



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](#)

약학박사 학위논문

**Studies on the Mechanism and Inhibitor
Discovery Based on the Structures of Two
Catalytic Enzymes: *Hp*KDO8PS and
NSDHL**

**촉매 효소(*Hp*KDO8PS 및 NSDHL)의 구조에 기반한 작용
기전 및 억제제 개발 연구**

2016년 8월

서울대학교 대학원

약학과 물리약학전공

조 수 진

Abstract

Studies on the Mechanism and Inhibitor Discovery Based on the Structures of Two Catalytic Enzymes: *Hp*KDO8PS and NSDHL

Sujin Cho

College of Pharmacy

Seoul National University

Structure-based drug design (SBDD) is the technique to design compounds whose chemical structures are fitted into the three dimensional structure of a protein and to optimize those 'hits' into clinical candidates. The knowledge of the protein structure can help to accelerate drug development and make it more cost-effective. For development novel antibiotics and lipid-lowering agents, KDO8PS [KDO8P (3-deoxy-D-manno-octulosonate-8-phosphate) synthase] from *Helicobacter pylori* (*Hp*KDO8PS) and NSDHL [NAD(P) dependent steroid dehydrogenase-like protein (sterol-4- α -carboxylate 3-dehydrogenase, decarboxylating)] from *Homo sapiens* were selected and detailed studies have been conducted on the protein structures and biophysical properties.

*Hp*KDO8PS is the enzyme that catalyzes the condensation reaction between arabinose 5-phosphate (A5P) and phosphoenolpyruvate (PEP) to synthesize KDO8P, the

precursor of the 8-carbon sugar 3-deoxy-D-*manno*-octulosonate (KDO). The crystal structure of *Hp*KDO8PS was determined alone and within various complexes, revealing an extra helix (HE) that is absent in the structures of KDO8PS from other organisms. In contrast to the metal coordination of the KDO8PS enzyme from *Aquifex aeolicus*, *Hp*KDO8PS is specifically coordinated with Cd²⁺ or Zn²⁺ ions, and isothermal titration calorimetry (ITC) and differential scanning fluorimetry (DSF) revealed that Cd²⁺ thermally stabilizes the protein structure more efficiently than Zn²⁺. In the substrate-bound structure, water molecules play a key role in fixing residues in the proper configuration to achieve a compact structure. Using the structures of *Hp*KDO8PS and API [arabinose 5-phosphate (A5P) and phosphoenolpyruvate (PEP) bisubstrate inhibitor], 21 compounds showing potential *Hp*KDO8PS-binding properties were generated via *in silico* virtual screening. The binding capacity to *Hp*KDO8PS of three compounds (avicularin, hyperin, and MC181) was confirmed through saturation transfer difference (STD) experiments, and binding mode of each compound was identified by combining competition experiments and docking simulation analysis. Hyperin was confirmed to bind to the A5P binding site, primarily via hydrophilic interaction, whereas MC181 bound to both the PEP and A5P binding sites through hydrophilic and hydrophobic interactions. These results were consistent with the epitope mapping by STD. The results are expected to provide clues for the development of *Hp*KDO8PS inhibitors.

NSDHL is one of the enzymes in the cholesterol biosynthesis that catalyzes NAD⁺-dependent oxidative decarboxylation of the C4 methyl groups from 4 α -carboxysterol to produce the corresponding 3-keto, C4-decarboxylated products. To determine the NSDHL crystal structure, the protein was purified, and crystallized. However, structural determination of NSDHL was not successful, due to the weak

anomalous signals from the SeMet data and the difficulty in indexing the data. Alternatively, homology modeling has been used to generate the 3D structure and identify the catalytic key residues of NSDHL. They adopt a Rossmann fold with six α -helices and parallel five β -strands. By comparing the NSDHL models with other Rossmann folds, the active site and the coenzyme binding site could be suggested. The coenzyme binding region is notably well conserved. Mutants causing genetic disorders (CHILD, CKS) were produced by site-directed mutagenesis and they were purified. A correlation between the mutants and thermal stability was investigated using DSF. The mutants are significantly less stable than the wild-type protein supporting the hypothesis that mutations can affect NSDHL folding and thermal stability. Also, ITC was performed to investigate the thermodynamics of the protein-coenzymes. The K_d values from the experiments revealed that NSDHL prefers NAD(H) to NADP(H) for its enzyme reaction. In addition, the STD spectra for NAD and NADH gave information of their binding modes to NSDHL. The spectra indicated NSDHL-coenzymes interaction. To understand better the relationship of the coenzyme binding modes and the disease-causing mutants, STD experiments of NAD⁺ or NADH with G205S and K232 Δ NSDHL were also performed. The mutants showed lower affinities to NADH compared to the wild-type protein. Also, the result showed slightly different binding modes of the coenzymes. These findings support the possibility that changes in binding modes could be relevant in certain disease states found in CHILD and CKS.

Keywords: Structure-based drug discovery, *Helicobacter pylori*, *Homo sapiens*, virtual screening, docking simulation, enzyme inhibitors, X-ray crystallography, NMR spectroscopy

Student number: 2010-21728

Contents

Abstract	i
Contents.....	v
List of tables	viii
List of figures	ix
Chapter 1. Studies on the inhibitor discovery based on the crystal structure of <i>Helicobacter pylori</i> 3-deoxy-D-manno-octulosonate 8-phosphate synthase (HpKDO8PS).....	1
1.1 Introduction	1
1.2 Experimental procedures.....	5
1.2.1 Cloning, expression, and protein purification	5
1.2.2 Crystallization and structure determination	8
1.2.3 Circular dichroism spectroscopy	11
1.2.4 Isothermal titration calorimetry.....	12
1.2.5 Differential scanning fluorimetry	13
1.2.6 Preparation of compounds for virtual screening.....	14
1.2.7 Preparation of protein structures	14
1.2.8 Virtual screening	15
1.2.9 Synthesis and preparation of ligands	16
1.2.10 NMR spectroscopy	19

1.3 Results	21
1.3.1 Protein expression and structure determination	21
1.3.2 Crystal structures of <i>Hp</i>KDO8PS	26
1.3.3 Thermal scanning for metal and metal-free <i>Hp</i>KDO8PS interaction	34
1.3.4 Validation of the docking method	37
1.3.5 <i>Hp</i>KDO8PS-targeted virtual screening	39
1.3.6 NMR analysis of interactions between hits and <i>Hp</i>KDO8PS	43
1.3.7 Docking analyses of <i>Hp</i>KDO8PS ligands	51
1.4 Discussion	56
Chapter 2. Studies on the mechanism and the protein folding based on the structure of <i>Homo sapiens</i> NAD(P) dependent steroid dehydrogenase (NSDHL) and its disease-related mutants	63
2.1 Introduction	63
2.2 Experimental procedures	68
2.2.1 Cloning, expression, and protein purification	68
2.2.2 Crystallization and preliminary structure determination	71
2.2.3 Differential scanning fluorimetry	73
2.2.4 Isothermal titration calorimetry	74
2.2.5 NMR spectroscopy	75
2.3 Results	77
2.3.1 Protein expression and structure determination	77
2.3.2 Homology modeling of NSDHL	82

2.3.3 Thermal stability of wild-type NSDHL and the mutants.....	86
2.3.4 Isothermal titration calorimetry.....	88
2.3.5 STD-NMR analysis of interaction modes between NSDHL and its coenzymes...	91
2.4 Discussion.....	104
Conclusion.....	110
References.....	111
국문초록.....	125

List of tables

- Table 1.** Primers for site-directed mutagenesis targeting the residues involved in metal binding within the active site of *Hp*KDO8PS
- Table 2.** Statistics of data collection and refinement
- Table 3.** ADME properties and docking scores of selected hits for virtual screening
- Table 4.** Docking scores of KDO8PS variants
- Table 5.** Primers used to clone the NSDHL expression plasmid
- Table 6.** Primers for site-directed mutagenesis targeting the residues involved in causing CHILD and CKS
- Table 7.** Conditions for crystallization of NSDHL (31–267) protein
- Table 8.** Results of NSDHL homology modeling

List of figures

- Figure 1.** Solubility improvement and purification result of *HpKDO8PS* after co-expression with various chaperones
- Figure 2.** *HpKDO8PS* crystals
- Figure 3.** 3D structure of apo*HpKDO8PS*wt
- Figure 4.** Interactions within the active sites in four *HpKDO8PS* X-ray crystal structures
- Figure 5.** Metal and PEP binding sites in *HpKDO8PS*
- Figure 6.** Cd²⁺ titration onto metal-free *HpKDO8PS* (*HpKDO8PS*_metal-free)
- Figure 7.** Thermal shift assay of *HpKDO8PS*_metal-free by differential scanning fluorimetry measured in the presence and absence of metal (Cd²⁺ or Zn²⁺)
- Figure 8.** Chemical structures of known *KDO8PS* inhibitors
- Figure 9.** Scheme of virtual screening
- Figure 10.** ¹H STD and WaterLOGSY NMR spectra in the aromatic region and docking conformation of hyperin
- Figure 11.** ¹H STD and WaterLOGSY NMR spectra in the aromatic region and docking conformation of MC181
- Figure 12.** ¹H STD NMR spectra of avicularin in the aromatic region
- Figure 13.** Docking conformation of API

- Figure 14.** Top-down view of the active site of *Aa*KDO8PS, *Hp*KDO8PS and *Ec*KDO8PS
- Figure 15.** CD spectra of a set of various forms of *Hp*KDO8PS
- Figure 16.** Cholesterol synthesis pathway
- Figure 17.** Sterol structures and reactions involved in C4-decarboxylation by NSDHL
- Figure 18.** NSDHL constructs
- Figure 19.** Expression and purification result of NSDHL
- Figure 20.** NSDHL (31–267) crystals
- Figure 21.** Three-dimensional structure of NSDHL derived from homology modeling method using SWISS-MODEL and molecular replacement approach
- Figure 22.** Sequence alignment of NSDHL (31–267) and the template proteins of homology modeling
- Figure 23.** Homology models of NSDHL
- Figure 24.** Differential Scanning Fluorimetry of NSDHL
- Figure 25.** Calorimetric data for the titration of 1 mM of the coenzymes to NSDHL wild-type
- Figure 26.** Calorimetric data for the titration of 1 mM of the coenzymes to NSDHL mutants
- Figure 27.** ^1H STD spectra (3.7 ppm ~ 10 ppm) of NAD^+

- Figure 28.** ^1H STD spectra (3.7 ppm ~ 10 ppm) of NADH
- Figure 29.** ^1H STD competition NMR spectrum of NAD^+ and NADH
- Figure 30.** ^1H STD spectra (3.7 ppm ~ 10 ppm) of NADP^+
- Figure 31.** ^1H STD spectra (3.7 ppm ~ 10 ppm) of NADPH
- Figure 32.** ^1H STD competition NMR spectrum of NADP^+ and NADPH
- Figure 33.** Comparison of ^1H STD competition NMR spectra
- Figure 34.** Normalized STD effect of NAD^+ protons in the presence of wild-type, G205S, or K232 Δ NSDHL
- Figure 35.** Normalized STD effect of NADH protons in the presence of wild-type, G205S, or K232 Δ NSDHL

Chapter 1. Studies on the inhibitor discovery based on the crystal structure of *Helicobacter pylori* 3-deoxy-D-manno-octulosonate 8-phosphate synthase (*Hp*KDO8PS)

1.1 Introduction

Helicobacter pylori (*H. pylori*) is a gram-negative, microaerophilic bacterium that colonizes the stomach. Marshall and Warren first identified this microbe from patients suffering from chronic gastritis and peptic ulcerations [1]. Currently, more than half of the general population is infected with *H. pylori*, which is linked to gastritis, duodenal ulcer, gastric cancer, gastric mucosa-associated lymphoid tissue lymphoma (MALT), and sudden infant death syndrome (SIDS) [2-7].

Triple therapy involving a proton pump inhibitor (omeprazole) and antibiotics (amoxicillin and clarithromycin) was initially recommended for treating *H. pylori* infection [8]; however, this conventional multi-therapy is no longer effective due to the prevalence of antibiotic resistance [9, 10]. Moreover, these antibiotics commonly disrupt the normal gastrointestinal flora, causing diarrhea as a side effect [11]. Therefore, it is

necessary to develop new anti-*H. pylori* agents based on the structure of novel cellular targets.

3-Deoxy-D-manno-octulosonate 8-phosphate synthase (KDO8PS) is the enzyme that catalyzes the condensation reaction between arabinose 5-phosphate (A5P) and phosphoenolpyruvate (PEP) to synthesize KDO8P, the precursor of the 8-carbon sugar 3-deoxy-D-manno-octulosonate (KDO). KDO is the essential component of gram-negative bacterial lipopolysaccharides (LPS) [12], functioning as the linkage between lipid A and O-antigen [13]. As the enzymes related to KDO synthesis and its incorporation within the LPS structure play an important role in the survival and growth of gram-negative bacteria, inhibition of KDO synthesis results in a loss of the LPS endotoxin component and reduced pathogenicity [14-16]. In addition, although endotoxin released from most pathogenic gram-negative bacteria induces a strong response from normal host immune systems, *H. pylori* LPS has significantly lower endotoxic and immuno-stimulatory effects due to its under-phosphorylated and under-acylated lipid A component that binds to immune receptors [17]. This biological activities may contribute to the prolongation of *H. pylori* infection and the local mucosal inflammatory reaction associated with the chronic infection [18, 19]. Because the enzymes that participate in bacterial LPS synthesis reactions are absent in mammalian systems, they are considered to be potential targets for the development of antibiotics or anti-endotoxin agents with fewer potentially undesirable cross-reactions [20]. Members of the KDO8PS family have been grouped into two classes according to their transition metal requirements [21]. *Escherichia coli* (*E. coli*) KDO8PS is metal independent (class I), whereas the *Aquifex aeolicus* (*A. aeolicus*) and *H. pylori* enzymes (*Hp*KDO8PS) require a transition metal (class II) [21, 22]. KDO8PS crystal structures, both alone and in complex with various ligands, have been previously reported

(e.g., KDO8PSs from *E. coli*, *A. aeolicus*, *Neisseria meningitidis*, *Burkholderia pseudomallei* and *Pseudomonas aeruginosa*) [23-27]. The KDO8PSs from *E. coli* (*Ec*KDO8PS) and *A. aeolicus* (*Aa*KDO8PS) are homotetramers in which each monomer forms a $(\beta/\alpha)_8$ barrel fold structure [23, 24]. In addition, similarities in the enzymatic reaction and active site structure between KDO8PS and 3-deoxy D-arabinoheptulosonate-7-phosphate synthase (DAH7PS) suggest that they share a common ancestor [24]. The active site cavity in the PEP- and A5P-bound forms is closed, whereas it is open in the substrate-free form [23]. It is possible that its binding site allows PEP to form its distorted C-2 geometry by adopting sp^3 instead of sp^2 hybridization, thereby facilitating the intermediate form prior to the condensation with A5P [23]. Based on the crystal structures of binary complexes of *Ec*KDO8PS with PEP and through the use of a mechanism-based inhibitor ($K_d = 0.4 \mu\text{M}$), the condensation reaction between the substrates was shown to be mediated by an intermediate form containing a transient oxocarbenium ion formed at the C-2 position in PEP [28]. Additionally, the structures of *Aa*KDO8PS in complex with R5P and PEP, along with the bisubstrate inhibitor API [arabinose 5-phosphate (A5P) and phosphoenolpyruvate (PEP) bisubstrate inhibitor], suggest that the water molecule coordinated to the metal ion on the *si* side of PEP is necessary to trigger the reaction [29]. These structures also support the hypothesis that a proton of the PEP-phosphate group is transferred to the A5P-aldehyde group to first form the hydroxyl group and that condensation is completed by a *syn* addition of water and A5P to the *si* side of PEP [29].

*Aa*KDO8PS Cys11, His185, Glu222, and Asp233 are the residues that interact with the aforementioned transitional metal ions. Among them, His185 in particular directs PEP to its active site and places a water molecule on the *si* side of PEP [30]. The Fe^{2+} or

Zn²⁺ ions bound to the *Aa*KDO8PS active site are replaced by other divalent metal ions, such as Cd²⁺ and Cu²⁺ [30-32], and these metal ion substitutions affect enzymatic activity by altering the environment of the active site [32]. For example, Cd²⁺-bound *Aa*KDO8PS exhibits maximal enzymatic activity over the other metal-bound proteins [30].

Based on the catalytic mechanism and *Aa*KDO8PS structures, Gatti and colleagues proposed novel inhibitors that mimic the intermediate form of the condensation reaction [29, 33]. In this study, we determined the crystal structures of *Hp*KDO8PS alone (apo*Hp*KDO8PSwt), in complex with PEP and zinc (*Hp*KDO8PS-PEP-Zn), and in complex with cadmium (*Hp*KDO8PS-Cd) in an effort to develop anti-*H. pylori* agents. The structure of an *Hp*KDO8PS mutant (*Hp*KDO8PS_H204A) was also determined.

Chemical compounds binding to the active site of *Hp*KDO8PS were evaluated by virtual screening using the determined *Hp*KDO8PS structure. Among the 21 compounds initially selected, three hits were selected and validated for binding to *Hp*KDO8PS via STD NMR spectroscopy and waterLOGSY experiments. We also performed docking simulations for *Hp*KDO8PS in complex with the three compounds. These compounds could serve as novel scaffolds for the development of antibiotics that inhibit the function of *Hp*KDO8PS.

1.2 Experimental procedures

1.2.1 Cloning, expression, and protein purification

The gene encoding HP0003 from *H. pylori* 26695 was amplified by PCR using *H. pylori* genomic DNA (strain ATCC 700392/26695) as a template. The PCR primers used to clone the *HpKDO8PS* expression plasmid, in which the restriction enzyme sites are underlined, were as follows: 5'-GGGAATTCCCATATGAAAACCTTCTAAAACAAAAACCCC-3' and 5'-CCGCTCGAGTTAAAATAAATTTGGATTTTAAACATGTCGG-3'. The PCR product and pCold I vector (Takara, Japan) were digested with *NdeI* and *XhoI* (NEB, UK) and ligated together. After confirming the sequence, the recombinant plasmid was overexpressed in chaperone-expressing *E. coli* cells pTF16/BL21 (Takara, Japan) grown in LB broth. When the culture media reached an OD₆₀₀ of 0.8, 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) was added to induce protein expression; cells were transferred to 15°C and grown for an additional 20 hours. The cells were collected by centrifugation at 4293 × *g* for 10 min, and the pellet was resuspended in lysis buffer [20 mM Tris, 500 mM NaCl, 5 mM imidazole, 0.5 mM tris(2-carboxyethyl)phosphine (TCEP), pH 7.8] and then sonicated at 4°C using a 30% duty cycle setting for 20 min (Cole Parmer Inc., USA). The lysate was centrifuged at 6708 × *g* for 60 min at 4°C, and the supernatant was loaded into an open Ni-NTA column (Qiagen, USA) pre-equilibrated with lysis buffer. The column was washed with a 30-fold excess volume of buffer

containing 150 mM imidazole; the protein was eluted when the buffer reached 400 mM imidazole using an imidazole gradient (from 150 to 500 mM). After concentrating the protein, the buffer was changed to 20 mM MES, 200 mM NaCl, and 0.5 mM TCEP at pH 6 by dialysis using an Amicon Ultra-15 Centrifugal Filter Unit with a 10k molecular weight cutoff (MWCO) (Millipore, USA). SDS-PAGE revealed purity up to 95%. The enzyme was concentrated to 0.12 mM for crystallization.

The *HpKDO8PS*_metal-free form used for circular dichroism (CD) and ITC experiments was prepared by dialyzing purified *HpKDO8PS*wt against a buffer containing 20 mM MES, 200 mM NaCl, 2 mM 1, 10-phenanthroline, and 2 mM EDTA at pH 6. Subsequently, dialysis with a buffer containing 20 mM MES, 200 mM NaCl, and 0.5 mM TCEP at pH 6 was conducted to reach a 1,000-fold dilution of the metal ion chelators. The resulting *HpKDO8PS*_metal-free form was concentrated to 0.03 mM for CD spectroscopy and 0.1 mM for ITC.

Mutagenesis primers (Table 1) targeting the four metal-binding residues were designed, and EZchange™ Site-Directed Mutagenesis Kit (Enzynomics, Korea) was used to generate point mutations according to the manufacturer's instructions. Briefly, the PCR conditions consisted of 25 repeated cycles including a 30-s melting step at 94°C, a 1-min annealing step at 55°C and a 5-min elongation step at 72°C. Template removal and ligation of the PCR products were performed using EZ-MIX buffer (Enzynomics, Korea). After transformation of the mutant plasmids into DH5 α -competent cells, the sequence of the plasmid DNA sequence was confirmed. BL21/pTF16 was used for overexpression of the mutant proteins, and the successfully expressed mutants (*HpKDO8PS*_C18A and *HpKDO8PS*_H204A) were purified.

Table 1. Primers for site-directed mutagenesis targeting the residues involved in metal binding within the active site of *HpKDO8PS*. In primer IDs, “F” represents forward, and “R” represents reverse. Mutated nucleotides are underlined in the primer sequences used for site-directed mutagenesis.

Primer ID	Primer Sequence	Mutated nucleotide
C18p-F	GTTTAAATCGCTGGGCCAG <u>CT</u> GTTCATTGAGAGCTTAG	C18A
C18p-R	CTAAGCTCTCAATGACAG <u>CT</u> TGGCCCAGCGATTAAAAC	C18A
H204p-F	GATTTTGGACGCTACCG <u>CT</u> AGCGTGCAAATGCCAG	H204A
H204p-R	CTGGCATTGTCACGCTAG <u>CG</u> GTAGCGTCAAAAATC	H204A
E241p-F	ATTGATGGGTTGTTTGCTG <u>CT</u> ACGCATGTTGATCCTAAA	E241A
E241p-R	TTAGGATCAACATGCGT <u>AG</u> CAGCAAACAACCCATCAAT	E241A
D252p-F	CTAAAAACGCCCTAAGCG <u>CT</u> TGGAGCAAACATGCTAAAAC	D252A
D252p-R	GTTTTAGCATGTTTGCTC <u>C</u> AGCGCTTAGGGCGTTTTTAG	D252A

1.2.2 Crystallization and structure determination

The most efficient crystallization conditions for each form of *Hp*KDO8PS were selected using the Crystal Screen™ and Index™ crystal screening kits (Hampton Research, USA) with the hanging drop vapor diffusion method. The protein solution and reservoir solution containing 18% polyethylene glycol (PEG) 3,350, 0.1 M HEPES, pH 7.5, and 0.25 M magnesium chloride hexahydrate were mixed together in a 1:1 drop ratio. Crystals were obtained through drop equilibration with the reservoir solution at 20°C for 14 days. Prior to data collection, the crystals were flash-cooled with liquid nitrogen with protection by the addition of 25% glycerol to the cryo-solution. An identical procedure was applied for *Hp*KDO8PS_H204A crystal preparation.

To obtain better diffraction-quality crystals of *Hp*KDO8PS-Cd and *Hp*KDO8PS-PEP-Zn, we performed streak seeding using pulverized apo*Hp*KDO8PSwt crystals as microseeds for nucleation. The protein solution contained the protein-binding partners (Cd²⁺, Zn²⁺ and PEP) at a 2-fold excess concentration over the protein molecules (*i.e.*, protein at 0.12 mM and protein binding partners at 0.24 mM). For the crystallization drop, 1 μL of prepared protein solution was added to the same volume of reservoir solution. The seed stock was prepared after washing an apo*Hp*KDO8PSwt crystal; the stock was continuously transferred to the original crystallization solution for stabilization and crushed using a glass homogenizer. This concentrated seed stock was serially diluted with the stabilizing buffer by a factor of 1,000. Each drop was streaked with the diluted solution using a clean whisker-like fiber. Crystals appeared in 7 to 14 days at 20°C after the start of the incubation and were additionally soaked in a cryo-solution containing 1

mM of Zn^{2+} , PEP or Cd^{2+} for 12 hours prior to the flash-freezing step.

The diffraction dataset for the apo*Hp*KDO8PSwt crystal was collected at 100 K at a resolution of 2.0 Å using an MAR225HE CCD detector at beamline BL44XU in the SPring-8 radiation facility (Hyogo, Japan). The dataset was processed and scaled with the HKL-2000 program package [34]. The apo*Hp*KDO8PSwt crystal was identified from the C2 space group with the following unit cell parameters: $a = 137.83 \text{ \AA}$, $b = 50.66 \text{ \AA}$, $c = 78.67 \text{ \AA}$, and $\beta = 110.5^\circ$. There were two monomers in each asymmetric unit. The calculated crystal volume per protein weight (V_M) was $2.12 \text{ \AA}^3 \text{ Da}^{-1}$, and the solvent content was 42.0% [35]. The structure of apo*Hp*KDO8PSwt was determined by molecular replacement using the PHENIX Phaser-MR program [36] based on the *Aa*KDO8PS structure as a search model (PDB: 1FX6) [23]. The initial structures were refined by alternately using the Refmac [37] and Phenix.refine [36] programs. The solvent molecules were inserted with Coot [38].

The X-ray diffraction data from a single *Hp*KDO8PS_H204A crystal were collected at PAL (Pohang, Korea) using an ADSC quantum 315r CCD detector on beamline 5C-SBII at 100 K and a resolution of 2.4 Å. The raw data were processed using the HKL-2000 program [34]. The *Hp*KDO8PS_H204A crystal belongs to the space group C2, with two monomers in each asymmetric unit (unit cell parameters: $a = 139.86 \text{ \AA}$, $b = 50.87 \text{ \AA}$, $c = 78.30 \text{ \AA}$, and $\beta = 104.73^\circ$). Its V_M was $2.22 \text{ \AA}^3 \text{ Da}^{-1}$, and the solvent content was 44.6% [35]. The structure of the *Hp*KDO8PS_H204A crystal was determined by molecular replacement using the PHENIX Phaser-MR program [36] based on the determined model of apo*Hp*KDO8PSwt. The refinement step for *Hp*KDO8PS_H204A was conducted in a manner identical to that for apo*Hp*KDO8PSwt.

Raw X-ray diffraction data for the *Hp*KDO8PS-Cd crystals were collected at a

resolution of 1.93 Å at 100 K using the same detector as for the *HpKDO8PS_H204A* crystal at PAL (Pohang, Korea). The data were processed and scaled using the HKL-2000 program [34]. The *HpKDO8PS-Cd* crystal was identified from the *C2* space group, with unit cell parameters of $a = 140.35$ Å, $b = 51.02$ Å, $c = 78.67$ Å, and $\beta = 104.7^\circ$. Two monomers of *HpKDO8PS-Cd* were found in each asymmetric unit, with a $V_M = 2.24$ Å³ Da⁻¹ and a solvent content of 45.2% [35]. The structure was determined by molecular replacement using apo*HpKDO8PSwt* as a model with the PHENIX Phaser-MR program [36]. Refinement of the *HpKDO8PS-Cd* model was performed by alternately using the Refmac [37] and Phenix.refine [36] programs, as mentioned above.

The *HpKDO8PS-PEP-Zn* diffraction dataset was collected at a resolution of 1.68 Å at 100 K using a Saturn A200 mosaic CCD detector at beamline 26B1 in SPring-8 (Hyogo, Japan). The dataset was processed and scaled with the HKL-2000 program [34]. The *HpKDO8PS-PEP-Zn* crystal was identified from the *C2* space group, with unit cell parameters of $a = 140.3$ Å, $b = 51.01$ Å, $c = 78.55$ Å, and $\beta = 104.37^\circ$. There were two monomers in each asymmetric unit. The V_M was 2.24 Å³ Da⁻¹, and the solvent content was 45.2% [35]. The structure of *HpKDO8PS-PEP-Zn* was determined by molecular replacement using the PHENIX Phaser-MR program [36] with the apo*HpKDO8PSwt* structure as a search model. The initial structures were refined by alternately using the Refmac [37] and Phenix.refine [36] programs.

1.2.3 Circular dichroism spectroscopy

For CD experiments, the *HpKDO8PS* forms (*HpKDO8PS_metal-bound*, *HpKDO8PS_metal-free*, *HpKDO8PS_C18A*, and *HpKDO8PS_H204A*) were used at a concentration of 0.03 mM in a buffer containing 20 mM MES, 200 mM NaCl, and 0.5 mM TCEP, pH 6. *HpKDO8PS_metal-bound* represents the sample used in the crystallization of apo*HpKDO8PS*_{wt}. CD spectra were collected using a ChirascanTM-plus CD spectrometer (Applied Photophysics, UK) with the following settings: scan rate of 120 nm min⁻¹, response time of 0.5 s, step size of 1 nm, and bandwidth of 1 nm. The path length was 0.2 mm using a quartz cell with detachable windows (Hellma, Germany). All scans were conducted in the far-UV range (190 – 260 nm). The data were processed by baseline subtraction and smoothing using the Pro-Data Viewer program (Applied Photophysics, UK).

1.2.4 Isothermal titration calorimetry

A 0.1 mM solution of *Hp*KDO8PS₂-metal-free was prepared as described above. Zinc chloride and cadmium chloride were dissolved to a concentration of 1.5 mM in the same buffer as the protein. The protein and metal ion solutions were degassed in a vacuum before measurement. All ITC experiments were performed using a MicroCal iTC200 microcalorimeter (GE healthcare, UK) at 25°C. The protein was added to the sample cell, and a metal solution (zinc or cadmium chloride) was charged in the injection syringe. During the titration, 20 aliquots of metal solution were injected into the sample chamber. The titrations began with a 60 s delay time and 0.4 μ L injection volume, followed by a 2 μ L injection-volume with 5 s delay times. Intervals of 150 s between injections were included at the end of each titration. The stirring speed in the sample chamber at 25 \pm 0.1°C was 1,000 rpm. Heat generation during the titration experiments was measured using the integrated data from the ITC calorimeter with the Origin 7.0 software package supplied by MicroCal, subtracting the heat generated by the buffer. K_a (binding constant), ΔH (enthalpy change) and N (stoichiometry) values were calculated by applying the one-site fitting model.

1.2.5 Differential scanning fluorimetry

*HpKDO8PS*_metal-free was diluted to 5 μM in a white 96-well plate using the identical buffer as mentioned above. A 5000X solution of the dye SYPRO Orange (Sigma Aldrich, USA) was added to each well to achieve a final concentration of 5X. Measurements were performed in 100 μL in the presence and absence of 75 μM metal ion (zinc or cadmium chloride). 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 (ThermoFisher Scientific, USA). The raw fluorescence data were plotted as a function of temperature, which generated a sigmoidal curve [Figure 7 (A)]. The inflection point of the transition curve (T_m , midpoint melting temperature) was calculated using Prism 5 (GraphPad software, USA) by applying the Boltzmann sigmoidal fitting [Figure 7 (B)].

1.2.6 Preparation of compounds for virtual screening

A total of 415 compounds from our in-house database were processed for docking with the Sybyl-X suite v. 1.3 (Certara, USA) [39]. The compounds were prepared by adding hydrogen atoms and rectifying wrong valences and then energy-minimized using the conjugated gradient method in the Tripos force field with the Gasteiger-Hückel charge method until a convergence value of $0.001 \text{ kcal } \text{Å}^{-1} \text{mol}^{-1}$ was reached.

1.2.7 Preparation of protein structures

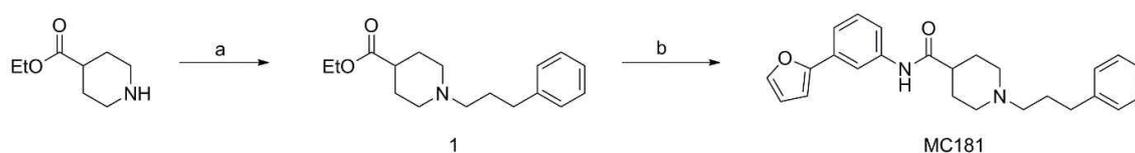
The crystal structure of *Aa*KDO8PS in complex with API (PDB: 1JCX) [29] was selected to gain insight into the ligand binding site. Superimposition of apo*Hp*KDO8PSwt and *Aa*KDO8PS in complex with API revealed that the C-terminal active sites of the two enzymes are highly conserved (motifs C18/11, K47/41, N54/48, R55/49, S56/50, A108/102, K130/124, H207/185, and E241/222). Therefore, the API molecule from the *Aa*KDO8PS-API structure was extracted and merged into our structure. The resulting API:*Hp*KDO8PS complex was energy-minimized. Both proteins were prepared with the Protein Preparation module of the Sybyl-X suite v. 1.3 [39] using the default parameters.

1.2.8 Virtual screening

Receptor-based virtual screening was performed to obtain new compounds harboring the desired activity profiles. The chemical database containing natural products and synthetic compounds was docked into the validated *Hp*KDO8PS binding site using Surflex-Dock [39]. First, a protomol was generated to define the active site, with a threshold of 0.4 and bloat set of 1 around the embedded ligand. Core interactions (Lys47, Asn54, Arg55, Ser56, Ala108, and Lys130) were also assigned in the protomol. The number of docking runs was set to 50, and other parameters were set as the Surflex-Dock Geom default settings. The final hit compounds were evaluated for binding by combining the consensus scoring function Cscore3 (Cscore >3) and Surflex-Dock total score ($-\log K_d$). Visual inspection considering the important interactions was necessary to evaluate binding.

