Doctoral Dissertation

Structural Basis for Substrate Recognition and Specificity of a Novel (*R*)-stereoselective ω-transaminase

（新規 ***R*** 体選択的ω-トランスアミナーゼの基質認識および基質特異性の構造基盤の解明）

Lijun Guan

管　立軍

CONTENTS

Abbreviations

DTT	dithiothreitol
DNA	deoxyribonucleic acid
DMA	dimethoxyamphetamine
HPLC:	high-performance liquid chromatography
IPTG:	isopropyl-β-D-thiogalactopyranoside
LB	lysogeny broth
MBA	methylbenzylamine
MOE	molecular operating environment
MWCO	molecular weight cut-off
PCR	polymerase chain reaction
PDB	protein data bank
PEG	polyethylene glycol
PF	photon factory
PLP	pyridoxal 5'-phosphate
PMP	pyridoxamine-5'-phosphate
RMSD	root mean square deviation
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
TAs	transaminases
TLC	Thin Layer Chromatography
U	unit

Chapter 1

Introduction

1.1 Application of biocatalysis

Recently enzymes are being utilized as biocatalysts in many industrial processes. A desired product raises the demand for a suitable biocatalyst. With proper enzymes, a detailed biochemical characterization study will help finding appropriate conditions to successfully use the enzymes as biocatalysts under the requirements of a specific process. If this characterization study should not make satisfying results, enzyme engineering might help creating enzyme variants. At this moment, the crystal structures of these enzymes are significantly important for rational engineering.

1.2 Chiral amines

Synthesis of chiral compounds is significant due to the side effects of undesired enantiomers. Pure chiral amines are also crucial building blocks in the synthesis of some pharmaceutical drugs, for example, antidiabetic (*S*)-Repaglinide [1], HIV-protease inhibitor (*S*)-Lopinavir [2], antihistamine (*R*)-Levocetirzine [3] and (*R*)-Cinacalcet [4] which is used to treat secondary hyperparathyroidism (Fig 1-1). During the synthesis of the chiral compounds, the generation of chirality is most often the actual challenge. Thus, optically pure amines became increasingly significant as building blocks in pharmaceutical industry.

1.3 Synthesis of chiral amines

Chemical transformations used to synthesize chiral amines have shortcomings such as harsh reaction conditions, effluent of toxic catalytic metal, and even insufficient enantiopurity. Several approaches to produce active chiral amines using biocatalysts, such as hydrolases, oxidoreductases and transferases [5] has been developed to overcome the shortcomings of chemical synthesis. However, these enzymes have their

own drawback such as low reactivity and reaction equilibrium issue [6]. As a solution to these problems, ω-transaminases (ω-TAs), trivially referring to amine-pyruvate transaminase, have been paid increased attention due to their unique ability for the asymmetric synthesis of pure chiral amines with high turnover rate, excellent stereoselectivity and stable catalytic activity [7]. Thus they are being seen as useful catalytic enzymes in industry production of chiral pharmaceutical drugs.

1.4 ω-Transaminases and their reaction mechanism

ω-TAs reversibly catalyze the transfer of an amino group from an amino compound to a ketone compound using the cofactor pyridoxal 5'-phosphate (PLP) (Fig. 1-2). The whole catalytic cycle of the transamination reaction consists of two half reactions following a typical ping pong bi bi kinetics [8]. In the first half reaction, the amino group of the amino donor is transferred to the PLP to form pyridoxamine-5'-phosphate (PMP) and the corresponding ketone product is released. In the second half reaction, the amino group is transferred from the PMP to the amino acceptor, while the cofactor is recycled to PLP and the product is released (Fig. 1-3). In detail, PLP forms a Schiff base with the active site lysine (internal aldimine). The amino group of the amino donor comes towards the internal aldimine to enable a transaldimination reaction between the internal aldimine and the amino donor, forming an external aldimine intermediate (Fig. 1-4). The external aldimine is further converted to a ketimine by a 1,3-prototropic shift. The ketimine is hydrolyzed to produce a PMP intermediate and release a ketone (Fig. 1-5). Then, in the second half reaction, the amino group will be transformed from PMP to the a amino acceptor.

1.5 Stereoselectivity of ω-transaminases

ω-Transaminases are divided into two groups, (*R*)- and (*S*)-ω-TAs, according to the chirality at the amino carbon atom of the substrate. Both enantiomers are indispensable for drug synthesis as described above, and therefore both of (*R*)-ω-TAs and (*S*)-ω-TAs increasingly become a top research. Up to now more than 20 (*S*)-ω-TAs have been identified and some of them have been biochemically characterized and display a similar substrate specificity [9]. Recently crystal structures of (*S*)-ω-TAs were solved and structural analyses were performed [10-12], such as (*S*)-ω-TAs from *Vibrio fluvialis JS17* (PDB ID:3NUI, 4E3Q and 4E3R), *Paracoccus denitrificans* (PDB ID:4GRX), *Pseudomonas putida* (PDB ID: 3A8U), *Pseudomonas aeruginosa* (PDB ID: 4BQ0, 4B9B and 4B98), and *Chromobacterium violaceum* (PDB ID: 4A6R, 4A6T, 4A6U, 4A72, 4AH3, 4BA4, 4BA5). These structures are useful to understand the mechanistic enzymology and substrate specificity of (*S*)-ω-TA, which also provide important information for their employment as industrial biocatalysts. However, the structure basis for the substrate specificity and stereoselectivity has not been clearly revealed yet.

In addition, although the kinetic resolution of racemic amine utilizing (*S*)-ω-TAs can produce (*R*)-enantiomers, (*R*)-ω-TAs are significant in the production of (*R*)-enantiomers by much more economical asymmetric synthesis (Fig. 1-6). In contrast to (*S*)-ω-TAs, there is relatively rare ω-TAs with (*R*)-stereoselectivity in the nature, and the information about them is also finite. Up to now, there has no crystal structures of (*R*)-ω-TA been reported yet.

1.6 (*R*)-ω-TAs in the nature and their application

The first (*R*)-stereoselective ω-TA identified from *Arthrobacter* sp.

KNK168(FERM-BP-5228) [(*R*)-TA] by enrichment screening shows activity on secondary amine and methyl ketones or small cyclic ketones [13] and was successfully applied to the asymmetric synthesis of several kinds of chiral amines, such as (R)-dimethoxyamphetamine (DMA), (*R*)-4-methoxyamphetamine, (*R*)-1-(3-hydroxyphenyl)ethylamine and (*R*)-1-(3-methoxyphenyl)ethylamine [14]. Recently, the report of purification, characterization, and gene cloning of it has been published [15].

A homolog of (*R*)-TA, ATA-117, was engineered by substrate walking approach to synthesize an active ingredient for type-2 diabetes, sitagliptin, on industry scale [16] (Fig. 1-7). The improved (*R*)-TA, designated as ATA-117-Rd11, gave a 13% increase in yield and a 19% reduction in total waste, which ultimately reduced the production cost when compared to the chemical approach used to synthesize sitagliptin (Fig. 1-8).

Attempts to create an (*R*)-ω-TA from other TAs were performed due to the rare of (*R*)-ω-TA. However, because of no crystal structure of any (*R*)-ω-TA, the rational protein design was thought to be very difficult and thus (*S*)-selective α-transaminases with solved crystal structure were firstly chosen and engineered the substrate-recognition site to create a mutant possessing (*R*)-ω-TA activity. However, it has not yet been successful. Subsequently, the researchers changed their train of thought. They identified 20 (*R*)-ω-TAs by an *in silico* search from public database [17] (Fig. 1-9). The substrate specificity and reaction conditions of these enzymes are under study.

1.7 Purpose of this research

Although the above three approaches to obtain an (*R*)-stereoselective ω-TA made breakthough, the structural information of (*R*)-stereoselective ω-TA should be

elucidated to further the development of biocatalytic application of the enzyme. Here I firstly determined eight crystal structures of (*R*)-ω-TA from *Arthrobacter* sp. KNK168(FERM-BP-5228) ((*R*)-TA); 1) the PLP-bound form of the wild-type (*R*)-TA ((*R*)-TA^{WT}), 2-3) two ketone substrate complexes of (*R*)-TA ((*R*)-TA^{PYR} and (*R*)-TA^{METH}), 4) the PMP-bound form of the wild-type (*R*)-TA ((*R*)-TA^{PMP}), 5) the K188Q mutant of (*R*)-TA complexed with the external aldimine, which is the reaction intermediate formed by the substrate *R*-MBA and cofactor PLP ((*R*)-TA^{K188Q_REX}), 6) the G136F mutant in complex with PLP ((*R*)-TA^{G136F}), 7) the mutant with 28 mutations (Mut28), which was identical to the previously engineered mutant, ATA-117-Rd11 ((*R*)-TA^{Mut28}) and 8) the K188Q mutant of (*R*)-TA complexed with the external aldimine, which is the intermediate formed by substrate cofactor PLP and nonreactive substrate (*S*)-α-MBA ((*R*)-TA^{K188Q_SEX}). Understanding of the structural features of (*R*)-ω-TAs will allow rational mutagenesis to redesign an enzyme to produce unusual chiral amines for specific industrial applications.

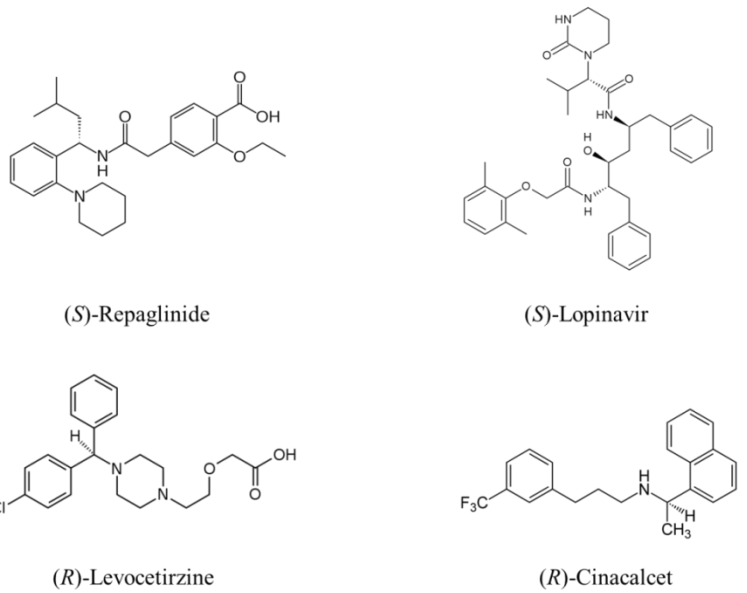


Fig 1-1 Example for chiral amines in pharmaceuticals.

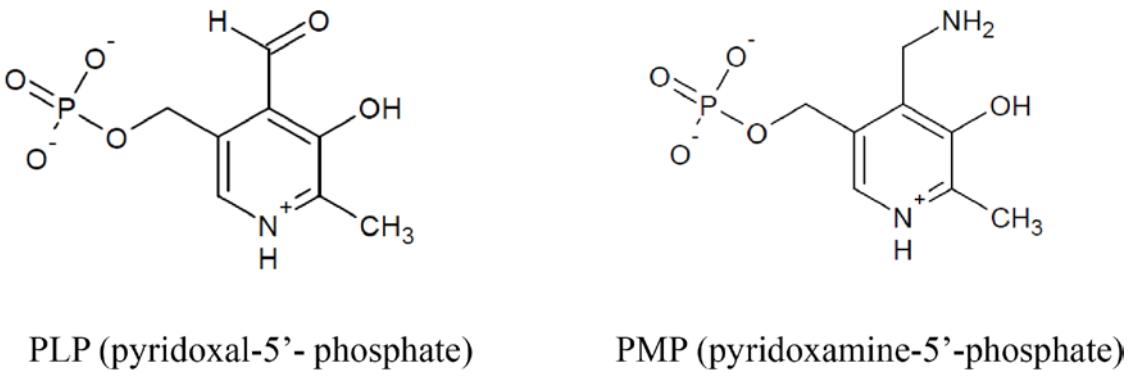


Fig.1-2 The chemical formulas of cofactor PLP and PMP.

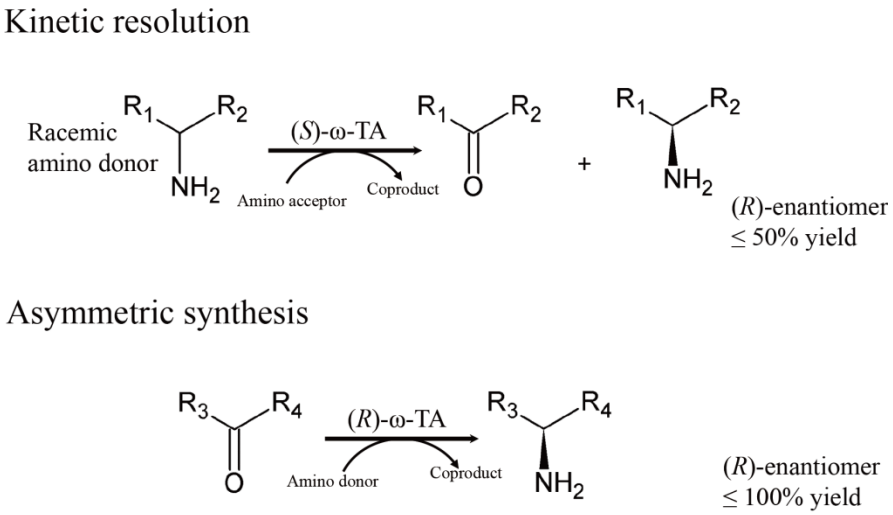


Fig. 1-3 Two steps of the transaminase reaction.

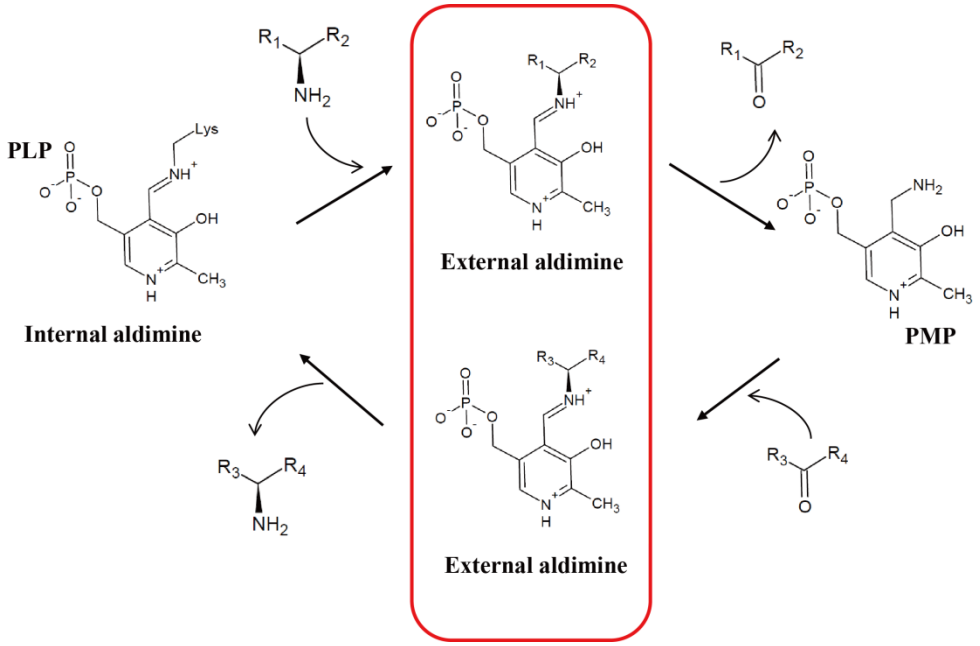


Fig. 1-4 A scheme of transforming between internal aldimne and external aldimine.

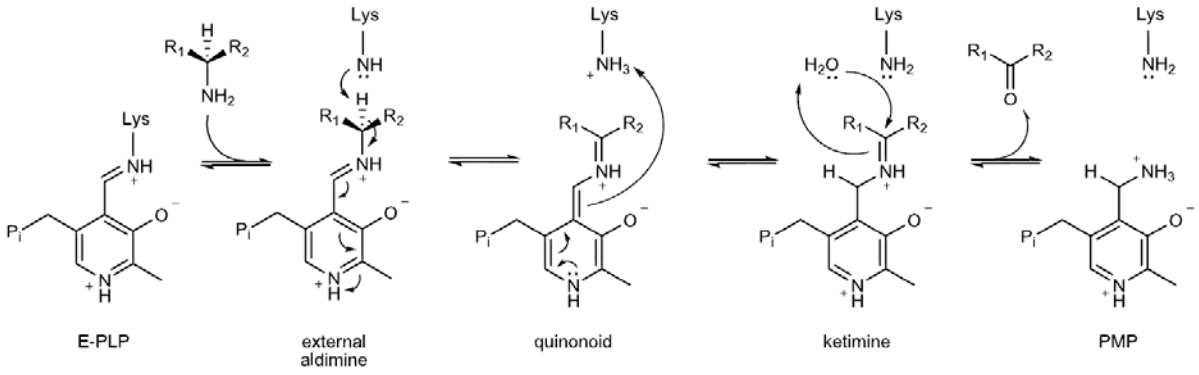


Fig. 1-5 A detailed first half reaction of the catalytic cycle.

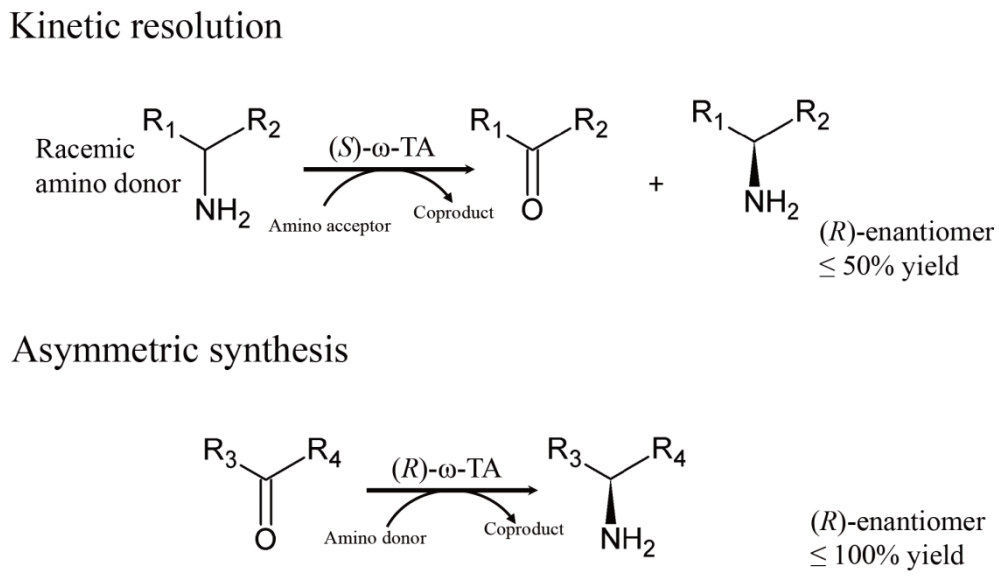
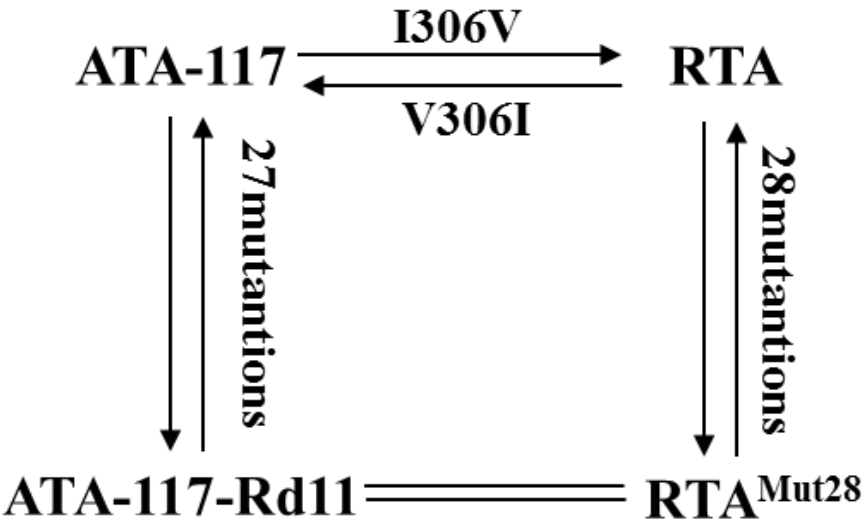


Fig. 1-6 Strategies for the synthesis of optically (*R*)-amines using ω-TAs.



S8P, Y60F, L61Y, H62T, V65A, V69T, D81G, M94I, I96L,
F122M, S124T, S126T, G136F, Y150S, V152C, A169L,
V199I, A209L, G215C, G217N, S223P, L269P, L273Y
T282S A284G P297S, S321P

Fig. 1-7 The differences and amino acid sequence comparison among (*R*)-TA, ATA-117, and ATA-117-Rd11.

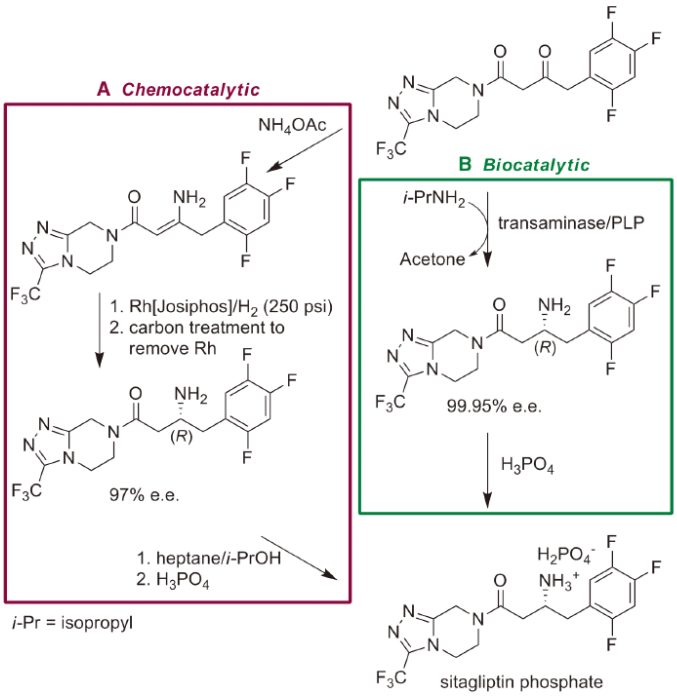


Fig.1-8 An example of successful biocatalytic application for the synthesis of (*R*)-amine [16].

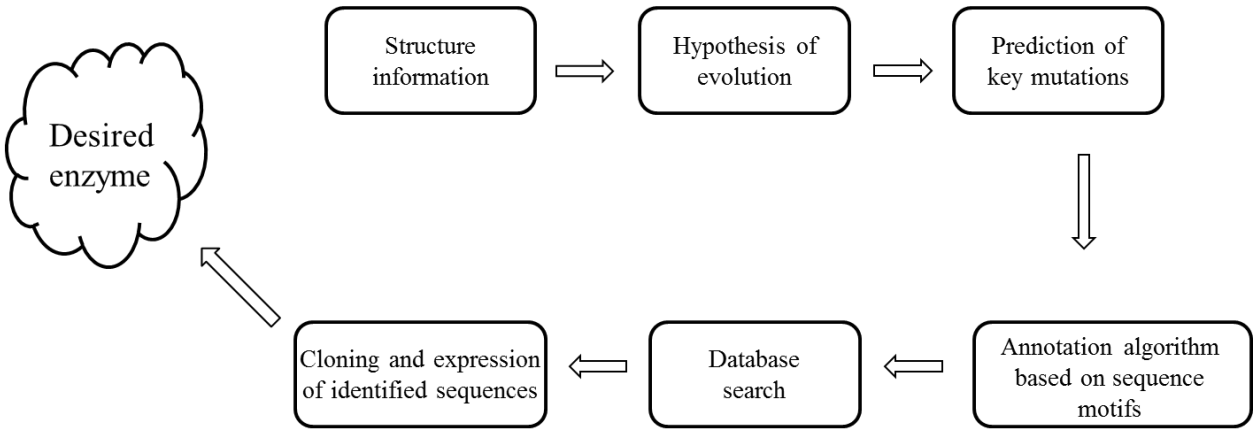


Fig. 1-9 Strategy to screen putative enzymes using *in silico* approach.

Chapter 2

Structural basis for substrate recognition and specificity of (*R*)-TA

2.1 Exordium

(*R*)-TA, identified from *Arthrobacter* sp. KNK168 (FERM-BP-5228), is a novel enzyme catalyzing the transfer of an amino group from an amino compound to a ketone compound using the cofactor PLP. In previous research, pyruvate and methoxyacetone (Fig. 2-1) were found as good amino acceptors.

In this chapter, I will discuss the expression, purification, crystallization and crystal structural analysis of (*R*)-TA. Furthermore, Structural basis for substrate recognition and specificity of (*R*)-TA was analyzed by the complex of (*R*)-TA and substrates. To investigate how the ketone compounds bind to the active site, I will describe the purification, crystallization and determination of the (*R*)-TA-pyruvate complex structure ((*R*)-TA^{PYR}) and (*R*)-TA-methoxyacetone complex structure ((*R*)-TA^{METH}). PMP is an important intermediate for catalyzing the transfer of an amino group from an amino compound to a ketone compound (Fig. 1-3). The (*R*)-TA^{PMP} structure was obtained in attempt to trap the external aldimine formed by PLP and *R*-MBA. In other words, (*R*)-TA had reacted with the donor substrate during incubation. In order to trap an external aldimine intermediate that could reveal the binding mode of the amino donor substrates, I made a K188Q mutant of (*R*)-TA and successfully used it to trap an external aldimine formed by PLP and *R*-MBA.

2.2 Materials and methods

2.2.1 Overexpression and purification

The (*R*)-TA expression plasmid constructed by inserting the (*R*)-amine transaminase gene (GenBank accession number AB638718) under the *lac* promoter of the pUCN vector [15] was used for transformation of *Escherichia coli* Rosetta (DE3) to

overexpress (*R*)-TA that contains 330 amino acids with no tag. The transformants were cultivated at 310 K in LB medium containing 100 μg ml^{-1} ampicillin. Overexpression was induced by adding 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) when the optical density at 600 nm reached 0.6, and the culture was continued 20 hours at 293 K. The cells were harvested by centrifugation at 5180 *g* at 277 K for 15 min and kept at 243K.

