2024

Structural and Functional Studies of the Human Members of the Macrophage Migration Inhibitory Factor Family

Andrew Parkins
University of the Pacific

Structural and Functional Studies of the Human Members of the Macrophage Migration Inhibitory Factor Family

By

Andrew Parkins

A Dissertation Submitted

In Partial Fulfillment of the

Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Thomas J. Long School of Pharmacy and Health Sciences
Pharmaceutical and Chemical Sciences

University of the Pacific
Stockton, California

2024

Structural and Functional Studies of the Human Members of the Macrophage Migration Inhibitory Factor Family

By

Andrew Parkins

APPROVED BY:

Dissertation Advisor: Georgios Pantouris, Ph.D.

Committee Member: Jerry Tsai, Ph.D.

Committee Member: Liang Xue, Ph.D.

Committee Member: Joseph Harrison, Ph.D.

Committee Member: Geoff Lin-Cereghino, Ph.D.

Department Chair: Jerry Tsai, Ph.D.

Structural and Functional Studies of the Human Members of the Macrophage Migration Inhibitory Factor Family

By

Andrew Parkins

Dedication

This dissertation is dedicated to my parents in honor of all the sacrifices that they made for me to have this opportunity and for the qualities that they passed down to me. Thank you to my mom for being patient and teaching me how to move processes along efficiently as well as the importance of education. Thank you to my dad for teaching me how to fix just about anything, how to troubleshoot, and how to work hard. I appreciate everything you've done for me, love you mom and dad.

So like it or not, I'll take those
Long nights, impossible odds
Keeping my back to the wall
If it takes all night to be just what I am
Well, I'm gonna be a blue collar man
~Styx "Blue Collar Man"

Acknowledgements

I wish to extend my heartfelt gratitude to Dr. Georgios Pantouris, my dedicated advisor, whose unwavering support and guidance have been vital in reaching this significant milestone. His mentorship has been not only professional but also deeply personal, he has been a brother to me.

I am deeply appreciative of the contributions made by Dr. Jerry Tsai, Dr. Liang Xue, Dr. Joe Harrison, and Dr. Geoff Lin-Cereghino as members of my dissertation committee. Their expertise has played a crucial role in shaping the direction of my academic pursuits.

Special thanks are also due to Dr. Brett Williams and Karen Johnson for their instrumental roles in ensuring the smooth operation of the chemistry department and facilitating the progression of my doctoral studies.

I am profoundly grateful to the members of the Pantouris lab, namely Chris Argueta, Aliyah Pilien, and Jasmine Vargas, whose steadfast support and collaborative efforts have greatly enhanced the quality of my research. Additionally, I extend my thanks to the undergraduate researchers who have generously contributed to our collective endeavors.

Acknowledgment is also owed to Dr. Andreas Franz, Dr. Jianhua Ren, the Xue lab, Dr. Banu Sankaran, the Thompson lab, the De Alba lab, the Lisi lab, and the Bhandari lab for their collaborative contributions, which have enriched the scope of my research projects.

Lastly, I express my deepest appreciation to my family and friends for their encouragement and support throughout my academic journey. Their belief in my abilities has been a constant source of motivation and inspiration.

Structural and Functional Studies of the Human Members of the Macrophage Migration Inhibitory Factor Family

Abstract

By Andrew Parkins

University of the Pacific
2024

Macrophage migration inhibitory factor (MIF) and D-dopachrome tautomerase (D-DT) are the two human members of the MIF superfamily, which are implicated in an array of autoimmune disorders, inflammatory diseases, and cancer via their pleiotropic functionality. Despite only sharing 34% sequence identity, MIF and D-DT have high structural homology and overlapping functional traits, including activation of the type II cell surface receptor CD74 and keto-enol tautomerase activity. The MIF and/or D-DT-induced activation of CD74 leads to signaling cascades pivotal for cell growth, proliferation, and inhibition of apoptosis. Such characteristics make MIF and D-DT attractive molecular targets for drug discovery.

Currently, all small molecule antagonists targeting the MIF/D-DT-CD74 axis primarily bind to the catalytic sites of these proteins. Nevertheless, the precise interplay between the catalytic residues and those crucial for CD74 activation remains enigmatic. Notably, alterations of catalytic residues, particularly the catalytic residue Pro1, have been shown to impede CD74 activation. Leveraging molecular dynamics simulations and nuclear magnetic resonance (NMR) spectroscopy, we explored the dynamic coupling between the catalytically active N-terminus of MIF and surface residues pivotal for CD74 activation. Our investigation exposed previously

unseen communication between the two sites and demonstrates the important role of MIF dynamics in the modulation of CD74 activation.

The keto-enol tautomerization assay utilizing 4-hydroxyphenylpyruvate (4-HPP) as a substrate has been instrumental in screening and characterization of MIF and D-DT variants as well as small molecule inhibitors. However, discrepancies between inhibition constant (K_i) values and Michaelis-Menten parameters raised concerns about the accuracy of results from this assay and the conclusions made from them. Our rigorous analysis identified that impurities present in substrate samples impacted the kinetic parameters of wild-type (WT) MIF as well as the K_i values of ISO-1, a well-studied inhibitor. Our findings, which were validated with multiple proteins, underscore the pronounced influence of substrate impurities on enzymatic activity. Thereby emphasizing the imperative of meticulously controlled experimental conditions for robust data interpretation.

While the majority of drug discovery efforts were focused on MIF, D-DT remains relatively underexplored in this regard. The identification of 4-(3-carboxyphenyl)-2,5-pyridinedicarboxylic acid (4-CPPC) as the first reversible and selective D-DT inhibitor opened new avenues of research for the protein. Structural analysis of D-DT – 4-CPPC revealed a ligand-induced conformational change of the C-terminal region that has mechanistic value. This observation is in stark contrast to MIF, which needs a rigid C-terminal for tertiary structure stability. In order to elucidate the impact of C-terminal conformational flexibility, we employed molecular dynamics simulations and NMR experiments. We found that while the binding of 4-CPPC did not alter the folding or thermostability of the protein, it drastically altered the protein's dynamics, allowing for the formation of new, long-range intersubunit communications.

Subsequent endeavors aimed at identifying highly selective D-DT inhibitors that did not cause a conformational change of the C-terminal region yielded 2,5-pyridinedicarboxylic acid (1). This molecule exhibits a low micromolar potency and a remarkable 79-fold specificity for D-DT over MIF. Crystallographic analysis of the D-DT-1 complex displayed that the C-terminal of D-DT was largely unperturbed by the binding of 1 and delineated structural disparities between D-DT and MIF active sites, underscoring the potential for rational drug design strategies. Further *in vivo* studies focusing on the cytokine activity of D-DT showed the efficacy of 1 as an inhibitor of D-DT induced activation of CD74. These findings show that 1 is a useful mechanistic tool for interrogating the pathophysiology of D-DT.

Despite these exciting discoveries, the role of the C-terminal region in the enzymatic activity and conformational flexibility of D-DT required further investigation. In-depth interrogation of seventeen protein variants and WT D-DT uncovered a previously unknown functional role of the C-terminal region. These insights deepen our comprehension of protein structure-function relationships and provides an invaluable foundation for future drug discovery studies targeting D-DT-mediated pathological conditions.

Overall, via our thorough experimental interrogations, we uncovered key structural and functional information about MIF and D-DT that will serve as the basis for future mechanistic and drug discovery projects.

Table of Contents

List of Tables

Table

List of Figures

Figure

List of Schemes

Scheme

List of Abbreviations

4-CPPC	4-(3-carboxyphenyl)-2,5-pyridinedicarboxylic acid
4-HPP	4-hydroxylphenylpyruvate
ACKR-3	atypical chemokine receptor-3
ALS	Advanced Light Source
AMPK	Adenosine monophosphate activated protein kinase
Akt	protein kinase B
CD	circular dichroism
CD74	cluster of differentiation 74
COX2	cyclooxygenase 2
CIF	crystallographic information file
D-DT	D-dopachrome tautomerase
DHI	5,6-dihydroxyindole
EDTA	ethylenediamine tetraacetic acid
ERK	extracellular signal-regulated kinase
HPPD	4-hydroxyphenylpyruvate dioxygenase
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LB	Luria-broth
MALDI-TOF MS	matrix-assisted laser desorption/ionization time-of-flight mass spectrometry
MAPK	mitogen-activated protein kinase
MD	molecular dynamics
MeCN	Acetonitrile

MIF	macrophage migration inhibitory factor
NF-κB	nuclear factor kappa B
NMR	nuclear magnetic resonance
NSCLC	non-small cell lung cancer
PCR	polymerase chain reaction
PDB	Protein Data Bank
PEG	polyethylene glycol
PGE_2	Prostaglandin E2
PI3K	phosphoinositide 3-kinase
PSF	protein structure file
RMSD	root mean squared deviation
RMSF	root mean squared fluctuation
SEC	size exclusion chromatography
TFA	trifluoroacetic acid
WT	wild type.

CHAPTER 1: INTRODUCTION

The Tautomerase Super Family (TSF) is composed of five families, 5-(carboxymethyl)-2-hydroxymuconate isomerase (CHMI) [1, 2], 4-oxalocrotonate tautomerase (4-OT)[3], malonate semialdehyde decarboxylase (MSAD)[4], cis-3-chloroacrylic acid dehalogenase (cis-CaaD)[5], and macrophage migration inhibitory factor (MIF)[6, 7]. Each family is named after its founding member and each have the following key features: β–α–β motifs and a catalytic N-terminal proline [1, 3]. CHMI, 4-OT, MSAD, and cis-CaaD are all bacterial proteins that have varying functions and substrates despite having similar topologies as well as the shared key features. These bacterial proteins are involved in pathways that are vital for survival, this ranges from being involved in the breakdown of aromatic molecules and amino acids for energy to the degradation of pesticides[2, 8]. Meanwhile, the MIF superfamily is composed of not only bacterial proteins, but also mammalian proteins, two of which, MIF and D-Dopachrome Tautomerase (D-DT), are found in humans. While both of these proteins have been identified as 4-hydroxyphenylpyruvate tautomerases that act upon 4-hydroxyphenylpyruvate [6, 9], a naturally occurring substrate for 4-hydroxyphenylpyruvate dioxygenase (HPPD) [10], neither protein has a known biological substrate. The catalytic activity of MIF and D-DT is of unknown biological importance, yet has been used widely as the first step towards characterizing potential inhibitors for their cytokine [11] and chemokine [12] activities that are key for immune system regulation, cell growth, and proliferation, which is a characteristic unique to the MIF superfamily.

MIF is a pleiotropic protein with cytokine [11], chemokine [12], tautomerase [6] , and nuclease [13] activities. First described in 1966, it was found to be released from lymphocytes where it inhibited the migration of macrophages [11]. Unlike many other cytokines, MIF is constitutively expressed

and stored in cytosolic compartments [14] before being released upon stimulation [15] through a non-canonical pathway as it lacks any signal sequences. Despite its early identification, MIF was not cloned [16] and its structure [17] solved until nearly three decades later. A monomeric subunit was discovered to be only ~12.5KDa in size, being composed of 115 residues including the starting methionine, and having a homotrimeric biological assembly. Each MIF monomer contains six β-strands (β1–β6) and two anti-parallel α-helices (α1-α2). The biological assembly features three active sites, each containing the catalytic residue Pro1 [18], at the interface of each monomer and a central solvent channel which is constructed by the β1-β4 strands of each subunit. Despite only sharing 34% sequence identity [9], human MIF's human homologue, D-DT, shares a very similar topology [19]. The MIF and D-DT research community labeled D-DT as a result of gene duplication due to the genes encoding for MIF and D-DT being located only ~80kb apart on chromosome 22 and being flanked by glutathione S-transferase genes [19]. This led to D-DT being largely ignored until recently as it has now been understood that MIF and D-DT bind different receptors or shared receptors with different proposed modes of binding [20] and affinity [9], resulting in differences in their roles in disease progression.

MIF binds and activates three chemokine receptors, CXCR2, CXCR4, and CXCR7 resulting in leukocyte arrest, chemotaxis of leukocytes and T-cells. The binding events of these receptors are critical for immune responses to foreign pathogens as well as the metastasis of cancer cells [12, 21, 22]. Activation of CXCR2 and CXCR4 by MIF is made possible by the presence of a pseudo-(E)LR [23] and RLR [24] motif respectively. This feature is not observed in D-DT as it lacks the structural characteristics previously mentioned to do so. MIF and D-DT are predicted to have different modes of binding [20] to cell surface receptor CD74 as well as different affinities, with Kds of approximately 1.4×10^{-9} and 5.4×10^{-9}, [9] respectively. Activation of CD74 by either protein leads

to the activation of cell survival and proliferation pathways [25, 26]. Despite the numerous activities of MIF and D-DT, the activation of CD74 is attributed for the majority of their tumorigenic and pro-inflammatory properties [27, 28, 29].

CD74, also known as Invariant chain (Ii), exists as a largely disordered homotrimer that is widely expressed in antigen presenting cells (APCs) where approximately 95% [27] of the time it acts as a chaperone for major histocompatibility class II (MHC-II) proteins [30]. Although it is also expressed in other cell types under inflammatory conditions, particularly epithelial cells [31]. All CD74 isoforms contain a small intracellular transmembrane domain, a CLIP (class II-associated li chain peptide) domain, and a trimerization domain where all three chains converge. The CLIP domain is a small region which is responsible for blocking the antigen binding cleft of MHC-II during transportation from the ER to endosomes, preventing premature antigen recognition [32]. Once CD74 and its attached MHC-II protein is embedded into the endosome, it is then digested, and the CLIP domain eventually leaves the antigen binding cleft. The MHC-II protein can then bind an antigen within the endosome, meanwhile the extracellular and transmembrane domains of CD74 are completely digested eventually leaving only the intracellular domain [33, 34]. The intracellular domain is left behind where it localizes in the nucleus in high quantities [34], although its role there is not completely understood. While CD74's main role is acting as a chaperone for MHC-II proteins, its role as a receptor for MIF and D-DT is crucial for cell survival and proliferation.

CD74 is capable of acting as a receptor for both MIF and D-DT in both its soluble [35] form within the cell, and on the cell surface complexed with co-receptor CD44 [36]. The interaction between soluble CD74 and MIF has been only recently described as blocking normal MIF activity, slowing the growth of tumor cells [35]. In contrast, activation of CD74 on the cell surface by MIF/D-

DT in the presence of the co-receptor, CD44 [36], leads to a cascade of cell growth and proliferation pathways, namely the PI3K/AKT, ERK 1/2, NF- kB, and AMPK pathways [25, 26]. Moreover, the activation of CD74 also results in the inhibition of p53 induced apoptosis [37]. The overexpression of the protein ligands, MIF and or D-DT, often leads to an increase in cell proliferation, resistance to cell death, and inflammation through these aforementioned pathways, resulting in a variety of diseases including cancers [38] and inflammatory diseases, such as rheumatoid arthritis [39]. While it would be ideal to directly study CD74, it has proven to have been difficult to study and has been deemed undruggable. These difficulties result from its high internalization rate [40] and lack of regular secondary structure [41] which is crucial for MHCII chaperone activity. As a result, the majority of the work done towards inhibiting the activation of CD74 has been focused on targeting the protein ligands, MIF and D-DT.

Both MIF and D-DT have been identified as being over expressed and promoting tumor growth in cancers, both of which are commonly reported in NSCLC [42, 43, 44] and glioblastomas [45]. Although MIF overexpression has been identified in more cancers, this is likely due to a lack of research focused on D-DT and not due to differences in activity. This is reflected in the discrepancy of total structures of MIF and D-DT available in the PDB, 101 versus 18, respectively. This is further emphasized by the fact that 61 of the MIF structures contain a small molecule inhibitor or substrate, meanwhile only 6 D-DT structures contain a small molecule. These MIF and D-DT structures have come from a wide variety of studies that began in the late 1990s, the majority of which aimed towards seeking to understand the MIF/D-DT-CD74 activation process. This has been achieved through the elucidation of important dynamic pathways in these proteins, permitting for communications between key residues, as well as attempting to identify and understand the mechanism of action of novel inhibitors [7, 21, 46-52].

A large amount of the work done to elucidate the dynamic pathways between functionally significant residues on MIF and D-DT focuses on the residues within and around the catalytic site as well as residues on the β sheets. While the catalytic and CD74 binding sites do not have any overlap in key residues or regions, mutations that impact catalytic activity often impact CD74 activation activity. It is thought that targeting β-sheet residues for mutagenesis may affect key communication pathways since catalytic residues as well as residues significant for the binding and activation of CD74 communicate using the β-sheets [51-53]. An amino acid, which is key to maintaining β-sheet rigidity and dynamic communications, in both MIF and D-DT was reported to be Y99 [51, 53] and F100 [52], respectively. In both cases, its mutation to alanine caused a loss of both activities, despite these residues not being directly involved in catalysis or CD74 activation. This effect that can be explained by an alteration in β-sheet dynamics [51].

A similar effect was also observed through several mutations of Pro-1, which is the catalytic residue for the keto-enol tautomerase activity of MIF and D-DT [7]. In the case of both proteins, mutation from Proline to any other residue causes a loss of catalytic activity; however, mutation to a residue that would increase the flexibility of the β-sheet, such as glycine, also caused a loss of CD74 activation activity [52, 54, 55]. In MIF, truncation of Pro-1 as well as mutation of Met-2 to alanine also lead to a loss of both catalytic and CD74 activation activity [48]. Meanwhile, a mutation from proline to a more rigid amino acid, such as methionine, maintained CD74 activation activity [48]. Although these results indicate the N-terminus is key for maintaining β-sheet rigidity and therefore communication pathways, its exact effect on the dynamics of the protein as well as regions important to activation of CD74 was not further explored.

Using a combination of molecular dynamics (MD) simulations and NMR experiments spanning from picoseconds to milliseconds, we delved into the potential implications of a MIF N-

terminal activation mechanism of CD74. Our investigation revealed a dynamic connection between the N-terminus and the CD74 activation site, situated on the surface of MIF. Communication between these functional sites occurs through the β-sheet of MIF, significantly influenced by N-terminal flexibility. Additionally, we provided a mechanistic explanation for the factors distinguishing MIF variants as CD74 agonists or antagonists.

with this in mind, the method commonly used as a first step to test the effects of a potential agonistic or antagonistic mutant/small molecule is an assay that measures the ability of MIF or D-DT to tautomerize a model substrate from its keto form to enol form. The most common substrate to use for both proteins is 4-hydroxylphenylpyruvate (4-HPP) which is commercially available from many different vendors. However, with several published Km values for 4-HPP with MIF, it is possible that distributors do not provide equal quality substrates, which may affect Km values for mutants and Ki values for inhibitors. While slight differences of these values are expected, atypically large discrepancies of Michalis-Menten parameters may be indicative of contaminants or differences in purities between different samples. This may lead to incorrect conclusions about different MIF/D-DT variants or about the potency and specificity of potential small molecule inhibitors.

Herein, we establish a standardized protocol for the 4-HPP keto-enol tautomerase assay as well as introduce a novel, spectrophotometric-based method for assessing the activity of HPPD, a protein that utilizes 4-HPP as its biological substrate. Using MIF as a model enzyme, we investigate the potential impact of impurities present in substrate samples on enzymatic activity. Employing a range of complementary techniques, our experimental approach demonstrates that the presence of underrepresented impurities in 4-HPP samples significantly affects enzyme activity. Validation of MIF's kinetic findings was conducted using D-DT and HPPD. These

findings highlight the crucial role of impurities in substrate samples in influencing the accuracy of kinetic data, emphasizing the need for more robust biochemical assays and high-throughput experiments.

Although the catalytic and CD74 binding site residues do not overlap, the majority of inhibitors that reduce MIF or D-DT's ability to bind and activate CD74 bind within the catalytic site [21, 44, 48, 56-59]. While these inhibitors are understood and well-defined crystallographically, their impact on the protein's dynamics is poorly understood. Whereas MIF has a large selection of commercially available and selective inhibitors, D-DT only has two reversible and selective inhibitors, being 4-CPPC [59] and (R)-5-Methyl-3-(1-(naphthalen-1-yl)ethyl)-6-(3-(tri-fluoromethyl)phenyl)thieno[2,3-d]pyrimidine-2,4(1H,3H)-dione, a derivative of thieno[2,3-d]pyrimidine-2,4(1H,3H)-dione [44]. However, the (R)-5-Methyl-3-(1-(naphthalen-1-yl)ethyl)-6-(3-(tri- fluoromethyl)phenyl)thieno[2,3-d]pyrimidine-2,4(1H,3H)-dione derivative was found to be more potent and selective than 4-CPPC while also inhibiting the proliferation of non-small cell lung cancer (NSCLC) cell, it lacks any sort of structural characterization. Meanwhile, first selective and reversible D-DT inhibitor, 4-CPPC, made a clear impact on D-DT dynamics as can be seen at the crystallographic level with the C-terminal lacking electron density after residue V113 [49]. This suggests that the C-terminal becomes highly flexible as part of the induced binding mechanism of this inhibitor. The increased flexibility was unexpected, considering that C-terminal rigidity for MIF is important for proper activity [48].

Previous investigations of the C-terminal of MIF reported that this region has a key role in stabilizing the protein's tertiary structure as well as modulating the protein's catalytic activity [60]. These conclusions were attributed to the ability of C-terminal residues to form intersubunit interactions that stabilized the biological assembly of MIF. Recently, it has been shown that this

region also has a functional role in the activation of CD74 [48, 51]. While these studies have enriched our understanding on the structure and function of MIF, the corresponding analyses for D-DT have not been performed.

Considering this information, we carried out a study to examine the impact of 4-CPPC binding on the dynamic profile of D-DT. MD simulations corroborated previous crystallographic findings and elucidated insights into the role of C-terminal flexibility. After experimentally ruling out structural instability, we demonstrated the dynamic effect of the C-terminal, which facilitates long-range intersubunit communications across the D-DT trimer. These intriguing discoveries offer a novel perspective on understanding the largely unexplored biological functionality of D-DT, particularly focusing on the CD74 activation mechanism. Furthermore, our analysis contributes vital information to the ongoing efforts aimed at elucidating the divergent functional profiles of MIF and D-DT. Considering MIF variants and MIF-ligand complexes as the sole molecules serving as CD74 antagonists, our findings provide crucial insights for the development of new anti-inflammatory and anticancer therapeutics modulating CD74 activity.

Based on these findings, our objective was to identify a new highly selective inhibitor targeting D-DT, intended to serve as a benchmark for future investigations. We present the discovery and characterization of 2,5-pyridinedicarboxylic acid (**1**) as a commercially available, bioactive, and exceptionally selective inhibitor of D-DT. Enzymatic assays demonstrate that compound **1** displays an impressive 79-fold inhibition selectivity for D-DT over MIF, establishing its position as the most discriminating among known D-DT inhibitors. Through crystallographic analysis of the D-DT-**1** complex, in conjunction with comparison to established MIF crystal structures, significant structural disparities within the active sites of MIF and D-DT are unveiled. These insights shed light on the regulation of ligand binding by both proteins and provide clarity

on inhibition selectivity criteria. The identification of compound 1 holds considerable promise for advancing our understanding of D-DT's biological activities in disease models and clinical contexts, while concurrently facilitating the development of next-generation D-DT modulators targeting CD74 activation.

Our examination of 2,5-pyridinedicarboxylic acid revealed that the displacement of the C-terminal, which occurs upon 4-CPPC binding, isn't a prerequisite for inhibiting D-DT-induced CD74 activation. To elucidate the role of the C-terminal in D-DT, we utilized high-resolution protein crystallography alongside various complementary biophysical, biochemical, and computational methods to unravel the structural foundation of D-DT-ligand complex formation. In this pursuit, we engineered seventeen D-DT variants and scrutinized their unique structural features to unveil the mechanistic intricacies governing ligand binding. Employing 4-HPP as the primary substrate and two selective inhibitors (4-CPPC and pyridine-2,5-dicarboxylate) with differing binding modalities, our study represents the first comprehensive analysis of the structural determinants influencing the formation of D-DT-ligand complexes. Through our investigation, we illustrate that the C-terminal region acts as a pivotal regulator of molecular recognition through a multi-faceted control mechanism.

CHAPTER 2: THE N-TERMINUS OF MIF REGULATES THE DYNAMIC PROFILE OF RESIDUES INVOLVED IN CD74 ACTIVATION

Published as: Parkins A, Skeens E, McCallum CM, Lisi GP, Pantouris G. The N-terminus of MIF regulates the dynamic profile of residues involved in CD74 activation. Biophys J. 2021 Sep 21;120(18):3893-3900. doi: 10.1016/j.bpj.2021.08.025. Epub 2021 Aug 24. PMID: 34437846; PMCID: PMC8511117.

Abstract

Macrophage migration inhibitory factor (MIF) is an immunomodulatory protein with a pathogenic activity in various inflammatory disorders, autoimmune diseases, and cancer. The majority of MIF-triggered pathological conditions are associated with activation of the cell surface receptor CD74. In the absence of small molecule antagonists that directly target CD74, MIF variants and MIF-ligand complexes have served as modulators of CD74 activity. These molecules have been reported to have either antagonistic or agonistic effects against the receptor, though the mechanistic parameters that distinguish the two groups are largely unknown. Through molecular dynamics (MD) simulations and nuclear magnetic resonance (NMR) experiments, we explored the relationship between MIF's catalytically active N-terminus and the surface residues important for the activation of CD74. We found that the two sites are connected *via* backbone dynamics that are propagated to the CD74 activation surface of MIF, from the β2 and β4 strands. Our results also provide mechanistic evidence that explain the functional characteristics of MIF variants, serving as CD74 agonists or antagonists. Such findings are of high importance for understanding the MIF-induced activation of CD74 as well as for the development of highly potent CD74 therapeutics.

Introduction

Macrophage migration inhibitory factor (MIF) is a pleiotropic protein with proinflammatory and tumorigenic properties, mainly associated with activation of the cell surface receptor CD74 [27, 28, 29]. The MIF-CD74 signaling axis has been an attractive target for drug development due to its involvement in activation of PI3K/AKT, ERK 1/2, NF-κB, and AMPK pathways that lead to inhibition of apoptosis, cell survival, and proliferation [25]. for decades, all efforts toward the discovery of small molecule CD74 antagonists have failed and the receptor is now considered "undruggable." This conclusion is associated with the intrinsically disordered structure [41] and fast cell internalization rate of CD74 [40] that complicate the in-depth study of the receptor.

The primary sequence of human MIF is composed of 115 amino acids. The crystal structure obtained in 1996 [17], revealed a homotrimeric biological assembly with a solvent channel along the 3-fold axis. Monomeric MIF contains two antiparallel α-helices (α1-α2) and six β-strands (β1-β6). The first four strands (β1-β4) form a b-sheet that defines the solvent channel while the remaining two (β5-β6) play a key role in the inter-subunit communications of the trimer. Homology modeling with two bacterial proteins led to the identification of its keto-enol tautomerase activity [61]. The keto-enol tautomerization site, which is located at the N-terminus of MIF and controlled by the catalytically active residue Pro1 [18], remains the sole target for identification and development of small molecule MIF ligands that serve as modulators of CD74 activity [48, 58, 62, 63]. Pro1 is buried in the catalytic pocket formed by the assembly of the trimer and its key position on the extension of β1 strand regulates the flexibility of the MIF β-sheet. Although the catalytic pocket is not part of the MIF-CD74 interface [48], MIF-ligand complexes and MIF variants possessing N-terminal modifications function as CD74 agonists or antagonists [48, 64

without fully understanding the parameters that distinguish their functional characteristics. Structural analyses of MIF-covalent inhibitor complexes have shown that MIF ligands with chemical moieties penetrating the MIF surface-solvent interface function as CD74 antagonists by interrupting the MIF-CD74 interface located outside the active site pocket [48]. According to the same study, surface residues located on the α1/β2 3.6 helix, β4/α2 loop and C-terminal play a key role in activation of CD74. Since this finding only explains the functional role of a fraction of MIF active site ligands, we still lack comprehensive understanding of the role of the enzymatic pocket and N-terminus in the MIF-induced activation of CD74.

The study of protein dynamics has been traditionally used to understand biological function [65]. for MIF, recent studies have shown the CD74 activation and catalytic sites to be dynamically connected with an allosteric region (Tyr99), located at one of the openings of the solvent channel [51, 53]. This allosteric region communicates with the two functional sites of MIF and influences their activities *via* unique, non-overlapping dynamic pathways. Dynamic communication between the CD74 activation and catalytic sites has yet to be reported.

Using molecular dynamics (MD) simulations and nuclear magnetic resonance (NMR) experiments spanning the picosecond-to-millisecond timescales, we investigated the potential implications of a MIF N-terminal activation mechanism of CD74. Our results reveal that the N-terminus and CD74 activation site, which is located on the surface of MIF, are dynamically connected. Communication between the two functional sites occurs *via* the β-sheet of MIF and is highly influenced by N-terminal flexibility. Here, we also provide a mechanistic explanation of the factors that distinguish MIF variants from functioning as CD74 agonists or antagonists. Considering MIF variants and MIF-ligand complexes are the only available molecules serving as

CD74 antagonists, such findings provide critical insights into the development of new anti-inflammatory and anticancer therapeutics that modulate CD74 activity.

Materials and Methods

Mutagenesis, Protein Expression and Purification.

Wild-type (WT) MIF was cloned in a pET-11b *Escherichia coli* expression vector and the produced plasmid was used as the template for production of the four N-terminal variants (P1M, P1G, ΔP1 and M2A). Mutagenesis for P1M, P1G, and M2A and truncation for ΔP1 were carried out according to the QuikChange site-directed mutagenesis protocol (Agilent). In all cases, the derived clones were confirmed by sequencing. The protein expression and purification processes were similar to what was previously described [53, 66]. WT MIF and MIF variants were grown in LB supplemented with 0.100mg/mL ampicillin, at 37°C, until OD_{600} reached 0.6-0.8. Protein expression was induced by the addition of 1mM isopropyl β-D-1-thiogalactopyranoside (IPTG). After a 4-hour incubation at 37°C, the cells were collected, washed with 20mM Tris pH 7.4, 20mM NaCl (lysis buffer), and stored at -80°C. for protein purification, the cells were resuspended in the lysis buffer and lysed by sonication. Cell debris was removed by centrifugation at 25,300 x g. The supernatant was filtered using a 0.22 μm syringe filter and loaded onto 5 ml and 120 ml Q-Sepharose columns (GE Healthcare) connected in series. The protein did not bind on the columns and were collected in the flow through with ~95% purity. The remaining contaminants were removed by size exclusion chromatography column (16/60 Superdex 75), running in the lysis buffer. Purity and protein concentration were determined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and the Pierce™ BCA protein assay kit (Thermo Fisher Scientific), respectively.

Circular Dichroism (CD) Spectroscopy.

Folding and thermal denaturation CD experiments were performed in a 1 mm quartz cuvette using a JASCO J-810 spectropolarimeter. The 20 μM MIF samples were prepared in 20 mM sodium phosphate pH 7.0, 1 mM EDTA (folding experiments) and 20 mM Tris pH 7.4, 20 mM NaCl (thermal denaturation experiments). Before each experiment, the samples and buffer were degassed with ultra-high purity nitrogen. The folding profiles of WT MIF and MIF variants were monitored between 260 nm and 195 nm, at 25°C. The thermal denaturation experiments were carried out by monitoring the variation of CD signal at 218 nm. CD spectra of MIF proteins were collected between 20°C to 90°C with a temperature change of 0.5 °C/min. All data were analyzed in ORIGIN 2019b. The CD experiments were repeated in triplicate.

MD Simulations.

The crystal structures of WT MIF (PDB entry:3DJH), P1M (PDB entry:4PKZ), M2A (PDB entry:4XX7), and ΔP1 (PDB entry:4XX8) were used as starting models for the 200 ns MD simulations. Mutation of Pro1 to Gly for the P1G starting model and addition of missing residues were carried out with Chimera [67]. The VMD plugin, psfgen, was used to add hydrogens and generate starting protein structure and coordinate files using the CHARMM36 [68] topology. The produced files were solvated with TIP3P waters and the charge of the system was neutralized with chloride ions using VMD's solvate and ionize modules respectively [69]. The structures were then minimized and heated to 300K before subsequent equilibration and production runs. Minimization and heating followed a three-step process: first the water molecules, second the amino acid sidechains, and finally the entire system was minimized and heated. Minimization took place over 20 ps and heating over 120 ps with a 2 fs time-step. The system was equilibrated at 300K for 1 ns. *G_correlation* [70], a GROMACS [71] plugin, was utilized to produce generalized cross correlation

data for α-carbon atoms. Correlation values associate the fluctuation of atoms belonging to a group (group = side chain, backbone, or residue) with another. Correlation values range from 0 to 1, with 0 indicating no correlation and 1 indicating that there is absolute correlation. Absolute correlation is only seen when comparing the residue of interest with itself. Large correlation values are expected when analyzing adjacent residues. On the other hand, large correlation values between non-adjacent residues can provide insights into dynamic signal transmission and regulation of protein functionality. RMSF data was generated for the α-carbon atoms using GROMACS. Root-mean-square deviation (RMSD) values for the crystal structures of MIF proteins, were obtained using the CCP4 supported program SUPERPOSE [72]. The resulting MD simulation models were analyzed in PyMOL [73]. All calculations were repeated in triplicate.

^{15}N-labeled protein expression and purification.

Isotopically labeled MIF samples for NMR studies were grown in M9 minimal media supplemented with ^{15}N-ammonium chloride (Cambridge Isotope Laboratories). The cultures were grown at 37°C to an OD_{600} of 0.8 – 1.0 and then induced with 1 mM IPTG, followed by an additional incubation at 20°C for 16 – 18 hours. for ΔP1, the cells were grown at 37°C to an OD_{600} of 0.8 and then induced with 1 mM IPTG at the same temperature, for 4 hours. Cells were subsequently harvested by centrifugation and stored at -80°C. ^{15}N-MIF samples were purified as described above.

NMR Spectroscopy.

for NMR experiments, the samples were dialyzed into a final buffer containing 20 mM sodium phosphate, 1 mM EDTA, and 7.5% D_2O at pH 7.0, and then concentrated to 0.5 – 1.0 mM. ^{1}H-^{15}N TROSY HSQCs and spin relaxation experiments were carried out on a Bruker Avance NEO 600 MHz spectrometer at 30°C. The ^{1}H and ^{15}N carrier frequencies were set to water

resonance and 120 ppm, respectively. All NMR spectra were processed with NMRpipe [74] and analyzed in Sparky [75]. Longitudinal and transverse relaxation rates were measured with relaxation times of 0, 20 (x2), 60 (x2), 100, 200, 600 (x2), 800, 1200 ms for R_1 and 16.9, 33.9 (x2), 67.8, 135.7 (x2), 169.6, 203.5 (x2) ms for R_2. Longitudinal and transverse relaxation rates were determined from peak intensities of each amide resonance at multiple delay points, as quantified in Sparky. Steady-state ^{1}H-[^{15}N] NOE were measured with a 6 second relaxation delay followed by a 3 second saturation (delay) for the saturated (unsaturated) experiments. Relaxation experiments were carried out in a temperature-compensated interleaved manner, were processed with in-house scripts, and were analyzed in GraphPad Prism 8.0 (GraphPad Software). Trimmed means for chemical shift perturbation analyses and relaxation data were determined by excluding the top and bottom 10% of the data sets. Standard deviations were also determined from these data sets.

Results and Discussion

Selection of MIF variants to dynamically characterize the N-terminus

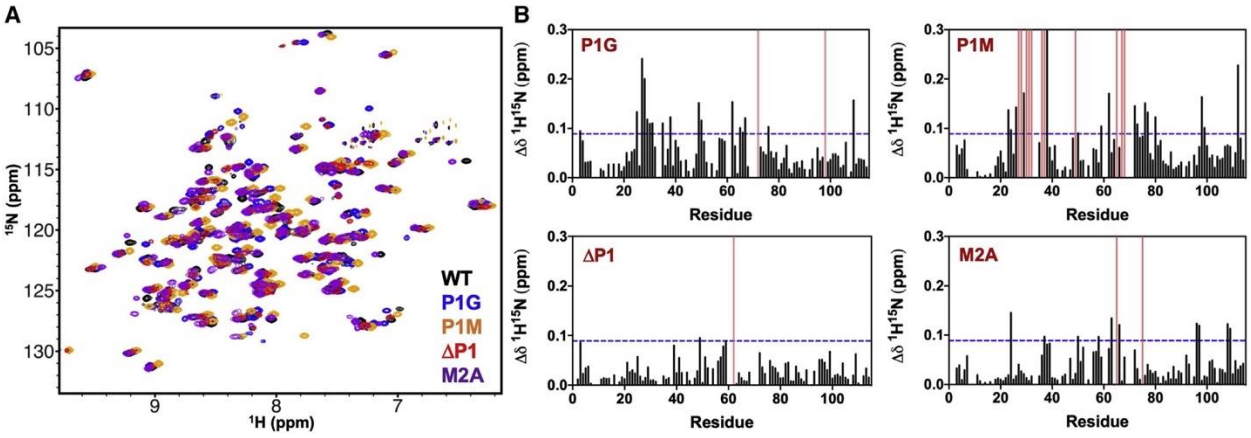
To investigate if the MIF N-terminus has a functional role in activation of CD74, we utilized selected MIF variants and employed MD simulations and NMR spin relaxation experiments. The experimental design was based on a successful approach we previously applied to obtain mechanistic insights into the dynamic signal transmission between distal MIF sites [51, 53]. for this study, we used four N-terminal MIF variants (P1M, P1G, ΔP1 and M2A) as well as the wild-type protein (**Supplemental Fig. 1**). We anticipated each of these variants would assign a unique dynamic identity to the MIF N-terminus. Proline is the least flexible amino acid, thus mutation to glycine was expected to escalate N-terminal flexibility (**Supplemental Fig. 1**). with P1G, our goal was to study how a highly dynamic N-terminus would influence the flexibility of

the MIF β-sheet and subsequently the surface responsible for activation of CD74 [48]. In contrast to P1G, ΔP1 would provide information on how the β1 strand of MIF was dynamically influenced by the absence of the rigid residue, Pro1 (**Supplemental Fig. 1**). M2A was used as an alternative means of altering the flexibility of the β1 strand, without removing Pro1. Superposition of multiple MIF-ligand crystal structures [53] has shown the side chain of Met2 to adopt multiple conformations and since Met2 is the first amino acid of the β1 stand, the conformational position of its side chain has a direct impact on the strand's flexibility. Finally, P1M was used as a positive control. In contrast to Met2, Met1 has reduced flexibility due to its position in the active site and it mimics Pro1 in WT MIF crystal structure (**Supplemental Fig. 1**).

Structural stability and folding profiles of MIF variants

Figure 2.1.

NMR fingerprinting of MIF variants.



Note. **A**. ^{1}H-^{15}N HSQC NMR spectra of P1G (blue), P1M (orange), ΔP1 (red), and M2A (purple) MIF overlaid with WT MIF (black). **B**. NMR chemical shift perturbations for MIF variants. The ^{1}H-^{15}N combined chemical shift is plotted for each residue in MIF for P1M, P1G, ΔP1, and M2A

(Figure 2.1 Continued)

MIF. In all cases, perturbations were determined by $\sqrt{\frac{1}{2}\left((\delta H)^2+\frac{((\delta N)^2)}{25}\right)}$, where δH and δN were calculated as ($\delta_{Variant}$ - δ_{WT}). Blue lines represent the 1.5σ above the 10% trimmed mean of all shifts. Pink bars denote residues that are broadened beyond detection.

We first investigated whether P1M, P1G, ΔP1 and M2A MIF proteins were properly folded and structurally stable (**Fig. 1**, **Supplemental Fig. 2**). This step is mandatory to confirm that data reporting on MIF dynamics were not influenced by the structural instability of the MIF variants. To achieve this, we utilized heteronuclear single quantum coherence (HSQC) and circular dichroism (CD) spectroscopy experiments. Despite significant chemical shift perturbations indicative of local structural alterations, the overall fold of the MIF protein is maintained, evidenced by the similar dispersion of resonances in ^{1}H-^{15}N HSQC NMR spectra (**Fig. 1A**). Independently, CD experiments further support the NMR findings demonstrating the overall fold of the MIF variants has not been significantly influenced by the N-terminal modifications (**Supplemental Fig. 2A**). Structural stability of MIF proteins was measured by temperature-dependent CD experiments, over a range of 20 - 90°C (**Supplemental Fig. 2B**). Using the melting profile of WT MIF as a reference (T_m= 74 ± 0.1°C), it was shown that the MIF variants have similar thermal stabilities, proving they are suitable for protein dynamic studies. The difference noted in the CD spectra between WT MIF and ΔP1 is related to the reduced β-stand percentage of the latter, due to Pro1 truncation. Pro1 is not part of β1-strand, though the rigidity it provides affects the flexibility of the N-terminus and thus the length of the β1 strand (**Supplemental Fig. 2C**). Crystallographic analysis of ΔP1 has shown the effect of this truncation is local, without any

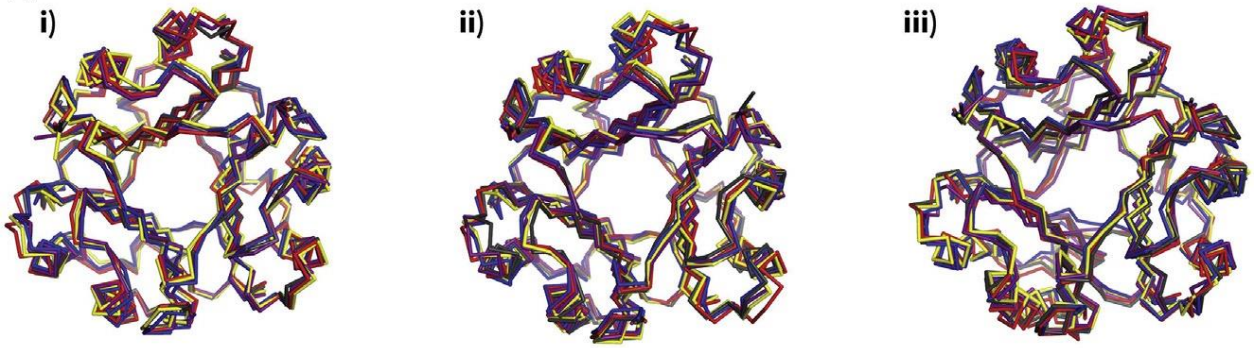
drastic impact on the stability of the MIF trimer (**Supplemental Fig. 2C**). Notably, the crystal structures of ΔP1 and the remaining P1G, P1M and M2A variants demonstrate high superposition accuracy to WT MIF with root-mean-square deviation (RMSD) values between 0.11Å and 0.25Å (**Supplemental Fig. 3**). The highest deviation, 0.25Å, observed for P1M, is related to minor displacements of amino acids surrounding Met1. Collectively, the NMR, CD, and crystallographic data confirm that the P1M, P1G, ΔP1 and M2A proteins are stable, properly folded and can therefore be used for protein dynamic studies.

Global and local dynamic fluctuations of MIF variants due to N-terminal alterations

Figure 2.2.

Dynamic profiles of MIF variants.

A



B

Global MIF Dynamics		Local MIF Dynamics		
	Trimer - RMSF (Å)		Amino Acid 1 - RMSF (Å)	Amino Acid 2 - RMSF (Å)
WTMIF	0.60 ± 0.16	WTMIF	0.69 ± 0.05	0.54 ± 0.01
P1M	0.60 ± 0.18	*P1M*	0.63 ± 0.02	0.53 ± 0.01
P1G	0.59 ± 0.16	*P1G*	0.85 ± 0.08 **(+22.9 % vs WTMIF)**	0.56 ± 0.04
ΔP1	0.60 ± 0.16	*ΔP1*	-	0.62 ± 0.09 **(+14.4 % vs WTMIF)**
M2A	0.61 ± 0.17	*M2A*	0.71 ± 0.04	0.54 ± 0.04

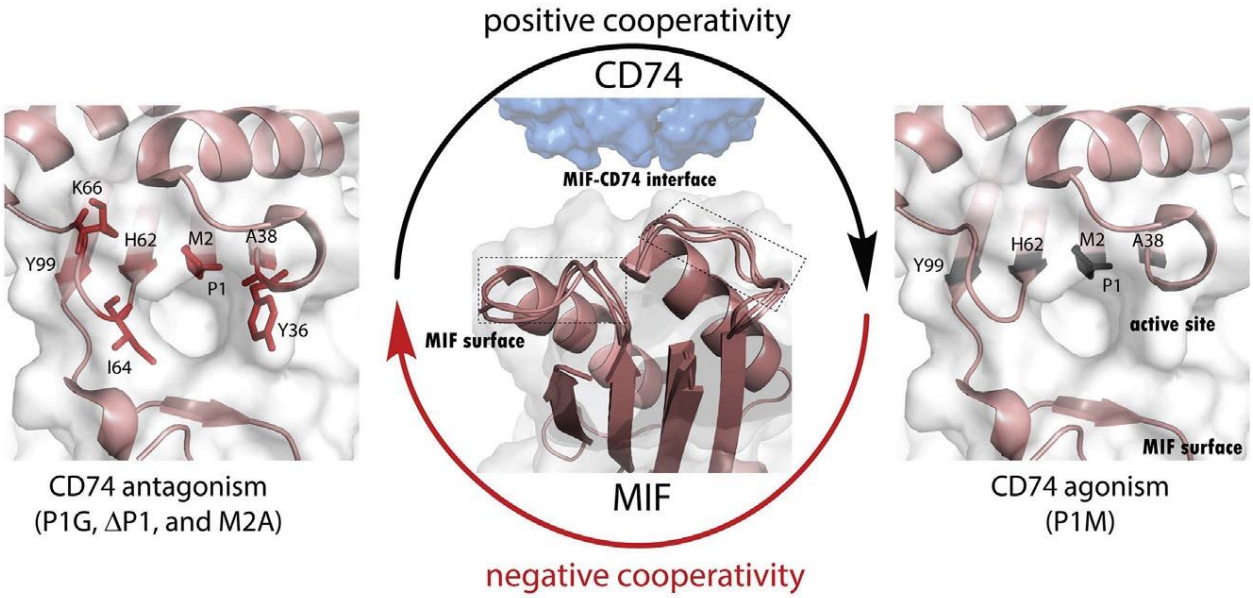
Note. **A.** The structural profiles of P1G (blue), P1M (yellow), ΔP1 (red), and M2A (purple) MIF were captured at (**i**) 0 ns, (**ii**) 100 ns and (**iii**) 200 ns simulation time points and overlaid on the corresponding profile of WT MIF (black). **B.** Global and local RMSF values for WT MIF and MIF variants. Amino acid 1 corresponds to Met, Gly, or Pro for P1M, P1G, and M2A, respectively. Amino acid 2 corresponds to Met for P1M, P1G and ΔP1 or Ala for M2A.

We evaluated the impact of P1M, P1G, ΔP1 and M2A on global (MIF trimer) and local (N-terminus) dynamics by performing 200 ns MD simulations. Alignment of the structural profiles of MIF variants onto WT MIF, at 100 ns intervals (**Fig. 2A**), and examination of the root mean square fluctuation (RMSF) values (**Fig. 2B**) revealed no drastic changes to the trimer. Together, these findings support the conclusion that the N-terminal alterations do not influence the global dynamic properties of MIF. The RMSF value of 0.60 ± 0.16 Å, for WT MIF, indicates the protein is not very flexible. Previously published 1μs MD simulations [51] are in agreement with this finding, confirming the limited flexibility of WT MIF. Shifting the focus to local dynamics, P1G and ΔP1 demonstrated a drastic increase in N-terminal fluctuation (**Fig. 2B**). In contrast to WT MIF, mutation of Pro to Gly, in P1G, increases the RMSF value of residue one by 22.9%, while truncation of the rigid residue Pro1, in ΔP1, increases flexibility of Met2 by 14.4%. As expected, the presence of Pro and Met at position one of M2A and P1M, respectively, retained the N-terminal rigidity of these mutants at RMSF values similar to that of WT MIF (**Fig. 2B**).

Correlation of N-terminal fluctuations with the MIF-induced activation of CD74

Figure 2.3.

Positive and negative cooperativity of MIF variants.



	aa	1*	2**	Tyr36	Ala38	His62	Ile64	Lys66	Tyr99
MIF variants	***P1M***	+1.41	+4.93	+9.87	**+12.4**	+5.44	-4.49	+3.97	**+11.0**
	P1G	**-25.4**	**-14.8**	**-21.3**	**-17.6**	**-17.0**	**-19.5**	-6.35	**-11.6**
	ΔP1	-	**-14.4**	**-23.3**	**-17.6**	**-14.3**	-4.49	**-11.5**	-6.85
	M2A	+0.71	0	-2.55	+0.69	-7.82	**-10.5**	-9.13	-6.85

Note. The percent changes, increase (+) or decrease (-), of correlation for MIF variants are shown in relation to WT MIF. Each of the selected residues has unique properties, as explained in the text. Data at values > ±10% are shown in bold. P1M is the only variant demonstrating positive cooperativity of key residues (Ala38 and Tyr99). Functional assays showed that P1M is a CD74 agonist [48]. The N-terminal alterations of P1G, ΔP1, and M2A are linked to a negative cooperativity (affected residues are shown in red) and CD74 antagonism [48]. The amino acids Tyr36, Ile64, and

(Figure 2.3 Continued)

Lys66 are surface exposed and important for activation of CD74. *Amino acid 1 corresponds to Met, Gly, or Pro for P1M, P1G, and M2A respectively. **Amino acid 2 corresponds to Met for P1M, P1G and ΔP1 or Ala for M2A.

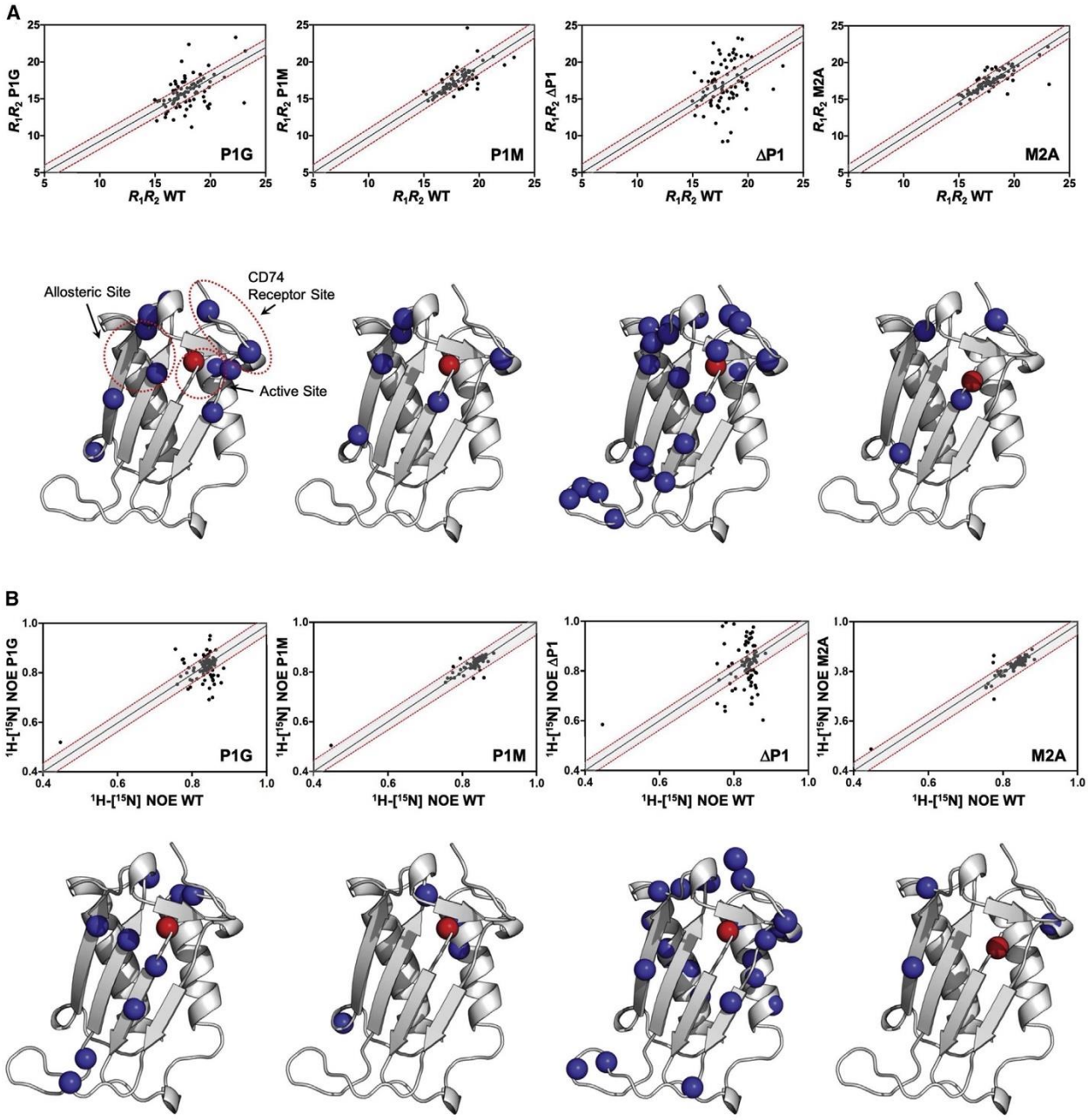
To probe how N-terminal fluctuations may affect the ability of MIF to activate CD74, we utilized g_correlation [70] to produce generalized cross correlation data for the C_α atoms of WT MIF and MIF variants (**Supplemental Fig. 4**). The global correlation features of MIF were previously described [51]. Considering the four variants do not influence the global flexibility of MIF (**Fig. 2B**), our analysis focused on the region around the N-terminus and more specifically the first two residues of MIF, the surface residues Tyr36, Ile64 and Lys66 and the first residues of β2 (Ala38) and β4 (His62) strands (**Fig. 3, Supplemental Table 1**). A MIF-induced CD74 activation assay has shown Tyr36, Ile64 and Lys66 are important for activation of the receptor *in vivo* (Pantouris, et al., 2015). The surface residue Tyr36 is located on the α1/β2 3.6 helix, which is connected to Ala38. The other two key residues, Ile64 and Lys66, are found on the β4/α2 loop, which is joined with His62. We anticipated an increase of β1 fluctuation, caused by the N-terminal alterations, would affect the backbone hydrogen bonding of the adjacent β2 (Ala38) and β4 (His62) strands, transmitting dynamic signals to the surface. If this hypothesis is correct, it could better explain how MIF variants with modified N-termini function as CD74 agonists or antagonists. Our data were interpreted in relation to WT MIF, demonstrating that the four variants can be distributed in two groups: the positive (P1M) and negative (P1G, ΔP1 and M2A) effector variants (**Fig. 3**). The P1M mutant increased the overall correlation of Ala38 and the allosteric residue Tyr99 [51, 53] with the MIF trimer by 12.4% and 11%, respectively. This positive cooperativity, observed for P1M,

has a functional explanation. MIF-induced CD74 activation assays revealed that P1M functions as CD74 agonist, having activity similar to that of WT MIF [48]. These findings provide the first piece of evidence to explain the functional features of a CD74 agonist. On the other hand, the negative effector group composed of P1G, ΔP1 and M2A reduced the overall correlation of key amino acids that are directly (part of MIF-CD74 interface) or indirectly (transmit dynamic signals to the MIF-CD74 interface) related to activation of CD74. for the P1G mutant, Tyr36, Ile64, and Tyr99 are affected the most, dropping their overall correlation with the MIF trimer by 21.3%, 17.6% and 11.6%, respectively (**Fig. 3**). Considering the crystal structure of P1G does not demonstrate any global or local structural features to explain this outcome (**Supplemental Fig. 3**), reduction in correlation must be related to the increased flexibility of the N-terminus. Similar to P1G, the ΔP1 variant reduced the overall correlation of Tyr36 and Lys66 by 23.3% and 11.5%, respectively, simultaneously disrupting the harmonic communication of β1 with the β2 and β4 strands (**Fig. 3**). The last member of the negative effector group, M2A, is affected the least due to the presence of Pro1, which controls the fluctuation of the N-terminus and preserves communication with the surface. Nevertheless, Ile64, found to be important for activation of CD74 [48], demonstrated reduction of its overall correlation by 10.5%. The negative effector role, assigned to P1G, ΔP1 and M2A is explained by functional assays, which showed that the three variants function as CD74 antagonists [48, 55].

NMR analyses of the MIF variants

Figure 2.4.

Relaxation parameters for WT MIF and MIF variants.



(Figure 2.4 Continued)

Note. **A.** R_1R_2 correlations between WT MIF and P1G, P1M, ΔP1, and M2A variants. **B.** ^{1}H-[^{15}N] NOE correlation between the same samples. In **A** and **B**, points that diverge from linear least squares fits (black lines) denote differences in NMR-detected relaxation parameters at specific residues in these samples. Gray shaded areas denote 1.5σ boundaries from the 10% trimmed mean of all relaxation data. Data points outside of these boundaries are deemed to be significant dynamic differences. Results of dynamic correlation analyses mapped onto the structure of the MIF monomer for (**A**) R_1R_2 and (**B**) ^{1}H-[^{15}N] NOE. Residues that fall outside of the 1.5σ boundaries are identified with blue spheres. The site of mutation or deletion is indicated by a red sphere. Critical regions of MIF identified in recent biophysical studies are highlighted with red dashed circles.

We next experimentally assessed the dynamic impact of N-terminal modifications, using NMR. We used spin relaxation experiments (T_1, T_2, ^{1}H-[^{15}N] NOE) to examine how global and local motions of MIF, which affect its function, were altered by these mutations. Per-residue plots of R_1, R_2, and ^{1}H-[^{15}N] NOE values (**Supplemental Fig. 5**) show that the backbone dynamics of MIF at the global level are not substantially altered, however correlation plots for each of these variants versus WT MIF detect clear differences in local regions of the protein (**Fig. 4**). Interestingly, each MIF variant has a unique impact on dynamics surrounding the N-terminus. We dissected these contributions across several timescales, where R_1R_2 relaxation rates, plotted as a product to suppress artifacts from anisotropic molecular tumbling, qualitatively report on micro-millisecond (μs-ms) protein motions (via elevated values driven by R_2 and R_{ex}) and ^{1}H-[^{15}N] NOEs better reflect pico-nanosecond motions. In both analyses, Pro1 modifications induced the greatest

change in MIF dynamics, evidenced by a stronger deviation from linearity in correlation plots. These data are consistent with increased line broadening in P1M and P1G NMR spectra (**Fig. 1B**) and indicate that disruption of Pro1 propagates dynamic changes to its surrounding residues. In the case of P1M, the broadened beyond detection points is related with conformational changes occurring due to the mutation of Pro1 to Met. Nevertheless, these conformational changes do not affect MIF folding, stability and CD74 activation potency, as shown by the CD, crystallographic and previously published *in vivo* functional data.

Consistent with our MD simulations, the P1G variant most strongly alters the dynamics of MIF. Interestingly, this effect is most pronounced in both R_1R_2 and ^{1}H-[^{15}N] NOE data, though the exact residues affected are different (**Fig. 4A and 4B**). Specifically, the P1G variant exhibits pronounced dynamic influence on residues 28-36, 44-49, and 60-68 (*via* elevated R_1R_2 and/or depressed ^{1}H-[^{15}N] NOE correlations), many of which fall on the β2 and β4 strands as highlighted by MD simulations. In addition, the N-terminus is unsurprisingly affected, as are other isolated regions that are mapped onto the MIF structure (**Fig. 4A and 4B)**. Weaker effects on R_1R_2 were observed in the P1M, ΔP1, and M2A variants, although residues 60-68 (P1M), 35-40, 64-66, 99-101 (ΔP1), and Pro1 and Lys66 (M2A) still display significant motional variation from WT MIF at sites of critical MD correlation. Changes in ^{1}H-[^{15}N] NOE correlations, relative to WT MIF, were less pronounced. However, subtle differences in the dynamics of these variants coupled with significant chemical shift perturbations (*i.e.*, local structural effects) at the C-terminal receptor binding region (**Supplemental Fig. 1B**) may help to explain their diminished CD74 activation.

Conclusions

for highly flexible proteins, conformational changes are often the guide for understanding the catalytic, folding, or activation mechanisms [76]. for CD74, direct dynamic interrogation is

unlikely to yield any mechanistic insights, due to its highly disordered structure [41]. Therefore, indirect dynamic approaches that include the protein ligand, MIF, is the only valid route. We have successfully utilized MIF dynamics to identify an allosteric site that modulates the MIF-induced activation of CD74 in the ns-μs timescale [51]. Using protein dynamic methods, we also highlighted the role of β-sheet in dynamic signal transmission between distal MIF sites [51, 53]. In this study, we applied a similar approach aiming to understand the role of the MIF N-terminus in activation of CD74 and clarify the mechanistic parameters that determine the agonistic or antagonistic properties of MIF variants. Using MD simulations and NMR experiments that explored the dynamic behavior of MIF variants, we determined how the N-terminus dynamically regulates the surface of MIF. Increasing the N-terminal flexibility disrupted the harmonic communication with the neighboring β2 and β4 strands, affecting key MIF surface residues involved in activation of CD74. from these findings, it becomes apparent that Pro1 serves as an anchor, controlling the flexibility of the β1 strand and subsequently the dynamics of the entire β-sheet. A previous study has shown that structural changes at the MIF N-terminus, due to inhibitor binding, impact the MIF-induced activation of PI3Ks in leucocytes (12). Our study provides a mechanistic explanation for the above observation. We hypothesize that MIF ligands can have a distinct regulatory effect on CD74 activity; agonistic or antagonistic, based on their ability to modulate the conformational flexibility of N-terminus.

Through a detailed correlation analysis of WT MIF and MIF variants, we identified the functional characteristics required by MIF proteins in order to serve as CD74 agonists or antagonists. A potential CD74 agonist requires significant enhancement of the overall correlation of amino acids that directly (e.g., Tyr36, Ile64, and/or Lys66) or indirectly (e.g., Ala38, His62, and/or Tyr99) affect the CD74 activation surface of MIF. A potential CD74 antagonist has the

opposite effect, reducing the overall correlation of functionally important residues. Regulation of biological activity by positive or negative cooperativity was also shown in other protein classes, including kinases and G-protein coupled receptors [77]. In these studies, MD simulations played a key role in understanding the connection between protein dynamics and biological activity. for this study, the MD simulations are supported by NMR experiments and previously published CD74 activation assays [48]. Such findings provide the first piece of evidence for understanding the mechanistic differences between MIF variants serving as CD74 agonists or antagonists and create a foundation for the development of highly potent CD74 modulators.

Supplemental Information

Supplemental Table 2.S1.

Correlation analysis of key residue for MIF variants.

