

國立臺灣大學生命科學院生命科學研究所

碩士論文

Graduate Institute of Life Science

College of Life Science

National Taiwan University

Master Thesis



依受質結合區探討 MTH1 與抑制劑之交互作用研究

Investigation of interaction between MTH1 and
synthesized inhibitors based on substrate binding site

李昱璇

Yu-Hsuan Lee

指導教授：鄭貽生 博士

Advisor: Yi-Sheng Cheng, Ph. D.

中華民國 108 年 6 月

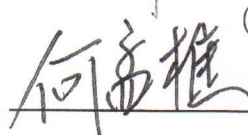
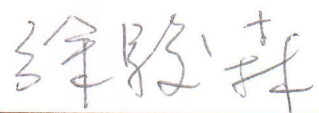
June, 2019

國立臺灣大學（碩）博士學位論文
口試委員會審定書

依受質結合區探討 MTH1 與抑制劑之交互作用
研究 Investigation of interaction between MTH1
and synthesized inhibitors based on substrate
binding site

本論文係李昱璇君（R05B21025）在國立臺灣大學生命
科學學系、所完成之碩（博）士學位論文，於民國 108 年 05
月 15 日承下列考試委員審查通過及口試及格，特此證明

口試委員：

		(簽名)
(指導教授)		
		
_____	_____	
_____	_____	
_____	_____	



致謝

能夠順利完成論文，首先要感謝指導老師鄭貽生老師的協助和建議，也很謝謝口試委員何孟樵老師和徐駿森老師撥空前來並很仔細的提出研究上的問題和指導。

念研究所的3年過程中雖然經過了一些事情而有了波折，碩一決定換實驗室後誤打誤撞進入完全不熟悉的領域。雖然那時候一直覺得自己因為一時間的情緒做出不理性的抉擇，並不斷地懷疑自己能不能順利的從研究所畢業。不過很幸運有遇到有很多幫忙和支援我的人，不管是實驗室的學長和同學們還是新竹同步輻射中心的 beamline staff 們，都作為我在實驗生活中的一大助力。

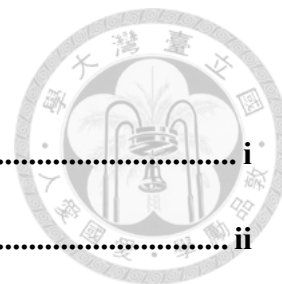
此外，在從研究所畢業的這個時刻，我想特別感謝我的家人讓我在上來台北念研究所的這段期間給予我經濟和心靈上最大的支持。

李昱璇 謹識於

台灣大學生命科學院

2019 年 6 月

Contents



口試委員會審定書	i
致謝	ii
Contents	iii
List of Figures	vi
List of Tables	viii
中文摘要	ix
Abstract	x
List of abbreviations	xii
Chapter 1 Introduction	1
1-1 Carcinogenesis: Distinguishing cancer cells from normal cells	1
1-1-1 ROS homeostasis in cancer development and apoptosis	1
1-1-2 Anticancer drug discovery exploiting ROS-mediated signaling pathways in cancer cell	2
1-2 Practical aspects in cancer therapeutic development	3
1-2-1 Identification and validation of molecular targets in cancer cell	3
1-2-2 High-throughput screening in searching for potential drug candidates and hit-to-lead optimization campaign	4
1-2-3 Biochemical and biophysical approaches in hit-to-lead optimization campaign	5
1-2-4 Thermodynamic profiles of protein-ligand complex formation	5
1-3 MutT homolog 1 (MTH1)- potential target for anticancer drug	7
1-3-1 Physiological function of MTH1 and ROS-modulated apoptosis	7
1-3-2 Protein structure of MTH1	8
1-3-3 MTH1 inhibitor development- perspectives and directions	9
1-4 Specific Aims and objectives	10

Chapter 2 Experimental section	12
2-1 Materials.....	12
2.2 Methods	12
2-2-1 Transformation.....	12
2-2-2 Protein Expression	13
2-2-3 Affinity Chromatography.....	13
2-2-4 IC ₅₀ Determination.....	14
2-2-5 Isothermal titration calorimetry	15
2-2-6 Size-exclusion chromatography.....	15
2-2-7 Crystallization and soaking.....	16
2-2-8 X-ray data collection.....	17
2-2-9 Model building and Refinement	17
Chapter 3 Results.....	19
3-1 MTH1 protein purification	19
3-1-1 Affinity Chromatography.....	19
3-1-2 Size-exclusion chromatography (SEC).....	19
3-2 Determination of binding affinity profiles and structure-activity relationship (SAR).....	20
3-2-1 Fragment-based drug screening and optimization performed with 2-aminopyrimidine-based scaffold	20
3-2-2 Binding affinity profile and SAR study of scaffold 1-derived compounds ..	21
3-2-3 Binding affinity profile of scaffold 2-derived compounds	22
3-3 Thermodynamic study on 2-aminopyrimidine-based compounds derived from scaffold 2	23
3-3-1 Enthalpy-driven binding of 2-aminopyrimidine-based compounds to MTH1	23
3-3-2 Phenomenon of Enthalpy-Entropy Compensation (EEC)	24
3-3-3 Rationalization of observed EEC phenomenon through insights provided by atomic model of MTH1 in complex with 2-aminopyrimidine-based compounds .	25
3-4 Structural analysis of MTH1 complex structure bound with 2-aminopyrimidine -	

based compounds.....	28
3-4-1 Conserve binding mode of pyrimidine-based compounds in the binding site of MTH1	28
3-4-2 Pyrimidine moiety.....	29
3-4-3 R1 and R2-substituent.....	30
3-5 Superposition of MTH1 complexed with product 8-oxo dGMP and compound 29, 30	30
3-5-1 Absence of affinity gain with extended substituent toward the entry of binding pocket	31
3-6 Notable difference in interaction pattern of heterocyclic substituent unraveled by deviation in position occupied by ligand	32
Chapter 4 Discussion	34
4-1 Functional mechanism of MTH1 protein and structural relevance to selectivity toward its substrate	34
4-2 The binding energetics underlying complex formation	35
4-2-1 Intrinsic phenomenon of EEC hinders the elevation of binding affinity	36
4-2-2 Proposed theories for explanation of contradictory effect of enthalpic and entropic component of Gibbs binding free energy	37
4-3 Structure-aided ligand design: perspectives and opportunities	38
4-3-1 Binding pattern observed in published MTH1-ligand complexes	39
4-3-2 Overview of available space left unexplored in MTH1 binding pocket in regard of information derived from explored regions	41
Chapter 5 Conclusion	43
References.....	45
Tables and Figures	51
Appendices	91

List of Figures

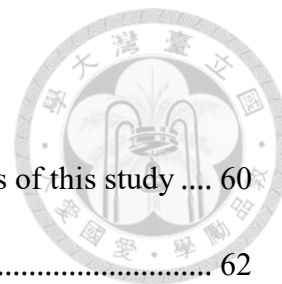


Figure 1. Schematic depiction of simplified work flow and objectives of this study	60
Figure 2. Purification of MTH1 protein from E.coli cell lysate	62
Figure 3. Binding affinity profile of scaffold 1 derivative compounds	64
Figure 4. Binding affinity profile of scaffold 2 derivative compounds	66
Figure 5. Thermodynamic study of scaffold 2 derivatives binding to MTH1 protein ...	68
Figure 6. Phenomenon of EEC observed in scaffold 2 derivatives binding to MTH1 protein	70
Figure 7. Comparison of thermodynamic profiles from structurally related compounds with different substituents incorporated in the R2-position	72
Figure 8. Comparison of thermodynamic profiles from structurally related compounds with different substituents incorporated in the R1-position	74
Figure 9. Phenomenon of enthalpy-entropy compensation gleaned from protein-ligand complex structures	76
Figure 10 Comparison of binding mode of 11 structurally related 2-aminopyrimidine- based compounds uncovered in x-ray crystallographic study.	78
Figure 11. Binding pose of MTH1 product 8-oxo dGMP, compound 29 and 30 in bound state gleaned from crystal structure complexed with MTH1	80
Figure 12. Mutual alignment of MTH1 complex structures bound with substrate, compound 29 and 30	82
Figure 13. Binding mode of compound 22, 24, 29 in MTH1 binding pocket and structural rearrangement of side chains from Phe27 and Met81	84

Figure 14. Binding mode of compound 12, 26, 30 in MTH1 binding pocket and structural rearrangement of side chains from Phe27 and Met81 86

Figure 15. Binding mode of compound 20, 24, 25, 26 in MTH1 binding pocket and structural rearrangement of side chains from Phe27 and Met81 88

Figure 16. Binding mode of compound 12, 22, 29, 30 in MTH1 binding pocket and structural rearrangement of side chains from Phe27 and Met81 90

List of Tables



Table 1. Structure and related biochemical data of scaffold 1-derived compounds ^a	51
Table 2. Structure and related biochemical data of scaffold 2-derived compounds ^a	52
Table 2. Structure and related biochemical data of scaffold 2-derived compounds ^a (continued).....	53
Table 3. Thermodynamic profiles of scaffold 2-derived compounds determined with iTC	54
Table 4. Data collection and refinement statistics for MTH1 crystal structure complexed with 11 structurally-related 2-aminopyrimidine-based compounds	55
Table 4. Data collection and refinement statistics for MTH1 crystal structure complexed with 11 structurally-related 2-aminopyrimidine-based compounds (continued).....	56
Table 4. Data collection and refinement statistics for MTH1 crystal structure complexed with 11 structurally-related 2-aminopyrimidine-based compounds (continued).....	57
Table 4. Data collection and refinement statistics for MTH1 crystal structure complexed with 11 structurally-related 2-aminopyrimidine-based compounds (continued).....	58

中文摘要

Mut T homolog 1 (MTH1) 為 Nudix 水解酶超家族的成員，在癌細胞的增生中扮演關鍵的角色。先前研究指出 MTH1 適合作為抗癌藥物的標靶，並已有 MTH1 抑制劑的體外以及體內成效被發表。本論文為深入探討 MTH1-抑制劑複合體的結合機制，以高通量藥物片段篩選具有抑制 MTH1 的小分子，以其中一組作為骨架，進一步設計修飾 2-氨基嘧啶衍生物共 14 個，依嘧啶環上 5 位碳乙基的有無分類為兩個群組，並分別解析 MTH1 與 11 個 2-氨基嘧啶衍生物複合體結構，輔以生化及熱力學分析，以探討 MTH1 與抑制物複合體結合作用之多方位資訊。

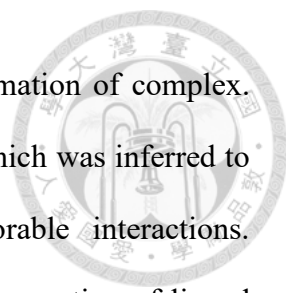
根據已解出之 MTH1 與受質複合體結構，可將 MTH1 之活性口袋區初步劃分為三個特定區域。MTH1 之受質 8-oxo dGTP (8-Oxo-2'-deoxyguanosine-5'-Triphosphate) 具有三部分官能基分別為鹼基，去氧核糖以及三磷酸，在鹼基部分主要由 Asp119, Asp120 及 Trp117 的側鏈鍵結，此區域也是這些抑制劑在結合口袋中的主要穩定作用力，旁側存在由多個疏水性胺基酸 Phe72, Phe74 及 Phe139 側鏈形成之空間，本實驗中將抑制劑的官能基延伸至由 Phe72, Phe74, Phe139 形成的空腔時，並無法提升抑制劑結合能力。從受質的 8 號位置氧原子位於 Phe27 與 Met81 的側鏈附近，且設計抑制劑亦與 Phe27 及 Met81 可形成凡得瓦力交互作用，本區將可提供與 MTH1 專一性結合。根據本研究系列合成抑制物之抑制效率分析，當抑制物延伸至受質之三磷酸的區域時，其結合能力降低三個級數，由 Lys23, Glu52, Glu55, Glu56 及 Glu100 組成的三磷酸結合區在催化水解反應時，會與金屬離子結合並形成水分子網絡，本區可作為未來延伸區域，以提高專一性結合。最後，本研究以受質結合區結構、熱力學分析及酵素動力學探討設計抑制劑的方式，將可作為未來在設計新穎抑制劑時參考依據。

Abstract

Mut T homolog 1 (MTH1) is a member of Nudix hydrolase superfamily and engages in proliferation of cancer cells as critical determinant. In previous study, it was proposed that MTH1 serves as ideal anticancer target for anticancer therapeutic development, and *in vitro* and *in vivo* efficacy were reported for several MTH1 inhibitors.

In this study, to elucidate the mechanism of MTH1-inhibitor complex formation extensively, high-throughput screening (HTS) was performed on fragment library intended to identify small molecule compound with moderate inhibitory activity against MTH1. From identified hits, one hit compound was utilized as scaffold and a series of 2-aminopyrimidine derivatives comprising 14 structurally-related compounds were subsequently synthesized based on HTS results. These 2-aminopyrimidine-based compounds are categorized into two groups based on the presence of ethyl substituent on pyrimidine ring. Explicitly, MTH1 complex structures bound with eleven 2-aminopyrimidine derivatives were obtained and supplemented with statistics from biochemical and thermodynamic analyses to complement the intricate binding interaction of MTH1-inhibitor complex from multiple aspects.

Based on MTH1-substrate complex structure, the substrate binding pocket can be divided into three subregions which are occupied by base, deoxyribose and triphosphate group of substrate 8-oxo dGTP, respectively. In the base binding region, the bonds are formed by side chains of Asp119, Asp120 and Trp117 which are essential in binding of synthesized inhibitors in the binding pocket. Adjacent to the base binding motif, a minor cavity formed by side chains of hydrophobic residues Phe72, Phe74, Phe139 allows further exploration with substituents. However, no enhanced binding affinity was observed in exploring this cavity. Additionally, in the analysis of MTH1-inhibitor complex structures, it was observed that residues Phe27 and Met81 form van der Waals



interactions selectively with synthesized compounds upon formation of complex. These residues sit in the proximity of 8-oxo group from substrate which was inferred to contribute to specificity of MTH1 through formation of favorable interactions. Intrinsically, as implied in structure-activity relationship study, incorporation of ligand moiety which occupies triphosphate binding region of MTH1 reduced binding affinity by three orders of magnitude. Structurally, triphosphate binding region constituted of Lys23, Glu52, Glu55, Glu56 and Glu100 facilitates catalysis of hydrolytic reaction through formation of water molecule network coordinated by metal ions. Therefore, it is indicated that this subregion can be further exploited for enhancement of specificity of inhibitors. In conclusion, through investigation of MTH1 substrate binding pocket with structural analysis, thermodynamic studies and enzyme kinetic analysis in this study, directions are provided for design of novel inhibitors in the future.

List of abbreviations



Asn	Asparagine
Asp	Aspartic acid
cDNA	complementary DNA
Cryo-EM	Cryogenic electron microscopy
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside triphosphates
DTT	Dithiothreitol
E.coli	Escherichia coli
EEC	Enthalpy-entropy compensation
FPLC	Fast protein liquid chromatography
Glu	Glutamic acid
Gly	Glycine
H-bonds	hydrogen bonds
HEPES	2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
IC50	The half maximal inhibitory concentration
IPTG	Isopropyl β -D-1-thiogalactopyranoside
ITC	isothermal titration calorimetry
kDa	kilodalton
K _d	dissociation constant
K _i	inhibitory constant
LB	Lysogeny broth
Leu	Leucine
LiSO ₄	Lithium sulfate
Lys	Lysine
Mg acetate	Magnesium acetate
MgCl ₂	Magnesium chloride
MTH1	MutT homolog-1
NaCl	Sodium chloride
NCS	Non-crystallographic symmetry
NMR	Nuclear Magnetic Resonance
NSRRC	National Synchrotron Radiation Research Center
PDB	Protein Data Bank
PEG	Polyethylene glycol
Phe	Phenylalanine

PMSF	Phenylmethane sulfonyl fluoride
PPi	Pyrophosphate
RNA	ribonucleic acid
ROS	Reactive oxygen species
SAR	Structure–activity relationship
TEMED	N,N,N',N'-tetramethylethane-1,2-diamine
Thr	Threonine
TLS	Translation–Libration–Screw-rotation
Tris	2-amino-2-(hydroxymethyl)propane-1,3-diol
Tris-HCl	2-amino-2-(hydroxymethyl)propane-1,3-diol hydrochloride
Trp	Tryptophan
Tween-20	Polysorbat 20
Tyr	Tyrosine
Val	Valine
2-oxo dATP	2'-deoxy-2-hydroxyadenosine
8-oxo dGTP	8-oxo-2'-deoxyguanosine-5'-triphosphate



Chapter 1 Introduction



1-1 Carcinogenesis: Distinguishing cancer cells from normal cells

Cell, the most basic functional unit of living organism, propagates by numerous biochemical reactions which are modulated by intricate network of intersecting signal transduction pathways (Evan and Vousden 2001). The signaling pathways which involve in cell growth, proliferation and apoptosis are integral modulators in maintaining physiological activities and fates of individual cell. As the fine-tuned modulation network is interrupted through dysregulation of participating biological molecules, carcinogenesis of normal cell is elicited as consequence.

In general, cancer is regarded as a group of immortal cells attributed to their dysregulated growth resulting mainly from defects in signal transduction pathways. In comparison to normal cells which undergo cell cycle and enter apoptosis under control of regulated molecular network. It is noteworthy that reprogramming of signaling network is essential in cancer as these immortal cells thrive in environment lethal to their normal counterparts (Harris et al., 2011).

1-1-1 ROS homeostasis in cancer development and apoptosis

The immortality of cancer cells is proposed to arise from altered molecular network modulating physiological events. To trigger and facilitate proliferation of cancer cells in microenvironment hostile to normal cells (Amoedo et al., 2013), adaptation of gene regulation in favor of the rigorous growth is pivotal for the survival of cancer cells.

Under elevated rate of metabolic activity common in fast-growing cancer cells, reactive oxygen species (ROS) is excessively generated as byproduct of metabolism.

In comparison to normal cells which grow and proliferate under regulation of intrinsic molecular network, homeostasis of intracellular ROS is critical in maintaining cellular function of cancer cells with elevated ROS level (Liou and Storz, 2010).

To sustain the physiological functions of cancer cells, redox balance is maintained through production of antioxidant proteins in response to increased levels of ROS (Kumari et al., 2018). Furthermore, to complement the understanding of impact of ROS on cancer initiation, progression and apoptosis, the role of ROS in cancer development has been extensively reviewed earlier from the perspective of cellular metabolism (Cairns et al., 2011).

1-1-2 Anticancer drug discovery exploiting ROS-mediated signaling pathways in cancer cell

Explicitly, to facilitate cancer progression, redox balance related to modulation of intracellular ROS level is regarded as key player in cancer physiology (Waris and Ahsan, 2006). Therefore, to exploit this well-characterized feature of cancer cells for anticancer agent development, it has been proposed in numerous studies that establishment of knowledge derived from ROS-mediated cellular events facilitates identification of molecular targets for drug discovery campaigns (Moloney and Cotter, 2018).

ROS-mediated signaling pathway for elimination of cancer cells, strategies were developed based on established knowledge of ROS-mediated cancer progression. Therefore, two opposite approaches featuring pro-oxidants and antioxidants which promote cell apoptosis through distinguished working mechanisms provide opportunities for anticancer agent development. Intended to elevate intracellular oxidative stress to induce cell apoptosis, pro-oxidants exert its effects through promoting accumulation of ROS in cancer cells. Oppositely, antioxidants inhibit signaling molecules from ROS-

mediated signal transduction pathway and suppress tumor growth via abrogation of ROS signaling (Trachootham et al., 2009).



1-2 Practical aspects in cancer therapeutic development

Attributed to intensive research on molecular mechanism for modulation of cellular events in cancer cells over decades, exploitation of established knowledges of cancer biology for anticancer agent development is fueled and several therapeutics including imatinib (Gleevec), gefitinib (Iressa), ceritinib (Zykadia), crizotinib (Xalkori) have been successfully developed in this period. In practical aspect, anticancer drug discovery is generally initiated from theoretical strategies with aims to restore programmed cell death in cancer cells (Stephen, 2005). As immortality of cancer cells stem from deregulation of apoptotic pathway, elimination of these immortal cells depend on intervention via therapeutic agents designed for targeting specific signaling molecules to elicit cell apoptosis.

Followed by concretization of abstract ideas derived from acquired knowledge of cancer biology, extensive search on druggable targets in molecular signaling pathway is performed as routine work in cancer drug discovery.

1-2-1 Identification and validation of molecular targets in cancer cell

In the initial stage of anticancer drug discovery campaign which spans various disciplines of scientific research, the most fundamental while influential issue is always concerned with identification of optimal molecular target with significant relevance to the disease. Recalls that the hypothesis which is supportive of the notion that modulating key molecular targets in cancer cells serves as dominant strategy in anticancer therapeutic development, through selection of specific targets for design of therapeutic agent several

problems emerge at this time point which may hinder further progress and success in anticancer drug discovery campaign.

To address the problems emerged from selection of exploitable targets in cancer cells, validation of identified molecular targets is the primary mission to be fulfilled to ensure that designed therapeutics is clinically effective through acting on selected target.

As paradigm of target validation, three lineages of approaches have been proposed and reviewed by John D. Benson et al. To investigate the formulated mechanism of therapeutic agent acting on target molecules which is related to establishment of its clinical efficacy, validation through setting up experimental trials and collection of data in genetic studies, cell-based and animal model systems plays a key role in the success of anticancer drug discovery campaign. As biological system is far more complex than simplified overview of signaling transduction pathways which modulate cellular events, rationalization of action of drug molecules designed for its intracellular targets in *in vitro* and *in vivo* systems is therefore fundamental in optimization of specificity, selectivity, physicochemical properties for drug-likeness and ADME profiles of designed drugs.

1-2-2 High-throughput screening in searching for potential drug candidates and hit-to-lead optimization campaign

To identify adequate starting point for development of anticancer drug, high-throughput screening (HTS) is routinely performed on established library comprised of synthesized chemicals in large amount. In regard of utility of identified hits from HTS in subsequent hit-to-lead optimization campaign, criteria for filtration of trivial fragment compounds in HTS library is set out to ensure the integrity of drug candidates. Pragmatically, to complement this context assessment of structure-activity relationship (SAR) and absorption, distribution, metabolism and excretion properties is conducted

(Konrad et al., 2003).

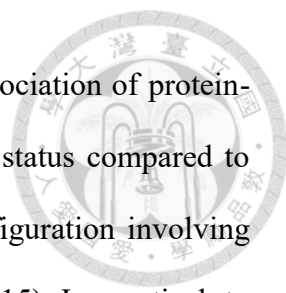
Based on results from HTS performed in conditioned biochemical assays, identified hits with moderate affinity toward a given target are further optimized. Therefore, investigation of hit compounds in preset experimental trials is performed to evaluate possibilities of creating lead series with favorable drug properties and optimal selectivity profile toward its targets. As outcome of hit-to-lead optimization campaign, modifications on initial scaffold compound may lead to compensation effects in biochemical affinity for target and drug-likeness properties. Despite of this dilemma, it is to recognize that establishment of complementary profile for ideal drug is comprised of balanced biological and physicochemical properties.

1-2-3 Biochemical and biophysical approaches in hit-to-lead optimization campaign

Throughout optimization program of hits identified from high-throughput screening on compound library, information related to interaction profiles of hit compounds with their target is on demand for development of more suitable drug candidate at this stage. As hit-to-lead optimization campaign usually initiates with structure-activity relationship (SAR) studies conducted in biochemical assays. Validation of binding event and elucidation of exact interaction profile are important in decision making during generation of optimized lead series.

In practical terms, determination of such interaction profiles is dependent on several well-established biophysical techniques including x-ray crystallography, nuclear magnetic resonance spectrometry, surface plasmon resonance spectrometry and isothermal titration calorimetry (Renaud et al., 2016).

1-2-4 Thermodynamic profiles of protein-ligand complex formation

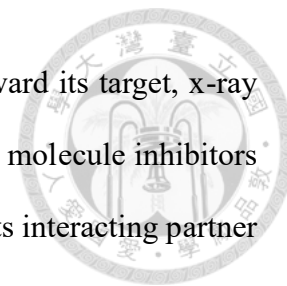


From the perspective of thermodynamics, association and dissociation of protein-ligand complex is always accompanied by fluctuation in energetic status compared to unbound state which is attributed to rearrangement of overall configuration involving formation and breakage of chemical bonds (Geschwindner et al., 2015). In practical, to gain more insight into details of protein-ligand association which is essentially highly dynamic in nature, through recording thermodynamic signal from binding event related to the formation of protein-ligand complex with isothermal titration calorimeter (iTC) which is praised for its convenience and simplicity in practical measurement (Schnapp et al., 2016). Furthermore, elucidation of driving forces facilitating the binding event is enabled through calculating the enthalpic and entropic component of Gibbs free energy which is given as $\Delta G = \Delta H - T\Delta S$. As illustrated in published research articles (Klebe, 2015), issues associated with thermodynamic analysis of protein-ligand association in various independent case studies came into spotlight which in certain degree further complicate the interpretation of measured iTC data (Sharp, 2001). Consequently, making judgement solely on thermodynamic profiles of each drug candidate might be biased in some occasion. To overcome this dilemma, interpretation of thermodynamic signatures of compound series binding to its target molecule derived from iTC data set is complemented by x-ray structure of protein-ligand complex which yields more reliable information for subsequent optimization program.

1-2-5 X-ray crystallography-static picture of protein-ligand complex

In searching for bioactive compounds which can be exploited as therapeutic agent for diseases, molecular recognition of drug target driven by physical forces which contribute to specificity and selectivity of the binding interaction serves as key element in structure-based drug design (Kuhn et al., 2011) . In support of developing drug

molecule which is characterized by high affinity and specificity toward its target, x-ray structure of target macromolecule complexed with an array of small molecule inhibitors sheds light on functional mechanism of how drug molecules act on its interacting partner (Deller et al., 2015).



1-3 MutT homolog 1 (MTH1)- potential target for anticancer drug

It is well-recognized that identification of disease-relevant targets plays a key role in drug discovery campaign. In the case of anticancer drug development, identified targets stem mainly from apoptosis-relevant biomolecules which include the BCL2 family of anti-apoptotic proteins, tumor necrosis factor (TNF)-related apoptosis-inducing ligand (TRAIL) receptors, inhibitor of apoptosis (IAP) proteins and MDM2 (Stephen, 2005). It is revealed that promoting apoptosis serves as main strategy in killing these immortal cells. Whereas, regarding the issue of intra-tumor heterogeneity proposed in previous studies, it is reasoned by Gad et al. that targeting genetic defects is limited and suggested that exploiting cancer phenotype may serve as more optimal strategy.

However, followed by launch of MTH1 inhibitor development undertaken by Gad et al., intensive debate on MTH1 as a validated anticancer target emerged as an issue for clinical evaluation of anticancer efficacy of drug candidates targeting MTH1 derived from lead optimization archives (Kawamura et al., 2016).

1-3-1 Physiological function of MTH1 and ROS-modulated apoptosis

As one of the members of Nudix hydrolase superfamily, MutT homolog 1 (MTH1) has become popular in cancer research recent years which is attributed to the fact that MTH1 has been validated as promising anticancer target in either *in vitro* or *in vivo* experimental assays from the results published in 2014 by Gad et al. However, to support

the statement that MTH1 is suitable to serve as an anticancer target, the function of MTH1 and its relevance in the pathogenesis of cancer should be further elucidated (Kettle et al., 2016).