To purify (*R*)-TA, the cells were resuspended in 50 mL lysis-buffer containing 20 mM potassium phosphate buffer (pH 6.8), 0.01% 2-mercaptoethanol, 1 mM PLP and 10% glycerol (v/v). The cells were disrupted by sonication for 6 min using BRANSON Digital Sonifier at power 7. After centrifugation at 40000 *g* at 277 K for 30 min, the supernatant was applied onto DEAE-Sepharose, and the (*R*)-TA protein was eluted with 400 mM NaCl dissolved in 20 mM potassium phosphate buffer (pH 6.8) containing 1 mM DTT and 10% glycerol (v/v). The eluted protein with 1 mM PLP additive was dialyzed against 20 mM potassium phosphate buffer (pH 6.8) containing 1 mM DTT and 10% glycerol (v/v). The dialyzate was applied to a ResourceQ column 6 ml (GE Healthcare) equilibrated with 20 mM potassium phosphate buffer (pH 6.8) containing 1 mM DTT and 10% glycerol (v/v), and the protein was eluted with a NaCl gradient ranging from 0 to 1 M. The active fractions obtained by ResourceQ chromatography were concentrated and loaded on a Superdex 200 HR 10/30 column (GE Healthcare) equilibrated with 20 mM potassium phosphate buffer containing 1 mM DTT and 5% glycerol. The active fractions were concentrated to 8 mg ml^{-1} by centrifugation at 4000 g using Vivaspin-20 (10000 molecular weight cutoff) prior to crystallization trials.

2.2.2 Preparing of mutant plasmids

The primers were designed according to the instruction manual of the PrimeSTAR^{TM} Mutagenesis Basal Kit (TaKaRa). The reaction mixture containing the wild-type (*R*)-TA plasmid, the mutagenic primers and PrimeSTAR Max Premix was put in to a temperature cycler. The thermal cycling parameters are as follows:

98 ºC, 60 sec (denaturation for template)

98 ºC, 10 sec (denaturation)
55 ºC, 15 sec (annealing)
72 ºC, 5 sec / kb of plasmid length (extension)
} 30 cycles

After temperature cycling, *Dpn*I endonuclease was added to the PCR product to digest the parental DNA template. The reaction was performed at 37 ºC for 1 h. Then, the amplified dsDNA were transformed into the XL1-blue competent cell (Novagen). Mutations were verified by Fasmac DNA Sequence Service. The mutants were overexpressed and purified in the same manner as wild type (*R*)-TA.

2.2.3 UV-visible spectroscopy of (*R*)-TA, (*R*)-TA mixed with *R*-MBA, K188Q mutant and K188Q mutant mixed with *R*-MBA

All UV-visible spectra were recorded on a UV2450 UV-visible spectrophotometer (Shimadzu). For UV-visible assays, the concentration of recombinant protein was 100 μM. The spectrophotometer was blanked with 20 mM potassium phosphate (pH 6.8) and spectra were collected from 500 nm to 300 nm. Cells had a sample volume of 200 μL. Spectra were collected at time intervals. Changes in the UV-visible spectrum were monitored.

2.2.4 Differential scanning fluorometry

The reaction system (20 μL) contained 25 μM purified protein, 20 mM Tris buffer (pH 8.0), 1×Sypro Orange, 10% glycerol (v/v) or 200 mM pyruvate within a 96 well PCR plate (FrameStar) and was incubated on ice with each condition repeated two times each for reliability. Then, the plate was sealed with optical sealing tape and placed within a quantitative PCR machine (CFX Connect^{TM}) and heated to 20ºC for 10 minutes. The temperature was then increased by 0.5ºC each minute with a reading taken for every degree Celsius change in temperature until 95ºC.

2.2.5 Enzyme assay and analysis

Transaminase activity was assayed by measuring the production of acetophenone, as described previously with minor modifications [15]. Briefly, the reaction mixture contained 10 mM R-MBA, 10 mM pyruvate or benzylacetone, 100 mM Tris-HCl (pH 8.5), 0.5 mM PLP and the enzyme solution in 20 mM potassium phosphate buffer (pH 6.8) containing 20% (v/v) glycerol and 1 mM DTT. The reaction mixture was incubated at 30 ºC for 60 min and stopped by adding HCl. The amount of acetophenone formed was determined by HPLC. For the determination of kinetic parameters, reaction mixtures containing 1.25-1000 mM pyruvate were used.

2.2.6 Preparation for cocrystallization

(*R*)-TA for cocrystallization with ketone substrates was overexpressed and purified in the same manner as described above, but glycerol was removed from the purifying buffer since sonication. Then equal volume of 400 mM pyruvate and methoxyacetone were added to the sample. The mixed sample was further incubated on the ice for 4h. Before crystallization, the yellow mixed sample was centrifuged at 16000 g for five min

to remove precipitates.

For the crystallization of the (*R*)-TA with substrate to obtain PMP intermediate, 10 mM amino donor substrate *R*-methylbenzylamine (*R*-MBA) was added to the sample and incubated on the ice for 2h. Before crystallization, the mixed sample was centrifuged at 16000 g for 5 min to remove impurities.

For the crystallization of the K188Q mutant of (*R*)-TA with substrate, 8 mM *R*-MBA was added to the protein solution and incubated at 293K for 6 hours. Before crystallization, the mixed sample was centrifuged at 16000 g for 5min to remove impurities.

2.2.7 Crystallization

All crystallization experiments were performed at 293 K by the sitting-drop vapor-diffusion method. In the initial screening, each drop was formed by mixing 0.4 μL of protein solution with 0.4 μL of reservoir solution, and equilibrated against 40 μL reservoir solution. In optimization step, the crystallization drops were prepared by mixing 2.0 μL of protein solution and 2.0 μL reservoir solution, and equilibrated against 500 μL reservoir solution in order to obtain larger crystals in 24-well plates (Hampton Research). Streak seeding method was used to obtain high quality crystals.

2.2.8 X-ray data collection and processing

The presence of glycerol and/or PEG3350 in the crystallization condition alleviated the need for subsequent cryoprotection, and thus crystals were picked up in a nylon loop (Hampton Research) and directly flash-cooled in a nitrogen cryostream (100 K). For all structures in this chapter, X-ray diffraction data were collected at the Photon Factory (Tsukuba, Japan). The data sets were composed of 360 images and collected using a

0.5° oscillation. The diffraction data was indexed, integrated and scaled with XDSme [18].

2.2.9 Phasing, model building and refinement

Initial phasing of (*R*)-TA^{WT} diffraction data was conducted by molecular replacement using MOLREP [19] in the CCP4 [20] and with a single molecule of branched-chain amino acid aminotransferase as the search model (PDB 1IYE). Model building was performed using Buccaneer [21] and arp/warp [22]. After model reconstruction, the structure was refined using refmac5 [23] and phenix [24]. PLP, glycerol molecule and water molecules were added to the model on the basis of $2F_o$-F_c and F_o-F_c electron density maps. The final manual fitting and structure refinement was completed using Coot [25].

For (*R*)-TA^{PYR}, (*R*)-TA^{METH}, (*R*)-TA^{PMP} and (*R*)-TA^{REX}, initial phasing of these diffraction data sets were conducted by molecular replacement using MOLREP in the CCP4. The search model was the crystal structure of (*R*)-TA^{WT}. Molecular replacement was performed using a resolution range from 20 Å to 3.0 Å. The initial phase was refined using Refmac5 then examined using Coot.

2.2.10 Structure analysis

Visualization and superposition of protein structures were achieved using PyMol and the RMSD values for superimposed structures were calculated using PDBeFold server. Interactions at the dimer interface were analyzed using the Protein Interfaces, Surfaces and Assemblies (PISA) server. Hydrogen bonds within 2.5-3.9 Å were selected. The dimer interface of (*R*)-TA^{WT} was determined by ΔG^{int}, the solvation-free energy gain up on formation of assembly. The 2D (*R*)-TA^{WT}-PLP interaction diagram was

generated using LigPlot^{+} program.

2.3 Results and discussion

2.3.1 Overexpression and purification of (*R*)-TA

(*R*)-TA^{WT} was overexpressed in *E. coli* Rosetta (DE3) and purified by DEAE-Sepharose, anion-exchange chromatography (Fig. 2-2) and gel filtration (Fig. 2-3). The yellow purified sample was applied to 12.5% SDS-PAGE and was confirmed as a single band (Fig. 2-4).

2.3.2 Crystallization

We have established methods for the preparation and crystallization of (*R*)-TA. Initial screening was performed using Wizard I and Wizard II (Emerald BioSystems), Crystal Screen HT and Index HT (Hampton Research). The reservoir-solution conditions that yielded protein crystals were optimized to produce crystals that were suitable for X-ray diffraction analysis. After the crystallization conditions had been refined, crystals suitable for X-ray analysis were obtained using a reservoir solution consisting of 0.2 M magnesium chloride, 0.1 M Bis-Tris pH 6.4, 22%(w/v) PEG 3350 (Fig. 2-5).

Initial screening for crystallization of (*R*)-TA mixed with 200 mM pyruvate or methoxyacetone was performed using Crystal Screen HT (Hampton Research), Index HT (Hampton Research) and Wizard I and II (Emerald BioSystems) kits, and crystals obtained in the reservoir solution similar as that of (*R*)-TA^{WT}. The crystallization conditions were further optimized by the sitting-drop vapor diffusion method and diffraction-quality crystals were grown to 0.1×0.2×0.1 mm^3 for three days at 293K. Purified sample without glycerol and ketone substrates was also crystalized in the same

condition. However, the quality of the crystals grown without glycerol and ketone substrates was much poorer than that grown with glycerol or ketone substrates (Fig. 2-6).

After incubation with the amino donor substrate *R*-MBA, (*R*)-TA changed from yellow to colourless (Fig.2-7) while K188Q changed from colourless to yellow. This would imply that the UV-visible spectrum of the cofactor has changed. It was supposed that (*R*)-TA had reacted with the donor substrate. By streak seeding, polyhedron-shaped crystals grew in 3 days with approximate dimensions of 0.15 ×0.20 × 0.15 mm^3 under the crystallization conditions consisting of 0.2 M magnesium chloride, 0.1 M Bis-Tris pH 6.6, 16%(w/v) PEG3350. The crystals of K188Q mutant mixed with *R*-MBA shown yellow, while the crystals of K188Q mutant without *R*-MBA showed colorless (Fig. 2-8).

2.3.3 UV-visible spectroscopy

Fig. 2-9 presents absorption spectra of (*R*)-TA and (*R*)-TA mixed with *R*-MBA. The sample of (*R*)-TA had an absorption maximum at 415 nm in solution, consistent with a protonated aldimine (internal aldimine in (*R*)-TA^{WT}). The sample of (*R*)-TA mixed with *R*-MBA had an absorption maximum at 330 nm in solution, consistent with a PMP state. It indicated that this sample probably equilibrated at PMP state.

Fig. 2-10 presents absorption spectra of K188Q mutant and K188Q mutant mixed with *R*-MBA. The sample of K188Q mutant had an absorption maximum at 340 nm in solution, consistent with a free PLP non-covalently bound in the active site. The sample of K188Q mutant mixed with *R*-MBA had an absorption maximum at 420 nm in solution, consistent with an external aldimine. It indicated that this sample probably

equilibrated at external aldimine state.

2.3.4 Differential scanning fluorometry

Protein crystallization results usually depend upon the stability of the protein. Poor protein crystals may be caused by loose floppy domains or loops that interfere the protein packing into ordered crystals. Efforts should be done to optimize the crystallization conditions in which the protein become more stable and hence reduce the number of different conformations. Denaturation temperature (Tm) at which half of the protein is in the folded state and the other half in the unfolded state could be taken into consideration. A higher Tm value is indicative of a more stable protein.

Differential scanning fluorometry allows the determination of a more thermodynamically stable protein. A number of factors such as buffers, salts, pH and ligands can be tested for their effects upon the stability of the protein. Sypro orange is used as an indicator of denatured proteins when it binds to hydrophobic regions of the unfolded protein. There is an increase in fluorescence at a wavelength of 610 nm with an excitation wavelength at 492 nm when bound by the Sypro orange. And the increase could be measured by a quantitative PCR machine. It makes rapid screening of conditions for stabilizing proteins. It also could investigate whether the ligand bound to the protein or not in that the Tm value of a protein would increase when a ligand bound to it and make it more stable.

Fig. 2-11 shows the Tm values of the differential scanning fluorometry trials with (*R*)-TA. The Tm value of sample containing glycerol and pyruvate were 2 ºC and 4 ºC higher than that of control, respectively. It indicated that the pyruvate had bound to (*R*)-TA. Hence the crystals grew in the present of glycerol and ketone substrates had a

much better quality than control.

2.3.3 Data collection and structure determination

X-ray diffraction data of (*R*)-TA^{WT} was collected to 1.65 Å. Figure 2-12 (A) shows an X-ray diffraction image of the (*R*)-TA^{WT} crystal. The crystal belonged to the space group $P4_22_12$ with unit-cell parameters of a = 80.62 Å, b = 80.62 Å, c = 93.88 Å. 37,724 unique reflections in the resolution range of 45-1.65 Å were collected with 99.9% completeness. The data-collection and processing statistics are summarized in Table 2-1. Matthews coefficient analysis suggested that there was one molecule per asymmetric unit in the (*R*)-TA^{WT} crystal. The initial solution with one molecule in the asymmetric unit had starting R_{factor} and R_{free} values in excess of 50%, which only dropped little respectively after refinement with refmac5. The model was therefore improved using Buccaneer and arp/warp which was successfully used to build most of the structure. This structure was refined using refmac5 and phenix to obtain a model containing water with R_{factor} and R_{free} values of 18.0% and 22.7%, respectively. After adding the PLP, glycerol and water molecules, the model was refined to R_{factor} 15.4% and R_{free} 17.2%.

Figure 2-12 (B) and Figure 2-12 (C) show X-ray diffraction images of the (*R*)-TA^{PYR} and (*R*)-TA^{METH} crystals, respectively. The crystals of (*R*)-TA^{PYR} and (*R*)-TA^{METH} also belonged to space group $P4_22_12$. The unit cell and the other crystallographic data collection statistics are shown in Table 2-2. Matthews coefficient value indicted that these two crystals contained one molecule per asymmetric unit, respectively. The refinement statistics are given in Table 2-2. The structures of (*R*)-TA^{PYR} and (*R*)-TA^{METH} were refined to R_{work}/R_{free} of 16.7%/21.8% and 15%/18%,

respectively.

Figure 2-12 (D) shows X-ray diffraction image of the (*R*)-TA^{PMP}. The crystals of (*R*)-TA^{PMP} belonged to space group *C*2. The unit cell and the other crystallographic data collection statistics are shown in Table 2-3. The Schiff linkage between the Lys188 and PLP appeared to be broken in the electron density (Fig. 2-13), which is consistent with the results of UV-visible spectroscopy. Thus, PMP was added in to the corresponding electron density. The refinement statistics are given in Table 2-3.

Figure 2-12 (E) shows X-ray diffraction image of the (*R*)-TA^{REX}. The crystals of (*R*)-TA^{REX} belonged to space group $P4_22_12$. The unit cell and the other crystallographic data collection statistics are shown in Table 2-4. The Schiff linkage between the PLP and *R*-MBA appeared in the electron density (Fig. 2-14), which is consistent with the results of UV-visible spectroscopy. Thus, external aldimine formed by PLP and *R*-MBA was added in to the corresponding electron density. The refinement statistics are given in Table 2-4.

2.3.4 Overall structure of (*R*)-TA^{WT}

The crystal structure of (*R*)-TA^{WT} was determined to 1.65 Å. The overall structure of (*R*)-TA showed a typical structural architecture of the fold-type IV class of PLP-dependent enzymes, which is constituted by a large and a small domain and the linker loop connecting these two domains. The small domain consists of six β-strands (β1 to β6) and three α-helices (α1 to α3) while the large domain is composed of nine β-strands (β7 to β15) and five α-helices (α4 to α8) (Fig.2-15). The fold-type IV class contains (*S*)-stereospecific aminotransferases which show similarity with (*R*)-TA, such as L-branched-chain amino acid aminotransferase from *Escherichia coli* (ex. 1IYE,

Z-score 36.6, r.m.s.d 1.7 Å, id 25%). (*R*)-TA also show similarity with the structure-known (*R*)-stereospecific aminotransferases belonging to the same fold class, such as D-amino acid aminotransferase from *Bacillus* sp. YM-1 (ex. 1A0G, Z-score 32.6, r.m.s.d 1.9 Å, id 28%).

2.3.5 Dimer interface of (*R*)-TA^{WT} and interaction between two protomers

The (*R*)-TA^{WT} molecules in the crystal structures forms a dimer in consistent with the results of size-exclusion chromatography (Fig. 2-16). The asymmetric unit of the crystal of (*R*)-TA^{WT} contained one molecule of the protein, which forms a dimer with the symmetry-related copy of itself. PISA server detected a lot of interactions among the dimer interface with the contact area of 2617.4 Å^2, the 17.6% of the total surface area of (*R*)-TA^{WT} (Table 2-5).

2.3.6 PLP and glycerol binding to (*R*)-TA^{WT}

The active-site cleft is formed by the two domains of a protomer and the loop consisting of residues 129-146 in the small domain of the other protomer (Fig.2-16). In the crystal structure of (*R*)-TA^{WT}, the cofactor PLP was in the internal aldimine form, in which PLP covalently bound to Lys188 forming a Schiff base (Fig. 2-17). Besides the carbon-nitrogen bond of the Schiff base, PLP was anchored to the active site pocket by a number of interactions with the amino acid residues, and no interaction was observed between PLP and the loop consisting of residue 129–146 (Fig. 2-18). The oxygen atom O1P of PLP was anchored by the main-chain amide and oxygen of Thr283 residue, while O2P and O3P were kept in place by the main-chain amide and oxygen of Thr247 residue and the side chain of Arg86 residue, respectively (Fig. 2-18). These three oxygen atoms of PLP were also anchored by hydrogen bond with three water molecules

which formed hydrogen bonds with Asn189, Trp192, Tyr67, Val69, Arg86, Asn226, Arg248, and Cys281 (Fig. 2-18). In addition, the PLP pyridine is sandwiched between Leu243 and Gly224/Phe225 which hinder any vertical movements of the pyridine ring plane. More importantly, Glu221 residue forms an electrostatic interaction with the pyridine nitrogen atom of PLP, causing stabilization of its positive charge and increasing the electrophilic character of the cofactor [26].

In the active site of (*R*)-TA^{WT}, a glycerol molecule forming hydrophilic interaction with Arg138 was also trapped at the position where the ketone, such as pyruvate, was expected to be bound (Fig. 2-19).

2.3.7 Binding mode of pyruvate and methoxyacetone

Figures 2-20 and 2-21 show the active sites of (*R*)-TA^{PYR} and (*R*)-TA^{METH}, respectively. The F_o-F_c omit electron density maps of pyruvate and methoxyacetone are shown in blue mesh contoured at 1.5 σ resolution. The two ketones bound noncovalently in the active site. In the crystal structure of (*R*)-TA^{PYR}, the α-carboxylate of pyruvate has a salt bridge interaction with the Arg138 of the other protomer and a hydrogen binding interaction to Try67, respectively. The keto oxygen atom of pyruvate forms a hydrogen bond with the Schiff linkage. In addition to these polar interactions, the methyl group of pyruvate is bound by hydrophobic interactions with Thr282 and Ala284. In the crystal structure of (*R*)-TA^{METH}, the ether oxygen atom of methoxyacetone forms a hydrogen bond with Tyr67 and Arg138 of the other protomer. The keto oxygen atom of methoxyacetone also forms a hydrogen bond with the Schiff linkage. Comparing to the methyl group of pyruvate bound in (*R*)-TA^{PYR}, the methyl group of methoxyacetone in (*R*)-TA^{METH} is bound by hydrophobic interactions with

Thr282, Thr283 and Ala284, nearly in the same binding site.

2.3.8 Activity assay of (*R*)-TA mutants

In the active site of (*R*)-TA^{PYR} and (*R*)-TA^{METH}, the keto oxygen atom of pyruvate and the ether oxygen atom of methoxyacetone have interaction with Arg138 of the other protomer. To verify the importance of Arg138 residue, R138G and R138A mutants were prepared and the activity measurements of these two mutants were performed (Table 2-6).

Both the two mutants have decreased transamination activity towards pyruvate and benzylacetone, but transamination activity towards benzylacetone decreased further drastically. It suggested the significance of Arg138 residue for the recognition of the substrate with a carboxylic acid group. This hypothesis is supported by the K_m values for pyruvate. R138Q and R138A mutants showed 51.2 and 70.9 mM, respectively, whereas the wild type showed 3.2 mM (Table 2-6). The lower K_m of R138Q than that of R138A may indicate a weak interaction between Gln138 and pyruvate. Additionally, the ratio of the specific activity for benzylacetone to that for pyruvate (B/P) was changed by the mutation of Arg138. The B/P of R138A and R138Q were higher than that of the wild type, indicating that the substrate preference was shifted from pyruvate to benzylacetone due to the mutations. R138A showed an even higher preference to benzylacetone probably because the side chain of Ala138 made extra space to accommodate benzylacetone.

2.3.9 Binding mode of PMP

(*R*)-TA^{WT} structure can be aligned with a subunit of (*R*)-TA^{PMP} with an r.m.s.d. of 0.238 Å for 321 Cα atoms. The N-termini was much better defined in (*R*)-TA^{PMP}

structure. No significant structural differences were observed in the active sites of the two structures except for a small displacement of the Lys188 residue and the cofactors. The PMP molecule non-covalently bound in the active site also forms an extensive network of hydrogen bonds in the same way as (*R*)-TA^{WT} (Fig. 2-22). A rotation of the cofactor pyridine ring of several degrees was observed between the internal aldimine of (*R*)-TA^{WT} and PMP-bound forms of (*R*)-TA^{PMP} (Fig. 2-23). The side chain of the conserved catalytic Lys188 residue is poised at a distance of 3.0 Å from the PMP amine nitrogen atom, a distance could catalyze the second half transamination (Fig. 2-24). The alignment of (*R*)-TA^{WT}, (*R*)-TA^{PYR}, (*R*)-TA^{METH} and (*R*)-TA^{PMP} structures showed that the keto-oxygen of pyruvate and methoxyacetone were close to the amine nitrogen atom of PMP (2.2 Å) and Lys188 (3.3 Å), respectively, in (*R*)-TA^{PMP} structure (Fig. 2-24).

2.3.10 External aldimine binding manner in (*R*)-TA^{REX}

In the crystal structure of (*R*)-TA^{REX}, the cofactor PLP was in the external aldimine form with the substrate *R*-MBA. The PLP portion of the external aldimine rotated approximately 25° relative to PLP molecules of (*R*)-TA^{WT} (Fig. 2-25). *R*-MBA moiety bound covalently to the C4A atom of PLP cofactor by its amino group, replacing the internal Schiff-base linkage. The methyl group of the *R*-MBA moiety was bound by Val69, Thr283 and Aal284. The aromatic ring was bound by the main chain of Gly224 and the side chains of Phe122, Trp192 and His62 of the other protomer via van der Waals interactions. The aromatic ring-binding pocket was covered by Arg138, forming a large space to accommodate the volumetric atomic group of *R*-MBA moiety (Fig. 2-26).

2.3.11 The Large and Small binding pockets

The crystal structure of (*R*)-TA^{REX} confirmed a hypothesis previously suggested that

the active site of ω-amine transminase consists of a small binding pocket and a large binding pocket, which determine the substrate specificity and stereoselectivity of ω-amine tranaminases. The small binding pocket consists of Val69, Thr283, and Ala284, while the large binding pocket consists of His62, Phe122, Arg138, Trp192, and Gly224, and, among these residues, His62 and Arg138 are from another protomer (Figs. 2-27, 2-28, and 2-29).

2.3.12 Substrate recognition

Up to now studies on the substrate structure-activity relationship about ω-amine transaminases have suggested some significant findings. Firstly, the most important property of the small pocket is the steric constraint and electrostatic repulsion between α-carboxylate group of the substrates and amino acid residues [8]. In addition, the active site residues forming the small pocket have much more influence on the substrate binding than catalysis [27]. Based on the crystal structure of (*R*)-TA^{REX}, the methyl group of *R*-MBA is directed away from the solvent side and indeed positioned in a small pocket filled with hydrophobic interactions. The size of this small pocket permits just an entry of a substituent up to a methyl group and could not accommodate an ethyl group, which could well structurally explain why (*R*)-TA has low activities towards with ethyl ketone [15].

The dual mode for the recognition of both hydrophobic and carboxylate groups in the large substrate-binding site is a great findings on the large pocket of ω-amine transaminases [8]. The dual substrate recognition mechanism of (*R*)-TA at the same or roughly the same site could be explained by the existence of Arg138 residue from the other protomer of the dimer in the large pocket. The comparison of (*R*)-TA^{WT} and

(*R*)-TA^{REX} showed that binding of *R*-MBA made the guanidinium group of Arg138 moving away from the active site (Fig. 2-30). It indicated that in order to recognize and bind the amino donors without carboxylic acid group like *R*-MBA, Arg138 moves away to provide larger space and uncharged environment. On the other hand, when (*R*)-TA recognize and bind the amino acceptor with carboxylic acid group like pyruvate, the guanidinium group of Arg138 is proposed to point to the active site (Fig. 2-20). The activities of the mutants of R138Q and R138A towards pyruvate verified this hypothesis. In addition, (*R*)-TA showed maximal activity towards pyruvate under weak alkaline conditions, and the optimum pH is in the range between 8.0 and 9.0 [15], suggesting that the deprotonation of pyruvate and coordinating Arg138 residue are probably important for the activity. Similar role of corresponding arginine residue in (*S*)-ω-TA was also proposed by docking simulations with different substrates [28].

Structural differences between (*R*)-TA^{WT} and (*R*)-TA^{REX} were localized to the active site. Besides the rotational difference of the pyridine ring of the PLP, the side-chain conformations of the residues of the other protomer that compose the binding pocket were significantly different. In (*R*)-TA^{REX} the side chain of His62 rotated a little to interact with the phenyl group of *R*-MBA and the Arg138 residue was oriented away from *R*-MBA resulting in a larger binding pocket and the entrance to the active site, when compared to (*R*)-TA^{WT} (Figs. 2-30 and 2-31).

Table 2-1 Data collection and refinement statistics of (*R*)-TA^{WT}

	(*R*)-TA^{WT}
Data collection	
Beamline	PF BL5A
Wavelength (Å)	1.0000
Space group	$P\,4_22_12$
Cell dimensions	80.62, 80.62, 93.88
Cell angles (°)	90.00, 90.00, 90.00
Resolution (Å)	45.0-1.65
Number of observations	485000
Number of unique reflections	37724
Completeness (%)	99.9 (98.4)
R_{merge} (%)	6.1 (70.1)
Redundancy	12.86
I/σ (I)	31.93 (2.53)
Refinement	
Resolution range (Å)	40.3-1.65
R_{work}/R_{free} (%)	15.4/17.2
Number of non-hydrogen atoms	
Protein	2566
Water	324
Ligand	PLP, glycerol
	21
RMSD bond length (Å)	0.007
RMSD bond angle (º)	1.209
Ramachandran plot (%)	
Preferred	97.6
Allowed	2.4
Outliers	0

The values in parentheses are for the highest resolution shell

$R_{work}=\Sigma||Fo|-|Fc||\ /\Sigma|Fo|$, where $|Fo|$ and $|Fc|$ are the observed and calculated structure factor amplitudes of a particular reflection.

R_{free} was calculated from 5% of reflections omitted from the refinement.