WTMIF								
aa	**Pro1**	**Met2**	**Tyr36**	**Ala38**	**His62**	**Ile64**	**Lys66**	**Tyr99**
Trial 1	0.30	0.30	0.33	0.31	0.29	0.28	0.26	0.31
Trial 2	0.28	0.27	0.30	0.27	0.28	0.25	0.25	0.28
Trial 3	0.27	0.28	0.31	0.29	0.31	0.27	0.25	0.29
Average	**0.28**	**0.28**	**0.31**	**0.29**	**0.29**	**0.27**	**0.25**	**0.29**
SD (±)	**0.02**	**0.02**	**0.02**	**0.02**	**0.02**	**0.01**	**0.01**	**0.01**
P1M								
aa	**Met1**	**Met2**	**Tyr36**	**Ala38**	**His62**	**Ile64**	**Lys66**	**Tyr99**
Trial 1	0.33	0.36	0.41	0.40	0.38	0.29	0.30	0.37
Trial 2	0.26	0.23	0.26	0.25	0.26	0.23	0.23	0.26
Trial 3	0.27	0.30	0.36	0.33	0.29	0.25	0.27	0.34
Average	**0.29**	**0.30**	**0.35**	**0.33**	**0.31**	**0.26**	**0.26**	**0.32**
SD (±)	**0.04**	**0.07**	**0.08**	**0.08**	**0.06**	**0.03**	**0.03**	**0.06**
P1G								
aa	**Gly1**	**Met2**	**Tyr36**	**Ala38**	**His62**	**Ile64**	**Lys66**	**Tyr99**
Trial 1	0.20	0.22	0.24	0.22	0.22	0.19	0.21	0.23
Trial 2	0.21	0.22	0.20	0.20	0.23	0.21	0.22	0.24
Trial 3	0.23	0.28	0.30	0.30	0.29	0.25	0.27	0.31
Average	**0.21**	**0.24**	**0.25**	**0.24**	**0.24**	**0.22**	**0.24**	**0.26**
SD (±)	**0.02**	**0.04**	**0.05**	**0.05**	**0.04**	**0.03**	**0.03**	**0.04**
□1P								
aa	**Pro1**	**Met2**	**Tyr36**	**Ala38**	**His62**	**Ile64**	**Lys66**	**Tyr99**

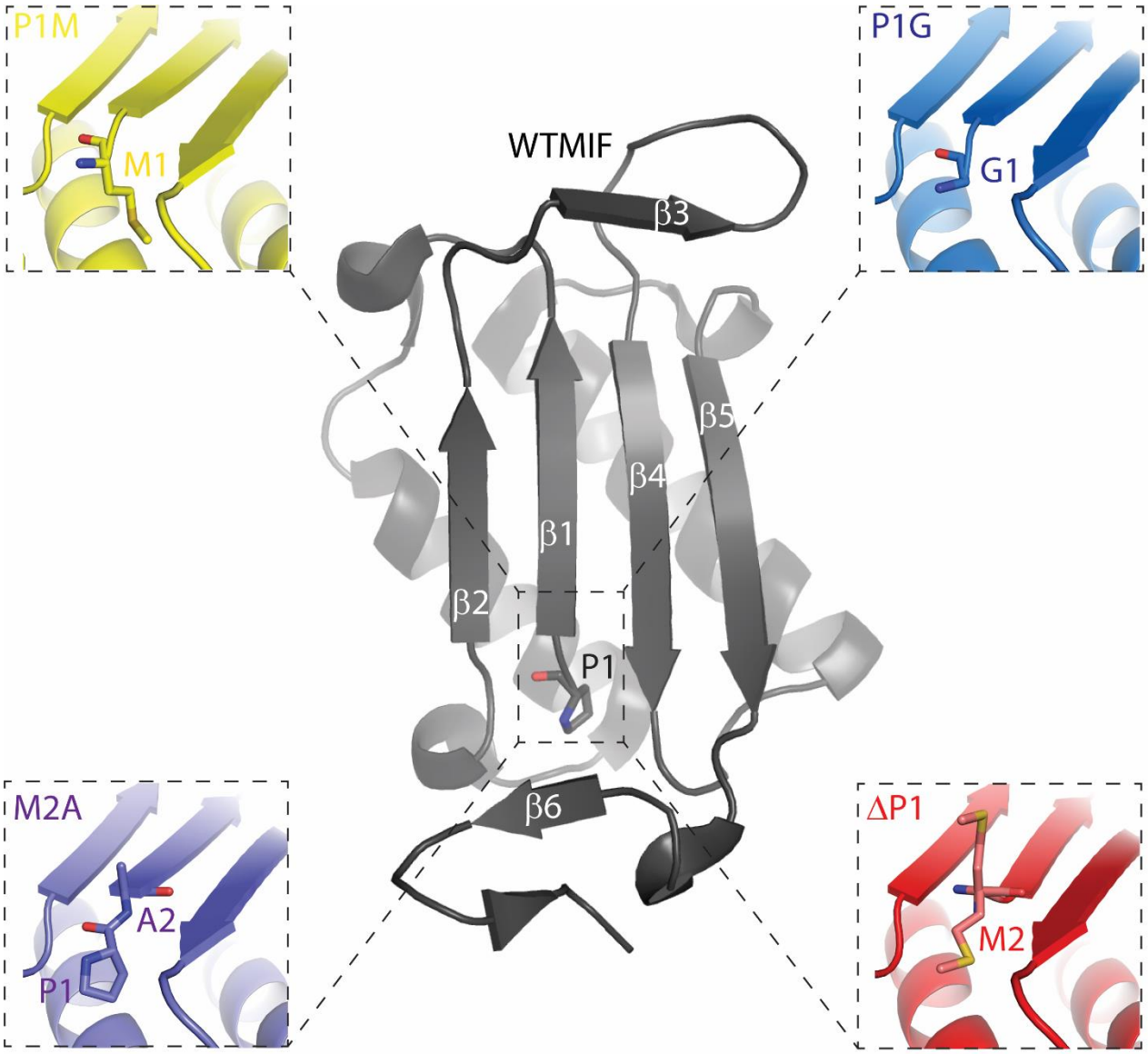
(Table 2.S1. continued)

Trial 1	-	0.24	0.22	0.23	0.27	0.25	0.24	0.29
Trial 2	-	0.25	0.26	0.25	0.27	0.24	0.24	0.29
Trial 3	-	0.24	0.24	0.23	0.22	0.21	0.12	0.24
Average	-	**0.24**	**0.24**	**0.24**	**0.25**	**0.23**	**0.22**	**0.27**
SD (±)	-	**0.003**	**0.02**	**0.01**	**0.02**	**0.02**	**0.03**	**0.03**
M2A								
aa	**Pro1**	**Ala2**	**Tyr36**	**Ala38**	**His62**	**Ile64**	**Lys66**	**Tyr99**
Trial 1	0.30	0.31	0.34	0.31	0.30	0.26	0.25	0.29
Trial 2	0.25	0.24	0.26	0.25	0.25	0.22	0.22	0.25
Trial 3	0.31	0.30	0.32	0.31	0.27	0.23	0.22	0.27
Average	**0.29**	**0.28**	**0.31**	**0.29**	**0.27**	**0.24**	**0.23**	**0.27**
SD (±)	**0.03**	**0.04**	**0.04**	**0.04**	**0.02**	**0.02**	**0.02**	**0.02**

Note. The table displays global correlation values of the reported residues, representing the average correlation of these residues with the MIF homotrimer. While there are a few residues that a single amino acid has a strong correlation with, the majority of amino acids have very little to none, which explains the small global correlation value.

Supplemental Figure 2.S1.

The MIF variants used in this study.



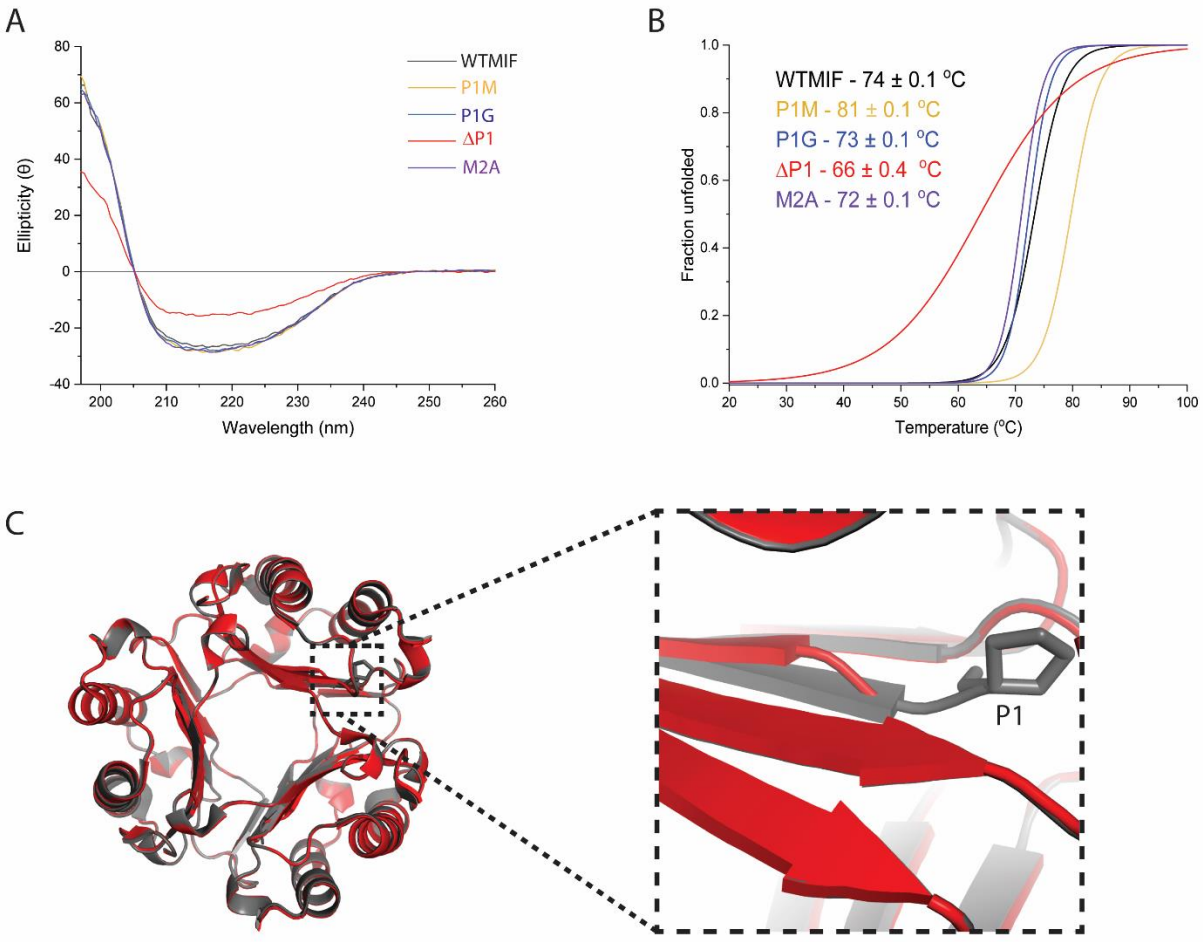
Note. The structure of monomeric WT MIF (PDB entry: 3DJH) is shown in black. The β-strands are labeled and the catalytically active N-terminus amino acid, Pro1, is shown as sticks. Although the biological assembly of MIF is homotrimeric, we illustrate only one monomer for simplicity. The N-terminal sites of MIF variants P1M (yellow), P1G (blue), M2A (purple) and ΔP1 (red) are

(Figure 2.S1 Continued)

shown at the four corners and obtained from the PDB entries 4PKZ, 1P1G, 4XX7, and 4XX8 for P1M, P1G, M2A, and ΔP1, respectively. Notably, the N-terminal residue of ΔP1, Met2, adopts two conformations.

Supplemental Figure 2.S2.

Analysis of protein folding and thermal stability of WT MIF and MIF variants.



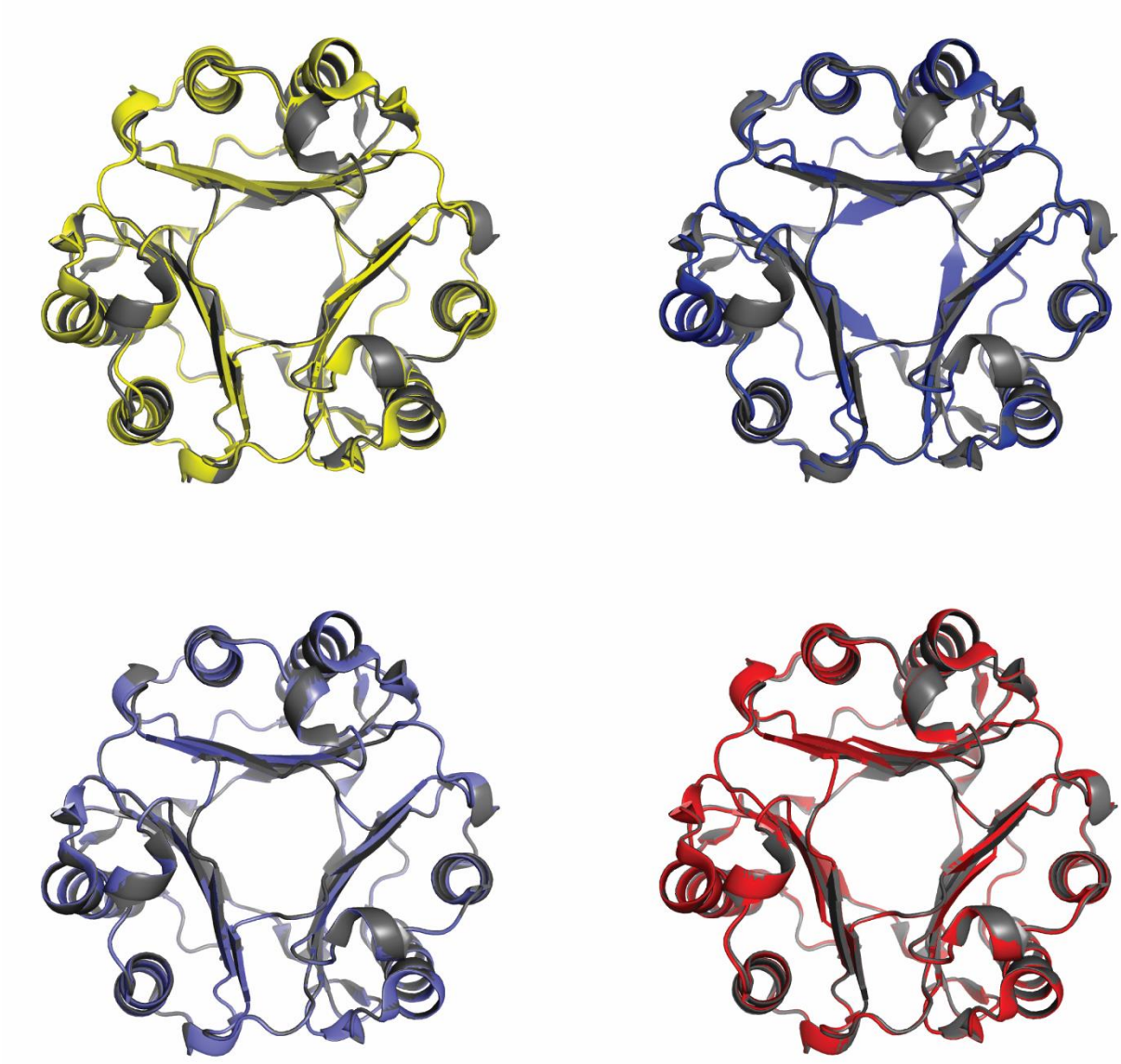
Note. **A.** The secondary structure profiles of WT MIF, P1M, P1G, ΔP1, and M2A are shown in black, yellow, blue, red, and purple, respectively. Folding profiles were monitored between 260nm and 195nm, at 25°C. **B.** Thermal denaturation curves of WT MIF and MIF variants. As before,

(Figure 2.S2 Continued)

WT MIF, P1M, P1G, ΔP1, and M2A are shown in black, yellow, blue, red, and purple, respectively. CD scans were obtained at 218nm. **C.** Alignment of WT MIF (black) with ΔP1 (red) crystal structures demonstrates no global differences between the two proteins. Nevertheless, ΔP1 reveals a shorter β1 strand due to Pro1 deletion (inset). The shorter β1 strand of ΔP1 explains why the CD spectra of WT MIF and ΔP1 differ.

Supplemental Figure 2.S3.

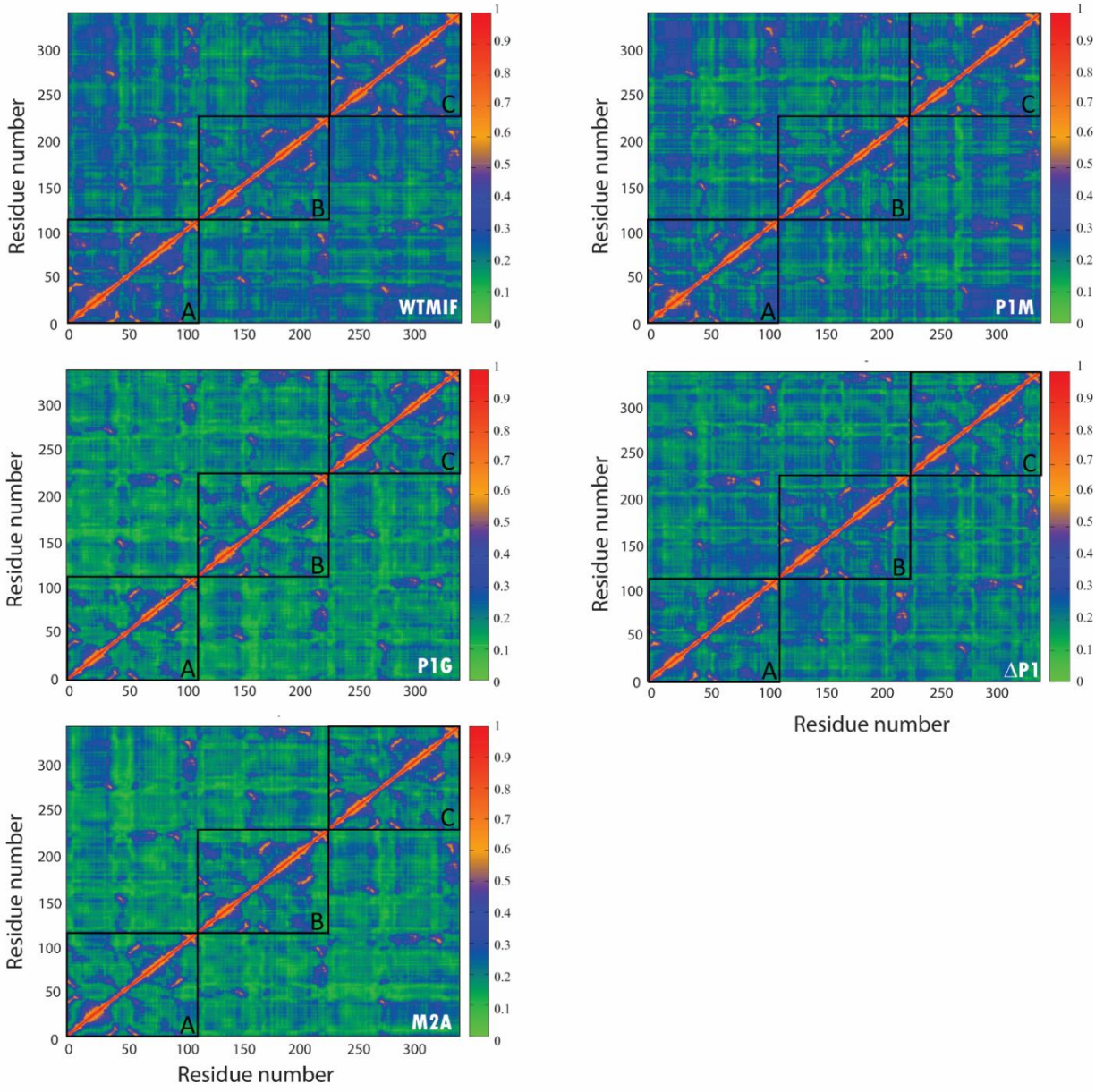
Superposition of WT MIF (black) crystal structure with the corresponding structures of P1M (yellow), P1G (blue), M2A (purple), and ΔP1 (red).



Note. All MIF variants show high superposition agreement to WT MIF with root-mean-square deviation (RMSD) values between 0.109Å and 0.245Å. The files used for alignment were obtained from Protein Data Bank as specified in methods.

Supplemental Figure 2.S4.

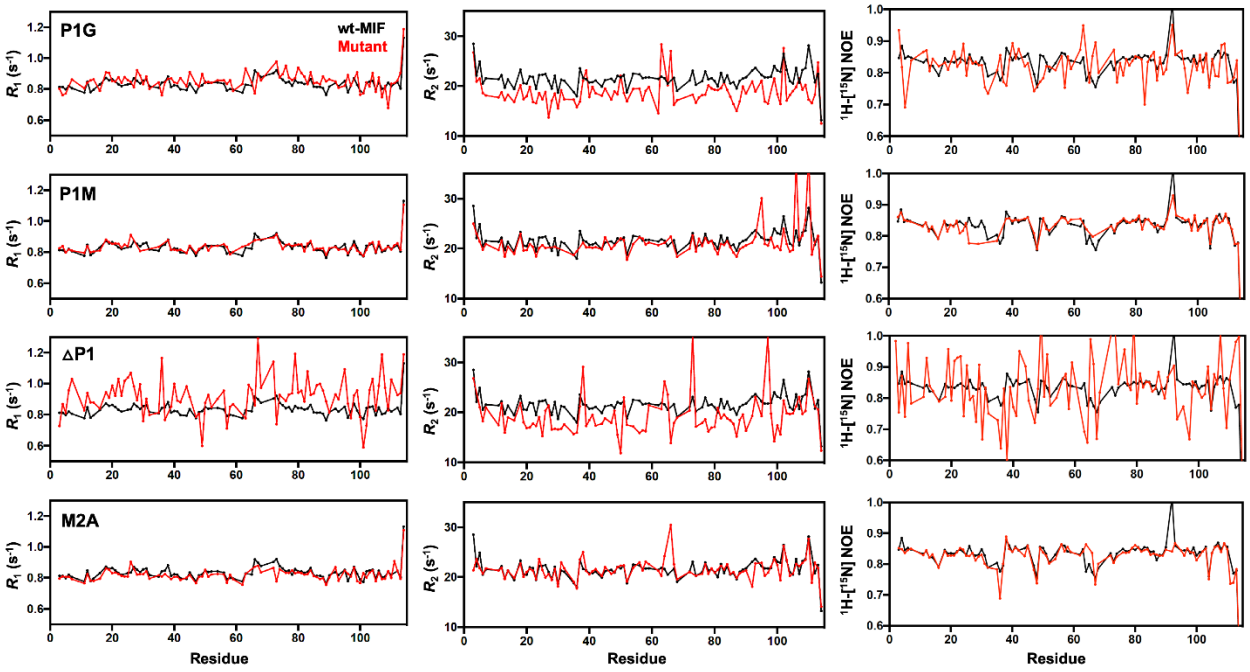
Cross-correlation analysis of C_α atoms for WT MIF and MIF variants.



Note. Correlation analyses of WT MIF and MIF variants were derived from 200 ns MD simulation runs, each one repeated in triplicate. The two axes demonstrate the amino acid number of MIF trimer. Correlation between the MIF residues increases from 0 (no correlation) to 1(absolute correlation). Boxes A, B and C highlight the three monomer of MIF trimer.

Supplemental Figure 2.S5.

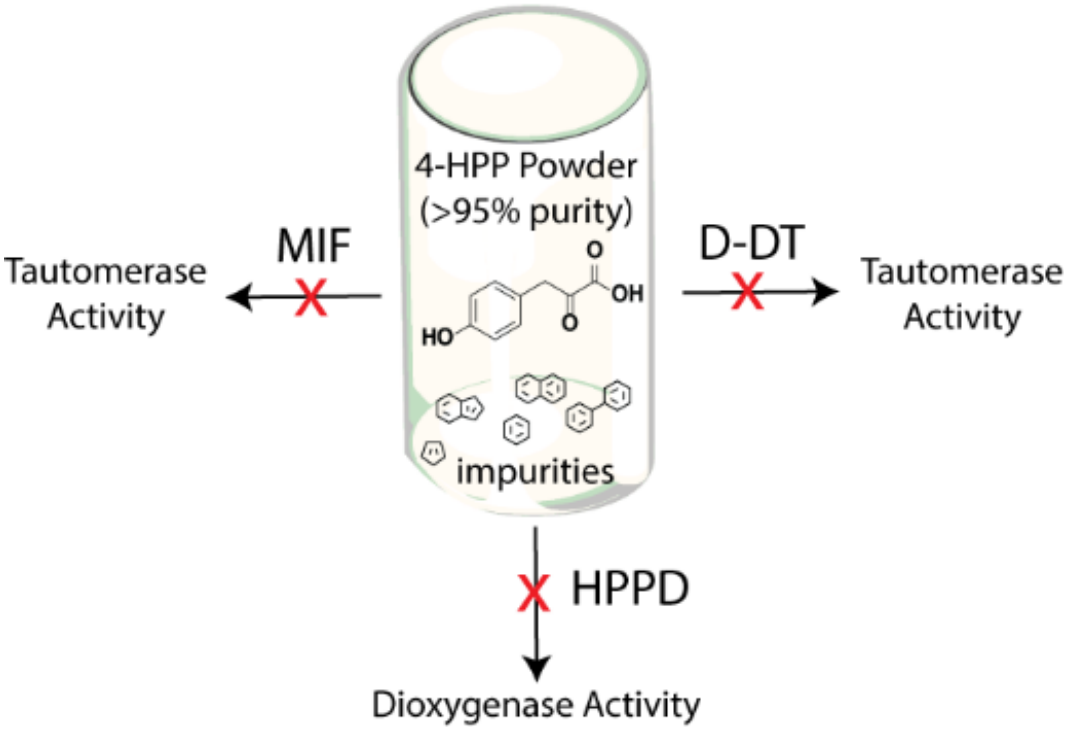
NMR spin relaxation parameters for WT MIF and MIF variants.



Note. The R_1 (left), R_2 (center) and ^{1}H-[^{15}N] NOE (right) values for WT MIF (black) are overlaid with those of P1G (top row), P1M (second row), ΔP1 (third row), and M2A (bottom row). All MIF variants are shown in red. These data comprise the correlation analysis shown in Figure 2.4.

CHAPTER 3: UNDERREPRESENTED IMPURITIES IN 4-HYDROXYPHENYLPYRUVATE AFFECT THE CATALYTIC ACTIVITY OF MULTIPLE ENZYMES

Published as: Parkins, Andrew et al. “Underrepresented Impurities in 4-Hydroxyphenylpyruvate Affect the Catalytic Activity of Multiple Enzymes.” Analytical chemistry vol. 95,11 (2023): 4957-4965. doi:10.1021/acs.analchem.2c04969



Abstract

Macrophage migration inhibitory factor (MIF) is a key immunostimulatory protein with regulatory properties in several disorders, including inflammation and cancer. All the reported inhibitors that target the biological activities of MIF have been discovered by testing against its keto/enol tautomerase activity. While the natural substrate is still unknown, model MIF substrates

are used for kinetic experiments. The most extensively used model substrate is 4-hydroxyphenyl pyruvate (4-HPP), a naturally occurring intermediate of tyrosine metabolism. Here, we examine the impact of 4-HPP impurities in the precise and reproducible determination of MIF kinetic data. To provide unbiased evaluation, we utilized 4-HPP powders from five different manufacturers. Biochemical and biophysical analyses showed that the enzymatic activity of MIF is highly influenced by underrepresented impurities found in 4-HPP. Besides providing inconsistent turnover results, the 4-HPP impurities also influence the accurate calculation of ISO-1's inhibition constant, an MIF inhibitor that is broadly used for in vitro and in vivo studies. The macromolecular NMR data show that 4-HPP samples from different manufacturers result in differential chemical shift perturbations of amino acids in MIF's active site. Our MIF-based conclusions were independently evaluated and confirmed by 4-hydroxyphenylpyruvate dioxygenase (HPPD) and D-dopachrome tautomerase (D-DT); two additional enzymes that utilize 4-HPP as a substrate. Collectively, these results explain inconsistencies in previously reported inhibition values, highlight the effect of impurities on the accurate determination of kinetic parameters, and serve as a tool for designing error-free in vitro and in vivo experiments.

Introduction

Commercially available compounds with a typical purity of ≥95 % are widely used in enzymatic assays as substrates or inhibitors. The remaining <5% consists of organic and/or inorganic impurities that are associated with the chemical synthesis process and may include precursor molecules, chemical reagents, metal ions, or reaction byproducts. Such chemical substances are occasionally reported to interfere with the accuracy of enzymatic assays, producing false positives and inconsistent kinetic data[43, 78, 79]. In principle, the percent of impurities found in these compounds is diminutive and should not hinder experimental set-up nor results.

Nevertheless, a typical enzymatic assay is performed at reagent and substrate concentrations that exceed the enzyme concentration by far. Under these conditions, chemical impurities may reach a concentration that actively interferes with the enzyme.

MIF is a versatile protein with a well-described functionality in adaptive and innate immunity[80]. The homotrimeric assembly of MIF activates the CD74/CD44 co-receptor pair[27, 36] and G-protein-coupled receptors (GPCRs) CXCR2 and CXCR4[12, 81], leading to signaling cascades that promote cell proliferation, migration, and inhibition of apoptosis. The biological pathways activated by the MIF/receptor complexes have regulatory effects on cardiovascular, autoimmune diseases, inflammation, and cancer[82-84]. Besides functioning as a cytokine and chemokine, MIF also possesses an endonuclease activity[13] and a keto-enol tautomerization activity[85]. The latter serves as the sole source for identification of MIF modulators[48, 58, 86]. Despite its early discovery[85], the keto-enol tautomerization activity of MIF is poorly understood and the natural substrate, is yet to be reported. Instead, several synthetic and naturally occurring molecules serve as model substrates. from those, D-dopachrome[85] and 4-hydroxyphenylpuruvic acid (4-HPP)[6] are the most widely used. Since D-dopachrome has several limitations, including photosensitivity and short linearity, 4-HPP is the preferred substrate for biochemical studies[48] and high-throughput screening efforts[87].

The keto-enol tautomerization of 4-HPP occurs in the active site formed between each subunit of the MIF trimeric assembly. In this structure, the 3 polypeptide subunits adopt a ring shape with a central opening known as the solvent channel[17]. The N-terminal residue, Pro1, serves as the catalytic base, while Lys32, Tyr36, from one subunit and Tyr95 and Asn97 from an adjacent subunit, stabilize the bound 4-HPP[7]. Mutations on the four active site residues or the catalytic base Pro1, reduce or abolish the catalytic activity of MIF, respectively. The hydrophobic active site is

known to be promiscuous as it binds to a diverse group of small molecules, some of which covalently modify Pro1[48, 58]. Besides small molecule inhibitors used in binding studies, crystallographic analysis has shown that the active site pocket of MIF also binds other reagents present in the experimental set-up that were not intended to bind (see PDB:3DJH)[46]. Recently, a regulatory site located at the opening of the solvent channel was reported to dynamically modulate functional regions of MIF, including surface amino acids responsible for the activation of CD74[51, 88] as well as the catalytic base Pro1[53].

Besides MIF, 4-HPP is a substrate for D-DT and HPPD, two enzymes with an emerging functional role in tumor cell survival and proliferation [44, 89]. D-DT is a homotrimer with a keto-enol tautomerization activity that is regulated by the N-terminal Pro, similar to MIF [90]. Although they possess the same enzymatic activity, MIF and D-DT share only 34% sequence identity [9]. The low amino acid identity has a direct impact on the composition, hydrophobicity, and ligand selectivity of each active site [57]. Thus far, D-DT has only two reversible inhibitors both of which were discovered via the tautomerization assay of 4-HPP [44, 59]. In contrast to MIF and D-DT, HPPD is a homodimeric enzyme that catalyzes dioxygenation of 4-HPP in the second step of tyrosine catabolism. Incorporation of molecular oxygen in 4-HPP yields homogentisate (HGA) and the reaction requires inorganic iron as a cofactor.

In this study, we use MIF as a model enzyme to examine how 4-HPP impurities may affect enzymatic activity. Our experimental approach involved an array of complementary techniques, and all together demonstrated that underrepresented impurities found in 4-HPP affect the enzymatic activity of this protein. D-DT and HPPD were employed to evaluate the MIF kinetic findings. In both cases, we noted that 4-HPP samples from different manufacturers produced inconsistent kinetic data. Altogether, our findings demonstrate the critical impact of 4-HPP

impurities on the accuracy of kinetic data and support the efforts for designing error-free biochemical assays and high-throughput experiments.

Materials and Methods

Materials

Information about the purity and suppliers of the reagents used is provided in the supplemental information.

Protein cloning, expression, and purification. Wild-type (WT) MIF, WT D-DT and WT HPPD were cloned in pET-11b, pET-22b and pET-21a vectors, respectively, and transformed into Escherichia coli BL21 gold (DE3) competent cells (Agilent Technologies, Inc.) for expression. Protein expression and purification were carried out according to previously described protocols [66,91]. More information is provided in supporting experimental section.

MIF and D-DT enzymatic assays. The kinetic experiments for MIF and D-DT were performed in a 96-well microplate, as previously described [51, 53]. A detailed protocol is described in supporting experimental section.

HPPD endpoint assay.

The HPPD endpoint assay was carried out in a 96-well microplate using 4-HPP samples from different manufacturers (TCI and Chemodex). First, a 30 mM stock solution of 4-HPP was prepared in 0.5 M ammonium acetate pH 6.2, similarly to what we described earlier for the MIF and D-DT assays. To prepare the 150 µL reaction solution, iron (II) sulfate, ascorbate, and 4-HPP were added at a final concentration of 1 mM, 4 mM, and 0-1.6 mM, respectively. The mixture, which also included 110 µL of 0.5 M ammonium acetate pH 6.2, was incubated at 30°C for 30 minutes and the reaction initiated by the addition of 10 µL HPPD, at a final concentration of 500 nM. After 5 min, the reaction was terminated by adding 150 µL MeOH. The microplate was

centrifuged at 2100 × g for 15 minutes at 4°C. A 50 µL sample from each well was transferred to a new 96-microplate that contained borate at a final concentration of 0.42 M. Borate was prepared in 100 % MeOH. The formation of enol-HPP-borate complex was monitored at 306 nm using a Tecan Infinite M-Plex microplate reader for 16 hours, taking readings every 30 min. Between reading the plate was sealed with an aluminum film to avoid evaporation of MeOH. The assay was performed in triplicate.

NMR and Mass Spectrometry analyses of the 4-HPP samples.

All NMR and MS spectra were correspondingly recorded on a JEOL-ECA600 instrument and triple quadrupole mass spectrometer interfaced with an electrospray ionization (ESI) source. Both analyses were performed following standard methods that are described in supporting experimental section.

Macromolecular NMR binding studies.

NMR experiments to study the interaction between 4-HPP samples from two different manufacturers (Chemodex and TCI) and MIF were acquired on a Bruker Avance III spectrometer operating at a 1H frequency of 600 MHz and equipped with a z-axis gradient cryogenic probe. NMR spectra for chemical shift assignment and for binding studies were collected at 303 K and 298 K, respectively. MIF forms a trimer of ~ 37.5 kDa, thus both 2D-[1H-15N]-TROSY[92] and 2D-[1H-15N]- HSQC were acquired; however, no significant improvement in the quality of the spectra was observed by using the TROSY sequence. Fast acquisition techniques, [1H-15N] SOFAST-HMQC[93], were used to achieve high 4-HPP/MIF ratios by keeping MIF concentration at low sub-millimolar concentration (50 µM). Two different buffers at pH values of 7.0 and 6.2 were used to test the effects of the impurities in 4-HPP samples. Buffer 1 is composed of 20 mM phosphate (pH 7.0), 1 mM EDTA. Buffer 2 contains 20 mM sodium acetate (pH 6.2). Both buffers contain 7.5%

D2O. Three samples were prepared in each buffer: a control sample without 4-HPP, 1:200 molar ratio of MIF:4-HPP (Chemodex), and 1:200 molar ratio of MIF:4-HPP (TCI). The concentration of 15N-uniformly labeled MIF stocks was determined to be 1.7 mM by ultraviolet–visible (UV-Vis) measurement at 280 nm and using a theoretical molar extinction coefficient of 13,075 cm-1 M-1 based on the amino acid sequence. The stock solutions were diluted to 50 μM for all experiments in the pertinent buffers. 4-HPP from Chemodex and TCI were prepared at a concentration of 600 mM and diluted to 10 mM for all experiments. Samples were prepared in advance and incubated overnight. Data were processed with NMRPipe[74] software and analyzed with SPARKY[75].

Results and Discussion

4-HPP samples from different manufacturers exhibit inconsistent kinetic results.

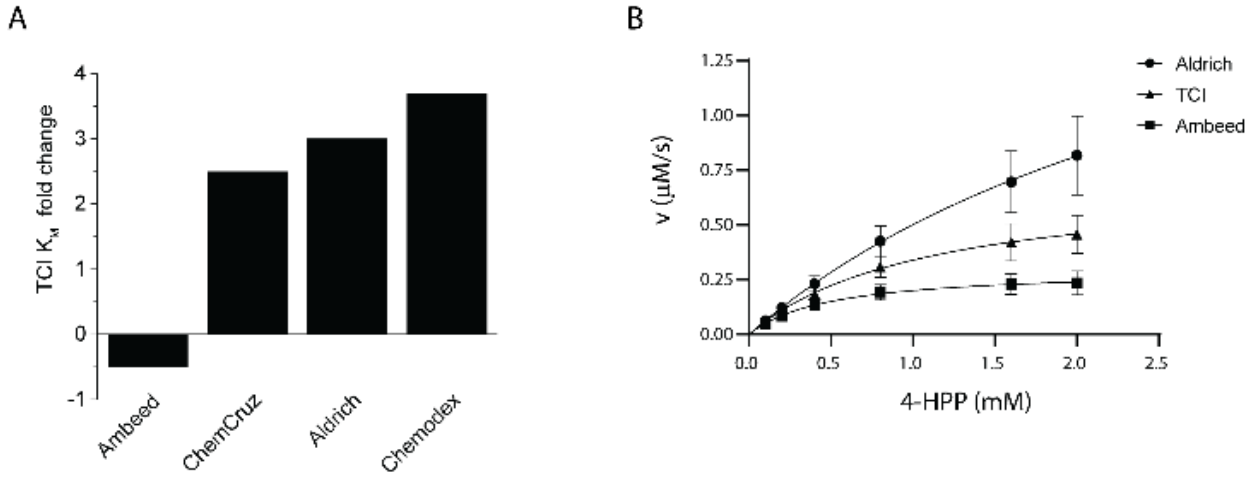
Previously published kinetic findings have demonstrated a deviation of ISO-1 inhibition constant (Ki) [21, 94, 95]; a molecule that serves as the reference inhibitor for many MIF biological and drug discovery studies.[96, 97] Inconsistent MIF kinetic findings are not restricted only to ISO-1 and it is believed that they are caused by substrate-related issues, such as the i) slow rate of 4-HPP tautomerization, ii) spectral interference between inhibitors and product (enol-HPP/borate) and iii) short linearity of product formation.[97]

To investigate the impact of 4-HPP in the irregularity of MIF enzymatic assay, we purchased the compound from five different manufacturers (TCI, Ambeed, ChemCruz, Aldrich, and Chemodex) and performed kinetic assays, having 4-HPP as the only variable. The stated purity on the 4-HPP vials was between >95% (TCI) to ≥98% (Chemodex). To ensure accuracy of our findings, we prepared the 4-HPP samples under identical conditions and performed the biological replicates in three different days. Kinetic analyses revealed a noticeable disagreement between the

produced Michaelis-Menten plots and corresponding parameters (Table 1, Fig. S1). Surprisingly, the Michaelis-Menten constant (KM) values varied between 0.5 ± 0.05 mM and 4.1 ± 1.0 mM for Ambeed and Chemodex, respectively. The 8-fold difference between these samples, indicates that the 4-HPP impurities and/or precursor molecules interfered with the assay and the formation of MIF/keto-HPP complex. The KM value of 1.1 ± 0.2 mM, obtained by the TCI sample, was found in agreement with several previous published studies [57, 94, 98]. Using this sample as reference, we compared the KM fold change between the various 4-HPP manufacturers. ChemCruz, Aldrich, and Chemodex exhibited the largest KM disagreement from the corresponding TCI value with 2.5, 3.0 and 3.7-fold change, respectively (Fig. 1A). We also examined how the MIF catalytic performance (kcat/KM) would be affected by the five 4-HPP samples. The highest (13.1 ± 1.4 s-1 mM-1) and lowest (10.1 ± 1.4 s-1 mM-1) kcat/KM values were obtained by TCI and Chemodex, respectively (Table 1). Significant deviation of their kinetic parameters was also noted between the TCI-Aldrich (same kcat/KM) and TCI-Ambeed (similar KM) sample pairs (Fig. 1B).

Figure 3.1.

Comparison of wild-type MIF kinetic data using 4-HPP samples from different vendors.



(Figure 3.1 Continued)

Note. A) The Michaelis-Menten constant (KM) of the most effective 4-HPP (TCI), was compared with the corresponding samples from Ambeed, ChemCruz, Aldrich, and Chemodex. The findings demonstrate a clear disagreement between the KM values that could be explained considering the inhibition effect of impurities found in each sample. B) Impurities found in the 4-HPP samples, obtained from Aldrich, TCI and Ambeed, have a drastic effect on the Michaelis-Menten kinetic parameters of wild-type MIF.

Table 3.1.
Michaelis-Menten parameters of wild-type MIF determined from keto-enol tautomerase assays of 4-HPP*

	K_M (mM)	V_{max} (μM s^{-1})	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1} mM^{-1})
Chemodex	**4.1 ± 1.0**	**2.1 ± 0.7**	**41.6 ± 14.7**	**10.1 ± 1.4**
ChemCruz	**2.7 ± 0.3**	**1.6 ± 0.3**	**31.1 ± 5.9**	**11.6 ± 1.9**
Ambeed	**0.5 ± 0.05**	**0.3 ± 0.07**	**6.0 ± 1.4**	**12.2 ± 1.6**
Aldrich	**3.3 ± 1.0**	**2.2 ± 0.8**	**44.6 ± 16.2**	**13.1 ± 1.5**
TCI	**1.1 ± 0.2**	**0.7 ± 0.2**	**14.3 ± 3.5**	**13.1 ± 1.4**

* The corresponding Michaelis-Menten plots are shown in Figure S1. The experiments were carried out in triplicate (n=3) and the error values are shown as standard deviations.

The 4-HPP composition affects the accurate calculation of inhibition constant.

It was previously documented that impurities found in commercially available compounds, cause a variation of KM and impact the accuracy of the resulting enzymatic data.[43, 79] for example, impurities in the commercially available 2,2′-azino-bis(3-ethylbenzothiazoline-6-sulfonate),

which serves as a substrate for peroxidase enzymes, were shown to inhibit the activity of horseradish peroxidase.

In the case of MIF, we were interested in exploring how potential impurities/precursor molecules found in 4-HPP may affect the accurate calculation of inhibition constant. We determined the Ki value of ISO-1, using 4-HPP from three different manufacturers: TCI - KM =1.1 ± 0.2 mM, ChemCruz - KM =2.7 ± 0.3 mM, and Chemodex - KM =4.1 ± 1.0 mM. The rationale for choosing the specific manufacturers was based on the idea that their differences in KM, would have an impact on the calculated Ki value of ISO-1. In general, a smaller KM value is indicative of an enzymatic assay that is less influenced by potential impurities and thus a smaller Ki is expected. TCI and Chemodex would be expected to result in the smallest and largest Ki values, respectively, while the ChemCruz value should be somewhere in between. All the experiments were carried out using the same ISO-1 sample (≥98% purity). Analogously to our turnover experiments, the only variable for the inhibition experiments was the origin of 4-HPP.

The experiments were repeated four times and the Michaelis-Menten plots were utilized to determine the inhibition constant of ISO-1. The average Ki values of 22.9 ± 5.6 μM, 44.4 ± 9.0 μM, and 64.4 ± 25.5 μM were obtained for TCI, ChemCruz, and Chemodex, respectively (Fig. S2, Table S1). As expected for ISO-1, the Lineweaver-Burk plots demonstrate a competitive-type of inhibition. A closer examination of the four inhibition trials, shows a high reproducibility rate for TCI and ChemCruz samples, while the Chemodex sample caused fluctuation of the Ki (Table S1). This fluctuation is likely to be caused by the Chemodex impurities, which serve as inhibitors of the MIF assay. Given that the commercially available 4-HPP samples have a minimum reported purity of 95% and the final concentration of 4-HPP in the enzymatic assay is up to 2 mM,

impurities can reach a final concentration of ≤100 µM. At this concentration range, underrepresented impurities can effectively inhibit MIF.

Noticeably, the inhibition findings are in agreement with the KM values obtained for TCI, ChemCruz, and Chemodex 4-HPP samples (Table 1). The previously published Ki values for ISO-1 are 24 µM,[21, 97] 44 µM,[94] and 120 µM.[95] Apart from being the most reproducible one, the Ki value of 24 µM is also in agreement with two previously reported binding constant (KD) values of ISO-1 obtained using fluorescence spectroscopy.[21, 97] Two out of three Ki values shown in this study (TCI - Ki=22.9 ± 5.6 µM and ChemCruz - Ki=44.4 ± 9.0 µM) match the previously published data, whereas the third (Chemodex - Ki=64.4 ± 25.5 µM) deviates noticeably from all known reported values. At first glance, the ~2-fold difference of TCI/ChemCruz and ~3-fold difference of TCI/Chemodex sample pairs, might not appear significant. However, these differences become critical in Ki calculations of highly potent MIF inhibitors and can lead to false positives in high-throughput screening efforts for identification of new drug-like agents.

NMR analysis of the 4-HPP samples indicates the presence of different impurities.

Our turnover and inhibition assays clearly indicate that the KM and Ki values are influenced by the impurities found in 4-HPP (Table 1, Table S1). To accurately assess 4-HPP samples' purity, we dissolved the five powders in deuterated dimethyl sulfoxide (DMSO-d6) and recorded their proton spectra. Prior studies have shown that the keto tautomer of 4-HPP is far more stable in aqueous solution,[99] while in DMSO the equilibrium favors formation of the enol-form by ~96%.[100] In agreement with these studies, our NMR findings illustrate the enol-HPP predominance in the equilibrium and it's detected at 12.9 ppm (Fig. S3).

Overlay of the five proton spectra disclose differences in purity that are well-aligned with our biochemical findings (Fig. S3, Fig. S4). Side-by-side comparison of the five spectra exhibits

that the TCI and Ambeed samples are the cleanest. Impurities around 1.9, 6.5, 7.7, and 9.8 ppm were detected only in Chemodex, Aldrich and ChemCruz samples and may account for the higher KM values observed in these three cases (Table 1, Fig. 1A). Several additional impurities at low abundance were detected in particular samples without providing a clear direction of their potential role in MIF enzymatic activity.

Unique 4-HPP contaminations are detected by mass spectrometry.

Visual comparison of the five 4-HPP powders revealed color differences that align with the kinetic findings (Fig. S5). The color of TCI powder, which accurately reproduces the published KM and Ki values for 4-HPP[57] and ISO-1,[21] respectively, noticeably differs from the colors of the remaining four. Although such observation cannot serve as a quantitative tool to evaluate the samples' purity or explain the enzymatic findings, it is indicative of differences associated with the synthetic pathway(s). Probing the chemical pathway(s) of 4-HPP synthesis, we identified three potential routes (Scheme S1). Pathways 1[101] and 2[102] share 4-hydroxybenzaldehyde (5) as a key intermediate and reach the final product in a fewer steps in comparison to pathway 3.[103]

We employed mass spectrometry to analyze the composition of TCI and Chemodex samples and detect ionizable precursor molecules or other contaminations that interfere with the activity of MIF. These two samples were chosen anticipating they would provide the clearest experimental outcome. The positive and negative ion modes were monitored in a mass-to-charge (m/z) ratio up to 600 Da and the TCI and Chemodex data were compared side-by-side. Analysis of the protonated adduct molecules demonstrated that the Chemodex sample is clearly more contaminated than the corresponding TCI one (Fig. S6). The peak at m/z 122.8 corresponds to 5 and it was detected in both TCI and Chemodex samples, as a consequence of a common manufacturing strategy. Elimination of hydroxyl group from the para position of 5 yields

benzaldehyde, which was previously reported to covalently interact with the catalytic residue Pro1 (PDB: 4PLU)[48]. Noteworthy, the Pro1-benzaldehyde chemical adduct was derived from the nucleophilic substitution reaction between MIF-Pro1 and a suicide inhibitor. We examined the inhibition potency of 5 at a single point concentration of 100 μM and found no inhibition activity against MIF (data not shown).

Besides the noticeable difference in purity, the positive ion mode analysis of Chemodex sample also revealed a unique contamination peak at m/z 85.9. This peak is not found in the TCI sample. Another unique contamination peak associated with the Chemodex sample was detected by the negative ion mode analysis at m/z 276.9 (Fig. S7). Neither of the two peaks was matched with the molecular weight (MW) of a precursor molecule (Scheme S1 - Pathways 1 or 2), while additional efforts to determine the identity of these impurities by high resolution accurate mass measurement (HRMS) did not yield any further insights. As expected, the main product (4-HPP) was detected in both positive and negative ion modes at m/z 181.0 and 178.9, respectively. Collectively, the MS findings further support the kinetic and NMR data and demonstrate that the impurities possessing inhibitory properties against MIF are not precursor molecules of 4-HPP (Scheme S1).

Inconsistencies between the 4-HPP samples are not lot/batch number specific. Comparison between the TCI and Chemodex samples may raise an argument that the kinetic, NMR, and MS findings are due to batch inconsistencies. To elucidate this point, two TCI and Chemodex samples with the same product number (H0294 - TCI and H0013 - Chemodex) but different lot numbers (TCI lot numbers: UJNCJ-SA and 6TM4L-YH and Chemodex lot numbers: P4006 and X8048) were evaluated side-by-side (Table 2, Fig. S8). The stated purity on the vials of TCI and Chemodex samples were >95% and ≥98%, respectively. The CAS number of 156-39-8, which was

shown on all vials, confirms the identity of 4-HPP. for the TCI samples, the findings were highly reproducible yielding the Km/Vmax values of 1.1 ± 0.2 mM/0.7 ± 0.2 μM s-1 (UJNCJ-SA) and 1.0 ± 0.0 mM/0.8 ± 0.1 μM s-1 (6TM4L-YH). The Km/Vmax values of 4.1 ± 1.0 mM/2.1 ± 0.7 μM s-1 (P4006) and 3.9 ± 0.1 mM/2.9 ± 0.1 μM s-1 (P4006) obtained for the two Chemodex samples were also reproducible confirming that the MIF conclusions shown above are not bias by any batch inconsistencies.

Table 3.2.
*Impact of lot number on the reproducibility of MIF kinetic results**

	Lot number	K_M (mM)	V_{max} (mM s^{-1})
Chemodex	P4006	4.1 ± 1.0	2.1 ± 0.7
	X8048	3.9 ± 0.1	2.9 ± 0.1
TCI	UJNCJ-SA	1.1 ± 0.2	0.7 ± 0.2
	6TM4L-YH	1.0 ± 0.0	0.8 ± 0.1

* The corresponding Michaelis-Menten plots are shown in Figure S8. The experiments were carried out in triplicate (n=3) and the error values are shown as standard deviations.

The impurities affecting the catalytic activity of MIF are not transition metal ions.

It was previously shown that transition metal ions, including zinc (II) or copper (II), inhibit the enzymatic activity of MIF with a half-maximal inhibitory concentration (IC50) of 1 μM [94]. These ions act as Lewis acids and are derived from the synthetic pathways followed for MIF inhibitors. Upon this finding, we probed whether transition metal ions may be responsible for the irregular kinetic results of Chemodex samples. Although, the precise synthetic pathway followed in this case is unknown, our MS analysis identified 5 as a key intermediate (Fig. S6). This result

suggests that 4-HPP was likely synthesized following either pathway 1 or 2 (Scheme S1). Literature review of the reactions involved in these two pathways did not demonstrate the usage of transition metal ions, as either catalysts or reagents. Nevertheless, we investigated it further by performing additional kinetic assays in the presence of 1 mM EDTA. for these experiments, the TCI 4-HPP (lot number: 6TM4L-YH) was used as a negative control to show that the chelation properties of EDTA have no effect of the produced kinetic data. Indeed, the Km and Vmax values of 1.2 ± 0.02 mM and 0.9 ± 0.01 μM s^{-1}, obtained in the presence of 1 mM EDTA, demonstrate no deviation from the previous kinetic findings (Table 3, Fig. S9). Upon testing the impact of 1mM EDTA on the kinetic results of Chemodex samples (Lots numbers: P4006 and X8048), we noted a minor deviation from the original values (no EDTA) that were within the statistical error (Table 3). Therefore, we concluded that EDTA had not impact on the assays involved the Chemodex samples and transition metal ions are not responsible for the irregular kinetic results.

Table 3.3.
*Impact of metal ions on the reproducibility of MIF kinetic results**

		(-) EDTA		**(+) EDTA**	
	Lot #	**K_M (mM)**	**V_{max} (μM s^{-1})**	**K_M (mM)**	**V_{max} (μM s^{-1})**
Chemodex	P4006	4.1 ± 1.0	2.1 ± 0.7	3.4 ± 0.3	2.7 ± 0.3
	X8048	3.9 ± 0.1	2.9 ± 0.1	4.5 ± 0.8	3.2 ± 0.6
TCI	6TM4L-YH	1.0 ± 0.0	0.8 ± 0.1	1.2 ± 0.02	0.9 ± 0.01

* The corresponding Michaelis-Menten plots are shown in Figure S9. The experiments were carried out in triplicate (n=3) and the error values are shown as standard deviations.

NMR binding studies between MIF and 4-HPP samples show different chemical shift perturbation patterns.

We have used protein NMR to gain a better understanding of the binding between MIF and 4-HPP samples from different manufacturers. To ensure that the impurities are present in significant amounts, we have prepared mixtures of 4-HPP (from Chemodex and TCI manufacturers) at 10 mM concentration and uniformly labeled 15N-MIF at 50 μM, thus resulting in a substrate/enzyme molar ratio of 200 and potential presence of impurities at ~ 500 μM. The keto-enol tautomerization rate of phenyl pyruvic acid and its derivatives is significantly enhanced (more than one order of magnitude) by pH increase in the 6.0 to 8.5 range in the absence of enzymatic catalysis[104]. However, rate enhancement by pH is only marginal in the presence of the enzyme[104]. Nonetheless, the NMR binding studies were done at pH 6.2 and 7.0 to study potential effects of the pH on the interaction between MIF and 4-HPP.

MIF assembles into a stable trimer[17] with a MW of ~37.5 kDa that might result in NMR signal broadening due to slow molecular rotational diffusion. Therefore, we tested the quality of spectra acquired with standard 2D-[1H-15N]- HSQC and the transverse relaxation-optimized sequence 2D-[1H-15N]-TROSY[92]. Both spectra show good signal-to-noise ratio. The 2D-[1H-15N]- HSQC experiment was finally used to obtain backbone amide 1H-15N chemical shift assignments (Fig. S10) using as reference the originally published data[105]. The assignments at both pH values, 6.2 and 7.0, are shown in Tables S2-S4. Nonetheless, NMR is a low-sensitivity technique, thus we used fast acquisition techniques (2D-[1H-15N]-SOFAST-HMQC)[93] to acquire sufficient NMR signal accumulation at the micromolar concentration needed for the binding studies. The 2D NMR experiments reveal clear chemical shift perturbations, and in some instances signal broadening, reflecting the interaction between MIF and the 4-HPP samples from Chemodex

and TCI. Chemical shift perturbations were obtained from the average change combining both 1H and 15N chemical shifts according to **equation 1**.

$$\Delta\delta^{av} = \left[(\delta_{1H\_MIF} - \delta_{1H\_MIF\_HPP})^2 + \left(\frac{\delta_{15N\_MIF} - \delta_{15N\_MIF\_HPP}}{5}\right)^2\right]^{1/2} \quad (\mathbf{1})$$

Where, δ_{1H_MIF} and δ_{15N_MIF} are the amide 1H and 15N chemical shifts, respectively, in the absence of 4-HPP; and $\delta 1_{H\_MIF\_HPP}$ and $\delta_{15N\_MIF\_HPP}$ are the amide 1H and 15N chemical shifts, respectively, in the presence of 4-HPP.

All chemical shift perturbations larger than one standard deviation were carefully inspected in the different spectra and assigned to the pertinent amino acids. A representative subset is shown in Figure S11, highlighting differential chemical shift perturbation or signal broadening. Upon binding to 4-HPP, some signals show significant changes in chemical shifts, for example A27 and T30, whereas others undergo severe broadening, such as the -NH of W108 indole side chain, I64 and K66 (Fig. S11). Interestingly, the extent of signal broadening and/or chemical shift perturbation differ for the same amino acids depending on whether the spectrum was acquired for the sample containing 4-HPP from TCI or from Chemodex (Table 4). In general, signal broadening upon binding to 4-HPP is more prominent at pH 6.2 compared to pH 7.0 (Table 4, Fig. S11). This result indicates that MIF-4-HPP interactions are affected by the pH-dependent keto-enol tautomerization. It is worth noting that the chemical shift perturbations are larger in the presence of the Chemodex 4-HPP compared to TCI for many amino acids; albeit the opposite is observed for several residues such as S63, K66 and I96 at pH 6.2 (Table 4). Therefore, we can conclude that there is no systematic deviation in the TCI or Chemodex 4-HPP binding data, which rules out sample preparation as a potential source of differential chemical shift perturbation.

Table 3.4.

Differential binding between MIF and 4-HPP from TCI and Chemodex based on NMR chemical shift perturbations

Amino Acid	Chemical shift perturbation (ppm) or broadness*			
	pH 7.0		pH 6.2	
	4-HPP Chemodex	4-HPP TCI	4-HPP Chemodex	4-HPP TCI
F3	0.074	0.039	**	**
L26	0.039	0.007	**	**
A27	0.022	0.009	0.058	0.055
T30	0.053	0.022	0.113	0.100
Q35	0.021	0.006	0.047	0.027
I37	0.052	0.023	**	**
V39	0.046	0.018	0.064	0.068
F49	0.038	0.014	**	**
G50	0.032 (broader)	0.028	**	**
S63	0.029	0.011	0.033	0.053
I64	0.041	0.027 (broader)	0.124	** (broader)
K66	0.065	0.046	0.063	0.1
I96	0.033	0.016	0.023	0.036
N102	0.028	0.023	0.046	0.052 (broader)
V106	0.025	0.024	0.025	0.026 (broader)
W108 side chain	0.048	0.024	0.058	0.061 (broader)
T112	0.027	0.012	0.014	* (broader)

* All residues have chemical shift perturbations larger than 1SD in at least one condition except T112, which is shown as an example of differential broadening. The largest differences in chemical shift perturbation between 4-HPP from Chemodex and TCI

The amino acids with chemical shift perturbations larger than one standard deviation (1 SD) upon addition of 4-HPP from Chemodex and TCI are located at or near MIF's trimer active sites in the interface between subunits (Fig. S12A, S12B). These amino acids match previously reported NMR data on the binding region of MIF including F3, V39, G50, K66, N102, W108, F113, A114 [18]. Few variations in the amino acids involved in the interaction are observed when comparing data for 4-HPP from Chemodex (Fig. S12A) and TCI (Fig. S12B). This result is likely due to the threshold of 1SD used for selecting the most significant perturbations, as binding to the different 4-HPP types clearly happens in the same interface according to our NMR data (Fig. S12A, S12B). Mapping onto the MIF structure amino acids with perturbations that differ depending on the 4-HPP manufacturer (Table 4) identified the structural motif involved in differential binding (Fig. S12C). Amino acids in the active site and nearby are perturbed differently depending on the origin of the 4-HPP. This result indicates that impurities also occupy the active site and cause different perturbations in the NMR signals depending on the type of 4-HPP. Altogether, the NMR data show clearly that MIF interacts differently with the two 4-HPP samples. This behavior could only be explained by the presence of impurities or other unwanted compounds, which modulate the interaction between MIF and 4-HPP via yet unknown mechanisms.

4-HPP impurities affect the catalytic activity of multiple enzymes.

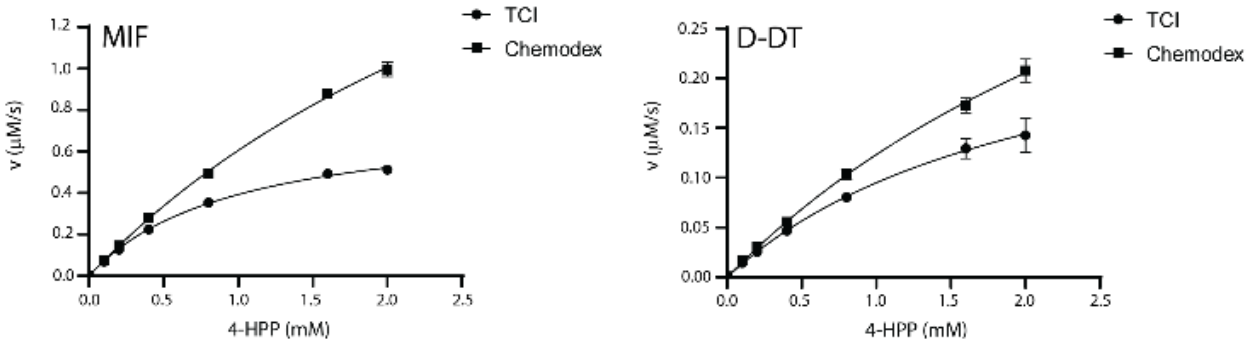
To obtain further insights into the functional role of 4-HPP impurities, we extended our studies on D-DT and HPPD; two enzymes that catalyze tautomerization and dioxygenation of 4-HPP, respectively. We anticipated that the D-DT findings would serve as a secondary evaluation point of the MIF results, since the two proteins share the same activity. Given HPPD is not a member of the tautomerase superfamily, our findings on this enzyme would provide conclusions with broader applications. Focusing first on D-DT, we performed kinetic assays using the TCI and

Chemodex samples, which were previously shown to have noticeable kinetic differences (Table 1). Our findings are in agreement with the MIF results and show that 4-HPP samples from different manufacturers yield diverse kinetic results that affect the accuracy and reproducibility of the Michaelis-Menten parameters (Fig. 2, Table S5). Notably, the Km values derived from the Chemodex and TCI samples, demonstrate a ~2-fold difference. Previous studies have shown that D-DT is less active that MIF [9, 44]. This conclusion is also confirmed herein and explains why 4-HPP impurities have significant but less drastic effect on D-DT (Fig. 2).

To analyze the effect of 4-HPP impurities on the catalytic activity of HPPD, we first had to overcome a major obstacle related to the assay itself. While the product (HGA) and substrate (4-HPP) cannot be differentiated by direct spectrophotometric measurements, monitoring of HPPD activity is currently accomplished via indirect or labor-intensive methods. One of them is a coupled assay, which employs homogentisate 1,2-dioxygenase (HGD) to convert HGA to maleylacetoacetate. The latter is quantified by measuring absorbance at 318 nm [106]. Other assays record the absorbance of high-performance liquid chromatography (HPLC)-isolated HGA, consumption of oxygen using an oxygen electrode [107] or formation of 14CO2 resulting from decarboxylation of the 14C-labeled substrate [108].

Figure 3.2.

Side-by-side comparison of MIF (left) and D-DT (right) Michaelis-Menten plots using 4-HPP from different manufacturers.



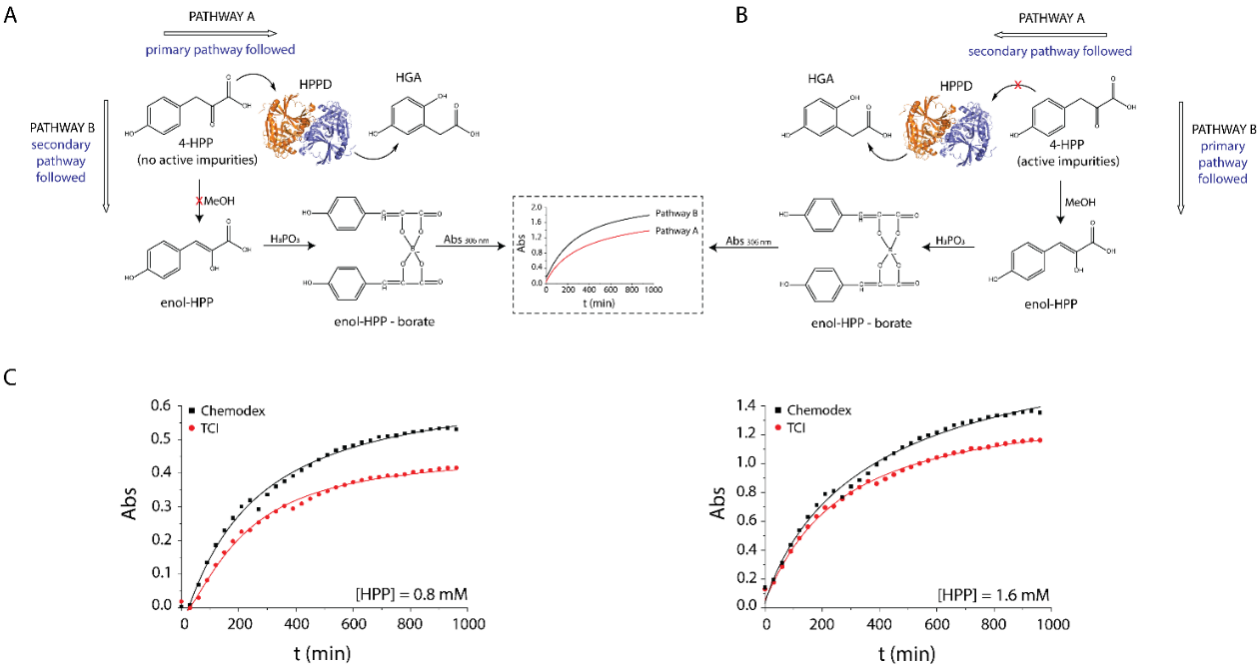
Note. The 4-HPP samples obtained from different manufacturers demonstrate distinguishable kinetic outcomes that have a direct impact on the accurate and reproducible calculation of Michaelis-Menten parameters.

for this study, we developed an endpoint assay that measures the relative concentration of 4-HPP left in the reaction mixture after inactivation of HPPD. Enzymatic dioxygenation of 4-HPP occurs for 5 min and the reaction is terminated by the addition of MeOH. Besides inactivation of HPPD, MeOH also converts keto-HPP to enol-HPP. MeOH was chosen among several organic solvents due to its ability to shifts the equilibrium point towards enol-HPP. Under this condition, the equilibrium percentage for enol-HPP is 100% and the tautomer remains stable for long periods (up to 1 month) [99, 100]. Enol-HPP is then chemically converted to enol-HPP-borate complex, which absorbs at 306 nm. Having 4-HPP as the only variant of this experiment, we anticipated two possible experimental outcomes. If 4-HPP contains impurities with no inhibitory action against HPPD, then pathway A would be mainly followed favoring the formation of HGA (Fig. 3A). On the other hand, if 4-HPP contains impurities with inhibitory action against HPPD, then pathway B would be mainly followed favoring the formation of enol-HPP-borate complex (Fig. 3B).

Therefore, pathway B would yield higher absorbance readings at 306 nm in relation to pathway A. Based on these two possible outcomes, a side-by-side comparison of Chemodex and TCI plots would tell us whether a substrate's impurities affect the enzymatic conversion of 4-HPP to HGA. The experiment ran for 16 hours, and absorbance readings were taken every 30 minutes. for accuracy of our findings, various controls were included. First, we confirmed that the assay monitors the formation of enol-HPP-borate complex by performing the experiment in the absence of borate (Fig. S13A) or 4-HPP (Fig. S13B). In both cases, we noted no significant changes in absorbance. The ability of HPPD to dioxygenate 4-HPP was also questioned. In the absence of enzyme, our experiment yielded higher absorbance readings that are explained by the presence of non-catalytically converted 4-HPP in the solution (Fig. S13C). from this control experiment, we noted that HPPD is more active in the case of TCI sample, while also demonstrating that impurities found in the Chemodex sample inhibit the enzyme.

Figure 3.3.

Impact of 4-HPP impurities on the catalytic activity of HPPD.



Note. (**A**) In the absence of impurities with inhibitory properties, 4-HPP will mainly follow pathway A producing HGA. (**B**) When 4-HPP contains impurities that inhibit the activity of HPPD, the substrate will mainly follow pathway B producing enol-HPP-borate complex. (**C**) Impurities found in 4-HPP inhibit the catalytic activity of HPPD. The curve with higher absorbance (black) is associated with lower HPPD activity, as described by Pathway B. The curve with lower absorbance (red) is associated with higher HPPD activity, as described by Pathway A.

