



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

약학 박사 학위논문

**Structure-based discovery of potential
inhibitors against two catalytic enzymes as
novel drug targets: human NSDHL and
PptT from *Mycobacterium tuberculosis***

**새로운 기전의 약물 타겟으로 알려진 NSDHL 및
PptT의 구조에 기반한 저해제 개발 연구**

2020년 8월

서울대학교 대학원
약학과 물리약학전공
김 동 균

Abstract

Structure-based discovery of potential inhibitors against two catalytic enzymes as novel drug targets: human NSDHL and PptT from *Mycobacterium tuberculosis*

Dong-Gyun Kim

College of Pharmacy

Seoul National University

Dissertation director: Professor Bong-Jin Lee

Structure-based drug design (SBDD) is one of the most efficient techniques to accelerate drug development and make it more cost-effective. The determination of the three dimensional structure of the protein facilitates the development of well-optimized compounds based on the structural information. For development of pharmacological agents for targeting novel enzymes using SBDD, we focused on two catalytic enzymes: (i) NAD⁺-dependent steroid dehydrogenase-like (NSDHL) from *Homo sapiens* and (ii) 4'-phosphopantetheinyl transferase (PptT) from *Mycobacterium tuberculosis*. These two enzymes have been reported as novel drug targets due to their essential function in biological metabolism of human and *M. tuberculosis*, respectively. NSDHL is an essential

enzyme in human cholesterol synthesis and a regulator of epidermal growth factor receptor (EGFR) trafficking pathways, has attracted interest as a therapeutic target due to its crucial relevance to cholesterol-related diseases and carcinomas. PptT is a crucial enzyme in intracellular survival and persistence of tuberculosis and covalently transfer the phosphopantethein moiety to a serine residue of carrier protein domain. In this study, we reported X-ray crystal structures of NSDHL, PptT and MSMEG_PPTase, which revealed a detailed description of the coenzyme-binding site and the binding mode. A structure-based virtual screening and biochemical evaluation were performed and identified novel inhibitors for NSDHL and PptT, respectively. Furthermore, further cell-based validation on the inhibitory activity revealed that our inhibitors were rationally developed. Overall, these findings could serve as good platforms for the development of therapeutic agents against NSDHL-related diseases and tuberculosis.

Keywords: Structure-based drug discovery, *Homo sapiens*, *Mycobacterium tuberculosis*, cholesterol synthesis pathway, virtual screening, docking simulation, enzyme inhibitors, X-ray crystallography

Student number: 2014-21973

Contents

Abstract	ii
Contents	iv
List of tables	vii
List of figures	viii
Chapter 1. Structure-based discovery of novel inhibitors against human NSDHL with the potential to suppress EGFR activity	1
1.1 Introduction	1
1.2 Materials and Methods	4
1.2.1 Gene cloning, protein expression, and purification.....	4
1.2.2 Crystallization and X-ray data collection	7
1.2.3 Structure determination, refinement, and analysis	7
1.2.4 Isothermal titration calorimetry (ITC).....	8
1.2.5 Size-exclusion chromatography with multiangle light scattering (SEC-MALS).....	9
1.2.6 Thermal shift assay (TSA).....	9
1.2.7 High-throughput virtual screening	10
1.2.8 NADH-based competitive binding assay for identifying inhibitors.....	11
1.2.9 Synthetic methods and characterization of a small molecule	12
1.2.10 Liquid Chromatography Mass Spectrometry (LCMS)	16
1.2.11 Kinetic solubility	17
1.2.12 Molecular docking.....	17

1.2.13 Surface plasmon resonance (SPR).....	18
1.2.14 Cellular viability assay	19
1.2.15 Flow cytometry	20
1.2.16 Immunoblotting.....	21
1.2.17 Public data analysis	23
1.2.18 Statistical analysis	23
1.2.19 Data availability	23
1.3 Results	24
1.3.1 Overall structures of human NSDHL	24
1.3.2 Diverse features of human NSDHL structures.....	31
1.3.3 NAD ⁺ binding site in NSDHL	35
1.3.4 Conformational change induced by the binding of NAD ⁺ allows the binding of a sterol precursor to NSDHL	38
1.3.5 Human NSDHL favorably employs NAD(H) as its coenzyme in the cholesterol synthesis pathway	43
1.3.6 Development of NSDHL inhibitors based on the structural information	46
1.3.7 Proposed binding mode between compound 9 and NSDHL.....	54
1.3.8 Therapeutic potential of NSDHL inhibition in EGFR-driven cancer	57
1.3.9 The NSDHL inhibitor accelerates EGFR degradation to suppress EGFR-dependent signaling	63
1.4 Discussion.....	66
Chapter 2. Structure-based discovery of selective inhibitors against PptT from <i>M. tuberculosis</i>	70

2.1 Introduction	70
2.2 Materials and Methods.....	72
2.2.1 Gene Cloning, protein expression, and purification	72
2.2.2 Crystallization and X-ray data collection	73
2.2.3 Structure determination, refinement, and analysis	75
2.2.4 High throughput virtual screening.....	75
2.2.5 BpsA-PptT coupled assay for searching inhibitors	76
2.2.6 2. Cytotoxicity in eukaryotic cell	77
2.2.7 Bactericidal activity in mycobacteria-infected murine macrophages.....	77
2.3 Results	79
2.3.1 Overall structures of mycobacteria PPTases	79
2.3.2 Structural features and active sites of mycobacteria PPTases	82
2.3.3 Development of small molecule inhibitors to suppress PptT activity.....	86
2.3.4 Identification of the selective scaffold inhibiting PptT from <i>M. tuberculosis</i>	92
2.3.5 PptT inhibitor reveals the bactericidal activity in mycobacteria-infected macrophages	94
2.4 Discussion.....	96
References	98
국문초록	108

List of tables

- Table 1.** Primers used in this study
- Table 2.** Statistics for data collection and model refinement
- Table 3.** Chemical structures and inhibitory potentials of NSDHL inhibitors
- Table 4.** Chemical structures and inhibitory potentials of PptT inhibitors

List of figures

- Figure 1.** Original images of immunoblots
- Figure 2.** Overall structures of the apo and holo forms of NSDHL
- Figure 3.** Conformation of the NSDHL structure and a topology diagram
- Figure 4.** A structural comparison of NSDHL and its homologs
- Figure 5.** Distinct dimeric interface of the NSDHL structure
- Figure 6.** NAD⁺ binding site and the unique conformational change in NSDHL induced by the binding of NAD⁺
- Figure 7.** B-factor distributions of two NSDHL structures
- Figure 8.** Thermal stability of NSDHL and its mutants
- Figure 9.** Gly205 and Lys232 in the NSDHL structure
- Figure 10.** ITC measurements for the coenzyme preference of NSDHL
- Figure 11.** Two parent hits selected from an initial set of 495 compounds
- Figure 12.** Inhibition curves of 11 analogs derived from the initial hits
- Figure 13.** The most potent inhibitor of NSDHL and the proposed docking model
- Figure 14.** TSA analysis of wild-type NSDHL and G47S and C104T mutants
- Figure 15.** Cytotoxicity of compound **9** in various types of cancer cell lines
- Figure 16.** Enhanced antitumor effect of erlotinib with the NSDHL inhibitor and suppression of EGFR activity by compound **9**

- Figure 17.** Genetic aberration of *NSDHL* and its clinical impact in EGFR-driven cancers
- Figure 18.** Proposed mode of action of compound **9** in the regulation of EGFR signaling
- Figure 19.** Overall structure of mycobacteria PPTases
- Figure 20.** Active site of Mtb_PptT
- Figure 21.** Active site of MSMEG_PptT
- Figure 22.** PptT-BpsA coupled assay system
- Figure 23.** Structures and inhibitory potentials of four selected inhibitors
- Figure 24.** MSMEG_PPTase-BpsA coupled assay for 4 selected compounds
- Figure 25.** Biological test of **Cpd1-9**

Chapter 1. Structure-based discovery of novel inhibitors against human NSDHL with the potential to suppress EGFR activity

1.1 Introduction

The cholesterol synthesis pathway is ubiquitous in animal cells and plays a pivotal role in producing cholesterol, which is not only a major component of all plasma membranes but also a precursor of steroid hormones [1]. The biosynthesis of cholesterol begins with acetylated coenzyme A, and cholesterol molecules are finally synthesized in more than 20 reactions, which are complicated and strictly regulated [2]. Dysregulation and imbalances in the cholesterol synthesis pathway have been considered major causes of human diseases, such as hypercholesterolemia, cardiovascular disease, cancer, and neuropathy, which lead to mortality and morbidity worldwide [3].

Several molecules that target enzymes involved in sterol biosynthesis have been developed. However, a majority of the agents are antifungal and target specific stages in the ergosterol synthesis pathway of fungi [4,5]. In addition, these agents are not considered as good medications due to their resistance and detrimental physiological side effects. Statins that inhibit 3-hydroxy-3-methylglutaryl-coenzyme A reductase (HMGCR) are the most successful drugs for treating hypercholesterolemia and cardiovascular diseases [6]. Nevertheless, it has been reported that the prolonged use of statins is strongly associated with side effects [7]. Thus,

the discovery of novel drugs that regulate the cholesterol synthesis pathway is needed to counter cholesterol-related diseases in a specific manner.

Enzymes, including HMGCR, that act upstream of the cholesterol synthesis pathway have been widely investigated, whereas many enzymes that participate in the downstream pathway from lanosterol to cholesterol have not been thoroughly investigated. NAD(P)-dependent steroid dehydrogenase-like (NSDHL, sterol-4- α -carboxylate 3-dehydrogenase, decarboxylating), which is an essential human enzyme in the downstream cholesterol synthesis pathway, catalyzes NAD⁺-dependent oxidative decarboxylation of the C4 methyl groups of 4 α -carboxysterols [8-10]. NSDHL is localized in the endoplasmic reticulum (ER) membrane, on which the enzymatic reaction occurs, and in lipid droplets, which are ER-derived cytoplasmic structures for storing lipids and cholesterols [11]. In addition to its association with cholesterol biosynthesis, this enzyme is also strongly related to cancer growth and the signaling of proto-oncogenes [12,13]. In particular, it has been reported that NSDHL regulates the expression and signaling of epidermal growth factor receptor (EGFR), and the loss of *NSDHL* gene expression sensitizes cancer cells to EGFR-targeting inhibitors [14]. Moreover, the accumulation of sterol metabolites as a result of NSDHL deficiency has been shown to induce the suppression of tumor growth [15]. Despite the multiple important roles of NSDHL in cholesterol-associated diseases, its specific inhibitors have not yet been developed because of a lack of structural information on human NSDHL protein.

This study aimed to elucidate the crystal structures of NSDHL and identify a structure-based inhibitor of NSDHL for novel and effective treatment against cholesterol imbalance and cancers. We described two crystal structures of NSDHL on the atomic level and identified key information regarding the active site of the enzyme, which was used to guide the discovery of the inhibitors. The inhibitory profiles of 12 analogs were determined through a competitive

inhibitor assay, and compound **9**, which exhibited a half-maximal inhibitory concentration (IC_{50}) of approximately 8 μ M, was found to be the most effective inhibitor. Furthermore, we verified that compound **9** altered EGFR protein turnover and suppressed EGFR signaling cascades, thus leading to sensitization to erlotinib, an EGFR kinase inhibitor, in both erlotinib-sensitive and erlotinib-resistant cancer cell lines. Taken together, these findings reveal considerable insights on the structure of human NSDHL that can lead to develop potent and novel inhibitors reducing EGFR activity and thus provide novel therapeutic strategies against EGFR-driven cancers.

1.2 Materials and Methods

1.2.1 Gene cloning, protein expression, and purification

The gene encoding NSDHL from *Homo sapiens* was amplified by polymerase chain reaction (PCR) using a clone provided by Korea Human Gene Bank (Medical Genomics Research Center, KRIBB, Korea) (Clone ID: hMU006554) as a template. The PCR primers used for the NSDHL crystallization and mutation studies are shown in Table 1. For the experiments, a truncated construct (31–267) of the full sequence (1–373) was employed. For the wild-type and two mutants (G205S and K232Δ) of NSDHL sequences (31–267), the applied procedures, from cloning to purification, were nearly identical. The PCR product and pET21a vector (Merck Millipore, Germany) were digested with both *NdeI* and *XhoI* (NEB, UK) and ligated. After verifying the DNA sequences, the recombinant protein was overexpressed in *E. coli* C41(DE3) (Sigma Aldrich, USA) and grown at 37°C in Luria-Bertani (LB) broth. When the cells reached an OD₆₀₀ of 0.5, 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG) was added to induce protein expression. Then, the cells were transferred to a 15°C incubator and grown for an additional 20 hours. The cells were harvested by centrifugation at 4,300 ×g for 10 minutes. The cell pellet was resuspended in lysis buffer (50 mM Tris, 500 mM NaCl, (pH 7.8), 5 mM imidazole, 0.5 mM TCEP, 10% (v/v) glycerol) and then sonicated at 4°C. The lysate was centrifuged at 18,000 ×g for 60 minutes at 4°C, and the supernatant was loaded into either an open Ni-NTA column (Qiagen, USA) or a HiTrap Chelating HP column (GE Healthcare) that was pre-equilibrated with lysis buffer. The column was washed with a 30-fold excess volume of buffer containing 30 mM imidazole. The protein was eluted at an imidazole concentration range of 100–250 mM. After concentrating the collected fractions, the buffer was changed to buffer A (40 mM HEPES, 150 mM NaCl, (pH 8.0), 0.5 mM TCEP, 5% (v/v) glycerol) by

dialysis using an Amicon Ultra-15 Centrifugal Filter Unit with a 10k molecular weight cut-off (Millipore, USA). The protein was further purified by gel filtration on a HiLoad 16/600 Superdex 200 prep-grade column (GE Healthcare). Crystallization of coenzyme-free NSDHL (NSDHL_{apo}) was carried out with a protein solution of 0.26 mM. To obtain the crystals of NAD(H)-bound NSDHL, 1 mM NAD⁺ or NADH was added to the final protein solution prior to crystallization. The NAD⁺-bound NSDHL crystal (NSDHL_{holo}) was solely obtained in this study.

Table 1. Primers used in this study. “F” represents forward, and “R” represents reverse. Restriction enzyme sites (*NdeI* and *XhoI*) are shown as colored underline (red and blue, respectively) in primer sequences. Mutated nucleotides are shown as black underline in the primer sequences used for site-directed mutagenesis.

Primer ID	Primer sequence (5'→3')
<i>NSDHL</i> -F	GGAATTCCATATGATGGAACCAGCAGTTAGCGAGC
<i>NSDHL</i> -R	CCGCCGCTCGAGGATGTGAAATGCCTTCCCAC
<i>NSDHL</i> G205S-F	CCTCATGGCATTTCAGCCCAAGGGACCCGCAG
<i>NSDHL</i> G205S-R	CTGCGGGTCCCTTGGGCTGAAAATGCCATGAGG
<i>NSDHL</i> K232Δ-F	TTCGTGATTGAAAATGGGAACCTGGTGGACTTCACC
<i>NSDHL</i> K232Δ-R	GGTGAAGTCCACCAAGTCCCATTTCGAATCACGAA
<i>NSDHL</i> G47S-F	GTG ATC GGG GGC TCT AGC TTC CTG GGG CAG CAC
<i>NSDHL</i> G47S-R	GTG CTG CCC CAG GAA GCT AGA GCC CCC GAT CAC
<i>NSDHL</i> C104T-F	GTA AAC ACA GTT TTC CAC ACC GCG TCA CCC CCA CCA TCC
<i>NSDHL</i> C104T-R	GGA TGG TGG GGG TGA CGC GGT GTG GAA AAC TGT GTT TAC

1.2.2 Crystallization and X-ray data collection

Crystallization was performed with manufactured screening kits using the sitting-drop vapor diffusion method. The best NAD⁺-bound NSDHL crystals were obtained in the presence of 1 mM NAD at 20°C by mixing equal volumes (0.3 µl each) of the protein solution (at a concentration of 10 mg/ml in buffer A) and reservoir solution, which consisted of 0.1 M Tris-HCl (pH 8.0), 0.2 M calcium chloride, and 44% (v/v) PEG400. CocrySTALLIZATION with NADH did not generate crystals, whereas many crystal hits were observed when NAD⁺ was used for cocrySTALLIZATION. The crystallization of NSDHL_{apo} (NAD⁺-free NSDHL) was carried out in the absence of NAD⁺ using the sitting-drop vapor diffusion method at 20°C by mixing equal volumes (0.5 µl each) of the protein solution (at a concentration of 10 mg/ml in buffer A) and reservoir solution, which consisted of 0.02 M glycine (pH 10.0), 0.5 M magnesium chloride, 0.02 M lithium chloride, and 33% (w/v) PEG1000. Prior to data collection, 20% (v/v) glycerol was added to each crystallization solution to protect the crystals from flash-cooling by liquid nitrogen. Diffraction datasets for the NSDHL_{holo} crystals were collected using an MAR300HE CCD detector at beamline BL44XU of SPring-8 (Japan). For the NSDHL_{apo} crystals, the data were collected using a Pilatus 6M detector at beamline 11C of Pohang Light Source (Republic of Korea). Both datasets were processed and scaled with the HKL-2000 software package [16].

1.2.3 Structure determination, refinement, and analysis

The crystal structure of NSDHL_{holo} was determined using the molecular replacement method, which used a monomer model of GOX2253 from *Gluconobacter oxydans* as a search model (PDB code: 3WJ7) [17], in the Phaser-MR program within the Phenix software suite [18]. The structure of NSDHL_{apo} was also determined by the molecular replacement method using a

monomer of the NSDHL_{holo} model as a template. In both models, the initial structures were further refined using Refmac [19] and Phenix.refine [18] software. The models were manually constructed, and ligands were added using the Coot software [20]. Five percent of the data were randomly set aside as the test data for the calculation of R_{free} [21]. The stereochemistry of the final structures was evaluated using MolProbity software [22]. The structural deviations were calculated using the secondary-structure matching (SSM) superpose option in Coot software [20]. The solvent-accessible surface areas were calculated using PISA [23], and the protein-ligand interactions were calculated using LigPlot [24].

1.2.4 Isothermal titration calorimetry (ITC)

ITC analysis was carried out using a MicroCal iTC200 instrument (GE Healthcare). All experiments were performed at 25°C under stirring at 1000 rpm in buffer A (40 mM HEPES, 150 mM NaCl, (pH 8.0), 0.5 mM TCEP, 5% (v/v) glycerol). All titrations of NSDHL into coenzymes (NAD⁺, NADH, NADP⁺, and NADPH (Santa Cruz Biotechnology, USA)) were performed using an initial injection volume of 0.4 μl followed by 20 identical 2 μl injections with a 5-second delay time per injection and intervals of 150 seconds between the injections. The stock solutions of the coenzyme and NSDHL were separately diluted in buffer A with a compound concentration of 1 mM and a protein concentration of 0.1 mM prior to the titrations. The NSDHL proteins were titrated into buffer A to estimate the heat background generated by the buffer, and all data were calculated by subtracting the background. To ensure reproducibility, all titrations were performed in triplicate. A one-site fitting model was applied, and a nonlinear least squares algorithm was used to obtain the binding constant (K_d), change in enthalpy (ΔH), and stoichiometry (N). Thermodynamic parameters were subsequently calculated with the following formula: $\Delta G = \Delta H - T\Delta S = -RT \ln K$, where ΔG , ΔH , ΔS , T, and R are the changes

in free energy, enthalpy, entropy of binding, experimental temperature, and gas constant, respectively. All heat generation data were collected and processed using the MicroCal Origin 7.0 software package.

1.2.5 Size-exclusion chromatography with multiangle light scattering (SEC-MALS)

SEC-MALS experiments for NSDHL_{apo} and NSDHL_{holo} were performed using a fast protein liquid chromatography (FPLC) system (GE Healthcare) connected to a Wyatt MiniDAWN TREOS MALS instrument and a Wyatt Optilab rEX differential refractometer (Santa Barbara, CA, USA). A Superdex 200 10/300 GL (GE Healthcare) gel filtration column pre-equilibrated with a buffer (50 mM HEPES, 500 mM NaCl, 500 mM imidazole, pH 8.0) was normalized using the ovalbumin protein. Proteins (5 mg/ml) were injected at a flow rate of 0.4 ml/min. The data were analyzed using the Zimm model for fitting static light scattering data and were graphed using EASI graph with an ultraviolet (UV) peak in ASTRA 6 software (Wyatt).

1.2.6 Thermal shift assay (TSA)

TSA for verifying the binding of compound **9** to NSDHL and the proposed docking model of compound **9** was performed in a QuantStudio 6 Real-Time PCR system (Applied BioSystems) using Sypro Orange (ThermoFisher) dye. A 5 μ M sample of wild-type NSDHL or its mutant (G47S or C104T), with or without 250 μ M ligand in 5% DMSO, was heat-denatured using a linear 25 $^{\circ}$ C to 99 $^{\circ}$ C gradient at a rate of 3 $^{\circ}$ C per minute. The denaturation curve and its derivative were obtained using QuantStudio Real-Time PCR software. The final reaction mixtures were prepared in 20 μ L volumes in at least triplicate in 384-well microplates. The

reaction was either 40 mM HEPES at pH 8.0, 150 mM NaCl (for the wild type) or the same buffer condition except with 5% glycerol and 0.5 mM TCEP-HCl (for G47S or C104T mutant protein). The average T_m values from each experiment for each variant were subtracted from compiled negative values for conditions with only 5% of DMSO to obtain ΔT_m D (Derivative T_m) using Protein Thermal Shift software (Applied BioSystems).

Protein solutions of wild-type NSDHL (31-267) and its mutants (G205S and K232 Δ) diluted in buffer A were added to each well of a MicroAmp 96-well reaction plate (Applied Biosystems) at a 5 μ M concentration. The assay was initiated with a final volume of 25 μ l after a 5x solution of SYPRO Orange (Sigma Aldrich) was added to each well. The temperature was ramped from 25 to 95°C at a rate of 1°C/min using an Applied Biosystems 7500 Fast Real-Time PCR Instrument System (Thermo Fisher Scientific). Fluorescence data were plotted as a sigmoidal curve. Midpoint melting temperature (T_m) values were calculated using Boltzmann sigmoidal fitting in GraphPad Prism 5 (GraphPad software). The data for each curve were normalized to the maximum and minimum fluorescence signals.

1.2.7 High-throughput virtual screening

Docking-based virtual screening was performed using Schrödinger Suite 2017-4 with the X-ray crystal structure of NAD⁺-bound NSDHL (PDB code: 6JKH). Protein preparation was revised using Protein Preparation Wizard in Maestro v11.4 software (Schrödinger Release 2017-4: Maestro 11.4, Schrödinger, LLC, New York, NY, USA, 2017), and the receptor grid box was generated with a 30 Å x 30 Å x 30 Å cube centered on complexed NAD⁺. The 388,852 molecule-containing KCB library was browsed for ligand preparation using LigPrep v4.4 (Schrödinger Release 2017-4: LigPrep 4.4, Schrödinger, LLC, New York, NY, USA, 2017) and the OPLS_2005 force field [25], and during the process, tautomer and ionization states at pH

7.0 ± 2.0 were generated using the Epik v4.2 module (Schrödinger Release 2017-4: Epik 4.2, Schrödinger, LLC, New York, NY, USA, 2017). Ligand docking was performed using Glide v7.7 (Schrödinger Release 2017-4: Glide 7.7, Schrödinger, LLC, New York, NY, USA, 2017) and the standard precision (SP) method. The top 5000 compounds were selected according to the Glide score. Then, rescoring of the top-ranking SP pose of each compound was conducted with the extra precision (XP) Glide scoring function. Based on the top 2000 compounds combined with visual inspection, 495 compounds were selected for biochemical testing. Among the 495 compounds that underwent in vitro biochemical screening, two compounds were identified as new chemical inhibitors. To find more analogs of active compounds, we performed a second round of virtual screening based on the fingerprint similarity method. The similarity between single pairs of compounds was calculated using Pipeline Pilot 2017 (Pipeline Pilot 2017: Dassault Systèmes, San Diego, CA, USA, 2017) for comparisons of the degree of similarity with two molecules. The database of KCB library compounds was screened to find analogs with a Tanimoto coefficient > 0.8 for comparisons between two active compounds, and 162 compounds were selected by visual inspection. Among the 162 compounds, 11 compounds were finally selected as potent inhibitors by in vitro biochemical screening.

1.2.8 NADH-based competitive binding assay for identifying inhibitors

To identify active hits from the extensive chemical pool, we conducted a high-throughput assay. The fluorescence signal of NADH increases when it binds to proteins, such as dehydrogenases [26], therefore, this property was used for the assay.

The assay was performed in a total volume of 100 µl in 96-well flat-bottom black

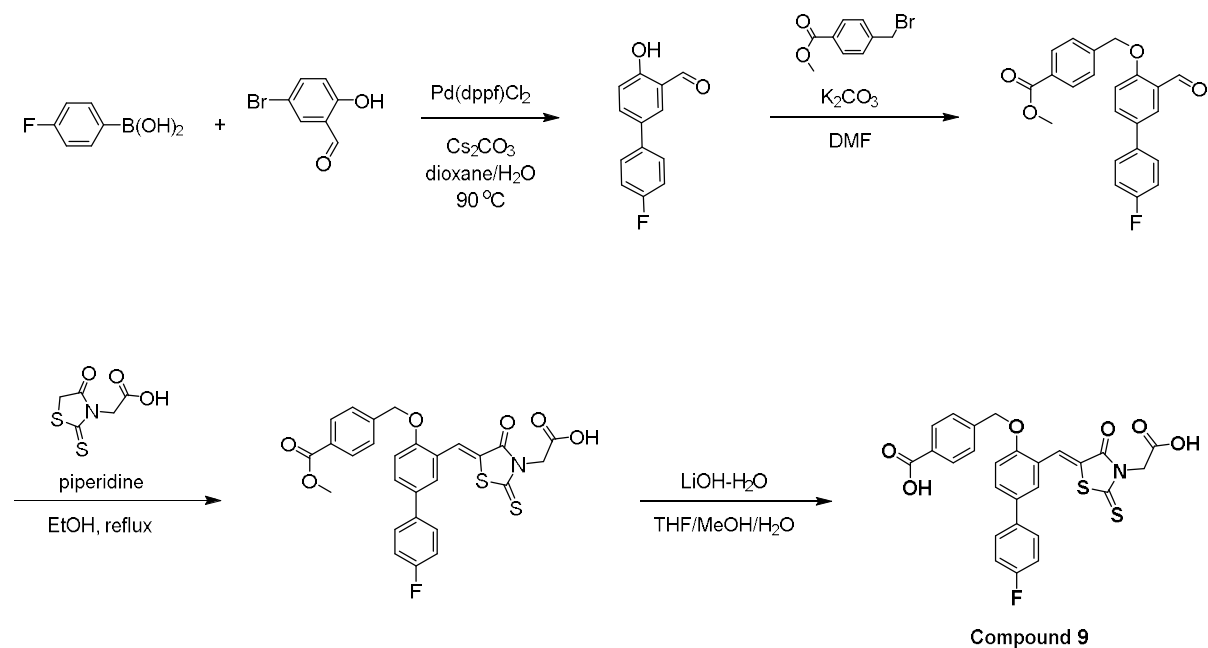
microplates (SPL Life Sciences). Two microliters of 100 μM library compounds were added to each well in duplicate. Forty-eight microliters of 32 μM NSDHL in assay buffer [50 mM HEPES (pH 8.0), 20% (v/v) glycerol] was added to each well. The plates were then incubated for 30 minutes at room temperature. The fluorescence intensity was measured using a SpectraMax M5 microplate reader (Molecular Devices) at Ex/Em = 340/460 nm at room temperature after adding 50 μl of 80 μM NADH solution in assay buffer. The inhibition (%) was calculated using the following equation: $\text{inhibition (\%)} = [1 - (F_{[\text{chemical}]} / F_{[\text{control}]})] \times 100$, where $F_{[\text{chemical}]}$ and $F_{[\text{control}]}$ are the fluorescence signals of the wells treated with the chemical and the dimethyl sulfoxide (DMSO) control, respectively. IC_{50} curves were fitted using the nonlinear regression function in GraphPad Prism5 (GraphPad software). The K_i values were converted from the IC_{50} and K_d values of NADH obtained from the ITC data. The data were generated using the Cheng-Prusoff equation for competitive inhibition.