Table. 2-2 Data collection and refinement statistics of (*R*)-TA^{PYR} and (*R*)-TA^{METH}

	(*R*)-TA^{PYR}	**(*R*)-TA^{METH}**
Data collection		
Beamline	AR-NE3A	AR-NW12A
Wavelength (Å)	1.0000	1.0000
Space group	$P\,4_22_12$	$P\,4_22_12$
Cell dimensions	81.04, 81.04, 94.17	80.63, 80.63, 93.99
Cell angles (°)	90.00, 90.00, 90.00	90.00, 90.00, 90.00
Resolution (Å)	50.0-1.81	50.0-1.92
Number of observations	395256	350033
Number of unique reflections	28920	24481
Completeness (%)	99.0 (97.0)	99.9 (99.6)
R_{merge} (%)	6.3 (23.7)	7.8 (45.0)
Redundancy	13.67	14.30
I/σ (I)	30.23 (9.34)	31.83 (7.03)
Refinement		
Resolution range (Å)	40.7-1.81	46.8-1.72
R_{work}/R_{free} (%)	16.7/21.8	15.0/18.0
Number of non-hydrogen atoms		
Protein	2531	2585
Water	429	435
Ligand	PLP, pyruvate	PLP, methoxyacetone
	21	24
RMSD bond length (Å)	0.019	0.009
RMSD bond angle (°)	1.839	1.252
Ramachandran plot (%)		
Preferred	98.1	98.1
Allowed	1.9	1.9
Outliers	0.0	0.0

The values in parentheses are for the highest resolution shell

$R_{work}=\Sigma||Fo|-|Fc||\ /\Sigma|Fo|$, where $|Fo|$ and $|Fc|$ are the observed and calculated structure factor amplitudes of a particular reflection.

R_{free} was calculated from 5% of reflections omitted from the refinement.

Table 2-3 Data collection and refinement statistics of (*R*)-TA^{PMP}

	(*R*)-TA^{PMP}
Data collection	
Beamline	PF BL1A
Wavelength (Å)	1.0000
Space group	*C*2
Cell dimensions	114.99, 108.28, 90.36
Cell angles (°)	90.00, 92.89, 90.00
Resolution (Å)	45.0-1.80
Number of observations	379532
Number of unique reflections	101853
Completeness (%)	99.7 (99.8)
R_{merge} (%)	10.2 (66.0)
Redundancy	3.73
I/σ (I)	8.85 (2.46)
Refinement	
Resolution range (Å)	39.7-1.71
R_{work}/R_{free} (%)	18.9/21.6
Number of non-hydrogen atoms	
Protein	10081
Water	307
Ligand	PMP
	64
RMSD bond length (Å)	0.02
RMSD bond angle (º)	1.80
Ramachandran plot (%)	
Preferred	98.0
Allowed	2.0
Outliers	0.0

The values in parentheses are for the highest resolution shell

$R_{work}=\Sigma||Fo|-|Fc||/\Sigma|Fo|$, where $|Fo|$ and $|Fc|$ are the observed and calculated structure factor amplitudes of a particular reflection.

R_{free} was calculated from 5% of reflections omitted from the refinement.

Table 2-4 Data collection and refinement statistics of (*R*)-TA^{REX}

	(*R*)-TA^{REX}
Data collection	
Beamline	AR-NE3A
Wavelength (Å)	1.0000
Space group	$P\,4_22_12$
Cell dimensions	80.74, 80.74, 93.51
Cell angles (°)	90.00, 90.00, 90.00
Resolution (Å)	50.0-1.72
Number of observations	474030
Number of unique reflections	33239
Completeness (%)	99.8 (98.8)
R_{merge} (%)	5.0 (30.5)
Redundancy	14.26
I/σ (I)	42.02 (9.41)
Refinement	
Resolution range (Å)	46.8-1.72
R_{work}/R_{free} (%)	15.0/18.0
Number of non-hydrogen atoms	
Protein	2585
Water	435
Ligand	external aldimine
	24
RMSD bond length (Å)	0.009
RMSD bond angle (º)	1.252
Ramachandran plot (%)	
Preferred	98.6
Allowed	1.4
Outliers	0.0

The values in parentheses are for the highest resolution shell

$R_{work}=\Sigma||Fo|-|Fc||\ /\Sigma|Fo|$, where $|Fo|$ and $|Fc|$ are the observed and calculated structure factor amplitudes of a particular reflection.

R_{free} was calculated from 5% of reflections omitted from the refinement.

Table 2-5. Interface of (*R*)-TA^{WT} dimer

chainA	chainB	Distance (Å)
Hydrogen bonds		
Arg52/N$^{\eta1}$	Glu42/O$^{\varepsilon2}$	3.23
Arg52/N$^{\eta2}$	Glu42/O$^{\varepsilon2}$	3.34
Ile55/N	Ala51/O	2.97
Arg52/N$^{\eta2}$	Arg52/O	2.86
Ile53/N	Ile53/O	2.84
Ser126/O$^{\gamma}$	Tyr60/O$^{\eta}$	2.74
Trp192/N	His62/O	3.69
Gln200/N$^{\varepsilon2}$	Met94/O	3.22
Tyr60/O$^{\eta}$	Ser126/O$^{\gamma}$	2.74
Arg138/N$^{\eta1}$	Gln155/O$^{\varepsilon1}$	2.88
Ser182/O$^{\gamma}$	Arg176/O	2.76
Arg197/N$^{\eta1}$	Arg180/O	2.82
Arg176/Nε	Ser181/O	2.81
Ser63/N	Phe190/O	3.11
Trp192/N	Gln191/O$^{\varepsilon1}$	3.35
Gly193/N	Gln191/O$^{\varepsilon1}$	2.86
Arg95/N$^{\eta1}$	Asp204/O$^{\delta1}$	2.95
Arg95/N$^{\eta2}$	Asp204/O$^{\delta2}$	2.76
Glu42/O$^{\varepsilon2}$	Arg52/N$^{\eta1}$	3.23
Glu42/O$^{\varepsilon2}$	Arg52/N$^{\eta2}$	3.34
Ala51/O	Ile55/N	2.97
Arg52/O	Arg52/N$^{\eta2}$	2.86
Ile53/O	Ile53/N	2.84
His62/O	Trp192/N	3.69
Met94/O	Gln200/N$^{\varepsilon2}$	3.22
Gln155/O$^{\varepsilon1}$	Arg138/N$^{\eta1}$	2.88
Arg176/O	Ser182/O$^{\gamma}$	2.76
Arg180/O	Arg197/N$^{\eta1}$	2.82
Ser181/O	Arg197/N	2.81
Phe190/O	Ser63/N	3.11
Gln191/ O$^{\varepsilon1}$	Trp192/N	3.35
Gln191/ O$^{\varepsilon1}$	Gly193/N	2.86

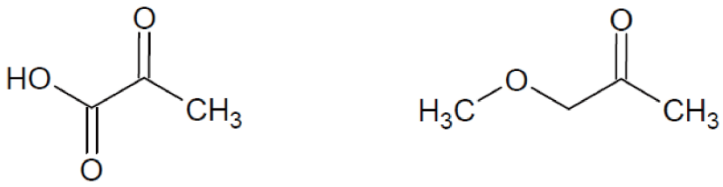
Asp204/$O^{\delta1}$	Arg95/$N^{\eta1}$	2.95
Asp204/$O^{\delta2}$	Arg95/$N^{\eta2}$	2.76
Salt bridges		
Arg52/$N^{\eta1}$	Glu42/$O^{\varepsilon2}$	3.23
Arg52/$N^{\eta2}$	Glu42/$O^{\varepsilon2}$	3.34
Arg95/$N^{\eta1}$	Asp204/$O^{\delta1}$	2.95
Arg95/$N^{\eta2}$	Asp204/$O^{\delta1}$	3.70
Arg95/$N^{\eta1}$	Asp204/$O^{\delta2}$	3.55
Arg95/$N^{\eta2}$	Asp204/$O^{\delta2}$	2.76
Glu42/$O^{\varepsilon2}$	Arg52/$N^{\eta1}$	3.23
Glu42/$O^{\varepsilon2}$	Arg52/$N^{\eta2}$	3.34
Asp204/$O^{\delta1}$	Arg95/$N^{\eta2}$	3.70
Asp204/$O^{\delta1}$	Arg95/$N^{\eta1}$	2.95
Asp204/$O^{\delta2}$	Arg95/$N^{\eta2}$	2.76
Asp204/$O^{\delta2}$	Arg95/$N^{\eta1}$	3.55

Table. 2-6 Kinetic parameters of (*R*)-TA and mutants

	Specific activity for Pyruvate (P) $U^{(a)}$/mg	Specific activity for Benzylacetone(B) U/mg	B/P[(b)]	K_m for Pyruvate (mM)	V_{max} for Pyruvate (U/mg)
WT	11.01	0.34	3.1%	3.2	15.4
R138Q	0.74	0.07	9.5%	51.2	4.9
R138A	0.83	0.17	20.5%	70.9	9.8

(a) One unit of (*R*)-TA^{WT} and mutants was defined as the amount of enzyme producing 1 μmol of D-alanine or (*R*)-amine per minute.

(b) The ratio of specific activity of benzylacetone to that of pyruvate.



pyruvate methoxyacetone

Fig.2-1 Structural formulae of pyruvate and methoxyacetone.

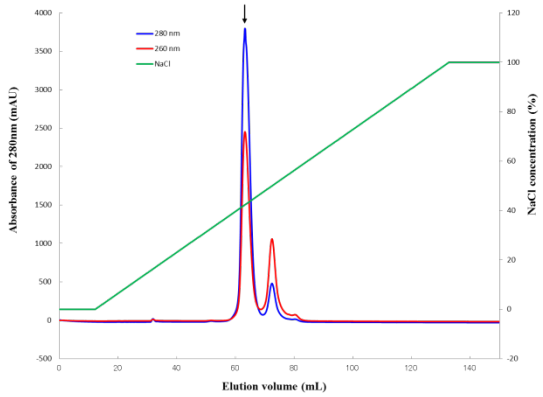


Fig. 2-2 Purification of (*R*)-TA^{WT} by anion-exchange chromatography (ResourseQ). Peak with arrow represents (*R*)-TA^{WT} which eluted at NaCl concentration of 0.35 M.

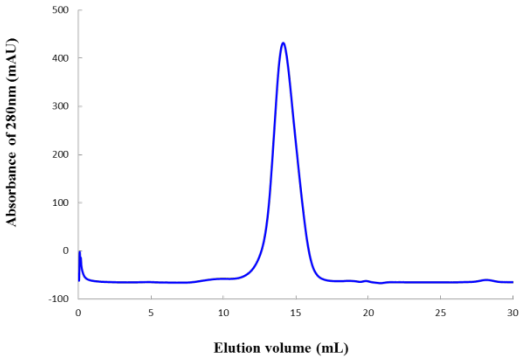


Fig. 2-3 Purification of (*R*)-TA^{WT} by size-exclusion chromatography.

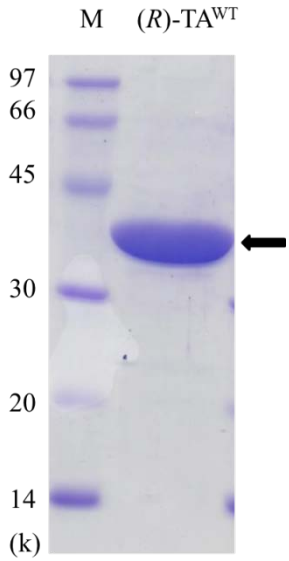


Fig. 2-4 SDS-PAGE of purified (*R*)-TA. Molecular-mass markers are labeled in k. The band of (*R*)-TA is indicated by an arrow.

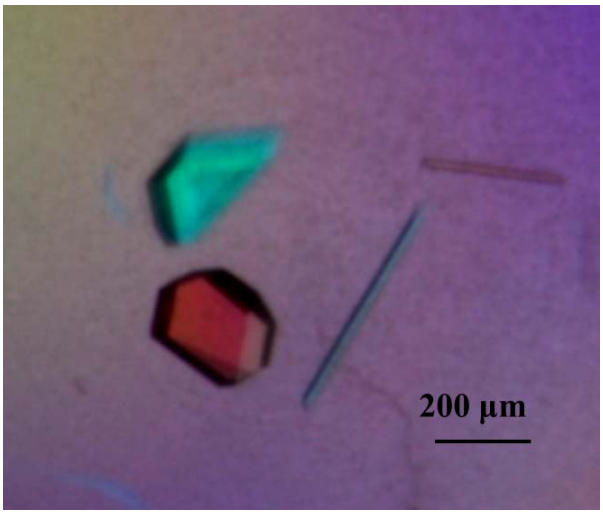
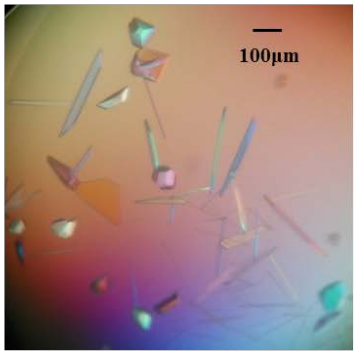
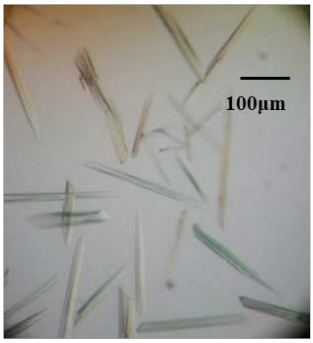


Fig. 2-5 Diffraction-quality crystal of (*R*)-TA^{WT} crystallized in the reservoir solution containing 0.2 M magnesium chloride, 0.1 M Bis-Tris pH 6.4, 22%(w/v) PEG 3350.

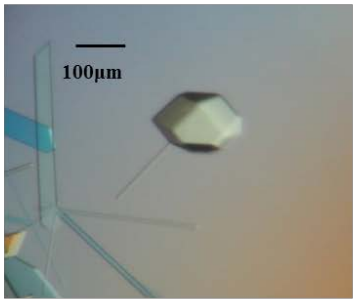
Sample with glycerol



Sample without glycerol



Sample without glycerol + 200 mM pyruvate



Sample without glycerol + 200 mM methoxyacetone

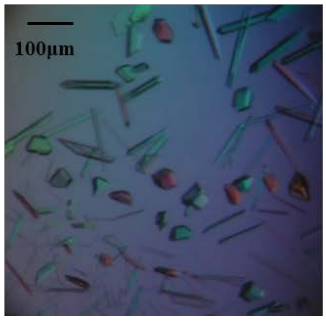


Fig. 2-6 Crystals grew with different samples in the similar crystallization condition.

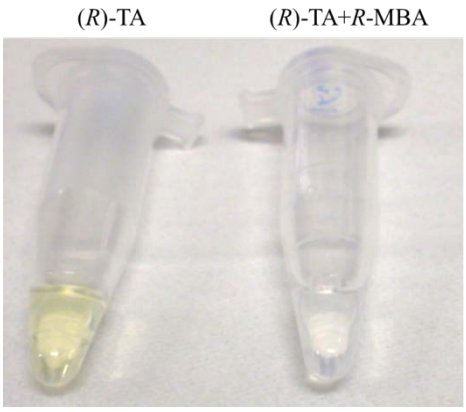


Fig. 2-7 The color of sample changed after incubation with the amino donor substrate *R*-MBA.

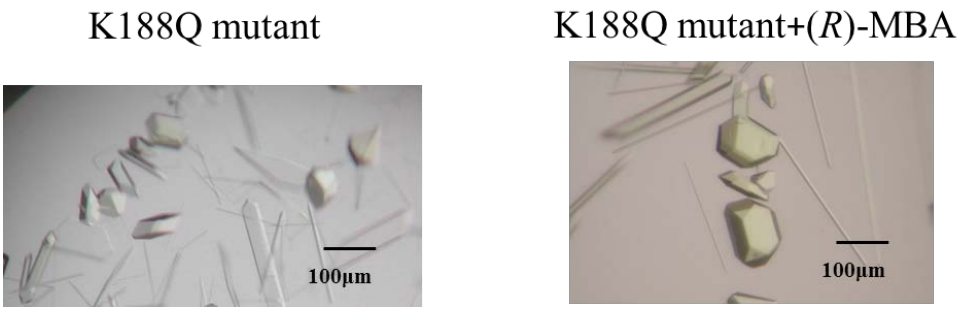


Fig. 2-8 Crystal pictures were took without a polarizing plate. The crystals of K188Q mutant mixed with *R*-MBA also shown yellow while the crystals of K188Q mutant shown colorless.

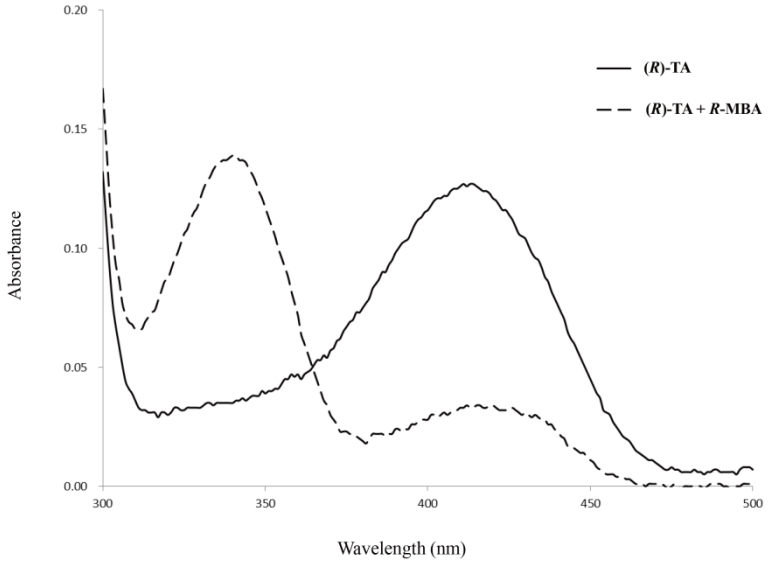


Fig. 2-9 UV-visual spectrum of (*R*)-TA and (*R*)-TA mixed with *R*-MBA showed typical peaks at 420 nm and 330 nm due to internal aldimine and cofactor PMP, respectively.

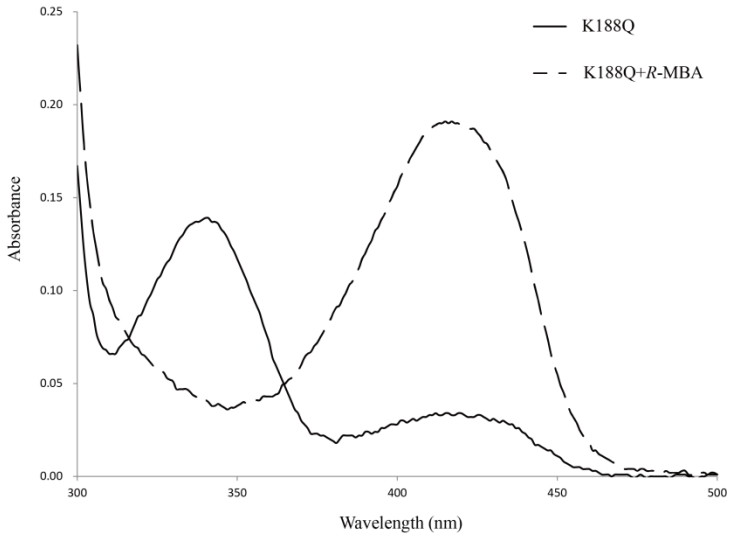
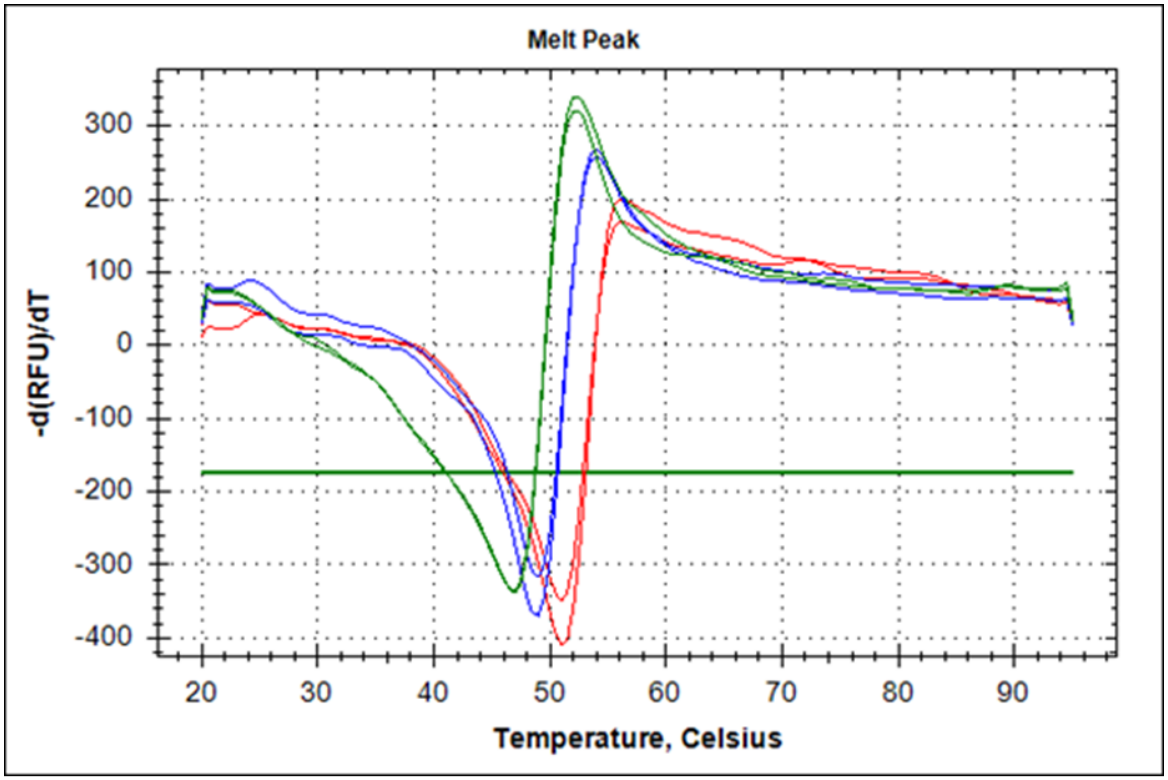


Fig. 2-10 UV-visual spectrum of K188Q mutant and K188Q mutant mixed with *R*-MBA showed typical peaks at 340 nm and 420 nm due to non-covalently bound PLP and external aldimine formed by PLP and *R*-MBA, respectively.



Green control

Blue 10% glycerol

Red 200 mM pyruvate

Sample	Melt Temp
cotrol	47.00
cotrol	47.00
10%glycerol	49.00
10%glycerol	49.00
200mM pyruvate	51.00
200mM pyruvate	51.00

Fig. 2-11 Melt temperature increased when glycerol or pyruvate was added to the sample. The figure shows the relationship between the negative derivative (–dRFU/dT) of the melting curves and temperature.

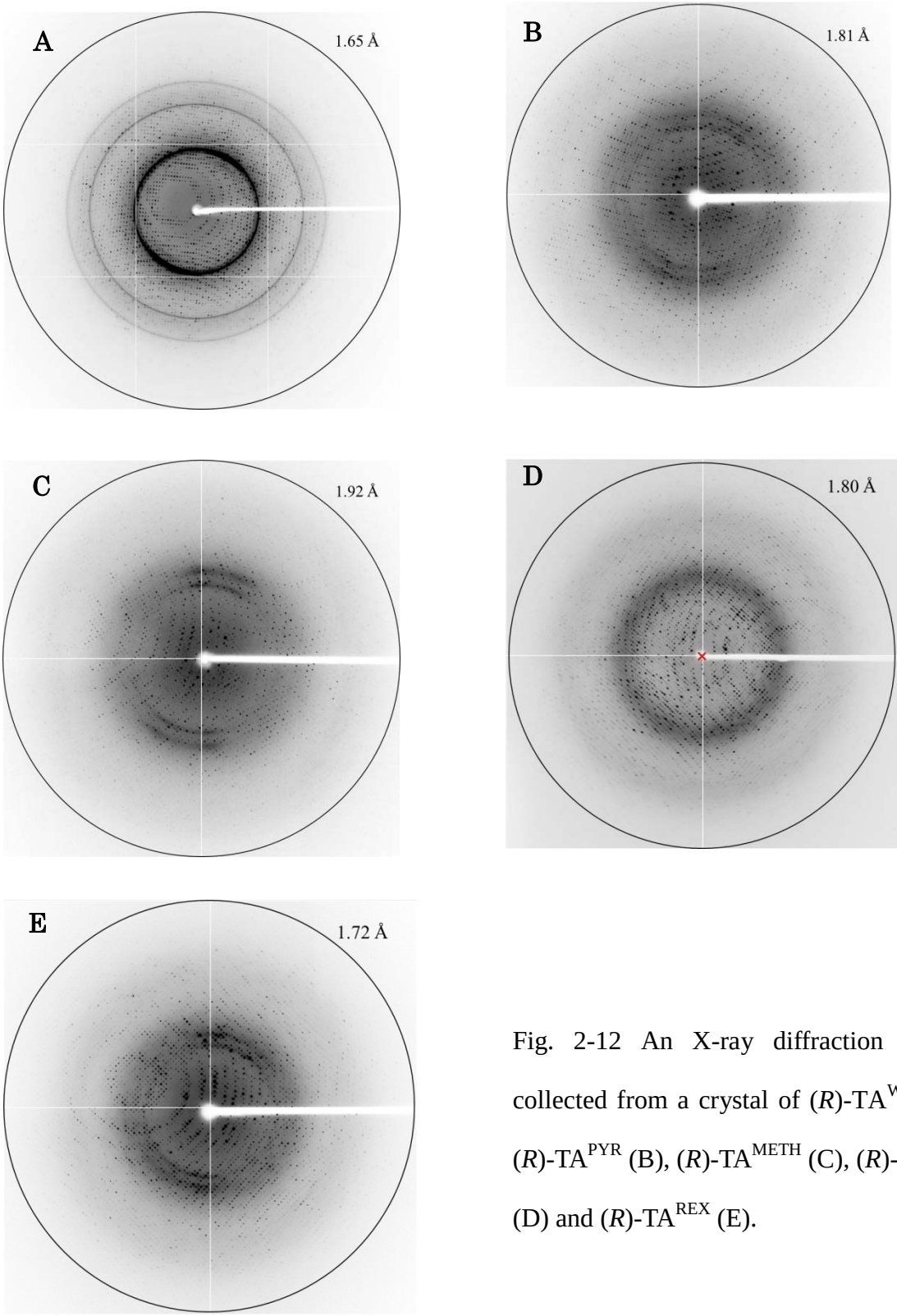


Fig. 2-12 An X-ray diffraction image collected from a crystal of (*R*)-TA^{WT} (A), (*R*)-TA^{PYR} (B), (*R*)-TA^{METH} (C), (*R*)-TA^{PMP} (D) and (*R*)-TA^{REX} (E).