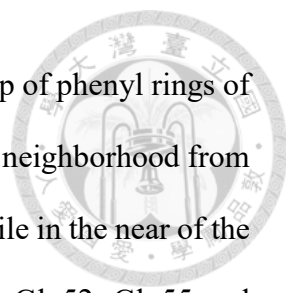


At the beginning of this chapter, ROS is referred to as important modulator in the physiological function of cell. In relation to the uncontrolled proliferation of cancer cells, elevated level of ROS leads to cell apoptosis in the absence of antioxidants such as housecleaning enzymes. This information reveals that inhibition of the activity of these cellular components maintaining the proliferation of cancer cells can be utilized in developing strategy for cancer therapies (Luo et al., 2010). In the case of MTH1 which is characterized as housecleaning enzyme in removing nucleobases oxidized by ROS to prevent cancer cell from entering programmed cell death, is regarded as promising target for development of small molecule inhibitors used in treatment of cancer (Gad et al., 2014).

1-3-2 Protein structure of MTH1

As a member of Nudix hydrolase superfamily, the overall structure of MTH1 is made up of typical Nudix fold α - β - α sandwich structure composed of two α -helices and eight β -strands, in which a mixed β -sheet is sandwiched between two α -helices (Bessman et al., 1996). Beside secondary structure building up the whole structural architecture, atomic detail of the hydrolase is also illustrated as followed.

Firstly, with insight into the substrate binding pocket of MTH1, key residues Asn-33, Asp-119 and Asp-120 participate in substrate recognition through hydrogen bonding interaction with base moiety of the substrate (Sakai et al., 2002) and therefore regarded as the most well-characterized feature for selectivity and specificity of MTH1 protein (Waz et al., 2017).

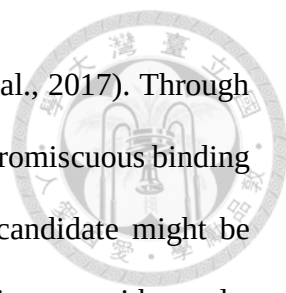


Besides, in the substrate binding pocket, a minor cavity made up of phenyl rings of residues Phe72, Phe74 and Phe139 is well-defined and located in the neighborhood from recognition element Asp119-Asp120 motif (Nissink et al., 2016). While in the near of the entry of MTH1 binding pocket, a Glu cluster composed of residues Glu52, Glu55 and Glu56, Glu100 which lies in the Nudix motif is featured as region for catalyzing the hydrolysis of the substrate 8-oxo dGTP and 2-oxo dATP via binding network made up of metal ion and water molecules (Mishima et al., 2004). In conclusion, structural features mentioned above is generally thought to be associated with substrate selectivity and specificity of MTH1.

1-3-3 MTH1 inhibitor development- perspectives and directions

From the time when the statement that hydrolase MTH1 serves as promising target for anticancer drug was proposed by Gad, H. et al., inhibitor development with aims to deplete MTH1 catalytic activity has been undertaken by several research groups during this period (Gad et al., 2014) (Huber et al., 2014) (Kettle et al., 2016) (Rahm et al., 2018). With the launch of anticancer drug discovery campaign targeting protein MTH1 in cancer cells which is non-essential in its normal counterparts, numerous structural motifs with moderate affinity toward MTH1 have been identified. Therefore, structure-based drug design is propagated with an abundance of crystal structures of MTH1 protein in either bound or unbound state accessible on the website of Protein data bank (PDB).

Over years, MTH1 crystal structure complexed with lead series optimized from structurally-irrelevant hit compounds provide insights into varied interaction pattern of MTH1-inhibitor complexes (Gad et al., 2014) (Huber et al., 2014) (Kettle et al., 2016) (Rahm et al., 2018). From these independently conducted optimization campaigns, design of potent inhibitor with desired selectivity profile is guided by key binding interactions



in MTH1-substrate complex structure (Nissink et al., 2016) (Waz et al., 2017). Through characterization of structural components in MTH1 binding pocket, promiscuous binding behavior is unraveled and implies that selectivity profile of drug candidate might be affected without establishing specific binding interactions. In consistence with results from off-target effect reported for MTH1 inhibitor TH588 in spite of reasonable selectivity profile from test on a panel of 87 enzymes, GPCRs, kinases, ion channels and transporters (Gad et al., 2014). Therefore, selectivity profile of MTH1 inhibitor has been augmented by mimicking interaction pattern adapted by substrate molecule in MTH1 binding pocket (Kettle et al., 2016) (Rahm et al., 2018).

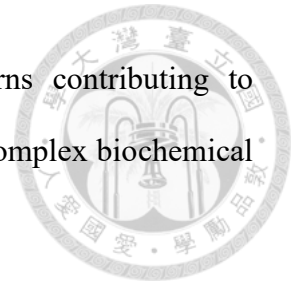
1-4 Specific Aims and objectives

To address the issue of promiscuous binding behavior observed for designed ligands targeting MTH1, key structural features in MTH1 binding pocket are explored with substituents to elucidate mechanism of protein-ligand complex formation.

In this study, a series of analogue compounds was synthesized based on 2-aminopyrimidine scaffold. Subsequently, structure-activity relationship (SAR) was established which provides information on affinity profiles of structurally-related compounds. In addition, visualization of atomic structure of MTH1-compound complexes and structural alignment with MTH1-substrate structure enables dissection of binding patterns observed for different ligands. Explicitly, underlying mechanism of formation of MTH1 complex structure is investigated by isothermal titration calorimetry (iTC) which reveals thermodynamic profiles of analyzed molecules (Figure 1).

Remarkably, research on binding interaction of MTH1 with a series of structurally related compounds unravels conformational changes related to accommodation of different ligands. To apply these acquired information in further optimization program, it

is important to formulate binding motifs and interaction patterns contributing to specificity and selectivity of protein toward its binding partners in complex biochemical environment.



Chapter 2 Experimental section



2-1 Materials

For expression of MTH1 protein, plasmid with cDNA encoding human MTH1 used in this experiment is provided by Ph.D. candidate Cheng Peng in our lab. Test compounds in this study were designed and synthesized by professor Ji-Wang Chern and graduate student Zhe-Hwa Cheng from graduate institute of pharmacy, National Taiwan University. Related experimental details regarding chemical design and synthesis is described in master thesis written by Zhe-Hwa Chen.

2.2 Methods

2-2-1 Transformation

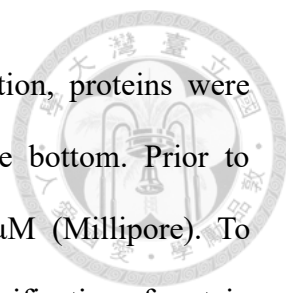
Rosetta competent cells used in this experiment were prepared previously according to protocol of our lab and stored at -80°C which serve as material for transformation experiment. Prior to each transformation, competent cells were thawed on ice for 10 minutes followed by reaction with $2\ \mu\text{l}$ of plasmid DNA (approximately $75\ \text{ng/nl}$) which should be mixed by pipetting several times then leave the mixture on ice for 2 minutes. The mixture was heat shocked in water bath at 42°C for 90 seconds followed by incubation for 2 minutes on ice. The mixture made up of competent cells and plasmid DNA was recovered in 1 ml lysogeny broth (without antibiotics) and further incubated at 37°C with shaking (180 rpm) for 1 hr which provide sufficient time for expression of antibiotic resistant genes. The competent cells were collected followed by incubation with centrifugation at 9,000 rpm for 3 minutes and resuspended in fresh LB medium. For plating the cells, LB-agar plate containing kanamycin and chloramphenicol were warmed up to room temperature previously and the plated cells are incubated at 37°C overnight.

2-2-2 Protein Expression

The cDNA encoding human MTH1 was expressed from the pET2.8a vector using *E. coli* strain Rosetta. For large-scale expression of MTH1 recombinant protein, colonies of transformants on LB agar plate were selected and left to grow in 50 ml LB medium containing kanamycin and chloramphenicol at 37°C for 8-12 hours. Prior to IPTG induction, 2 liters of LB medium should be freshly prepared and autoclaved. Subsequently, 5 ml of LB medium containing *E. coli* expressing the target protein is added to 500 ml autoclaved LB medium with appropriate antibiotics added and incubated at 37°C for 2 hours with shaking at (180-200 rpm) until the O.D. value has reached 0.4-0.5. To induce the expression of MTH1 recombinant protein 1 M IPTG was added to a final concentration of 1 mM followed by incubation for 2.5 hours at 37°C with vigorous shaking (200-220 rpm). The cells were harvested with centrifugation at 8,000 rpm for 20 minutes and during the process the temperature was kept at 4 °C. Collected cell pellet was stored at 20°C.

2-2-3 Affinity Chromatography

All buffer used in purification were prepared and filtered using Stericup (Millipore) with pore size of 0.22 µm in advance. Prior to cell lysis, the cell pellet stored at 20 °C was thawed on ice for 10-20 minutes followed by dissolving in lysis buffer (20 mM Tris-HCl pH 7.4, 500 mM NaCl, 10% Glycerol, 10 mM Imidazole, 2 mM TCEP) which is supplied with freshly prepared stock solution of 0.5 M PMSF and protease inhibitor cocktail (Roche). The process of cell lysis was conducted on sonicator (Misonix Sonicator 3000) under the condition of preset program (pulse-on time= 10 s, pulse-off time= 10 s, total=5 min). Lysis solution was kept on ice during the procedure to maintain the stability of target protein. To remove cell debris, cell lysis was centrifugated at 12,500 rpm for 25



minutes with temperature set at 4 °C. As outcome of centrifugation, proteins were dissolved in supernatant with pellet formed by cell debris at the bottom. Prior to purification, Supernatant was filtered using Syringe Filter 0.22 µM (Millipore). To acquire purified protein solution containing only the target protein, purification of protein solution was performed using HisTrap FF column (GE healthcare) to capture his-tagged recombinant protein which was followed by elution with buffer containing 500 mM imidazole (Appendix 7). Purification of target protein was operated on ÄKTA Protein Purification Systems (GE healthcare), detailed settings in purification program are listed as below. flow rate=5 ml/min, pressure limit= 0.3 mpa, Equilibrium volume= 5 ml, Sample injection volume= 120 ml, wash volume= 15 ml, elution volume = 200 ml.

2-2-4 IC₅₀ Determination

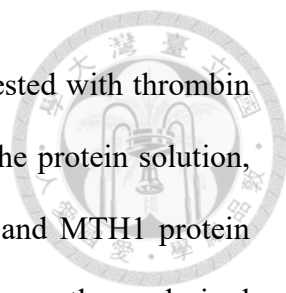
The inhibition activity of synthesized compounds against MTH1 protein is measured in the presence of substrate 8-oxo dGTP (Trilink Biotechnologies) and IC₅₀ value of test compound is derived from dose-response curve fitted by data from inhibition assay. Firstly, compounds to be analyzed were dissolved in dimethyl sulfoxide (DMSO) followed by serial dilution in 1:4 starting from an initial concentration of 20 mM. For a single measurement, 12-15 dilutions of test compound were prepared prior to mixing with MTH1 protein and substrate 8-oxo dGTP which were diluted to final concentration of 4.8 nM and 100 µM in reaction buffer (10 mM Tris-acetate pH7.5, 40 mM NaCl, 10 mM Mg acetate, 1 mM DTT and 0.005% Tween20), respectively. Followed by sample preparation, reaction mixture was incubated at 37 °C for 75 minutes. Released PPi was detected with PPiLight inorganic pyrophosphate assay kit from Lonza according to manufacturer's instruction. Through several cycles of trial and error, adjustments were made to compounds with suboptimal fitting curve which should be diluted in a 1:3 or 1:2 dilution

series to give more optimal dose-response curve. As MTH1 enzyme activity is correlated to the amount of pyrophosphate released during the reaction catalyzed by MTH1 (Appendix 8), monitoring luminescence produced from PPi in a wide concentration range of inhibitor yields primary data which can be fitted into a dose-response curve using the analytical method non-linear regression supplied in GraphPad Prism software.

2-2-5 Isothermal titration calorimetry

Thermodynamic analysis of test compounds was performed on MicroCal 200 ITC (GE healthcare) under constant pressure at 25°C. For each measurement, MTH1 protein solvated in iTC buffer which is prepared according to the method of Gad et al. (100 mM Tris pH 8.0, 40 mM NaCl, 10 mM MgCl₂, 0.005% Tween-20 and 1 mM Tris(2-carboxyethyl)phosphine hydrochloride) with DMSO proportional to the concentration of ligand solution added was loaded into 280 µl sample cell and no bubbles should be present to avoid signal produced by noise. To prepare ligand solution, ligand dissolved in DMSO (20-100 mM) was further diluted in iTC buffer to final concentration of 600-860 µM. According manufacturer's protocol, the concentration of DMSO should not surpass 5% as thermodynamic signal would be disturbed. For generation of a binding isotherm with thermodynamic parameters including the dissociation constant (K_d), the enthalpy of binding (ΔH), the entropy of binding ($T\Delta S$) and stoichiometry (n), 20-30 injections for a single measurement is required with the solution in the sample cell stirred during the entire measurement process at 1000 rpm. Analysis of data set was carried out using Origin 7 software. For curve fitting, one-site model was applied and calculated related statistics from the fitted binding isotherm (Appendix 10).

2-2-6 Size-exclusion chromatography



Prior to crystallization, His-tagged MTH1 proteins were digested with thrombin to remove His-tag at 4°C. To eliminate cleaved His-tag present in the protein solution, size-exclusion chromatography was performed to separate His-tag and MTH1 protein according to their molecular size. Before separation of target protein from other undesired substances in the solution with size-exclusion chromatography, buffer (20 mM Tris pH 7.4, 150 mM NaCl, 2 mM TCEP, 5% Glycerol) used in the experiment was prepared and filtered using Stericup (Millipore) with pore size of 0.22 μ m. Experimental details regarding settings in FPLC machine is described as below. For size exclusion chromatography, pressure limit is set to 0.3 mpa followed by flow rate= 1ml/min; equilibrium volume= 15 ml, elution volume= 120 ml, fraction size= 2ml. Protein sample was preloaded into syringe with a volume of 5 ml followed by injection into superdex75 gel filtration column (GE healthcare) during equilibrium.

2-2-7 Crystallization and soaking

Prior to crystallization, purity of MTH1 protein with his-tag removed was examined using SDS-PAGE on which no extra visible band should be present. For sample preparation, 50 mM $MgCl_2$ was added to concentrated MTH1 protein (8-10 mg ml⁻¹) to a final concentration of 5 mM. To acquire MTH1 crystals, hanging drop experiment was set up which is composed of drops (2 μ l) containing 1:1 mixture of protein solution and mother liquid (0.2 M $LiSO_4$, 0.1 M sodium acetate pH 3.75, 30% PEG6000) to be diffused against reservoir solution (~1 ml) in the well at 22°C. Needle-like crystals appeared after 1-2 days. For soaking test compounds into MTH1 crystal, diffraction quality crystals were selected and transferred into reservoir supplemented with 5 mM ligand solution. In this case, crystals should be soaked for at least 16-24 hours to ensure that ligands are bound.

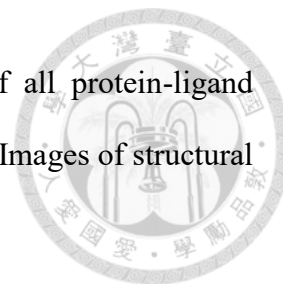
2-2-8 X-ray data collection

Soaked crystals were first sorted according to cracks appeared on the surface to avoid diffraction image with defects which directs to data set of bad quality. As qualifying criterium for ideal diffraction quality crystals, they should appear flawless without significant cracks and thick instead needle-like in their conformation. All crystals were flash-frozen in liquid nitrogen and mounted using cryo-loops with size of 0.1-0.2 mm. Data collection was performed at BL13B1 and 13C1, National Synchrotron Radiation Research Center, Taiwan. Experimental settings for collection of a single data set is described as below: detector distance = 200-250 mm, exposure time= 2-60 s, frame width= 0.3-0.5°, total *oscillation* angle= 80-120°, wavelength=1.000, 0.976 Å for 13B1 and 13C1, respectively. Data sets were integrated and scaled with HKL2000 (Otwinowski et al., 1997).

2-2-9 Model building and Refinement

Processed x-ray data was primarily solved via molecular replacement using *apo* MTH1 structure available in Protein Databank (PDB ID: 3ZR1). To place ligand into reasonable position in the solved structure, coordinate and restraint files for ligand were first generated using eLBOW (Moriarty et al., 2009) which is a convenient tool built in Phenix (Adams et al., 2010) for preparing ligand files applied in subsequent structure determination procedure. Followed by ligand placement, several cycles of refinement were performed on current model to optimize the position of placed atom in corresponding calculated electron density map. In further stage of model building, placement of water molecules in solved structure was performed in Autobuild (Terwilliger et al., 2008). As water molecules were placed, a few rounds of refinement with TLS parameter and NCS restraint should be sufficient to complete the final model

with refined R values. Structure determination and refinement of all protein-ligand complex structures solved in the study were conducted with Phenix. Images of structural models were generated using software Pymol (Schrödinger).



Chapter 3 Results



3-1 MTH1 protein purification

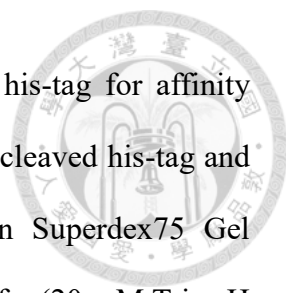
The purification of target protein from cell lysate is intended to obtain research material for subsequent experimental trials with high purity. The complete procedure is composed of three individual parts which involves affinity column purification of his-tagged target protein followed by removing purification tag with thrombin digestion at interval stage before the final purification procedure of size-exclusion chromatography is performed.

3-1-1 Affinity Chromatography

The purified target protein MTH1 with his-tag was eluted from his-tag affinity column during wash volume 550-700 ml (Figure 2A). The purity of his-tagged MTH1 protein was examined in the following SDS-PAGE analysis (Figure 2B). In regard of band shown on SDS-PAGE gel which is indicative of protein degradation, experimental condition including incubation time and temperature during large-scale protein expression was revised and modified.

As a result, three independent experimental trials for large-scale protein expression were performed under condition of incubation temperature 22°C for 14、16 hours and 37°C for 2 hour. From Figure 2B it is demonstrated that protein purified from lysate of *E. coli* cells incubated at 37°C retains the integrality of his-tagged MTH1 protein compared to 22°C which indicates incubation time and temperature might be the influential factors in protein degradation for this case.

3-1-2 Size-exclusion chromatography (SEC)



After purified his-tagged MTH1 protein has been obtained, his-tag for affinity chromatography was removed with thrombin digestion. To separate cleaved his-tag and MTH1 protein, size-exclusion-chromatography was performed on Superdex75 Gel Filtration column (GE healthcare) equilibrated with gel filtration buffer (20 mM Tris pH 7.4, 150 mM NaCl, 2 mM TCEP, 5% Glycerol). As the separation of MTH1 protein from other undesired substances present in sample solution depends on size of molecules, the standardized elution volume corresponding to representative molecular weight is given in the form of referential parameter as shown in Figure 2C. After the final stage of purification, the separated MTH1 protein was examined on SDS-PAGE gel which shifts from the position of molecular weight 20 kDa to approximately 18 kDa compared to his-tagged MTH1 protein (Figure 2D).

Consequently, purified MTH1 protein with his-tag removed (Figure 2E) serves as experimental material in later crystallization trial for x-ray diffraction experiment (Figure 2F).

3-2 Determination of binding affinity profiles and structure-activity relationship (SAR)

3-2-1 Fragment-based drug screening and optimization performed with 2-aminopyrimidine-based scaffold

In searching for scaffold compound with moderate binding affinity toward MTH1, high-throughput screening (HTS) on 2313 structurally irrelevant fragments was performed by Ph.D. candidate Cheng Peng from our lab. Consequently, four compounds were identified to be micromolar binders in the enzyme inhibition assay. From these identified hits, compound 6077639 was selected as structural skeleton for further optimization campaign in regard of its inhibitory activity on prostate cancer cell lines.

Provided that x-ray structure of MTH1 bound with compound 6077639 has been

solved by Ph.D. candidate Cheng Peng, notable interaction pattern which may contribute to favorable binding interaction are characterized. Attributed to observation from x-ray complex structure, systematical exploration of binding pocket using 2-aminopyrimidine scaffold was further performed to elucidate the correlation between interaction pattern of individual functionality and its contribution to binding affinity profile of inhibitor candidates.

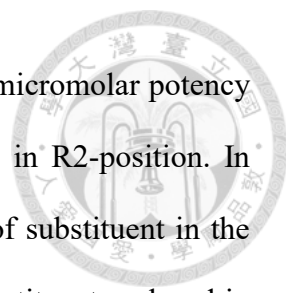
For simplicity of description, the synthesized compound series in this study is categorized into two separate groups according to the presence or absence of ethyl group at C5 of the pyrimidine moiety. The parent compounds of these 2-aminopyrimidine derivative compound series are referred to as scaffold 1 and 2 in the following text.

3-2-2 Binding affinity profile and SAR study of scaffold 1-derived compounds

For compounds optimized using scaffold 1 which has ethyl group incorporated at C5 of the pyrimidine ring (Figure 3A), dramatic loss in binding affinity was observed compared to hit compound 6077639 (Figure 3A).

In the six compounds derived from scaffold 1 with biochemical potency distributes unevenly in sub-millimolar IC_{50} range (Fig 3B-E), variation in binding affinity profile is noteworthy which approximately differentiates by a factor of 200 for the weakest and the most potent binder in this compound series.

In the 2-aminopyrimidine-based compound series optimized using scaffold 1, exploration of structure-activity relationships (SARs) was performed on the substituents in the R1- and R2-position. To complement this objective, compound 1, 3, 6, 13, 14 and 16 was designed and synthesized by our colleague Zhe-Hwa Cheng from graduate institute of pharmacy. In biochemical assay, compound 14 with 1-methylpiperazinyl incorporated in the R2-position exhibits millimolar activity toward MTH1 and is the



weakest binder in this structurally-related compound series. Whilst micromolar potency is determined for compound 16 substituted with piperidinyl group in R2-position. In summary, the observation described above suggests that influence of substituent in the R2-position is profound (Fig 3B-E). Nevertheless, in the case of substituent replaced in the R1-position which are N-methyl and N-cyclopropyl in this compound series, the impact on binding affinity profile is rather insignificant as observed in compounds pairs 1 and 13; 3 and 14; 6 and 16 which differentiate by a factor of 4 to 15 depending on substituent incorporated in the R2-position.

3-2-3 Binding affinity profile of scaffold 2-derived compounds

Followed by biochemical assay for characterization of binding affinity profiles of the 2-aminopyrimidine-based compound series derived from scaffold 1. It was notified that using scaffold 1 for optimization appears to hinder the enhancement of binding affinity toward MTH1 (Figure 3). In regard of this obstacle in the optimization campaign, ethyl group at C5 of the pyrimidine moiety was speculated to be the underlying cause for the dramatic loss in biochemical potency. To confirm the assumption that ethyl group at C5 of the pyrimidine moiety should be the underlying cause for weak activity of scaffold 1-derived compounds toward MTH1, A series of structurally-related compounds based on scaffold 2 (Figure 4A-B) with ethyl group at C5 of the pyrimidine moiety removed was designed and synthesized.

Expectedly, the removal of ethyl group indeed contributed to enhancement of the binding affinity profiles of 2-aminopyrimidine-based compound series derived from scaffold 2 (Figure 4C-F). The overall binding affinity profile of the compound series was determined in the nanomolar IC_{50} range. Nevertheless, despite of improved binding affinity toward MTH1, significant deviation in binding affinity profile within the

compound series is notified which differentiates by three orders of magnitude.

Among the substituents incorporated at R2, 1-(oxetan-3-yl) piperazinyl and 1-methylpiperazinyl group apparently prohibits the elevation of binding affinity of compound 12, 22, 29 and 30 which are replaced with substituents referred above. While compounds incorporated with either piperidinyl or morpholine substituent were found to be nanomolar binders in enzyme inhibition assay (Figure 4D,F).

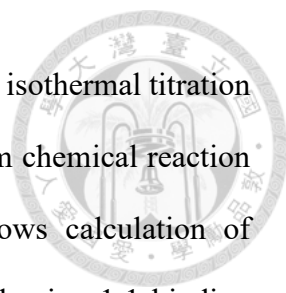
Consequently, the comparison of binding affinity profiles in regard of the influence of substituents at either R1 or R2 reveals that the modulation of binding affinity is more correlated to substituents incorporated in the R2-position rather than R1-position. Of note, this finding is in coherence with the observation made in the 2-aminopyrimidine derivative compound series optimized based on scaffold 1 discussed previously (Fig 3).

3-3 Thermodynamic study on 2-aminopyrimidine-based compounds derived from scaffold 2

For establishment of complementary affinity profiles of optimized compound series, binding affinity of ligand toward MTH in the absence of substrate was measured using isothermal titration calorimetry (Figure 5A). In consideration of the detection limit of isothermal titration calorimetry (ITC) which lies within the range of 10^{-2} - 10^{-9} M according to instruction of manufacturer, the determination of dissociation constant (K_d) including related thermodynamic parameters was performed on 2-aminopyrimidine-based compound series optimized using scaffold 2 (Figure 4A,B) with weak binders derived from scaffold 1 excluded.

3-3-1 Enthalpy-driven binding of 2-aminopyrimidine-based compounds to MTH1

To determine the thermodynamic profile of the present compound series modified

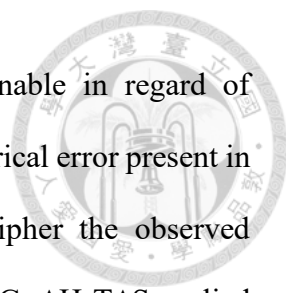


using scaffold 2 with profound improvement in biochemical potency, isothermal titration calorimetry (ITC) was applied to record signals of heat released from chemical reaction involving breakage and formation of chemical bonds which allows calculation of thermodynamic parameters ΔH and $-T\Delta S$ from titration curve fitted using 1:1 binding model (Appendix 10) . Consequently, an array of thermodynamic profiles for the 2-aminopyrimidine-based compound series binding to MTH1 was established (Figure 5B-I). Meanwhile, it is notable that binding affinity profile measured using iTC is in consistency with previously determined biochemical potency in enzyme inhibition assay (Figure 5J). In spite of similar structural skeleton, deviated binding mechanism is revealed in compound 12, 22, 29 and 30 (Figure 5K). While for compound 20, 24, 25 and 26 which are the most potent binding partners for MTH1 in this study, deviation in binding mechanism is absent (Figure 5L).

The plot of thermodynamic signatures demonstrates that the binding interaction between target molecule MTH1 and 2-aminopyrimidine based compound series is dominated by favorable enthalpic component of Gibbs free energy (Figure 6A). From another aspect, despite of significant enthalpic signal measured directly in biophysical analysis using iTC, the enthalpic contribution is always counteracted by a minor entropic component which eventually leads to similar binding affinity profile observed across the investigated compound series.

3-3-2 Phenomenon of Enthalpy-Entropy Compensation (EEC)

Within the 2-aminopyrimidine-based compound series which bind to MTH1 in an enthalpy-driven fashion, profound enthalpy-entropy compensation (EEC) was unraveled (Figure 6A,B) when enthalpic component (ΔH) was plotted against entropic component ($-T\Delta S$) with correlation coefficient value $R^2=0.94$ (Figure 6B). Whereas, the correlation



which is indicative of enthalpy-entropy compensation is questionable in regard of numerous published research articles debating over the issue of empirical error present in the calculated thermodynamic parameters ΔH and $-T\Delta S$. To decipher the observed phenomenon which is often masked by errors rooted in the equation $\Delta G = \Delta H - T\Delta S$ applied in calculation of the entropic component ΔS which could not be measured experimentally, it was suggested that thermodynamic data should be handled with carefulness in respect to susceptible error in the experimental workflow described above.