1.2.9 Synthetic methods and characterization of a small molecule

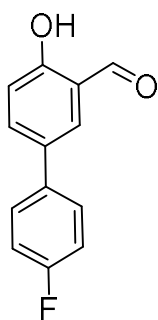
Unless otherwise stated, all reactions were performed under an inert (N_2) atmosphere. The reagents and solvents were reagent grade and purchased from Sigma-Aldrich, Alfa Aesar, Combi-Blocks and TCI Korea. ^1H and ^{13}C NMR spectra were recorded on Bruker Avance (300 MHz), Bruker Avance III (400 MHz) or Bruker Ascend (500 MHz) spectrometers at 25°C. Chemical shifts are reported in parts per million (ppm). Data for ^1H NMR are reported as follows: chemical shift (δ ppm) (integration, multiplicity, coupling constant (Hz)). Multiplicities are reported as follows: s = singlet, d = doublet, t = triplet, q = quartet, dd = double of doublet, and m = multiplet. Data for ^{13}C are reported in terms of chemical shifts (δ

ppm). The residual solvent peak was used as an internal reference. The mass spectra were obtained on an Electrospray TOF (ESI-TOF) mass spectrometer (BrukeramaZon X).

Synthetic schemes



4'-fluoro-4-hydroxy-[1,1'-biphenyl]-3-carbaldehyde.

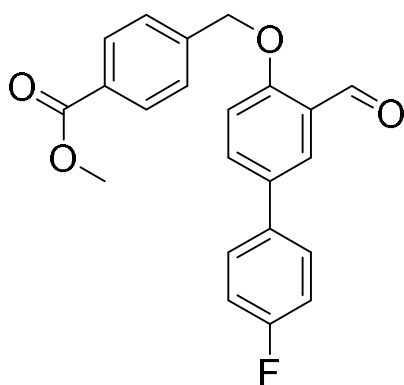


5-bromo-2-hydroxybenzaldehyde (4.30 g, 21.39 mmol), 4-fluorophenylboronic acid (3.59 g, 25.67 mmol) and Cs_2CO_3 (10.45 g, 32.08 mmol) were added to the reaction flask containing dioxane/ H_2O (60 ml, 1:1 ratio). Nitrogen gas was applied to the reaction mixture through a needle for 5 minutes, and $\text{Pd}(\text{dppf})\text{Cl}_2$ was then added to the reaction mixture. The reaction mixture was stirred at 90°C for 3 hours. The reaction mixture was diluted with EtOAc and washed with water followed by brine. The organic layer was collected, dried over anhydrous

sodium sulfate, filtered, concentrated, and purified by silica gel column chromatography to generate the product (4.21 g, 90.9%).

^1H NMR (300 MHz, CDCl_3): δ 10.99 (1H, s), 9.97 (1H, s), 7.70 (2H, d, $J = 6.0$ Hz), 7.50 (2H, d, $J = 6.0$ Hz), 7.16–7.05 (3H, m); MS (ESI) m/z Calcd for $\text{C}_{13}\text{H}_9\text{FO}_2(\text{M}^+)$: 216.0, Found, 217.1 ($\text{M}+\text{H}^+$).

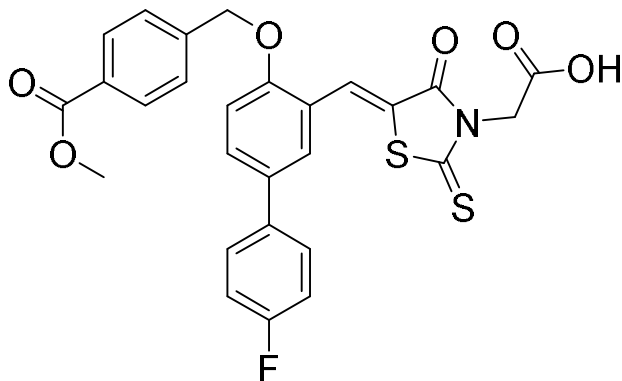
methyl 4-(((4'-fluoro-3-formyl-[1,1'-biphenyl]-4-yl)oxy)methyl)benzoate.



4'-fluoro-4-hydroxy-[1,1'-biphenyl]-3-carbaldehyde (1.00 g, 4.62 mmol), methyl 4-(bromomethyl)benzoate (1.16 g, 5.09 mmol) and K_2CO_3 (1.28 g, 9.25 mmol) were added to the reaction flask containing anhydrous DMF. The reaction mixture was stirred at room temperature for 2 hours. The reaction mixture was diluted with EtOAc and washed with water (50 ml x 3 times) followed by brine. The organic layer was collected, dried over anhydrous sodium sulfate, filtered, concentrated and purified by silica gel column chromatography to generate the product (1.64 g, 97.4%).

^1H NMR (500 MHz, CDCl_3): δ 10.61 (1H, s), 8.10 (2H, d, $J = 8.3$ Hz), 8.05 (1H, d, $J = 2.5$ Hz), 7.72 (1H, dd, $J = 8.6$ Hz, 2.5 Hz), 7.54 (2H, d, $J = 8.3$ Hz), 7.50 (2H, m), 7.13 (2H, t, $J = 8.6$ Hz), 7.07 (1H, d, $J = 10.5$ Hz), 5.30 (2H, s), 3.93 (3H, s); MS (ESI) m/z Calcd for $\text{C}_{22}\text{H}_{17}\text{FO}_4(\text{M}^+)$: 364.1, Found, 365.0 ($\text{M}+\text{H}^+$).

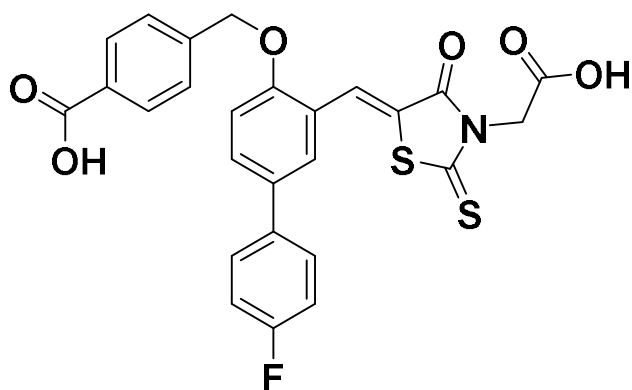
(Z)-2-(5-((4'-fluoro-4-((4-(methoxycarbonyl)benzyl)oxy)-[1,1'-biphenyl]-3-yl)methylene)-4-oxo-2-thioxothiazolidin-3-yl)acetic acid



methyl 4-(((4'-fluoro-3-formyl-[1,1'-biphenyl]-4-yl)oxy)methyl)benzoate (200 mg, 0.55 mmol), 2-(4-oxo-2-thioxothiazolidin-3-yl)acetic acid (94.46 mg, 0.49 mmol) and piperidine (11 μ l, 0.11 mmol) were added to the reaction flask containing EtOH (3 ml). The reaction mixture was refluxed at 90°C for 2 hours. The yellow precipitate in the reaction mixture was obtained by filtration and washed with cold EtOH to generate the product (239 mg, 81.1%).

^1H NMR (500 MHz, DMSO- d_6): δ 9.23 (1H, s), 8.02 (2H, d, J = 8.2 Hz), 8.00 (1H, d, J = 2.5 Hz), 7.78 (1H, dd, J = 8.4 Hz, 2.1 Hz), 7.70 (2H, m), 7.63 (3H, m), 7.34 (3H, m), 5.46 (2H, s), 4.74 (2H, s), 3.94 (3H, s); MS (ESI) m/z Calcd for $\text{C}_{27}\text{H}_{20}\text{FNO}_6\text{S}_2$ (M^+): 537.0, Found, 538.1 ($\text{M}+\text{H}^+$).

(Z)-4-(((3-((3-(carboxymethyl)-4-oxo-2-thioxothiazolidin-5-ylidene)methyl)-4'-fluoro-[1,1'-biphenyl]-4-yl)oxy)methyl)benzoic acid (compound 9)



(Z)-2-(5-((4'-fluoro-4-((4-(methoxycarbonyl)benzyl)oxy)-[1,1'-biphenyl]-3-yl)methylene)-4-oxo-2-thioxothiazolidin-3-yl)acetic acid (200 mg, 0.37 mmol) was dissolved in THF/MeOH/H₂O (4 ml, 3:1:1 ratio), and LiOH-H₂O (31 mg, 0.75 mmol) was added to the reaction vessel. The reaction mixture was stirred at room temperature for 5 hours. The reaction mixture was adjusted to pH 2.0 by adding 1 N HCl aqueous solution. The precipitate in the reaction mixture was obtained by filtration and washed with cold EtOH to generate compound **9** (187 mg, 95.9%).

mp: 100–102 °C (lit.ref 99–101 °C); TLC (CHCl₃:MeOH, 98:2 v/v): R_f = 0.23; [α]_D = –21.5 (0.1 M in n-hexane); ¹H NMR (400 MHz, CDCl₃): δ 9.30 (1H, s), 7.55–7.41 (6H, m), 5.61 (1H, d, *J* = 5.5 Hz), 5.40 (1H, d, *J* = 5.5 Hz), 4.93 (1H, m), 4.20 (2H, q, *J* = 8.5 Hz), 2.11 (3H, s), 1.25 (3H, t, *J* = 8.5 Hz); ¹³C NMR (125 MHz, CDCl₃): δ 165.4, 165.0, 140.5, 138.7, 131.5, 129.2, 118.6, 84.2, 75.8, 66.7, 37.9, 20.1; IR (Nujol): 1765 cm⁻¹; UV/Vis: λ_{max} 267 nm; HRMS (m/z): [M]⁺ calcd. for C₂₀H₁₅Cl₂NO₅, 420.0406; found, 420.0412; Elemental analysis (calcd., found for C₂₀H₁₅Cl₂NO₅): C (57.16, 57.22), H (3.60, 3.61), Cl (16.87, 16.88), N (3.33, 3.33), O (19.04, 19.09).

1.2.10 Liquid Chromatography Mass Spectrometry (LCMS)

LCMS experiment was performed using Waters ACQUITY (UPLC), SQD2 with PDA Detector (ACQUITY). LC Column was BEH C18 1.7 μm (2.1 \times 50 mm) and mobile phases were following condition:

Time	DW (0.2% TFA) (%)	ACN (%)	Flow rate (ml/min)
0	90	10	0.4
0.5	90	10	0.4
2.0	10	90	0.4
2.8	10	90	0.4
3.5	90	10	0.4
5	90	10	0.4

1.2.11 Kinetic solubility

The solubility of compound **9** was determined in Dulbecco's phosphate-buffered saline (DPBS) (Welgene Inc., Daegu, Korea). To determine the kinetic solubility, a stock solution of the compound was prepared at a concentration of 50 mM in DMSO. After a 100-fold dilution of the stock solution to DPBS (i.e., at a final concentration of 500 μM in 1% DMSO), the mixture was agitated for 2 hours at room temperature. The mixture was then centrifuged at 13,200 rpm for 10 minutes and filtered with a 0.2 μm syringe filter (Minisart RC15, Sartorius Stedim Biotech, Goettingen, Germany), and the resulting filtrate was analyzed by an LC/MS/MS system (Applied Biosystems 3200 Qtrap MS/MS system with Alliance Waters e2695 LC system) to determine the compound concentration in the medium.

1.2.12 Molecular docking

The docking of compound **9** was performed using Schrödinger Suite 2017-4 with the X-ray crystal structure of NAD⁺-bound NSDHL. Protein preparation was revised using Protein

Preparation Wizard in Maestro v11.4 (Schrödinger Release 2017-4: Maestro 11.4, Schrödinger, LLC, New York, NY, USA, 2017), and the receptor grid box was generated with a 30 Å x 30 Å x 30 Å cube centered on complexed NAD⁺. The ligands were minimized using a Merck molecular force field (MMFF) with a dielectric constant of 80.0 in MacroModel v11.8 (Schrödinger Release 2017-4: MacroModel 11.8, Schrödinger, LLC, New York, NY, USA, 2017). Ligand docking was performed using Glide v7.7 software (Schrödinger Release 2017-4: Glide 7.7, Schrödinger, LLC, New York, NY, USA, 2017) with the SP method. The proposed binding models of the compounds were visualized using PyMOL v1.5 (Schrödinger, LLC, 2012) [27].

1.2.13 Surface plasmon resonance (SPR)

The direct binding affinity of NSDHL with compound **9** was investigated by SPR analysis. SPR measurements were performed at 25°C using the Biacore T200 apparatus (GE Healthcare). For immobilization, the Amine Coupling Kit containing 0.1 M N-hydroxysuccinimide and 0.4 M 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride was applied to a CM5 sensor chip (GE Healthcare) with a buffer containing 10 mM HEPES, 200 mM NaCl, pH 8.0, and 0.005% (v/v) Tween20 according to the manufacturer's protocol. Subsequently, 50 µg/mL NSDHL dissolved in 10 mM sodium phosphate at pH 7.4 was injected at a rate of 5 µl/min at regular intervals until the immobilization level reached approximately 8,000 response units (RU). To deactivate the remaining activated carboxyl groups, 1.0 M ethanolamine at pH 8.5 was applied to the surface of the sensor chip for 400 s. The control experiment was treated as described above without protein, and the response of the control was subtracted from each sample dataset. Compound **9** at 2-fold diluted concentrations of 0.625–20.0 µM in PBS buffer containing 5% DMSO and 0.005% Tween 20 was injected over the NSDHL-immobilized chip

at a rate of 30 $\mu\text{l}/\text{min}$ for 120 s, followed by dissociation for 300 s in multicycle reactions. The sensor chip surface was regenerated for 10 s with 10 mM NaOH between cycles. The SPR response data were fitted to the simple bimolecular 1:1 Langmuir isotherm binding model to determine the equilibrium dissociation constant (K_D) using Biacore T200 evaluation software 3.0 (GE Healthcare).

1.2.14 Cellular viability assay

A431 (epidermoid carcinoma cell line), A549 and H1299 (non-small cell lung cancer cell lines), and K562 (myelogenous leukemia cell line) cells were obtained from the Korean Cell Line Bank. The A431 cells were cultured in Dulbecco's modified Eagle's medium (DMEM), and A549, H1299, and K562 cells were cultured in Roswell Park Memorial Institute (RPMI)-1640 medium, which were both supplemented with 10% (v/v) fetal bovine serum (FBS), L-glutamine, and antibiotics (penicillin and streptomycin). The cells were cultured at 37°C in a humidified incubator with 5% CO₂. To determine the cytotoxicity of the inhibitors, A431, A549, and H1299, and K562 cells were plated in 96-well flat-bottom plates at a concentration of 5×10^3 and 1×10^5 cells per well, respectively. After 24 hours, the medium was removed, and compound **9** (0.16–500 μM) or the vehicle (DMSO) was added to each well. After an additional incubation for 72 hours, the plates were treated with a stock solution of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl-tetrazolium bromide (MTT) reagent, followed by additional incubation for 2 hours. After the removal of the MTT solution, DMSO was added to the plates to dissolve the purple crystals. The absorbance was measured using a SpectraMax M5 multi-detection microplate reader (Molecular Devices, USA) at a wavelength of 570 nm.

The cytotoxicity assay of the NSDHL inhibitor with the EGFR-targeting drug erlotinib was performed in an erlotinib-sensitive cell line (A431) and erlotinib-resistant cell line (A549

and H1299). Then, each cell line was plated in 96-well cell culture plates at a concentration of 5×10^3 cells per well. After 24 hours, A431 cells were treated with either DMSO (0.2%), erlotinib (0.5 μ M), compound **9** (200 or 500 μ M), or compound **9** and erlotinib (either 200 or 500 μ M, and 0.5 μ M, respectively). A549 and H1299 cells were subjected to the same treatment conditions except for 5 μ M of erlotinib. All plates were incubated for 48 hours after the treatment. Cell viability was measured using MTT reagent as described above, and the absorbance of the DMSO-treated groups was normalized to the average viability. We used the sensitization index (SI) [28,29] to determine whether our inhibitor enhances the cytotoxicity of the EGFR-targeting drug. The SI was calculated for each individual well on a 96-well plate as follows: $SI = (X_E / D_E) / (X_D / D_D)$, where X and E are the viability of each well treated with compound **9** and erlotinib, respectively. D is the viability of cells treated with DMSO as a negative control to compound **9** or erlotinib. $SI = 1$ indicates no interaction, and $SI < 1$ indicates the enhanced cytotoxic effect. We verified that DMSO had no effect on the cell viability assay.

1.2.15 Flow cytometry

A431 cells were plated in 6-well plates at 50~60% confluency. Each well was treated with either compound **9** (100 μ M) or the vehicle (the corresponding volume of DMSO) at 24 hours after cell plating and was incubated for 72 hours. After washing twice with phosphate-buffered saline (PBS), the cells were starved for an additional 24 hours in serum-free DMEM at 37°C. At the end of the incubation, the medium was changed to serum-free DMEM containing 500 ng/ml EGF, and the plates were then incubated for 8, 24 and 48 hours. After washing the cells with PBS twice, the cells for all conditions were harvested simultaneously and then washed with fluorescence-activated cell sorting (FACS) buffer (1% FBS and 0.01% NaN₃ in PBS) and stained with either phycoerythrin (PE)-conjugated anti-EGFR (Abcam, Cambridge, UK,

ab130738) or its isotype control at 4°C for 20 minutes in the dark. Fluorescence was measured using a FACS Aria III flow cytometer (BD Biosciences), and the data were analyzed by FlowJo software (Tree Star). The EGFR level of each dataset was determined using geometric mean fluorescence intensity (gMFI), and the RFI was calculated using the following formula: gMFI of PE-EGFR / gMFI of the isotype control antibody. Three independent experiments were carried out to ensure reproducibility.

1.2.16 Immunoblotting

A431 cells were lysed with radioimmunoprecipitation assay buffer (50 mM Tris (pH 7.4), 150 mM NaCl, 1 mM EDTA, 1% NP-40, 0.25% sodium deoxycholate) supplemented with protease (Roche) and phosphatase inhibitors (Sigma). For immunoblotting, cell lysates were separated using 7% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to nitrocellulose membranes. The membranes were blocked with 5% nonfat dry milk, incubated with primary antibodies overnight at 4°C and subsequently reacted with horseradish peroxidase-conjugated secondary antibodies (GTX213110-01 and GTX213111-01, GeneTex) for 1 hour at room temperature. Bands were visualized using enhanced chemiluminescence with a ChemiDoc™ MP imager (Bio-Rad).

Antibodies specific for the following factors were used for immunoblotting at the indicated dilutions: phospho-EGFR (Tyr1068) (1: 5000, #2234, Cell Signaling Technology), EGFR (1:5000, SC-03, Santa Cruz), phospho-AKT (Ser473) (1:2000, #4060, Cell Signaling), AKT (1:1000, #9272, Cell Signaling Technology), phospho-ERK1/2 (Thr202/Tyr204) (1:2000, #9101, Cell Signaling Technology), ERK1/2 (1:1000, #9102, Cell Signaling Technology) and β -actin (1:10000, MAB1501R, Millipore). Original images of all blots are shown in Figure 1.

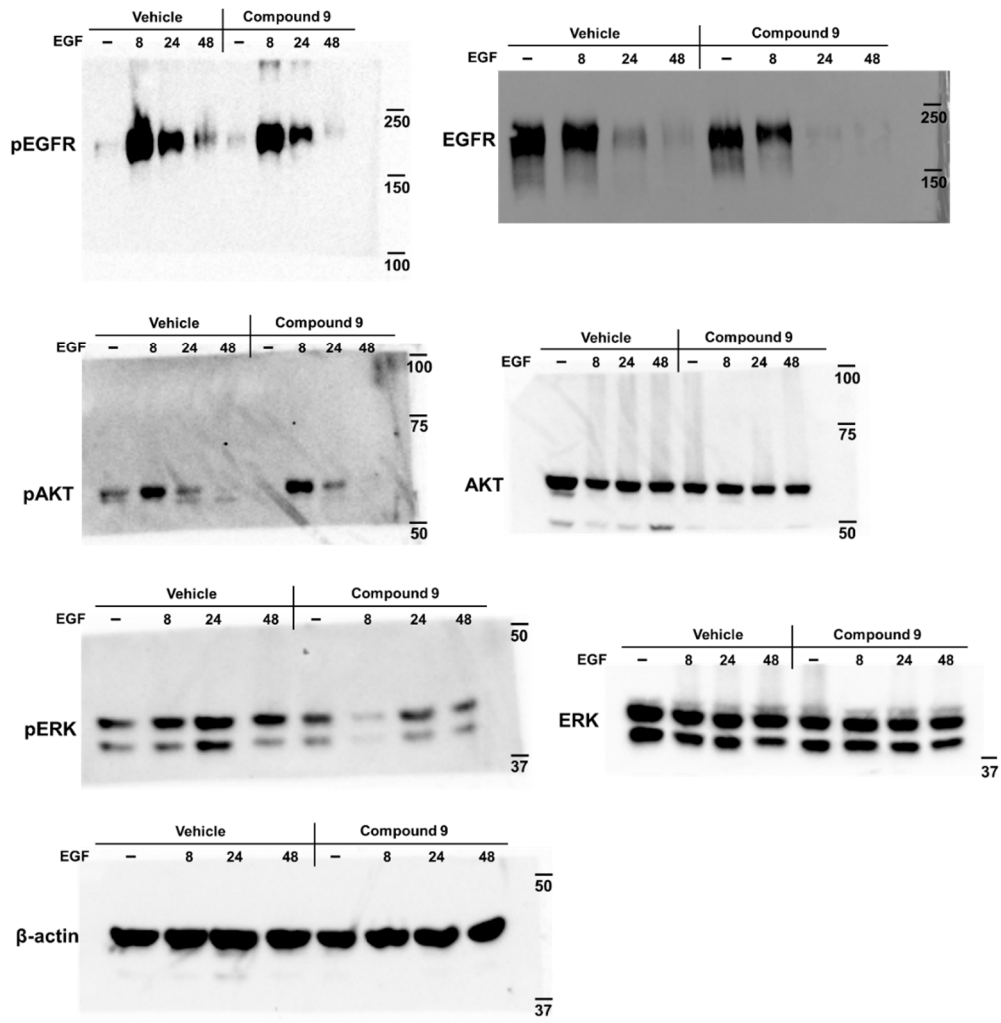


Figure 1. Original images of immunoblots.

1.2.17 Public data analysis

Genomic alterations of *NSDHL* in EGFR-driven human cancers in the TCGA PanCancer Atlas datasets were analyzed using cBioportal (<http://www.cbioportal.org/>) [30]. The analysis of overall survival (OS) of patients with lung cancer were performed using the Kaplan-Meier (KM) plotter online database (<http://www.kmplot.com>) [31]. Three public datasets of human lung cancers (GSE31210 [32], GSE30219 [33], and caArray [34]) were selected in the KM plotter, and the mean expression of two *NSDHL* probes (209 279_s_at and 215093_at) was used in the analyses. To stratify the patients into two groups with high or low expression levels of *NSDHL*, the cutoff value of *NSDHL* expression was determined based on the computed best-performing thresholds.

1.2.18 Statistical analysis

Statistical analysis was performed using GraphPad Prism5 (GraphPad software). Statistically significant values were compared using unpaired two-tailed Student's t-test or one-way analysis of variance (ANOVA) with Bonferroni's post hoc test. $P < 0.05$ was used to indicate a statistically significant difference, and the data are presented as the mean \pm standard error of the mean (SEM).

1.2.19 Data availability

The coordinates and structural factors were deposited in the Research Collaboratory for Structural Bioinformatics (RCSB) PDB under the accession codes 6JJKG for human *NSDHL*_{apo} (NAD⁺-free form) and 6JKH for human *NSDHL*_{holo} (NAD⁺-bound form).

1.3 Results

1.3.1 Overall structures of human NSDHL

In this study, we obtained two crystal structures, including the single domain, from an extensive number of trials aimed at crystallizing various NSDHL constructs as well as the full-length construct including the transmembrane region. The two crystal structures consisted of a ligand-free form (NSDHL_{apo}) and an NAD⁺-bound form (NSDHL_{holo}) from the truncated construct (31–267) (Figure 2).

NSDHL_{apo} and NSDHL_{holo} were refined to $R_{\text{work}} / R_{\text{free}} = 22.8\% / 29.1\%$ at a 2.9 Å resolution and $R_{\text{work}} / R_{\text{free}} = 17.3\% / 23.8\%$ at a 3.0 Å resolution, respectively. NSDHL_{apo} and NSDHL_{holo} consist of 432 and 429 amino acid residues in two monomers (two NAD⁺ molecules), respectively, and 3 and 18 water molecules in the asymmetric unit, respectively (Table 2). The monomers of NSDHL_{apo} and NSDHL_{holo} comprise five α -helices and nine β -strands (Figure 3). In both models, we observed disordered regions, including (i) N-terminal residues [Lys31–Gln34 (chain A of NSDHL_{apo}) and Lys31–Asn35 (chain B of NSDHL_{apo} and chain A and B of NSDHL_{holo})], (ii) a loop between β 4 and α 3 [Pro109–Asn112 (chain B of NSDHL_{apo})], (iii) loops between β 5 and α 4 [Gly152–Lys167 (chain A and B of NSDHL_{apo}), Val153–Met166 (chain A of NSDHL_{holo}), and Asp154–Pro168 (chain B of NSDHL_{holo})], (iv) loops between β 6 and β 7 [Phe204–Asp208 (chain A and B of NSDHL_{apo}) and Gly202–Pro209 (chain A and B of NSDHL_{holo})], and (v) a loop between β 8 and α 5 [Gly229–Lys232 (chain B of NSDHL_{apo})].

