active site from (top view). C) Cross-sectional view of the surface representation of the active site.

Methods

Comparative structure modeling of hCNT3 A comparative protein structure model of hCNT3 was created using MODELLER 9.14.[31] The 2.4 Å co-crystal structure of vcCNT (39% sequence identity) bound to uridine in the inward-facing occluded conformation served as the template (PDB ID 3TIJ).[14] The sequence alignment was obtained by a manual refinement of gaps in the output from the PROMALS3D[95] server. The ligand and water molecules from the crystal structure were treated as the BLK residue type in the alignment and were essentially copied from the template structure into a model as a rigid body, retaining proper stereochemistry in the surrounding protein-binding site and protein-ligand interactions. One hundred models were created using the automodel class with default settings. The models had acceptable protein orientation-dependent statistically optimized atomic potential (SOAP-Protein) scores.[96] The top-scoring model was used for further analysis.

Virtual screening The model was evaluated for ligand enrichment from a set of decoys based on enrichment curves and corresponding logAUC values.[97,98] Docking was performed using UCSF DOCK 3.6.[99] Decoys of high physicochemical but low topological similarity were generated with the aid of the directory of useful decoys-enhanced (DUD-E) based on a set of known ligands.[100] Different combinations of water molecules were tested to determine which arrangement resulted in the highest enrichment. The binding pocket with no waters produced the highest logAUC of 35.26, suggesting that it is suitable for predicting novel ligands.[97,101,102] A random selection of compounds from a mixed set of decoys and ligands yields a logAUC of 14.5; twice as many ligands as random selection yields a logAUC of 24.5; and an enrichment that selects 10 times as many ligands as random yields a logAUC of 47.7.[97] With a validated model in hand, two different ZINC compound libraries were then screened: KEGG-DRUG[103] and a ZINC library of "fragment-like" compounds.[104] Next, docking poses were ranked by the DOCK scoring function, which is a sum of van der Waals interactions, electrostatic potential, and ligand desolvation penalty terms. Finally, the top 500 poses were inspected by eye to prioritize compounds for experimental validation.

Generation of cell lines Porcine kidney tubular epithelium nucleoside transporter deficient cells (PK15NTD)[34] were donated by Dr. Chung-Ming Tse (The Johns Hopkins University School of Medicine, Baltimore, MD, USA). Stably transfected PK15NTD cells were created by transfecting pcDNA5/FTR (Invitrogen) vector containing the full-length human CNT3 cDNA (CNT3) and the

empty vector (EV) using Lipofectamine 2000 (Invitrogen) per manufacturer's instructions. They were then cultured in Eagle's minimal essential medium with Earle's balanced salt solution with 1 mM sodium pyruvate, 0.1 mM non-essential amino acids, 10% FBS, 100 I.U./mL penicillin, 100 μg/mL streptomycin, and 200 μg/mL hygromycin B at 37°C and 5% CO_2, as reported previously.[35,36]

Uptake experiments Uptake studies were performed as described previously.[38] Briefly, PK15NTD cells were seeded on non-coated polystyrene 24-well plates (Corning). After 48-hour, the cells were incubated for 10 minutes in a 37°C sodium-free buffer (5 mM HEPES, 10 mM glucose, 1 mM $CaCl_2$, 140 mM N-methyl-D-glucamine, 5 mM KH_2PO_4, 1 mM $MgCl_2$, pH 7.4). The uptake was initiated by the addition of 33.3 nM [5,6-^{3}H]-uridine (Moravek MT 799) in a sodium-containing buffer (5 mM HEPES, 10 mM glucose, 1 mM $CaCl_2$, 5 mM KCl, 135 mM NaCl, 1 mM $MgCl_2$, 0.8 mM Na_2HPO_4, 3.3 mM NaH_2PO_4, pH 7.4) and terminated by washing the cells twice with 4°C sodium-free buffer. Cells were lysed by the addition of lysis buffer (0.1% SDS vol/vol, 0.1 N NaOH). Intracellular radioactivity was measured by scintillation counting and normalized per well of protein content using the BCA Protein Assay Kit (Pierce 23225). [^{3}H]5-amino-1-phenyl-1H-1,2,3-triazole-4-carbimidothioic acid (ZINC000000383878) was custom synthesized by Moravek Biochemicals.

Results and discussion

A comparative structure model of hCNT3 based on the structure of vcCNT[14] was created for the purposes of virtual screening. Two separate libraries were screened. The first library was the KEGG-DRUG database containing 7,624 approved drugs and drug metabolites approved in Japan, USA, and Europe.[103] The goal of this screen was to discover whether hCNT3 participates in the tissue distribution or elimination of existing drugs on the market or is a target of off-target inhibition. The second one was a library of 847,909 "fragment-like" compounds, defined as molecules with logP $\leq$ 3.5, molecular weight $\leq$ 250, and number of rotatable bonds $\leq$ 5.[104] In theory, fragments are non-functionalized compounds suitable for optimization with low target affinity but high ligand efficiency.[104] This screen was performed to discover compounds that might be transported by hCNT3 but are dissimilar to nucleosides.

Following a visual inspection of the top ranking molecules, 8 drugs and 6 fragments were purchased for experimental validation. The set represents a wide variety of chemical classes (**Figure 3.63** and **Figure 3.65**). To test the compounds, a PK15NTD cell line stably overexpressing hCNT3 was created.

An uptake assay using radiolabeled uridine validated the proper function of hCNT3 and inhibition by phloridzin at potencies previously reported in the literature (experimental IC_{50} = 7.40 ± 1.43 μM, literature[36] IC_{50} = 25 ± 3.5 mM) (**Figure 3.65** and **Figure 3.67**). Each of the purchased compounds was screened at two concentration, high (500 or 250 μM) and low (5 or 2.5 μM). Some of the molecules were dissolvable only in a mixture of water and DMSO or HCl; the highest used concentration of these solvents served as a negative control.

In the drug screen, of the 8 compounds tested, ticagrelor and sonedenoson significantly inhibited the uptake of uridine (**Figure 3.66**). Ticagrelor is an oral, reversibly binding $P2Y_{12}$ receptor antagonist that blocks ADP-induced platelet aggregation.[105] It is marketed in the United States by AstraZeneca under the name Brilinta. It has replaced Plavix (clopidogrel) as the more efficacious drug for patients with acute coronary syndrome.[106] Sonedenoson is an $A_{2A}AR$ agonist that was clinically evaluated by King Pharmaceuticals as a prospective new topical drug for the treatment of chronic, neuropathic, diabetic foot ulcers.[107] Its development has since been terminated due to efficacy problems.[108] With that in mind, we have chosen not to follow up on sonedenoson.

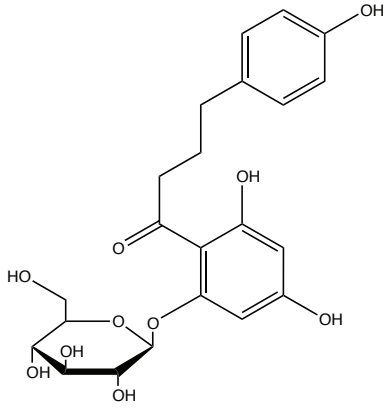
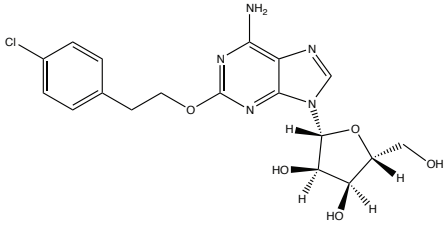
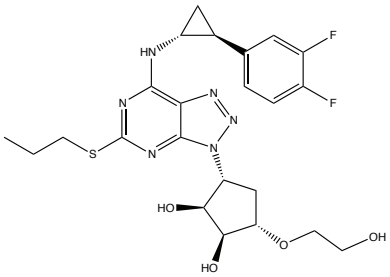


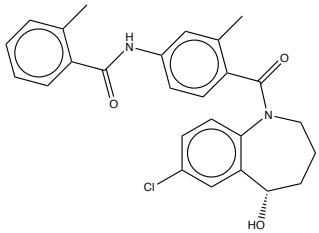
Figure 3.62 Molecular structure of phloridzin.



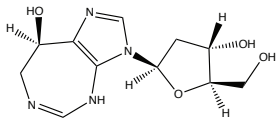
(2*R*,3*R*,4*S*,5*R*)-2-(6-amino-2-(4-chlorophenethoxy)-9*H*-purin-9-yl)-5-(hydroxymethyl)tetrahydrofuran-3,4-diol
sonedenoson
ZINC000003966398



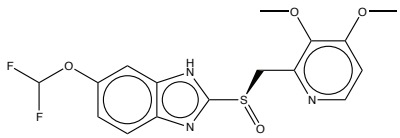
(1*S*,2*S*,3*R*,5*S*)-3-(7-(((1*R*,2*S*)-2-(3,4-difluorophenyl)cyclopropyl)amino)-5-(propylthio)-3*H*-[1,2,3]triazolo[4,5-*d*]pyrimidin-3-yl)-5-(2-hydroxyethoxy)cyclopentane-1,2-diol
ticagrelor
ZINC000028957444



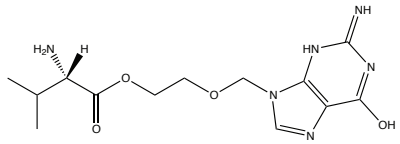
(*S*)-*N*-(4-(7-chloro-5-hydroxy-2,3,4,5-tetrahydro-1*H*-benzo[*b*]azepine-1-carbonyl)-3-methylphenyl)-2-methylbenzamide
tolvaptan
ZINC000001490477



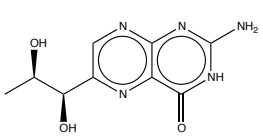
(*R*)-3-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-3,4,7,8-tetrahydroimidazo[4,5-*d*][1,3]diazepin-8-ol
pentostatin
ZINC000003806262



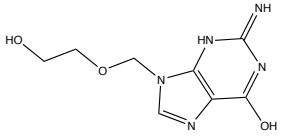
(*R*)-6-(difluoromethoxy)-2-(((3,4-dimethoxypyridin-2-yl)methyl)sulfinyl)-1*H*-benzo[*d*]imidazole
pantoprazole
ZINC000004676424



(*S*)-2-((6-hydroxy-2-imino-2*H*-purin-9(3*H*)-yl)methoxy)ethyl 2-amino-3-methylbutanoate
valacyclovir
ZINC000001530713

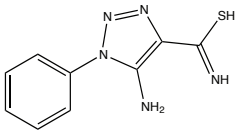


2-amino-6-((1*R*,2*R*)-1,2-dihydroxypropyl)pteridin-4(3*H*)-one
biopterin
ZINC000017129259

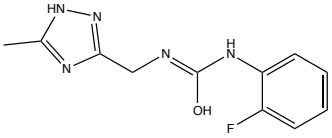


9-((2-hydroxyethoxy)methyl)-2-imino-3,9-dihydro-2*H*-purin-6-ol
acyclovir
ZINC000001530555

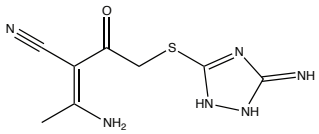
Figure 3.63 Molecular structure of compounds selected from the virtual screen of the KEGG DRUG database.



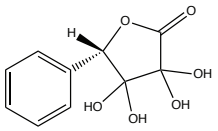
5-amino-1-phenyl-1*H*-1,2,3-triazole-4-carbimidothioic acid
ZINC000000383878



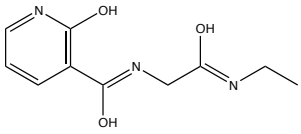
(*Z*)-*N*-(2-fluorophenyl)-*N'*-((5-methyl-1*H*-1,2,4-triazol-3-yl)methyl)carbamimidic acid
ZINC000075139316



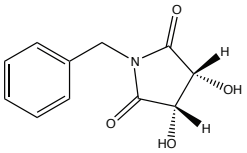
(*Z*)-3-amino-2-(2-((5-imino-2,5-dihydro-1*H*-1,2,4-triazol-3-yl)thio)acetyl)but-2-enenitrile
ZINC000006312152



(*S*)-3,3,4,4-tetrahydroxy-5-phenyldihydrofuran-2(3*H*)-one
ZINC000004114169



(*Z*)-*N*-((*Z*)-2-(ethylimino)-2-hydroxyethyl)-2-hydroxynicotinimidic acid
ZINC000040039580



(3*S*,4*R*)-1-benzyl-3,4-dihydroxypyrrolidine-2,5-dione
ZINC000012637565

Figure 3.64 Molecular structures of compounds selected from the virtual screen of ZINC library of fragments.

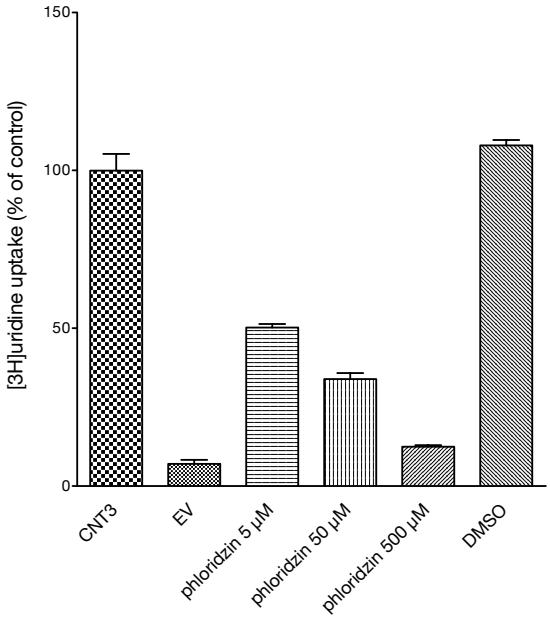


Figure 3.65 Cellular uptake results testing the stable PK15NTD cell line overexpressing hCNT3. Phloridzin, a known hCNT3 inhibitor, represents the positive control. The DMSO negative control is also shown. Error bars represent s.e.m. Each condition was replicated 3 times.

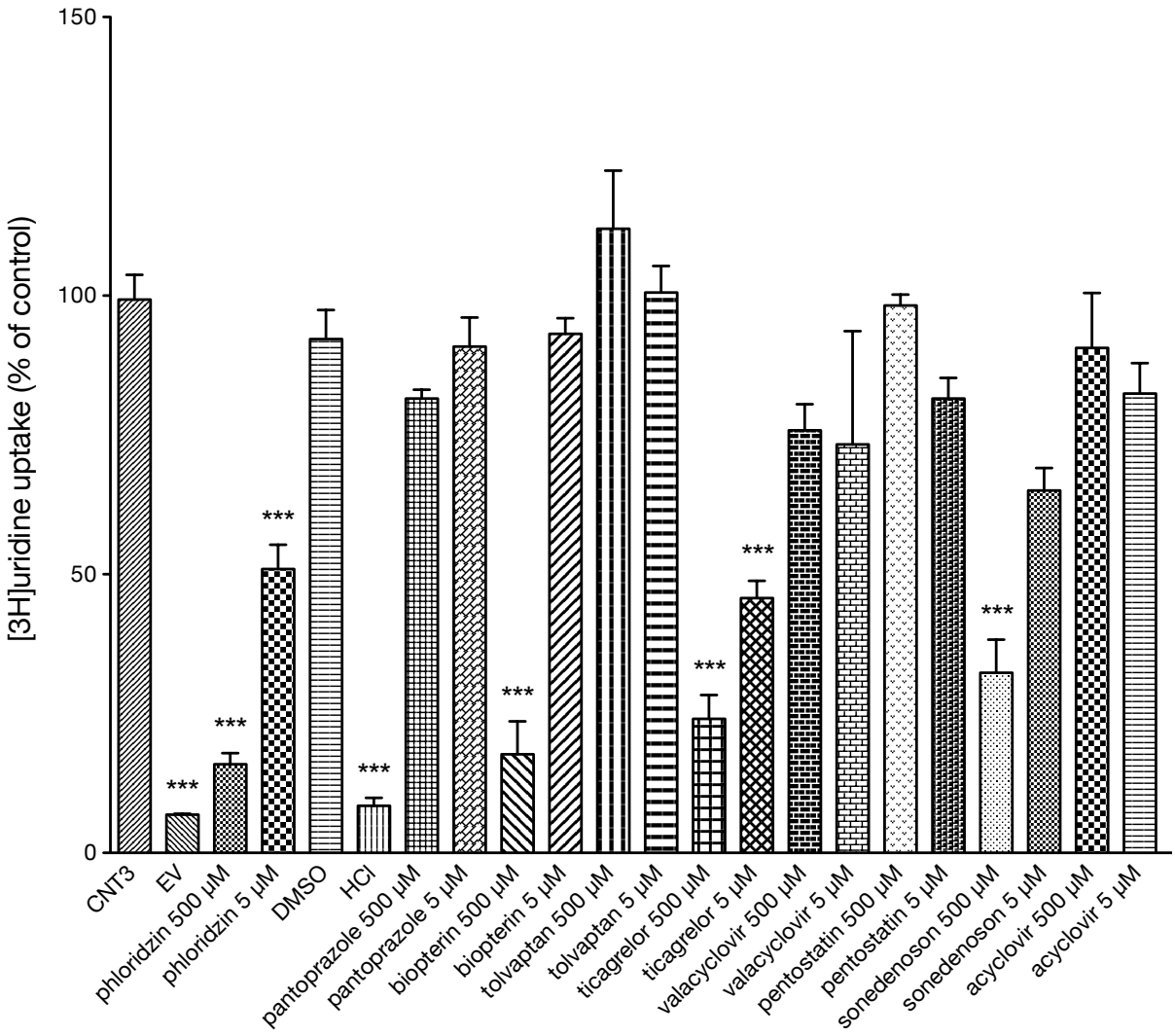


Figure 3.66 Cellular uptake results testing the compounds purchased from the KEGG DRUG screen. Normalized [3H]uridine uptake in PK15NTD cells stably transfected with hCNT3 with and without the addition of compounds predicted to bind from the virtual screen. Phloridzin, a known hCNT3 inhibitor, represents the positive control. *** represents p-value < 0.001. Error bars represent s.e.m. Each condition was replicated 3 times.