1.2.9 Synthesis and preparation of ligands

Hyperin was prepared as reported by Lee and colleagues [40], and avicularin was purchased (Jinan Boss Chemical Industry, China). MC181 was synthesized using the following 2-step method.



Scheme 1. (a) 3-Phenylpropylbromide, potassium carbonate, DMF, 70°C, 2 hours; (b) i. 3N-HCl, microwave 150°C, 20 min; ii. 3-(2-furyl)aniline hydrochloride, Et₃N, EDC, DMAP, room temperature, 3.5 hours.

Step 1) ethyl 1-(3-phenylpropyl)piperidine-4-carboxylate (Compound 1)

A mixture of ethyl-4-piperidinecarboxylate (500 mg, 3.18 mmol), 3-phenylpropylbromide (576.2 μ L, 3.82 mmol), and potassium carbonate (in excess) prepared in DMF (1 mL) was stirred at 70°C for 2 hours. The reaction mixture was cooled to room temperature and then diluted with ethyl acetate, washed with water and brine, dried on anhydrous MgSO₄, and concentrated *in vacuo*. The residue was purified by flash chromatography on silica gel (MeOH : CH₂Cl₂ = 1 : 20) to generate ethyl 1-(3-phenylpropyl)piperidine-4-carboxylate (863.9 mg, 98.6%), referred to as Compound 1.

¹H NMR (600 MHz, CD₃OD): δ 7.26-7.23 (m, 2H), 7.19-7.17 (m, 2H), 7.16-7.13 (m, 1H), 4.13-4.09 (m, 2H), 2.89 (d, J = 11.46 Hz, 2H), 2.61 (t, J = 15.24 Hz, 2H), 2.38-2.36 (m, 2H), 2.35-2.30 (m, 1H), 2.07 (t, J = 22.02 Hz, 2H), 1.91-1.87 (m, 2H), 1.84-1.81

(m, 2H), 1.75-1.68 (m, 2H), 1.23(t, $J = 14.28$, 3H); ^{13}C NMR (150 MHz, CDCl_3): δ 175.07, 142.10, 128.37, 128.27, 125.72, 60.25, 58.14, 52.94, 41.16, 33.70, 28.57, 28.20, 14.21, 14.20; LC/MS (ESI⁺) m/z : 276.3 [$M + \text{H}$]⁺.

Step 2) N-(3-(furan-2-yl)phenyl)-1-(3-phenylpropyl)piperidine-4-carboxamide (MC181)

A mixture of Compound 1 (79.6 mg, 0.289 mmol) and 3N-HCl (2 mL) was subjected to microwave synthesis (Monowave 300) at 150°C for 20 min. The reaction mixture was concentrated *in vacuo* to produce the corresponding acid, and the resulting residue was used for the next reaction without further purification. A mixture of the acid (20 mg, 0.070 mmol), 3-(2-furyl)aniline hydrochloride (12.9 mg, 0.081 mmol), and triethylamine (113 μL , 0.81 mmol) in CH_2Cl_2 (10 mL) was added to 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (77.7 mg, 0.405 mmol) in a dropwise manner and stirred at room temperature for 1.75 hours. The mixture was added to 4-dimethylaminopyridine (3.5 mg, 0.03 mmol) and stirred at room temperature for 75 min. The reaction mixture was diluted with CH_2Cl_2 , washed with water and brine, dried on anhydrous MgSO_4 , and concentrated *in vacuo*. The crude product was purified by flash chromatography on silica gel ($\text{MeOH} : \text{CH}_2\text{Cl}_2 = 1 : 20$) to generate MC181 (10 mg, 37%).

^1H NMR (600 MHz, CD_3OD): δ 7.94 (t, $J = 4.2$ Hz, 1H), 7.54 (dd, $J = 1.8, 0.66$ Hz, 1H), 7.43-7.41 (m, 1H), 7.31 (t, $J = 15.8$ Hz, 1H), 7.28-7.25 (m, 2H), 7.21-7.20 (m, 2H), 7.18-7.15 (m, 1H), 6.73 (dd, $J = 3.36, 0.66$ Hz, 1H), 6.50 (dd, $J = 3.39, 1.92$ Hz, 1H), 3.22 (d, $J = 12.1$ Hz, 2H), 2.66 (t, $J = 11.94$ Hz, 2H), 2.67-2.62 (m, 2H), 2.53-2.48 (m, 1H), 2.40 (t, $J = 20.7$ Hz, 2H), 1.96-1.89 (m, 6H); ^{13}C NMR (150 MHz, CD_3OD): δ 174.08, 153.51, 142.11, 142.09, 141.15, 138.90, 131.45, 128.06, 127.99, 125.7, 119.08, 118.64,

114.94, 114.93, 111.32, 111.29, 104.99, 104.97, 57.37, 52.26, 42.16, 32.90, 27.35, 27.12;

LC/MS (ESI⁺) *m/z*: 389.3 [*M* + H]⁺.

1.2.10 NMR spectroscopy

NMR experiments [41, 42] were performed at 298 K using a Bruker Avance DRX 600 MHz spectrometer equipped with a 5-mm TXI ($^1\text{H}/^{13}\text{C}/^{15}\text{N}$) probe. The NMR sample was prepared as a mixture of 5 μM *Hp*KDO8PS and 0.2 mM ligand, which were dissolved in a solution containing 98% D_2O and 2% DMSO. On- and off-resonance irradiations were applied at chemical shifts of 0 ppm and -30 ppm, respectively. *Hp*KDO8PS was saturated using a train of Gaussian-shaped 50-ms-long pulses. The total length of the saturation train was set to 2 s.

Prior to acquisition, a 15-ms spin-lock pulse with a 1-W strength (T_2 filter) was applied to remove protein signals from the STD spectrum. The ^1H NMR spectrum was acquired with 32K real points and 6,320 scans.

Single pseudo-2D data from two serial free-induction decays were divided into two separate 1D data (on- and off-resonance). To increase the signal-to-noise ratio, the 1D raw data were processed with a 0.5 Hz line broadening and exponential window function prior to Fourier transformation. For comparison, the STD effects of individual peaks were quantified by the simple equation $(I_{\text{off}} - I_{\text{on}})/I_{\text{off}}$, where I_{on} and I_{off} represent the absolute intensities of the on- and off-resonance spectra peaks, respectively.

The off-resonance spectrum is identical to a conventional ^1H NMR spectrum. The largest STD value was set to 100%, and the other STD values were normalized to the largest value [43, 44].

WaterLOGSY NMR experiments [45] were conducted to confirm the results of the STD experiments. All of the WaterLOGSY spectra were recorded at 298 K using a

Jeol ECA 600 MHz spectrometer equipped with a 5-mm triple resonance inverse probe. The ligand (0.2 mM) was dissolved in the presence and absence of 5 μ M *Hp*KDO8PS in a solution containing 10% D₂O and 90% H₂O in a total volume of 300 μ L. To selectively excite water, Gaussian-shaped 20-ms pulses were irradiated at approximately 4.7 ppm. The mixing time for the magnetization transfer was set at 2 s. The NMR spectra were acquired with 256 transients and 8K data points. WATERGATE pulse sequences were interleaved in the WaterLOGSY sequences to suppress water signals. To increase the signal-to-noise ratio, free-induction decays were processed with a 0.5 Hz line broadening and exponential window function prior to Fourier transformation.

For STD-NMR competition experiments, PEP or A5P was added to the NMR samples, which contained either *Hp*KDO8PS and hyperin or *Hp*KDO8PS and MC181. The same acquisition and processing parameters mentioned above were used to obtain the STD NMR spectra.

The peaks of the 1D ¹H NMR spectra were assigned using a server (www.acdlabs.com/resources/ilab) for the assignment of peaks and analysis of STD and WaterLOGSY data. All STD and WaterLOGSY spectra were analyzed in a spectral range of approximately 6 ppm to 8 ppm, where aromatic protons appear.

1.3 Results

1.3.1 Protein expression and structure determination

Several *HpKDO8PS* constructs were prepared using the pCold I vector (Takara, Japan). Five types of different chaperone-expressing BL21 competent cells (pG-KJE8, pGRO7, pKJE7, pG-TF2, and pTF16, Takara, Japan) were tested for the most efficient production of soluble protein extracts. Among them, the pTF16 chaperone-expressing BL21 competent cells provided the best results, and soluble protein was prepared (Figure 1).

We determined four crystal structures, namely, apo*HpKDO8PS*wt, *HpKDO8PS*_H204A, *HpKDO8PS*-Cd, and *HpKDO8PS*-PEP-Zn (Figure 2). The data and refinement statistics are presented in the Table 2. The electron density map of apo*HpKDO8PS*wt showed some poorly resolved regions. Several residues between residues 212 to 218 in chain A and 211 to 217 in chain B exhibited poor electron density. In the N-terminal region, loop-forming residues 1 to 9 were also invisible on the electron density map, possibly due to their level of structural disorder.

In the refined model of *HpKDO8PS*_H204A, *HpKDO8PS*-Cd, and *HpKDO8PS*-PEP-Zn, two monomers were observed in each asymmetric unit, and they were well superimposed with the apo*HpKDO8PS*wt model, with root mean square (r.m.s.) deviation values of 0.50 Å, 0.36 Å, and 0.36 Å, respectively, for the 513, 514, and 514 C_α atom pairs. In these structures, two loop regions involving residues 1 to 9 and 210 to 220, which

correspond to the similar sites in apoHpKDO8PSwt, were also invisible.

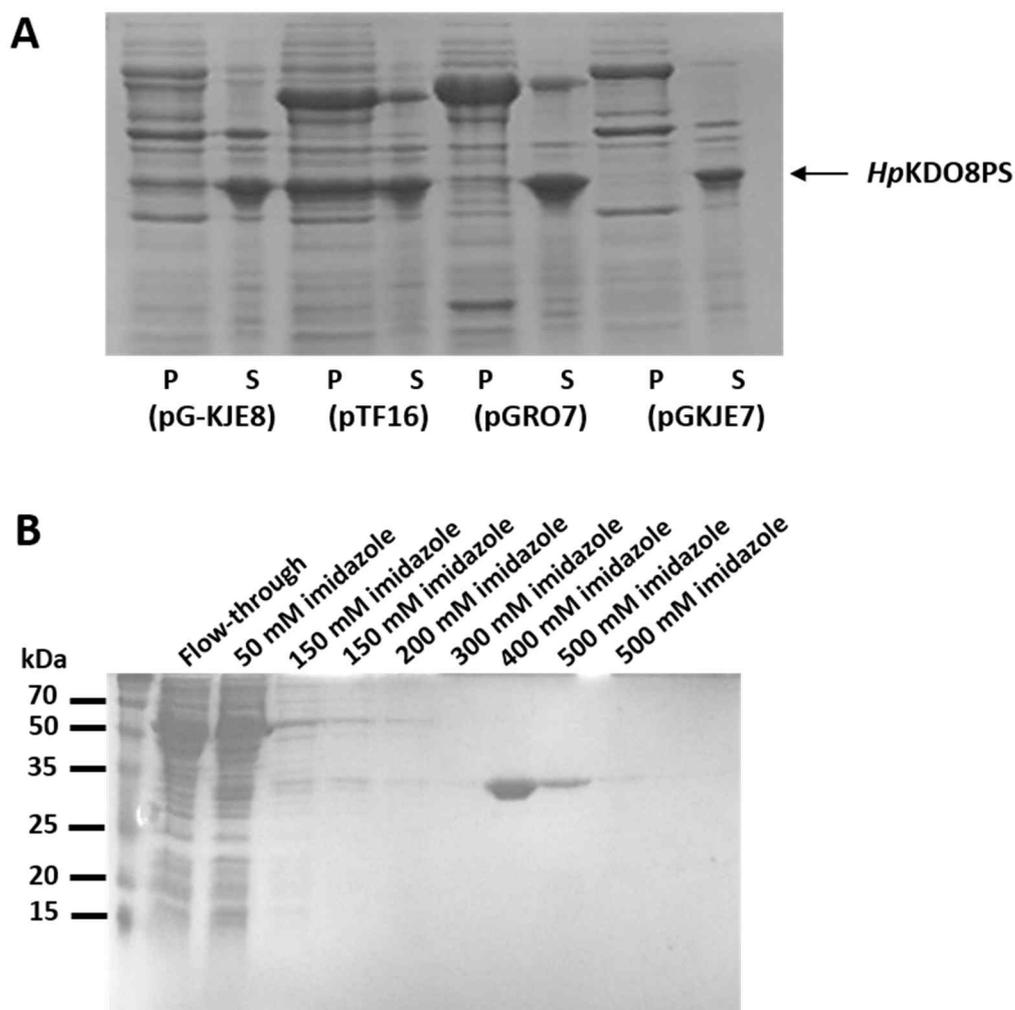


Figure 1. Solubility improvement and purification result of *HpKDO8PS* after co-expression with various chaperones. (A) Solubility-enhancing effects of chaperones are compared by SDS-PAGE. The chaperone-expressing plasmids are denoted below the electrophoresis data. “P” represents pellet, and “S” represents supernatant of the cell lysate. (B) Purification by immobilized metal affinity chromatography (IMAC).

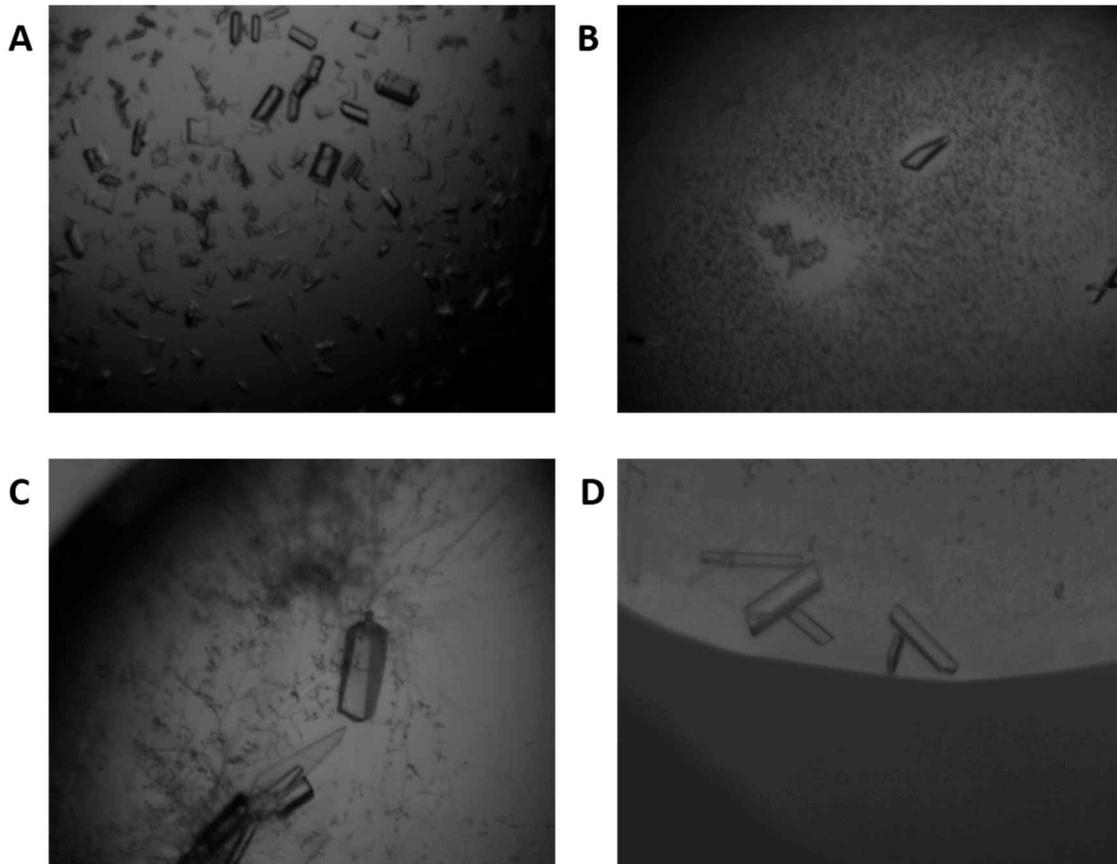


Figure 2. *HpKDO8PS* crystals: (A) apo*HpKDO8PS*, (B) *HpKDO8PS_H204A*, (C) *HpKDO8PS*-Cd, and (D) *HpKDO8PS*-PEP-Zn. The crystallization solution is consist of 18% v/v polyethylene glycol 3,350, 0.1 M HEPES at pH 7.5, and 0.25 M magnesium chloride hexahydrate.

Table 2. Statistics of data collection and refinement.

Data set	apoHpKDO8PSwt	HpKDO8PS_ H204A	HpKDO8PS- Cd	HpKDO8PS- PEP-Zn
Data collection				
Wavelength (Å)	0.90	1.00	1.00	1.00
Space group	C2	C2	C2	C2
Cell dimensions				
a (Å)	137.83	139.86	140.35	140.30
b (Å)	50.66	50.87	51.02	51.01
c (Å)	78.67	78.30	78.67	78.55
β (°)	110.5	104.73	104.7	104.37
Resolution (Å)	30-2.0	50-2.4	50-1.93	50-1.68
R_{merge} (%) ^[a]	8.4 (34.2) ^[e]	6.2 (15.5) ^[e]	7.4 (20.3) ^[e]	4.9 (27.7) ^[e]
I/σ(I)	46.45 (11.93) ^[e]	38.45 (13.41) ^[e]	32.85 (5.65) ^[e]	34.31 (2.12) ^[e]
Redundancy ^[b]	4.9 (4.4) ^[e]	4.1 (4.2) ^[e]	3.5 (1.9) ^[e]	4.0 (3.0) ^[e]
Completeness (%)	99.0 (93.3) ^[e]	96.1 (99.9) ^[e]	92.2 (67.7) ^[e]	99.2 (90.1) ^[e]
Unique reflections	34675	22173	41488	65566
Model refinement				
R_{work} ^[c] / R_{free} ^[d] (%)	18.69/23.91	18.29/23.83	19.17/21.18	18.02/20.10
Reflections	32093	21294	36113	61335
Protein atoms	4033	3984	4002	3990
Water molecules	293	180	352	528
Ligand/metal atoms	-	-	2	22
B factor (Å ²)				
Protein	22.39	25.15	24.47	19.04
Ligand/metal ion	-	-	23.95	18.28
Water	34.61	29.68	34.38	33.20
R.m.s derivations ^[f]				
Bond lengths (Å)	0.009	0.011	0.009	0.006
Bond angles (°)	1.270	1.469	1.362	1.168
R.m.s Z score ^[e]				
Bond lengths	0.45	0.58	0.55	0.32
Bond angles	0.64	0.74	0.72	0.57
Ramachandran plot (%)				
Favorable region	98.63	98.42	98.42	99.40
Outliers	0	0	0	0
MolProbity score	1.25	1.25	1.36	0.93
PDB code	4Z1A	4Z1B	4Z1C	4Z1D

[a] $R_{merge} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$, where $I(hkl)$ is the intensity of reflection hkl , \sum_{hkl} is the sum over all reflections and \sum_i is the sum over i measurements of reflection hkl .

[b] N_{obs}/N_{unique} .

[c] $R_{work} = \frac{\sum_{hkl} ||F_{obs} - k|F_{calc}||}{\sum_{hkl} |F_{obs}|}$.

[d] R_{free} was calculated same way as R_{work} , but with the 5% of the reflections excluded from the refinement.

[e] Values in parentheses indicate the highest resolution shell.

[f] Root-mean-square (r.m.s) deviation was calculated with REFMAC [46].

1.3.2 Crystal structures of *Hp*KDO8PS

*Hp*KDO8PS is a 30.3 kDa enzyme with two monomers in each asymmetric unit. The monomers adopt the $(\beta/\alpha)_8$ barrel topology [Figure 3 (A)], similar to their previously reported homologs. The secondary structures were assigned using the STRIDE server [48]. Both apo*Hp*KDO8PSwt monomeric structures (*i.e.*, the metal- and substrate-free forms of wild-type *Hp*KDO8PS), as well as *Ec*KDO8PS and *Aa*KDO8PS [23], are similar to each other, with r.m.s. deviations of 1.38 Å and 1.08 Å, respectively, with 463 and 478 equivalent C α s and harboring sequence identity levels of 46.9% and 51.5%, respectively. An extra α -helix (HE, residues 160–166) is found between the H5 α -helix and S6 β -strand in *Hp*KDO8PS (Figure 3); a 3_{10} -helix between S8 and H8 (residues 246–249) was also found, as in *Aa*KDO8PS. However, due to poor electron density, the 3_{10} -helix was not observed in the structure of *Ec*KDO8PS (not shown). No hairpins were observed at the N-terminus of *Hp*KDO8PS and *Ac*KDO8PS [Figure 3 (C)], in contrast with *Ec*KDO8PS.

The crystal structure of the H204A mutant (*Hp*KDO8PS_H204A) was similar to that of apo*Hp*KDO8PSwt. However, several residues located primarily at the active site (Lys52, Asn54, Arg55, Gln133, Asp252, and Asn255) were oriented in different directions in *Hp*KDO8PS_H204A compared with the corresponding residues in apo*Hp*KDO8PSwt [Figure 4 (A) and (B)]. These residues are charged amino acids or have similar polar groups that can form hydrogen bonds. Asn255 in *Hp*KDO8PS_H204A is flipped into the active site, forming a hydrogen bond with a water molecule (water-7) [Figure 4 (B)]. In addition, water-7 and water-8 are linked together by a hydrogen bond. In apo*Hp*KDO8PSwt, Glu241, Asp252, and Asn255 are moved toward the space

previously filled with His204 [Figure 4 (B)]. The distance between the residues in *HpKDO8PS_H204A* is shorter compared with apo*HpKDO8PSwt*, with the Glu241-Asp252 distance changing from 9.0 Å to 8.2 Å, the Glu241-Asn255 distance changing from 7.2 Å to 5.3 Å, and the Asp252-Asn255 distance from 7.6 Å to 5.0 Å. In the *HpKDO8PS_H204A* structure, a hydrogen bond between Arg55 and Ser251 is absent; therefore, Arg55 moves toward the outside of the active site [Figure 4 (A) and (B)]. Lys52, which hydrogen bonds with water-5 in apo*HpKDO8PSwt*, interacts with water-12 and Asn54 through multiple hydrogen bonds, moving closer to Asn54 by a distance of 5.0 Å compared with the Lys52-N-Asn54-O distance of 9.3 Å observed in apo*HpKDO8PSwt* [Figure 4 (A) and (B)].

To analyze the metal geometry in the active site, crystals of the Cd²⁺-bound form (*HpKDO8PS-Cd*) were obtained. A Cd²⁺ ion is bound to the metal binding residues (Cys18, His204, Glu241, and Asp252) of each protomer, forming a distorted octahedral geometry with a water molecule (*i.e.*, water-19) [Figures 4 (C) and 5 (A)]. In the active site, several residues and water molecules are linked to each other by hydrogen bonds making space for substrate binding [23]. Compared with the apo*HpKDO8PSwt* structure, the *HpKDO8PS-Cd* Asp252 residue moves toward Asn255 when it is bound to a Cd²⁺ ion [Figure 4 (A) and (C)]; the distance between the two residues then changes from 7.6 Å to 4.8 Å by linking a water molecule (*i.e.*, water-17).

Instead of forming a hydrogen bond with Ser251, the Arg55-NH residue is bound to water-24, and the Arg55-Nε residue is linked to Thr56 via water-23, favoring a rigid conformation [Figure 4 (A) and (C)]. The link between Ser49-water-5-Lys52 [Figure 4 (A)] is complex in *HpKDO8PS-Cd* due to water-18 and water-20, which mediate the interaction with Asn54 via two additional water molecules (*i.e.*, water-21 and water-22)

[Figure 4 (C)]. The overall structure of the active site could become compact through polar interactions. Consequently, the distances between Ser49, Lys52, and Asn54 become shorter. More precisely, between apo*Hp*KDO8PSwt and *Hp*KDO8PS-Cd, the Ser49-Lys52 distance changes from 4.0 Å to 3.3 Å, the Ser49-Asn54 from 12.0 Å to 9.7 Å, and the Lys52-Asn54 from 9.3 Å to 7.0 Å.

We attempted to crystallize *Hp*KDO8PS in complex with substrates to determine which residues participate in the binding of specific substrates, and *Hp*KDO8PS-PEP-Zn crystals were obtained using various substrates and metal ion combinations. In *Hp*KDO8PS-PEP-Zn, a Zn²⁺ ion is bound to a site identical to the one bound by the Cd²⁺ ion in *Hp*KDO8PS-Cd. This binding forms a distorted square pyramidal geometry involving neighboring residues and a water molecule (*i.e.*, water-**34**) [Figures 4 (D) and 5 (B)]. The overall folding conformation in *Hp*KDO8PS-PEP-Zn crystals appears to be similar to that of the model apo*Hp*KDO8PSwt, though some residues generate differences in the active site responsible for substrate binding. Indeed, several residues interact directly with PEP. Notably, Ser49 and Lys52 are linked to PEP O2', whereas Lys130 captures PEP O2P, O2, and O1 [Figure 4 (C)]. More water molecules are observed in the *Hp*KDO8PS-PEP-Zn active site, as they form hydrogen bonds with PEP and amino acids. Among them, water-**27** and water-**31** form a direct link with PEP O1P, whereas water-**28** and water-**30** bind to PEP O3P [Figure 4 (C)]. In particular, as in *Ec*KDO8PS and *Aa*KDO8PS, water-**31** is consistently found on the PEP *si* side, which is expected to trigger the condensation reaction [28, 29]. As the structure of *Hp*KDO8PS-PEP-Zn was obtained at higher resolution compared to apo*Hp*KDO8PSwt, it shows a more detailed water network together with Ser49, Lys52, Asn54, Arg55, Gln133, Ser251, Asp252, and Zn²⁺ [Figure 4 (D)].

In addition, structural comparison between *Hp*KDO8PS-Cd and *Hp*KDO8PS-PEP-Zn shows that the water molecules are located in similar positions in both crystals. Water-**18** in *Hp*KDO8PS-Cd is substituted to PEP O2' in linking Ser49 and Lys52 in *Hp*KDO8PS-PEP-Zn. Additionally, the connections found in *Hp*KDO8PS-Cd are consistently observed in *Hp*KDO8PS-PEP-Zn, starting at the trigonal link between Ser49, Lys52, and water-**20** (or water-**42** in *Hp*KDO8PS-PEP-Zn) and extending to Asn54 via water-**21** and water-**22** (or water-**40** and water-**43**) [Figure 4 (C) and (D)].

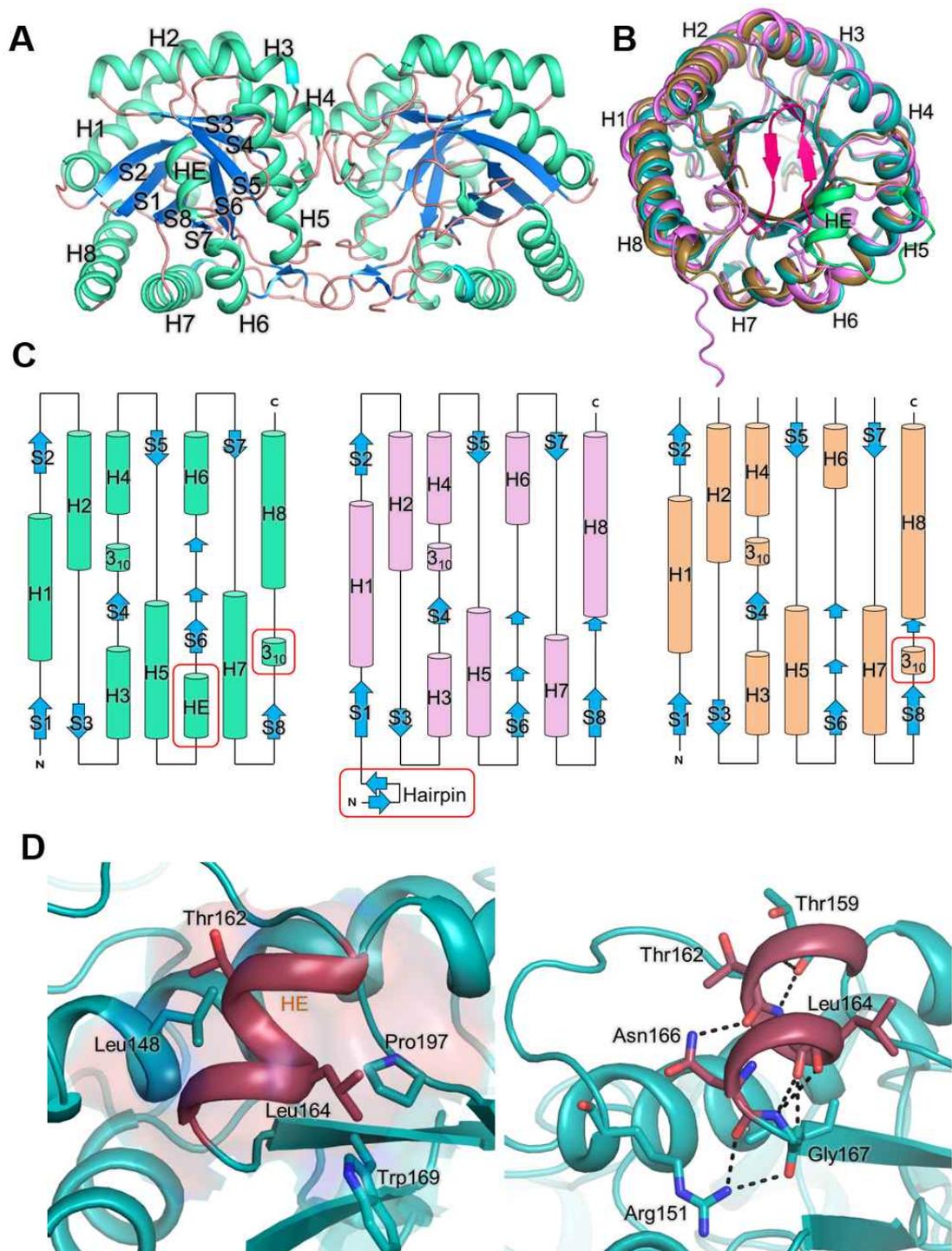


Figure 3. 3D structure of ApoHpKDO8PSwt (PDB: 4Z1A). (A) The crystal structure of apoHpKDO8PSwt exhibited a $(\beta/\alpha)_8$ topology. The secondary structures are drawn in a cartoon diagram with distinctive structures in different colors (α -helix, *green*, β -strand, *cyan*, and loop, *orange*). (B) ApoHpKDO8PSwt is superimposed onto EcKDO8PS (PDB: 1D9E) and

*Aa*KDO8PS (PDB: 1FX6) for comparison. Both structures are drawn in a cartoon diagram (*Ec*KDO8PS, *pink*, *Aa*KDO8PS, *brown*, and *Hp*KDO8PS, *green*). A hairpin in *Ec*KDO8PS is highlighted in *hot pink*. An extra helix (HE) in *Hp*KDO8PS is highlighted in *lime green*. (C) Topological models of *Hp*KDO8PS, *Ec*KDO8PS, and *Aa*KDO8PS. Cylinders represent α -helices (*Ec*KDO8PS, *pink*, *Aa*KDO8PS, *orange*, and *Hp*KDO8PS, *green*), and *cyan* arrows represent β -strands. The differences between models are highlighted in *red* boxes. (D) Interactions of HE (*raspberry*). HE interacts with neighboring residues in H5 and S6 as well as loops through both hydrophobic interaction (left) and hydrogen bonding (right). The interactions are shown in the surface diagram by *black* dashed lines.