3-3-3 Rationalization of observed EEC phenomenon through insights provided by atomic model of MTH1 in complex with 2-aminopyrimidine-based compounds

Assuming that behind the phenomenon of EEC there is certain physical mechanism modulating such intrinsic interplay between binding partners and solvent molecules in surrounding environment as proposed in numerous research works, the counteracting effect between thermodynamic parameters enthalpy and entropy (Figure 6C) may somehow be investigated with insights from atomic structure of protein-ligand complex which uncovers interaction pattern of binding partners.

Thus, to clarify the complex formation process described using thermodynamic parameters ΔH , $-T\Delta S$ and Gibbs free energy ΔG which is the summation of enthalpy (ΔH) and entropy ($-T\Delta S$), schematic illustration is presented in Figure 6D-F. Elaborately, enthalpic component of binding free energy involves formation of H-bonds and van der Waals contacts while entropic component is involved in conformational change and hydrophobic interaction which can hardly be measured using present experimental techniques.

3-3-3-1 Thermodynamic profiles of structurally-related compounds with substituent

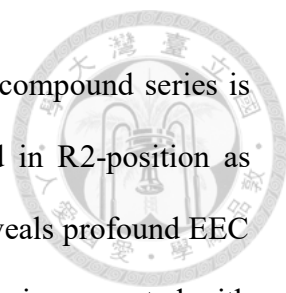
explored on R2-position

As revealed in bar diagram from Figure 6A which demonstrates typical phenomenon of EEC observed frequently in compound series stem from optimization campaign. The relation between structural feature of structurally-related compounds and the effect of EEC is further investigated as the present compound series is systematically explored on R1 and R2-position with substituents.

Firstly, thermodynamic profiles of scaffold 2-derived compounds binding to MTH1 was categorized according to R1-substituent, which is underrepresented by compound 12, 20, 26, 30 and compound 22, 24, 25, 29 depicted as bar diagram displayed in Figure 7A-B. Subsequently, the thermodynamic parameters ΔH and $-T\Delta S$ was plotted in diagram with correlation coefficient R^2 indicated (Figure 7B-C). From the results of linear regression analysis on thermodynamic parameters of scaffold 2-derived compounds categorized based on substituent incorporated in R1-position, the R^2 value for compounds with N-methyl group incorporated in R1-position is 0.99 while the presence of additional cyclopropane ring lead to decreased correlation coefficient value of $R^2 = 0.86$. For rationalization of this paradox, visualization of atomic details of protein-ligand complex is performed with x-ray crystallography. In Figure 7D, structural alignment between compound 20 (colored magenta) and 30 (colored yellow) unraveled deviated positioning of R2-substituent in MTH1 binding pocket. Meanwhile, this tendency was also observed in structural alignment of compound 25 (colored blue) and 29 (colored violet) in Figure 7E.

3-3-3-2 Thermodynamic profiles of structurally-related compounds with substituent explored on R1-position and structural relevance of EEC phenomenon

Opposed to arrangement of the scaffold-2 derived compounds in Figure 7 according



to substituent incorporated in R1-position as described above. The compound series is categorized into four groups according to substituent incorporated in R2-position as illustrated in Figure 8. Rearrangement of thermodynamic profiles reveals profound EEC upon replacement of R1-substituent in scaffold-2 derived compounds incorporated with four variants of substituents in R2-position. As plausible explanation for this intrinsic phenomenon, structural alignment of compounds from Figure 8 A-D provides insights into binding mode adapted by these compounds in MTH1 binding pocket which suggests more extensive hydrophobic interaction between R1-substituent and phenyl cluster made up of Phe27, Phe72, Phe74, Phe139 upon replacement of N-methyl group with N-cyclopropyl group (Figure 8E-H).

3-3-3-3 The influence of size of substituent in the aspect of Enthalpy-Entropy Compensation

For scaffold-2 derived compounds investigated in thermodynamic study using iTC. Comprehensive overview of thermodynamic profiles of representative compounds in combination with structural insights is presented in Figure 9 which is compiled of three individual parts. In this compact figure, thermodynamic profiles illustrated in bar diagram for compound 24, 29, 12, 30 (Fig 9E-H) reveal deviated pattern of thermodynamic signatures within the structurally-related compound series. In attempt to uncover the correlation between structural features of compound and deviated thermodynamic profiles, insight into the structural architecture of MTH1 binding pocket (shown in surface representation) bound with compound 24, 29, 12, 30 (shown in stick representation) is presented in Figure 9 A-D. To formulate conclusion for the compensation effect based on information from structural details of protein-ligand complex and related thermodynamic profiles, it should be first notified that as the size of

R2-substituent increases, less enthalpically favorable and more entropy-driven is the binding interaction which is observed in compound 29 and 30 in comparison to compound 24 and 12 (Figure 9I-L). Therefore, as already demonstrated in Figure 8, incorporation of N-cyclopropyl group in R1-position leads to more entropically favorable binding interaction.

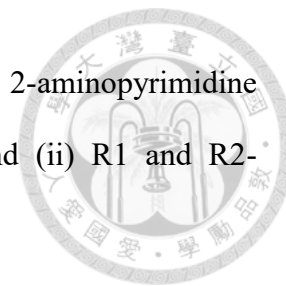
3-4 Structural analysis of MTH1 complex structure bound with 2-aminopyrimidine - based compounds

Structural analysis of binding interaction was performed on eleven MTH1-inhibitor complex structures obtained via soaking preformed MTH1 crystals in ligand dissolved in DMSO and further diluted in reservoir solution (1:10). Similar binding pattern in above-mentioned protein-ligand structures is unraveled. For elucidation of underlying mechanism of the varying binding affinity profiles revealed in enzyme inhibition assay for determination of inhibitory activity of 2-aminopyrimidine-based compound series comprising fourteen structurally-related compounds, thorough inspection of interaction pattern in an array of structurally-related ligands binding to target molecule MTH1 protein provides plausible explanations for the structure-activity relationship (SAR) uncovered in biochemical assay (Figure 10).

3-4-1 Conserve binding mode of pyrimidine-based compounds in the binding site of MTH1

Upon inspection of protein-ligand complexes solved by x-ray crystallography, it was found that conserve binding mode is adopted by 2-aminopyrimidine-based compound series bound in the binding pocket of MTH1 (Figure 10B). To illustrate the details of general binding geometry and orientation of the 2-aminopyrimidine-based

compounds, common features of structural architecture of the 2-aminopyrimidine compound series are decomposed into (i) pyrimidine moiety and (ii) R1 and R2-substituent for ease of description.



3-4-2 Pyrimidine moiety

From previous results of high-throughput screening (HTS) performed on fragment library composed of 2313 compounds of divergent structural architecture purchased from CHEMBRIDGE, one bioactive hit compound was identified which is referred to as compound 6077639 in respect of ID code assigned by manufacturer. Followed by the discovery of this HTS hit compound, cocrystal structure of MTH1 bound with compound 6077639 determined to a resolution of 1.8 Å was resolved by Ph.D. candidate Cheng Peng which sheds light on the intrinsic pattern of protein-ligand interaction (Figure 10A).

At first glance, the binding mode is well-characterized which is mainly attributed to the pyrimidine moiety of ligand molecule which is held firmly in the bottom of binding pocket by the indole ring of Trp117 via π - π stacking interaction. H-bonds formed between recognition element Asp119-Asp120 binding motif and amino group of hit compound 6077639 also contribute to the stabilization forces in the protein-ligand complex (Figure 10A).

In the present optimized compound series, the pyrimidine moiety serves as anchor group which strengthens and rigidifies the orientation of ligand molecule which is related to restriction on bond angle of H-bonds. Simultaneously, systematical exploration with substituents on the pyrimidine moiety allows tracing of structural changes upon binding of the pyrimidine derivative compounds in the binding pocket of MTH1 in a rational fashion.

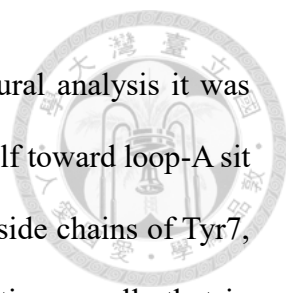
3-4-3 R1 and R2-substituent

Chosen for the subsequent optimization campaign from HTS, compound 6077639 serves as structural skeleton for the present 2-aminopyrimidine-based compound series (Figure 10A). For modification of hit compound 6077639, C4 and C6 of the pyrimidine moiety are assigned as R1 and R2-substituent in the following passages and explored with several functionalities. In the case of R1-substituent, which occupies the cavity formed by phenyl rings of Phe27, Phe72, Phe74 and Phe139, interacts with these side chains through hydrophobic contacts. On the other hand, from the pyrimidine moiety anchored by indole ring of Trp117, substituent at R2-position extends the skeleton of small molecule compound toward the entrance of binding pocket and forms either hydrophobic interaction or van der Waals contact with several rigid components of binding pocket made up of side chains of Tyr7, Thr8, Leu9, Lys23 and Asn33. (Figure 11C).

After thoroughly inspecting the structural details revealed in binding pocket of MTH1 bound with different ligand molecules obtained via x-ray crystallography. The high similarity of binding mode observed in protein-ligand complexes determined in this study further hampers the formulation of rationales for the largely deviated binding affinity profiles of the 2-aminopyrimidine-based compound series. Nevertheless, there remains some notable features to be explored and may provide plausible explanation for the varying inhibitory activity of the present 2-aminopyrimidine-based compounds against MTH1.

3-5 Superposition of MTH1 complexed with product 8-oxo dGMP and compound 29, 30

In the binding pocket of MTH1, compound 29 and 30 overlay well with MTH1 product 8-oxo dGMP which serves as natural ligand exploited for drug design (Figure 14A-B). To dissect the interaction pattern of above-mentioned ligands, contact between



ligand moieties and nearby residues was analyzed. From the structural analysis it was noted that oxetane ring incorporated at N4 of compound 30 lends itself toward loop-A sit in S2' pocket (Figure 14C) and consequently lesser interaction with side chains of Tyr7, Thr8, Leu9, Met81 and Val83 is formed. In regard of this observation, recalls that in MTH1 x-ray structure complexed with product 8-oxo dGMP the α -phosphate is overlaid with oxetane ring from both compound 29 and 30 which otherwise adopted an alternative binding pattern in catalytic region with loop-A involved. With information regarding binding pattern of triphosphate group in MTH1 binding pocket which is bound in coordination with water molecule network intended for hydrolysis of substrate 8-oxo dGTP. Modification of functionality in this position should take this point into consideration. Further, based on this observation, exploration of region occupied by deoxyribose and α -phosphate with substituents requires more attention to be paid in respect to the hydrophobic side chains and several structural water molecules which are likely the key factors in modulation of optimal ligand binding.

3-5-1 Absence of affinity gain with extended substituent toward the entry of binding pocket

The initiative of growing structural skeleton selected from high-throughput screening (HTS) toward entry of binding pocket is complemented by design and synthesis of compound 29 and 30. However, enhancement in binding affinity is not observed in these enlarged compounds which stay in low micromolar IC₅₀ range. Since x-ray structures of MTH1-29 and 30 complexes are available, analysis of atomic details in protein-ligand binding interaction is allowed to dissect the underlying cause for absence of improvement in potency (Figure 15A). To understand the structural arrangement of ligand in MTH1 binding pocket in pragmatical way, structural alignment of substrate 8-

oxo dGTP and compound 29, 30 is illustrated in Figure 15 B-D with 8-oxo group of substrate and residue Phe27, Met81 highlighted. In addition, distance between side chains of residue Phe27, Met81 and atom on ligand is also compared in substrate and compound 29, 30. 8-oxo group of substrate is placed in close distance to side chains of Phe27 and Met81 which is regarded as crucial structural feature in selectivity and substrate recognition.

Finally, a schematic view (Figure 15F) was illustrated for MTH1 binding pocket characterized with functionalities from substrate 8-oxo dGTP. Notable is that the oxetane ring of compound 29 and 30 is overlapped with α -phosphate lied in catalytic site of MTH1 (Figure 15E) which is proposed to be coordinated with a well-organized water molecule network via metal ion during hydrolysis as mentioned previously.

3-6 Notable difference in interaction pattern of heterocyclic substituent unraveled by deviation in position occupied by ligand

Through analysis of hydrophobic contacts between heterocyclic substituent and nearby side chains, similarity in binding interaction in 2-aminopyrimidine-based compound series and host molecule is unraveled (Figure 13, 14, 16).

To decipher the paradox in deviated binding affinity profiles regarding the similarity of binding mode adopted by 2-aminopyrimidine-based compound series, conformational change accompanied by ligand binding was observed and identified in side chains surrounding R2-substituent. Explicitly, residues Phe27 and Met81 which has been reported to involve in substrate recognition and selectivity appear to undergo structural rearrangement upon binding of ligand. Further, in contrast to the rigidity of most interacting side chains, superposition of compounds with different heterocyclic substituent incorporated in R2-position reveals that R2-substituent in compound 22 and

29 was placed in position located closer to loop-A (L-A) featuring key residues Lys23, Phe27, Asn33 and Gly34 in the binding pocket of MTH1 (Figure 11C). Moreover, interaction pattern of heterocyclic substituent in R2-position from x-ray structure of MTH1 bound with compounds derived from scaffold 2 are elaborated in more detail as followed.

3-6-1 Influence of compounds incorporated with different substituents in R2-position on the configuration of key residues Phe27 and Met81

As implied in Figure 12, selectivity of MTH1 for its substrate is proposed to correlate with favorable van der Waals contacts between residues Phe27, Met81 and 8-oxo group of substrate 8-oxo dGTP. 2-aminopyrimidine based compound series including compound 12, 20, 22, 24, 25, 26, 29 and 30 which deviate in their binding affinity profiles (Figure 13-16). It is speculated that mobile residues Phe27 and Met81 which experience structural rearrangement might contribute to selectivity and affinity toward oncoming ligand. Through alignment of solved complex structures, fluctuating structural features were formulated as below. Enlarged R2-substituents which are incorporated into compound 12, 22, 29 and 30 appear to induce conformational change of Phe27 (Figure 13,14, 16D-E), while this feature is not observed in compound 20, 24, 25 and 26 which may attribute to smaller size of R2-substituent (Figure 15D-E). On the other side, Met81 is a flexible residue which seems to play a role in accommodating oncoming ligands, while from present information it is hard to find any correlation between binding affinity and structural configuration of Met81 and adjacent ligand moieties (Figure 15D-E).

Chapter 4 Discussion



4-1 Functional mechanism of MTH1 protein and structural relevance to selectivity toward its substrate

Attribute to significant role in eliminating oxidatively damaged nucleoside triphosphates played by MTH1 protein, intricate structural architecture of the nucleotide pool sanitizing enzyme in relation to its selectivity toward substrates 8-oxo dGTP and 2-oxo dATP is regarded as one of the most fundamental issues in related research works devoted to investigation of functional mechanism of MTH1 protein (Sakai et al., 2002).

To present a comprehensive overview, observations derived from published MTH1 structure in complex with substrates are elaborated as follows. Predominantly, the Asp-Asp binding motif situated at the bottom of MTH1 binding pocket is a subject of research interest and acts as anchoring group for base moiety of substrates through formation of hydrogen bonding interaction (Nakamura, T. et al., 2010). Therefore, as described in research article by Svensson et al., residues Asp119 and Asp120 of which the Asp-Asp binding motif is constituted serves as key element in recognition and binding of substrates 8-oxo dGTP and 2-oxo dATP. With more insights into the pattern of hydrogen bonding interaction between Asp-Asp binding motif and base moiety of substrates, difference in interaction pattern is unraveled (Waz et al., 2017) which propagated further investigation on recognition mechanism of these structurally similar substrate molecules in consideration of change in protonation state of residues Asp119 and Asp120 which position in close neighborhood. Nonetheless, another perspective in regard of the issue of protonation state in Asp-Asp binding motif has been earlier illustrated by Nissink et al. with observation made in comparison of hydrogen bonding patterns adopted by MTH1 product 8-oxo dGMP and several MTH1 inhibitors. In respect to implications from these

observations described by the authors, it is otherwise pointed out that the explicit recognition mechanism for MTH1 may lie elsewhere in MTH1 binding pocket.

In more pragmatic aspect, assuming that the structural architecture of MTH1 binding pocket is adaptively manufactured for binding of oxidatively damaged nucleoside triphosphates followed by hydrolysis of phosphate group, there has to be specialized recognition mechanism for differentiating substrates from numerous potential binding partners present in environment under usual physiological circumstances. Therefore, binding mode of substrate 8-oxo dGTP was inspected and analyzed in previous studies by separate research groups (Nissink et al., 2016) (Waz et al., 2017) for elucidation of mechanism underlying MTH1 substrate recognition. Explicitly, both research groups hold the same viewpoint on the significance of side chains on residues Phe27 and Met81 which come into close contact with 8-oxo group of substrate 8-oxo dGTP and this interaction mode is further postulated to be determining in recognition of substrate by MTH1.

4-2 The binding energetics underlying complex formation

Universally, formation of complex is driven by dynamic procedure in which binding partners and surrounding solvents undergo rearrangement of interaction patterns and steric properties of receptor and ligand are adaptively rearranged for shape complementarity of formed complex. To shed light on the elaborately coordinated process of complex formation which is difficult to observe with present experimental techniques, measurement of Gibbs free energy (ΔG) which is further partitioned into individual components of enthalpy (ΔH) and entropy ($-T\Delta S$) by means of either computational calculation or thermodynamic experiments provides insights into binding energetics of complex formation. Nonetheless, in the measurement, parameterization of contributions from separate binding interactions such as formation of hydrogen bonds, van der Waals contacts, hydrophobic interactions and conformational change is precluded which implies

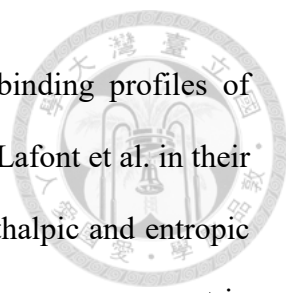
that statistics derived from the measurement of binding energetics is a rather elusive without addressing the details of binding event.

Thus, perceiving the fact that interaction and complex formation of binding partners retains certain degree of complexity, there remains elusive points in respect to enthalpic and entropic contributions in binding event to be clarified.

4-2-1 Intrinsic phenomenon of EEC hinders the elevation of binding affinity

To name the most prominent while unclarified effect in binding event characterized by measured thermodynamic profiles (Klebe G, 2015), the widely reported phenomenon of Enthalpy-Entropy Compensation (EEC) in hit-to-lead optimization archive has been uncovered and substantially frustrated all attempts to improve binding affinity of designed ligands (King NM et al., 2012). Notably, from thermodynamic analysis of 2-aminopyrimidine based compounds binding to MTH1 protein in this study, profound EEC phenomenon was unraveled which might consequently hinder any progress in the improvement of binding affinity profile.

As gleaned from published results of thermodynamic analysis conducted on compound series from optimization campaign (Lafont V, 2007), binding of an array of structurally-related compounds to target molecule appears to be frequently confronted with counteracting effects posed by EEC phenomenon (Klebe G., 2015). Though, recalling that the objective of hit-to-lead optimization archive is manifested by enhancement of binding affinity toward target molecule, overriding the obstacle of compensating energetic contributions from enthalpy and entropy is of urgent necessity as the establishment of binding profile featured by combination of favorable enthalpic and entropic contribution is more preferred in terms of elevation of binding affinity (Velazquez et al., 2001).



Pragmatically, to abrogate the paradoxical phenomenon in binding profiles of compound series from optimization campaign, it is demonstrated by Lafont et al. in their research article that such inevitable effect which is in relation to enthalpic and entropic penalties resulted from introduced functionalities is directly linked to rearrangement in structural architecture of binding partners and adaptation of deviated interaction patterns. Therefore, formulation of underlying causes for absence of net gain in binding free energy by means of structure-based analysis reveals information which is more straightforward under the limitation of observing the dynamic binding event with present experimental techniques.

4-2-2 Proposed theories for explanation of contradictory effect of enthalpic and entropic component of Gibbs binding free energy

Prior to formulation of proposed mechanisms for EEC phenomenon, fundamental concepts regarding complex formation is illustrated as follows.

Firstly, to elicit binding event under a wide spectrum of circumstances in which influential factors such as salt concentrations (Bissantz et al., 2010), pH in solvent environment (Onufriev et al., 2013) extensively exert its modulating effects on interaction pattern of involved binding partners, physical driving forces of distinct origins which occasionally appears to be mutually counteracted act as critical player in process during complex formation and binding interaction between involved partner molecules in adaptation to either intentionally adjusted buffer condition or physiological environment in living organisms.

Nevertheless, aside from compiled interaction network with binding partners and solvent molecules involved dynamically which is propelled by physical driving forces stemming from enthalpic and entropic component of binding free energy, several

mechanisms for intrinsic phenomenon of EEC are proposed in physical terms despite of the complexity of the biophysical system.

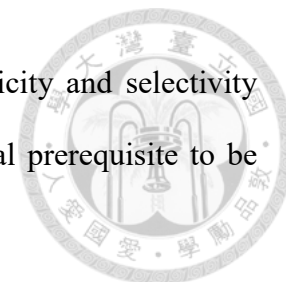
In regard of the popularity of this issue which is fundamental in decision making for optimization of binding affinity profile, practical aspects provided by numerous research groups (Chodera et al., 2013) have inspired distinct approaches for handling and interpreting observed EEC phenomenon during investigation of binding interactions. Among various proposed mechanisms extensively reviewed by Chodera et al. in their review article, several proposals which are most advocated were highlighted by the authors. Of these foremost theories, the key points are mainly focused on solvent reorganization upon binding (Grunwald E et al., 1995), restriction of conformation in bound state (Dunitz JD, 2003) and liberation of solvent molecules through extreme tight-fitting of ligand molecules (Qian H, 1998). In pragmatical aspect, these proposals appear to complement some observations made earlier, while the topic remains under debate with opposing opinions possessed by researchers. It is of great importance that interpretation of the intrinsic phenomenon of EEC observed in this study should be independent from other purposed perspectives to retain the originality of characteristics exhibited by unique system composed of protein and ligand molecules investigated in this study.

4-3 Structure-aided ligand design: perspectives and opportunities

In perspective of structure-aided ligand design promoted by numerous research groups participated in fast-paced industrial development of pharmaceuticals intended to meet the needs for revolutionary therapies against diseases which are held incurable, investigation of biological molecules engaged in physiological processes by means of techniques dedicated to reveal atomic structure of target molecule of medicinal interest is preferential and make intuitive sense.

Practically, exploiting these intricate structural features in development of potential

drug candidates with optimal binding profiles of improved specificity and selectivity toward target molecule MTH1 is regarded as the most fundamental prerequisite to be fulfilled during the stage of hit-to-lead optimization campaign.



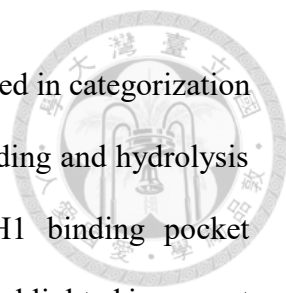
Recalling that the integrity of formed protein-ligand complex is considered to be complementarity between two binding partners in terms of steric and electrostatic properties (Geschwindner et al., 2015), it is virtually compilation of investigations with structural analysis and modification of hit compounds to give rise to a lead compound with optimal binding profile. Further, it is demonstrated in exploration of MTH1 binding pocket that a variety of structural architecture can be accommodated which display several distinct binding configurations. Explicitly, integral structural features in MTH1 binding pocket explored with attempts to gain more binding affinity for optimized compounds are elaborated subsequently in the following passage.

4-3-1 Binding pattern observed in published MTH1-ligand complexes

As scrutinization of mechanism underlying recognition and selectivity of target molecule is comprised of investigation conducted on molecular structure of target molecule in its bound and unbound state, it is widely perceived that success in ligand design could be achieved provided that 3D-structure of the target molecule is resolved to atomic resolution which enables visualization of structural architecture and explicit features to be exploited through growing hit compound into larger lead molecule with favorable structural features incorporated.

Therefore, with 3D-structure of target molecule in hand, elucidation of the underlying mechanism of binding interactions gleaned from crystal structures of numerous MTH1-ligand complexes is propagated.

Firstly, for simplifying illustration of binding mode adopted by ligand molecules,



MTH1 structure complexed with substrate 8-oxo dGTP is applied in categorization and characterization of the specifically functionalized region for binding and hydrolysis of oxidized nucleoside triphosphates. Three subregions in MTH1 binding pocket occupied by base, ribose and triphosphate group of 8-oxo dGTP are highlighted in respect to their unique structural features adapted for binding and interacting with substrate molecules.

Initiated from compiled research by Gad et al. devoted to validation of MTH1 protein as promising anticancer target, extensive investigation spanning multiple disciplines has provided insights from various aspects in the functional mechanism and physiological role of MTH1 which gave rise to series of compounds with unique features. As revealed in previous investigation utilizing high-throughput screening to identify small molecule binders for target molecule MTH1 biochemically, it is well-characterized that fragments with diverse structural skeleton were found to be accommodated by MTH1 at the bottom of substrate binding pocket. Among these fragment compounds, it is noted that universal binding mode is adopted which is attributed to anchoring group constituted of key residues Trp117, Asp119 and Asp120. Therefore, as base moiety of substrate is bound firmly through formation of hydrogen bonding interaction with Asp119-Asp120 binding motif and stabilized by indole ring of Trp117 via π - π stacking, binding pattern of small molecule compounds is substantially determined by the configuration of ligand moiety and side chains of residues Trp117, Asp119 and Asp120. Besides, in addition to the anchoring group where hit compounds are generally harbored, a minor hydrophobic pocket adjacent to anchoring group denoted as pocket 1 was explored by Rahm et al. According to the authors, filling pocket 1 which is surrounded by hydrophobic residues Leu11, Leu20, Met116, and Phe124 indeed elevated the biochemical potency of synthesized compounds. Distinctively, 2-aminopyrimidine based compounds investigated

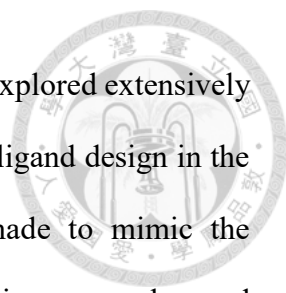
in this study which target the hydrophobic pocket formed by phenyl rings of Phe72, Phe74 and Phe139 on the opposite side appear to be fitted while significant improvement is absent.

Provided that region in MTH1 binding pocket for accommodating base moiety of substrate is well-characterized which is mainly composed of three binding elements explored earlier including anchoring group of Trp117, Asp119 and Asp120 in combination with two hydrophobic cavities on each side, establishment of optimal binding interaction in ribose binding region is of great concern regarding selectivity of designed ligands. To complement this objective, proposals have been presented by Kettel et al. and Rahm et al. in their published research articles previously. By means of straightforward approaches with aims to mimic the binding pattern of deoxyribose group of substrate, hydrogen bonding contacts to the backbone of T8 and Gly34 were probed with intention to gain more selectivity toward MTH1.

4-3-2 Overview of available space left unexplored in MTH1 binding pocket in regard of information derived from explored regions

In regard of information derived from published results of compiled researches thriving to elucidate the interaction patterns of MTH1-ligand complex structures through exploration of MTH1 binding pocket with series of structurally-related compounds adaptively modified from HTS hit compounds, formulation of integral structural features of target molecule MTH1 for optimal binding of ligand molecules featuring diverse structural architecture is enlightened. Therefore, with exception of the well-characterized binding regions in MTH1 binding pocket there remains unexplored regions which reserve opportunities for further optimization archive.