The potency of ticagrelor was further examined by determining the IC_{50} values for uridine and adenosine accumulation (**Figure 3.67** and **Figure 3.68**). It proved to be equally as potent at uridine uptake inhibition as the standard phloridzin, at IC_{50} of 6.47 ± 1.27 μM vs 7.40 ± 1.43 μM, respectively. For adenosine uptake, ticagrelor was significantly less effective as an inhibitor, with an IC_{50} of 40.2 ± 1.4 μM vs an IC_{50} of 16.2 ± 1.3 μM for phloridzin. This finding is in agreement with the K_M values for hCNT3, with adenosine having a lower K_M (15.1 ± 1.8 μM adenosine vs 21.6 ± 5.4 μM uridine),[4] which indicates that hCNT3 has a higher affinity for adenosine than for uridine. The difference in affinity stems from a difference in interaction between the purine and pyrimidine nucleobases and the CNT3 binding pocket residues. Also, ticagrelor does not appear to be selective for hCNT3, because it also inhibits hCNT1- and hCNT2-mediated uridine uptake (**Figure 3.69**).

Recent work involving ticagrelor has independently discovered its adenosine-mediated effects through the interaction with nucleoside transporters, in addition to its interaction with the primary target, the $P2Y_{12}$ receptor.[106] Ticagrelor was shown to inhibit adenosine uptake in human erythrocytes in a dose-dependent manner.[109] Further work has implicated ENT1 as the nucleoside transporter most potently inhibited by the drug, with an adenosine uptake IC_{50} of 0.26 μM. The inhibition of ENT1 on platelets leads to an accumulation of extracellular adenosine and a subsequent activation of adenosine A_{2A} receptors, culminating in an increase in basal cAMP and vasodilator-stimulated phosphoprotein phosphorylation (VASP-P).[110] Its interaction with hCNT3 could therefore be of secondary importance, as clinically relevant concentrations of ticagrelor never reach hCNT3 inhibition levels.[111] Based on its size and large functional groups, it is unlikely to be a hCNT3 substrate.

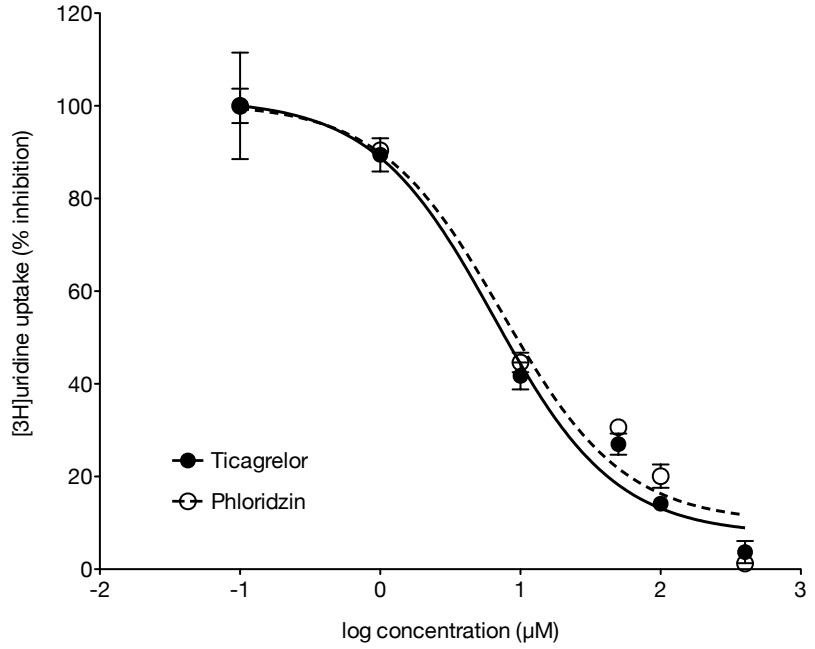


Figure 3.67 The inhibitory effect of ticagrelor and phloridzin on CNT3-mediated [3H]uridine uptake. Each data point was replicated 3 times (IC_{50} ticagrelor 6.47 ± 1.27 μM, IC_{50} phloridzin 7.40 ± 1.43 μM).

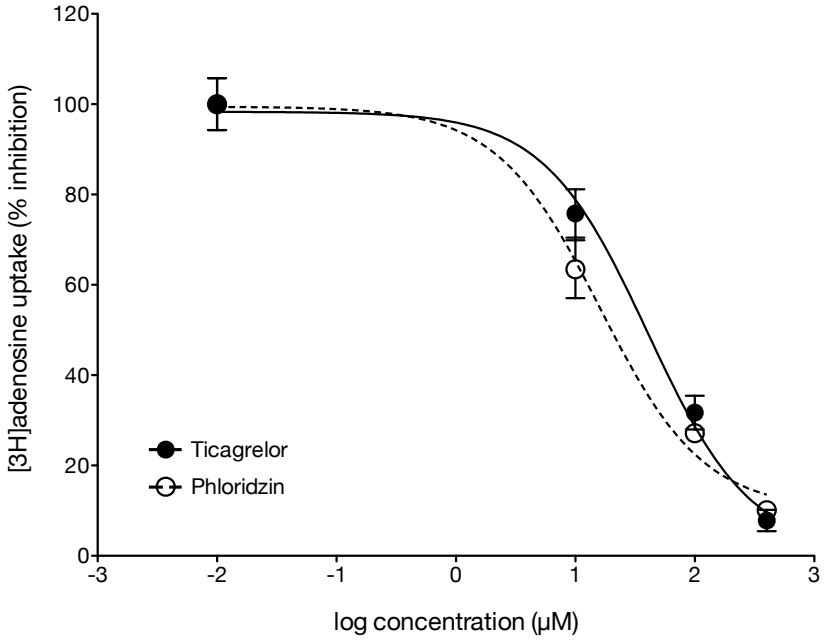


Figure 3.68 The inhibitory effect of ticagrelor and phloridzin on CNT3-mediated [3H]adenosine uptake. Each data point was replicated 3 times (IC_{50} ticagrelor = 40.19 ± 1.40 μM, IC_{50} phloridzin = 16.15 ± 1.33 μM).

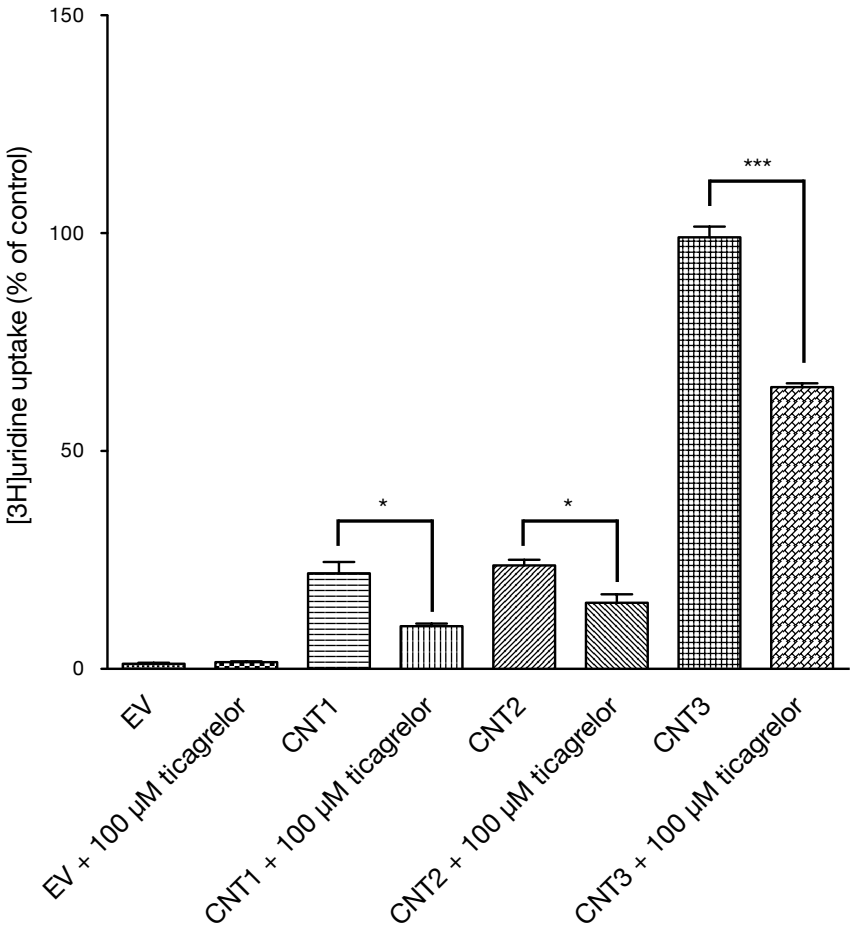


Figure 3.69 Cellular uptake results testing the specificity of ticagrelor against human concentrative nucleoside transporters. Normalized [3H]uridine uptake in PK15NTD cells stably transfected with hCNT1, hCNT2, and hCNT3 with and without the addition of 100 µM ticagrelor. *** represents p-value < 0.001. * represents p-value < 0.05. Error bars represent s.e.m. Each condition was replicated 3 times.

In the fragment screen, of the 6 compounds tested, only ZINC000000383878, a 1,2,3-triazole, exhibited mild uridine inhibition (**Figure 3.70**). 1,2,3-triazoles and their derivatives are important in medicinal chemistry for combinatorial library synthesis and have been shown to be effective as antivirals, antibacterials, antifungals, anticonvulsants, antidepressants, and antineoplastics.[112,113] As mentioned above, fragments are non-functionalized compounds and therefore are typically weak ligands.[104] Due to the small size of the compound, we hypothesized that it could be a substrate, which would explain the mild, but significant inhibition of uridine uptake. To test this hypothesis, we radiolabeled the compound. Despite a promising virtual screen pose and preliminary data, the hCNT3-mediated uptake of ZINC000000383878 was not different between PK15NTD cells overexpressing hCNT3 and negative control (**Figure 3.71**). Thus, due to weak inhibition at a high *in vitro* concentration of 500 μM and no inhibition at a physiologically high concentration of 5 μM, we decided not to pursue the compound as an inhibitor.

Additional hits from the virtual screen await purchasing and testing.

Conclusion

Only compounds with the top 500 poses from each screen were examined, and of those, only 14 in total were purchased. Despite being largely unsuccessful in finding high impact substrates or inhibitors so far, the screen shows promise, as it correctly identified ticagrelor as an inhibitor. This work is part of an ongoing collaboration between the Giacomini and Sali labs with the goal of finding novel ligands for human SLCs, and will be continued further.

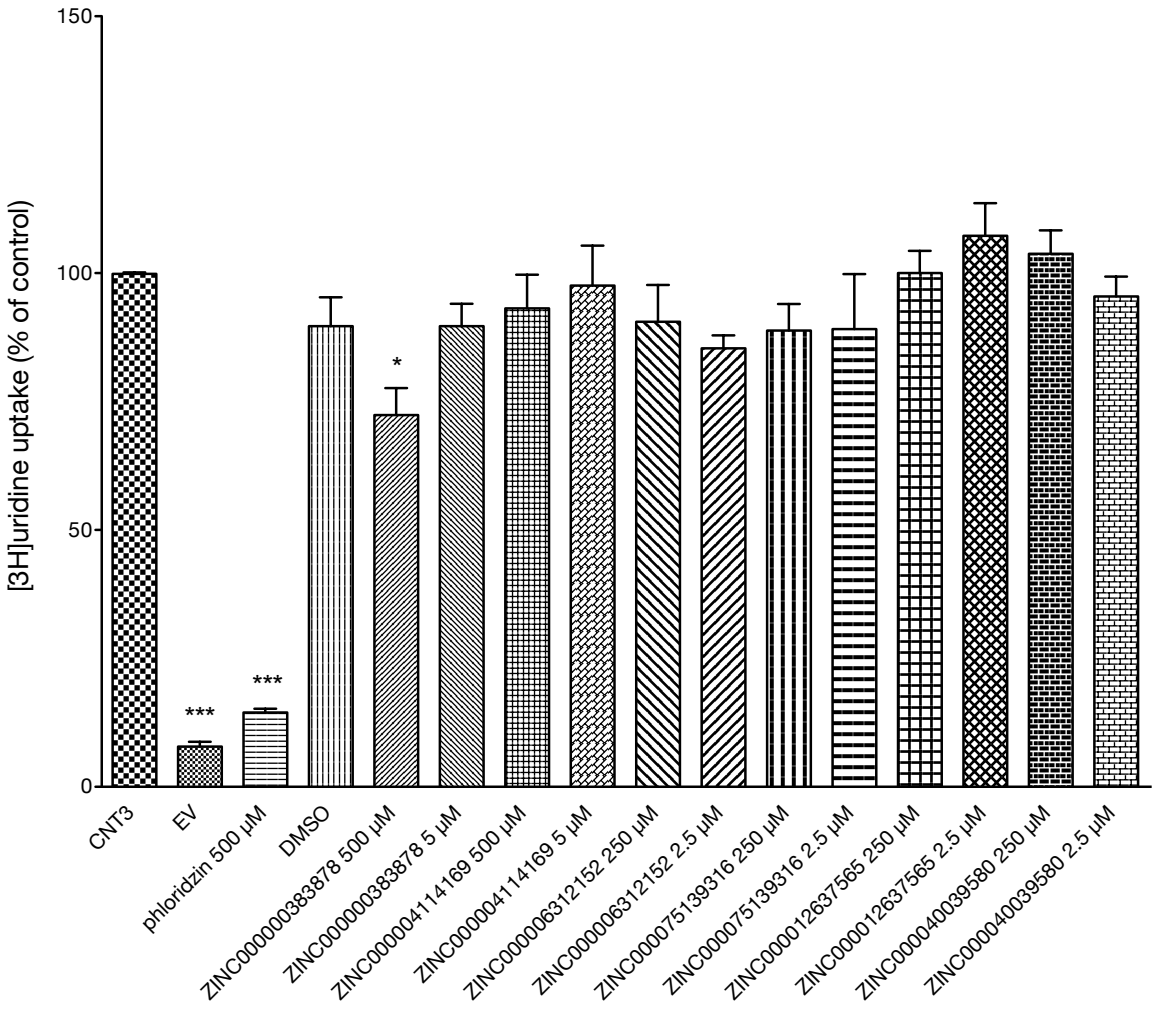


Figure 3.70 Cellular uptake results testing the compounds purchased from the ZINC fragment screen. Normalized [3H]uridine uptake in PK15NTD cells stably transfected with hCNT3 with and without the addition of compounds predicted to bind from the virtual screen. Phloridzin, a known hCNT3 inhibitor, represents the positive control. * represents p-value < 0.05. *** represents p-value < 0.001. Error bars represent s.e.m. Each condition was replicated 3 times.

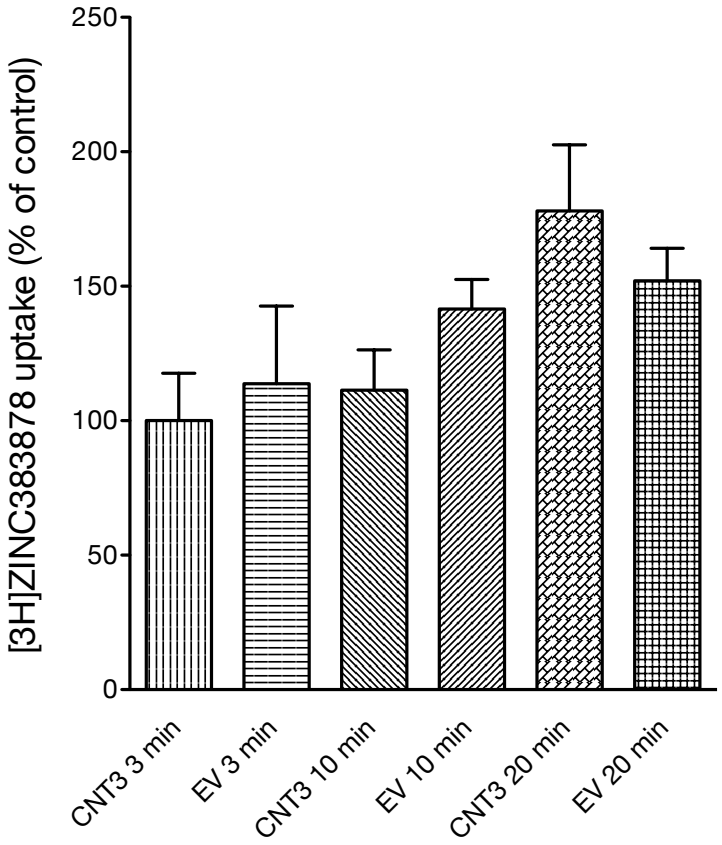


Figure 3.71 Cellular uptake results testing the radiolabeled ZINC000000383878. Normalized [3H]ZINC000000383878 uptake in PK15NTD cells (EV) and PK15NTD cells stably transfected with hCNT3 (CNT3). Error bars represent s.e.m. Each condition was replicated 3 times.