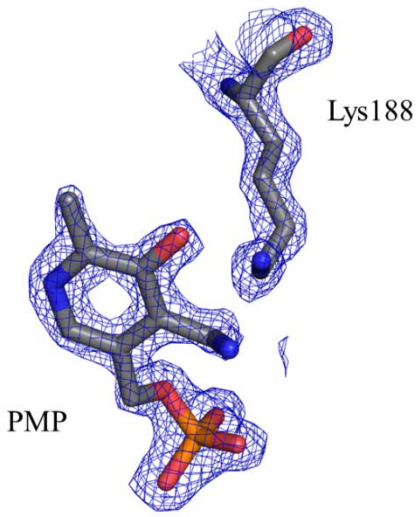


Fig. 2-13 F_o-F_c omit density map of Lys188 and PMP, contoured at 2.0 σ.

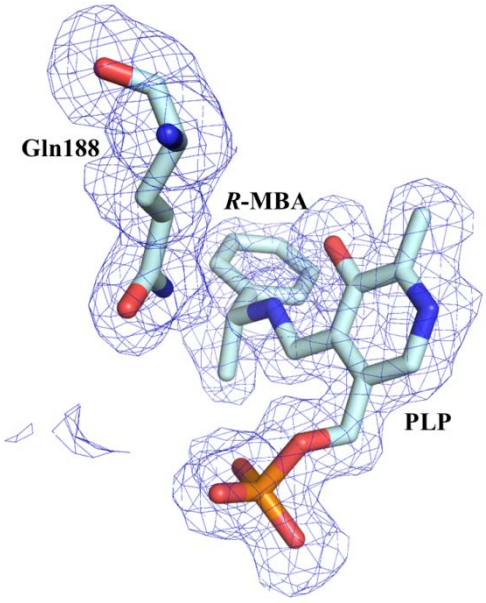


Fig. 2-14 F_o-F_c omit map of Gln188 and the external aldimine formed by PLP and *R*-MBA, contoured at 2.0 σ.

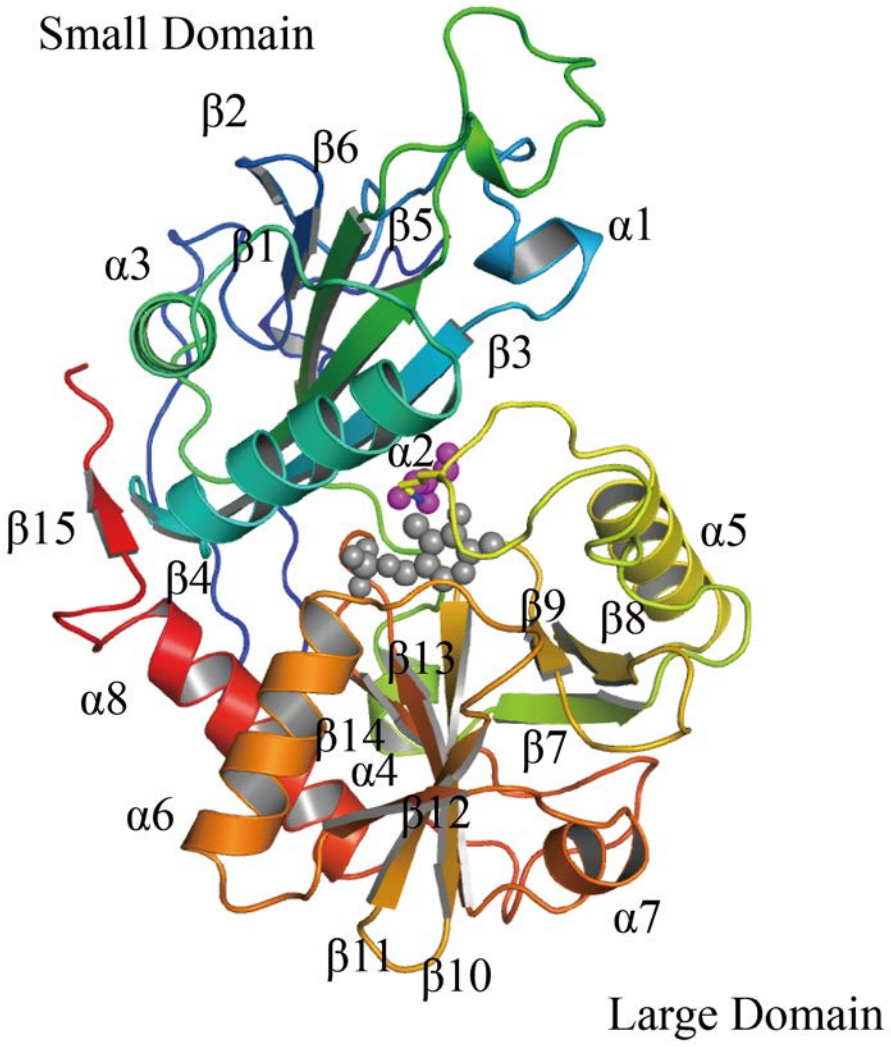


Fig. 2-15 The (*R*)-TA^{WT} monomer is colored in rainbow representation from the N-terminus in blue to the C-terminus in red. Secondary structure elements are labeled. PLP (grey) and glycerol (magenta) among the large domain and small domain are shown as spheres.

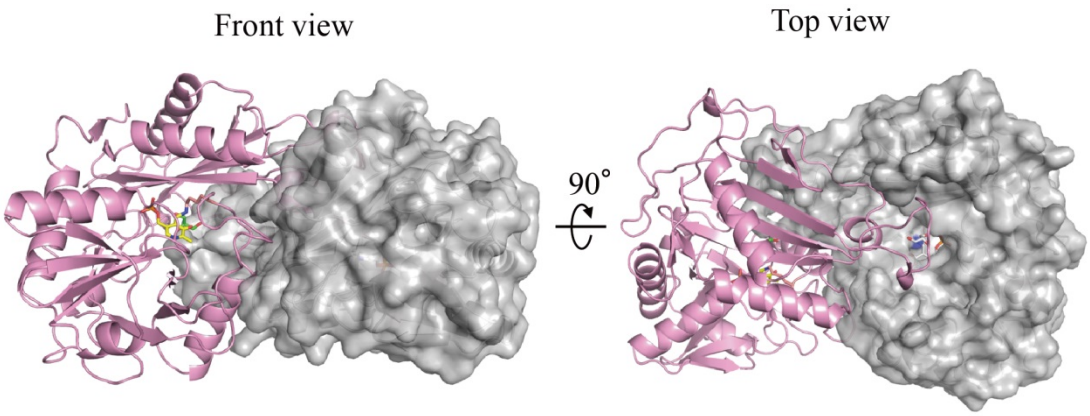


Fig. 2-16 The (*R*)-TA^{WT} dimer, with subunits shown as cartoon in pink and surface in grey. The two active sites are located in the dimer interface and are indicated by the bound PLP cofactor and glycerol in each site shown as stick. Front view is rotated 90° to top view to better show the dimer.

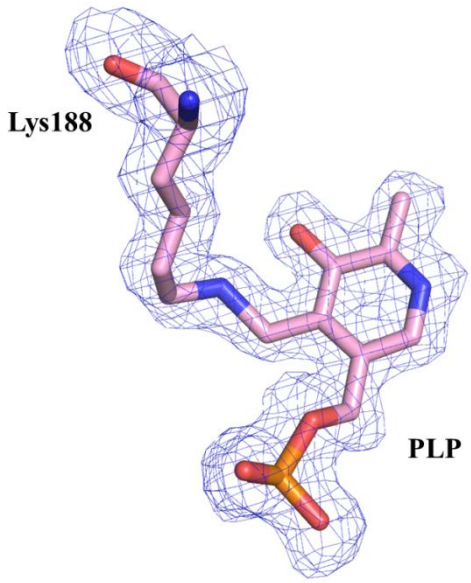


Fig. 2-17 The F_o-F_c omit map of cofactor PLP and residue Lys188 is contoured at a 2σ.

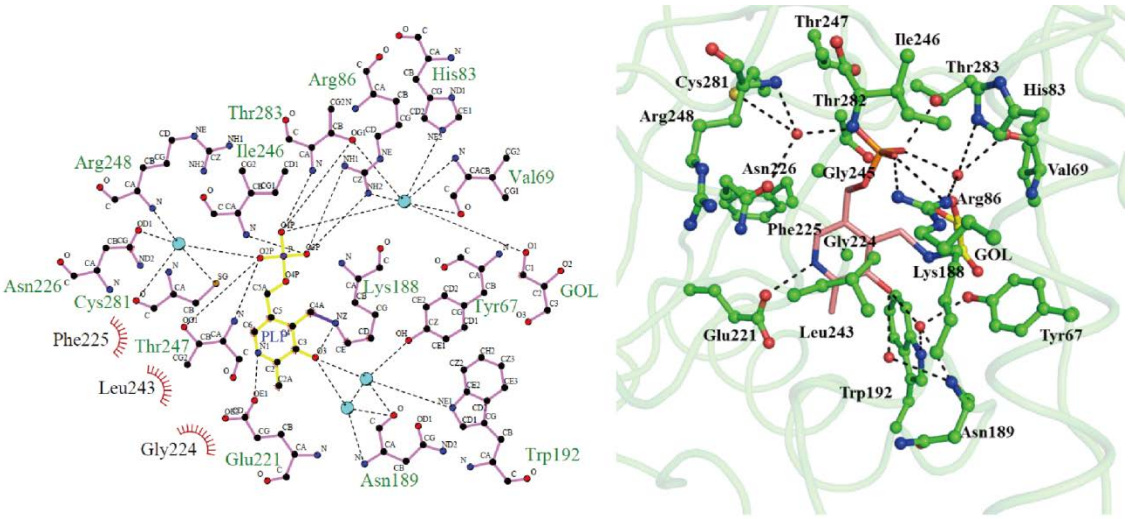


Fig. 2-18 Detailed structural representation of the interactions between PLP and (*R*)-TA^{WT}. (A) is LIGPLOT representation of the PLP and (*R*)-TA^{WT}. Hydrophilic interacting atoms are connected by black dashed lines; nonligand residues involved in direct hydrophobic contacts with PLP are shown as red semicircles with radiating spokes. The Schiff base between PLP and Lys188 residue of (*R*)-TA^{WT} are shown in purple. Water molecules are shown as cyan circles. (B) There-dimensional residue interaction map illustrating the amino acids within 4 Å of PLP. PLP is colored as pink carbon sticks. Residues within the binding pocket are colored with green carbons in ball-and-stick representation.

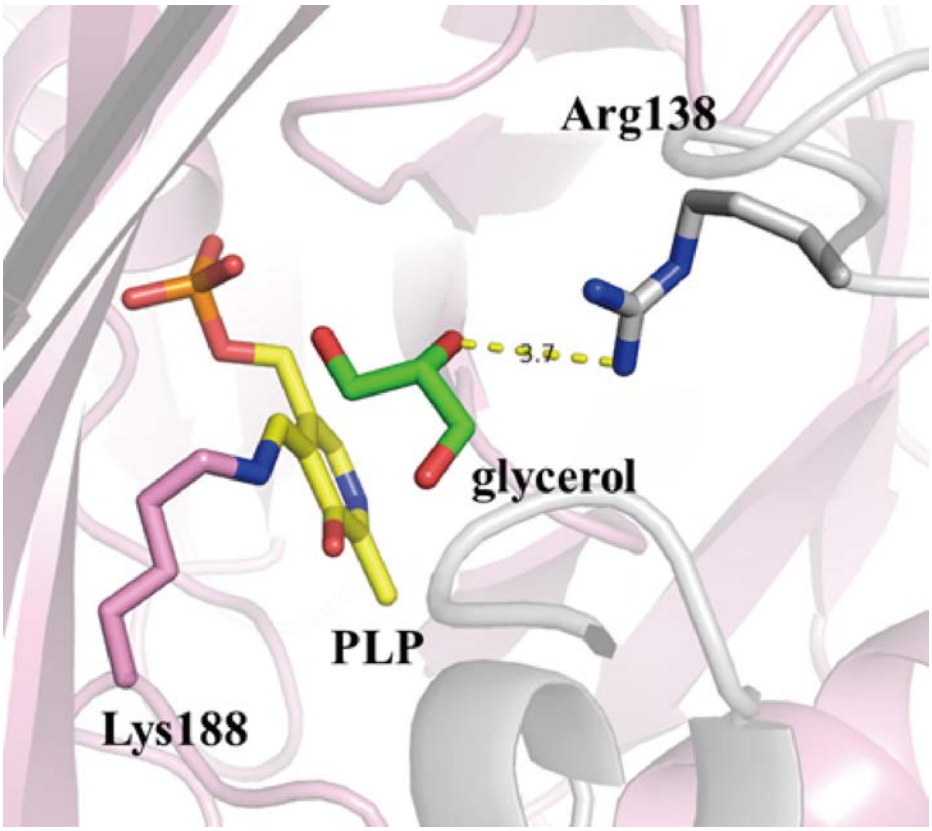


Fig. 2-19 View of the glycerol molecule among the two protomers of (*R*)-TA^{WT}. The Lys188, PLP, glycerol molecule and the residue Arg138 from adjacent protomer are shown as stick in pink, yellow, green and white, respectively. The distance between glycerol and Arg138 is shown by black dashed line and labeled (3.7 Å).

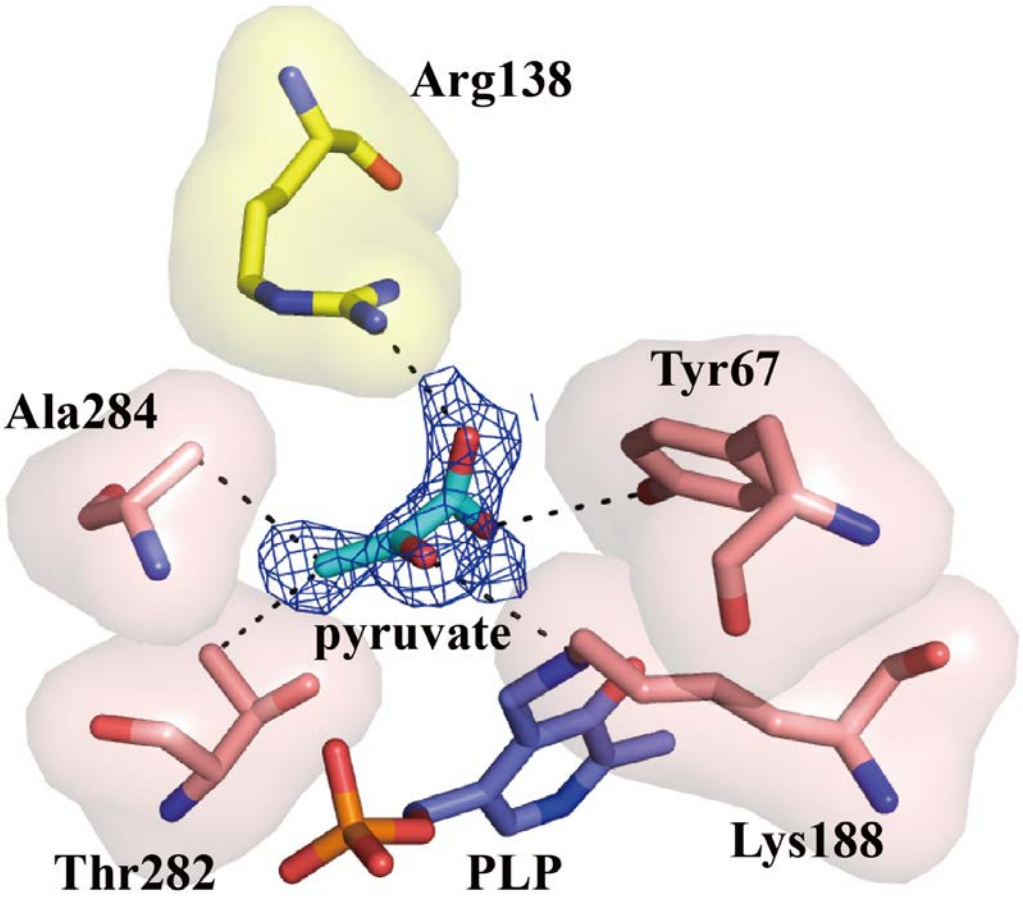


Fig. 2-20 Active site of (*R*)-TA^{PYR}. Pyruvate is shown in both a stick model and F_o-F_c omit density map, contoured at 1.5 σ resolution. The residues which interact with pyruvate are shown in both stick and surface model. Arg138 from the other protomer is colored in yellow.

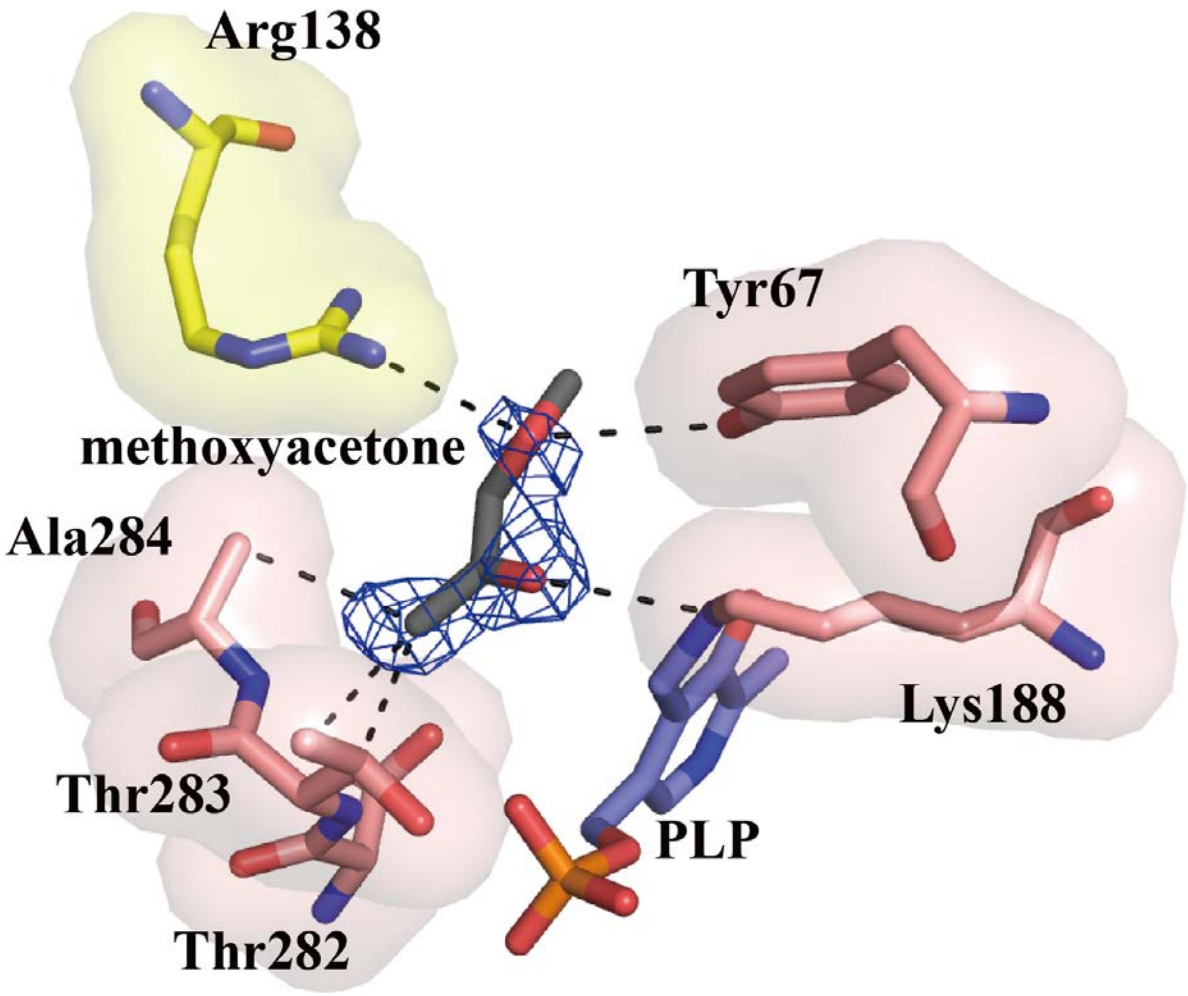


Fig. 2-21 Active site of (*R*)-TA^{METH}. Methoxyacetone is shown in both a stick model and F_o-F_c omit density map, contoured at 1.5 σ resolution. The residues which interact with methoxyacetone are shown in both stick and surface model. Arg138 from the other protomer is colored in yellow.

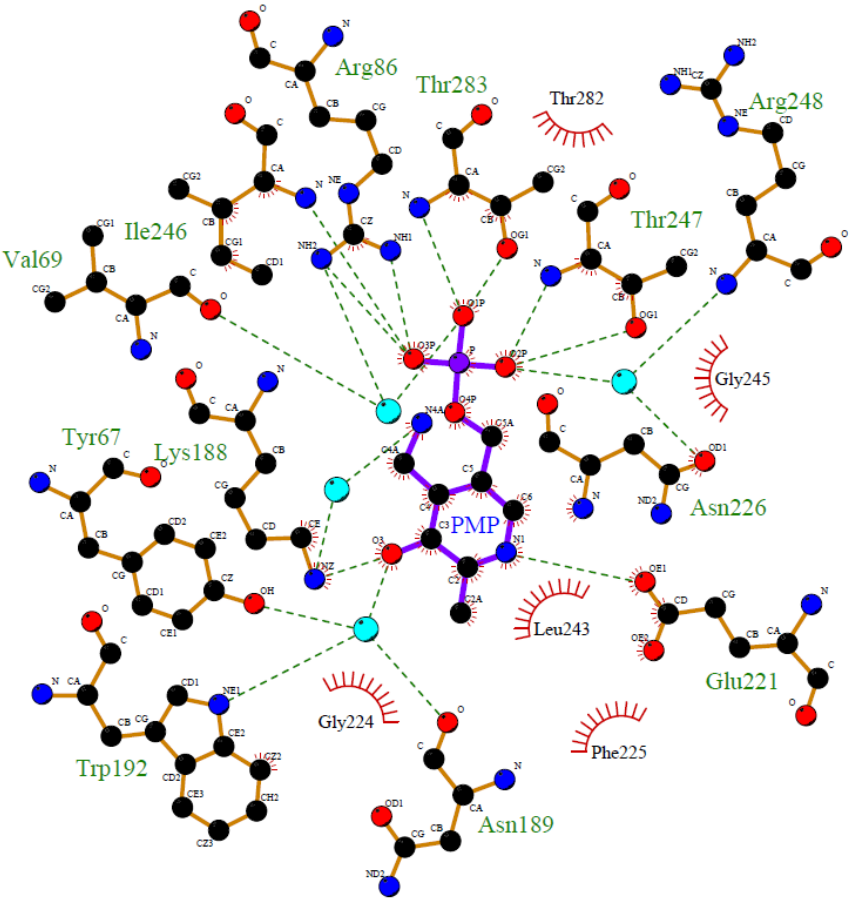


Fig. 2-22 LIGPLOT representation of the PMP in (*R*)-TA^{PMP} structure. Hydrophilic interacting atoms are connected by black dashed lines; non-ligand residues involved in direct hydrophobic contacts with PMP are shown as red semicircles with radiating spokes. Water molecules are shown as cyan circles.

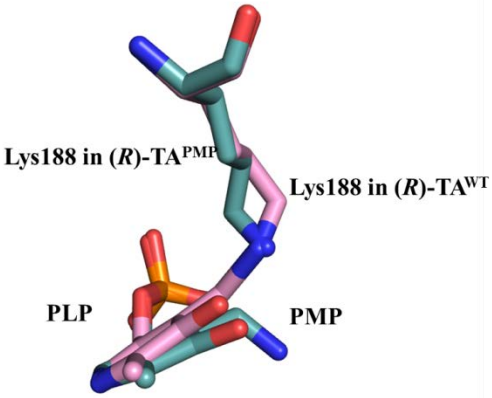


Fig. 2-23 Comparison of cofactor and Lys188 in (*R*)-TA^{WT} and (*R*)-TA^{PMP}, which are colored in pink and light teal, respectively.

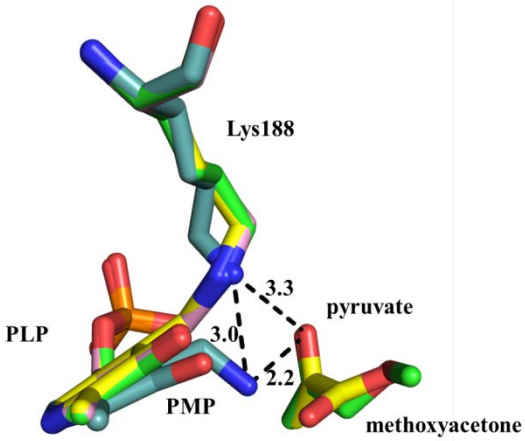


Fig. 2-24 Alignment of the Lys188 and and cofactor in (*R*)-TA^{WT} (pink), (*R*)-TA^{PYR} (yellow), (*R*)-TA^{METH} (green) and (*R*)-TA^{PMP} (light teal). Ketones in (*R*)-TA^{PYR} and (*R*)-TA^{METH} are also shown in the same color.

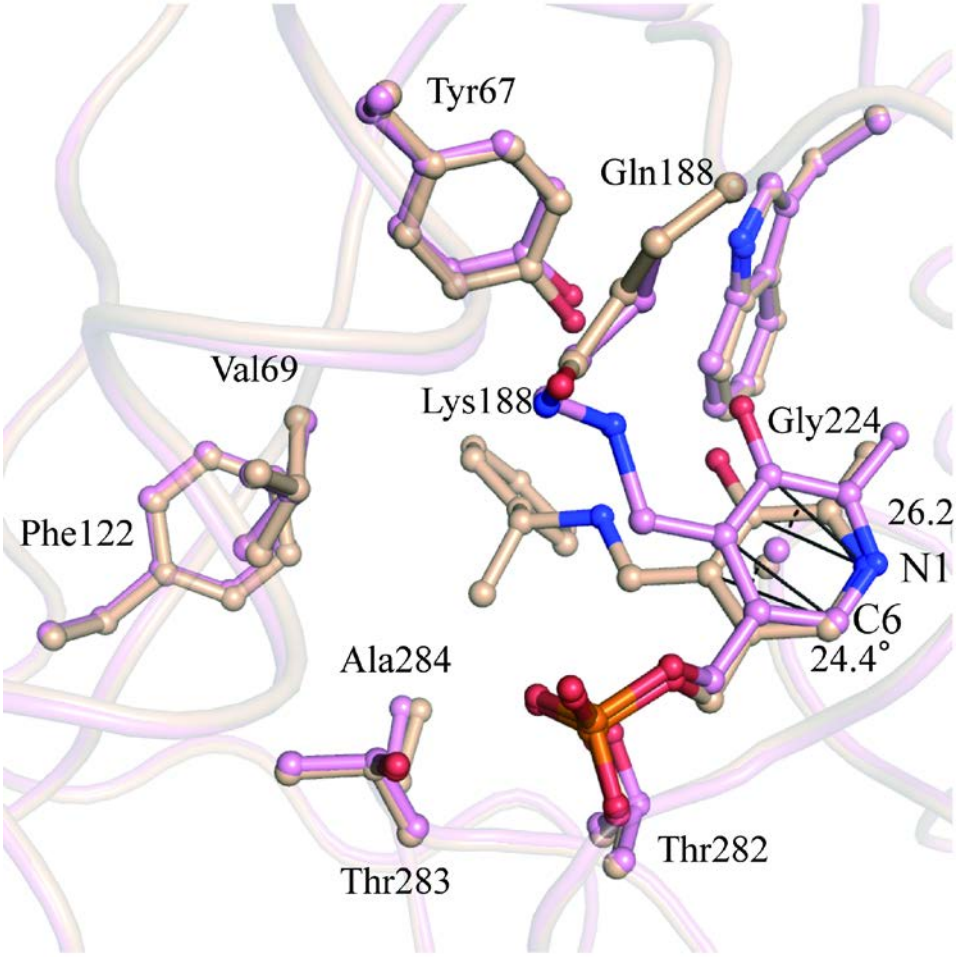


Fig. 2-25 Superposition of ligand binding pockets in one protomer of (*R*)-TA^{WT} (pink) and (*R*)-TA^{REX} with external aldimine (wheat). The PLP portion of the external aldimine rotated approximately 25° relative to PLP molecules of (*R*)-TA^{WT}.