As briefly elaborated in previous section, MTH1 binding pocket is partialized according to regions occupied by base, ribose and triphosphate group of substrate. In



respect to regions harboring base and ribose group which have been explored extensively earlier, phosphate binding region is less discussed and exploited for ligand design in the frame of MTH1 inhibitor development. Though, attempt was made to mimic the interaction formed between α -phosphate group and K23 terminal amino group observed in product 8-oxo dGMP by Kettel et al., barely would catalytic region constituted of Glu cluster formed by Glu52, Glu55, Glu56 and Glu100 be mentioned in ligand design for MTH1 in previous studies. This might be attributed to the intricate binding pattern observed in the catalytic region which is coordinated by metal ions with water molecule network to hold the substrate in place and catalyze hydrolytic reaction of triphosphate group.

Nonetheless, the critical part in design of potent and highly selective ligand is considered to lie in elucidation of the recognition mechanism of MTH1 for oxidatively damaged nucleoside triphosphates. As stated by Nissink et al. and Waz et al. that 8-oxo group of 8-oxo dGTP interacts with side chains of Phe27 and Met81 favorably in comparison with its counterpart dGTP without the presence of 8-oxo group.

In support of this notion, profiling of synthesized compounds against extended panel of kinases was performed by Kettle et al. and Rahm et al. in their research works to reduce the probability of off-target effect and secure that the selectivity is retained during the optimization archive at early stage. Therefore, it is suggested that substructure commonly seen in kinase inhibitors is readily accommodated in MTH1 binding pocket (Nissink et al., 2016) (Rahm et al., 2018) which implies improvement of selectivity of ligand is substantially confronted with the issue posed by kinases and its inhibitors.

Chapter 5 Conclusion

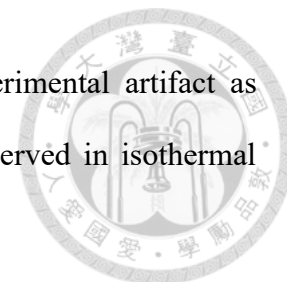


The biological machinery- protein is intrinsically fine-tuned by molecular recognition driven from physical forces. To date, it is well-known that such biological macromolecules participate in signaling pathway synchronizing fundamental physiologic function of all living organisms on earth, capability of predicting such behavior of specifically functionalized entities irreplaceable for cooperative biological function network has become one of the most integral part in developing strategies for handling diseases with identified molecular targets.

In this study, a series of 2-aminopyrimidine-based compounds is designed and synthesized for exploring and probing available space in binding pocket of MTH1 protein. Firstly, binding affinity profile of the compound series is determined in enzyme inhibition assay which unraveled tremendous deviation by a factor of 10^6 in biochemical potency of 2-aminopyrimidine-based compounds. Propagated by this experimental observation, protein-ligand complex structures determined to atomic resolution is further obtained via x-ray crystallography. Consequently, counterintuitive results were found in crystal structure of MTH1 in complex of an array of 2-aminopyrimidine-based compounds. Thus, it was confirmed that similar binding mode is adopted by protein-ligand complex structures determined in this study.

From above-mentioned experimental results, it appears to be difficult to rationalize the deviated binding affinity profiles determined for 2-aminopyrimidine-based compound series. To gain more insights into how two binding partners may interact, isothermal titration calorimetry was performed to characterize thermodynamic profiles of scaffold 2-derived compounds. Intriguingly, profound EEC (Enthalpy-Entropy compensation) phenomenon was notified which may be interpreted in terms of previously proposed

physical theory. Nevertheless, it could be somehow just an experimental artifact as described in research articles with focus on the issue of EEC observed in isothermal titration calorimetry.



Drawing a conclusion on current experimental results would be compromising partly attributed to the complexity of protein-ligand interaction and surrounding solvent environment. Nevertheless, in experimental analysis performed on protein-ligand complex structure and related binding profiles several notable points were found which is formulated as followed, 1. The impact of replacement of substituent in R2-position is far more significant than R1-position of scaffold, 2. Compound series derived from scaffold 1 is less active in enzyme inhibition assay against MTH1 which is attributed to incorporation of ethyl group at C5 of the pyrimidine moiety, 3. Overlap of R2-substituent with α -phosphate of substrate in the catalytic site for hydrolysis of triphosphate group appears to be related to loss of affinity in compound 12, 22, 29 and 30.

In order to complement the initial objective of optimizing biochemical potency of hit compound from HTS, substituents were utilized to probe the adaptability of MTH1 toward an array of structurally-related compounds. For optimal filling of minor pocket adjacent to R1-substituent, extensive hydrophobic interaction with phenyl cluster in the pocket is established with incorporation of N-cyclopropyl group. While for region occupied by R2-substituent, divergence in binding interaction is accompanied by increase in the size of substituent and this trend is also observed in thermodynamic study.

Opportunities for further optimization campaign is provided based on the proposition that integral element in binding pocket is thoroughly characterized in terms of molecular dynamic properties which involves solvation and desolvation during complex formation. To this point, more compiled research works are left to be conceptualized and accomplished.

References



Baum, B., Muley, L., Smolinski, M., Heine, A., Hangauer, D., and Klebe, G. (2010).

Non-additivity of functional group contributions in protein-ligand binding: a comprehensive study by crystallography and isothermal titration calorimetry. *J Mol Biol* 397(4): 1042-1054.

Berman, H., Henrick, K. and Nakamura, H. (2003). Announcing the worldwide Protein Data Bank. *Nat Struct Biol* 10(12): 980.

Bessman, M. J., Frick, D. N. and O'Handley, S. F. (1996). The MutT Proteins or "Nudix" Hydrolases, a Family of Versatile, Widely Distributed, "housecleaning" Enzymes. *J Biol Chem* 271(41): 25059-25062.

Bissantz, C., Kuhn, B. and Stahl, M. (2010). A medicinal chemist's guide to molecular interactions. *J Med Chem* 53(14): 5061-5084.

Brandt, T., Holzmann, N., Muley, L., Khayat, M., Wegscheid-Gerlach, C., Baum, B., Heine, A., Hangauer, D. and Klebe, G. (2011). Congeneric but still distinct: how closely related trypsin ligands exhibit different thermodynamic and structural properties. *J Mol Biol* 405(5): 1170-1187.

Chodera, J. D. and Mobley, D. L. (2013). Entropy-enthalpy compensation: role and ramifications in biomolecular ligand recognition and design. *Annu Rev Biophys* 42: 121-142.

Dullweber, F. D., Stubbs, M. T., Musil, D., Sturzebecher, J. and Klebe, G. (2001). Factorising ligand affinity: a combined thermodynamic and crystallographic study of trypsin and thrombin inhibition. *J Mol Biol* 313(3): 593-614.

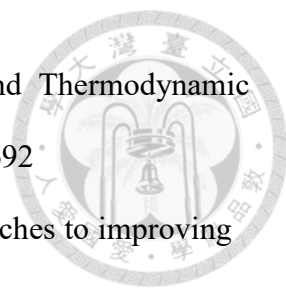
Dunitz, J.D. (2003). Win some, lose some: enthalpy-entropy compensation in weak intermolecular interactions. *Chem Biol* 2(11): 709-712.

Emsley, P., Lohkamp, B., Scott, W. G. and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4): 486-501.

Rahm, F., Viklund, J., Tresaugues, L., Ellermann, M., Giese, A., Ericsson, U., Forsblom, R., Ginman, T., Gunther, J., Hallberg, K., Lindstrom, J., Persson, L.B., Silvander, C., Talagas, A., Diaz-Saez, L., Fedorov, O., Huber, K.V.M., Panagakou, I., Siejka, P., Gorjanacz, M., Bauser, M. and Andersson, M. (2018). Creation of a Novel Class of Potent and Selective MutT Homologue 1 (MTH1) Inhibitors Using Fragment-Based Screening and Structure Based Drug Design. *J. Med. Chem.* 2018, 61, 2533-2551.

Gad, H., Koolmeister, T., Jemth, A. S., Eshtad, S., Jacques, S. A., Ström, C. E., Svensson, L. M., Schultz, N., Lundbäck, T., Einarsdottir, B. O., Saleh, A., Göktürk, C., Baranczewski, P., Svensson, R., Berntsson, R. P. A., Gustafsson, R., Strömberg, K., Sanjiv, K., JacquesCordonnier, M. C., Desroses, M., Gustavsson, A. L., Olofsson, R., Johansson, F., Homan, E. J., Loseva, O., Bräutigam, L., Johansson, L., Höglund, A., Hagenkort, A., Pham, T., Altun, M., Gaugaz, F. Z., Vikingsson, S., Evers, B., Henriksson, M., Vallin, K. S. A., Wallner, O. A., Hammarström, L. G. J., Wiita, E.; Almlöf, I., Kalderén, C., Axelsson, H., Djureinovic, T., Puigvert, J. C., Häggblad, M., Jeppsson, F., Martens, U., Lundin, C., Lundgren, B., Granelli, I., Jensen, A. J., Artursson, P., Nilsson, J. A., Stenmark, P., Scobie, M., Berglund, U. W. and Helleday, T. (2014). MTH1 inhibition eradicates cancer by preventing sanitation of the dNTP pool. *Nature* 508(7495): 215-221.

Geschwindner, S., Ulander, J. and Johansson, P. (2015). Ligand Binding Thermodynamics in Drug Discovery: Still a Hot Tip? *J Med Chem* 58(16): 6321-6335.

- 
- Grunwald, E. and Steel, C. (1995).** Solvent Reorganization and Thermodynamic Enthalpy-Entropy Compensation. *J Am Chem Soc.* 117:5687–5692.
- Huggins, D.J., Sherman, W. and Tidor, B. (2012).** Rational approaches to improving selectivity in drug design. *J Med Chem* 55(4): 1424-1444.
- Kamiya, H. (2006).** Recognition of nucleotide analogs containing the 7,8-dihydro-8-oxo structure by the human MTH1 protein. *J. Biochem.* 140, 843-849
- Kettle, J. G., Alwan, H., Bista, M., Breed, J., Davies, N. L., Eckersley, K., Fillery, S., Foote, K. M., Goodwin, L., Jones, D. R., Kack, H., Lau, A., Nissink, J. W., Read, J., Scott, J. S., Taylor, B., Walker, G., Wissler, L. and Wylot, M. (2016).** Potent and Selective Inhibitors of MTH1 Probe Its Role in Cancer Cell Survival. *J. Med. Chem.* 59, 2346–2361.
- King, N. M., Prabu-Jeyabalan, M., Bandaranayake, R. M., Nalam, M. N., Nalivaika, E. A., Ö zen, A., Haliloğlu, T., Yılmaz, N. K. and Schiffer, C. A. (2012).** Extreme Entropy–Enthalpy Compensation in a Drug-Resistant Variant of HIV-1 Protease. *ACS Chem. Biol.* 120702155441004.
- Klebe, G. (2015).** Applying thermodynamic profiling in lead finding and optimization. *Nat Rev Drug Discov* 14(2): 95-110.
- Koshland, Jr., E., D. (1994).** The key-lock theory and the induced fit theory. *Angew. Chem. Int. Ed. Engl.* 1994, 33, 2375–2378
- Kuhn, B., Fuchs, J.E., Reutlinger, M., Stahl, M. and Taylor, N.R. (2011).** Rationalizing tight ligand binding through cooperative interaction networks. *J Chem Inf Model* 51(12): 3180-3198.
- Kuhn, B., Guba, W., Banner, D., Bissantz, C., Ceccarelli, S., Haap, W., Körner, M., Kuglstatter, A., Lerner, C., Mattei, P., Neidhart, W., Pinard, E., Rudolph, M., Schulz-gasch, T., Woltering, T. and Stahl, M. (2016).** A Real-World Perspective

on Molecular Design. *J Med Chem* 59(9): 4087-4102.

Lafont, V., Armstrong, A.A., Ohtaka, H., Kiso, Y., Mario, Amzel, L. and Freire, E.

(2007). Compensating enthalpic and entropic changes hinder binding affinity optimization. *Chem Biol Drug Disc.* 69:413–422.

Liou, G. Y. and Storz, P. (2010). Reactive oxygen species in cancer. *Free Radic Res* 44(5): 479-496.

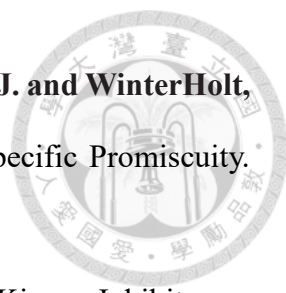
Luo, M., He, H., Kelley, M.R. and Georgiadis, M.M. (2010). Redox regulation of DNA repair: implications for human health and cancer therapeutic development. *Antioxid. Redox Signaling* 12, 1247– 1269.

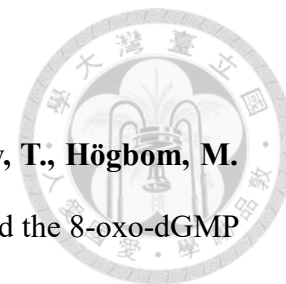
Maki, H., and Sekiguchi, M. (1992). MutT protein specifically hydrolyses a potent mutagenic substrate for DNA synthesis. *Nature* 355,273–275

Meyer, E. (1992). Internal water molecules and H-bonding in biological macromolecules: a review of structural features with functional implications. *Protein Sci.* 1, 1543-1 562.

Mishima, M., Sakai, Y., Itoh, N., Kamiya, H., Furuichi, M., Takahashi, M., Yamagata, Y., Iwai, S., Nakabeppu, Y. and Shirakawa, M. (2004). Structure of human MTH1, a Nudix family hydrolase that selectively degrades oxidized purine nucleoside triphosphates. *J Biol Chem* 279(32): 33806-33815.

Nakamura, T., Meshitsuka, S., Kitagawa, S., Abe, N., Yamada, J., Ishino, T., Nakano, H., Tsuzuki, T., Doi, T., Kobayashi, Y., Fujii, S., Sekiguchi, M. and Yamagata, Y. (2010). Structural and dynamic features of the MutT protein in the recognition of nucleotides with the mutagenic 8-oxoguanine base. *J.Biol.Chem.*285,444–452

- 
- Nissink, J. W., Bista, M., Breed, J., Carter, N., Embrey, K., Read, J. and WinterHolt, J. J.** (2016). MTH1 Substrate Recognition--An Example of Specific Promiscuity. *PLoS One* 11(3): e0151154.
- Noble, M.E., Endicott, J.A. and Johnson, L.N.** (2004). Protein Kinase Inhibitors: Insights into Drug Design from Structure. *Science* 2004, 303, 1800–1805.
- Olsson, T. S. G., Ladbury, J. E., Pitt, W. R. and Williams, M. A.** (2011). Extent of enthalpy-entropy compensation in protein-ligand interactions. *Protein Science*. 2011; 20:1607–1618.
- Onufriev, A.V. and Alexov, E.** (2013). Protonation and pK changes in protein-ligand binding. *Q. Rev. Biophys.* 46, 181–209.
- Qian, H.** (1998). Entropy-enthalpy compensation: Conformational fluctuation and induced-fit. *J Chem Phys.* 109:10015–10017
- Russo, M.T., Blasi, M.F., Chiera, F., Fortini, P., Degan, P., Macpherson, P., Furuichi, M., Nakabeppu, Y., Karran, P., Aquilina, G. and Bignami, M.** (2004). The oxidized deoxynucleoside triphosphate pool is a significant contributor to genetic instability in mismatch repair-deficient cells. *Mol.Cell.Biol.* 24,465–474.
- Sakai, Y., Furuichi, M., Takahashi, M., Mishima, M., Iwai, S., Shirakawa, M. and Nakabeppu, Y.** (2002). A molecular basis for the selective recognition of 2-hydroxy-dATP and 8-oxo-dGTP by human MTH1. *J. Biol. Chem.* 277,8579–8587
- Schnapp, G., Klein, T., Hoevels, Y., Bakker, R.A. and Nar, H.** (2016). Comparative Analysis of Binding Kinetics and Thermodynamics of Dipeptidyl Peptidase-4 Inhibitors and Their Relationship to Structure. *J Med Chem* 59(16): 7466-7477.
- Sharp, K.** (2001). Entropy-enthalpy compensation: fact or artifact? *Protein Sci* 10(3): 661-667.
- Storz, P.** (2005). Reactive oxygen species in tumor progression. *Front Biosci* 10:1881–



Svensson, L.M., Jemth, A.S., Desroses, M., Loseva, O., Helleday, T., Högbom, M.

and Stenmark, P. (2011). Crystal structure of human MTH1 and the 8-oxo-dGMP product complex. *FEBS Lett* 585(16): 2617-2621.

Tellinghuisen, J. A. (2003). study of statistical error in isothermal titration calorimetry.

Anal Biochem. 2003; 321:79–88.

Waz, S., Nakamura, T., Hirata, K., Koga-Ogawa, Y., Chirifu, M., Arimori, T.,

Tamada, T., Ikemizu, S., Nakabeppu, Y. and Yamagata, Y. (2017). Structural and Kinetic Studies of the Human Nudix Hydrolase MTH1 Reveal the Mechanism for Its Broad Substrate Specificity. *J Biol Chem* 292(7): 2785-2794.

wwPDB consortium. Protein Data Bank: The single global archive for 3D macromolecular structure data jointly managed by the Worldwide Protein Data Bank.

Nucleic Acid Res. 2019; doi:10.1093/nar/gky949.

Velazquez-Campoy, A., Kiso, Y. and Freire, E. (2001). The binding energetics of first- and second-generation HIV-1 protease inhibitors: implications for drug design. *Arch Biochem Biophys*;390:169–175.

Tables and Figures



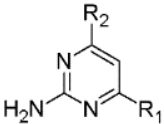
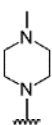
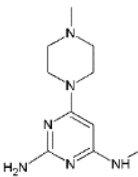
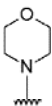
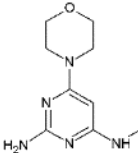
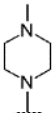
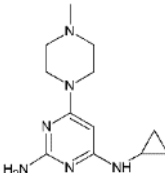
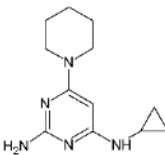
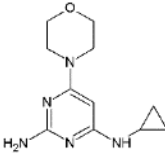
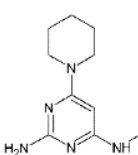
Table 1. Structure and related biochemical data of scaffold 1-derived compounds^a

Scaffold 1	ID	R1=	R2=	Structure	IC ₅₀ (nM) ¹	K _i (nM) ²
	1				57300	36200
	3				261000	165200
	6				71100	45000
	13				263900	167000
	14				>10 ⁶	675300
	16				4800	3000

^{a1}Biochemical potency of scaffold 1 derivatives was determined in enzyme inhibition assay using PPiLight inorganic pyrophosphate assay kit from Lonza which detects inorganic phosphate released from hydrolysis of substrate 8-oxo dGTP catalyzed by MTH1.

²Inhibitory constant K_i was calculated from IC₅₀ determined in enzyme inhibition assay using formula IC₅₀=(1 + [S]/K_m) × K_i.

Table 2. Structure and related biochemical data of scaffold 2-derived compounds^a

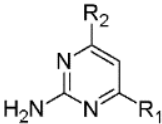
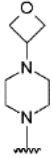
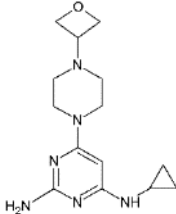
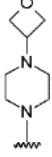
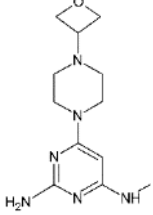
Scaffold 2	ID	R1=	R2=	Structure	IC ₅₀ (nM) ¹	K _i (nM) ²
	12				310	190
	20				10	6.3
	22				1500	949.3
	24				6.3	4.0
	25				41	25.9
	26				22.6	14.3

^{a1}Biochemical potency of scaffold 2 derivatives was determined in enzyme inhibition assay using PPiLight inorganic pyrophosphate assay kit from Lonza which detects inorganic phosphate released from hydrolysis of substrate 8-oxo dGTP catalyzed by MTH1.

²Inhibitory constant K_i was calculated from IC₅₀ determined in enzyme inhibition assay using formula $IC_{50} = (1 + [S]/K_m) \times K_i$.

Table 2. Structure and related biochemical data of scaffold 2-derived compounds^a

(continued)

Scaffold 2	ID	R1=	R2=	Structure	IC ₅₀ (nM) ¹	K _i (μM) ²
	29				2030	1284.8
	30				140	88.6

^{a1}Biochemical potency of scaffold 2 derivatives was determined in enzyme inhibition assay using PPiLight inorganic pyrophosphate assay kit from Lonza which detects inorganic phosphate released from hydrolysis of substrate 8-oxo dGTP catalyzed by MTH1.

²Inhibitory constant K_i was calculated from IC₅₀ determined in enzyme inhibition assay using formula $IC_{50} = (1 + [S]/K_m) \times K_i$.

Table 3. Thermodynamic profiles of scaffold 2-derived compounds determined with iTC

ID	N (sites)	K _d [× 10 ⁹ M ⁻¹]*	ΔG [kcal/mol]	ΔH [kcal/mol]	-TΔS [kcal/mol]
6077639	1.69	199 ± 12	-9.16 ± 0.16	-7.34 ± 0.08	-1.8 ± 0.08
20	0.89	6.2 ± 4.8	-11.3 ± 0.69	-20.66 ± 1.6	9.3 ± 2.2
22	0.78	64.9 ± 16.2	-9.76 ± 0.25	-15.43 ± 0.05	5.7 ± 0.3
24	0.92	2.9 ± 0.1	-11.69 ± 0.07	-18.7 ± 0.28	4.7 ± 0.2
25	1.01	9.6 ± 5.3	-11.05 ± 0.41	-14.83 ± 0.86	3.8 ± 0.5
26	1.01	3.1 ± 0.5	-11.59 ± 0.21	-18.3 ± 0.42	6.7 ± 0.6
29	1.25	129 ± 53	-9.75 ± 0.81	-7.95 ± 2.47	-1.8 ± 1.7
30	1.12	16.1 ± 3.9	-10.66 ± 0.15	-11.7 ± 0.36	1.0 ± 0.2

*Dissociation constant K_d of test compounds was calculated from equation $K_d = 1/K_a$ where K_a is the association rate constant given in M⁻¹s⁻¹.

Table 4. Data collection and refinement statistics for MTH1 crystal structure complexed with 11 structurally-related 2-aminopyrimidine-based compounds

Complex structure	MTH1-6	MTH1-12	MTH1-14
Beamline	NSRRC13B1	NSRRC13B1	NSRRC13B1
Wavelength (Å)	1	1	1
Data collection			
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Cell parameters			
<i>a</i> , <i>b</i> , <i>c</i> (Å)	59.0, 67.5, 79.3	59.0, 67.8, 79.6	59.2, 67.2, 79.6
Resolution (Å)	30.0-2.3 (2.38-2.30)*	30.0-2.1 (2.18-2.10)	30.0-1.9 (1.97-1.90)
R _{merge}	0.052 (0.157)	0.066 (0.262)	0.057 (0.317)
I/σ(I)	25.3 (10.1)	21.1 (5.4)	21.3 (4.4)
Completeness (%)	99.3 (100.0)	99.7 (98.9)	99.6 (99.5)
Redundancy	4.7 (4.9)	4.6 (4.4)	4.0 (3.9)
Reflections (total/unique)	218220/14689	474823/19253	359134/25855
Refinement			
R _{work} /R _{free}	19.2/26.0	19.0/25.3	19.3/24.7
Reflections (factor/test)	14317/1347	18634/1738	25104/2316
No. of protein atoms	2476	2476	2476
No. of ligand atoms	17	16	20
No. of water	120	274	260
Model quality			
R.M.S.D.			
Bond lengths (Å)	0.008	0.008	0.007
Bond angles (°)	1.01	0.99	0.94
Average B-factor (Å ²)	28.0	24.6	28.3
B-factor (macromolecules)	27.8	24.1	27.3
B-factor (ligands)	29.8	23.1	48.5
B-factor (solvent)	30.2	29	36.4
Ramachandran plot			
Most favored (%)	99.7	99	99.3
Allowed (%)	0.3	1	0.7
Disallowed (%)	0	0	0

* values for highest-resolution shell are shown in parentheses.

Table 4. Data collection and refinement statistics for MTH1 crystal structure complexed with 11 structurally-related 2-aminopyrimidine-based compounds (continued)

Complex structure	MTH1-16	MTH1-20	MTH1-22
Beamline	NSRRC13C	NSRRC 13C1	NSRRC 13B1
Wavelength (Å)	0.976	0.976	1
Data collection			
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Cell parameters			
<i>a</i> , <i>b</i> , <i>c</i> (Å)	59.1, 67.7, 79.9	59.2, 67.5, 80.0	59.0, 68.2, 78.9
Resolution (Å)	30.0-2.1 (2.18-2.10)	30.0-2.1 (2.18-2.10)	30.0-1.9 (1.97-1.90)
R _{merge}	0.086 (0.358)	0.051 (0.106)	0.062 (0.289)
I/σ(I)	10.5 (3.5)	23.2 (12.6)	19.2 (5.3)
Completeness (%)	99.6 (99.8)	99.5 (99.5)	99.8 (99.8)
Redundancy	3.7 (3.7)	3.6 (3.7)	4.4 (4.3)
Reflections (total/unique)	111364/19234	219912/19298	363668/25653
Refinement			
R _{work} /R _{free}	19.8/26.3	19.0/24.6	19.0/24.8
Reflections (factor/test)	18526/1726	18970/1842	25018/2329
No. of protein atoms	2476	2475	2476
No. of ligand atoms	38	30	36
No. of water	187	242	246
Model quality			
R.M.S.D.			
Bond lengths (Å)	0.008	0.008	0.007
Bond angles (°)	0.98	0.96	0.95
Average B-factor (Å ²)	32.6	27.6	30.0
B-factor (macromolecules)	32.2	27.0	29.3
B-factor (ligands)	41.0	17.2	23.6
B-factor (solvent)	35.8	34.5	38.5
Ramachandran plot			
Most favored (%)	99	99	99.7
Allowed (%)	1	1	0.3
Disallowed (%)	0	0	0

* values for highest-resolution shell are shown in parentheses.