For both TCI and Chemodex samples, early absorbance readings show negligible differences from each other; however, after 1.5 hours differences in the concentrations of remaining 4-HPP begin to be apparent (Fig. 3C). As the readings continue, the difference becomes greater until the complete conversion of keto-HPP to enol-HPP at 16 hours. The Chemodex 4-HPP

sample has a great deal more remaining 4-HPP than the TCI sample, which is apparent by the higher absorbance at 306 nm (Fig. 3C). These findings suggest that there is a difference in the contaminants found in the two 4-HPP samples, which result in the inhibition of HPPD.

Conclusions

In this study, we followed an unbiased approached to provide answers regarding the role of impurities in 4-HPP. We purchased and kinetically interrogated 4-HPP samples from five different manufacturers to observe inconsistent turnover and inhibition results of our model enzyme, MIF. Whereas the only variant in these experiments was the origin of 4-HPP, our findings were associated with the composition of the substrate and more specifically the underrepresented contaminations (<5%) that are produced during chemical synthesis of 4-HPP. Under the experimental design that is usually followed for enzymatic assays, impurities in the range of <5% become significant and can actively interfere with enzymatic activity. The presence of contaminations in 4-HPP, was independently examined and confirmed by NMR and MS. Macromolecular NMR analysis of the kinetically best (TCI) and worst (Chemodex) 4-HPP samples demonstrates that the differential chemical shift perturbations of amino acids in MIF's active site can only be explained by the presence of impurities found in the substrate. The findings on MIF are independently supported by kinetic results of D-DT and HPPD and altogether support the design of faultless *in vitro* and *in vivo* experiments.

Supplemental Information

Supporting Experimental Section

Materials.

The 4-HPP samples were obtained from TCI (lot numbers: UJNCJ-SA and 6TM4L-YH), Ambeed (lot number: A632541-002), ChemCruz (lot number: F1917), Aldrich (lot number:

BCBW9189), and Chemodex (lot numbers: P4006 and X8048) and stored according to the manufacturer recommendations. The reported purities were >95% (TCI), 96% (Ambeed), ≥96% (ChemCruz), 98% Aldrich, and ≥98% (Chemodex). ISO-1 (≥98% purity) was purchased from Millipore-Sigma (lot number: 3548358). 4-Hydroxybenzaldehyde (98% purity) was purchased from Alfa Aesar (lot number: 10222703). Sodium L-ascorbate (>98%) was acquired from TCI. Ethylenediaminetetraacetic acid (EDTA >99%) and methanol (MeOH) with the corresponding purities of >99% and ≥99.8% were obtained from VWR chemicals. Deuterated DMSO (99.9 % *d*) and ^{15}N-ammonium chloride (99%) were purchased from Cambridge Isotope Laboratory. All the buffers were of the highest analytical grade (≥97% purity) and used without further purification.

Protein cloning, expression, and purification.

Cells were grown at 37°C in sterilized Luria Broth (LB) containing 0.100 mg/ml of ampicillin, until optical density at 600nm (OD_{600}) of 0.6-0.8 was reached. Protein expression was induced by the addition of 1mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Induction of MIF and D-DT occurred at 37°C for 4 hours, while in the case of HPPD, the culture was maintained at 16°C for 16 hours. Following, the cells were collected by centrifugation, washed with the corresponding lysis buffer (20 mM NaCl, 20 mM Tris.HCl, pH 7.4 - MIF; 20 mM NaCl, 20 mM Tris.HCl, pH 8.5 - D-DT; 1mM EDTA, 50 mM Tris.HCl, pH 7.5 - HPPD), and stored at -80°C. The ^{15}N-labeled MIF protein was produced as described above with some modifications. Instead of using LB, the isotopically labeled WT MIF was grown in M9 minimal media supplemented with ^{15}N-ammonium chloride. Cells were grown and induced as described above. Purification of MIF proteins was performed as following. The cell pellet containing the protein of interest was resuspended in the lysis buffer and lysed *via* sonication before being centrifuged to remove cell

debris. The supernatant was then filtered with a 0.22 µm PES syringe filter (Bioland Scientific LLC) and loaded onto 5 ml and 120 ml Q-Sepharose columns connected in series. MIF did not bind to either of the columns and collected in the flow-through. D-DT was purified in a similar manner using only a 5 ml Q-Sepharose column. At pH 8.5, the protein bound to the resin and eluted with 5% of 1 M NaCl, 20 mM Tris.HCl, pH 8.5. The remaining impurities of MIF and D-DT were removed by size exclusion chromatography using a 16/60 Superdex 75 column (Cytiva, Marlborough, MA). The MIF and D-DT running buffers, used for the final purification step, were the lysis buffers reported above. for HPPD purification, the cell lysate was prepared in a similar manner to what we described above for MIF and loaded onto a 120 mL Q-Sepharose column. HPPD did not bind to the column and eluted in the flow though. Following, the protein was further purified by size exclusion chromatography using a 16/60 Superdex 75 column that was equilibrated with 50mM Tris.HCl, pH 7.5. Protein purity and concentration for MIF, D-DT, and HPPD were determined by sodium dodecyl sulfate–polyacrylamide (SDS-PAGE) gel electrophoresis and Pierce BCA assay (Thermo Fisher), respectively.

MIF and D-DT enzymatic assays.

A 30 mM stock solution of 4-HPP was prepared in 0.5 M ammonium acetate pH 6.2 and incubated at room temperature under rocking for ≥16 hours to allow the formation of keto-HPP tautomer. for the assay, 4-HPP was added first at a final concentration range of 0 to 2 mM, followed by 0.420 M of borate, and 50 nM MIF or 250 nM D-DT. To examine the impact of ions in the enzymatic activities of MIF and D-DT, the above process repeated, including 1 mM EDTA. for the MIF inhibition studies, DMSO (1% final) or ISO-1 dissolved in DMSO were also added to the reaction mixture. The final concentration of ISO-1 varied between 0-75 µM. In all experiments, the total reaction volume was 150µL. The formation of enol-HPP/borate complex (ε_{306} = 11400

M^{-1} cm^{-1}) was measured at 306 nm using a Tecan Infinite M-Plex microplate reader (TECAN). Each experiment was carried out for either 180 seconds (MIF) or 300 seconds (D-DT), using 10-second intervals. Analysis was performed in GraphPad Prism 9.3.1 and all experiments were performed at least in triplicate.

NMR analysis of the 4-HPP samples.

The NMR spectra of 4-HPP were recorded on a JEOL-ECA600 instrument at room temperature with spinning (15 Hz). Commercial deuterated dimethyl sulfoxide (DMSO-d_6) solvent was used without additional treatment. The proton spectrum was recorded with an x-offset of 8 ppm and a spectral sweep width of 16 ppm. The data was digitized with 32 k points and zero-filled four times prior to Fourier transformation including a 0.1 Hz single exponential apodization function. The pulse field gradient heteronuclear multiple quantum coherence (pfg-HMQC) spectrum was recorded with an x-offset at 5 ppm (sweep width of 15 ppm, 1024 points in f1, zero-filled 2x) and y-offset at 100 ppm (sweep width of 250 ppm, 256 points in f2, zero-filled 4x). The one-bond coupling constant was set to 140 Hz. The three gradients were G1 and G2 (1 ms, 0.18 T/m), G3 (1 ms, 90.54 mT/m) and a gradient recovery period of 0.1 ms. The pulse field gradient heteronuclear multiple bond correlation (pfg-HMBC) spectrum was acquired with an x-offset at 5 ppm (sweep width of 15 ppm, 2048 points in f1, zero-filled 2x) and y-offset at 100 ppm (sweep width of 250 ppm, 256 points in f2, zero-filled 4x). The one-bond coupling constant was set to 140 Hz and the long-range coupling constant was set at 8 Hz. The three gradients were G1 and G2 (1 ms, 0.18 T/m), G3 (1 ms, 90.54 mT/m) and a gradient recovery period of 0.1 ms.

Analysis of 4-HPP impurities by Mass Spectrometry.

The 4-HPP powders, obtained from TCI and Chemodex, were dissolved in a mixed solvent of high-performance liquid chromatography (HPLC) grade methanol and water (1:1 volume per

volume) to make a stock solution of ~10^{-3} M. The stock solution was following diluted in the same solvent to achieve a concentration of ~10^{-5} M and introduced into the mass spectrometer via loop injection through a 6-port injection valve. Mass spectrometry measurements were performed on a triple quadrupole mass spectrometer interfaced with an electrospray ionization (ESI) source (Varian 320L, Agilent Technologies, Santa Clara, CA). A separate syringe pump was connected to the ESI source with a flow rate set at 25 μL/min, which pushed the mixed solvent to the ESI source via the injection valve. The full scan was collected using the profile mode with a mass range of *m/z* 30-600. The standard instrumental parameters were employed throughout the experiments. The parameters include the ESI needle onset voltage of 4.5 kV, capillary voltage of 40 V, nitrogen nebulizing gas pressure of 50 psi, drying gas pressure of 18 psi and drying gas temperature of 180 °C. Collision-induced dissociation (CID) experiments were performed on selected ions with higher *m/z* values than the compound in the sample. Argon gas was used as the collision gas with a pressure of 0.5 m Torr. The collision energy was set around 10-15 V.

Scheme 3.S1.

Possible synthetic routes of 4-HPP.

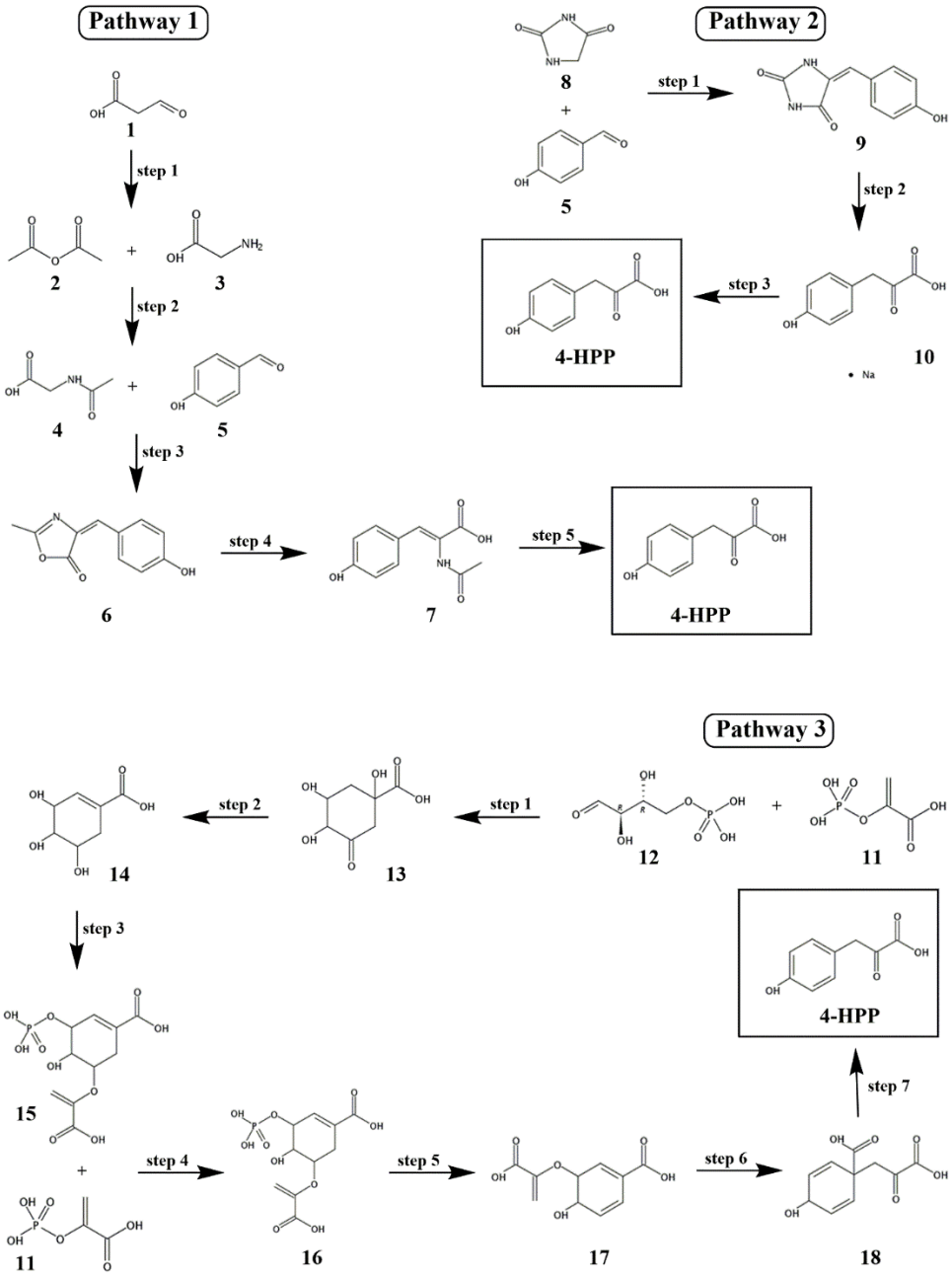


Table 3.S1.

Inhibition potency of ISO-1 against wild-type MIF using 4-HPP from different manufacturers.

	Trial 1	***Trial 2***	***Trial 3***	***Trial 4***	**Average K_i (μM)***
Chemodex	47.5	85.7	86.7	37.8	64.4 ± 25.5
ChemCruz	40.4	40.4	57.8	39.0	44.4 ± 9.0
TCI	19.5	20.9	31.2	19.8	22.9 ± 5.6

* Error values are shown as standard deviations.

Table 3.S2.

*MIF chemical shifts for the amide backbone (ppm) **

	pH 7.0		pH 6.2	
	Amide Nuclei			
Residue	^{15}N	^{1}H	^{15}N	^{1}H
F3	125.291	8.971	125.775	8.967
V5	122.437	8.364	122.445	8.35
N6	126.594	8.397	126.535	8.364
T7	117.341	9.165	117.592	9.177
R11	124.098	9.095	124.081	9.092
S13	112.578	7.908	112.577	7.901
V14	123.65	7.565	123.627	7.562
D16	123.827	8.825	123.804	8.828
F18	121.65	7.848	121.697	7.852
L19	116.096	8.356	116.163	8.369
S20	118.113	7.993		
E21	124.496	8.194	124.546	8.199
L22	117.616	8.3	117.571	8.307
T23	114.467	7.454	114.571	7.477
Q24	119.409	7.836		
Q25	117.647	8.569	117.571	8.563
L26	119.113	8.43	118.907	8.394
A27	127.351	8.264	127.41	8.264
Q28	116.403	7.558	116.313	7.532
A29	119.413	7.991	119.419	8.016
T30	104.083	8.022	103.977	8.016
G31	110.122	7.69	110.14	7.659

(Table 3.S2 Continued)

K32	121.356	7.649		
Q35	114.211	8.47	114.472	8.475
Y36	113.647	7.397	113.093	7.439
I37	121.616	7.699		
A38	134.896	9.105	134.088	8.753
V39	116.859	8.772	117.223	8.814
H40	126.935	8.666	126.797	8.604
V41	126.359	8.684	126.549	8.692
V42	129.338	9.273	129.291	9.241
D44	112.473	8.684	112.502	8.677
Q45	115.307	8.687	115.234	8.662
M47	122.482	7.704	122.482	7.743
A48	122.004	8.994	121.739	8.955
F49	121.264	8.38		
G50	120.121	9.104		
G51	106.429	9.636	106.233	9.555
S52	114.126	7.559	114.011	7.541
E54	119.92	7.644	119.919	7.658
C56	122.506	9.619	122.396	9.609
A57	120.027	8.839	119.97	8.824
L58	119.455	8.653		
C59	121.929	9.148	121.707	9.132
H62	125.781	9.147	125.476	9.123
S63	115.097	8.563	115.681	8.596
I64	128.61	7.927	128.72	7.983
G65	118.546	7.97		
K66	110.173	8.37	110.71	8.384
I67	117.642	7.335	117.632	7.335
G68	114.06	7.634	114.025	7.65
N72	117.002	8.106	117.114	8.094
R73	122.17	7.475	122.072	7.474
S74	115.87	7.162	115.765	7.156
S76	116	8.149	116.049	8.177
K77	124.277	7.567	124.306	7.593
L78	120.062	7.564	120.062	7.571
L79	118.17	9.037	118.129	9.016
C80	115.991	8.626	116.023	8.649
G81	107.832	7.718	107.888	7.733
L82	122.964	7.846	122.959	7.841
L83	118.958	8.368	118.901	8.369
A84	122.921	8.17	122.959	8.169
E85	117.529	8.183	117.487	8.15

(Table 3.S2 Continued)

R86	111.609	8.5	111.535	8.481
L87	113.734	7.091	113.74	7.099
R88	113.003	7.016	113.023	7.018
I89	119.206	6.889	119.208	6.864
S90	124.478	9.004	124.512	9.006
R93	123.915	8.097	123.926	8.121
Y95	119.449	7.857		
I96	119.959	8.918		
N97	126.22	8.287	125.966	8.289
Y99	119.528	8.732		
D100	124.54	8.631	124.613	8.645
M101	125.815	9.014	125.744	9.016
N102	123.093	8.771	122.878	8.755
A103	130.617	9.11	130.624	9.117
A104	114.479	8.3	114.528	8.293
N105	114.295	8.232	114.167	8.213
V106	119.819	7.33	119.945	7.343
W108 sc**	129.175	9.804	129.214	9.879
N109	126.432	7.145	126.486	7.143
N110	104.721	7.137	104.721	7.137
S111	113.925	7.604	113.977	7.614
T112	111.118	8.573	111.439	8.614
F113	118.4	8.106		
A114	127.224	7.346	127.176	7.361

* Blank cell implies that the chemical shift could not be determined due to either signal overlap or broadening. ** sc denotes side chain.

Table 3.S3.

*Chemical shifts for the amide backbone of MIF:Chemodex at 200:1 molar ratio (ppm)**

	pH 7		pH 6.2	
	Amide Nuclei			
Residue	^{15}N	^{1}H	^{15}N	^{1}H
F3	125.662	8.967		
V5	122.367	8.366	122.346	8.354
N6	126.534	8.395	126.542	8.376
T7	117.403	9.168	117.64	9.185
R11	124.111	9.096	124.126	9.101

(Table 3.S3 Continued)

S13	112.595	7.906	112.581	7.903
V14	123.669	7.564	123.659	7.565
D16	123.803	8.823	123.796	8.822
F18	121.66	7.852	121.671	7.86
L19	116.117	8.357	116.196	8.37
S20	118.166	8.002		
E21	124.521	8.192	124.545	8.195
L22	117.656	8.291	117.568	8.296
T23	114.491	7.454	114.549	7.459
Q25	117.646	8.554	117.604	8.555
L26	118.995	8.399		
A27	127.315	8.243	127.31	8.21
Q28	116.41	7.55	116.394	7.542
A29	119.372	7.994	119.253	7.979
T30	103.917	7.98	103.607	7.931
G31	110.165	7.683	110.24	7.666
K32	121.27	7.615		
Q35	114.296	8.483	114.563	8.518
Y36	113.619	7.398	113.232	7.416
I37	121.84	7.726		
A38			133.769	8.529
V39	117.088	8.77	117.517	8.838
H40	126.941	8.635	126.844	8.603
V41	126.398	8.688	126.56	8.695
V42	129.329	9.261	129.248	9.244
D44	112.484	8.68	112.475	8.678
Q45	115.283	8.673	115.194	8.655
M47	122.441	7.713	122.485	7.74
A48	121.942	8.99	121.724	8.956
F49	121.415	8.403		
G50	120.008	9.08		
G51	106.355	9.604	106.186	9.527
S52	114.055	7.552	113.945	7.528
E54	119.922	7.644	119.927	7.657
C56	122.485	9.615	122.4	9.608
A57	120.035	8.827	120.064	8.815
L58	119.434	8.663		
C59	121.938	9.159	121.759	9.128
H62	125.715	9.146		
S63	115.165	8.589	115.646	8.628
I64	128.411	7.916	128.107	7.962
G65	118.574	7.962		

(Table 3.S3 Continued)

K66	110.498	8.376	111.022	8.396
I67	117.675	7.328	117.763	7.343
G68	114.05	7.64	114.061	7.65
N72	117.067	8.088	117.111	8.088
R73	122.107	7.469	121.959	7.467
S74	115.836	7.161	115.786	7.155
S76	116.045	8.165	116.118	8.186
K77	124.281	7.569	124.261	7.576
L78	120.061	7.565	120.078	7.561
L79	118.191	9.038	118.191	9.015
C80	115.99	8.632	115.93	8.638
G81	107.837	7.728	107.914	7.738
L82	122.967	7.844	122.949	7.845
L83	118.95	8.369		
A84	122.934	8.166	122.981	8.164
E85	117.52	8.181	117.493	8.142
R86	111.592	8.488	111.518	8.476
L87	113.765	7.098	113.754	7.106
R88	113.013	7.014	113.013	7.018
I89	119.184	6.869	119.207	6.865
S90	124.465	8.997	124.436	8.995
R93	123.931	8.103	124.047	8.137
I96	120.126	8.918		
N97	126.127	8.3	125.816	8.326
Y99	119.416	8.71		
D100	124.529	8.635	124.605	8.665
M101	125.729	9.036		
N102	122.956	8.762	122.649	8.753
A103	130.601	9.104	130.61	9.111
A104	114.498	8.287	114.55	8.279
N105	114.334	8.22	114.263	8.207
V106	119.857	7.306	119.894	7.294
W108 sc**	129.209	9.852		
N109	126.404	7.115		
N110	104.772	7.142	104.797	7.144
S111	113.935	7.599	113.921	7.594
T112	111.242	8.584	111.453	8.6
F113	118.169	8.127		
A114	127.05	7.335	126.865	7.33

* Blank cell implies that the chemical shift could not be determined due to either signal overlap or broadening. ** sc denotes side chain.

Table 3.S4.

*Chemical shifts for the amide backbone of MIF: TCI at 200:1 molar ratio (ppm)**

	pH 7		pH 6.2	
	Amide Nuclei			
Residue	^{15}N	^{1}H	^{15}N	^{1}H
F3	125.47	8.956		
V5	122.391	8.371		
N6	126.619	8.396	126.576	8.387
T7	117.401	9.179	117.597	9.183
R11	124.095	9.094	124.099	9.1
S13	112.582	7.905	112.588	7.901
V14	123.668	7.562	123.666	7.562
D16	123.822	8.823	123.796	8.822
F18	121.66	7.848	121.684	7.861
L19	116.097	8.348	116.199	8.364
S20	118.147	7.968		
E21	124.514	8.191	124.525	8.194
L22	117.631	8.295	117.604	8.294
T23	114.511	7.453	114.528	7.464
Q24	119.427	7.838		
Q25	117.631	8.569	117.592	8.555
L26	119.081	8.426	118.836	8.387
A27	127.356	8.255	127.319	8.212
Q28	116.412	7.551	116.395	7.536
A29	119.405	7.987	119.269	7.99
T30	104.04	8.002	103.712	7.931
G31	110.159	7.687	110.244	7.667
K32	121.278	7.627		
Q35	114.239	8.467	114.538	8.498
Y36	113.643	7.39	113.262	7.42
I37	121.725	7.707		
V39	116.95	8.775	117.5	8.854
H40	126.936	8.642	126.856	8.604
V41	126.378	8.695	126.595	8.697
V42	129.319	9.265	129.235	9.241
D44	112.497	8.677	112.485	8.683
Q45	115.279	8.677	115.207	8.656
M47	122.451	7.709	122.522	7.75
A48	121.942	8.986	121.729	8.96
F49	121.322	8.388		
G50	120.08	9.077		

(Table 3.S4 Continued)

G51	106.413	9.604	106.259	9.543
S52	114.105	7.554	113.962	7.534
E54	119.94	7.644	119.927	7.642
C56	122.506	9.619	122.415	9.591
A57	120.045	8.842	120.046	8.808
L58	119.451	8.653		
C59	121.85	9.147	121.755	9.132
S63	115.112	8.574	115.637	8.648
I64	128.61	7.954		
G65	118.147	7.968		
K66	110.402	8.376	111.231	8.406
I67	117.642	7.335	117.681	7.362
G68	114.045	7.64	114.008	7.648
N72	117.034	8.099	117.119	8.057
R73	122.125	7.477	121.969	7.47
S74	115.87	7.162	115.835	7.165
S76	116.015	8.156	116.137	8.186
K77	124.252	7.568	124.254	7.578
L78	120.047	7.561	120.021	7.559
L79	118.139	9.034	118.137	9.02
C80	115.965	8.621	115.988	8.648
G81	107.824	7.723	107.888	7.733
L82	122.95	7.843	122.95	7.838
L83	118.975	8.375	118.863	8.362
A84	122.929	8.164	122.95	8.16
E85	117.513	8.181	117.46	8.14
R86	111.61	8.49	111.529	8.482
L87	113.69	7.087	113.76	7.1
R88	113.011	7.01	113.055	7.019
I89	119.206	6.889	119.163	6.863
S90	124.469	8.999	124.417	8.992
R93	123.899	8.094	124.004	8.127
Y95	119.449	7.864		
I96	120.021	8.928		
N97	126.089	8.294	125.802	8.312
Y99	119.515	8.72		
D100	124.525	8.643	124.661	8.66
M101	125.758	9.025		
N102	122.989	8.76	122.62	8.749
A103	130.598	9.101	130.562	9.101
A104	114.508	8.291	114.595	8.282
N105	114.332	8.226	114.387	8.207

(Table 3.S4 Continued)

V106	119.768	7.308	119.83	7.331
W108 sc**	129.213	9.827	129.393	9.928
N109	126.389	7.135		
N110	104.769	7.141	104.805	7.141
S111	113.94	7.6	113.929	7.596
T112	111.177	8.571		
A114	127.091	7.336	126.827	7.327

* Blank cell implies that the chemical shift could not be determined due to either signal overlap or broadening. ** sc denotes side chain.

Table 3.S5.

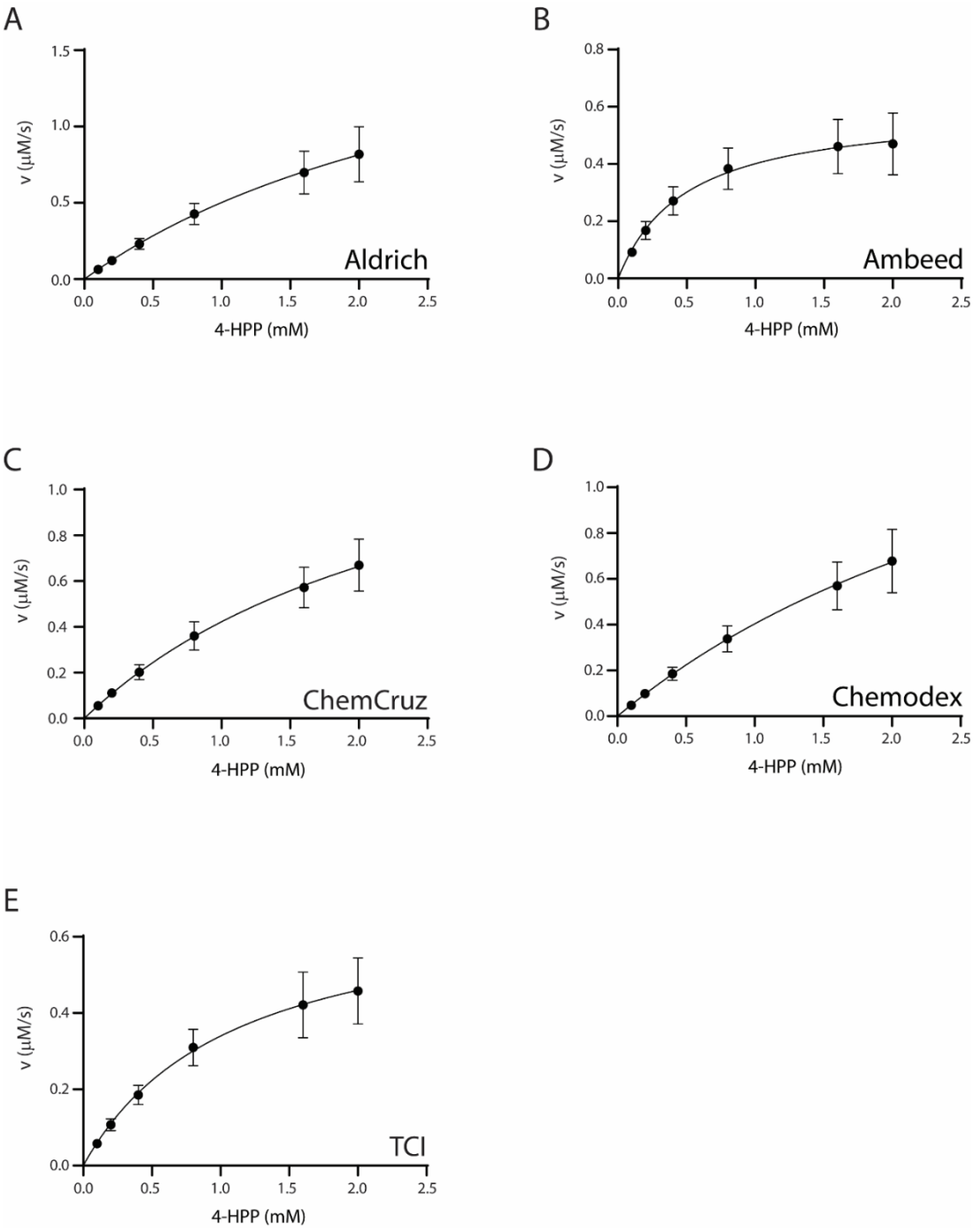
Michaelis-Menten parameters of wild-type MIF and D-DT using 4-HPP from different manufacturers

		MIF*		**D-DT ***	
	Lot number	**K_M (mM)**	**V_{max} (μM s^{-1})**	**K_M (mM)**	**V_{max} (μM s^{-1})**
Chemodex	X8048	3.9 ± 0.1	2.9 ± 0.1	4.1 ± 0.6	0.6 ± 0.1
TCI	6TM4L-YH	1.0 ± 0.0	0.8 ± 0.1	2.1 ± 0.5	0.3 ± 0.1

* Overlay of the MIF and D-DT Michaelis-Menten plots is shown in Figure 2. The experiments were carried out in triplicate (n=3) and the error values are shown as standard deviations.

Figure 3.S1

Michaelis-Menten plots of MIF using 4-HPP from different manufacturers.

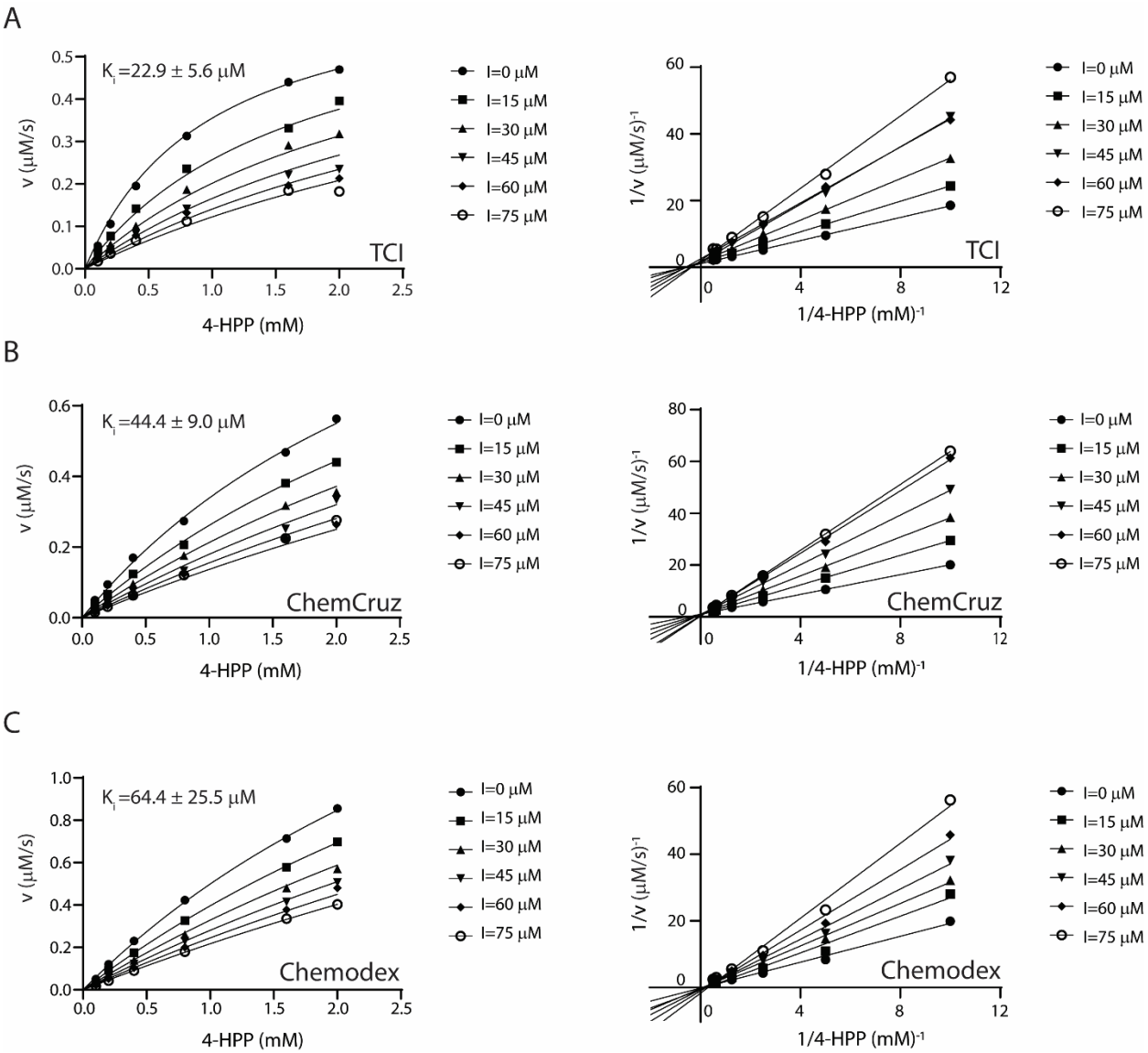


(Figure 3.S1 Continued)

Note. The 4-HPP samples were obtained from (**A**) Aldrich (lot number: BCBW9189), (**B**) Ambeed (lot number: A632541-002), (**C**) ChemCruz (lot number: F1917), (**D**) Chemodex (lot numbers: P4006), and (**E**) TCI (lot numbers: UJNCJ-SA).

Figure 3.S2.

Michaelis-Menten (left) and Lineweaver-Burk (right) plots of ISO-1, an MIF inhibitor.

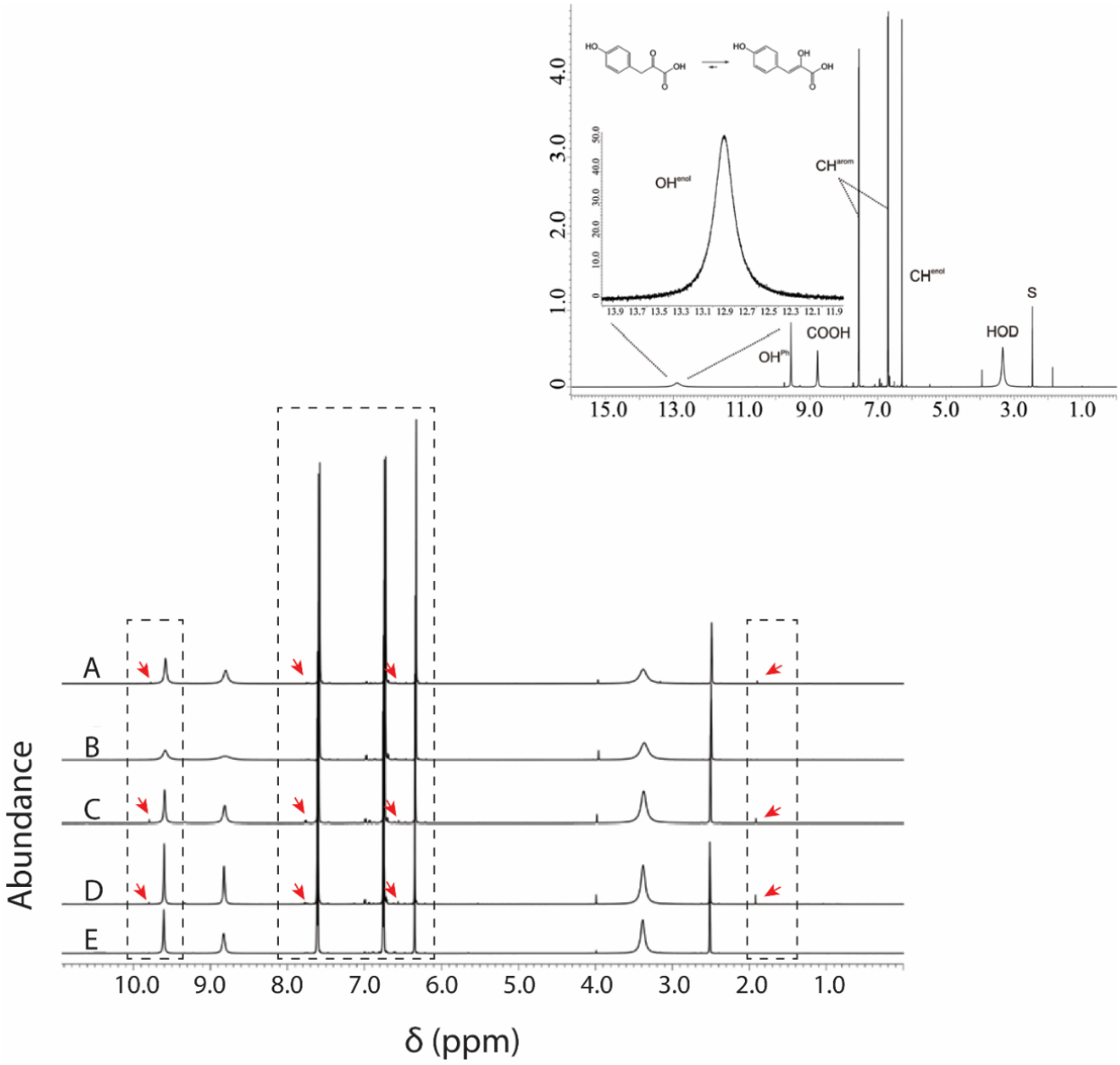


(Figure 3.S2 Continued)

Note. The tautomerase activity of MIF was measured using 4-HPP from different vendors; **A:** TCI, **B:** ChemCruz, and **C:** Chemodex, respectively. Both non-linear and linear plots demonstrate a competitive type of inhibition for ISO-1. Differences in the inhibition constant, suggest 4-HPP contains impurities that interfere with the assay.

Figure 3.S3.

NMR analysis of 4-HPP samples from different commercially available sources.



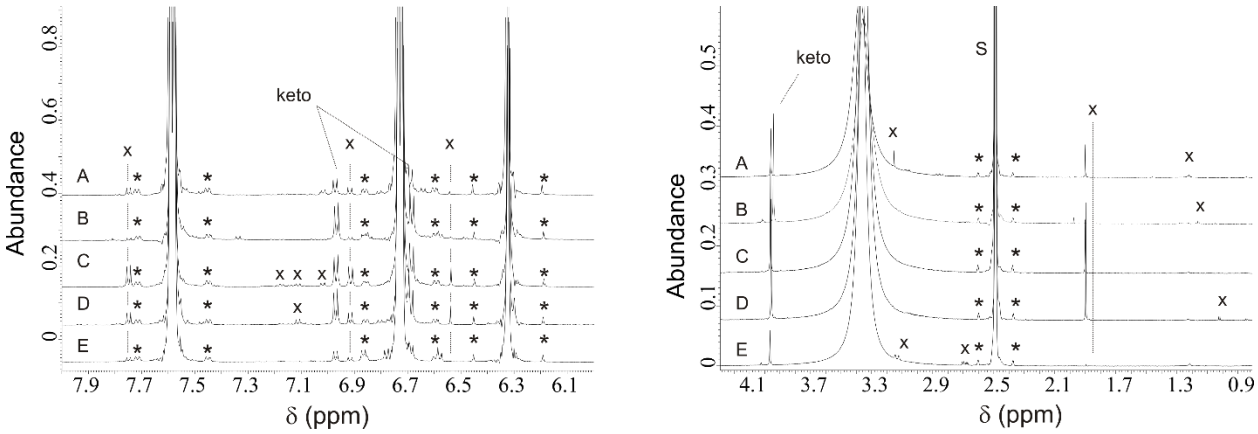
Note. The upper right panel (zoom out) shows the ^{1}H-NMR spectrum of 4-HPP in DMSO-d_6. The enol hydroxyl proton was observed at 12.9 ppm. Deuterium-exchanged water and solvent residual peak were labeled. There was < 0.5 % of the keto-form present. Conjugation with the enol resulted in a remarkable chemical shift of 8.8 ppm for the carboxylic acid proton, different from typical

(Figure 3.S3 Continued)

values around 12 ppm. Main panel (zoon in): from the top to the bottom, the spectra represent 4-HPP samples obtained from **A:** Aldrich, **B:** Ambeed, **C:** ChemCruz, **D:** Chemodex, **E:** TCI, respectively. Unique impurities that are found only in particular samples are indicated with red arrows. The dotted rectangles mark the regions where the impurities are found.

Figure 3.S4.

Details of 4-HPP ^{1}H-NMR spectra.



4-HPP was obtained from **A**: Aldrich, **B**: Ambeed, **C**: ChemCruz, **D**: Chemodex, **E**: TCI and dissolved in DMSO-d_6. Impurities found in the regions of 7.9-6.1 ppm (left) and 4.1-0.9 ppm (right) were labeled with x. The asterisks denote ^{13}C-satellites of the pure compound. Some impurities were present in all samples, whereas other impurities were unique to some samples. It must be clearly stated that NMR is inadequate as an instrumental technique for trace analysis. Additional impurities might be present in the baseline.

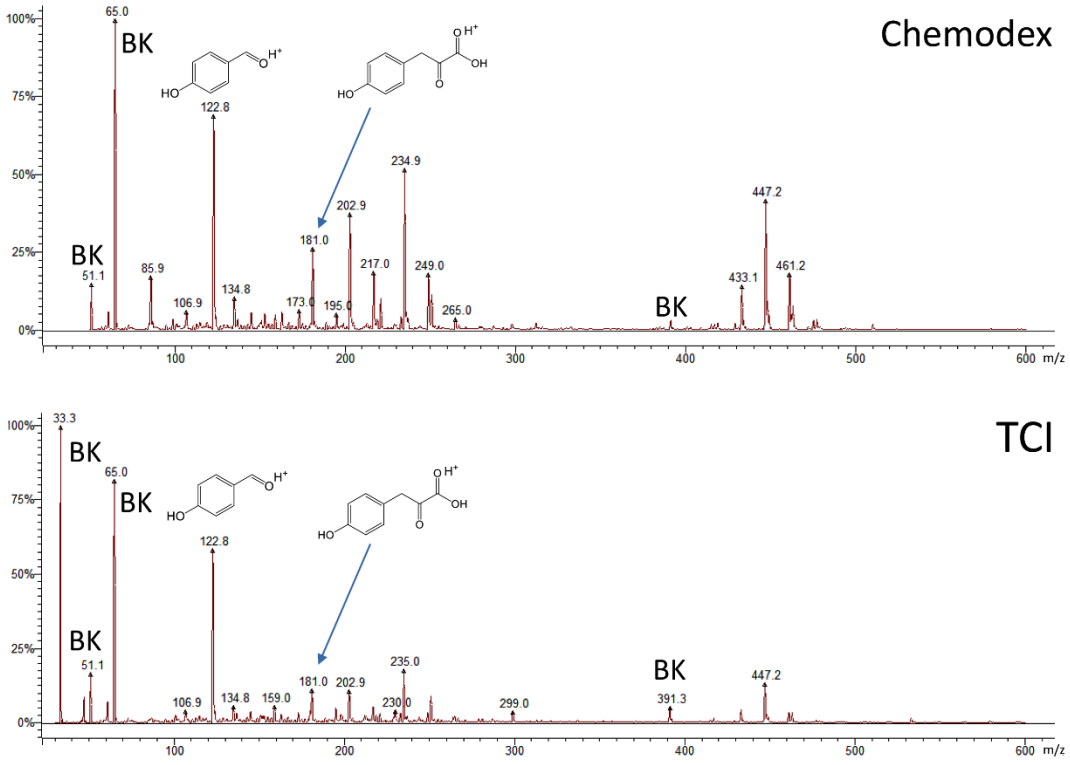
Figure 3.S5.

Comparison of the five 4-HPP powders.

Note. Visual inspection of the five powders demonstrates a clear difference between the color of TCI powder and the remaining four. Of note, the TCI 4-HPP was the one which most accurately reproduced the previously published turnover and inhibition data for 4-HPP and ISO-1, respectively.

Figure 3.S6.

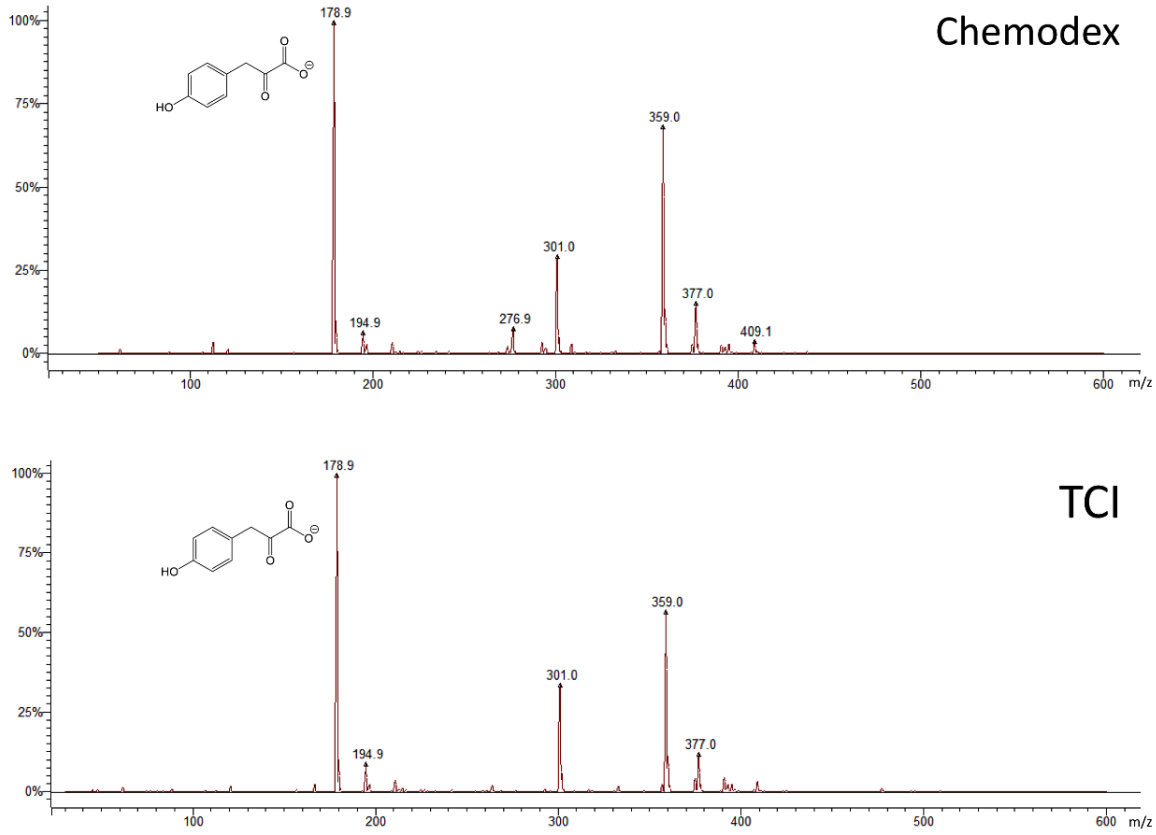
Positive ion mode analysis of 4-HPP samples obtained from Chemodex and TCI.



Note. The protonated 4-HPP and the key precursor molecule 4-hydroxybenzaldehyde were detected at *m/z* 181.0 and 122.8, respectively. A unique contamination peak for Chemodex was detected at *m/z* 85.9. BK stands for background peaks.

Figure 3.S7.

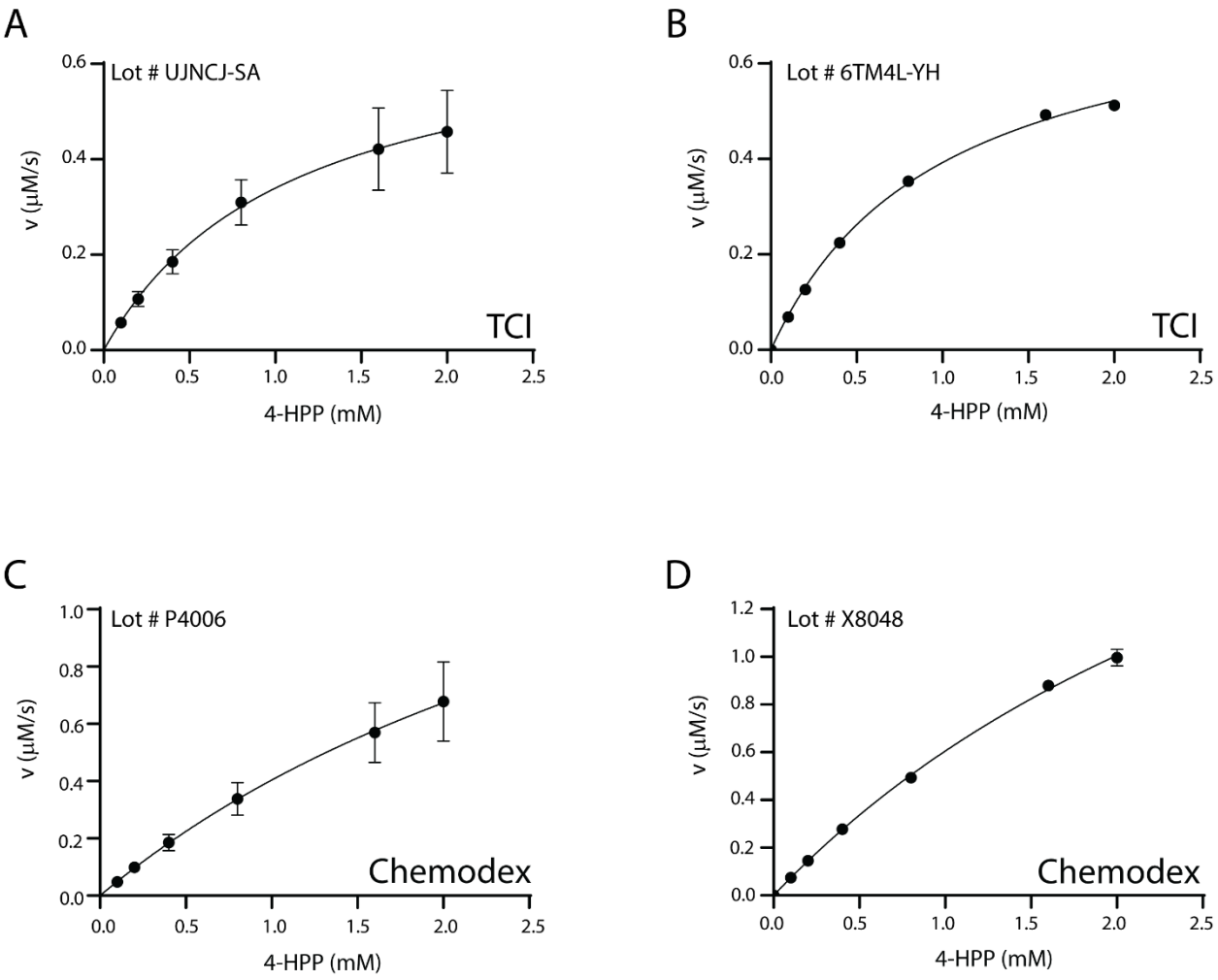
Negative ion mode analysis of 4-HPP samples obtained from Chemodex and TCI.



Note. The negatively charged 4-HPP was detected at *m/z* 178.9. A unique contamination peak for Chemodex was detected at *m/z* 276.9.

Figure 3.S8.

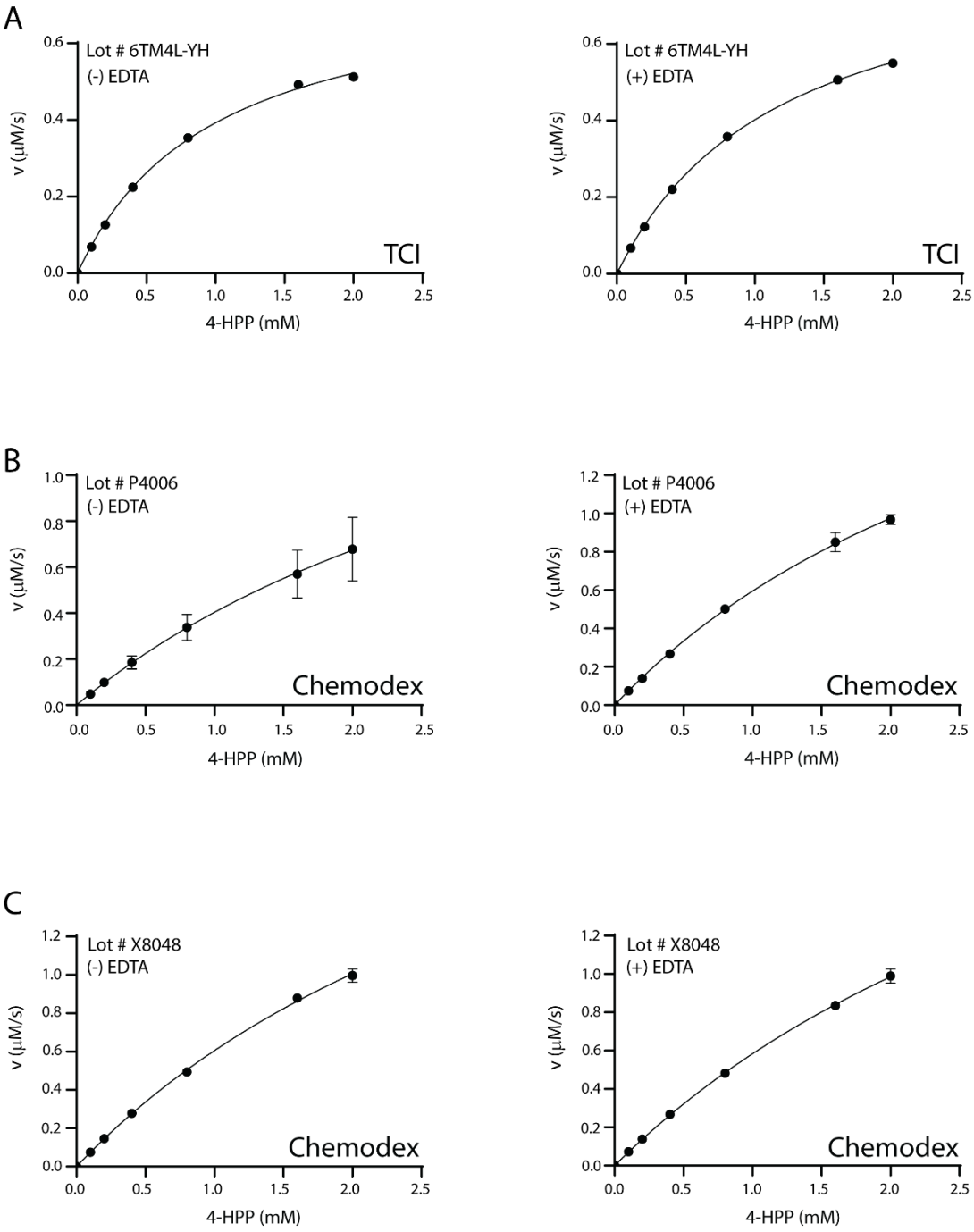
Impact of 4-HPP lot number on the reproducibility of Michaelis-Menten plots.



Note. The Michaelis-Menten plot of MIF was examined using 4-HPP from TCI (A and B) and Chemodex (C and D). from each manufacturer, we purchased and tested two different batches of 4-HPP. The outcome of this experiment demonstrates reproducible kinetic findings that are independent to the compound's Lot number.

Figure 3.S9.

Impact of metal ions on the reproducibility of Michaelis-Menten plots.



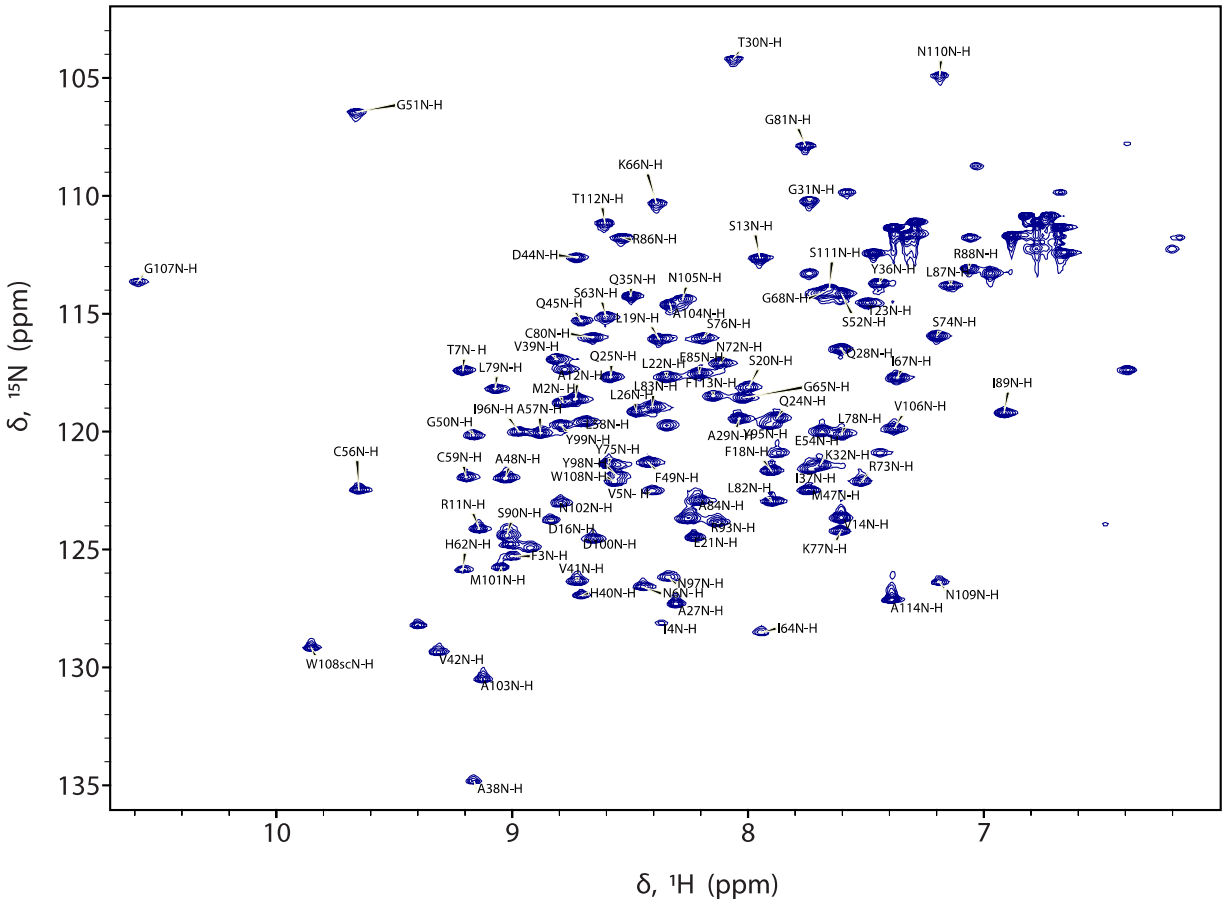
Note. The Michaelis-Menten plot of MIF was examined in the presence or absence of EDTA. (A)The TCI 4-HPP (Lot # 6TM4L-YH) was used as a control. (B-C) In the presence of EDTA,

(Figure 3.S9 Continued)

the Chemodex 4-HPPs (Lots # P4006 and X8048) do not demonstrate improvement of the kinetic parameters. Such findings show that metal ions are not responsible for the kinetic inconsistencies observed between TCI and Chemodex samples.

Figure 3.S10.

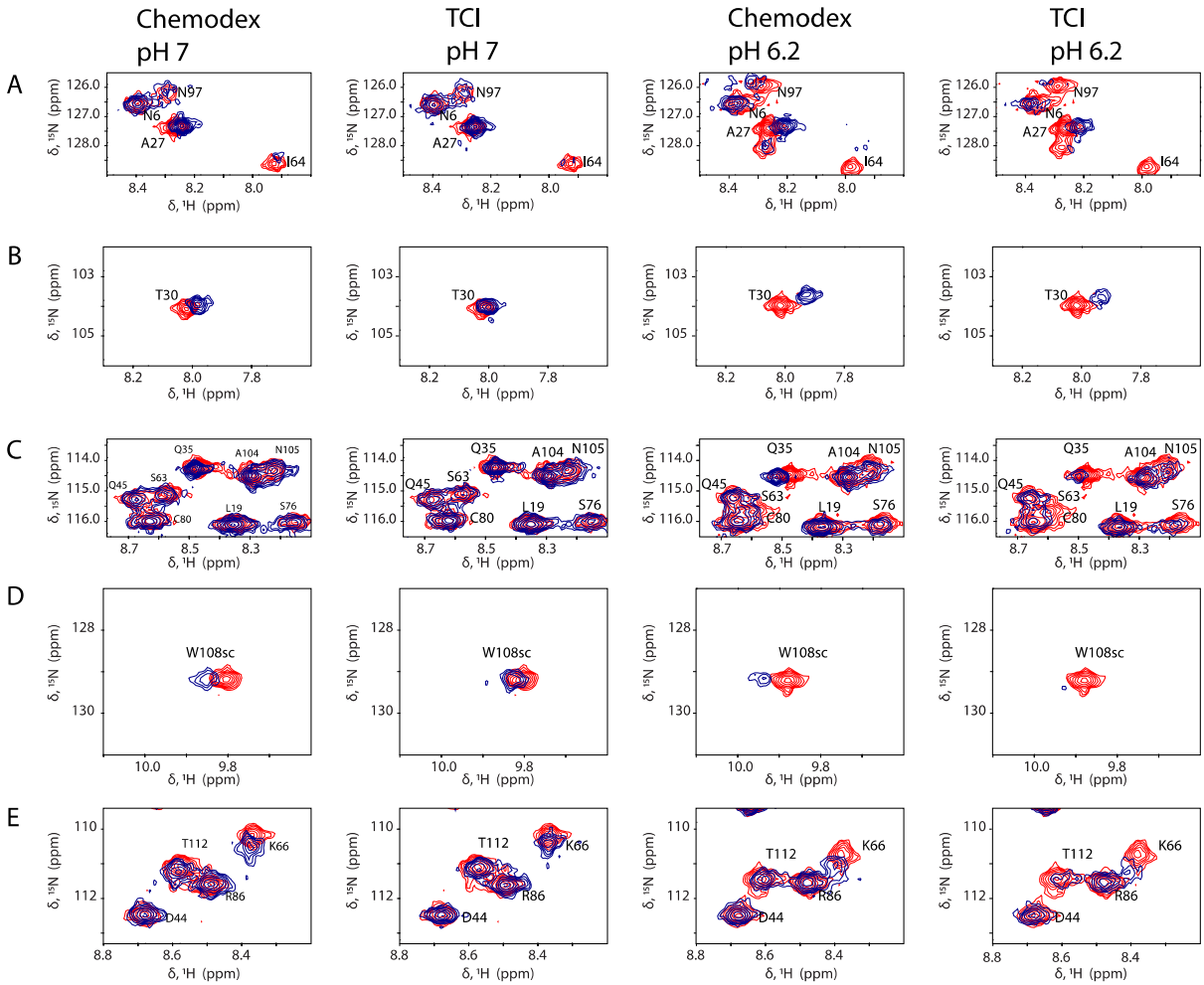
Backbone amide ^{1}H-^{15}N chemical shift assignment of MIF.



Note. 2D-[^{1}H-^{15}N]- HSQC spectrum of ^{15}N uniformly labeled MIF at 200 μM in 20 mM phosphate buffer, 1 mM EDTA, pH 7.0, 7.5% D_2O acquired at 30 ºC.

Figure 3.S11.

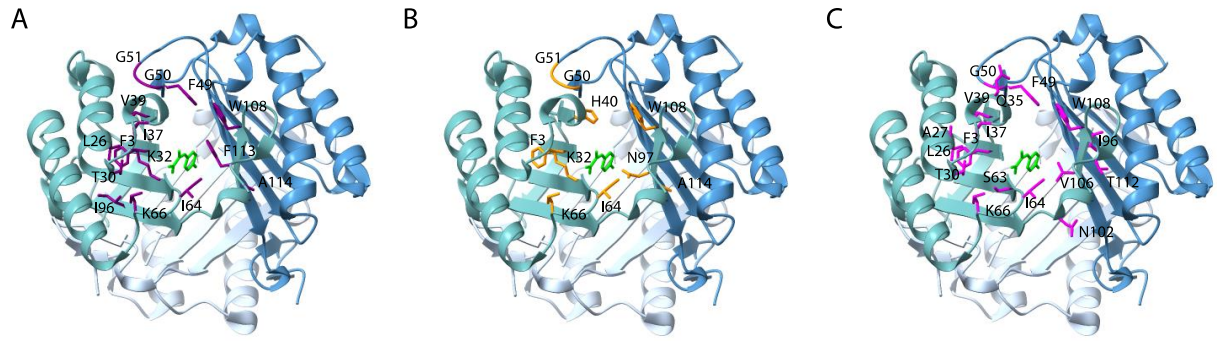
Chemical shift perturbations indicate differential binding between MIF and 4-HPP from different manufacturers.