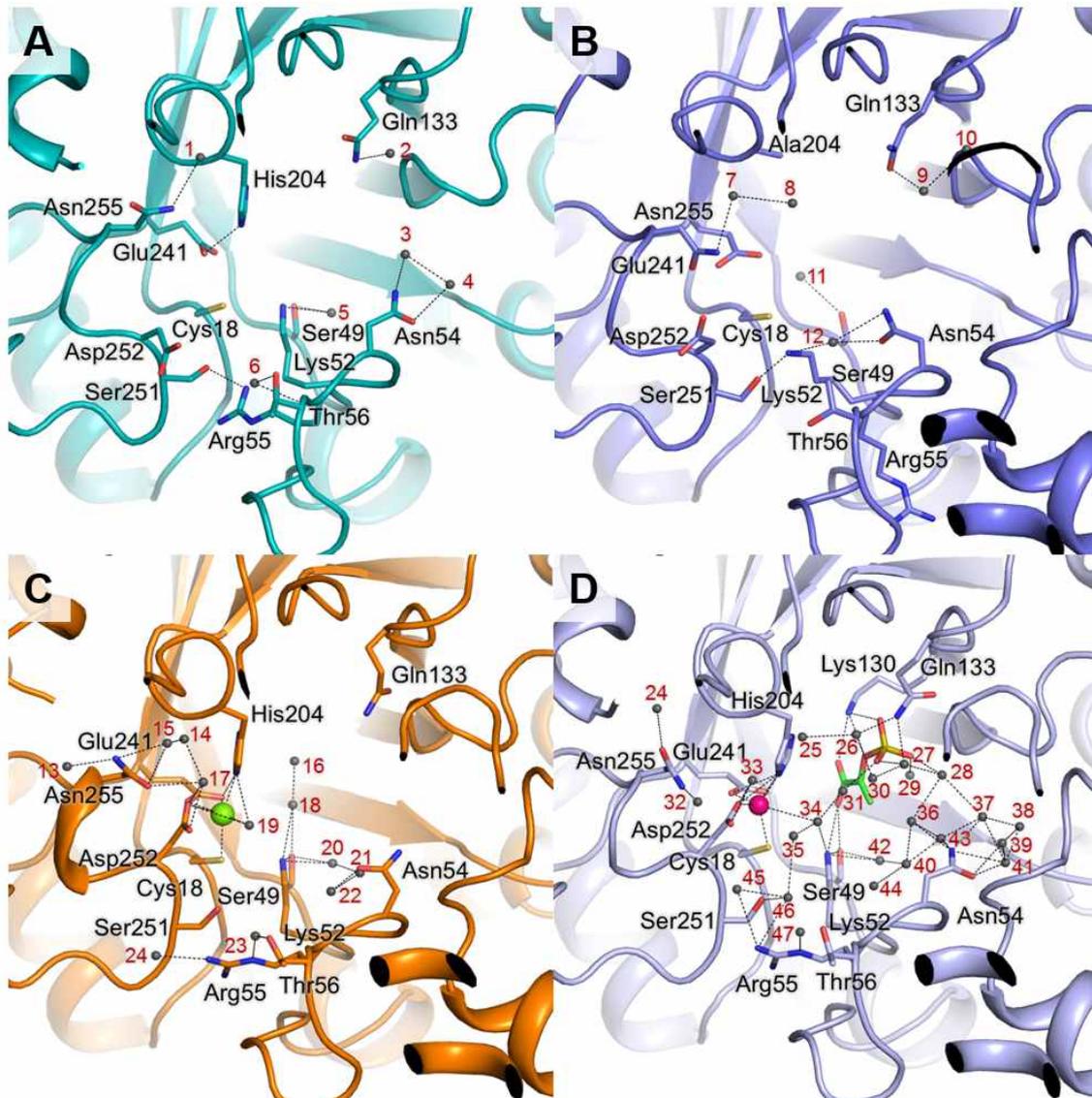


Figure 4. Interactions within the active sites in four *HpKDO8PS* X-ray crystal structures. The residues related to the interactions are shown as sticks; water molecules (numerically ordered) are drawn as *gray* balls; hydrogen bonds are drawn as *black* dashed lines. (A) Interactions between apo*HpKDO8PS*wt (*green*) and water molecules. (B) Interactions between *HpKDO8PS_H204A* (*cyan*) and water molecules. (C) Interactions between *HpKDO8PS-Cd* (*orange*) and water molecules. The cadmium ion (Cd^{2+}) is shown as a *green* ball. (D) Interactions between *HpKDO8PS-PEP-Zn* (*light blue*) and water molecules. The zinc ion (Zn^{2+}) is shown as a *pink* ball, and PEP is shown as a stick (carbon, *green*; oxygen, *red*; phosphorus, *olive*).

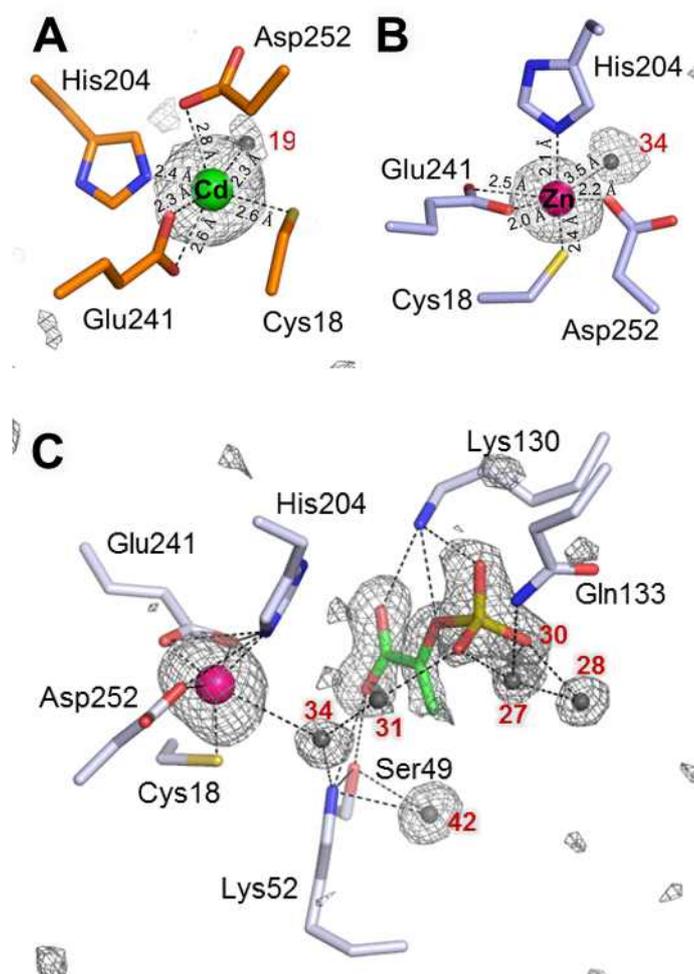


Figure 5. Metal and PEP binding sites in *HpKDO8PS*. Omit difference maps ($F_o - F_c$) contoured at 3.0σ (gray) are shown around the metal, water and PEP. (A) Cd^{2+} (green ball) has a distorted octahedral coordination in *HpKDO8PS*, in contrast with the geometry in gas phase or *AaKDO8PS*, which exhibit a tetrahedral or square pyramidal coordination, respectively. (B) Zn^{2+} (pink ball) shows a distorted square pyramidal geometry. The metal geometry of the Cd^{2+} ion and Zn^{2+} ion in *HpKDO8PS* is remarkably opposite to that in *AaKDO8PS*. (C) Zn^{2+} (pink ball) and PEP binding sites in *HpKDO8PS*. PEP (carbon, green; oxygen, red; phosphorus, olive) interacts directly with Ser49, Lys52, and Lys130. Gln133 also participates in PEP binding via water molecules (gray balls).

1.3.3 Thermal scanning for metal and metal-free *HpKDO8PS* interaction

The thermostabilizing effects of Cd^{2+} and Zn^{2+} on *HpKDO8PS* were measured by ITC and DSF experiments. The ITC data with Cd^{2+} injection were well fitted to the single-site binding in the isotherm model (Figure 6). However, Zn^{2+} ion titration with the enzyme caused an exothermic reaction up to the 6th injection and an endothermic reaction in the following injections (from the 7th to the 20th injections), generating data that could not be fitted to the single-site binding isotherm model. Indeed, *HpKDO8PS* contains a single site for a Cd^{2+} ion, and the binding affinity (K_d) is 460.5 ± 75.5 nM, with a heat change (ΔH) of -5741 ± 61.83 cal/mol.

The DSF experiments were conducted as described in the Experimental procedures section. The fluorescent dye (SYPRO Orange) binds to the hydrophobic residues of the protein and fluoresces. During the thermal unfolding process, hydrophobic regions of the protein are exposed, and fluorescence intensity increases with enhanced dye binding [49]. Thus, thermostability can be evaluated from the shift in T_m . In the absence of metal, T_m was up to 48°C, suggesting destabilization. In contrast, T_m was observed to shift to 63°C in the presence of Cd^{2+} and was 59°C in the presence of Zn^{2+} (Figure 7). This result supported the hypothesis that Cd^{2+} contributes more to the thermostability of *HpKDO8PS* more than Zn^{2+} .

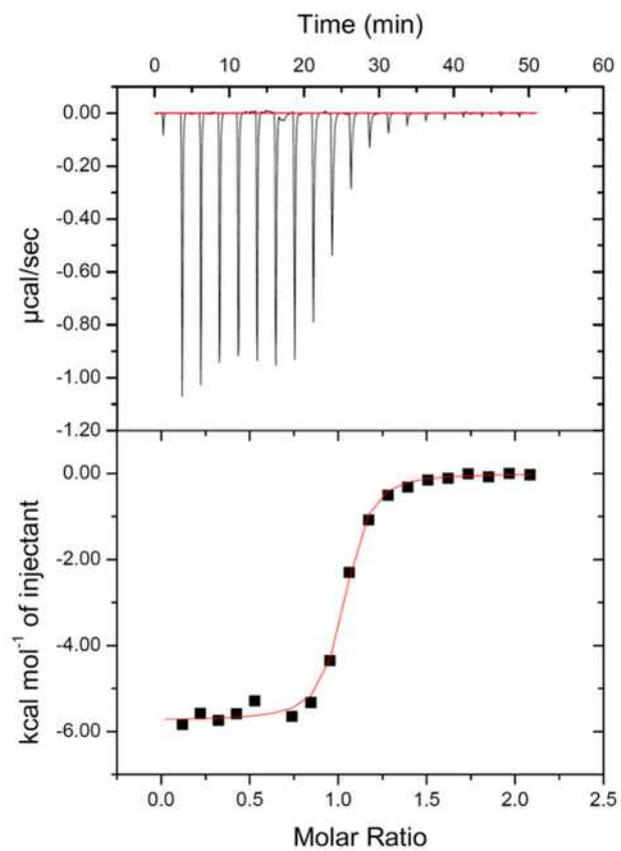


Figure 6. Cd²⁺ titration onto metal-free *HpKDO8PS* (*HpKDO8PS*_metal-free). Cd²⁺ injection induces an exothermal reaction and stabilizes the enzyme. The ITC data were well fitted in a single-site binding isotherm model.

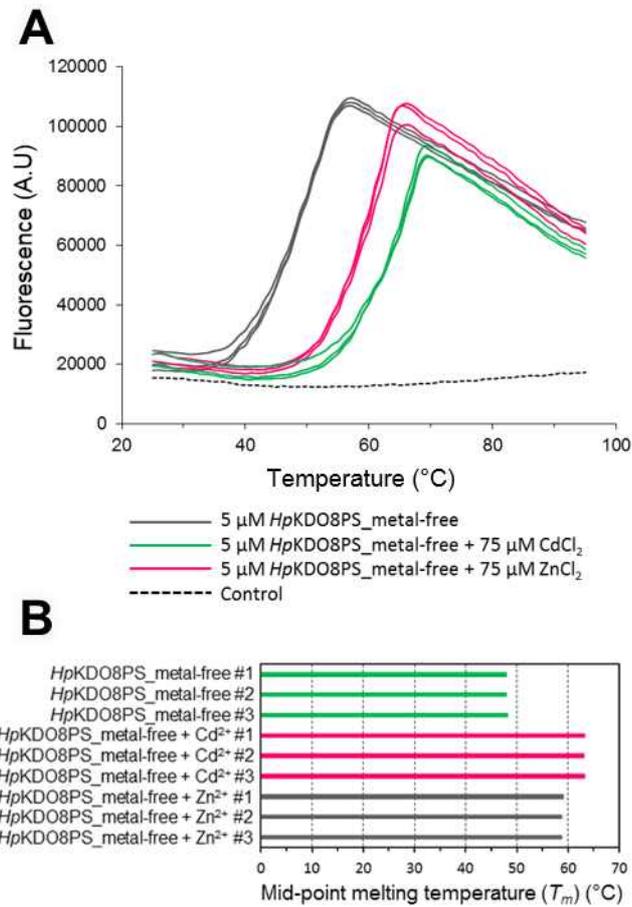
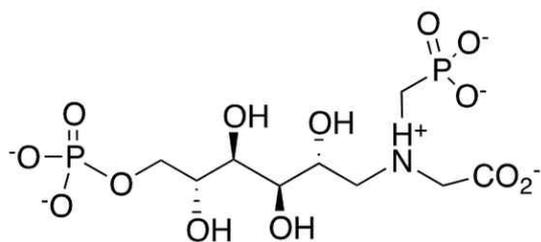


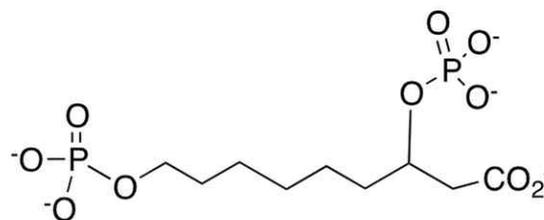
Figure 7. Thermal shift assay of *HpKDO8PS_metal-free* by differential scanning fluorimetry measured in the presence and absence of metal (Cd^{2+} or Zn^{2+}). (A) Protein thermal denaturation is depicted by a fluorescence versus temperature plot in the presence of SYPRO Orange. (B) Mid-point melting temperatures (T_m) calculated from the melting curve inflections shown in Figure 7 (A).

1.3.4 Validation of the docking method

Before virtually screening our in-house chemical database with the Surflex-Dock program [39], the docking protocol was evaluated for its ability to reproduce the binding modes of known KDO8PS inhibitors [API and 2,8-bis(phosphonoxy)-octanoic acid] (Figure 8). API [29], which mimics the intermediate form of the KDO8PS substrate condensation reaction, was used as a reference molecule by re-docking it into the active site of apo*Hp*KDO8PSwt. The API phosphate and phosphonate groups formed a hydrogen bond network with Ala108, Lys130 and Asn54, which was highly consistent with the reported co-crystal structure (not shown) [29], thereby validating our docking protocol.



API



2,8-Bis(phosphonoxy)-octanoic acid

Figure 8. Chemical structures of known KDO8PS inhibitors.

1.3.5 *Hp*KDO8PS-targeted virtual screening

A step-wise strategy for virtual screening was employed in our study to identify novel *Hp*KDO8PS inhibitors, and various criteria were applied in combination to select hit compounds, as outlined in Figure 9. First, 65 top-ranked compounds (out of 415 docked) were selected based on their Surflex-Dock energy score and Cscore (Cscore >3) [39]. Then, their interactions with the receptor active site were then analyzed to select compounds with the required hydrogen bond interactions. Compounds that made contact with residues that have the desired hydrogen bonds (residues Asn54/Arg55 and Lys130/Ala108) were considered to be real hits. The third filter was a visual inspection of the binding pose considering the diversity of ligand scaffolds, such as the number of hetero atoms, hydrogen bond donor/acceptor, different types of aromatic, and non-aromatic ring and alkyl groups. Finally, 21 compounds, all of which have drug-like profiles obeying Lipinski's rule-of-five: no more than 5 hydrogen bond donors (the total number of nitrogen–hydrogen and oxygen–hydrogen bonds); no more than 10 hydrogen bond acceptors (all nitrogen or oxygen atoms); a molecular mass less than 500 daltons; and an octanol-water partition coefficient (log P) not greater than 5 [50], were selected for further bioactivity evaluation. For comparison, the 21 compounds were docked into *Hp*KDO8PS-Cd, *Ec*KDO8PS, and *Aa*KDO8PS; the docking scores and ranks of the 21 compounds are listed in Table 3. In addition, virtual screening using those structures was conducted (Table 4). The ranking of Surflex-Dock scores is slightly worse for the *Hp*KDO8PS-Cd docking results and falls significantly when the compounds were docked into KDO8PS from other species (*E. coli* and *A. aeolicus*).

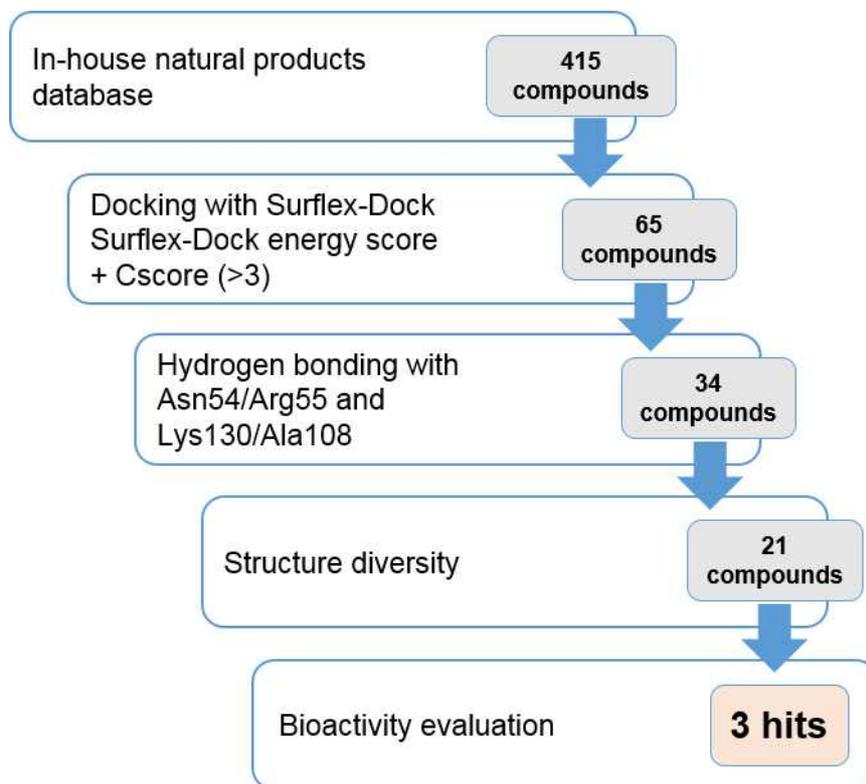


Figure 9. Scheme of virtual screening.

Table 3. ADME properties and docking scores of selected hits for virtual screening.

Compound	M.W. [a]	LogP [b]	TPSA [c]	HBD [d]	HBA [e]	RB [f]	Surflex-Dock Score (-log Ka)				Rank of Surflex-Dock Score				similarity Score (predicted by Surflex Sim)	Rank of Similarity Score
							Metal- free	Cd	Ec KDO8PS	Aa KDO8PS	Metal- free	Cd	Ec KDO8PS	Aa KDO8PS		
AX_33	332.39	0.45	119.61	5	7	8	10.87	7.56	5.64	7.59	4	52	76	46	6.64	29
MC159	455.59	4.73	52.65	1	5	9	10.64	6.69	5.83	6.89	6	95	57	87	5.38	268
SP_08	466.53	1.7	131.88	5	8	6	10.29	8.65	5.41	8.52	8	11	71	17	6.2	109
MC009	394.47	3.51	80.49	1	7	7	9.77	8.06	7.02	7.15	15	33	11	68	5.46	259
Hyperin	464.38	-0.02	206.6	8	12	4	9.75	7.96	5.37	5.91	16	37	83	128	6.34	77
MC162	421.53	3.59	66.59	2	6	8	9.62	7.71	6.37	7.26	18	46	28	60	6.44	65
Avicularin	434.35	0.45	186.37	7	11	3	9.55	7.63	5.13	5.43	19	49	97	148	6.08	135
MC101	388.5	4.75	45.48	1	4	7	9.5	6.65	6.28	8.15	20	98	30	26	5.29	315
MC181	388.5	4.75	45.48	1	4	7	9.47	8.35	5.29	7.08	21	23	89	72	5.97	146
HV_03	388.45	-0.15	141.37	5	8	5	9.46	6.83	5.66	7.65	22	89	65	42	6.88	12
MC014	434.53	4.46	76.58	1	7	6	9.4	7.88	5.38	8.14	23	40	82	29	5.87	213
PC_16	388.45	0.03	141.37	5	8	5	8.93	6.93	5.36	7.64	41	85	84	44	6.6	34
BK_13	360.4	2.12	88.38	3	6	6	8.9	7.35	6.15	6.3	45	60	37	110	7.25	2
MC144	403.52	3.86	63.05	1	6	7	8.87	8.45	5.87	6.98	47	18	52	79	5.44	260
BK_16	360.4	1.88	88.38	3	6	7	8.86	7.72	5.4	6.73	48	44	81	92	6.58	37
MC116	467.52	3.47	96.45	2	8	9	8.85	7.56	5.91	9.51	50	50	50	6	5.62	237
MC104	424.51	3.76	59.51	1	6	6	8.79	8.4	6.9	8.43	54	21	13	19	5.49	256
MC173	477.56	2.45	113.38	1	7	7	8.78	7.31	4.29	6.97	55	63	165	82	5.47	258
PN_10	384.38	-0.35	134.91	4	9	8	8.76	7.1	4.51	8.95	56	78	144	11	7.13	3
MC105	383.46	3.59	71.57	3	5	7	8.7	7.16	7.54	6.79	61	75	9	90	6.1	133
MC106	430.54	5.16	50.8	1	5	7	8.7	7.17	5.49	7.14	62	74	70	70	5.31	301
API (reference)	414.22	-11.21	266.57	10	14	12	11.35	11.46	7.67	11.03	2	1	8	2	-	1

[a] Molecular Weight

[b] Partition coefficient

[c] Topological polar surface area

[d] The number of hydrogen bond donor

[e] The number of hydrogen bond acceptor

[f] The number of rotatable bonds

Table 4. Docking scores of KDO8PS variants.

(A) <i>Hp</i> KDO8PS-Cd			(B) <i>Ec</i> KDO8PS			(C) <i>Aa</i> KDO8PS		
Compound	Surflex-Dock Score (-log K _d)	Rank of Surflex-Dock Score	Compound	Surflex-Dock Score (-log K _d)	Rank of Surflex-Dock Score	Compound	Surflex-Dock Score (-log K _d)	Rank of Surflex-Dock Score
MC007	9.86	4	HV_15	8.52	3	BM_37	13.00	1
HV_7	8.98	8	PGS_3	8.17	5	SP_6	10.76	3
SP_8	8.65	11	MC105	7.54	9	SP_9	9.18	8
MC144	8.45	18	PN_16	7.15	10	PC_6	8.69	15
PC_15	8.45	19	MC009	7.02	11	MC104	8.43	19
HV_10	8.41	20	PGA_4	6.79	15	MC102	8.15	25
MC104	8.40	21	PGS_1	6.55	19	HV_9	8.14	28
HV_14	8.38	22	SP_14	6.55	20	MC014	8.14	29
MC181	8.35	23	HV_9	6.48	23	MC016	8.13	30
HV_8	8.08	31	MC011	6.47	24	SP_5	7.87	35
HV_11	8.07	32	HV_14	6.46	25	PN_9	7.87	36
MC009	8.06	33	SP_4	6.43	26	MC012	7.60	45
Hyperin	7.96	37	MC008	5.97	46	AX_33	7.59	46
BM_41	7.95	38	SS_4	5.93	47	BK_30	7.57	48
SP_2	7.94	39	TR_1	5.81	58	PGA_5	7.49	51
MC014	7.89	40	PC_6	5.60	68	HV_5	7.47	53
BK_16	7.72	44	CT_3	5.47	74	PGA_6	7.42	55
Avicularin	7.63	49	CW_4	5.42	80	MC004	7.41	56
AX_33	7.56	52	API (reference)	7.67	8	API (reference)	11.03	2
PC_4	7.53	53						
API (reference)	11.46	1						

1.3.6 NMR analysis of interactions between hits and *HpKDO8PS*

To validate the results from our virtual screening, we applied saturation transfer difference (STD) NMR spectroscopy and water-ligand observed via gradient spectroscopy (waterLOGSY) to evaluate whether the individual hit compounds bind to *HpKDO8PS* [41, 42, 45, 51].

STD NMR spectroscopy is a useful approach to detect a reduction in a ligand's NMR signal, which is caused by the nuclear Overhauser effect (NOE) between the protein and its ligand [42]. The on-resonance spectrum records a 1D ¹H NMR signal for each compound, which is reduced by the magnetization transfer from *HpKDO8PS*, whereas the off-resonance spectrum serves as a reference [Figures 10 (B), 11 (B), and 12 (B)] that is identical to a standard 1D ¹H NMR spectrum for ligands [52]. The STD spectrum represents the difference between the on- and off-resonance signals [Figures 10 (D), 11 (D), and 12 (D)], and the signals on the STD spectrum reflect the interaction between the protein and its ligand. Figures 10 (C), 11 (C) and 12 (C) show representative STD spectra for each compound in the absence of *HpKDO8PS*.

Among the 21 potential ligands derived from our *in silico* virtual screening, three ligands, avicularin (quercetin-3-*O*- α -L-arabinofuranoside) (Figure 12), hyperin (quercetin-3-*O*- β -D-galactopyranoside) (Figure 10), and MC181 (*N*-(3-(furan-2-yl)phenyl)-1-(3-phenylpropyl)piperidine-4-carboxamide) (Figure 11), produced STD spectra that indicated interactions with *HpKDO8PS*. Because hyperin and avicularin [Figures 10 (A) and 12 (A)] have identical aromatic moieties involved in *HpKDO8PS*

binding and exhibited only a minor difference in their auxiliary sugar moieties (β -D-galactopyranose for hyperin and α -L-arabinofuranose for avicularin), the following analysis was performed only for hyperin.

All of the five protons of hyperin exhibited STD signals, among which proton-2 showed the largest reduction in signal ratio between the on- and off-resonance spectra [Figure 10 (D)]. For comparison, the levels of STD signals of the other protons were normalized to that of proton-2, as shown in Figure 10 (A). Proton-1 exhibited the second-largest STD signal, which represented 75 percent of that of proton-2. The protons from the sugar moieties could not be observed because of buffer signals, though the buffer demonstrated no STD signals as an internal reference.

For MC181, proton-6 exhibited the largest STD effect, and the STD effects of protons-1, -3, -5, -8, and -10 were quantified and normalized to that of proton-6, demonstrating relatively low STD effects [Figure 11 (A)]. To corroborate the STD data, waterLOGSY experiments were performed for hyperin and MC181. This technique uses magnetization transfer from water to the free ligand and the ligand-bound protein via intermolecular NOE [53, 54].

We compared the waterLOGSY spectra of the ligand in the absence [Figures 10 (G) and 11 (G)] and presence [Figures 10 (H) and 11 (H)] of *HpKDO8PS*, showing the peaks in the opposite signs. The protons corresponding to these peaks are believed to interact with *HpKDO8PS* via water molecules within the active site, and they were observed only at 6 ppm to 7.5 ppm (hyperin) and 6.5 ppm to 8 ppm (MC181). These results were highly consistent with those from the STD experiments.

To identify the *HpKDO8PS* ligand-binding site, STD competition experiments

were conducted using the natural substrates of KDO8PS, PEP [Figures 10 (E) and 11 (E)] and A5P [Figures 10 (F) and 11 (F)], as competitors. These substrates are well known to bind to different parts of the same active site in KDO8PS [23, 55].

The hyperin competition STD data showed that the STD peak intensities or hyperin in the presence of A5P were significantly decreased but that PEP was less affected in 6 ppm to 7.5 ppm of the hyperin STD spectrum compared with A5P [Figure 10 (E) and (F)]. These results indicate that hyperin competes with A5P for *Hp*KDO8PS binding, suggesting that hyperin binds to the A5P binding site (A5P-subsite) of *Hp*KDO8PS.

For MC181, the overall intensities of the STD spectra were significantly decreased in the presence of PEP or A5P [Figure 11 (E) and (F)], indicating that MC181 binds to a broad range of PEP and A5P binding sites within the *Hp*KDO8PS active site. This result is in agreement with the fact that MC181 is a long molecule.

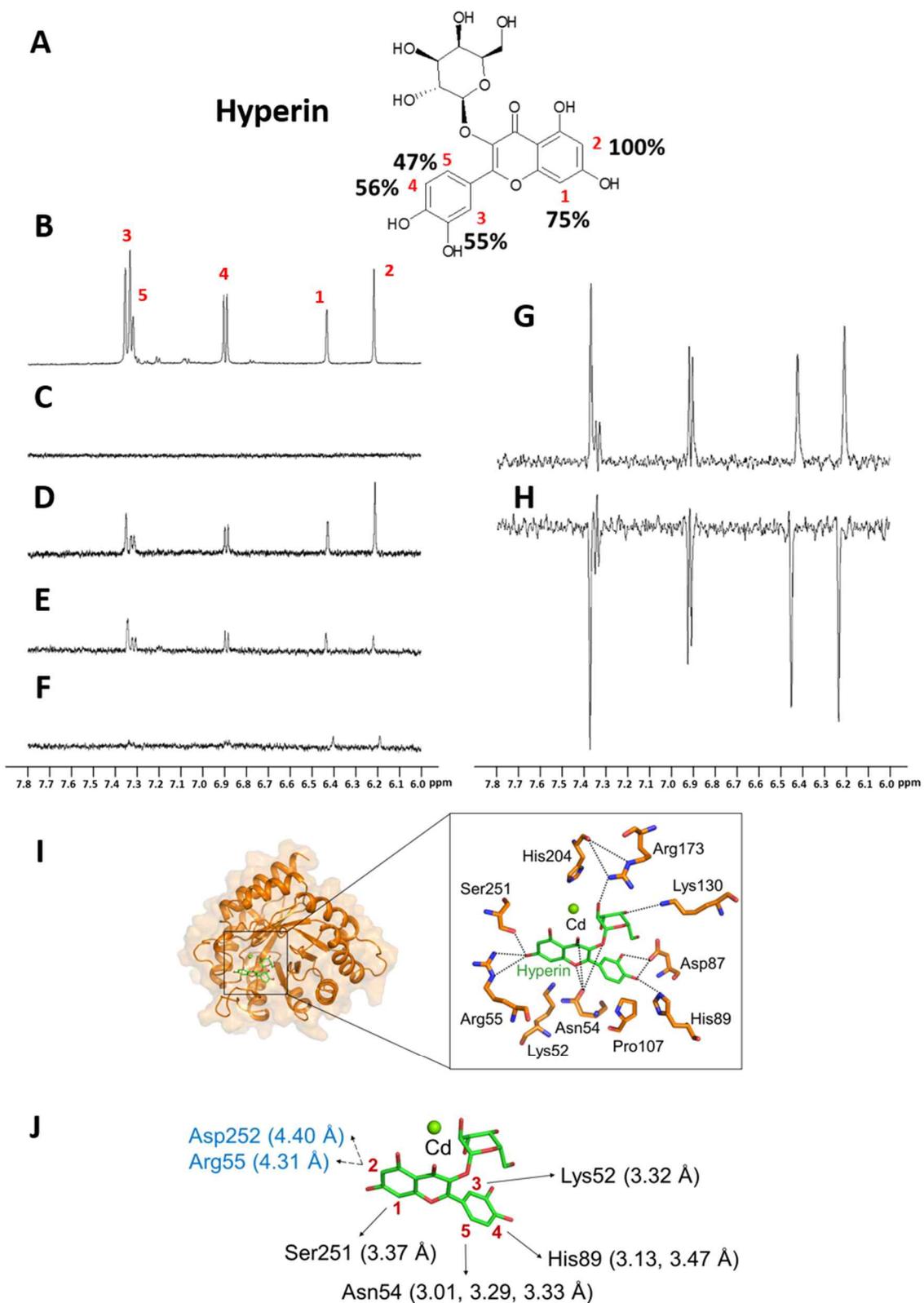


Figure 10. ^1H STD and WaterLOGSY NMR spectra in the aromatic region and docking conformation of hyperin. (A) Chemical structure of hyperin and epitope mapping (each proton is

numbered in *red*, and the values of the normalized STD effect are presented as percentages). (B) Reference $^1\text{H-NMR}$ spectrum of hyperin. (C) STD spectrum of hyperin in the absence of *HpKDO8PS*. (D) STD spectrum of hyperin in the presence of *HpKDO8PS*. (E) STD spectrum of hyperin in the presence of *HpKDO8PS* and PEP. (F) STD spectrum of hyperin in the presence of *HpKDO8PS* and A5P. (G) WaterLOGSY spectrum of hyperin in the absence of *HpKDO8PS*. (H) WaterLOGSY spectrum of hyperin in the presence of *HpKDO8PS*. (I) Docking conformation of hyperin. The distances between residues (carbon atoms, *orange*), shown as sticks, and the compound are less than 3.5 Å. Hyperin (carbon atoms, *green*) forms hydrogen bonds (*black* dashed lines) with Asn54, Arg55, Asp87, His89, Lys130, Arg173, and Ser251. (J) Distances between the protons of hyperin and residues. The protons exhibiting an STD effect are numbered as in Figure 10 (A). Although proton-2 exhibited the largest STD effect, it is not consistent with the distances between the proton and residues (dashed arrows and indicated in *blue*).