Table 4. Data collection and refinement statistics for MTH1 crystal structure complexed with 11 structurally-related 2-aminopyrimidine-based compounds (continued)

Complex structure	MTH1-24	MTH1-25	MTH1-26
Beamline	NSRRC 13C1	NSRRC 13C1	NSRRC 13B1
Wavelength (Å)	0.976	0.976	1
Data collection			
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Cell parameters			
<i>a</i> , <i>b</i> , <i>c</i> (Å)	59.1, 68.9, 78.3	59.1, 68.8, 78.9	59.2, 68.2, 78.8
Resolution (Å)	30.0-2.2 (2.28-2.20)	30.0-2.2 (2.28-2.20)	30.0-2.3 (2.38-2.30)
R _{merge}	0.100 (0.282)	0.095 (0.331)	0.074 (0.451)
I/σ(I)	13.0 (4.9)	12.8 (4.8)	15.0 (2.6)
Completeness (%)	99.5 (99.4)	97.3 (99.4)	97.6 (92.8)
Redundancy	4.3 (4.0)	3.3 (3.1)	3.6 (3.3)
Reflections (total/unique)	403919/16640	463061/16988	453854/14867
Refinement			
R _{work} /R _{free}	19.3/25.6	19.5/26.6	19.8/26.6
Reflections (factor/test)	16173/1483	16090/1517	13688/1147
No. of protein atoms	2476	2476	2476
No. of ligand atoms	34	34	15
No. of water	151	147	90
Model quality			
R.M.S.D.			
Bond lengths (Å)	0.007	0.007	0.009
Bond angles (°)	0.99	0.97	1.15
Average B-factor (Å ²)	30.7	39.2	43.6
B-factor (macromolecules)	30.6	39.1	43.6
B-factor (ligands)	20.0	30.6	34.5
B-factor (solvent)	34.8	43.6	46.8
Ramachandran plot			
Most favored (%)	99.3	98.7	98
Allowed (%)	0.7	1.3	2
Disallowed (%)	0	0	0

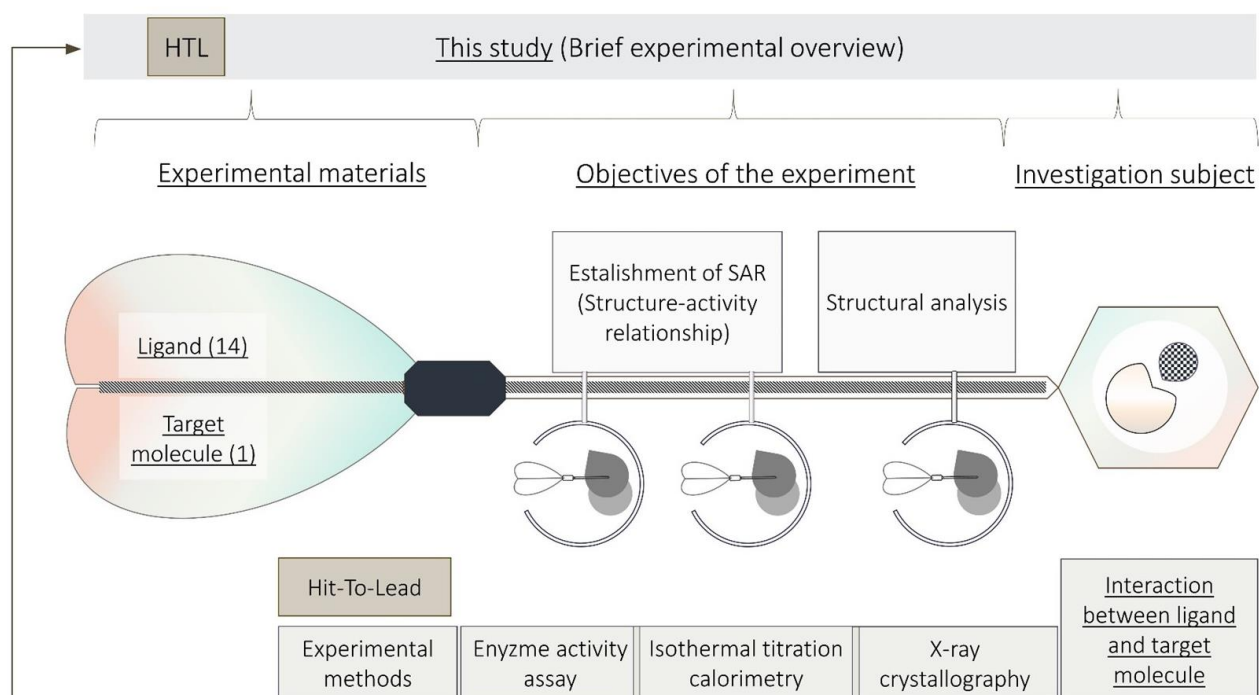
* values for highest-resolution shell are shown in parentheses.

Table 4. Data collection and refinement statistics for MTH1 crystal structure complexed with 11 structurally-related 2-aminopyrimidine-based compounds (continued)

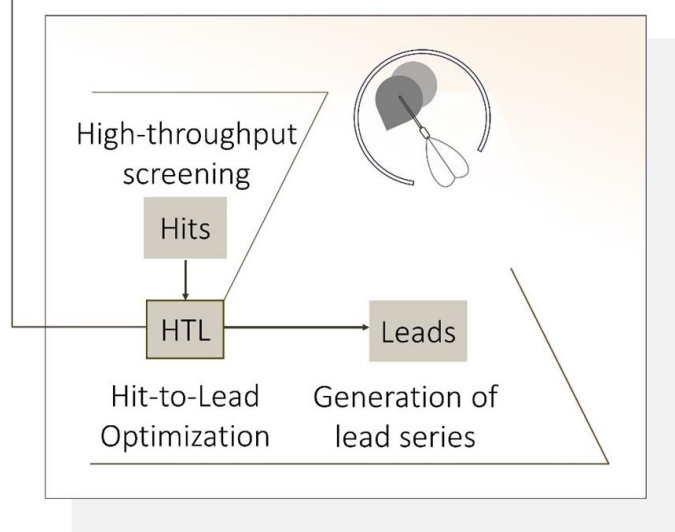
Complex structure	MTH1-29	MTH1-30
Beamline	NSRRC 13C1	NSRRC 13C1
Wavelength (Å)	0.976	0.976
Data collection		
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Cell parameters		
<i>a</i> , <i>b</i> , <i>c</i> (Å)	59.1, 68.4, 79.4	59.2, 67.7, 79.6
Resolution (Å)	30.0-2.1 (2.18-2.10)	30.0-1.8 (1.86-1.80)
R _{merge}	0.050 (0.132)	0.067 (0.462)
I/σ(I)	23.6 (10.3)	19.5 (2.9)
Completeness (%)	99.7 (100.0)	99.5 (97.6)
Redundancy	3.9 (4.1)	4.3 (3.9)
Reflections (total/unique)	332824/19399	342060/30306
Refinement		
R _{work} /R _{free}	19.9/25.2	18.8/23.1
Reflections (factor/test)	19046/1802	29190/2580
No. of protein atoms	2476	2476
No. of ligand atoms	21	38
No. of water	187	325
Model quality		
R.M.S.D.		
Bond lengths (Å)	0.008	0.007
Bond angles (°)	0.95	0.95
Average B-factor (Å ²)	37.4	22.4
B-factor (macromolecules)	37.0	21.1
B-factor (ligands)	35.9	30.0
B-factor (solvent)	41.9	31.5
Ramachandran plot		
Most favored (%)	99	99.3
Allowed (%)	1.0	0.7
Disallowed (%)	0.0	0.0

* values for highest-resolution shell are shown in parentheses

A



B



C

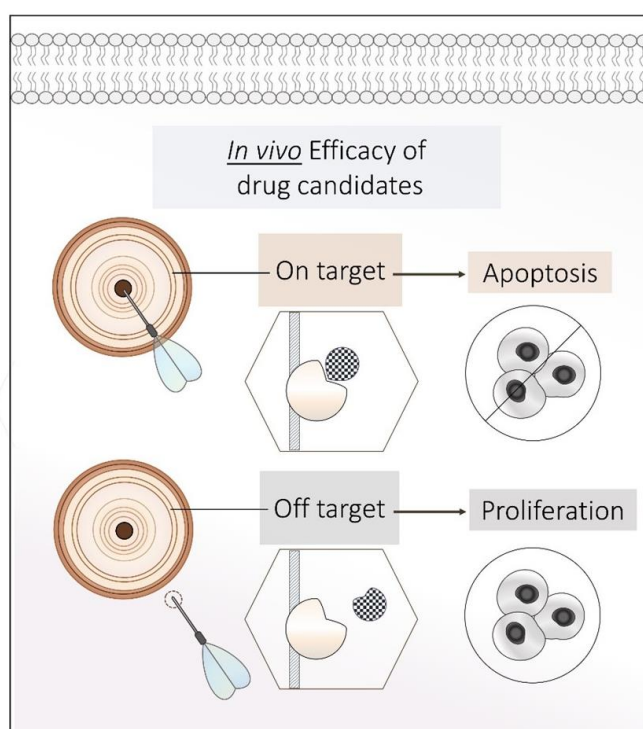


Figure 1. Schematic depiction of simplified work flow and objectives of this study

(A) Brief overview of experimental work flow and objectives in this study (the concept of the study is depicted illustratively which resembles designed ligand which plays as a key (referred to lock and key model) and interacts with a given target molecule.

In this study, to investigate interaction pattern of 14 small molecule compounds (at the stage of hit-to-lead optimization) in the binding pocket of target molecule, biochemical and biophysical assays such as enzyme activity analysis and isothermal titration calorimetry (iTTC) were performed. For visualization of ligand molecule in binding pocket, crystals of target molecule in complex with ligand were obtained via soaking and complex structures were subsequently determined with x-ray diffraction.

(B) Simplified work flow of drug discovery campaign which is comprised of three integral stages. Starting from high-throughput screening performed on compound libraries, hits with moderate affinity toward a given target are identified. Through optimization of hit compounds by means of biochemical and biophysical techniques, leads with elevated binding affinity and improved selectivity profiles are generated and serve as drug candidates for further evaluation of *in vitro* and *in vivo* efficacy.

(C) Off-target effects which is related to lack of specificity and selectivity of a given ligand have to be taken into consideration which is influential in establishment of efficacy of designed drug candidates.

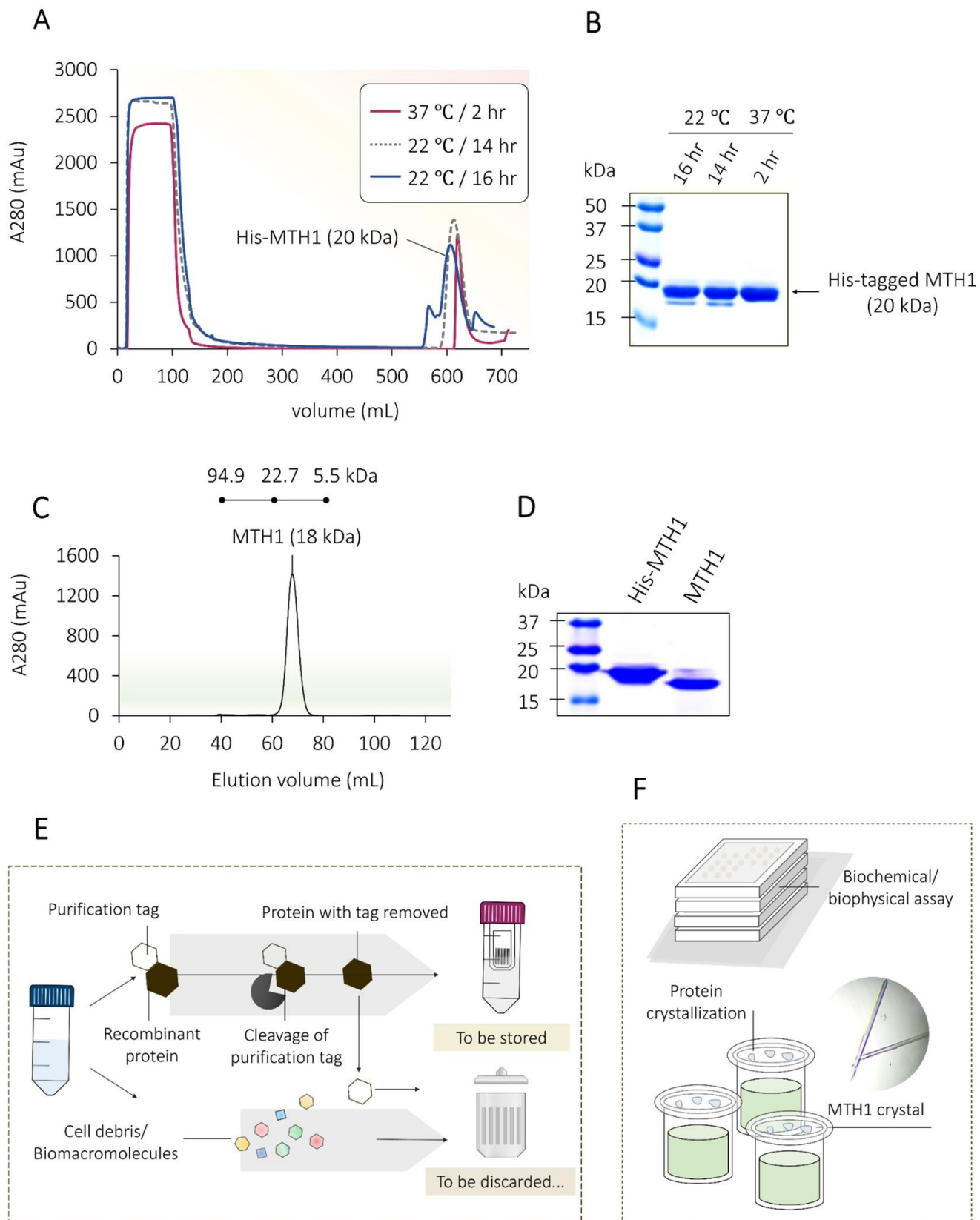
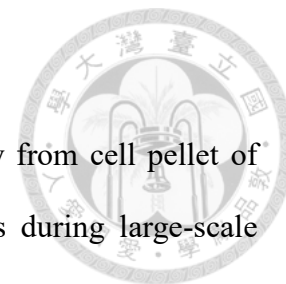


Figure 2. Purification of MTH1 protein from E.coli cell lysate



(A) purification of his-tagged MTH1 with affinity chromatography from cell pellet of *E.coli* incubated under three independent experimental conditions during large-scale expression of target protein.

(B) the quality of purified his-MTH1 protein was subsequently checked on SDS-PAGE gel which suggests protein degradation under condition of incubation time 14 and 16 hour at 22°C. Whilst his-MTH1 protein purified from cell lysate of *E.coli* incubated at 37°C for two hours remains intact after purification using affinity column.

(C) Size-exclusion chromatography is applied for separation of his-tag and MTH1 protein. MTH1 protein (18 kDa) was eluted out of column at elution volume 60-80 ml.

(D) SDS-PAGE analysis for confirming separation of MTH1 protein from cleaved his-tag.

(E) Schematic illustration of work flow of recombinant protein purification

(F) Purified protein sample serves as experimental material in biochemical, biophysical assays, and protein crystallization.

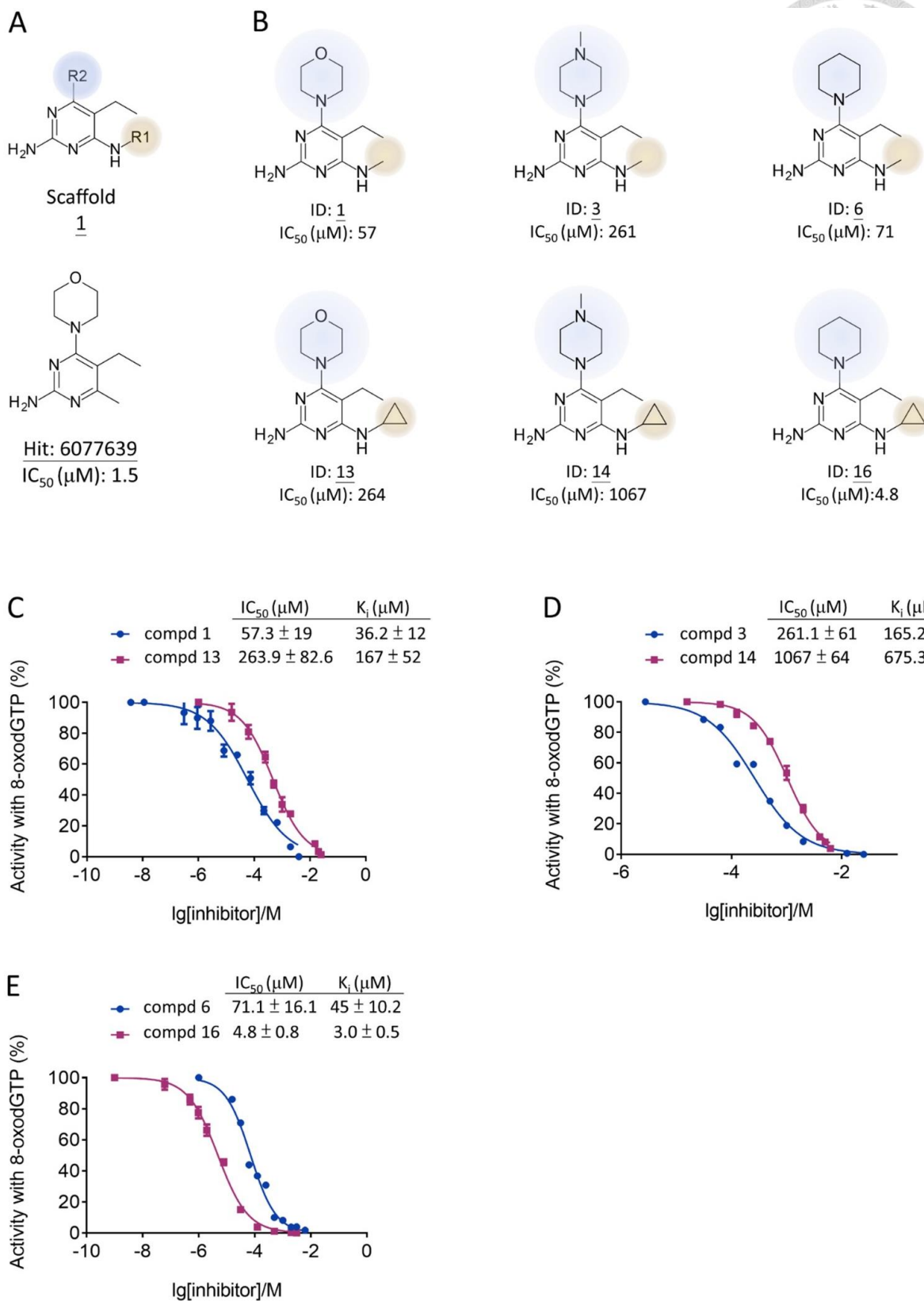


Figure 3. Binding affinity profile of scaffold 1 derivative compounds

- (A) Structural architecture of scaffold 1 with R1 and R2-substituent marked
- (B) Structural architecture of six structurally-related compounds derived from scaffold 1 with ID indicated below

(C-E) dose-response curve of scaffold 1 derivative compounds measured in enzyme inhibition assay, IC_{50} value is calculated from fitting curve and K_i inhibition constant is then empirically converted from IC_{50} value. It is remarkable that comparison of K_i of scaffold 1 derivatives according to substituents incorporated in the R1 and R2-position demonstrates significant fluctuation in biochemical potency upon replacement of substituents in R2-position. In opposite, the effect of R1-substituent is less prominent.



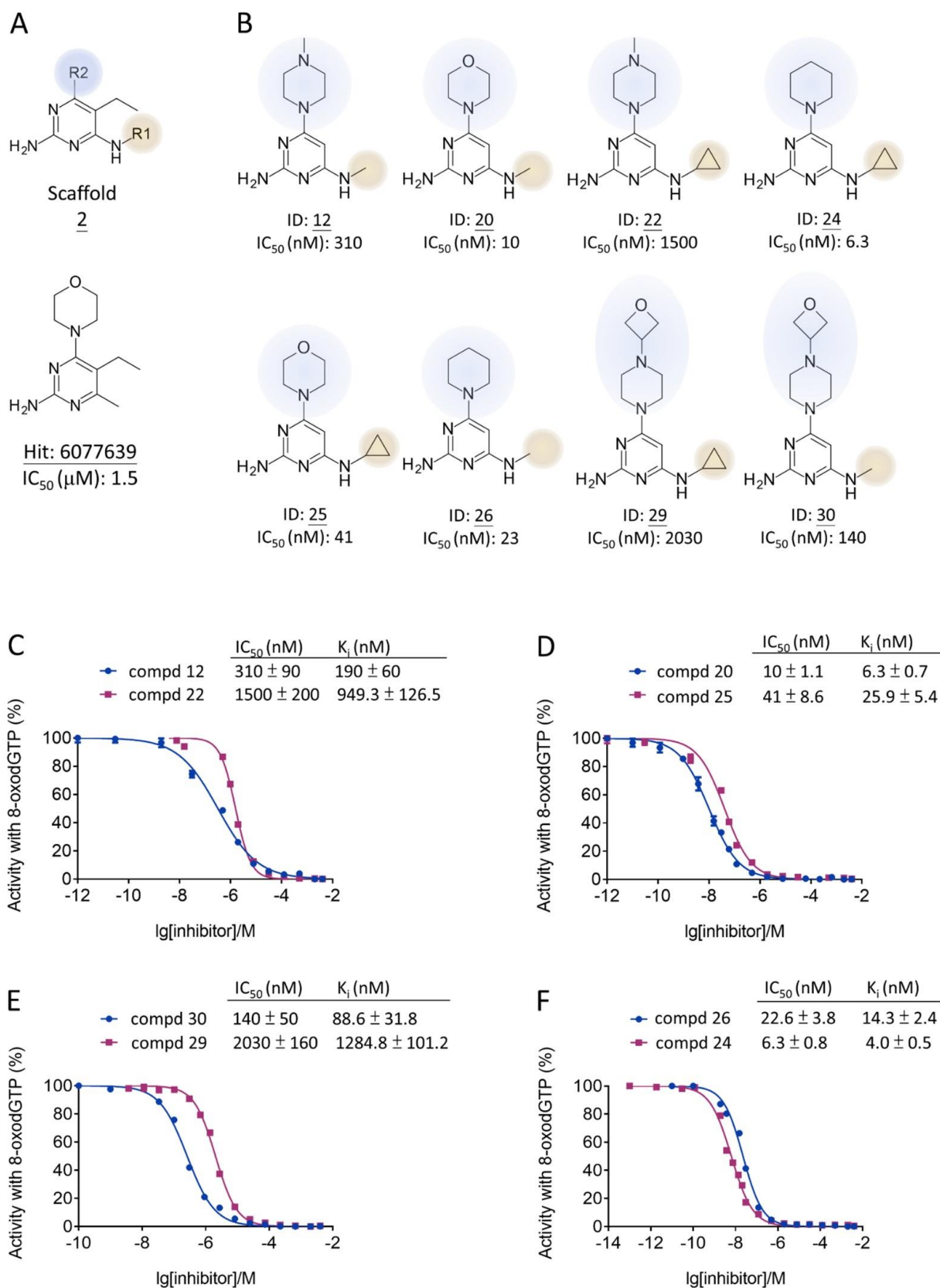


Figure 4. Binding affinity profile of scaffold 2 derivative compounds

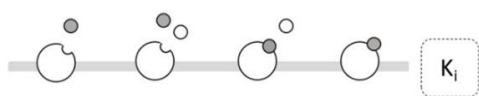
- (A) Structural architecture of scaffold 2 with R1 and R2-substituent marked
- (B) Structural architecture of eight structurally-related compounds derived from scaffold 2 with ID indicated below

(C-F) dose-response curve of scaffold 2 derivative compounds measured in enzyme inhibition assay, IC_{50} value is calculated from fitting curve and K_i inhibition constant is then empirically converted from IC_{50} value. It is remarkable that comparison of K_i of scaffold 2 derivatives according to substituents incorporated in the R1 and R2-position demonstrates significant fluctuation in biochemical potency upon replacement of substituents in R2-position. In opposite, the effect of R1-substituent is less prominent.



A

Enzyme Inhibition assay



Isothermal Titration Calorimetry

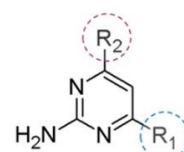


○ Substrate

● Ligand

○ Protein (unbound state)

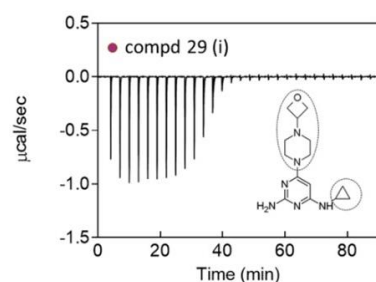
○ Protein-ligand complex



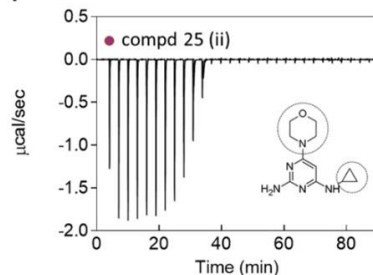
Scaffold 2



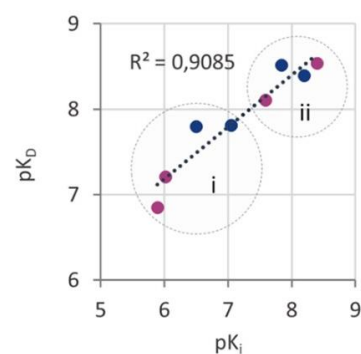
B



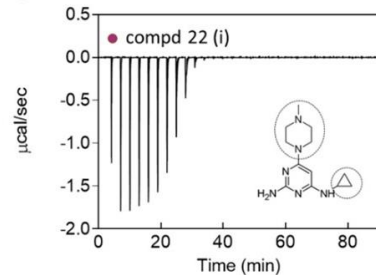
F



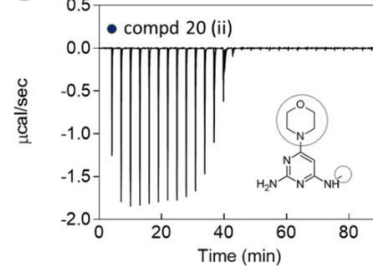
J



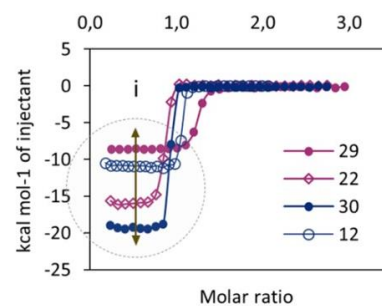
C



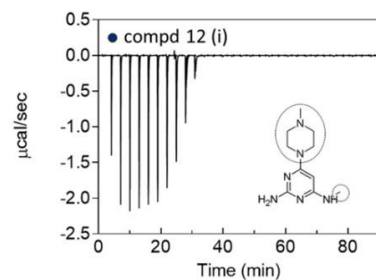
G



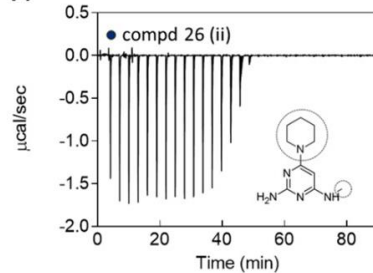
K



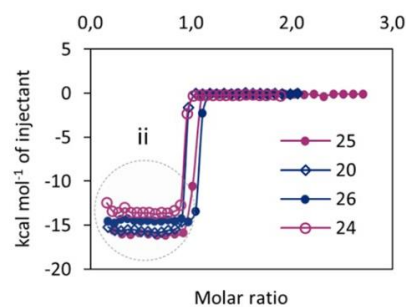
D



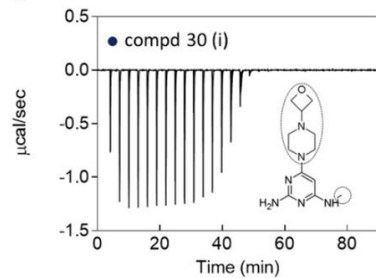
H



L



E



I

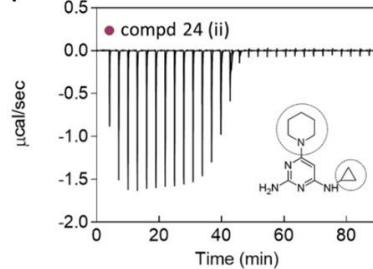


Figure 5. Thermodynamic study of scaffold 2 derivatives binding to MTH1 protein

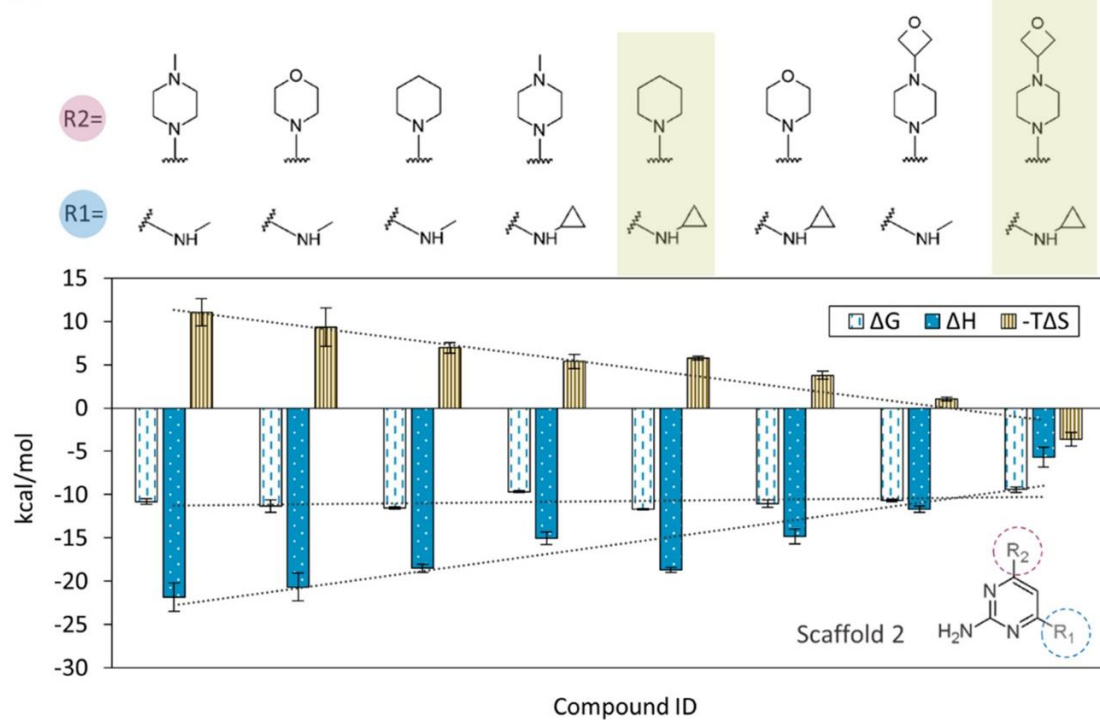
(A) Schematic representation of two analytical methods for characterization of binding affinity profiles of 2-amino pyrimidine-based compound series in this study. In enzyme inhibition assay, ligand and substrate compete for the same binding site on MTH1 protein. Oppositely, binding affinity between protein and ligand is measured in thermodynamic study using isothermal titration calorimetry (iTC).