Two NSDHL_{apo} (and NSDHL_{holo}) monomers in the asymmetric unit were found to be highly similar to each other, with the root mean square (r.m.s.) deviation of 0.78 Å (0.67 Å for

NSDHL_{holo}) for 210 equivalent C α (207 equivalent C α for NSDHL_{holo}) atom pairs. Two monomer structures of NSDHL_{apo} were nearly identical to those of NSDHL_{holo}, with r.m.s. deviations of 0.61–0.83 Å for 203–208 equivalent C α pairs. Despite the structural similarities between NSDHL_{apo} and NSDHL_{holo}, greater r.m.s. deviations were observed in the following several regions: the loops between β 2 and β 3 and between β 7 and β 8. The loop between β 2 and β 3 showed a maximum deviation of 2.56 Å when two C α in Asp74 were compared with each other, whereas the loop between β 7 and β 8 deviated with a maximum deviation of 2.79 Å when two C α in Gly222 were compared with each other. A structural similarity search using the Dali server [35] revealed other structures that are similar to the monomer structures of NSDHL_{apo} and NSDHL_{holo} (Figure 4).

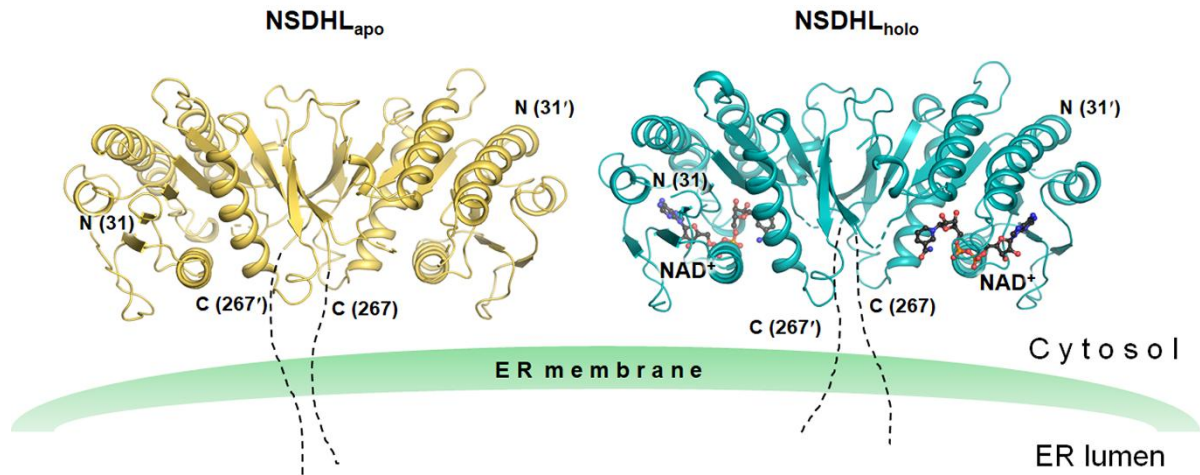


Figure 2. Overall structures of the apo and holo forms of NSDHL. Two structures of NSDHL (31–267) are anchored in the ER membrane. NSDHL_{apo} and NSDHL_{holo} are shown as ribbon diagrams and colored in yellow and cyan, respectively. NAD⁺ of NSDHL_{holo} (black carbon) is shown as a ball and stick, and disordered loops are shown as dashed lines. Truncated C-terminal residues are shown with dotted lines. The structures were constructed using PyMOL [27].

Table 2. Statistics for data collection and model refinement.

Crystals	NSDHL _{apo} (Ligand-free form)	NSDHL _{holo} (NAD ⁺ -bound form)
A. Data collection		
X-ray wavelength (Å)	0.9794	0.9000
Space group	<i>P2</i> ₁	<i>P2</i> ₁
Unit cell length (<i>a</i> , <i>b</i> , <i>c</i> , Å)	49.48 58.53 75.60	50.02 77.37 53.99
Unit cell angle (α , β , γ , °)	90.00 108.34 90.00	90.00 96.80 90.00
Resolution range (Å)	30.00–2.9 (2.95–2.90) ^a	30–3.0 (3.05–3.00) ^a
Total / unique reflections	29,802 / 9,057	32,821 / 8,240
Completeness (%)	98.5 (96.6) ^a	99.8 (100) ^a
CC _{1/2} ^b	0.964 (0.884) ^a	0.916 (0.691) ^a
<i>I</i> / σ _{<i>I</i>}	22.2 (2.7) ^a	22.2 (5.3) ^a
<i>R</i> _{merge} ^c	0.076 (0.287) ^a	0.146 (0.865) ^a
B. Model refinement		
<i>R</i> _{work} / <i>R</i> _{free} ^d	0.228 / 0.291	0.173 / 0.238
No. / average <i>B</i> -factor (Å ²)		
Protein atoms	3,335 / 85.7	3,293 / 47.1
Water oxygen atoms	3 / 69.3	18 / 39.35
Ligand atoms		
Nicotinamide adenine dinucleotide	N/A	88 / 56.8
R.m.s. deviations from ideal geometry		
Bond lengths (Å) / bond angles (°)	0.006 / 0.907	0.008 / 1.49
Ramachandran plot (%)		
Most favorable	93.46	95.13
Allowed	6.54	4.87
Disallowed	0.00	0.00

^a Values in parentheses refer to the highest resolution shell.

^b CC_{1/2} is described in ref [36].

^c $R_{\text{merge}} = \frac{\sum_h \sum_i |I(h)_i - \langle I(h) \rangle|}{\sum_h \sum_i I(h)_i}$, where *I*(*h*) is the intensity of reflection *h*, \sum_h is the sum over all reflections, and \sum_i is the sum over *i* measurements of reflection *h*.

^d $R = \frac{\sum | |F_{\text{obs}}| - |F_{\text{calc}}| |}{\sum |F_{\text{obs}}|}$, where *R*_{free} is calculated for a randomly chosen 5% of reflections, which were not used for structure refinement and *R*_{work} is calculated for the remaining reflections.

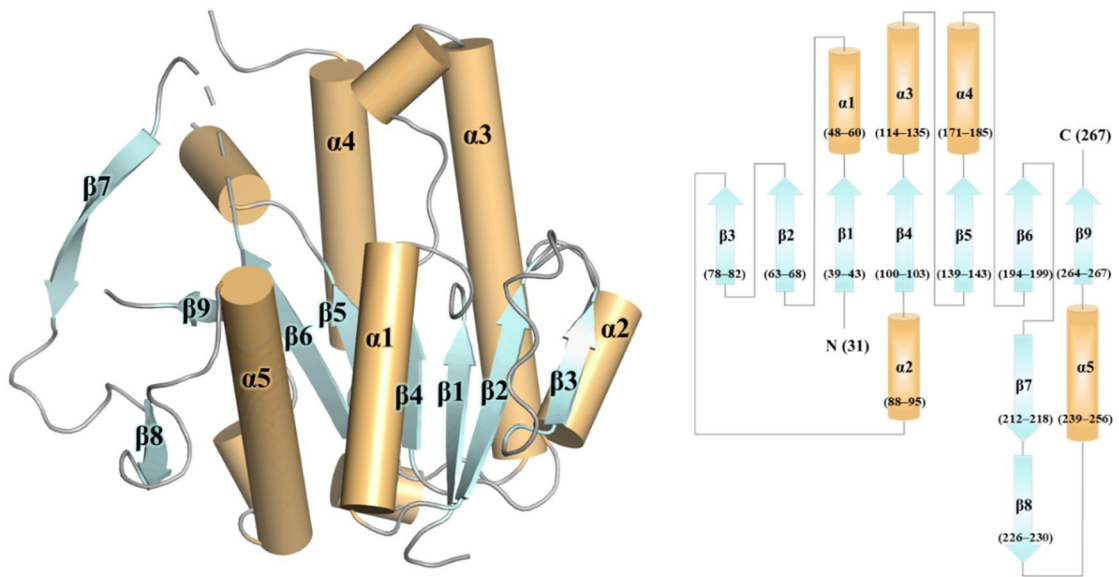


Figure 3. Conformation of the NSDHL structure and a topology diagram. Helices (light-orange) and strands (cyan) are shown as a cylindrical and ribbon diagram. Loops are shown in gray, and disordered residues are shown as dashed lines. The topology diagram of NSDHL is shown in the right panel with the residue numbers corresponding to the secondary structure elements. The NSDHL_{apo} model was used in this figure.

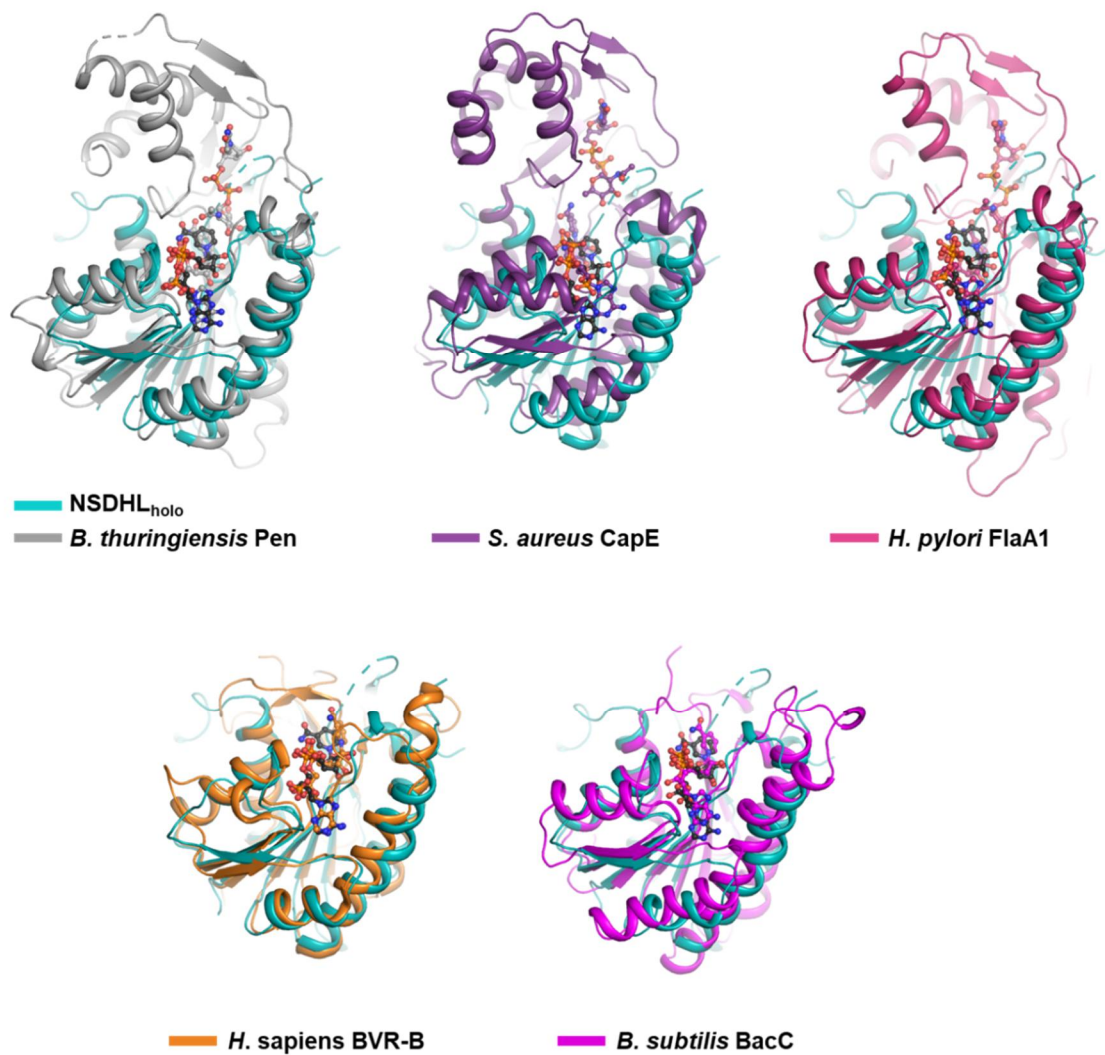


Figure 4. A structural comparison of NSDHL and its homologs. Superimposed views of the monomer of NSDHL_{holo} and its homologs identified on the Dali server: (i) *Bacillus thuringiensis* Pen (gray) [37] (PDB code 6BWC; Z-scores of 22.3–22.6, r.m.s. deviations of 1.9–2.0 Å, and a sequence identity of 26–27% for 177–179 equivalent C α pairs), (ii) *Staphylococcus aureus* CapE (purple) [38] (PDB code 4G5H; Z-scores of 22.2–22.4, r.m.s. deviations of 2.0–2.1 Å, and a sequence identity of 21% for 179 equivalent C α pairs), (iii) *Helicobacter pylori* FlaA1 (red pink) [39] (PDB code 2GN4; Z-scores of 21.5–21.6, r.m.s. deviations of 1.9–2.0 Å, and a sequence identity of 22% for 178 equivalent C α pairs), (iv) *Homo sapiens* biliverdin IX β reductase (BVR-B) (orange) [40] (PDB code 1HDO; Z-scores of 20.7–21.0, r.m.s. deviations of 2.1–2.2 Å, and a sequence identity of 18% for 173 equivalent C α pairs), and (v) *Bacillus subtilis* oxidoreductase BacC (magenta) [41] (PDB code 5TIV; Z-scores of 20.5–20.6, r.m.s. deviations of 2.2 Å, and a sequence identity of 17% for 177–178 equivalent C α

pairs). All coenzymes and substrates are shown as ball and sticks. Oxygen, nitrogen, and phosphorus atoms are shown in red, blue, and orange, respectively. The carbon atoms are shown as colors identical to the proteins except for NAD⁺ (black) in NSDHL.

1.3.2 Diverse features of human NSDHL structures

Although the full-length structure of NSDHL has two domains (a substrate-binding domain and a coenzyme-binding domain) along with a short transmembrane domain, the two structures of NSDHL (31–267) in this study had a single domain architecture consisting of an α/β doubly wound scaffold of the short-chain dehydrogenase/reductase (SDR) family fold [42]. The structures of NSDHL consisted of a seven-stranded parallel β -sheet, which was located between two arrays of two or three α -helices ($\alpha 1$ and $\alpha 6$ on one side and $\alpha 2$, $\alpha 3$, and $\alpha 4$ on the other side). The central β -sheet and the two groups of helices were mainly held together through hydrophobic interactions. The coenzyme-binding site was located within two $\beta\alpha\beta\alpha\beta$ Rossmann folds [43] of $\beta 1$ - $\alpha 1$ - $\beta 2$ -(loop)- $\beta 3$ and $\beta 4$ - $\alpha 3$ - $\beta 5$ - $\alpha 4$ - $\beta 6$ (Figure 3). The dimeric interfaces in NSDHL_{apo} and NSDHL_{holo} were composed of a pair of three-stranded antiparallel β -sheets ($\beta 8$ - $\beta 7'$ - $\beta 9$ and $\beta 8'$ - $\beta 7$ - $\beta 9'$) (Figure 5A). The interface areas of NSDHL_{apo} and NSDHL_{holo} per dimer were 4,450 and 4,050 Å², respectively. The dimeric structures of the NSDHL crystals were consistent with the size-exclusion chromatography with multiangle light scattering (SEC-MALS) results, which showed the dimeric forms in solution (Figure 5B). The residues involved in the hydrophilic and hydrophobic interactions at the interfaces of the NSDHL structures were nearly identical. The central $\beta 7$ (and $\beta 7'$) had hydrophilic interactions with $\beta 8'$ and $\beta 9'$ ($\beta 8$ and $\beta 9$ for $\beta 7'$) of the other monomer (Figure 5C), and the residues involved in the hydrophobic interactions of the dimeric interface are shown in Figure 5D.

Despite substantial structural similarities between NSDHL and other homologs (Figure 4), strikingly different structural features of NSDHL were observed. Compared with the Gly168–Thr204 region of *Bacillus thuringiensis* Pen (Protein Data Bank (PDB) code: 6BWC) [37], which is composed of one α -helix and two short β -strands, the

corresponding region (His201–Thr238) of NSDHL was found to have two long β -strands (β 7 and β 8) (Figure 5E). In addition, while the Gly168–Thr204 region of Pen forms the binding pocket for its sugar substrate, the His201–Thr238 region of NSDHL is positioned within the coenzyme-binding region and forms a dimeric interface with C-terminal β 9 (Figure 5D).

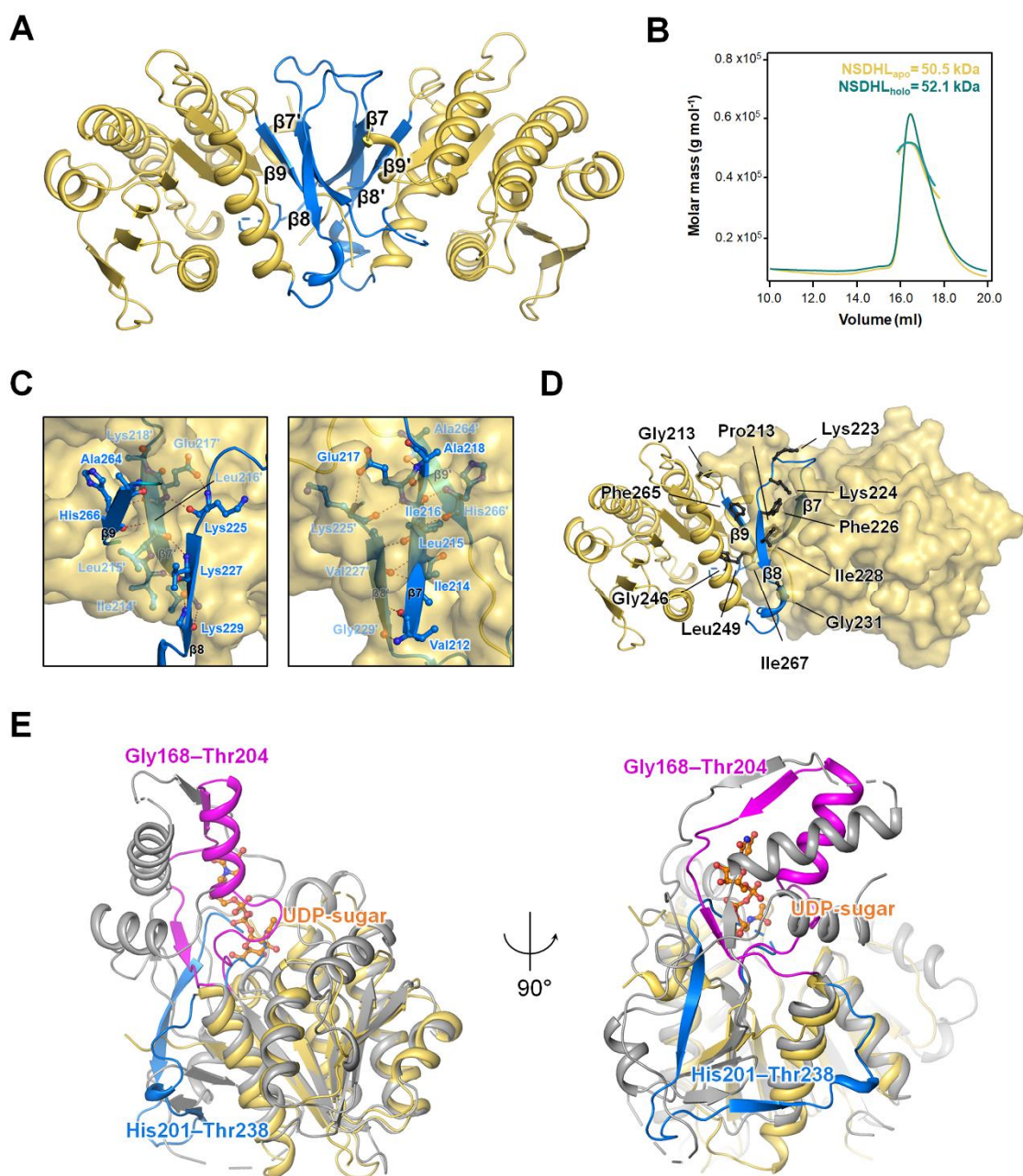


Figure 5. Distinct dimeric interface of the NSDHL structure. **A** Representation (blue) of the dimeric interface consisting of a pair of three-stranded antiparallel β -sheets ($\beta 8$ - $\beta 7'$ - $\beta 9$ and $\beta 8'$ - $\beta 7$ - $\beta 9'$). **B** NSDHL_{apo} (yellow line) and NSDHL_{holo} (cyan line) were analyzed by SEC-MALS. The thick line indicates the measured molecular mass. **C** Detailed hydrophilic interactions are shown as red dotted lines, and the interacting residues are shown as a ball and stick with the corresponding β -strands of chain B. **D** Residues contributing to the hydrophobic interactions at the interface are shown as black balls and sticks. All residues of chain A and chain B are shown as a ribbon

diagram and a surface view, respectively. **E** Superimposed views of NSDHL and *B. thuringiensis* Pen (PDB code 6BWC). The region (His201–Thr238, covering β 7– β 8) of NSDHL and the corresponding region (Gly168–Thr204) of Pen are shown in blue and magenta, respectively. The sugar substrate of Pen (orange carbon) is shown as a ball and stick. All figures are shown using NSDHL_{apo} due to no major difference between NSDHL_{apo} and NSDHL_{holo} in this description.

1.3.3 NAD⁺ binding site in NSDHL

In the NAD⁺ binding site, eleven hydrogen bonds were found between the surface residues of NSDHL and NAD⁺, in addition to many other kinds of hydrophobic interactions (Figure 6A, B). In the adenosine ribose moiety of NAD⁺, Asp84 and the backbone amides of Leu85 and Ile69 showed hydrogen bonds with the adenine ring, and the ring was stabilized by the hydrophobic interaction with Ile69 and Gly83 (Figure 6A, B). A backbone amide of Ser106 formed a hydrogen bond with the oxygen of the adenine-ribose ring. Furthermore, the proximal residues Cys104 and Ala105 had hydrophobic interactions with the ribose ring (Figure 6A, B). The bridging of the acidic Asp68 with the two hydroxyl groups of adenine-ribose implies that NSDHL may prefer NAD⁺ to NADP⁺ as a coenzyme (Figure 6A). This finding is consistent with the general concept that a single coenzyme is chosen among more than two coenzymes involved in enzyme-catalyzed reactions. The glycine-rich motif GGxGxxG (Gly44–Gly50), which is the characteristic sequence of the complex SDR family [44], was observed in the NAD⁺ binding site (Figure 6B). In addition, this motif was also found to accommodate the adenine-ribose and pyrophosphate moieties by hydrophilic (backbone amides of Phe48 and Leu49) and hydrophobic interactions (Gly44, Ser46, and Gly47) (Figure 6A, B). In the nicotinamide-ribose moiety, the catalytic residues Tyr172 and Lys176, which are conserved as the YxxxK motif in the SDR family [44], formed hydrogen bonds with two hydroxyl groups of the ribose moiety (Figure 6A). An additional hydrogen bond between a water molecule and a hydroxyl group was also observed in chain A of NSDHL_{holo} (Figure 6A). All moieties of NAD⁺ were clearly defined, except for the nicotinamide rings. The absence of a sterol substrate may have resulted in the poor observation of the nicotinamide rings in both monomers of NSDHL_{holo}. This structural information on the

interactions between NSDHL and its coenzyme served as the basis for the design of the inhibitors.

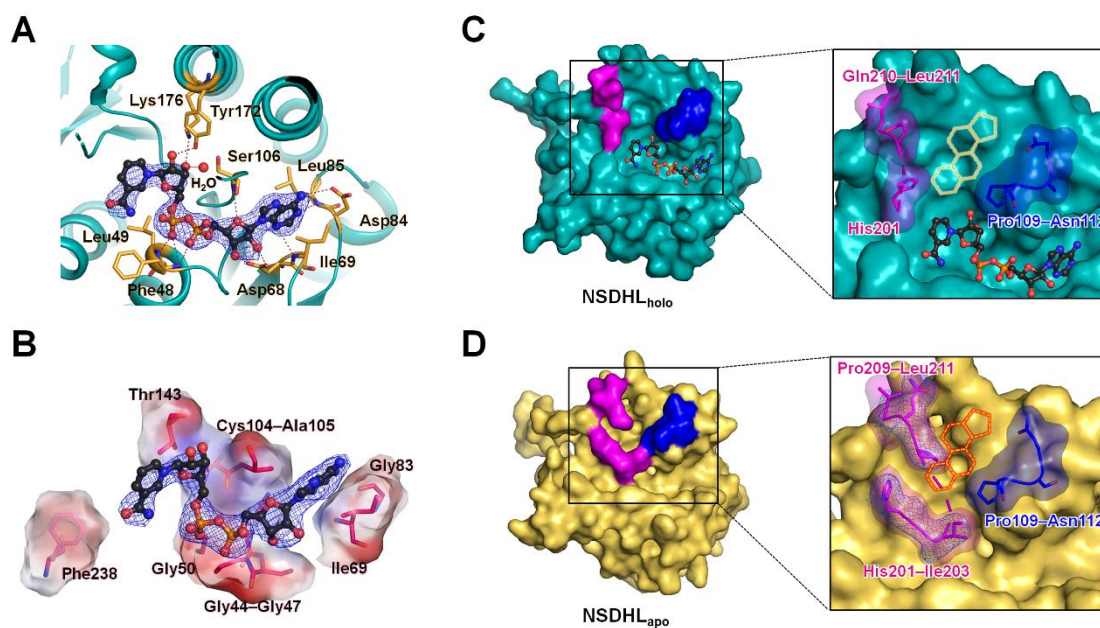


Figure 6. NAD⁺ binding site and the unique conformational change in NSDHL induced by the binding of NAD⁺. **A, B** Detailed views of the NAD⁺ binding site shown with hydrophilic and hydrophobic interactions. The hydrophilic interactions are shown as red dotted lines. The interacting residues and a water molecule are labeled and shown as orange sticks and a red sphere, respectively (**A**). The residues contributing to hydrophobic interactions are labeled and shown as pink sticks with the electrostatic potential surface view (**B**). NAD⁺ is shown with a $2mF_o - DF_c$ electron density map, which is contoured at 1.0σ in blue meshes. **C, D** Unique open and closed forms observed in NSDHL_{holo} and NSDHL_{apo}, respectively. Both structures are shown as surface views, and two regions relevant to the conformational change are highlighted in magenta and blue, respectively, in the black rectangular boxes. In the detailed views (right) enlarged from the black rectangular boxes, the key residues are also shown as sticks in the transparent surface views. The sterol precursor in the expected substrate-binding site is shown as a yellow dashed model. The red-colored sterol precursor model indicates possible blockage of a substrate in the substrate-binding site. The key residues observed in NSDHL_{apo} are shown with a $2mF_o - DF_c$ map of 1.0σ in blue meshes (**D**).