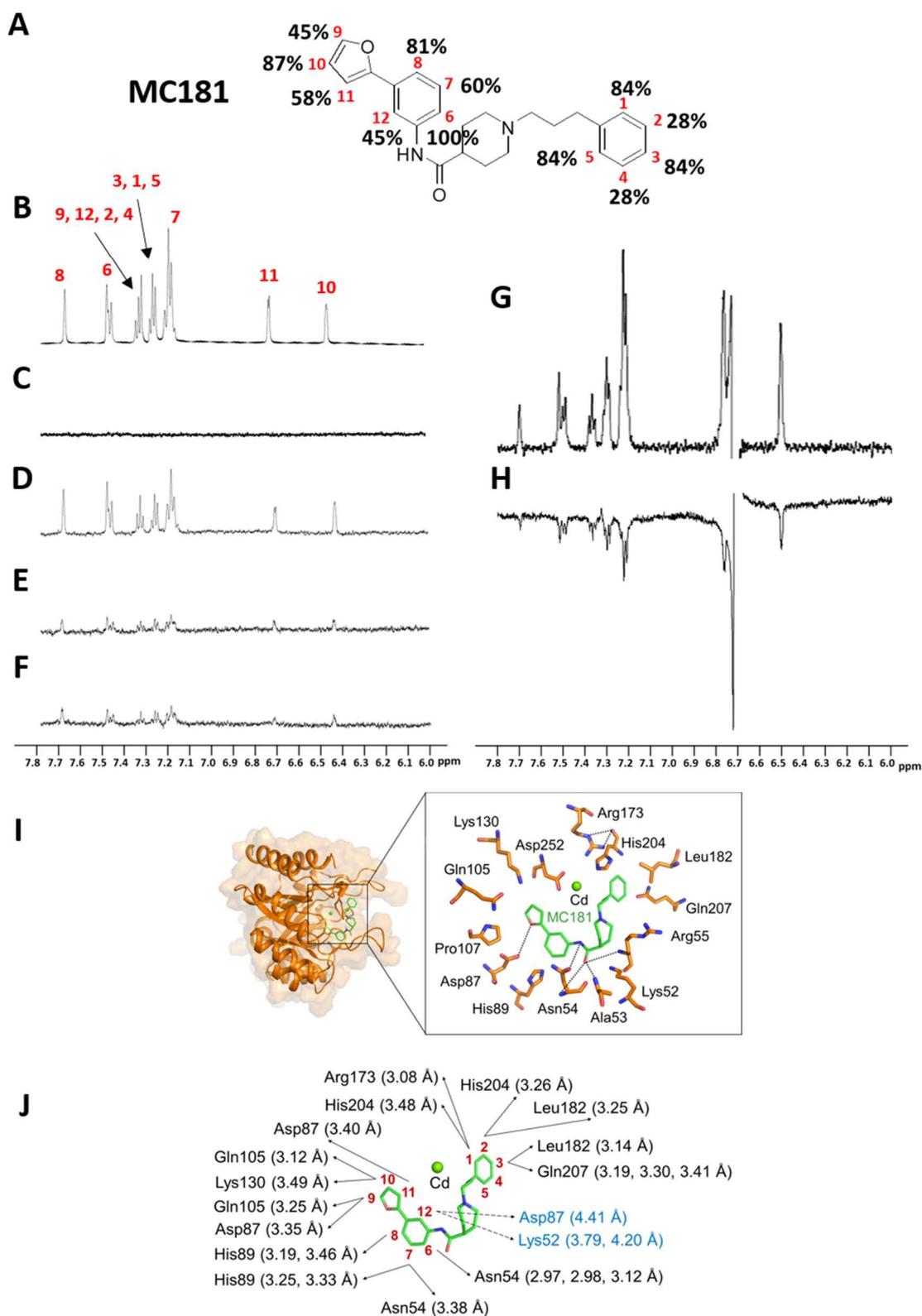


Figure 11. ^1H STD and WaterLOGSY NMR spectra in the aromatic region and docking conformation of MC181. (A) Chemical structure of MC181 and epitope mapping (each proton is

numbered in *red*, and the values of the normalized STD effect are presented as percentages). (B) Reference $^1\text{H-NMR}$ spectrum of MC181. (C) STD spectrum of MC181 in the absence of *HpKDO8PS*. (D) STD spectrum of MC181 in the presence of *HpKDO8PS*. (E) STD spectrum of MC181 in the presence of *HpKDO8PS* and PEP. (F) STD spectrum of MC181 in the presence of *HpKDO8PS* and A5P. (G) WaterLOGSY spectrum of MC181 in the absence of *HpKDO8PS*. (H) WaterLOGSY spectrum of MC181 in the presence of *HpKDO8PS*. (I) Docking conformation of MC181. The distances between residues (carbon atoms, *orange*), shown as sticks, and the compound are less than 3.5 Å. MC181 (carbon atoms, *green*) forms hydrogen bonds (*black* dashed lines) with Ala53, Asn54, Arg55, and Asp87. (J) Distances between the protons of MC181 and residues. The protons exhibiting an STD effect are numbered as in Figure 11 (A). The epitope mapping result is consistent with the distances between the compound and enzyme, except for proton-12 (dashed arrows and indicated in *blue*).

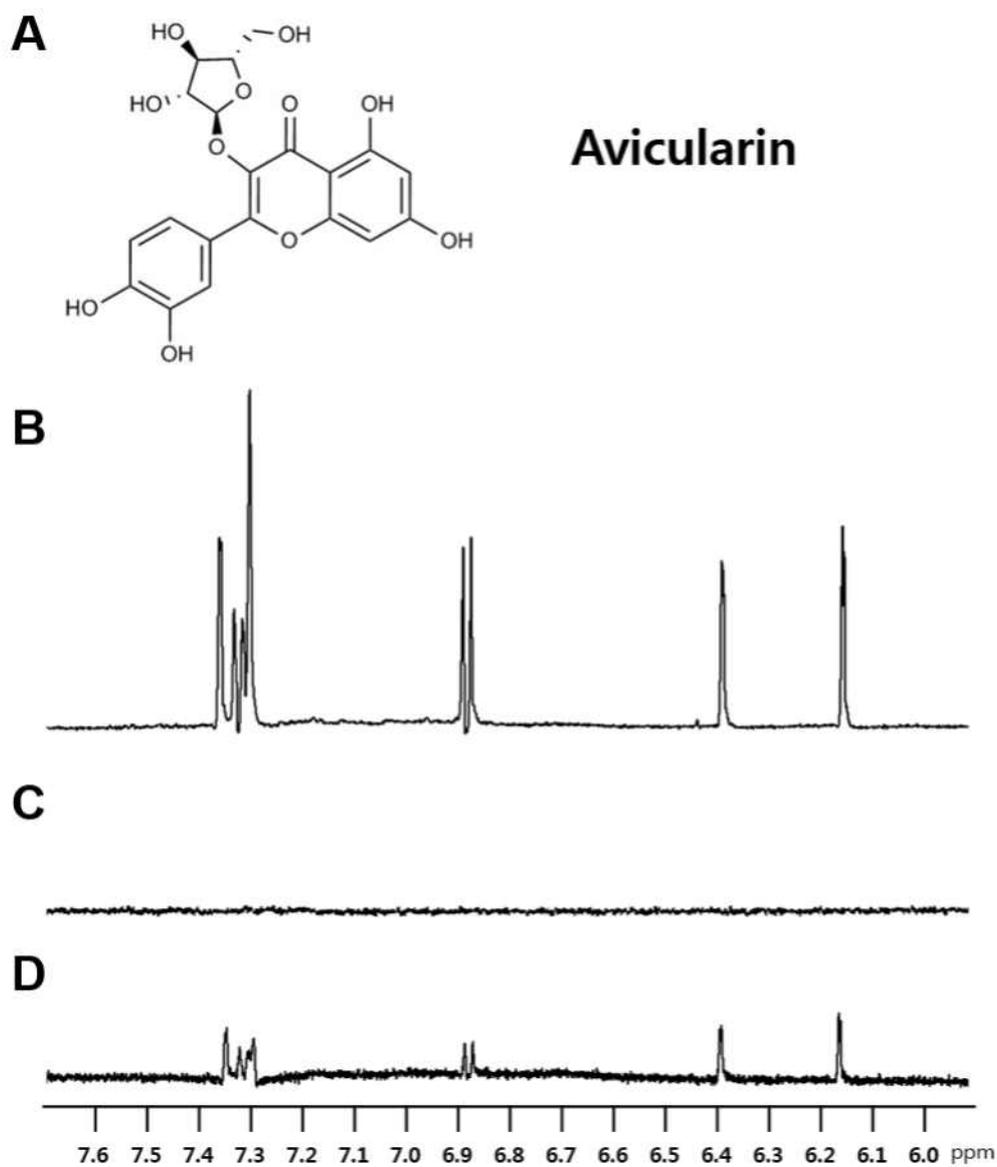


Figure 12. ^1H STD NMR spectra of avicularin in the aromatic region. (A) Chemical structure of avicularin. (B) Reference ^1H -NMR spectrum of avicularin. (C) STD spectrum avicularin in the absence of *HpKDO8PS*. (D) STD spectrum of avicularin in the presence of *HpKDO8PS*.

1.3.7 Docking analyses of *Hp*KDO8PS ligands

The binding modes of representative hit compounds were predicted by Surflex-Dock docking [39]. As shown in Figures 10 (I) and 11 (I), these results revealed that the hit compounds fit well in the area covering the combined binding sites for PEP and A5P, the natural KDO8PS substrates. The compounds were docked into the *Hp*KDO8PS-Cd structure, which is presumably more similar to the Zn²⁺-bound *Hp*KDO8PS structure than its apo-form. The binding modes of hyperin and MC181 were compared with that of the known inhibitor API [28] (Figure 13), which was used as the reference compound for our docking analysis.

According to the hyperin-bound models, the hydroxyl groups of the sugar moiety are docked into the PEP binding site (PEP-subsite) [23, 28] with hydrogen bonds to Lys130 and Arg173, which interact with nearby His204 [Figure 10 (I)]. However, hyperin exhibits no interaction with the metal because it is placed farther from the bound Cd²⁺ ion compared with the API binding site. In addition, the hydrogen bonds with Lys47, Lys52, and Ala108 in the API-bound model disappear in the hyperin-docked conformation [Figures 10 (I) and 13]. The hyperin benzopyran ring occupies the site where A5P binds (A5P-subsite) in the *Aa*KDO8PS structure [23, 28] [Figure 10 (I)], forming multiple hydrogen bonds with Asn54, Arg55 and Ser251, which are not observed in the API-bound model. The hyperin dihydroxyphenyl ring is placed near Asp87, His89, and Pro107 via hydrogen bonds and hydrophobic interactions [Figure 10 (I)].

In the MC181 binding mode analysis [Figure 11 (I)], the PEP-subsite [23] is

occupied by the MC181 furan ring, which interacts with Gln105, Pro107, and Lys130 at a short distance. Although the furan ring of MC181 is farther away from the Cd²⁺ ion than the PEP moiety in the API-bound model, the phenyl group of MC181 is well matched to API [Figures 11 (I) and 13]. Indeed, it binds to Leu182 via a hydrophobic interaction and also interacts with His204, Gln207, and Asp252. MC181 covers the A5P-subsite with its piperidine ring, and the peptide oxygen forms hydrogen bonds with Ala53, Asn54, and Arg55 [Figure 11 (I)]. Asn54 also forms a hydrogen bond with the peptide's nitrogen atom.

MC181 was also docked into *Ec*KDO8PS and *Aa*KDO8PS, and the results showed that the key residues for MC181 binding are different from those in *Hp*KDO8PS. MC181 binds to Asn62, Arg63, Ser64, Ala116, Lys138, His202, and Gln205 in *Ec*KDO8PS, or Asn48, Arg49, Ser50, Ala102, Lys124, His185, and Gln188 in *Aa*KDO8PS. Considering the protein sequence differences between *Hp*KDO8PS, *Ec*KDO8PS, and *Aa*KDO8PS, the binding key residues are 45% (*Ec*KDO8PS) or 52% (*Aa*KDO8PS) identical compared with those of *Hp*KDO8PS.

Analysis of the three catalytic pockets (*Hp*KDO8PS, *Ec*KDO8PS and *Aa*KDO8PS) revealed that although the active sites are highly conserved, the dimensions of the catalytic channel rim differ greatly [Figure 14 (A)]. Figure 14 (A) shows four regions, A–D, that represent the boundaries of the catalytic channel rim. Region A corresponds to Asn54 in *Hp*KDO8PS, Ser232 in *Aa*KDO8PS and Asn26 in *Ec*KDO8PS, region B to Arg55 in *Hp*KDO8PS and Arg49, Ser50 in *Aa*KDO8PS, region C to Gln207 in *Hp*KDO8PS, Phe103 in *Aa*KDO8PS and Phe117 in *Ec*KDO8PS, and region D to Phe134 in *Hp*KDO8PS and Pro190 in *Aa*KDO8PS. Variation at the A, B and C regions produces a significantly wider channel rim in the crystal structure of *Hp*KDO8PS

compared to the other two KDO8PSs. The structural manifestation of this variation is that the measured B-D distance is 6.21 Å in *Aa*KDO8PS compared to 11.7 Å in *Hp*KDO8PS, whereas the measured A-C distance is 10.06 Å in *Ec*KDO8PS compared to 20.6 Å in *Hp*KDO8PS.

The binding poses of MC181 in the active site of *Aa*KDO8PS or *Ec*KDO8PS revealed that the catalytic channel is too narrow to accommodate the entire MC181 molecule, with the furan ring and phenyl group of MC181 protruding out of the binding pocket [Figure 14 (B)]. Nonetheless, MC181 can bind deeply at the bottom of the *Hp*KDO8PS catalytic channel, which occupies both the PEP- and A5P-sites.

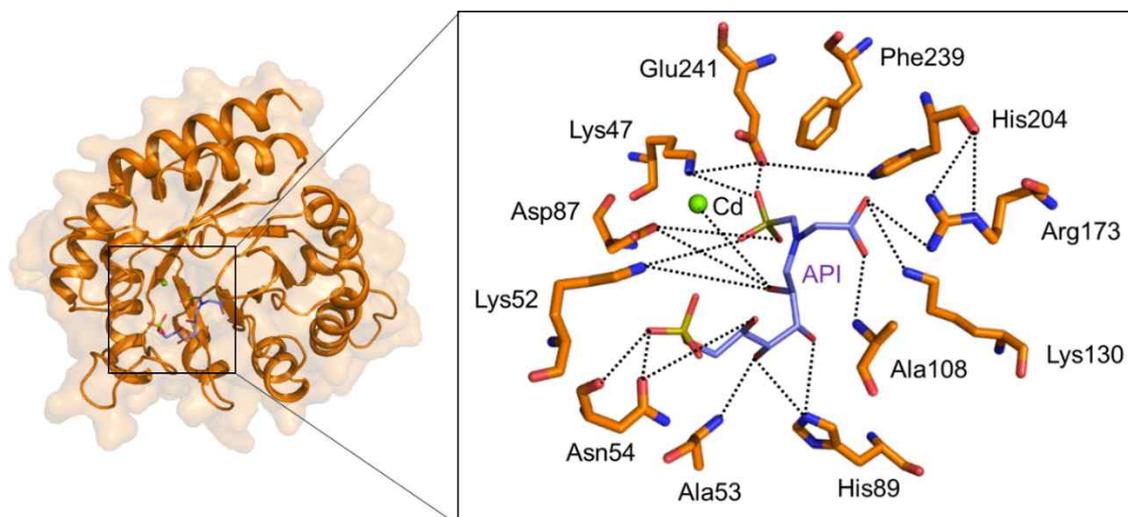


Figure 13. Docking conformation of API. The distances between residues (carbon atoms, *orange*), shown as sticks, and the compound are less than 3.5 Å. API (carbon atoms, *violet*) forms hydrogen bonds (*black dashed lines*) with Lys47, Lys52, Asn54, Ala53, Asp87, His89, Ala108, Lys130, Arg173, and Glu241.

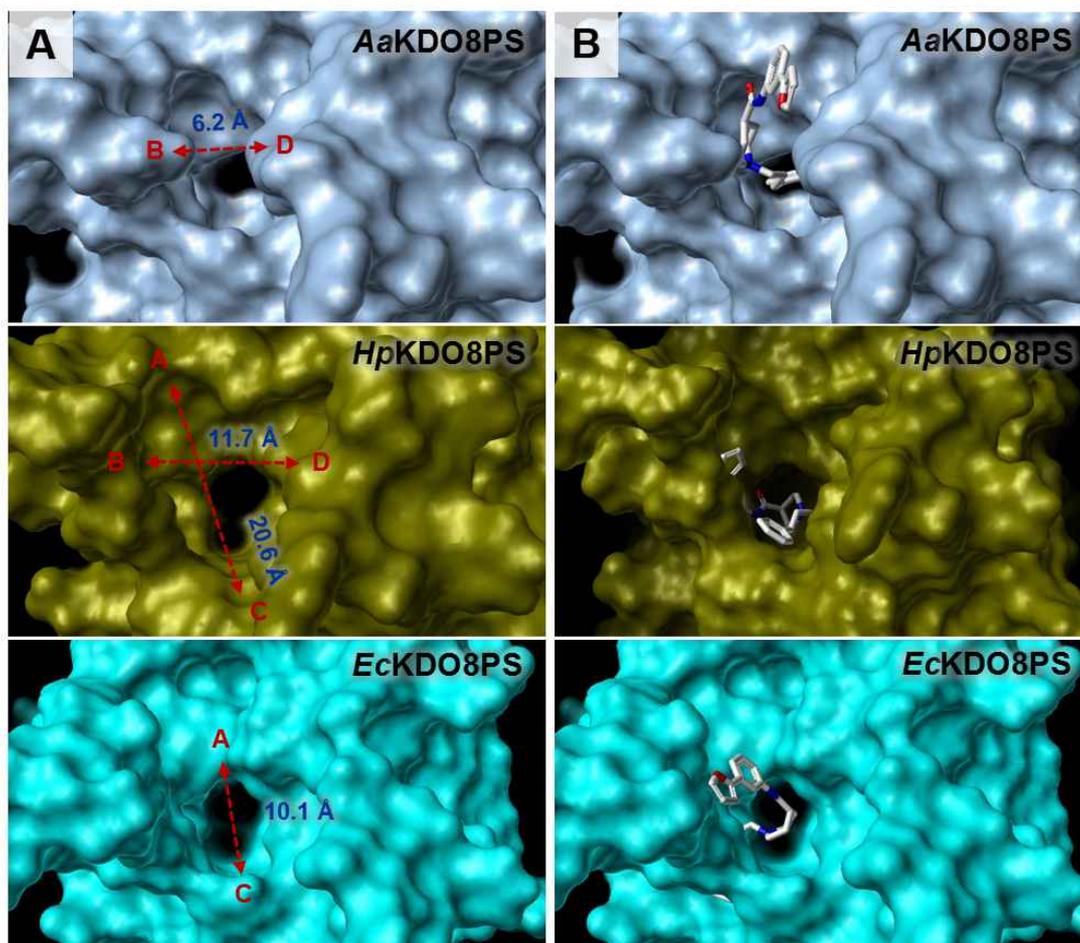


Figure 14. Top-down view of the active site of *AaKDO8PS* (top), *HpKDO8PS* (middle) and *EcKDO8PS* (bottom). (A) The distances between the boundaries of the catalytic channel rim are shown with red arrows. Letters A–D denote four boundary areas surrounding the channel rim. (B) The binding poses of MC181 in the active sites of *AaKDO8PS* (top), *HpKDO8PS* (middle) and *EcKDO8PS* (bottom).

1.4 Discussion

*Hp*KDO8PS possesses an extra helix (HE) (Figure 3) that is not observed in the structures of KDO8PS from *E. coli*, *A. aeolicus*, *B. cenocepacia* [26], *P. aeruginosa* [27], and *N. meningitidis* [25]. This HE interacts with its neighboring H5 α -helix, S6 β -strand, and loops (H6–S7 and HE–S6) through hydrogen bonds and hydrophobic interactions [Figure 3 (D)]. Arg151 interacts with Gly167 at the HE–S6 loop [Figure 3 (D), right panel], an interaction that affects the PEP-binding residues Lys130 and Gln133, which are located at the N-terminus of the loop. Moreover, Leu164, which is located in this same loop, forms a hydrophobic interaction with Pro197 in the H6–S7 loop [Figure 3 (D), left panel], which in turn affects the S7–H7 loop, where Gln207 and His204 form an A5P and a metal ligand-binding site, respectively. Based on these observations, the additional HE may play a role in indirectly stabilizing the conformation of the active site. However, further studies are needed to evaluate the possible role of the HE.

*Hp*KDO8PS is a metal-dependent enzyme that has four metal-coordinating residues, namely, Cys18, His204, Glu241, and Asp252, in the active site, similar to *Aa*KDO8PS [23]. These residues were mutated to determine whether they are related to protein stability and active site coordination. Of the four mutants (C18A, H204A, E241A, and D252A), the crystal structure of the H204A (*Hp*KDO8PS_H204A) was determined. As shown in our CD spectroscopy studies, *Hp*KDO8PS_H204A demonstrated a better stability in its secondary structure than the other mutants, which presumably led to crystallization success (Figure 15). The two other mutants, E241A and D252A, were unavailable due to low protein stability during purification.

The structure of *HpKDO8PS_H204A* [Figure 4 (B)] showed that the residues adjacent to His204 in the active site became closer to each other. Moreover, the conformations of Asn54, Lys52, and Asn255 were significantly altered with respect to our other wild-type structures. Because the residues adjacent to His204 are related to substrate and metal binding, His204 is believed to not only serve as a metal-binding residue but also as a contributor to active site formation. These results are consistent with a previous report showing that the His185 residue in the active site of *AaKDO8PS* is responsible for forming the substrate binding space [30].

The structures of *HpKDO8PS* in complex with Cd^{2+} or $\text{PEP}/\text{Zn}^{2+}$ (*HpKDO8PS*-*Cd* and *HpKDO8PS*-*PEP*-*Zn*) [Figure 4 (C) and (D)] were determined, and different hydrogen bond networks with water molecules were observed within the active site. In particular, the *HpKDO8PS*-*PEP*-*Zn* active site appeared rigid and therefore stabilized via its complicated hydrogen bond network [Figure 4 (D)]. By facilitating the condensation process, the water molecules in the structures appear to be involved in the formation of a stable complex between the protein and ligands. The idea of a protein-ligand interaction mediated by water was supported by waterLOGSY-NMR experiments [56].

Interestingly, in these structures, *HpKDO8PS* binds to Zn^{2+} and Cd^{2+} ions in distorted square pyramidal and octahedral geometries [Figure 5 (A) and (B)], respectively, in contrast to the *AaKDO8PS* enzyme structure [32]. Moreover, in contrast with Cd^{2+} , the ITC data (Figure 6) for Zn^{2+} , the naturally occurring metal ion binding to *HpKDO8PS*, did not fit to the single binding site isothermal model, possibly for the following reasons: (i) multiple binding modes could exist during the equilibration process between Zn^{2+} and *HpKDO8PS*; (ii) inevitable interactions could occur with the buffer; or (iii) metal ion precipitation could generate heat [57] (data not shown). The DSF data supported the idea

that Cd^{2+} enhances *HpKDO8PS* thermostability to a greater degree than Zn^{2+} (Figure 7). According to Krosky and colleagues, the K_{cat} value increases by approximately 2-fold, and the A5P K_{m} value decreases by approximately 6.5-fold upon replacement of the Zn^{2+} ion by Cd^{2+} in Zn^{2+} -bound *HpKDO8PS* [22]. Consistently, our metal coordination results indicated that the metal geometries of Cd^{2+} and Zn^{2+} could affect enzyme activity. Indeed, the properties of metal ions as Lewis acids play a particularly important role in biology: metal ions can activate coordinated ligands for reactivity by affecting either bond length, bond angles, or coordination site number [58]. In *HpKDO8PS*, Cd^{2+} may favor a more ideal coordination than Zn^{2+} for the enzymatic reaction. The metal binding properties will contribute not only to structure-based drug design by providing molecular insights, but also to development of metal complex inhibitors which directly binds to the active site of *HpKDO8PS*. Metal complexes have been applied to design matrix metalloproteinase inhibitors [59]. The inhibitors contain two components: an organic backbone that interact with the protein surface, and a metal-binding head group that coordinates the active-site metal ion [60].

To date, only a few inhibitors of KDO8PS have been reported (Figure 8), all of which were designed to mimic the intermediate form of the condensation reaction between A5P and PEP, and the lack of known ligands has limited the use of conventional ligand-based screening methods to identify novel KDO8PS inhibitors. Instead, structural information for the crystal structure of *HpKDO8PS* determined for the first time in our present study was utilized to identify novel scaffolds that specifically bind to *HpKDO8PS* using structure-based virtual screening.

This step-wise virtual screening approach successfully identified three novel chemotypes (avicularin, hyperin, and MC181) as *HpKDO8PS* inhibitors. To the best of

our knowledge, this is the first *in silico* study on the identification of novel KDO8PS inhibitors. Though the size of our in-house database is relatively small, all compounds in the database have drug-like properties for the following reasons: they are all lead-like; most compounds are newly identified natural product; and some compounds have anti-cancer or anti-bacterial activities (but their targets are unclear). For comparison, an API-based 3D similarity search against the same in-house database was conducted using the SurflexSim program [39]. Notably, three active compounds that directly bind to *Hp*KDO8PS were highly ranked in the output of Surflex-Dock scoring (docking ranks ranged from 16 to 21), and the ranking was significantly lower when they were docked into *Ec*KDO8PS and *Aa*KDO8PS, suggesting that the compounds may specifically bind to *Hp*KDO8PS. However, a simple 3D similarity search failed to select the most active compounds (similarity ranks ranged from 77 to 146), as shown in the Table 3. Taken together, the results demonstrated that the virtual screening approach was successful for the discovery of novel KDO8PS inhibitors.

The *Hp*KDO8PS binding modes of hyperin and MC181 were investigated using STD-NMR experiments and docking simulations. According to the docking results, the A5P-subsite is spatially occupied by the benzopyran ring of hyperin, which is consistent with the epitope mapping results from STD experiments [Figure 10 (A)], even though the STD values could not be easily converted to the distance between the protein and ligand [Figure 10 (J)]. This result from the docking experiments was also confirmed by competition data, which indicated that hyperin competes with A5P for the ligand binding site [Figure 10 (F)]. Although the sugar moiety of hyperin binds to the PEP-subsite, according to the docking simulation results, it was difficult to confirm these data by NMR experiments because the protons of the sugar moiety could not be detected on the NMR

spectra.

According to docking simulations, MC181 and API share similar binding *HpKDO8PS* modes, despite their different chemical structures (Figures 11 and 13). MC181, through hydrogen bonds and hydrophobic interactions, binds to and broadly covers the PEP- and the A5P-sites in *HpKDO8PS*. These results were highly consistent with the results obtained from STD competition experiments against PEP or A5P [Figure 11 (E) and (F)], which showed a decline in the STD spectral intensities. The epitope mapping results [Figure 11 (A)] were also consistent with the observed distances between the compounds and enzyme, as shown in Figure 11 (J). MC181 was also docked into *EcKDO8PS* and *AaKDO8PS* to investigate whether it specifically binds to *HpKDO8PS* (Figure 14). The active sites are highly conserved, yet the catalytic channel rims of the enzymes differ significantly [Figure 14 (A)], with *HpKDO8PS* having a wider channel rim, which is consistent with the MC181 docking results showing that MC181 binds well to the *HpKDO8PS* PEP- and A5P- sites. Conversely, due to their narrow channel rims, MC181 juts out from the active sites of *EcKDO8PS* and *AaKDO8PS* [Figure 14 (B)]. Based on the results, MC181 can be considered a promising scaffold for developing new antibiotics that specifically act against *H. pylori*.

The interactions observed for MC181 involved both hydrophobic and hydrophilic binding modes, whereas hyperin primarily interacted with *HpKDO8PS* via hydrophilic interactions. Furthermore, the compound docking results provided clues for the modification of these inhibitors. Hyperin and avicularin are derivatives of quercetin, representing a typical subclass of flavonoids [61], which have been reported to exert various biological effects, including antimicrobial, antihypertensive, neuroprotective, and chemoprotective effects [61-63]. Among those biological activities, the antimicrobial

effect is supported by the interactions between *Hp*KDO8PS and these compounds (hyperin [64, 65] and avicularin [66]).

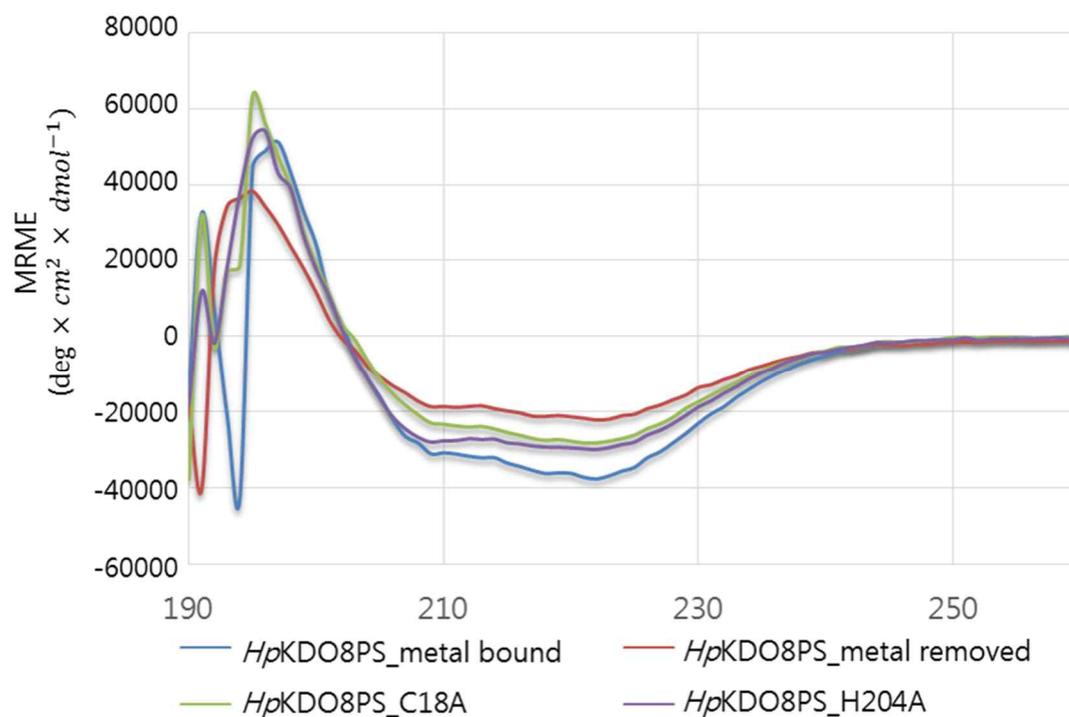


Figure 15. CD spectra of a set of various forms of *HpKDO8PS*. CD spectrum of *HpKDO8PS_metal-bound* was recorded for reference. Metal chelator-treated (*e.g.*, using EDTA and 1, 10-phenanthroline) *HpKDO8PS_metal-free* resulted in the worst folding among all of the protein forms tested. Two soluble mutants (*HpKDO8PS_C18A*, and *HpKDO8PS_H204A*) exhibited a protein folding conformation that was less affected compared with *HpKDO8PS_metal-free*. The spectrum of *HpKDO8PS_H204A* exhibited a slightly less damaged state compared with *HpKDO8PS_C18A*.

Chapter 2. Studies on the mechanism and the protein folding based on the structure of *Homo sapiens* NAD(P) dependent steroid dehydrogenase (NSDHL) and its disease-related mutants

2.1 Introduction

The cholesterol synthesis is an important cellular pathway to generate cholesterol which is an essential component of all animal plasma membranes and a precursor of steroid hormones [67]. The enzymes of cholesterol synthesis should be controlled carefully for the following reasons: (i) it is energetically costly to make cholesterol; (ii) its dysregulation causes human diseases [68, 69]. Specifically, the imbalance of the pathway is related to hypercholesterolemia, cardiovascular disease, neuropathy, and cancer [68-71]. To treat those human diseases, many drugs have been developed which target cholesterol biosynthesis enzymes [69, 72]. Especially, statins that control 3-hydroxy-3-methylglutaryl-CoA reductase (HMGCR) have been regarded as successful drugs for treating hypercholesterolemia but their significant adverse effects have been reported [68, 69, 73]. The effect of statins have been unraveled that they regulate HMGCR which catalyzes the third step of the mevalonate pathway (Figure 16).

Unlike HMGCR, the other enzymes in the pathway still remain less studied. Study on the enzymes could help to understand regulation of cholesterol homeostasis and suggest new strategies for therapy of cholesterol-related diseases [68]. NSDHL (sterol-4- α -carboxylate 3-dehydrogenase, decarboxylating) is one of them that catalyzes NAD⁺-dependent oxidative decarboxylation of the C4 methyl groups from 4 α -carboxysterol to produce the corresponding 3-keto, C4-decarboxylated products [74, 75] (Figure 17).

Like other enzymes in the post-HMGCR pathway, NSDHL has not only been emphasized for cholesterol synthesis, but also attracted attention due to accumulation of the intermediates related to cancer growth and signaling of proto-oncogenes [70, 76, 77]. In particular, NSDHL has been reported to make epidemic growth factor receptor (EGFR) reduce the expression and signaling [78]. And inhibition of NSDHL is regarded as a good strategy against tumor with activated EGFR-KRAS signaling [70]. Furthermore, NSDHL is related to breast cancer metastasis due to its translocation to the plasma membrane from the intracellular compartment, and its increasing expression in cell lines representing the metastatic progression [79].

NSDHL is an X-linked gene and its mutations cause Congenital Hemidysplasia with Ichthyosiform erythroderma and Limb Defects (CHILD) and CK syndrome (CKS) [80]. NSDHL with each missense mutation (A105V, A182P, or G205S) is causative of typical CHILD syndrome which is characterized by asymmetric skin and abnormal limb. Although wild-type NSDHL have been reported to be localized at the surface of lipid droplets, G205S could no longer be localized on the droplets [81]. On the other hand, CKS is caused by missing of K232 whose characteristics include intellectual disability, brain malformation, and dysmorphic features [80]. The CKS-causing NSDHL mutant has altered protein folding and it is characterized as temperature-sensitive protein [82].

Moreover, a missense mutation of NSDHL causing CHILD syndrome resulted in proteins that no longer localized on lipid droplets and could not function as a cholesterologenic enzyme [81].