(B-E) Experimental data of scaffold 2 derivative compounds with pK_D value lower than 8.0 (classified as group i shown in diagram (J)) recorded using isothermal titration calorimetry.

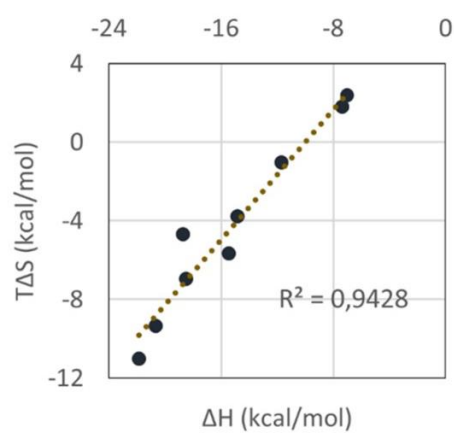
(F-I) Experimental data of scaffold 2 derivative compounds with pK_D value higher than 8.0 (classified as group ii shown in diagram (J)) recorded using isothermal titration calorimetry.

(J) Plotting pK_i value determined in enzyme inhibition assay against thermodynamically measured pK_D gives a regression coefficient value of 0.92 (K) binding isotherms calculated with one-site fitting model using raw data measured in iTC experiment from (B-E), significant deviation in the magnitude of exothermic signal recorded is notified (circled and indicated with arrows). (L) binding isotherms calculated with one-site fitting model using raw data from (F-I), relatively averaged magnitude of exothermic signal is characteristic for compounds from group ii (circled).

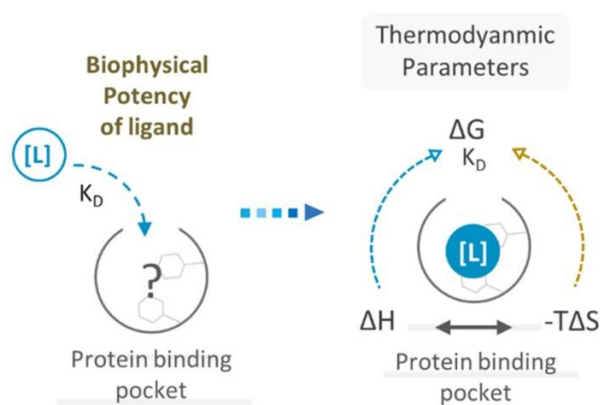
A



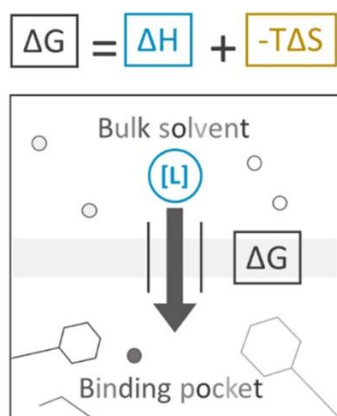
B



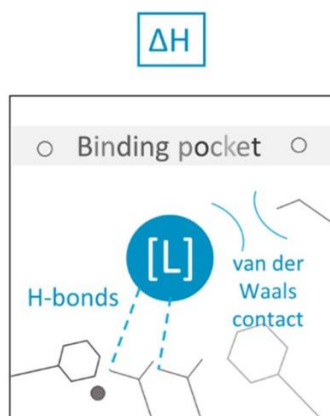
C



D



E



F

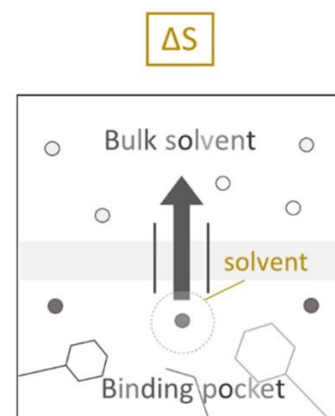


Figure 6. Phenomenon of EEC observed in scaffold 2 derivatives binding to MTH1 protein



(A) Diagram of thermodynamic profiles determined with isothermal titration calorimetry for 2-aminopyrimidine-based compounds derived from scaffold 2. Profound enthalpy-entropy compensation (EEC) is observed in the structurally related compound series. In the diagram, thermodynamic parameters ΔG , ΔH and $-T\Delta S$ are depicted as bar in color white, blue and orange, respectively. Substituents incorporated in R1 and R2-position of each scaffold 2 derivative compounds are displayed at top of the bar diagram.

(B) Plot of correlation between two thermodynamic parameters ΔH and $-T\Delta S$. ($r^2 = 0.94$)

(C) Schematic draw of working mechanism underlying protein-ligand complex formation which is analyzed using isothermal titration calorimeter. On the right side of arrow, mutually counteracting thermodynamic parameters ΔH and $-T\Delta S$ which contribute to binding free energy (ΔG) are illustrated schematically.

(D-F) Schematic representation of protein-ligand complex formation process illustrated with thermodynamic parameters notified. Firstly, as the binding partners are both solvated before the formation of complex, (D) driving force (ΔG , binding free energy) of complex formation propagate ligand molecule to enter the binding pocket.

The binding of two independent binding partners generally involves (E) formation of h-bonds and van der Waals contacts which are detected in the form of heat signal from iTC experiment. (F) While almost impossible to be observed in present experimental techniques, the conformational change of protein and ligand molecules during formation of complex or release of solvent molecules into bulk solvent and hydrophobic interaction which does not involve formation and breakage of chemical bonds would be attributed to entropic component ($-T\Delta S$) of binding free energy (ΔG).

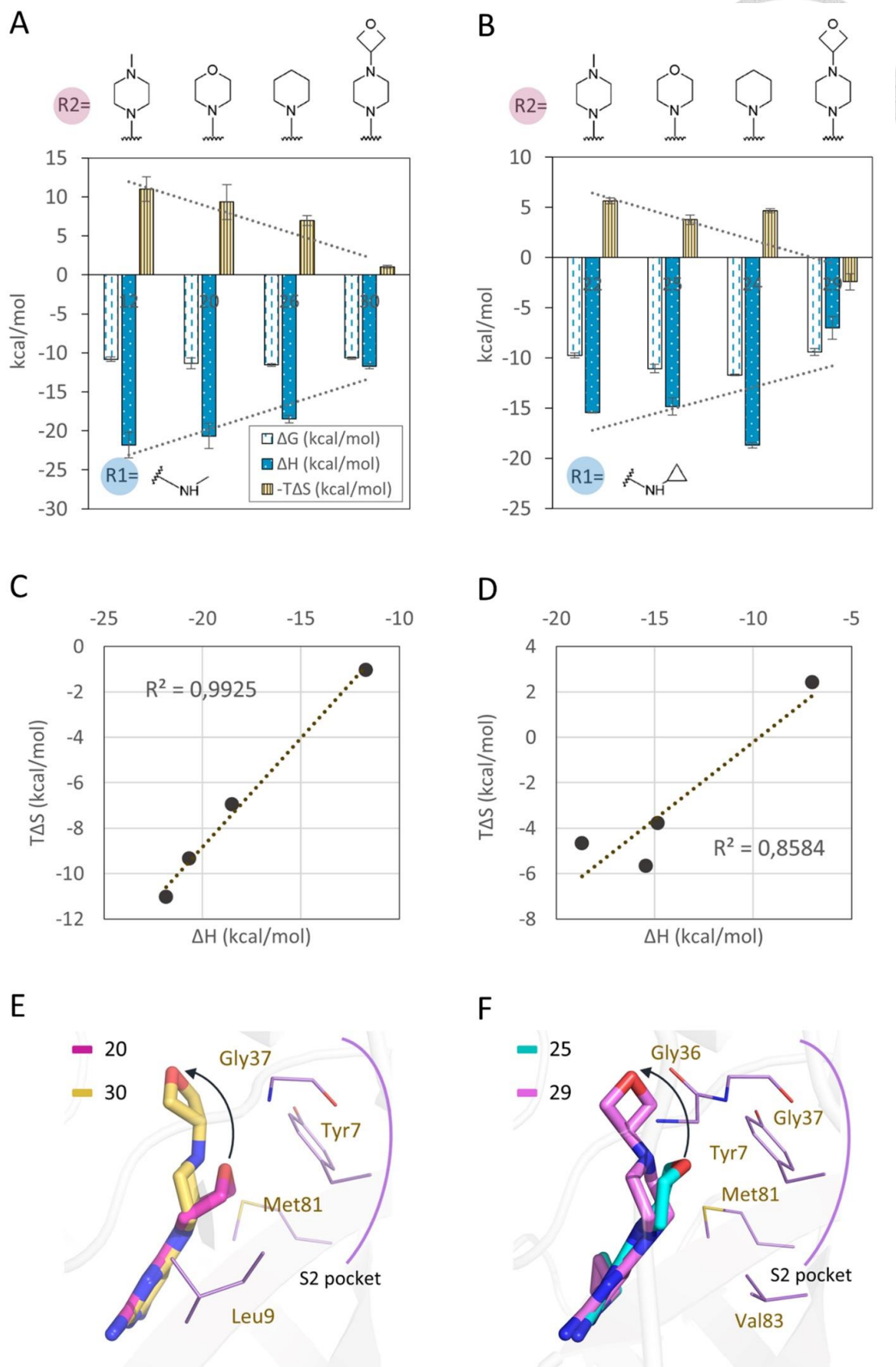


Figure 7. Comparison of thermodynamic profiles from structurally related compounds with different substituents incorporated in the R2-position

(A,B) eight structurally-related compounds are grouped according to substituents incorporated in R1-position as shown in bar diagram (A) and (B). Subsequently, thermodynamic profiles of the 2-aminopyrimidine-based compound series are plotted in bar diagram which stand for calculated thermodynamic parameters ΔG , ΔH and $-T\Delta S$. Four variants of R2-substituents corresponding to each compound are displayed above the bar diagram (A) and (B).

(C), (D), profound enthalpy-entropy compensation (EEC) is demonstrated in (C) which represents result from correlation analysis of thermodynamic parameters ΔH and $-T\Delta S$ derived from structurally-related compound series illustrated in (A) with a R^2 value of 0.993. While with N-cyclopropyl incorporated in R1-position, compound series with an array of identical R2-substituents from (B) result in a less prominent value of R^2 which dropped to 0.858 in the correlation analysis of thermodynamic parameters ΔH and $-T\Delta S$ as indicated in diagram (D).

(E), (F), Representative compounds with an array of substituents explored in the R2-position are superimposed in (E) and (F), separately. In (E), compound 20 and 30 from (A) are shown in stick representation colored in magenta and yellow, respectively. While with the same R2-substituents as compounds referred above, compound 25 and 29 from (B) with different R1-substituent incorporated are shown in stick representation colored in cyan and violet, respectively.

Side chains of adjacent residues in S2 pocket are shown in line representation colored in purple.

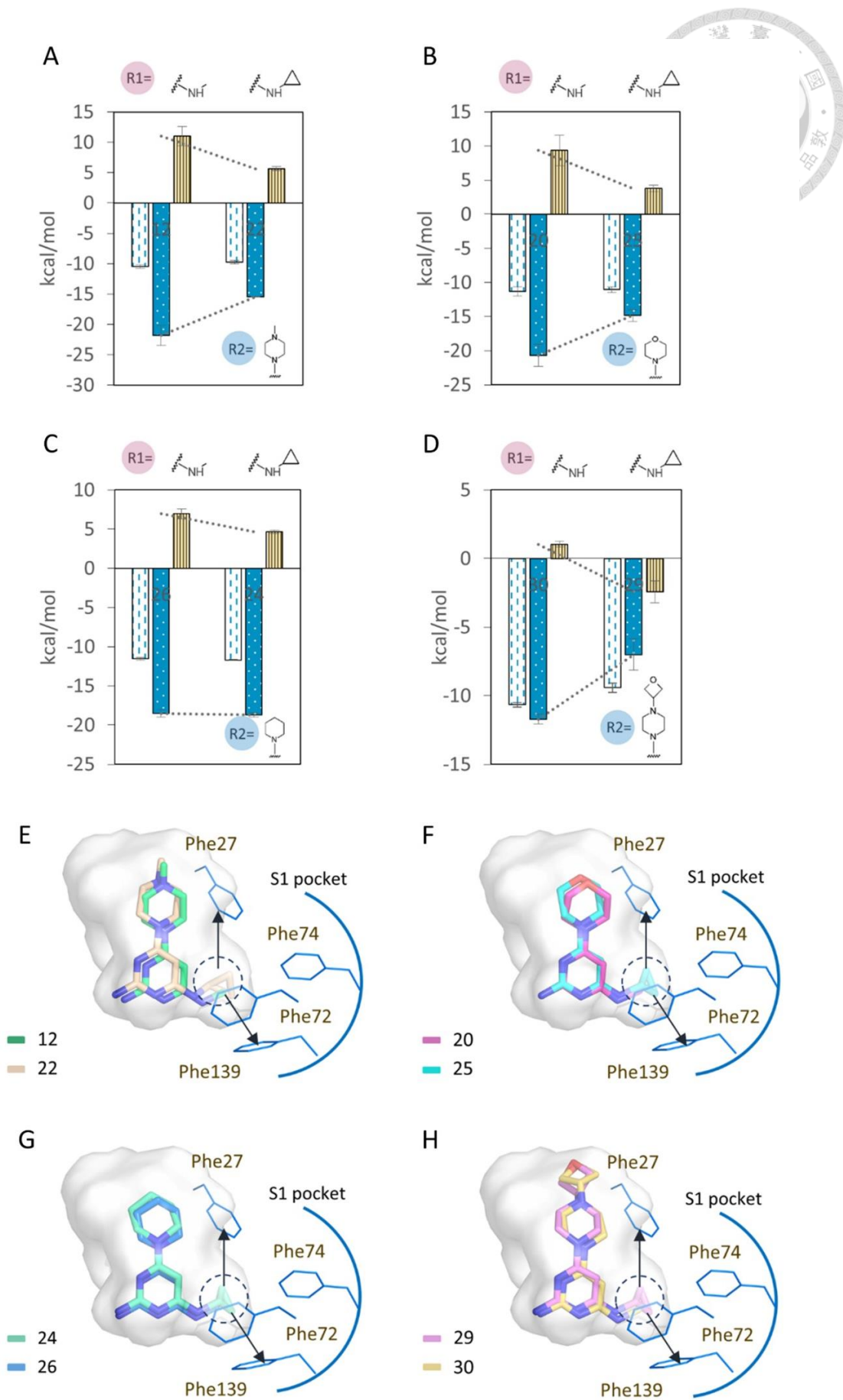


Figure 8. Comparison of thermodynamic profiles from structurally related compounds with different substituents incorporated in the R1-position

(A-D), thermodynamic parameters are plotted as bars colored in white, blue and orange for ΔG , ΔH and $-T\Delta S$, respectively. In each bar diagram, thermodynamic profiles of structurally related compounds which differentiate in R1-substituent are categorized into four independent groups according to R2-substituent of which incorporated into the 2-aminopyrimidine-based compounds.

(E-H), Superposition of compounds referred in (A-D) unraveled notable structural features which might be related to the observed phenomenon of enthalpy-entropy compensation (EEC) in R1-substituted counterpart compounds with the R1-substituent circled in black-dashed lines. Key residues Phe27, Phe72, Phe74 and Phe139 surrounding the R1-substituent are shown in line and labeled. Arc in blue is indicative of S1 pocket while arrows represent additional hydrophobic contact with surrounding residues from cyclopropane ring.

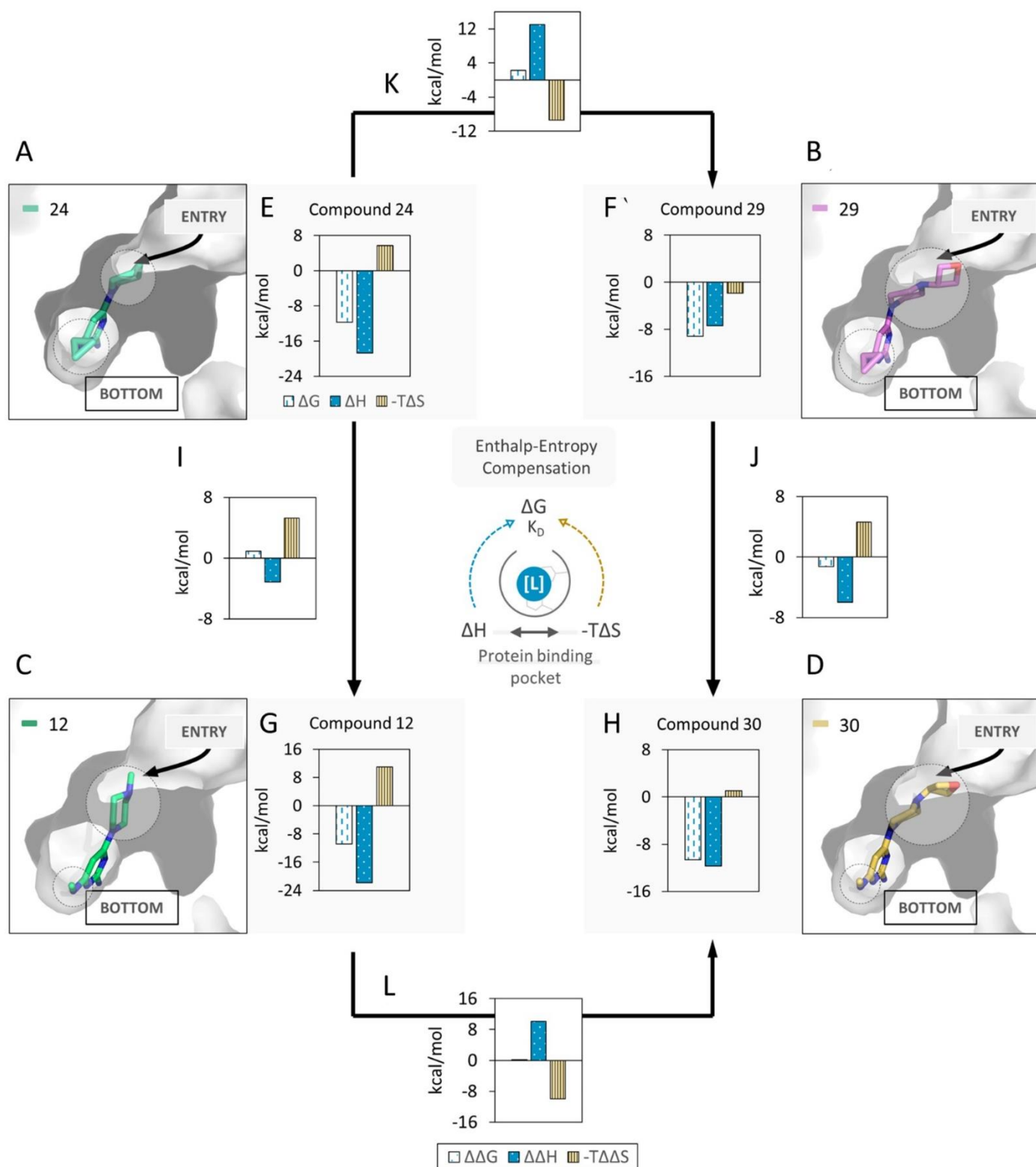


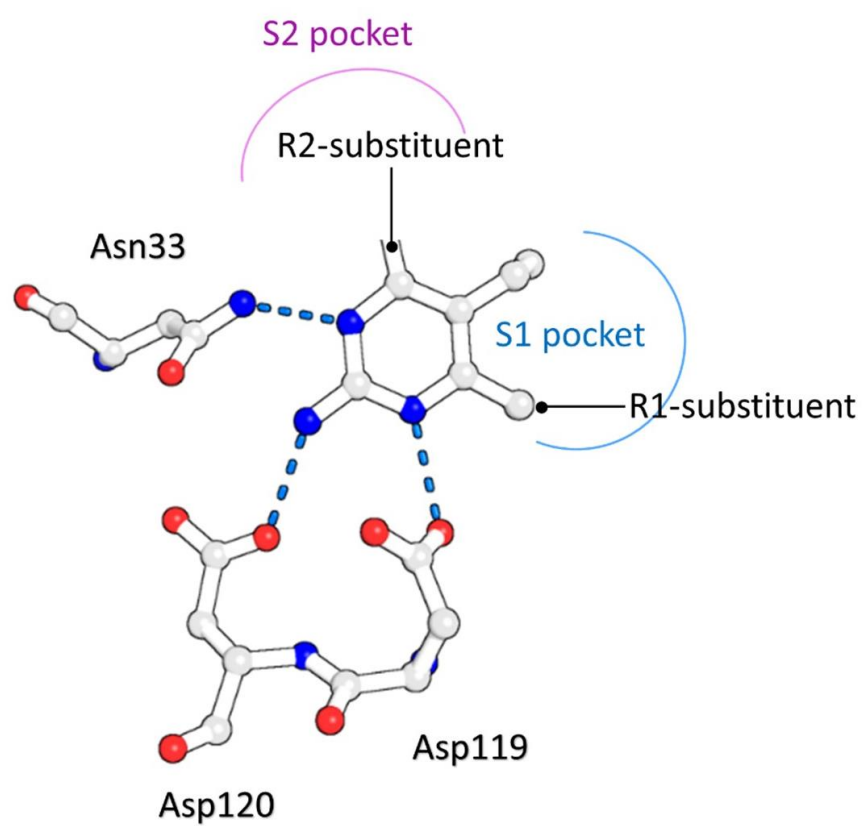
Figure 9. Phenomenon of enthalpy-entropy compensation gleaned from protein-ligand complex structures

(A-D) thermodynamic profiles of compound 24, 29, 12 and 30 with chemical structure and biochemical potency determined in enzyme inhibition assay listed.

(E-H) MTH1 binding pocket shown in surface representation with residues nearby displayed. The entry and bottom of MTH1 binding pocket are also indicated. To indicate the positioning of R1 and R2-substituent from separate ligand molecules in the MTH1 binding pocket, $\textcircled{\text{R1}}$ and $\textcircled{\text{R2}}$ are utilized as symbol for substituents at C4 and C6 of the pyrimidine moiety of compound 24, 29, 12 and 30 in the images.

(I-L) bar diagrams for $\Delta\Delta G$, $\Delta\Delta H$ and $-T\Delta\Delta S$ calculated from two sets of thermodynamic profiles which indicate the deviation in thermodynamic parameters between (I) compound 29 and 24 (J) compound 30 and 29 (K) compound 30 and 12 (L) compound 12 and 24. The counteracting effect of two thermodynamic parameters ΔH and $-T\Delta S$ is significant as demonstrated in diagrams i-l.

A



B

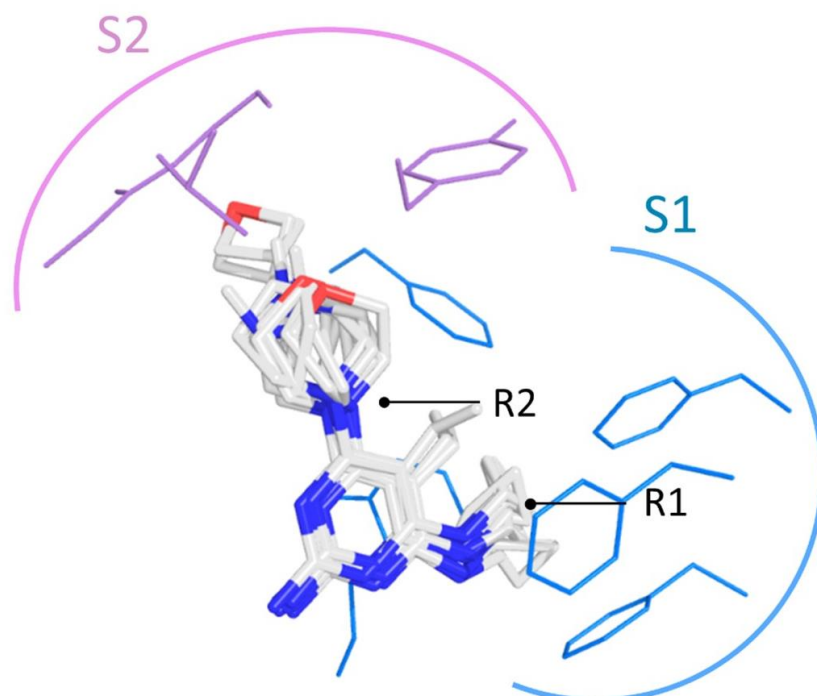
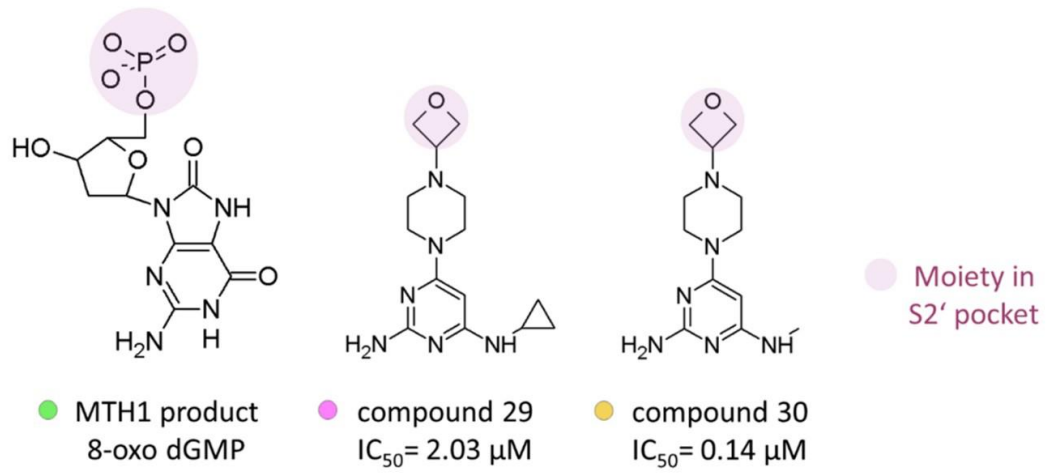


Figure 10 Comparison of binding mode of 11 structurally related 2-aminopyrimidine-based compounds uncovered in x-ray crystallographic study.