References

1. Wang J, Su SF, Dresser MJ, Schaner ME, Washington CB, Giacomini KM. Na(+)-dependent purine nucleoside transporter from human kidney: cloning and functional characterization. *Am J Physiol*. 1997;273(6 Pt 2):F1058-F1065.
2. Wang J, Giacomini KM. Molecular determinants of substrate selectivity in Na+-dependent nucleoside transporters. *J Biol Chem*. 1997;272(46):28845-28848.
3. Ritzel MW, Yao SY, Huang MY, Elliott JF, Cass CE, Young JD. Molecular cloning and functional expression of cDNAs encoding a human Na+-nucleoside cotransporter (hCNT1). *Am J Physiol*. 1997;272(2 Pt 1):C707-C714.
4. Ritzel MW, Ng AM, Yao SY, et al. Molecular identification and characterization of novel human and mouse concentrative Na+-nucleoside cotransporter proteins (hCNT3 and mCNT3) broadly selective for purine and pyrimidine nucleosides (system cib). *J Biol Chem*. 2001;276(4):2914-2927. doi:10.1074/jbc.M007746200.
5. Baldwin SA, Beal PR, Yao SYM, King AE, Cass CE, Young JD. The equilibrative nucleoside transporter family, SLC29. *Pflugers Arch*. 2004;447(5):735-743. doi:10.1007/s00424-003-1103-2.
6. Barnes K, Dobrzynski H, Foppolo S, et al. Distribution and functional characterization of equilibrative nucleoside transporter-4, a novel cardiac adenosine transporter activated at acidic pH. *Circ Res*. 2006;99(5):510-519. doi:10.1161/01.RES.0000238359.18495.42.
7. Gray JH, Mangravite LM, Owen RP, et al. Functional and genetic diversity in the concentrative nucleoside transporter, CNT1, in human populations. *Mol Pharmacol*. 2004;65(3):512-519. doi:10.1124/mol.65.3.512.
8. Gray JH, Owen RP, Giacomini KM. The concentrative nucleoside transporter family, SLC28. *Pflugers Arch*. 2004;447(5):728-734. doi:10.1007/s00424-003-1107-y.
9. Zhang J, Smith KM, Tackaberry T, et al. Uridine binding and transportability determinants of human concentrative nucleoside transporters. *Mol Pharmacol*. 2005;68(3):830-839. doi:10.1124/mol.105.012187.
10. Pastor-Anglada M, Cano-Soldado P, Errasti-Murugarren E, Casado FJ. SLC28 genes and concentrative nucleoside transporter (CNT) proteins. *Xenobiotica*. 2008;38(7-8):972-994. doi:10.1080/00498250802069096.
11. Aymerich I, Duflot S, Fernández-Veledo S, et al. The concentrative nucleoside transporter family (SLC28): new roles beyond salvage? *Biochem Soc Trans*. 2005;33(Pt 1):216-219.

References (continued)

doi:10.1042/BST0330216.

12. Errasti-Murugarren E, Pastor-Anglada M. Drug transporter pharmacogenetics in nucleoside-based therapies. *Pharmacogenomics*. 2010;11(6):809-841. doi:10.2217/pgs.10.70.
13. Galmarini CM, Mackey JR, Dumontet C. Nucleoside analogues: mechanisms of drug resistance and reversal strategies. *Leukemia*. 2001;15(6):875-890.
14. Johnson ZL, Cheong C-G, Lee S-Y. Crystal structure of a concentrative nucleoside transporter from Vibrio cholerae at 2.4 Å. *Nature*. 2012;483(7390):489-493. doi:10.1038/nature10882.
15. Veenhoff LM, Heuberger EHML, Poolman B. Quaternary structure and function of transport proteins. *Trends Biochem Sci*. 2002;27(5):242-249.
16. Santos-Rodrigues dos A, Grañé-Boladeras N, Bicket A, Coe IR. Nucleoside transporters in the purinome. *Neurochemistry International*. April 2014. doi:10.1016/j.neuint.2014.03.014.
17. Young JD, Yao SYM, Baldwin JM, Cass CE, Baldwin SA. The human concentrative and equilibrative nucleoside transporter families, SLC28 and SLC29. *Molecular Aspects of Medicine*. 2013;34(2-3):529-547. doi:10.1016/j.mam.2012.05.007.
18. Hediger MA, Clémençon B, Burrier RE, Bruford EA. The ABCs of membrane transporters in health and disease (SLC series): introduction. *Molecular Aspects of Medicine*. 2013;34(2-3):95-107. doi:10.1016/j.mam.2012.12.009.
19. Mackey JR, Galmarini CM, Graham KA, et al. Quantitative analysis of nucleoside transporter and metabolism gene expression in chronic lymphocytic leukemia (CLL): identification of fludarabine-sensitive and -insensitive populations. *Blood*. 2005;105(2):767-774. doi:10.1182/blood-2004-03-1046.
20. Okazaki T, Javle M, Tanaka M, Abbruzzese JL, Li D. Single nucleotide polymorphisms of gemcitabine metabolic genes and pancreatic cancer survival and drug toxicity. *Clin Cancer Res*. 2010;16(1):320-329. doi:10.1158/1078-0432.CCR-09-1555.
21. Slugoski MD, Smith KM, Mulinta R, et al. A conformationally mobile cysteine residue (Cys-561) modulates Na+ and H+ activation of human CNT3. *J Biol Chem*. 2008;283(36):24922-24934. doi:10.1074/jbc.M801793200.
22. Slugoski MD, Smith KM, Ng AML, et al. Conserved glutamate residues Glu-343 and Glu-519 provide mechanistic insights into cation/nucleoside cotransport by human concentrative nucleoside transporter hCNT3. *J Biol Chem*. 2009;284(25):17266-17280. doi:10.1074/jbc.M109.009613.

References (continued)

23. Slugoski MD, Ng AML, Yao SYM, et al. Substituted Cysteine Accessibility Method Analysis of Human Concentrative Nucleoside Transporter hCNT3 Reveals a Novel Discontinuous Region of Functional Importance within the CNT Family Motif (G/A)XKX3NEFVA(Y/M/F). *J Biol Chem*. 2009;284(25):17281-17292. doi:10.1074/jbc.M109.009704.
24. Slugoski MD, Ng AML, Yao SYM, et al. A proton-mediated conformational shift identifies a mobile pore-lining cysteine residue (Cys-561) in human concentrative nucleoside transporter 3. *J Biol Chem*. 2008;283(13):8496-8507. doi:10.1074/jbc.M710433200.
25. Slugoski MD, Loewen SK, Ng AML, et al. Specific mutations in transmembrane helix 8 of human concentrative Na+/nucleoside cotransporter hCNT1 affect permeant selectivity and cation coupling. *Biochemistry*. 2007;46(6):1684-1693. doi:10.1021/bi061692s.
26. Forrest LR. Structural Symmetry in Membrane Proteins. *Annu Rev Biophys*. 2015;44(1):311-337. doi:10.1146/annurev-biophys-051013-023008.
27. Perez C, Khafizov K, Forrest LR, Krämer R, Ziegler C. The role of trimerization in the osmoregulated betaine transporter BetP. *EMBO Rep*. 2011;12(8):804-810. doi:10.1038/embor.2011.102.
28. Perez C, Koshy C, Yildiz Ö, Ziegler C. Alternating-access mechanism in conformationally asymmetric trimers of the betaine transporter BetP. *Nature*. 2012;490(7418):126-130. doi:10.1038/nature11403.
29. Tsai C-J, Khafizov K, Hakulinen J, et al. Structural asymmetry in a trimeric Na+/betaine symporter, BetP, from Corynebacterium glutamicum. *J Mol Biol*. 2011;407(3):368-381. doi:10.1016/j.jmb.2011.01.028.
30. Kim J, Wu S, Tomasiak TM, et al. Subnanometre-resolution electron cryomicroscopy structure of a heterodimeric ABC exporter. *Nature*. November 2014. doi:10.1038/nature13872.
31. Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol*. 1993;234(3):779-815. doi:10.1006/jmbi.1993.1626.
32. Pei J, Kim B-H, Grishin NV. PROMALS3D: a tool for multiple protein sequence and structure alignments. *Nucleic Acids Research*. 2008;36(7):2295-2300. doi:10.1093/nar/gkn072.
33. Shen M-Y, Sali A. Statistical potential for assessment and prediction of protein structures. *Protein Sci*. 2006;15(11):2507-2524. doi:10.1110/ps.062416606.
34. Aran JM, Plagemann PG. High-affinity, equilibrative nucleoside transporter of pig kidney cell line (PK-15). *Biochim Biophys Acta*. 1992;1108(1):67-74.

References (continued)

35. Ward JL, Sherali A, Mo ZP, Tse CM. Kinetic and pharmacological properties of cloned human equilibrative nucleoside transporters, ENT1 and ENT2, stably expressed in nucleoside transporter-deficient PK15 cells. Ent2 exhibits a low affinity for guanosine and cytidine but a high affinity for inosine. *J Biol Chem*. 2000;275(12):8375-8381.
36. Wang C, Pimple S, Buolamwini JK. Interaction of fused-pyrimidine nucleoside analogs with human concentrative nucleoside transporters: High-affinity inhibitors of human concentrative nucleoside transporter 1. *Cell*. 2011;81(1):82-90. doi:10.1016/j.bcp.2010.09.009.
37. Hastrup H, Karlin A, Javitch JA. Symmetrical dimer of the human dopamine transporter revealed by cross-linking Cys-306 at the extracellular end of the sixth transmembrane segment. *Proc Natl Acad Sci USA*. 2001;98(18):10055-10060. doi:10.1073/pnas.181344298.
38. Hoque KM, Chen L, Leung GPH, Tse CM. A purine-selective nucleobase/nucleoside transporter in PK15NTD cells. *AJP: Regulatory, Integrative and Comparative Physiology*. 2008;294(6):R1988-R1995. doi:10.1152/ajpregu.00016.2008.
39. Mangravite LM, Giacomini KM. Sorting of Rat SPNT in Renal Epithelium Is Independent of N-Glycosylation. *Pharm Res*. 2003;20(2):319-323. doi:10.1023/A:1022247826750.
40. Keller T, Egenberger B, Gorboulev V, et al. The large extracellular loop of organic cation transporter 1 influences substrate affinity and is pivotal for oligomerization. *J Biol Chem*. 2011;286(43):37874-37886. doi:10.1074/jbc.M111.289330.
41. Hughson AG, Lee GF, Hazelbauer GL. Analysis of protein structure in intact cells: Crosslinking in vivo between introduced cysteines in the transmembrane domain of a bacterial chemoreceptor. *Protein Sci*. 1997;6(2):315-322. doi:10.1002/pro.5560060206.
42. Careaga CL, Falke JJ. Thermal motions of surface alpha-helices in the D-galactose chemosensory receptor. Detection by disulfide trapping. *J Mol Biol*. 1992;226(4):1219-1235.
43. Bass RB, Butler SL, Chervitz SA, Gloor SL, Falke JJ. Use of site-directed cysteine and disulfide chemistry to probe protein structure and dynamics: applications to soluble and transmembrane receptors of bacterial chemotaxis. *Meth Enzymol*. 2007;423:25-51. doi:10.1016/S0076-6879(07)23002-2.
44. Careaga CL, Falke JJ. Structure and dynamics of Escherichia coli chemosensory receptors. Engineered sulfhydryl studies. *Biophys J*. 1992;62(1):209–16–discussion217–9. doi:10.1016/S0006-3495(92)81806-4.
45. Stansfeld PJ, Goose JE, Caffrey M, et al. MemProtMD: Automated Insertion of Membrane

References (continued)

Protein Structures into Explicit Lipid Membranes. *Structure*. 2015;23(7):1350-1361. doi:10.1016/j.str.2015.05.006.

46. Johnson ZL, Lee J-H, Lee K, et al. Structural basis of nucleoside and nucleoside drug selectivity by concentrative nucleoside transporters. *Elife*. July 2014:e03604. doi:10.7554/eLife.03604.
47. Vergara-Jaque A, Fenollar-Ferrer C, Kaufmann D, Forrest LR. Repeat-swap homology modeling of secondary active transporters: updated protocol and prediction of elevator-type mechanisms. *Front Pharmacol*. 2015;6:183. doi:10.3389/fphar.2015.00183.
48. Ovchinnikov S, Kamisetty H, Baker D. Robust and accurate prediction of residue-residue interactions across protein interfaces using evolutionary information. *Elife*. 2014;3:e02030. doi:10.7554/eLife.02030.
49. Krishnamurthy H, Piscitelli CL, Gouaux E. Unlocking the molecular secrets of sodium-coupled transporters. *Nature*. 2009;459(7245):347-355. doi:10.1038/nature08143.
50. Capra JA, Singh M. Predicting functionally important residues from sequence conservation. *Bioinformatics*. 2007;23(15):1875-1882. doi:10.1093/bioinformatics/btm270.
51. Yernool D, Boudker O, Jin Y, Gouaux E. Structure of a glutamate transporter homologue from Pyrococcus horikoshii. *Nature*. 2004;431(7010):811-818. doi:10.1038/nature03018.
52. Levy ED, Pereira-Leal JB, Chothia C, Teichmann SA. 3D Complex: A Structural Classification of Protein Complexes. *PLoS Comp Biol*. 2006;2(11):e155. doi:10.1371/journal.pcbi.0020155.
53. André I, Strauss CEM, Kaplan DB, Bradley P, Baker D. Emergence of symmetry in homooligomeric biological assemblies. *Proc Natl Acad Sci USA*. 2008;105(42):16148-16152. doi:10.1073/pnas.0807576105.
54. Goodsell DS, Olson AJ. Structural symmetry and protein function. *Annu Rev Biophys Biomol Struct*. 2000;29:105-153. doi:10.1146/annurev.biophys.29.1.105.
55. Monod J, Wyman J, Changeux J-P. On the Nature of Allosteric Transitions: A Plausible Model. *J Mol Biol*. 1965;12:88-118.
56. Vangone A, Spinelli R, Scarano V, Cavallo L, Oliva R. COCOMAPS: a web application to analyze and visualize contacts at the interface of biomolecular complexes. *Bioinformatics*. 2011;27(20):2915-2916. doi:10.1093/bioinformatics/btr484.
57. Wells JA, McClendon CL. Reaching for high-hanging fruit in drug discovery at protein–protein interfaces. *Nature*. 2007;450(7172):1001-1009.

References (continued)

58. Almén MS, Nordström KJV, Fredriksson R, Schiöth HB. Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin. *BMC Biol*. 2009;7(1):50. doi:10.1186/1741-7007-7-50.
59. Pieper U, Schlessinger A, Kloppmann E, et al. Coordinating the impact of structural genomics on the human α-helical transmembrane proteome. *Nat Struct Mol Biol*. 2013;20(2):135-138. doi:10.1038/nsmb.2508.
60. Stroud RM. New tools in membrane protein determination. *F1000 Biol Rep*. 2011;3(8):8. doi:10.3410/B3-8.
61. Sud M, Fahy E, Cotter D, et al. LMSD: LIPID MAPS structure database. *Nucleic Acids Research*. 2007;35(Database issue):D527-D532. doi:10.1093/nar/gkl838.
62. van Meer G, Voelker DR, Feigenson GW. Membrane lipids: where they are and how they behave. *Nature Reviews Molecular Cell Biology*. 2008;9(2):112-124. doi:10.1038/nrm2330.
63. Trometer C, Falson P. Mammalian membrane protein expression in baculovirus-infected insect cells. *Methods Mol Biol*. 2010;601(Chapter 7):105-117. doi:10.1007/978-1-60761-344-2_7.
64. Jarvis DL, Finn EE. Modifying the insect cell N-glycosylation pathway with immediate early baculovirus expression vectors. *Nature Biotechnology*. 1996;14(10):1288-1292. doi:10.1038/nbt1096-1288.
65. Stroud RM, Schertler GFX. Membranes. *Curr Opin Struct Biol*. 2011;21(4):495-496. doi:10.1016/j.sbi.2011.08.001.
66. Chaudhary S, Pak JE, Pedersen BP, et al. Efficient expression screening of human membrane proteins in transiently transfected Human Embryonic Kidney 293S cells. *Methods*. 2011;55(4):273-280. doi:10.1016/j.ymeth.2011.08.018.
67. Chaudhary S, Pak JE, Gruswitz F, Sharma V, Stroud RM. Overexpressing human membrane proteins in stably transfected and clonal human embryonic kidney 293S cells. *Nature Protocols*. 2012;7(3):453-466. doi:10.1038/nprot.2011.453.
68. Long SB, Tao X, Campbell EB, MacKinnon R. Atomic structure of a voltage-dependent K+ channel in a lipid membrane-like environment. *Nature*. 2007;450(7168):376-382. doi:10.1038/nature06265.
69. Thompson AA, Liu JJ, Chun E, et al. GPCR stabilization using the bicelle-like architecture of mixed sterol-detergent micelles. *Methods*. 2011;55(4):310-317. doi:10.1016/j.ymeth.2011.10.011.
70. Chae PS, Rasmussen SGF, Rana RR, et al. Maltose-neopentyl glycol (MNG) amphiphiles for

References (continued)

solubilization, stabilization and crystallization of membrane proteins. *Nat Methods*. 2010;7(12):1003-1008. doi:10.1038/nmeth.1526.