1.3.4 Conformational change induced by the binding of NAD⁺ allows the binding of a sterol precursor to NSDHL

The structural analysis of NSDHL_{apo} and NSDHL_{holo} showed that the B-factors of the key residues in the NAD⁺ binding site compared with the overall B-factors were lower for NSDHL_{holo} (36 / 47 Å²) than for NSDHL_{apo} (85 / 85 Å²) (Supplementary Figure S3). Similarly, the key residues and the overall structure of NSDHL_{holo} showed a clearer electron density map than those of NSDHL_{apo}, suggesting that the binding of NAD⁺ induces the rigid conformation of NSDHL. The structural comparison of NSDHL_{apo} and NSDHL_{holo} also revealed strikingly different conformations. In the structure of NSDHL_{holo}, the region surrounded by the loops (His201–Leu211, where Gly202–Pro209 were unclear, and Pro109–Asn112), which is likely to constitute the substrate-binding site and the active site, had an open conformation (Figure 6C). On the other hand, the loop (His201–Leu211) of NSDHL_{apo}, where three additional residues (Gly202, Ile203, and Pro209) were newly observed, occupied the position of the nicotinamide ring located in the active site and obstructed the expected substrate-binding region, implying a closed conformation in the absence of NAD⁺ (Figure 6D).

The structural importance of the loop (His201–Leu211) was also confirmed in TSA of two NSDHL mutants (G205S and K232Δ), which are the most common causes of congenital hemidysplasia with ichthyosiform erythroderma and limb defects (CHILD) and CK syndromes [45], respectively. G205S and K232Δ had a midpoint melting temperature (T_m) of 39.1 and 37.8°C, respectively, whereas wild-type NSDHL had a T_m of 49.6°C, indicating that NSDHL mutants have unstable folding and are degraded near the physiological temperature (Figure 8). Interestingly, Gly205, which constitutes the loop (His201–Leu211), and Lys232', which is present in the other monomer of the dimeric

interface, were observed near the loop (Figure 9). Taken together, the conformational change between two NSDHL structures may facilitate the uptake of a sterol substrate and may suggest that the catalytic reaction of NSDHL is the result of a precisely ordered and sequential mechanism [46]. Moreover, we identified that the loop (His201–Leu211) plays a pivotal role in the conformational change and stability for binding its substrate and coenzyme.

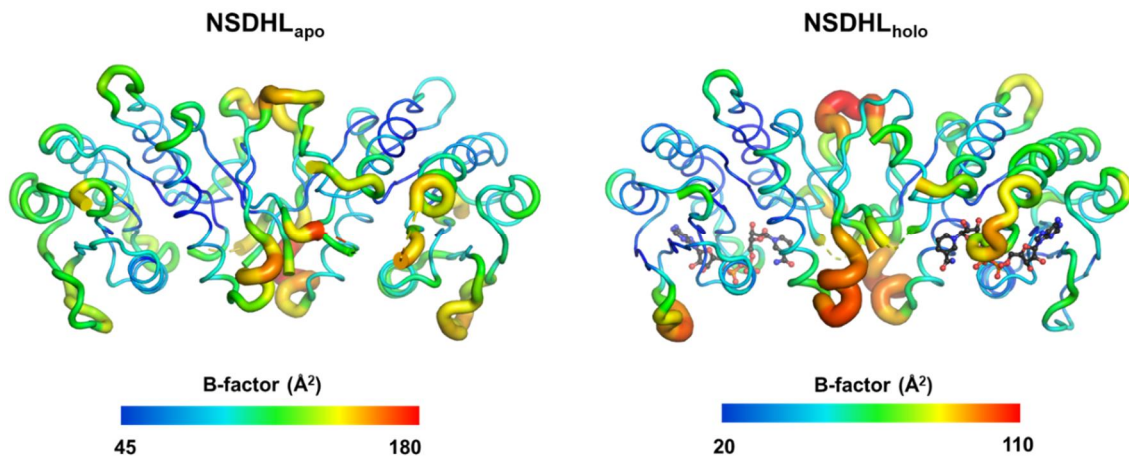


Figure 7. B-factor distributions of two NSDHL structures. The NSDHL structures are shown as a ribbon diagram, and NAD⁺ is shown as a ball and stick model. Each B-factor (Å²) range of both structures is shown as a gradient marker colored from deep blue to red.

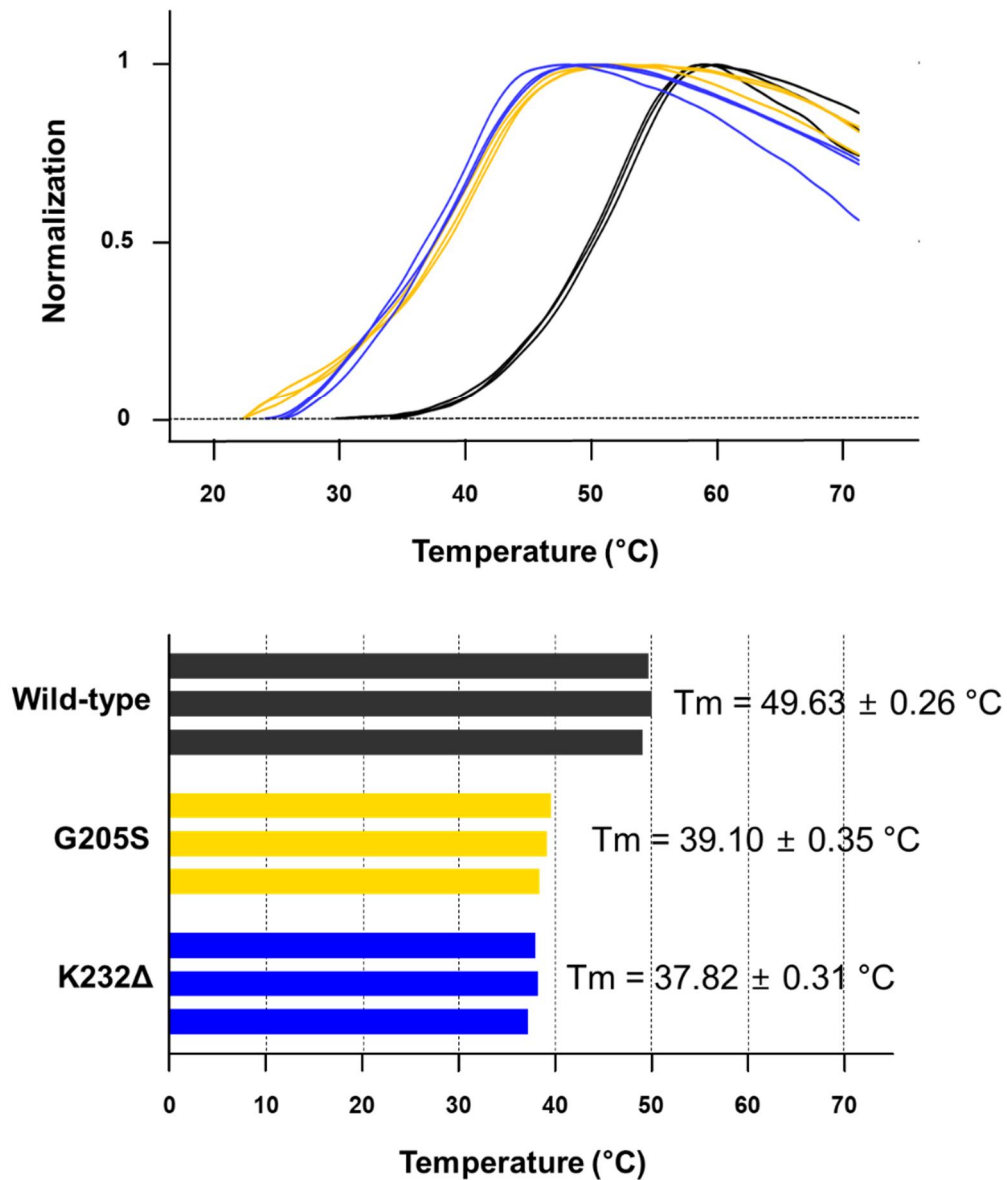


Figure 8. Thermal stability of NSDHL and its mutants. Wild-type, G205S (involved in CHILD syndrome), and K232Δ (involved in CK syndrome) of NSDHL are shown in black, yellow, and blue, respectively. The melting curves are normalized to each maximum value as a fraction. For comparison, a column diagram (below) is shown with the midpoint melting temperature (T_m). The melting points were determined as described in the methods section. The T_m values are presented as the mean \pm SEM of three independent experiments.

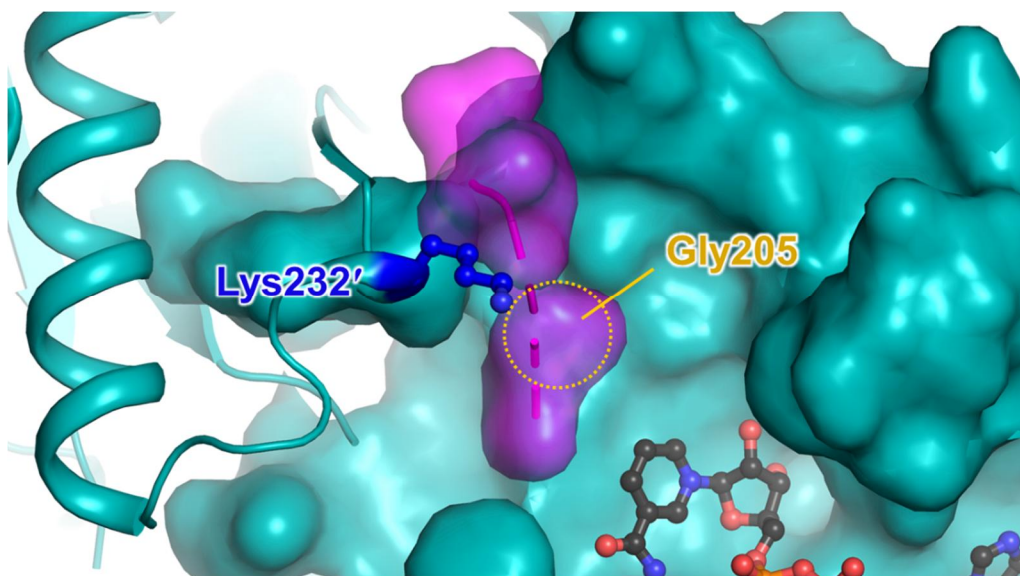


Figure 9. Gly205 and Lys232 in the NSDHL structure. The NSDHL_{holo} (cyan) dimer is shown as the same surface view (A chain) as that shown in Fig. 6C and a ribbon diagram (B chain). The dashed loop (magenta) indicates the disordered regions (Gly202–Pro209). Gly205 is located in the disordered loop and shown as a yellow dotted circle. Lys232' of the other monomer is located proximal to the loop and shown as a ball and stick (blue).

1.3.5 Human NSDHL favorably employs NAD(H) as its coenzyme in the cholesterol synthesis pathway

An in silico prediction of the coenzyme specificity of NSDHL has been reported [47]. As mentioned above, the major factor influencing coenzyme specificity is the presence of the acidic residue Asp68, which is conserved in the NAD(H)-dependent SDR family [44]. To better understand the coenzyme-binding mechanism of NSDHL, we explored the preference of coenzymes by analyzing their binding affinity to NSDHL.

In ITC experiment for the quantification of the binding affinity, the titration curves of NAD⁺ or NADH to NSDHL were fitted well to the single-site binding model. Interestingly, NADH ($K_d = 21.4 \pm 1.1 \mu\text{M}$) showed a 7-fold greater binding affinity to NSDHL than NAD⁺ ($K_d = 151.5 \pm 70.2 \mu\text{M}$) (Figure 10A, B). However, the NADP⁺ and NADPH titrations did not show any significant differences (Figure 10C, D), indicating weak binding affinity beyond the limit (high concentration). We concluded that NSDHL employs NAD⁺ as an electron acceptor for decarboxylation of the C4-methyl sterol in the cholesterol synthesis pathway.

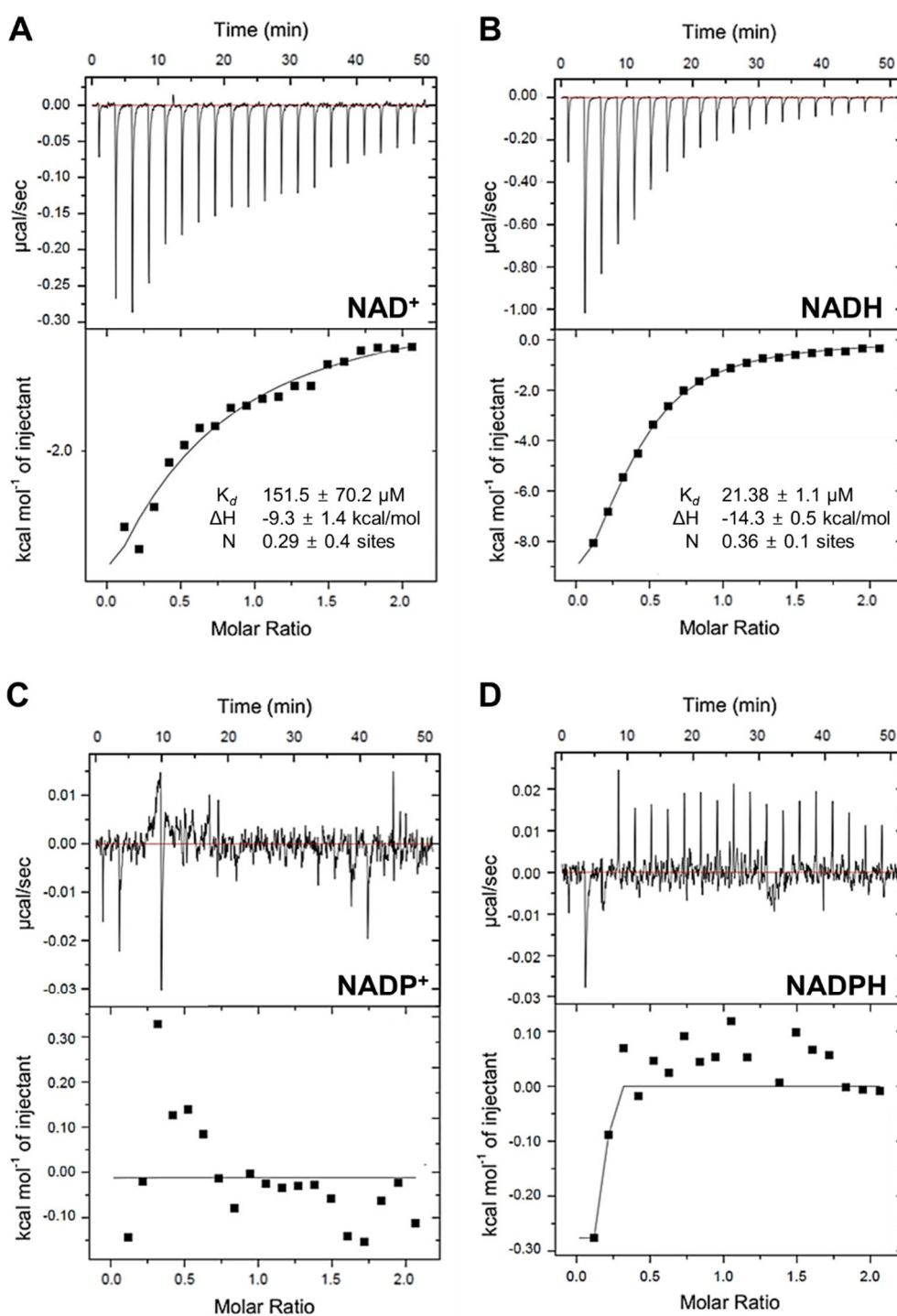


Figure 10. ITC measurements for the coenzyme preference of NSDHL. **A–D** Calorimetric data for the titration of 1 mM of NAD⁺ (**A**), NADH (**B**), NADP⁺ (**C**), and NADPH (**D**) into 0.1 mM NSDHL at 25°C. The calorimetric data for NADP(H) (**C**, **D**) were not determined. A one-site fitting model and a nonlinear least squares algorithm were used to obtain the binding constant (K_d), change in enthalpy (ΔH), and

stoichiometry (N). All values are presented as the mean \pm SEM of three independent experiments.

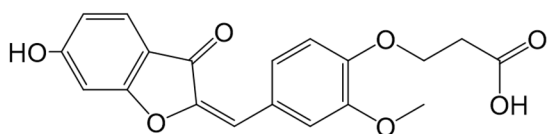
1.3.6 Development of NSDHL inhibitors based on the structural information

Detailed information on the overall structures and their coenzyme-binding site provided the basis for the design of the NSDHL inhibitors. We selected 495 putative compounds from chemical libraries of the Korea Chemical Bank (KCB), which consists of 388,852 molecules, by a structure-based virtual screening approach. Two compounds (**1** and **5**) (Figure 11) in the initial set of 495 compounds were identified to moderately inhibit NSDHL by the NADH-based competitive binding assay. Between these two compounds, compound **5** showed more potent inhibitory activity, with IC_{50} and K_i values of 39.33 μ M and 13.70 μ M, respectively (Figure 11 and Table 3).

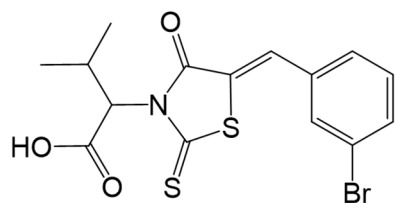
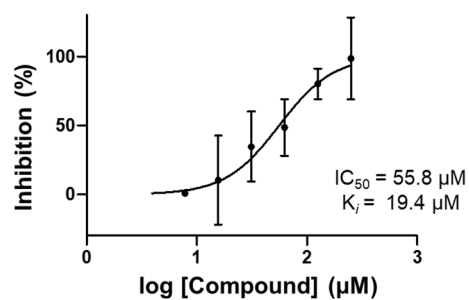
To obtain a potent NSDHL inhibitor, we further conducted a fingerprint similarity search and selected 162 compounds. The NADH-based competitive binding assay indicated that 11 analogs (compounds **6–38**), which were derived from compound **5**, showed various inhibitory profiles (Table 3 and Figure 12). To comprehend the chemical structure of each biochemical property, structure-activity relationship (SAR) analysis was performed. The main scaffold ((*Z*)-5-benzylidene-2-thioxothiazolidin-4-one), and structures and inhibitory potentials of all the 12 obtained analogs are shown in Table 3. In this study, compound **9** was identified as the most potent analog, with an IC_{50} of 8.42 μ M and a K_i of 2.93 μ M, out of all the selected derivatives (Table 3 and Figure 13A). The inhibitory activity of compound **9** was 5-fold greater than that of the parent compound **5**.

Except for compounds **5** and **38**, all analogs have $-CH_2COOH$ groups on R_1 . More hydrophobic substitutions on R_1 , such as $-C(CH_3)_2COOH$ of compound **5** and the benzyl group of compound **38**, may have resulted in no significant change in the activity. In most

analogs, the substitutions on R₂, R₃, R₄, and R₅ with small groups, such as -OH, Br, and OCH₃, rarely influenced inhibitory activity. However, the positioning of a phenylmethanol moiety on R₂ showed increased inhibitory activity. In particular, the presence of a carboxyl group on the phenylmethanol moiety (compound **9**) showed the strongest inhibitory potential. Furthermore, despite the structural similarity with the phenylmethanol moiety, a phenol moiety on R₂ (compounds **13** and **14**) showed a moderately lower inhibitory activity than the phenylmethanol moiety. In addition, when R₃ was substituted with either 4-fluorobenzene (compound **9**) or 1-chloro-2-methylbenzene (compound **6**), a marked difference in the inhibitory potentials was observed. Compound **9**, which was the most potent inhibitory analog in this study, showed 8-fold stronger inhibitory activity than compound **6**, which was the poorest inhibitory analog among the derivatives. However, substituting a benzene ring of the main scaffold with either a pyrazole (compound **32**) or furan ring (compound **35**) still showed a moderate effect on the inhibitory activity.



Compound 1



Compound 5

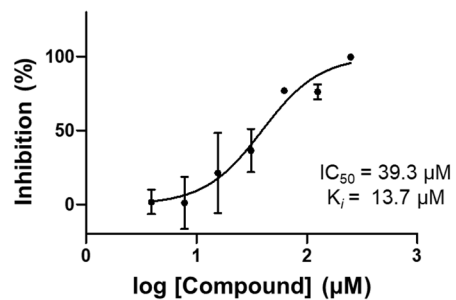
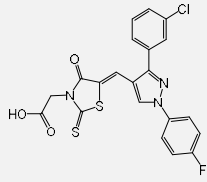
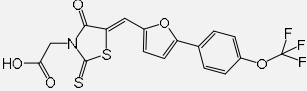


Figure 11. Two parent hits selected from an initial set of 495 compounds. Chemical structures and inhibition curves of compounds **1** and **5**. The points and error bars in the curve indicate the average values and SEMs, respectively. All values are presented as the mean \pm SEM of two independent experiments.

Table 3 Chemical structures and inhibitory potentials of NSDHL inhibitors.

Cpd	Structure	R ₁	R ₂	R ₃	R ₄	R ₅	IC ₅₀ (μ M)	LogIC ₅₀ \pm S.E. (μ M)	K _i (μ M)
5			-	Br	-	-	39.33	1.59 \pm 0.08	13.70
38			OH	-	OH	-	30.71	1.48 \pm 0.03	10.70
6			OH		-	-	70.59	1.85 \pm 0.11	24.59
9					-	-	8.42	0.93 \pm 0.10	2.93
12				Br	-	-	20.99	0.96 \pm 0.09	7.31
16				Br	-	-	27.13	1.43 \pm 0.03	9.45
17				-	OH	-	20.79	1.32 \pm 0.08	7.24
31				-	-		19.85	1.30 \pm 0.05	6.91
13				Br	-	-	46.95	1.77 \pm 0.07	16.35
14				-	Br	-	36.96	1.57 \pm 0.07	12.87

32		11.56	1.11 ± 0.03	4.03
35		24.49	1.39 ± 0.07	8.53

*Data represent the mean of two or three independent experiments. S.E. indicates standard error of LogIC₅₀. Cpd **32** and **35** belong to the other chemical scaffold although they possess the identical 2-thioxothiazolidin-4-one moiety.

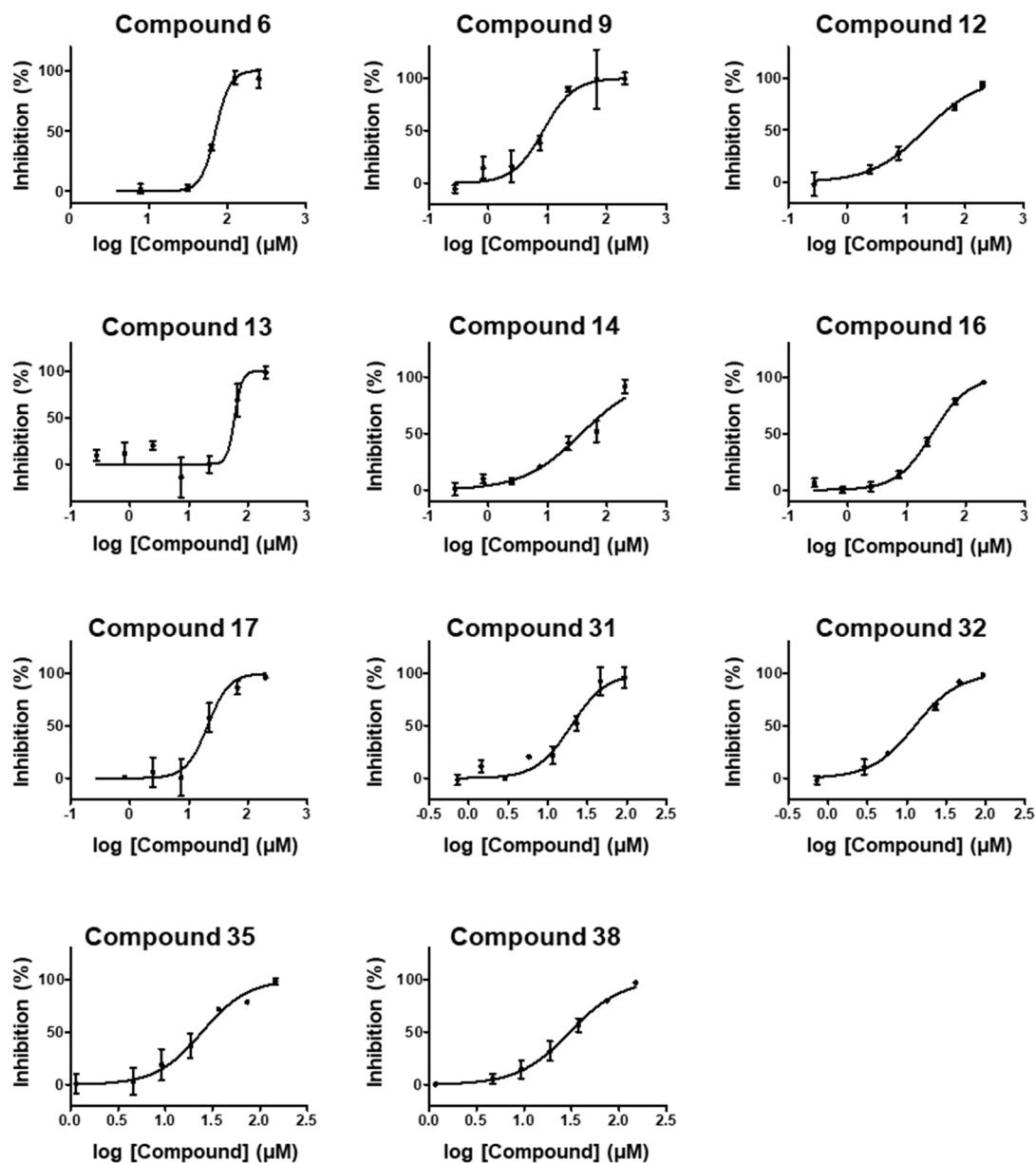


Figure 12. Inhibition curves of 11 analogs derived from the initial hits. Points and bars in the curve indicate the average values and SEMs, respectively. All values are presented as the mean \pm SEM of three independent experiments.