Note. (**A-E**) Selected regions of 2D-[^{1}H-^{15}N]-SOFAST-HMQC spectra of MIF in the presence (blue) and absence (red) of 4-HPP from different manufacturers (Chemodex and TCI) at two pH values 6.2 and 7.0 as indicated. Pertinent amino acids are labeled.

Figure 3.S12.

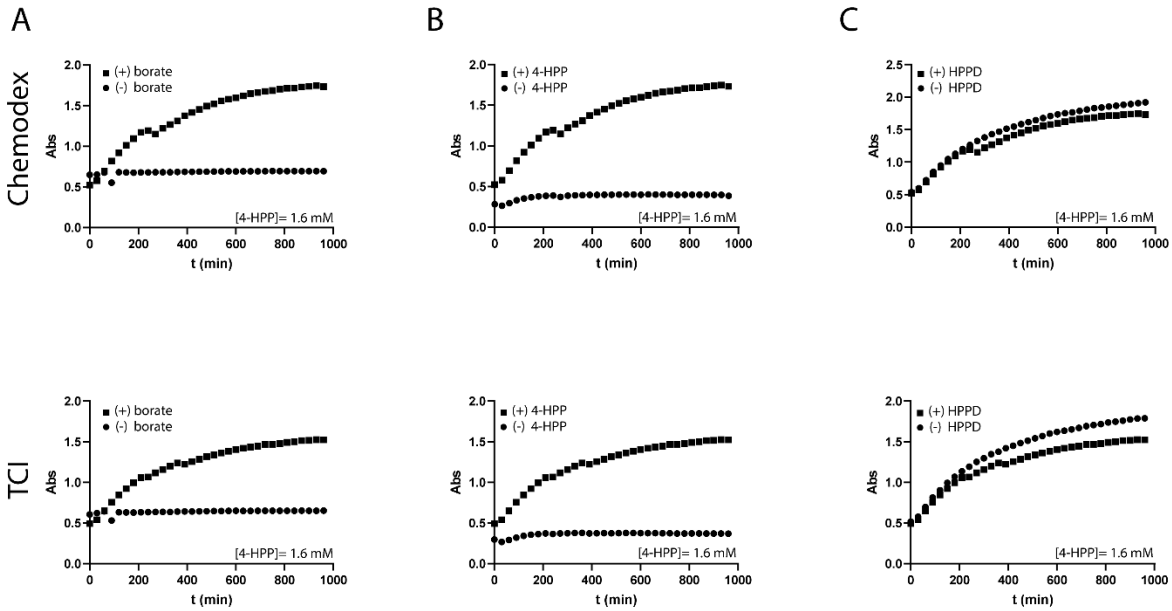
Chemical shift perturbation mapping onto the three-dimensional structure of MIF.



Note. Each monomer in the three-dimensional structure of MIF trimer bound to 4-HPP (PDB: 1CA7) is shown in different shades of blue. Only one 4-HPP molecule is depicted (green). Amino acids with chemical shift perturbations larger than 1SD for the binding of MIF with 4-HPP are highlighted in (**A**) purple for 4-HPP from Chemodex, pH 7.0; and (**B**) orange for 4-HPP from TCI, pH 7.0. (**C**) Mapping of residues that undergo differential chemical shift perturbation or broaden due to the binding of 4-HPP from TCI and Chemodex.

Figure 3.S13.

Control experiments to confirm the accuracy of HPPD assay.



Note. (**A-B**) In the absence of borate (**A**) or 4-HPP (**B**), the absorbance of the reaction mixture remains constant over time. This finding clearly shows that the substance absorbing at 306 nm is the enol-HPP-borate complex. (**C**) The absorbance reduction noted for the reaction mixture containing HPPD versus the one without HPPD, is associated with the catalytic conversion of 4-HPP to HGA. HPPD is more active in the case of TCI sample, demonstrating at the same time that impurities found in the Chemodex sample inhibit the enzyme.

CHAPTER 4: LIGAND-INDUCED CONFORMATIONAL CHANGES ENABLE INTERSUBUNIT COMMUNICATIONS IN D-DOPACHROME TAUTOMERASE

Published as: Parkins A, Chen E, Rangel VM, Singh M, Xue L, Lisi GP, Pantouris G. Ligand-induced conformational changes enable intersubunit communications in D-dopachrome tautomerase. Biophys J. 2023 Apr 4;122(7):1268-1276. doi: 10.1016/j.bpj.2023.02.019. Epub 2023 Feb 17. PMID: 36804669; PMCID: PMC10111345.

Abstract

D-Dopachrome tautomerase (D-DT or MIF-2) is a multifunctional protein with immunomodulatory properties and a documented pathogenic role in inflammation and cancer that is associated with activation of the cell surface receptor CD74. Alongside D-DT, macrophage migration inhibitory factor (MIF) is also known to activate CD74, promoting pathogenesis. While the role of the MIF/CD74 axis has been extensively studied in various disease models, the late discovery of the D-DT/CD74 axis has led to a poor investigation into the D-DT-induced activation mechanism of CD74. A previous study has identified 4-CPPC as the first selective and reversible inhibitor of D-DT and reported its potency to block the D-DT induced activation of CD74 in a cell-based model. In this study, we employ molecular dynamics (MD) simulations and nuclear magnetic resonance (NMR) experiments to study 4-CPPC induced changes to the dynamic profile of D-DT. We found that binding of the inhibitor remarkably promotes the conformational flexibility of C-terminal without impacting the structural stability of the biological assembly. Consequently, long-range intrasubunit (>11 Å) and intersubunit (>30 Å) communications are enabled between distal regions. Communication across the three subunits is accomplished via 4-

CPPC, which serves as a communication bridge after Val113 is displaced from its hydrophobic pocket. This previously unrecognized structural property of D-DT is not shared with its human homolog MIF, which exhibits an impressive C-terminal rigidity, even in the presence of an inhibitor. Considering the previously reported role of MIF's C-terminal in the activation of CD74, our results break new ground for understanding the functionality of D-DT in health and disease.

Introduction

Ligand binding often prompts conformational changes that inform the biological activity of a protein. Depending on the ligand, these structural events may have either a salutary or deleterious functional effect, providing, at the same time, mechanistic insight into the development of novel modulators [109, 110]. for transmembrane proteins, binding of an agonist triggers conformational changes that are essential for signaling cascades inside the cell [111]. On the other hand, binding of an antagonist promotes structural changes that induce an inactive receptor state [112]. Ligand-induced conformational changes have also been observed in soluble proteins, such as enzymes, and are coupled to their activities [113]. These protein motions have been studied over a broad timescale range by a variety of methods, including NMR, cryogenic electron microscopy (cryo-EM), protein crystallography, spectrophotometry, and MD simulations [66, 114].

While several proteins undergo loop and large domain conformational changes to perform biological activities, others accomplish their tasks with minimal flexibility. A representative example of the latter is MIF. MIF is a well-known immunomodulatory protein that was first described in the 1960s [11, 115]. The MIF trimer possesses several non-overlapping biological activities, including cytokine [27], chemokine [12], endonuclease [13], and catalytic [7]. Considering the homotrimeric biological assembly is only ~37 kDa, it is of great interest to

understand how a small protein like MIF coordinates multiple biological activities. Recent studies exploring the dynamic features of MIF have shown that its central □-sheet is responsible for regulating dynamic signal transmission between distal functional sites [53, 91]. While no major conformational changes have been observed, dynamic regulation of these activities occur through correlated backbone and sidechain motions on the fast timescale.

D-DT is a homolog of MIF and the second human member of the MIF superfamily. D-DT was originally isolated from rat liver and named after its ability to catalyze tautomerization of D-dopachrome [116]. Despite the 34% sequence identity, MIF and D-DT share endonuclease, tautomerase, and certain cytokine activities. The cytokine activity is the most notable as it is related to activation of the CD74/CD44 co-receptor complex [9, 27]. Activation of the receptor complex is associated with a cascade of MAP kinase signaling events resulting in various downstream effects, such as cell proliferation and inhibition of apoptosis [36].

The shared keto/enol tautomerase activity is of great interest as it is responsible for the identification of all reported small molecule MIF and D-DT modulators [21, 48, 57-59, 63, 117]. The enzymatic pocket is formed between two subunits and is controlled by the preserved N-terminus Proline. In the absence of a defined biological substrate, various naturally occurring (4-hydroxyphenylpyruvate (4-HPP) and phenylpyruvate (PP)) and synthetic (D-dopachrome) model substrates have been used. The MIF and D-DT active sites share several key residues important for the binding and tautomerization of these substrates, though overall, the catalytic site of MIF is appreciably more hydrophobic than that of D-DT and contains several aromatic residues that D-DT lacks [57]. These differences have culminated in numerous selective inhibitors for the better studied MIF and a shared covalent inhibitor with D-DT, known as 4-iodo-6-phenylpyrimidine (4-IPP). Meanwhile, only two reversible D-DT inhibitors have been described thus far, with the

first, 4-(3-Carboxyphenyl)-2,5-pyridinedicarboxylic acid (4-CPPC), being reported recently [15, 49, 59]. Crystallographic analysis of the D-DT/4-CPPC structure shows the ligand binds *via* an induced fit mechanism, forcing the C-terminal to adopt an open conformation. Consequently, the C-terminal becomes highly flexible, which explains the lack of electron density for its last four residues, Met114-Leu117 [49]. Considering the previously reported rigidity of the MIF trimer [51], the role of C-terminal residues in the MIF-induced activation of CD74 [48], and the similar structural topology between MIF and D-DT, it is of great importance to understand the potential biological significance of D-DT's enhanced C-terminal flexibility.

In this study, we investigate the effect of 4-CPPC binding on the dynamic profile of D-DT. MD simulations confirmed the previously reported crystallographic findings and exposed mechanistic insights into the role of C-terminal flexibility. After we experimentally excluded the scenario of structural instability, we showed the effect on the C-terminal to be clearly dynamic, enabling long-range intersubunit communications across the D-DT trimer. These exciting findings offer a new avenue for understanding the vastly unexplored biological functionality of D-DT, with focus on the CD74 activation mechanism. At the same time, our analysis provides key information, which aids the efforts made thus far to explain the divergent functional profiles of MIF and D-DT.

Materials and Methods

Materials and General Methods.

All chemicals used for buffers and synthesis were purchased from Sigma-Aldrich or Fisher Scientific and used without further purification. ^{1}H NMR and ^{13}C NMR spectra of 4-CPPC were collected on a JEOL ECA 600 MHz FT-NMR spectrometer. Protein NMR experiments were performed on a Bruker Avance NEO 600 MHz spectrometer at 30 ^{o}C. Mass spectrometry (MS)

spectra for small molecules were recorded on a JEOL Accu TOF LC^{TM} time-of-flight (TOF) mass spectrometer equipped with a Direct Analysis in Real Time (DART) ion source. Isothermal microcalorimetric measurements were performed on an affinity isothermal titration calorimetry (ITC) instrument (TA instruments).

Synthesis of 2,5-Dimethyl-4-(3-methylphenyl)pyridine (1).

In a 25 mL round-bottomed flask, m-Tolylboronic acid (219 mg, 1.61 mmol) and 4-bromo-2,5-dimethylpyridine (200 mg, 1.08 mmol) were dissolved in a mixture of 1:1 iPrOH: water (4 mL). After adding Na_3PO_4,12H_2O (1.43 g, 3.76 mmol), and 10 % Pd/C (40 mg, 3.5 mol %) to the flask, the reaction mixture was stirred for 12 h at 80 ^oC under Argon. The reaction mixture was diluted with water (10 mL) and Et_2O (10 mL) and filtered through celite. The filtrate was separated into two layers and the aqueous layer was extracted with Et_2O (2 x 10 mL). The combined organic layers were washed with brine (10 mL), dried over anhydrous Na_2SO_4, and concentrated under vacuum. Gradient chromatography of the residue (silica gel) with 0-2% MeOH in dichloromethane yielded compound **1** (115 mg, 54%). ^{1}H NMR (600 MHz, $CDCl_3$) δ 8.39 (s, 1H), 7.31 (m, 1H), 7.18-7.21p (m, 1H), 7.08-7.12 (m, 2H), 7.01 (s, 1H). 2.55 (s, 3H), 2.40 (s, 3H), 2.22 (s, 3H). ^{13}C NMR (150 MHz, $CDCl_3$) δ 155.61, 150.24, 149.96, 139.27, 139.10, 129.21, 128.64, 128.29,127.63, 123.77, 125.6, 23.79, 21.50, 16.90. HRMS (DART) calcd 198.1277 for $C_{14}H_{16}N$ $[M+H]^+$, found 198.1266.

Synthesis of 4-(3-Carboxyphenyl)-2,5-pyridinedicarboxylic acid (4-CPPC).

In a 25 mL round-bottomed flask, compound **1** (115 mg, 0.583 mmol) and $KMnO_4$ (507 mg, 3.21 mmol) were added to water (5 mL). After the reaction mixture was refluxed for 4 h, the excess of $KMnO_4$ was reduced with Na_2SO_3 (74.0 mg, 0.587 mmol), filtered through celite, and washed with hot water (95 ^oC). The volume of the filtrate was reduced to 7 mL under vacuum. The

filtrate was then acidified with 6M HCl in an ice-water bath to yield white precipitates. The precipitates were collected under vacuum filtration, washed with ice-cold water (25 mL), dried under vacuum overnight to yield **4-CPPC** (70 mg, 42 %). ^{1}H NMR (600 MHz, deuterated dimethyl sulfoxide (DMSO-d_6)) δ 9.10 (s, 1H), 8.07 (dt, *J* = 6.0 Hz, 1.8, 1H), 8.05 (s, 1H), 8.01 (t, *J* = 1.8 Hz, 1H), 7.77 (m, 1H), 7.67 (t, *J* = 6 Hz, 1H). ^{13}C NMR (150 MHz, DMSO-d_6) δ 167.27, 166.91, 165.48, 150.19, 149.96, 148.61, 137.87, 132.59, 131.04, 129.97, 129.61, 129.03, 128.93, 125.51. HRMS (DART) calcd 288.0503 for $C_{14}H_{10}NO_6$ $[M+H]^+$: found 288.0486.

Protein expression and purification.

Wild-type (WT) D-DT for kinetic experiments, ITC and NMR studies was expressed and purified as previously described [49, 52]. for kinetic assays and ITC, a pET-22b plasmid encoding D-DT was transformed into BL21-Gold (DE3) *E. coli* cells and grown in Luria-Bertani (LB) medium at 37 ºC to an OD_{600} 0.6 - 0.8. The ^{15}N-labeled D-DT was expressed in a similar way but instead of LB, the cells were grown in M9 minimal medium with $^{15}NH_4Cl$ to an OD_{600} 0.8 - 1.0. Protein expression was induced with 1 mM IPTG and incubated at 37 ºC for 4 h (kinetics and ITC) or 20 ºC for 18 h (NMR). The cell culture was harvested by centrifugation and resuspended in 20 mM Tris, 20 mM NaCl, pH 8.5 supplemented with 1 mM phenylmethylsulfonyl fluoride (PMSF). Cells were lysed on ice by sonication and then centrifuged to remove cell debris. The supernatant was filtered through a 0.22-μm filter, loaded onto a Q-Sepharose (anion-exchange) column, washed with 20 mM Tris, 20 mM NaCl, pH 8.5 and eluted with 5% of 20 mM Tris, 1 M NaCl, pH 8.5. Additional contaminants were removed by size-exclusion chromatography (HiLoad 16/600 Superdex 75 pg). Protein purity was assessed by sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE). The concentration of D-DT was determined by ultraviolet–visible

spectroscopy (UV-Vis) utilizing ε_{280} = 5500 $M^{-1}cm^{-1}$ and Pierce Bicinchoninic Acid (BCA) Protein Assay Kit (Thermo Fisher Scientific, Waltham, MA).

Isothermal Titration Calorimetry.

Before the ITC experiment, all samples were degassed for 5 min, at room temperature. The titration experiment was carried out with the ligand (4-CPPC) in a 264 μL rotating syringe and WT D-DT in the isothermal sample cell. The final concentrations of ligand and protein were 500 μM and 100 μM, respectively. 4-CPPC was injected into the sample cell as 50 aliquots of 3.5 μL. The control run was carried out under identical conditions with the D-DT/4-CPPC titration experiment but instead of protein, the sample cell contained buffer (20mM Tris, 20 mM NaCl, pH 8.50). The syringe rotational speed was 125 rpm, each injection lasted 4 s, and the delay between injections was 300 s. Before the first injection, the instrument equilibrated for 1800 s. Data generated for ligand-to-buffer and ligand-to-protein experiments were processed in the NanoAnalyze software (version 3.11.0). The corrected heat plot of D-DT/4-CPPC was fitted in the independent model.

Tautomerase Inhibition Assay.

Tautomerase inhibition assay was carried out in a similar manner as previously described [49]. A 30 mM stock solution of 4-HPP was prepared in 0.5 M ammonium acetate pH 6.20 and incubated overnight, under rocking, to favor the formation of the keto isoform. Serial dilutions of 4-HPP were then made yielding a final concentration range between 0-2 mM. Following 4-HPP, borate was added at a final concentration of 0.420 M. The inhibitor (4-CPPC) was dissolved in 100% DMSO and added to the mixture providing a final concentration range of 0-75μM. In all wells, the final concentration of DMSO was 1%. D-DT was then added at a final concentration of 250 nM. The reaction was measured in a Tecan Infinite M-Plex microplate reader at 306 nm for

300 s using 10 s intervals (ε_{306} = 11400 M^{-1} cm^{-1}). The kinetic experiments were performed in triplicate.

NMR titrations.

All NMR data were processed using NMRPipe [74] and analyzed in Sparky [75]. The DMSO tolerance of the D-DT protein was tested in a buffer of 20 mM Tris, 20 mM NaCl at pH 7.4. DMSO-d_6 (Sigma-Aldrich) was titrated into the D-DT sample to final (*v*/*v*) percentages of 3.2, 10, 18.9, and 26.8%, after which the integrity of the structure was monitored through ^{1}H-1D and 2D experiments. Based on this analysis, $^{1}H^{15}N$ transverse relaxation optimized spectroscopy (TROSY)- heteronuclear single quantum coherence (HSQC) spectra of D-DT were subsequently collected in a buffer of 20 mM Tris, 20 mM NaCl and 15% DMSO-d_6 at pH 7.4. 4-CPPC was dissolved in 100% DMSO-d_6 to create a stock solution at a concentration of 50 mM. Titrations were performed by adding small aliquots of 4-CPPC to the sample with gentle mixing by pipette before the acquisition of spectra.

MD Simulations and Analysis.

The starting models for our 1 μs simulations were obtained from the crystal structures of apo D-DT (Protein Data Bank (PDB): 1DPT) and D-DT/4-CPPC (PDB: 6C5F). The biological assembly of D-DT was generated in PyMOL [73] and the missing residues were added with Chimera [67]. Protein structure files (PSF) with hydrogens were generated using the Visual Molecular Dynamics (VMD) plugin, psfgen [69]. The PSF for D-DT was produced using the CHARMM36 topology [118], meanwhile a separate topology for 4-CPPC was generated using the CGENFF webserver. The resulting PSF and PDB pair were then used to generate a new set of PSF/PDB pair that placed the protein in a TIP3P water box. TIP3P waters and the system's net-charge were set to 0 by the addition of sodium and chloride ions. Minimization, heating to 300 K, and equilibration

processes were carried out with a 2 fs timestep, analogous to what has been previously described [91]. Simulations were then run using NAMD 2.12 [119]. Generalized cross-correlation analysis was performed for the α-carbons using g_correlation [70], a GROMACS [71] plug-in. Root mean square fluctuation (RMSF) analysis was performed in GROMACS. Analysis of the simulations was carried out in PyMOL. All simulations and subsequent analyses were completed in triplicate.

Results and Discussion

Apo and 4-CPPC bound D-DT demonstrate distinct dynamic profiles at the C-terminal.

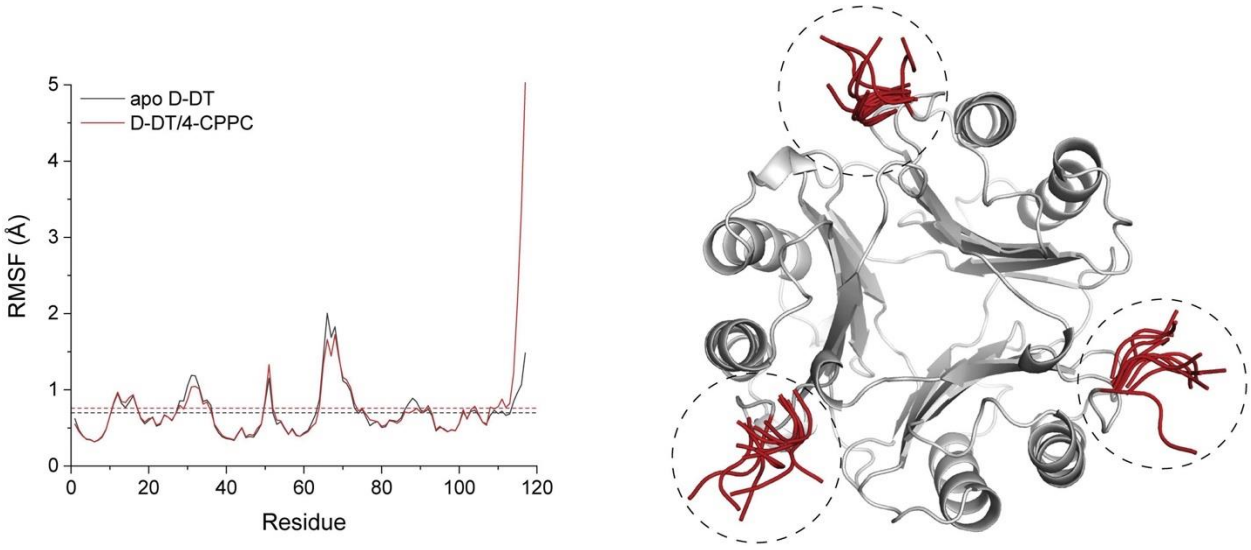
In a previous study, we reported the co-crystal structure of D-DT in complex with 4-CPPC and showed that ligand binding is associated with conformational changes of the C-terminal [49]. The ligand-induced displacement of the C-terminal drastically increased the flexibility of the last four amino acids, Met114-Leu117, which in protein crystallography is expressed as lack of electron density. It is of great interest to note that increased flexibility of the C-terminal is an uncommon feature of the MIF superfamily members. Besides D-DT, only *Plasmodium falciparum* (*Pf*) MIF has previously revealed a flexible C-terminal upon ligand binding [117].

To explore the significance of this finding, we performed 1μs MD simulations of the apo D-DT and 4-CPPC bound structures. The length of these trajectories was within the timescale of loop motion and thus appropriate to examine the dynamic profile of the C-terminal [114]. In agreement with the crystallographic observations, RMSF analyses of the apo D-DT and D-DT/4-CPPC structures showed noticeable differences, which are associated with 4-CPPC binding. Upon ligand binding, the C-terminal residues Lys109-Val113 are moderately affected whilst Met114-Leu117 demonstrate a sharp increase of their dynamic activity (**Fig. 1**). A careful examination of the RMSF plots reveals that 4-CPPC has only a local dynamic effect, perturbing the C-terminal and to a lesser extent the α2/β4 loop (Val66-Gly68). Both regions are proximal to the 4-CPPC

binding site. The average RMSF values of 0.70±0.3 Å and 0.76±0.6 Å for the corresponding structures of apo D-DT (black dash line) and D-DT/4-CPPC, further support that 4-CPPC does not alter the global dynamic profile of D-DT (**Fig. 1**). Interestingly, the previously reported RMSF value of apo MIF, which was determined at 0.90 Å by 1μs MD simulations, is similar with the RMSF value of apo D-DT, reported herein [51].

Figure 4.1.

Root mean square fluctuation (RMSF) of apo D-DT and D-DT/4-CPPC.



Note. The average RMSF values of apo D-DT (0.70±0.3 Å) and D-DT/4-CPPC (0.76±0.6 Å) are shown as black and red dashed lines, respectively. Snapshots of the D-DT/4-CPPC structure, at 100 ns intervals, demonstrate the intense dynamic activity (black dashed circles) of the C-terminal. The C-terminal residues Met114-Leu117 are highlighted in red.

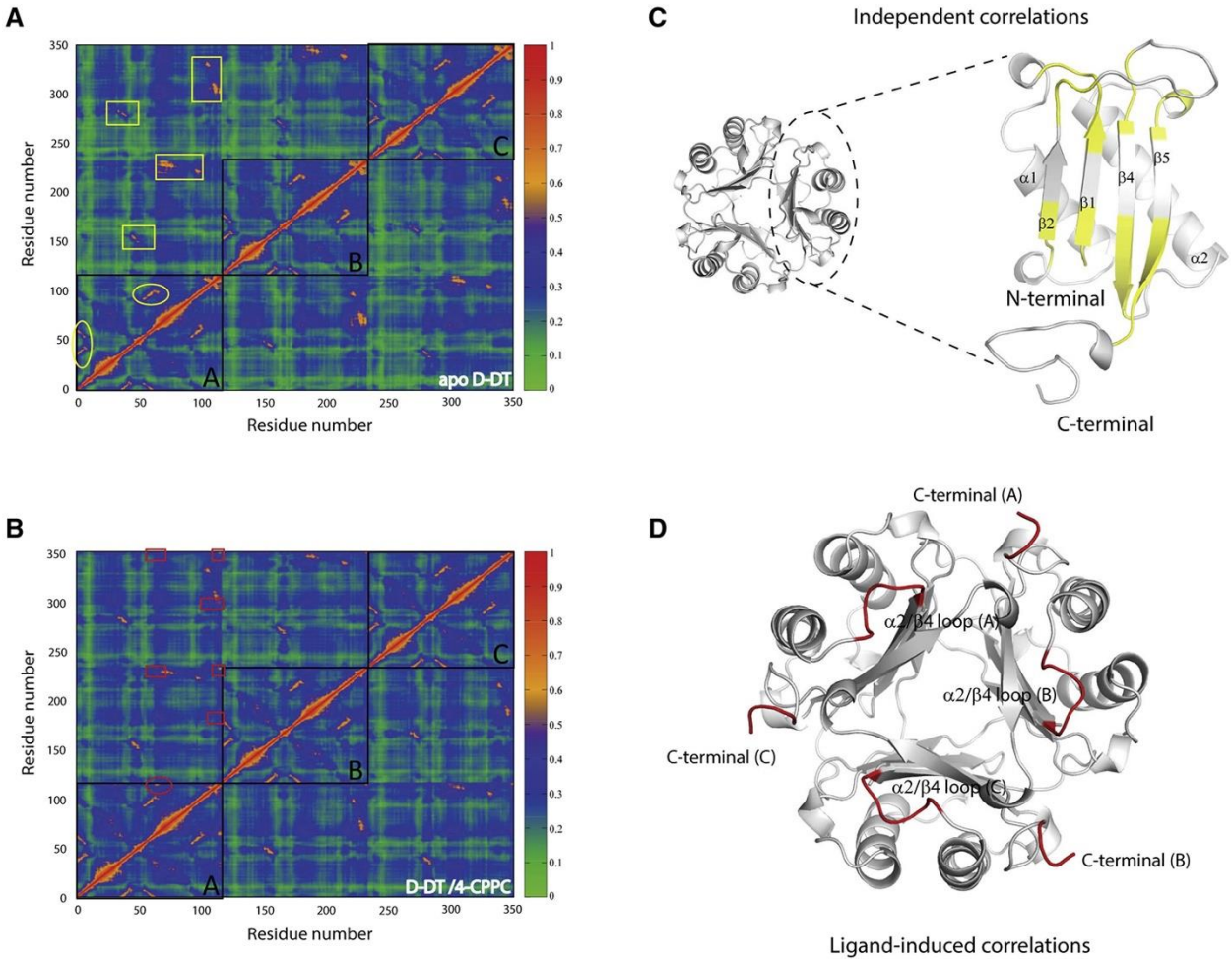
4-CPPC binding enables novel correlations across the D-DT trimer.

Correlation analysis is a tool employed to study physical motions of atoms within a protein, under different stimuli such as mutations, protein-protein interactions, or ligand binding. Through

this method, the dynamic characteristics of a protein are analyzed and information about correlated motions are extracted [120]. In the case of D-DT, the cross-correlation analyses of C_α fluctuations for apo D-DT and D-DT/4-CPPC were plotted and compared (**Fig. 2A, B**). Correlation values range between 0 - 1, with 0 and 1 indicating no correlation and absolute correlation, respectively. Absolute correlation is only observed when comparing a residue with itself. Large correlation values are expected when analyzing adjacent residues. However, such values are not useful for understanding protein dynamics because fluctuations of residue "X" are expected to influence the adjacent residues (X-1 and X+1). In contrast, the large correlation values between distal D-DT residues provide important structural and functional insights.

Figure 4.2

Correlation Analyses of Apo D-DT and D-DT/4-CPPC Structures.



Note. **(A-B)** Correlation of C_α atoms was derived from 1 μs MD simulations using the apo (PDB: 1DPT) and 4-CPPC bound (PDB: 6C5F) crystal structures of D-DT. Each black box (A, B, and C) represents the three monomers of D-DT. Correlation values vary between 0 - 1, indicating no correlation and absolute correlation, respectively. The Y and X axes show the 351 amino acids of the D-DT trimer (Pro1-Leu117). The correlation plots of **(A)** apo D-DT and **(B)** D-DT/4-CPPC demonstrate noticeable differences, which are associated with binding of 4-CPPC. Correlations within a single monomer (intrasubunit) or between monomers (intersubunit) are marked in ellipses

(Figure 4.2 Continued)

and boxes, respectively. The apo D-DT correlation plot shows the intra- and intersubunit correlations found in both structures. These correlations are marked in yellow **(A)** and called "independent" because they are not influenced by 4-CPPC binding. The intra- and intersubunit correlations associated with 4-CPPC binding **(B)** are shown in red. **(C)** The independent correlations (yellow) are illustrated on the D-DT monomer. **(D)** The ligand-induced correlations (red) are presented on the D-DT trimer.

The D-DT correlations within a single monomer (intrasubunit) or between monomers (intersubunit) are marked in ellipses and boxes, respectively (**Fig. 2A, B**). for clarity, independent and ligand-induced correlations are shown on different plots. The independent intra- and intersubunit correlations are marked in yellow, on the apo D-DT plot (**Fig. 2A**). Such correlations are called "independent" because they are formed in both apo D-DT and D-DT/4-CPPC structures, and thus are not influenced by 4-CPPC binding. On the other hand, the ligand-induced intra- and intersubunit correlations are enabled by 4-CPPC binding. These correlations are presented on the D-DT/4-CPPC plot (**Fig. 2B**).

Analysis of the independent correlations highlights the importance of the core β-sheet in coordinating communications across the D-DT trimer (**Fig. 2C**). The N-terminal residues 1-3 form strong correlations with residues 37-39 and 61-63. At the same time, residues 61-63 are correlated with residues 99-101. A similar pattern of correlated motion was previously reported for MIF [51]. While the two proteins also possess similar global flexibility, D-DT$_{RMSF}$= 0.70 Å (reported herein) and MIF$_{RMSF}$= 0.90 Å [51], our findings demonstrate common dynamic features that can be utilized for further elucidating their functionalities. Additional correlations within the D-DT monomer involve the following: residues 6-8 with 56-57, 7-10 with 43-45, 55-57 with 93-95 and 63-64 with

101-103. All the reported intrasubunit correlations are illustrated on the D-DT monomer (**Fig. 2C**). As expected, the intersubunit correlations, which are marked in yellow boxes (**Fig. 2A**), are located at the three dimeric interfaces and are responsible for subunit-subunit communications.

Binding of 4-CPPC in the active site of D-DT enables new intra- and intersubunit communications, which are highlighted on the D-DT/4-CPPC plot (**Fig. 2B**). The ligand-induced intrasubunit correlation, marked in the ellipse, associates residues 64-68 with 115-117. Residues 64-68 are located on the α2/β4 loop, whilst residues 115-117 are found at the C-terminal (**Fig. 2D**). Intersubunit correlations between these two regions were noted across the D-DT trimer. At the same time, the C-terminal residues 115-117 of subunit A form strong correlations with the C-terminal residues 115-117 from subunits B and C. These intra- and intersubunit communications are all shown in the biological assembly of D-DT (**Fig. 2D**).

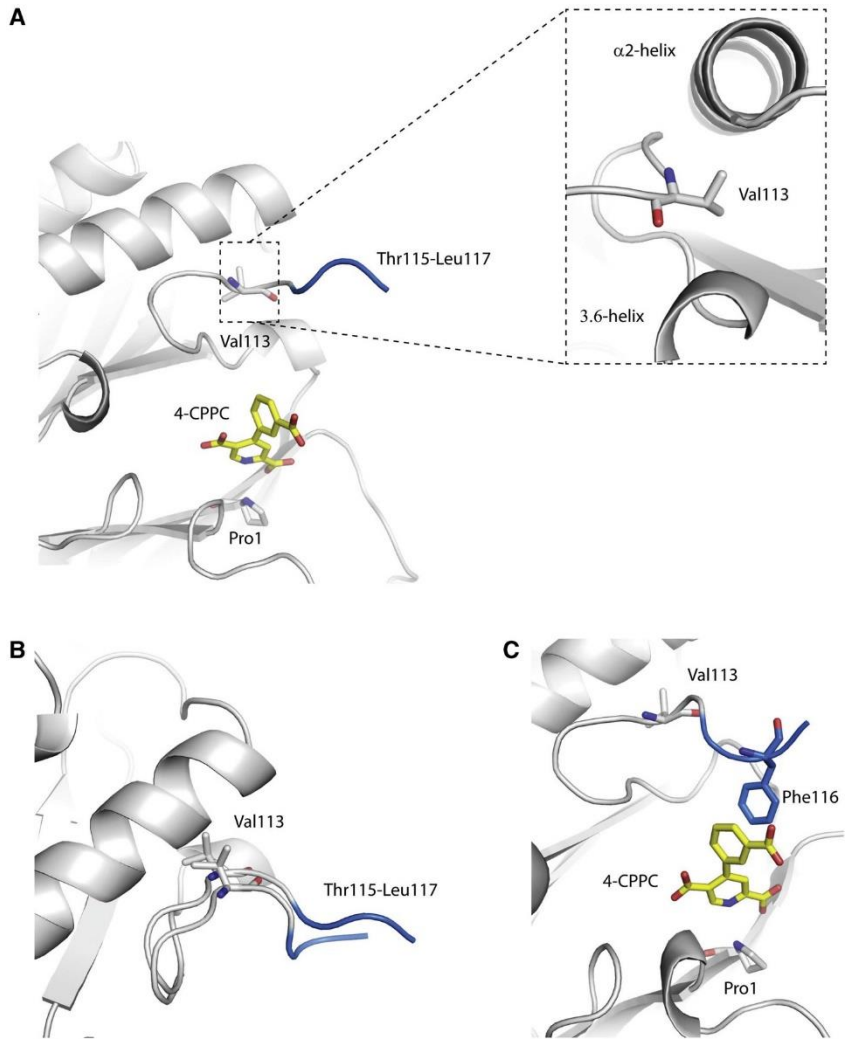
Interactions between the C-terminal and 4-CPPC trigger the formation of new intra- and intersubunit communications.

Structural analysis of D-DT demonstrates the ligand-induced correlated regions are found at distances beyond the physical limits of direct communication. for example, the C-terminal residues 115-117 of chain A are found >30 Å away from the corresponding C-terminal segments of chains B and C. Similarly, residues 64-68 and 115-117 of subunit A, are located >11 Å apart. To explain these findings, we closely examined the MD simulations of D-DT/4-CPPC. At the beginning of the simulations, the C-terminal is largely flexible and Thr115-Leu117 are extended out into the solvent unable to form interactions with either the protein or 4-CPPC (**Fig. 3A**). At the same time, Val113 is hidden in a hydrophobic pocket formed between α2-helix and β5/C-terminal 3.6 helix. In that pocket, Val113 possesses the same orientation as seen in the crystal structures of apo D-DT and D-DT/4-CPPC (**Fig. 3A**). As the simulations progress and the C-terminal retains

its high flexibility, Val113 is displaced from the hydrophobic pocket (**Fig. 3B**). This leads to a C-terminal conformational change that favors the formation of an aromatic-aromatic interaction between Phe116 and the benzoic moiety of 4-CPPC (**Fig. 3C**). This interaction enables communication between the C-terminal and α2/β4 loop of the same monomer. In turn, intersubunit communications between the three C-termini are enabled via the adjacent α2-helix. Residues 64-68 of the α2/β4 loop, as well as the α2-helix, play a key role in the intersubunit correlations. Via this pathway, the ligand induces novel intra- and intersubunit communications across the biological assembly of D-DT.

Figure 4.3.

Intra- and cross-subunit communication are triggered via interactions between the C-terminal and 4-CPPC.



Note. **(A**) Upon binding of 4-CPPC (yellow sticks) the C-terminal adopts a conformation that exposes The115-Leu117 (blue) into the solvent. Val113 (grey sticks) is hidden in a hydrophobic pocket, formed between the α2-helix and β5/C-terminal 3.6 helix. **(B)** As the simulations progress, Val113 is displaced from the hydrophobic pocket, allowing C-terminal residues Thr115-Leu117 to adopt a new conformation. **(C)** Due to the new conformation of the C-terminal, 4-CPPC forms

(Figure 4.3 Continued)

aromatic-aromatic interactions with Phe116 (blue sticks) and enables several long-range communications across the D-DT trimer.

Experimental interrogation of the D-DT/4-CPPC complex confirms the computational findings.

To experimentally verify the computational findings, we employed solution NMR; a method that is regularly paired with MD simulations to examine protein dynamics under various stimuli. As part of this effort, we first synthesized 4-CPPC via a two-step pathway (**Scheme S1**). The structures o 4-CPPC and its biaryl precursor **1** were confirmed by ^{1}H and ^{13}C NMR (**Fig. S1**).

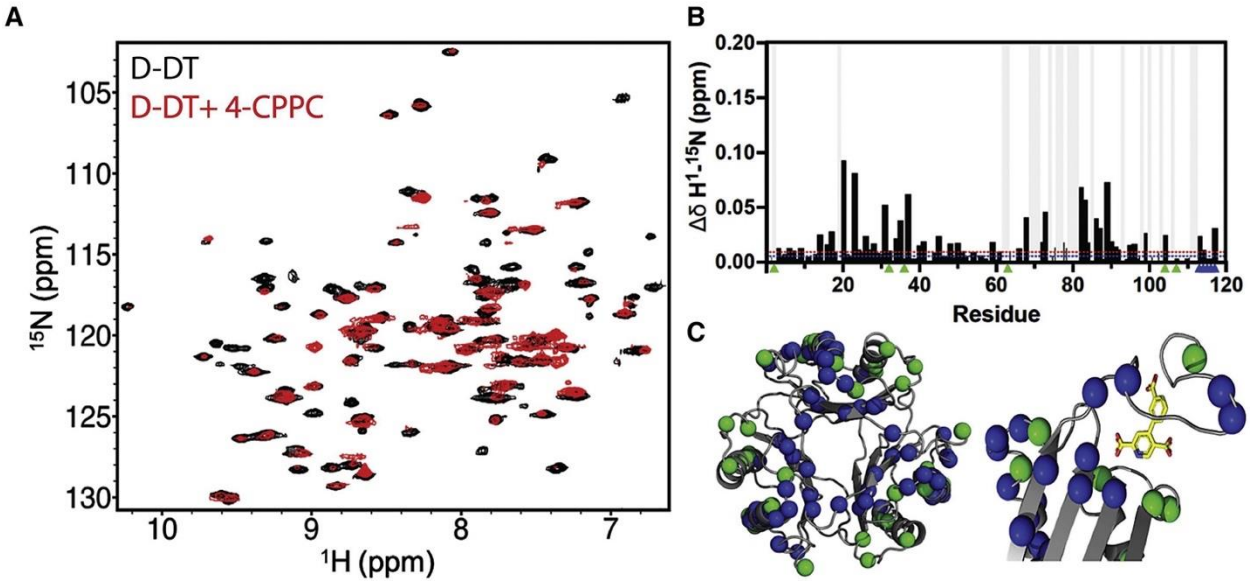
We performed ITC and kinetic experiments to confirm that the newly synthesized molecule binds D-DT as a competitive inhibitor. for the ITC experiments, the heat emitted from the ligand-to-buffer titration was used to correct our calorimetric data. The corrected data points of 4-CPPC-to-protein titration were fitted in the independent model, providing a K_D of 51 μM (**Fig. S2**). The binding mode of 4-CPPC was then examined using the keto/enol enzymatic assay of D-DT. Our kinetic findings were in agreement with the previously published data, providing a competitive type of inhibition with an inhibition constant (K_i) of 38 μM (**Fig. S3**) [49].

To assist solubility of 4-CPPC and minimize NMR spectral artifacts during titration of D-DT with 4-CPPC, a background of 15% (*v*/*v*) DMSO-d_6 was used without significantly altering the D-DT structure (**Fig. S4**). ^{1}H ^{15}N TROSY-HSQC spectral overlays of apo D-DT with 0% vs 15% DMSO-d_6 showed only modest chemical shift perturbations in a consistent spectral pattern and direction, localized mainly to flexible loop regions. At first glance, the structural stability of D-DT in 15% DMSO-d_6 may be unexpected. However, precipitation of a macromolecule in

various organic solvents is highly influenced by the protein's amino acid composition, pH, and polarity [121]. In the case of D-DT, at low concentrations of DMSO, the protein has a higher affinity for water than DMSO, supporting its retention of natural structure. Protein precipitation and denaturation analyses of ribonuclease, lysozyme, β-lactoglobulin, and chymotrypsinogen have shown that these proteins retain their natural structures at DMSO concentrations higher than 15% [121].

Figure 4.4.

Solution interactions of 4-CPPC with D-DT.



Note. **(A)** ^{1}H-^{15}N HSQC NMR spectral overlay of apo D-DT (black) with 4-CPPC-bound D-DT (red). Sample conditions were 0.5 mM D-DT in 20 mM Tris, 20 mM NaCl and 15% (*v*/*v*) DMSO-d_6 at pH 7.4. **(B)** NMR chemical shift perturbations caused by 4-CPCC. Gray vertical bars highlight sites of NMR line broadening. Blue and red dashed lines denote the 10% trimmed mean +1.5σ of all shifts, respectively. Green triangles correspond to the 4-CPPC-interacting residues, while blue triangles correspond to residues that lack electron density in the 4-CPPC-bound X-ray crystal

(Figure 4.4 Continued)

structure [49]. **(C)** Chemical shifts greater than 1.5σ of the 10% trimmed mean (green spheres) are mapped onto the D-DT trimer and monomer structures (PDB: 7MSE), while blue spheres indicate sites of line broadening. The 4-CPPC ligand is also shown with the D-DT monomer, using PDB: 6C5F as the reference for alignment.

According to the K_D value of 4-CPPC for D-DT (**Fig. S2**), we designed two titration experiments. The first experiment was performed with the ligand and protein at 1:1 molar ratio (**Fig. 4A**) and the second with the ligand in a five-fold excess (**Fig. S5A**). for the first titration experiment, we noticed that 4-CPPC induced many changes in the C-terminal face of the protein, including moderate chemical shift perturbations and a large number of line broadened peaks, which is a qualitative signature of millisecond dynamics (**Fig. 4B, 4C**). Chemical shift perturbations map to both α-helices and loop regions, while line broadening localizes mostly around the α2 helix, residues spanning the opening of the solvent cavity, and the C-terminal residues 111 and 112, which immediately precede the site involved in the induced fit of 4-CPPC [49]. Residues known to be involved in the D-DT/4-CPPC interaction, or that become flexible upon binding of D-DT to 4-CPPC, are structurally or dynamically affected in the NMR experiments. Most notably, residues surrounding the 4-CPPC binding pocket become highly flexible, consistent with the MD simulations.

The second titration experiment was performed having the ligand in a high excess. Under these conditions, 4-CPPC has a drastic impact on the global flexibility of D-DT (**Fig. S5A, S5B**). Regions located distant from the binding site of 4-CPPC demonstrate elevated dynamic activity, which is consistent with the formation of intersubunit communications. While it is difficult to

distinguish intra- and intersubunit communication in a symmetric molecule by NMR, the locations and types of the NMR-detectable perturbations provide some clues. NMR revealed chemical shift perturbations and extensive line broadening around the α2 helix and β4 sheet at the monomer-monomer interface, suggestive of intersubunit crosstalk. Likewise, most of the line broadened resonances surround the D-DT solvent channel, which prior studies of both MIF and D-DT demonstrated to facilitate intermonomer communication [52, 53]. Other portions of D-DT, namely the α1 helix and loops decorating the periphery of the structure, display minimal evidence of line broadening and local chemical shift perturbations. These sites are also distal to the subunit interfaces, suggesting only intrasubunit effects. Given the low binding affinity of 4-CPPC, the chemical shift perturbation differences between the two titration experiments are attributed to the saturation level of D-DT. At stoichiometric quantities (1:1 molar ratio), the three active sites are not fully occupied and thus the ligand has a moderate impact on the dynamics of the trimer. On the other hand, at a five-fold excess all active sites are occupied to enable displacement of all three C-termini, which in turn enables the global communication between the three subunits. The limited solubility of 4-CPPC at NMR-level concentrations in >80% aqueous buffer precluded longer spin relaxation experiments to quantitate rates of conformational exchange and altered structural populations. Future work will focus on optimizing 4-CPPC solubility, perhaps through chemical modification, to facilitate further NMR studies.

Conclusions

In the absence of large-amplitude collective displacements, such as domain movements, proteins mediate fundamental biochemical processes via correlated motions of residues. Correlated motion is a fundamental property of the MIF β-sheet [122] that is used to finely tune at least two of the protein's biological activities [51, 53, 91]. Computational and experimental evidence demonstrate

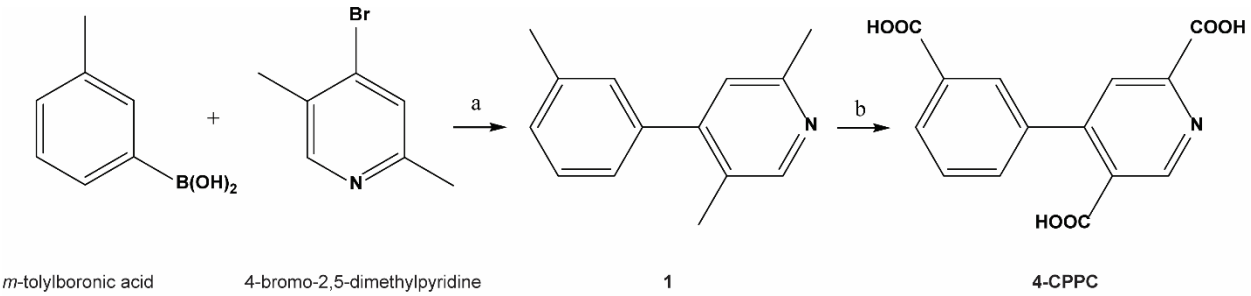
that MIF can execute a fascinating variety of functions without undergoing any major conformational changes. Instead, local flexibility is sufficient to modulate its functionality [51, 53].

In contrast to MIF, D-DT reveals an unexpected mobility of its C-terminal. Conformational changes of the C-terminal resulting from phosphorylation [123], ion [124] or ligand [125] binding, or protein-protein interactions [126], are known to regulate a variety of biological tasks, the most notable being channel gating [127], signaling [128] and catalysis [125]. for D-DT, the biological significance of C-terminal flexibility is enigmatic. A previous study comparing the D-DT/CD74 and MIF/CD74 interaction features, showed that D-DT has a higher binding rate to CD74 but also a faster dissociation, in comparison to MIF [9]. These findings suggest that the D-DT-induced activation mechanism of CD74 may be noticeably different from the corresponding mechanism of MIF, involving distinct binding surfaces and/or conformations. Further studies on D-DT's C-terminal are required to elucidate the potential functionality of this site. Such studies may include deletion(s) and/or single point mutation(s) of the residues (Met114-Leu117) previously shown to have enhanced flexibility upon 4-CPPC binding. The structural alterations to the C-terminal as a consequence of these modifications may increase or reduce the affinity to a substrate/inhibitor or even point the importance of C-terminal for activation of CD74. Collectively, the findings presented herein offer a fresh perspective for exploring and understanding the D-DT-induced activation mechanism of CD74.

Supporting Information

Scheme 4.S1.

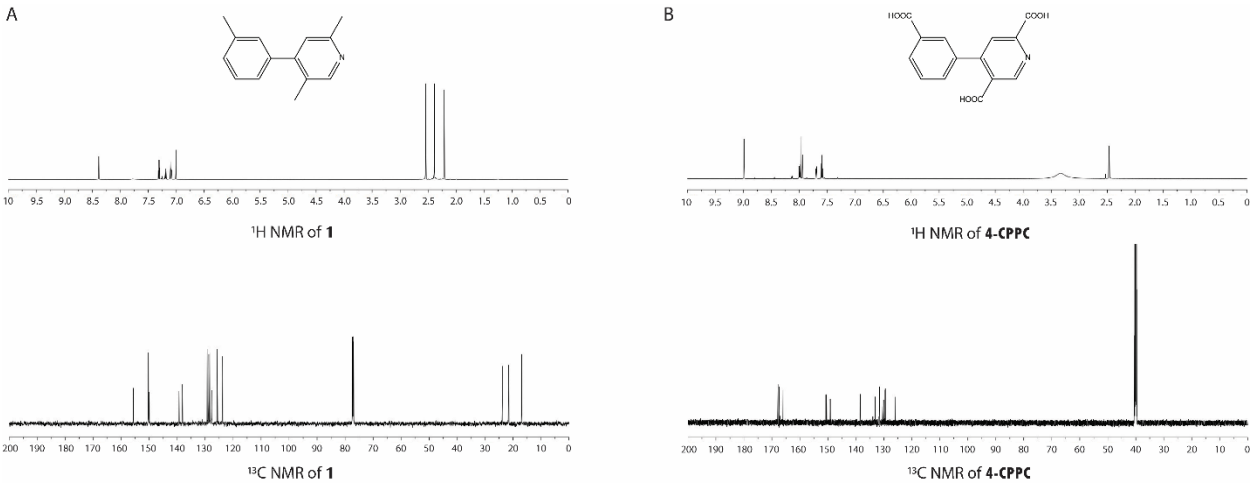
The synthetic pathway of 4-CPPC.



Note. Reagents and conditions: a) 10% Pd/C, $Na_3PO_4 \bullet 12H_2O$, 55% iPrOH in water, 80 °C, 12 h, under Argon, 53%; b) $KMnO_4$, H_2O, reflux, 4 h followed by Na_2SO_3, 42%.

Figure 4.S1.

NMR analysis of 1 and 4-CPPC.

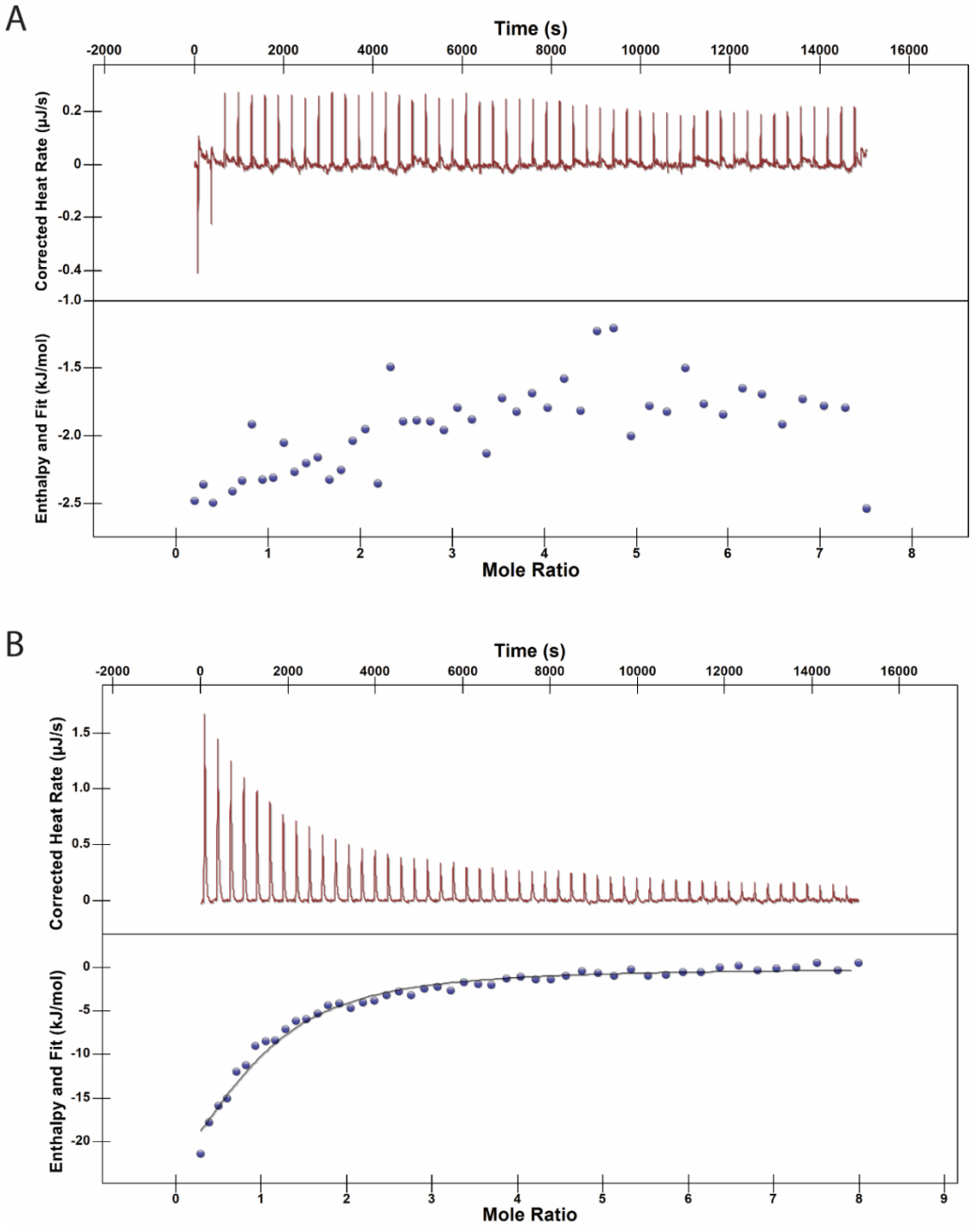


(Figure 4.S1 Continued)

Note. The structures of **(A) 1** and **(B) 4-CPPC** were confirmed by the recorded ^{1}H-NMR (top) and ^{13}C-NMR (bottom) spectra.

Figure 4.S2.

Representative ITC profiles for the titration of (A) 4-CPPC to buffer, and (B) 4-CPPC to D-DT.



Note. Both experiments were carried out at 25°C. A detailed analysis of the experimental conditions is provided in methods.

Figure 4.S3.

Dixon plot for competitive inhibition of D-DT by 4-CPPC

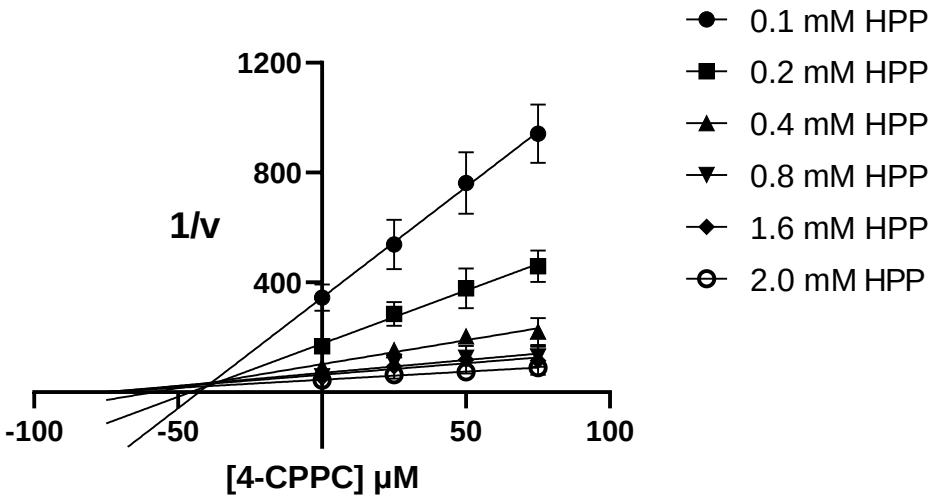
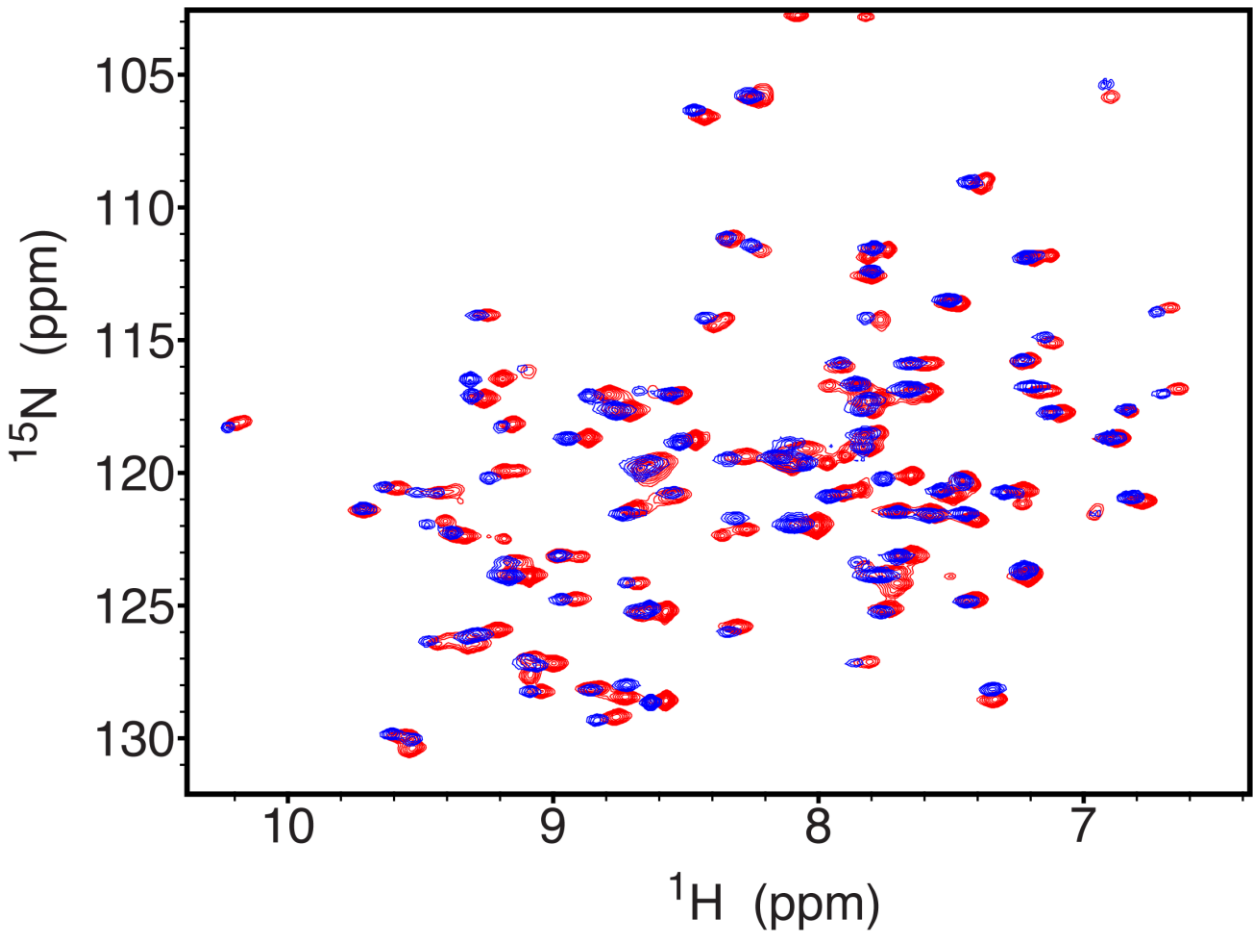


Figure 4.S4.

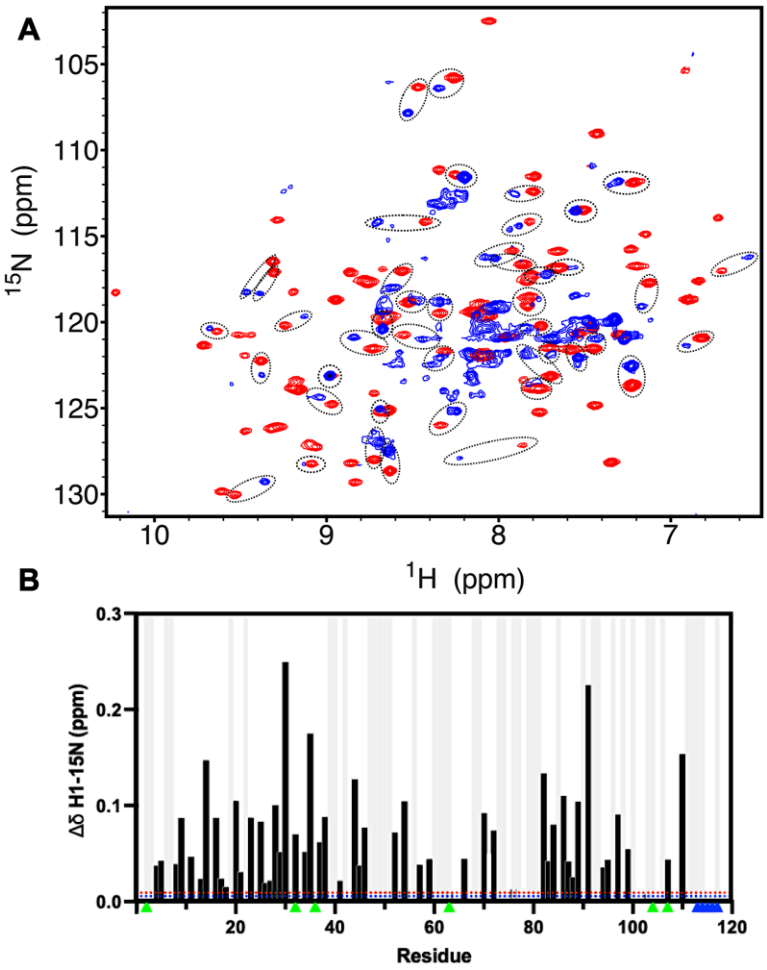
^{1}H-^{15}N HSQC NMR spectral overlay of apo D-DT in the absence (red) and presence (blue) of 15% DMSO.



Note. The addition of DMSO to the NMR sample buffer, which significantly aided the solubility of the 4-CPPC ligand, did not affect the overall distribution of chemical shifts in the NMR fingerprint.

Figure 4.S5.

Solution interactions of 4-CPPC with D-DT at 5:1 molar ratio



Note. **(A)** ^{1}H-^{15}N HSQC NMR spectral overlay of apo D-DT (red) with D-DT/4-CPPC (blue). Sample conditions were 0.5 mM D-DT in 20 mM Tris, 20 mM NaCl, 15% (*v*/*v*) DMSO-d_6, pH 7.4. Large chemical shift perturbations (dashed circles) and high degrees of line broadening are observed. **(B)** NMR chemical shift perturbations caused by a larger excess of 4-CPCC. Gray vertical bars highlight sites of NMR line broadening. Blue and red dashed lines denote the 10% trimmed mean +1.5σ of all shifts, respectively. Green triangles correspond to the 4-CPPC-

(Figure 4.S5 Continued)

interacting residues, while blue triangles correspond to residues that are missing electron density in the 4-CPPC-bound X-ray structure.

CHAPTER 5: 2,5-PYRIDINEDICARBOXYLIC ACID IS A BIOACTIVE AND HIGHLY SELECTIVE INHIBITOR OF D-DOPACHROME TAUTOMERASE

Published as: Parkins A, Das P, Prahaladan V, Rangel VM, Xue L, Sankaran B, Bhandari V, Pantouris G. 2,5-Pyridinedicarboxylic acid is a bioactive and highly selective inhibitor of D-dopachrome tautomerase. Structure. 2023 Mar 2;31(3):355-367.e4. doi: 10.1016/j.str.2023.01.008. Epub 2023 Feb 17. PMID: 36805127.

Abstract

Macrophage migration inhibitory factor (MIF) and D-dopachrome tautomerase (D-DT) are two pleotropic cytokines, which are coexpressed in various cell types to activate the cell surface receptor CD74. Via the MIF/CD74 and D-DT/CD74 axes, the two proteins exhibit either beneficial or deleterious effect on human diseases. In this study, we report the identification of 2,5-pyridinedicarboxylic acid (a.k.a. **1**) that effectively blocks the D-DT induced activation of CD74 and demonstrates an impressive 79-fold selectivity for D-DT over MIF. Crystallographic characterization of D-DT-**1** elucidates the binding features of **1** and reveals previously unrecognized differences between the MIF and D-DT active sites that explain the ligand's functional selectivity. The commercial availability, low cost, and high selectivity make **1** the ideal tool for studying the pathophysiological functionality of D-DT in disease models. At the same time, our comprehensive biochemical, computational, and crystallographic analyses serve as a guide for generating highly potent and selective D-DT inhibitors.

Introduction

Understanding the correlation between protein expression and disease progression is laborious as it requires in depth investigation of mechanistic pathway(s) using a well-designed experimental study. If the same activity is shared among multiple proteins, then the task is even more challenging and demands the use of selective modulators. A representative example of such proteins is MIF and D-DT; the two human members of the MIF superfamily that are coexpressed in various cell types and tissues and share a cytokine [9, 27], tautomerase [85, 129], and endonuclease [13, 15] activity.