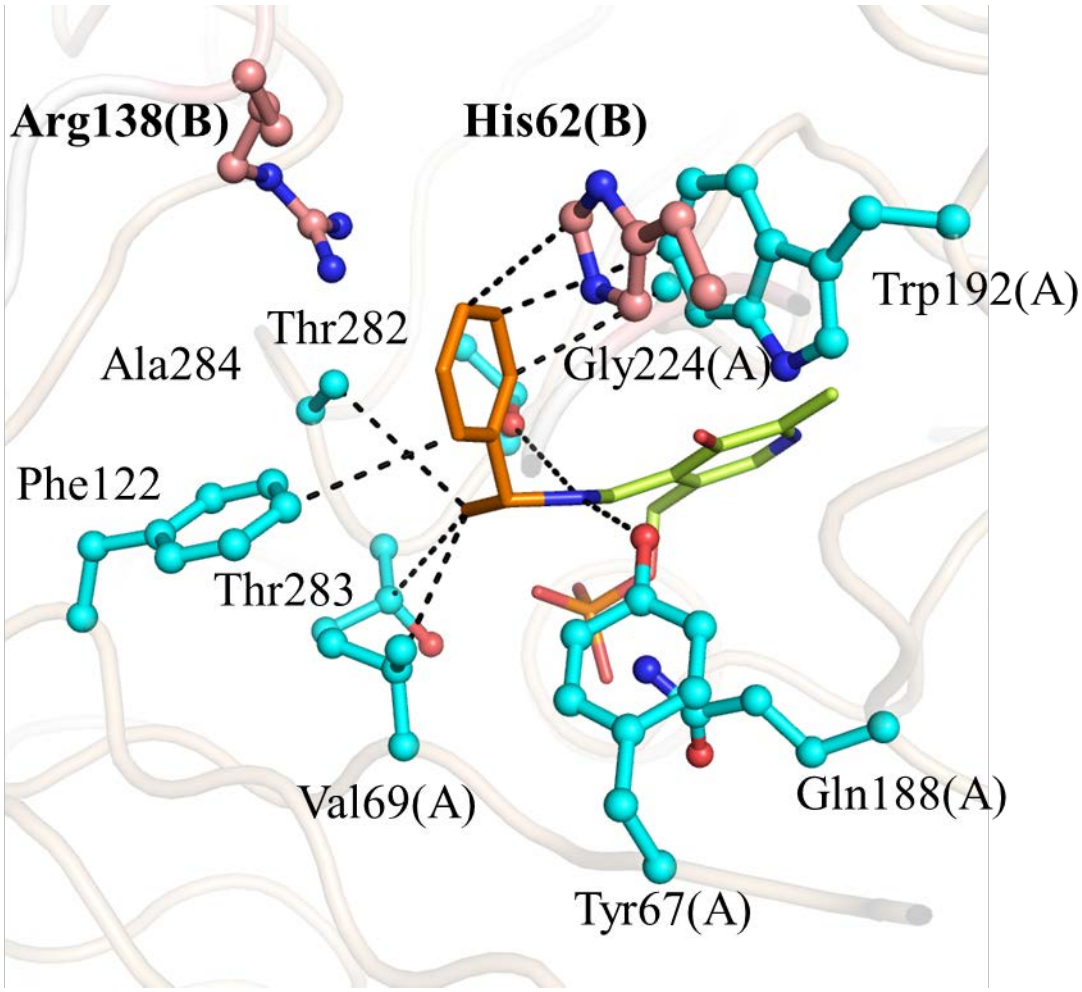


Fig. 2-26 Binding mode of external aldimine in (*R*)-TA^{REX} consisted of PLP (stick in yellow) and *R*-MBA (stick in orange). The interactions are shown as dashed lines. Residues binding the external aldimine are shown as ball-stick model. The His62 and Arg138 residues from adjacent protomer are shown in pink.

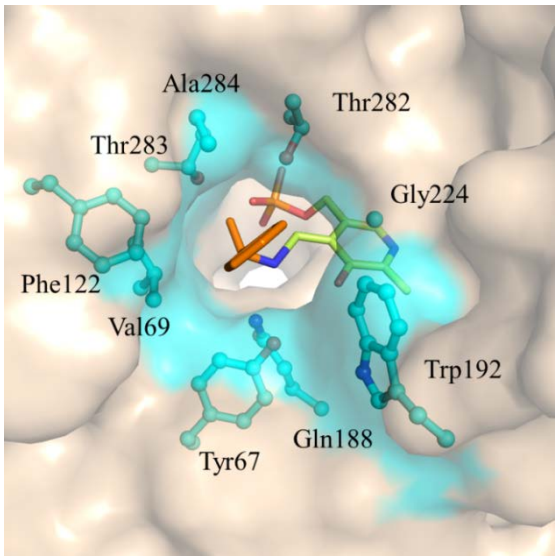


Fig. 2-27 Active site residues of the protomer (surface in wheat) containing external aldimine formed by PLP and *R*-MBA in (*R*)-TA^{REX} are shown as ball-stick and surface in cyan.

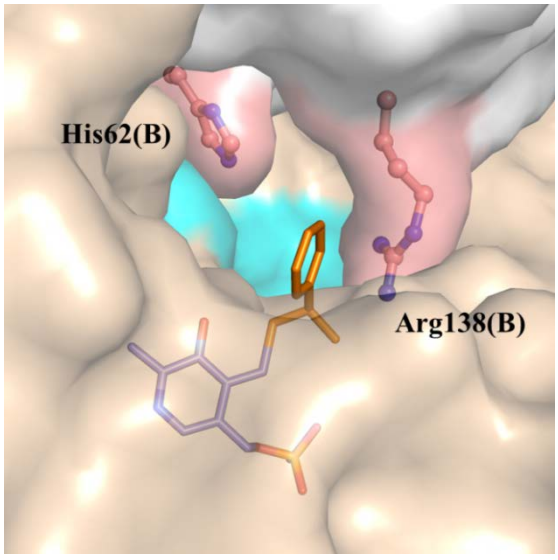


Fig. 2-28 Active site residues of the adjacent protomer (surface in white) are shown as ball-stick and surface in pink.

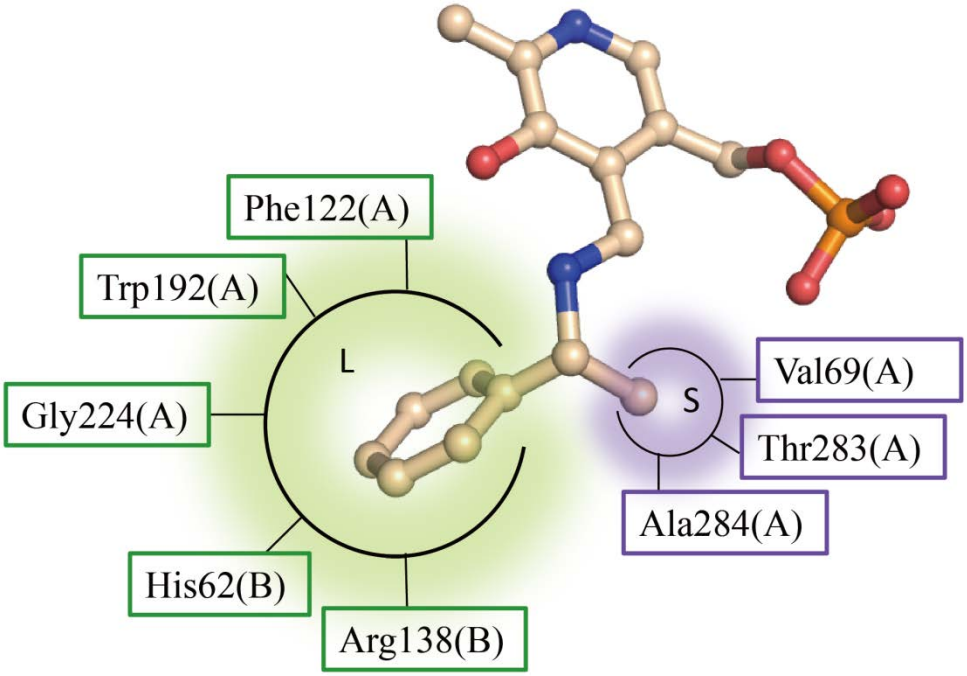


Fig. 2-29 Diagram representation of large and small binding pockets. The external aldimine is shown as ball-stick model. Residues in the large binding pocket are shown in green boxes, and small binding pocket residues are shown in purple boxes.

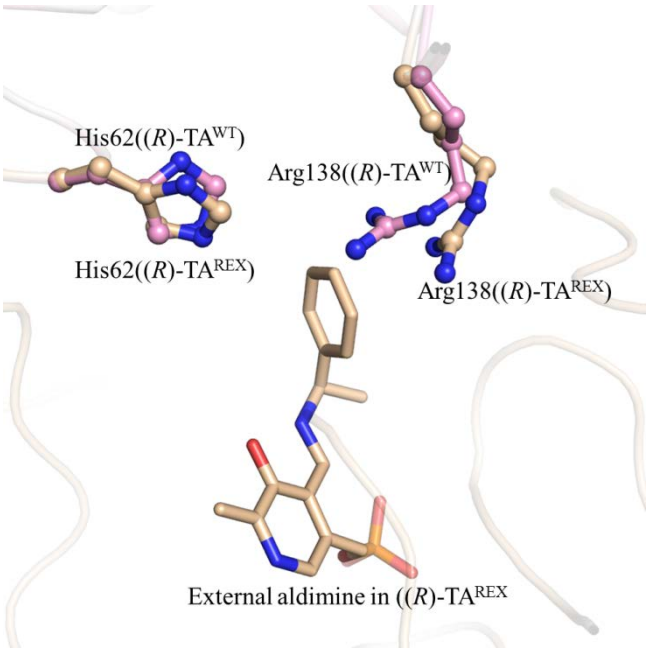


Fig.2-30 Superposition of ligand binding pockets in another protomer of (*R*)-TA^{WT} (pink) and (*R*)-TA^{REX} with external aldimine (wheat).

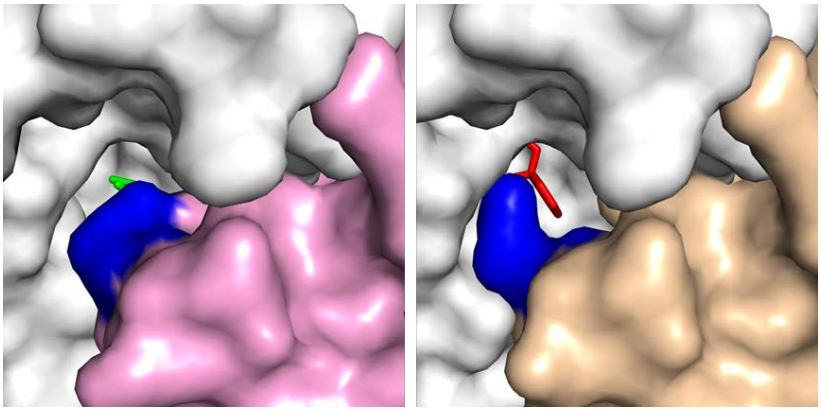


Fig. 2-31 The surface interfaces between the dimer of (*R*)-TA^{WT} (left) and between the dimer of (*R*)-TA^{REX} with external aldimine (right). The Arg138 residues are highlight in dark blue.

Chapter 3

Structural elucidation of the successful molecular engineering

3.1 Exordium

In this chapter, I determined the crystal structure of (*R*)-TA^{Mut28} and an alternated loop in it comparing to the (*R*)-TA^{WT} structure. The corresponding loop in single point mutant (*R*)-TA^{G136F} showed a similar conformation as that of (*R*)-TA^{Mut28}. The crystal structures we determined here revealed the structural details of the previous successful research to engineer ATA-117-11Rd from ATA-117, using a combination of variety of enzyme engineering approaches [16]. In the previous work, they improved 1) the substrate preference for a methyl ketone analog of sitagliptin (referred to as Substrate 1 (Fig. 3-1)), 2) the substrate preference for sitagliptin ketone (referred to as Substrate 2 (Fig. 3-1)), and then, 3) the stability in the conditions with high concentration of organic solvent. Because the difference between (*R*)-TA and ATA-117 is only an amino acid residue the position 306 (Val and Ile in (*R*)-TA and ATA-117, respectively) far from the active sites, and because our mutant enzyme (*R*)-TA^{Mut28} and their improved enzyme are the exactly the same protein, the results of the docking simulations of Substrate 1 and Substrate 2 using our crystal structures explain how their original and improved enzyme recognize these compounds. These crystal structures indicated that the mutation introduced into the 136th residue of (*R*)-TA altered the conformation of a loop besides the active site, resulting in the drastically modified substrate-binding pocket in volume, shape, and surface property, to accommodate the large sitagliptin ketone molecule. These findings explained in detail the previous successful molecular engineering and proposed a new target region for mutagenesis study to change the substrate specificity of ω-transaminases.

3.2 Materials and methods

3.2.1 Preparation of Mut28 mutant and G136F/Y/H/W mutant sample

The methods of primer design and PCR for making G136F/Y/H/W mutant plasmid are same as that were described in Chapter 2. The DNA fragment containing the gene of Mut28 was synthesized (GenScript USA Inc) and inserted into the pUC57. The Mut28 gene was amplified with KOD-plus DNA polymerase (Toyobo, Tokyo, Japan) and inserted into the *Nde*I-*Sac*I site of pET28a (+) to overexpress fusion protein with an N-His tag. After being verified by Fasmac DNA Sequence Service, the Mut28 mutant and G136F/Y/H/W mutant were overexpressed in the same way as (*R*)-TA.

The G136F/Y/H/W mutants were purified in the same manner as wild type (*R*)-TA. To purify Mut28, the supernatant after disrupting and centrifugating was applied onto Ni-Sepharose column, and Mut28 was eluted with 200 mM imidazole dissolved in 20 mM potassium phosphate buffer (pH 6.8) containing 1 mM DTT and 10% glycerol (v/v). Before dialyzing against 20 mM potassium phosphate buffer (pH 6.8) containing 1 mM DTT and 10% glycerol (v/v) overnight at 277 K, 1 mM PLP and 500U Thrombin protease (GE Healthcare) were added into the eluted protein. The dialyzed protein was applied onto Ni-Sepharose column once again to remove the uncleaved protein. Subsequent ion exchange chromatography and size exclusion chromatography were performed as described in Chapter 2. Finally the purified fractions were concentrated to 15 mg ml^{-1} by centrifugation at 4000 g using Vivaspin-20 (10000 molecular weight cutoff).

3.2.2 Sample preparation for crystallization

Before crystallization of Mut28, protein buffer was changed to 20 mM Bis-Tris

buffer (pH6.8) containing 1 mM DTT and 10% glycerol (v/v). Initially, crystallization was performed under the same crystallization conditions described in former Chapters. However, crystals in diffraction quality were not obtained. Then screening and refinement were performed in the same way as that described in Chapter 2.

3.2.3 X-ray data collection and processing

The X-ray diffraction data of (*R*)-TA^{Mut28} and (*R*)-TA^{G136F} crystals were collected at the AR-NW12A and BL5A beamlines at Photon Factory (Tsukuba, Japan), respectively. The datasets were collected with a wavelength of 1.0000 Å, a rotation angle of 0.5º, an exposure time of 1 s for each image. All data sets were composed of 360 images and collected. The diffraction data were indexed, integrated and scaled with XDSme.

3.2.4 Structure determination and refinement

The structures of (*R*)-TA^{Mut28} and (*R*)-TA^{G136F} were solved by molecular replacement. The manual fitting and structure refinement was completed in Coot and validated with the Ramachandran plot drawn by the *PROCHECK* program.

3.2.5 Enzyme assay and analysis

Transaminase activity was assayed by measuring the production of acetophenone, as described in Chapter 2.

3.2.6 The method of molecular docking

The docking simulations were performed using the software package MOE2011.10 (Chemical Computing Group, Montreal, Canada). The dimer of the crystal structures of (*R*)-TA^{WT}, (*R*)-TA^{Mut28} and (*R*)-TA^{G136F} were used as the receptor model. The missing hydrogen atoms of (*R*)-TA^{Mut28} dimer were generated and energy minimized using the

Merck molecular force field 94x (MMFF94x) force field with distance dependent dielectric electrostatics. The possible conformations of sitagliptin were generated using the systematic mode and used for the simulation. Each conformation was posed on the possible active sites, found by the function Site Finder, using an Induced Fit protocol. After the following energy minimization, a solution was selected from the candidates and used for discussion.

3.3 Results and discussion

3.3.1 Crystallization

By streak seeding, crystals of Mut28 grew in 1 week. The crystallization drops were prepared by mixing 1.5 μL sample (15mg/ml) and 1.5 μL reservoir solution (0.2 M magnesium chloride, 0.1 M Tris-HCl pH 8.3, 18% PEG 3350) (Fig. 3-2A).

Crystals of the G136F mutant of (*R*)-TA grew in 5 days containing 1.0 μL sample (10mg/ml) and 1.0 μL reservoir solution (0.2 M magnesium chloride, 0.1 M HEPES-HCl pH 7.8, 16% PEG 3350) (Fig. 3-2B).

3.3.2 Data collection, processing of (*R*)-TA^{Mut28} and (*R*)-TA^{G136F}

Fig.3-3 show X-ray diffraction images of the (*R*)-TA^{Mut28} (A) and (*R*)-TA^{G136F} (B) crystals, respectively. The crystals of (*R*)-TA^{Mut28} and (*R*)-TA^{G136F} belonged to space group $P2_1$. The unit cell and the other crystallographic data collection statistics are shown in Table 3-1. Matthews coefficient value indicted that these two crystals contained twelve molecules per asymmetric unit.

3.3.3 Structure determination of (*R*)-TA^{Mut28} and (*R*)-TA^{G136F}

Initial phasing of (*R*)-TA^{Mut28} diffraction data using (*R*)-TA^{WT} as search model was

failed. Then the search model was modified by removing residues 2-33, 131-144 and 320-330 from (*R*)-TA^{WT} structure which are forming loop in it. Modified model made phasing of (*R*)-TA^{Mut28} succeeded. Removed residues were manually rebuilt using Coot and refined using refmac5. (*R*)-TA^{G136F} was solved by molecular replacement using the structure of the (*R*)-TA^{Mut28} as search model. The refinement statistics are given in Table 3-1. The structures of (*R*)-TA^{Mut28} and (*R*)-TA^{G136F} were refined to R_{work}/R_{free} of 16.0%/21.3% and 20.4%/21.8%, respectively. The Ramachandran plots of (*R*)-TA^{Mut28} and (*R*)-TA^{G136F} showed that 97.1%/2.8% and 98.1%/1.9% of φ-ψ pairs lie in the most favored and additionally allowed regions, respectively.

3.3.4 Cofactor binding manner in (*R*)-TA^{Mut28} and (*R*)-TA^{G136F}

In the crystal structure of (*R*)-TA^{Mut28}, the PLP molecule is recognized by nearly the same amino acid residues in a similar manner to that of (*R*)-TA^{WT} with some exceptions (Fig. 3-4). There is a lack of continuous omit map for the carbon-nitrogen bond between PLP and Lys188 (Fig. 3-5). Lys188 lies on the *re*-face of the PLP and the amino nitrogen atom forms a hydrogen bond to the aldehyde oxygen atom of PLP with the interaction distance of 2.6 Å, apparently too long for Schiff-base linkage. The PLP in (*R*)-TA^{Mut28} rotated approximately 25° about the N1-C6 bond relative to the PLP molecules of (*R*)-TA^{WT} (Fig. 3-6A). In the crystal structure of (*R*)-TA^{G136F}, the binding mode of the cofactor PLP was nearly the same as that in (*R*)-TA^{WT} (Fig. 3-6B).

3.3.5 Loops 129-146 in (*R*)-TA^{Mut28} and (*R*)-TA^{G136F} show significantly different conformation

The loop 129-146 showed significantly different conformations between (*R*)-TA^{WT} and (*R*)-TA^{Mut28} probably depending on the amino acid residue at the position 136 (Fig.

3-7A). Gly136 was located near Val 152 in (*R*)-TA^{WT}, whereas Phe136 was distant from the residue Val152 in (*R*)-TA^{Mut28}, accompanied by the long-range alteration of the loop's backbone conformation (Fig. 3-7B). It suggests that the probable steric hindrance between the bulky side chain of Phe136 and the residue Val152 was the main cause of the conformation difference (Fig. 3-7B). The crystal structure of single point mutant (*R*)-TA^{G136F} showed a similar loop 129-146 as that of (*R*)-TA^{Mut28} (Fig. 3-8), which confirmed the above hypothesis. As a result, the key residue to recognize ketone substrate, Arg138, has been lost from the substrate-binding pocket (Fig. 3-9) and this altered loop in the dimer interface led to a considerable open entrance to the active site (Fig. 3-10).

3.3.6 Activity assay of (*R*)-TA mutants

Activity measurements of four mutants of (*R*)-TA, G136F/Y/H/W, were performed. Each mutant showed an extremely high K_m towards pyruvate (Table 3-2), indicating that these mutants could not recognize pyruvate properly. It can be explained that the introduction of a bulky side chain into the residue 136 altered the conformation of the loop 129-146 and, as the result, moved out Arg138, the key residue to pyruvate recognition, of the active site. (Fig. 3-9) The activity ratio (B/P) of the four mutants also increased a lot (Table 3-2), indicated that the property of the substrate-binding pocket of these mutants favor bulky and hydrophobic substrate. These mutational analyses can be explained by our models mentioned above that the introduced bulky side chain of 136th residue causes conformational alteration of loop 129-146.

3.3.7 Docking simulation

In the previous work to engineer ATA-117-Rd11 from ATA-117, they improved 1)

the substrate preference for a methyl ketone analog of sitagliptin (Substrate 1; Fig.3-1), 2) the substrate preference for sitagliptin ketone (Substrate 2; Fig.3-1), and then, 3) the stability under the conditions with high concentration of organic solvent. Because the difference between the wild-type (*R*)-TA and ATA-117 is only an amino acid residue at the position 306 (Val and Ile in (*R*)-TA and ATA-117, respectively) far from the active sites, crystal structures we determined here provided us the enough structural information to elucidate the previous engineering of substrate specificity.

In the docked model of (*R*)-TA^{WT} complexed with Substrate 1, Substrate 1 is anchored almost by hydrophobic interactions except for the hydrogen bond with Ser223, His62 and Arg138 of the other protomer (Fig. 3-11A). The methyl group and the cyclic group of Substrate 1 are bound in the small pocket and large pocket that determined by the crystal structure of (*R*)-TA^{K188Q}, respectively. When Ser223 of ATA-117 was mutated to proline, the activity towards Substrate 1 showed an 11-fold improvement [16]. It could be presumed that the interaction between pyrrolidine ring of Pro223 and cyclic group of Substrate 1 was stronger than the hydrogen bond between Ser223 and the nitrogen atom of Substrate 1(Fig. 3-11A).

When Substrate 1 docked in (*R*)-TA^{G136F}, it has different binding mode from that docked in (*R*)-TA^{WT}. In this docking model, the methyl group of Substrate 1 has the similar orientation and is bound by the same residues as that of *R*-MBA in the crystal structure of (*R*)-TA^{K188Q} and that of Substrate 1 docked with (*R*)-TA^{WT} (Fig. 3-11B). However, the cyclic group of Substrate 1 in this result rotated a lot comparing to that of Substrate 1 docked with (*R*)-TA^{WT}. In addition, the cyclic group of Substrate 1 in this result was anchored mainly by the residue from the other protomer of the dimer, which is different from that docked in (*R*)-TA^{WT} (Figs. 3-11A and 3-11B).

The cyclic group of Substrate 2 docked in (*R*)-TA^{Mut28} and that of Substrate 1 docked in (*R*)-TA^{G136F} are almost bound in the same large pockets, respectively (Fig. 3-11C). The 1,2,4-3- fluorine phenyl group of Substrate 2, however, could not be trapped by the small binding pocket of (*R*)-TA^{G136F} in the same manner as docked in the (*R*)-TA^{Mut28}, because of the steric interference with Phe122, which is mutated to Met122 in (*R*)-TA^{Mut28} (Fig. 3-11D). It could be supposed that the mutation 122M might be important to accommodate the 1,2,4-3- fluorine phenyl group of Substrate 2.

3.3.8Altered binding pockets in (*R*)-TA^{Mut28}

(*R*)-TA showed amino donor specificity toward some aliphatic and arylic chiral amines. Among the amines tested, *R*-MBA was a good amino donor [15]. It seems reasonable to suppose that R-MBA is most suitable for the architecture of the active site, and is tightly bound by hydrophobic interactions between MBA and residues in the active site (Fig. 2-26). The aliphatic amino donors have a lower activity than arylic amino donors and the chain length seems to require at least 6 carbon atoms. That could be explained by that the hydrophobic interactions between the substrate and the enzyme may become weak or disappear when phenyl group was substituted by arylic group. The study on substrate structure-reactivity also indicated that hydrophobic interaction in the large pocket with an aryl group is stronger than that with aliphatic. Comparing to (*R*)-TA^{WT}, the small binding pockets of (*R*)-TA^{mut28} was enlarged by the mutations of V69T, F122M, T282S and A284G (Fig. 3-12). The enlargement of large binding pocket was realized by the alteration of the loop 129-146. Furthermore, the large pocket is even enlarged by the mutation of H62T (Fig. 3-9).

Moreover, the Phe122 residue binds to the methyl group of Substrate 1 in the small

pocket of (*R*)-TA^{G136F}, while the corresponding residue binds phenyl group of *R*-MBA in the large pocket of (*R*)-TA^{REX}. It indicated that the residues constituting the small pocket may vary according to the size of substrate.

As depicted above, along with the alteration of loop 129-146, Phe136 residue came towards large pocket. Actually the room for the cyclic group of Substrate 1 docked with (*R*)-TA^{WT} is occupied by Phe136 residue in (*R*)-TA^{G136F} model so that the cyclic group of Substrate 1 rotates to the space vacated by the shifting of the loop. In a word, the large pocket would shift from L1 to L2 as shown in Fig. 3-13 due to the site mutation of G136F. In addition, in the (*R*)-TA^{Mut28} docked with Substrate 2, the pockets consist of different residues comparing to that of the Substrate 1 docked in the crystal structure of (*R*)-TA^{G136F} (Fig. 3-14).