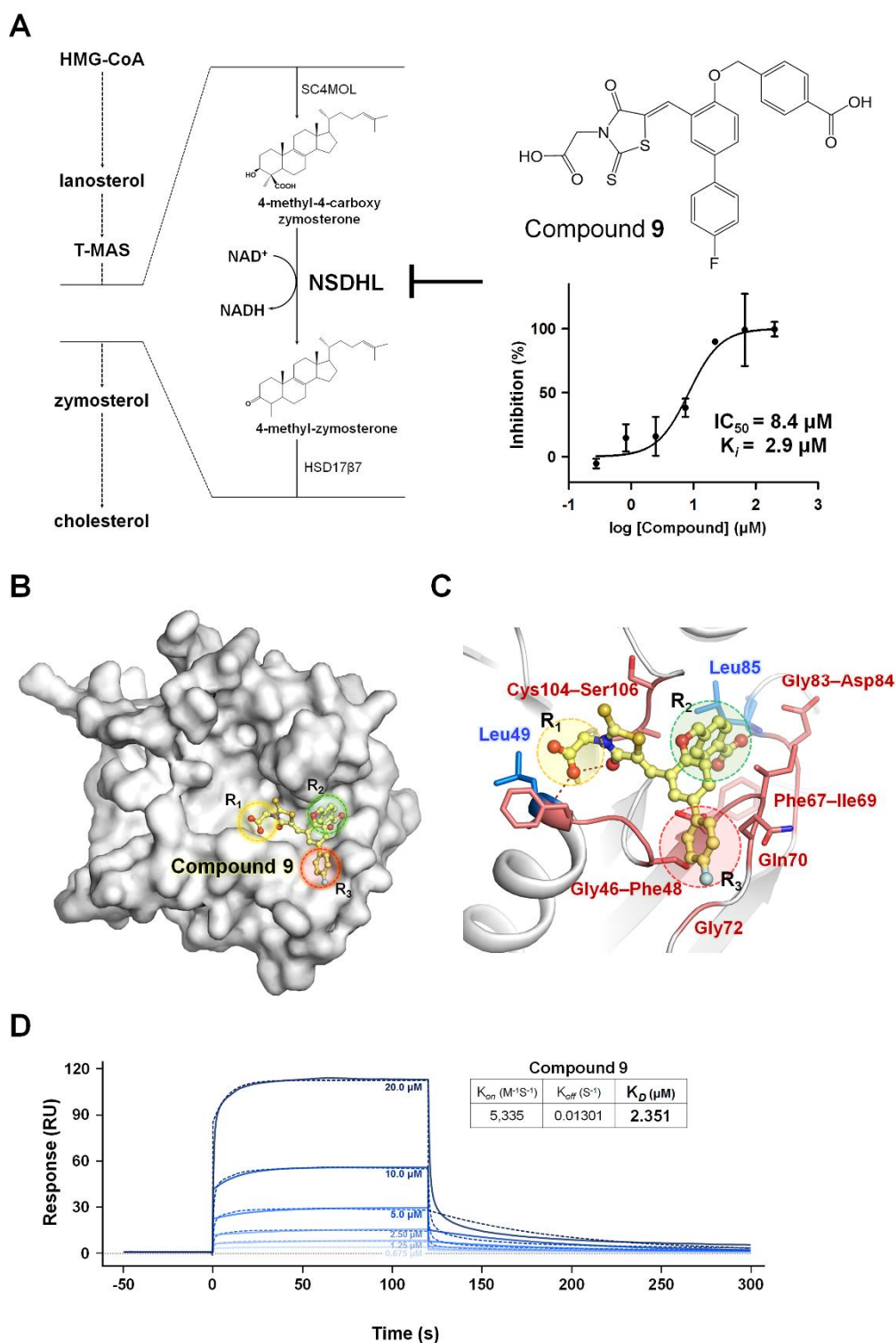


Figure 13. The most potent inhibitor of NSDHL and the proposed docking model. **A** Catalytic mechanism of NSDHL employing NAD^+ as a coenzyme in the cholesterol synthesis pathway, and the chemical structure and inhibition curve of compound **9**. The inhibition of NSDHL by compound **9** is shown as a T line. The data points in the

curve are presented as the average value of three independent experiments and the error bars indicate the SEM. **B** The overall view of a docking model of compound **9** based on the NSDHL_{holo} structure. NSDHL and compound **9** are shown as a surface view (white) and a ball and stick (pale-yellow carbon), respectively. **C** A detailed view showing the proposed binding mode of compound **9**. The hydrophilic interactions are shown as red dotted lines, and the key residues are shown as blue sticks. The residues contributing to hydrophobic interactions are shown as salmon-colored sticks. R₁, R₂, and R₃, indicating key substitution sites, are shown as dashed line circles (yellow, green, and red, respectively). **D** SPR analyses for measurement of the binding affinity between NSDHL and compound **9**. An SPR sensorgram shows the direct binding of with increasing concentrations of compound **9** (0.625, 1.25, 2.50, 5.00, 10.0, and 20.0 μ M) to immobilized NSDHL. The response data and the fitted data are presented as lines and dotted lines, respectively. The calculated K_{on} , K_{off} , and K_D values are also shown on the right side.

1.3.7 Proposed binding mode between compound **9** and NSDHL

To better understand the correlation between the structures and activities of the compounds, molecular docking of the most potent inhibitor, compound **9**, was performed using Schrödinger Suite with the X-ray crystal structure of NSDHL_{holo}. The docking model of compound **9** bound to NSDHL showed the predictive binding mode of compound **9** (Figure 13B). Two polar interactions were predicted in binding between NSDHL and compound **9** as follows: (i) the backbone amide of Leu49 forming a hydrogen bond with the carboxyl oxygen of the -CH₂COOH group positioned on R₁ and (ii) the backbone amide of Leu84 forming a hydrogen bond with the carboxyl oxygen of phenylmethanol positioned on R₂ (Figure 13C). In addition, residues Gly46–Phe48 of the glycine-rich motif GGxGxxG, which is involved in the binding of nucleotides in the SDR family, were found to constitute the binding site cavity for the main core body of compound **9** with Cys104–Ser106 residues. The residues Phe67–Ile69, Gly83, and Asp84 were found to form the hydrophobic pocket for the phenylmethanol group on R₂. Furthermore, Gln70 and Gly72 stabilized the hydrophobic fluorobenzene group on R₃ of compound **9** (Figure 13C).

We further validated that compound **9** directly binds to NSDHL. The sensorgram of SPR showed the binding affinity of compound **9** to NSDHL to be $K_D = 2.351 \mu\text{M}$ (Figure 13D), verifying the direct binding of compound **9** to NSDHL. In addition, we confirmed that the binding of compound **9** to wild-type or mutant NSDHL was also associated with a significant increase in T_m , as determined by fluorescent thermal shift analysis (Figure 14). The temperature-induced unfolding of wild-type NSDHL was largely shifted by a mean of 4.7 °C upon ligand binding. However, the G47S and C104T mutants that were

designed to hinder the binding of compound **9** on the basis of the docking prediction (Figure 13C) showed mean ΔT_m values of 0.6 and 2.0 °C, respectively, much lower than that of the wild type (Figure 14). These results reveal that Gly47 and Cys104 affect the binding of compound **9** to NSDHL, which is consistent with the proposed binding mode of compound **9**.

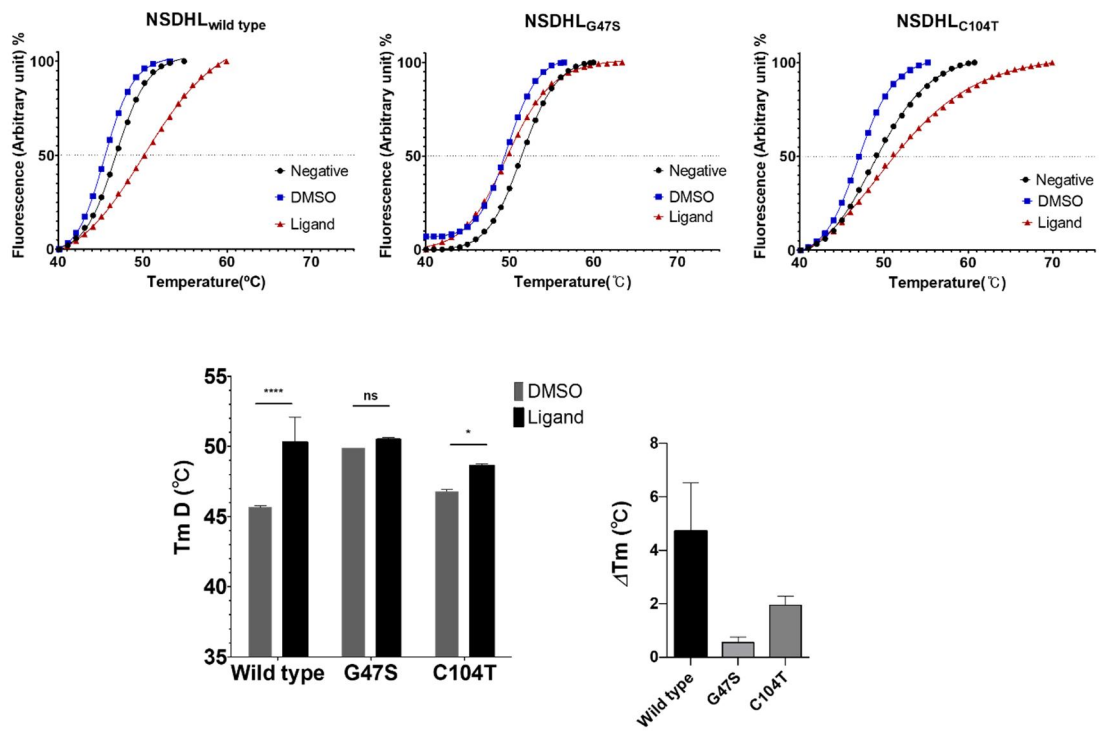


Figure 14. TSA analysis of wild-type NSDHL and G47S and C104T mutants. Values that are significantly different are indicated by a bar and asterisks as follows: ****, P value of ≤ 0.0001 ; *, P value = 0.030; unpaired two-tailed Student's *t*-test.

1.3.8 Therapeutic potential of NSDHL inhibition in EGFR-driven cancer

Drugs targeting the EGFR signaling pathway are strongly linked to the internalization and degradation of EGFR [48]. Considering that the downregulation of EGFR expression followed by depletion of *NSDHL* leads to the sensitization of cancer cells to EGFR-targeting drugs [14], we postulated that our compound can inhibit EGFR activity and increase the antitumor efficacy of EGFR kinase inhibitors. By measuring the cytotoxicity of compound **9** in various epithelial cancer cells expressing EGFR with different response rates to erlotinib, which is an EGFR kinase inhibitor, we confirmed that compound **9** alone did not affect the viability of erlotinib-sensitive and erlotinib-resistant cell lines (Figure 15A) except at the highest concentration (500 μ M). The compound also did not have a cytotoxic effect on three other mammalian cell lines, K562, HepG2, and Vero, except at very high concentrations (500 μ M) (Figure 15B). Considering the weak interaction of nonoptimized compound **9**, we used a high concentration (200 and 500 μ M) of compound **9** to identify the biological effect despite the cytotoxic concentration of compound **9**. Notably, compared to treatment with compound **9** alone, the combined treatment with erlotinib and compound **9** augmented the growth-inhibitory effect of erlotinib in both erlotinib-sensitive and erlotinib-resistant cell lines (sensitization index (SI) < 1) (Figure 16A), indicating that despite the cytotoxic effect of the high concentration of compound **9**, it could induce sensitization toward erlotinib in cancer cells harboring EGFR. To further elucidate the potential utility of NSDHL inhibitors in EGFR-driven human cancers, we additionally investigated the clinical relevance of NSDHL in these types of cancers. The Cancer Genome Atlas (TCGA) datasets showed a high frequency of genomic alterations in *NSDHL* in various types of *EGFR* mutation- or

amplification-harboring human cancers, including lung adenocarcinoma (n = 503), lung squamous cell carcinoma (n = 466), and head and neck squamous cell carcinoma (n = 488) (Figure 17A). In addition, high *NSDHL* expression levels caused by its gene amplification are associated with poor overall survival in lung cancer patients (Figure 17B). No mutations occurred at the key residues of the NAD⁺ binding site in NSDHL in these cancers. Overall, this evidence indicates that the specific inhibition of NSDHL may represent a valuable approach for the treatment of EGFR-driven carcinomas.

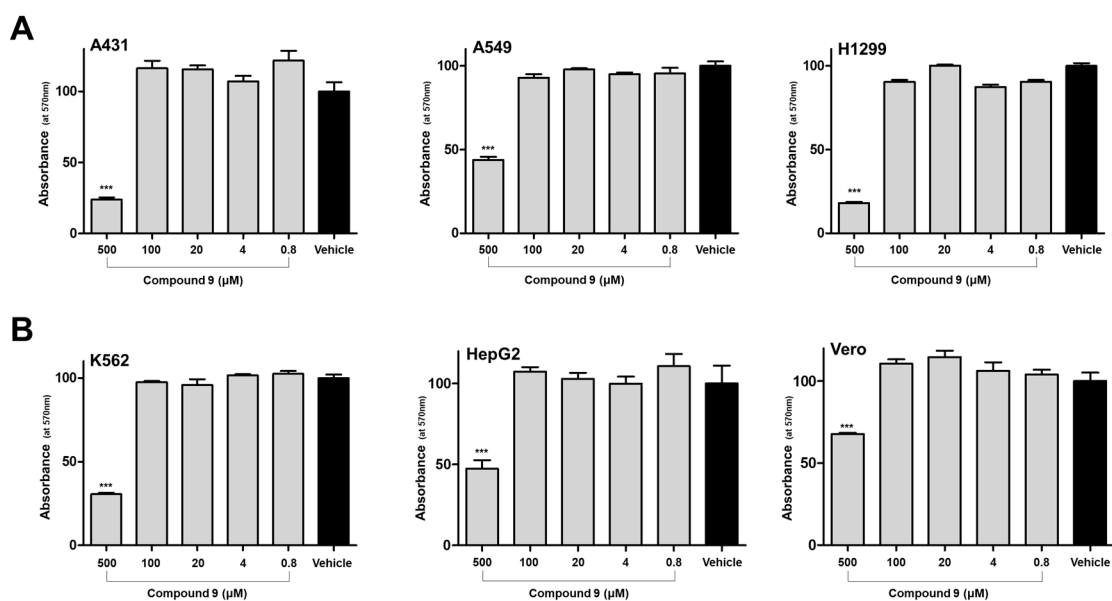


Figure 15. Cytotoxicity of compound **9** in various types of cancer cell lines. **A** Cell viability in the presence of compound **9** in erlotinib-sensitive (A431) and erlotinib-resistant (A549 and H1299) cell lines. **B** Cell cytotoxicity of compound **9** in K562, HepG2, and Vero cells. All cell lines were treated either compound **9** at a concentration of 0.16–500 μM or with vehicle (DMSO). Each column indicates the average value, and the error bars indicate the SEMs (**P < 0.01; ***P < 0.001; unpaired two-tailed Student's *t*-test). The data were determined by three or four independent experiments.

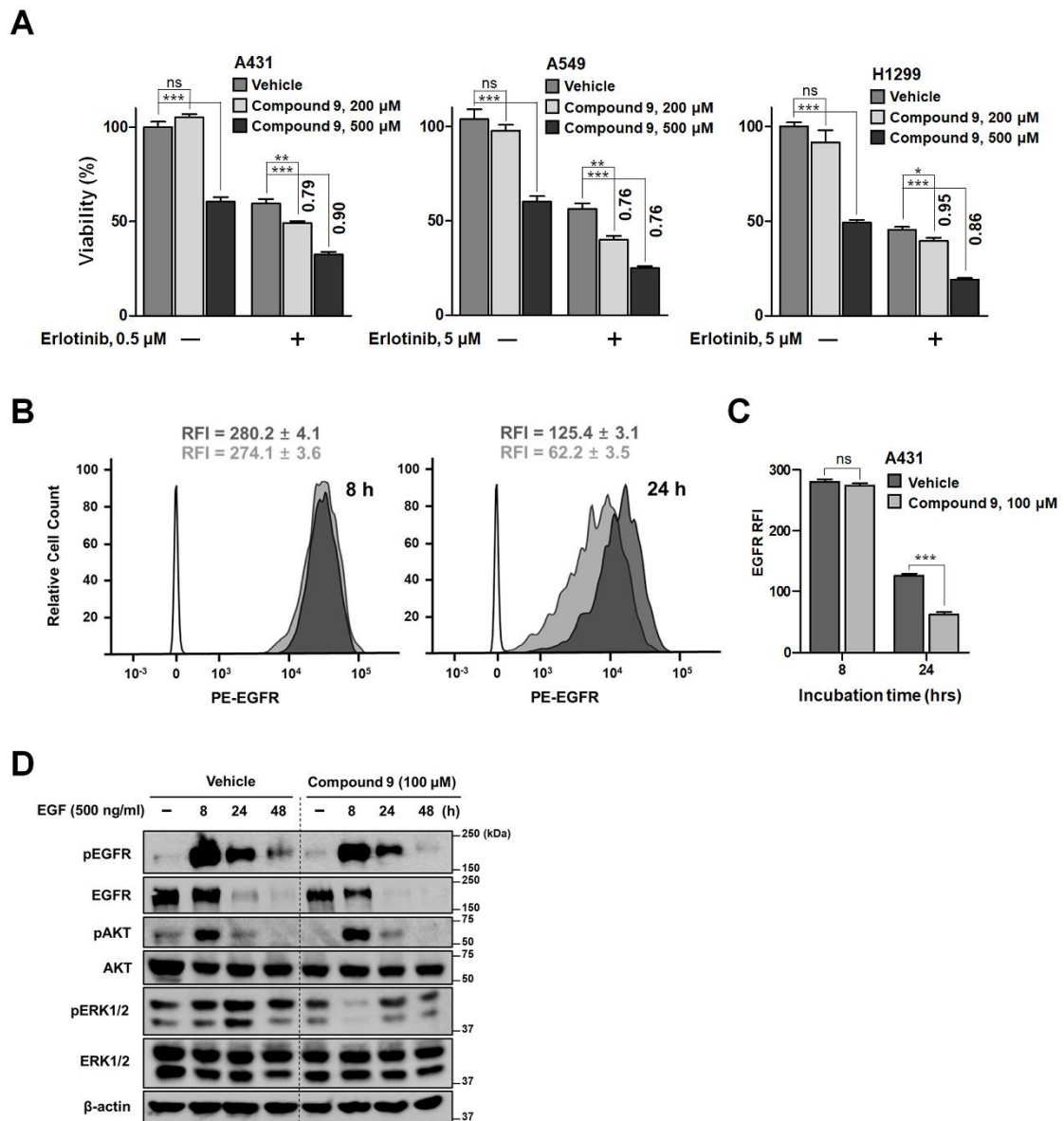


Figure 16. Enhanced antitumor effect of erlotinib with the NSDHL inhibitor and suppression of EGFR activity by compound **9**. **a** Enhanced growth inhibition of erlotinib with compound **9** on erlotinib-sensitive cell line (A431) and erlotinib-resistant cell lines (A549 and H1299) at 48 hours. Numbers above the columns indicate the SI that shows relative increase in erlotinib cytotoxicity in SI < 1 (the equation is described in the Methods section). **b** Flow cytometry analysis of cell surface EGFR expression in response to NSDHL inhibition. A431 cells pretreated with either the vehicle or compound **9** (100 μM) for 72 hours were serum-starved for 24 hours and then stimulated with EGF (500 ng/ml) for the indicated times. The histograms show the amount of phycoerythrin (PE)-conjugated EGFR in the indicated

group of cell lines. The empty peaks indicate the negative control (isotype control antibody), and the filled peaks indicate PE-conjugated EGFR in the vehicle-treated (dark gray) and compound **9**-treated groups (light gray). **c** Comparison of the RFI between the vehicle-treated and compound **9**-treated groups. **d** Lysates from the indicated cells that were prepared as mentioned above (**b**) were subjected to immunoblotting for the analysis of the NSDHL inhibition effect on the EGFR-dependent signaling pathway. All columns represent the average value of three or four independent experiments, and the error bars indicate the SEM (ns, not significant. *P < 0.05. **P < 0.01. ***P < 0.001; ANOVA with Bonferroni's post hoc test and unpaired two-tailed Student's t-test).

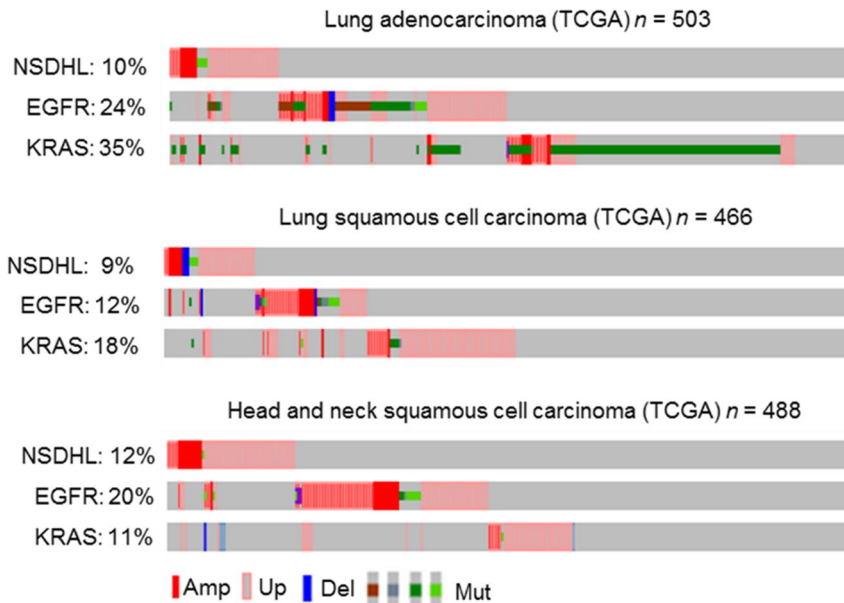
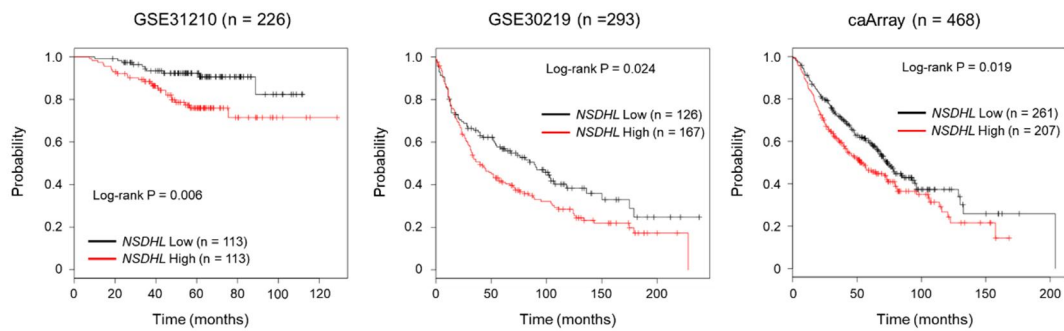
A**B**

Figure 17. Genetic aberration of *NSDHL* and its clinical impact in EGFR-driven cancers. **A** OncoPrints showing the genomic aberrations of *NSDHL* in EGFR-driven human cancers in the TCGA datasets. Amp, amplification; Up, mRNA high; Del, deep deletion; Mut, mutation. **B** KM survival analysis of the OS of patients with lung cancer according to *NSDHL* expression in the indicated datasets. The data were analyzed by the KM method with the log-rank test.

1.3.9 The NSDHL inhibitor accelerates EGFR degradation to suppress EGFR-dependent signaling

The binding of a ligand (i.e., epidermal growth factor (EGF) or TGF α) to the cell surface EGFR naturally induces its internalization, recycling, and degradation [49]. In this endocytosis pathway, NSDHL has been known to play a crucial role in the regulation of the endocytic trafficking of EGFR [14]. In addition, loss of NSDHL activity induces the acceleration of EGFR degradation pathway to lysosomes and rapidly terminates EGFR-dependent signaling [14]. Thus, we hypothesized that our inhibitor may suppress EGFR expression and its signaling pathway, which would subsequently enhance the antitumor effect of erlotinib. To investigate the effect of compound **9** on EGFR endocytic trafficking, we performed a flow cytometry analysis to evaluate the expression levels of cell-surface EGFR in response to treatment with compound **9** in A431 cells, which show an abnormally high-level expression of EGFR. No major shift was observed in the cytometric peak between compound **9** and the vehicle groups at 8 hours after EGF treatment of A431 cells (Figure 16B). At 24 hours, a left-shift of the peaks (Figure 16B) and a reduction in the relative fluorescence intensity (RFI) (Figure 16C) were observed in both groups, indicating a reduction in the EGFR expression level. Interestingly, compared with the vehicle-treated group, the compound **9**-treated group showed substantially decreased fluorescence values ($p = 0.0002$) (Figure 16C), suggesting that the rate of EGFR degradation in lysosomes was increased while the rate of EGFR recycling to the cell surface was decreased by compound **9**.

To further clarify the inhibitory effect of compound **9** on the EGFR signaling pathway, we performed immunoblotting for analysis of the activities of EGFR and its key downstream signaling effectors, such as protein kinase B (PKB, also known as AKT) and

extracellular regulated kinases (ERKs). The EGF-induced phosphorylation levels of EGFR were more rapidly decreased by treatment with compound **9** than in the control. Furthermore, the total EGFR expression began to decrease after 8 hours of EGF treatment in the compound **9**-treated group, while prolonged EGFR expression was observed at that time of EGF stimulation in the control group (Figure 16D). In the absence of EGF, the compound **9**-treated group showed the suppressed phosphorylation of AKT, but comparable phosphorylation levels of AKT in the control group were observed at 8, 24, and 48 hours of EGF stimulation (Figure 16D). Interestingly, compound **9** markedly reduced the phosphorylation of ERK1/2 at 8 hours of EGF stimulation. Then, ERK signaling recovered, but still showed a weaker signal in the compound **9**-treated group than in the control group (Figure 16D). These data indicated that compound **9** induces EGFR degradation and subsequently suppresses the EGFR/ERK signaling pathway. Based on these findings, we concluded that compound **9** may be useful for inhibiting EGFR signaling and improving the antitumor effect of EGFR-targeting drugs on cancer cells (Figure 18).