To understand better the relationship of the protein misfolding and the diseases, NSDHL and its mutants (G205S and K232 Δ) have been studied on the protein thermal stability and the coenzyme regulation. To determine NSDHL structure, NSDHL have been purified, crystallized, and the diffraction data has been collected. By comparing to other Rossmann folds, key amino acids in NSDHL active site have been revealed. The thermal stability of the wild-type, G205S, and K232 Δ NSDHL has been measured using differential scanning fluorimetry (DSF). In addition, via isothermal titration calorimetry (ITC), NAD(H) has had stronger affinity than that of NADP(H), implying that NAD(H) is more reliable coenzyme than NADP(H) in the enzyme reaction. By analyzing the data acquired from saturation transfer difference (STD) NMR spectroscopy, the binding modes of NSDHL to the coenzymes were determined in an atomic resolution. And the NSDHL mutants have been observed that the coenzyme binds to them in different manners. As hypothesized, the mutants have different characteristics implying that the protein misfolding and the diseases are closely connected.

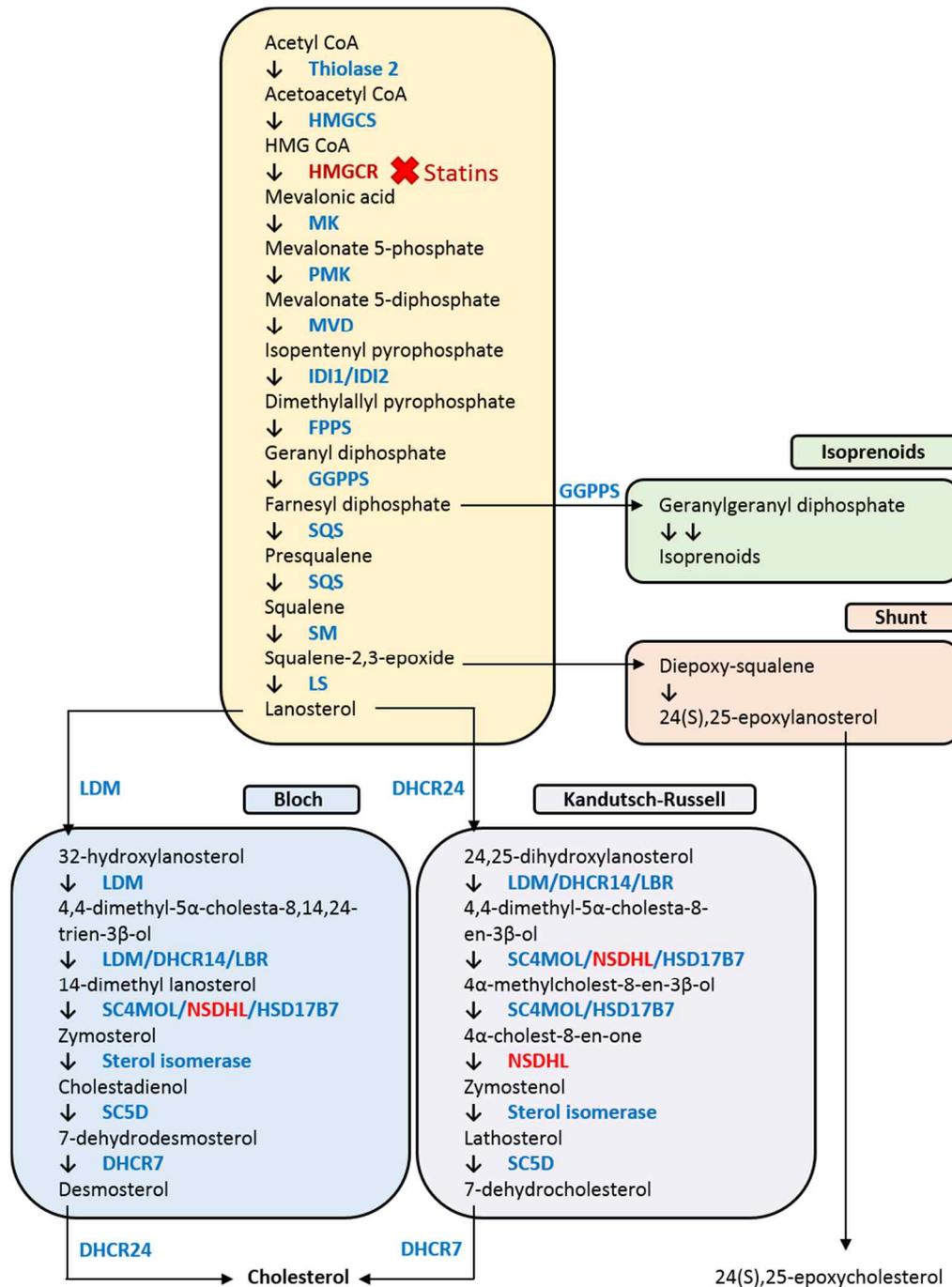


Figure 16. Cholesterol synthesis pathway. HMGCR is the target of the statins, important drugs for treating high cholesterol levels. C4-methylsterols are demethylated by NSDHL and its partners. MK, mevalonate kinase; PMK, phosphomevalonate kinase; MVD, diphosphomevalonate decarboxylase; FPPS, farnesyl-pyrophosphate synthase; SQS, squalene synthase; LDM, lanosterol 14 α -demethylase; SC5D, sterol C5-desaturase. (Edited from L.J. Sharpe, A.J. Brown, *J Biol Chem*, 288 (2013) 18707-18715)

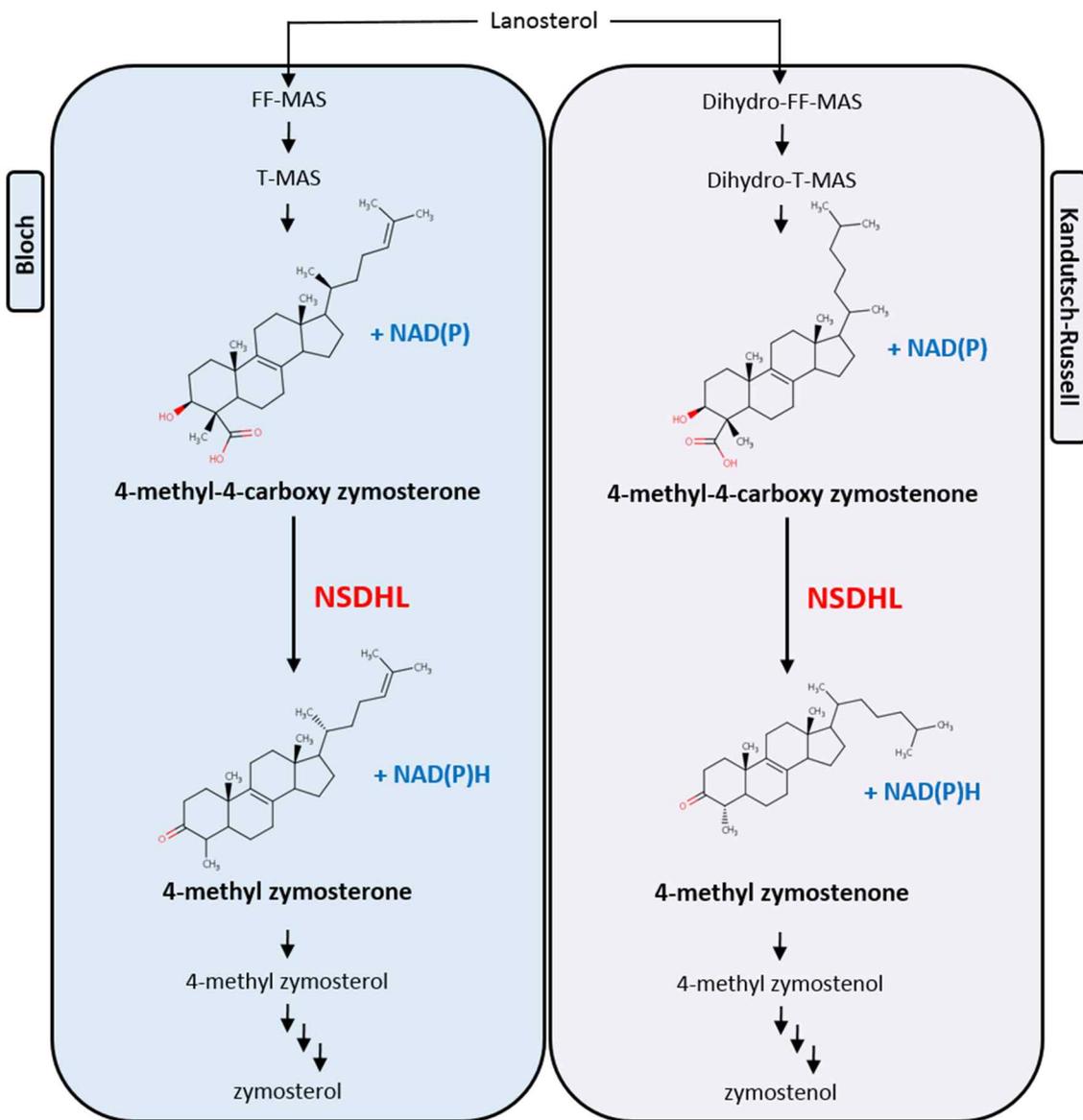


Figure 17. Sterol structures and reactions involved in C4-decarboxylation by NSDHL.

2.2 Experimental procedures

2.2.1 Cloning, expression, and protein purification

The gene encoding NSDHL from *Homo sapiens* was amplified by PCR using the clone provided from Korea Human Gene Bank (Medical Genomics Research Center, KRIBB, Korea) (Clone ID: hMU006554) as a template. The PCR primers used to clone the NSDHL expression plasmid, in which the restriction enzyme sites are underlined, were shown in Table 5. The PCR product and pET21a vector (Merck Millipore, Germany) were digested with *NdeI* and *XhoI* (NEB, UK) and ligated together. After confirming the sequence, the recombinant plasmid was overexpressed in expressing *E. coli* cells C41 (DE3) (Sigma Aldrich, USA) grown both in LB broth and M9 broth containing selenomethionine (SeMet) [Figure 19 (A)]. When the culture media reached an OD₆₀₀ of 0.5, 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG) was added to induce protein expression; cells were transferred to 15°C and grown for an additional 20 hours. The cells were collected by centrifugation at 4293 × *g* for 10 min, and the pellet was resuspended in lysis buffer (20 mM Tris, 500 mM NaCl, 5 mM imidazole, 0.5 mM TCEP, pH 7.8) and then sonicated at 4°C using a 15% duty cycle setting for 20 min (Cole Parmer Inc., USA). The lysate was centrifuged at 6708 × *g* for 60 min at 4°C, and the supernatant was loaded into an open Ni-NTA column (Qiagen, USA) 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 when the buffer reached 70 mM imidazole using an imidazole

gradient (from 30 to 500 mM) [Figure 19 (B)]. After concentrating the protein, the buffer was changed to 40 mM HEPES, 150 mM NaCl, 5% v/v glycerol and 0.5 mM TCEP at pH 8 by dialysis using an Amicon Ultra-15 Centrifugal Filter Unit with a 10k molecular weight cutoff (MWCO) (Millipore, USA). The protein was run on a gel filtration column (HiLoad 16/600 Superdex 75 pg, GE Healthcare sciences, Germany) to achieve good purity [Figure 19 (C)]. SDS-PAGE revealed purity up to 95%. The protein was concentrated to 0.26 mM for crystallization.

Mutagenesis primers (Table 6) targeting the four key residues related to CHILD and CKS were designed, and EZchange™ Site-Directed Mutagenesis Kit (Enzymomics, Korea) was used to generate point mutations according to the manufacturer's instructions. Briefly, the PCR conditions consisted of 25 repeated cycles including a 30-s melting step at 94°C, a 1-min annealing step at 55°C and a 5-min elongation step at 72°C. Template removal and ligation of the PCR products were performed using EZ-MIX buffer (Enzymomics, Korea). After transformation of the mutant plasmids into DH5 α -competent cells, the sequence of the plasmid DNA sequence was confirmed. C41 (DE3) was used for overexpression of the mutant proteins, and the successfully expressed mutants (NSDHL G205S and NSDHL K232 Δ) were purified.

Table 5. Primers used to clone the NSDHL expression plasmid.

Primer ID	Primer sequence (5'→3')	Restriction enzyme
nsdhl-1	GGGAATTCC <u>CATATG</u> ATGGAACCAGCAGTTAGCGAGC	<i>NdeI</i>
nsdhl-31	GGGAATTCC <u>CATATG</u> AAGGTTAACCAGAATCAGGCC	<i>NdeI</i>
nsdhl-216	CCGCCGCT <u>CGAGG</u> ATGAGGATGGGTACCAACTG	<i>XhoI</i>
nsdhl-254	CCGCCGCT <u>CGAGG</u> AGCTGCTCTGCCGCCA	<i>XhoI</i>
nsdhl-262	CCGCCGCT <u>CGAGC</u> CCACCCAGTGTCGAGTC	<i>XhoI</i>
nsdhl-267	CCGCCGCT <u>CGAGG</u> ATGTGAAATGCCTTCCAC	<i>XhoI</i>
nsdhl-272	CCGCCGCT <u>CGAGG</u> GGCTCATCATTGGTGATGTGAAA	<i>XhoI</i>
nsdhl-373	CCGCCGCT <u>CGAGC</u> TTGACCCTCCGCAGGTG	<i>XhoI</i>

Table 6. Primers for site-directed mutagenesis targeting the residues involved in causing CHILD and CKS. In primer IDs, “F” represents forward, and “R” represents reverse. Mutated nucleotides are underlined in the primer sequences used for site-directed mutagenesis.

Primer ID	Primer sequence (5'→3')	Mutated nucleotide
A105p-F	ACAGTTTTCCACTGT <u>GTG</u> TCACCCCCACCATCC	A105V
A105p-R	GGATGGTGGGGGTG <u>ACAC</u> ACAGTGGAAAACCTGT	A105V
A182p-F	ATCTTACAGGAGAGG <u>CCG</u> GTTCTGGGCGCCAAC	A182P
A182p-R	GTTGGCGCCCAGAAC <u>CGG</u> CCTCTCCTGTAAGAT	A182P
G205p-F	CCTCATGGCATTTC <u>AGC</u> CAAGGGACCCGCAG	G205S
G205p-R	CTGCGGGTCCCTTGGG <u>CT</u> GAAAATGCCATGAGG	G205S
K232p-F	TTCGTGATTGGAATGGGA <u>ACT</u> TGGTGGACTTCACC	K232Δ
K232p-R	GGTGAAGTCCACCAAGTCCCATTTCCAATCACGAA	K232Δ

2.2.2 Crystallization and preliminary structure determination

The most efficient crystallization conditions for NSDHL were selected using the Crystal Screen™, Index™ (Hampton Research, USA), Wizard™ classic (Rigaku Reagents, USA), Proplex™, MemGold™ (Molecular Dimensions, UK) crystal screening kits with the sitting drop vapor diffusion method. The protein solution and reservoir solution were mixed together in a 1:1 drop ratio. Crystals were obtained through drop equilibration with the reservoir solutions at 20°C for 7 days (Table 7). Prior to data collection, the crystals were flash-cooled with liquid nitrogen with protection by the addition of 20% glycerol to the cryo-solution.

The diffraction dataset for the NSDHL (31–267) crystals were collected at PAL (Pohang, Korea) using an ADSC quantum 315r CCD detector on beamline 5C-SBII at 100 K and a resolution of 2.2 ~ 2.4 Å. The dataset were processed and scaled with the HKL-2000 program package [34]. The NSDHL crystal belongs to the space group $P2_1$ (unit cell parameters: $a = 70.229$ Å, $b = 71.226$ Å, $c = 75.828$ Å, and $\beta = 92.178^\circ$). Determination of the structure of NSDHL (31–267) was tried by methods including molecular replacement (MR) and single wavelength anomalous dispersion (SAD). MR was performed using the PHENIX Phaser-MR program [36] based on the GOX2253 from *Gluconobacter oxydans* structure as a search model (PDB: 3WJ7) or SWISS-MODEL modeling structure obtained via NSDHL protein sequence [83]. The initial structures were refined by alternately using the Refmac [37] and Phenix.refine [36] programs. The solvent molecules were inserted with Coot [38].

Table 7. Conditions for crystallization of NSDHL (31–267) protein.

Precipitation reagent	Buffer	pH	Salt	Additive
25% w/v Polyethylene glycol 3,350	0.1 M Citric acid	3.5		
20% w/v Polyethylene glycol 3,350			0.2M Sodium thiocyanate	
45% v/v (+/-)-2-Methyl-2,4-pentanediol	0.1M Bis-tris	5.5	0.2M Calcium chloride dihydrate	
30% v/v Pentaerythritol ethoxylate (15/4 EO/OH)	0.05M Bis-tris	6.5	0.05M Ammonium sulfate	
10% w/v Polyethylene glycol 8,000	0.1M Imidazole	8.0	0.2M Calcium acetate	
2.0M Ammonium sulfate (dibasic)	0.1M Tris	7.0	0.2M Lithium sulfate	
15% w/v Polyethylene glycol 10,000	0.1M Sodium citrate	5.5	2% v/v Dioxane	
15% w/v Polyethylene glycol 20,000	0.1 M HEPES	7.0		
15% v/v Ethanol	0.1M Tris	8.0	0.1M Sodium chloride	5% v/v MPD
50% v/v Polyethylene glycol 400	0.1M Sodium acetate	4.5	0.2M Lithium sulfate	
25% w/v Polyethylene glycol 1,500	0.1M PCB	7.0		
44% v/v Polyethylene glycol 400	0.1M Tris	8.0	0.2M Calcium chloride dihydrate	
7% w/v Polyethylene glycol 3,000	0.1M Tricine	7.4	0.05M Lithium sulfate	

2.2.3 Differential scanning fluorimetry

NSDHL (31–267) wild-type, G205S, K232Δ were diluted to 5 μM in a white 96-well plate using the identical buffer as mentioned above. A 5000X solution of the dye SYPRO Orange (Sigma Aldrich, USA) was added to each well to achieve a final concentration of 5X. Measurements were performed in 100 μL. 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 (ThermoFisher Scientific, USA). The raw fluorescence data were plotted as a function of temperature, which generated a sigmoidal curve. The inflection point of the transition curve (T_m , midpoint melting temperature) was calculated using Prism 5 (GraphPad software, USA) by applying the Boltzmann sigmoidal fitting.

2.2.4 Isothermal titration calorimetry

A 0.1 mM solution of NSDHL was prepared as described above. Coenzymes [NAD⁺, NADH, NADP⁺, and NADPH (Santa Cruz Biotechnology, USA)] were dissolved to a concentration of 1 mM in the same buffer as the protein. The protein and coenzyme solutions were degassed in a vacuum before measurement. All ITC experiments were performed using a MicroCal iTC200 microcalorimeter (GE healthcare, UK) at 25°C. The protein was added to the sample cell, and a coenzyme solution was charged in the injection syringe. During the titration, 20 aliquots of metal solution were injected into the sample chamber. The titrations began with a 60 s delay time and 0.4 μL injection volume, followed by a 2 μL injection-volume with 5 s delay times. Intervals of 150 s between injections were included at the end of each titration. The stirring speed in the sample chamber at 25 ± 0.1°C was 1,000 rpm. Heat generation during the titration experiments was measured using the integrated data from the ITC calorimeter with the Origin 7.0 software package supplied by MicroCal, subtracting the heat generated by the buffer. K_a (binding constant), ΔH (enthalpy change) and N (stoichiometry) values were calculated by applying the one-site fitting model.

2.2.5 NMR spectroscopy

NMR experiments were performed at 298 K using a Bruker Ascend™ 800 MHz spectrometer equipped with a 5-mm cryogenic CP-TCI z-gradient probe and SampleJet. The NMR sample was prepared as a mixture of 10 μ M NSDHL and 0.5 mM ligand, which were dissolved in 99.9% D₂O (Sigma Aldrich, USA). On- and off-resonance irradiations were applied at chemical shifts of 0 and -30 ppm, respectively. NSDHL was saturated using a train of Gaussian-shaped 50-ms-long pulses. The total length of the saturation train was set to 2 s. Prior to acquisition, a 15 ms spin-lock pulse with a 1-W strength (T_2 filter) was applied to remove protein signals from the STD spectrum. The ¹H NMR spectrum was acquired with 32K real points and 512 scans. Single pseudo-2D data from two serial free-induction decays were divided into two separate 1D data (on- and off-resonance). To increase the signal-to-noise ratio, the 1D raw data were processed with a 0.5-Hz line broadening and exponential window function prior to Fourier transformation. For comparison, the STD effects of individual peaks were quantified by the simple equation $(I_{off} - I_{on})/I_{off}$, where I_{on} and I_{off} represent the absolute intensities of the on- and off-resonance spectra peaks, respectively. The off-resonance spectrum is basically identical to a conventional ¹H NMR spectrum. The largest STD value was set to 100%, and the other STD values were normalized to the largest value [43, 44]. For STD-NMR competition experiments, [NAD⁺ and NADH] or [NADP⁺ and NADPH] were added to the NMR samples, which contained NADHL. The same acquisition and processing parameters mentioned above were used to obtain the STD NMR spectra.

The peaks of the 1D ¹H NMR spectra were assigned by referencing spectra in

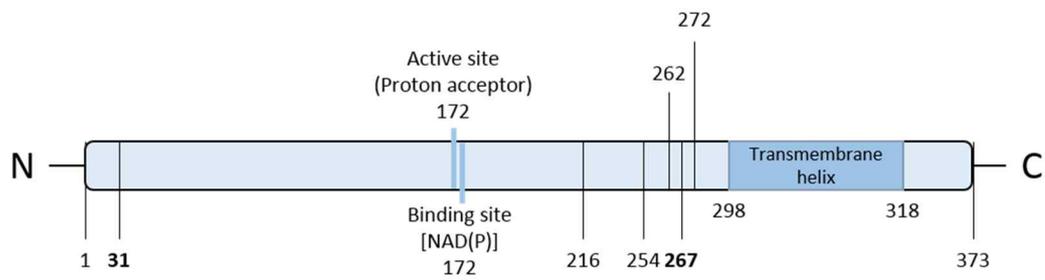
Biological Magnetic Resonance Data Bank (BMRD) [84] for the assignment of peaks and analysis of STD. All STD spectra were analyzed in a spectral range of approximately 3.7 ppm to 10 ppm, in which peaks of the ligands appear.

2.3 Results

2.3.1 Protein expression and structure determination

To produce soluble and homogeneous NSDHL, twelve truncated constructs have been designed (Figure 18) with reference to structure prediction results from Intrinsically Unstructured/disordered Proteins Prediction server (IUPred) [85] and Protein Secondary Structure Prediction server (JPred) [86]. Among them, the NSDHL (31–267) construct with the pET-21a vector (Merck, Millipore, Germany) was highly soluble and stable.

From the purified construct, several crystals were successfully made using the commercial crystallization solutions (Figure 20) and the diffraction data of both native NSDHL and SeMet-derivatized NSDHL were collected. However, structural determination of NSDHL was not successful, due to the weak anomalous signals of SeMet data and difficulty in indexing the data. When performed the molecular replacement (MR), a homolog structure (PDB: 3WJ7) or a structural model from SWISS-MODEL [87] (see 2.3.2 *Homology modeling of the 3D structure of NSDHL*) were used as an initial search model in each trial [Figure 21 (A)]. A better MR result was obtained from the model by SWISS-MODEL with the good agreement to the electron density difference map of NSDHL crystals [Figure 21 (B)]. However, it was hard to find another structural pair using the previously found one in the $P2_1$ space group, which did not allow to refine the crystal structure further.



Construct	Expression	Solubility	Construct	Expression	Solubility
1-373	-	-	1-262	++	-
31-373	-	-	31-262	++	-
1-272	++	-	1-254	++	-
31-272	++	-	31-254	++	-
1-267	+++	++	1-216	++	-
31-267	+++	+++	31-216	++	-

Figure 18. NSDHL constructs. Based on the results from IUPred and Jpred, 12 constructs was designed and test. Among them 1–267 and 31–267 constructs were successfully expressed, purified and used for further experiments.

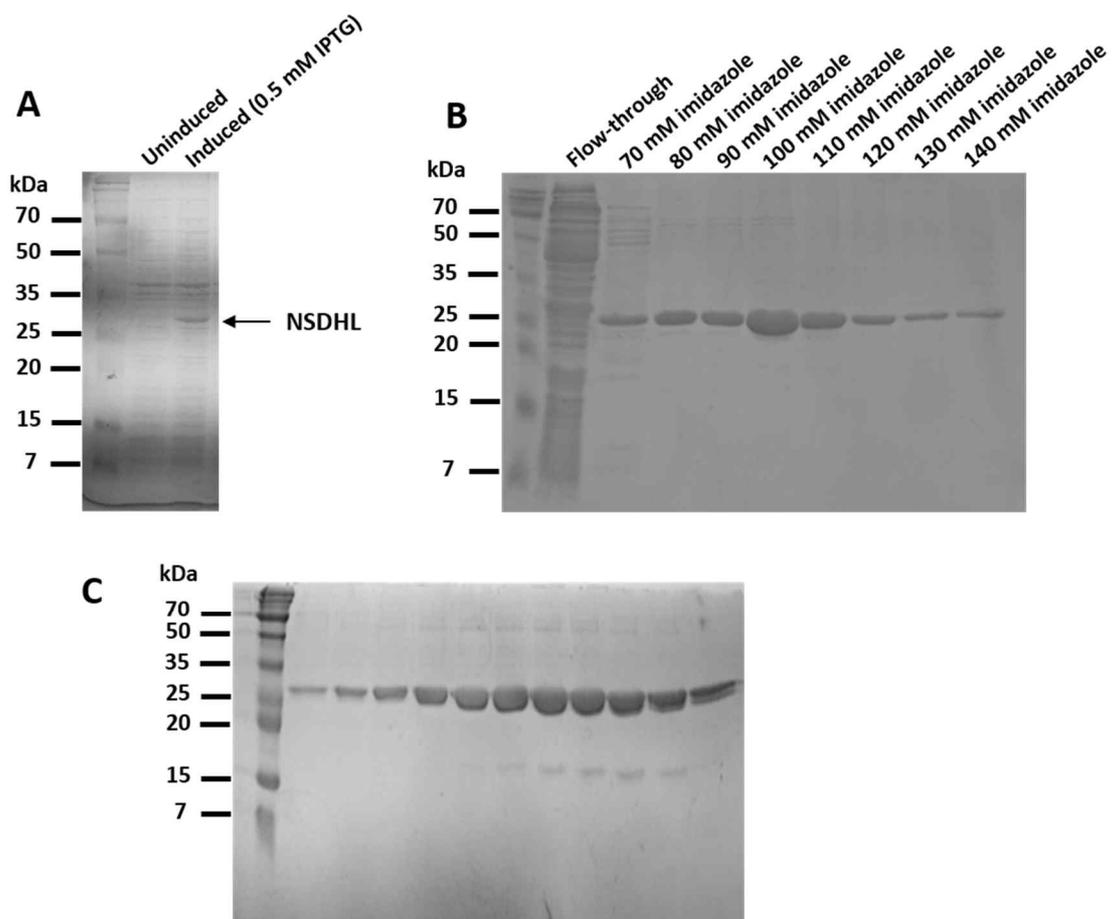


Figure 19. Expression and purification result of NSDHL. (A) Coomassie-stained SDS-PAGE gel of the overexpression. (B) Purification by immobilized metal affinity chromatography (IMAC). (C) Purification by size-exclusion chromatography.

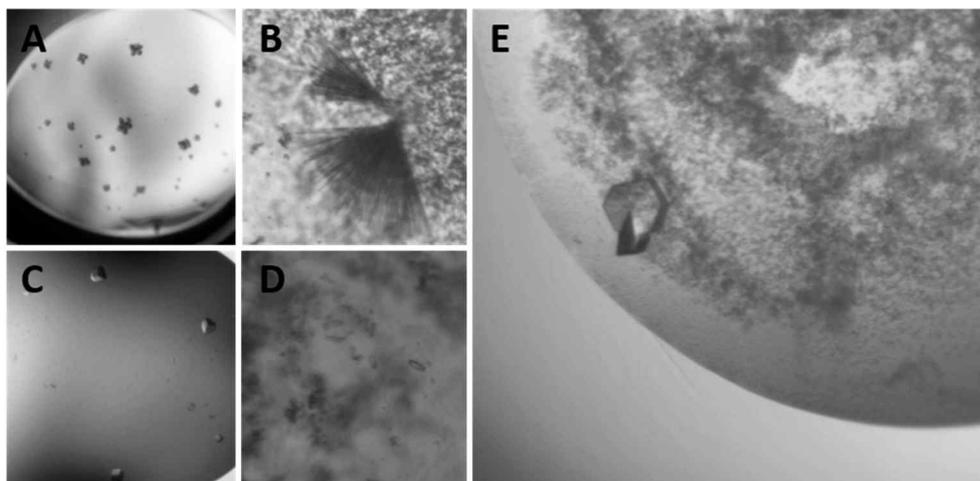


Figure 20. NSDHL (31–267) crystals. Each crystallization solution is consist of (A) 30% v/v pentaerythritol ethoxylate (15/4 EO/OH), 0.05 M ammonium sulfate, 0.05 M bis-tris at pH 6.5 (B) 15% w/v polyethylene glycol 20,000, 0.1 M HEPES at pH 7, (C) 20% w/v Polyethylene glycol 3,350, 0.2 M potassium nitrate, (D) 45% v/v (+/-)-2-methyl-2,4-pentadiol, 0.2 M calcium chloride dehydrate, 0.1 M bis-tris at pH 6.5, and (E) 25% w/v polyethylene glycol 1,500, 0.1 M PCB at pH7.

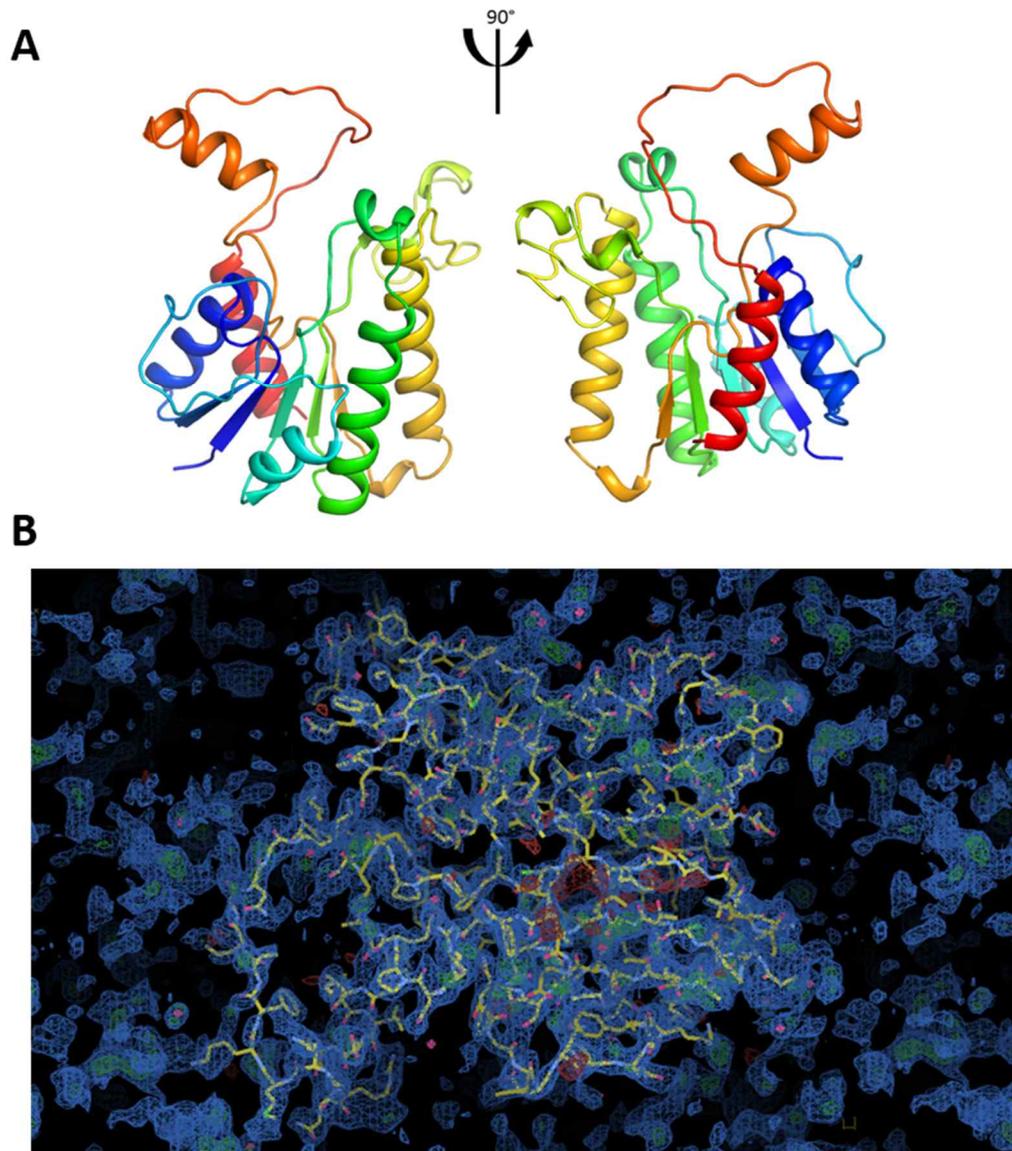


Figure 21. Three-dimensional structure of NSDHL derived from (A) homology modeling method using SWISS-MODEL and (B) molecular replacement approach. The model was further refined using X-ray crystallographic data of the NSDHL (31–267) crystal. Building more residues was enabled by the *F_o-F_c* map.