MIF was originally reported in the 1960s as an "inhibitory factor" of random macrophage migration produced by T-cells [11], but was later on also detected in a variety of immune cells and nearly every tissue [130]. The MIF signal transduction network is complicated and involves biological partners found on the cell surface as well as intracellularly. The most well-studied protein partners are the CD74/CD44 complex [27, 36], which serves as the canonical MIF receptor, and non-canonical chemokine receptors CXCR2, CXCR4, and CXCR7 [12, 131]. The MIF induced activation of CD74/CD44, chemokine receptors, or the CD74/CXCR heterogenic complexes (CD74/CXCR2 [12], CD74/CXCR4 [132], and CD74/CXCR4/CXCR7 [131]) occurs in either an autocrine or paracrine fashion, depending on the stimulus and cell type. Subsequently, the mitogen-activated protein kinase (MAPK)/extracellular signal-regulated kinase (ERK), phosphoinositide 3-kinase (PI3K)-protein kinase B (AKT), and/or AMP-activated protein kinase (AMPK) signaling pathways are activated with downstream effects that regulate cell growth, proliferation, and inhibition of apoptosis. Under pathological conditions, these signal transduction events will advance a variety of disorders, including inflammatory diseases [82] and cancer [133]. The expression level of MIF in these conditions is highly influenced by the gene's promoter polymorphisms. for

example, a 5-8 repeated sequence of CATT tetranucleotide at position -794 of the human *MIF* gene has been associated with disease severity for patients with rheumatoid arthritis [134]. Besides the expression levels of MIF, recent clinical findings obtained from patients with inflammatory diseases [135] and cancer [136] have shown that the pathogenic action of MIF is also related to a unique structural isoform of the protein found only in oxidized environment. The oxidized structural isoform of MIF [88], also known as oxMIF, is exclusively detected in inflammation and cancer and serves as an attractive molecular target and diagnostic tool for these pathological conditions.

Structurally, MIF is composed of only 114 amino acids forming a homotrimeric biological assembly. The MIF trimer possesses a solvent channel at the center, and an enzymatic pocket in the interface of two monomers [17]. Pro1, which is located at the edge of this pocket, is responsible for the tautomerase activity of MIF. Via this activity, a large number of MIF modulators have been discovered and characterized including competitive [50], non-competitive [21], and covalent inhibitors [48]. Among the numerous MIF inhibitors reported thus far, only 4-iodo-6-phenylpyrimidine (4-IPP) is known to inhibit the activity of D-DT, by creating a covalent adduct with the catalytic residue Pro1 [57].

D-DT is the second human member of the MIF superfamily that shares a high structural homology [90] yet a low sequence identity (only 34%) with MIF [9]. Although D-DT was cloned and crystallized in the 1990s [90], its functionality has not been extensively studied due to the belief that it is the product of *MIF* gene duplication [19]. Nevertheless, growing evidence has shown that D-DT expression has a non-redundant functionality in inflammation, autoimmune diseases, cardiovascular disorders, and cancer [137]. In some pathological conditions D-DT may act cooperatively/synergistically with MIF, while in others the two proteins have different or even oppositional activities. Cooperative action of MIF and D-DT has been reported in various types of

cancer [138], inflammatory disorders including sepsis [9] and chronic obstructive pulmonary disease [139], and heart conditions such as cardiac ischemia/reperfusion injury [140, 141]. In contrast, studies on adipose tissue inflammation [142] and systemic sclerosis [143] demonstrate that the expression of D-DT and MIF serve two distinct functional roles.

Whereas the MIF functionality is associated with activation of multiple receptors and receptor complexes, D-DT is mainly mediated by the CD74/CD44 receptor complex [9]. The D-DT/CD74 interface has yet to be defined, but previous results have shown that a D-DT tautomerase inhibitor can effectively block activation of the receptor [59]. Upon the formation of the homotrimeric assembly, the active site pocket of D-DT is positioned in the interface of two monomers, similar to MIF [90]. Despite sharing the tautomerase activity, MIF and D-DT have noticeable differences in the amino acid composition of their active sites. Consequently, the large number of MIF inhibitors are ineffective on D-DT and only two reversible inhibitors are so far known for the latter protein [44, 59]. The first inhibitor, called 4-(3-carboxyphenyl)-2,5-pyridinedicarboxylic acid (4-CPPC), was identified recently by an *in-silico* screen of 1.6 million compounds [59]. The second D-DT inhibitor, (R)-5-Methyl-3-(1-(naphthalen-1-yl)ethyl)-6-(3-(trifluoromethyl)phenyl)thieno[2,3-d]pyrimidine-2,4(1H,3H)-dione was discovered via the screening of 305 compounds from an in house collection [15]. In both cases, the inhibitors revealed micromolar inhibition potencies and low to moderate selectivity for D-DT over MIF.

In this study, we describe the identification and characterization of 2,5-pyridinedicarboxylic acid (**1**) as a commercially available, bioactive, and highly selective inhibitor of D-DT. Enzymatic assays demonstrate that **1** has a 79-fold inhibition selectivity for D-DT over MIF. Notably, this compound has the highest selectivity among the known D-DT inhibitors. Crystallographic analysis of D-DT-**1** and side-by-side comparison with known MIF crystal

structures expose key structural differences between the MIF and D-DT active sites that were previously unseen. These new findings provide insights into the regulation of ligand binding by MIF and D-DT and clarify the inhibition selectivity criteria. The identification of **1** would aid the efforts towards understanding the biological activity of D-DT in disease models and clinical samples as well as the development of new generation D-DT modulators that target activation of CD74.

Materials and Methods

Table 5.1

Key Resources Table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and Virus Strains		
BL21 (DE3) E. coli/pET-11b (human WT MIF)	This study	N/A
BL21 (DE3) E. coli/pET-22b (human WT D-DT)	This study	N/A
Chemicals, Peptides, and Recombinant Proteins		
2,5-Pyridinedicarboxcylic acid (>98% purity)	Acros Organics	Cat# 131870050
4-CPPC (99.4% purity)	Axon Medchem	Cat# 3189
4-hydroxyphenylatepyruvate (>95% purity)	TCI	Cat# H0294
4-Phenylpyridine (>98% purity)	TCI	Cat# P0162
Ammonium Acetate (97.0 % purity)	VWR	Cat# BDH9204
Ampicillin Sodium Salt	VWR	Cat# 97061-442
Boric Acid (≥99.5% purity)	VWR	Cat# BDH9222
DMSO (≥99.9% purity)	VWR	Cat# 97061-250
Isopropyl b-D-1-thiogalactopyranoside (IPTG)	Bioland Scientific	Cat# CI01
Luria Broth (LB)	VWR	Cat# 97064-114
Pyridine-2-carboxylic acid (>99% purity)	TCI	Cat# P0421
Nicotinic acid (99.5% purity)	Acros Organics	Cat# 128290050
Terephthalic acid (>99% purity)	TCI	Cat# T0166
Triton X-114 (proteomics grade)	VWR	Cat#M144-1L
WT D-DT	This study	N/A
WT MIF	This study	N/A
Critical Commercial Assays		
Pierce BCA Protein Assay Kit	Thermo-Fisher	Cat# 23225
Hema III Staining Solutions	Thermo-Fischer	Cat# 122911

(Table 5.1 Continued)

Deposited Data		
WT MIF	Crichlow et al., 2009	PDB: 3DJH
WT DDT	Sugimoto et al., 1999	PDB: 1DPT
D-DT-1	This Paper	PDB: 8DBB
Recombinant DNA		
pET11b- WT MIF	This study	N/A
pET22b-WT D-DT	This study	N/A
Experimental Models: Cell Lines		
BL21(DE3) Competent Cells	Agilent Technologies	Cat# 200131
Experimental Models: Organisms		
C57BL6/J male mice	Jackson Laboratory	RRID: IMSR_JAX:000664
Software and Algorithms		
Aimless	CCP4	https://www.ccp4.ac.uk/html/aimless.html
CHARMM36 Force Field	University of Maryland School of Pharmacy	http://mackerell.umaryland.edu/charmm_ff.shtml
COOT	MRC Laboratory of Molecular Biology (LMB)	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/
CHARMM General Force Field (CGenFF)	University of Maryland	https://cgenff.umaryland.edu/
GROMACS	Abraham et al. (2015)	https://www.gromacs.org/index.html
G_Correlation	O. Lange et al. (2005)	https://www.mpinat.mpg.de/grubmueller/g_correlation
MATLAB	MathWorks	https://www.mathworks.com/products/matlab.html
MolGpKa	Xundrug	https://xundrug.cn/molgpka
Molrep	CCP4	https://www.ccp4.ac.uk/html/molrep.html
Nanoanalyze 3.11.0	TA Instruments	https://www.tainstruments.com/
Nanoscale Molecular Dynamics	University of Illinois at Urbana-Champaign	http://www.ks.uiuc.edu/Research/namd/
PRODRG	CCP4	https://www.ccp4.ac.uk/html/cprodrg.html
Prism	Graphpad	https://www.graphpad.com/scientific-software/prism/
PyMOL	DeLano Scientific LLC	http://www.pymol.org/

(Table 5.1 Continued)

PyVOL	Smith et al. (2019)	https://schlessinger-lab.github.io/pyvol/index.html
Refmac5	CCP4	https://www.ccp4.ac.uk/html/refmac5/description.html
SUPERPOSE	CCP4	http://www.ccp4.ac.uk/html/superpose.html
Visual Molecular Dynamics	University of Illinois at Urbana-Champaign	http://www.ks.uiuc.edu/Research/vmd/
Origin 2019b	OriginLab	https://www.originlab.com/
XDS	Kabsch (2010)	https://xds.mr.mpg.de/
Other		
Advanced Light Source beamline 8.2.1	Lawrence Livermore National Labs	https://als.lbl.gov/
Circular dichroism spectropolarimeter	Jasco Inc.	Model# J-810
Affinity ITC calorimeter	TA Instruments	Model# Affinity ITC
Microplate reader	TECAN	Model# Infinite M-Plex

Experimental Model and Subject Details

Cell lines

Recombinant MIF and D-DT proteins were produced and purified from bacterial cells, E. coli strain BL21-(DE3) (Agilent Technologies).

Organisms

Activation of CD74 by D-DT or D-DT-1 complex was evaluated *in vivo* using 10–12 weeks old C57BL6/J male mice (Jackson Laboratory - Bar Harbor, ME). The mice were housed in a pathogen-free animal facility, at Cooper University Hospital, Camden, NJ, USA (IACUC protocol approved by Cooper University Hospital, Camden, NJ, USA).

Method Details

Protein Expression and Purification.

MIF and D-DT were expressed and purified as previously described [49, 66]. Briefly, pET11b and pET22b plasmids expressing MIF and D-DT, respectively, were transformed into BL21 (DE3)

competent cells according to the heat shock method. The cells were grown in LB broth enriched with 100 μg/mL ampicillin, at the optimal temperature of 37°C. When the solution's optical density at 600 nm became 0.6-0.8, the cell cultures were induced with 1 mM IPTG for 4 hours at 37°C. Following, the cells were collected by centrifugation, washed with the corresponding MIF or D-DT lysis buffer (20 mM Tris.HCl, 20 mM NaCl, pH 7.4 for MIF and 20 mM Tris.HCl, 20 mM NaCl, pH 8.5 for D-DT), and stored at -80°C. for protein purification, the cells were lysed via sonication, centrifuged to remove cell debris, and filtered with a 0.22 μm PES syringe filter (Bioland Scientific LLC). The MIF lysate was loaded onto a 5 ml Q-Sepharose column connected in series with a second 120 ml Q-Sepharose column. MIF did not bind to either column and eluted in the flow through with a ~95% purity. for D-DT purification, the lysate was loaded onto a 5mL Q-Sepharose column and eluted with 5% of 20 mM Tris.HCl, 1 M NaCl, pH 8.5. The two proteins were further purified by size exclusion chromatography using a 16/60 Superdex 75 column (Cytiva, Marlborough, MA). for this step, the lysis buffers of MIF and D-DT were used as corresponding running buffers. Protein purity and concentration were determined by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-Page) and bicinchoninic acid (BCA) protein assay, respectively.

Kinetic Experiments.

The multipoint kinetic experiments were carried out in a 96-well microplate according to previously published protocols [49, 53]. A 30 mM stock solution of 4-HPP was prepared in 0.5 M ammonium acetate pH 6.2 and the keto tautomer was formed after an overnight incubation at room temperature. This stock was used to create a working concentration gradient of 0 to 2 mM. Compound **1** was prepared in 100% DMSO and used at final concentration ranges of 0-50 μM and 0-3000 μM for D-DT and MIF, respectively. In all experiments, the final concentration of DMSO

was 1%. for the MIF assays, EDTA was also added at a final concentration of 1 mM. Borate was added at a final concentration of 0.420 M and the reaction was initiated by adding MIF or D-DT at a final concentration of 50 nM or 250 nM, respectively. The formation of enol-borate complex was monitored at 306 nm (ε_{306} = 11400 M^{-1} cm^{-1}) in a Tecan Infinite M-Plex microplate reader (TECAN). Each experiment was conducted for either 180 seconds (MIF) or 300 seconds (D-DT) and the data was analyzed in GraphPad Prism 9.4.1. for the single point kinetic experiments, 4-HPP and **1-5** were used at final concentrations of 1 mM and 100 μM, respectively, while the procedure was similar to what we described above. All kinetic experiments were performed in triplicate.

Isothermal calorimetry.

The ITC experiments were carried out in an affinity ITC calorimeter (TA instruments) with all samples being prepared in 20 mM phosphate buffer, pH 6.2 and 1% DMSO and degassed prior to each experiment. for the titration experiment, **1** was deposited into a 264 μL rotating syringe, while D-DT was placed in the isothermal sample cell. D-DT and **1** were tested at a final concentration of 500 μM and 100 μM, respectively. The control run utilized 20 mM phosphate buffer, pH 6.2 and 1% DMSO in the sample cell rather than protein. The contents of the syringe were then injected into the sample cell over 50 steps, each step releasing 3.5 μL. Each injection took place over 4 seconds with a 300-second delay. The syringe speed was set to 125 revolutions per minute (RPM) to ensure proper mixing of the samples. Before the first injection and subsequent data collection, the calorimeter was equilibrated for 1800 s. Data were analyzed using the NanoAnalyze software (version 3.11.0). The heat plot for D-DT-**1** was corrected against the Buffer-**1** heat plot and fitted in the independent model.

Circular dichroism spectroscopy.

Thermal denaturation experiments were performed in a Jasco J-810 circular dichroism spectropolarimeter following the protocol described before [91]. D-DT and **1** were prepared in 20 mM Tris.HCl, 20 mM NaCl, pH 7.4 at final concentrations of 20 μM and 0-200 μM, respectively. Before each experiment, all solutions were degassed with ultra-high purity grade nitrogen. The CD spectra of apo D-DT or D-DT-**1** were recorded from 20°C to 90°C at a rate of 0.5°C/min. The fraction of unfolded D-DT over temperature was analyzed in Origin 2019b.

Protein crystallography and data analysis.

Crystallization of D-DT-**1** was performed in a 24-well plate (Hampton research) using the hanging drop method. D-DT was concentrated to 12 mg/ml and mixed at 1:1 (v/v) ratio with well solutions containing 24-30% (w/v) PEG 4000, 0.1 M sodium citrate (pH 5.2-6.2), and 0.2 M ammonium sulfate. The triangular shaped crystals of apo D-DT were formed within 1-2 weeks and soaked into equilibrated drops of the mother liquor enriched with 100 mM of **1**. Due to the high concentration of PEG 4000, the crystals were flash frozen without a cryoprotectant. Single wavelength (1.00004 Å) diffraction trials were performed at Lawrence Berkeley National Laboratory using the Advanced Light Source (ALS) beamline 8.2.1. All diffraction trials were performed at 100 K and a complete data set was collected to 1.30 Å. Data reduction, including integration and scaling was accomplished using XDS [144] while final merging was achieved with aimless [145]. The crystal structure of D-DT-**1** was solved by molecular replacement using Molrep, a CCP4-supported program [72]. Apo D-DT (PDB entry: 1DPT) was used as the search structure, producing an initial model in a P3 space group. The model was further refined using Refmac5 [146] and COOT [147] and visualized in PyMOL [73]. Finally, the electron density of **1** was clearly seen in all active sites. The coordinates and crystallographic information files of **1** were generated by

PRODRG (CCP4-supported program). Structural alignment was accomplished using Superpose [72]. The final model has been deposited in worldwide Protein Data Bank with the identification number 8DBB.

Molecular Dynamics Simulations.

MD simulations of apo D-DT and D-DT-**1** were performed as previously reported [91]. The models used for the simulations were obtained from the crystal structures of apo D-DT (PDB entry: 1DPT) and D-DT-**1** (PDB entry: 8DBB). Using PyMOL, the protein's biological assembly was generated, and non-protein atoms were removed. A protein structure file (PSF) and PDB pair that included the missing hydrogens was generated utilizing the autoPSF plug-in of Visual Molecular Dynamics (VMD) program [69]. The PSF/PDB pair for D-DT was produced from CHARMM36 forcefield topology [68]. The ionization state of **1** was assigned using MolGpka [109], while its topology file was generated using CHARMM general force field (CGenFF) [118]. Solvated PSF and PDB were produced by the VMD solvate plugin with TIP3P waters. The system was then minimized and heated to 300K in a stepwise fashion. Before initiating the 200 ns run, the system was equilibrated at 300K for 1ns. Generalized cross-correlation analysis of Cα-atoms was performed using g_correlation [70], a GROMACS plug-in [71, 148]. RMSF analysis was performed by GROMACS. All simulations were run in triplicate.

Removal of bacterial lipopolysaccharides (LPS).

Prior the neutrophil recruitment assay, D-DT was treated for removal of bacterial LPS following a previously established protocol [149]. The procedure was initiated by mixing 30 mL of ice-cold LPS-free phosphate buffered saline (PBS) pH 7.40, with 200 μL D-DT, and 1% Triton X-114 (VWR life science). The mixture of PBS/D-DT/triton X-114 was subsequently incubated at 4^{o}C and 37^{o}C for 30 and 10 minutes, respectively. The first incubation step was performed under

gentle mixing while second incubation step was undisrupted. At the end of the 10-minute incubation, the mixture was transferred to eppendorf tubes and centrifuged to form two phases. The lower phase, which contained triton X-114 and LPS were disposed, while the upper phase (D-DT) was used for two additional rounds of purification. At the end of the third purification, D-DT was dialyzed for 16 hours in a 4 L LPS-free PBS (pH 7.40) solution. The next day, protein concentration was determined by the BCA assay and D-DT aliquots were stored in -80^{o}C.

Neutrophil Recruitment Assay.

Wild-type male mice (10–12 weeks old) of C57BL6/J genetic background were purchased from Jackson Laboratory (Bar Harbor, ME) and housed in a pathogen-free animal facility, at Cooper University Hospital, Camden, NJ, USA. Mice were administered a one-time intratracheal instillation of 50 µl PBS (vehicle) with or without LPS-free D-DT and/or **1**. The final concentration of D-DT and **1** was 20 µg/ml and 0-40 µM, respectively. Intratracheal instillation was performed following the methodology previously described [51] with some modifications. Six hours after the administration of experimental solutions, the mice were sacrificed to collect BAL fluid. The total protein content in the BAL fluid was measured by BCA assay. Approximately 1ml of the BAL fluid was pelleted by centrifugation at 1000 rpm for 10 min at 4°C. Following, the cell pellet was resuspended in 200 µl 1x PBS and centrifugation was repeated one more time at room temperature. The slides were air dried and stained with Hema III Differential Quick Stain. Neutrophils were counted manually in the smears to calculate their total percentage. Briefly, the cytosmear was randomly divided into 4-5 arbitrary areas that housed the maximum number of neutrophils. Two hundred different inflammatory cell types, including neutrophils, were counted in these areas from which the percentage of neutrophils was calculated. At least 4 mice were used for each experimental group. The animal study protocol was approved by the Institutional Animal Care and

Use Committee (IACUC) of Cooper University Hospital, Camden, NJ, USA. Statistical analysis was done with GraphPad Prism 9.4.1 by one-way ANOVA with Tukey's post-hoc correction.

Quantification and Statistical Analysis

All single point and kinetic experiments were carried out in triplicate (n=3) with the data being analyzed in ORIGIN 2019b. for both **Figures 1** and **2**, the error is represented as standard deviation. The D-DT-**1** crystal structure was refined using the R_{free} value along with a Ramachandran plot to fix angle outliers before submission to PDB/RCSB (https://www.rcsb.org). All statistics for the structure are present in **Table 1** and the "Full wwPDB X-ray Structure Validation Report", which contains a variety of structure metrics, is available in the RCSB database. All MD data was obtained in triplicate (n=3), root mean square fluctuation (RMSF) was averaged, presented with an error value obtained from the standard deviation, and plotted (**Figure S5**) using ORIGIN 2019b. In vivo studies utilizing mice had a minimum group size of 4 (n=4), the data was analyzed using GraphPad Prism 9.4.1 by one-way ANOVA with Tukey's post-hoc correction. The data is presented in **Figure 8** as the mean with standard error of the mean (SEM).

Results and Discussion

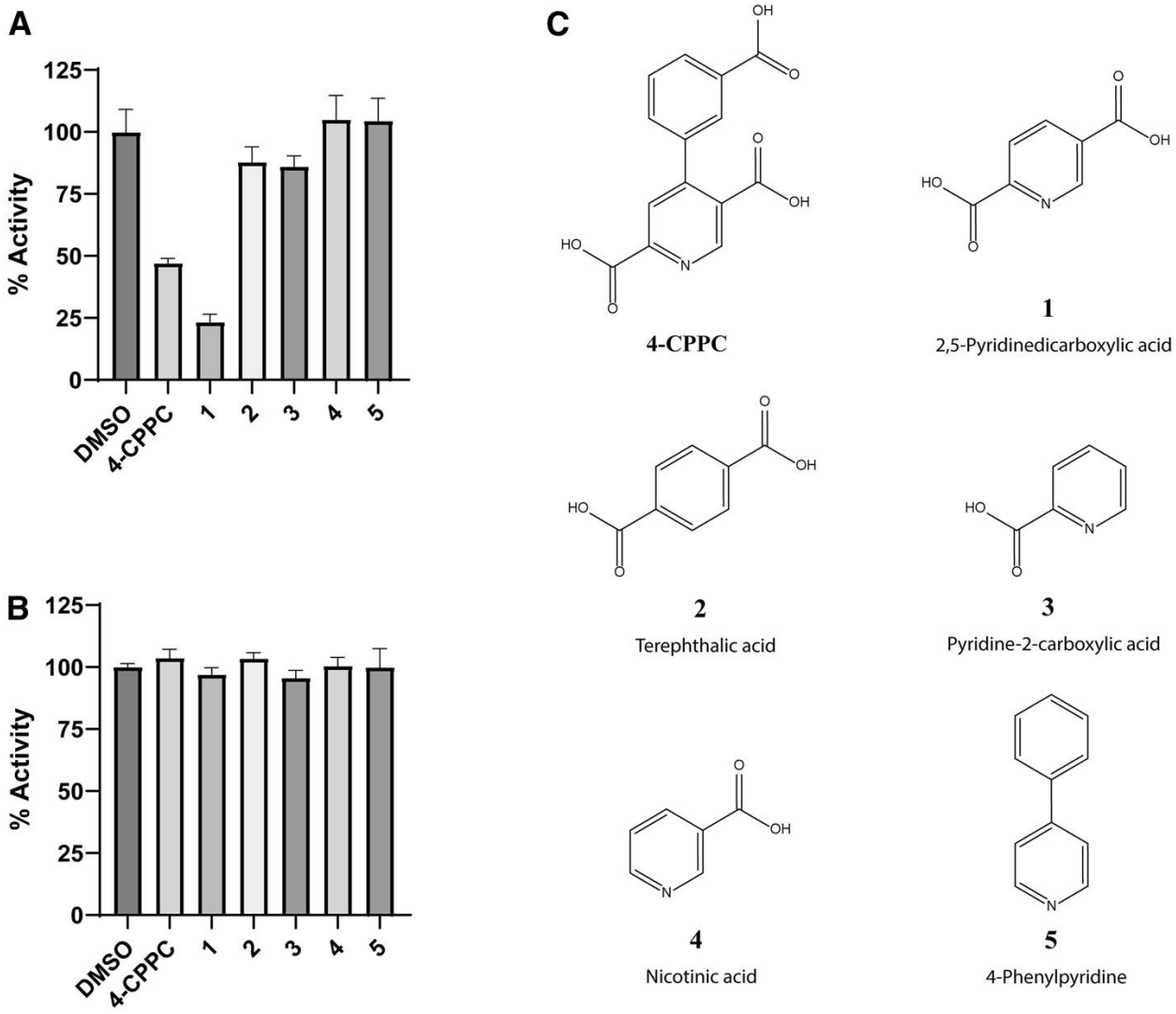
Identification of 1 as a highly selective inhibitor of D-DT.

Analysis of the D-DT-4-CPPC co-crystal structure suggested the benzoic acid moiety was redundant due to the lack of stabilizing interactions with active site residues [49]. To experimentally verify this structural observation and gain mechanistic insights into inhibition of D-DT, we purchased selected 4-CPPC derivatives and performed single point kinetic assays on both D-DT and MIF (**Figs. 1A, 1B**). **1** was generated by deletion of the m-Toluic acid group in 4-CPPC, while a nitrogen to carbon substitution in the pyridine ring produced **2** (terephthalic acid) (**Fig. 1C**). Removal of either 2- or 5-carboxylate group produced **3** (pyridine-2-carboxylic acid) and **4** (nicotinic acid), respectively. Finally, **5** (4-phenylpyridine) was produced by elimination of the

three carboxylate groups in 4-CPPC. We projected that these structural modifications of 4-CPPC, would provide key information for the development of new D-DT inhibitors. The single point screening strategy, at a testing concentration of 100 μM, was chosen over a multipoint kinetic assay due to the moderate inhibition potency of 4-CPPC [49, 59]. At the same time, we anticipated that the stepwise removal of functional groups from 4-CPPC (**Fig. 1C**) would only have a negative impact on the inhibition potency of the produced molecule, thus a high testing concentration of each compound would be required.

Figure 5.1.

Percent activity of D-DT and MIF by 4-CPPC derivatives.



(Figure 5.1 Continued)

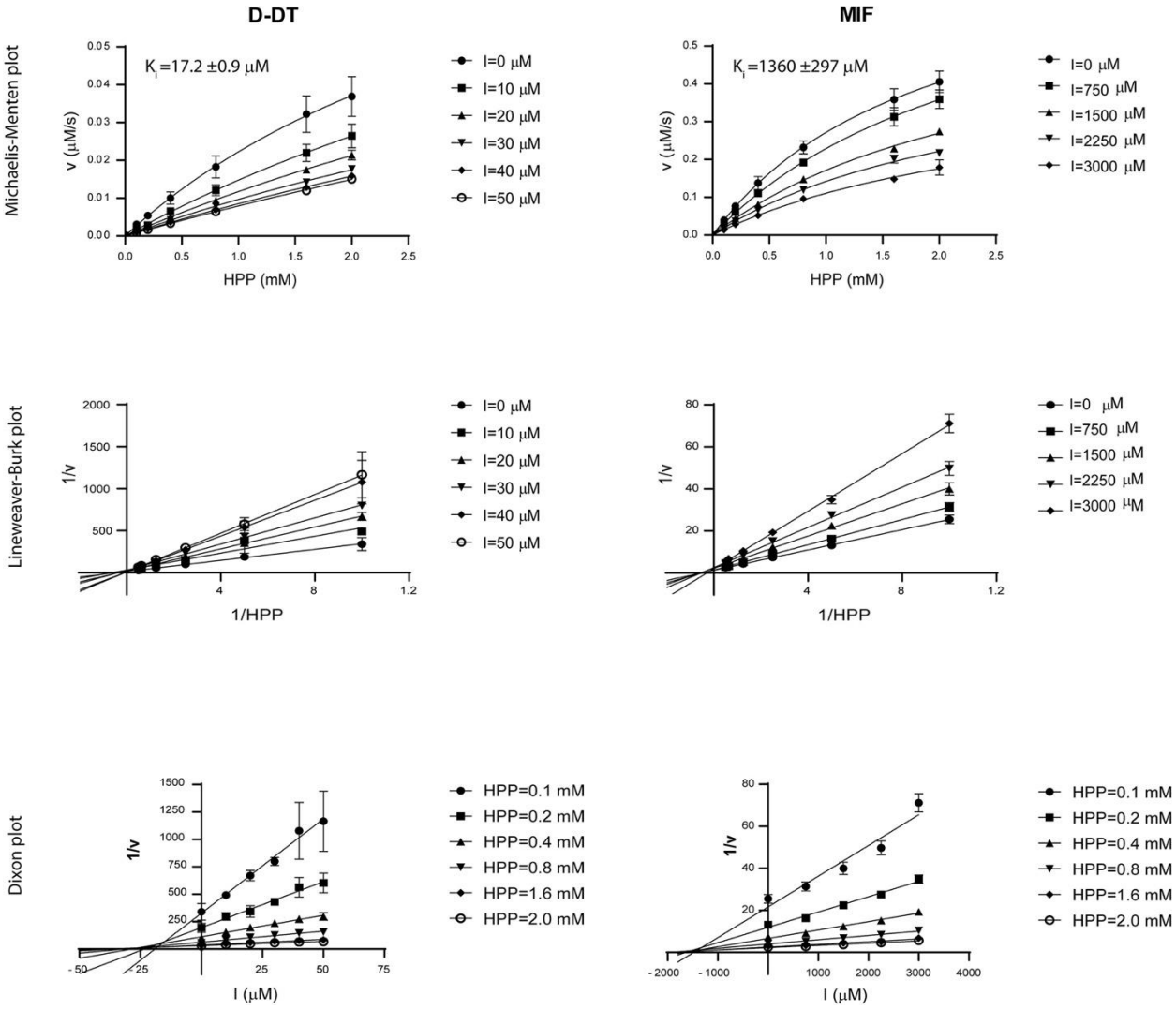
Note. The tautomerase activities of (**A**) D-DT and (**B**) MIF were tested in the presence of 4-CPPC derivatives (**C**) at the single concentrations of 100 μM and 1 mM for the 4-CPPC derivatives and 4-HPP, respectively. 100% activity corresponds to the average value of the non-inhibited reaction. The experiments were carried out in triplicate (n=3) and the error values are shown as standard deviations (SD).

The five compounds along with 4-CPPC were tested against D-DT (**Fig. 1A**) and MIF (**Fig. 1B**) to determine inhibition potency and selectivity. for our experiments 4-CPPC, with a previously reported K_i of 33 ± 0.7 μM [49], served as a reference D-DT inhibitor (positive control). As a negative control, we used dimethyl sulfoxide (DMSO) at a final concentration of 1%. In comparison to 4-CPPC, **1** had an improved inhibition potency and retained selectivity for D-DT (**Figs. 1A, 1B**). At the single concentration of 100 μM, 4-CPPC and **1** inhibit the activity of D-DT by 53% and 77%, respectively. The negligible inhibition potencies of **2-5** against D-DT, demonstrate that the nitrogen atom and two carboxylate groups of the pyridine ring serve a key role in the inhibition potency of **1** (**Fig. 1A)**. Similar to 4-CPPC and **1**, compounds **2-5** are inactive against MIF (**Fig. 1B**).

We carried out multipoint kinetic experiments against D-DT and MIF to quantitatively determine the inhibition potency and selectivity of **1** (**Fig. 2**). The single point kinetic data (**Fig. 1A, B**) guided the design of these experiments and determine the appropriate concentration ranges of **1**. for D-DT, **1** was tested at a concentration range of 0-50 μM. In the case of MIF, we had to do several preliminary kinetic trials before we finalized the concentration range between 0-3000 μM. We noticed absorbance inconsistencies, when **1** tested at concentrations >600 μM. This finding aligns with the previously described property of **1** to form complexes with solution ions (e.g. Zn, Ni, Cu, and Fe) [150], and it was resolved by the use of 1 mM ethylenediaminetetraacetic acid (EDTA). The K_i values of 17.2±0.9 μM and 1360±297 μM were obtained for D-DT and MIF, respectively (**Fig. 2**). **1** demonstrates an impressive 79-fold selectivity for D-DT over MIF. In contrast to 4-CPPC [49], the inhibition potency and selectivity of **1** for D-DT was improved by 2- and 6-fold, respectively.

Figure 5.2.

Kinetic characterization of 1 against D-DT and MIF.



Note. The inhibition activity of **1** against D-DT and MIF was examined at the concentration ranges of 0-50 μM and 0-3000 μM, respectively. The experiments were carried out in triplicate (n=3) and the error values are shown as standard deviations (SD).

Structural characterization of D-DT-1.

We probed crystallographic characterization of the D-DT-**1** complex, to provide structural evidence of the compound's inhibition activity and selectivity. We employed isothermal titration

calorimetry (ITC) and circular dichroism (CD) spectroscopy, prior to setting up crystallization trays, in order to determine the binding affinity of **1** and thermal stability of D-DT-**1**, respectively. Our ITC findings demonstrate a 1:1 protein to ligand stoichiometry with a dissociation constant (K_D) of 36 μM (**Fig. S1**). Next, we ran thermal denaturation experiments of apo D-DT and D-DT-**1**. To ensure all active sites were fully occupied, we recorded the melting profile of D-DT-**1** three times at ligand concentrations multiple folds above the K_D value (**Fig. S1**). The melting temperature of apo D-DT was determined at 71 ^{o}C, whereas the ligand did not have a noticeable effect on the stability of the protein (**Fig. S2**).

Given the modest K_D value, we did not attempt co-crystallization of D-DT with **1**. Instead, we produced apo D-DT crystals and soaked them in 100 mM of **1**. Through this approach, the D-DT-**1** crystal structure was obtained (**Table 1**) and the electron density of **1** was clearly detected in all three active sites (**Fig. 3A, Fig. S3**). The structural impact of D-DT-**1** complex formation was examined at a global (trimeric D-DT) and local (active site) level with the known structure apo D-DT (Protein Data Bank (PDB) entry: 1DPT [90]) as the point of reference. At a global level, D-DT-**1** and apo D-DT demonstrated negligible structural differences, with a root-mean-square deviation (RMSD) value of 0.33 Å (**Fig. 3B**).

Table 5.2.

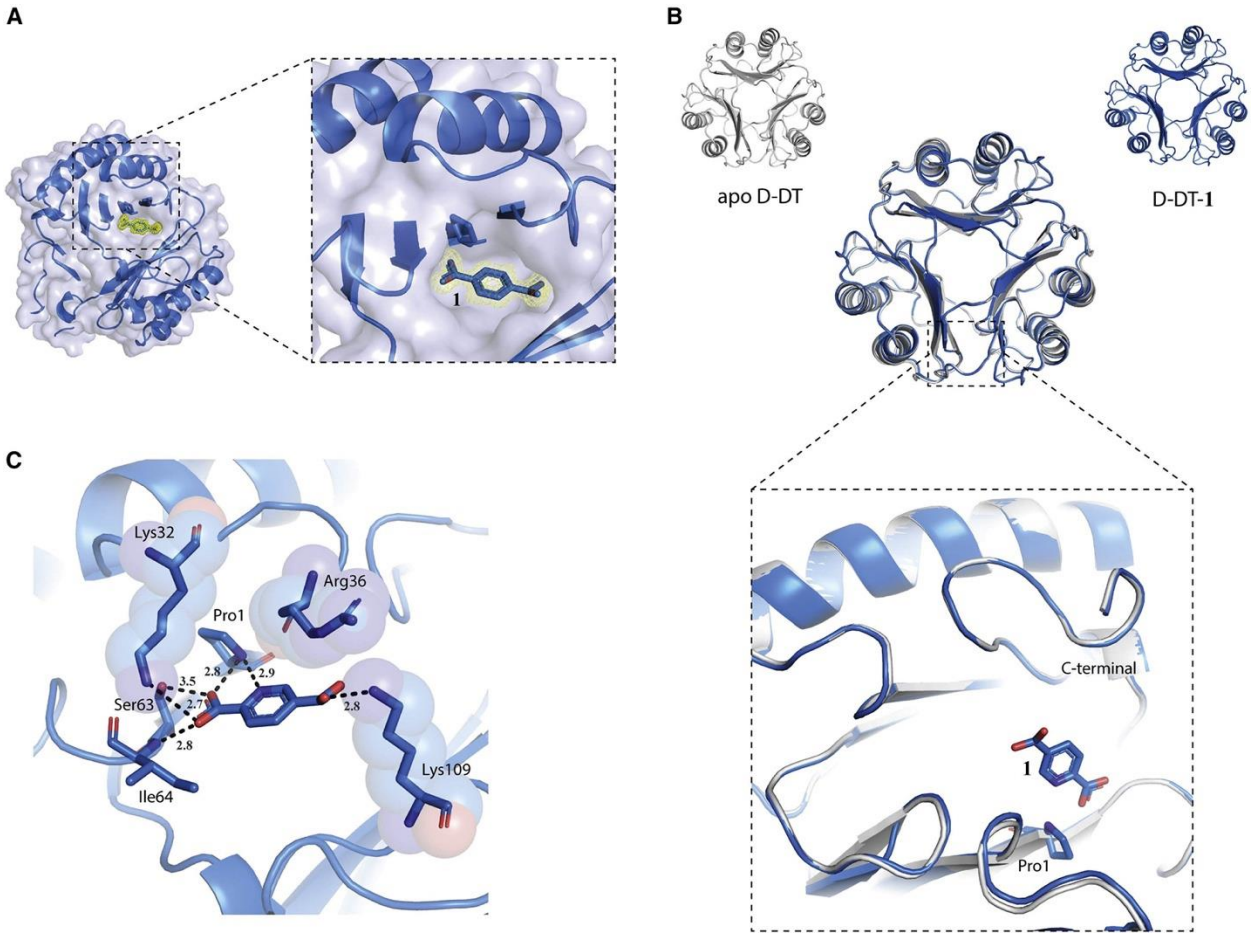
Data collection and refinement statistics for D-DT-1

	D-DT-1
Data collection	
Space group	P 3
Cell dimensions	
a, *b*, *c* (Å)	83.72, 83.72, 40.63
α, β, γ (°)	90.00, 90.00, 120.00
Resolution (Å)	41.90–1.30 (1.33–1.30)[a]
R_{sym} or R_{merge}	0.094 (0.173)
$I/\sigma I$	20.3 (8.6)
Completeness (%)	95.8 (93.7)
Redundancy	6.7 (6.2)
Refinement	
Resolution (Å)	41.90–1.30
No. reflections	71,400
R_{work}/R_{free}	0.12/0.16
No. atoms	
Protein	2,638
Ligand/ion	49
Water	505
B-factors (Å^2)	
Protein	8.6
Ligand/ion	13.9
Water	26.6
RMS deviations	
Bond lengths (Å)	0.012
Bond angles (°)	1.766

[a] Values in parentheses are for highest-resolution shell.

Figure 5.3.

Crystallographic characterization of D-DT-1.



Note. **(**A) The *2Fo-Fc* electron density of **1** was clearly observed in the three active sites of D-DT. The electron density of **1** is shown in yellow, while the inhibitor is illustrated as blue sticks. (**B**) The D-DT-**1** crystal structure (blue) demonstrated high superposition agreement to the corresponding structure of apo D-DT (grey). Binding of **1** does not induce any major conformational changes in the active site environment of D-DT, including the C-terminus. The catalytic residue Pro1 and **1** are shown as blue sticks. (**C**) The inhibitor forms a number of hydrogen bonding and electrostatic interactions with active site residues. The D-DT-**1** hydrogen bonds are shown as black dotted lines and the bond's distance is measured in Å. The three residues

(Figure 5.3 Continued)

participating in electrostatic interactions with the carboxylate groups of **1** are presented as transparent spheres. Hydrophobic interactions play a minor but supporting role in stabilizing D-DT-**1** complex.

Binding of **1** did not trigger any major conformational changes in the active site as we have previously seen with 4-CPPC [49] and the C-terminal adopted a structural orientation similar to that of apo D-DT (**Fig. 3B**). Upon comparison with apo D-DT (**Fig. S4A**), we found that the active site opening of D-DT-**1** was noticeably enlarged (**Fig. S4B**). Our observation agrees with previous crystallographic findings [59] and is associated with the ligand-induced conformational change of Arg36. This gating residue is located at the opening of the active site, positioned in two distinct conformations depending on the protein's state. The "closed" conformation of Arg36 was noted in the apo D-DT structure and resulted in a narrow active site opening that does not physically allow any molecules to enter the pocket (**Fig. S4A**). However, in the presence of an active site binder such as **1**, the side chain of Arg36 adopts a different conformation that increases the opening of the pocket and permits ligand binding (**Fig. S4B**).

Hydrogen bonding and electrostatic interactions are the main stabilizing forces of D-DT-**1** (**Fig. 3C**). Hydrophobic interactions also contribute to the stability of the complex albeit to a lesser extent. The two carboxylate groups of **1** form hydrogen bonds with Pro1, Lys32, Ser63, Ile64, and Lys109 and electrostatic interactions with Lys32, Arg36 and Lys109 (**Fig. 3C**). The carboxylate group, at position two, plays a central role in binding by forming four hydrogen bonding interactions with active site residues. The second carboxylate group, at position five, promotes binding of **1** due to electrostatic interactions for the most part (**Fig. 3C**). Our crystallographic

analysis and single point kinetic results support that neither of the carboxylate groups would be able to orient correctly in the active site of D-DT if the pyridine nitrogen is missing (**Figs. 1A, 3C**). Hydrophobic interactions involving Pro1, Ser63 and Ile64, play a weak but supportive role in stabilizing D-DT-**1**.

Analysis of D-DT-1 profile by MD simulations.

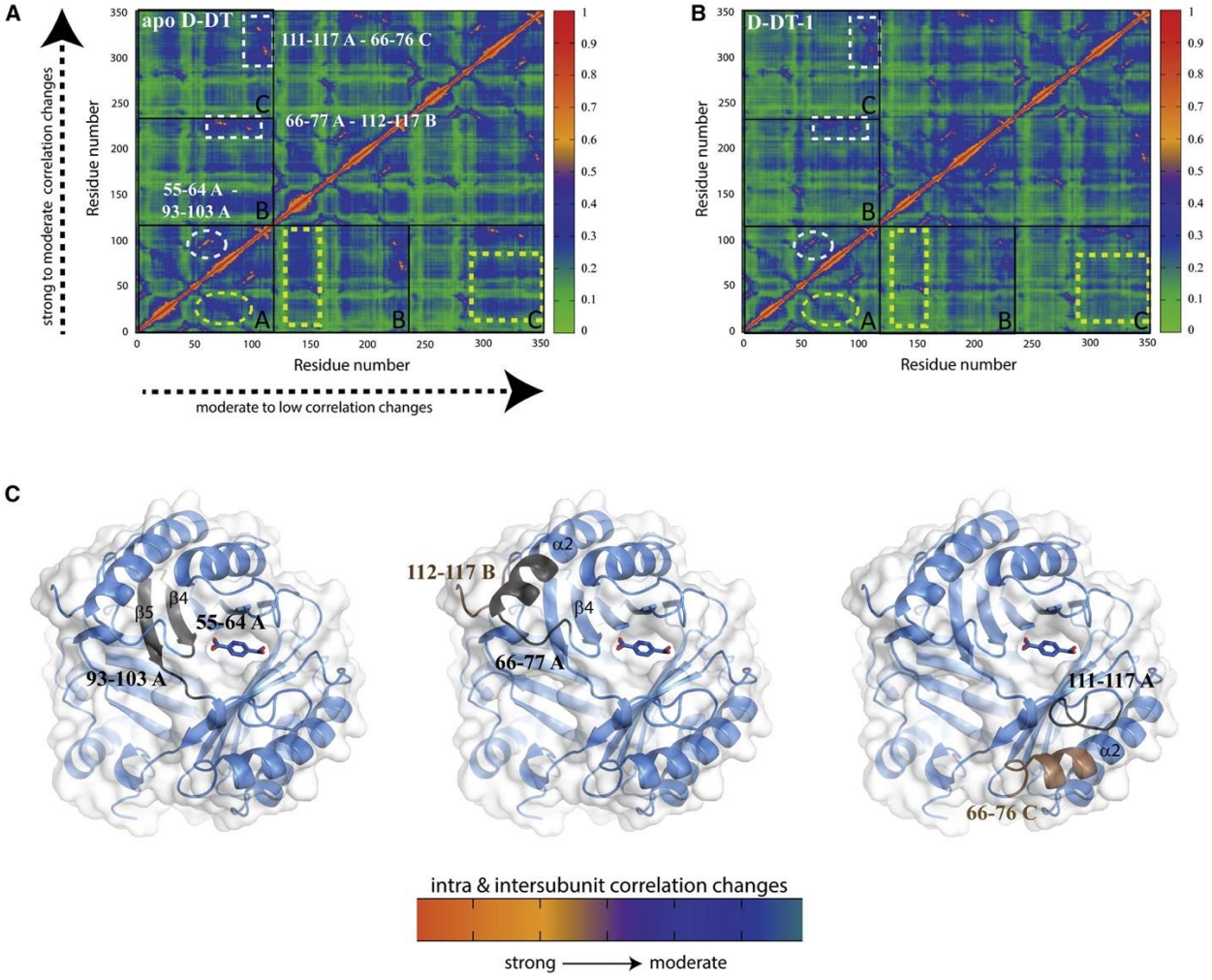
Along with the conformational change of Arg36, which controls ligand binding via regulation of the active site opening, we have previously brought to light an induced-fit mechanism that involved C-terminal conformational changes upon 4-CPPC binding [49]. Consequently, the residues M114-L117 became unusually flexible resulting in elimination of their electron densities. These findings uncover structural features of D-DT that noticeably differ from what we know for MIF and aid the efforts towards understanding the structure-function aspects of the MIF superfamily members.

We employed MD simulations to examine how **1** influences the intra- (within monomer A) and intersubunit (monomer A with monomers B and C) communications of D-DT. The duration of each calculation was 200 ns and the produced correlation plots of apo D-DT and D-DT-**1** were compared side-by-side (**Figs. 4A, 4B**). The extent of correlation between two given regions is graphically measured on a scale of 0 to 1. Green, blue, orange, and red colors represent regions with none too low, low to moderate, moderate to strong, and strong to absolute correlation, respectively (**Figs. 4A, 4B**). The absolute correlation value equates to 1 and is obtained only when a residue is compared to itself. Distal correlated regions, illustrated in orange, are the most interesting for further assessment. from this analysis, we found **1** has an overall negative impact on the intra- and intersubunit correlations of D-DT. Graphically, this is illustrated in the apo D-DT (before binding) and D-DT-**1** (after binding) correlation plots as a color change from either

orange to blue (white circle and boxes) or blue to green (yellow circle and boxes) (**Figs. 4A, 4B**). Moderate to low correlation changes (yellow circle and boxes) have only a qualitative value of global reduction of communications, thus, these regions would not undergo further analysis. In contrast, the strong to moderate correlation changes (white circle and boxes) highlight interruption of key intra- and intersubunit communications. Such correlation changes would be analyzed in detail because they serve as a tool to understand the inhibition features of **1**.

Figure 5.4.

Correlation analysis of apo D-DT and D-DT-1.



(Figure 5.4 Continued)

Note. The Cα correlation analyses were obtained from 200 ns MD simulations using the high-resolution crystal structures of apo D-DT (PDB entry: 1DPT) and D-DT-**1** (PDB entry: 8DBB). Side-by-side examination of the (**A**) apo D-DT and (**B**) D-DT-**1** correlation plots demonstrate that binding of the ligand resulted in reduction of intra- and intersubunit communications. The white circle and boxes, shown vertically, highlight the strong to moderate correlation changes occurring within monomer A (circle) or between monomer A and the other two monomers (boxes), as a consequence of ligand binding. The yellow circle and boxes, shown horizontally, highlight the moderate to low correlation changes occurring within monomer A (circle) or between monomer A and the other two monomers (boxes), as a consequence of ligand binding. (**C**) The strong to moderate correlation changes are also mapped on the structure of D-DT.

One intrasubunit (55-64 A with 93-103 A) and two intersubunit (66-77 A with 112-117 B and 111-117 A with 66-76 C) correlations were significantly affected by the formation of D-DT-**1** complex (**Figs. 4A, 4B**). After mapping these regions on the crystal structure of D-DT-**1** (**Fig. 4C**), we probed for crystallographic evidence to explain the correlation results. We found that Ser63, Ile64, and Lys109, which form hydrogen bonding interactions with **1** (**Fig. 3C**), possess key positions on the affected segments. Ser63 and Ile64 are located on β4 strand and α2/β4 loop, respectively. Restriction in their molecular motions, due to hydrogen bonding with **1**, has a direct impact on the conformational flexibility of β4 strand. Consequently, the intrasubunit correlation between segments 55-64 and 93-103 of subunit A, is also affected (**Fig. 3C**). Lys109 is located on a key position of the C-terminal, like Ser63 and Ile64. Hindering the mobility of Lys109 due to

hydrogen bonding with the ligand, influences the ability of the C-terminal to effectively communicate with the α2 helix from the adjacent monomer.

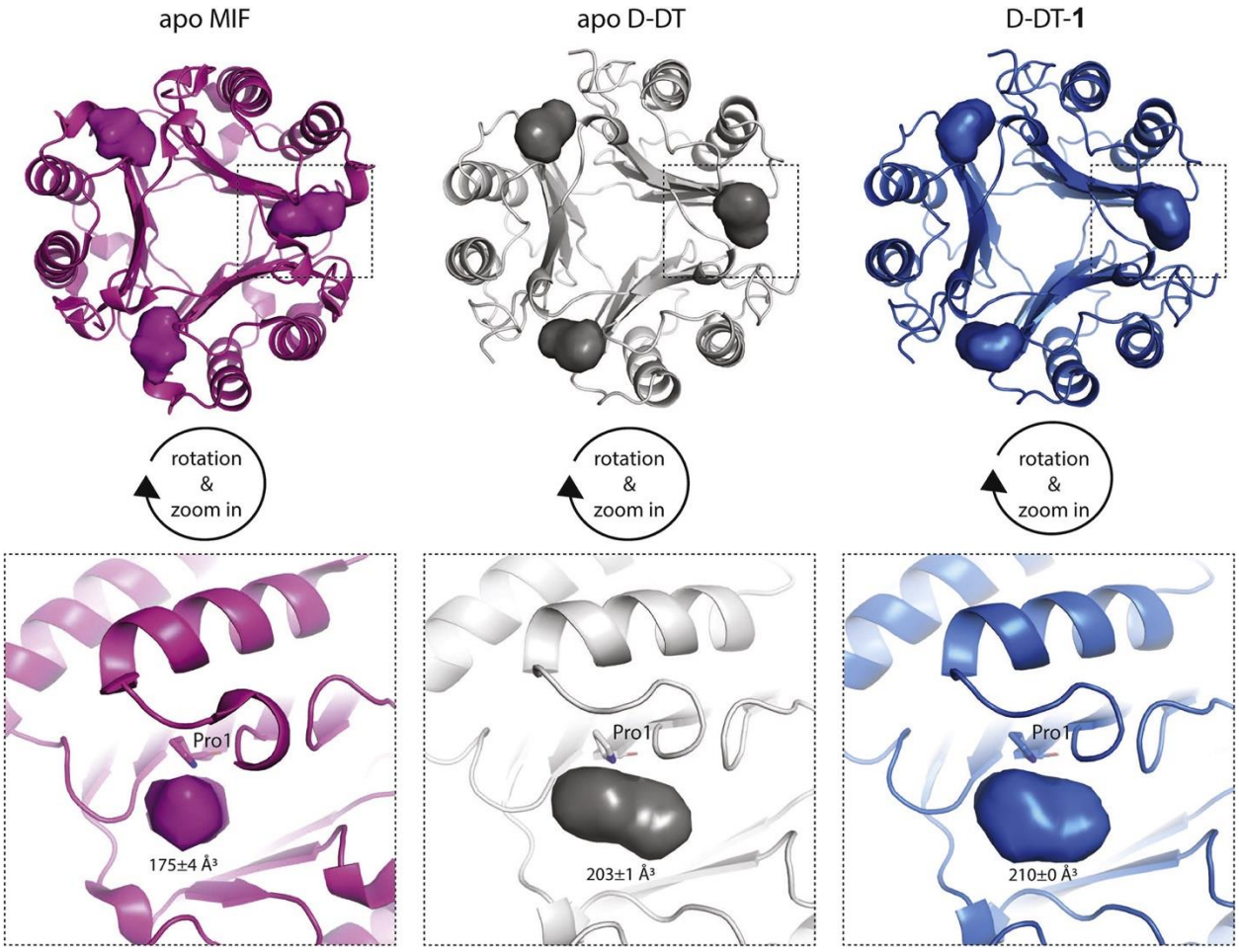
To assess the impact of ligand binding in protein dynamics, we calculated root mean square fluctuation (RMSF) of apo D-DT and D-DT-**1** and overlaid the two plots (**Fig. S5**). At a global level (D-DT trimer), we noted negligible differences with an average RMSF value of 0.62±0.2 Å and 0.58±0.2 Å for D-DT and D-DT-**1**, respectively. At a local level (active site pocket), we detected noticeable shifting in the fluctuation activity of two regions. After structural mapping, we found that the two affected segments, 24-36 and 64-69, are located proximal to the binding site of **1**. Upon complex formation, Lys32, Arg36, and Ile64 establish electrostatic (Lys32 and Arg36) and hydrogen bonding (Il64) interactions with **1** (**Fig. 3C**), leading to reduction of 24-36 and 64-69 segments' flexibility.

Structural basis of ligand recognition and functional selectivity.

The multipoint kinetic results showed an impressive selectivity of **1** for D-DT (**Fig. 2**). To clarify the parameters of this selectivity, we performed a structural analysis focusing on the ligand recognition features of D-DT and MIF. The active site volumes of apo MIF (magenta), apo D-DT (grey), and D-DT-**1** (blue) (**Fig. 5**) were calculated and compared using pyVOL, a PyMOL plugin for visualization, comparison, and volume calculation of ligand binding sites [151]. Numerically, the active site volume of apo MIF is the smallest of all three, with an average value of 175±4 $Å^3$. In the absence of a ligand, the average volume of D-DT was calculated at 203±1 $Å^3$, while in the presence of **1**, it increased to 210±0 $Å^3$ due to the conformational change of Met114 (**Fig. S6**). Conformational changes of Lys32 and Arg36 enabled stabilizing interactions with **1** (**Fig. 3**), but they are not accountable for the increase of volume.

Figure 5.5.

Illustration MIF and D-DT active site volumes.



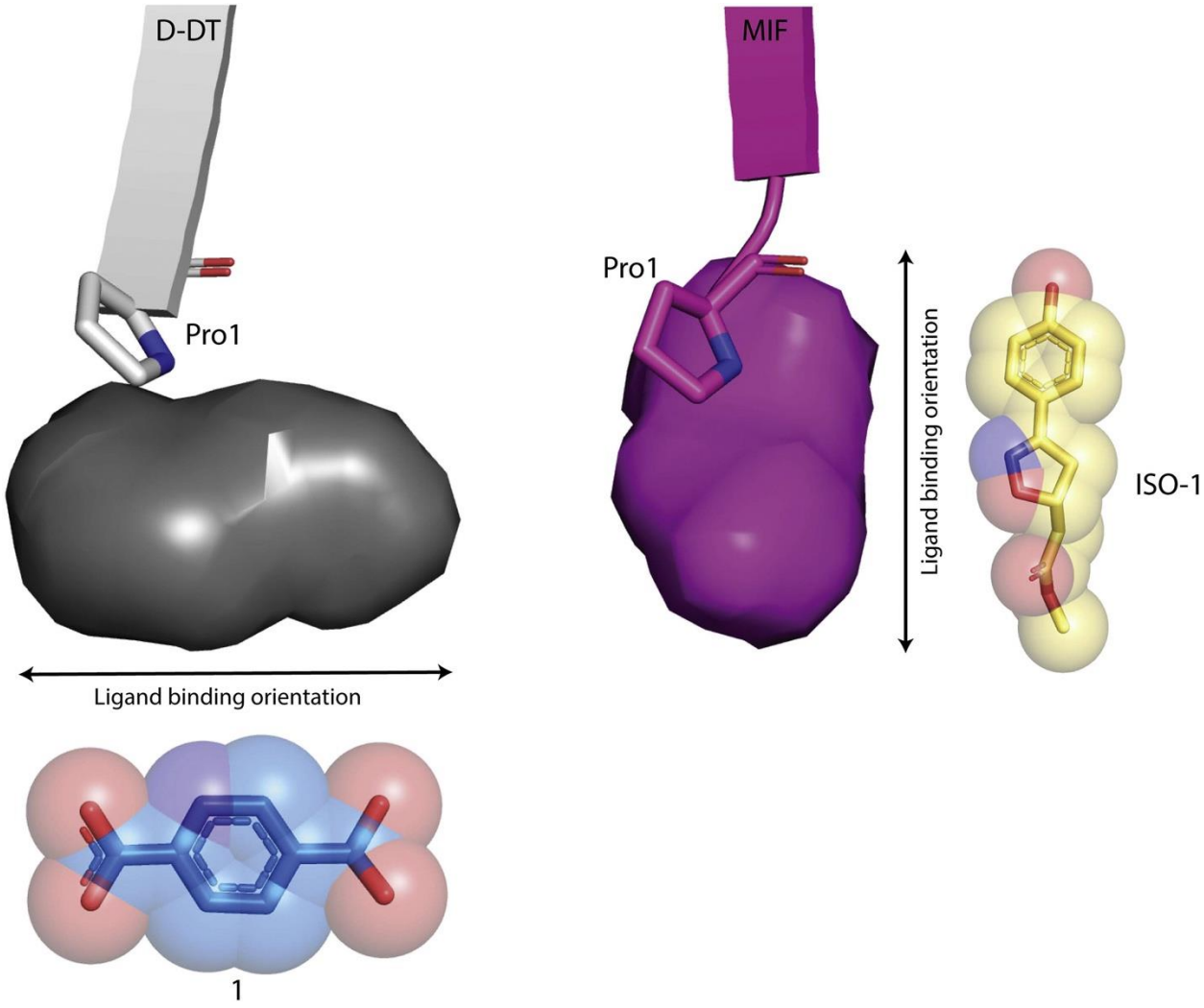
Note. PyVOL, a PyMOL plugin, was used to visualize, compare, and calculate the active site volumes of apo MIF (magenta), apo D-DT (grey), and D-DT-**1** (blue). for this analysis, the crystal structures of apo MIF (PDB entry: 3DJH), apo D-DT (PDB entry: 1DPT), and D-DT-**1** (PDB entry: 8DBB) were utilized. The active site volumes are shown as blobs and Pro1 as stick.

Structural alignment of apo MIF and D-DT displayed a perpendicular orientation of the two active site pockets (**Fig. S7**). This finding shows that the structural basis of ligand recognition for each protein is significantly different. The significance of this discovery is better understood

by comparing the binding orientations of **1** and 4,5-Dihydro-3-(4-hydroxyphenyl)-5-isoxazoleacetic acid methyl ester (ISO-1) in the corresponding active site pockets of D-DT and MIF (**Fig. 6**). ISO-1 is widely used for *in vitro* and *in vivo* experiments as a proof-of-concept MIF therapeutic [152]. In this study, ISO-1 was chosen as a representative MIF inhibitor based on our crystallographic analysis, which showed this molecule accurately represents the binding orientation of MIF ligands with diverse chemical scaffolds (**Fig. S8**).

Figure 5.6.

Side-by-side comparison of the D-DT and MIF active site pockets.



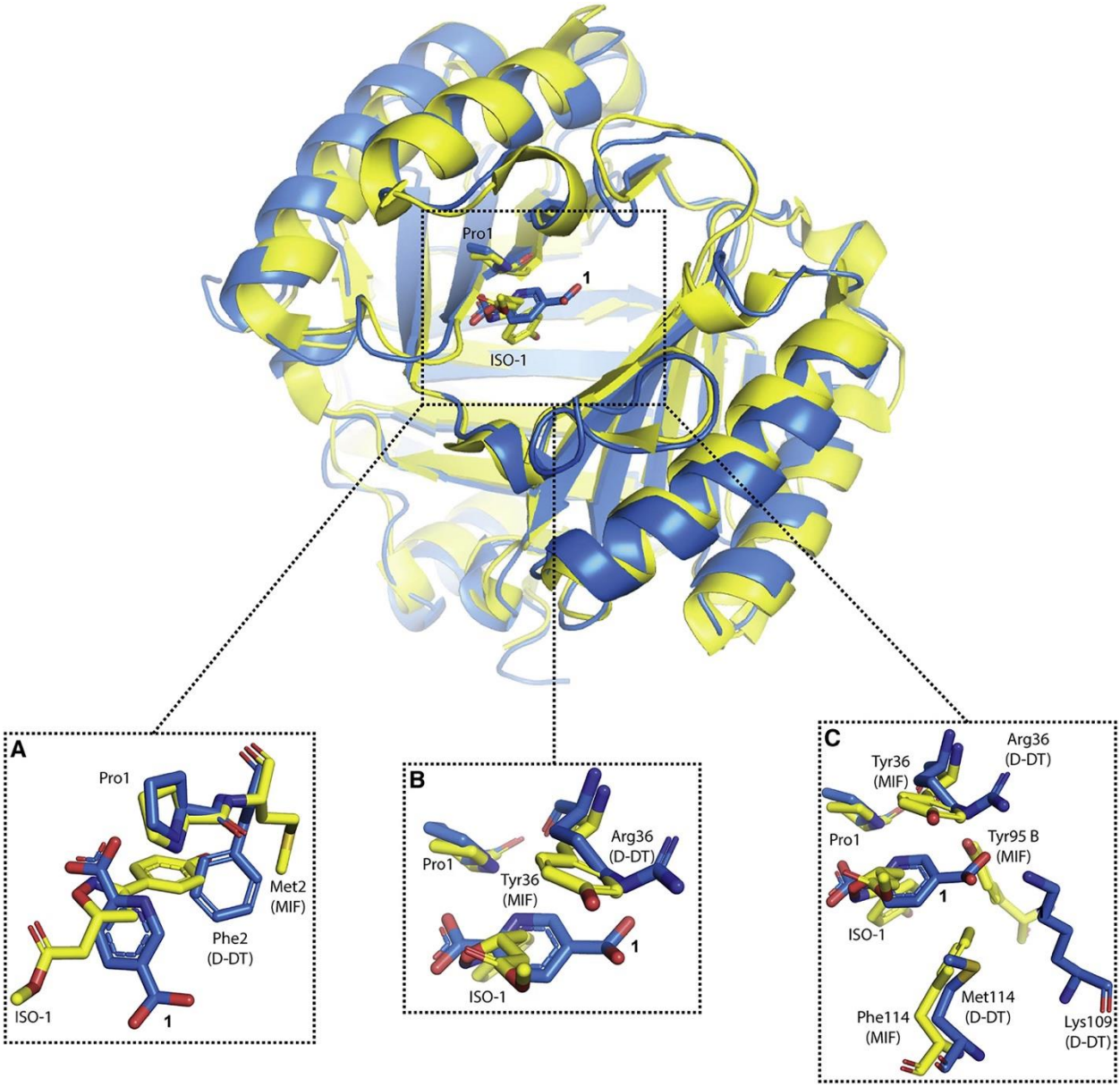
(Figure 5.6 Continued)

Note. In reference to Pro1, the active site pockets of D-DT (grey) and MIF (magenta) have a perpendicular and parallel orientation, respectively. This explains why 1 is able to bind D-DT but not MIF. MIF active site favors binding of compounds that adopt orientations similar to that of ISO-1.

Upon comparing the binding features of **1** and ISO-1 (**Fig. 7**), we found key structural differences that explain the high selectivity of **1** for D-DT; the first being the conformational flexibility of residue two. for MIF and D-DT, residue two is a Met and Phe, respectively (**Fig. 7A**). The crystal structures of apo MIF [46] and MIF-ligands [53] demonstrate that the side chain of Met2 is highly flexible and adopts multiple conformations. The enhanced flexibility of this residue allows MIF to accommodate ligands in binding orientations like that of ISO-1 (**Fig. 6, Fig S8**). for D-DT, this is excluded due to the presence of Phe2 (**Fig. 7A**). The conformational flexibility of residue 36 is another regulatory point, which provides mechanistic evidence in relation to MIF and D-DT ligand selectivity (**Fig. 7B**). This residue possesses a dual functional role of ligand recognition and stabilization. for MIF, the conformational changes of Tyr36 do not have a noticeable impact on ligand binding (**Fig S8**). Instead, the primary role of such movements is to enable interactions with the ligand and promote stabilization of the MIF-ligand complex [153]. Arg36 also contributes to ligand stability in D-DT (**Fig. 3C**), however, the predominant functional role of its conformational change is to enable ligand binding by enlarging the size of the active site opening (**Fig. S2**). The MIF pocket contains multiple aromatic residues that not only increase the hydrophobicity of the pocket, but also define the shape of the pocket (**Fig. 7C**). Thus, in relation to Pro1, a ligand like **1** would clash due to the requirement of binding in a vertical geometry.

Figure 5.7.

Structural basis of ligand recognition and functional selectivity for MIF and D-DT.



Note. A side-by-side comparison of D-DT-**1** (blue) and MIF-ISO-1 (yellow) crystal structures illustrates the structural differences associated with the functional selectivity of **1**. (**A**) Residue 2 plays a key role in orienting the ligand in the active sites of MIF and D-DT. (**B**) Residue 36 has a dual functional role, which includes ligand recognition and stabilization. for D-DT, ligand recognition is the primary functional role for Arg36, while Tyr36 of MIF serves as stabilizing

(Figure 5.7 Continued)

factor of ISO-1 after ligand binding. (**C**) The hydrophobic amino acids found in the active site of MIF, but not in D-DT, define the topology of the active site pocket, which drastically differs from that of D-DT (**Fig. 5.S7**).