71. Popot J-L, Althoff T, Bagnard D, et al. Amphipols from A to Z. *Annu Rev Biophys*. 2011;40(1):379-408. doi:10.1146/annurev-biophys-042910-155219.
72. Liao M, Cao E, Julius D, Cheng Y. Structure of the TRPV1 ion channel determined by electron cryo-microscopy. *Nature*. 2013;504(7478):107-112. doi:10.1038/nature12822.
73. Li X, Mooney P, Zheng S, et al. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat Methods*. 2013;10(6):584-590. doi:10.1038/nmeth.2472.
74. Nogales E, Scheres SHW. Cryo-EM: A Unique Tool for the Visualization of Macromolecular Complexity. *Mol Cell*. 2015;58(4):677-689. doi:10.1016/j.molcel.2015.02.019.
75. Wu S, Avila-Sakar A, Kim J, et al. Fabs enable single particle cryoEM studies of small proteins. *Structure*. 2012;20(4):582-592. doi:10.1016/j.str.2012.02.017.
76. Tomasiak TM, Pedersen BP, Chaudhary S, et al. General qPCR and Plate Reader Methods for Rapid Optimization of Membrane Protein Purification and Crystallization Using Thermostability Assays. - PubMed - NCBI. *Curr Protoc Protein Sci*. 2001;77:29.11.1-29.11.14. doi:10.1002/0471140864.ps2911s77.
77. Loewen SK, Ng AM, Yao SY, Cass CE, Baldwin SA, Young JD. Identification of amino acid residues responsible for the pyrimidine and purine nucleoside specificities of human concentrative Na(+) nucleoside cotransporters hCNT1 and hCNT2. *J Biol Chem*. 1999;274(35):24475-24484.
78. Xiao G, Wang J, Tangen T, Giacomini KM. A novel proton-dependent nucleoside transporter, CeCNT3, from Caenorhabditis elegans. *Mol Pharmacol*. 2001;59(2):339-348.
79. Possee RD. Baculoviruses as expression vectors. *Curr Opin Biotechnol*. 1997;8(5):569-572.
80. Deng D, Xu C, Sun P, et al. Crystal structure of the human glucose transporter GLUT1. *Nature*. 2014;510(7503):121-125. doi:10.1038/nature13306.
81. Paulsen CE, Armache J-P, Gao Y, Cheng Y, Julius D. Structure of the TRPA1 ion channel suggests regulatory mechanisms. *Nature*. April 2015. doi:10.1038/nature14367.
82. Seddon AM, Curnow P, Booth PJ. Membrane proteins, lipids and detergents: not just a soap opera. *Biochimica et Biophysica Acta (BBA) - Biomembranes*. 2004;1666(1-2):105-117. doi:10.1016/j.bbamem.2004.04.011.

References (continued)

83. Strop P, Brunger AT. Refractive index-based determination of detergent concentration and its application to the study of membrane proteins. *Protein Sci*. 2005;14(8):2207-2211. doi:10.1110/ps.051543805.
84. McGuffin LJ, Bryson K, Jones DT. The PSIPRED protein structure prediction server. *Bioinformatics*. 2000;16(4):404-405.
85. Ward JJ, McGuffin LJ, Bryson K, Buxton BF, Jones DT. The DISOPRED server for the prediction of protein disorder. *Bioinformatics*. 2004;20(13):2138-2139. doi:10.1093/bioinformatics/bth195.
86. Sachdev D, Chirgwin JM. Properties of soluble fusions between mammalian aspartic proteinases and bacterial maltose-binding protein. *J Protein Chem*. 1999;18(1):127-136.
87. Terpe K. Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Appl Microbiol Biotechnol*. 2003;60(5):523-533. doi:10.1007/s00253-002-1158-6.
88. Elwi AN, Damaraju VL, Kuzma ML, et al. Human concentrative nucleoside transporter 3 is a determinant of fludarabine transportability and cytotoxicity in human renal proximal tubule cell cultures. *Cancer Chemother Pharmacol*. 2009;63(2):289-301. doi:10.1007/s00280-008-0739-1.
89. Wang J, Giacomini KM. Serine 318 is essential for the pyrimidine selectivity of the N2 Na+-nucleoside transporter. *J Biol Chem*. 1999;274(4):2298-2302.
90. Yao SYM, Ng AML, Slugoski MD, et al. Conserved glutamate residues are critically involved in Na+/nucleoside cotransport by human concentrative nucleoside transporter 1 (hCNT1). *J Biol Chem*. 2007;282(42):30607-30617. doi:10.1074/jbc.M703285200.
91. Cass CE, Gaudette LA, Paterson AR. Mediated transport of nucleosides in human erythrocytes. Specific binding of the inhibitor nitrobenzylthioinosine to nucleoside transport sites in the erythrocyte membrane. *Biochim Biophys Acta*. 1974;345(1):1-10.
92. Crespy V, Aprikian O, Morand C, et al. Bioavailability of phloretin and phloridzin in rats. *J Nutr*. 2001;131(12):3227-3230.
93. Gupte A, Buolamwini JK. Synthesis and biological evaluation of phloridzin analogs as human concentrative nucleoside transporter 3 (hCNT3) inhibitors. *Bioorg Med Chem Lett*. 2009;19(3):917-921. doi:10.1016/j.bmcl.2008.11.112.
94. Damaraju VL, Smith KM, Mowles D, et al. Interaction of fused-pyrimidine nucleoside analogs

References (continued)

with human concentrative nucleoside transporters: High-affinity inhibitors of human concentrative nucleoside transporter 1. *Biochem Pharmacol*. 2011;81(1):82-90. doi:10.1016/j.bcp.2010.09.009.

95. Pei J, Kim B-H, Grishin NV. PROMALS3D: a tool for multiple protein sequence and structure alignments. *Nucleic Acids Research*. 2008;36(7):2295-2300. doi:10.1093/nar/gkn072.
96. Dong GQ, Fan H, Schneidman-Duhovny D, Webb B, Sali A. Optimized atomic statistical potentials: assessment of protein interfaces and loops. *Bioinformatics*. 2013;29(24):3158-3166. doi:10.1093/bioinformatics/btt560.
97. Fan H, Irwin JJ, Webb BM, Klebe G, Shoichet BK, Sali A. Molecular docking screens using comparative models of proteins. *J Chem Inf Model*. 2009;49(11):2512-2527. doi:10.1021/ci9003706.
98. Fan H, Irwin JJ, Sali A. Virtual ligand screening against comparative protein structure models. *Methods Mol Biol*. 2012;819:105-126. doi:10.1007/978-1-61779-465-0_8.
99. Coleman RG, Carchia M, Sterling T, Irwin JJ, Shoichet BK. Ligand pose and orientational sampling in molecular docking. Romesberg F, ed. *PLoS ONE*. 2013;8(10):e75992. doi:10.1371/journal.pone.0075992.
100. Mysinger MM, Carchia M, Irwin JJ, Shoichet BK. Directory of Useful Decoys, Enhanced (DUD-E): Better Ligands and Decoys for Better Benchmarking. *J Med Chem*. 2012;55(14):6582-6594. doi:10.1021/jm300687e.
101. Schlessinger A, Geier E, Fan H, et al. Structure-based discovery of prescription drugs that interact with the norepinephrine transporter, NET. *Proc Natl Acad Sci USA*. 2011;108(38):15810-15815. doi:10.1073/pnas.1106030108.
102. Fan H, Schneidman-Duhovny D, Irwin JJ, Dong G, Shoichet BK, Sali A. Statistical Potential for Modeling and Ranking of Protein-Ligand Interactions. *J Chem Inf Model*. 2011;51(12):3078-3092. doi:10.1021/ci200377u.
103. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Research*. November 2016. doi:10.1093/nar/gkw1092.
104. Carr RAE, Congreve M, Murray CW, Rees DC. Fragment-based lead discovery: leads by design. *Drug Discov Today*. 2005;10(14):987-992. doi:10.1016/S1359-6446(05)03511-7.
105. Husted S, van Giezen JJJ. Ticagrelor: The First Reversibly Binding Oral P2Y 12Receptor

References (continued)

Antagonist. *Cardiovascular Therapeutics*. 2009;27(4):259-274. doi:10.1111/j.1755-5922.2009.00096.x.

106. Parker WAE, Storey RF. Ticagrelor: agonising over its mechanisms of action. *Blood*. 2016;128(23):2595-2597. doi:10.1182/blood-2016-10-743930.
107. Gao Z-G, Jacobson KA. Emerging adenosine receptor agonists: an update. *Expert Opin Emerg Drugs*. 2011;16(4):597-602. doi:10.1517/14728214.2011.644786.
108. de Lera Ruiz M, Lim YH, Zheng J. Adenosine A2A Receptor as a Drug Discovery Target. *J Med Chem*. 2013:131115115724005. doi:10.1021/jm4011669.
109. van Giezen JJJ, Sidaway J, Glaves P, Kirk I, Björkman J-A. Ticagrelor inhibits adenosine uptake in vitro and enhances adenosine-mediated hyperemia responses in a canine model. *J Cardiovasc Pharmacol Ther*. 2012;17(2):164-172. doi:10.1177/1074248411410883.
110. Aungraheeta R, Conibear A, Butler M, et al. Inverse agonism at the P2Y12 receptor and ENT1 transporter blockade contribute to platelet inhibition by ticagrelor. *Blood*. 2016;128(23):2717-2728. doi:10.1182/blood-2016-03-707844.
111. Armstrong D, Summers C, Ewart L, Nylander S, Sidaway JE, van Giezen JJJ. Characterization of the Adenosine Pharmacology of Ticagrelor Reveals Therapeutically Relevant Inhibition of Equilibrative Nucleoside Transporter 1. *J Cardiovasc Pharmacol Ther*. January 2014. doi:10.1177/1074248413511693.
112. Pokhodylo N, Shyyka O, Matiychuk V. Synthesis and anticancer activity evaluation of new 1,2,3-triazole-4-carboxamide derivatives. *Medicinal Chemistry Research*. 2014;23(5):2426-2438. doi:10.1007/s00044-013-0841-8.
113. Haider S, Alam MS, Hamid H. 1, 2, 3-Triazoles: scaffold with medicinal significance. *Inflammation and Cell Signaling*. 2014. doi:10.14800/ics.95.

Chapter 4: Expanding the Druggable Proteome Through Discovery of Mutant-Specific Binding Pockets

The contents of this chapter have been submitted in expanded form as part of a grant application for the NIH Transformative R01 program. The project is in its proof-of-concept phase and will continue to be funded for the next several years. Key to the design, execution, and ultimate success of this project is the unconventional team that has been assembled and the innovative, fertile environment in which the project has been and will be conducted. The expertise of medical oncologist/cancer biologist Trever Bivona, structural and systems biologists James Fraser and Andrej Sali, genome engineering pioneer Jonathan Weissman, the UCSF Small Molecule Discovery Center (SMDC) led by Michelle Arkin, and small molecule discovery pioneer Kevan Shokat are combined to leverage the UCSF institutional strengths in genetics, pharmacology, structural, computational, and cancer biology.

In the first section, the details of the proposal are outlined. We strive to expand the druggable proteome by identifying mutations that enable us to pharmacologicaly modulate the cancer-driving mutant form of a protein, while avoiding modulating the native form or related family members present in non-cancerous cells. In the second section, we illustrate why this ambitious project requires unconventional integrative, interdisciplinary, and cutting-edge approaches by examining phosphatase PTEN R130G as a case study.

The author's contribution included the setup of a structural bioinformatics framework for Part B of the proposal, to predict proteins with cancer mutations that may lead to a mutant-specific cryptic binding site. In addition, as a proof-of-concept of the robustness of the approach, the author created a comparative structure model of the PTEN R130G mutant and subsequently screened a library of 41 million small molecules, to identify potential selective activators of the functionally inactive mutant for experimental validation.

Project aims and research strategy

Approach overview

The goal of this proposal is to create an intellectual foundation and experimental platform that will transform the precision treatment of cancer patients broadly and improve their survival. We propose a conceptually and technically innovative, systematic, and multidisciplinary approach to efficiently design effective targeted cancer therapies that act selectively against the critical cancer-driving mutant proteins present in tumors. We will identify mutations that convert undruggable proteins into therapeutic targets and mutations that enable us to selectively pharmacologically modulate the cancer-driving mutant form a protein, while sparing the native form or related family members present in non-cancerous cells.

Our underlying hypothesis is that mutations in cancer genes alter the thermodynamic stability of the mutated protein,[1] creating new cryptic small molecule binding pockets[2] otherwise undruggable proteins and allowing for selective targeting of the mutant but not native form of the cancer-driving protein. Our approach is transformative because it has the potential to destroy human tumors on a broad scale and with minimal toxicity. The innovative nature of this proposal is due to its ability to simultaneously: 1) re-prioritize critical cancer mutations for therapeutic targeting in patients, 2) uncover the biochemical, biological, and functional effects of these therapeutically untapped cancer mutations, 3) establish a new, iterative platform to accelerate the development of maximally effective genetically-targeted cancer therapies at unprecedented scale and breadth, and 4) provide mutant-selective activators (correctors) that would unlock the therapeutic potential of intractable tumor suppressor targets such as PTEN. This innovation would transform the field, as there is no established or systematic approach to accomplish this goal and no mutant-selective activator therapy is approved for use in cancer patients (**Figure 4.1**). Because of the interdisciplinary nature of this application, we have assembled a team of investigators with synergistic and deep domain expertise, which will allow us to successfully complete the high-risk, high-reward project we propose. The payoff of these studies will be both an improved understanding, at the basic level, of the effects and functions of the genetic alterations present in human cancers and, at the translational level, of a new model to design effective therapies against any relevant cancer mutation and match these therapies with appropriate cancer patient populations. If successful, this

approach will improve the survival of cancer patients and, more broadly, propel the field of genomic medicine forward to improve human health in the near term.

Background and rationale

We propose a high risk, high impact project to improve our understanding of the genetic basis of disease (focusing initially on cancer) and enhance the survival of cancer patients through the systematic development of unconventional precision, genetically-targeted cancer therapies. The identification of specific somatic genetic alterations that drive tumor growth and the development of specific targeted therapies that act against them has revolutionized the treatment of patients suffering from a wide spectrum of advanced stage tumors.[3] Prominent examples of the success of genetically-targeted cancer therapy include the first-line use of ABL kinase inhibitors (e.g., imatinib, nilotinib) in chronic myeloid leukemia patients, of EGFR kinase inhibitors (e.g., erlotinib, afatinib, osimertinib) in patients with EGFR mutant lung cancer, of ALK kinase inhibitors (e.g., crizotinib, ceritinib, alectinib) in patients with ALK gene rearrangement positive lung cancer, and of BRAF kinase inhibitors (vemurafenib, dabrafenib) in patients with BRAF mutant melanoma.[4] The increased efficacy and decreased toxicity of these targeted therapies has led to a move away from traditional cytotoxic chemotherapy and toward genetically informed cancer treatment for the relevant cancer patient subsets. However, despite these notable paradigm-defining successes and the widespread recognition that virtually all cancers are driven by specific genetic alterations present in the tumor, targeted therapies are not available for the vast majority of cancer patients. This is because only ~5% of the recurrent genetic alterations present in human tumors are currently functionally characterized and therapeutically targeted with specific drugs.[5] Most of the 5% of cancer alterations that are currently targeted in patients with clinically approved molecular therapies are kinases, with a few important exceptions such as nuclear hormone receptors.[5,6] Thus, ~95% of the cancer genome (and a multitude of non-kinase targets) remains therapeutically untapped as patients nearly uniformly succumb to lethal cancer progression. Furthermore, almost all of the small fraction of patients overall who derive an initial benefit from current targeted therapies suffer from drug-resistant disease progression, as the tumors eventually escape from treatment.[7-11] Because of this resistance, targeted therapies are not curative in cancer patients (with some notable exceptions). Thus, targeted therapies currently are not available for the majority of cancer patients and not curative in the majority of patients in which they are deployed. This stark reality is true even for emerging immunotherapies such as anti-programmed death-ligand 1 (PD-L1)

and anti-programmed death 1 (PD-1) antibodies that show remarkable activity in some patients but yet are ineffective in the majority of cancer patients.[12] Furthermore, responses to these promising immunotherapies often are similarly limited by the development of resistance that is driven by emergent genetic alterations.[13,14] Our project has the potential to help address this emerging problem and improve immunotherapy response as well.