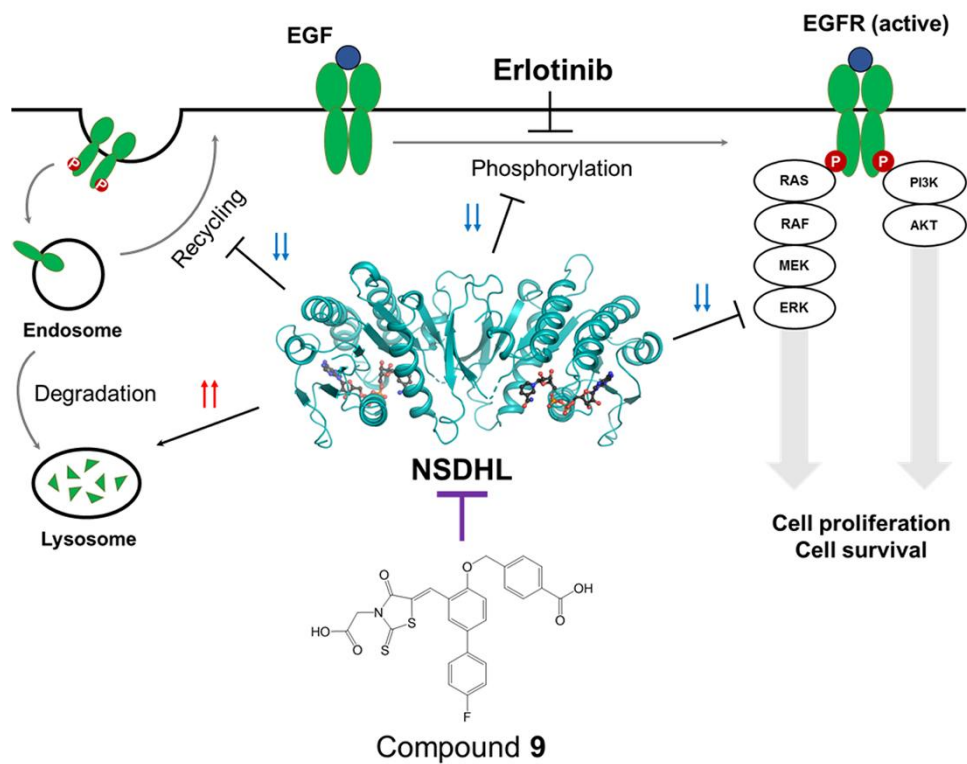


Figure 18. Proposed mode of action of compound 9 in the regulation of EGFR signaling.

1.4 Discussion

Contrary to the research interest in the cholesterol synthesis pathway, investigations of either the NSDHL structure or NSDHL inhibitors were limited. This study provides a foundation for the investigation of NSDHL inhibitors by clearly identifying the structures of NSDHL on the atomic level in both apo and holo forms. This study also identified a novel inhibitor, compound **9**, which was the most potent inhibitor among the 12 analogs included in this study. The biological potential of compound **9** was investigated, and the results showed that it not only induces the suppression of EGFR expression and signaling but also has the enhanced antitumor effect when it is combined with an EGFR kinase inhibitor.

The structural analysis of the two forms of NSDHL identified an early part of the ordered sequential mechanism of NSDHL. The transition from the closed form to the open form, which is induced by the binding of NAD^+ , suggests that NSDHL allows for the uptake of a sterol substrate via a unique structural intermediate. The loop (His201–Leu211), which plays an important role in this conformational change, may contribute to the formation of the active site cavity and regulate specific substrate recognition, which is similar with other SDR family [41]. The disorder of the loop region in our structures may be attributed to the absence of the substrate-binding domain, which was consistently reported in other studies [50]. When the thermal stability of the two NSDHL mutants (G205S and K232 Δ), which cause CHILD and CK syndromes, respectively [45], was investigated, the T_m values of G205S and K232 Δ were similar to physiological temperature and corresponded to previously reported T_m values in a study that demonstrated protein degradation of NSDHL-K232 Δ at 37°C [51]. These two mutation sites were closely related to the loop (His201–Leu211) in our structures, suggesting that

a single mutation of Gly205 or Lys232 may influence the loop orientation (His201–Leu211), which may cause a substantial instability of the NSDHL structure.

Our study revealed that NSDHL strongly prefers NAD(H) over NADP(H). Although NSDHL is expected to employ NAD⁺ as a coenzyme for its catalytic reaction, NADH showed the strongest binding affinity to NSDHL in this study. We suggest that this unexpected affinity may have been attributed to the strict regulation of the cholesterol synthesis pathway to maintain a steady flow of sterol intermediates. After the catalytic reaction, NADH, which is a reactant of NAD⁺, may serve as a product inhibitor with its strong binding affinity and prevent the unnecessary depletion of a sterol precursor, such as isocitrate dehydrogenase [52,53]. By revealing the unique conformation and reaction mechanism of NSDHL, this study provides a foundation for the development of inhibitors. Furthermore, the good fit of NAD⁺ within the coenzyme-binding pocket of NSDHL serves as a rationale for developing small-molecule inhibitors and as an attractive starting platform.

The inhibition of NSDHL and the consequent accumulation of C4-methyl sterol can cause severe X-linked genetic disorders, such as CHILD and CK syndromes [45]. Concerns have been raised on the therapeutic potential of targeting NSDHL for these phenotypes that are bound to be generated in the process of inhibiting NSDHL. Nevertheless, a previous study has demonstrated that a deficiency of *NSDHL* is tolerated in adult animals, suggesting that the therapeutic exploitation of NSDHL could be a viable strategy in adults with minimal toxicity [54]. To the best of our knowledge, the natural compound FR171456 [55], which has been identified as an inhibitor of ERG26/NSDHL, is the only molecule that has been reported for the inhibition of NSDHL. In this study, a combined analysis that included NSDHL-NAD⁺ complex structure-based virtual

screening and in vitro biochemical screening identified compound **9** as the most potent inhibitor among the 12 selected analogs of a novel scaffold that is different with the previous natural compound. The docking model based on the NSDHL-NAD⁺ complex structure suggested that a specific substitution on R₁, R₂, and R₃ of the main scaffold leads to the potent inhibitory activity of compound **9**. Considering that the -CH₂COOH group on R₁ was found to be located near the spacious hydrophilic pocket in the docking model, future studies may identify more potent inhibitors by optimizing and modifying R₁. Identification of the direct binding affinity between NSDHL and compound **9** verified that compound **9** was rationally designed as a competitive inhibitor for the coenzyme binding sites of NSDHL. Furthermore, despite solubility and stability problems of mutants selected near the R₂ or R₃ region, the thermal stability deviation of two mutants selected near the R₁ region with compound **9** indicated that the proposed docking model of compound **9** is reliable.

One of the significant challenges in this study was the therapeutic availability of the NSDHL inhibitor as a novel anti-cancer drug in EGFR-driven human cancer. Consistent with the previous report showing the close connection between NSDHL and EGFR signaling [14], our findings demonstrated that NSDHL inhibition alone did not overcome EGFR-induced cancer cell growth while the combined inhibition with NSDHL and EGFR could promote the sensitivity of cancer cells to an EGFR-targeted drug. This sensitization effect induced by the high concentration of compound **9** in EGFR-driven cancer cells might be considered an off-target effect. However, the validation of the direct binding of compound **9** to NSDHL and the consistent tendency of the biological data induced by compound **9** comparable to the results by gene depletion of NSDHL [14] suggest that compound **9** can be a potential lead compound targeting EGFR-driven cancer for future

study. Considering the clinical evidence on the frequent genomic alteration of *NSDHL* in EGFR-driven cancers and a previous finding that EGFR/KRAS-dependent tumor growth was suppressed by the accumulation of a sterol intermediate [54], a level of NSDHL activity high enough to cause an imbalance in cholesterol synthesis might contribute to the development of EGFR-driven cancer by altering the activity of EGFR. Although further evaluations are required to support the link between NSDHL-mediated cholesterol synthesis and EGFR signaling in the progression of cancer, our findings indicated that NSDHL inhibition has an inhibitory effect against EGFR expression, activity, and downstream signaling, particularly the RAS/ERK pathway, and these results were consistent with observations in previous studies [54],[14]. These previous results and our data indicate that NSDHL inhibitors can be considered a reasonable effector when used with other anti-EGFR drugs that target a broad class of tumors with activated EGFR and RAS oncogenes.

In this study, the structures of NSDHL were unraveled for the first time with the description of the unique conformational transition upon binding of a coenzyme and identified a novel inhibitor based on the elucidated structures. Furthermore, this study verified that the inhibition of NSDHL by compound **9** can suppress EGFR expression and EGFR-dependent signaling and induce sensitization to erlotinib. Therefore, our findings provide valuable insights into the therapeutic potential of targeting NSDHL and serves as a basis for developing therapeutic agents against NSDHL-mediated diseases.

Chapter 2. Structure-based discovery of selective inhibitors against PptT from *M. tuberculosis*

2.1 Introduction

Tuberculosis (TB) is a major global health problem causing ill-health among millions of people each year worldwide. In 2018, there were an estimated 10 million new TB cases and there were also 1.5 million TB deaths [56]. Currently, the emergence of multidrug-resistant (MDR) tuberculosis and extensively drug-resistant (XDR) tuberculosis makes it more difficult to treat TB with existing drugs [57,58]. In addition, Latent TB infection (LTBI) is another problem in treatment of TB. LTBI has no signs or symptoms of acute TB infection before activation. However, it carries a high risk of subsequently developing active TB disease during chronic infection [56,59]. This LTBI makes it more difficult to early detect TB and prolongs a period of treatment [60]. For these reasons, there has so far been urgent necessity to develop the novel therapeutic agents to resolve drug-resistance and the chronic infection problem of TB. Our group has been focusing on 4'-phosphopantetheinyl transferase (PptT) from *Mycobacterium tuberculosis*, the causative agent of TB, to establish innovative drug design for TB treatment.

The mycobacterial envelope plays a significant role in viability and drug resistance of the bacterium by blocking the antibiotics and protecting bacterium from hostile environment [61,62]. This cell envelope contains high lipid contents with uncommon structures such as mycolic acid which are known as contributing to pathogenicity and

persistence of bacterium [62-64]. In biosynthesis of these mycobacterial cell envelope, a number of enzymes contribute to production of these uncommon lipids: polyketide synthases (PKSs), two fatty acid synthases (FASs), and non-ribosomal peptide synthetases (NRPSs) [65]. All these enzymes require posttranslational modification of their acyl carrier protein (ACPs) or peptidyl carrier protein (PCPs) domain to be activated. The inactive form (apo) is converted to be active form (holo) by Mg^{2+} -dependent transfer of the 4'-phosphopantetheinyl (Ppt) moiety of coenzyme A (CoA). PptT, Sfp-type PPTase from *M. tuberculosis*, is the crucial enzyme that mediate this posttranslational modification [66] and essential for the persistence and survival of *M. tuberculosis* during the chronic phase of infection, which was demonstrated in vitro and in vivo [67]. Thus, discovery of novel inhibitors of this enzyme will make it possible to develop new therapeutic drugs for the treatment of mycobacterial infections.

In this study, we determined the crystal structures of PptT from *M. tuberculosis* complexed with CoA and PPTases from *Mycobacterium smegmatis* complexed with CoA or adenosine-3',5'-diphosphate (3',5'-ADP). Based on structural information, the rational virtual screening and PptT-BpsA coupled assay were carried out to find the efficient candidates inhibiting PptT. The selected compounds were further validated for the selectivity and biological effect. Taken together, our findings herein provide considerable insights on the structures of mycobacteria PPTases, which resulted in the development of a novel and selective inhibitor.

2.2 Materials and Methods

2.2.1 Gene Cloning, protein expression, and purification

The gene, *PptT*, encoding *M. tuberculosis* PPTase, was amplified from *M. tuberculosis* (H37RV strain) genomic DNA by the polymerase chain reaction (PCR) using the forward and reverse primers, 5'-GCCG CAT ATG ACG GTA GGC ACG CTG GTG G-3' and 5'-CCGCCG CTC GAG TAG CAC GAT CGC GGT CAG CAC-3'. The *MSMEG_2648* encoding *M. smegmatis* PPTase, was amplified by PCR using *M. smegmatis* (*mc(2)155* strain) genomic DNA as a template. The forward/reverse primers of *MSMEG_2648* are 5'-GGAATTC CAT ATG ACG GAC AGC CTG CTT TCC C-3'/5'-CCGCCG CTC GAG CAG GAC GAT CGC CGT CAC GG-3'. The underlined sequences in all primers are restriction enzyme sites, NdeI and XhoI, respectively. All genes were inserted into vector pET-21a(+) (Novagen) including eight-residue tag (LEHHHHHH). The identity of all inserts was confirmed by commercial sequencing service. The recombinant PptT protein was overexpressed in *E. coli* BL21(DE3) cells using Luria-Bertani (LB) culture medium with plasmid-appropriate antibiotics. Cells were grown at 37°C until an OD600 0.7 and cooled in 15°C incubator for 30 min before addition of 0.5 mM isopropyl 1-thio-β-D-galactopyranoside (IPTG). After induction, cells were incubated for additional 24 hours at 15°C. The cells were harvested by centrifugation at 4,300 × g for 10 min. The cell pellet was resuspended in lysis buffer [50 mM MES, pH 5.8, 500 mM NaCl, 10 mM imidazole, 10 mM MgCl₂ and 10% (v/v) glycerol] and then sonicated at 4°C. The lysate was centrifuged at 18,000 × g for 1 hours at 4°C to remove cell debris. The supernatant was filtered using 0.45μM syringe filter (Millipore) and loaded to a Hi-Trap Chelating

Column (GE Healthcare life sciences) that was pre-equilibrated with binding buffer [50 mM MES, pH 5.8, 500 mM NaCl, and 10% (v/v) glycerol]. The soluble fractions were eluted from a range of 200–300 mM Imidazole. After concentrating the collected fractions, the protein was further purified by gel filtration on a HiLoad 16/600 Superdex 200 prep-grade column (GE Healthcare life sciences) that was pre-equilibrated with Buffer A [20 mM MES, pH 5.8, 150 mM NaCl, 5 mM Dithiothreitol (DTT) and 1 mM MgCl₂]. The final protein was collected and concentrated to 11 mg/ml using an Amicon Ultra-15 centrifugal filter unit (Millipore). Prior to crystallization, 1 μM of CoA tri-lithium salt (Sigma) was added to final protein solution.

The recombinant *MSMEG_2648* protein was also overexpressed in *E. coli* BL21(DE3) cells using LB culture medium. Purification was performed in the same method as described for PptT. In case of *MSMEG_2648* protein, crystallization was carried out with 10 mg/ml protein solutions containing 1 μM of CoA and not containing, respectively.

For BpsA-PptT coupled assay, pCDFDUET vector containing *BpsA* gene, encoding indigoidine synthase, was provided from David F. Ackerley [68]. BpsA protein was obtained using a Hi-Trap Chelating Column (GE Healthcare life sciences) that was pre-equilibrated with binding buffer [50 mM Tris, pH 8.0, 500 mM NaCl, 10% (v/v) glycerol] without additional purification steps and stored at -80°C with 25 % (v/v) glycerol.

2.2.2 Crystallization and X-ray data collection

Crystals of PptT were grown using the sitting-drop vapor diffusion method on empty reservoir at 4°C over 6 days. The qualitative crystal was grown in a 1.8 : 1.2 μl drop

volume consisting of 1.8 μl protein solution and 1.2 μl crystallization buffer [30% (w/v) PEG 8000, 100 mM Sodium acetate/acetic acid, pH 4.5, 200 mM Lithium sulfate] in Wizard 1 (Emerald Biosystem). Crystals of MSMEG_2648 were grown in both presence and absence of 1 μM CoA and not containing CoA. Crystallizations were carried out in the sitting drop vapor diffusion method by mixing equal volumes (0.3 μl each) of protein solution and the reservoir solution. Crystals in the presence of 1 μM CoA were grown in a variety of crystallization solutions and the qualitative crystals were produced in crystallization buffer [100 mM Tris, pH 8.5, 250 mM magnesium chloride, 40% PEG 1000] in MemGold2TM (Molecular Dimensions) at 20°C. Crystals of MSMEG_2648 in the absence of 1 μM CoA were grown in crystallization buffer [0.2 M MES, pH 6.5, 0.2 M calcium chloride, 33%(v/v) PEG4000] in MemGold1TM (Molecular Dimensions) at 4°C. Mixture of crystallization buffer solution and 15% (v/v) ethylene glycol was used as a cryo-protectant solution prior to data collection under nitrogen gas stream. The data of PptT crystals were collected using ADSC Quantum Q270r CCD detector at the beamline 7A of Pohang Light Source, Republic of Korea. The data of MSMEG_2648 crystals obtained in presence and absence of CoA were collected using ADSC Quantum 315r CCD detector at the beamline 5C of Pohang Light Source, Republic of Korea.

The crystals of PptT belong to the orthorhombic space group $C222_1$ with unit cell parameters of $a = 97.77 \text{ \AA}$, $b = 121.44 \text{ \AA}$, $c = 49.07 \text{ \AA}$, $\alpha = 90^\circ$, $\beta = 90^\circ$, and $\gamma = 90^\circ$. The crystals of MSMEG_2648 with 1 μM CoA and without 1 μM CoA belong to the orthorhombic space group $P2_12_12_1$ with unit cell parameters of $a = 67.55 \text{ \AA}$, $b = 78.78 \text{ \AA}$, $c = 81.88 \text{ \AA}$, $\alpha = 90^\circ$, $\beta = 90^\circ$, and $\gamma = 90^\circ$, and the monoclinic space group $C2_1$ with unit cell parameters of $a = 117.35 \text{ \AA}$, $b = 73.98 \text{ \AA}$, $c = 65.39 \text{ \AA}$, $\alpha = 90^\circ$, $\beta = 95^\circ$, and $\gamma = 90^\circ$, respectively. All data was processed and scaled with HKL2000 [16].

2.2.3 Structure determination, refinement, and analysis

The structure of *M. tuberculosis* PptT (referred to Mtb-PptT in this study) was determined by the molecular replacement (MR) method in Molrep [69] using the structure of *M. tuberculosis* PptT (PDB code 4QJK) [70] as a search model. The two structures of *M. smegmatis* (referred to Ms-PPTase1 and Ms-PPTase2 in this study) were also solved by MR method in Molrep [69] using the refined Mtb-PptT as a template. In all models, the initial structures were further refined using the Refmac program [19]. The models were manually constructed and the library parameters for ligand, CoA, was generated using The PRODRG Server [71]. Water molecules were added using the Coot program [20] and manually inspected by identifying hydrogen bond interaction with surrounding environment. Five percent of the data were randomly set aside as the test data for the calculation of R_{free} [21]. The stereochemistry of the final structures was evaluated using MolProbity [22]. The structural deviations were calculated using SSM superpose option in Coot program [20]. The solvent-accessible surface areas were calculated using PISA [23]. The protein-ligand interactions were calculated using the LigPlot [24].

2.2.4 High throughput virtual screening

Docking-based virtual screening was performed using Schrödinger Suite 2017-4 with the X-ray crystal structure of PptT. Protein preparation was revised using Protein Preparation Wizard in Maestro v11.4 software (Schrödinger Release 2017-4: Maestro 11.4, Schrödinger, LLC, New York, NY, USA, 2017), and the receptor grid box was generated

with a 30 Å x 30 Å x 30 Å cube centered on complexed coenzyme A. The 388,852 molecule-containing KCB library was browsed for ligand preparation using LigPrep v4.4 (Schrödinger Release 2017-4: LigPrep 4.4, Schrödinger, LLC, New York, NY, USA, 2017) and the OPLS_2005 force field [25], and during the process, tautomer and ionization states at pH 7.0 ± 2.0 were generated using the Epik v4.2 module (Schrödinger Release 2017-4: Epik 4.2, Schrödinger, LLC, New York, NY, USA, 2017). Ligand docking was performed using Glide v7.7 (Schrödinger Release 2017-4: Glide 7.7, Schrödinger, LLC, New York, NY, USA, 2017) and the standard precision (SP) method. The top compounds were selected according to the Glide score. Then, rescoring of the top-ranking SP pose of each compound was conducted with the extra precision (XP) Glide scoring function. Based on the top 2000 compounds combined with visual inspection, 495 compounds were selected for biochemical testing. Among the 495 compounds that underwent in vitro biochemical screening, seven compounds were identified as new chemical inhibitors. To find more analogs of active compounds, we performed a second round of virtual screening based on the fingerprint similarity method, and 286 compounds were selected by visual inspection. Among the 286 compounds, 4 compounds were additionally selected as potent inhibitors by in vitro biochemical screening.

2.2.5 BpsA-PptT coupled assay for searching inhibitors

Indigoidine, a blue colored product of BpsA reaction, was unstable and water insoluble, but has known to be stable in DMSO. The assay system was optimized by finely tuning DMSO concentration. 6-nitroso-1,2-benzopyrone (6-NOBP) that has been previously reported as an inhibitor of Sfp-type PPTase was established in clear bottom 96-well microplates (Greiner Bio-One). Reaction Mixtures (2 µM PptT, 20 µM CoA, 50 mM Tris,

pH 8.0, 8 mM L-glutamine, 5 mM ATP, 10mM MgCl₂) containing three different concentration (20%, 32%, or 36%, respectively) of DMSO were then added in 150 µl. Reaction was initiated by adding 50 µl of Initiation Mixture (10 µM BpsA, 50 mM Tris, pH 8.0, 8 mM L-glutamine, 5 mM ATP, 10 mM MgCl₂ and 20 % DMSO). Reaction value in each DMSO concentration was measured using SpectraMax M5 (Molecular Devices) at 590 nm at room temperature. The assay of the selected inhibitors for PptT and MSMEG_PPTase was performed as described above. The reaction was the most stable in 36% concentration of DMSO. IC₅₀ curves were fitted using the non-linear regression function of GraphPad Prism5 (GraphPad software). The data were normalized were generated using the Cheng–Prusoff equation for competitive inhibition.

2.2.6 2. Cytotoxicity in eukaryotic cell

To examine cellular toxicity, Vero cells were seeded on a 96-well plate at 2×10^4 cells per well in the RPMI medium containing 10% FBS. Drug was first added at 1000 µM and then two-fold serially diluted eight times. Following incubation at 37°C for two days, cellular viability was measured using MTT dye and a commercial kit (Promega G4100) following its instructions. Absorbance of dyed cellular solution was measured at 570 nm.

2.2.7 Bactericidal activity in mycobacteria-infected murine macrophages

J774.1 cells (Murine macrophage cell line) were seeded on a 96-well microplate at 2.5×10^5 cells per well in Roswell Park Memorial Institute (RPMI)-1640 medium containing

10% FBS. The macrophages were infected with *M. tuberculosis* Erdman at a multiplicity of infection (MOI) of 2.5 for 4 hours at 37°C. Extracellular bacteria were removed by washing three times with the same medium, and then tested drug, Cpd1-9, was added at 0 to 50 μ M following dilutions by the media. Media was changed every sampling day (0, 1, 3 days). Intracellular bacteria were quantified by lysing the cells with 0.1% saponin at the indicated time points and plating dilutions on 7H10 agar.

2.3 Results

2.3.1 Overall structures of mycobacteria PPTases

In this study, we obtained three crystal structures to guide the design of PPTase inhibitors as follows: (i) A structure of PptT from *M. tuberculosis* complexed with CoA (Mtb_PptT) (Figure 19A); (ii) A structure of PPTase from *M. smegmatis* complexed with CoA (MSMEG_PPTase I) (Figure 19B); (iii) A structure of PPTase from *M. smegmatis* complexed with adenosine 3',5'-diphosphate (3',5'-ADP) which is a part of CoA (MSMEG_PPTase II) (Figure 19B).

Mtb_PptT was refined to $R_{\text{work}} / R_{\text{free}} = 19.1\% / 22.8\%$ at a 1.78 Å resolution, and MSMEG_PPTase I and MSMEG_PPTase II were refined to $R_{\text{work}} / R_{\text{free}} = 22.8\% / 29.1\%$ at 1.59 Å resolution and $R_{\text{work}} / R_{\text{free}} = 17.5\% / 20.1\%$ at 1.87 Å resolution, respectively. Mtb_PptT was found to exist a monomer consisting of 226 amino acid residues, one CoA, one Mg^{2+} ion, four SO_4^{2-} and 94 water molecules in the asymmetric unit. Three residues (Met1–Val3) and the C-terminal hexa-histidine residues are disordered. MSMEG_PPTase I was found to exist two monomers consisting of 444 amino acid residues, six Ca^{2+} ions, two CoA, two pentaethylene glycols and 414 water molecules in the asymmetric unit. In this model, two Ca^{2+} exists in the active site of the monomer. One Ca^{2+} bound to different site, which is separated from the active site of the monomer. Two residues (Met1–Thr2) and the C-terminal hexa-histidine residues are disordered in all chain. Two monomers of MSMEG_PPTase I in the asymmetric unit are highly similar to each other, with an r.m.s. deviation of 0.1325 Å for 221 C α atom pairs. MSMEG_PPTase II was found to exist two monomers consisting of 440 amino acid residues, two 3',5'-ADP molecules, two Mg^{2+} , and 163 water molecules in the asymmetric unit. Three residues (Met1-Asp3) and the C-

terminal fusion residues are disordered in all chains. Two monomers of MSMEG_PPTase II in the asymmetric unit are highly similar to each other, with r.m.s. deviations of 0.4102 Å for 220 C α atom pairs.

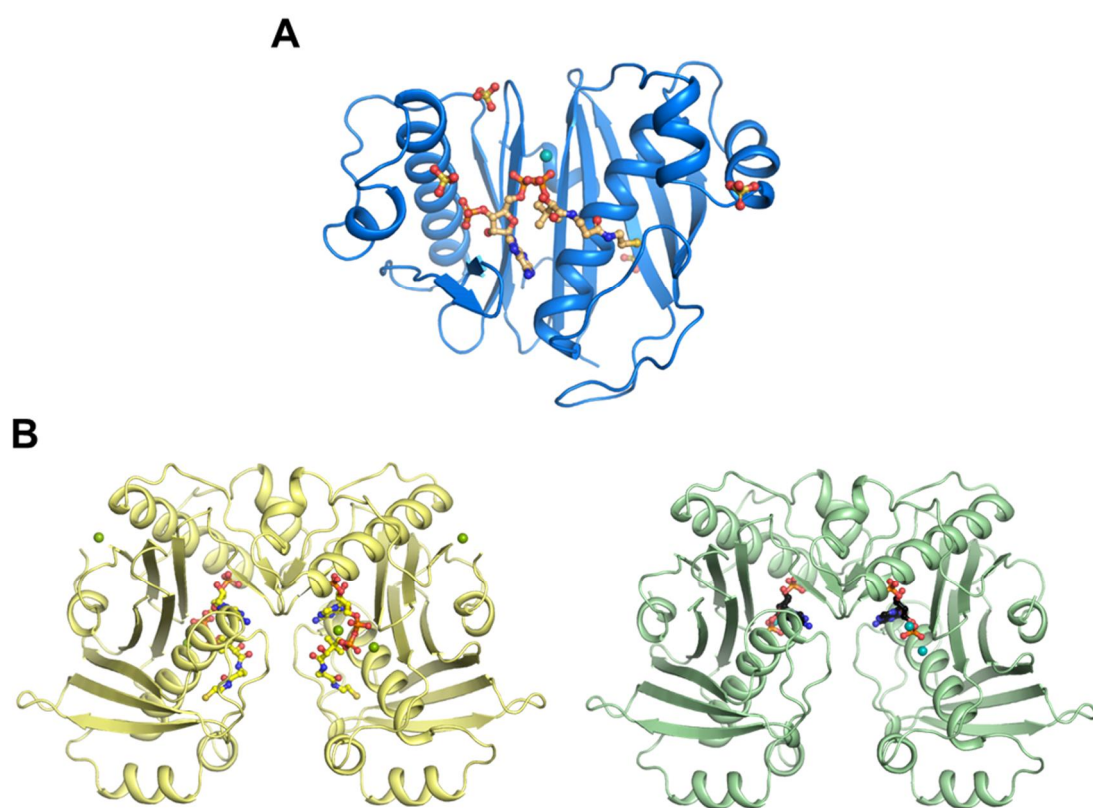


Figure 19. Overall structure of mycobacteria PPTases. **A** Crystal structure of Mtb_PptT complexed with CoA represented as ribbon diagram colored blue. **B** Crystal structure of MSMEG_PPTase I and II complexed with CoA and 3',5'-ADP represented as ribbon diagram colored yellow and lightgreen, respectively.