2.3.2 Homology modeling of NSDHL

Due to the difficulty of phasing the X-ray diffraction data, homology modeling has been performed to reveal the structure of NSDHL. For the purpose, SWISS-MODEL server was applied to build more realistic models of NSDHL, using 31–267 residues of NSDHL. The initial calculation filtered 50 templates from total 3849 templates. Among them, three models were successfully built (Table 8). The sequences of NSDHL (31–267) and the template proteins used in homology modeling were aligned using ClustalW (Figure 22).

The models of NSDHL are homo-dimers and consist of six α -helices and parallel five β -strands which forms a sheet [Figure 23 (A)]. They reveals a modified Rossmann fold, which are generally found in nucleotides binding protein. By comparing NSDHL model to template structures (PDB: 1Z45 and 4ZRN) [88, 89], the putative active site and coenzyme binding site could be suggested. Key amino acids for coenzyme binding are identified by secondary structure alignment with the templates [Figure 23 (B)]: Phe48, Leu49, Asp68, Gln70, Asp84, Cys104, Val120, Tyr172, and Lys176 in NSDHL. The residues whose side chains are involved in coenzyme binding are well conserved.

The substrate-binding site of NSDHL was predicted by comparing the two modeling templates to NSDHL model. In both templates, UDP-glucose binds to a cavity which is constituted by the 5th (S5), 6th (S6), and 7th (S7) β -strands and the random coil [between 6th (H6) α -helix and S7 β -strand] of the Rossmann fold. The corresponding region is the most plausible substrate binding site in NSDHL [Figure 23 (C)]: Pro108–Lys114, Ser145–Ser147, Gly205–Gln210, Lys223–Ile228, and Lys232–Phe237.

Table 8. Results of NSDHL homology modeling.

No.	Name	Title	Coverage	Found by	Sequence identity (%)	Sequence similarity (%)	Oligo state	*GMQE
1	1Z45.1.A	GAL 10 bifunctional protein	7-224	BLAST	29.2	36	homo-dimer	0.61
2	1Z45.1.A	GAL 10 bifunctional protein	7-237	HHBlits	24.5	33	homo-dimer	0.64
3	4ZRN.1.A	UDP-glucose 4-epimerase	8-237	HHBlits	26.7	33	homo-dimer	0.64

*GMQE (Global model quality estimate) is scored in a number between zero and one, reflecting the expected accuracy of a model built with that alignment and template. Higher numbers indicate higher reliability.

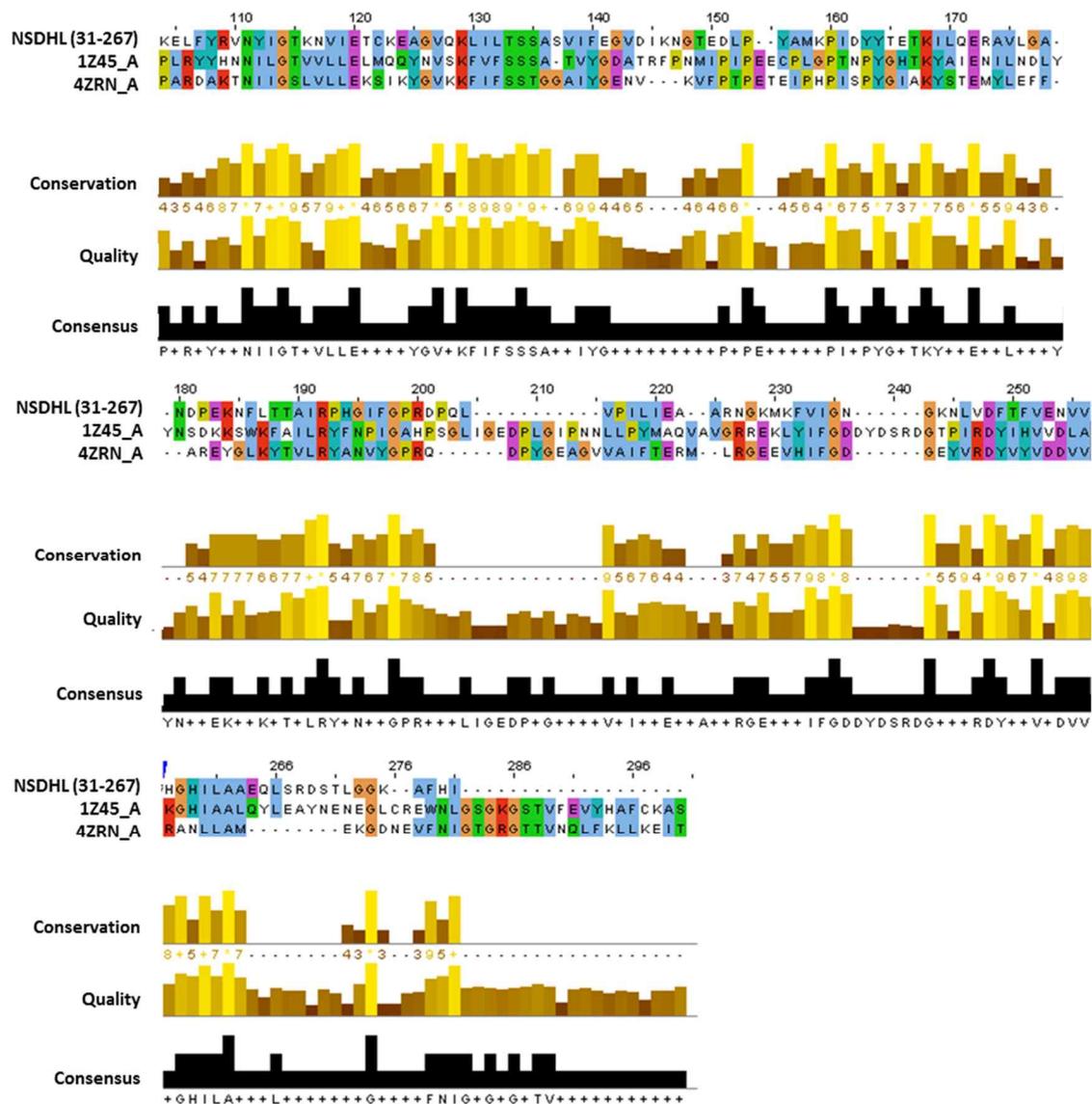


Figure 22. Sequence alignment of NSDHL (31–267) and the template proteins of homology modeling (PDB: 1Z45 and 4ZRN). The sequences were aligned using ClustalW. Conserved positions in the alignment are colored. Plots indicate the level of conservation, quality and consensus.

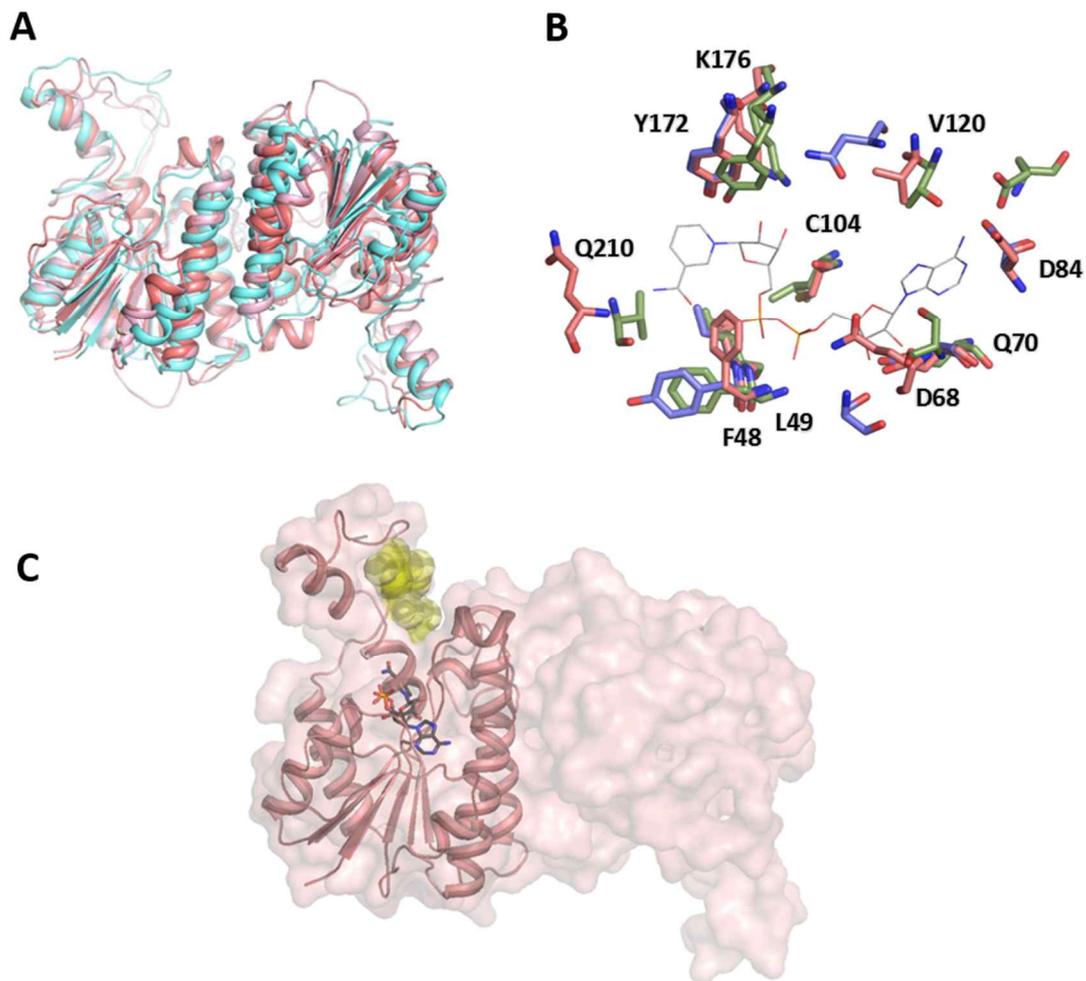


Figure 23. Homology models of NSDHL. (A) Three models which are obtained from SWISS-MODEL are superimposed. (B) Key residues (stick, *salmon pink*) in the NSDHL for binding to NAD⁺ are identified by alignment with the templates [PDB: 1Z45 (*blue*) and 4ZRN (*green*)]. NAD⁺ is shown as *gray* lines. (C) The putative coenzyme and substrate binding sites are determined by comparing to the template structures. NSDHL is shown as cartoon and surface diagrams, *yellow* balls indicates the substrate binding region, and NAD⁺ is shown as sticks in *gray*.

2.3.3 Thermal stability of wild-type NSDHL and the mutants

The effects of site-directed mutagenesis to thermal stability were measured by DSF experiments. The DSF experiments were performed as described in the Experimental procedure section. In the case of NSDHL wild-type, T_m was up to 49.5°C. In contrast, T_m was observed to shift to 39.5°C for NSDHL G205S mutant and was 36.5°C for the NSDHL K232Δ mutant (Figure 24). These results supported the hypothesis that the mutant which is found in patients suffering from CHILD syndrome (G205S) and CKS (K232Δ) affect the protein stability and folding.

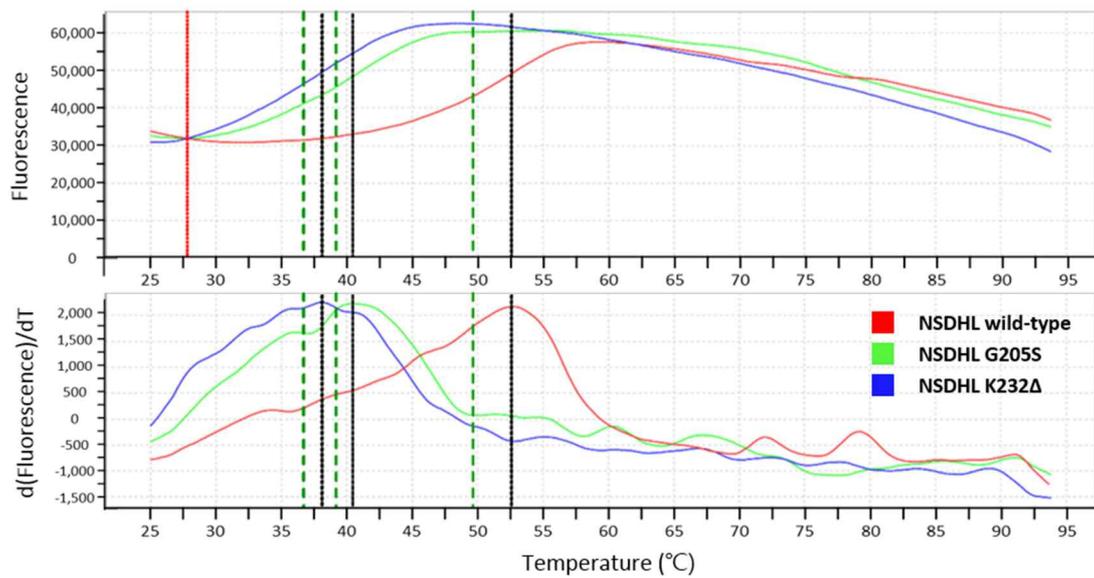


Figure 24. Differential Scanning Fluorimetry of NSDHL. Thermal denaturation curves obtained for NSDHL wild-type, G205S, and K232Δ. The red curve is wild-type NSDHL, melting temperature (T_m) = 49.5°C; green curve is NSDHL G205S, T_m = 39.0°C; and blue curve is NSDHL K232Δ, T_m = 36.5°C.

2.3.4 Isothermal titration calorimetry

ITC has been used to investigate the thermodynamics of NSDHL-coenzyme interactions. The ITC data with NAD⁺ or NADH injection have been well fitted to the single-site binding in the isotherm model [Figures 25 (A) and (B)]. However, NADP⁺ or NADPH titration with NSDHL have generated data that could not be fitted to the single-site binding isotherm model. Indeed, NSDHL has NAD⁺ binding affinity (K_d) is 151.51 ± 70.21 μM , with a heat change (ΔH) of -9287 ± 1400 cal/mol, NADH K_d is 21.38 ± 1.1 μM , with a ΔH of -14320 ± 547.4 cal/mol.

Interactions between NSDHL mutants and coenzymes also have been investigated by ITC. NAD⁺ or NADH have been titrated to each ITC sample cell containing NSDHL G205S and K232 Δ which are related to CHILD and CKS, respectively (Figure 26). All the ITC data with NAD⁺ or NADH injection into the mutants have been well fitted to the single-site binding isotherm model. Binding affinity of the mutants have been calculated: G205S has NAD⁺ K_d is 43.82 ± 12.10 μM , and NADH K_d is 30.25 ± 1.20 μM ; K232 Δ has NAD⁺ K_d is 100.68 ± 43.41 μM , and NADH K_d is 56.86 ± 6.23 μM .

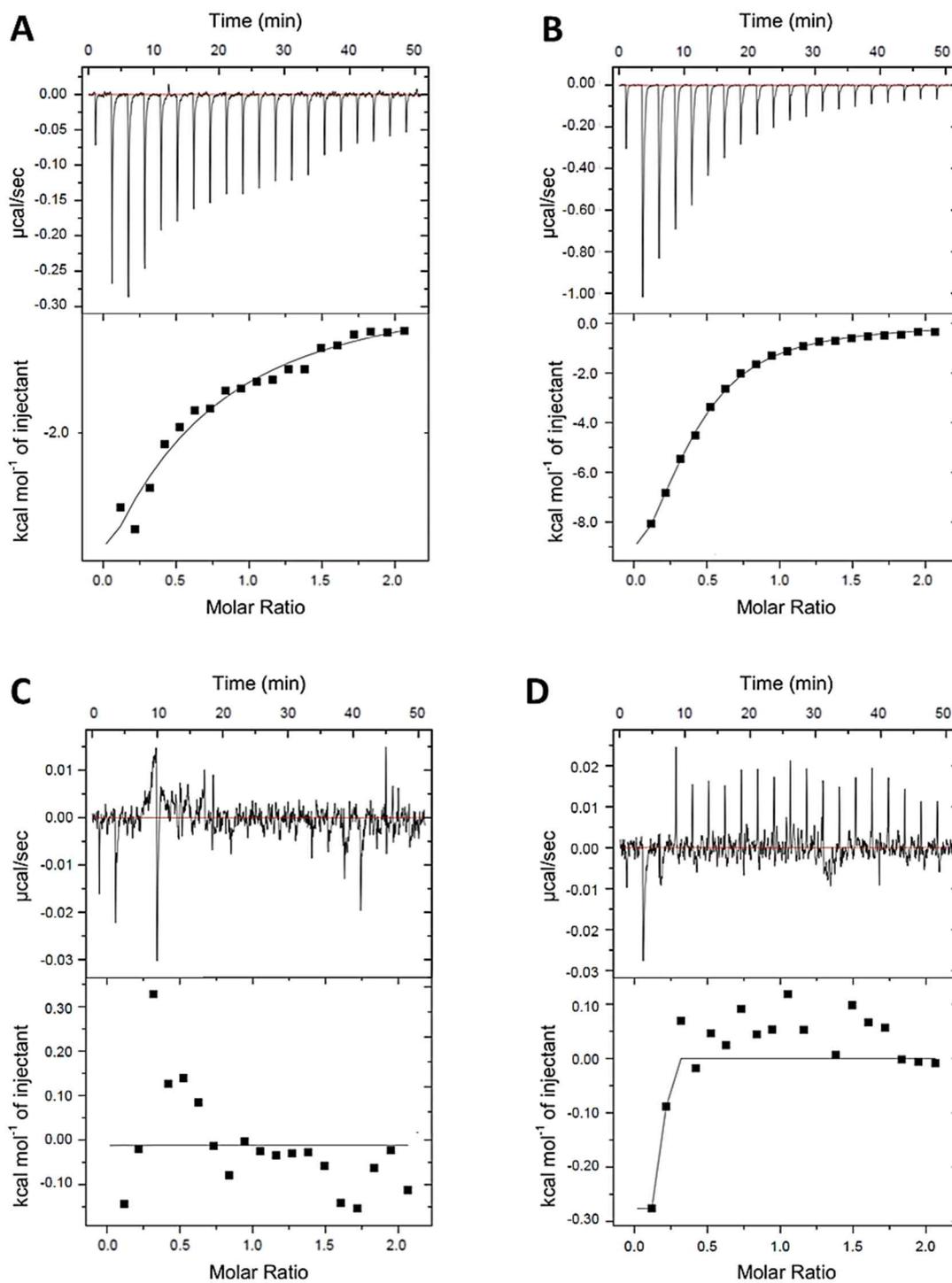


Figure 25. Calorimetric data for the titration of 1 mM of the coenzymes to NSDHL wild-type: (A) NAD^+ , (B) NADH , (C) NADP^+ and (D) NADPH into 0.1 mM of NSDHL at 25°C.

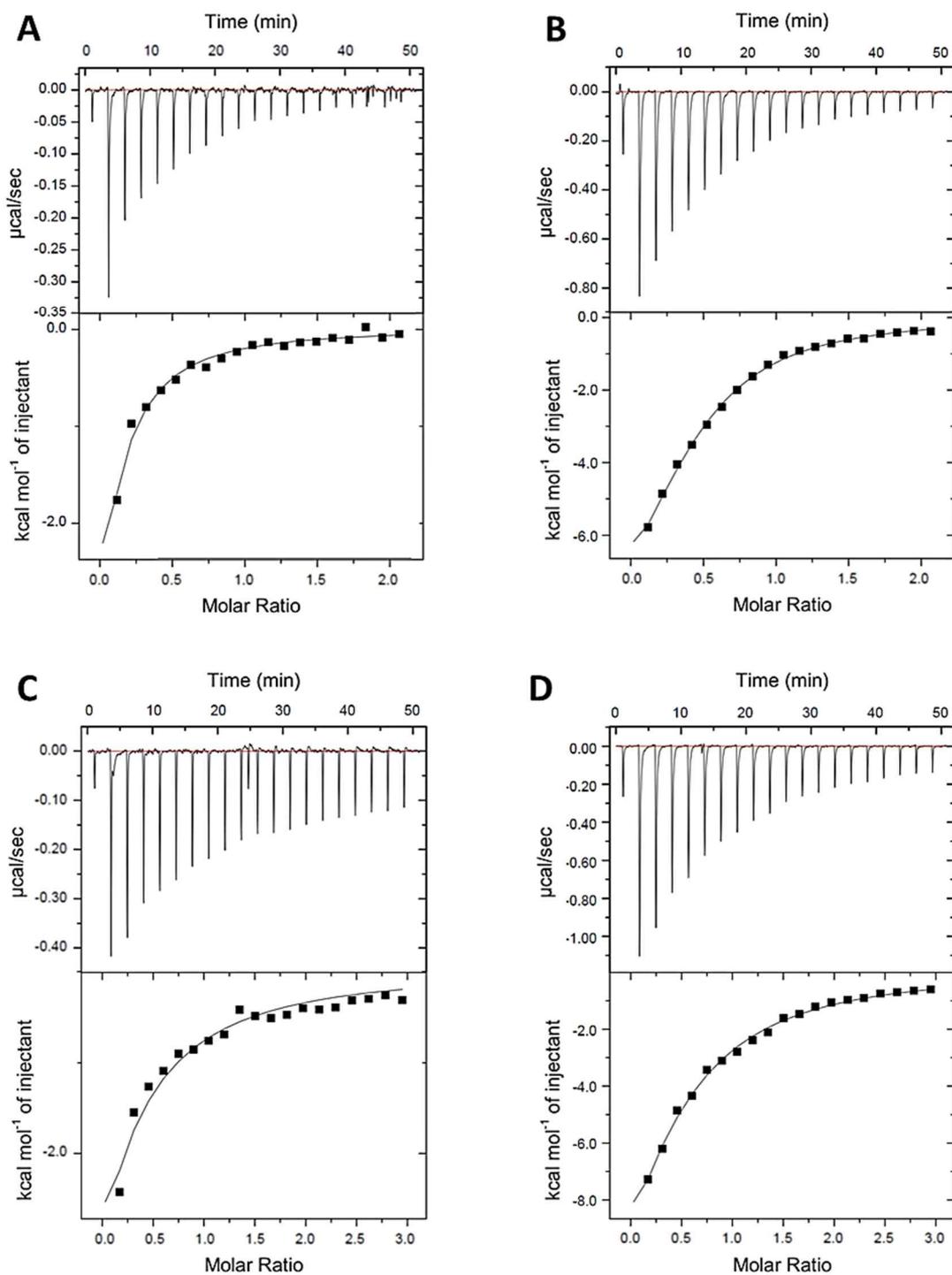


Figure 26. Calorimetric data for the titration of 1 mM of the coenzymes to NSDHL mutants: (A) NAD^+ , (B) NADH into 0.1mM of NSDHL G205S mutant, and (C) NAD^+ , (D) NADH into 0.1 mM of NSDHL K232 Δ at 25°C.

2.3.5 STD-NMR analysis of interaction modes between NSDHL and its coenzymes

STD NMR spectroscopy was used to examine interaction between NSDHL and its coenzymes by quantifying reductions in their NMR signals, which are caused by STD effects, the transfer of nuclear Overhauser effect (NOE) from NSDHL to its coenzyme [42]. In STD-NMR experiments, 1D ^1H NMR peaks of on-resonance spectrum for individual ligands were recorded. The signals could be reduced by the magnetization transfer from NSDHL to ligands, compared to those of off-resonance spectrum [Figures 27 (B), 28 (B), 30 (B), and 31 (B)] that is basically identical to a standard 1D ^1H NMR spectrum for ligands [52]. The STD spectrum represents the difference between signal intensities of the on- and off-resonance spectra, the latter is shown in Figures 27 (D), 28 (D), 30 (D), and 31 (D). The STD spectra could quantitatively reflect contributions of ligand protons in the binding to the NSDHL. Figures 27 (D), 28 (D), 30 (D), and 31 (D) show reference STD spectra for each compound in the absence of NSDHL.

NAD^+ (Figure 27), NADH (Figure 28), NADP^+ (Figure 30) and NADPH (Figure 31), produced STD spectra that indicated interactions with NSDHL. All of the STD spectra have been analyzed in spectra range from 3.7 to 10 ppm. The nine protons of NAD^+ exhibited STD signals, among which proton-5 showed the largest reduction in signal ratio between the on- and off-resonance spectra [Figure 27 (D)]. For comparison, STD effects of the other protons were normalized to that of proton-5, as shown in Figure 27 (D). Proton-7 exhibited the second-largest STD signal, which represented 84 percent of that of proton-5. For NADH, proton-2 exhibited the largest STD effect, and the STD

effects of protons-**1**, -**3**, -**4**, -**5**, -**7** and -**8** were quantified and normalized to that of proton-**2**, demonstrating relatively large differences of STD values between proton-**2** and others [Figure 28 (D)].

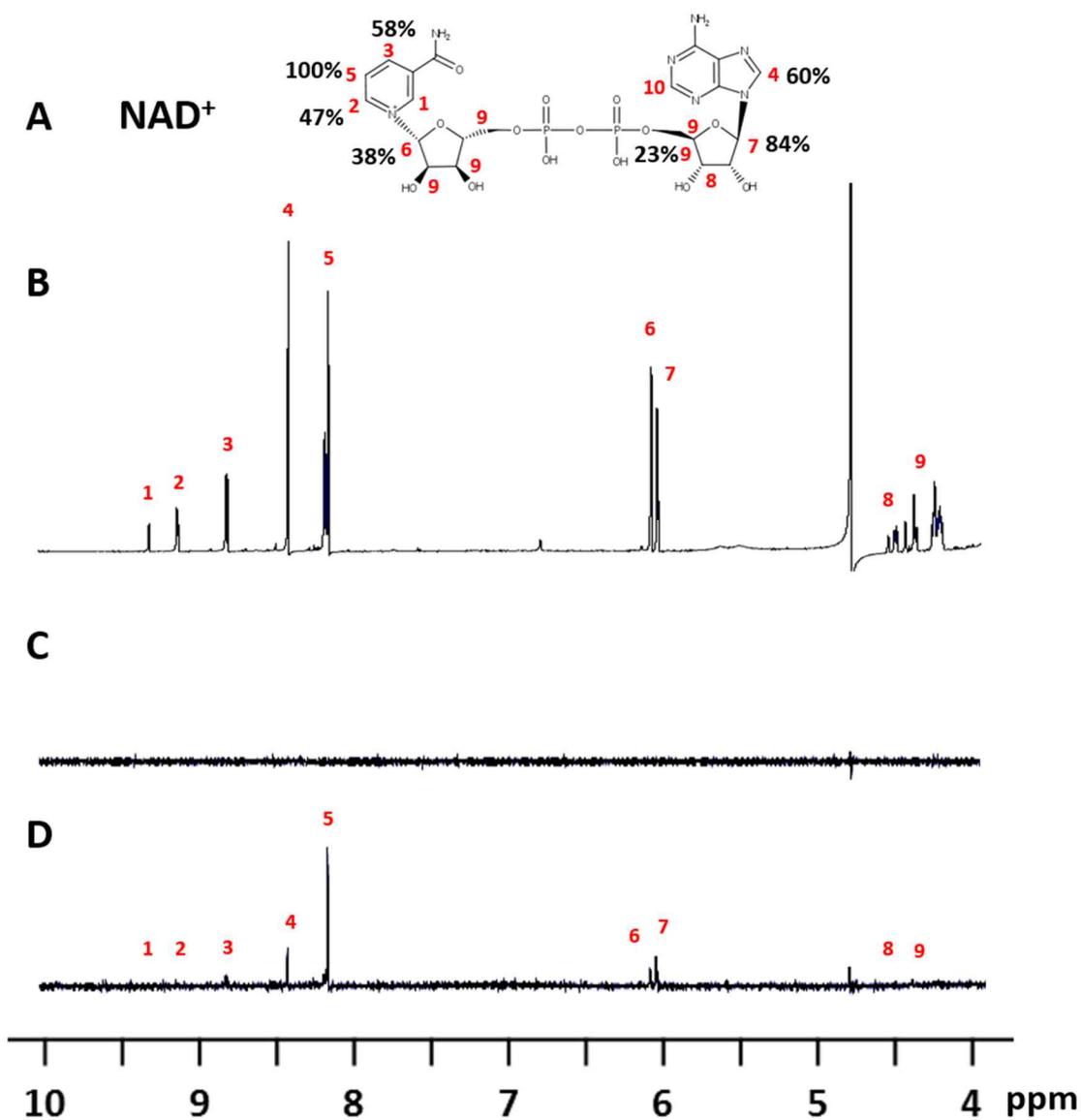


Figure 27. ¹H STD spectra (3.7 ppm ~ 10 ppm) of NAD⁺. (A) Chemical structure of NAD⁺ (each proton is numbered in *red*). Proton-10 did not appear in the spectrum at the low NAD⁺ concentration. The normalized STD values are presented as percentages. (B) Reference ¹H NMR spectrum of NAD⁺. (C) Reference STD spectrum of NAD⁺ in the absence of NSDHL. (D) STD spectrum of NAD⁺ in the presence of NSDHL.

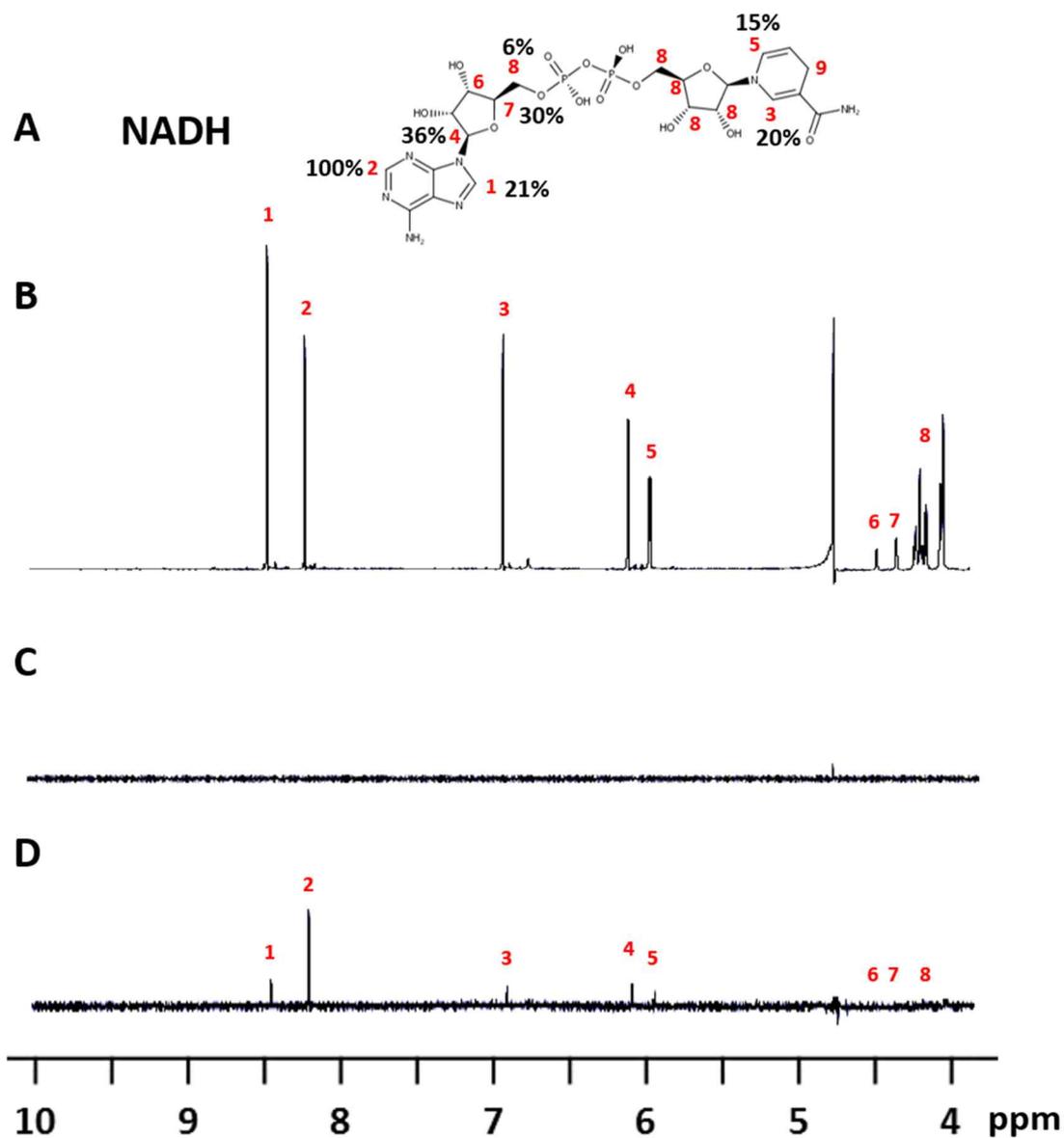


Figure 28. ^1H STD spectra (3.7 ppm ~ 10 ppm) of NADH. (A) Chemical structure of NADH (each proton is numbered in *red*). Proton-9 appeared in a spectral range of 2.5 ppm to 3.0 ppm (not shown). The normalized STD values are presented as percentages. (B) Reference ^1H NMR spectrum of NADH. (C) Reference STD spectrum of NADH in the absence of NSDHL. (D) STD spectrum of NADH in the presence of NSDHL.