Several obstacles have prevented progress in the field. First, we are currently unable to clinically modulate the vast majority of cancer-driving genetic alterations present in tumors, including recurrent mutant forms of many proteins that are potentially druggable.[15] Progress has been hindered because the biological function of many of these specific recurrent cancer variants remains poorly characterized.[15] Moreover, the extent to which potent and selective pharmacologic modulation of many non-traditional cancer targets (such as individual epigenetic regulators and tumor suppressors) can be achieved with conventional drug design approaches remains unclear. Furthermore, many of the alterations in tumors are in genes other than kinases and often inactivate rather than activate the function of the relevant protein.[15] Hence, many of the ~95% of unexploited cancer targets that are inactivating variants have been thought to be less amenable to direct pharmacologic modulation and considered undruggable.[6] However, the success of small molecules that "re-activate" inactivated proteins in other genetically-based diseases provides a potential path forward. The recent clinical introduction of ivacaftor, a mutant-selective CFTR G551D potentiator that restores the function of the mutant CFTR, in cystic fibrosis patients provides an important counterpoint to the concept that some proteins are undruggable.[16,17] The success of ivacaftor offers proof of concept of the biological and clinical utility of pharmacologic restoration of disease-driving mutant proteins. Additional proof-of-concept "pharmacological chaperones" are emerging for other diseases.[18-20] Initial successes against p53[21-23] suggest the broader potential of small molecule correctors to reactivate tumor suppressors in cancer, but a systematic approach to achieve this goal is lacking.

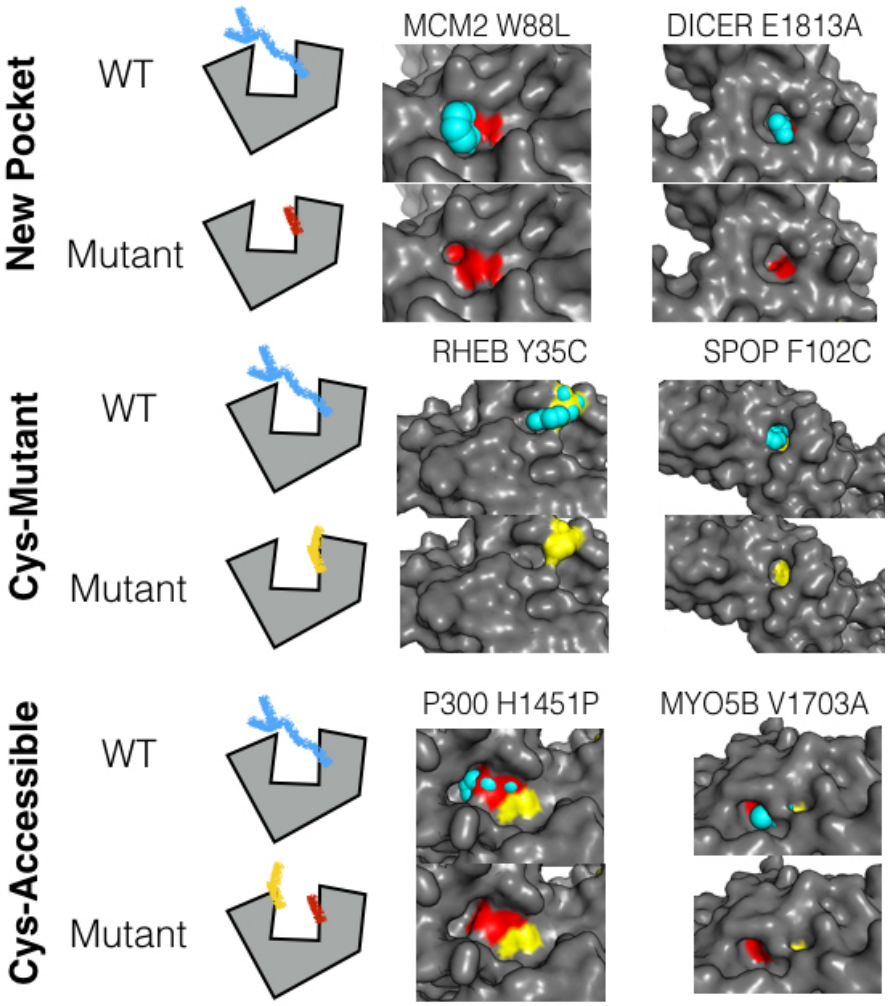


Figure 4.1 Example changes in druggable pockets created by mutations. Top) In the simplest case, mutations (red residues) remove a wild type side chain (blue) that presents a steric block to a pocket (mapped CryptoSite and FTMap). Middle) Mutations to cysteine (yellow) present new druggable opportunities, extending the concept validated by KRAS G12C molecules. Bottom) Some mutations change the pocket or accessibility of a cysteine side chain, suggesting a strategy for mutant-specific covalent modulation. In addition to these illustrative static examples of mutant pocket changes, CryptoSite also incorporates flexibility at other sites, which will increase the opportunities for allosteric mutant-specific small molecule modulation.

Second, most of the available targeted cancer therapies do not selectively modulate the activity of the mutant cancer-driving protein.[3] The success of both targeted inhibition of activated cancer-driving proteins (e.g., kinases such as RAF or MEK) and the potential targeted enhancement of inactivated cancer-promoting proteins (e.g., tumor suppressors such as PTEN) is severely compromised without selectivity against the mutant protein target. This is because of the associated toxicity that results from modulation of the native form of the protein (or its family members) in non-cancer cells.[8] Therefore, non-mutant selective agents must be dosed in patients to limit toxicity in non-cancer cells that harbor the native form of the protein (or its family members) targeted by the drug. This limitation results in less complete biological impact on the intended target in the tumor cells, contributing to the immediate or eventual escape of the tumor cells from the therapy and drug-resistant tumor progression in patients.[8] Thus, even the druggable cancer genome is not optimally leveraged to specifically, completely, and most safely destroy tumors in patients.

One recent notable exception to this historically limited approach to the development of targeted cancer therapy is the design and deployment of third generation covalent, mutant-selective, wild type sparing EGFR kinase inhibitors in EGFR mutant lung cancer patients (e.g., osimertinib/AZD9291).[24-26] Because these agents inhibit the EGFR mutant variants driving tumor growth and more modestly inhibit wild type EGFR, they have demonstrated improved anti-tumor effects and less toxicity compared to first (erlotinib) and second (afatinib) generation EGFR kinase inhibitors used in patients. Moreover, the mutant specific inhibitors are effective in erlotinib-resistant lung cancers.[24-26] Together, these proof of principle data demonstrate the utility of mutant-selective targeted therapies to improve outcomes for cancer patients. However, the cancer mutant-specific agents currently available (e.g., osimertinib targeting EGFR) are not devoid of activity against the wild type protein (e.g., EGFR) and clinical toxicity remains a challenge.[26] Mutant selective preclinical agents targeting KRAS G12C have also been reported but have yet to undergo clinical testing.[27] Unfortunately, these efforts, while significant, remain focused on only a very small fraction of the already established somatic variants present in the cancer genome. This approach to develop covalent inhibitors of cancer-driving proteins such as mutant EGFR and mutant RAS is limited in scale because it depends upon covalent interaction of a drug with a critical cysteine residue in the target protein (e.g., EGFR C797 and KRAS G12C). Most cancer targets, which do not have a cysteine residue that is critical for activity, are not druggable by this approach.

We propose to overcome these obstacles to systematically and substantially expand the 'druggable' cancer genome as our central goal. Our approach will not only convert undruggable mutant cancer proteins into druggable targets but also increase our ability to selectively target only the critical cancer-driving mutant protein to more effectively and specifically kill tumor cells and spare non-cancerous cells. Key to accomplishing our goal is the fact that most mutations only marginally reduce protein thermodynamic stability and lead to increased fluctuations around the native state.[28] Both the structural and dynamic effects of mutations can therefore increase the propensity for moderately sized hydrophobic pockets to open in the protein **Figure 4.1**. These pockets present new opportunities to discover novel binding sites for potent and selective pharmacologic agents to either therapeutically inhibit or, conversely, even rescue mutant protein function, in a selective manner. There are also additional opportunities to target cysteine residues, that are transiently exposed only in the mutant, with covalently labeling molecules **Figure 4.1**. While most approaches prioritize molecules that directly bind to and inhibit the active sites of enzymes, such as kinases, we will instead exploit mutant specific conformational effects to prioritize and target the most prevalent and biologically and clinically important cancer mutations. Our work will directly attack targets that are likely to be ignored by conventional strategies using extracellular domain directed antibodies or active site directed small molecules. We will use our novel, integrated approach to catalyze a sea change in the design and use of genetically targeted cancer therapies. Our studies are potentially paradigm shifting and will propel forward the fields of cancer biology, genetics, and pharmacology, cancer therapy, and genomic medicine. If successful, our efforts will create a new armamentarium of more effective and safe targeted therapies to improve the survival of cancer patients broadly in the near term.

Project-enabling technology innovations and platform

To accomplish our goal and develop a new paradigm for the functional annotation and therapeutic targeting of cancer mutations, we will use and integrate four complementary, innovative strategies to create a systematic platform to accelerate the identification of selective genetically targeted cancer therapies. Our platform utilizes four components, outlined in **Figure 4.2**.

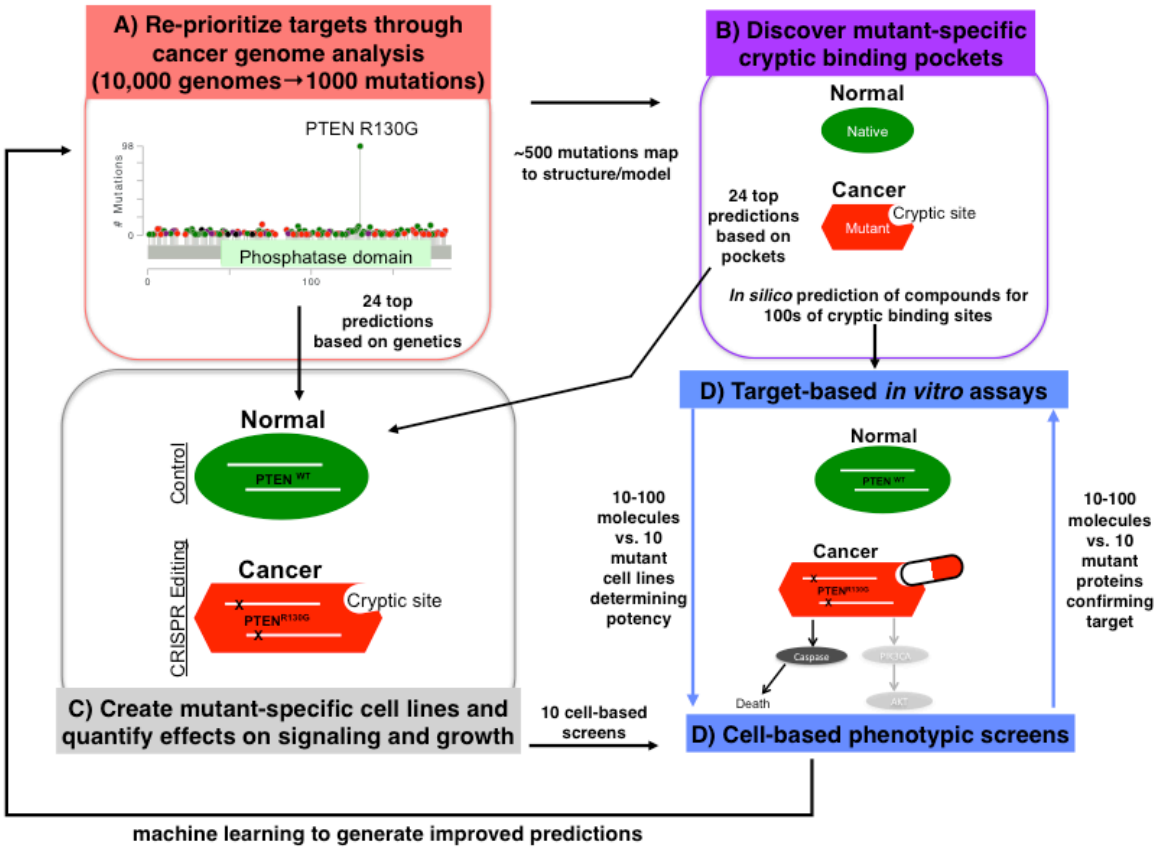


Figure 4.2 A novel iterative platform to create targeted cancer therapies. A) To prioritize critical cancer mutations that are likely rational targets for therapeutic intervention, we will perform an unbiased analysis of cancer genomes and exomes. Based on our preliminary analysis of >10,000 available human cancer genomes across all tumor types, this re-prioritization of recurrent somatic cancer point mutations consists of approximately 1,000 distinct mutations. Of these, we will select the 24 most significantly enriched mutations for systematic biological and functional annotation in relevant human cellular models (C). B) To prioritize mutations that convert undruggable proteins into druggable targets, we will determine the potential for mutations to create new "cryptic" binding pockets in target proteins. Based on preliminary results, we expect that half of the mutations (~500) will map to regions of the protein where an experimental structure or high quality homology model exists and may impact the potential for forming new cryptic sites in these regions. Based on statistical machine learning of validated cryptic sites, we will select mutations inducing the 24 most promising cryptic pockets for systematic biological and functional annotation (C). All new pockets large enough to accommodate a >150 MW small molecule will be subjected to virtual screening to identify potential small molecule modulators (D). C) To study the 48 mutations identified by genome analysis (A) and biophysical properties (B), we will use CRISPR/CAS9 to create mutant specific cell lines. These cell lines will serve two purposes. First, we will quantify the effects of each cancer variant on cell signaling, proliferation, growth, survival, and tumor growth in immunocompromised mice (tumor xenograft studies) using our established approaches and the unedited isogenic cellular model for control. The information from this stage will feed back into (A) and (B), both of which incorporate machine learning and statistical enrichment metrics to identify the most promising mutations. Second, the cell lines will be leveraged for cell-based small molecule drug efficacy screening (D), with the wild type reference cell lines serving as a control. D) We will take a two-pronged approach, guided by functional studies (C), to discover inhibitors or "correctors" of mutated proteins. First, on the top ten validated cell lines (C) we will determine the effects of small molecule perturbation screening based on cell viability and proliferative signals observed via biochemical and microscopic analysis. Second, we will perform virtual screening on all pockets identified (B). For the top predictions, we will perform functional assays in cell lines and *in vitro* using purified proteins. The information from functional studies (C) and screening, will feed back into our genome analysis and cryptic site prediction algorithm, both of which incorporate machine learning and statistical enrichment metrics to identify through a reinforcing, iterative learning model the most promising mutations for therapeutic intervention.