Inhibition of D-DT induced activation of CD74 by 1.

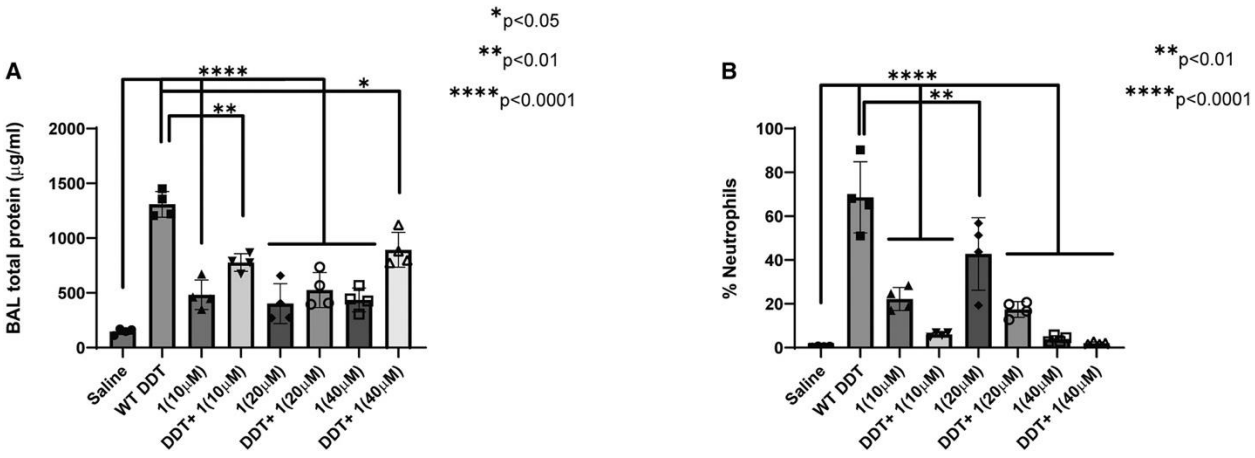
Following the biochemical, computational, and structural characterization of D-DT-**1**, we probed whether **1** inhibits the D-DT induced activation of CD74. To avoid *in vitro* complications related to the stability of the receptor [41], we utilized a previously established *in vivo* assay that accurately measures the MIF or D-DT induced activation of CD74 [29, 48, 51, 52, 57, 58, 88]. The principle of this neutrophil recruitment assay is based on the ability of lipopolysaccharides (LPS)-free MIF or D-DT to activate CD74 on the cell surface of bronchoalveolar lavage (BAL) macrophages. Upon activation of CD74, the proinflammatory proteins macrophage-inflammatory protein 2 (MIP-2) and keratinocyte chemoattractant (KC) are recruited, which in turn trigger accumulation of neutrophils within the period of 4-6 hours [29]. The number of neutrophils found in the alveolar space determines the extent of CD74 activation.

For this assay, **1** was intratracheally installed in the lungs of mice at a concentration range of 10-40 μM. This range was chosen considering the K_i (**Fig. 2**) and K_D (**Fig. S1**) values of D-DT-**1**, as well as the potential lethality risk in the case of using higher concentrations of **1**. The inhibitor was installed in the presence or absence of D-DT, having saline as a negative and apo D-DT as a positive control. Testing **1** in the absence of D-DT, would determine if the compound itself has inflammatory properties. Intratracheal installation of apo D-DT resulted in an increase of BAL proteins that is significantly different from the remaining experimental groups (**Fig. 8A**). In

contrast, the saline group did not have a noticeable impact on recruiting BAL proteins. Together, these findings show the experiment followed the anticipated mechanistic pathway [29], which was the accumulation of proinflammatory proteins upon the D-DT-induced activation of CD74. While the concentration of BAL proteins is not an accurate measurement of CD74 activation, we also calculated neutrophils (**Fig. 8B**). As shown by the saline group, neutrophils are not present in the alveolar space; instead, they are recruited upon installation of D-DT. The significant elevation of neutrophils, noted in the case of apo D-DT (**Fig. 8B**), was blocked in the presence of 10-40 μM of **1**. At the same time, the compound itself has minor inflammatory properties.

Figure 5.8.

Neutrophil recruitment assay monitoring the D-DT induced activation of CD74.



Note. (**A**) Intratracheal installation of LPS free D-DT triggered recruitment of proinflammatory proteins in the bronchoalveolar lavage fluid of mice. In contrast, saline had not a noticeable effect. These findings agree with a previously established mechanism in which it was shown the action of D-DT is associated with activation of macrophage CD74 [29]. The inflammatory effect of D-DT is suppressed when adding **1**, showing that the compound blocks the D-DT-induced activation of CD74. (**B**) Upon activation of macrophage CD74 by D-DT, neutrophils are recruited in the

(Figure 5.8 Continued)

alveolar space. This action is blocked when **1** is added at a concentration range of 10-40 μM. Data were plotted using four replicates and expressed as mean ± SEM.

Highly promising protein modulators with impressive therapeutic results at preclinical models, often hit a roadblock in clinical trials [154]. Among the reasons contributing to the negative outcome are the incomplete understanding of the protein's molecular mechanisms and the compensatory expression of homologs that share the same biological activity [117]. The two human members of the MIF superfamily, D-DT and MIF, share the most notably cytokine activity [9, 27] as well as endonuclease [13, 15] and tautomerase [85, 129] activities. The early identification of MIF as an immunomodulatory protein garnered much interest towards resolving its biological activity, ignoring at the same time the existence of D-DT. Several studies have shown D-DT has a regulatory function in immunity [9, 155], though an in-depth understanding of its functionality, in human physiology and pathology, is still missing. One could attribute that this is partially due to the lack of highly selective D-DT modulators that can be utilized alongside the abundant MIF inhibitors.

In this study, we report the identification and characterization of **1** as a highly selective inhibitor of D-DT, providing at the same time the structural basis of this selectivity. Kinetic experiments revealed a 79-fold selectivity of **1** for D-DT over its human homolog MIF, while a neutrophil recruitment assay demonstrated that the compound is an effective *in vivo* inhibitor of the CD74 activation. The micromolar inhibition potency (17.2±0.9 μM) of **1** for D-DT leaves room for improvement, but without shadowing its importance as the D-DT inhibitor with the highest selectivity reported thus far. A careful examination of the literature shows that ISO-1,

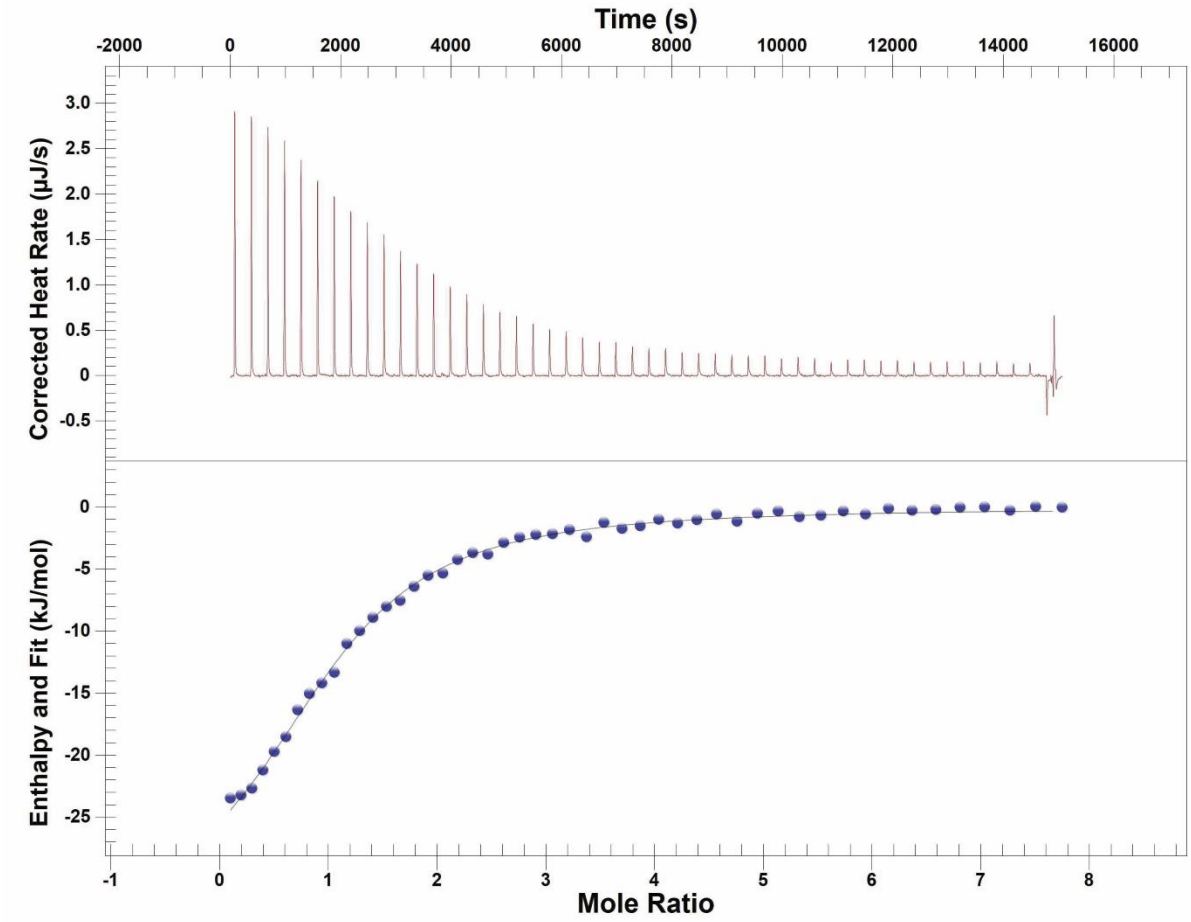
which is the most utilized MIF inhibitor for *in vitro*/*in vivo* MIF studies, has inhibition (K_i) and dissociation constants (K_D) equal to 24.1±1.7 μM and 14.4±2.3 μM, respectively [21]. ISO-1 and **1** may concurrently be used to elucidate the precise biological functionality of MIF and D-DT, in cell types and tissues that the two proteins are coexpressed.

A side-by-side comparison of **1** with the two previously reported, selective inhibitors of D-DT demonstrate the clear advantages of this compound. Starting with 4-CPPC, **1** has a 2- and 6-fold improved inhibition potency and selectivity, respectively [49]. While neither the inhibition constant of 4-CPPC (33±0.7 μM) nor **1** (17.2±0.9 μM) make the two molecules drug candidates, the improved D-DT/MIF selectivity of the latter (4-CPPC:13-fold selectivity versus **1**:79-fold selectivity), make it more suitable for mechanistic studies. The conformational changes observed at the C-terminal of D-DT as a consequence of 4-CPPC binding are of unknown significance, but it is important to note that in the case of **1**, the C-terminal remains unaffected. Besides the functional aspects, **1** has additional advantages related to its cost and availability. **1** is broadly available in the market at a much lower price than 4-CPPC, which is offered only by selected manufacturers. In regards to the second known inhibitor of D-DT [44], (R-5-Methyl-3-(1-(naphthalen-1-yl)ethyl)-6-(3-(trifluoromethyl)phenyl)thieno[2,3-d]pyrimidine-2,4(1H,3H)-dione), **1** also has multiple advantages. Besides the fact that the compound is not commercially available, it has also a much lower selectivity for D-DT (>50-fold selectivity). Unfortunately, the inhibition potency of this inhibitor cannot be directly compared to the corresponding value of **1** due to usage of different substrates. Collectively, our findings demonstrate that **1** is an attractive tool for interrogating the biological activity of D-DT in various disease models and clinical samples.

Supplemental Information

Figure 5.S1.

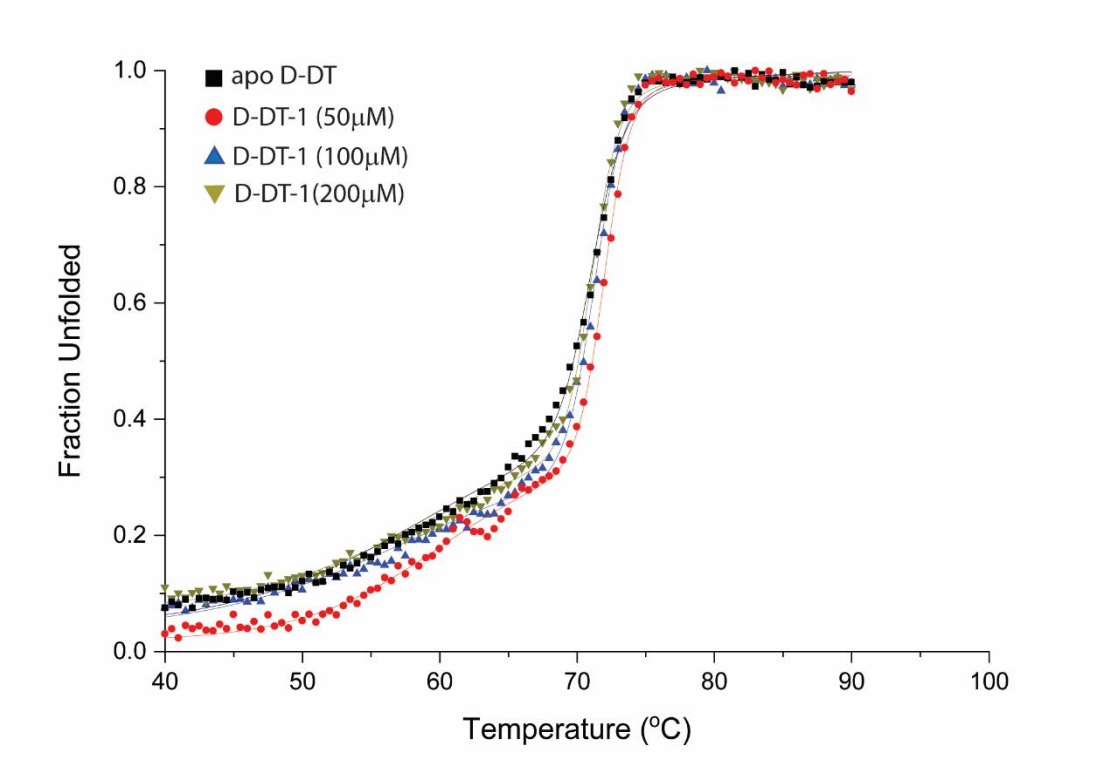
Characterization of D-DT-1 interactions by isothermal titration calorimetry, Related to Figure 5.1.



Note. The dissociation constant (K_D) and binding stoichiometry (n) between D-DT and **1** were determined at 36 μM and 1.19, respectively.

Figure 5.S2.

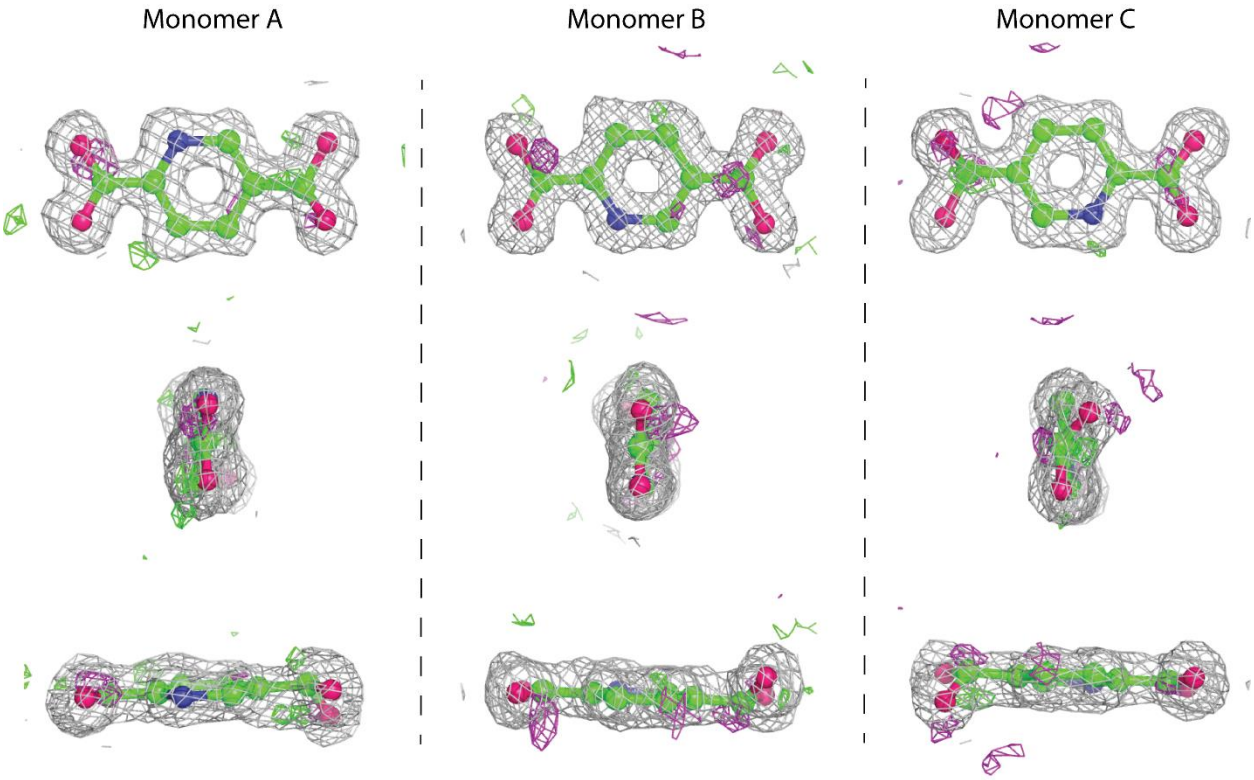
Thermal stability profiles of D-DT in the presence or absence of 1, Related to Figure 2.



Note. The impact of compound's binding on the structural stability of D-DT was measured by circular dichroism (CD), using **1** at a concentration range of 0-200 μM. The melting temperature of apo D-DT was determined at 71°C. The ligand has negligible effect on altering the thermal stability of D-DT.

Figure 5.S3.

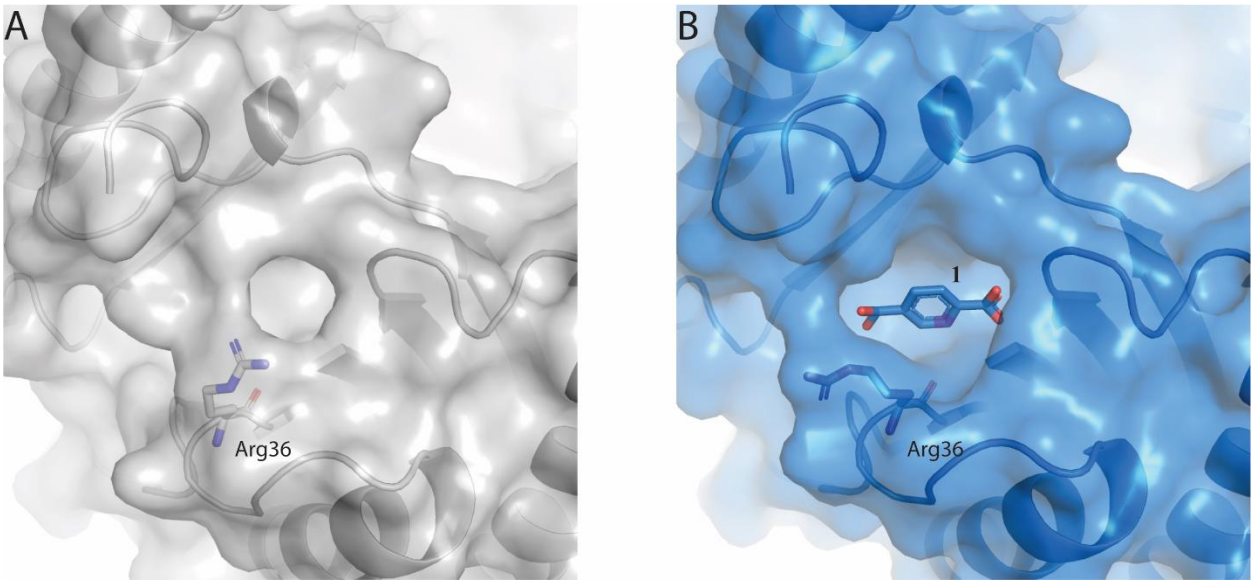
Electron density of 1 in the three active sites of D-DT, Related to Table 1.



Note. The $2mF_o$-DF_c map (grey) of **1** was measured at 0.7 rmsd. The mF_o-DF_c map was measured at 3 rmsd with the negative and positive electron density to be illustrated in purple and green, respectively.

Figure 5.S4.

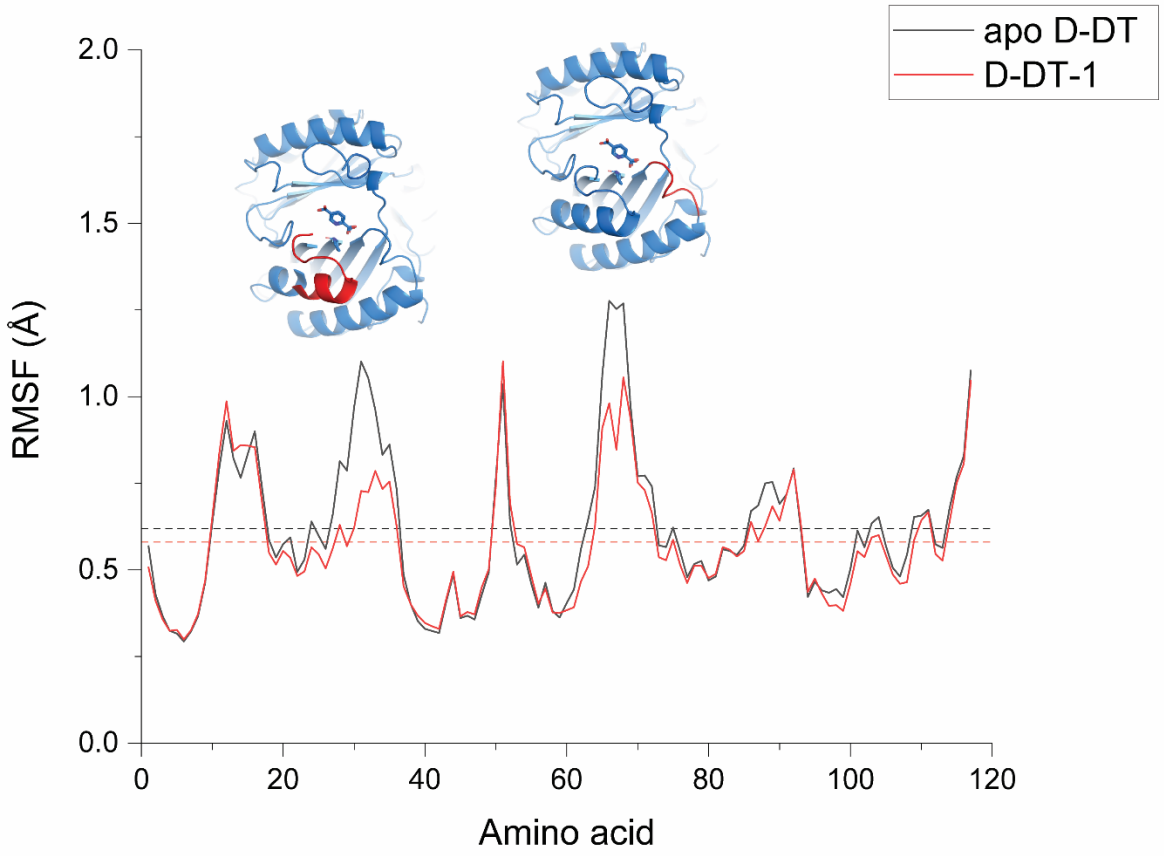
Gating role of Arg36, Related to Figure 3.



Note. Arg36 serves as a gating residue for the active site opening of D-DT. (**A**) In the absence of a ligand (apo D-DT), the side chain of Arg36 adopts a "closed" conformation that makes the active site non-accessible. (**B**) In the presence of an active site binder (D-DT-**1**), Arg36 adopts an "open" conformation that leads to expansion of the active site opening.

Figure 5.S5.

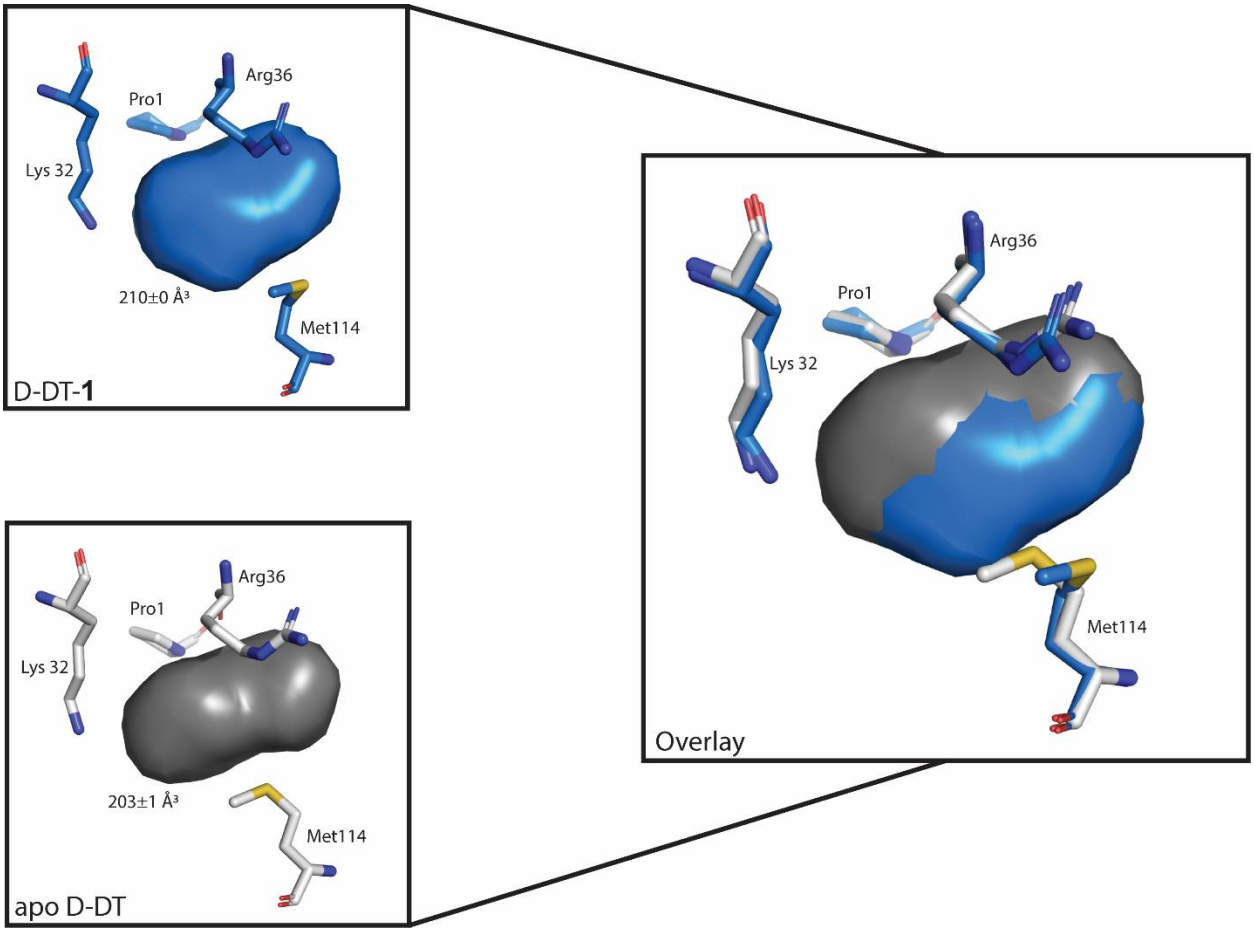
Root mean square fluctuation (RMSF) plots of apo D-DT and D-DT-1, Related to Figure 4.



Note. Binding of **1** reduced the flexibility of two regions found in close proximity to the ligand. These regions are highlighted (red) in the structure of D-DT. The average flexibility of D-DT trimer, before and after ligand's binding is shown as black and red dotted lines, respectively. Binding of **1** dropped the average RMSF value from 0.62 ±0.2Å (apo D-DT) to 0.58 ±0.2Å (D-DT-**1**).

Figure 5.S6.

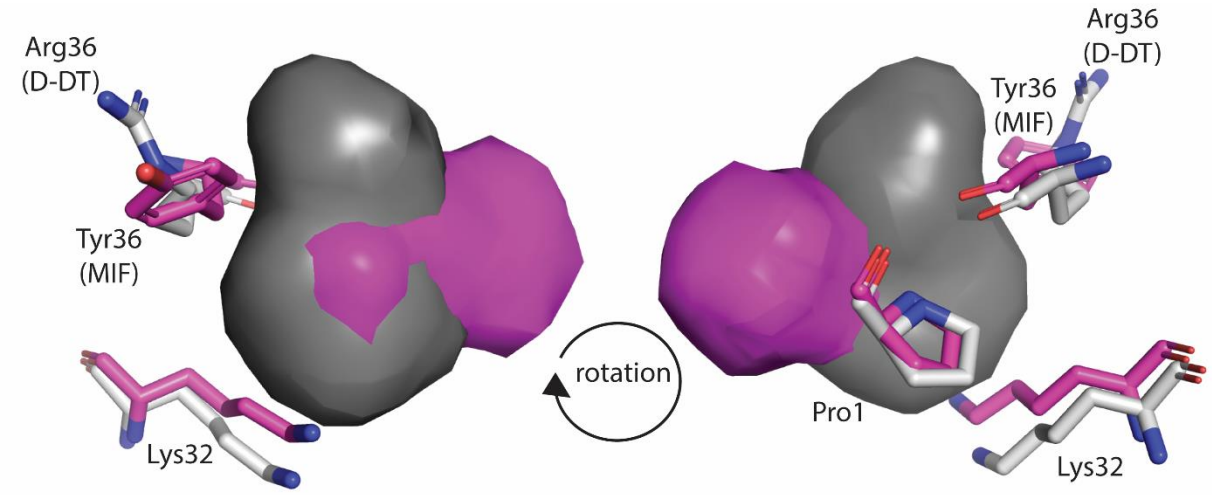
Conformational change of Met114 affects the active site volume of D-DT, Related to Figure 5.



Note. Upon binding of **1**, Lys32 and Arg36 undergo side chain conformational changes in order to form stabilizing interactions with the ligand. The conformational change of Met114 is responsible for increasing the active site volume from 203±1 $Å^3$ (apo D-DT) to 210±0 $Å^3$ (D-DT-**1**). The active site volumes of apo D-DT and D-DT-**1** are illustrated as grey and blue blobs, respectively. Pro1 is used as an orientation point. All residues are shown as sticks.

Figure 5.S7.

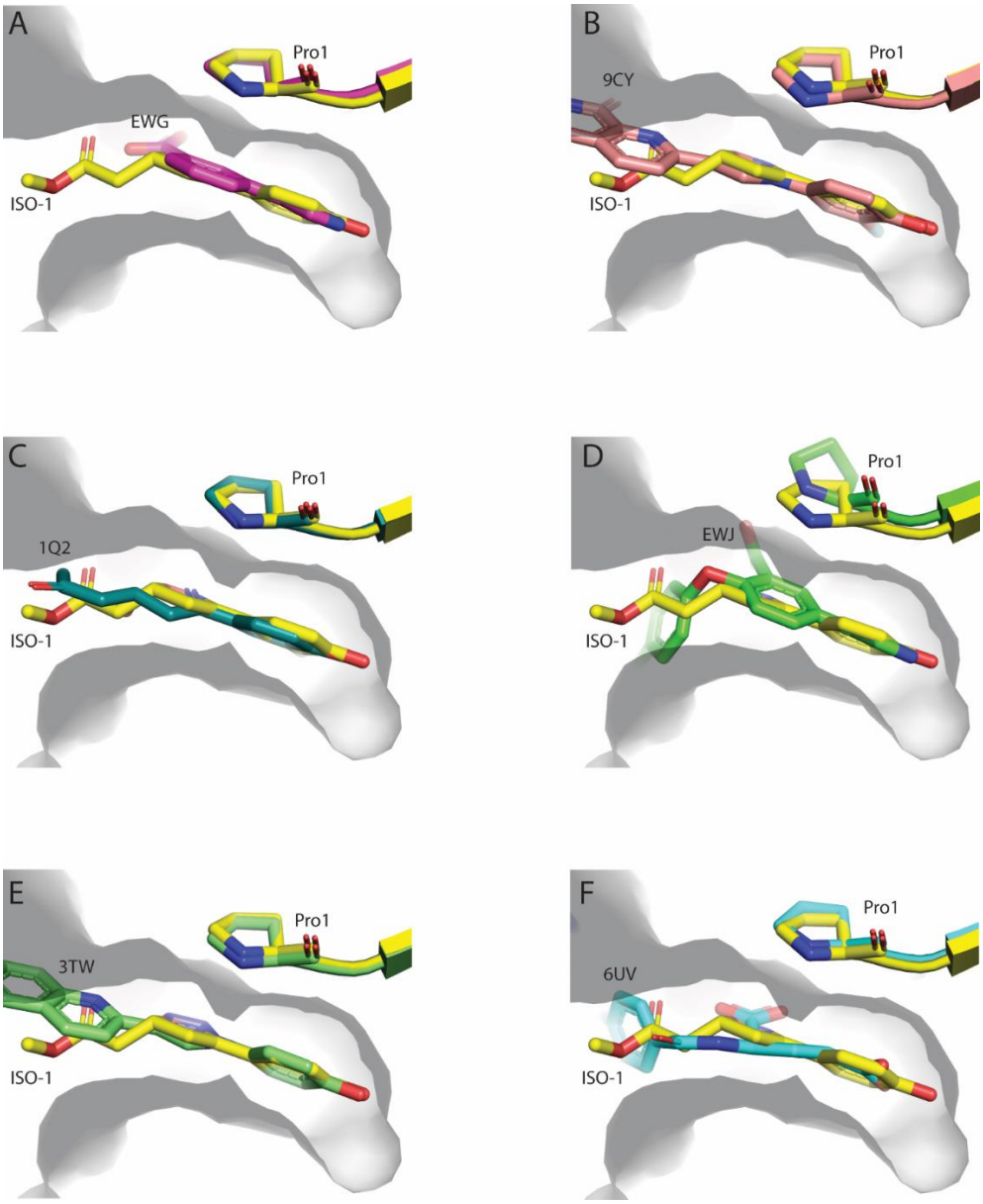
Topology and side-by-side comparison of the apo MIF and apo D-DT active site pockets, Related to Figure 6.



Note. Structural alignment of apo MIF (magenta) and apo D-DT (grey) demonstrates the two pockets are perpendicular to each other. Residues 1, 32, and 36, shown as sticks, are used as orientation points.

Figure 5.S8.

Binding orientation of MIF ligands with diverse chemical scaffolds, Related to Figure 7.



Note. (**A-F)** Having MIF-ISO-1 crystal structure as the reference point, we compared the binding orientation of six MIF ligands with diverse chemical scaffolds. Each ligand is illustrated with the three-letter code assigned to it by protein data bank (PDB). The MIF active site pocket is shown as a grey cavity. ISO-1(PDB entry: 1LJT, [50]), EWG (PDB entry: 6CBG, [156]), 9CY (PDB entry:

(Figure 5.S8 Continued)

6B2C, [157]), 1Q2 (PDB entry: 4K9G, [158]), EWJ (PDB entry: 6CBH, [156]), 3TW (PDB entry: 4WRB, [95]), 6UV (PDB entry: 5XEJ, [159]) and Pro1 are shown as sticks.

CHAPTER 6: THE C-TERMINAL REGION OF D-DT REGULATES MOLECULAR RECOGNITION FOR PROTEIN-LIGAND COMPLEXES

Published as: Andrew Parkins, Aliyah Veronica R. Pilien, Alexander M. Wolff, Christopher Argueta, Jasmine Vargas, Shahrzad Sadeghi, Andreas H. Franz, Michael C. Thompson, Georgios Pantouris; The C-terminal Region of D-DT Regulates Molecular Recognition for Protein-Ligand Complexes. *Journal of Medicinal Chemistry* ***2024*** *67 (9), 7359-7372*

DOI: 10.1021/acs.jmedchem.4c00177

Abstract

Systematic analysis of molecular recognition is critical for understanding the biological function of macromolecules. for the immunomodulatory protein D-dopachrome tautomerase (D-DT), the mechanism of protein-ligand interactions is poorly understood. Here, seventeen carefully designed protein variants and wild type (WT) D-DT were interrogated with an array of complementary techniques to elucidate the structural basis of ligand recognition. Utilization of a substrate and two selective inhibitors with distinct binding profiles offered previously unseen mechanistic insights into D-DT – ligand interactions. Our results demonstrate that the C-terminal region serves a key role in molecular recognition via regulation of the active site opening, protein-ligand interactions, and conformational flexibility of the pocket's environment. While our study is the first comprehensive analysis of molecular recognition for D-DT, the findings reported herein promote understanding of protein functionality and enable the design of new structure-based drug discovery projects.

Introduction

D-DT, also known as MIF-2, is a small proinflammatory cytokine[9] and enzyme[129] that regulates immune system responses. Encoded by chromosome 22, D-DT is primarily localized in the cytosol of several tissues and cells,[14, 137] however in response to stimulation, it is trafficked to the extracellular space or nucleus.[9, 15] The protein was originally described in 1993 as a tautomerase enzyme that converts D-dopachrome into 5,6-dihydroxyindole (DHI).[116] Crystallographic studies performed a few years later showed that D-DT shares structural homology with the previously identified tautomerase enzyme called macrophage migration inhibitory factor (MIF).[90] The structural and functional overlap between MIF and D-DT, as well as their adjacent gene localization, led to the premature conclusion that D-DT was a redundant protein. for this reason, all the attention of the community was given to MIF for almost two decades. However, the accumulating evidence collected over the last decade suggests a distinctive role of D-DT in human pathophysiology[82, 83, 137, 138] and promotes the need for novel mechanistic studies to understand molecular recognition of D-DT complexes.

Similar to MIF,[27, 36, 131] D-DT binds and activates the chemokine receptor CXCR7 (ACKR3)[139] and the type II cell surface receptor CD74.[9] Activation of CXCR2 by MIF, is shown to be associated with the pseudo Glu-Leu-Arg (ELR) motif of the protein,[12, 23] which in the case of D-DT is missing. Experimental findings also showed MIF to activate CXCR4[12, 24, 81] but such knowledge is missing in the case of D-DT. Depending on the cell type, D-DT-induced signaling is associated with multiple pathways including MAPK/ERK,[9, 160] NF-κB,[161] COX2/PGE_2,[162] and AMPK.[141] Activation of these pathways has a downstream effect on inflammation, cardiovascular disorders, autoimmune diseases, and cancer. While neither the structure of MIF – CD74 nor D-DT – CD74 complexes are available, structure-activity, protein dynamics, and

mutational studies performed on MIF have provided key information on the MIF-induced activation of CD74.[48, 51, 53, 88, 91] In contrast, D-DT experimental and computational findings related to CD74 activation are limited, and obtained in consideration of the protein's homology with MIF.[20, 52, 163]

Omission of the initial methionine results in a D-DT monomer that is composed of only 117 amino acids. The secondary structure of D-DT forms a four stranded b-sheet packed against two a helices.[90] Two additional, identical subunits come together to form a globular homotrimer as the biological assembly. A pocket between adjacent monomers makes up the active site of this enzyme. The tautomerase activity of D-DT is controlled by the N-terminus Proline (Pro1).[116, 129] While the biological substrate of D-DT remains unknown, D-dopachrome and 4-hydroxyphenylpyruvate (4-HPP) have been primarily used as model substrates.[44, 116, 129, 164, 165] Via this enzymatic activity, four tautomerase inhibitors of D-DT have been identified thus far. The suicide inhibitor 4-iodo-6-phenylpyrimidine (4-IPP) was the first reported inhibitor of D-DT that also inhibits MIF.[57, 166] 4-IPP covalently binds to Pro1, forming a 6-phenylpyrimidine (6-PP) adduct. The first selective and reversible inhibitor of D-DT, known as 4-(3-carboxyphenyl)-2,5-pyridinedicarboxylic acid (4-CPPC), was published a few years later.[59] 4-CPPC binds via an induced fit mechanism that displaces the C-terminal residues Met114-Leu117, whilst increasing their conformational flexibility.[49] The conformational flexibility of this region was further investigated with molecular dynamics (MD) simulations and nuclear magnetic resonance (NMR) spectroscopy revealing structural features of D-DT that had never been observed in any prior MIF studies.[167] While the enhanced conformational flexibility of the C-terminal residues does not destabilize the biological assembly of D-DT, 4-CPPC may have a mechanistic value for studying the potential functional role of the C-terminal tail. Another selective and reversible

inhibitor of D-DT, a derivative of thieno[2,3-d]pyrimidine-2,4(1H,3H)-dione, was reported to have higher potency and selectivity than 4-CPPC and was shown to inhibit the proliferation of non-small cell lung cancer (NSCLC) cells.[44] Pyridine-2,5-dicarboxylate is the most recently reported inhibitor of D-DT, and similar to the previous two, reversibly binds the active site of the protein.[165] Besides the fact that this molecule inhibits the D-DT-induced activation of CD74 *in vivo,* it also demonstrates the highest reported inhibition selectivity, 79-fold, for D-DT over MIF. The high selectivity, micromolar potency, and bioactivity make pyridine-2,5-dicarboxylate a suitable molecule for D-DT functional studies.

In this study, we utilized high resolution protein crystallography along with several complementary biophysical, biochemical, and computational methods to provide the structural basis of D-DT – ligand complex formation. Towards this goal, we generated seventeen D-DT variants and via their unique structural properties, investigated the mechanistic details that control ligand binding. As ligands, we employed the most commonly reported substrate, 4-HPP, and two selective inhibitors (4-CPPC and pyridine-2,5-dicarboxylate) that bind to D-DT in a distinct fashion. Our study provides the first comprehensive analysis of the structural parameters that control formation of D-DT – ligand complexes and demonstrates that the C-terminal region is a key regulator of molecular recognition via a multi-tier control mechanism.

Materials and Methods

General procedures.

Production of D-DT variants was accomplished following the QuikChange site-directed mutagenesis protocol (Agilent Technologies, Santa Clara, CA). The DNA oligos used for mutagenesis are described in supporting information (**Table S9**). with the exception of L117G/F116G/T115G/M114G, all the remaining variants were synthesized using pET-22b(+)-

WT D-DT as a template. The quadruple glycine mutant was produced, utilizing pET-22b(+)-L117G/F116G/T115G as a template. The polymerase chain reaction (PCR) products were digested with Dpn I (New England Biolabs) and transformed into XL10-Gold ultracompetent cells (Agilent Technologies, Santa Clara, CA). All D-DT variants were confirmed by sequencing. Protein purity was analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-Page). Protein concentration was determined by the Pierce BCA Protein Assay Kit (Thermo Fisher Scientific) following the manufacturer's protocol. CD experiments were carried out in a JASCO J-815 spectropolarimeter. The mass spectrometry experiments were performed in a Kratos-Shimadzu Axima-CFR MALDI-TOF mass spectrometer. Kinetic assays were carried out in a Tecan Infinite M-Plex microplate reader.

Materials.

All solvents and chemicals were of the highest analytical grade and used without further purification. 4-HPP (>95% purity) was purchased from TCI America. 4-CPPC (99% purity) was obtained from Axon Medchem. Sinapinic acid (>99% purity) was bought from Fluka. Dimethyl sulfoxide (DMSO) and Luria Broth (LB) media were purchased from VWR. Polyethylene glycol (PEG) 3350 and 4000 were obtained from Sigma-Aldrich and TCI America, respectively. Isopropyl β-D-1-thiogalactopyranoside (IPTG) and ampicillin sodium salt were purchased from Gold Biotechnology. All tools and consumables required for protein crystallography were purchased from Hampton Research.

Protein expression and purification.

The D-DT variants were expressed and purified in a similar manner as previously described.[165, 168] The plasmid encoding the protein of interest was transformed into BL21(DE3) competent *E. coli* cells (Agilent Technologies). The cells were then grown in 1 L of LB media

enriched with 100 μg/mL of ampicillin, under shaking at 37 ˚C, until they reached an optical density at 600 nm (OD_{600}) of 0.6-0.8. At this point, protein expression was induced by the addition of 1 mM IPTG, and the cells were incubated for an additional 4 hours at the same temperature. The cells were then collected by centrifugation, washed with the lysis buffer (20 mM Tris.HCl pH 8.5, 20 mM NaCl), and stored in 50 mL centrifuge tubes. Cell pellets of the poly-glycine mutants were lysed in the lysis buffer adjusted to pH 8.0, while all the remaining variants, including WT D-DT, were lysed in the regular lysis buffer (20 mM Tris.HCl pH 8.5, 20 mM NaCl). WT D-DT, single-point alanine mutants, V113 mutants, and poly-glycine variants were loaded onto a 5 mL Q-Sepharose column that was equilibrated in their respective lysis buffers. The poly-glycine variants came off in the flow-through after the lysate peak, meanwhile WT D-DT and the other variants bound to the column and eluted with 5% of the elution buffer (20 mM Tris.HCl pH 8.5, 1 M NaCl). The truncation variants were loaded onto a 5 mL Q-Sepharose column connected in series to a 120 mL Q-Sepharose column. Once the protein was loaded, 10% of the elution buffer was introduced, and the protein came off in the flow though. In all cases, fractions containing the protein of interest were collected, and concentrated for further purification by size exclusion chromatography (SEC) using a 16/60 Superdex 75 column (Cytiva, Marlborough, MA). for this step, 20 mM Tris.HCl pH 7.4, 20 mM NaCl was always used as a running buffer. Monodisperse fractions containing >95% pure protein were collected, concentrated, and stored at -80 ˚C for further use.

Circular dichroism (CD) spectroscopy.

All experiments were performed in a 1 mm quartz cuvette, while the folding profiles of D-DT variants (20 mM) were recorded between 260 and 195 nm, as described before.[91] Prior to each experiment, the running buffer composed of 20 mM sodium phosphate pH 7.0 and 1mM

EDTA was degassed with ultrahigh purity nitrogen. The triplicate data sets were analyzed in GraphPad Prism 10.

Kinetic experiments.

Keto-enol tautomerase assays were performed in a similar manner to what has been previously published.[168] Briefly, a 30 mM stock of 4-HPP stock was prepared in 0.5 M ammonium acetate pH 6.2 (4-HPP buffer) and incubated overnight at room temperature while rocking. The stock was then diluted with the 4-HPP buffer to create a final concentration range of 0-2 mM. The substrate solution was added to a 96-well microplate and a borate solution with a working concentration of 0.42 M was introduced. The reaction was then initiated by the addition of protein at a final concentration of 250 nM. Formation of the enol-borate complex (ε_{306}=11400 $M^{-1}cm^{-1}$) was monitored at 306 nm, for 300 seconds and having 10 second intervals. The data was analyzed using GraphPad Prism 10. All experiments were performed in triplicate.

Protein crystallization.

Crystallization of D-DT variants was performed in 24-well hanging drop plates. All proteins were concentrated to 12 mg/mL except for T115A, which was concentrated to 10 mg/mL. Proteins and mother liquor were mixed at a 1:1 (v/v) ratio. The crystallization conditions of D-DT variants is as follows: WT D-DT: 24-30% PEG 4000, 0.1 M sodium citrate pH 5.6-6.0, 0.2 M ammonium acetate; V113N: 26-32% PEG 4000, 0.1 M sodium citrate pH 5.2-5.8, 0.2 M ammonium acetate; D114-117: 26-32% PEG 3350, 0.25-0.275 M ammonium acetate, D109-117: 26-28 % PEG 4000, 0.1 M sodium citrate pH 6.0-6.2, 0.2 M ammonium acetate, T115A: 28-30% PEG 4000, 0.1 M sodium citrate pH 5.8-6.2, 0.2 M ammonium acetate. Crystals, which varied in shapes, were formed at different time points, and reached their full size within two weeks. Co-crystal structures of D-DT variants with the selective inhibitor 4-CPPC were acquired by soaking

the apo crystals into pre-equilibrated drops of mother liquor enriched with 100 mM 4-CPPC. Similar to what had been previously used,[49] the high concentration of PEG 4000 served as the cryoprotectant for the D-DT crystals.

X-ray data collection and crystal structure determination.

Single-crystal X-ray diffraction data was collected by the Advanced Light Source (ALS) beamline 8.3.1 (wavelength of 1.11583 Å) of the Lawrence Berkeley National Lab, which is equipped with a Dectris Pilatus3 S 6M detector. Data reduction was carried out using DIALS[169] as implemented via the xia2 software package.[170] The crystal structures of D-DT variants were obtained via molecular replacement using the CCP4-supported program, Molrep.[72] In all cases, WT D-DT (PDB entry: 1DPT) served as the search structure. After obtaining the initial model, each variant underwent further refinement using Refmac5[146] and COOT,[147] while the structures were visualized in PyMOL.[73] 4-CPPC coordinates and crystallographic information file (CIF) were generated by the CCP4-supported program, PRODRG.[171] RMSD of finalized crystal structures were determined by Superpose (CCP4 supported). In the cases of V113N, Δ114-117, and Δ109-117, RMSD calculations were performed per monomer to obtain the most accurate alignment and averaged in order to obtain the final values for the homotrimeric assembly. The crystallographic tables are provided in the supplemental material. Ramachandran analyses showed 0% outliers and all residues in favored regions.

MALDI-TOF MS analysis.

All experiments were carried out in linear mode with a 120 cm flight tube in a turbo molecular vacuum of approximately 10^{-6} torr. for ionization, an N_2-laser (l = 337 nm, 120 mJ/pulse) was used along with manual X-Y-stage adjustment for optimal ion formation. The laser attenuator was adjusted to a threshold where ion formation was just beginning. Ions were

extracted with a 20 kV potential after a 350 ns collisional cooling delay. for instrument calibration, protein standards were used (ProteoMass^{TM}, Sigma-Aldrich). The calibration sample (1 mL, 20 pmol of each protein) was crystallized on the plate with sinapinic acid (2 mg/mL in MeCN/0.1% TFA 1:1, 1 mL). The same matrix was used for analysis of D-DT protein samples whose concentrations were at 40 mg/mL (1 mL spotted = 40 ng of protein, approximately 3-5 pmol on plate). Data files were processed with Shimadzu Biotech Launchpad software (ver. 2.9.3.2011624). Both calibration spectra and experimental spectra were smoothed to show average *m/z* values of all ion peaks.

MD Simulations.

MD simulations were carried out similarly to what was previously described.[91, 167] All structures for the simulations were prepared from the WT D-DT structure (PDB entry: 1DPT). The additional, non-protein atoms were first removed from the structure and the biological assembly was generated using PyMOL. Mutations and truncations were made using UCSF Chimera.[67] The structures were used as inputs for psfgen, a VMD plug-in,[69] to generate a new PDB/PSF pair containing hydrogens. The PSFs and PDBs for each structure were input into VMD's solvate plug-in to create another PSF/PDB pair with the protein inside a water box. The charge of the system was checked and balanced using the autoionize plug-in. Each system was minimized, heated to 300K, and equilibrated before the 200 ns production run. The RMSF data of the α-carbons was produced by GROMACS.[71] The simulations and subsequent analyses were performed in duplicate.

Statistics.

The experiments shown in this study were performed in triplicates unless otherwise stated. Error is shown as ±SD (standard deviation) from the mean value. Statistical analysis was

performed using the unpaired *t* test (GraphPad Prism 10), while statistical significance is expressed as * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Results and Discussion

Selection of Protein Variants.

Our previous work, utilizing MD simulations and NMR, revealed an enhanced flexibility of D-DT's C-terminal residues when bound to the selective and reversible inhibitor, 4-CPPC. As binding of this ligand did not compromise the biological assembly of D-DT, it raises the question whether C-terminal conformational changes are a key aspect for molecular recognition. Thus, various mutations were made on the C-terminal residues to understand the biological significance of this flexibility.

Single point alanine (Ala) mutations were first considered. An overlay of WT D-DT (PDB entry: 1DPT) crystal structure onto the corresponding structure of D-DT – 4-CPPC (PDB entry: 6C5F) revealed that Val113 was the last C-terminal residue to be drastically affected by 4-CPPC binding (**Fig. S1**). for this reason, the single point Ala mutations were extended from the C-terminus Leu117 to Val113 (L117A, F116A, T115A, M114A, and V113A).

While the Ala mutations would provide information about the potential functional role of side chains, the C-terminal glycine (Gly) mutations would aid in understanding how backbone flexibility affects ligand recognition and dynamic coupling of residues across the biological assembly. Our previously published correlation analysis of D-DT – 4-CPPC revealed that the structural flexibility of the C-terminal tail enables long-range intra-/intersubunit communications across D-DT.[167] Whether these findings are associated with the functionality of D-DT, it remains unclear. Thus, the Gly variants would offer a valuable tool for exploring the relationship between protein dynamics and catalytic activity. Towards this goal, we produced four variants (L117G,

L117G/F116G, L117G/F116G/T115G, and L117G/F116G/T115G/M114G). Production of poly-Gly variants stopped at Met114 considering the crystal structure of D-DT – 4-CPPC, which pointed that this was the last residue with enhanced flexibility (**Fig. S1**).[49]

In addition to Ala and Gly variants, which provide information on the functionality of side chains and backbone, respectively, we also produced three C-terminal truncations (Δ114-117, Δ109-117, and Δ104-117). These variants were carefully selected to avoid truncating any major secondary structure of D-DT (**Fig. S2A**). The rationale for producing these variants was that a stepwise deletion of C-terminal residues would gradually increase the size of the active site opening, allowing unrestricted accessibility to the catalytic pocket. Therefore, if conformational changes are required for ligand recognition, it would become apparent. We employed the crystal structure of WT D-DT to model the active site openings for Δ114-117, Δ109-117, and Δ104-117 (**Fig. S2B**). A side-by-side comparison of WT D-DT, and the three modeled truncations, confirmed that the size of the active site opening progressively increases, from WT D-DT to Δ104-117. Considering these alterations would not drastically impact the stability of the D-DT trimer, they would offer an alternative approach to explore the potential implementation of an induced-fit binding mechanism for D-DT.

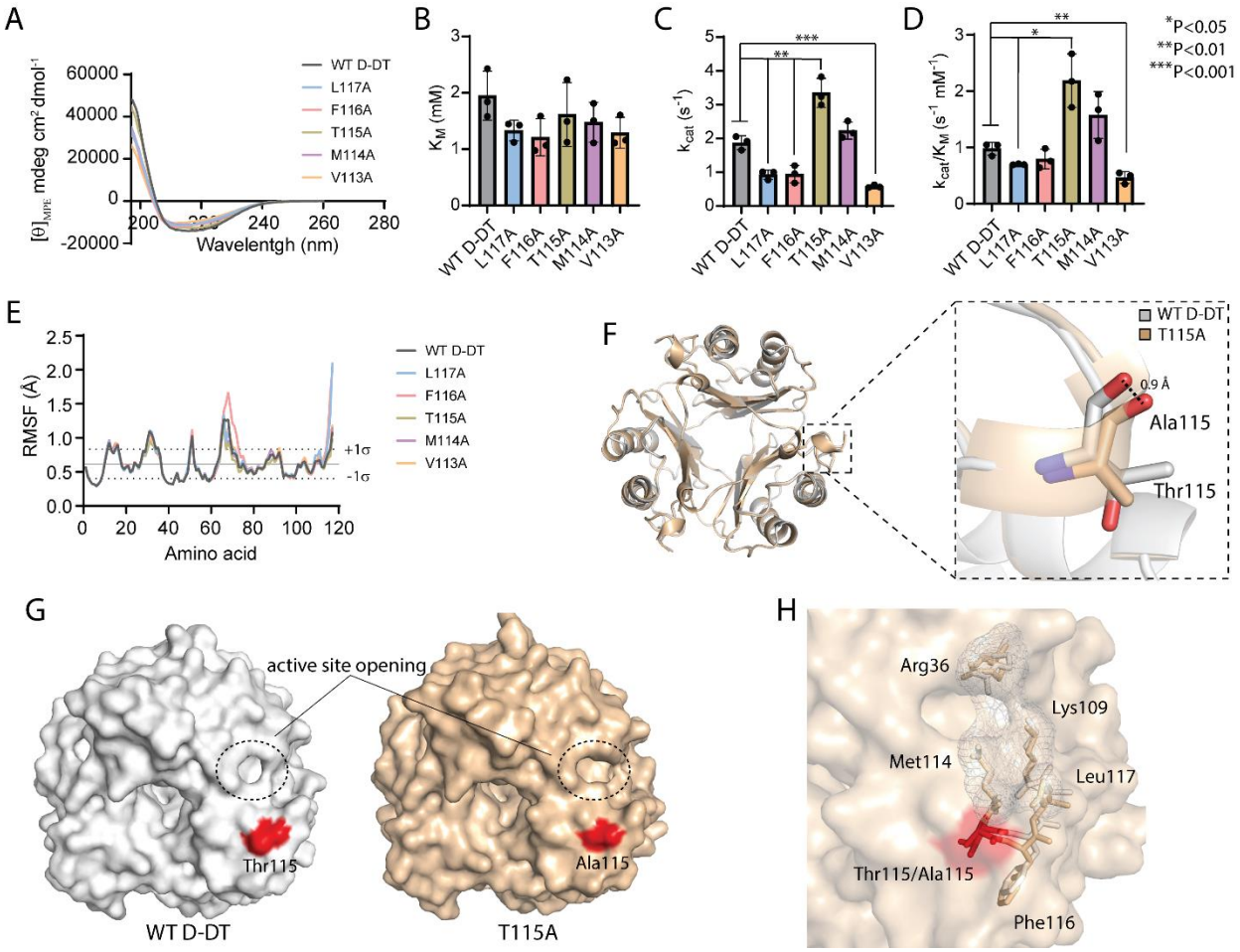
Characterization of Ala variants.

Prior to the kinetic characterization of single point Ala mutants, we performed folding experiments using circular dichroism (CD) spectroscopy. The five variants (L117A, F116A, T115A, M114A, and V113A) and WT D-DT were analyzed side-by-side, demonstrating similar secondary structural profiles (**Fig. 1A**). Following, we employed the keto-enol tautomerization assay to assess the impact of these mutations on the enzymatic activity of D-DT. for our experiments, 4-HPP was used as a substrate.[168] Three biological replicates were performed and

the Michaelis-Menten parameters were analyzed (**Figs. 1B-D**). Truncation of side chains for residues 113-117 did not have any drastic impact on the binding affinity of 4-HPP, as shown by the Michaelis constant (K_M) values (**Fig. 1B**, **Table S1**). In contrast, the turnover numbers (also known as k_{cat}) were noticeably impacted (**Fig. 1C**, **Table S1**). Having WT D-DT (k_{cat}=1.88±0.20 s^{-1}) as the point of reference, we observed two trends. While L117A, F116A, and V113A demonstrated a reduction of their turnover numbers, T115A showed a surprising increase. for simplicity, we converted the k_{cat} values to percent (%) change using the following formula: [(value $_{variant}$ - value $_{WT\ D\text{-}DT}$) / value $_{WT\ D\text{-}DT}$] × 100. L117A and F116A with corresponding k_{cat} values of 0.93±0.14 s^{-1} and 0.95±0.26 s^{-1} exhibited a similar drop of ~50%. V113A, with a k_{cat} value of 0.58±0.04 s^{-1}, was the variant being affected the most, as shown by the 69% reduction of its turnover value. In contrast, T115A (k_{cat}=3.35±0.43 s^{-1}) increased its ability to turnover 4-HPP by 78% (**Fig. 1C**, **Table S1**). The catalytic efficiencies (k_{cat}/K_M) of WT D-DT (0.98±0.11 s^{-1} mM^{-1}) and the Ala variants were also investigated (**Fig. 1D**, **Table S1**). V113A (0.46±0.10 s^{-1} mM^{-1}) and T115A (2.18±0.48 s^{-1} mM^{-1}) were the two mutants that demonstrated the highest deviation from the k_{cat}/K_M value of WT D-DT. The side chain mutation of residue 113 reduced the k_{cat}/K_M value of V113A by 53%, whereas in the case of T115A, mutagenesis created a more efficient enzyme yielding an impressive increase of 122%.

Figure 6.1.

Biophysical, biochemical, and structural characterization of Ala variants.



Note. **A)** The folding profiles of WT D-DT and the five Ala variants (L117A, F116A, T115A, M114A, and V113A) were monitored by CD spectroscopy. **B-D)** Steady-state kinetic parameters of D-DT variants using 4-HPP as a substrate. The experiments were carried out in triplicate (n=3) and the error values are shown as standard deviations. **E)** RMSF profiles of D-DT variants produced by 200 ns trajectories. The average RMSF value of WT D-DT (0.62±0.22 Å) is shown as a solid black line. The black dashed lines represent one standard deviation (±1σ) from the mean value of WT D-DT. **F)** Superposition analysis of WT D-DT (light grey) and T115A (light brown) crystal structures. The point mutation caused a 0.9 Å conformational shift of residue 115.

(Figure 6.1 Continued)

Thr115 and Ala115 are shown as light grey and light brown sticks, respectively. **G)** Comparison of the active site openings for WT D-DT (light grey) and T115A (light brown). The position of Thr115 and Ala115 in the corresponding crystal structures of WT D-DT and T115A are highlighted in red. **H)** Conformational changes, induced by the T115A mutation (red), alter the size of the active site opening. The surface area of Arg36, Lys109, and Met114 that contributes to defining the active site opening of WT D-DT is presented as light grey mesh. The surface area of T115A is shown in light brown. Affected residues are shown as sticks. The crystal structure of WT D-DT (PDB entry: 1DPT) served as the search model for T115A (PDB entry: 8VG7).

We performed MD simulations to explore the effect of side chain truncations on the global (biological assembly) and local (around the point mutation) conformational flexibility of D-DT. Our analysis showed that the Ala mutants and WT D-DT have similar root-mean-square-fluctuation (RMSF) values (**Fig. 1E**, **Table S2**). While the average RMSF value of WT D-DT was determined at 0.62±0.22 Å, the corresponding values of the five variants ranged from 0.59±0.18 Å (T115A) to 0.67±0.31 Å (F116A) (**Table S2**). Upon inspection of the V113A (Avg. $_{RMSF}$ = 0.63±0.21 Å) and T115A (Avg. $_{RMSF}$ = 0.59±0.18 Å) fluctuation profiles, we did not notice any substantial differences that could potentially explain the kinetic findings for these proteins (**Fig. 1D**).

Besides the global analysis, we also explored the local dynamic effect of these mutations. Noticeable changes in local dynamics were only observed for L117A and F116A (**Fig. 1E**). Specifically, the side chain truncations of Leu117 and Phe116 drastically increased the conformational flexibility of the C-terminal region, which was not surprising considering the

position of these residues. In the case of F116A, the point mutation had an additional effect on residues 66-76, which are located proximal to the C-terminal tail (**Fig. 1E**). Despite these differences, the kinetic findings show that L117A and F116A are two variants of low interest (**Fig. 1D**). Collectively, our findings support that alteration of protein dynamics, due to side chain truncation, did not have any drastic effect on the catalytic activities of Ala variants.

Role of residue 115 in molecular recognition.

T115A was crystallized and structurally inspected considering that it was the C-terminal variant with the most substantial impact on the catalytic activity of D-DT (**Table S3**). Overlay of T115A crystal structure onto WT D-DT (PDB entry: 1DPT) revealed a high superposition agreement with a root-mean-square deviation (RMSD) value of 0.33 Å (**Fig. 1F**, **Table S4**). Closer inspection of the mutated residue exposed a conformational change of 0.9 Å (**Fig. 1F**), which at first glance appears insignificant. However, surface analysis showed that this small shift of residue's 115 position had a notable impact on the size and shape of active site's opening (**Fig. 1G**). In turn, this is explained by conformational changes of residues found surrounding Ala115 (**Fig. 1H**). Upon Ala115 backbone movement, residues Arg36, Lys109, and Met114 adopt distinct side chain conformations that increase the size of the active site opening (**Fig. 1H**). Phe116 and Leu117 are also affected by the T115A mutation, but their conformational changes do not have any direct impact on the opening of the active site.