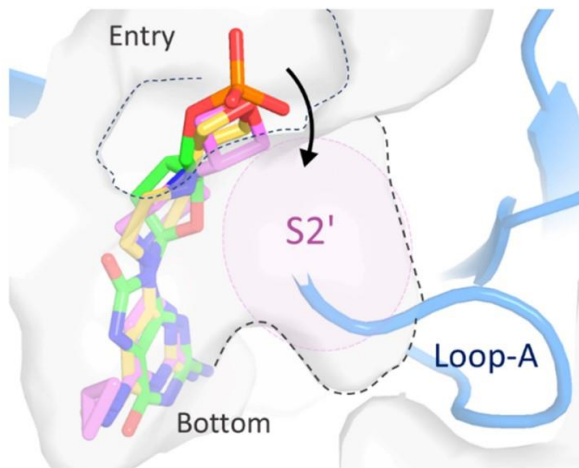
(A) compound 6077639 identified from high-throughput drug screening (HTS) serves as scaffold for the present 2-aminopyrimidine compound series. Hydrogen bonding interaction between pyrimidine moiety and side chains of residues Asn33, Asp119, Asp120 is depicted as dashed lines. For investigation of binding interaction of protein-ligand complex, systematical exploration with substituents was probed on C4 and C6 of the pyrimidine moiety which is noted as R1 and R2-substituent, respectively.

(B) Structural alignment of 11 ligands from 2-aminopyrimidine-based compound series uncovered binding mode of structurally related ligands in the binding pocket of MTH1. Arc adjacent to side chains of key residues signifies subregion in the binding pocket assigned according to space occupied by R1 and R2-substituent of ligand molecule. Two subregions are defined as S1 and S2 in this study which encompass residues interacting with R1 and R2-substituent in the binding pocket, respectively.

A



B



C

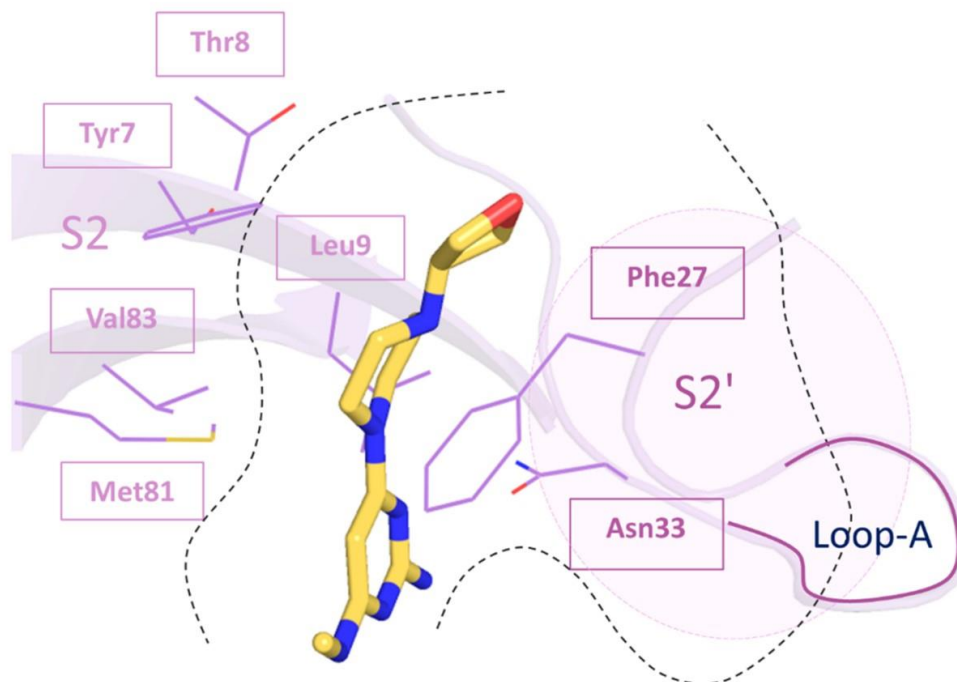


Figure 11. Binding pose of MTH1 product 8-oxo dGMP, compound 29 and 30 in bound state gleaned from crystal structure complexed with MTH1

(A) Structure of MTH1 product 8-oxo dGMP (PDB ID: 3ZR0), compound 29 and 30 with biochemical potency IC_{50} listed below. Moieties in close distance to S2' pocket are colored light purple.

Further, spots in green, dark pink and yellow symbolize ligand molecules displayed in diagram (B) and (C).

(B) Flexible loop-A is illustrated as cartoon which positions adjacent to substrate binding site modulating binding of triphosphate group of substrate and inhibitors. In the diagram, alignment of ligands of similar structural skeleton reveals that positioning of moieties near the entry of binding pocket overlaps well which is in close distance to S2' pocket as indicated by arrow.

(C) Close-up view of compound 30 bound in the binding pocket of MTH1. Residues engage in interaction with ligand are displayed which are categorized into S2 and S2' according to the spatial arrangement related to the positioning of R2-substituent in the binding pocket. Apparently, moiety near the entry of binding pocket tends to move toward S2'.

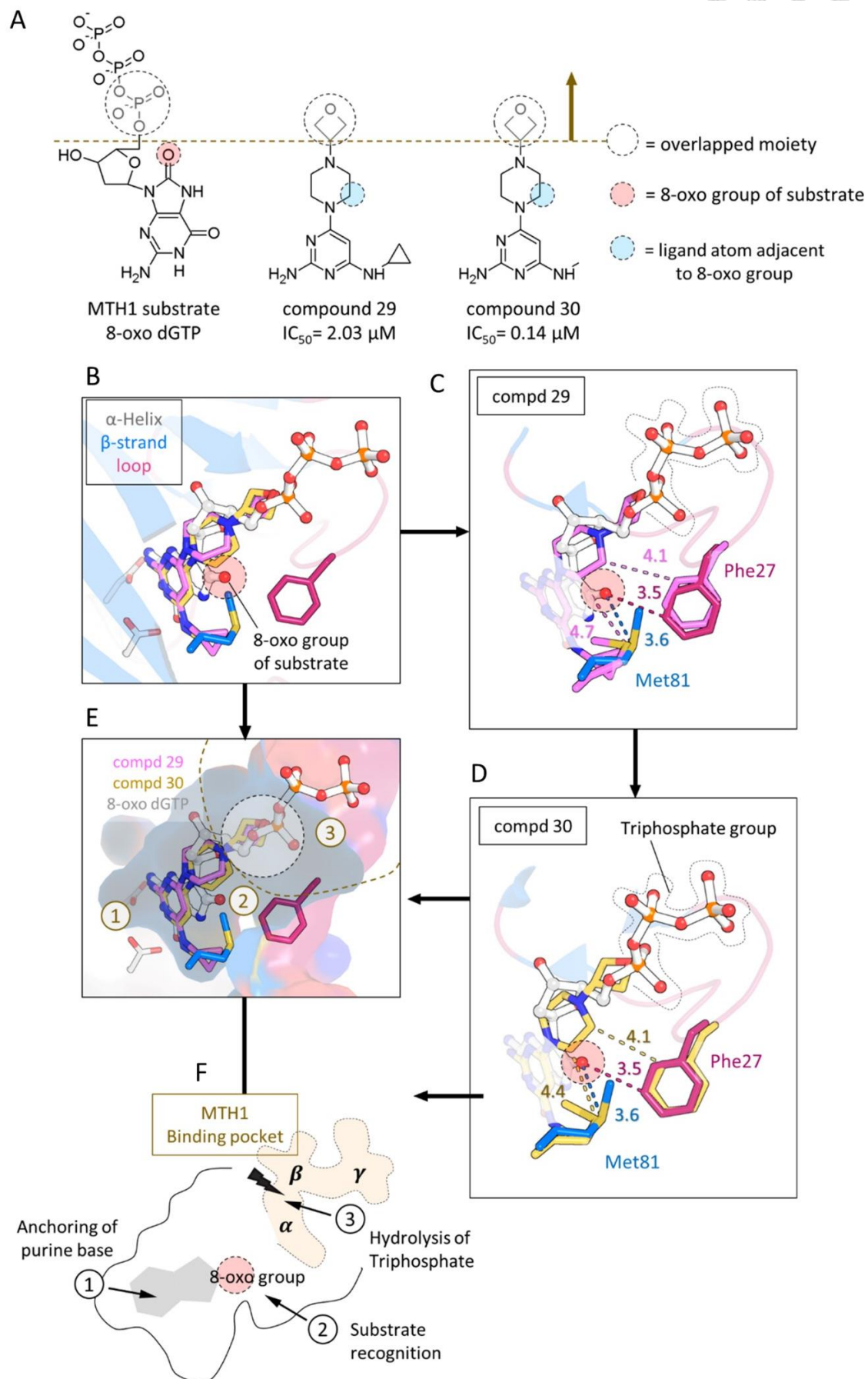


Figure 12. Mutual alignment of MTH1 complex structures bound with substrate, compound 29 and 30

(A) Comparison of structures of MTH1 substrate 8-oxo dGTP (PDB ID: 5FNI), compound 29 and 30 with overlapped region marked in circle. Dashed line in brown separates the catalytic site for triphosphate group of 8-oxo dGTP from other subregions in the binding pocket. Aside from triphosphate group to be hydrolyzed, 8-oxo group of substrate which has been proposed to be involved in recognition of substrate by MTH1 is highlighted in circle colored in red while ligand atom adjacent to 8-oxo group is notified with circle in blue.

(B) Superimposition of MTH1 complex structures bound with MTH1 substrate 8-oxo dGTP, compound 29 and 30 with 8-oxo group of substrate highlighted in circle colored in red.

(C-D) Distance between side chains of key residue Phe27, Met81 and 8-oxo group of substrate in comparison to distance between side chains of key residue Phe27, Met81 and atoms of ring substituent incorporated in R2-position of (C) compound 29 (shown in stick representation colored in purple) and (D) compound 30 (shown in stick representation colored in yellow).

(E) MTH1 binding pocket shown in surface representation, dashed line indicates the catalytic site for hydrolysis of triphosphate group. ①, ②, ③ stand for subregions assigned according to the positioning of purine base, deoxyribose and triphosphate group from substrate 8-oxo dGTP.

(F) Schematic representation of MTH1 binding pocket indicated with labels ①, ②, ③ which are in coherence with labeling from (e).

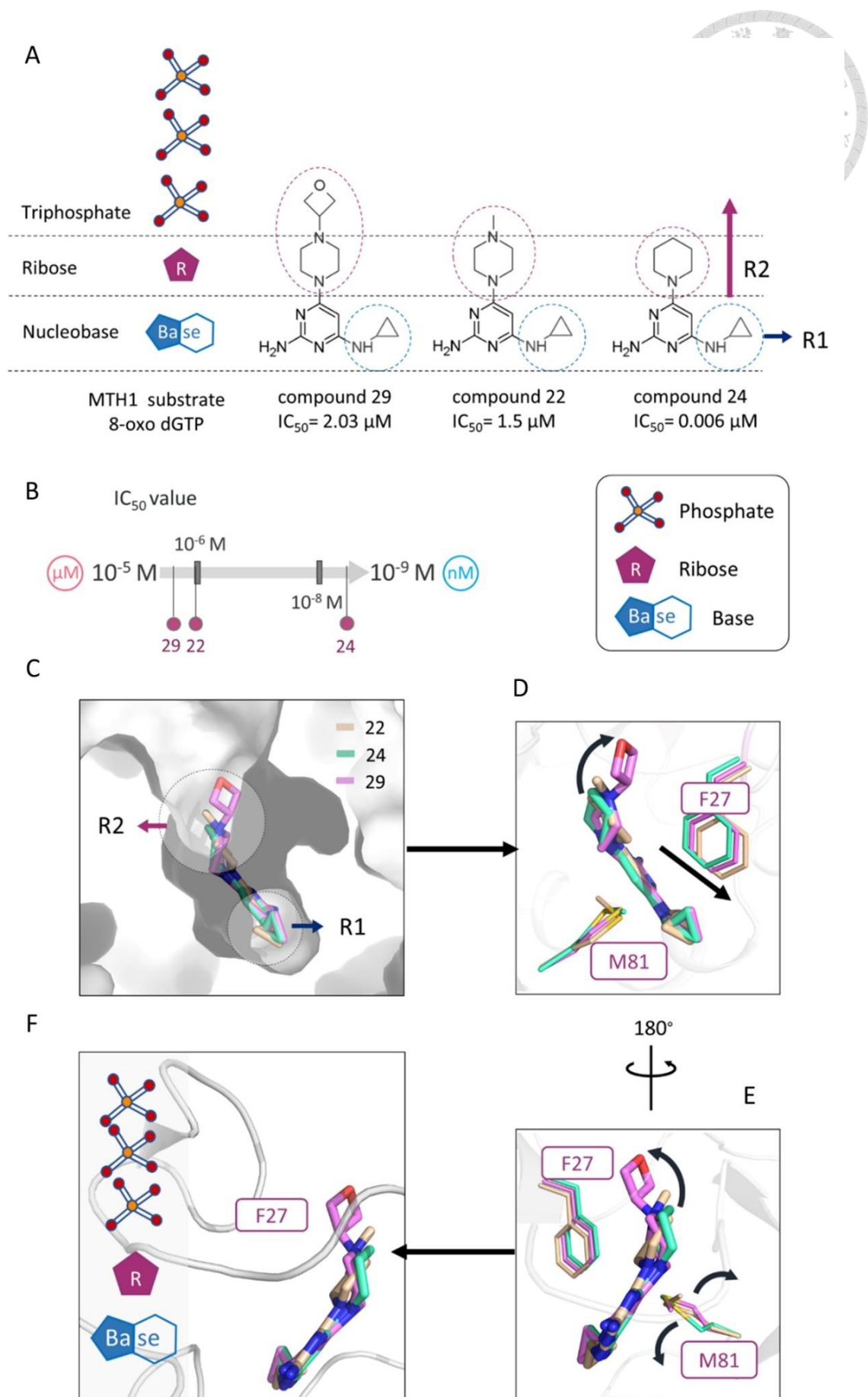


Figure 13. Binding mode of compound 22, 24, 29 in MTH1 binding pocket and structural rearrangement of side chains from Phe27 and Met81

- (A) Structure and affinity profiles of compound 22, 24, 29 with position corresponding to MTH1 substrate 8-oxo dGTP (composed of nucleobase, ribose and triphosphate group).
- (B) IC₅₀ value of compound 22, 24 and 29 is arranged on scale bar according to their affinity toward MTH1.
- (C) MTH1 binding pocket (shown in surface representation) bound with ligand (superimposition of compound 22, 24 and 29)
- (D-E) key residues Phe27 and Met81 are relatively mobile upon binding of compound 22, 24 and 29.
- (F) Flexible loop on which Phe27 locates, position of ligand binding site is indicated with substrate 8-oxo dGTP which occupies MTH1 binding pocket with triphosphate group sitting near the entry.

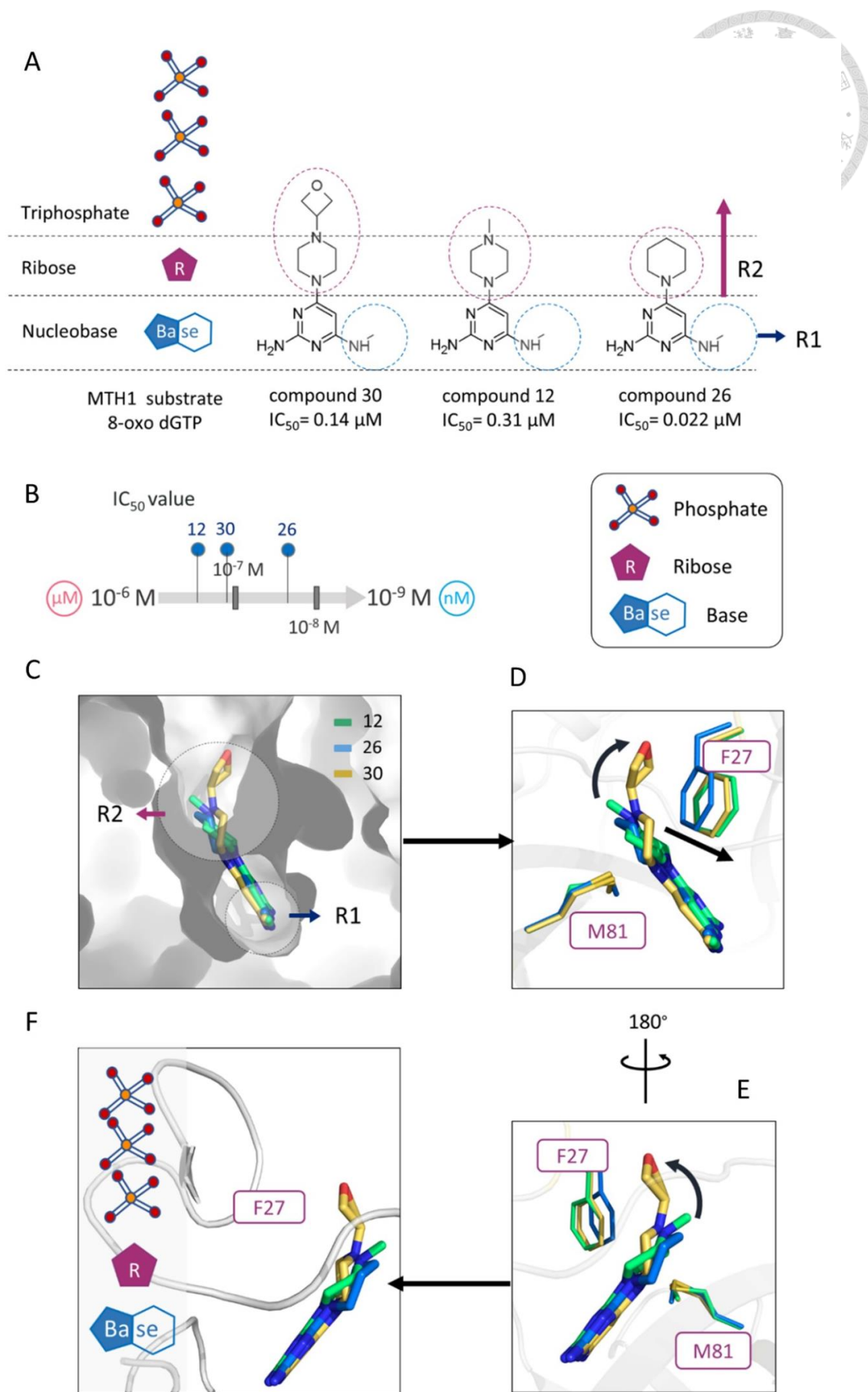


Figure 14. Binding mode of compound 12, 26, 30 in MTH1 binding pocket and structural rearrangement of side chains from Phe27 and Met81

- (A) Structure and affinity profiles of compound 12, 26, 30 with position corresponding to MTH1 substrate 8-oxo dGTP (composed of nucleobase, ribose and triphosphate group).
- (B) IC₅₀ value of compound 12, 26 and 30 is arranged on scale bar according to their affinity toward MTH1.
- (C) MTH1 binding pocket (shown in surface representation) bound with ligand (superimposition of compound 12, 26 and 30)
- (D-E) key residue Phe27 is relatively mobile upon binding of compound 12, 26 and 30, while structural rearrangement is not observed in the side chain of Met81.
- (F) Flexible loop on which Phe27 locates, position of ligand binding site is indicated with substrate 8-oxo dGTP which occupies MTH1 binding pocket with triphosphate group sitting near the entry.

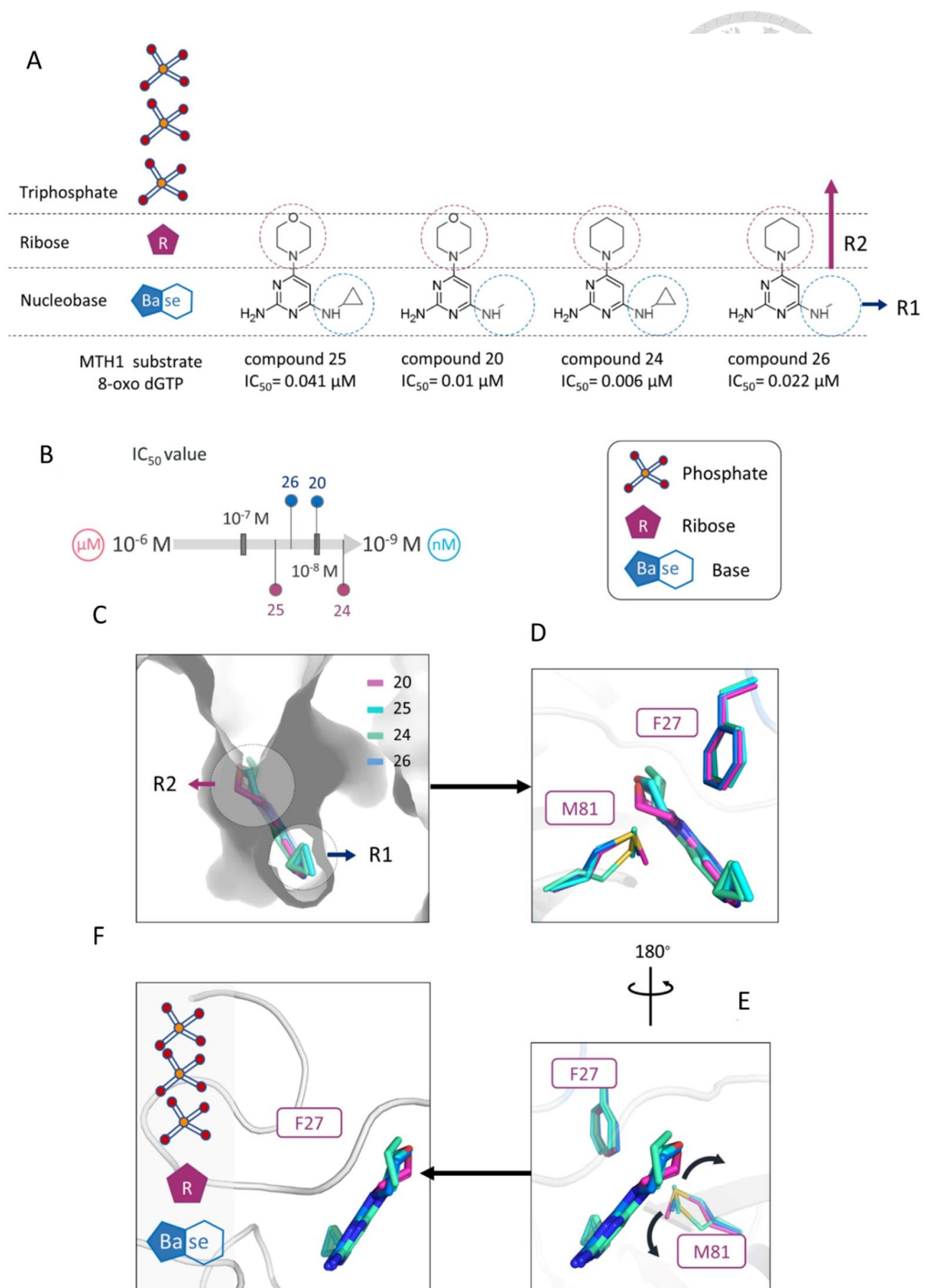


Figure 15. Binding mode of compound 20, 24, 25, 26 in MTH1 binding pocket and structural rearrangement of side chains from Phe27 and Met81

(A) Structure and affinity profiles of compound 20, 24, 25, 26 with position corresponding to MTH1 substrate 8-oxo dGTP (composed of nucleobase, ribose and triphosphate group).

(B) IC₅₀ value of compound 20, 24, 25, 26 is arranged on scale bar according to their affinity toward MTH1.

(C) MTH1 binding pocket (shown in surface representation) bound with ligand (superimposition of compound 20, 24, 25, 26)

(D-E) key residue Met81 is relatively mobile upon binding of compound 20, 24, 25, 26, while structural rearrangement is not observed in the side chain of Phe27.

(F) Flexible loop on which Phe27 locates, position of ligand binding site is indicated with substrate 8-oxo dGTP which occupies MTH1 binding pocket with triphosphate group sitting near the entry.

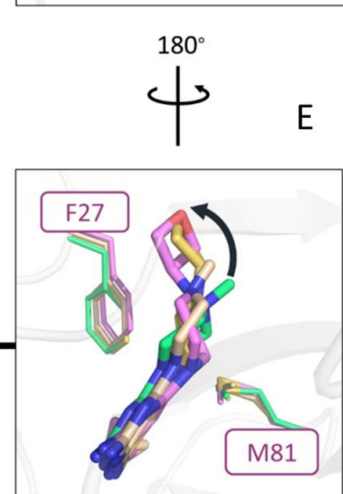
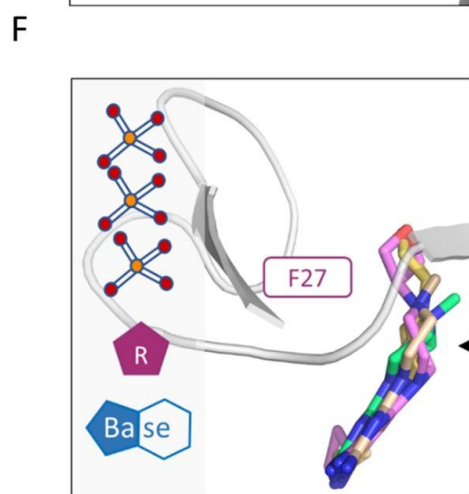
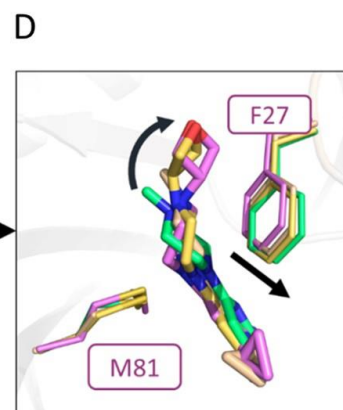
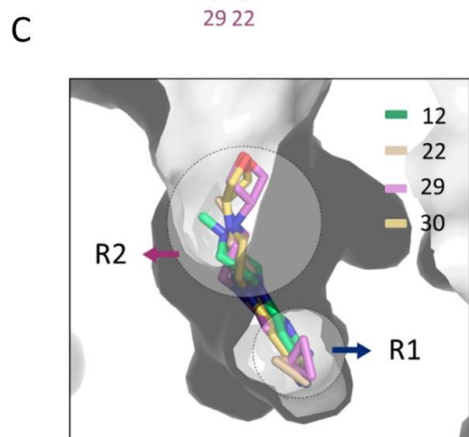
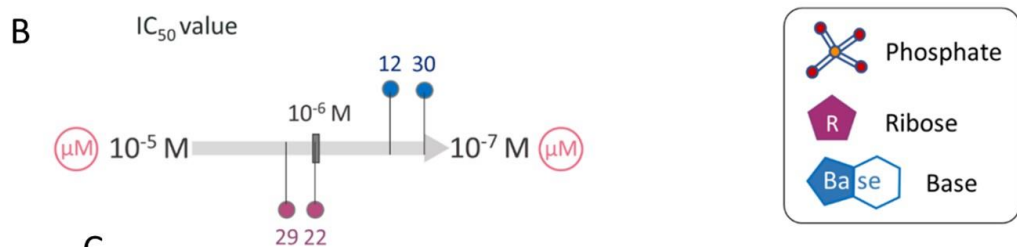
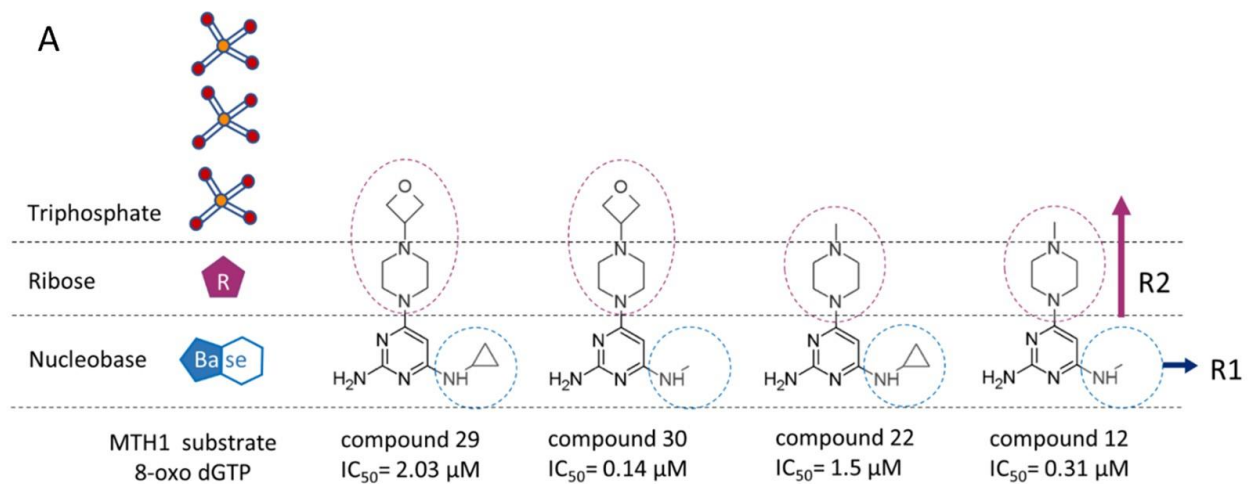


Figure 16. Binding mode of compound 12, 22, 29, 30 in MTH1 binding pocket and structural rearrangement of side chains from Phe27 and Met81

(A) Structure and affinity profiles of compound 12, 22, 29, 30 with position corresponding to MTH1 substrate 8-oxo dGTP (composed of nucleobase, ribose and triphosphate group).