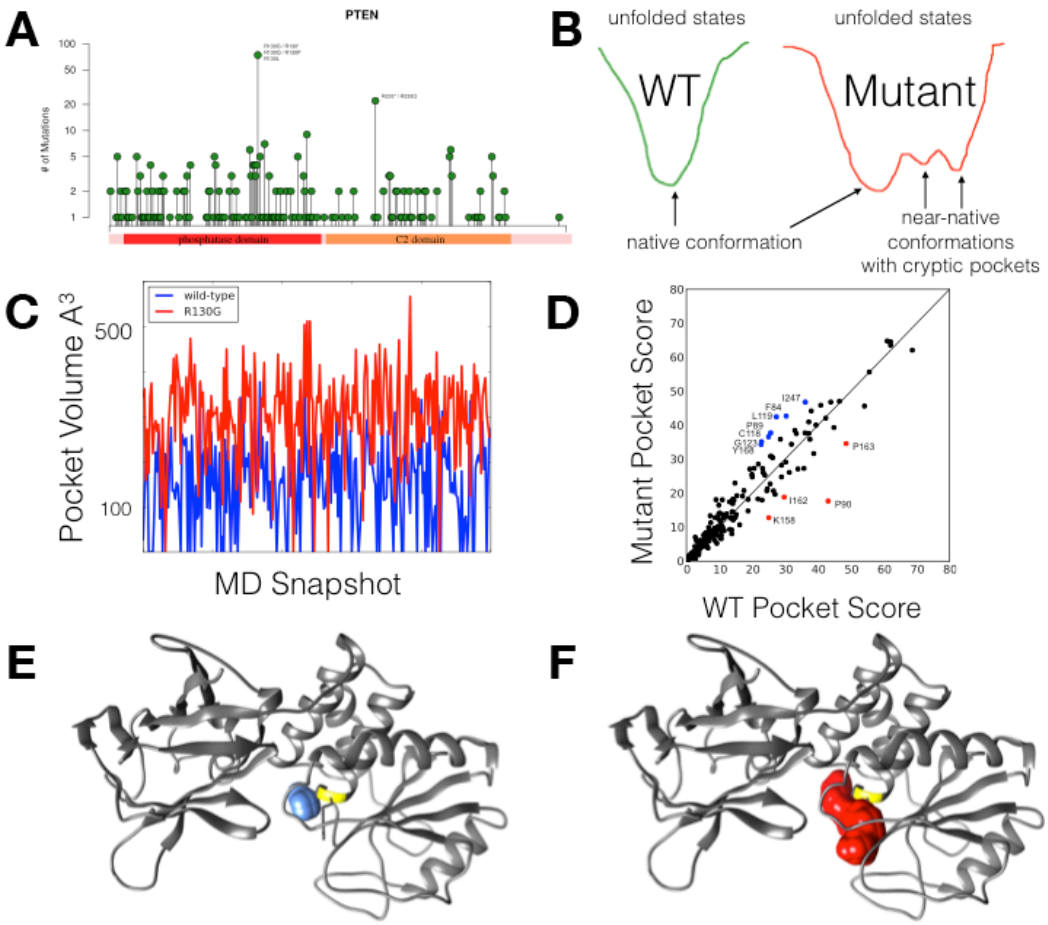


Figure 4.3 A mutation adjacent to the PTEN active site is a candidate mutant-specific cryptic pocket and target for a small molecule re-activator of PTEN function. A) PTEN domain structure is shown with the number of mutations identified at each residue shown on the Y-axis. R130 is a hotspot for mutations identified from diverse tumor types. B) Mutations can shift the energy landscape by introducing near-native states with only minor destabilization of the native state. These near-native states can include exposed concave pockets, which provide an increased surface area that in turn maximizes intra-molecular interactions. Mutations emerging in cancer are more likely to induce these pockets since most mutations are mildly destabilizing. C) Many proteins have small molecule-binding pockets that are not easily detectable in the ligand-free structures. Sometimes, however, a binding site is flat in the absence of a ligand and only forms in the presence of a ligand; such binding sites are called cryptic sites. During AllosMod simulations, we find transient pockets formed in the mutant (red) and wild-type (blue) proteins. D) Some clusters of residues have increased propensity for pocket formation in the mutant PTEN (blue dots), while other residues have decreased propensity in the mutant (red dots). This result suggests that mutations will induce specific pockets that can be targeted with small molecules. E) The active-site adjacent pocket of PTEN is quite small in the WT protein. R130 is shown in sticks and the backbone position is highlighted in yellow. F) In contrast, the elimination of the R130 side chain opens access to a cryptic hydrophobic pocket of over 500 cubic Å. This pocket likely destabilizes the active site architecture and leads to a decrease in catalytic activity. The binding interactions with the substrate and the steric interactions with the other active site residues would need to be compensated for by a rescue molecule.

Part A - Re-prioritization of cancer targets through biologically unbiased cancer genome analysis Here, the key innovation is the nomination of cancer variants of interest in a truly unbiased manner, independent of the biological function of the particular protein involved or the presumed function of the specific cancer variant (e.g., a function-inactivating versus activating mutation). Our approach, based on the use of custom scripts and an extensively developed novel bioinformatics pipeline for cancer genome analysis, stands in contrast to conventional strategies to prioritize cancer variants for therapeutic modulation. These classical approaches have prioritized for therapeutic investigations mostly proteins in which activating variants are present and proteins that are considered readily inhibited with antibody or small molecule based therapies, such as kinases.[3] This approach, while generating some successful targeted therapies such as EGFR and BRAF kinase inhibitors, has minimized or neglected the vast majority (~95%) of potential drug targets present in the cancer genome.[3] Strikingly, almost the entire cancer genome that is potentially targetable therefore remains therapeutically unexplored.[5] Our unbiased approach not only allows us to re-prioritize genetic cancer targets but also to immediately expand the target landscape across the majority of human cancers.

We do include an important filter in our cancer-genome based target re-prioritization. Importantly, we will incorporate current biological and clinical knowledge of the role of a particular protein of interest in which an identified mutation resides during prioritization for follow up studies in Parts B, C, D. This feature of our approach allow us to identify mutations of interest in an unbiased manner while also prioritizing the study of the individual mutant proteins that are more likely to be cancer driver versus passenger events and hence more biologically and clinically relevant. Indeed, we will seek to prioritize those mutants that are more frequently selected for across tumors or in a particular tumor type with unmet need and those that are present in a known cancer-driving protein (criteria that are suggestive of a cancer driving function). We will further prioritize proteins that have been therapeutically inaccessible to date, such as clinically relevant tumor suppressors including PTEN in order to maximize the potential biological and clinical impact of our efforts.

In preliminary studies, we used our new approach to mine publically available cancer genome data to generate a list of potentially druggable cancer variants consisting of 1000 somatic point mutations across >10,000 human tumors and all tumor types (via The Cancer Genome Atlas - TCGA - datasets and cBIO portal). These cancer variants include many recurrent mutations in proteins of known function as well as variants of unknown significance both in proteins of known

function and proteins whose functions are poorly characterized. Critically, we identified many variants present in tumor suppressors with broad relevance in many tumor types, including PTEN. PTEN is a particularly illustrative candidate nominated by our analysis. This is because, although functional inactivation of PTEN contributes to oncogenesis in a wide spectrum of human cancers, PTEN has been considered an undruggable target.[29,30] This notion is based on the fact that anti-cancer therapeutic modulation of PTEN would require a function-restoring drug and that significant toxicity may result from the systemic enhancement of wild type PTEN function and consequent growth suppressive effects in non-cancer cells. Yet, our approach reprioritizes the PTEN mutant R130G, the most frequent recurrent somatic variant of PTEN present in human cancers, for functional and structural studies to enable mutant-specific therapeutic modulation selectively in cancer cells **Figure 4.3**. We found that the PTEN R130G variant is present in approximately 3% of all human tumors that lack another rational therapeutic target and are particularly aggressive malignancies, including many endometrial carcinomas (~36%), gliomas (~14%), melanomas (~10%), prostate adenocarcinomas (~10%), and non-small cell lung cancers (~2.5%). Thus, this singular PTEN mutant variant (PTEN R130G) is present in over 100,000 cancer patients. This preliminary analysis demonstrates the enormous potential of our unbiased re-prioritization of cancer targets to iteratively and substantially expand the target landscape that could be leveraged to broadly deploy targeted therapies in cancer patients who currently are without genetically-based therapeutic options.

Part B - Discovery of mutations that convert cancer variants and undruggable proteins into druggable targets by inducing new cryptic binding pockets The vast majority of the mutations identified by our bioinformatics analysis (Part A) will occur in proteins without currently available small molecules that can therapeutically modulate their function. The goal of this project component is to determine how recurrent mutations in these undruggable proteins can influence the propensity to form small molecule binding pockets. Our underlying hypothesis is that thermodynamically destabilizing cancer mutations will induce “cryptic” binding pockets that are not significantly populated by the wild type protein.[2] One benefit of discovering pockets that are only present in the mutant, but not wild type protein, is the potential to increase the specificity of small molecules modulation. More intriguingly, molecules that interact, at least in part, with such pockets could rescue the function of the protein, converting a target from undruggable to druggable.[19,31] For example, in the recurrent R130G PTEN mutation the loss of the arginine side chain opens access to a hydrophobic pocket adjacent to the active site **Figure 4.3**. This mutation

likely destabilizes the precise conformations required for catalysis and coordinating the scissile phosphate, which presents an immediate hypothesis for chemical rescue of the disabled mutant PTEN.

To identify mutations with the potential to create ligand binding pockets on a broad scale, we will create a structural bioinformatics pipeline that incorporates mutational data into our recently developed prediction tool CryptoSite.[2] CryptoSite analyzes sequence, structure, and dynamics using crystal structures or homology models and molecular dynamics simulations on simplified energy landscapes. Currently, CryptoSite predicts cryptic sites, which are ligand-binding sites that are not populated in the ground state of the protein but are energetically accessible and large enough to bind ligands. This novel approach uses machine learning to predict cryptic sites and performs with relatively high accuracy (for our benchmark, the true positive and false positive rates are 73% and 29%, respectively). Importantly, cryptic sites are predicted in many proteins that are currently considered undruggable. The key innovation here is to identify cryptic sites that are specifically induced by recurrent cancer mutations. We will modify the code to incorporate mutations, post-translational modifications, and novel metrics of electron density flexibility that we have recently shown increase the potential of virtual screening for identifying novel compounds[33] at pockets that can be mapped using tools such as FTMap.[32] Our analysis will prioritize the recurrent cancer variants uncovered through the bioinformatics analysis (Part A) for functional characterization (Part C) based on the potential for mutant-specific binding pockets. We will follow these predictions with cell-based screens and with *in vitro* biochemistry of candidate molecules identified by virtual screening (Part D). These studies will identify molecules with the potential to modulate the thermodynamics and function of important, but presently undruggable, proteins that are commonly mutated in cancer (e.g., PTEN R130G).

Part C - Systematic biological and functional annotation of cancer variants using cutting-edge CRISPR/CAS9 genetic perturbation in relevant human cancer models In parallel to the computational approaches to discover mutant specific druggable opportunities, we will perform systematic experiments to determine the mutations with the most significant biological effects. Here, the key innovation is the use of CRISPR/CAS9 genetic perturbation technology[33,34] to define the biological effects and function of specific cancer variants of interest in appropriate human cancer model systems. The precise biological effects of many of cancer variants present in tumors and identified in our preliminary studies (e.g., PTEN R130G) on key cellular phenotypes, such as

intracellular signaling, proliferation, growth and survival are incompletely characterized. Furthermore, knowledge of the effects of many well-studied cancer mutations has arisen largely from genetic overexpression or knockdown studies or pharmacologic perturbation using drugs that are not selective for the cancer mutation but rather target the protein of interest. These standard approaches can provide some insight but, in many cases, may not faithfully reproduce the genetic context or biological effects of the specific cancer mutant (e.g., they to not allow for the study of mutant homozygosity versus heterozygosity or accurate gene ploidy or for the elucidation of unanticipated non-enzymatic functions of the variant in question). Hence, it is clear that the effects of a particular cancer variant may not be equivalent to surrogate non-specific genetic or pharmacologic modulation of the relevant protein in which the mutation resides. Therefore, conventional approaches to investigate the effects of specific cancer variants may generate incomplete, misleading, or inaccurate information. Moreover, our structural and pharmacologic strategy to target cancer variants is based on a deep knowledge of mutation-specific effects on the functional properties of the protein of interest. Hence, we have devised an alternative approach that differs from the standard and even contemporary approach to cancer mutation functional annotation and that capitalizes on emerging technology and deep domain expertise present in our collaborative team.

We will use the CRISPR/CAS9 system for genome editing and functional dissection of the cancer mutants prioritized for study in Part A, in parallel with our studies of these variants in Part B. We will genetically introduce the top prioritized cancer mutants and functionally define their effects using orthogonally informative, complementary human cellular models that will include: i) patient-derived cancer cell lines matched to the particular tumor type(s) in which the mutation is present but that do not endogenously harbor the somatic variant of interest; ii) a genetically controlled system consisting of tissue-matched non-cancerous cells that are immortalized but do not harbor the somatic variant of interest; iii) in each cell line system, massively parallel intragenic suppressor screens to restore proper protein function. Furthermore, we will use patient-derived cancer cells that harbor the variant of interest, where available commercially or through our active and successful IRB-approved tumor biopsy program for patient-derived xenograft models (and *ex vivo* derived cell cultures),[35] to study the effects of reverting the mutant to wild type using CRISPR/CAS9 genetic editing. We will ensure robust functional characterization of each cancer variant by creating at least two cellular models from each class above (therefore using at least six distinct cellular models for the functional study of each variant). **Figure 4.4** shows this experimental

paradigm using the PTEN R130G mutant as an illustrative case. We have established the CRISPR/CAS9 system in our laboratory in collaboration with genome-editing pioneer Jonathan Weissman (HHMI/UCSF).[33,36,37] This powerful system allows us to introduce and select for individual (or multiple) cancer variants in cells to recapitulate the zygosity or copy number present in the tumor. We will define the effects of each cancer variant on cell signaling, proliferation, growth, survival, cellular transformation, and tumor growth in immunocompromised mice (tumor xenograft studies) using our established approaches and the unedited isogenic cellular model for a baseline control.[38-43] The analysis of signaling parameters will consist of both candidate-based approaches (e.g., examining cell cycle profiles or signaling via the canonical RAF-MEK-ERK or PI3K-AKT pathways) and unbiased approaches (e.g., using high-throughput transcriptome deep sequencing or phosphoprotein arrays to delineate unanticipated mutant-specific biological effects and programs), as we have done previously to characterize putative functions of cancer targets of interest.[38-43] In preliminary studies, we have engineered the PTEN R130G into human bronchial epithelial cells, resulting in homozygous incorporation (mirroring the human tumor specimen findings) and increased levels of phosphorylated AKT as expected for a loss-of-function PTEN mutation. Once established by the CRISPR/CAS9 system, we will validate the phenotypic and biological effects of individual variants using conventional genetic over-expression or knockdown studies (cDNA over-expression or siRNA/shRNA knockdown) and also CRISPR/dCAS9 endogenous gene activation and inhibition methods recently developed by the Weissman group (CRISPRa/i),[44,45] as appropriate for each variant based on the biological results of the genome editing experiments. To probe the hypothesis that mutant-induced conformational defects can be corrected and validate our mutant pocket drug-targeting strategy, we will also screen for "second site suppressor mutations" using CRISPR mutagenesis approaches. The connection between suppressor mutation and small molecule rescue is well established for systems such as CFTR, where suppressor mutations de-risked drug discovery efforts.[46] For mutations with a screenable phenotype (e.g., FACS-based sorting of reporter gene activation), we will use CRISPR mutagenesis[47] targeting of the mutant gene to identify suppressor mutations that reverse the phenotype. Collectively, these studies will allow us to establish the full spectrum of biological effects of each cancer variant, whether activating or inactivating, in an appropriate genetic and tissue context and confirm the probability of success for achieving mutant specific drug modulation of each cancer mutant protein studied.

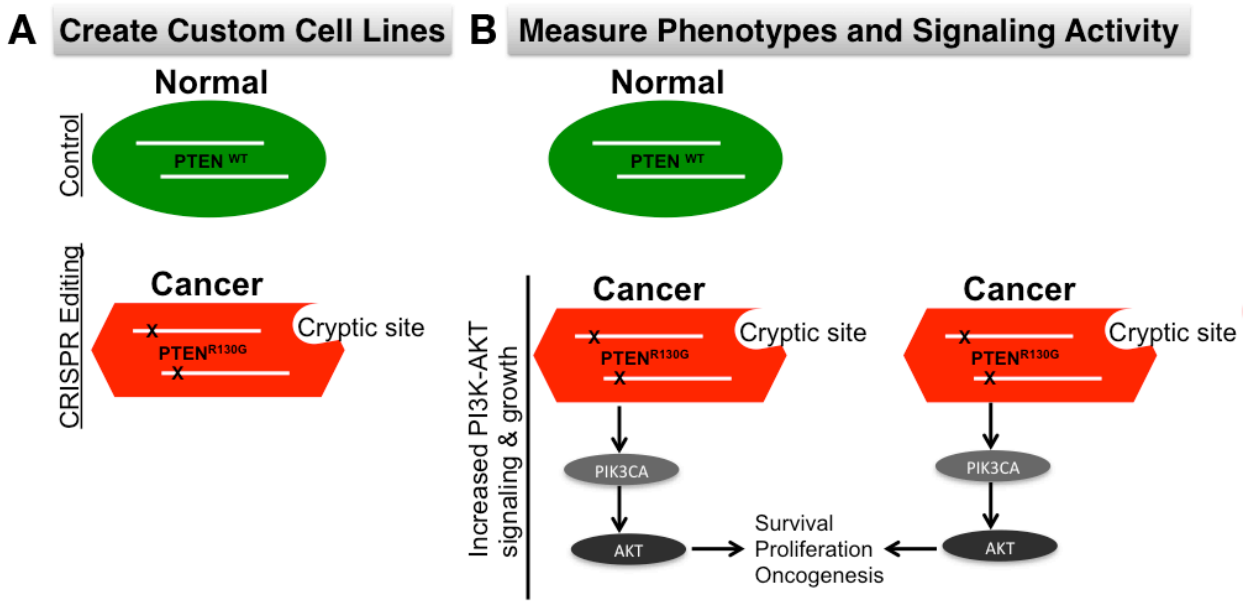


Figure 4.4 Functional annotation of the biological and phenotypic effects of specific cancer variants identified and prioritized in Part A and B. A) Creation of PTEN R130G isogenic cellular systems. B) Phenotypic and biological characterization of the PTEN R130G variant in the models generated in (A). PTEN R130G is shown for illustrative purposes, as we envision this as a next-generation scalable platform for high fidelity cancer variant functional characterization.