To extend our understanding, we employed two selective D-DT inhibitors, 4-CPPC and pyridine-2,5-dicarboxylate, to carry out inhibition assays for WT D-DT and T115A. for reference, the previously reported inhibition constant (K_i) values of 4-CPPC and pyridine-2,5-dicarboxylate for WT D-DT are 33±0.7 μM[49] and 17.2±0.9 mM,[165] respectively. Notably, pyridine-2,5-dicarboxylate is a derivative of 4-CPPC that demonstrates enhanced inhibition

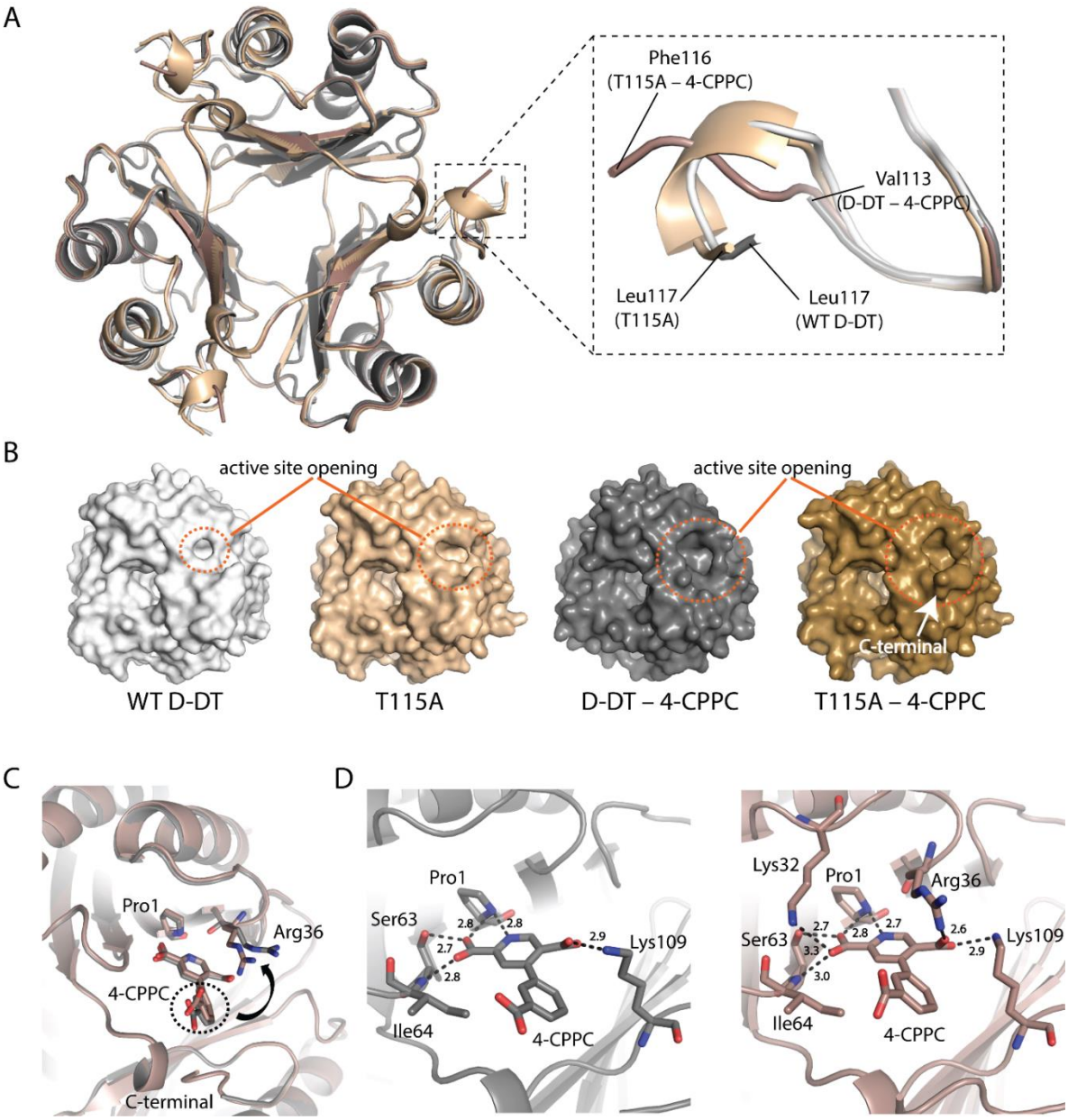
selectivity and potency, over the maternal molecule.[165] Despite their structural similarities, we anticipated that utilizing both inhibitors would be valuable for understanding the mechanism of recognition by D-DT, since they each have a distinct impact on the C-terminal residues. for 4-CPPC, formation of the protein-inhibitor complex is associated with major conformational changes of the C-terminal region.[49] Contrary to 4-CPPC, binding of pyridine-2,5-dicarboxylate does not affect the conformation of the C-terminal residues.[165] In the presence of 4-CPPC, the corresponding K_i values of WT D-DT and T115A were determined at 37.7±5.6 μM and 36.4±3.3 μM (**Fig. S3**). for pyridine-2,5-dicarboxylate, the K_i values of 17.1±0.2 μM and 18.5±1.2 μM were obtained for WT D-DT and T115A, respectively (**Fig. S4**). Our findings confirmed previously published K_i values of 4-CPPC[49] and pyridine-2,5-dicarboxylate[165] against WT D-DT and demonstrated that the single point mutation of Thr115 had neither a beneficial, nor detrimental, effect on the inhibition potencies of these two inhibitors.

We co-crystallized T115A – 4-CPPC (**Table S3**) to examine the structural details of these interactions. with the crystallographic findings of D-DT – 4-CPPC in mind,[49] we considered this ligand as the best one to investigate the molecular recognition features of D-DT. Alignment of T115A – 4-CPPC onto the crystal structures of WT D-DT, T115A, and D-DT – 4-CPPC (PDB entry: 6C5F) demonstrated high superposition agreement with RMSD values of 0.17 Å, 0.29 Å, and 0.23 Å, respectively (**Fig. 2A**, **Table S4**). In the absence of global structural differences, we focused on the local impact that 4-CPPC had on the C-terminal residues and the opening of the active site. A close inspection of T115A – 4-CPPC C-terminal region revealed electron density up to Phe116. Upon binding of 4-CPPC, the C-terminal tail of T115A adopts a unique conformation that noticeably differs from those of WT D-DT, apo T115A, and D-DT – 4-CPPC (**Fig. 2A**). Surface analysis, with emphasis on the catalytic pocket, demonstrated similar active

site openings for D-DT – 4-CPPC and T115A – 4-CPPC yet remarkably different surface scaffolds (**Fig. 2B**). These differences are associated with the C-terminal residues 114-116, which are missing from the D-DT – 4-CPPC crystal structure but are present in the structure of T115A – 4-CPPC (**Fig. 2A**).

Figure 6.2.

Crystallographic analyses of T115A and T115A – 4-CPPC.



Note. **A)** Evaluation of the superposition agreement between WT D-DT (light grey), T115A (light brown), D-DT – 4-CPPC (dark grey), and T115A – 4-CPPC (dark brown). The C-terminal region illustrates differences in the conformational flexibility of the four proteins. The C-terminus residues are pointed with solid lines. **B)** Side-by-side comparison of the active site

(Figure 6.2 Continued)

openings (orange dashed circles), of the four D-DT variants. The C-terminal region of T115A – 4-CPPC is pointed with a white arrow. **C)** Conformational changes of 4-CPPC's benzoic ring (dashed circle) affect the conformation of the active site residue Arg36 (black arrow). Pro1, Arg36 and 4-CPPC from the crystal structures of D-DT – 4-CPPC (dark grey), and T115A – 4-CPPC (dark brown) are shown as sticks. **D)** Hydrogen bonding interactions (dashed lines) between active site residues and 4-CPPC, as noted in the crystal structures of D-DT – 4-CPPC (left), and T115A – 4-CPPC (right). Hydrogen bonding distances are shown in angstrom (Å). The crystal structure of WT D-DT (PDB entry: 1DPT) served as the search model for T115A (PDB entry: 8VG7) and T115A – 4-CPPC (PDB entry: 8VG8).

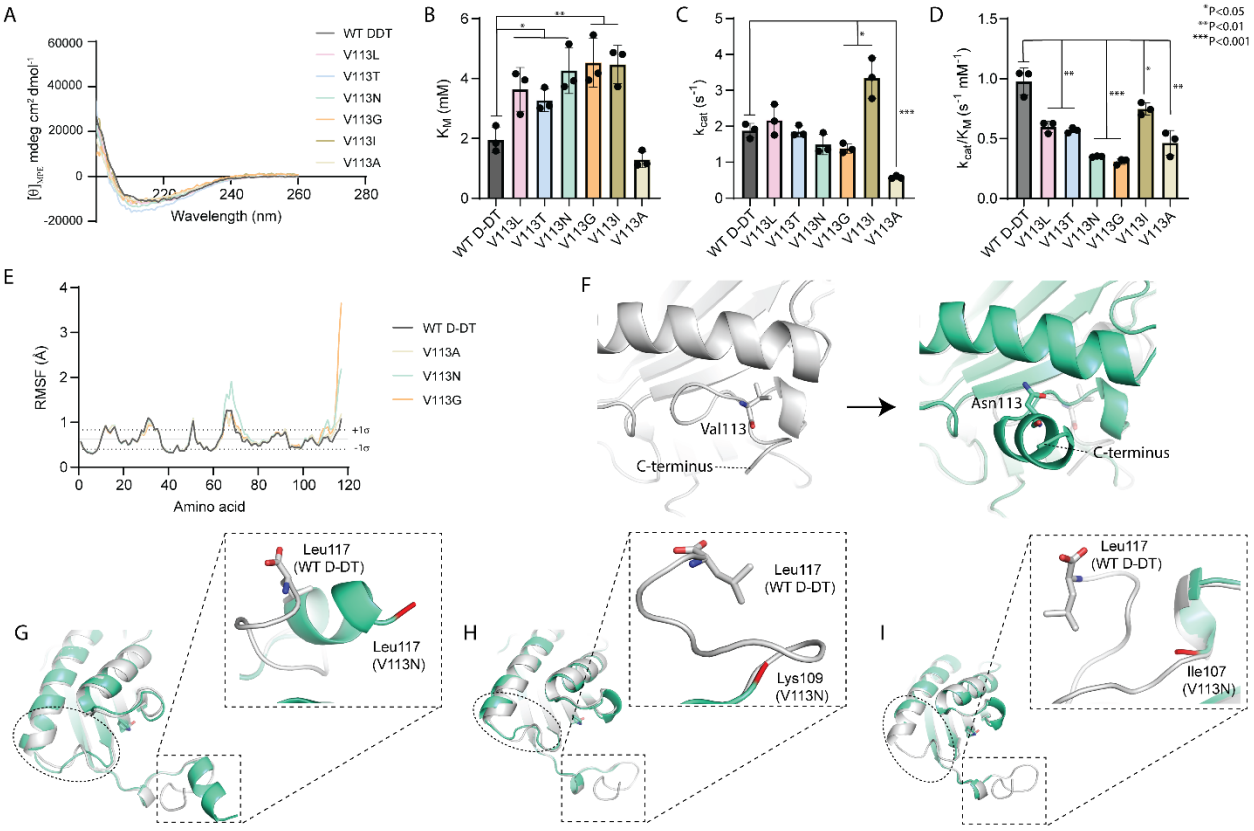
Analyzing the binding motif of 4-CPPC in the active sites of D-DT – 4-CPPC and T115A – 4-CPPC, we observed key differences that expose previously unseen molecular recognition features of D-DT. The C-terminal residues 114-116 (T115A – 4-CPPC) force the benzoic moiety of 4-CPPC to alter its conformation (**Fig. 2C**). Upon this movement, Arg36 undergoes a conformational shift to form a strong hydrogen bonding interaction (H-bond) with the 5-carboxylate group of the molecule's pyridine ring. Notably, the H-bond of 4-CPPC with Arg36 is not observed in the D-DT – 4-CPPC structure (**Fig. 2D**). Considering the role of Arg36 in regulating the size of the active site opening, we conclude that formation of a 2.6 Å H-bond "locks" 4-CPPC in the active site of D-DT. In addition to Arg36, Lys32 was also noted to form a 3.3 Å H-bond with the 2-carboxylate group of the pyridine ring. While all the hydrogen bonding interactions observed in the crystal structure of D-DT – 4-CPPC are retained in T115A – 4-CPPC, the additional H-bonds observed with Arg36 and Lys32 are clearly C-terminal induced.

Role of residue 113 in molecular recognition.

Following up on the kinetic data of V113A, we probed answers regarding the poor turnover number and catalytic efficiency of this specific mutant (**Figs. 1C-D**). While Val113 is packed in a hydrophobic pocket formed between 2 helix and C-terminal 3.6 helix, we considered selective mutations that would potentially provide mechanistic insights. Firstly, we mutated valine to leucine (V113L), isoleucine (V113I), and asparagine (V113N). The rationale behind these mutations was that they would alter the packing properties of residue 113. Besides packing, V113N would also offer information regarding the role of side chain's hydrophobicity. Following, we produced V113T. This variant would alter the hydrophobic character on the side chain, without altering the packing properties of the residue. Lastly, we produced V113G anticipating that this variant would provide information about the role of backbone flexibility.

Figure 6.3.

Impact of Val113 mutation on the structure and function of D-DT.



Note. **A)** The folding profiles of WT D-DT and the six Val113 variants (V113L, V113T, V113N, V113G, V113I, and V113A) were plotted and compared. **B-D)** Impact of Val113 mutagenesis on the Michaelis-Menten parameters of D-DT. The experiments were carried out in triplicate (n=3) and the error values are shown as standard deviations. **E)** Comparison of the RMSF profiles for WT D-DT and the three Val113 variants (V113A, V113N, and V113G) with the smallest k_{cat}/K_M values. The black dashed lines represent one standard deviation ($\pm 1\sigma$) from the mean value of WT D-DT (black solid line). **F)** The crystal structures of WT D-DT (left - grey) and V113N (right - green) were examined, side-by-side, to evaluate the structural impact of V113N mutation. On the right, an overlay of WT D-DT (transparent grey) onto V113N demonstrates the original

(Figure 6.3 Continued)

position of residue 113, C-terminal tail, and the surrounding environment before the mutation. Val113 and Asn113 are shown as sticks. **G-I)** The three subunits of WT D-DT (grey) and V113N (green) were superimposed and compared to observe structural differences associated with the V113N mutation. for each subunit of V113N, the terminus residue is highlighted in red. Leu117, which is the terminus amino acid of WT D-DT, is shown as sticks. The dashed ovals and boxes illustrate conformational differences noted, between WT D-DT and V113N crystal structures, in the 64-77 segment (β4/α2 loop and α2 helix) and C-terminal region, respectively. The crystal structure of WT D-DT (PDB entry: 1DPT) served as the search model for V113N (PDB entry: 8VFW).

The folding profiles of V113G, V113L, V113I, V113T, V113N, and V113A were examined side-by-side with WT D-DT (**Fig. 3A**). While this analysis did not demonstrate any outliers, we proceeded with the kinetic characterization of all variants. Side chain modifications of residue 113 had a noticeable impact on the affinity of D-DT for 4-HPP, as shown by the K_M results (**Fig. 3B, Table S1**). Apart from V113A, all the remaining variants displayed an increase of their K_M values. The highest deviation from the K_M value of WT D-DT was noted with V113G, V113I, and V113N, which exhibited an increase in the range of 120-130% (**Table S1**). These results support that alterations of the packing, flexibility, and/or hydrophobicity of residue 113 have a negative impact on the affinity of D-DT for its substrate. In contrast to the K_M values, the turnover numbers of Val113 variants were not drastically affected except for V113G (26% decrease), V113I (78% increase), and V113A (69% decrease) (**Fig. 3C, Table S1**). Overall, the k_{cat}/K_M values of all variants reduced (**Fig. 3D, Table S1**). V113N and V113G demonstrated the

highest deviation from the k_{cat}/K_M value of WT D-DT with corresponding reductions of 64% and 68%.

Selecting the three variants with the smallest k_{cat}/K_M values (V113N, V113G, and V113A), we performed 200 ns MD simulations and plotted their RMSF profiles (**Fig. 3E, Table S2**). Our approach aimed to display dynamic features of these variants that would potentially explain the kinetic findings. While the average RMSF values of these variants were similar to that of WT D-DT (**Table S2**), we focused on the local environment of the C-terminal region. Mutation of Val113 to either asparagine or glycine enhanced the conformational flexibility of the C-terminal residues 108-117, while V113A did not display noticeable differences from WT D-DT (**Fig. 3E**). In the case of V113N, we also observed an elevated fluctuation for residues 64-77. These residues are located on the β4/α2 loop and α2 helix, which in turn are found in close proximity to the C-terminal region of the adjacent monomer.

From the kinetic (**Figs. 3B-D**) and protein dynamics (**Fig. 3E**) analyses, V113N appears as the variant with the most interesting features. for this reason, we performed protein crystallography to obtain structural insights on this variant (**Table S5**). Superposition analysis of V113N onto the crystal structure of WT D-DT yielded an RMSD value of 0.54 Å (**Table S4**). Upon closer inspection, the Val113 to Asn mutation displaced the residue from its hydrophobic pocket (**Figs. 3F-G**) and increased the C-terminal flexibility (**Figs. 3G-I**). from a protein crystallography perspective, this statement is supported by either the lack of electron density for the affected region or a structural conformation that significantly differs from the corresponding conformation of WT D-DT. Conformational changes of the C-terminal region were observed in one of the three subunits of V113N (**Fig. 3G**). for this subunit, the electron density was clearly observed up to the C-terminus residue, Leu117. In contrast, the remaining two subunits

possessed incomplete electron densities of their C-terminal regions. for subunits two and three, Lys109 (**Fig. 3H**) and Ile107 (**Fig. 3I**) were the last amino acids with complete electron density. In agreement with the MD simulations, WT D-DT and V113N exhibited remarkable differences in the dynamic profile of residues 64-77 (**Figs. 3G-I**). In one of the subunits of V113N, the electron density of residues 65-72 was missing (**Fig. 3I**).

To associate the crystallographic findings of V113N with functional insights, we performed inhibition studies of WT D-DT and V113N in the presence of either 4-CPPC or pyridine-2,5-dicarboxylate. As shown by the inhibition constants of 37.7±5.6 mM (WT D-DT) and 58.1±3.4 mM (V113N), the Val to Asn point mutation did not have a strong impact on the inhibition potency of 4-CPPC (**Fig. S5**). However, the inhibition studies of pyridine-2,5-dicarboxylate led to an opposite conclusion. with reference to WT D-DT findings (17.1±0.2 mM), the inhibition potency of pyridine-2,5-dicarboxylate against V113N essentially vanished at 397±31 mM (**Fig. S6**).

We crystallized V113N with 4-CPPC (**Table S5**) and aligned it onto the crystal structures of WT D-DT, V113N, and D-DT – 4-CPPC (**Fig. 4A**). Superposition agreement between the various structures ranged from 0.48 Å to 0.63 Å (**Table S4**). The highest superposition agreement of 0.48 Å was noted between the crystal structures of V113N – 4CPPC and V113N, while the lowest agreement of 0.63 Å was observed between V113N – 4CPPC and D-DT – 4CPPC. Diverse dynamic profiles of the C-terminal tail were noted upon inspection of the four crystal structures with V113N, V113N – 4CPPC, and D-DT – 4CPPC being characterized by increased conformational flexibility in comparison to WT D-DT (**Fig. 4A**).

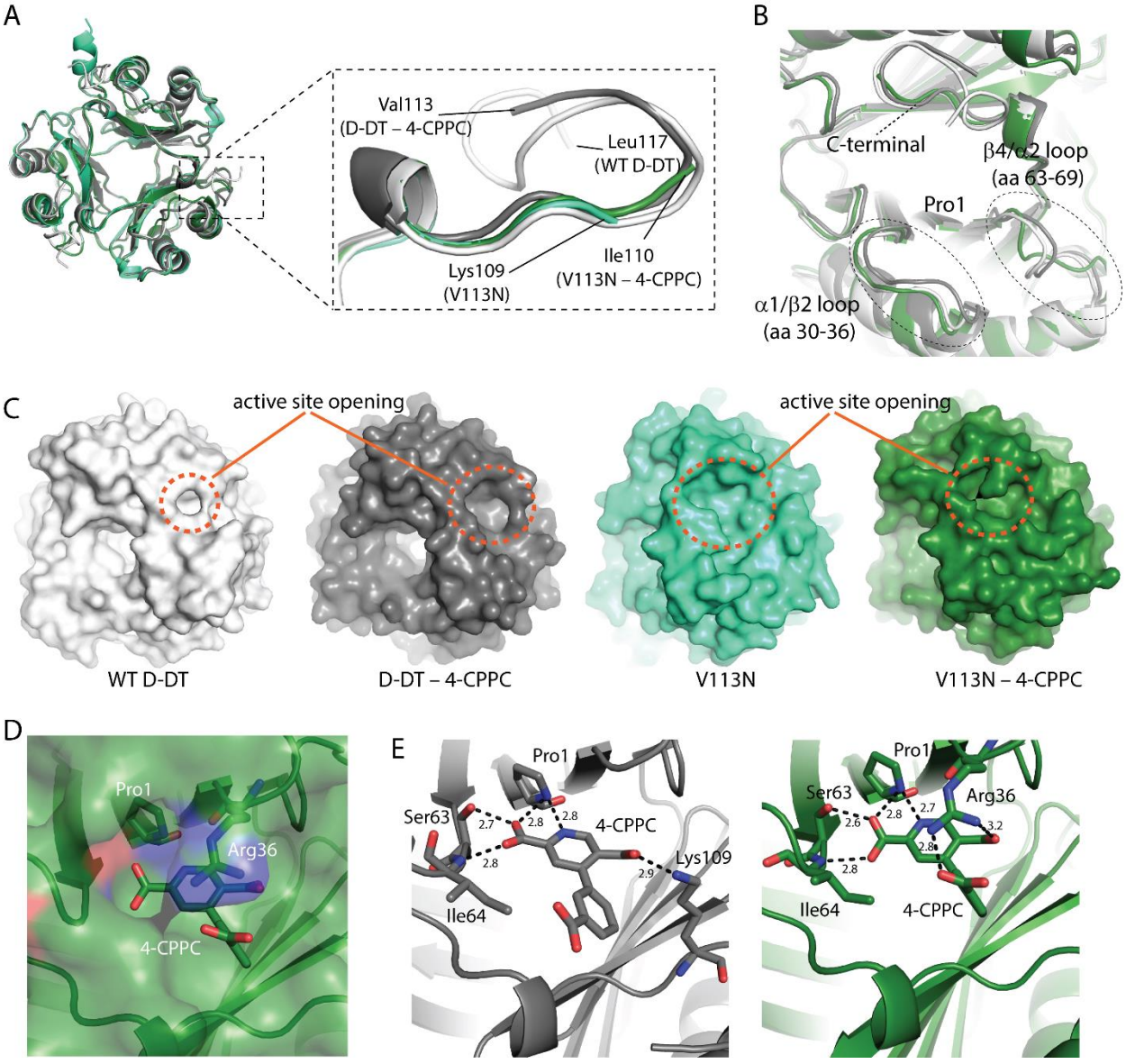
This enhanced flexibility had a clear impact on the $\alpha 1/\beta 2$ (residues 30-36) and $\beta 4/\alpha 2$ (residues 63-69) loops, which are located proximal to the C-terminal tail (**Fig. 4B**). Comparison

of the active site opening of the four proteins fully explains the inhibition findings obtained for 4-CPPC and pyridine-2,5-dicarboxylate. The highly flexible C-terminal region, which is enabled by the V113N mutation, alters the surface scaffold around the active site and converts the binding pocket to a groove (**Fig. 4C**). Such architectural alteration cannot effectively retain pyridine-2,5-dicarboxylate bound to D-DT. In addition, the increased flexibility of Lys109 did not promote stabilization of the protein-ligand complex. As noted in the crystal structure of D-DT – pyridine-2,5-dicarboxylate,[165] Lys109 forms a 2.8 Å hydrogen bonding interaction with the 5-carboxylate group of the molecule. This bond is likely abolished in the case of V113N variant.

In-depth inspection of V113N – 4CPPC revealed previously unrecognized structural features of D-DT. When the inhibitor binds to V113N, Arg36 adopts a novel conformation that "locks" 4-CPPC in the active site pocket (**Fig. 4D**). This conformation of Arg36 enables hydrogen bonding interactions with two of the carboxylate groups of 4-CPPC and supports stabilization of protein-ligand complex (**Fig. 4E**). The 2.8 Å H-bond formed between the benzoic moiety of 4-CPPC and Arg36 has not been observed in any of the earlier analyzed protein-inhibitor crystal structures and offers a fresh perspective of molecular recognition.

Figure 6.4.

Crystallographic analyses of V113N and V113N – 4-CPPC.



Note. **A)** The crystal structures of WT D-DT (light grey), D-DT – 4-CPPC (dark grey), V113N (green), and V113N – 4-CPPC (dark green) were superimposed and compared. Structural differences at the C-terminal region are illustrated. The terminus residue for each protein is pointed by solid lines. **B)** The increased conformational flexibility of the C-terminal tail has an impact on the α1/β2 and β4/α2 loops of D-DT (dashed ovals). **C)** Comparison of the active site

(Figure 6.4 Continued)

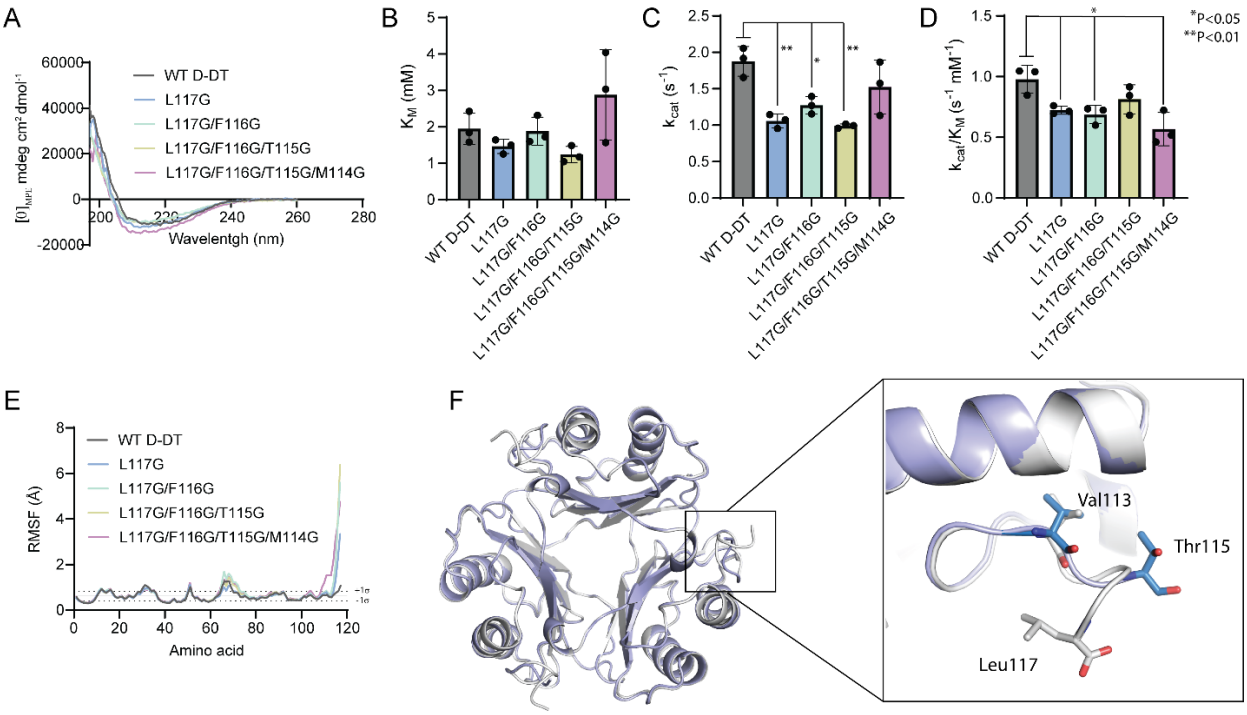
opening (orange dashed circles) of WT D-DT, D-DT – 4-CPPC, V113N, and V113N – 4-CPPC. **D)** Upon binding of 4-CPPC, Arg36 adopts a new conformation that "locks" the ligand in the active site of V113N. Pro1, Arg36, and 4-CPPC are shown as sticks. **E)** Analysis of the hydrogen bonding interactions (dashed lines) between D-DT and 4-CPPC (sticks), as noted in the crystal structures of D-DT – 4-CPPC (left) and V113N – 4-CPPC (right). Hydrogen bonding distances are shown in angstrom (Å). The active site residues involved in hydrogen bonding interactions are shown as sticks. The crystal structure of WT D-DT (PDB entry: 1DPT) served as the search model for V113N (PDB entry: 8VFW) and V113N – 4-CPPC (8VG5).

Characterization of Gly variants.

To evaluate the role of backbone flexibility in molecular recognition, we first examined whether any regions of D-DT, including the C-terminal tail, are inherently flexible. The crystal structures of WT D-DT, obtained at 290 K and 310 K (**Table S6**), were aligned onto the previously published, cryo structure of WT D-DT (PDB entry: 7MSE - 100 K). The RMSD values of 0.21 Å (WT D-DT – 290 K) and 0.24 Å (WT D-DT – 310 K) demonstrate high superposition agreement with WT D-DT (**Fig. S7**, **Table S4**). Neither the C-terminal tail nor any other region of D-DT revealed spontaneous flexibility at higher temperatures. In light of these findings, we induced backbone flexibility by introducing Gly mutations. Single (L117G), double (L117G/F116G), triple (L117G/F116G/T115G) and quadruple (L117G/F116G/T115G/M114G) Gly variants were produced, keeping the crystallographic findings of D-DT – 4-CPPC in mind.[49]

Figure 6.5.

Characterization of Gly variants.



Note. **A)** Folding profiles of WT D-DT and the four Gly variants (L117G, L117G/F116G, L117G/F116G/T115G, and L117G/F116G/T115G/M114G). **B-D)** Michaelis-Menten parameters of WT D-DT and the Gly variants. The experiments were carried out in triplicate (n=3) and the error values are shown as standard deviations. **E)** RMSF profiles of WT D-DT and the Gly variants. The black dashed lines represent one standard deviation (±1σ) from the mean value of WT D-DT (black solid line). **F)** Superposition analysis of the crystal structures of WT D-DT (light grey) and L117G (blue). The C-terminus of WT D-DT and L117G are shown as sticks. The crystal structure of WT D-DT (PDB entry: 1DPT) served as the search model for L117G (PDB entry: 8VFO).

CD experiments for WT D-DT and the four Gly variants produced consistent profiles (**Fig. 5A**). Kinetic analyses showed that the enhanced backbone flexibility of the C-terminal

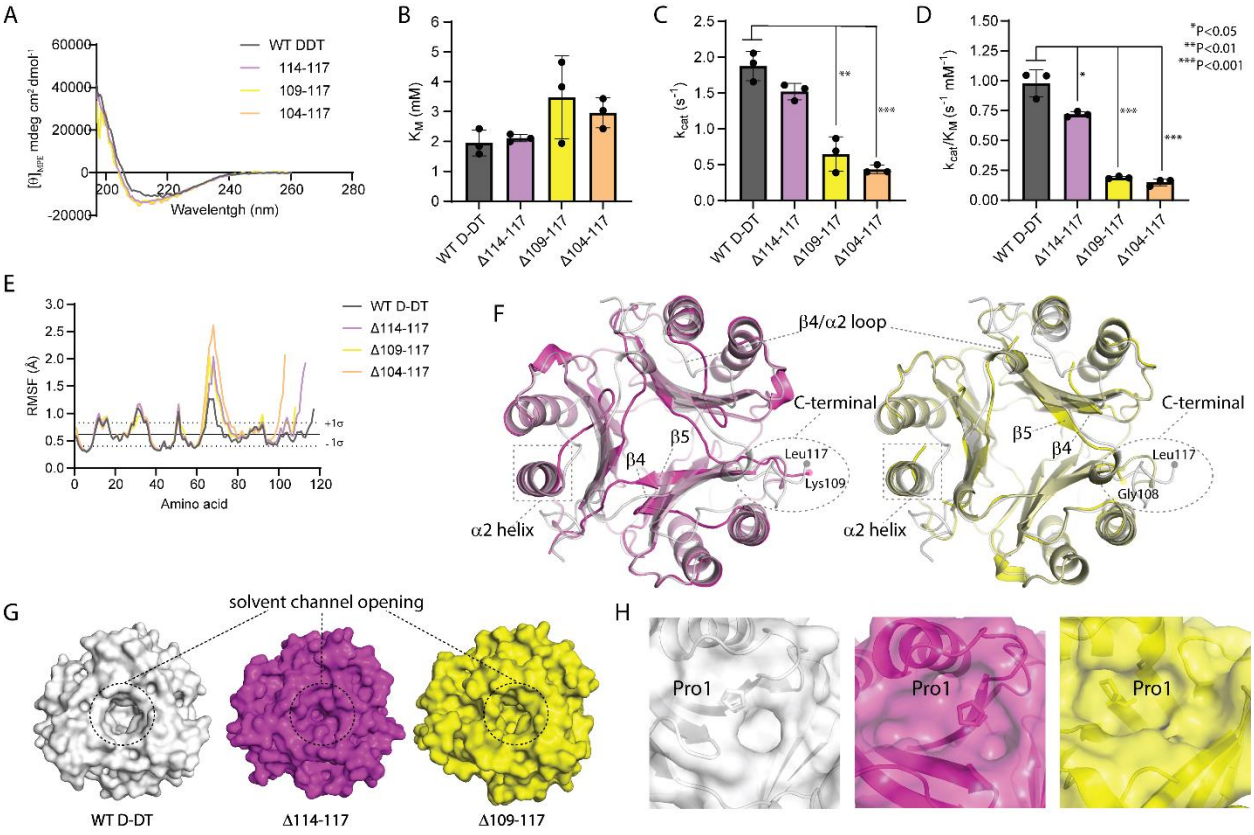
region did not significantly impact the K_M values of Gly variants, something that cannot be said for the k_{cat} and k_{cat}/K_M values (**Figs. 5B-D**, **Table S1**). To couple the kinetic findings with structural insights, we performed MD simulations and analyzed the dynamic profile of these variants. from the single (L117G) to quadruple (L117G/F116G/T115G/M114G) variant, we noted a gradual increase of the C-terminal's flexibility (**Fig. 5E**). The extent (number of residues being affected) and magnitude (RMSF value) of this flexibility are directly correlated with the number of Gly residues being introduced. for example, in the case of L117G, high conformational flexibility was noted for the C-terminal residues Met114-Leu117. When three additional amino acids are mutated to Gly, the number of residues drastically affected increases to eleven (Ile107-Leu117). Interestingly, this enhanced flexibility of the C-terminal region does not have a noticeable effect on the α1/β2 and β4/α2 loops, which are found in proximity (**Fig. 5E**). This can only be explained by accepting that the flexible segment of the C-terminal residues spans outwards from the biological assembly of D-DT to the solvent.

L117G was also crystallized, and its structural features analyzed (**Table S7**). The RMSD value of 0.46 Å (**Table S4**), which was obtained by superposing L117G onto WT D-DT, shows that the two crystal structures have an overall satisfactory agreement with an exception to the C-terminal region (**Fig. 5F**). In agreement with the MD simulations, the single point mutation of Leu117 to Gly amplified the flexibility of residues 114-117 (**Fig. 5E**). Phe116 and the C-terminus Leu117 lacked electron density, while Met114 and Thr115 adopted a distinctively structural conformation from the corresponding amino acids of WT D-DT (**Fig. 5F**). Of great interest is the effect of L117G on Val113. According to the crystallographic findings, the side chain of Val113 remains "locked" in the hydrophobic pocket formed between the α2 helix and C-terminal 3.6 helix (**Fig. 5F**). This finding evidently shows that the conformational flexibility of

the C-terminal region, which is induced by the Gly mutation, extends towards the solvent without affecting the biological assembly of D-DT.

Figure 6.6.

Structural and functional analysis of C-terminal truncation variants.



Note. **A)** The folding profiles of WT D-DT, Δ114-117, Δ109-117, and Δ104-117 were examined by CD spectroscopy. **B-D)** Bar-chat analysis of the Michaelis-Menten parameters for WT D-DT and the C-terminal truncation variants. The experiments were carried out in triplicate (n=3) and the error values are shown as standard deviations. **E)** RMSF profiles of WT D-DT, D114-117, D109-117, and Δ104-117. The black dashed lines represent one standard deviation (±1σ) from the mean value of WT D-DT (black solid line). **F)** The crystal structures of Δ114-117 (left –

(Figure 6.6 Continued)

magenta) and Δ109-117 (right - yellow) were superposed onto the corresponding structure of WT D-DT (transparent light grey). Structural differences at the α2 helix, β4/α2 loop, β4 and β5 strands, and C-terminal region are highlighted. The C-terminus residues of WT D-DT, Δ114-117, Δ109-117 are pointed with filled circles. **G)** Conformational changes of the β4 and β5 strands alter the opening of the solvent channel (black dashed circles). **H)** The C-terminal truncations modify the active site opening and volume, impacting to some extent the accessibility of the catalytic residue Pro1 (sticks). The crystal structure of WT D-DT (PDB entry: 1DPT) served as the search model for Δ114-117 (PDB entry: 8VDY) and Δ109-117 (8VFK).

Characterization of the C-terminal truncation variants.

To complete our investigation, we designed, expressed, and purified three C-terminal truncation variants of D-DT: Δ114-117, Δ109-117, and Δ104-117. Considering the potentially detrimental effect of truncations on protein folding, we first performed CD experiments and analyzed the data side-by-side with WT D-DT. Our findings demonstrated negligible differences across the four samples (**Fig. 6A**). Similar to the Ala and Gly mutants, we then carried out kinetic experiments to explore the impact of C-terminal truncation on the K_M, k_{cat}, and k_{cat}/K_M values of D-DT. Based on our findings, the K_M values of Δ114-117, Δ109-117, and Δ104-117 did not show a significant deviation from the corresponding value of WT D-DT (**Fig. 6B**, **Table S1**). In contrast, the turnover number displayed a clear reduction, with Δ109-117 and Δ104-117 being affected the most. While Δ114-117 reduced its k_{cat} value by only 19%, Δ109-117 and Δ104-117 exhibited the corresponding drops of 65% and 77% (**Fig. 6C**, **Table S1**). On the same trend with the k_{cat} values, the catalytic efficiency reduced from WT D-DT to D104-117. with WT

D-DT as the point of reference, Δ109-117 and Δ104-117 revealed a sharp decline of their k_{cat}/K_M values by 81% and 85%, respectively (**Fig. 6D**, **Table S1**). RMSF analysis of WT D-DT and the three truncation variants, showed diverse dynamic profiles on the C-terminal region and residues 63-79, which are located on the β4/α2 loop (residues 63-69) and α2 helix (residues 70-79) (**Fig. 6E**, **Table S2**). As shown by our data, the conformational flexibility at this region is variant dependent. Residues 70-79 were affected due to their proximity to the C-terminal tail of the adjacent monomer. An increase in the conformational flexibility of residues 70-79, in turn, promotes the conformational flexibility of the β4/α2 loop. Due to the low catalytic efficiency of the truncation variants, inhibition studies were not an option. for this reason, we utilized matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) to probe pyridine-2,5-dicarboxylate binding. Of note, MALDI-TOF MS has been successfully used in protein-ligand studies and it is applicable even for high-throughput screening projects.[172] Following this approach, WT D-DT, Δ114-117, and Δ109-117 were experimentally analyzed showing that only WT D-DT and Δ114-117 bind the inhibitor (**Fig. S8**).

Upon crystallization of Δ114-117 and D109-117, we examined their structural features (**Table S8**). Superposition analyses of Δ114-117 and Δ109-117 onto WT D-DT yielded the corresponding RMSD values of 1.11 Å and 0.54 Å (**Table S4**). Between them, Δ114-117 and Δ109-117 yielded an RMSD value of 0.96 Å. The corresponding electron densities of Δ114-117 and Δ109-117 C-termini were clearly observed up to Lys109 (two out of three subunits) and Gly108 (all three subunits), respectively (**Fig. 6F**). The structural disagreement, noted on superposing the two truncation variants onto WT D-DT, was mainly associated with conformational changes (observed in Δ114-117 and Δ109-117) and/or lack of electron density (observed only in Δ109-117) at three regions: the α2 helix, β4/α2 loop, and C-terminal (**Fig. 6F**).

Due to the increased flexibility of the C-terminal tail, two out of the three subunits of Δ109-117 lacked the electron density of residues found on the α2 helix and/or β4/α2 loop. The enhanced mobility of the β4/α2 loop and C-terminal region, also impacted portions of the β4 and β5 strands, changing their conformations (**Fig. 6F**). The conformational changes observed on the β5 strand altered the opening of the solvent channel, as shown by the comparison of WT D-DT with the two variants (**Fig. 6G**). To associate the crystallography findings of Δ114-117 and Δ109-117 with our original hypothesis (**Fig. S2**), we also examined the surface surrounding the active site. Δ114-117 and Δ109-117 revealed noticeable differences from WT D-DT, as well as distinctions between them (**Fig. 6H**). These differences are evident in the opening and volume of active site as well as the accessibility of the catalytic residue Pro1. Although both Δ114-117 and Δ109-117 increased their active site openings, the volume of the cavity was altered in a distinct way impacting the accessibility of Pro1 (**Fig. 6H**). In Δ114-117, Pro1 appears accessible comparable to WT D-DT. In contrast, Δ109-117 presents Pro1 in a conformation that does not favor catalysis. Together these observations explain the catalytic findings of Δ114-117 and Δ109-117 (**Figs. 6B-D**)

Conclusion

Previous studies focused on MIF [60] reported that the C-terminal region has a key role in stabilizing the protein's tertiary structure as well as modulating the enzymatic activity. These conclusions were attributed to the ability of C-terminal residues to form intersubunit interactions that are important for stabilizing the biological assembly of MIF. More recent studies have shown that this region also has a functional role in the MIF-induced activation of CD74. [48, 51] While these studies enriched our understanding on MIF's structure and function, the corresponding analyses on its human homolog, D-DT, are yet to be reported.

Using a systematic approach, we probed understanding of the role of C-terminal region in D-DT functionality. Seventeen protein variants were interrogated with high-resolution protein crystallography, biophysical assays, turnover and inhibition experiments, MD simulations, and mass spectrometry yielding previously unseen structural features of D-DT that control protein-ligand recognition. Conformational changes of Thr115 serve a key role in the catalytic tautomerization of 4-HPP. Although the binding affinity of 4-HPP is not influenced, threonine-induced movements of the C-terminal region assign distinct conformations to residues that control the opening of the active site; primarily to Arg36. Upon altering the opening of the active site, accommodation of substrate and product release becomes more effective, something that is evident by the impressive k_{cat} and k_{cat}/K_M values of T115A.

for the first time, Arg36 was observed to have multiple functional tasks. Besides regulating ligand admission in the pocket, it is capable of adopting diverse conformations in order to stabilize the protein-ligand complex. This was experimentally observed twice in the crystal structures of T115A – 4-CPPC and V113N – 4-CPPC. Conformational changes of Arg36, as viewed by the crystal structure of V113N – 4-CPPC, protect the protein-ligand complex from the harsh dynamic events occurring proximal to the binding site. Notably, the diverse functionality of Arg36 is regulated by the conformational flexibility of the C-terminal region. Val113, found in the hydrophobic pocket formed between the α2 helix and C-terminal 3.6 helix, appears to be a key residue for controlling the mobility of the C-terminal region. Mutation of Val to Asn displaces residue 113 from its hydrophobic pocket and increases the flexibility of C-terminal residues. Although this region is not inherently flexible, the induced flexibility of the C-terminal by Asn113 can negatively affect catalysis via altering the conformational flexibility of the α2 helix and β4/α2 loop. A highly dynamic C-terminal tail primarily affects the α2 helix of

the adjacent monomer due to their proximity. Consequently, the β4/α2 loop elevates its dynamic activity, influencing ligand binding and/or catalysis (e.g., active site opening/volume and accessibility of Pro1). Future studies focused on elucidating the catalytic mechanism of 4-HPP tautomerization would further explain the effect of C-terminal variants in the k_{cat} and K_M values, while QM/MM calculations may add value by exposing any potential transition states. [173] Collectively, our work offers the first comprehensive analysis of the structural features that regulate molecular recognition in D-DT and may be used to better understand the catalytic mechanism of this protein as well as to promote drug discovery efforts.

Supplemental Information

Table 6.S1.

Steady-state kinetic parameters of D-DT variants[*].

Variant	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1} mM^{-1})
WT D-DT	1.95±0.43	1.88±0.20	0.98±0.11
L117A	1.33±0.18	0.93±0.14	0.70±0.01
F116A	1.21±0.33	0.95±0.26	0.79±0.17
T115A	1.62±0.57	3.35±0.43	2.18±0.48
M114A	1.48±0.35	2.24±0.25	1.58±0.42
V113A	1.29±0.27	0.58±0.04	0.46±0.10
V113N	4.27±0.77	1.50±0.28	0.35±0.01
V113L	3.64±0.73	2.17±0.43	0.60±0.05
V113I	4.47±0.64	3.34±0.56	0.75±0.05
V113T	3.27±0.38	1.86±0.15	0.57±0.02
V113G	4.53±0.81	1.39±0.13	0.31±0.02
Δ114-117	2.11±0.12	1.52±0.11	0.72±0.02
Δ109-117	3.47±1.39	0.65±0.24	0.19±0.01
Δ104-117	2.96±0.51	0.43±0.06	0.15±0.03
L117G	1.47±0.20	1.06±0.10	0.72±0.03
L117G/F116G	1.88±0.39	1.27±0.12	0.69±0.07
L117G/F116G/T115G	1.24±0.23	0.99±0.03	0.81±0.12
L117G/F116G/T115G/M114G	2.88±1.24	1.52±0.37	0.57±0.14

* Experiments were carried out in biological triplicate (n=3) and the error values are shown as standard deviations.

Table 6.S2.

Root-mean-square-fluctuation (RMSF) analysis of D-DT variants[*].

Variant	RMSF (Å)
WT D-DT	0.62±0.22
L117A	0.64±0.25
F116A	0.67±0.31
T115A	0.59±0.18
M114A	0.61±0.20
V113A	0.63±0.21
V113N	0.70±0.34
V113G	0.68±0.41
Δ114-117	0.72±0.35
Δ109-117	0.69±0.32
Δ104-117	0.78±0.46
L117G	0.64±0.35
L117G/F116G	0.74±0.61
L117G/F116G/T115G	0.73±0.66
L117G/F116G/T115G/M114G	0.79±0.60

[*] The 200 ns MD simulations were carried out in duplicates (n=2). Error is expressed as standard deviations.

Table 6.S3.

Crystallographic data collection and refinement statistics for T115A and T115A – 4-CPPC.

	T115A	**T115A – 4-CPPC**
Data collection		
Space group	H 3	H 3
Cell dimensions		
a, *b*, *c* (Å)	82.28, 82.28, 40.73	81.95, 81.95, 40.81
α, β, γ (°)	90.00, 90.00, 120.00	90.00, 90.00, 120.00
Resolution (Å)	41.14-1.24 (1.26-1.24)*	40.98-1.33 (1.35-1.33)*
R_{sym} or R_{merge}	0.070 (0.269)	0.046 (0.265)
I / σI	14.2 (3.1)	22.4 (5.7)
Completeness (%)	99.8 (96.5)	97.2 (93.9)
Redundancy	4.7 (2.6)	5.1 (5.1)
Refinement		
Resolution (Å)	41.14-1.24 (1.27-1.24)	40.98-1.33 (1.37-1.33)
No. reflections	27635	21599
R_{work} / R_{free}	0.10/0.12	0.11/0.14
No. atoms		
Protein	883	874
Ligand/ion	0	21
Water	99	83
B-factors		
Protein	12.66	13.43
Ligand/ion	-	26.66
Water	31.61	31.08
R.m.s. deviations		
Bond lengths (Å)	0.020	0.016
Bond angles (°)	1.948	1.729

* Values in parentheses are for highest-resolution shell.

Table 6.S4.

Root-mean-square-deviation (RMSD) analysis of D-DT variants.

	Moving Structure	Fixed Structure	RMSD (Å)
WT D-DT at multi-temperature	WT D-DT at 290 K	WT D-DT* at 100 K	0.21
	WT D-DT at 310 K	WT D-DT* at 100 K	0.24
Ala mutations	T115A – 4CPPC	WT D-DT**	0.17
	T115A – 4CPPC	D-DT – 4CPPC	0.23
	T115A – 4CPPC	T115A	0.29
	T115A	WT D-DT**	0.33
Val113 mutations	V113N – 4CPPC	V113N	0.48
	WT D-DT**	V113N – 4CPPC	0.52
	V113N	WT D-DT**	0.54
	V113N – 4CPPC	D-DT – 4CPPC	0.63
Gly mutation	L117G	WT D-DT**	0.46
C-terminal truncations	Δ109-117	WT D-DT**	0.54
	Δ109-117	Δ114-117	0.96
	Δ114-117	WT D-DT**	1.11

* PDB entry: 7MSE, ** PDB entry: 1DPT

Table 6.S5.

Crystallographic data collection and refinement statistics for V113N and V113N – 4-CPPC.

	V113N	**V113N – 4-CPPC**
Data collection		
Space group	P 63	P 63
Cell dimensions		
a, *b*, *c* (Å)	111.24, 111.24, 65.97	111.10, 111.10, 66.10
α, β, γ (°)	90.00, 90.00, 120.00	90.00, 90.00, 120.00
Resolution (Å)	96.34-1.31 (1.33-1.31)*	55.6-1.50 (1.53-1.50)*
R_{sym} or R_{merge}	0.073 (0.881)	0.233 (1.974)
$I / \sigma I$	12.4 (0.2)	6.5 (0.1)
Completeness (%)	99.9 (98.9)	99.1 (95.2)
Redundancy	17.1 (7.7)	17.0 (9.1)
Refinement		
Resolution (Å)	96.34-1.31(1.34-1.31)	55.61-1.50 (1.54-1.50)
No. reflections	105609	69924
R_{work} / R_{free}	0.19/0.22	0.23/0.26
No. atoms		
Protein	3240	3304
Ligand/ion	13	34
Water	554	539
B-factors		
Protein	23.30	26.49
Ligand/ion	31.30	46.14
Water	36.51	37.57
R.m.s. deviations		
Bond lengths (Å)	0.011	0.009
Bond angles (°)	1.511	1.450

* Values in parentheses are for highest-resolution shell.

Table 6.S6.

Crystallographic data collection and refinement statistics for WT D-DT (290K) and WT D-DT (310K).

	WT D-DT (290K)	**WT D-DT (310K)**
Data collection		
Space group	P 3	P 3
Cell dimensions		
a, *b*, *c* (Å)	84.35, 84.35, 41.03	84.29, 84.29, 49.06
α, β, γ (°)	90.00, 90.00, 120.00	90.00, 90.00, 120.00
Resolution (Å)	73.05-1.23 (1.25-1.23)*	73.00-1.29 (1.31-1.29)*
R_{sym} or R_{merge}	0.086 (0.879)	0.091 (1.423)
$I / \sigma I$	8.1 (0.4)	11.4 (0.5)
Completeness (%)	92.1 (44.0)	97.9 (78.8)
Redundancy	4.3 (1.8)	9.1 (4.5)
Refinement		
Resolution (Å)	73.05 -1.23 (1.26-1.23)	73.00 -1.29 (1.32-1.29)
No. reflections	83207	76386
R_{work} / R_{free}	0.14/0.18	0.16/0.17
No. atoms		
Protein	2669	2660
Ligand/ion	0	0
Water	293	234
B-factors		
Protein	14.94	16.67
Ligand/ion	-	-
Water	34.00	34.08
R.m.s. deviations		
Bond lengths (Å)	0.015	0.013
Bond angles (°)	1.851	1.682

* Values in parentheses are for highest-resolution shell.

Table 6.S7.

Crystallographic data collection and refinement statistics for L117G.

	L117G
Data collection	H 3
Space group	
Cell dimensions	
a, *b*, *c* (Å)	80.59, 80.59, 40.12
α, β, γ (°)	90.00, 90.00, 120.00
Resolution (Å)	40.29-1.35 (1.37-1.35)*
R_{sym} or R_{merge}	0.056 (0.252)
$I / \sigma I$	20.4 (3.0)
Completeness (%)	99.6 (94.8)
Redundancy	4.5 (2.8)
Refinement	
Resolution (Å)	40.29-1.35 (1.39-1.35)
No. reflections	20211
R_{work} / R_{free}	0.17/0.19
No. atoms	
Protein	856
Ligand/ion	13
Water	171
B-factors	
Protein	9.33
Ligand/ion	14.70
Water	28.86
R.m.s. deviations	
Bond lengths (Å)	0.013
Bond angles (°)	1.801

* Values in parentheses are for highest-resolution shell.

Table 6.S8.

Crystallographic data collection and refinement statistics for Δ114-117 and Δ109-117.

	Δ114-117	**Δ109-117**
Data collection		
Space group	P 1 21 1	P 61
Cell dimensions		
a, *b*, *c* (Å)	99.53, 59.53, 129.98	75.50, 75.50, 96.72
α, β, γ (°)	90.00, 107.93, 90.00	90.00, 90.00, 120.00
Resolution (Å)	89.84-2.44 (2.48-2.44)*	65.38-1.59 (1.62-1.59)*
R_{sym} or R_{merge}	0.267 (1.363)	0.247 (1.879)
$I / \sigma I$	4.5 (1.0)	4.9 (0.7)
Completeness (%)	99.8 (99.7)	98.5 (97.2)
Redundancy	3.4 (3.5)	11.6 (11.8)
Refinement		
Resolution (Å)	89.84-2.44 (2.50-2.44)	65.38-1.59 (1.63-1.59)
No. reflections	51582	39321
R_{work} / R_{free}	0.25/0.30	0.18/0.18
No. atoms		
Protein	9383	2374
Ligand/ion	0	27
Water	7	212
B-factors		
Protein	29.90	20.20
Ligand/ion	-	48.76
Water	15.12	36.49
R.m.s. deviations		
Bond lengths (Å)	0.008	0.013
Bond angles (°)	1.567	1.569

* Values in parentheses are for highest-resolution shell.

Table 6.S9.

DNA oligos used for mutagenesis[*].

Variant	Primer	Primer sequence (5′ to 3′)
L117A	Forward	GTCATGACTTTTGCCTAGTAGCTCGAG
	Reverse	CTCGAGCTACTAGGCAAAAGTCATGAC
F116A	Forward	ACGGTCATGACTGCCTTATAGTAGCTC
	Reverse	GAGCTACTATAAGGCAGTCATGACCGT
T115A	Forward	GGGACGGTCATGGCCTTTTTATAGTAG
	Reverse	CTACTATAAAAAGGCCATGACCGTCCC
M114A	Forward	ATAGGGACGGTCGCCACTTTTTTATAG
	Reverse	CTATAAAAAAGTGGCGACCGTCCCTAT
V113A	Forward	AAGATAGGGACGGCCATGACTTTTTTA
	Reverse	TAAAAAAGTCATGGCCGTCCCTATCTT
V113N	Forward	GGCAAGATAGGGACGAACATGACTTTTTTATAG
	Reverse	CTATAAAAAAGTCATGTTCGTCCCTATCTTGCC
V113L	Forward	GGCAAGATAGGGACGCTCATGACTTTTTTATAG
	Reverse	CTATAAAAAAGTCATGAGCGTCCCTATCTTGCC
V113I	Forward	GGCAAGATAGGGACGATCATGACTTTTTTATAG
	Reverse	CTATAAAAAAGTCATGATCGTCCCTATCTTGCC
V113T	Forward	GGCAAGATAGGGACGACCATGACTTTTTTATAG
	Reverse	CTATAAAAAAGTCATGGTCGTCCCTATCTTGCC
V113G	Forward	GGCAAGATAGGGACGGGCATGACTTTTTTATAG
	Reverse	CTATAAAAAAGTCATGCCCGTCCCTATCTTGCC

(Table 6.S9 Continued)

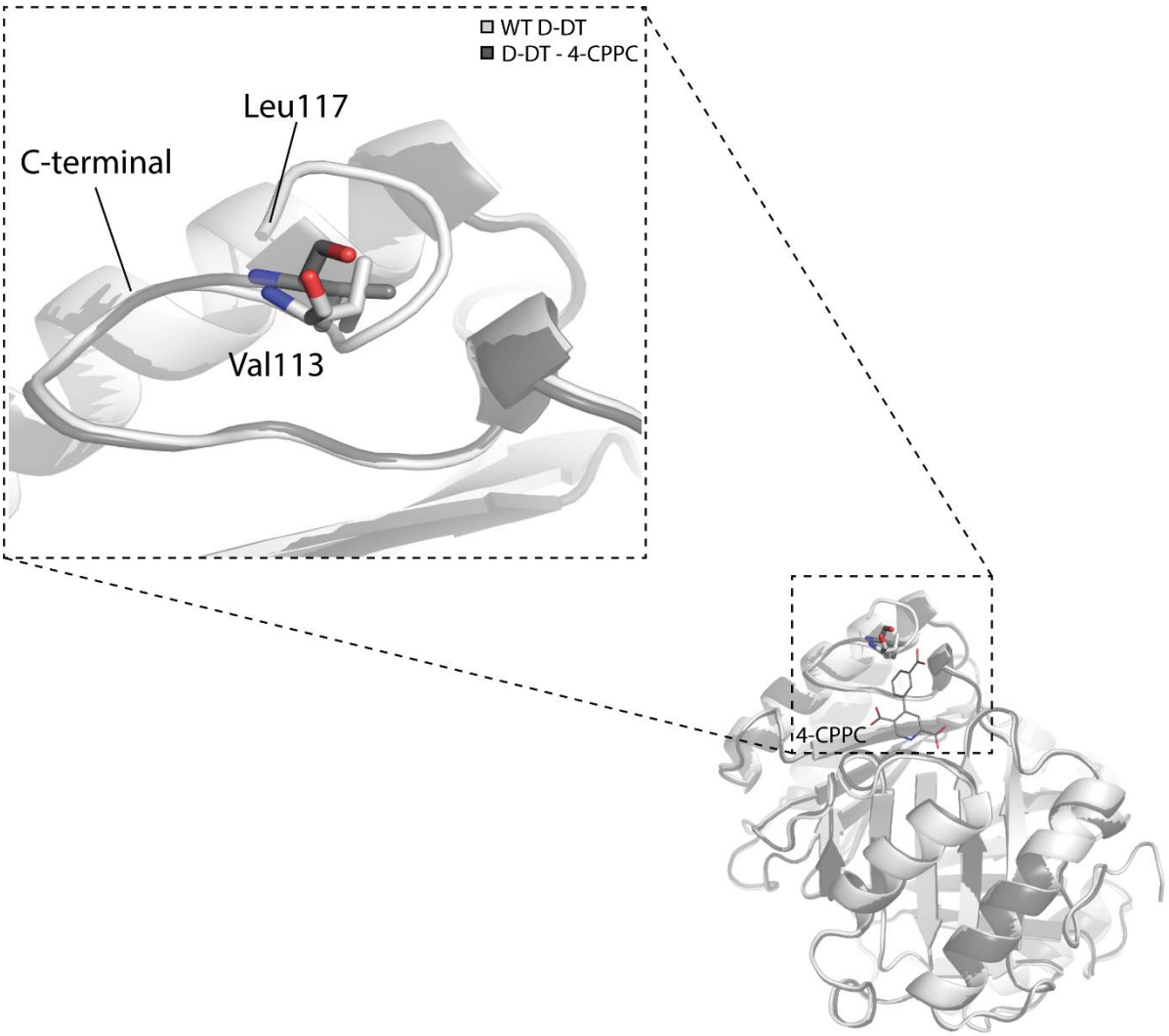
Δ114-117	Forward	AAGATAGGGACGGTCTAGTAGCTCGAGCAC
	Reverse	GTGCTCGAGCTACTAGACCGTCCCTATCTT
Δ109-117	Forward	TCCTGGCAGATTGGCTAGTAGCTCGAGCAC
	Reverse	GTGCTCGAGCTACTAGCCAATCTGCCAGGA
Δ104-117	Forward	TTTTTCCCCTTGGAGTAGTAGCTCGAGCAC
	Reverse	GTGCTCGAGCTACTACTCCAAGGGGAAAAA
L117G	Forward	GTCATGACTTTTGGGTAGTAGCTCGAG
	Reverse	CTCGAGCTACTACCCAAAAGTCATGAC
L117G/F116G	Forward	ACGGTCATGACTGGAGGGTAGTAGCTCGAG
	Reverse	CTCGAGCTACTACCCTCCAGTCATGACCGT
L117G/F116G/T115G	Forward	GGGACGGTCATGGGTGGAGGGTAGTAGCTCGAG
	Reverse	CTCGAGCTACTACCCTCCACCCATGACCGTCCC
L117G/F116G/T115G/M114G**	Forward	AAGATAGGGACGGTCGGAGGTGGAGGGTAGTAG
	Reverse	CTACTACCCTCCACCTCCGACCGTCCCTATCTT

* WT D-DT was used as a template for PCR with exception of **

** The confirmed plasmid of L117G/F116G/T115G was used as a template for PCR

Figure 6.S1.

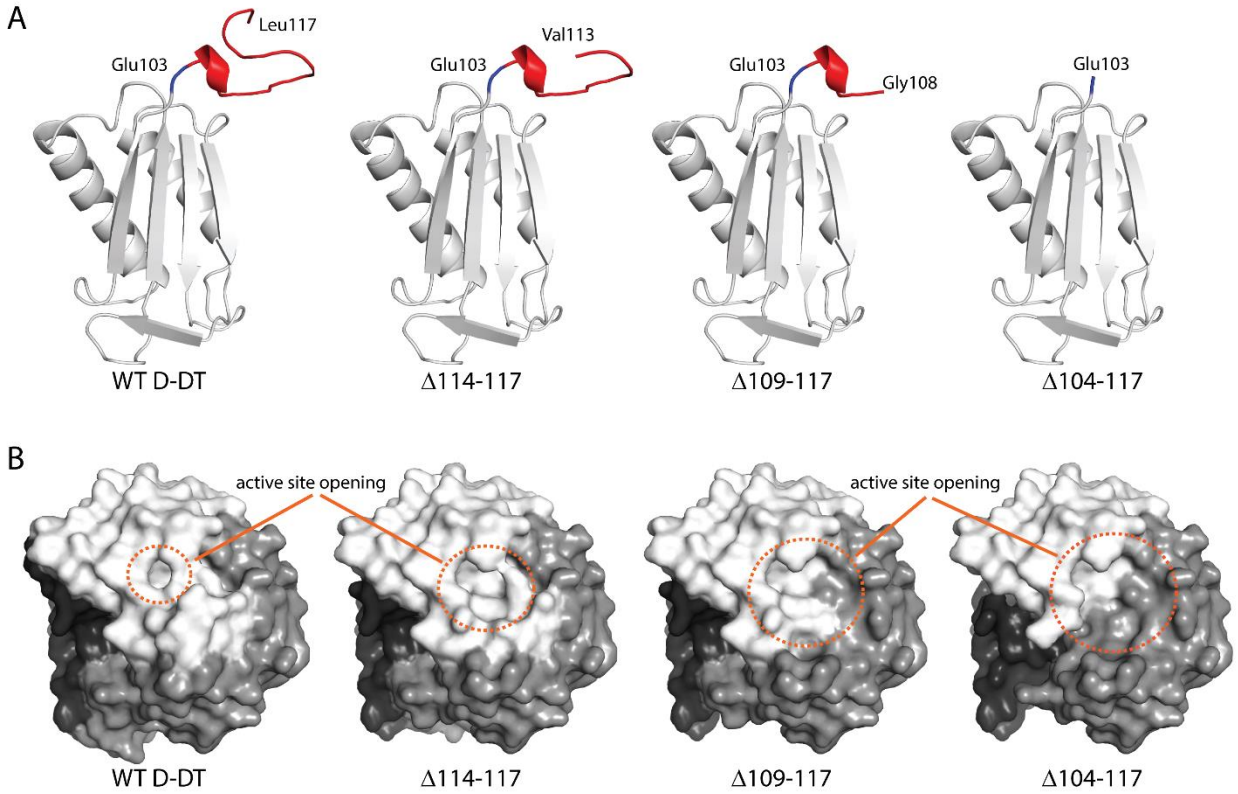
Conformational changes at the C-terminal region of D-DT due to 4-CPPC binding.



Note. Superposition of WT D-DT (light grey - PDB entry: 1DPT) onto D-DT – 4-CPPC (dark grey - PDB entry: 6C5F) demonstrates that 4-CPPC (dark grey lines) binding increases the mobility of the C-terminal residues. Val113, shown as sticks, was the last residue affected.

Figure 6.S2.

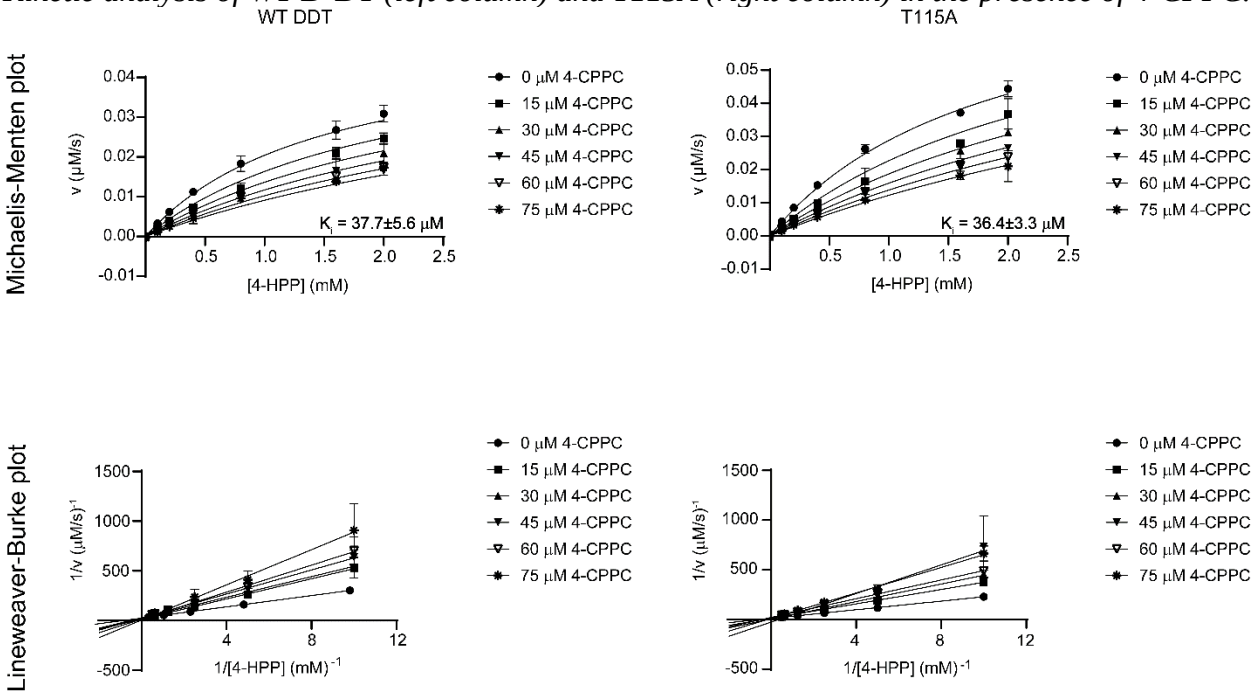
Modeling the effect of C-terminal truncations on the structure of D-DT.