Variable binding pockets were observed due to ligand-binding, different size of substrates and mutations. Based on the above information, it may become easy to rationally design the pockets to accommodate a desired chiral amine compound.

3.3.9 Elucidation of the Sitagliptin-producing Transaminase, ATA-117-rd11

Based on the crystal structures, their previous work can be divided into three structural modifications; 1) alteration of the conformation of the loop 129-146, 2) clipping or replacement of the side chains of the residues composing the large- or small-binding pockets, and 3) improvement of the stability in the conditions with high concentration of organic solvent (Fig.3-14). Among these three modifications, the 1st and 2nd points can be explained by our crystal structures.

The ATA-117 hardly showed activity toward Substrate 1 because of the limited volume of the large-binding site, and only a single mutation, G136Y, improved the

activity toward Substrate 1 [16]. The introduction of the bulky side chain into the position 136 would alter the conformation of the loop 129-146 as demonstrated by the crystal structure of the G136F mutant (Fig.3-7B). This structural change would move Arg138, which were positioned between the large- and small-binding pockets and narrowed the volume of the substrate-binding site in ATA-117, away from the active site, resulting in the enlarged substrate-binding pocket (Fig.3-15) and opened entrance to the active site (Fig.3-10). As the result, the large-binding pocket would be extended by the mutation G136Y to accommodate the Substrate 1. The mutation at the position 136 was finally refined to be G136F.

The 2nd modification step was achieved by cumulative mutations to widen the substrate-binding pocket. It was reported that introduction of mutations of H62T, V69G, F122I, G136Y, E137I, V199I, A209L, S223P, S282S, and A284G improved the substrate preference toward Substrate 2. As for the mutation G136Y, we have already described above the effect on the substrate specificity. The mutation E137I is not pertinent in this paper, because the position 137 was back-mutated to Glutamate during the 3rd engineering step for stability. Among the other seven mutation sites, positions 62, 69, 122, 282, and 284 are located in the substrate-binding pocket. The five mutations introduced into these positions were intended to reduce its volume of the side chain, which enlarged the space of the pocket.

The conformational difference of the loop 129-146 provided a suggestion about the reason for the improvement caused by V199I and A209L mutations. During the previous engineering of ATA117, the introduction of V199I, A209L, and G136Y mutations (together with H62T, E137I, and T282S mutations) was improved the reaction rate for Substrate 2 by 75-fold over its parent enzyme (including S223P

mutation). Phe136 formed van der Waals interactions with Ile199, Leu209, and Pro223 in ATA117-Rd11 (Fig.3-16). This interaction seemed to stabilize the position of Phe136 and, as the result, stabilize the conformation of the loop 129-146 in the state suitable for large substrates. If Ile199 and Leu209 were back-mutated to Val and Ala, respectively, the van der Waals interactions and stabilization of the loop would be weaker. Based on this model, the improvement caused by V199I, A209L, and G136Y mutations can be explained that these mutations improved stability of the loop's conformation to accommodate Substrate 2.

3.3.10 Putative role of the loop 129-146

The large substrate binding pocket is a key determinant to substrate specificity and stereoselectivity of ω-amine transaminase [8]. Based on the crystal structure of (*R*)-TA^{REX} and docking models, the loop 129-146 consists of large pocket on the dimer interface. In the crystal structures of (*R*)-TA^{Mut28} and (*R*)-TA^{G136F}, the loop 129-146 displays an obviously different conformation from (*R*)-TA^{WT} due to the steric hindrance caused by a mutation of G136F. The significantly different substrate specificity between ATA-117 and ATA-117-11Rd comes from, at least in part, the different conformation of this loop. Comparing to (*R*)-TA, the altered substrate specificity of single mutation of G136F/Y/H/W and R138A/Q in this loop could confirm above hypothesis. Based on these results, we propose that the loop 129-146 mainly determines the substrate specificity of (*R*)-TA.

From the alignment of (*R*)-TA and seven *R*-setereoselective ω-transaminases recently isolated [29], most of the residues that constitute the substrate-binding pockets of (*R*)-TA are conserved in the other seven enzymes, except for the residues in the loop

129-146 (Figs. 3-17, 3-18 and 3-19). It is reasonable to assume that the diversity of the amino acid sequence of the loop causes the different substrate specificity among the eight enzymes, in accordance with our model that the loop 129-146 mainly determines the substrate specificity of (*R*)-TA. In addition, mutation of the residues Phe3, Pro48, Try150, Gln155, Val199, and His203, which exist around the loop 129-146 (Fig.3-17), are also expected to change the conformation of this loop. Thus, the loop 129-146 functions as a substrate specificity determinant and should be given priority to study, including the residues around it, when (*R*)-selective amine transaminases are expected to be engineered to change the substrate specificity.

Table. 3-1 Data collection and refinement statistics of (*R*)-TA^{Mut28} and (*R*)-TA^{G136F}

	(*R*)-TA^{Mut28}	**(*R*)-TA^{G136F}**
Data collection		
Beamline	AR- NW12A	BL5A
Wavelength (Å)	1.0000	1.0000
Space group	$P\,2_1$	$P\,2_1$
Cell dimensions	82.01, 133.54, 195.54	81.96, 134.28, 196.04
Cell angles (°)	90.00, 100.41, 90.00	90.00, 100.35, 90.00
Resolution (Å)	20.0-2.20	50.0-2.27
Number of observations	719999	615286
Number of unique reflections	201959	161915
Completeness (%)	96.3 (90.2)	98.8 (92.2)
R_{merge} (%)	6.2 (28.6)	8.5 (65.5)
Redundancy	3.57	3.80
I/σ (I)	13.99 (3.83)	11.6(2.00)
Refinement		
Resolution range (Å)	20.0-2.20	47.88-2.27
R_{work}/R_{free} (%)	17.0/21.3	20.4/21.8
Number of non-hydrogen atoms		
Protein	30232	30291
Water	1803	1863
Ligand	PLP	PLP
	192	192
RMSD bond length (Å)	0.009	0.006
RMSD bond angle (º)	1.345	1.062
Ramachandran plot (%)		
Preferred	97.1	98.1
Allowed	2.8	1.9
Outliers	0.1	0.0

The values in parentheses are for the highest resolution shell

$R_{work}=\Sigma||Fo|-|Fc||\ /\Sigma|Fo|$, where $|Fo|$ and $|Fc|$ are the observed and calculated structure factor amplitudes of a particular reflection.

R_{free} was calculated from 5% of reflections omitted from the refinement.

Table. 3-2 Kinetic parameters of (*R*)-TA and mutants

	Specific activity for Pyruvate (P) U[(a)]/mg	Specific activity for Benzylacetone(B) U/mg	B/P[(b)]	K_m for Pyruvate (mM)	V_{max} for Pyruvate (U/mg)
WT	11.01	0.34	3.1%	3.2	15.4
G136Y	0.48	0.38	79.2%	354.8	13.5
G136F	0.23	1.08	469.6%	934.6	18.4
G136H	1.09	0.40	36.7%	124.8	15.5
G136W	0.72	0.18	25.0%	414.5	19.8

(a) One unit of (*R*)-TA^{WT} and mutants was defined as the amount of enzyme producing 1 μmol of D-alanine or (*R*)-amine per minute.

(b) The ratio of specific activity of benzylacetone to that of pyruvate.

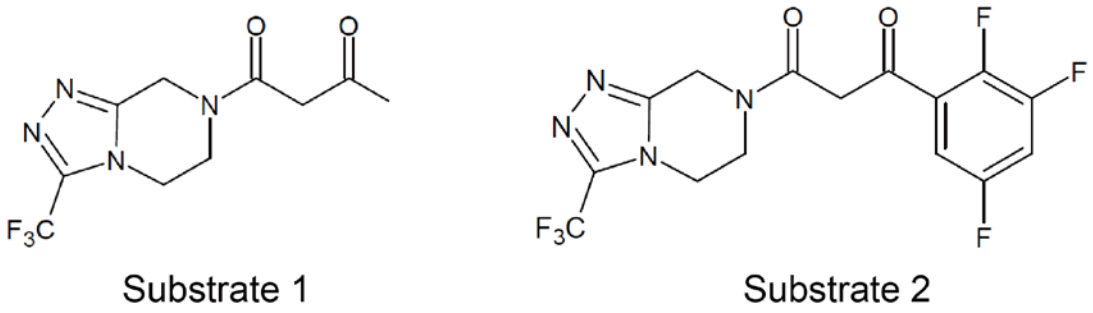
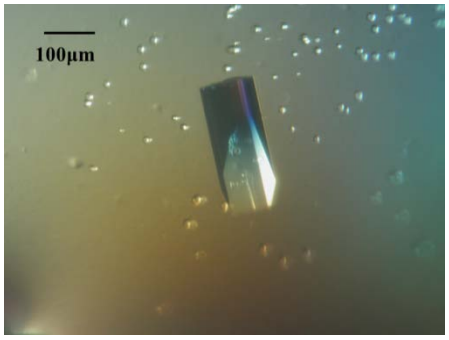


Fig. 3-1 Structural formulae of Substrate 1 and Substrate 2.

A



B

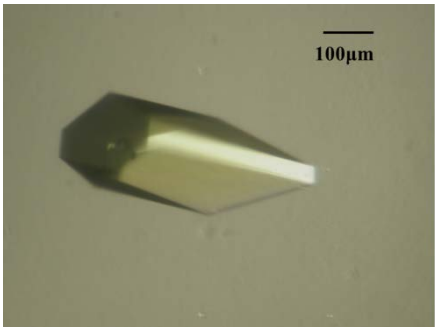


Fig.3-2 The crystal of Mut28 mutant and G136F mutant.

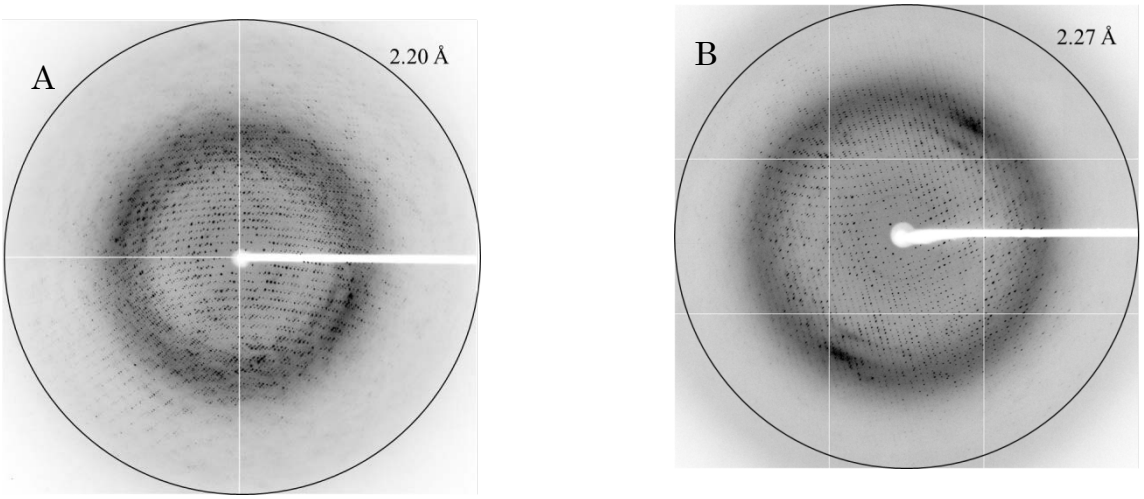


Fig. 3-3 An X-ray diffraction image collected from a crystal of (*R*)-TA^{Mut28} and (*R*)-TA^{G136F}.

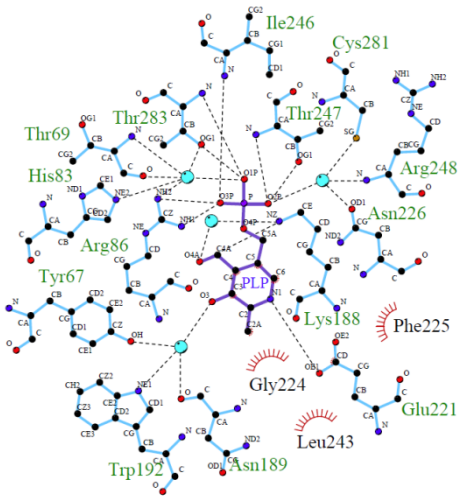


Fig. 3-4 The LIGPLOT representation of the PLP in (*R*)-TA^{Mut28}. The hydrophilic interactions are represented by black dashed lines; non-ligand residues involved in direct hydrophobic contacts with PLP are shown as red semicircles with radiating spokes. The amino acid residues of (*R*)-TA^{Mut28} are colored cyan. Water molecules are shown as cyan circles.

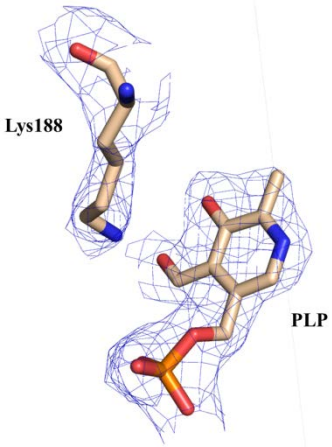


Fig. 3-5 F_o-F_c omit map of PLP and residue188 contoured at a 2σ.

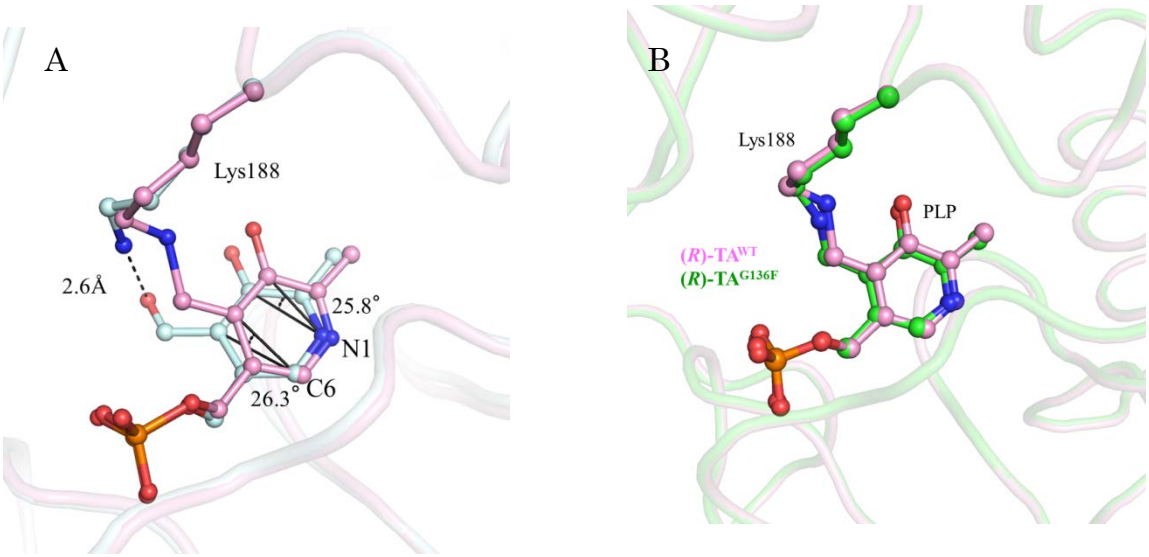


Fig. 3-6 Comparison of cofactor PLP in (*R*)-TA^{WT} and (*R*)-TA^{Mut28} or (*R*)-TA^{G136F}. The Lys188, cofactor PLP of (*R*)-TA^{Mut28} are colored in light blue, and the corresponding Schiff base in (*R*)-TA^{WT} and (*R*)-TA^{G136F} are is colored pink and green, respectively.

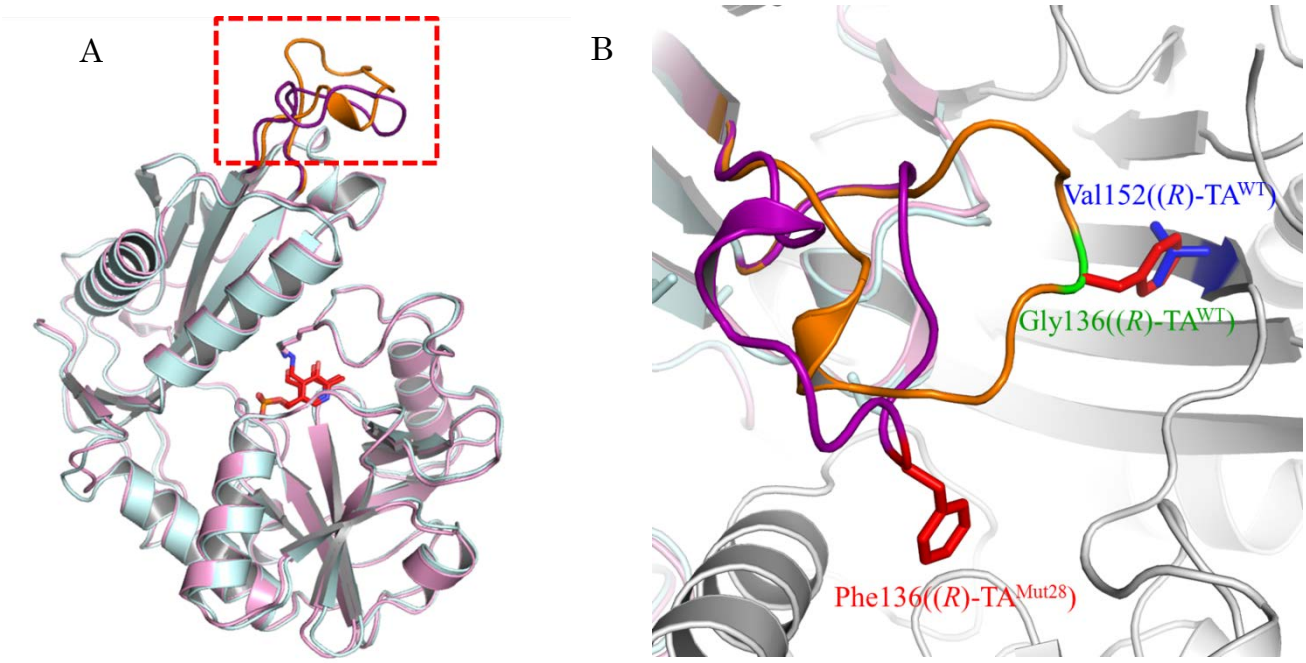


Fig. 3-7 (A) Overall structure comparison of (*R*)-TA^{WT} (pink) and (*R*)-TA^{Mut28} (pale cyan). The structurally different loop 129-146 of (*R*)-TA^{WT} and (*R*)-TA^{Mut28} are highlighted in a red box in orange and purple, respectively. (B) Predicted steric interference between the side chain of residue 136 and Val152. When Gly136 (green) in the loop 129-146 of (*R*)-TA^{WT} (orange) was mutated to phenylalanine (red) in its energy-favorable rotamer, steric hindrance is expected between the side chain and Val152. The conformation of the loop 129-146 of (*R*)-TA^{Mut28} is colored purple.

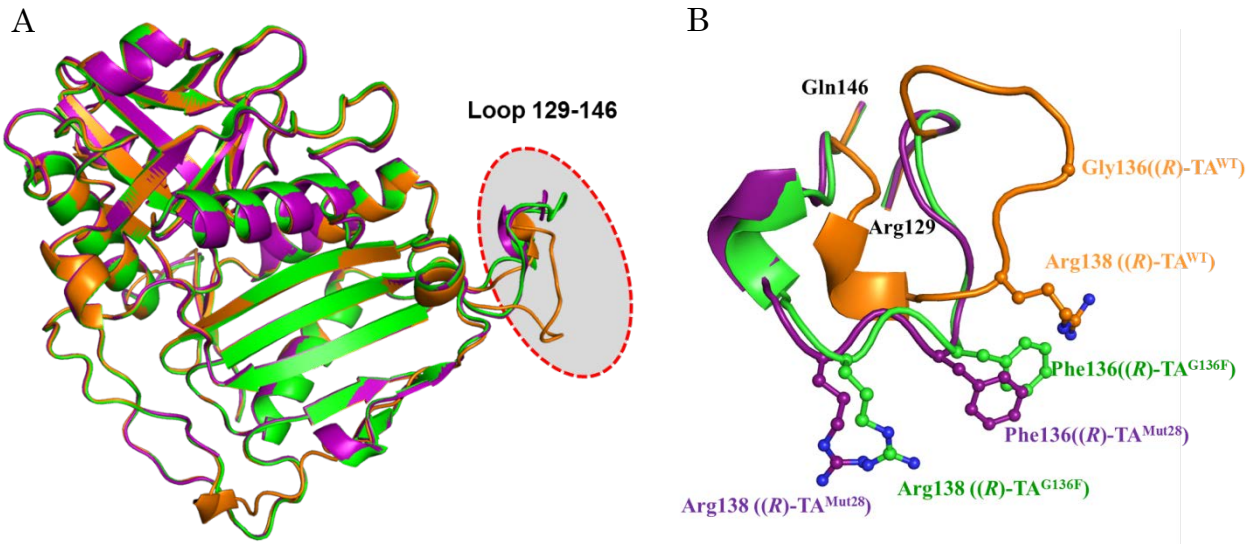


Fig. 3-8 (A) Overall structure alignment of (*R*)-TA^{WT}, (*R*)-TA^{G136F} and (*R*)-TA^{Mut28}. The monomer of (*R*)-TA^{WT}, (*R*)-TA^{G136F} mutant and (*R*)-TA^{Mut28} is shown as cartoon in orange, green and purple, respectively. The three crystal structures are nearly the same except for the loop 129-146 which is highlighted in a red dashed oval. The R.M.S.D for (*R*)-TA^{WT} and (*R*)-TA^{G136F}, (*R*)-TA^{WT} and (*R*)-TA^{Mut28} are 0.164 Å and 0.334 Å, respectively. (B) A close-view of the comparison of the three loops in (*R*)-TA^{WT}, (*R*)-TA^{G136F} and (*R*)-TA^{Mut28}.

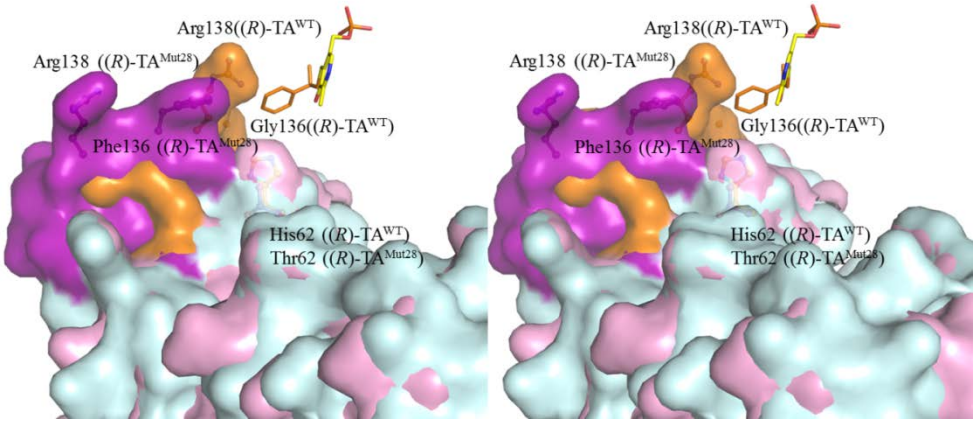


Fig. 3-9 Stereographic close-up view of comparison of (*R*)-TA^{K188Q} and (*R*)-TA^{Mut28}. The external aldimine in (*R*)-TA^{K188Q} is shown in stick model.

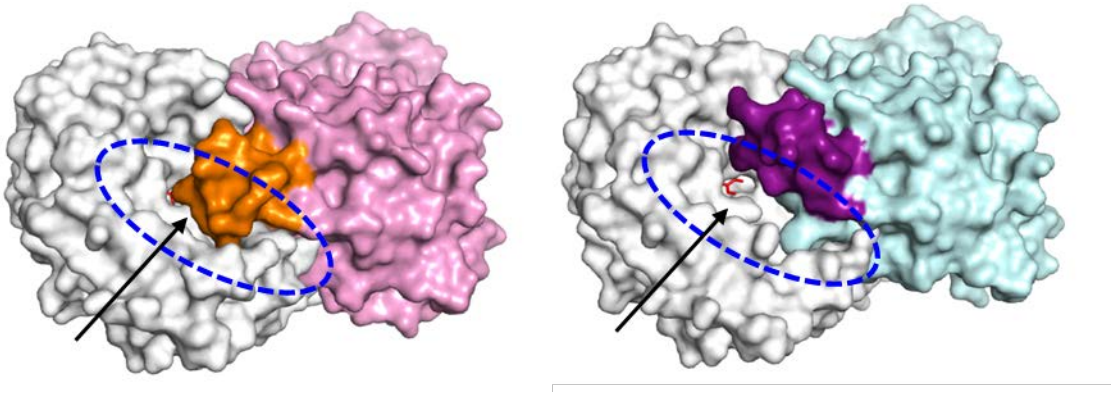


Fig. 3-10 The enlargement of the entrance of the substrate-binding sites caused by the mutation G136F. The dimer surfaces of (*R*)-TA^{WT} (left) and (*R*)-TA^{Mut28} (right) indicated a larger substrate entrance in (*R*)-TA^{Mut28}. A protomer is colored, while the other is colorless. The loop 129-146 of the colored protomer of (*R*)-TA^{WT} and (*R*)-TA^{Mut28} are highlighted in orange and magenta, respectively. The substrate entrances are circled in dashed oval.