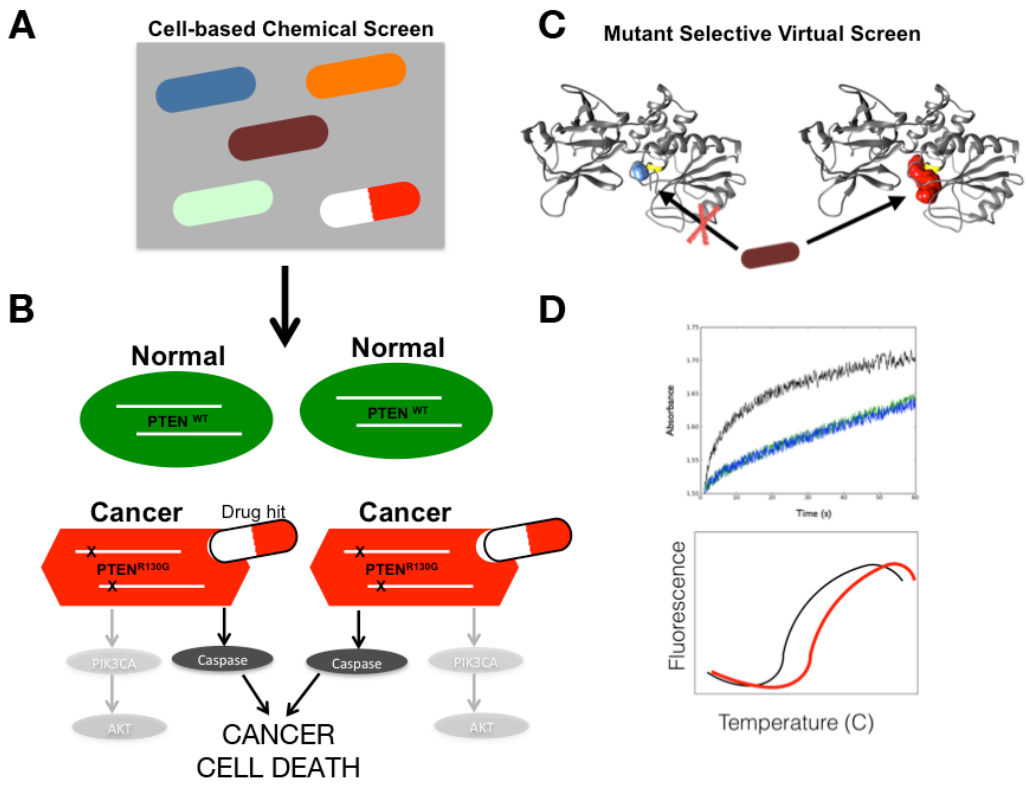


Figure 4.5 Iterative discovery and testing of mutant-specific small molecule modulators of cancer variant function to suppress cancer cell viability and growth. A,B) Pharmacologic screen design and initial validation to complete cellular and biological characterization of the drug hits identified. We will iterate based on the hits that are identified and validated. Targeting PTEN R130G is shown for illustrative purposes, as we envision this as a next-generation scalable platform for high fidelity cancer variant functional characterization of not only this common PTEN mutant but also other common, currently undruggable cancer mutants. C,D) Virtual screening of promising mutant specific pockets (shown here in red) will identify candidate molecules. These molecules will be tested with enzyme assays, where available, and thermal shift assays.

Critically, our findings here will inform the characterization and therapeutic exploitation of the mutation-specific conformational effects of these variants, as described below. We expect to conduct comprehensive functional studies on up to 48 cancer variants prioritized and studied in Parts A, B over the duration of the project (~9-10 variants/year over 5 years), ensuring feasibility of completing the work during the funding period.

Part D - Identification of pharmacologic modulators to perturb cancer variant function to selectively suppress growth and survival of tumors that harbor the targeted variant To identify small molecule modulators of presently undruggable cancer targets, we will employ two complementary strategies. Both approaches exploit the potential for mutations to introduce new cryptic sites for small molecule binding. In the first strategy, the key innovation is the use of cellular models containing the mutations of interest that are generated in Part C to screen for small molecules that show therapeutic effects, including decreased cell proliferation, growth, and survival, and induction of the relevant signaling pathway changes **Figure 4.5.** Although the cell-based nature of this screen does not necessitate direct action on the mutated target, the prioritization of pocket-containing targets gives multiple paths to success by either direct or indirect modulation of the deleterious phenotypes. In the second strategy, we will focus directly on the cryptic pockets by predicting and testing candidate molecules. This combination of cell- and target-based methods will increase our success of identifying small molecule modulators of phenotypes caused by recurrent mutations in cancer from the >100,000 molecules in the UCSF SMDC diversity (drug screening) library.

For the cell-based approach, we will use the isogenic cell line without the cancer variant as a negative control. The primary goal will be to identify small molecule agents that significantly suppress the growth of the cell line with the mutant protein but do not impact the growth of the cell line with the native form of the protein. Small molecules meeting these criteria will then be tested in independent validation studies that include: 1) re-testing in the original isogenic pair of cell lines and 2) testing in three additional isogenic pairs of cell lines (Part C). Hits that score positively in these validation studies will be prioritized for further detailed analysis to determine the mechanism of cell growth suppression in the sensitive, mutant expressing cell line models. For example, we will determine whether the drug hit induces cell cycle arrest or apoptosis as assessed by standard molecular biomarker studies including cell cycle profiling, immunoblot analysis to measure cell cycle checkpoint proteins such as p21, p27, and apoptotic markers including cleaved Parp or

caspases 3, 7, 8, as we have done previously.[38-43] We hypothesize that the pockets prioritized in Part B will provide a likely target for these molecules and we will examine the effects of drug treatment on cellular signaling biomarkers relevant to the pathway of interest. For example, for a drug targeting PTEN R130G we will measure the levels of total and phosphorylated forms of signaling proteins in the PI3K and AKT pathway (e.g., phospho-AKT and total AKT levels by immunoblot) and expect that a 'corrector' small molecule targeting PTEN R130G will decrease PI3K-AKT signaling (e.g., indicated by decreased phospho-AKT). The isogenic cell lines expressing the native form of the target (e.g., wild type PTEN) will be used as negative controls in these cellular and signaling experiments. These studies will be conducted initially *in vitro* and then *in vivo* (tumor xenografts) where appropriate (e.g., upon demonstration of positive *in vitro* effects and the suitability and sufficiency of drug for *in vivo* dosing in mice), using our established approaches to measure both the cellular and biological effects as outlined in Part C. We expect to conduct initial screening studies on the ten most promising specific cancer variants prioritized and functionally characterized in Parts A-C (~two per year) to ensure the feasibility of completing the proposed work during the funding period.

To complement the cell-based screen, we will take a parallel target-centric approach. The output of our CryptoSite predictions is a series of potential small molecule binding pockets that are formed in the mutant, but not wild type protein. We will perform a virtual screen of all purchasable compounds in the ZINC database against each of these modeled binding pockets using methods established in the Sali lab.[48] We anticipate that this process will identify promising molecules for potentially thousands of new cryptic sites, which will be carefully manually curated. There is precedence for a small molecule rescue of protein folding in the recent work on p53.[49] Based on previously published protein expression protocols used for determination of X-ray crystal structures, we will express and purify the minimal functional units of the mutant and wild type proteins and purchase a small set (10-100) of candidate compounds. While the specific targets will dictate the potential for enzymatic or protein-protein interaction assays, the expertise in the UCSF SMDC for generally applicable thermal shift assays will allow us to test the idea that the molecules tested will compensate for thermodynamic defects of the protein. We will be opportunistic to employ structure-based drug design approaches by crystallizing the mutant proteins complexed with lead ligands and take advantage of recent advances by our groups in exploiting protein flexibility in design.[50] A major advantage is the synergy between the two approaches: we will use the cell-based assays to test the cellular potency of the molecules identified in *in vitro* studies and we

will use the target-based assays to determine whether the molecules identified in the cell-based screens are active due to direct modulation of their intended targets.

To summarize, the payoff of these studies will be an improved understanding, at the basic level, of the effects and functions of the genetic alterations present in human cancers and, at the translational level, a new model to design more effective therapies against any cancer driving mutant and match these drugs with appropriate patient populations to improve human health.

Spotlight on PTEN R130G

PTEN, or phosphatase and tensin homolog, is a 403 residue, dual-specificity phosphatase that can dephosphorylate both polypeptides and phosphoinositide substrates.[51] It is a known tumor suppressor by acting on phosphatidylinositol (3,4,5)-trisphosphate [PtdIns(3,4,5)P3], a potent secondary messenger that binds pleckstrin-homology (PH) domains of AKT family of kinases and PDK1. Deregulation of PtdIns(3,4,5)P3, due to the loss of proper PTEN function activates the PI 3-kinase pathway, leading to oncogenesis.[30] Germline PTEN mutations have been found in Cowden syndrome, Bannayan-Riley-Ruvalcaba, Proteus, and Proteus-like syndrome patients.[52] Somatic PTEN mutations play a major role in the pathogenesis of almost every cancer type examined.[53] In fact, the frequency of PTEN mutations in tumors is high, rivaling only that of p53.[54]

As mentioned in Chapter 4.1, we found that the PTEN R130G variant is prevalent, present in approximately 3% of all human tumors that lack another rational therapeutic target and are particularly aggressive malignancies, including many endometrial carcinomas (~36%), gliomas (~14%), melanomas (~10%), prostate adenocarcinomas (~10%), and non-small cell lung cancers (~2.5%). Thus, this singular PTEN mutant variant (PTEN R130G) is present in over 100,000 cancer patients.

With many of the alterations in tumors in genes other than kinases, mutations often inactivate rather than activate the function of the relevant protein.[15] Functional phosphoinositol phosphatase activity assays have shown that the R130G mutation is inactivating.[55] In line with the broad aims of this project, mainly, to tap into the large portion of the unexploited cancer targets that are inactivating variants and have been thought to be less amenable to direct pharmacologic modulation and considered undruggable,[6] we will strive to find mutant-selective activator molecules to rescue the proper PTEN function. The success of small molecules that "re-activate" inactivated proteins in other genetic-based diseases provides a potential path forward. As mentioned in Chapter 4.1, one such example is ivacaftor, a mutant-selective CFTR G551D potentiator that restores the function of the mutant CFTR in cystic fibrosis patients.[16,17] The success of ivacaftor offers a proof of concept of the biological and clinical utility of pharmacologic restoration of disease-driving mutant proteins. Our strategy for restoring PTEN function in the R130G mutant consists of *in silico* screening for small molecules capable of specifically binding into the mutant active site, which would then help restore proper substrate binding.

Methods

Comparative structure modeling of PTEN R130G We created comparative protein structure models of PTEN R130G with MODELLER 9.17 (San Francisco, CA, USA),[56] using the co-crystal structure of PTEN with L(+)-tartrate (PDB ID 1D5R) as a template. Sequence alignment was computed with PROMALS3D.[57] The L(+)-tartrate ligand was copied from the template structure into the models as a rigid body, using the BLK residue type in MODELLER. One hundred models were generated, using the automodel class with default settings. The models were evaluated using the protein orientation-dependent statistically optimized atomic potential (SOAP-Protein) score.[58] The top-scoring model was used for analysis.

Virtual screening Docking was performed using the UCSF DOCK infrastructure set up by the Shoichet laboratory. The top scoring R130G model was used for the prediction of putative binding sites with FTMap web server.[32] The only predicted binding site cavity was then used for sphere and grid setup within DOCK 3.7.[59] L(+)-tartrate molecule was treated as a coenzyme. Because it is a nonstandard residue type that is undefined within DOCK, the assignment of partial charge distribution was performed with PRODRG.[60] As a preliminary screen, a library of 42 million small molecules of molecular weight 200-300 Da (standard reactivity, wait OK purchasibility) was selected from ZINC 15.[61]

Results

PTEN is composed of 2 domains. The 179-residue N-terminal phosphatase domain performs the catalysis function, while the 166-residue C-terminal C2 domain participates in the recruitment of signaling proteins.[51] The signature, PTEN-specific motif HCKAGKGR forms one side of the binding cavity and contains C124, which is the catalytic residue. R130 is critical for the capture and positioning of the ligands, as indicated by the four hydrogen bonds that it makes in the crystal structure with the L(+)-tartrate molecule, a PTEN inhibitor (PDB ID 1D5R) (**Figure 4.6A**).

As seen in the R130G model, the mutation eliminates key interactions stabilizing the ligand within the active site, and creates access to a large, hydrophobic cavity adjacent to it (**Figure 4.6**). While the PTEN active site is already considered to be larger than those of other phosphatases (e.g., VHR,

PTP1B) to accommodate phosphoinositide substrate,[51] the total cavity created by the addition of the adjacent pocket is further enlarged by 268 $Å^3$.[62]

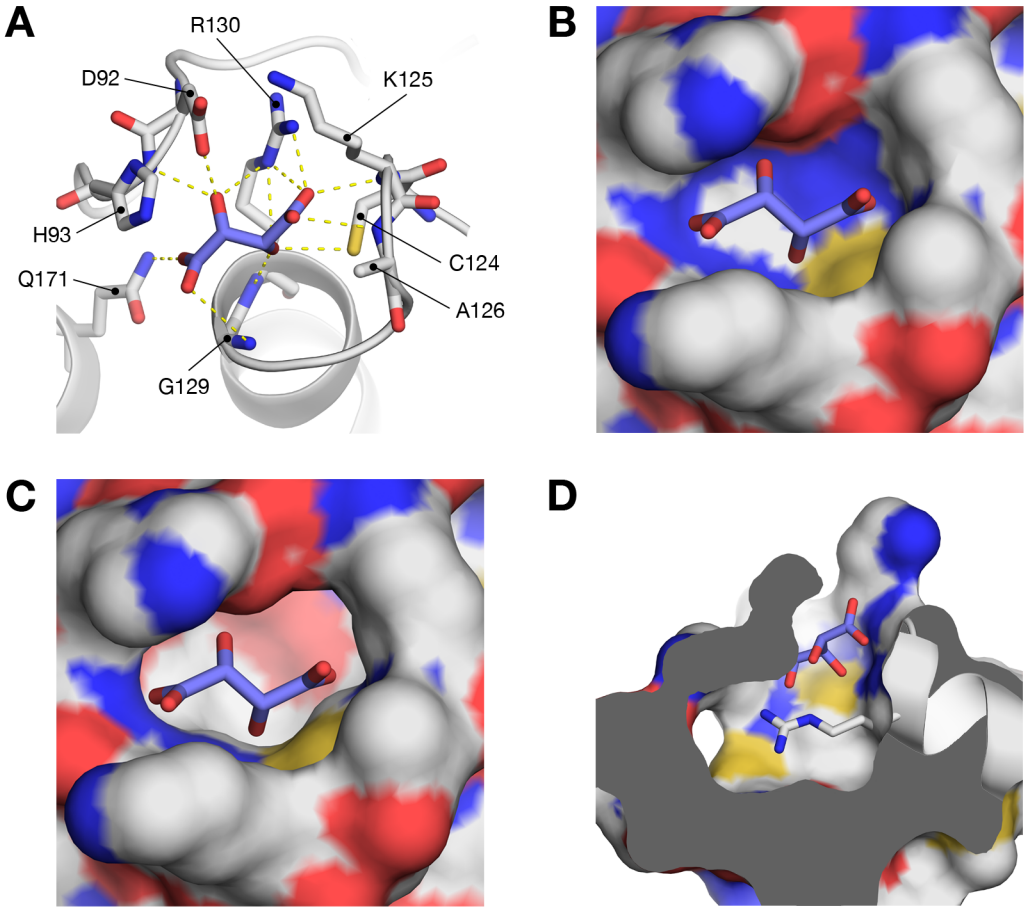
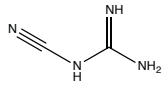


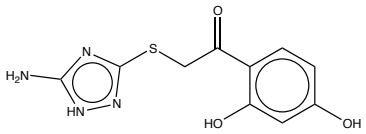
Figure 4.6 Cartoon and surface representations of the PTEN WT and R130G active sites. A) Close-up of the PTEN active site showing the hydrogen bond network (shown in yellow dashed lines) formed between L(+)-tartrate (shown in marine), a PTEN inhibitor, and the active site residues (PDB ID 1D5R). The HCXXGXXR motif present in protein tyrosine phosphatases spans residues 123-130. B) Surface representation of the active site in the WT protein. C) Surface representation of the active site in the R130G mutant. The lack of the arginine side chain eliminates critical interactions stabilizing the ligand within the active site and creates access to a hydrophobic pocket adjacent to the active site. D) Cross-sectional view of the large R130G pocket created by the absence of the R130 side chain. Also shown is an overlay of the R130 side chain and the pα4 helix in the WT structure.

Our strategy for restoring PTEN function in the R130G mutant consists of *in silico* screening for small molecules capable of binding into the newly accessible hydrophobic pocket next to the active site. Such binding would likely prevent substrates from accessing that pocket by reconstituting an active site of similar volume as the wild type, while also providing polar groups necessary for proper ligand capture, orientation, and stabilization within the active site.