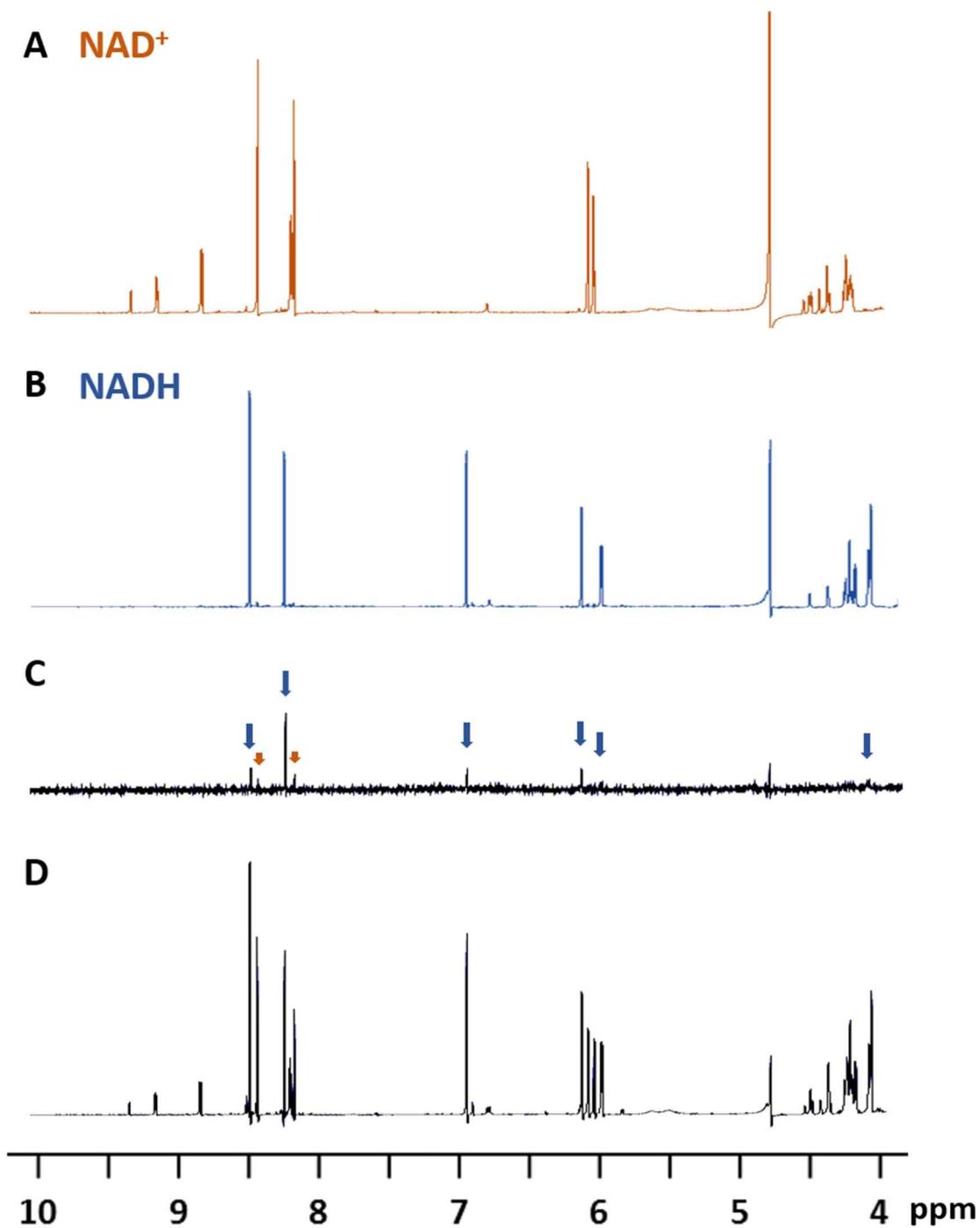


Figure 29. ^1H STD competition NMR spectrum of NAD^+ and NADH . (A) Reference ^1H NMR spectrum of NAD^+ (*orange*). (B) Reference ^1H NMR spectrum of NADH (*blue*). (C) STD spectrum of NAD^+ and NADH mixed solution in the presence of NSDHL. NADH shows stronger STD effect than NAD^+ . (D) Reference ^1H NMR spectrum of NAD^+ and NADH mixed solution.

Similar STD-NMR experiments were carried out for NADP⁺ and NADPH. In STD spectrum of NADP⁺, proton-6 indicated the largest STD effect, and the STD effects of protons -4, -7, and -8 were normalized to that of proton-6, as shown in Figure 30 (D). NADPH STD experiment showed that proton-2 had the largest STD effect, the levels of STD signals of the other protons were normalized to that of proton-2, as shown in Figure 31 (D). Be taken altogether, the reduced form of coenzymes (NADH or NADPH) demonstrated relatively larger STD effect than that of the corresponding oxidized ones. It is consistent with the ITC results.

To confirm the preference of NSDHL for binding to NADH or NADPH, the competition experiments were carried out. STD competition experiments were conducted adding both NAD⁺ and NADH [Figure 29 (C)] or NADP⁺ and NADPH [Figure 32 (C)]. The NAD⁺-NADH competition STD data showed that the STD peak intensities of NAD⁺ in the presence of NADH were significantly decreased when they were compared to the NAD⁺ STD spectra [Figures 27 (D) and 29 (C)]. The STD peak intensities of NADH were less affected by NAD⁺ [Figures 29 (B) and (C)]. These results indicated that NADH more competent for NSDHL binding than NAD⁺. For NADP⁺-NADPH competition STD data, higher intensities shown in STD spectra of NADPH indicated that NADPH more strongly binds to NSDHL [Figure 32 (C)]. Comparing of two STD competition spectra [Figures 33 (A) and (B)] gave the information that NSDHL has the relatively strong preference for binding to NAD(H) than NADP(H).

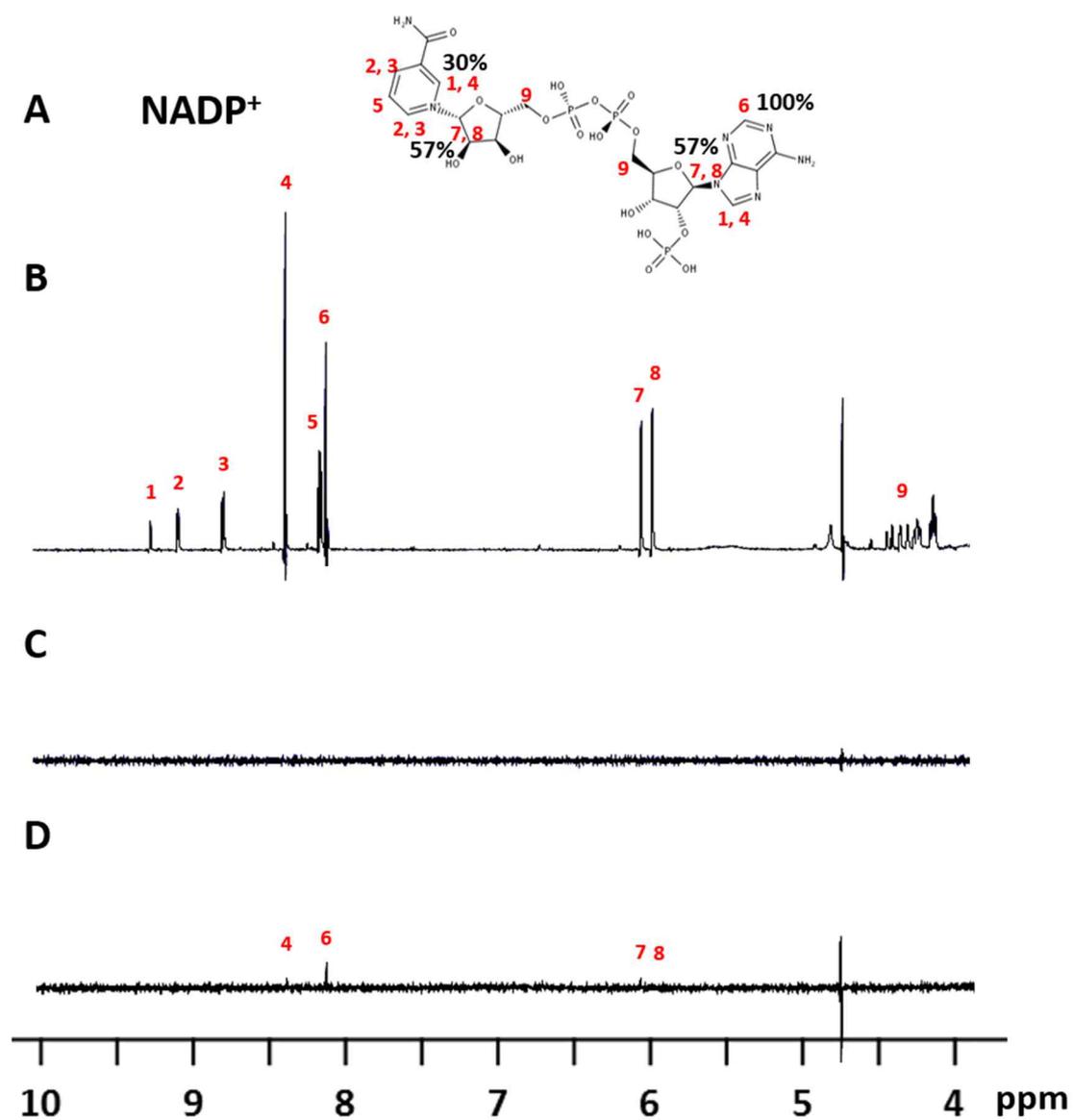


Figure 30. ^1H STD spectra (3.7 ~ 10 ppm) of NADP^+ . (A) Chemical structure of NADP^+ (each proton is numbered in red). The normalized STD values are presented as percentages. (B) Reference ^1H NMR spectrum of NADP^+ . (C) Reference STD spectrum of NADP^+ in the absence of NSDHL. (D) STD spectrum of NADP^+ in the presence of NSDHL.

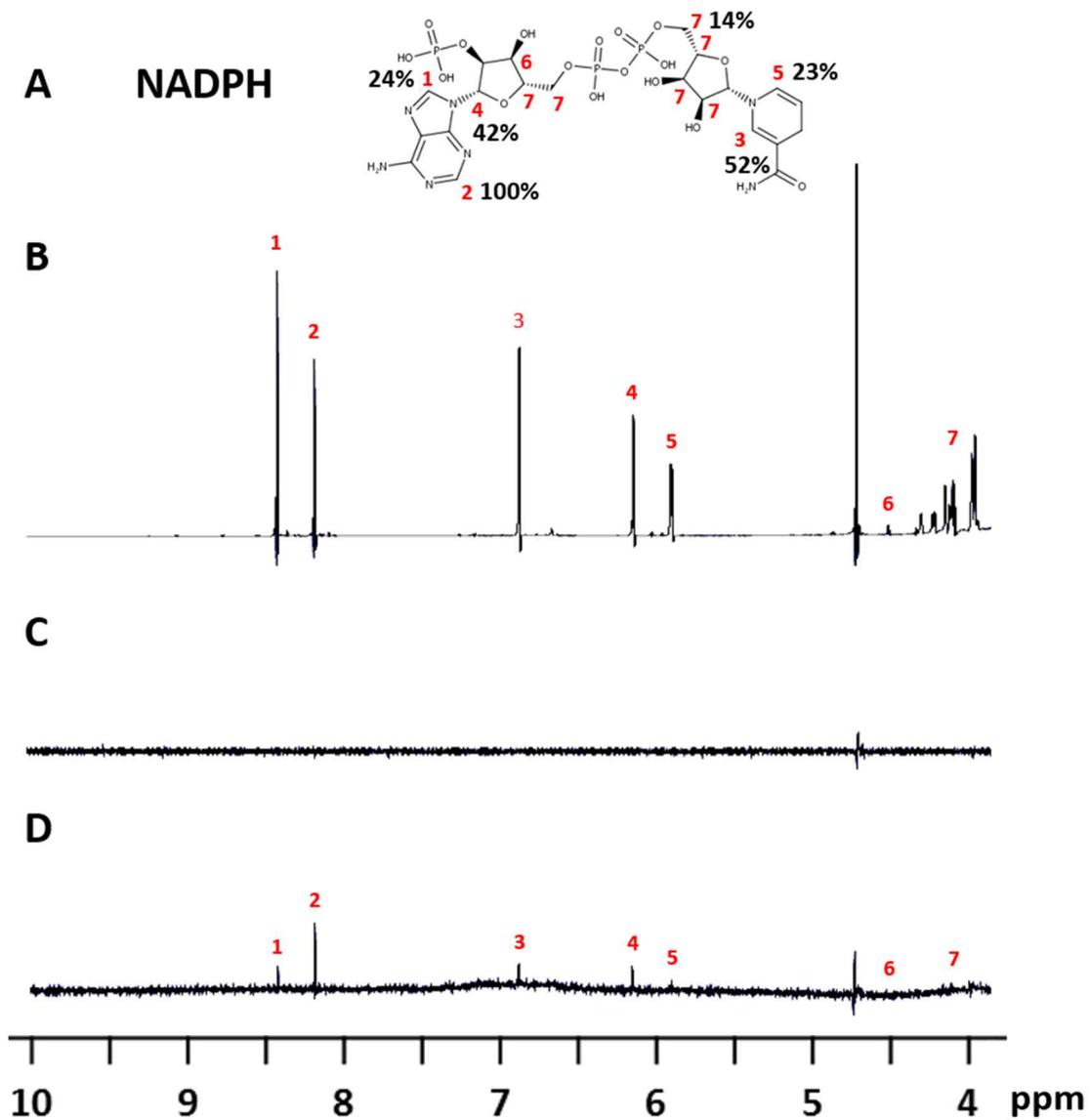


Figure 31. ^1H STD spectra (3.7 ~ 10 ppm) of NADPH. (A) Chemical structure of NADPH (each proton is numbered in *red*). The normalized STD values are presented as percentages. (B) Reference ^1H NMR spectrum of NADPH. (C) Reference STD spectrum of NADPH in the absence of NSDHL. (D) STD spectrum of NADPH in the presence of NSDHL.

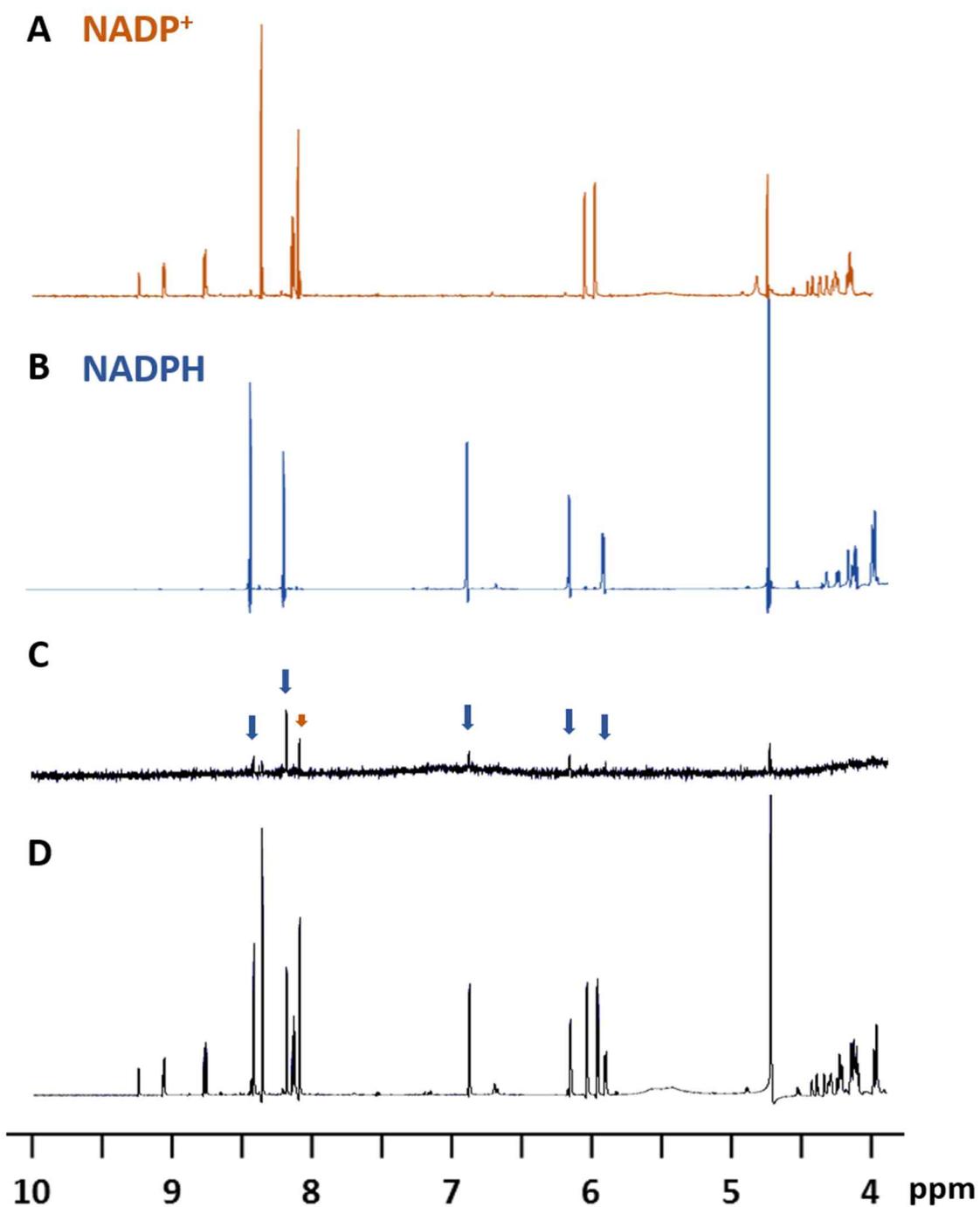


Figure 32. ^1H STD competition NMR spectrum of NADP⁺ and NADPH. (A) Reference ^1H NMR spectrum of NADP⁺ (*orange*). (B) Reference ^1H NMR spectrum of NADPH (*blue*). (C) STD spectrum of NADP⁺ and NADPH mixed solution in the presence of NSDHL. NADPH shows stronger STD effect than NADP⁺. (D) Reference ^1H NMR spectrum of NADP⁺ and NADPH mixed solution.

To better understand better the relationship of coenzyme binding modes and disorder-related NSDHL mutants, we performed STD experiments of NAD⁺ or NADH with two mutants (G205S and K232Δ NSDHL). Proton-5 in the NAD⁺ reference ¹H NMR spectrum showed the largest STD effect in NSDHL wild-type STD result, whereas proton-7 had the largest value in the G205S mutant. Although K232Δ mutant presented the largest STD value in proton-5 as the wild-type, the STD effect of proton-3 and proton-4 became smaller than in NSDHL wild-type (Figure 34).

In all the NADH STD spectra with wild-type NSDHL or the mutants, proton-2 exhibited the largest STD effect. However, the mutants' proton-7 had the second largest values differently from the wild-type where proton-4 was the second (Figure 35).

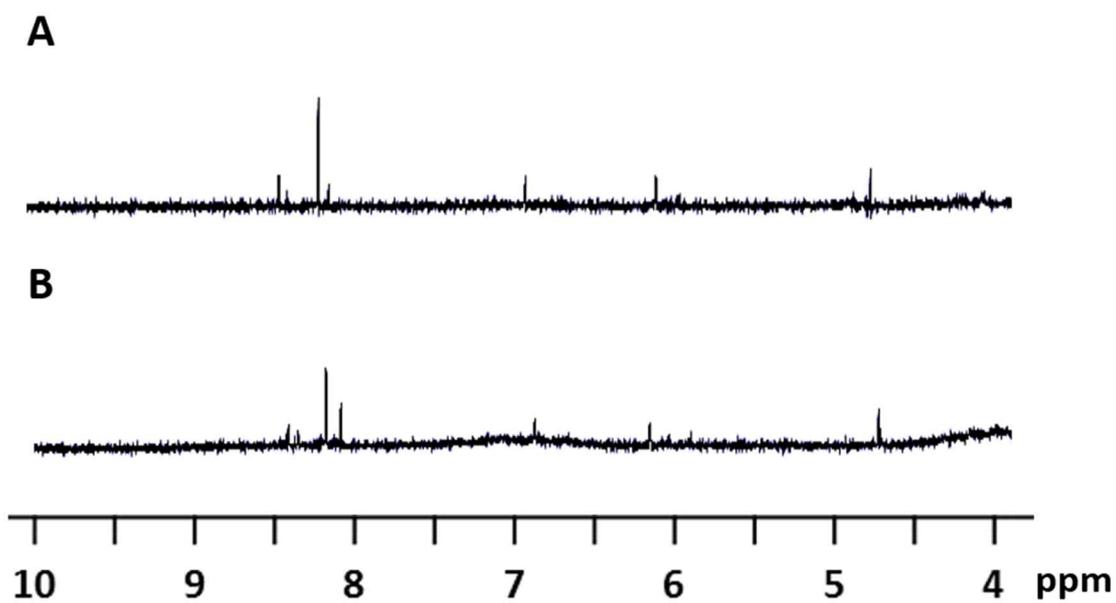


Figure 33. Comparison of ¹H STD competition NMR spectra. (A) STD spectrum of NAD⁺ and NADH mixed solution in the presence of NSDHL. (B) STD spectrum of NADP⁺ and NADPH mixed solution in the presence of NSDHL.

A

Proton No.	Wild type		G205S		K232Δ	
	STD effect (%)	Rank	STD effect (%)	Rank	STD effect (%)	Rank
1	0	8	12	8	31	6
2	47	5	26	7	24	7
3	58	4	47	4	42	4
4	60	3	57	3	48	3
5	100	1	79	2	100	1
6	38	6	44	5	33	5
7	84	2	100	1	77	2
8	0	8	0	9	0	9
9	23	7	34	6	7	8

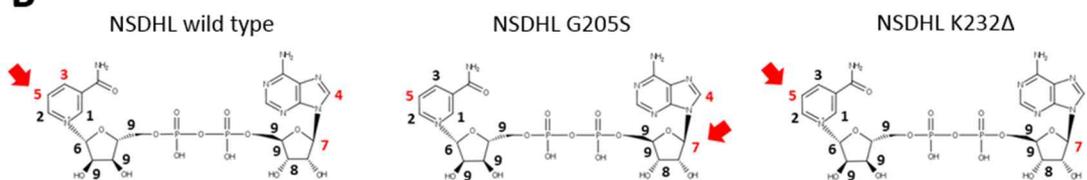
B

Figure 34. Normalized STD effect of NAD⁺ protons in the presence of wild-type, G205S, or K232Δ NSDHL. (A) Each peak in the reference ¹H NMR spectrum [Figure 27 (B)] has been numbered, and the STD effect has been calculated by normalizing to the largest value. (B) Based on the ranks presented in (A), protons which have STD effect more 50% is numbered in *red*. The proton presenting the largest STD effect in each result is indicated by *red* arrows.

A

Proton No.	Wild type		G205S		K232Δ	
	STD effect (%)	Rank	STD effect (%)	Rank	STD effect (%)	Rank
1	21	4	23	4	19	5
2	100	1	100	1	100	1
3	20	5	21	5	22	4
4	36	2	37	3	34	3
5	15	6	18	6	17	7
6	0	8	0	8	0	8
7	30	3	42	2	52	2
8	6	7	8	7	17	6

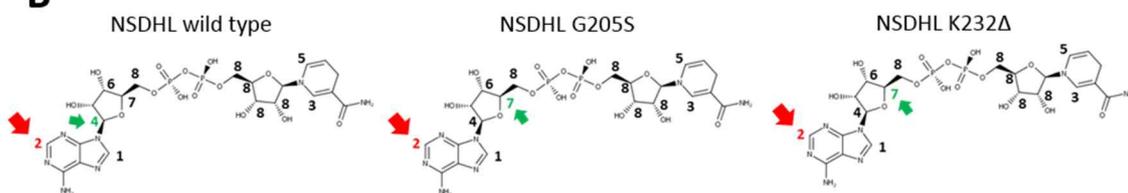
B

Figure 35. Normalized STD effect of NADH protons in the presence of wild-type, G205S, or K232Δ NSDHL. (A) Each peak in the reference ^1H NMR spectrum [Figure 28 (B)] has been numbered, and the STD effect has been calculated by normalizing to the largest value. (B) Based on the ranks presented in (A), protons which have the largest STD effect is numbered in *red* and indicated by *red* arrows. The proton presenting the second largest STD effect is numbered in *green* and indicated by *green* arrows.

2.4 Discussion

Due to the structural similarities of intermediates in the cholesterol synthesis and broad substrate specificity of the enzymes in cholesterol synthesis pathway, the specific reaction mechanism of NSDHL still remains undefined [90]. The structural and biophysical studies for NSDHL have been conducted for the following reasons: structural information is essential to identify key amino acids in the active site and it would provide valuable insight for designing novel drugs; and biophysical studies on the residue–ligand interaction would propose the reaction mechanism. Understanding of the mechanism is critical in early discovery and development of NSDHL inhibitors with potential utility in hypercholesterolemia and cancer.

Initially, a construct of NSDHL (31–267) was crystallized for determination of the 3D structure. Although the crystals diffracted to a good resolution (~ 2.2 Å), structure refinement was not successful from a collected data set. Because only when the structure is satisfactorily refined can the chosen space group be considered established and correct, the difficulty in space group determination was one of the main reasons that drives into a failure in phasing. Twinning may make difficult to identify the right space group. Close-packed crystals with low mosaic spreads were needed to solve the problem, therefore, extensive trials including additive/detergent screening, additional methionine introduction, and site-directed mutagenesis. However, it is hard to solve the problem by the extra methods. Currently, reductive methylation of NSDHL has been planned.

It is essential to identify the putative active site in NSDHL for the reason as mentioned above. Thus, homology modeling was conducted instead. In the models,

coenzymes [NAD(H) or NADP(H)] possibly bound to a cavity in the C-terminal regions which is formed with all β -strands. The substrate is expected to be located on the active site cavity near the cavity entrance which is composed of the H6 α -helix and the random coil between the H6 and the S6 β -strand. The substrate would bind closely to the nicotinamide ring of the coenzyme in a similar way as UDP-glucose binds in the template structures [88, 89]. Interestingly, Gly205 and Lys232 are located at the substrate binding site. Because CHILD and CKS are caused by point mutations of the residues, it is expected that the occurrences of the disease resulting from the altered substrate binding properties.

For measurement of the thermal stability, DSF experiment was performed which is an inexpensive, simple and quick method that probes the refolding of a protein as it experiences progressive denaturation [91, 92]. The hypothesis was confirmed that the mutant proteins are structurally perturbed comparing to its wild-type protein. The mutants presented notably lower melting temperatures (T_m), implying the mutations affect proteins to reduce their structural stability. The result is consistent with the prior study about K232 Δ folding that was altered at 37°C. [82]. Because the K232 Δ T_m is measured at 36.5°C, like the temperature with the human body, the CKS patients possibly have defective NSDHL protein and it will cause the accumulation of cholesterol intermediates. G205S also has lower T_m as 39.5°C, similar to the body temperature, and its unstable structure could be one of the main causes of CHILD syndrome. Indeed, NSDHL is localized on lipid droplet and redistributed to the endoplasmic reticulum in depletion of fatty acid, but G205S could no longer be localized on lipid droplet [81]. This result suggests that the defective protein folding causes the altered localization of NSDHL, and it finally causes the skin lesion and limb anomalies in CHILD syndrome.

NSDHL has been annotated as NAD(P) dependent steroid dehydrogenase. To confirm the exact coenzyme which NSDHL uses in the reaction, ITC and STD techniques were applied. Because all the experiments were conducted with the NSDHL 31–267 construct, the results could be different from those from the full-length protein. When NAD and NADH was titrated to NSDHL 31–267 construct, well-fitted graphs were obtained by ITC and each K_d was calculated in the micromolar range with relatively low affinity. In contrast, ITC data of NADP and NADPH were hardly fitted to the single-site binding isotherm model possibly due to the low affinities. Also, in STD result (Figure 33), NAD(H) showed stronger STD effect than NADP(H). Consequently, it has been revealed that NSDHL prefers NAD(H) to NADP(H) for its enzyme reaction. In the ITC studies, NADH showed a 7.5 fold higher affinity to NSDHL than that of NAD^+ , which is consistent with the STD competition data (Figure 29). These results are significantly different from a common belief that a substrate has higher affinity than that of a corresponding product. Possibly the following reasons of the result could be suggested: (i) the exact reaction mechanism of NSDHL needs to be more defined due to the complex role of the enzymes in cholesterol synthesis pathway, as mentioned before; (ii) NSDHL could be an example of the protein that NADH binds to stronger rather than NAD^+ , though the protein needs NAD^+ as coenzyme for its reaction. Carboxyl-terminal binding protein (CtBP) is the case, that shows a >100 fold higher affinity of NADH than NAD^+ [93]. Based on the evidence, it is concluded that NADH binds to NSDHL more strongly than NAD^+ for efficiency of enzyme reaction in the multi-enzyme pathway. In the cytoplasm, the concentration of free NAD^+ remarkably exceeds that of NADH in a ratio of 700 : 1. If the NSDHL reaction occurs, the product NADH easily diffuses to the cytoplasm due to the gradient of concentration. To prevent NADH to spread out and make the next enzyme

utilize NADH easily, NSDHL might hold NADH strongly until the next reaction happens; or (iii) for the tight regulation of the pathway, NADH might bind to NSDHL as a case of product inhibition [94]. Isocitrate dehydrogenase, one of Rossmann fold proteins, is carefully regulated to avoid unnecessary depletion of isocitrate by the product inhibition (by NADH or NADPH) mechanism [95]. However, the proposed mechanism should be obtained from further kinetics studies.

The STD experiment is a popular and highly versatile screening method which could not only be used for the identification of protein-ligand interactions over an extraordinarily broad affinity range, but also give quantitative information about the binding, through the epitope mapping [96]. By using the method, the interactions between NSDHL and the coenzymes were observed in STD spectra and the epitope mapping results gave information of their binding modes. The epitope mapping was obtained by normalizing the STD effects against the strongest STD signal that was assigned to 100%. Also, the shorter the protein-ligand proton-proton distance (bound state), the stronger the intensity of the corresponding STD signal. The protons with the largest STD effect are different when NAD^+ or NADH bind to NSDHL [Figures 34 (B) and 35(B)], which implied that the distance between the protons of the protein and the ligands have changed. NAD^+ binds to the protein more closely by its nicotinamide part, in contrast, the adenine ring of NADH strongly binds to NSDHL. From the result, it can be expected that conformations of NSDHL are dependent on the type of the enzyme-bound coenzymes. It is highly related to the report regarding human UDP-galactose 4-epimerase which has both open and closed forms depending on coenzymes [97]. In the NAD^+ STD spectrum, the proton-5 showed the largest STD effect implying the part is placed relatively near NSDHL (Figure 27). However, in the other NAD^+ -protein complex structures (PDB:

1Z45 and 4ZRN) [88, 89], the protons corresponding to proton-5 are not the closest to NSDHL. Similarly, proton-2 (Figure 28) with the largest STD value in the NADH STD spectrum is not the closest in the other NADH–protein complex structure (PDB: 1EK6) [97]. The disagreement between the STD result and the complex structures possibly is due to the STD experiments conducted without the substrate. In the complex structures, the coenzymes bind to the substrates as well as the neighboring residues of NSDHL. Through this interactions, the coenzymes could fit into the active sites. In contrast, in our study, NAD⁺ binds to NSDHL without any interaction with the substrate, therefore the binding mode is different from that of the complex structures. In contrast to the ITC results from NADP(H) titration, NADP(H) were identified to bind to NSDHL via STD experiments. The issue was related to the detection limit of each method: ITC whose K_d values from the nano- to the micromolar range, and STD's K_d values to the millimolar limit.

The genetic disease-related mutants showed different results when they were compared to the wild-type NSDHL: they showed lower apparent thermal stability (Figure 24); they had lower affinities to NADH in ITC data; and the each binding mode of the coenzymes was found slightly changed through STD epitope mapping (Figures 34 and 35). These findings suggests that the mutants have altered protein folding. A large number of diseases in humans result from protein misfolding and aggregation. Protein misfolding is believed to be the primary cause of Alzheimer's disease, Parkinson's disease, Huntington's disease, Creutzfeldt–Jakob disease, cystic fibrosis, and Gaucher's disease [98]. Like the diseases, CHILD and CKS were confirmed again as protein misfolding diseases based on the evidences from this work along with other studies demonstrating that the mutants' altered folding [81, 82]. To the best of our knowledge, this is the first

work that shows that the disease-causing residues are near the putative active site of NSDHL. It is possible that the mutants have unusual active site structure, which causes different arrangement of the ligands compared to that of the wild-type NSDHL. These findings support the possibility that changes in binding mode could be relevant in certain disease states found in CHILD and CKS.