2.3.2 Structural features and active sites of mycobacteria PPTases

Mtb_PptT and MSMEG_PPTase share the sequence identity of 76% and include two common α/β domains with pseudo 2-fold symmetry, which is the characteristics of Sfp-type PPTases. In the structure of Mtb_PptT, the N-terminal domain has a three-stranded antiparallel β -sheet with a long α -helix, and the C-terminal domain has a five-stranded antiparallel β -sheet with a long α -helix. Both domains are decorated with loops and short helices and between two domains, the binding site for the CoA and associated Mg^{2+} ion is observed (Figure 20). In the CoA binding site, eleven hydrogen bonds were found between the surface residues of Mtb_PptT and CoA, in addition to the interaction of Mg^{2+} ion with Mtb_PptT and CoA (Figure 20). In the 3',5'-ADP moiety of CoA, Tyr160 and a backbone amide of Lys78 showed hydrogen bonds with the adenine ring, and Arg48 and Arg56 formed hydrogen bonds with the 5'-phosphate. In the phosphopantethein moiety, the flexible moiety was stabilized by the hydrophilic interactions with Glu157 and Leu171 (Figure 20). The bridging of two phosphopates by a Mg^{2+} ion was observed consistently with the catalytic reaction of this enzyme that breaks the phosphophate bond using Mg^{2+} for transferring a phosphopantethein moiety. Furthermore, the interaction residues of His93, Asp114, Ala115, and Glu116 with Mg^{2+} were identified to play a crucial role in the catalysis of PptT (Figure 20).

The overall structures of MSMEG_PPTases are highly similar to that of Mtb_PptT. In addition, the key residues involved in the binding of CoA in MSMEG_PPTase I did not show the significant difference from that in Mtb_PptT. However, MSMEG_PPTase I contained two Ca^{2+} ions instead of Mg^{2+} in the active site (Figure 21A), which may be due to the crystallization solution containing 0.2 M calcium chloride. Interestingly, the

MSMEG_PPTase II showed the complexed structure with 3',5'-ADP (MSMEG_PPTase II) which is the catalytic product of CoA (Figure 21B). In addition, this structure possessed two Mg²⁺ ions interacting with the oxygen atoms of 3'-phosphate of 3',5'-ADP unlike MSMEG_PPTase I. Two Mg²⁺ ions formed the interactions with His88 and Asp109 corresponding to His93 and Asp114 of Mtb_PptT, and additional waters (Figure 21C). Although the high sequence identity, this structural analysis suggests that PptT from pathogenic *M. tuberculosis* and PPTase from non-pathogenic *M. smegmatis* may have a distinct catalytic mechanism and the difference may allow to develop the selective inhibitor against pathogenic bacteria.

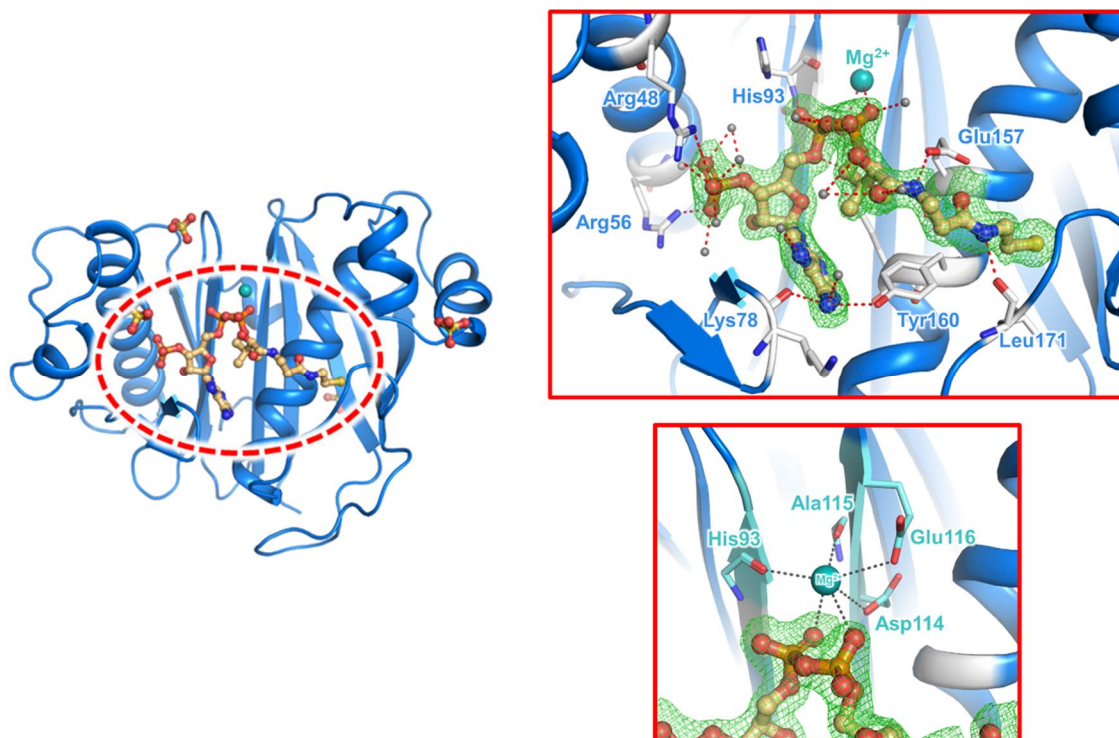


Figure 20. Active site of Mtb_PptT. CoA binding site is enlarged into red box (right-side). The key residues for the binding of CoA and Mg²⁺ are labeled and the interaction bonds are shown as red- or black-dotted lines. Gray spheres indicate water molecules.

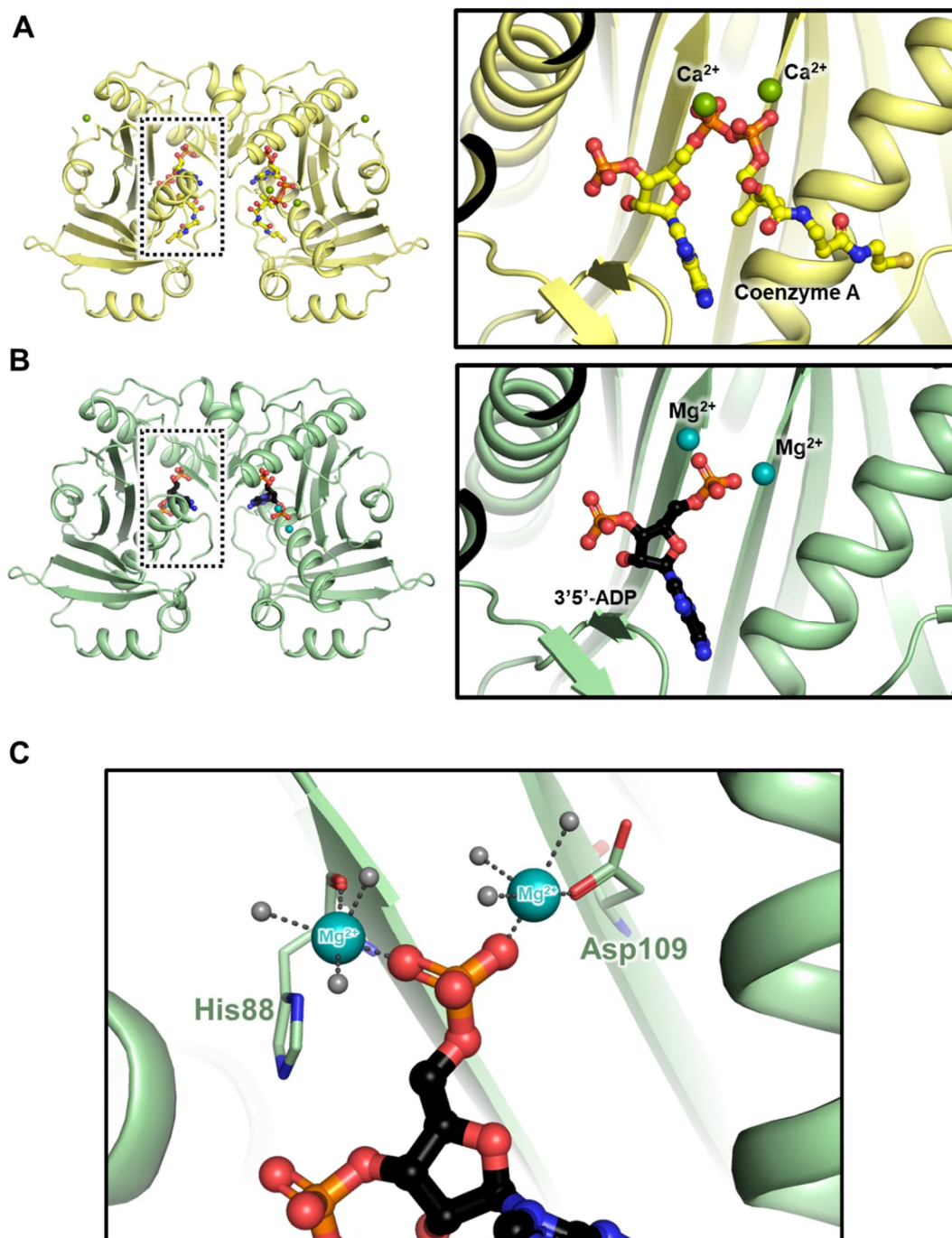


Figure 21. Active site of MSMEG_PptT. **A,B** CoA binding site is enlarged into red box (right-side). **C** The key residues for the binding of CoA and Mg^{2+} are labeled and the interaction bonds are shown as red- or black-dotted lines. Gray spheres indicate water molecules.

2.3.3 Development of small molecule inhibitors to suppress PptT activity.

Before the design of compounds inhibiting PptT, we had to establish the high-throughput enzymatic assay. From the previous research, we adapted the extensive PptT inhibition assay coupled with BpsA which catalyses the conversion of two molecules of L-glutamine into indigoidine, a blue pigment that can be readily detected at 590 nm absorbance [68]. To confirm the stability of the assay system, we performed PptT-BpsA coupled assay using 6-nitroso-1,2-benzopyrone (6-NOBP), which has been reported as the inhibitor against Sfp-type PptT, and the result showed the significant inhibition of PptT (Figure 22A). However, the unstability and insolubility of the indigoidine, which is blue pigment and a product of BpsA, resulted in the early peak time to carry out the high-throughput screening. To stabilize the indigoidine, we tuned the concentration of DMSO in the assay system. Indeed, as increasing the DMSO concentration, the assay was stabilized and then we adapted 32% DMSO concentration (Figure 22B).

An initial set of 495 putative compounds was selected from chemical libraries of the Korea Chemical Bank (KCB) by a structure-based virtual screening approach. In the first round, PptT-BpsA coupled assay identified seven compounds (**Cpd1-2**, **1-3**, **1-6**, **1-7**, **1-9**, **2-3**, and **2-8**) inhibiting the enzymatic activity of PptT (Table 4). To obtain more potent PptT inhibitors, we further conducted a fingerprint similarity search and selected 286 compounds. The PptT-BpsA coupled assay identified the additional four derivatives (**Cpd1-4**, **2-1**, **2-2**, and **2-7**) showing various inhibitory profiles (Table 4). To comprehend the chemical structure of each biochemical property of total eleven analogs, structure-activity relationship (SAR) analysis was performed. According to a type of the main scaffold, the potential inhibitors were classified into 2 groups (A and B) shown in Table

4. Among these compounds, we selected four compounds (**1-3**, **1-4**, **1-6**, and **1-9**) as the most potent analogs according to the inhibitory potential (IC₅₀ values of 0.635, 0.540, 0.637, and 1.975 μ M, respectively) (Figure 23).

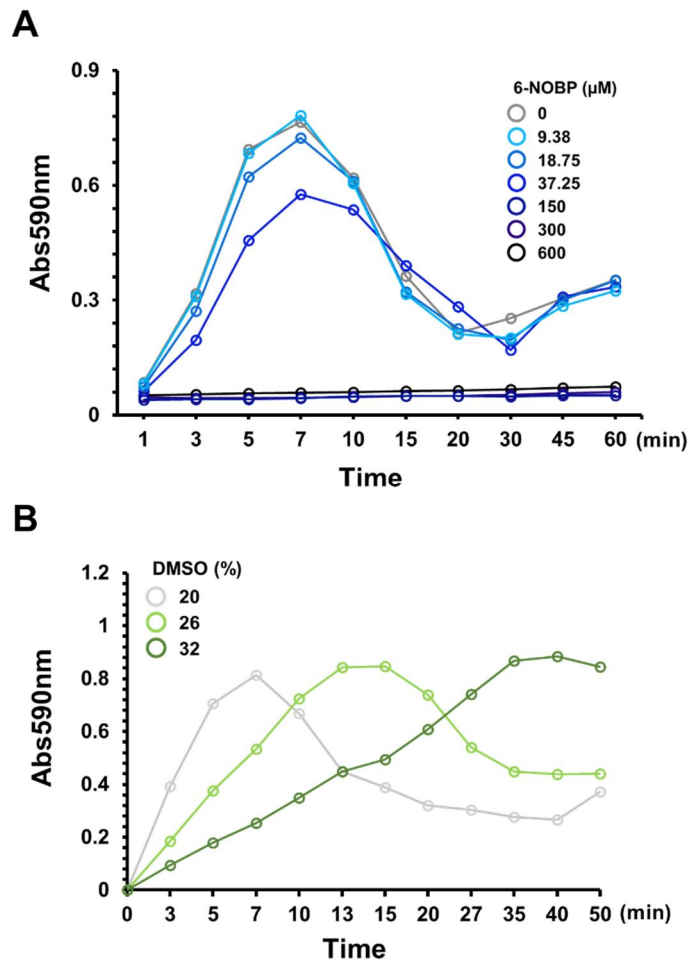
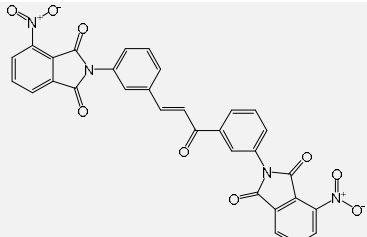
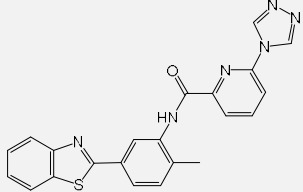
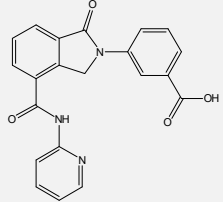
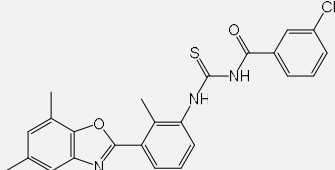
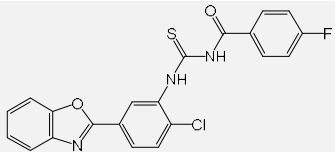
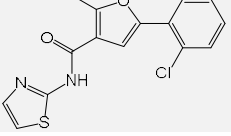
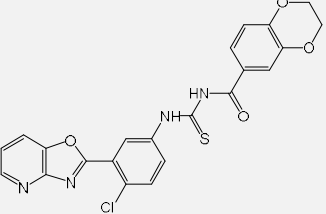
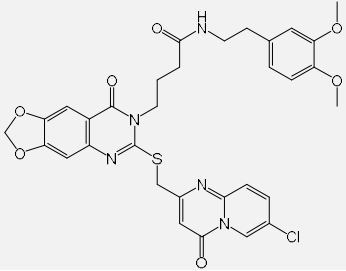
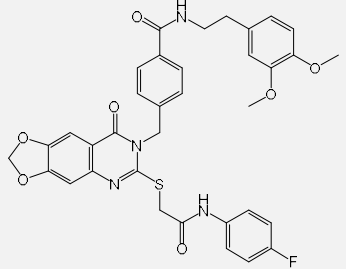
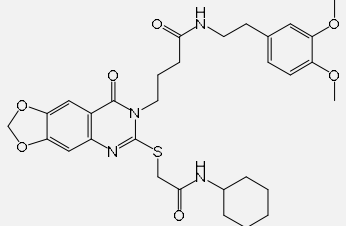
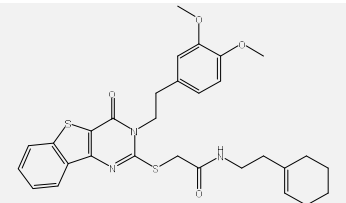


Figure 22. PptT-BpsA coupled assay system. **A** Inhibitory test using 6-NOBP (0–600 μM). **B** DMSO test (20–32%) for stabilizing the assay.

Table 4 Chemical structures and inhibitory potentials of PptT inhibitors.

Group	Cpd	Structure	IC ₅₀ (μM)
A	1-2		2.288
	1-3		0.635
	1-6		0.637
	1-7		2.332
	1-9		1.975
	2-3		3.378
	2-8		2.817

B	1-4		0.540
	2-1		4.741
	2-2		2.486
	2-7		2.523

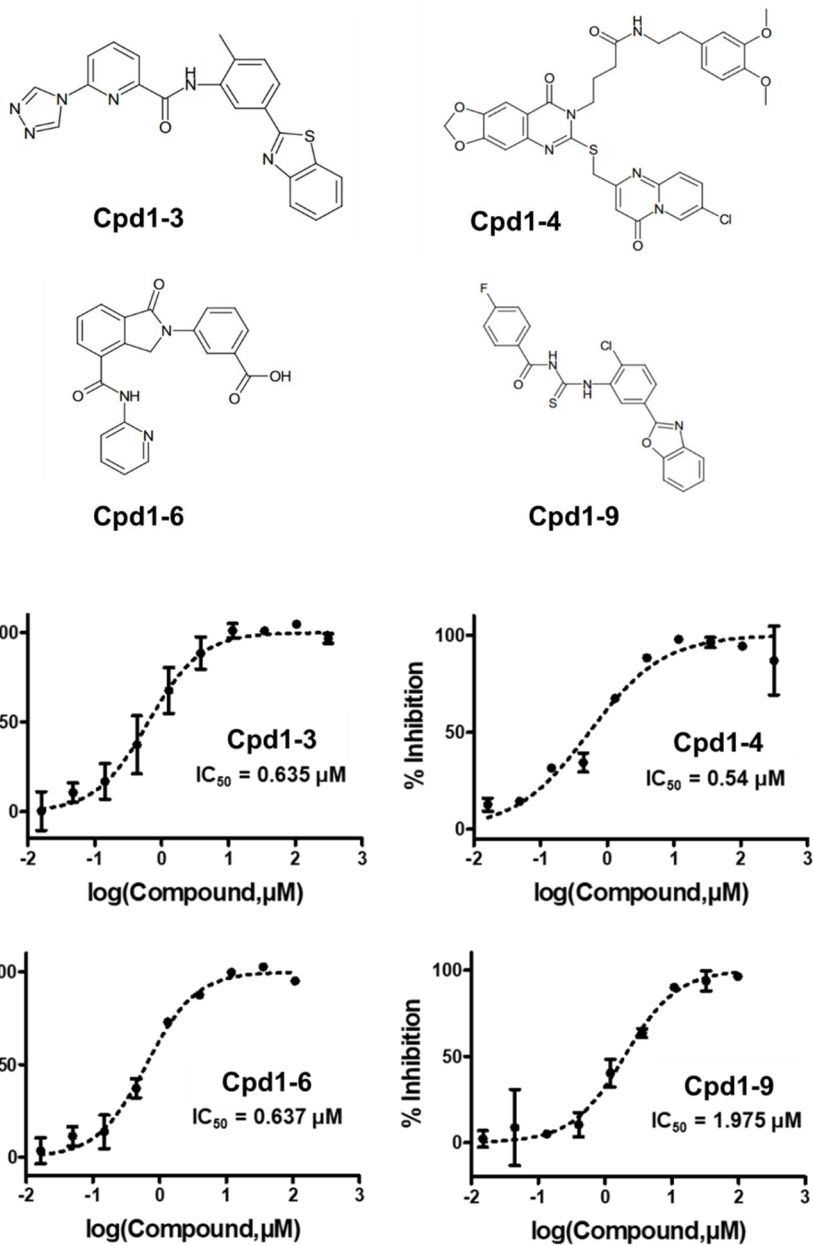


Figure 23. Structures and inhibitory potentials of four selected inhibitors. All chemical structures are drawn using chemdraw and each inhibition curve is presented below.

2.3.4 Identification of the selective scaffold inhibiting PptT from *M. tuberculosis*.

Despite no significant difference between the active sites of Mtb_PptT and MSMEG_PPTase, we hypothesized that the selected inhibitors may show the different tendency against the proteins from the pathogenic and non-pathogenic species. To validate this hypothesis, we performed PptT-BpsA coupled assay for the four selected inhibitors against PptT in the presence of MSMEG_PPTase. As expected, we confirmed that two inhibitors (**Cpd1-3** and **Cpd1-9**) did not show the inhibitory potential in contrast to the others (**Cpd1-4** and **Cpd1-6**) showing the inhibitory activity comparable to that of PptT (Figure 23). Remarkably, **Cpd1-3** and **Cpd1-9** possessed the common scaffold (group **A**). The result indicates that the scaffold of group **A** possesses Mtb-specific and selective inhibitory activity.

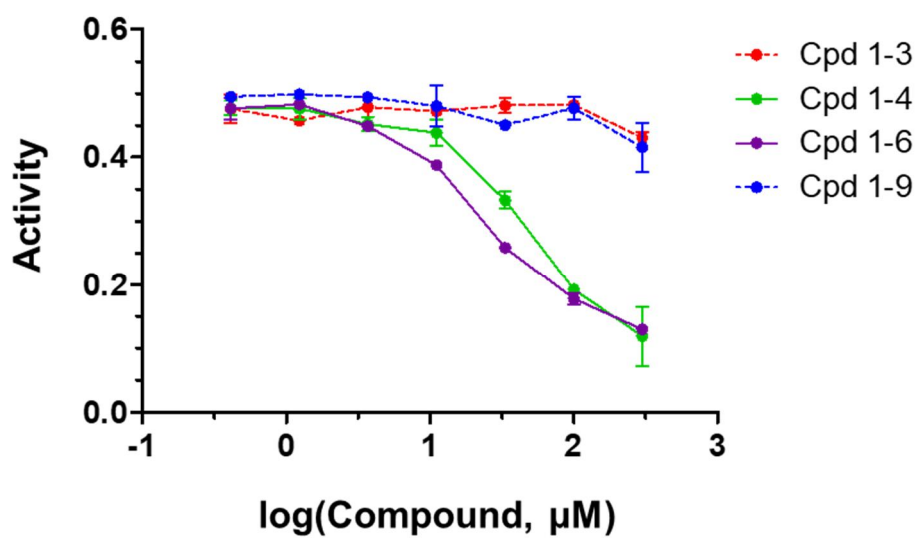


Figure 24. MSMEG_PPTase-BpsA coupled assay for 4 selected compounds. The data were determined by three independent experiments.

2.3.5 PptT inhibitor reveals the bactericidal activity in mycobacteria-infected macrophages

PptT from *M. tuberculosis* has been reported for playing a crucial role in the viability and persistence of tuberculosis under the latency stage (ref.). Therefore, we postulated that if our inhibitors reduce the activity of PptT, they can show the bactericidal activity toward mycobacteria-infected macrophages. To identify this biological effect of the inhibitors, we treated the inhibitors (0–50 μ M) into murine macrophage cell line infected by tuberculosis. Among the four inhibitors, we confirmed that **Cpd1-9** shows the significant bactericidal activity (Figure 24A). **Cpd1-9** exhibited the concentration-dependent suppression of the cell viability at both 1 and 3 days. Considering the cytotoxicity of

C

p

d

1

-

9

i

n

v

e

r

o

c

e

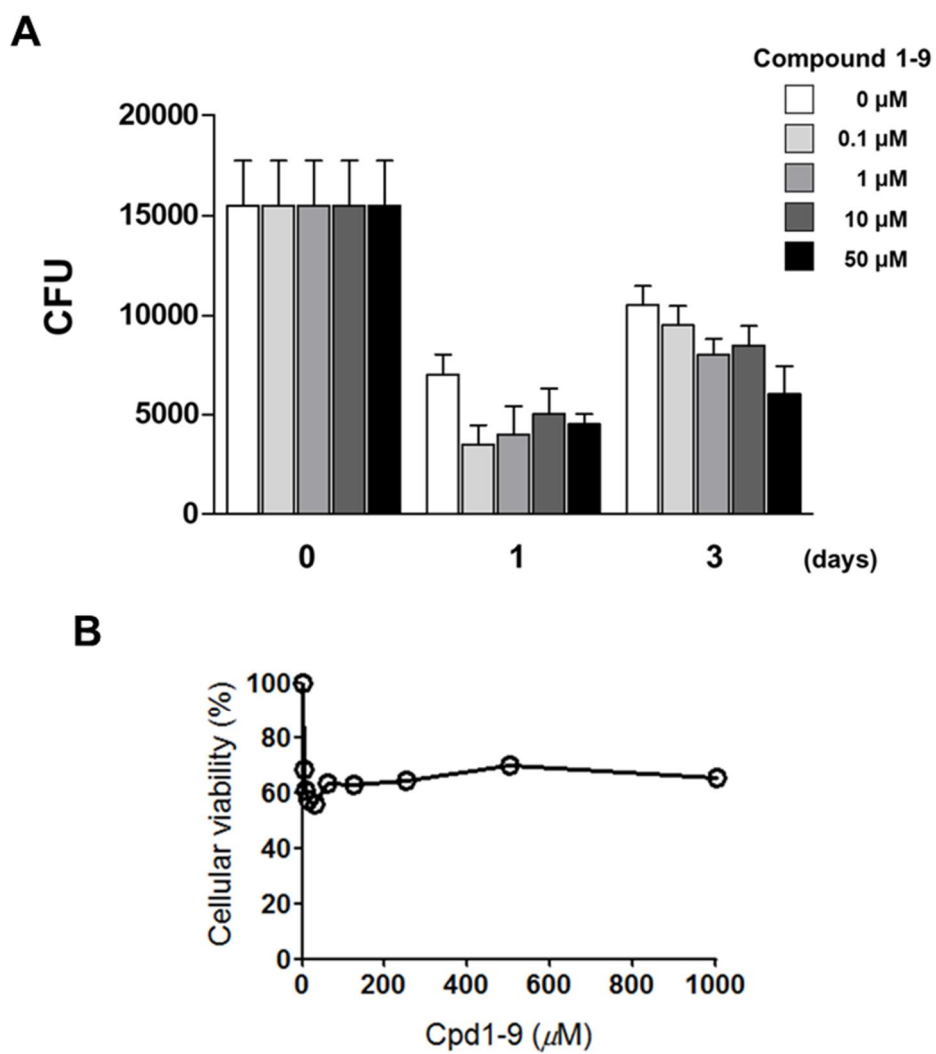


Figure 25. Biological test of **Cpd1-9**. **A** Bactericidal activity of **Cpd1-9** in the mycobacteria-infected macrophages. **B** Cytotoxicity assay of **Cpd1-9** in Vero cell line.