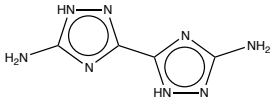
We screened a ZINC subset library of 42 million small molecules of molecular weight 200-300 Da. Based on the hits, we further explored the chemical space surrounding ZINC000001939756, ZINC000002016764, ZINC000002566399, ZINC000004658576, and ZINC000108280248 (**Figure 4.7**). In total, 12 compounds were purchased. In close collaboration with the Bivona and Fraser laboratories, we are actively working to express and purify both the PTEN wild type and the PTEN R130G mutant for functional validation of these compounds. In parallel, we plan to crystallize both isoforms with and without the compounds. Early expression studies have made us optimistic as to the feasibility of the study but a significant amount of work remains to be done on all fronts.



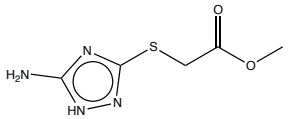
1-cyanoguanidine
ZINC000005112486



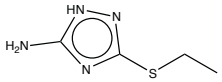
2-((5-amino-1*H*-1,2,4-triazol-3-yl)thio)-1-(2,4-dihydroxyphenyl)ethanone
ZINC000006132438



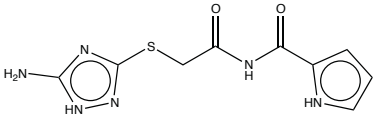
1*H*,2'*H*-[3,3'-bi(1,2,4-triazole)]-5,5'-diamine
ZINC000001063069



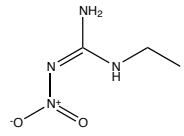
methyl 2-((5-amino-1*H*-1,2,4-triazol-3-yl)thio)acetate
ZINC000005577500



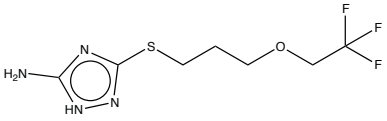
3-(ethylthio)-1*H*-1,2,4-triazol-5-amine
ZINC000001939756



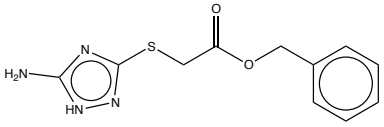
N-(2-((5-amino-1*H*-1,2,4-triazol-3-yl)thio)acetyl)-1*H*-pyrrole-2-carboxamide
ZINC000032809261



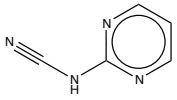
(*Z*)-1-ethyl-2-nitroguanidine
ZINC000016033785



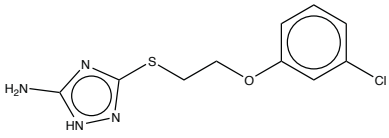
3-((3-(2,2,2-trifluoroethoxy)propyl)thio)-1*H*-1,2,4-triazol-5-amine
ZINC000042374095



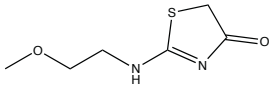
benzyl 2-((5-amino-1*H*-1,2,4-triazol-3-yl)thio)acetate
ZINC000005647138



N-(pyrimidin-2-yl)cyanamide
ZINC000013638829



3-((2-(3-chlorophenoxy)ethyl)thio)-1*H*-1,2,4-triazol-5-amine
ZINC000008184616



2-((2-methoxyethyl)amino)thiazol-4(5*H*)-one
ZINC000036613065

Figure 4.7 Compounds selected from an in silico screen for experimental validation.

References

1. Nussinov R, Tsai C-J. Allostery in disease and in drug discovery. *Cell*. 2013;153(2):293-305. doi:10.1016/j.cell.2013.03.034.
2. Cimermancic P, Weinkam P, Rettenmaier TJ, et al. CryptoSite: Expanding the Druggable Proteome by Characterization and Prediction of Cryptic Binding Sites. *J Mol Biol*. 2016;428(4):709-719. doi:10.1016/j.jmb.2016.01.029.
3. Gonzalez de Castro D, Clarke PA, Al-Lazikani B, Workman P. Personalized cancer medicine: molecular diagnostics, predictive biomarkers, and drug resistance. *Clinical Pharmacology & Therapeutics*. 2013;93(3):252-259. doi:10.1038/clpt.2012.237.
4. AACR Cancer Progress Report Writing Committee, Sawyers CL, Abate-Shen C, et al. AACR Cancer Progress Report 2013. *Clin Cancer Res*. 2013;19(20 Suppl):S4-S98. doi:10.1158/1078-0432.CCR-13-2107.
5. Workman P, Al-Lazikani B. Drugging cancer genomes. *Nat Rev Drug Discov*. 2013;12(12):889-890. doi:10.1038/nrd4184.
6. Sawyers CL. The cancer biomarker problem. *Nature*. 2008;452(7187):548-552. doi:10.1038/nature06913.
7. Sawyers CL. The 2011 Gordon Wilson Lecture: overcoming resistance to targeted cancer drugs. *Trans Am Clin Climatol Assoc*. 2012;123:114–23–discussion123–5.
8. Workman P, Al-Lazikani B, Clarke PA. Genome-based cancer therapeutics: targets, kinase drug resistance and future strategies for precision oncology. *Current Opinion in Pharmacology*. 2013;13(4):486-496. doi:10.1016/j.coph.2013.06.004.
9. Rosell R, Bivona TG, Karachaliou N. Genetics and biomarkers in personalisation of lung cancer treatment. *Lancet*. 2013;382(9893):720-731. doi:10.1016/S0140-6736(13)61715-8.
10. Hrustanovic G, Lee BJ, Bivona TG. Mechanisms of resistance to EGFR targeted therapies. *Cancer Biol Ther*. 2013;14(4):304-314. doi:10.4161/cbt.23627.
11. Bivona TG, Doebele RC. A framework for understanding and targeting residual disease in oncogene-driven solid cancers. *Nat Med*. 2016;22(5):472-478. doi:10.1038/nm.4091.
12. Pitt JM, Vétizou M, Daillère R, et al. Resistance Mechanisms to Immune-Checkpoint Blockade in Cancer: Tumor-Intrinsic and -Extrinsic Factors. *Immunity*. 2016;44(6):1255-1269. doi:10.1016/j.immuni.2016.06.001.
13. Zaretsky JM, Garcia-Diaz A, Shin DS, et al. Mutations Associated with Acquired Resistance to PD-1 Blockade in Melanoma. *New Engl J Med*. 2016;375(9):819-829.

References (continued)

doi:10.1056/NEJMoa1604958.

14. Gao J, Shi LZ, Zhao H, et al. Loss of IFN-γ Pathway Genes in Tumor Cells as a Mechanism of Resistance to Anti-CTLA-4 Therapy. *Cell*. 2016;167(2):397-404.e399. doi:10.1016/j.cell.2016.08.069.
15. Patel MN, Halling-Brown MD, Tym JE, Workman P, Al-Lazikani B. Objective assessment of cancer genes for drug discovery. *Nat Rev Drug Discov*. 2013;12(1):35-50. doi:10.1038/nrd3913.
16. Pellowska M, Merk D, Schubert-Zsilavecz M. Advances in personalized medicine - medicinal chemistry and pharmacology of vemurafenib and ivacaftor. *Pharmazie*. 2013;68(7):484-491.
17. Molloy K, McElvaney NG. Ivacaftor: from bench to bedside... and back again. *Am J Respir Crit Care Med*. 2014;190(2):128-129. doi:10.1164/rccm.201406-1122ED.
18. Froese DS, Michaeli A, McCorvie TJ, et al. Structural basis of glycogen branching enzyme deficiency and pharmacologic rescue by rational peptide design. *Hum Mol Genet*. 2015;24(20):5667-5676. doi:10.1093/hmg/ddv280.
19. Mecozzi VJ, Berman DE, Simoes S, et al. Pharmacological chaperones stabilize retromer to limit APP processing. *Nat Chem Biol*. 2014;10(6):443-449. doi:10.1038/nchembio.1508.
20. Lieberman RL, Wustman BA, Huertas P, et al. Structure of acid beta-glucosidase with pharmacological chaperone provides insight into Gaucher disease. *Nat Chem Biol*. 2007;3(2):101-107. doi:10.1038/nchembio850.
21. Bykov VJN, Issaeva N, Shilov A, et al. Restoration of the tumor suppressor function to mutant p53 by a low-molecular-weight compound. *Nat Med*. 2002;8(3):282-288. doi:10.1038/nm0302-282.
22. Joerger AC, Bauer MR, Wilcken R, et al. Exploiting Transient Protein States for the Design of Small-Molecule Stabilizers of Mutant p53. *Structure*. 2015;23(12):2246-2255. doi:10.1016/j.str.2015.10.016.
23. Wassman CD, Baronio R, Demir Ö, et al. Computational identification of a transiently open L1/S3 pocket for reactivation of mutant p53. *Nat Commun*. 2013;4:1407. doi:10.1038/ncomms2361.
24. Cross DAE, Ashton SE, Ghiorghiu S, et al. AZD9291, an irreversible EGFR TKI, overcomes T790M-mediated resistance to EGFR inhibitors in lung cancer. *Cancer Discov*. 2014;4(9):1046-1061. doi:10.1158/2159-8290.CD-14-0337.
25. Walter AO, Sjin RTT, Haringsma HJ, et al. Discovery of a mutant-selective covalent inhibitor

References (continued)

of EGFR that overcomes T790M-mediated resistance in NSCLC. *Cancer Discov*. 2013;3(12):1404-1415. doi:10.1158/2159-8290.CD-13-0314.

26. Jänne PA, Yang JC-H, Kim D-W, et al. AZD9291 in EGFR inhibitor-resistant non-small-cell lung cancer. *New Engl J Med*. 2015;372(18):1689-1699. doi:10.1056/NEJMoa1411817.
27. Ostrem JM, Peters U, Sos ML, Wells JA, Shokat KM. K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions. *Nature*. 2013;503(7477):548-551. doi:10.1038/nature12796.
28. Wang Z, Moult J. SNPs, protein structure, and disease. *Hum Mutat*. 2001;17(4):263-270. doi:10.1002/humu.22.
29. Simpson L, Parsons R. PTEN: life as a tumor suppressor. *Exp Cell Res*. 2001;264(1):29-41. doi:10.1006/excr.2000.5130.
30. Sulis ML, Parsons R. PTEN: from pathology to biology. *Trends Cell Biol*. 2003;13(9):478-483.
31. Xia Y, DiPrimio N, Keppel TR, et al. The designability of protein switches by chemical rescue of structure: mechanisms of inactivation and reactivation. *J Am Chem Soc*. 2013;135(50):18840-18849. doi:10.1021/ja407644b.
32. Kozakov D, Grove LE, Hall DR, et al. The FTMap family of web servers for determining and characterizing ligand-binding hot spots of proteins. *Nature Protocols*. 2015;10(5):733-755. doi:10.1038/nprot.2015.043.
33. Qi LS, Larson MH, Gilbert LA, et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*. 2013;152(5):1173-1183. doi:10.1016/j.cell.2013.02.022.
34. Lo T-W, Pickle CS, Lin S, et al. Precise and heritable genome editing in evolutionarily diverse nematodes using TALENs and CRISPR/Cas9 to engineer insertions and deletions. *Genetics*. 2013;195(2):331-348. doi:10.1534/genetics.113.155382.
35. Blakely CM, Pazarentzos E, Olivas V, et al. NF-κB-activating complex engaged in response to EGFR oncogene inhibition drives tumor cell survival and residual disease in lung cancer. *Cell Rep*. 2015;11(1):98-110. doi:10.1016/j.celrep.2015.03.012.
36. Larson MH, Gilbert LA, Wang X, Lim WA, Weissman JS, Qi LS. CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nature Protocols*. 2013;8(11):2180-2196. doi:10.1038/nprot.2013.132.
37. Gilbert LA, Larson MH, Morsut L, et al. CRISPR-mediated modular RNA-guided regulation of

References (continued)

transcription in eukaryotes. *Cell*. 2013;154(2):442-451. doi:10.1016/j.cell.2013.06.044.

38. Zhang Z, Lee JC, Lin L, et al. Activation of the AXL kinase causes resistance to EGFR-targeted therapy in lung cancer. *Nat Genet*. 2012;44(8):852-860. doi:10.1038/ng.2330.
39. Bivona TG, Hieronymus H, Parker J, et al. FAS and NF-κB signalling modulate dependence of lung cancers on mutant EGFR. *Nature*. 2011;471(7339):523-526. doi:10.1038/nature09870.
40. Lin L, Asthana S, Chan E, et al. Mapping the molecular determinants of BRAF oncogene dependence in human lung cancer. *Proc Natl Acad Sci USA*. 2014;111(7):E748-E757. doi:10.1073/pnas.1320956111.
41. Pazarentzos E, Giannikopoulos P, Hrustanovic G, et al. Oncogenic activation of the PI3-kinase p110β isoform via the tumor-derived PIK3Cβ(D1067V) kinase domain mutation. *Oncogene*. 2016;35(9):1198-1205. doi:10.1038/onc.2015.173.
42. Hrustanovic G, Olivas V, Pazarentzos E, et al. RAS-MAPK dependence underlies a rational polytherapy strategy in EML4-ALK-positive lung cancer. *Nat Med*. 2015;21(9):1038-1047. doi:10.1038/nm.3930.
43. Lin L, Sabnis AJ, Chan E, et al. The Hippo effector YAP promotes resistance to RAF- and MEK-targeted cancer therapies. *Nat Genet*. 2015;47(3):250-256. doi:10.1038/ng.3218.
44. Gilbert LA, Horlbeck MA, Adamson B, et al. Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell*. 2014;159(3):647-661. doi:10.1016/j.cell.2014.09.029.
45. Horlbeck MA, Gilbert LA, Villalta JE, et al. Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation. *Elife*. 2016;5:914. doi:10.7554/eLife.19760.
46. Rabeh WM, Bossard F, Xu H, et al. Correction of both NBD1 energetics and domain interface is required to restore ΔF508 CFTR folding and function. *Cell*. 2012;148(1-2):150-163. doi:10.1016/j.cell.2011.11.024.
47. Findlay GM, Boyle EA, Hause RJ, Klein JC, Shendure J. Saturation editing of genomic regions by multiplex homology-directed repair. *Nature*. 2014;513(7516):120-123. doi:10.1038/nature13695.
48. Fan H, Irwin JJ, Sali A. Virtual ligand screening against comparative protein structure models. *Methods Mol Biol*. 2012;819:105-126. doi:10.1007/978-1-61779-465-0_8.
49. Boeckler FM, Joerger AC, Jaggi G, Rutherford TJ, Veprintsev DB, Fersht AR. Targeted rescue of a destabilized mutant of p53 by an in silico screened drug. *Proc Natl Acad Sci USA*.

References (continued)

2008;105(30):10360-10365. doi:10.1073/pnas.0805326105.

50. Fischer M, Coleman RG, Fraser JS, Shoichet BK. Incorporation of protein flexibility and conformational energy penalties in docking screens to improve ligand discovery. *Nat Chem*. 2014;6(7):575-583. doi:10.1038/nchem.1954.
51. Lee JO, Yang H, Georgescu MM, et al. Crystal structure of the PTEN tumor suppressor: implications for its phosphoinositide phosphatase activity and membrane association. *Cell*. 1999;99(3):323-334.
52. Eng C. PTEN: one gene, many syndromes. *Hum Mutat*. 2003;22(3):183-198. doi:10.1002/humu.10257.
53. Keniry M, Parsons R. The role of PTEN signaling perturbations in cancer and in targeted therapy. *Oncogene*. 2008;27(41):5477-5485. doi:10.1038/onc.2008.248.
54. Cantley LC, Neel BG. New insights into tumor suppression: PTEN suppresses tumor formation by restraining the phosphoinositide 3-kinase/AKT pathway. *Proceedings of the National Academy of Sciences*. 1999;96(8):4240-4245.
55. Kato H, Kato S, Kumabe T, et al. Functional Evaluation of p53 and PTEN Gene Mutations in Gliomas. *Clin Cancer Res*. 2000;6(10):3937-3943. doi:10.1080/00313029800169236.
56. Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol*. 1993;234(3):779-815. doi:10.1006/jmbi.1993.1626.
57. Pei J, Kim B-H, Grishin NV. PROMALS3D: a tool for multiple protein sequence and structure alignments. *Nucleic Acids Research*. 2008;36(7):2295-2300. doi:10.1093/nar/gkn072.
58. Deng D, Xu C, Sun P, et al. Crystal structure of the human glucose transporter GLUT1. *Nature*. 2014;510(7503):121-125. doi:10.1038/nature13306.
59. Fan H, Irwin JJ, Webb BM, Klebe G, Shoichet BK, Sali A. Molecular docking screens using comparative models of proteins. *J Chem Inf Model*. 2009;49(11):2512-2527. doi:10.1021/ci9003706.
60. Schüttelkopf AW, van Aalten DMF. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr D Biol Crystallogr*. 2004;60(Pt 8):1355-1363. doi:10.1107/S0907444904011679.
61. Sterling T, Irwin JJ. ZINC 15 – Ligand Discovery for Everyone. *J Chem Inf Model*. 2015;55(11):2324-2337. doi:10.1021/acs.jcim.5b00559.
62. Dundas J, Ouyang Z, Tseng J, Binkowski A, Turpaz Y, Liang J. CASTp: computed atlas of

References (continued)

surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Research*. 2006;34(Web Server issue):W116-W118. doi:10.1093/nar/gkl282.

Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

Author Signature  Date 12/31/16