Conclusion

As molecular devices, protein performs dynamic biological functions; life maintenance, replication, defense and reproduction. And these functions are highly related to sequence, dynamics, and protein 3D structure. Therefore, protein structural studies are essential to understand mechanism of functions, interaction with other protein and small molecules, and disease occurrence. In this study, two proteins have been studied on the structures and biophysical properties for the final purpose of structure-based drug design. Through the study on *HpKDO8PS*, four types of X-ray crystal structures were determined and the structures were used to identify the protein-binders via virtual screening and ligand-observed NMR experiments. The hits have potential to be new drugs with further lead optimization and biological testing. Also, the properties related to metal binding were determined by ITC, DSF experiments introducing in-depth discussion into the contribution of metal ions to the structural stability. In the case of NSDHL, it was possible to expect the specific binding sites of substrate and coenzyme via sequence homology modeling. In addition, the disease-causing mutants were revealed that they have different coenzyme binding modes and affinities compared to those of wild-type NSDHL. Based on the results, it was concluded that the mutants have defective folding and have difficulties in binding to the substrate and coenzymes with favorable binding modes. Therefore, they lead to malfunction in the cholesterol biosynthetic pathway, and cause the pathological symptoms. In this thesis, structure-based drug design was demonstrated by the studies on *HpKDO8PS*, and the close link between protein misfolding and diseases was confirmed again through the work of NSDHL.

References

- [1] B.J. Marshall, J.R. Warren, Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration, *Lancet*, 1 (1984) 1311-1315.
- [2] A.A. Ciociola, D.J. McSorley, K. Turner, D. Sykes, J.B. Palmer, *Helicobacter pylori* infection rates in duodenal ulcer patients in the United States may be lower than previously estimated, *Am J Gastroenterol*, 94 (1999) 1834-1840.
- [3] J.C. Atherton, The pathogenesis of *Helicobacter pylori*-induced gastro-duodenal diseases, *Annu Rev Pathol*, 1 (2006) 63-96.
- [4] Y. Yamaoka, *Helicobacter pylori: molecular genetics and cellular biology*, Caister Academic, *Wymondham*, 2008.
- [5] X.C. Wu, P. Andrews, V.W. Chen, F.D. Groves, Incidence of extranodal non-Hodgkin lymphomas among whites, blacks, and Asians/Pacific Islanders in the United States: anatomic site and histology differences, *Cancer epidem*, 33 (2009) 337-346.
- [6] K.M. Mohandas, *Helicobacter pylori* and gastric lymphoma, *N Engl J Med*, 331 (1994) 746.
- [7] J.R. Kerr, A. Al-Khattaf, A.J. Barson, J.P. Burnie, An association between sudden infant death syndrome (SIDS) and *Helicobacter pylori* infection, *Arch Dis Child*, 83 (2000) 429-434.
- [8] P. Malfertheiner, F. Megraud, C.A. O'Morain, J. Atherton, A.T. Axon, F. Bazzoli, G.F. Gensini, J.P. Gisbert, D.Y. Graham, T. Rokkas, E.M. El-Omar, E.J. Kuipers,

- Management of *Helicobacter pylori* infection--the Maastricht IV/ Florence Consensus Report, *Gut*, 61 (2012) 646-664.
- [9] W. Wu, Y. Yang, G. Sun, Recent Insights into Antibiotic Resistance in *Helicobacter pylori* Eradication, *Gastroenterol Res Pract*, 2012 (2012) 723183.
- [10] E.A. Cameron, K.U. Powell, L. Baldwin, P. Jones, G.D. Bell, S.G. Williams, *Helicobacter pylori*: antibiotic resistance and eradication rates in Suffolk, UK, 1991-2001, *J Med Microbiol*, 53 (2004) 535-538.
- [11] D. Carcanague, Y.K. Shue, M.A. Wuonola, M. Uria-Nickelsen, C. Joubran, J.K. Abedi, J. Jones, T.C. Kuhler, Novel structures derived from 2-[[2-(2-pyridyl)methyl]thio]-1H-benzimidazole as anti-*Helicobacter pylori* agents, Part 2, *J Med Chem*, 45 (2002) 4300-4309.
- [12] D.H. Levin, E. Racker, Condensation of arabinose 5-phosphate and phosphorylenol pyruvate by 2-keto-3-deoxy-8-phosphooctonic acid synthetase, *J Biol Chem*, 234 (1959) 2532-2539.
- [13] A.P. Moran, M.M. Prendergast, B.J. Appelmelk, Molecular mimicry of host structures by bacterial lipopolysaccharides and its contribution to disease, *FEMS Immunol Med Microbiol*, 16 (1996) 105-115.
- [14] C.R. Raetz, C. Whitfield, Lipopolysaccharide endotoxins, *Annu Rev Biochem*, 71 (2002) 635-700.
- [15] T.J. Wyckoff, C.R. Raetz, J.E. Jackman, Antibacterial and anti-inflammatory agents that target endotoxin, *Trends Microbiol*, 6 (1998) 154-159.

- [16] C.R. Raetz, C.M. Reynolds, M.S. Trent, R.E. Bishop, Lipid A modification systems in gram-negative bacteria, *Annu Rev Biochem*, 76 (2007) 295-329.
- [17] S. Manocha, D. Feinstein, A. Kumar, A. Kumar, Novel therapies for sepsis: antiendotoxin therapies, *Expert Opin Investig Drugs*, 11 (2002) 1795-1812.
- [18] A.P. Moran, The role of endotoxin in infection: *Helicobacter pylori* and *Campylobacter jejuni*, *Subcell Biochem*, 53 (2010) 209-240.
- [19] N. Semeraro, P. Montemurro, C. Piccoli, V. Muolo, M. Colucci, G. Giuliani, D. Fumarola, S. Pece, A.P. Moran, 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 (1996) 1255-1260.
- [20] M. Sarkar, L. Maganti, N. Ghoshal, C. Dutta, In silico quest for putative drug targets in *Helicobacter pylori* HPAG1: molecular modeling of candidate enzymes from lipopolysaccharide biosynthesis pathway, *J Mol Model*, 18 (2012) 1855-1866.
- [21] M.R. Birck, R.W. Woodard, *Aquifex aeolicus* 3-deoxy-D-manno-2-octulosonic acid 8-phosphate synthase: a new class of KDO 8-P synthase?, *J Mol Evol*, 52 (2001) 205-214.
- [22] D.J. Krosky, R. Alm, M. Berg, G. Carmel, P.J. Tummino, B. Xu, W. Yang, *Helicobacter pylori* 3-deoxy-D-manno-octulosonate-8-phosphate (KDO-8-P) synthase is a zinc-metalloenzyme, *Biochim Biophys Acta*, 1594 (2002) 297-306.

- [23] H.S. Duewel, S. Radaev, J. Wang, R.W. Woodard, D.L. Gatti, Substrate and metal complexes of 3-deoxy-D-manno-octulosonate-8-phosphate synthase from *Aquifex aeolicus* at 1.9-Å resolution. Implications for the condensation mechanism, *J Biol Chem*, 276 (2001) 8393-8402.
- [24] S. Radaev, P. Dastidar, M. Patel, R.W. Woodard, D.L. Gatti, Structure and mechanism of 3-deoxy-D-manno-octulosonate 8-phosphate synthase, *J Biol Chem*, 275 (2000) 9476-9484.
- [25] F.C. Cochrane, T.V. Cookson, G.B. Jameson, E.J. Parker, Reversing evolution: re-establishing obligate metal ion dependence in a metal-independent KDO8P synthase, *J Mol Biol*, 390 (2009) 646-661.
- [26] L. Baugh, L.A. Gallagher, R. Patrapuvich, M.C. Clifton, A.S. Gardberg, T.E. Edwards, B. Armour, D.W. Begley, S.H. Dieterich, D.M. Dranow, J. Abendroth, J.W. Fairman, D. Fox, 3rd, B.L. Staker, I. Phan, A. Gillespie, R. Choi, S. Nakazawa-Hewitt, M.T. Nguyen, A. Napuli, L. Barrett, G.W. Buchko, R. Stacy, P.J. Myler, L.J. Stewart, C. Manoil, W.C. Van Voorhis, Combining functional and structural genomics to sample the essential *Burkholderia* structome, *PloS one*, 8 (2013) e53851.
- [27] S.K. Nelson, A. Kelleher, G. Robinson, S. Reiling, O.A. Asojo, Structure of 2-keto-3-deoxy-D-manno-octulosonate-8-phosphate synthase from *Pseudomonas aeruginosa*, *Acta Crystallogr Sect F Struct Biol Cryst Commun*, 69 (2013) 1084-1088.
- [28] O. Asojo, J. Friedman, N. Adir, V. Belakhov, Y. Shoham, T. Baasov, Crystal structures of KDOP synthase in its binary complexes with the substrate

- phosphoenolpyruvate and with a mechanism-based inhibitor, *Biochemistry*, 40 (2001) 6326-6334.
- [29] J. Wang, H.S. Duewel, R.W. Woodard, D.L. Gatti, Structures of *Aquifex aeolicus* KDO8P synthase in complex with R5P and PEP, and with a bisubstrate inhibitor: role of active site water in catalysis, *Biochemistry*, 40 (2001) 15676-15683.
- [30] J. Wang, H.S. Duewel, J.A. Stuckey, R.W. Woodard, D.L. Gatti, Function of His185 in *Aquifex aeolicus* 3-deoxy-D-manno-octulosonate 8-phosphate synthase, *J Mol Biol*, 324 (2002) 205-214.
- [31] H.S. Duewel, R.W. Woodard, A metal bridge between two enzyme families. 3-deoxy-D-manno-octulosonate-8-phosphate synthase from *Aquifex aeolicus* requires a divalent metal for activity, *J Biol Chem*, 275 (2000) 22824-22831.
- [32] F. Kona, P. Tao, P. Martin, X. Xu, D.L. Gatti, Electronic structure of the metal center in the Cd²⁺, Zn²⁺, and Cu²⁺ substituted forms of KDO8P synthase: implications for catalysis, *Biochemistry*, 48 (2009) 3610-3630.
- [33] X. Xu, J. Wang, C. Grison, S. Petek, P. Coutrot, M.R. Birck, R.W. Woodard, D.L. Gatti, Structure-based design of novel inhibitors of 3-deoxy-D-manno-octulosonate 8-phosphate synthase, *Drug Des Discov*, 18 (2003) 91-99.
- [34] Z. Otwinowski, W. Minor, Processing of X-ray diffraction data collected in oscillation mode, *Method Enzymol*, 276 (1997) 307-326.
- [35] B.W. Matthews, Solvent content of protein crystals, *J Mol Biol*, 33 (1968) 491-497.

- [36] P.D. Adams, P.V. Afonine, G. Bunkoczi, V.B. Chen, I.W. Davis, N. Echols, J.J. Headd, L.W. Hung, G.J. Kapral, R.W. Grosse-Kunstleve, A.J. McCoy, N.W. Moriarty, R. Oeffner, R.J. Read, D.C. Richardson, J.S. Richardson, T.C. Terwilliger, P.H. Zwart, PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallogr D Biol Crystallogr*, 66 (2010) 213-221.
- [37] G.N. Murshudov, A.A. Vagin, E.J. Dodson, Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr D Biol Crystallogr*, 53 (1997) 240-255.
- [38] P. Emsley, B. Lohkamp, W.G. Scott, K. Cowtan, Features and development of Coot, *Acta Crystallogr D Biol Crystallogr*, 66 (2010) 486-501.
- [39] C. L.P., Sybyl–X,2.1 ed, SYBYL molecular modeling software, in, St.Louis, MO, 2013.
- [40] S.Y. Lee, K.H. Kim, I.K. Lee, K.H. Lee, S.U. Choi, K.R. Lee, A new flavonol glycoside from *Hylomecon vernalis*, *Arch Pharm Res*, 35 (2012) 415-421.
- [41] M. Mayer, B. Meyer, Characterization of ligand binding by saturation transfer difference NMR spectroscopy, *Angew Chem Int Edit*, 38 (1999) 1784-1788.
- [42] B. Meyer, T. Peters, NMR Spectroscopy techniques for screening and identifying ligand binding to protein receptors, *Angew Chem Int Edit*, 42 (2003) 864-890.
- [43] A. Viegas, J. Manso, F.L. Nobrega, E.J. Cabrita, Saturation-Transfer Difference (STD) NMR: A Simple and Fast Method for Ligand Screening and Characterization of Protein Binding, *J Chem Educ*, 88 (2011) 990-994.

- [44] A. Viegas, J. Manso, M.C. Corvo, M.M.B. Marques, E.J. Cabrita, Binding of Ibuprofen, Ketorolac, and Diclofenac to COX-1 and COX-2 Studied by Saturation Transfer Difference NMR, *J Med Chem*, 54 (2011) 8555-8562.
- [45] C. Dalvit, G. Fogliatto, A. Stewart, M. Veronesi, B. Stockman, WaterLOGSY as a method for primary NMR screening: Practical aspects and range of applicability, *J Biomol Nmr*, 21 (2001) 349-359.
- [46] G.N. Murshudov, A.A. Vagin, E.J. Dodson, Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr D Biol Crystallogr*, 53 (1997) 240-255.
- [47] V.B. Chen, W.B. Arendall, J.J. Headd, D.A. Keedy, R.M. Immormino, G.J. Kapral, L.W. Murray, J.S. Richardson, D.C. Richardson, MolProbity: all-atom structure validation for macromolecular crystallography, *Acta Crystallogr D Biol Crystallogr*, 66 (2010) 12-21.
- [48] M. Heinig, D. Frishman, STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins, *Nucleic Acids Res*, 32 (2004) W500-502.
- [49] F.H. Niesen, H. Berglund, M. Vedadi, The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability, *Nat Protoc*, 2 (2007) 2212-2221.
- [50] C.A. Lipinski, F. Lombardo, B.W. Dominy, P.J. Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, *Adv Drug Deliv Rev*, 46 (2001) 3-26.

- [51] C. Dalvit, P. Pevarello, M. Tato, M. Veronesi, A. Vulpetti, M. Sundstrom, Identification of compounds with binding affinity to proteins via magnetization transfer from bulk water, *J Biomol Nmr*, 18 (2000) 65-68.
- [52] J.L. Wagstaff, S.L. Taylor, M.J. Howard, Recent developments and applications of saturation transfer difference nuclear magnetic resonance (STD NMR) spectroscopy, *Mol Biosyst*, 9 (2013) 571-577.
- [53] G. Otting, NMR studies of water bound to biological molecules, *Prog Nucl Mag Res Sp*, 31 (1997) 259-285.
- [54] C. Ludwig, P.J.A. Michiels, X. Wu, K.L. Kavanagh, E. Pilka, A. Jansson, U. Oppermann, U.L. Gunther, SALMON: Solvent accessibility, ligand binding, and mapping of ligand orientation by NMR Spectroscopy, *J Med Chem*, 51 (2008) 1-3.
- [55] R. Vainer, V. Belakhov, E. Rabkin, T. Baasov, N. Adir, Crystal structures of *Escherichia coli* KDO8P synthase complexes reveal the source of catalytic irreversibility, *J Mol Biol*, 351 (2005) 641-652.
- [56] L. Unione, S. Galante, D. Diaz, F.J. Canada, J. Jimenez-Barbero, NMR and molecular recognition. The application of ligand-based NMR methods to monitor molecular interactions, *Medchemcomm*, 5 (2014) 1280-1289.
- [57] N.E. Grosseohme, A.M. Spuches, D.E. Wilcox, Application of isothermal titration calorimetry in bioinorganic chemistry, *J Biol Inorg Chem*, 15 (2010) 1183-1191.

- [58] K.L. Haas, K.J. Franz, Application of Metal Coordination Chemistry To Explore and Manipulate Cell Biology, *Chem Rev*, 109 (2009) 4921-4960.
- [59] D.T. Puerta, S.M. Cohen, A bioinorganic perspective on matrix metalloproteinase inhibition, *Curr Top Med Chem*, 4 (2004) 1551-1573.
- [60] K.L. Haas, K.J. Franz, Application of metal coordination chemistry to explore and manipulate cell biology, *Chem Rev*, 109 (2009) 4921-4960.
- [61] A.W. Boots, G.R. Haenen, A. Bast, Health effects of quercetin: from antioxidant to nutraceutical, *Eur J Pharmacol*, 585 (2008) 325-337.
- [62] S.C. Bischoff, Quercetin: potentials in the prevention and therapy of disease, *Curr Opin Clin Nutr Metab Care*, 11 (2008) 733-740.
- [63] K.A. Lee, S.H. Moon, K.T. Kim, A.F. Mendonca, H.D. Paik, Antimicrobial Effects of Various Flavonoids on *Escherichia coli* O157:H7 Cell Growth and Lipopolysaccharide Production, *Food Sci Biotechnol*, 19 (2010) 257-261.
- [64] M. Kalegari, Antibacterial, allelopathic and antioxidant activity of extracts and compounds from *Rourea induta* Planch. (Connaraceae), *J Appl Pharm Sci*, (2012).
- [65] S. Lee, D.S. Shin, K.B. Oh, K.H. Shin, Antibacterial compounds from the leaves of *Acanthopanax senticosus*, *Arch Pharm Res*, 26 (2003) 40-42.
- [66] A.M. Metwally, A.A. Omar, F.M. Harraz, S.M. El Sohafy, Phytochemical investigation and antimicrobial activity of *Psidium guajava* L. leaves, *Pharmacogn Mag*, 6 (2010) 212-218.
- [67] I. Hanukoglu, Steroidogenic enzymes: structure, function, and role in regulation of steroid hormone biosynthesis, *J Steroid Biochem Mol Biol*, 43 (1992) 779-804.

- [68] L.J. Sharpe, A.J. Brown, Controlling cholesterol synthesis beyond 3-hydroxy-3-methylglutaryl-CoA reductase (HMGCR), *J Biol Chem*, 288 (2013) 18707-18715.
- [69] S.B. Helliwell, S. Karkare, M. Bergdoll, A. Rahier, J.R. Leighton-Davis, C. Fioretto, T. Aust, I. Filipuzzi, M. Frederiksen, J. Gounarides, D. Hoepfner, A. Hofmann, P.E. Imbert, R. Jeker, R. Knochenmuss, P. Krastel, A. Margerit, K. Memmert, C.V. Mialt, N.R. Movva, A. Muller, H.U. Naegeli, L. Oberer, V. Prindle, R. Riedl, S. Schuierer, J.A. Sexton, J.S. Tao, T. Wagner, H. Yin, J. Zhang, S. Roggo, S. Reinker, C.N. Parker, FR171456 is a specific inhibitor of mammalian NSDHL and yeast Erg26p, *Nat Commun*, 6 (2015).
- [70] L. Gabitova, D. Restifo, A. Gorin, K. Manocha, E. Handorf, D.H. Yang, K.Q. Cai, A.J. Klein-Szanto, D. Cunningham, L.E. Kratz, G.E. Herman, E.A. Golemis, I. Astsaturov, Endogenous Sterol Metabolites Regulate Growth of EGFR/KRAS-Dependent Tumors via LXR, *Cell Rep*, 12 (2015) 1927-1938.
- [71] S. Ghosh, S. Mukherjee, A. Basu, Chandipura virus perturbs cholesterol homeostasis leading to neuronal apoptosis, *J Neurochem*, 135 (2015) 368-380.
- [72] E.I. Mercer, Inhibitors of Sterol Biosynthesis and Their Applications, *Prog Lipid Res*, 32 (1993) 357-416.
- [73] S. Lim, P.C. Oh, I. Sakuma, K.K. Koh, How to balance cardiorenometabolic benefits and risks of statins, *Atherosclerosis*, 235 (2014) 644-648.
- [74] K. Baudry, E. Swain, A. Rahier, M. Germann, A. Batta, S. Rondet, S. Mandala, K. Henry, G.S. Tint, T. Edlind, M. Kurtz, J.T. Nickels, The effect of the erg26-1 mutation on the regulation of lipid metabolism in *Saccharomyces cerevisiae*, *J Biol Chem*, 276 (2001) 12702-12711.

- [75] C. Mo, M. Valachovic, S.K. Randall, J.T. Nickels, M. Bard, Protein-protein interactions among C-4 demethylation enzymes involved in yeast sterol biosynthesis, *Proc Natl Acad Sci U S A*, 99 (2002) 9739-9744.
- [76] M.A. He, L.E. Kratz, J.J. Michel, A.N. Vallejo, L. Ferris, R.I. Kelley, J.J. Hoover, D. Jukic, K.M. Gibson, L.A. Wolfe, D. Ramachandran, M.E. Zwick, J. Vockley, Mutations in the human SC4MOL gene encoding a methyl sterol oxidase cause psoriasiform dermatitis, microcephaly, and developmental delay, *J Clin Invest*, 121 (2011) 976-984.
- [77] X.Y. Liu, A.W. Dangel, R.I. Kelley, W. Zhao, P. Denny, M. Botcherby, B. Cattanaach, J. Peters, P.R. Hunsicker, A.M. Mallon, M.A. Strivens, R. Bate, W. Miller, M. Rhodes, S.D.M. Brown, G.E. Herman, The gene mutated in bare patches and striated mice encodes a novel 3 beta-hydroxysteroid dehydrogenase, *Nat Genet*, 22 (1999) 182-187.
- [78] A. Sukhanova, A. Gorin, I.G. Serebriiskii, L. Gabitova, H. Zheng, D. Restifo, B.L. Egleston, D. Cunningham, T. Bagnyukova, H.Q. Liu, A. Nikonova, G.P. Adams, Y. Zhou, D.H. Yang, R. Mehra, B. Burtness, K.Q. Cai, A. Klein-Szanto, L.E. Kratz, R.I. Kelley, L.M. Weiner, G.E. Herman, E.A. Golemis, I. Astsaturov, Targeting C4-Demethylating Genes in the Cholesterol Pathway Sensitizes Cancer Cells to EGF Receptor Inhibitors via Increased EGF Receptor Degradation, *Cancer Discov*, 3 (2013) 96-111.
- [79] T. Xue, Y. Zhang, L. Zhang, L. Yao, X. Hu, L.X. Xu, Proteomic analysis of two metabolic proteins with potential to translocate to plasma membrane associated

- with tumor metastasis development and drug targets, *J Proteome Res*, 12 (2013) 1754-1763.
- [80] M. Morimoto, C. Souich, J. Trinh, K.W. McLarren, C.F. Boerkoel, G. Hendson, Expression profile of NSDHL in human peripheral tissues, *J Mol Histol*, 43 (2012) 95-106.
- [81] M. Ohashi, N. Mizushima, Y. Kabeya, T. Yoshimori, Localization of mammalian NAD(P)H steroid dehydrogenase-like protein on lipid droplets, *J Biol Chem*, 278 (2003) 36819-36829.
- [82] K.W. McLarren, T.M. Severson, C. du Souich, D.W. Stockton, L.E. Kratz, D. Cunningham, G. Hendson, R.D. Morin, D. Wu, J.E. Paul, J.H. An, T.N. Nelson, A. Chou, A.E. DeBarber, L.S. Merkens, J.L. Michaud, P.J. Waters, J.Y. Yin, B. McGillivray, M. Demos, G.A. Rouleau, K.H. Grzeschik, R. Smith, P.S. Tarpey, D. Shears, C.E. Schwartz, J. Gecz, M.R. Stratton, L. Arbour, J. Hurlburt, M.I. Van Allen, G.E. Herman, Y.J. Zhao, R. Moore, R.I. Kelley, S.J.M. Jones, R.D. Steiner, F.L. Raymond, M.A. Marra, C.F. Boerkoel, Hypomorphic Temperature-Sensitive Alleles of NSDHL Cause CK Syndrome, *Am J Hum Genet*, 87 (2010) 905-914.
- [83] B. Yin, D.B. Cui, L.J. Zhang, S.Q. Jiang, S. Machida, Y.A. Yuan, D.Z. Wei, Structural insights into substrate and coenzyme preference by SDR family protein Gox2253 from *Gluconobater oxydans*, *Proteins*, 82 (2014) 2925-2935.
- [84] E.L. Ulrich, H. Akutsu, J.F. Doreleijers, Y. Harano, Y.E. Ioannidis, J. Lin, M. Livny, S. Mading, D. Maziuk, Z. Miller, E. Nakatani, C.F. Schulte, D.E. Tolmie, R.K. Wenger, H.Y. Yao, J.L. Markley, BioMagResBank, *Nucleic Acids Res*, 36 (2008) D402-D408.

- [85] Z. Dosztanyi, V. Csizmok, P. Tompa, I. Simon, IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content, *Bioinformatics*, 21 (2005) 3433-3434.
- [86] A. Drozdetskiy, C. Cole, J. Procter, G.J. Barton, JPred4: a protein secondary structure prediction server, *Nucleic Acids Res*, 43 (2015) W389-W394.
- [87] M. Biasini, S. Bienert, A. Waterhouse, K. Arnold, G. Studer, T. Schmidt, F. Kiefer, T.G. Cassarino, M. Bertoni, L. Bordoli, T. Schwede, SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information, *Nucleic Acids Res*, 42 (2014) W252-W258.
- [88] J.B. Thoden, H.M. Holden, The molecular architecture of galactose mutarotase/UDP-galactose 4-epimerase from *Saccharomyces cerevisiae*, *J Biol Chem*, 280 (2005) 21900-21907.
- [89] S.M. Shin, J.M. Choi, E. di Luccio, Y.J. Lee, S.J. Lee, S.J. Lee, S.H. Lee, D.W. Lee, The structural basis of substrate promiscuity in UDP-hexose 4-epimerase from the hyperthermophilic Eubacterium *Thermotoga maritima*, *Arch Biochem Biophys*, 585 (2015) 39-51.
- [90] J. Acimovic, D. Rozman, Steroidal Triterpenes of Cholesterol Synthesis, *Molecules*, 18 (2013) 4002-4017.
- [91] J.V. Rodrigues, V. Prosinecki, I. Marrucho, L.P.N. Rebelo, C.M. Gomes, Protein stability in an ionic liquid milieu: on the use of differential scanning fluorimetry, *Phys Chem Chem Phys*, 13 (2011) 13614-13616.

- [92] Y.C. Fan, J.L. Yan, S.L. Zhang, J.H. Li, D. Chen, P.G. Duan, Fluorescence Spectroscopic Analysis of the Interaction of Papain with Ionic Liquids, *Appl Biochem Biotechnol*, 168 (2012) 592-603.
- [93] C.C. Fjeld, W.T. Birdsong, R.H. Goodman, Differential binding of NAD⁺ and NADH allows the transcriptional corepressor carboxyl-terminal binding protein to serve as a metabolic sensor, *Proc Natl Acad Sci U S A*, 100 (2003) 9202-9207.
- [94] B.F. Cooper, F.B. Rudolph, Product inhibition applications, *Methods Enzymol*, 249 (1995) 188-211.
- [95] M. I, G. Schmidt, H. Sahm, K.P. Stahmann, Role of a peroxisomal NADP-specific isocitrate dehydrogenase in the metabolism of the riboflavin overproducer *Ashbya gossypii*, *J Mol Catal B-Enzym*, 10 (2000) 335-343.
- [96] C. Aguirre, O. Cala, I. Krimm, Overview of Probing Protein-Ligand Interactions Using NMR, *Curr Protoc Protein Sci*, 81 (2015) 17 18 11-24.
- [97] J.B. Thoden, T.M. Wohlers, J.L. Fridovich-Keil, H.M. Holden, Crystallographic evidence for Tyr 157 functioning as the active site base in human UDP-galactose 4-epimerase, *Biochemistry*, 39 (2000) 5691-5701.
- [98] T.K. Chaudhuri, S. Paul, Protein-misfolding diseases and chaperone-based therapeutic approaches, *FEBS J*, 273 (2006) 1331-1349.

국문초록

촉매 효소(*Hp*KDO8PS 및 NSDHL)의 구조에 기반한 작용 기전 및 억제제 개발 연구

서울대학교 대학원
약학과 물리약학 전공
조수진

Structure-based drug design (SBDD)는 단백질의 삼차원 구조에 기반을 둔 신약개발 방법으로 단백질에 직접 결합하는 물질을 디자인하고, 신약 후보 물질로서 적절한 성질을 가지도록 최적화하는 과정으로 이루어져있다. 이 기술은 구조적 정보를 이용하여 생물학적 표적 단백질과 약물간의 상호작용을 연구하고, 궁극적으로 더욱 빠르고 경제적인 신약개발이 가능하도록 한다. 새로운 항생제와 고지혈증 치료제의 개발을 위하여 *Helicobacter pylori* 유래의 KDO8PS (3-deoxy-D-manno-octulosonate-8-phosphate synthase, *Hp*KDO8PS)와 *Homo sapiens* 유래의 NSDHL (sterol-4- α -carboxylate 3-dehydrogenase, decarboxylating)을 선택하여 이의 구조적, 생물물리학적 연구를 수행하였다.

*Hp*KDO8PS은 arabinose 5-phosphate (A5P)와 phosphoenolpyruvate (PEP)의 중합반응을 매개하여 lipopolysaccharide 합성의 중요물질인 3-deoxy-D-manno-octulosonate (KDO)의 전구체 KDO8P를 생산한다. 이 연구를 통하여

*HpKDO8PS*의 단백질 결정 구조를 세계 최초로 규명하였다. 다른 생물에서 발견되는 KDO8PS와 달리 *HpKDO8PS*는 extra helix (HE)를 가지고 있으며, 이 단백질의 Cd^{2+} 나 Zn^{2+} 금속이온이 결합한 구조는 *Aquifex aeolicus* 종의 KDO8PS와 다른 금속 결합 양상을 보인다. Isothermal titration calorimetry (ITC)와 differential scanning fluorimetry (DSF) 실험을 통해 Zn^{2+} 이온이 단백질을 열역학적으로 안정화시킨다는 사실을 발견하였다. 또한 이 단백질의 기질 결합 구조를 통해 기질의 결합에 물 분자가 중요한 역할을 하는 것을 알 수 있었다. *HpKDO8PS*와 API [arabinose 5-phosphate (A5P)과 phosphoenolpyruvate (PEP)의 구조를 합친 저해제]의 도킹 결과를 이용하여, 단백질에 결합할 것으로 예상되는 21개의 화합물을 *in silico* virtual screening을 통해 선정하였다. 그 중, avicularin, hyperin, MC181의 세 가지 화합물이 실제로 *HpKDO8PS*에 결합하는 것을 NMR saturation transfer difference (STD) 실험을 통해 규명하였고, 이 세 화합물을 단백질 구조에 도킹하여 단백질과의 상호작용을 면밀히 분석하였다. Hyperin의 경우, 단백질의 A5P 결합자리에 친수성 상호작용을 통해 결합하는 반면, MC181은 친수성, 소수성 상호작용 모두의 두 가지 방법을 통해 PEP와 A5P의 결합자리에 결합하는 것을 확인하였다. 이는 STD 실험을 통한 epitope mapping 결과와 일치한다. 본 연구는 *HpKDO8PS*의 저해제 개발에 중요한 단서를 제공할 것으로 기대한다.

NSDHL은 콜레스테롤 합성과정에서 4 α -carboxysterol의 NAD(H)⁺-dependent oxidative decarboxylation를 유도하여 3-keto, C4-decarboxylated 산물을 만드는 역할을 한다. NSDHL의 삼차원 구조를 규명하기 위하여 클로닝과 단

백질 정제 과정을 통해 고농도의 NSDHL을 얻었고, 이를 결정화하여 X선 회절 실험을 진행하였다. 회절 데이터를 성공적으로 수집하였으나, 약한 selenomethionine의 신호와 결정 packing의 문제로 위상결정에는 실패하였다. 이를 대신해 homology modeling을 시도하여 NSDHL의 모델 구조를 얻었다. homodimer로서 6개의 α -helix와 평행한 5개의 β -strand로 이루어져 있으며 전체적으로 Rossmann fold를 가지고 있다. 다른 Rossmann fold를 가진 단백질과 비교를 통해 단백질 활성자리와 조효소 결합자리를 예측하였다. NSDHL은 유전 질환과도 밀접한 관계가 있는데 이 질병(CHILD, CKS)을 유발하는 돌연변이에 관한 연구도 진행하였다. DSF 실험을 통해 돌연변이 단백질의 열역학적 안정성을 측정하였다. 돌연변이는 wild-type 단백질에 비해 낮은 녹는점을 보였고, 이를 통해 돌연변이 단백질이 상대적으로 불완전한 folding을 가지고 있다는 가설을 증명하였다. 또한, ITC를 통해 NSDHL과 조효소의 상호작용, 돌연변이와 조효소의 상호작용을 연구하였다. 그 결과 NSDHL은 효소반응에서 NADP(H)보다는 NAD(H)를 주로 이용한다는 것을 알 수 있었고, 각 상호작용의 해리상수를 도출하였다. NMR STD 실험을 통해 NAD와 NADH의 NSDHL 결합 시 가까이 위치할 부분을 알아내고, 돌연변이에서는 다르게 결합함을 밝힐 수 있었다. 이 발견을 통해 CHILD syndrome과 CKS 등의 유전 질환의 발병과 NSDHL의 구조가 밀접하게 연관되어 있음을 유추하였고, 이는 향후 콜레스테롤 합성 저해제 개발에 도움을 줄 것으로 예상된다.

주요어 : Structure-based drug discovery, *Helicobacter pylori*, *Homo sapiens*, virtual

screening, docking simulation, enzyme inhibitors, X-ray crystallography, NMR spectroscopy.

학 번 : 2010-21728