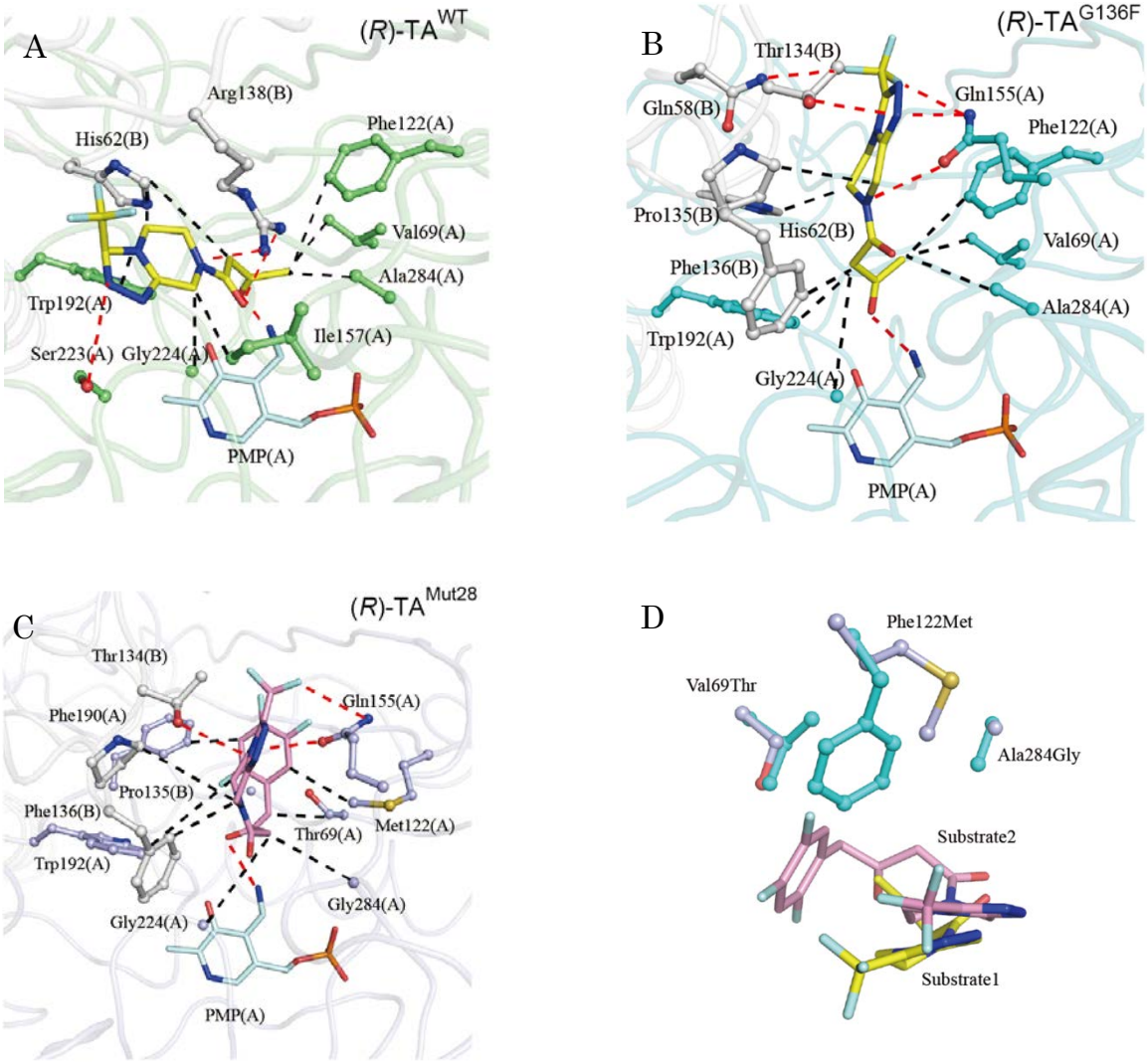


Fig. 3-11 Docking of Substrate 1 and Substrate 2 into the pockets of (*R*)-TA^{WT}, (*R*)-TA^{Mut28} and (*R*)-TA^{G136F}. (A) Docking of Substrate 1 in (*R*)-TA^{WT}. (B) Docking of Substrate 1 in (*R*)-TA^{G136F}. (C) Docking of Substrate 2 in (*R*)-TA^{Mut28}. In (A) and (B), Substrate 1is colored yellow and Substrate 2 in (C) is colored pink. In (A), (B) and (C), the polar interactions between ligand and residue are shown as red dash lines while nonpolar interactions are shown in black dash lines. (D) Portion alignment of Substrate 1 (yellow) docked in (*R*)-TA^{G136F} (cyan) and Substrate 2 (pink) docked in (*R*)-TA^{Mut28} (light blue).

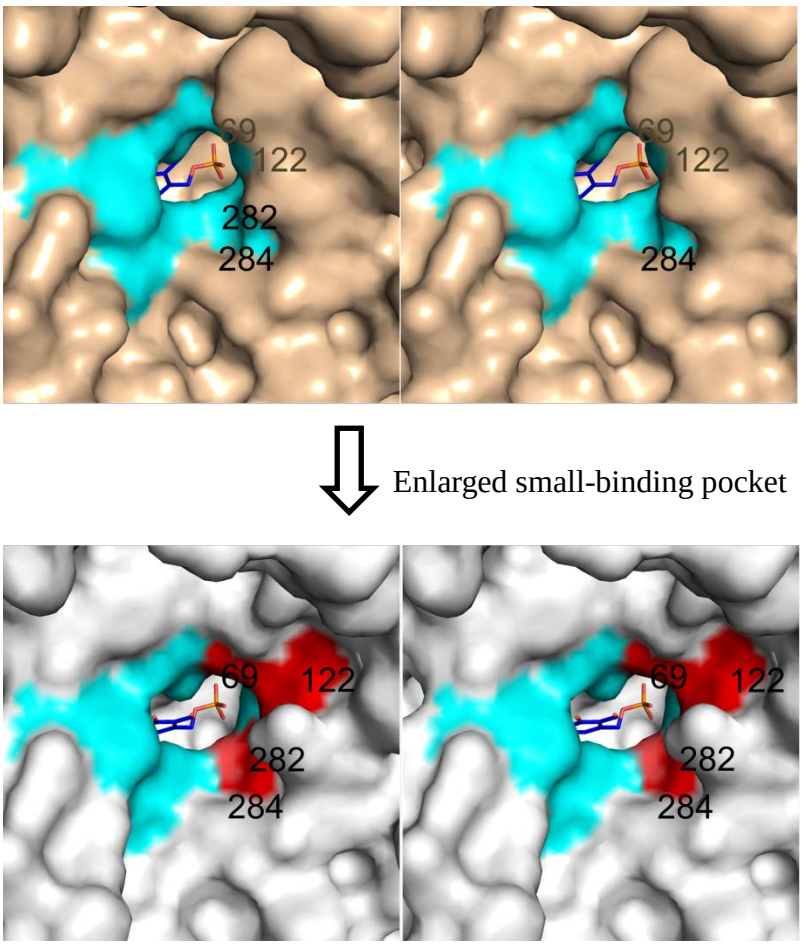


Fig. 3-12 The enlargement of the small-binding pocket caused by the mutations V69T, F122M, T282S and A284G. A protomer of (*R*)-TA^{WT} (up) and (*R*)-TA^{Mut28} (below) are shown as surface. The residues in the active site are shown in cyan or red. The location of the residues 69, 122, 282, and 284 are labeled.

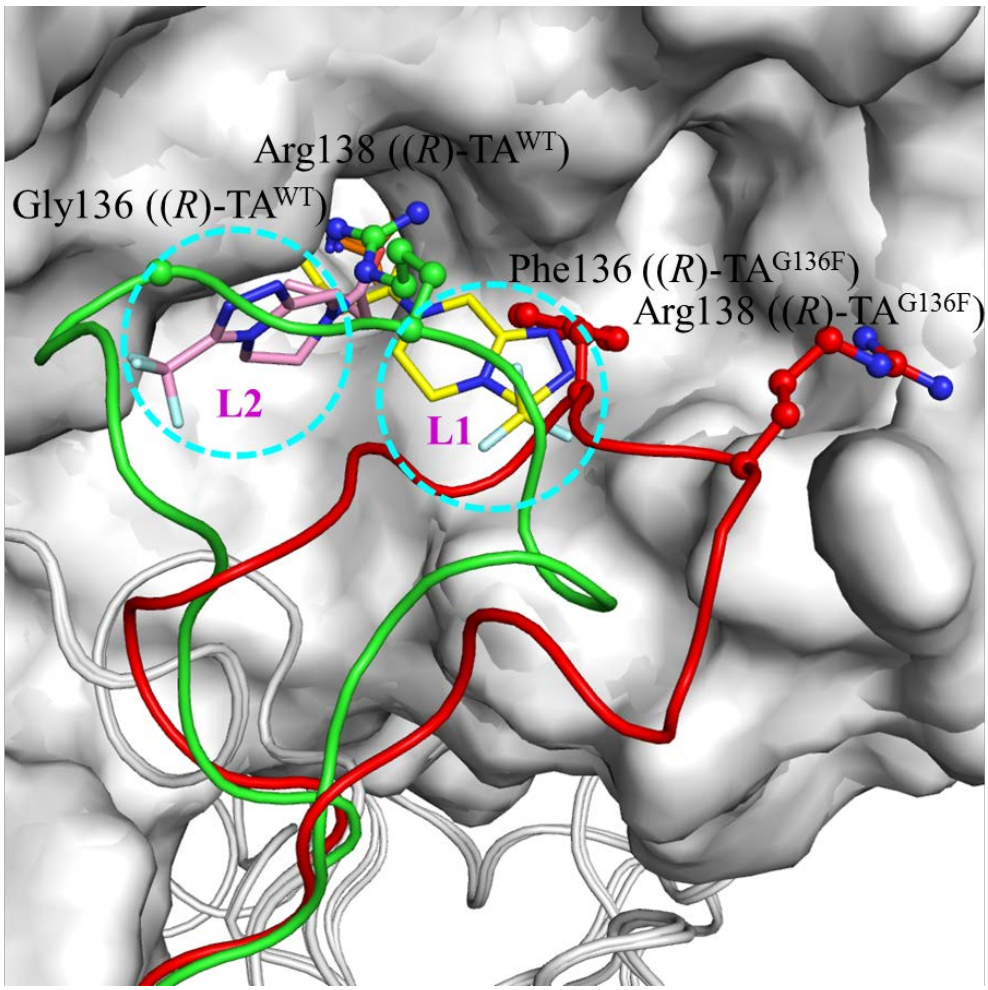


Fig.3-13 A stereograph of the comparison between Substrate 1 docked in (*R*)-TA^{WT} (yellow) and (*R*)-TA^{G136F} (pink). The loop129-146 in the other protomer of (*R*)-TA^{WT} and (*R*)-TA^{G136F} are colored in green and red, respectively. The large binding pockets of (*R*)-TA^{WT} (L1) and (*R*)-TA^{G136F} (L2) are represented as cyan dashed circle and labeled in purple.

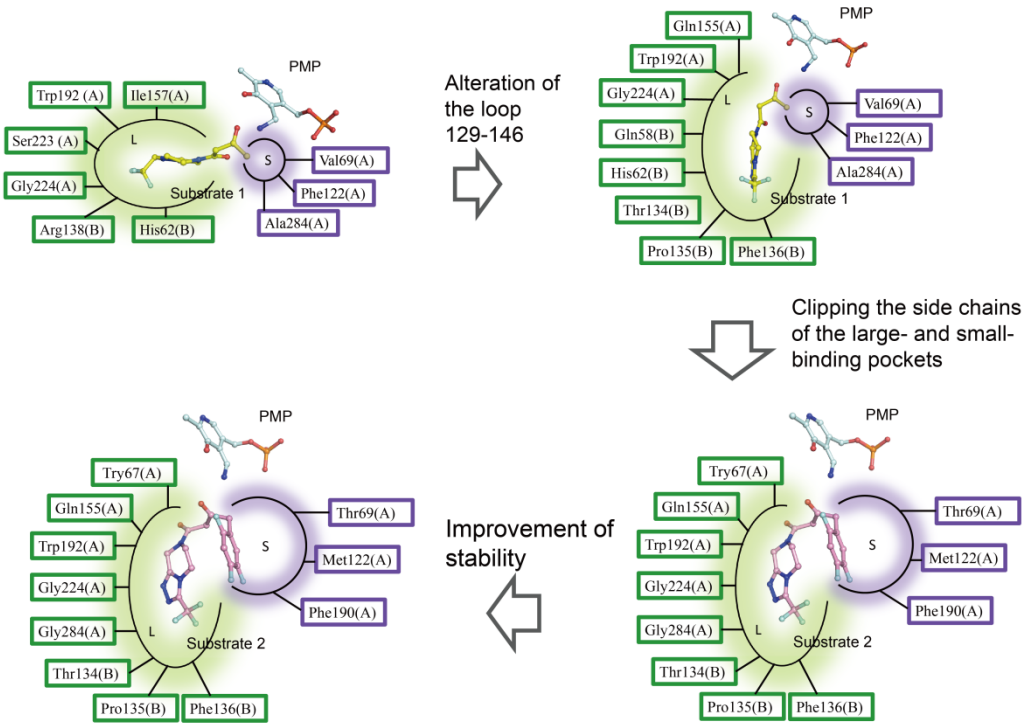


Fig. 3-14 A schematic diagram of the previously succeeded molecular engineering of ATA-117. The amino acid residues consisting of the large binding pocket (L) and small binding pocket (S) are represented in green text boxes and purple text boxes, respectively. Cofactor PMP and docked ligands are shown as sticks.

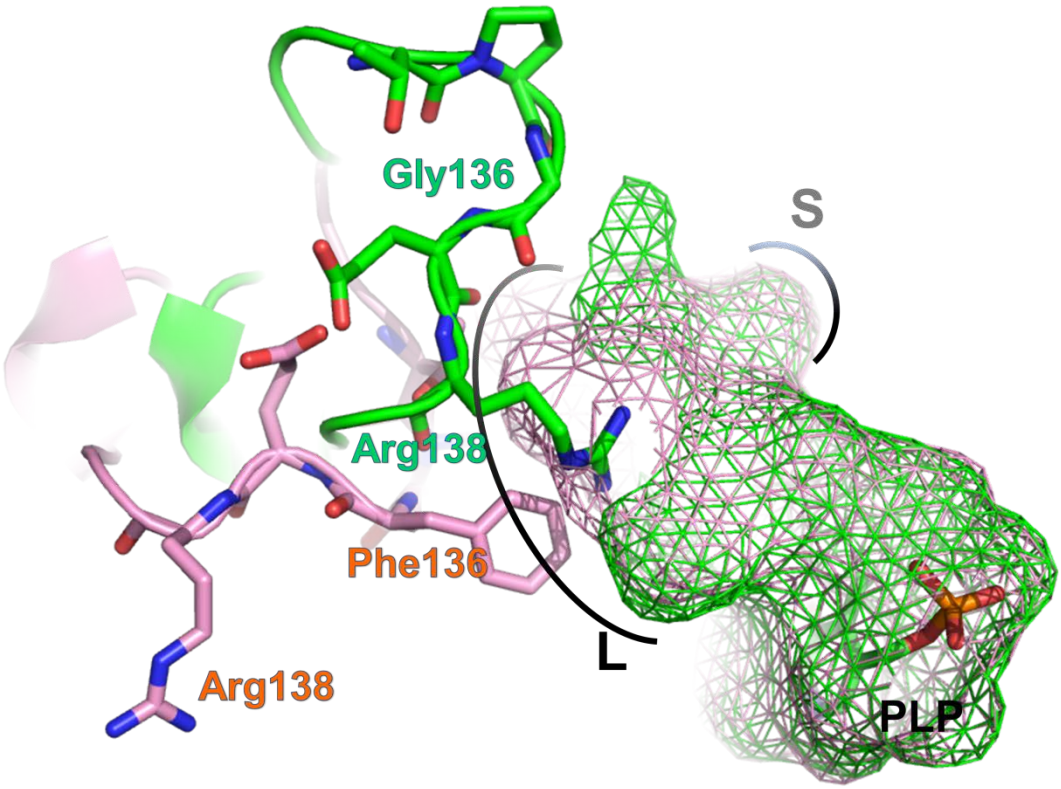


Fig. 3-15 The enlargement of the large-binding pocket (L) caused by the mutation of G136F. The active site cavity of (*R*)-TA^{WT} and (*R*)-TA^{G136F} are shown as mesh in green and pink, respectively. Cofactor PLP and residues of loop 129-146 are shown as stick. The small-binding pockets (S) are conserved between the two structures.

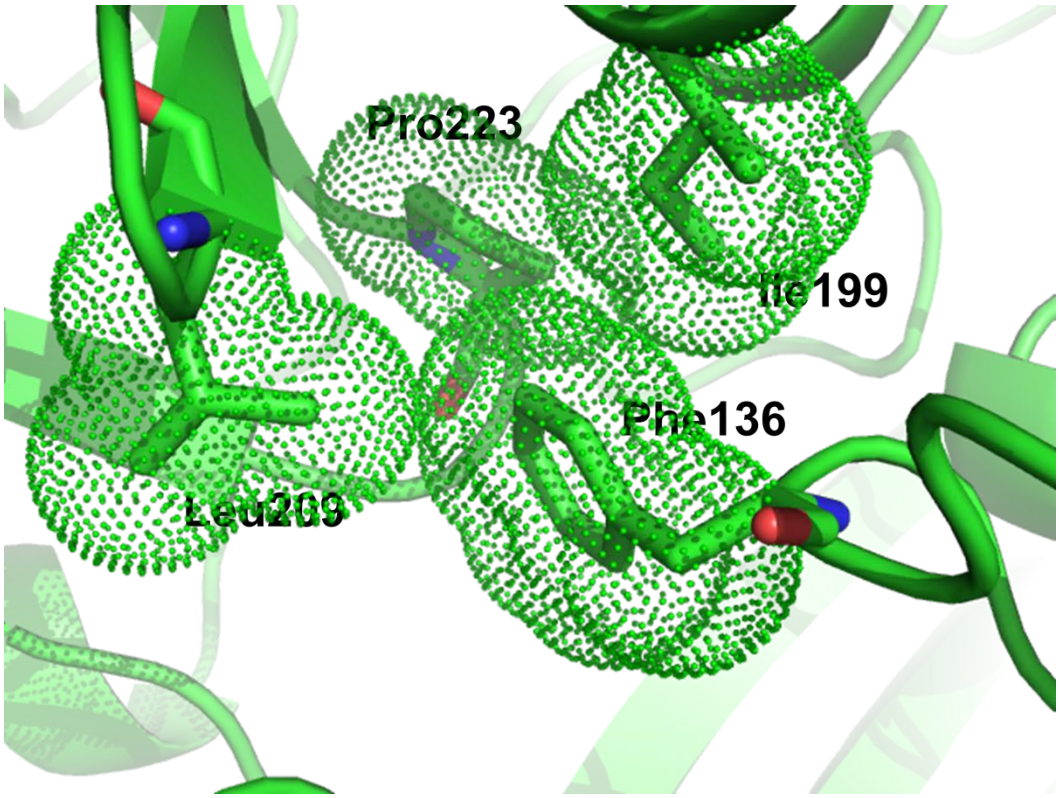


Fig. 3-16 Close-up view of the residue Phe136 in (*R*)-TA^{Mut28}. The residues Phe136, Ile199, Leu209 and Pro223 are shown as sticks. The van der Waals radii of the side chains of these four residues are represented as green dots.

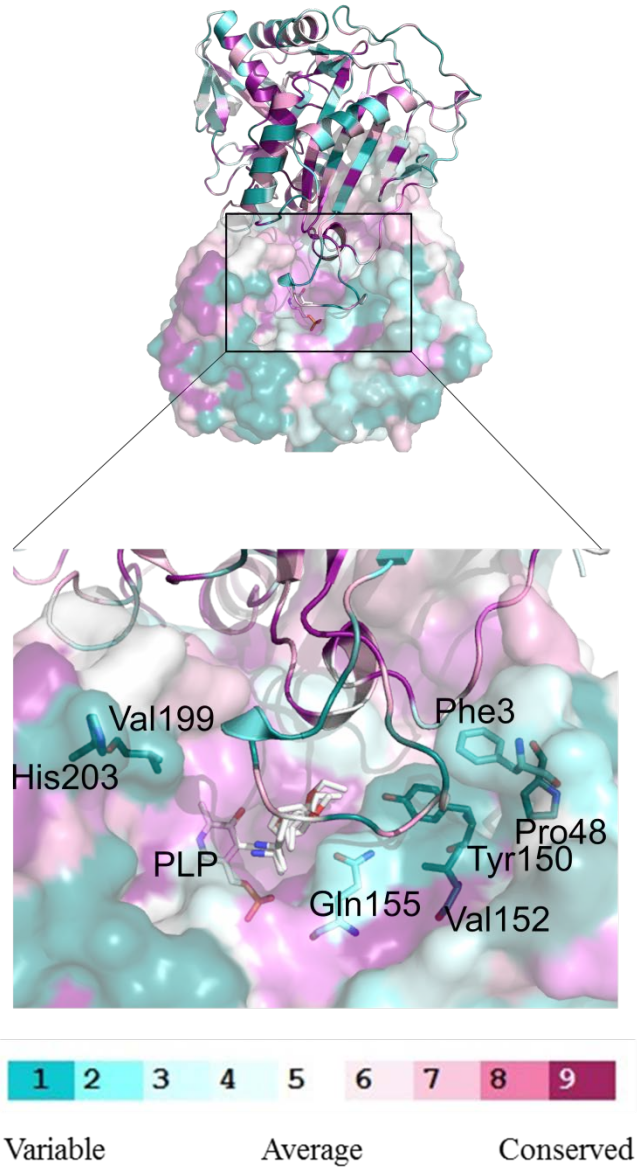


Fig. 3-17 Conservativeness of the residues in the active site and loop129-146. The color representation of conservativeness of the amino acid residues on the molecular surface of (*R*)-TA^{WT} with turquoise through maroon, indicating variable-through-conserved. The conservativeness was evaluated by ConSurf server using the aligned sequence of (*R*)-TA^{WT} and its seven homologs.

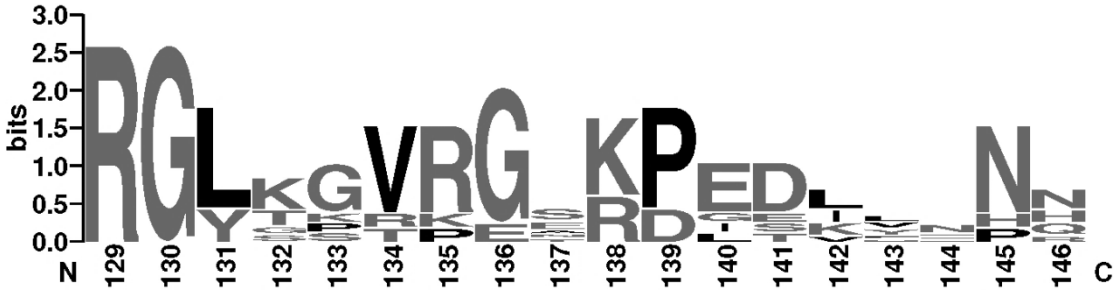


Fig. 3-18 Logographical representations of the amino acid sequence conservation within loop 129-146 in (*R*)-TA^{WT} and the seven homologs. The height of each stack corresponds to the degree of sequence conservation at that position. The height of letters represents the relative frequency of each amino acid at the position. Logo representations were obtained by using Weblogo. The seven homologs of (*R*)-TA are from *Aspergillus terreus* (GI_115385557), *Penicillium chrysogenum* (GI_211591081), *Aspergillus oryzae* (GI_169768191), *Aspergillus fumigatus* (GI_70986662), *Neosartorya fischeri* (GI_119483224), *Gibberella zeae* (GI_46109768) and *Mycobacterium vanbaalenii* (GI_120405468). These homologs are also used in Figs. 3-17 and 3-19.

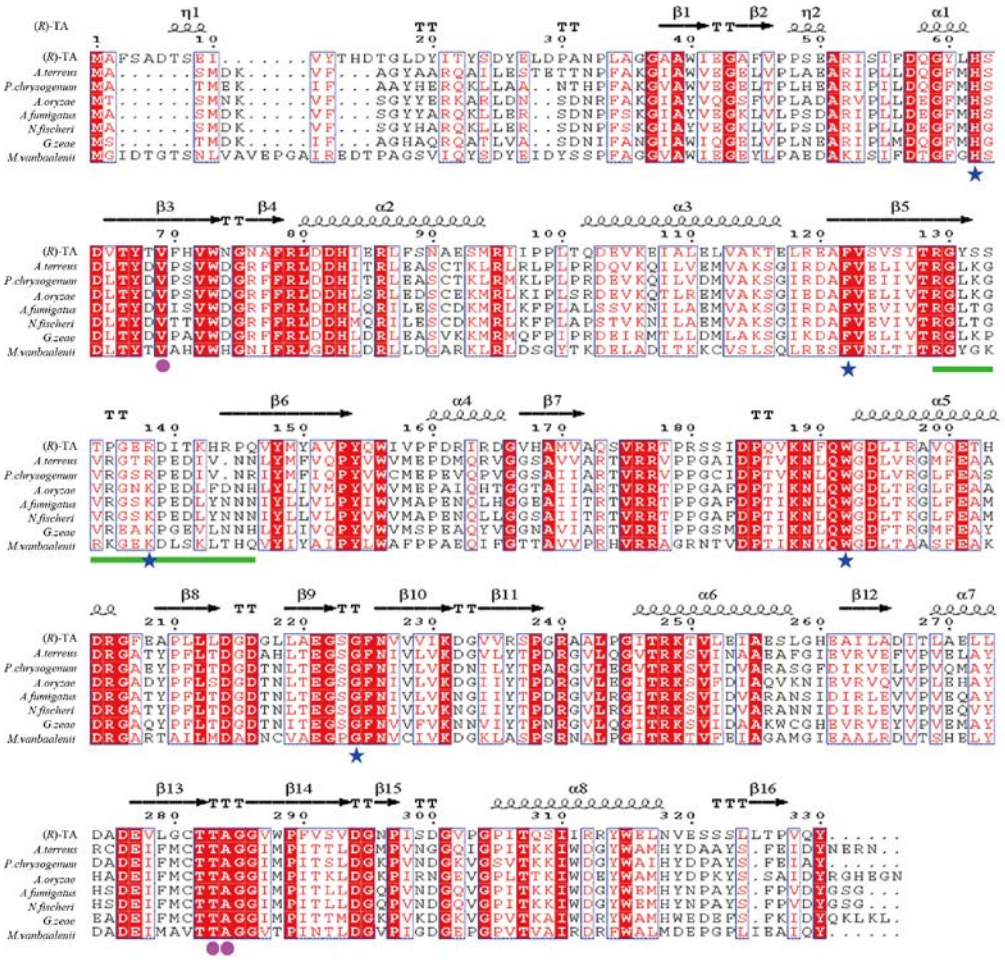


Fig. 3-19 Sequence alignment of eight (*R*)-stereoselectivity ω-transaminases. Strictly conserved residues are boxed in white on a red background and highly conserved residues are boxed in red on a white background. At the top of the sequences, schematic representations of the secondary structure elements of (*R*)-TA^{WT} are shown. The α helix is depicted by a coil and the β strand by an arrow. Alignment was generated by ClustalW and the figure was generated by ESPript. The residues in the small binding pocket, large binding pocket and loop 126-149 are highlighted by filled magenta circle, blue star and green line, respectively.

Chapter 4

Stereoselectivity of (*R*)-TA

4.1 Exordium

(*R*)-TA possesses excellent stereoselectivity toward (*R*)-enantiomers. One of the main goals of structural analysis of (*R*)-TA is to reveal the mechanism of chiral discrimination for it. *R*-MBA is a favorable amino donor of (*R*)-TA. In chapter 2, I trapped an external aldimine intermediate formed by PLP and *R*-MBA and identified the large- and small- binding pockets which may determine the stereoselectivity of (*R*)-TA. In this Chapter, I determined an external aldimine intermediate ((*R*)-TA^{SEX}) formed by PLP and non-reactive *S*-MBA using K188Q mutant of (*R*)-TA and reveal the further molecular basis of the stereoselectivity of (*R*)-TA.

4.2 Materials and methods

4.2.1 UV-visible spectroscopy of K188Q mutant, K188Q mutant mixed with *S*-MBA or *R*-MBA

The K188Q mutant preparation and detection method are same as that described in Chapter 2.