Note. **(A)** Using monomeric D-DT, which was obtained from the crystal structure of WT D-DT (PDB entry: 1DPT), the desired truncation variants were modeled. for each truncation, the terminus amino acid is marked on the structure. In red, the entire C-terminal segment that was gradually removed is shown. Glu103, the terminus residue of Δ104-117, is shown in blue. **(B)** Projection of the impact of C-terminal truncations on the active site opening of D-DT. The opening is marked within the orange dashed circles.

Figure 6.S3.

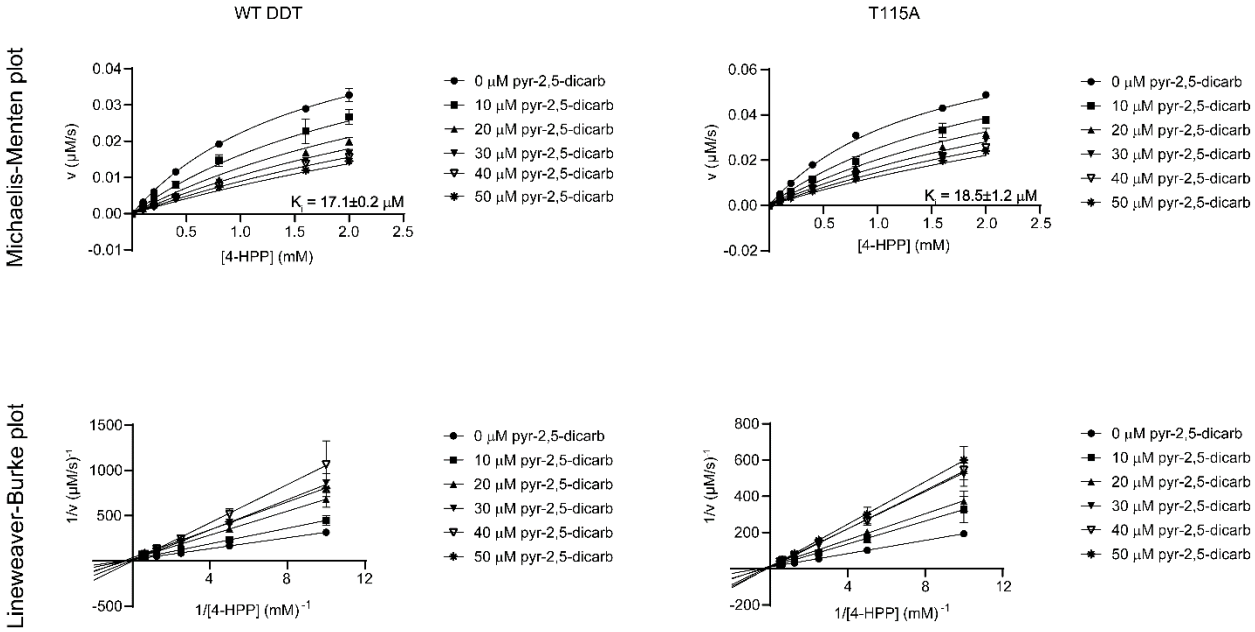
Kinetic analysis of WT D-DT (left column) and T115A (right column) in the presence of 4-CPPC.



Note. The Michaelis-Menten and Lineweaver-Burk plots are shown on the top and bottom panels, respectively. The tautomerase activity of D-DT was measured using 4-HPP as a substrate. The error values are shown as standard deviations. Each experiment was repeated in triplicate.

Figure 6.S4.

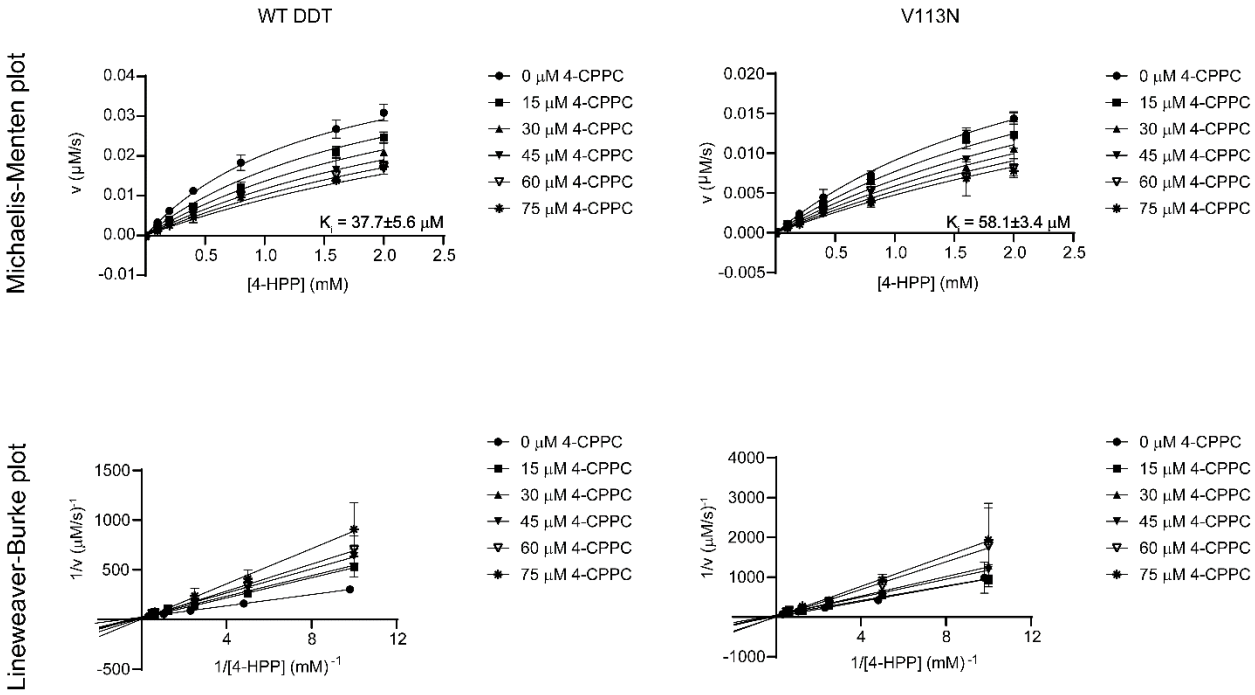
Kinetic analysis of WT D-DT (left column) and T115A (right column) in the presence of pyridine-2,5-dicarboxylate.



Note. The Michaelis-Menten and Lineweaver-Burk plots are shown on the top and bottom panels, respectively. The tautomerase activity of D-DT was measured using 4-HPP as a substrate. The error values are shown as standard deviations. Each experiment was repeated in triplicate.

Figure 6.S5.

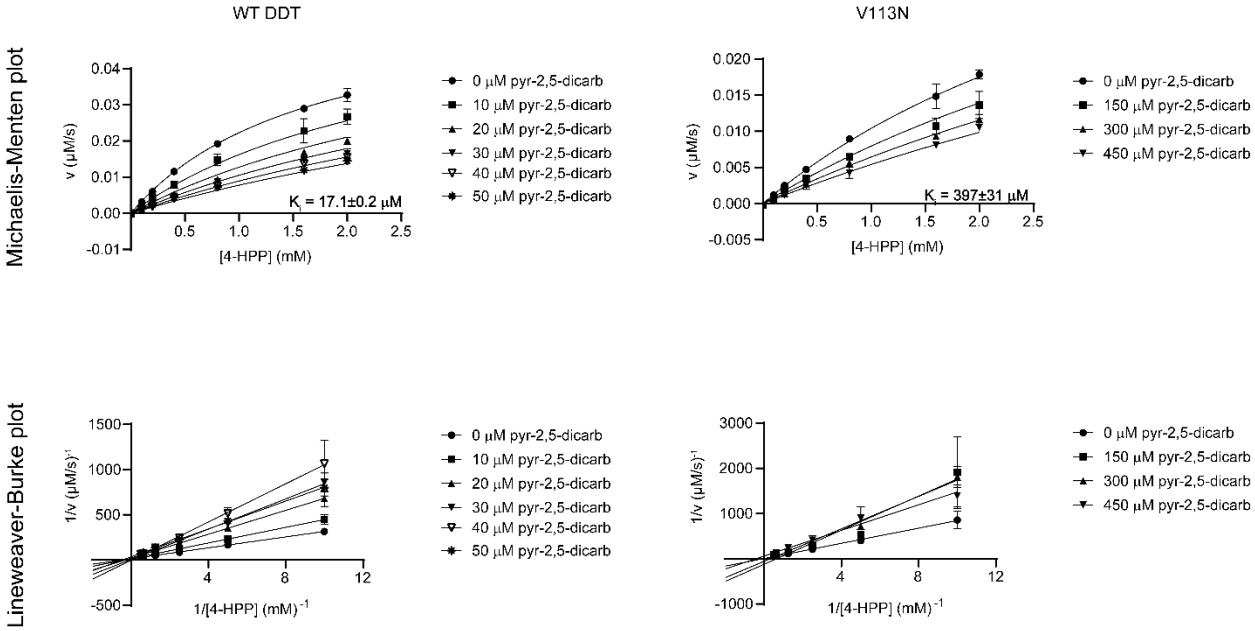
Kinetic analysis of WT D-DT (left column) and V113N (right column) in the presence of 4-CPPC.



Note. The Michaelis-Menten and Lineweaver-Burk plots are shown on the top and bottom panels, respectively. The tautomerase activity of D-DT was measured using 4-HPP as a substrate. The error values are shown as standard deviations. Each experiment was repeated in triplicate.

Figure 6.S6.

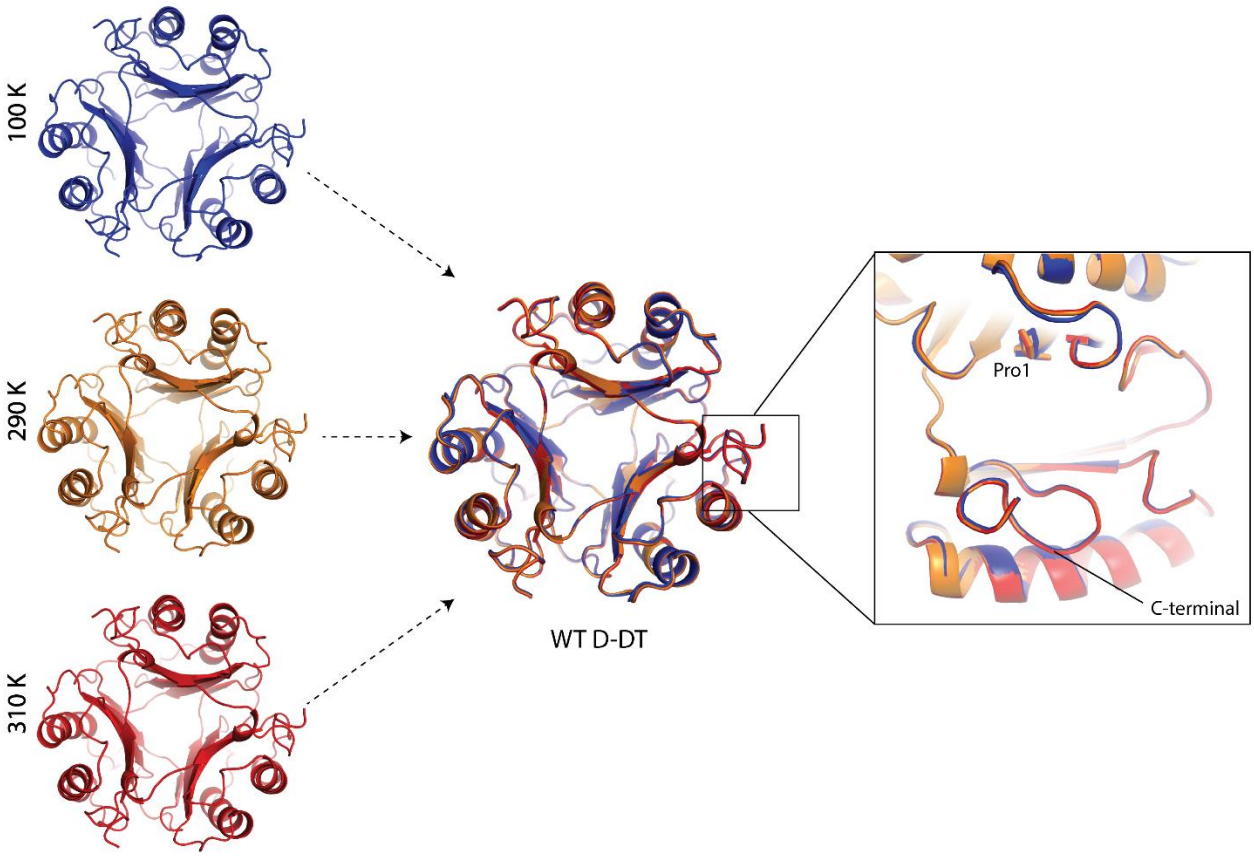
Kinetic analysis of WT D-DT (left column) and V113N (right column) in the presence of pyridine-2,5-dicarboxylate.



Note. The Michaelis-Menten and Lineweaver-Burk plots are shown on the top and bottom panels, respectively. The tautomerase activity of D-DT was measured using 4-HPP as a substrate. The error values are shown as standard deviations. Each experiment was repeated in triplicate.

Figure 6.S7.

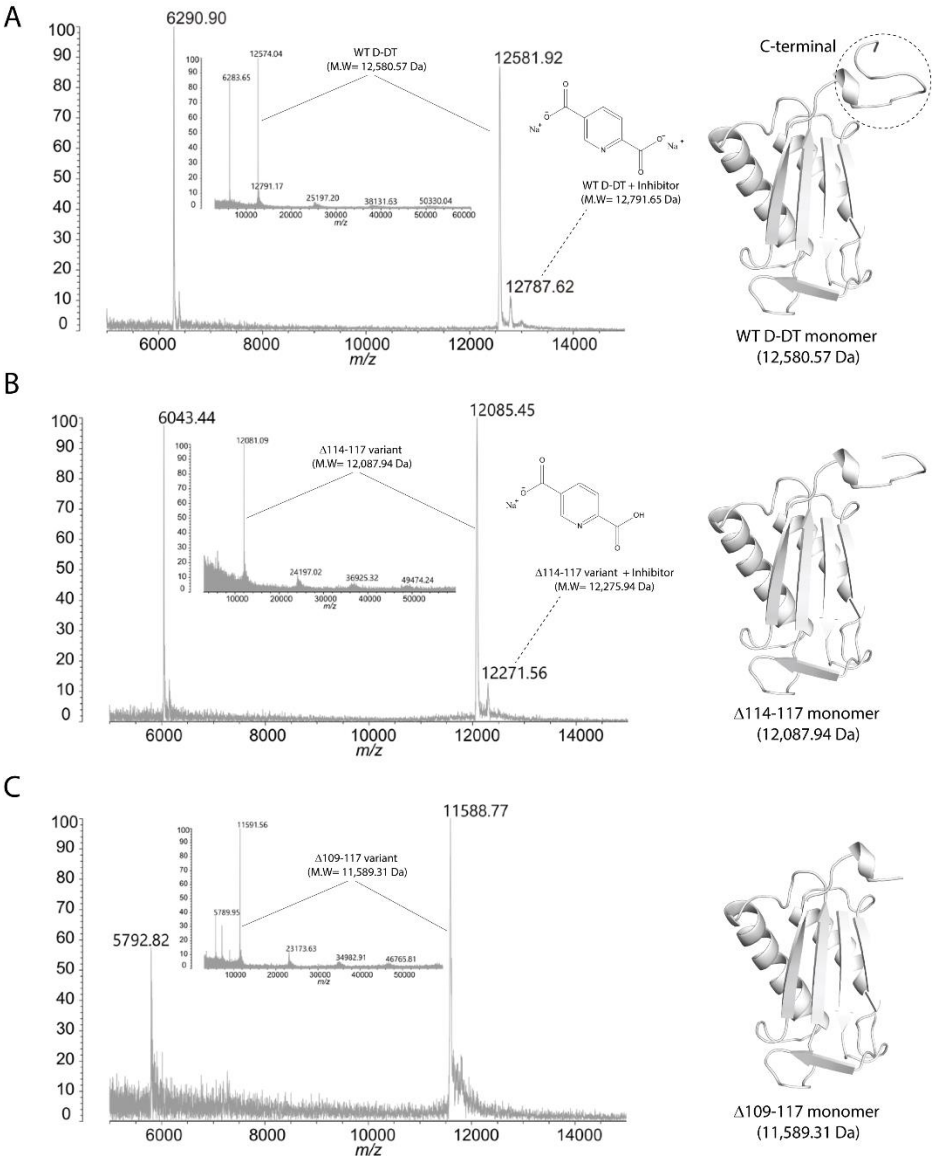
Superposition analysis of the crystal structure of WT D-DT at different temperatures.



Note. The crystal structures of WT D-DT at 290K (orange) and 310 K (red) were obtained in this study. The crystal structure of WT D-DT at 100 K was previously published (PDB entry: 7MSE). Pro1 is shown as sticks while the C-terminal region is pointed with a black solid line.

Figure 6.S8.

MALDI-TOF MS analysis of A) WT D-DT, B) Δ114-117, and C) Δ109-117.



Note. On the left, the MS spectra of D-DT - pyridine-2,5-dicarboxylate complex (Zoom in) and apo D-DT (Zoom out) are presented. The monomeric structure of D-DT variant under study along with its theoretical molecular weight (MW) are shown on the right. Detection of the D-DT - pyridine-2,5-dicarboxylate complex was noted only in the cases of WT D-DT and Δ114-117

(Figure 6.S8 Continued)

(dashed lines in panels A and B). The experimental molecular weights of WT D-DT, Δ114-117, and Δ109-117 are pointed with solid lines.

CONCLUSION

MIF has been the main focus of the scientific community and as a result, is thoroughly characterized, even being used to teach scientific methodology to high-school students[174]. By comparison, until our work, D-DT was missing in-depth structural investigations as well as the proper tool for investigation of its disease pathology. Overall, our findings enhance the understanding of the complex protein dynamics and structure of MIF and D-DT while also elucidating key differences between the two cytokines. These data will be used to guide the design of new specific and bioactive inhibitors of the proteins.

Through detailed analysis, we found that specific amino acid correlations in MIF proteins determine their function as CD74 agonists or antagonists. CD74 agonists require enhanced correlations among amino acids critical for the binding of CD74 either directly or indirectly, while antagonists show reduced correlations. This enhanced provides a foundation for developing potent CD74 modulators.

In a separate study, we examined how impurities in 4-HPP influence results from kinetic assays that are used for characterization of inhibitors and protein variants with MIF. By sourcing 4-HPP from different suppliers and analyzing enzymatic activity, we linked inconsistent results to impurities under 5%, confirmed by NMR and MS. This underscores the importance of substrate purity for the in enzymatic assays.

Additionally, we explored the unique C-terminal flexibility of D-DT, brought to light by the conformational change induced upon the binding of 4-CPPC[49]. Although the current role of this flexibility is unknown, it may affects its interaction with CD74, resulting in a different binding affinity than MIF [9]. The mobility of the C-terminal and its potential roles in biological

functions such as signaling [128] and catalysis [175] warrant further investigation through targeted mutations.

We also identified and characterized **1**, as a highly selective inhibitor of D-DT that does not cause significant changes to the conformational of the C-terminal, with *in vivo* efficacy. Its superior 79-fold selectivity for D-DT over MIF sets a new benchmark for D-DT inhibitors. To explain this high selectivity, crystallographic analysis of MIF and D-DT were performed and revealed key differences between the catalytic sites of the two proteins, providing key insight for the generation of new specific inhibitors. Further *in vivo* and *in vitro* studies utilizing **1** and ISO-1, a well-studied specific MIF inhibitor [21], will help elucidate the distinct biological roles of MIF and D-DT.

Finally, interrogation of C-terminal region of D-DT utilizing 17 unique variants and WT revealed that the C-terminal is key for ligand recognition. One of the key residues for this function was found to be Thr115. Through conformational changes, this residue influences the conformations changes of active site residues which are critical for ligand binding. This was demonstrated by the enhanced kinetics of a T115A variant, whose truncated sidechain allows for easier conformational changes of catalytic site residues, most notably Arg36, which is responsible for the opening and closing of the active site.

References

(1) Whitman, C. P. *Arch Biochem Biophys* **2002**, *402* (1), 1-13. DOI: 10.1016/S0003-9861(02)00052-8.
(2) Jenkins, J. R.; Cooper, R. A. *J Bacteriol* **1988**, *170* (11), 5317-5324. DOI: 10.1128/jb.170.11.5317-5324.1988.
(3) Poelarends, G. J.; Whitman, C. P. *Bioorg Chem* **2004**, *32* (5), 376-392. DOI: 10.1016/j.bioorg.2004.05.006.
(4) Poelarends, G. J.; Johnson, W. H., Jr.; Murzin, A. G.; Whitman, C. P. *J Biol Chem* **2003**, *278* (49), 48674-48683. DOI: 10.1074/jbc.M306706200.
(5) Poelarends, G. J.; Serrano, H.; Person, M. D.; Johnson, W. H., Jr.; Murzin, A. G.; Whitman, C. P. *Biochemistry* **2004**, *43* (3), 759-772. DOI: 10.1021/bi0355948.
(6) Rosengren, E.; Aman, P.; Thelin, S.; Hansson, C.; Ahlfors, S.; Bjork, P.; Jacobsson, L.; Rorsman, H. *FEBS Lett* **1997**, *417* (1), 85-88. DOI: 10.1016/s0014-5793(97)01261-1.
(7) Lubetsky, J. B.; Swope, M.; Dealwis, C.; Blake, P.; Lolis, E. *Biochemistry* **1999**, *38* (22), 7346-7354. DOI: 10.1021/bi990306m.
(8) Whitman, C. P.; Aird, B. A.; Gillespie, W. R.; Stolowich, N. J. *Journal of the American Chemical Society* **1991**, *113* (8), 3154-3162. DOI: 10.1021/ja00008a052. Hartmans, S.; Jansen, M. W.; van der Werf, M. J.; de Bont, J. A. *J Gen Microbiol* **1991**, *137* (8), 2025-2032. DOI: 10.1099/00221287-137-8-2025.
(9) Merk, M.; Zierow, S.; Leng, L.; Das, R.; Du, X.; Schulte, W.; Fan, J.; Lue, H.; Chen, Y.; Xiong, H.; et al. *Proc Natl Acad Sci U S A* **2011**, *108* (34), E577-585. DOI: 10.1073/pnas.1102941108.
(10) Moran, G. R. *Arch Biochem Biophys* **2005**, *433* (1), 117-128. DOI: 10.1016/j.abb.2004.08.015.
(11) Bloom, B. R.; Bennett, B. *Science* **1966**, *153* (3731), 80-82. DOI: 10.1126/science.153.3731.80.
(12) Bernhagen, J.; Krohn, R.; Lue, H.; Gregory, J. L.; Zernecke, A.; Koenen, R. R.; Dewor, M.; Georgiev, I.; Schober, A.; Leng, L.; et al. *Nat Med* **2007**, *13* (5), 587-596. DOI: 10.1038/nm1567.
(13) Wang, Y.; An, R.; Umanah, G. K.; Park, H.; Nambiar, K.; Eacker, S. M.; Kim, B.; Bao, L.; Harraz, M. M.; Chang, C.; et al. *Science* **2016**, *354* (6308). DOI: 10.1126/science.aad6872.
(14) Jankauskas, S. S.; Wong, D. W. L.; Bucala, R.; Djudjaj, S.; Boor, P. *Cell Signal* **2019**, *57*, 76-88. DOI: 10.1016/j.cellsig.2019.01.006.
(15) Xiao, Z.; Chen, D.; Mulder, F.; Song, S.; van der Wouden, P. E.; Cool, R. H.; Melgert, B. N.; Poelarends, G. J.; Dekker, F. J. *Chemistry* **2022**, *28* (1), e202103030. DOI: 10.1002/chem.202103030.
(16) Bernhagen, J.; Calandra, T.; Mitchell, R. A.; Martin, S. B.; Tracey, K. J.; Voelter, W.; Manogue, K. R.; Cerami, A.; Bucala, R. *Nature* **1993**, *365* (6448), 756-759. DOI: 10.1038/365756a0.
(17) Sun, H. W.; Bernhagen, J.; Bucala, R.; Lolis, E. *Proc Natl Acad Sci U S A* **1996**, *93* (11), 5191-5196. DOI: 10.1073/pnas.93.11.5191.
(18) Stamps, S. L.; Fitzgerald, M. C.; Whitman, C. P. *Biochemistry* **1998**, *37* (28), 10195-10202. DOI: 10.1021/bi9806955.

(19) Merk, M.; Mitchell, R. A.; Endres, S.; Bucala, R. *Cytokine* **2012**, *59* (1), 10-17. DOI: 10.1016/j.cyto.2012.03.014.
(20) Meza-Romero, R.; Benedek, G.; Jordan, K.; Leng, L.; Pantouris, G.; Lolis, E.; Bucala, R.; Vandenbark, A. A. *Cytokine* **2016**, *88*, 62-70. DOI: 10.1016/j.cyto.2016.08.024.
(21) Cho, Y.; Crichlow, G. V.; Vermeire, J. J.; Leng, L.; Du, X.; Hodsdon, M. E.; Bucala, R.; Cappello, M.; Gross, M.; Gaeta, F.; et al. *Proc Natl Acad Sci U S A* **2010**, *107* (25), 11313-11318. DOI: 10.1073/pnas.1002716107 from NLM.
(22) Vera, P. L.; Iczkowski, K. A.; Wang, X.; Meyer-Siegler, K. L. *PLoS One* **2008**, *3* (12), e3898. DOI: 10.1371/journal.pone.0003898. Tarnowski, M.; Grymula, K.; Liu, R.; Tarnowska, J.; Drukala, J.; Ratajczak, J.; Mitchell, R. A.; Ratajczak, M. Z.; Kucia, M. *Mol Cancer Res* **2010**, *8* (10), 1328-1343. DOI: 10.1158/1541-7786.MCR-10-0288.
(23) Weber, C.; Kraemer, S.; Drechsler, M.; Lue, H.; Koenen, R. R.; Kapurniotu, A.; Zernecke, A.; Bernhagen, J. *Proc Natl Acad Sci U S A* **2008**, *105* (42), 16278-16283. DOI: 10.1073/pnas.0804017105.
(24) Kontos, C.; El Bounkari, O.; Krammer, C.; Sinitski, D.; Hille, K.; Zan, C.; Yan, G.; Wang, S.; Gao, Y.; Brandhofer, M.; et al. *Nat Commun* **2020**, *11* (1), 5981. DOI: 10.1038/s41467-020-19764-z.
(25) Su, H.; Na, N.; Zhang, X.; Zhao, Y. *Inflamm Res* **2017**, *66* (3), 209-216. DOI: 10.1007/s00011-016-0995-1.
(26) Brock, S. E.; Rendon, B. E.; Xin, D.; Yaddanapudi, K.; Mitchell, R. A. *PLoS One* **2014**, *9* (6), e99795. DOI: 10.1371/journal.pone.0099795.
(27) Leng, L.; Metz, C. N.; Fang, Y.; Xu, J.; Donnelly, S.; Baugh, J.; Delohery, T.; Chen, Y.; Mitchell, R. A.; Bucala, R. *J Exp Med* **2003**, *197* (11), 1467-1476. DOI: 10.1084/jem.20030286.
(28) Penticuff, J. C.; Woolbright, B. L.; Sielecki, T. M.; Weir, S. J.; Taylor, J. A., 3rd. *Nat Rev Urol* **2019**, *16* (5), 318-328. DOI: 10.1038/s41585-019-0171-9. Bozzi, F.; Mogavero, A.; Varinelli, L.; Belfiore, A.; Manenti, G.; Caccia, C.; Volpi, C. C.; Beznoussenko, G. V.; Milione, M.; Leoni, V.; et al. *J Exp Clin Cancer Res* **2017**, *36* (1), 16. DOI: 10.1186/s13046-016-0475-z.
(29) Takahashi, K.; Koga, K.; Linge, H. M.; Zhang, Y.; Lin, X.; Metz, C. N.; Al-Abed, Y.; Ojamaa, K.; Miller, E. J. *Respir Res* **2009**, *10*, 33. DOI: 10.1186/1465-9921-10-33.
(30) Lindner, R. *Cells* **2017**, *6* (1). DOI: 10.3390/cells6010006.
(31) Tanese, K.; Hashimoto, Y.; Berkova, Z.; Wang, Y.; Samaniego, F.; Lee, J. E.; Ekmekcioglu, S.; Grimm, E. A. *J Invest Dermatol* **2015**, *135* (11), 2775-2784. DOI: 10.1038/jid.2015.204. Kursunel, M. A.; Esendagli, G. *Cytokine Growth Factor Rev* **2016**, *31*, 73-81. DOI: 10.1016/j.cytogfr.2016.07.005. Guda, M. R.; Rashid, M. A.; Asuthkar, S.; Jalasutram, A.; Caniglia, J. L.; Tsung, A. J.; Velpula, K. K. *Am J Cancer Res* **2019**, *9* (12), 2760-2773.
(32) Bijlmakers, M. J.; Benaroch, P.; Ploegh, H. L. *J Exp Med* **1994**, *180* (2), 623-629. DOI: 10.1084/jem.180.2.623.
(33) Pieters, J.; Horstmann, H.; Bakke, O.; Griffiths, G.; Lipp, J. *J Cell Biol* **1991**, *115* (5), 1213-1223. DOI: 10.1083/jcb.115.5.1213.
(34) Matza, D.; Kerem, A.; Medvedovsky, H.; Lantner, F.; Shachar, I. *Immunity* **2002**, *17* (5), 549-560. DOI: 10.1016/s1074-7613(02)00455-7.
(35) Fukuda, Y.; Bustos, M. A.; Cho, S. N.; Roszik, J.; Ryu, S.; Lopez, V. M.; Burks, J. K.; Lee, J. E.; Grimm, E. A.; Hoon, D. S. B.; et al. *Cell Death Dis* **2022**, *13* (2), 117. DOI: 10.1038/s41419-022-04552-y.
(36) Shi, X.; Leng, L.; Wang, T.; Wang, W.; Du, X.; Li, J.; McDonald, C.; Chen, Z.; Murphy, J. W.; Lolis, E.; et al. *Immunity* **2006**, *25* (4), 595-606. DOI: 10.1016/j.immuni.2006.08.020.

(37) Lue, H.; Thiele, M.; Franz, J.; Dahl, E.; Speckgens, S.; Leng, L.; Fingerle-Rowson, G.; Bucala, R.; Luscher, B.; Bernhagen, J. *Oncogene* **2007**, *26* (35), 5046-5059. DOI: 10.1038/sj.onc.1210318.
(38) Mora Barthelmess, R.; Stijlemans, B.; Van Ginderachter, J. A. *Cancers (Basel)* **2023**, *15* (2). DOI: 10.3390/cancers15020395.
(39) Morand, E. F.; Leech, M.; Bernhagen, J. *Nat Rev Drug Discov* **2006**, *5* (5), 399-410. DOI: 10.1038/nrd2029.
(40) Ong, G. L.; Goldenberg, D. M.; Hansen, H. J.; Mattes, M. J. *Immunology* **1999**, *98* (2), 296-302. DOI: 10.1046/j.1365-2567.1999.00868.x.
(41) Jasanoff, A.; Wagner, G.; Wiley, D. C. *EMBO J* **1998**, *17* (23), 6812-6818. DOI: 10.1093/emboj/17.23.6812.
(42) Coleman, A. M.; Rendon, B. E.; Zhao, M.; Qian, M. W.; Bucala, R.; Xin, D.; Mitchell, R. A. *J Immunol* **2008**, *181* (4), 2330-2337. DOI: 10.4049/jimmunol.181.4.2330. Guo, Y.; Hou, J.; Luo, Y.; Wang, D. *Cancer Cell Int* **2013**, *13* (1), 28. DOI: 10.1186/1475-2867-13-28.
(43) Zhang, H.; Duan, J.; Wu, O. *Saudi J Biol Sci* **2020**, *27* (6), 1527-1532. DOI: 10.1016/j.sjbs.2020.04.027.
(44) Xiao, Z.; Osipyan, A.; Song, S.; Chen, D.; Schut, R. A.; van Merkerk, R.; van der Wouden, P. E.; Cool, R. H.; Quax, W. J.; Melgert, B. N.; et al. *J Med Chem* **2022**, *65* (3), 2059-2077. DOI: 10.1021/acs.jmedchem.1c01598.
(45) Otvos, B.; Silver, D. J.; Mulkearns-Hubert, E. E.; Alvarado, A. G.; Turaga, S. M.; Sorensen, M. D.; Rayman, P.; Flavahan, W. A.; Hale, J. S.; Stoltz, K.; et al. *Stem Cells* **2016**, *34* (8), 2026-2039. DOI: 10.1002/stem.2393. Alban, T. J.; Bayik, D.; Otvos, B.; Rabljenovic, A.; Leng, L.; Jia-Shiun, L.; Roversi, G.; Lauko, A.; Momin, A. A.; Mohammadi, A. M.; et al. *Front Immunol* **2020**, *11*, 1191. DOI: 10.3389/fimmu.2020.01191. Ghoochani, A.; Schwarz, M. A.; Yakubov, E.; Engelhorn, T.; Doerfler, A.; Buchfelder, M.; Bucala, R.; Savaskan, N. E.; Eyupoglu, I. Y. *Oncogene* **2016**, *35* (48), 6246-6261. DOI: 10.1038/onc.2016.160. Lee, S. H.; Kwon, H. J.; Park, S.; Kim, C. I.; Ryu, H.; Kim, S. S.; Park, J. B.; Kwon, J. T. *PLoS One* **2021**, *16* (9), e0257375. DOI: 10.1371/journal.pone.0257375.
(46) Crichlow, G. V.; Lubetsky, J. B.; Leng, L.; Bucala, R.; Lolis, E. J. *Biochemistry* **2009**, *48* (1), 132-139. DOI: 10.1021/bi8014423.
(47) McLean, L. R.; Zhang, Y.; Li, H.; Choi, Y. M.; Han, Z.; Vaz, R. J.; Li, Y. *Bioorg Med Chem Lett* **2010**, *20* (6), 1821-1824. DOI: 10.1016/j.bmcl.2010.02.009.
(48) Pantouris, G.; Syed, M. A.; Fan, C.; Rajasekaran, D.; Cho, T. Y.; Rosenberg, E. M., Jr.; Bucala, R.; Bhandari, V.; Lolis, E. J. *Chem Biol* **2015**, *22* (9), 1197-1205. DOI: 10.1016/j.chembiol.2015.08.006.
(49) Pantouris, G.; Bucala, R.; Lolis, E. J. *Biochemistry* **2018**, *57* (26), 3599-3605. DOI: 10.1021/acs.biochem.8b00344.
(50) Lubetsky, J. B.; Dios, A.; Han, J.; Aljabari, B.; Ruzsicska, B.; Mitchell, R.; Lolis, E.; Al-Abed, Y. *J Biol Chem* **2002**, *277* (28), 24976-24982. DOI: 10.1074/jbc.M203220200.
(51) Pantouris, G.; Ho, J.; Shah, D.; Syed, M. A.; Leng, L.; Bhandari, V.; Bucala, R.; Batista, V. S.; Loria, J. P.; Lolis, E. J. *Angew Chem Int Ed Engl* **2018**, *57* (24), 7116-7119. DOI: 10.1002/anie.201803191.
(52) Chen, E.; Reiss, K.; Shah, D.; Manjula, R.; Allen, B.; Murphy, E. L.; Murphy, J. W.; Batista, V. S.; Bhandari, V.; Lolis, E. J.; et al. *J Biol Chem* **2021**, *297* (3), 101061. DOI: 10.1016/j.jbc.2021.101061.

(53) Pantouris, G.; Khurana, L.; Ma, A.; Skeens, E.; Reiss, K.; Batista, V. S.; Lisi, G. P.; Lolis, E. J. *Cell Chem Biol* **2020**, *27* (6), 740-750 e745. DOI: 10.1016/j.chembiol.2020.05.001.
(54) Swope, M.; Sun, H. W.; Blake, P. R.; Lolis, E. *EMBO J* **1998**, *17* (13), 3534-3541. DOI: 10.1093/emboj/17.13.3534.
(55) Fingerle-Rowson, G.; Kaleswarapu, D. R.; Schlander, C.; Kabgani, N.; Brocks, T.; Reinart, N.; Busch, R.; Schutz, A.; Lue, H.; Du, X.; et al. *Mol Cell Biol* **2009**, *29* (7), 1922-1932. DOI: 10.1128/MCB.01907-08.
(56) Pantouris, G.; Rajasekaran, D.; Garcia, A. B.; Ruiz, V. G.; Leng, L.; Jorgensen, W. L.; Bucala, R.; Lolis, E. J. *J Med Chem* **2014**, *57* (20), 8652-8656. DOI: 10.1021/jm501168q.
(57) Rajasekaran, D.; Zierow, S.; Syed, M.; Bucala, R.; Bhandari, V.; Lolis, E. J. *FASEB J* **2014**, *28* (11), 4961-4971. DOI: 10.1096/fj.14-256636.
(58) Singh, A. K.; Pantouris, G.; Borosch, S.; Rojanasthien, S.; Cho, T. Y. *J Cell Mol Med* **2017**, *21* (1), 142-153. DOI: 10.1111/jcmm.12949.
(59) Tilstam, P. V.; Pantouris, G.; Corman, M.; andreoli, M.; Mahboubi, K.; Davis, G.; Du, X.; Leng, L.; Lolis, E.; Bucala, R. *J Biol Chem* **2019**, *294* (49), 18522-18531. DOI: 10.1074/jbc.RA119.009860.
(60) Mischke, R.; Gessner, A.; Kapurniotu, A.; Juttner, S.; Kleemann, R.; Brunner, H.; Bernhagen, J. *FEBS Lett* **1997**, *414* (2), 226-232. DOI: 10.1016/s0014-5793(97)01039-9. El-Turk, F.; Cascella, M.; Ouertatani-Sakouhi, H.; Narayanan, R. L.; Leng, L.; Bucala, R.; Zweckstetter, M.; Rothlisberger, U.; Lashuel, H. A. *Biochemistry* **2008**, *47* (40), 10740-10756. DOI: 10.1021/bi800603x.
(61) Subramanya, H. S.; Roper, D. I.; Dauter, Z.; Dodson, E. J.; Davies, G. J.; Wilson, K. S.; Wigley, D. B. *Biochemistry* **1996**, *35* (3), 792-802. DOI: 10.1021/bi951732k.
(62) Hare, A. A.; Leng, L.; Gandavadi, S.; Du, X.; Cournia, Z.; Bucala, R.; Jorgensen, W. L. *Bioorg Med Chem Lett* **2010**, *20* (19), 5811-5814. DOI: 10.1016/j.bmcl.2010.07.129.
(63) Cournia, Z.; Leng, L.; Gandavadi, S.; Du, X.; Bucala, R.; Jorgensen, W. L. *J Med Chem* **2009**, *52* (2), 416-424. DOI: 10.1021/jm801100v.
(64) Jorgensen, W. L.; Gandavadi, S.; Du, X.; Hare, A. A.; Trofimov, A.; Leng, L.; Bucala, R. *Bioorg Med Chem Lett* **2010**, *20* (23), 7033-7036. DOI: 10.1016/j.bmcl.2010.09.118.
(65) Austin, R. H.; Beeson, K. W.; Eisenstein, L.; Frauenfelder, H.; Gunsalus, I. C. *Biochemistry* **1975**, *14* (24), 5355-5373. DOI: 10.1021/bi00695a021. Dajnowicz, S.; Cheng, Y.; Daemen, L. L.; Weiss, K. L.; Gerlits, O.; Mueser, T. C.; Kovalevsky, A. *ACS Omega* **2020**, *5* (30), 18787-18797. DOI: 10.1021/acsomega.0c01900.
(66) Khurana, L.; ElGindi, M.; Tilstam, P. V.; Pantouris, G. *Methods Enzymol* **2019**, *629*, 307-360. DOI: 10.1016/bs.mie.2019.05.053.
(67) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J Comput Chem* **2004**, *25* (13), 1605-1612. DOI: 10.1002/jcc.20084.
(68) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmuller, H.; MacKerell, A. D., Jr. *Nat Methods* **2017**, *14* (1), 71-73. DOI: 10.1038/nmeth.4067.
(69) Humphrey, W.; Dalke, A.; Schulten, K. *J Mol Graph* **1996**, *14* (1), 33-38, 27-38. DOI: 10.1016/0263-7855(96)00018-5.
(70) Lange, O. F.; Grubmuller, H. *Proteins* **2006**, *62* (4), 1053-1061. DOI: 10.1002/prot.20784.
(71) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. *Journal of Computational Chemistry* **2005**, *26* (16), 1701-1718. DOI: https://doi.org/10.1002/jcc.20291.

(72) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G.; McCoy, A.; et al. *Acta Crystallogr D Biol Crystallogr* **2011**, *67* (Pt 4), 235-242. DOI: 10.1107/S0907444910045749.
(73) Delano, W. L. The PyMOL Molecular Graphics System. 2002.
(74) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J Biomol NMR* **1995**, *6* (3), 277-293. DOI: 10.1007/BF00197809.
(75) Lee, W.; Tonelli, M.; Markley, J. L. *Bioinformatics* **2015**, *31* (8), 1325-1327. DOI: 10.1093/bioinformatics/btu830.
(76) Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y.; et al. *Science* **2010**, *330* (6002), 341-346. DOI: 10.1126/science.1187409.
(77) Foda, Z. H.; Shan, Y.; Kim, E. T.; Shaw, D. E.; Seeliger, M. A. *Nat Commun* **2015**, *6*, 5939. DOI: 10.1038/ncomms6939. Wang, Y.; V, S. M.; Kim, J.; Li, G.; Ahuja, L. G.; Aoto, P.; Taylor, S. S.; Veglia, G. *Nat Commun* **2019**, *10* (1), 799. DOI: 10.1038/s41467-019-08655-7. Dalton, J. A. R.; Pin, J. P.; Giraldo, J. *Sci Rep* **2017**, *7* (1), 4944. DOI: 10.1038/s41598-017-05095-5.
(78) Yapel, A.; Han, M.; Lumry, R.; Rosenberg, A.; Shiao, D. F. *Journal of the American Chemical Society* **1966**, *88* (11), 2573-2584. DOI: 10.1021/ja00963a036. Hermann, J. C.; Chen, Y.; Wartchow, C.; Menke, J.; Gao, L.; Gleason, S. K.; Haynes, N. E.; Scott, N.; Petersen, A.; Gabriel, S.; et al. *ACS Med. Chem. Lett.* **2013**, *4*, 197.
(79) Kröplin, T.; Fischer, C.; Iven, H. *Eur. J. Clin Pharmacol* **1999**, *55*, 285.
(80) Calandra, T.; Spiegel, L. A.; Metz, C. N.; Bucala, R. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 11383.
(81) Lacy, M.; Kontos, C.; Brandhofer, M.; Hille, K.; Groning, S.; Sinitski, D.; Bourilhon, P.; Rosenberg, E.; Krammer, C.; Thavayogarajah, T.; et al. *Sci Rep* **2018**, *8* (1), 5171. DOI: 10.1038/s41598-018-23554-5.
(82) Tilstam, P. V.; Schulte, W.; Holowka, T.; Kim, B. S.; Nouws, J.; Sauler, M.; Piecychna, M.; Pantouris, G.; Lolis, E.; Leng, L.; et al. *J. Clin. Invest.* **2021**, *131*, na.
(83) Stoppe, C.; Rex, S.; Goetzenich, A.; Kraemer, S.; Emontzpohl, C.; Soppert, J.; Averdunk, L.; Sun, Y.; Rossaint, R.; Lue, H.; et al. *Antioxid Redox Signal* **2015**, *23* (11), 865-879. DOI: 10.1089/ars.2014.6243.
(84) Costa-Silva, B.; Aiello, N. M.; Ocean, A. J.; Singh, S.; Zhang, H.; Thakur, B. K.; Becker, A.; Hoshino, A.; Mark, M. T.; Molina, H.; et al. *Nat. Cell Biol.* **2015**, *17*, 816. Xiong, C.; Huang, B.; Cun, Y.; Aghdasi, B. G.; Zhou, Y. *Clin Orthop Relat Res.* **2014**, *472*, 1943. Tilstam, P. V.; Qi, D.; Leng, L.; Young, L.; Bucala, R. *Expert Opin Ther Targets* **2017**, *21*, 671.
(85) Rosengren, E.; Bucala, R.; Aman, P.; Jacobsson, L.; Odh, G.; Metz, C. N.; Rorsman, H. *Mol. Med.* **1996**, *2*, 143.
(86) El Turk, F.; Fauvet, B.; Ouertatani-Sakouhi, H.; Lugari, A.; Betzi, S.; Roche, P.; Morelli, X.; Lashuel, H. A. *Bioorg. Med. Chem.* **2010**, *18*, 5425.
(87) Zapatero, M. C.; Pérez, P.; Vázquez, M. J.; Colmenarejo, G.; de Los Frailes, M.; Ramón, F. *J. Biomol Screen* **2016**, *21*, 446.
(88) Skeens, E.; Pantouris, G.; Shah, D.; Manjula, R.; Ombrello, M. J.; Maluf, N. K.; Bhandari, V.; Lisi, G. P.; Lolis, E. J. *Front Mol. Biosci* **2022**, *9*, 783669.
(89) Shan, C.; Lu, Z.; Li, Z.; Sheng, H.; Fan, J.; Qi, Q.; Liu, S.; Zhang, S. *Cell Death Dis* **2019**, *10*, 525.
(90) Sugimoto, H.; Taniguchi, M.; Nakagawa, A.; Tanaka, I.; Suzuki, M.; Nishihira, J. *Biochemistry* **1999**, *38*, 3268.

(91) Parkins, A.; Skeens, E.; McCallum, C. M.; Lisi, G. P.; Pantouris, G. *Biophys J* **2021**, *120* (18), 3893-3900. DOI: 10.1016/j.bpj.2021.08.025.
(92) Pervushin, K.; Riek, R.; Wider, G.; Wüthrich, K. *J. Am. Chem. Soc.* **1998**, *120*, 6394.
(93) Schanda, P.; Kupce, E.; Brutscher, B. *J. Biomol NMR* **2005**, *33*, 199.
(94) Xiao, Z.; Fokkens, M.; Chen, D.; Kok, T.; Proietti, G.; van Merkerk, R.; Poelarends, G. J.; Dekker, F. J. *Eur. J. Med. Chem.* **2020**, *186*, 111849.
(95) Dziedzic, P.; Cisneros, J. A.; Robertson, M. J.; Hare, A. A.; Danford, N. E.; Baxter, R. H.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2015**, *137*, 2996.
(96) Al-Abed, Y.; Dabideen, D.; Aljabari, B.; Valster, A.; Messmer, D.; Ochani, M.; Tanovic, M.; Ochani, K.; Bacher, M.; Nicoletti, F.; et al. *J. Biol. Chem.* **2005**, *280*, 36541.
(97) Cisneros, J. A.; Robertson, M. J.; Valhondo, M.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2016**, *138*, 8630.
(98) Cho, Y.; Jones, B. F.; Vermeire, J. J.; Leng, L.; DiFedele, L.; Harrison, L. M.; Xiong, H.; Kwong, Y. K.; Chen, Y.; Bucala, R.; et al. *J. Biol. Chem.* **2007**, *282*, 23447.
(99) Chi, Y.; Duan, J.; Qi, X.; Chen, G. *Bioelectrochemistry* **2003**, *60*, 37.
(100) Cassidei, L.; Dell'Atti, A.; Sciacovelli, O. *Zeitschrift für Naturforschung C* **1980**, *35*, 1.
(101) Nguyen, D. P.; Sladek, R. N.; Do, L. H. *Tetrahedron Lett.* **2020**, *61*, 152196.
(102) Hong. China 2020.
(103) Mosin, O. V.; Shvets, V. I.; Skladnev, D. A.; Ignatov, I. *Biomed Khim* **2014**, *60*, 448.
(104) Knox, W. E.; Pitt, B. M. *J. Biol. Chem.* **1957**, *225*, 675.
(105) Mühlhahn, P.; Bernhagen, J.; Czisch, M.; Georgescu, J.; Renner, C.; Ross, A.; Bucala, R.; Holak, T. A. *Protein Sci.* **1996**, *5*, 2095.
(106) Xu, Y. L.; Lin, H. Y.; Cao, R. J.; Ming, Z. Z.; Yang, W. C.; Yang, G. F. *Bioorg. Med. Chem.* **2014**, *22*, 5194.
(107) Lin, J. F.; Sheih, Y. L.; Chang, T. C.; Chang, N. Y.; Chang, C. W.; Shen, C. P.; Lee, H. J. *PLoS One* **2013**, *8*, e69733.
(108) Crouch, N. P.; Lee, M. H.; Iturriagagoitia-Bueno, T.; MacKinnon, C. H. *Methods Enzymol* **2000**, *324*, 342.
(109) Pan, M.; Yu, Y.; Ai, H.; Zheng, Q.; Xie, Y.; Liu, L.; Zhao, M. *Nat Struct Mol Biol* **2021**, *28* (7), 614-625. DOI: 10.1038/s41594-021-00617-2.
(110) Chen, C. Y.; Chang, Y. C.; Lin, B. L.; Huang, C. H.; Tsai, M. D. *J Am Chem Soc* **2019**, *141* (51), 19983-19987. DOI: 10.1021/jacs.9b10687.
(111) Zhao, P.; Truong, T. T.; Merlin, J.; Sexton, P. M.; Wootten, D. *Biochem Pharmacol* **2022**, *199*, 114985. DOI: 10.1016/j.bcp.2022.114985.
(112) Wu, B.; Chien, E. Y.; Mol, C. D.; Fenalti, G.; Liu, W.; Katritch, V.; Abagyan, R.; Brooun, A.; Wells, P.; Bi, F. C.; et al. *Science* **2010**, *330* (6007), 1066-1071. DOI: 10.1126/science.1194396.
(113) Karthikeyan, S.; Zhou, Q.; Osterman, A. L.; Zhang, H. *Biochemistry* **2003**, *42* (43), 12532-12538. DOI: 10.1021/bi035450t.
(114) Henzler-Wildman, K.; Kern, D. *Nature* **2007**, *450* (7172), 964-972. DOI: 10.1038/nature06522.
(115) David, J. R. *Proc Natl Acad Sci U S A* **1966**, *56* (1), 72-77. DOI: 10.1073/pnas.56.1.72.
(116) Odh, G.; Hindemith, A.; Rosengren, A. M.; Rosengren, E.; Rorsman, H. *Biochem Biophys Res Commun* **1993**, *197* (2), 619-624. DOI: 10.1006/bbrc.1993.2524.
(117) Pantouris, G.; Serys, M.; Yuasa, H. J.; Ball, H. J.; Mowat, C. G. *Amino Acids* **2014**, *46* (9), 2155-2163. DOI: 10.1007/s00726-014-1766-3.

(118) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; et al. *J Comput Chem* **2010**, *31* (4), 671-690. DOI: 10.1002/jcc.21367.
(119) Phillips, J. C.; Hardy, D. J.; Maia, J. D. C.; Stone, J. E.; Ribeiro, J. V.; Bernardi, R. C.; Buch, R.; Fiorin, G.; Henin, J.; Jiang, W.; et al. *J Chem Phys* **2020**, *153* (4), 044130. DOI: 10.1063/5.0014475.
(120) Luo, J.; Bruice, T. C. *Proc Natl Acad Sci U S A* **2002**, *99* (26), 16597-16600. DOI: 10.1073/pnas.262667599.
(121) Arakawa, T.; Kita, Y.; Timasheff, S. N. *Biophys Chem* **2007**, *131* (1-3), 62-70. DOI: 10.1016/j.bpc.2007.09.004.
(122) Fenwick, R. B.; Orellana, L.; Esteban-Martin, S.; Orozco, M.; Salvatella, X. *Nat Commun* **2014**, *5*, 4070. DOI: 10.1038/ncomms5070.
(123) Lee, N. Y.; Koland, J. G. *Protein Sci* **2005**, *14* (11), 2793-2803. DOI: 10.1110/ps.051630305.
(124) Schumann, F.; Hoffmeister, H.; Bader, R.; Schmidt, M.; Witzgall, R.; Kalbitzer, H. R. *J Biol Chem* **2009**, *284* (36), 24372-24383. DOI: 10.1074/jbc.M109.025635.
(125) Mechaly, A. E.; Haouz, A.; Miras, I.; Barilone, N.; Weber, P.; Shepard, W.; Alzari, P. M.; Bellinzoni, M. *FEBS Lett* **2012**, *586* (11), 1606-1611. DOI: 10.1016/j.febslet.2012.04.034.
(126) Deprez, C.; Lloubes, R.; Gavioli, M.; Marion, D.; Guerlesquin, F.; Blanchard, L. *J Mol Biol* **2005**, *346* (4), 1047-1057. DOI: 10.1016/j.jmb.2004.12.028.
(127) Celic, A. S.; Petri, E. T.; Benbow, J.; Hodsdon, M. E.; Ehrlich, B. E.; Boggon, T. J. *J Biol Chem* **2012**, *287* (21), 17232-17240. DOI: 10.1074/jbc.M112.354613.
(128) Kisselev, O. G.; Downs, M. A.; McDowell, J. H.; Hargrave, P. A. *J Biol Chem* **2004**, *279* (49), 51203-51207. DOI: 10.1074/jbc.M407341200.
(129) Nishihira, J.; Fujinaga, M.; Kuriyama, T.; Suzuki, M.; Sugimoto, H.; Nakagawa, A.; Tanaka, I.; Sakai, M. *Biochem Biophys Res Commun* **1998**, *243* (2), 538-544. DOI: 10.1006/bbrc.1998.8123.
(130) Calandra, T.; Roger, T. *Nat Rev Immunol* **2003**, *3* (10), 791-800. DOI: 10.1038/nri1200.
(131) Alampour-Rajabi, S.; El Bounkari, O.; Rot, A.; Muller-Newen, G.; Bachelerie, F.; Gawaz, M.; Weber, C.; Schober, A.; Bernhagen, J. *FASEB J* **2015**, *29* (11), 4497-4511. DOI: 10.1096/fj.15-273904.
(132) Schwartz, V.; Lue, H.; Kraemer, S.; Korbiel, J.; Krohn, R.; Ohl, K.; Bucala, R.; Weber, C.; Bernhagen, J. *FEBS Lett* **2009**, *583* (17), 2749-2757. DOI: 10.1016/j.febslet.2009.07.058.
(133) Yang, S.; He, P.; Wang, J.; Schetter, A.; Tang, W.; Funamizu, N.; Yanaga, K.; Uwagawa, T.; Satoskar, A. R.; Gaedcke, J.; et al. *Cancer Res* **2016**, *76* (13), 3838-3850. DOI: 10.1158/0008-5472.CAN-15-2841.
(134) Baugh, J. A.; Chitnis, S.; Donnelly, S. C.; Monteiro, J.; Lin, X.; Plant, B. J.; Wolfe, F.; Gregersen, P. K.; Bucala, R. *Genes Immun* **2002**, *3* (3), 170-176. DOI: 10.1038/sj.gene.6363867.
(135) Thiele, M.; Kerschbaumer, R. J.; Tam, F. W.; Volkel, D.; Douillard, P.; Schinagl, A.; Kuhnel, H.; Smith, J.; McDaid, J. P.; Bhangal, G.; et al. *J Immunol* **2015**, *195* (5), 2343-2352. DOI: 10.4049/jimmunol.1500572.
(136) Schinagl, A.; Thiele, M.; Douillard, P.; Volkel, D.; Kenner, L.; Kazemi, Z.; Freissmuth, M.; Scheiflinger, F.; Kerschbaumer, R. J. *Oncotarget* **2016**, *7* (45), 73486-73496. DOI: 10.18632/oncotarget.11970.
(137) Zan, C.; Yang, B.; Brandhofer, M.; El Bounkari, O.; Bernhagen, J. *FASEB J* **2022**, *36* (11), e22601. DOI: 10.1096/fj.202201213R.

(138) Guo, D.; Guo, J.; Yao, J.; Jiang, K.; Hu, J.; Wang, B.; Liu, H.; Lin, L.; Sun, W.; Jiang, X. *Int J Cancer* **2016**, *139* (9), 2056-2067. DOI: 10.1002/ijc.30278.
(139) Song, S.; Liu, B.; Habibie, H.; van den Bor, J.; Smit, M. J.; Gosens, R.; Wu, X.; Brandsma, C. A.; Cool, R. H.; Haisma, H. J.; et al. *EBioMedicine* **2021**, *68*, 103412. DOI: 10.1016/j.ebiom.2021.103412.
(140) Miller, E. J.; Li, J.; Leng, L.; McDonald, C.; Atsumi, T.; Bucala, R.; Young, L. H. *Nature* **2008**, *451* (7178), 578-582. DOI: 10.1038/nature06504.
(141) Qi, D.; Atsina, K.; Qu, L.; Hu, X.; Wu, X.; Xu, B.; Piecychna, M.; Leng, L.; Fingerle-Rowson, G.; Zhang, J.; et al. *J Clin Invest* **2014**, *124* (8), 3540-3550. DOI: 10.1172/JCI73061.
(142) Kim, B. S.; Tilstam, P. V.; Arnke, K.; Leng, L.; Ruhl, T.; Piecychna, M.; Schulte, W.; Sauler, M.; Frueh, F. S.; Storti, G.; et al. *FASEB J* **2020**, *34* (3), 4219-4233. DOI: 10.1096/fj.201901511R.
(143) Vincent, F. B.; Lin, E.; Sahhar, J.; Ngian, G. S.; Kandane-Rathnayake, R.; Mende, R.; Hoi, A. Y.; Morand, E. F.; Lang, T.; Harris, J. *Clin Transl Immunology* **2018**, *7* (12), e1042. DOI: 10.1002/cti2.1042.
(144) Kabsch, W. *Acta Crystallogr D Biol Crystallogr* **2010**, *66* (Pt 2), 125-132. DOI: 10.1107/S0907444909047337.
(145) Evans, P. R.; Murshudov, G. N. *Acta Crystallogr D Biol Crystallogr* **2013**, *69* (Pt 7), 1204-1214. DOI: 10.1107/S0907444913000061.
(146) Murshudov, G. N.; Skubak, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A. *Acta Crystallogr D Biol Crystallogr* **2011**, *67* (Pt 4), 355-367. DOI: 10.1107/S0907444911001314.
(147) Emsley, P.; Cowtan, K. *Acta Crystallogr D Biol Crystallogr* **2004**, *60* (Pt 12 Pt 1), 2126-2132. DOI: 10.1107/S0907444904019158.
(148) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. *SoftwareX* **2015**, *1-2*, 19-25. DOI: https://doi.org/10.1016/j.softx.2015.06.001.
(149) Aida, Y.; Pabst, M. J. *J Immunol Methods* **1990**, *132* (2), 191-195. DOI: 10.1016/0022-1759(90)90029-u.
(150) Norkus, E.; Gaidamauskas, E.; Stalnionienė, I.; Crans, D. C. *Heteroatom Chemistry* **2005**, *16* (4), 285-291. DOI: 10.1002/hc.20123.
(151) Smith, R. H. B., Dar, A.C., and Schlessinger, A. *bioRxiv* **2019**.
(152) Al-Abed, Y.; VanPatten, S. *Future Med Chem* **2011**, *3* (1), 45-63. DOI: 10.4155/fmc.10.281.
(153) Orita, M.; Yamamoto, S.; Katayama, N.; Aoki, M.; Takayama, K.; Yamagiwa, Y.; Seki, N.; Suzuki, H.; Kurihara, H.; Sakashita, H.; et al. *J Med Chem* **2001**, *44* (4), 540-547. DOI: 10.1021/jm000386o.
(154) Van den Eynde, B. J.; van Baren, N.; Baurain, J.-F. *Annual Review of Cancer Biology* **2020**, *4* (Volume 4, 2020), 241-256. DOI: https://doi.org/10.1146/annurev-cancerbio-030419-033635.
(155) Edwards, C. J.; Feldman, J. L.; Beech, J.; Shields, K. M.; Stover, J. A.; Trepicchio, W. L.; Larsen, G.; Foxwell, B. M.; Brennan, F. M.; Feldmann, M.; et al. *Mol Med* **2007**, *13* (1-2), 40-58. DOI: 10.2119/2006-000056.Edwards.
(156) Trivedi-Parmar, V.; Robertson, M. J.; Cisneros, J. A.; Krimmer, S. G.; Jorgensen, W. L. *ChemMedChem* **2018**, *13* (11), 1092-1097. DOI: 10.1002/cmdc.201800158.

(157) Dawson, T. K.; Dziedzic, P.; Robertson, M. J.; Cisneros, J. A.; Krimmer, S. G.; Newton, A. S.; Tirado-Rives, J.; Jorgensen, W. L. *ACS Med Chem Lett* **2017**, *8* (12), 1287-1291. DOI: 10.1021/acsmedchemlett.7b00384.
(158) Ioannou, K.; Cheng, K. F.; Crichlow, G. V.; Birmpilis, A. I.; Lolis, E. J.; Tsitsilonis, O. E.; Al-Abed, Y. *Int J Oncol* **2014**, *45* (4), 1457-1468. DOI: 10.3892/ijo.2014.2551.
(159) Fukushima, K.; Furuya, M.; Kamimura, T.; Takimoto-Kamimura, M. *Acta Crystallogr D Struct Biol* **2021**, *77* (Pt 3), 293-299. DOI: 10.1107/S2059798321000474.
(160) Ishimoto, K.; Iwata, T.; Taniguchi, H.; Mizusawa, N.; Tanaka, E.; Yoshimoto, K. *Cytokine* **2012**, *60* (3), 772-777. DOI: 10.1016/j.cyto.2012.07.037.
(161) Li, H.; He, B.; Zhang, X.; Hao, H.; Yang, T.; Sun, C.; Song, H.; Wang, Y.; Zhou, Y.; Zhu, Z.; et al. *Cell Biosci* **2022**, *12* (1), 128. DOI: 10.1186/s13578-022-00867-7.
(162) Ji, H.; Zhang, Y.; Chen, C.; Li, H.; He, B.; Yang, T.; Sun, C.; Hao, H.; Zhang, X.; Wang, Y.; et al. *J Neuroinflammation* **2021**, *18* (1), 130. DOI: 10.1186/s12974-021-02186-z.
(163) Chen, E.; Widjaja, V.; Kyro, G.; Allen, B.; Das, P.; Prahaladan, V. M.; Bhandari, V.; Lolis, E. J.; Batista, V. S.; Lisi, G. P. *J Biol Chem* **2023**, *299* (6), 104729. DOI: 10.1016/j.jbc.2023.104729.
(164) Parkins, A.; Sandin, S. I.; Knittel, J.; Franz, A. H.; Ren, J.; de Alba, E.; Pantouris, G. *Analytical Chemistry* **2023**, *95* (11), 4957-4965. DOI: 10.1021/acs.analchem.2c04969.
(165) Parkins, A.; Das, P.; Prahaladan, V.; Rangel, V. M.; Xue, L.; Sankaran, B.; Bhandari, V.; Pantouris, G. *Structure* **2023**, *31*, 355.
(166) Winner, M.; Meier, J.; Zierow, S.; Rendon, B. E.; Crichlow, G. V.; Riggs, R.; Bucala, R.; Leng, L.; Smith, N.; Lolis, E.; et al. *Cancer Res* **2008**, *68* (18), 7253-7257. DOI: 10.1158/0008-5472.CAN-07-6227.
(167) Parkins, A.; Chen, E.; Rangel, V. M.; Singh, M.; Xue, L.; Lisi, G. P.; Pantouris, G. *Biophys J* **2023**, *122* (7), 1268-1276. DOI: 10.1016/j.bpj.2023.02.019.
(168) Parkins, A.; Pantouris, G. *STAR Protoc* **2023**, *4* (3), 102375. DOI: 10.1016/j.xpro.2023.102375.
(169) Winter, G.; Waterman, D. G.; Parkhurst, J. M.; Brewster, A. S.; Gildea, R. J.; Gerstel, M.; Fuentes-Montero, L.; Vollmar, M.; Michels-Clark, T.; Young, I. D.; et al. *Acta Crystallogr D Struct Biol* **2018**, *74* (Pt 2), 85-97. DOI: 10.1107/S2059798317017235.
(170) Winter, G. *Journal of Applied Crystallography* **2010**, *43* (1), 186-190. DOI: doi:10.1107/S0021889809045701.
(171) Schuttelkopf, A. W.; van Aalten, D. M. *Acta Crystallogr D Biol Crystallogr* **2004**, *60* (Pt 8), 1355-1363. DOI: 10.1107/S0907444904011679.
(172) Haslam, C.; Hellicar, J.; Dunn, A.; Fuetterer, A.; Hardy, N.; Marshall, P.; Paape, R.; Pemberton, M.; Resemannand, A.; Leveridge, M. *J Biomol Screen* **2016**, *21* (2), 176-186. DOI: 10.1177/1087057115608605.
(173) Berraud-Pache, R.; Garcia-Iriepa, C.; Navizet, I. *Front Chem* **2018**, *6*, 116. DOI: 10.3389/fchem.2018.00116.
(174) Argueta, C.; Vargas, J.; Parkins, A.; Ren, J.; Pantouris, G. *Journal of Chemical Education* **2023**, *100* (6), 2312-2319. DOI: 10.1021/acs.jchemed.3c00141.
(175) Mechaly, A. E.; Haouz, A.; Miras, I.; Barilone, N.; Weber, P.; Shepard, W.; Alzari, P. M.; Bellinzoni, M. *FEBS Letters* **2012**, *586* (11), 1606-1611. DOI: https://doi.org/10.1016/j.febslet.2012.04.034.