(B) IC₅₀ value of compound 12, 22, 29, 30 is arranged on scale bar according to their affinity toward MTH1.

(C) MTH1 binding pocket (shown in surface representation) bound with ligand (superimposition of compound 12, 22, 29, 30)

(D-E) key residue Phe27 is relatively mobile upon binding of compound 12, 22, 29, 30, while structural rearrangement is not observed in the side chain of Met81.

(F) Flexible loop on which Phe27 locates, position of ligand binding site is indicated with substrate 8-oxo dGTP which occupies MTH1 binding pocket with triphosphate group sitting near the entry.

Appendices



Appendix 1. Composition of buffer in SDS-PAGE analysis

(i) 5X Running buffer

Reagent	Stock	volume/weight
Tris base	Powder	15.1 g
Glycine	Powder	72.0 g
SDS	Powder	5.0 g

Dissolved in 1 liter of ddH₂O

(ii) 5X SDS sample loading dye

Reagent	Stock	Final concentration
Tris-HCl (pH 6.8)	1M	250 mM
SDS	Powder	10%
Glycerol	100%	50%
Bromophenol blue	Powder	0.5%
β-mercaptoethanol	100%	500 mM

(iii) SDS-PAGE gels

Stacking gel (5%)

Reagent	Stock	Volume*
ddH ₂ O	-	4.1 ml
Polyacrylamide/Bis	30%	1.0 ml
Tris-HCl (pH 6.8)	1.0 M	750 μl
SDS	10%	60 μl
Ammonium Persulfate (APS)	10%	60 μl
TEMED	100%	6 μl

* 2 gel recipe for 5% stacking gel



Resolving gel (12%)

Reagent	Stock	Volume*
ddH ₂ O	-	3.96 ml
Polyacrylamide/Bis	30%	4.8 ml
Tris-HCl (pH 8.8)	1.5 M	3.0 ml
SDS	10%	120 µl
Ammonium Persulfate (APS)	10%	120 µl
TEMED	100%	4.8 µl

* 2 gel recipe for 12% resolving gel

Resolving gel (15%)

Reagent	Stock	Volume*
ddH ₂ O	-	2.76 ml
Polyacrylamide/Bis	30%	6.0 ml
Tris-HCl (pH 8.8)	1.5 M	3.0 ml
SDS	10%	120 µl
Ammonium Persulfate (APS)	10%	120 µl
TEMED	100%	4.8 µl

* 2 gel recipe for 15% resolving gel

Coomassie blue stain

Reagent	volume/weight
Coomassie brilliant blue R-250 (GE healthcare)	5 table
Methanol	600 ml
Acetic	200 ml

Add RO water to final volume of 2 L

Destaining buffer

Reagent	Volume
Methanol	600 ml
Acetic	200 ml

Add RO water to final volume of 1 L

Appendix 2. Composition of buffer used in affinity chromatography



(i) Lysis buffer

Reagent	Stock	Volume/weight	Final concentration
Tris-HCl (pH 8.0)	1.0 M	4 ml	20 mM
NaCl	Powder	5.844 g	500 mM
Glycerol	100%	20 ml	10 %
Imidazole	Powder	0.136 g	10 mM
TCEP	Powder	0.114 g	2 mM

Dissolved in 0.2 liter of ddH₂O with pH adjusted to 7.4

(ii) Binding buffer

Reagent	Stock	Volume/weight	Final concentration
Tris-HCl (pH 8.0)	1.0 M	20 ml	20 mM
NaCl	Powder	29.22 g	500 mM
Glycerol	100%	100 ml	10 %
Imidazole	Powder	0.68 g	10 mM

Dissolved in 1 liter of ddH₂O with pH adjusted to 7.4

(iii) Elution buffer

Reagent	Stock	Volume/weight	Final concentration
Tris-HCl (pH 8.0)	1.0 M	10 ml	20 mM
NaCl	Powder	14.61 g	500 mM
Glycerol	100%	50 ml	10 %
Imidazole	Powder	17.02 g	500 mM

Dissolved in 0.5 liter of ddH₂O with pH adjusted to 7.4

Appendix 3. Composition of buffer used in size-exclusion chromatography

(i) Gel filtration buffer

Reagent	Stock	Volume/weight	Final concentration
Tris-HCl (pH 8.0)	1.0 M	20 ml	20 mM
NaCl	Powder	8.76 g	150 mM
Glycerol	100%	50 ml	5 %
TCEP	Powder	0.57 g	2 mM

Dissolved in 1 liter of ddH₂O with pH adjusted to 7.4

Appendix 4. Composition of buffer for storage of purified protein sample

(i) Storage buffer

Reagent	Stock	Volume/weight	Final concentration
HEPES	Powder	2.3 g	20 mM
NaCl	Powder	8.76 g	300 mM
Glycerol	100%	50 ml	10 %
TCEP	Powder	0.125 g	1 mM

Dissolved in 0.5 liter of ddH₂O with pH adjusted to 7.5



Appendix 5. Composition of buffer for MTH1 activity assay

(i) MTH1 Activity buffer ^a

Reagent	Stock	Final concentration
Tris-acetate	1M	100 mM
NaCl	Powder	40 mM
Mg acetate	Powder	10 mM
DTT	1M	1 mM
Tween-20	100%	0.005%

Dissolved in 0.5 liter of ddH₂O with pH adjusted to 7.5

^abuffer recipe from Gad et al. (2014)



Appendix 6. Composition of buffer for iTC study

(i) MTH1 iTC buffer ^a

Reagent	Stock	Final concentration
Tris-acetate	1M	100 mM
NaCl	Powder	40 mM
MgCl ₂	Powder	10 mM
DTT	1M	1 mM
Tween-20	100%	0.005%

Dissolved in 0.5 liter of ddH₂O with pH adjusted to 7.5

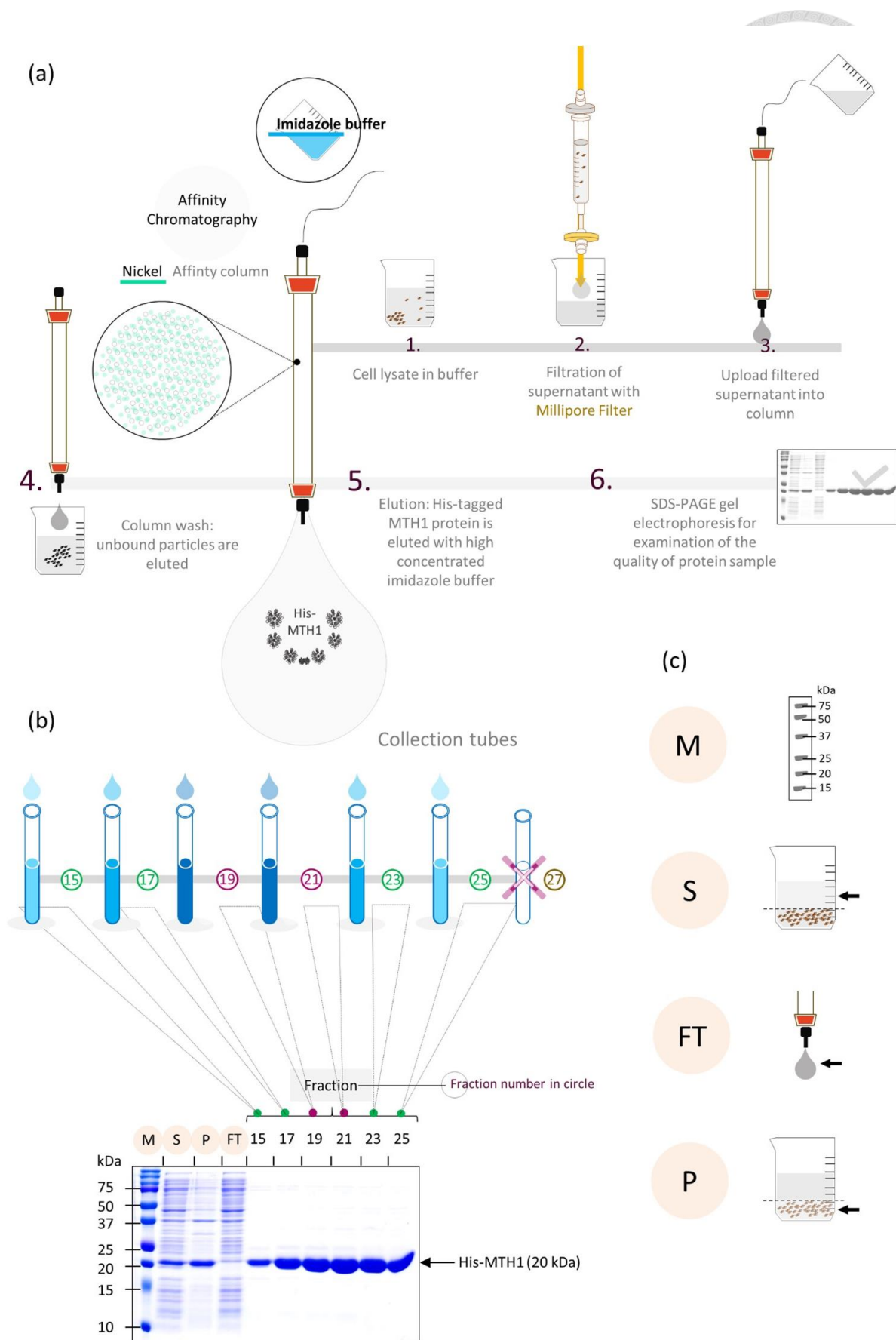
^a buffer recipe from Gad et al. (2014)

(ii) MTH1 iTC buffer ^b

Reagent	Stock	Final concentration
Tris-acetate	1M	100 mM
NaCl	Powder	40 mM
MgCl ₂	Powder	10 mM
TCEP	Powder	1 mM
Tween-20	100%	0.005%

Dissolved in 0.5 liter of ddH₂O with pH adjusted to 7.5

^b buffer recipe from Gad et al. (2014) modified



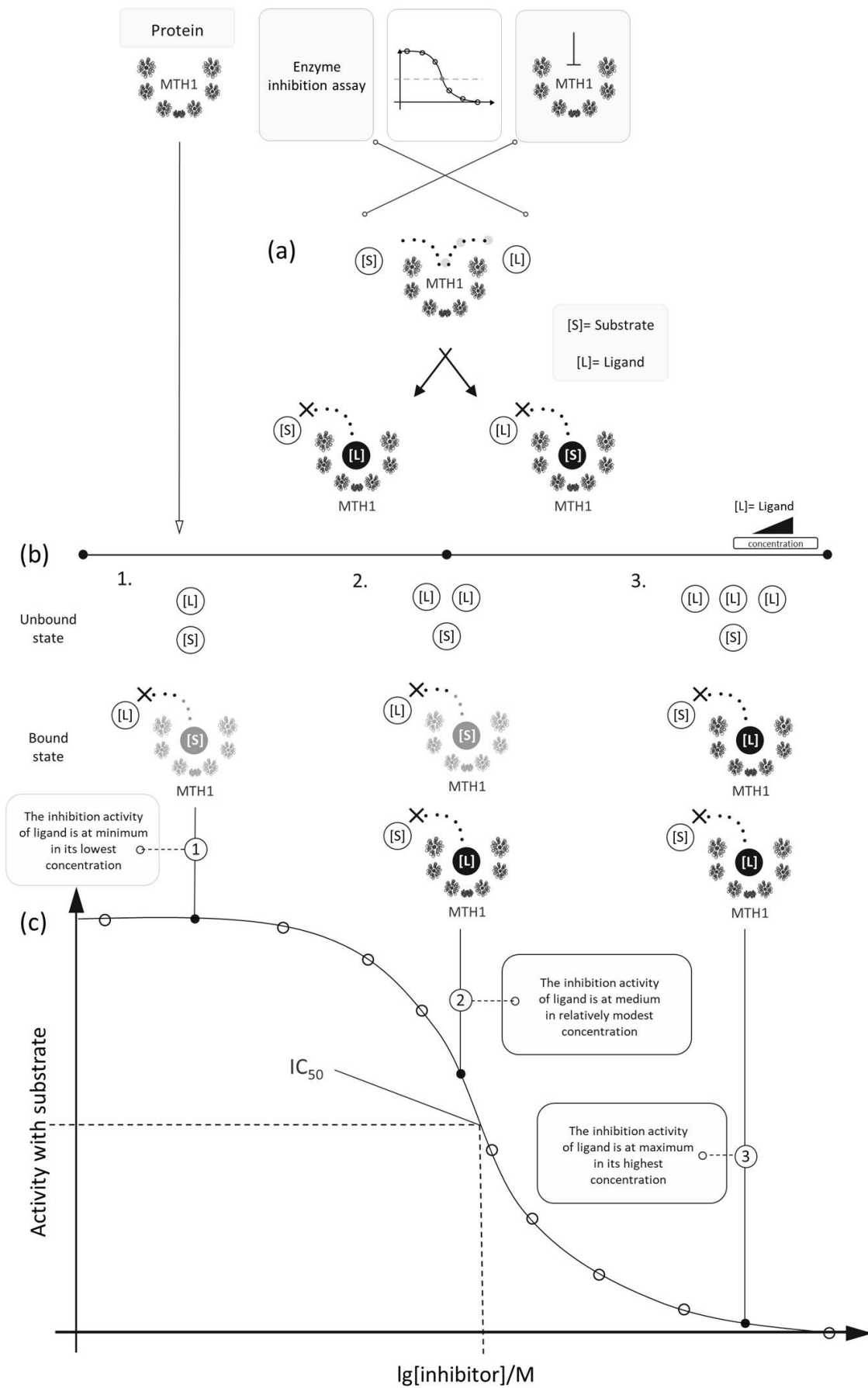
Appendix 7. Workflow of protein purification

(a) Overview of purification procedure from separation of cell lysate to inspection of purified protein on SDS-PAGE gel.

(b) An array of collection tubes is illustrated with fraction number given next to each collection tube. For confirmation of the purity of protein sample contained in collection tubes, protein solution from fraction 15-25 is further checked on SDS-PAGE gel. Sample solution including S= supernatant, P= pellet, FT= Flow-through collected from primary stage of purification is also checked to evaluate the efficiency of protein purification.

(c) Schematic representation of M=marker, S= supernatant, FT= Flow-through, P= pellet.



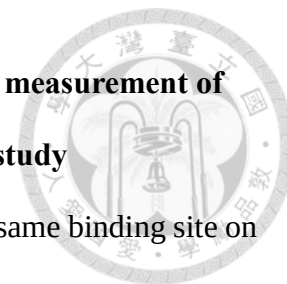


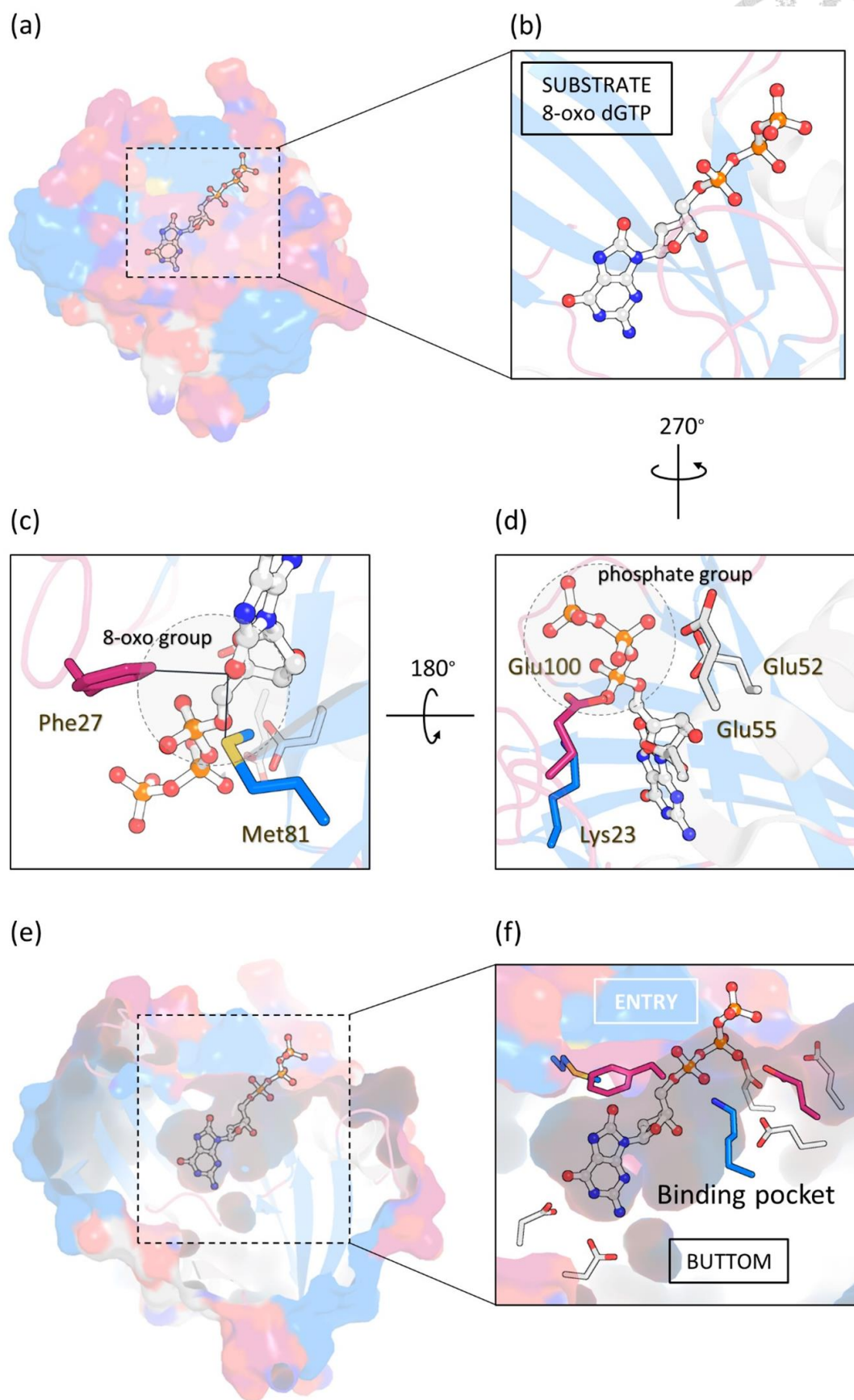
Appendix 8. Experimental details of enzyme inhibition assay for measurement of IC_{50} value of 2-aminopyrimidine-based compound series in this study

(a) Schematic illustration of substrate and ligand competing for the same binding site on MTH1 protein.

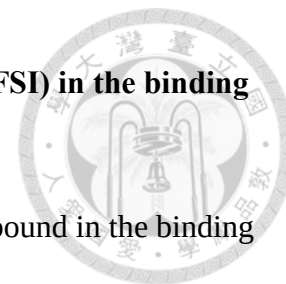
(b) Schematic representation of experimental details of data points from dose-response curve illustrated in (c). Inhibitory activity of ligand to be analyzed is measured in buffer solution containing fixed amount of MTH1 protein and substrate 8-oxo dGTP with serially diluted ligand solution added.

(c) Dose-response curve fitted using non-linear regression analysis. Data points colored in black are elaborated for experimental details in text box placed nearby. The IC_{50} value which stands for concentration of inhibitor as the activity of target enzyme is half-inhibited is therefore calculated from the fitting curve.





Appendix 9. Binding mode of substrate 8-oxo dGTP (PDB ID: 5FSI) in the binding pocket of MTH1



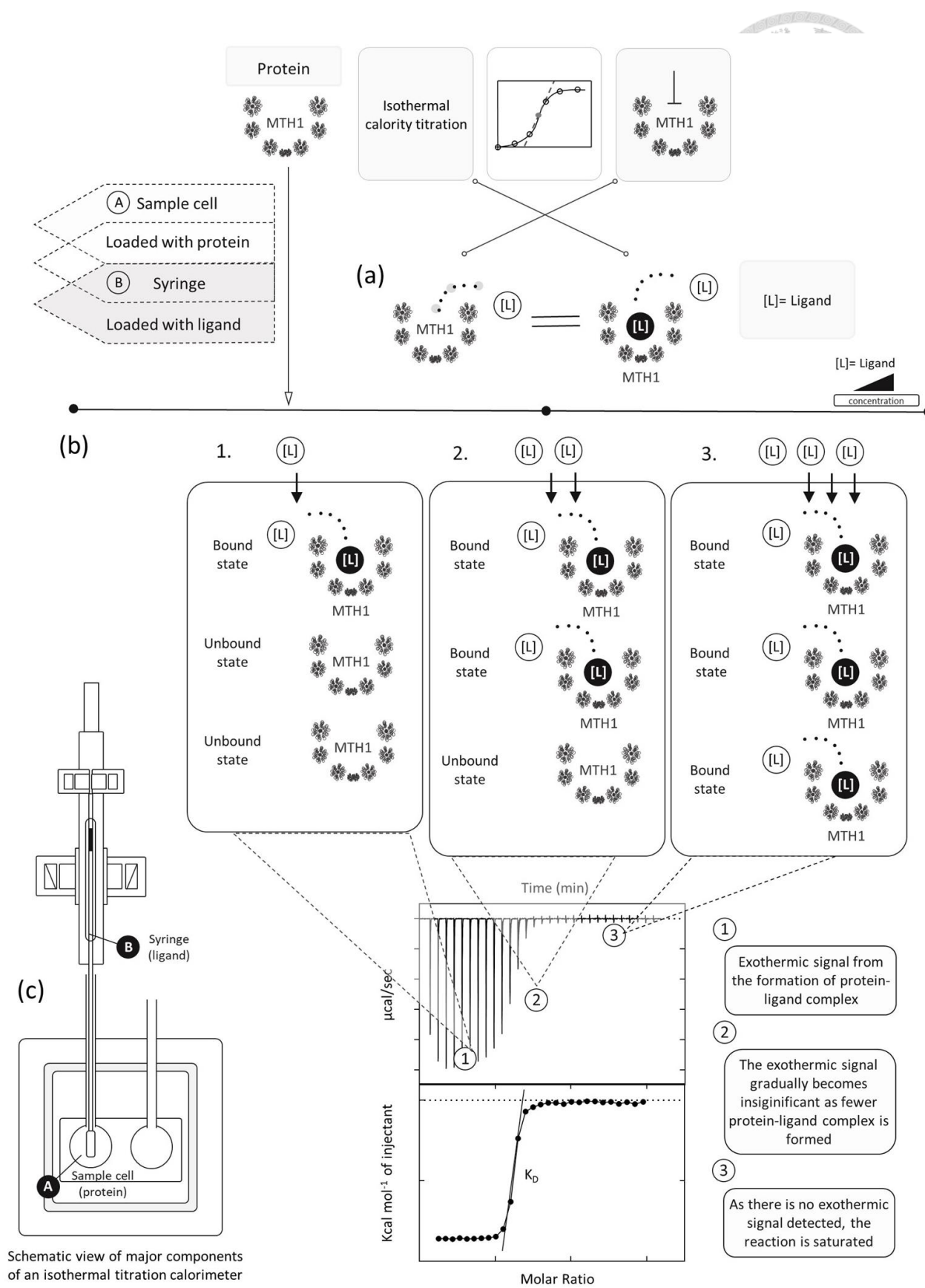
(a) Substrate 8-oxo dGTP (shown in ball and stick representation) bound in the binding pocket of MTH1 (shown in surface representation).

(b) Close-up view of 8-oxo dGTP in the binding pocket. (MTH1 protein is displayed in cartoon representation as background. colors are assigned to secondary structure which is composed of α -helix (colored grey), β -strand (colored blue) and loop (colored magenta).

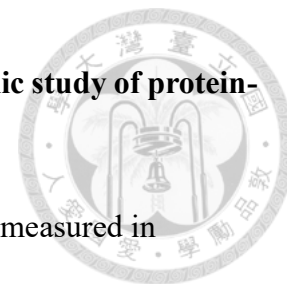
(c) Key residues Phe27 and Met81 which are positioned in close distance to 8-oxo group of 8-oxo-dGTP, are regarded as critical element in recognition of substrate by MTH1.

(d) Catalytic site for hydrolysis of triphosphate group of 8-oxo dGTP is made up of Glu cluster which coordinates with water molecule network via metal ions. Key residues Lys23, Glu52, Glu55 and Glu100 are shown in stick representation.

(e-f) Insight into binding mode of substrate 8-oxo dGTP in MTH1 binding pocket. The purine base is bound firmly in the bottom through formation of h-bonds with side chains of Asn33, Asp119 and Asp120. Opposite to the purine base, triphosphate group of 8-oxo dGTP is placed in the catalytic site adjacent to the entry of binding pocket.



Appendix 10. Dissection of mechanism underlying thermodynamic study of protein-ligand formation using isothermal titration calorimetry



- (a) In the experiment, binding affinity between protein and ligand is measured in the absence of substrate.
- (b) The dissociation constant (K_D) is measured through titration of ligand solution into protein diluted in buffer which is loaded in sample cell labeled as B in (c). The gradually increasing concentration of ligand in the sample cell containing fixed amount of protein lead eventually to saturation of the reaction which is elaborated in detail in schematic representation above the typical isotherm from isothermal titration calorimetry experiment.
- (c) Schematic illustration of an iTC machine with syringe (in which ligand solution is usually loaded) and sample cell (in which protein sample is usually loaded) labeled as A and B, respectively.