2.4 Discussion

A central concept of structure-based drug discovery (SBDD) is to develop the small molecule compounds fitting well into active site of the target enzyme, based on the structure of the specific biological targets [72]. In the present study, we determined crystal structures of PptT and MSMEG_PPTase which guided the inhibitor development of PptT which is a novel and long-sought drug target. No structural differences between PptT and MSMEG_PPTases were observed. However, in substrate, there are notable differences between them. Unlike PptT, MSMEG_PPTases revealed the unusual substrate binding mode with unusual Ca^{2+} ions. Most structures of Sfp-type PPTases have been reported to require Mg^{2+} for the catalysis of CoA [64]. Despite the general concept and Sfp-type PPTase, in MSMEG_PPTase I, we observed two Ca^{2+} ions instead of Mg^{2+} , similar to AcpS, AcpS-type from *Bacillus subtilis*. This observation may be resulted from the crystallization condition containing calcium chloride. Indeed, regardless of the presence of the Mg^{2+} ions in the structure, MSMEG_PPTase need Mg^{2+} for the catalysis, not Ca^{2+} corresponding to the previous works [73]. The structure of MSMEG_PPTase II showed the unusual 3',5'-ADP, a catalytic product of CoA with two Mg^{2+} , which have not been reported before. Since we did not add any 'carrier protein' which can obtain the phosphopantethein moiety of CoA, we could not elucidate the reason of 3',5'-ADP binding mode. However, the unique binding mode of 3',5'-ADP visualized the end of catalytic mechanism of PPTase and the presence of Mg^{2+} ions supported that MSMEG_PPTase employs Mg^{2+} ions for the catalysis. These structural features of the active sites in mycobacteria PPTase guided the development of PptT inhibitors through the virtual screening. Our renewal PptT-BpsA assay facilitated to validate the

inhibitory efficacy of a large number of compounds and the selected compounds rationally classified into two groups. Among them, the four most potent compounds are confirmed and further validated in the presence of MSMEG_PPTases. This assay allowed to select two compounds possessing the common scaffold that was only efficient in PptT, not in MSMEG_PPTase, indicating the Mtb-selectivity of the inhibitor. Furthermore, corresponding to the previous work that PptT is required for the intracellular survival of mycobacteria in macrophages [67], **Cpd1-9** has a significant effect on the viability of mycobacteria in macrophages, suggesting that this compound may be a potent and selective inhibitor possessing the bactericidal activity on the latent tuberculosis.

In conclusion, we uncovered the crystal structures of macrobacteria PPTases with the description of the different substrates and identified a novel inhibitor based on the elucidated structures. In addition, this study identified that the selective inhibition of PptT by **Cpd1-3** and **1-9** that did not show the inhibitory activity on MSMEG_PPTase with the bactericidal activity on the TB within the macrophage. Our findings can provide the valuable starting point into the therapeutic drugs targeting PptT, ultimately tuberculosis.

References

1. Hanukoglu I (1992) Steroidogenic enzymes: structure, function, and role in regulation of steroid hormone biosynthesis. *J Steroid Biochem Mol Biol* 43 (8):779-804. doi:10.1016/0960-0760(92)90307-5
2. Acimovic J, Rozman D (2013) Steroidal triterpenes of cholesterol synthesis. *Molecules* 18 (4):4002-4017. doi:10.3390/molecules18044002
3. Malfertheiner P, Megraud F, O'Morain CA, Atherton J, Axon AT, Bazzoli F, Gensini GF, Gisbert JP, Graham DY, Rokkas T, El-Omar EM, Kuipers EJ (2012) Management of *Helicobacter pylori* infection--the Maastricht IV/ Florence Consensus Report. *Gut* 61 (5):646-664. doi:10.1136/gutjnl-2012-302084
4. Wu W, Yang Y, Sun G (2012) Recent Insights into Antibiotic Resistance in *Helicobacter pylori* Eradication. *Gastroenterology research and practice* 2012:723183. doi:10.1155/2012/723183
5. Cameron EA, Powell KU, Baldwin L, Jones P, Bell GD, Williams SG (2004) *Helicobacter pylori*: antibiotic resistance and eradication rates in Suffolk, UK, 1991-2001. *J Med Microbiol* 53 (Pt 6):535-538
6. Carcanague D, Shue YK, Wuonola MA, Uria-Nickelsen M, Joubran C, Abedi JK, Jones J, Kuhler TC (2002) Novel structures derived from 2-[[[(2-pyridyl)methyl]thio]-1H-benzimidazole as anti-*Helicobacter pylori* agents, Part 2. *Journal of medicinal chemistry* 45 (19):4300-4309
7. Ramkumar S, Raghunath A, Raghunath S (2016) Statin Therapy: Review of Safety and Potential Side Effects. *Acta Cardiol Sin* 32 (6):631-639
8. Raetz CR, Whitfield C (2002) Lipopolysaccharide endotoxins. *Annual review of biochemistry* 71:635-700. doi:10.1146/annurev.biochem.71.110601.135414
9. Wyckoff TJ, Raetz CR, Jackman JE (1998) Antibacterial and anti-inflammatory agents that target endotoxin. *Trends in microbiology* 6 (4):154-159

10. Raetz CR, Reynolds CM, Trent MS, Bishop RE (2007) Lipid A modification systems in gram-negative bacteria. *Annual review of biochemistry* 76:295-329. doi:10.1146/annurev.biochem.76.010307.145803
11. Caldas H, Herman GE (2003) NSDHL, an enzyme involved in cholesterol biosynthesis, traffics through the Golgi and accumulates on ER membranes and on the surface of lipid droplets. *Hum Mol Genet* 12 (22):2981-2991. doi:10.1093/hmg/ddg321
12. Moran AP (2010) The role of endotoxin in infection: *Helicobacter pylori* and *Campylobacter jejuni*. *Subcell Biochem* 53:209-240. doi:10.1007/978-90-481-9078-2_10
13. Semeraro N, Montemurro P, Piccoli C, Muolo V, Colucci M, Giuliani G, Fumarola D, Pece S, Moran AP (1996) Effect of *Helicobacter pylori* lipopolysaccharide (LPS) and LPS derivatives on the production of tissue factor and plasminogen activator inhibitor type 2 by human blood mononuclear cells. *J Infect Dis* 174 (6):1255-1260
14. Sukhanova A, Gorin A, Serebriiskii IG, Gabitova L, Zheng H, Restifo D, Egleston BL, Cunningham D, Bagnyukova T, Liu H, Nikonova A, Adams GP, Zhou Y, Yang DH, Mehra R, Burtness B, Cai KQ, Klein-Szanto A, Kratz LE, Kelley RI, Weiner LM, Herman GE, Golemis EA, Astsaturov I (2013) Targeting C4-demethylating genes in the cholesterol pathway sensitizes cancer cells to EGF receptor inhibitors via increased EGF receptor degradation. *Cancer Discov* 3 (1):96-111. doi:10.1158/2159-8290.CD-12-0031
15. Sarkar M, Maganti L, Ghoshal N, Dutta C (2012) In silico quest for putative drug targets in *Helicobacter pylori* HPAG1: molecular modeling of candidate enzymes from lipopolysaccharide biosynthesis pathway. *Journal of molecular modeling* 18 (5):1855-1866. doi:10.1007/s00894-011-1204-3
16. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Method Enzymol* 276:307-326. doi:Doi 10.1016/S0076-6879(97)76066-X

17. Yin B, Cui DB, Zhang LJ, Jiang SQ, Machida S, Yuan YA, Wei DZ (2014) Structural insights into substrate and coenzyme preference by SDR family protein Gox2253 from *Gluconobater oxydans*. *Proteins* 82 (11):2925-2935. doi:10.1002/prot.24603
18. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66 (Pt 2):213-221. doi:10.1107/S0907444909052925
19. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53 (Pt 3):240-255. doi:10.1107/S0907444996012255
20. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66 (Pt 4):486-501. doi:10.1107/S0907444910007493
21. Brunger AT (1992) Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* 355 (6359):472-475
22. Chen VB, Arendall WB, 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66 (Pt 1):12-21. doi:10.1107/S0907444909042073
23. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372 (3):774-797. doi:10.1016/j.jmb.2007.05.022
24. Wallace AC, Laskowski RA, Thornton JM (1995) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng* 8 (2):127-134
25. Jorgensen WL, Maxwell DS, Tirado-Rives J (1996) Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of

Organic Liquids. *Journal of the American Chemical Society* 118 (45):11225-11236.
doi:10.1021/ja9621760

26. Lakowicz JR, Szymanski H, Nowaczyk K, Johnson ML (1992) Fluorescence lifetime imaging of free and protein-bound NADH. *Proc Natl Acad Sci U S A* 89 (4):1271-1275
27. Schrodinger, LLC (2015) The PyMOL Molecular Graphics System, Version 1.8.
28. Viegas A, Manso J, Nobrega FL, Cabrita EJ (2011) Saturation-Transfer Difference (STD) NMR: A Simple and Fast Method for Ligand Screening and Characterization of Protein Binding. *J Chem Educ* 88 (7):990-994. doi:Doi 10.1021/Ed101169t
29. Viegas A, Manso J, Corvo MC, Marques MMB, Cabrita EJ (2011) Binding of Ibuprofen, Ketorolac, and Diclofenac to COX-1 and COX-2 Studied by Saturation Transfer Difference NMR. *J Med Chem* 54 (24):8555-8562. doi:Doi 10.1021/Jm201090k
30. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E, Cerami E, Sander C, Schultz N (2013) Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* 6 (269):p11. doi:10.1126/scisignal.2004088
31. Györfy B, Surowiak P, Budczies J, Lanczky A (2013) Online survival analysis software to assess the prognostic value of biomarkers using transcriptomic data in non-small-cell lung cancer. *PLoS One* 8 (12):e82241. doi:10.1371/journal.pone.0082241
32. Okayama H, Kohno T, Ishii Y, Shimada Y, Shiraishi K, Iwakawa R, Furuta K, Tsuta K, Shibata T, Yamamoto S, Watanabe S, Sakamoto H, Kumamoto K, Takenoshita S, Gotoh N, Mizuno H, Sarai A, Kawano S, Yamaguchi R, Miyano S, Yokota J (2012) Identification of genes upregulated in ALK-positive and EGFR/KRAS/ALK-negative lung adenocarcinomas. *Cancer Res* 72 (1):100-111. doi:10.1158/0008-5472.CAN-11-1403

33. Rousseaux S, Debernardi A, Jacquiau B, Vitte AL, Vesin A, Nagy-Mignotte H, Moro-Sibilot D, Brichon PY, Lantuejoul S, Hainaut P, Laffaire J, de Reynies A, Beer DG, Timsit JF, Brambilla C, Brambilla E, Khochbin S (2013) Ectopic activation of germline and placental genes identifies aggressive metastasis-prone lung cancers. *Sci Transl Med* 5 (186):186ra166. doi:10.1126/scitranslmed.3005723
34. Director's Challenge Consortium for the Molecular Classification of Lung A, Shedden K, Taylor JM, Enkemann SA, Tsao MS, Yeatman TJ, Gerald WL, Eschrich S, Jurisica I, Giordano TJ, Misek DE, Chang AC, Zhu CQ, Strumpf D, Hanash S, Shepherd FA, Ding K, Seymour L, Naoki K, Pennell N, Weir B, Verhaak R, Ladd-Acosta C, Golub T, Gruidl M, Sharma A, Szoke J, Zakowski M, Rusch V, Kris M, Viale A, Motoi N, Travis W, Conley B, Seshan VE, Meyerson M, Kuick R, Dobbin KK, Lively T, Jacobson JW, Beer DG (2008) Gene expression-based survival prediction in lung adenocarcinoma: a multi-site, blinded validation study. *Nat Med* 14 (8):822-827. doi:10.1038/nm.1790
35. Holm L, Rosenstrom P (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res* 38 (Web Server issue):W545-549. doi:10.1093/nar/gkq366
36. Karplus PA, Diederichs K (2012) Linking crystallographic model and data quality. *Science* 336 (6084):1030-1033. doi:10.1126/science.1218231
37. Delvaux NA, Thoden JB, Holden HM (2018) Molecular architectures of Pen and Pal: Key enzymes required for CMP-pseudaminic acid biosynthesis in *Bacillus thuringiensis*. *Protein Sci* 27 (3):738-749. doi:10.1002/pro.3368
38. Miyafusa T, Caaveiro JM, Tanaka Y, Tsumoto K (2013) Dynamic elements govern the catalytic activity of CapE, a capsular polysaccharide-synthesizing enzyme from *Staphylococcus aureus*. *FEBS Lett* 587 (23):3824-3830. doi:10.1016/j.febslet.2013.10.009
39. Ishiyama N, Creuzenet C, Miller WL, Demendi M, Anderson EM, Harauz G, Lam JS, Berghuis AM (2006) Structural studies of FlaA1 from *Helicobacter pylori* reveal the mechanism for inverting 4,6-dehydratase activity. *The Journal of biological chemistry* 281 (34):24489-24495. doi:10.1074/jbc.M602393200

40. Pereira PJ, Macedo-Ribeiro S, Parraga A, Perez-Luque R, Cunningham O, Darcy K, Mantle TJ, Coll M (2001) Structure of human biliverdin IXbeta reductase, an early fetal bilirubin IXbeta producing enzyme. *Nature structural biology* 8 (3):215-220. doi:10.1038/84948
41. Perinbam K, Balaram H, Guru Row TN, Gopal B (2017) Probing the influence of non-covalent contact networks identified by charge density analysis on the oxidoreductase BacC. *Protein Eng Des Sel* 30 (3):265-272. doi:10.1093/protein/gzx006
42. Orengo CA, Jones DT, Thornton JM (1994) Protein superfamilies and domain superfolds. *Nature* 372 (6507):631-634. doi:10.1038/372631a0
43. Rao ST, Rossmann MG (1973) Comparison of super-secondary structures in proteins. *J Mol Biol* 76 (2):241-256
44. Kavanagh KL, Jornvall H, Persson B, Oppermann U (2008) Medium- and short-chain dehydrogenase/reductase gene and protein families : the SDR superfamily: functional and structural diversity within a family of metabolic and regulatory enzymes. *Cell Mol Life Sci* 65 (24):3895-3906. doi:10.1007/s00018-008-8588-y
45. Morimoto M, Souich C, Trinh J, McLarren KW, Boerkoel CF, Henderson G (2012) Expression profile of NSDHL in human peripheral tissues. *J Mol Histol* 43 (1):95-106. doi:10.1007/s10735-011-9375-x
46. Cook PF, Cleland WW (2007) *Enzyme kinetics and mechanism*. Garland Science, London
47. Bhatia C, Oerum S, Bray J, Kavanagh KL, Shafqat N, Yue W, Oppermann U (2015) Towards a systematic analysis of human short-chain dehydrogenases/reductases (SDR): Ligand identification and structure-activity relationships. *Chem Biol Interact* 234:114-125. doi:10.1016/j.cbi.2014.12.013
48. Okada Y, Kimura T, Nakagawa T, Okamoto K, Fukuya A, Goji T, Fujimoto S, Sogabe M, Miyamoto H, Muguruma N, Tsuji Y, Okahisa T, Takayama T (2017) EGFR Downregulation after Anti-EGFR Therapy Predicts the Antitumor Effect in

Colorectal Cancer. *Mol Cancer Res* 15 (10):1445-1454. doi:10.1158/1541-7786.MCR-16-0383

49. Tomas A, Futter CE, Eden ER (2014) EGF receptor trafficking: consequences for signaling and cancer. *Trends Cell Biol* 24 (1):26-34. doi:10.1016/j.tcb.2013.11.002
50. Yamamoto K, Kusunoki M, Urabe I, Tabata S, Osaki S (2000) Crystallization and preliminary X-ray analysis of glucose dehydrogenase from *Bacillus megaterium* IWG3. *Acta Crystallogr D Biol Crystallogr* 56 (Pt 11):1443-1445
51. McLarren KW, Severson TM, du Souich C, Stockton DW, Kratz LE, Cunningham D, Henderson G, Morin RD, Wu D, Paul JE, An J, Nelson TN, Chou A, DeBarber AE, Merkens LS, Michaud JL, Waters PJ, Yin J, McGillivray B, Demos M, Rouleau GA, Grzeschik KH, Smith R, Tarpey PS, Shears D, Schwartz CE, Gecz J, Stratton MR, Arbour L, Hurlburt J, Van Allen MI, Herman GE, Zhao Y, Moore R, Kelley RI, Jones SJ, Steiner RD, Raymond FL, Marra MA, Boerkoel CF (2010) Hypomorphic temperature-sensitive alleles of NSDHL cause CK syndrome. *Am J Hum Genet* 87 (6):905-914. doi:10.1016/j.ajhg.2010.11.004
52. Cooper BF, Rudolph FB (1995) Product inhibition applications. *Methods Enzymol* 249:188-211
53. Maeting I, Schmidt G, Sahm H, Stahmann KP (2000) Role of a peroxisomal NADP-specific isocitrate dehydrogenase in the metabolism of the riboflavin overproducer *Ashbya gossypii*. *Journal of Molecular Catalysis B: Enzymatic* 10 (1):335-343. doi:[https://doi.org/10.1016/S1381-1177\(00\)00135-1](https://doi.org/10.1016/S1381-1177(00)00135-1)
54. Gabitova L, Restifo D, Gorin A, Manocha K, Handorf E, Yang DH, Cai KQ, Klein-Szanto AJ, Cunningham D, Kratz LE, Herman GE, Golemis EA, Astsaturov I (2015) Endogenous Sterol Metabolites Regulate Growth of EGFR/KRAS-Dependent Tumors via LXR. *Cell Rep* 12 (11):1927-1938. doi:10.1016/j.celrep.2015.08.023
55. Helliwell SB, Karkare S, Bergdoll M, Rahier A, Leighton-Davis JR, Fioretto C, Aust T, Filipuzzi I, Frederiksen M, Gounarides J, Hoepfner D, Hofmann A, Imbert PE, Jeker R, Knochenmuss R, Krastel P, Margerit A, Memmert K, Miault CV,

- Movva NR, Muller A, Naegeli HU, Oberer L, Prindle V, Riedl R, Schuierer S, Sexton JA, Tao J, Wagner T, Yin H, Zhang J, Roggo S, Reinker S, Parker CN (2015) FR171456 is a specific inhibitor of mammalian NSDHL and yeast Erg26p. *Nat Commun* 6:8613. doi:10.1038/ncomms9613
56. World Health Organization. (2019) Global tuberculosis report 2019. World Health Organization, Geneva
 57. West K (1992) MDR (multidrug-resistant) tuberculosis. *Emergency medical services* 21 (9):54, 57-58
 58. van Ingen J, de Lange WC, Boeree MJ, Iseman MD, Daley CL, Heifets LB, Bottger EC, van Soolingen D (2011) XDR tuberculosis. *The Lancet Infectious diseases* 11 (8):585. doi:10.1016/S1473-3099(11)70200-6
 59. Barry CE, 3rd, Boshoff HI, Dartois V, Dick T, Ehrt S, Flynn J, Schnappinger D, Wilkinson RJ, Young D (2009) The spectrum of latent tuberculosis: rethinking the biology and intervention strategies. *Nature reviews Microbiology* 7 (12):845-855. doi:10.1038/nrmicro2236
 60. Lewis K (2010) Persister cells. *Annual review of microbiology* 64:357-372. doi:10.1146/annurev.micro.112408.134306
 61. Silhavy TJ, Kahne D, Walker S (2010) The bacterial cell envelope. *Cold Spring Harbor perspectives in biology* 2 (5):a000414. doi:10.1101/cshperspect.a000414
 62. Glickman MS, Cox JS, Jacobs WR, Jr. (2000) A novel mycolic acid cyclopropane synthetase is required for cording, persistence, and virulence of *Mycobacterium tuberculosis*. *Molecular cell* 5 (4):717-727
 63. Glickman MS, Jacobs WR, Jr. (2001) Microbial pathogenesis of *Mycobacterium tuberculosis*: dawn of a discipline. *Cell* 104 (4):477-485
 64. Chalut C, Botella L, de Sousa-D'Auria C, Houssin C, Guilhot C (2006) The nonredundant roles of two 4'-phosphopantetheinyl transferases in vital processes of

Mycobacteria. Proc Natl Acad Sci U S A 103 (22):8511-8516.
doi:10.1073/pnas.0511129103

65. Portevin D, De Sousa-D'Auria C, Houssin C, Grimaldi C, Chami M, Daffe M, Guilhot C (2004) A polyketide synthase catalyzes the last condensation step of mycolic acid biosynthesis in mycobacteria and related organisms. Proceedings of the National Academy of Sciences of the United States of America 101 (1):314-319. doi:10.1073/pnas.0305439101
66. Lambalot RH, Gehring AM, Flugel RS, Zuber P, LaCelle M, Marahiel MA, Reid R, Khosla C, Walsh CT (1996) A new enzyme superfamily - the phosphopantetheinyl transferases. Chemistry & biology 3 (11):923-936
67. Leblanc C, Prudhomme T, Tabouret G, Ray A, Burbaud S, Cabantous S, Mourey L, Guilhot C, Chalut C (2012) 4'-Phosphopantetheinyl transferase PptT, a new drug target required for Mycobacterium tuberculosis growth and persistence in vivo. PLoS Pathog 8 (12):e1003097. doi:10.1371/journal.ppat.1003097
68. Owen JG, Copp JN, Ackerley DF (2011) Rapid and flexible biochemical assays for evaluating 4'-phosphopantetheinyl transferase activity. The Biochemical journal 436 (3):709-717. doi:10.1042/BJ20110321
69. Collaborative Computational Project N (1994) The CCP4 suite: programs for protein crystallography. Acta crystallographica Section D, Biological crystallography 50 (Pt 5):760-763. doi:10.1107/S0907444994003112
70. Vickery CR, Kosa NM, Casavant EP, Duan S, Noel JP, Burkart MD (2014) Structure, biochemistry, and inhibition of essential 4'-phosphopantetheinyl transferases from two species of Mycobacteria. ACS Chem Biol 9 (9):1939-1944. doi:10.1021/cb500263p
71. van Aalten DM, Bywater R, Findlay JB, Hendlich M, Hooft RW, Vriend G (1996) PRODRG, a program for generating molecular topologies and unique molecular descriptors from coordinates of small molecules. J Comput Aided Mol Des 10 (3):255-262

72. Anderson AC (2003) The process of structure-based drug design. *Chem Biol* 10 (9):787-797. doi:10.1016/j.chembiol.2003.09.002
73. Beld J, Sonnenschein EC, Vickery CR, Noel JP, Burkart MD (2014) The phosphopantetheinyl transferases: catalysis of a post-translational modification crucial for life. *Nat Prod Rep* 31 (1):61-108. doi:10.1039/c3np70054b

국문초록

새로운 기전의 약물 타겟으로 알려진 NSDHL 및 PptT의 구조에 기반한 저해 제 개발 연구

김 동 균

서울대학교 대학원

약학과 물리약학 전공

지도교수 이 봉 진

Structure-based drug design (SBDD)는 약물을 빠르고 적은 비용으로 효율적으로 개발할 수 있는 가장 효과적인 기술 중 하나이다. 이 기술에 있어서 해당 단백질의 삼차원 구조의 규명은 구조에 최적화된 화합물의 개발을 가능하게 한다. 이 기술을 이용하여 유망한 효소를 표적으로하는 의약품의 개발을 위해, 본 연구진은 호모사피엔스 유래 NAD⁺-dependent steroid dehydrogenase-like (NSDHL)와 결핵균 유래 4'-phosphopantetheinyl transferase (PptT)의 두가지 촉매 효소에 주목하였다. 이러한 두 효소는 사람과 결핵균 주의 생물학적 대사에 있어 중요한 기능을 하는 것으로 인해 새로운 약물 개발 표적으로 주목받았

다. NSDHL은 인간 콜레스테롤 합성에 필수적인 효소이며 표피 성장 인자 수용체 (EGFR) trafficking pathway의 조절 인자이며, 콜레스테롤 관련 질환 및 암종에 대한 중대한 관련성으로 인해 새로운 표적 단백질로서 관심을 끌어들였다. PptT는 코엔자임A의 phosphopantethein 부분을 carrier protein 도메인의 세린 잔기에 공유결합으로 전달하며, 결핵의 세포 내 생존 및 persistence에 중요한 역할을 하여, 기존의 알려진 결핵 약물과는 다른 새로운 기작의 약물 타겟으로 알려져 있다. 본 연구에서는 기질의 결합 부위 및 결합형태에 대한 세부적인 묘사와 함께 NSDHL과 PptT 및 MSMEG_PPTase의 X-선 삼차원 결정 구조를 보고하였다. 이를 토대로, 구조 기반의 가상 스크리닝 및 생화학적 평가를 수행하여 NSDHL 및 PptT에 대한 새로운 억제제를 개발하였다. 또한, 저해능에 대한 추가 세포기반 검증은 우리의 억제제가 합리적으로 개발되었음을 밝혀냈다. 이러한 연구는 NSDHL 관련 질병 및 결핵에 대항하는 치료제 개발을 위한 좋은 플랫폼으로서 작용할 수 있다.

주요어: Structure-based drug discovery, *Homo sapiens*, *Mycobacterium tuberculosis*, cholesterol synthesis pathway, virtual screening, docking simulation, enzyme inhibitors, X-ray crystallography

학 번: 2014-21973