4.2.2 (*R*)-TA activity assay following inhibited by *S*-MBA

In this Chapter, (*R*)-TA activity was assayed by measuring the production of D-alanine using Thin Layer Chromatography (TLC). 0-100 mM *S*-MBA was added to the sample and incubated for 2 hours prior to activity measurement. The reaction mixture contained 10 mM *R*-MBA, 10 mM pyruvate, 100 mM Tris-HCl (pH 8.5), 0.5 mM PLP and the enzyme solution in 20 mM potassium phosphate buffer (pH 6.8) containing 20% (v/v) glycerol and 1 mM DTT. The reaction mixture was incubated at 30 ºC for 60 min and stopped by adding HCl. The amount of alanine formed was determined by TLC. The plate (Merck, TLC Silica gel 60 F_{254}) was developed using

n-butanol/acetic acid/water [4:1:1(v/v/v)] and followed by ninhydrin stain.

4.2.3 Sample preparation for crystallization

The sample of K188Q mutant was overexpressed and purified in the same way as that described in Chapter 2. The purified sample was also concentrated to 15 mg mL^{-1} by centrifugation at 4000 g using Vivaspin-20 (10000 molecular weight cutoff).

For the crystallization of the K188Q mutant of (*R*)-TA with *S*-MBA, 10 mM *S*-MBA was added to the protein solution and incubated at 293K for 6 hours. Before crystallization, the mixed sample was centrifuged at 16000 g for 5 min to remove impurities.

4.2.4 Crystallization

Cocrystallization was performed under the same crystallization conditions described in former Chapters. Streak seeding method was also used to obtain high quality crystals.

4.2.5 Data collection at Synchrotron

The X-ray diffraction data of (*R*)-TA^{SEX} complex crystals were collected at the Ar-NE3A beamline at Photon Factory (Tsukuba, Japan). The dataset was collected with a wavelength of 1.0000 Å, a rotation angle of 0.5º, an exposure time of 1.0 s for each image.

4.2.6 Data processing, determination of initial phase

The X-ray diffraction data of (*R*)-TA^{SEX} was also indexed, integrated and scaled with XDSme. Initial phasing of the diffraction data set was also conducted by molecular replacement in the same way as that described in Chapter 2.

4.2.7 Structure determination and refinement

The structure determination and refinement were performed using the same way as that described in Chapter 2. The three-dimensional coordinates of *S*-MBA was generated in MOE2011.10 (Chemical Computing Group, Montreal, Canada). The restraints dictionary of the three-dimensional coordinates of *S*-MBA was generated by Refmac5. *S*-MBA was added on the basis of $2F_o$-F_c and F_o-F_c electron density maps. The final manual fitting and structure refinement was also completed in Coot and validated with the Ramachandran plot drawn by the *PROCHECK* program.

4.3 Results and discussion

4.3.1 Sample preparation for crystallization

After incubation with the amino donor substrate *S*-MBA, K188Q mutant changed from colourless to yellow, but it is lighter than that occurred when donor substrate *R*-MBA was added into the sample. This may also imply that the UV-visible spectrum of the cofactor has changed. And it was supposed that the external aldimine had been formed.

4.3.2 UV-visible spectroscopy

Figure 4-1 showed absorption spectra of K188Q derived species. The spectrum of K188Q mutant shows a major absorption maximum at the 340 nm region and a small shoulder in the 420 nm region, indicating that the PLP is mainly in the non-covalent binding form. This was confirmed by the crystal structure of K188Q mutant (data not shown).

After the sample with *R*-MBA was incubated for 6 hours, the spectrum of K188Q mutant with *R*-MBA transferred predominantly to an absorption maximum at 420 nm,

which reflected the formation of an external aldimine formed by PLP and *R*-MBA. However, there was only a partial conversation for K188Q mutant with *S*-MBA, indicating that it is a harder or slower process for *S*-MBA to form external aldimine with PLP in K188Q mutant.

4.3.3 Crystallization

By streak seeding, polyhedron-shaped crystals grew in 3 days with approximate dimensions of 0.15 ×0.20 × 0.15 mm^3 under the crystallization conditions consisting of 0.2 M magnesium chloride, 0.1 M Bis-Tris pH 6.4, 15%(w/v) PEG3350. The crystals of K188Q mutant mixed with *S*-MBA shown yellow but lighter than that of K188Q mutant mixed with *R*-MBA, while the crystals of K188Q mutant showed colorless (Fig. 4-2).

4.3.4 Data collection, processing and determination of (*R*)-TA^{SEX} structure

Figure 4-3 shows X-ray diffraction image of the (*R*)-TA^{SEX}. The crystals of (*R*)-TA^{SEX} belonged to space group $P4_22_12$. The unit cell and the other crystallographic data collection statistics are shown in Table 4-1. Matthews coefficient value indicted that there is one molecule per asymmetric unit. The initial phase was also determined by molecular replacement using the crystal structure of (*R*)-TA^{WT} as search model. The Schiff linkage between the PLP and *S*-MBA appeared in the electron density (Fig. 4-2), which is consistent with the results of UV-visible spectroscopy. Thus, external aldimine formed by PLP and *S*-MBA was added in to the corresponding electron density. However, there was only a partial electron density for the phenyl group of *S*-MBA. Manual structure construction and refinement were finished using Refmac5 and Coot. The refinement statistics are given in Table 4-1. The Ramachandran plot of (*R*)-TA^{REX} showed that 97.8% and 2.2% of φ-ψ pairs lie in the most favored and additionally

allowed regions, respectively.

4.3.5 External aldimine binding manner in (*R*)-TA^{SEX}

In the crystal structure of (*R*)-TA^{SEX}, the cofactor PLP was in the external aldimine form with the substrate *S*-MBA, which was generated during the co-crystallization of the mutant K188Q of (*S*)-TA and the substrate. The PLP portion of the external aldimine rotated approximately 24° relative to PLP molecules of (*R*)-TA^{WT}, nearly the same degree as that in (*R*)-TA^{REX} (Figs. 4-5). There is no obvious difference of the residues around the external aldimine formed by PLP and *S*-MBA among (*R*)-TA^{WT} and (*R*)-TA^{SEX}. *S*-MBA moiety bound covalently to the C4A atom of PLP cofactor by its amino group, replacing the internal Schiff-base linkage. The methyl group of the *S*-MBA moiety was mainly bound by Val69. The aromatic ring was bound only by the main chain of Gly224 and the side chains of Ala284 via van der Waals interactions. Comparing to the external aldimine formed by PLP and R-MBA in (*R*)-TA^{REX}, the external aldimine formed by PLP and *S*-MBA in (*R*)-TA^{SEX} had fewer interactions with the around residues, especially for the phenyl group (Fig. 4-6). That may be the reason why there was a partial lack of electron density for the phenyl group of *S*-MBA.

4.3.6 *S*-MBA is an inhibitor of (*R*)-TA

Usually, ω-TAs can be developed as either kinetic resolution or asymmetric synthesis of chiral amine. In this study, crystallographical and UV-visible spectrum studies indicated that *S*-MBA could bind to the active site of (*R*)-TA. The result of TLC showed that *S*-MBA is an inhibitor of (*R*)-TA at high concentration (Fig. 4-7). The similar phenomenon was also observed at an *S*-specific ω-transaminase from *Vibrio fluvialis* JS17 which was also inhibited by *R*-enantiomer [30]. From a practical

viewpoint, the information obtained above should be given adequate attention to determine optimal reaction conditions in kinetic resolution process.

4.3.7 Stereoselectivity of (*R*)-TA

In the nature, there are some enantiocomplementary enzyme pairs that catalyze the same reaction but favor opposite enantiomers. Although they are not mirror-image molecules; they contain active sites that are functionally opposite. Previous study, according to the survey of X-ray crystal structures of enantiocomplementary enzymes, classified the enantiocomplementary enzymes into four groups as 1) different Protein folds recreate mirror-image active sites; 2) different protein folds with exchanged locations of binding sites; 3) same protein fold with exchanged locations of binding sites; and 4) same protein fold with different locations of a catalytic group (Fig. 4-9) [31]. Thus, the fold type and the binding mode of the substrate decide the stereoselectivity.

For transaminases, there are two cases. D-amino acid aminotransferase (DATA) and branched chain L-amino acid aminotransferase (BCAT) are enantiocomplementary pair. Both enzymes have the same protein fold (type IV) and position the substrate on the *re*-face of PLP. An exchange of the binding sites for the α-carboxylate and the side chain reverses the enantiopreference. In contrast, for ω-transaminases, they belong to the subgroup 1). Lysine 188 in (*R*)-TA^{WT} is located at the same position as in the fold class IV transaminases, such as BCATs and DATAs, leading proton transfer on the *re*-face of PLP. In contrast, the lysine of (*S*)-TAs is positioned on the other side (*si*-face) of the PLP in fold I and fold IV enzymes (Fig. 4-10). The substrate and catalytic lysine residue lie on opposite faces of the PLP in the two enzymes, so that the combinations of active

site and cofactor are mirror images (Fig. 4-11).

For ω-transaminases a dual recognition mode has been postulated. The bulky substituent of the amine or ketone substrate binds in the large binding pocket and there is a small binding pocket, which prohibits the binding of substituents larger than an ethyl group. The large- and small-binding pockets of (*R*)-TA identified in this study, to some extent, could determine the stereoselectivity of (*R*)-TA.

Recently, homology modelling of the (*R*)-selective-ω-TAs from *A. terreus* and *A. fumigatus* was carried out to study the active site of the (*R*)-selective-ω-TAs [32]. Two external aldimine intermediates formed between *S*-MBA or *R*-MBA and the active site PLP were modeled to the active site followed by energy minimization. In contrast to the productive external aldimine ((*R*)-external aldimine), the phenyl group of *S*-MBA in the (*S*)-external aldimine has a severe steric clash with W148 in the active site. It was supposed that this steric clash caused by *S*-MBA determines the (*R*)-stereoselectivity of the (*R*)-selective-ω-TA. However, in fact, crystallographical studies and UV-visible spectroscopy have shown that the *S*-MBA could bind to the active site of (*R*)-TA in this thesis. A different theory must be submitted to explain the stereoselectivity of (*R*)-TA.

Alignment of external aldimines formed by PLP and *R*-MBA and *S*-MBA, respectively, showed that the PLP portion of these two external aldimines overlapped well, while the MBA portion deviated (Fig. 4-8). The Schiff linkage of the external aldimine, formed by PLP and *R*-MBA, and the pyridine ring of PLP were on the same plane while several degrees was observed in the external aldimine formed by PLP and *S*-MBA. In addition, the phenyl group of *R*-MBA and *S*-MBA had similar orientation while the methyl groups of *R*-MBA and *S*-MBA were vertical. However, based on the structural information described above, good ideas were not obtained to reveal the

mechanism of stereoselectivity of (*R*)-TA.

The mechanism of stereospecificity for the hydrogen transfer between C-4’ of a bound coenzyme and the C-2 of a substrate during transamination catalyzed by pyridoxal enzyme had been previously proposed on the basis of the biochemical results (Fig. 4-12) [33]. Here I firstly confirmed the theory of stereospecificity for hydrogen transfer using crystallographic analysis (Fig. 4-13). The hydrogen at C-2 position of *R*-MBA pointed to the Try67 and Gln188 (Lys188 in wild type (*R*)-TA) residues that are important catalysis residue in transamination. This orientation is much more suitable for hydrogen transfer than that of the hydrogen at C-2 position of *S*-MBA, which had a reverse direction. This could also explain the low conversation of external aldimine for *S*-MBA. Based on the conclusion proposed above, the orientation of the hydrogen at C-2 position of MBA is a key factor to determine the chiral discrimination of (*R*)-TA, and the reaction of *S*-MBA would stop at external aldimine state.

Thus, the stereoselectivity of ω-TA occurs at the substrate binding step or external aldmine step if the substrate could bind to the active site.

Table 4-1 Data collection and refinement statistics of (*R*)-TA^{SEX}

	(*R*)-TA^{SEX}
Data collection	
Beamline	AR-NE3A
Wavelength (Å)	1.0000
Space group	$P\,4_22_12$
Cell dimensions	81.07, 81.07, 93.67
Cell angles (°)	90.00, 90.00, 90.00
Resolution (Å)	50.0-1.72
Number of observations	479155
Number of unique reflections	33616
Completeness (%)	99.8 (99.0)
R_{merge} (%)	6.3 (39.7)
Redundancy	14.25
I/σ (I)	35.84 (8.10)
Refinement	
Resolution range (Å)	48.76-1.72
R_{work}/R_{free} (%)	15.1/18.7
Number of non-hydrogen atoms	
Protein	2518
Water	476
Ligand	external aldimine
	24
RMSD bond length (Å)	0.020
RMSD bond angle (º)	1.965
Ramachandran plot (%)	
Preferred	97.8
Allowed	2.2
Outliers	0.0

The values in parentheses are for the highest resolution shell

$R_{work}=\Sigma||Fo|-|Fc||\ /\Sigma|Fo|$, where $|Fo|$ and $|Fc|$ are the observed and calculated structure factor amplitudes of a particular reflection.

R_{free} was calculated from 5% of reflections omitted from the refinement.

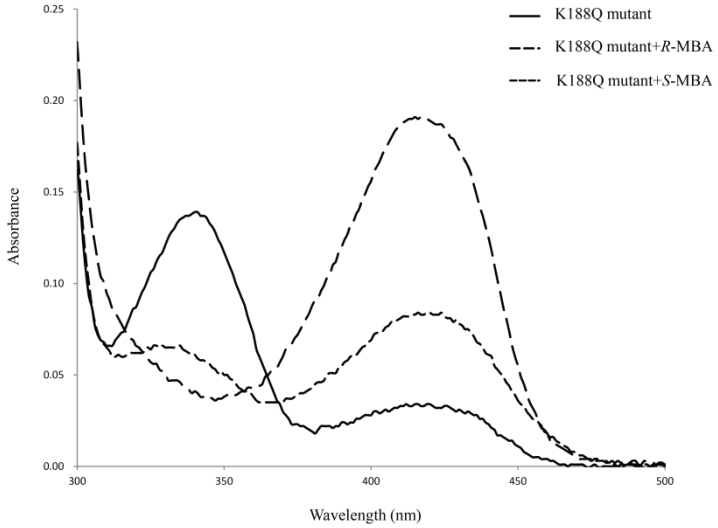


Fig. 4-1 Absorption spectra changes of K188Q derivatives (A) and (*R*)-TA (B) with *R*-MBA or *S*-MBA.

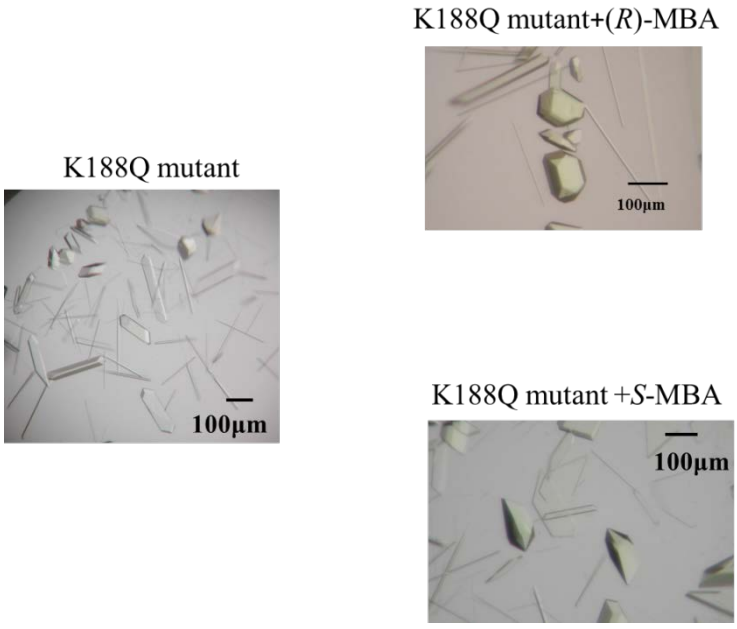


Fig. 4-2 Crystal pictures were took without a polarizing plate. The crystals of K188Q mutant mixed with *S*-MBA also shown yellow while the crystals of K188Q mutant shown colorless.

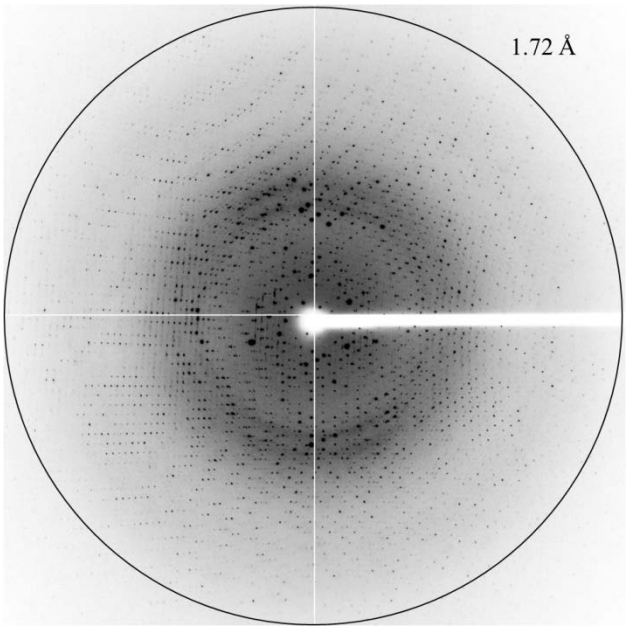


Fig. 4-3 An X-ray diffraction image collected from a crystal of (*R*)-TA^{SEX}.

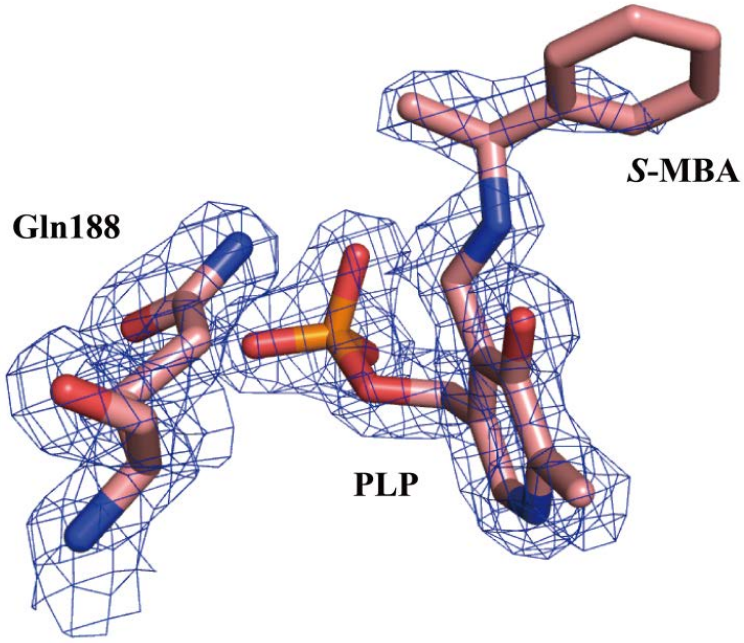


Fig. 4-4 F_o-F_c omit density map of Gln188 and the external aldimine formed by PLP and *S*-MBA, contoured at 1.5 σ resolution.

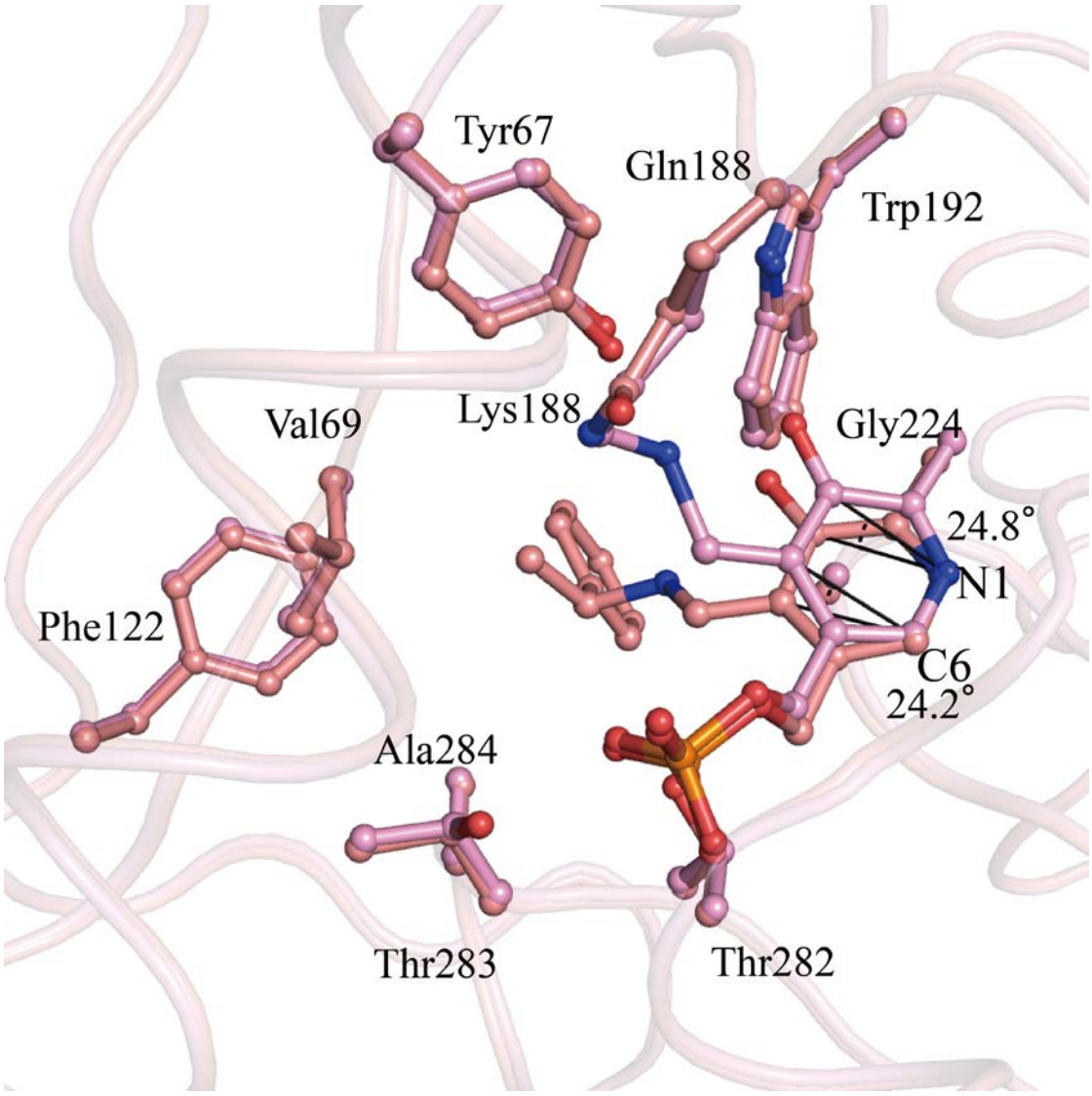


Fig. 4-5 Superposition of ligand binding pockets in one protomer of (*R*)-TA^{WT} (pink) and (*R*)-TA^{SEX} with external aldimine (salmon). The PLP portion of the external aldimine rotated approximately 25° relative to PLP molecules of (*R*)-TA^{WT}.

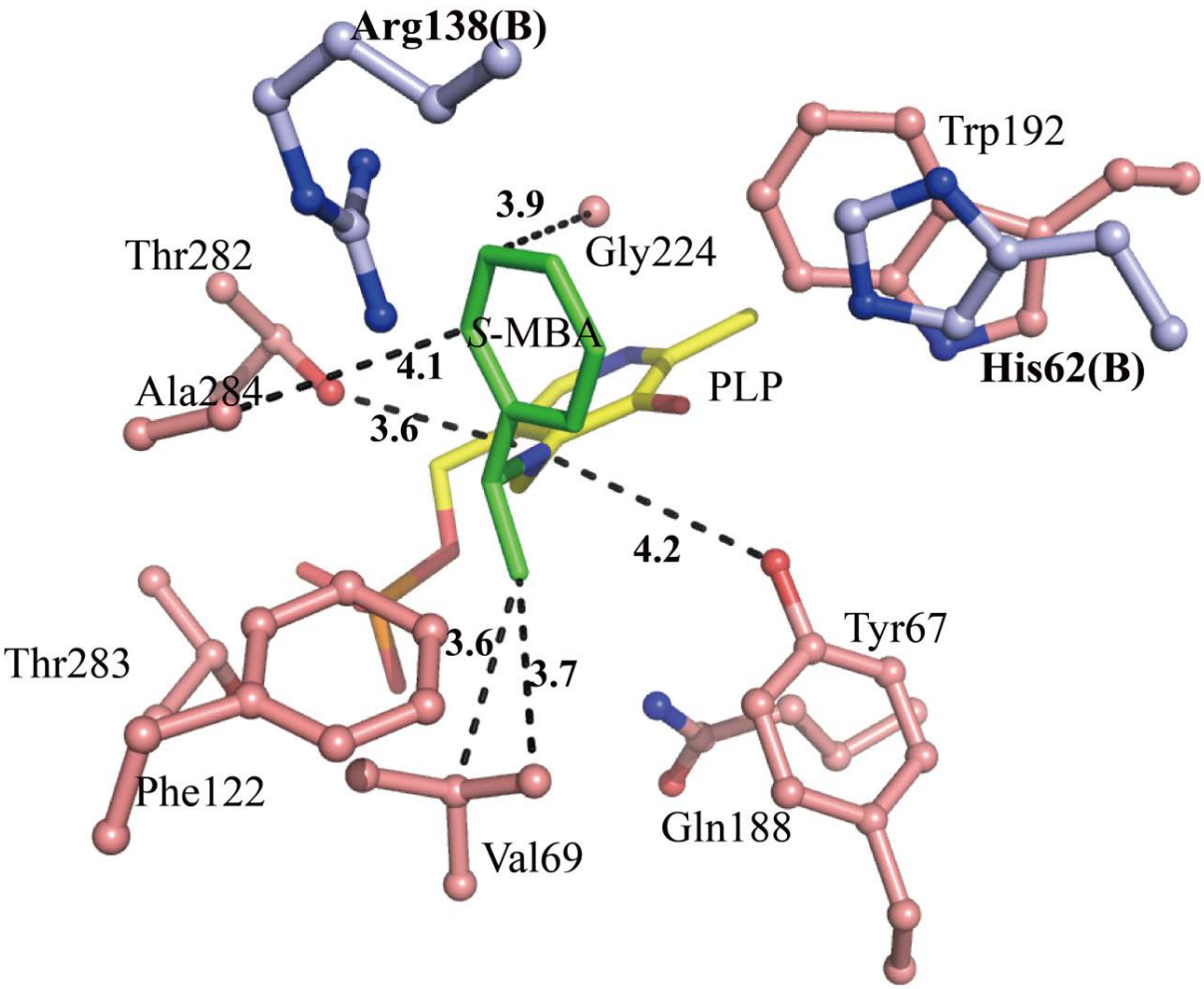


Fig. 4-6 Binding mode of external aldimine in (*R*)-TA^{SEX} consisted of PLP (stick in yellow) and *S*-MBA (stick in green). The interactions are shown as black dashed lines. Residues binding the external aldimine are shown as ball-stick model. The His62 and Arg138 residues from adjacent protomer are shown in light blue.

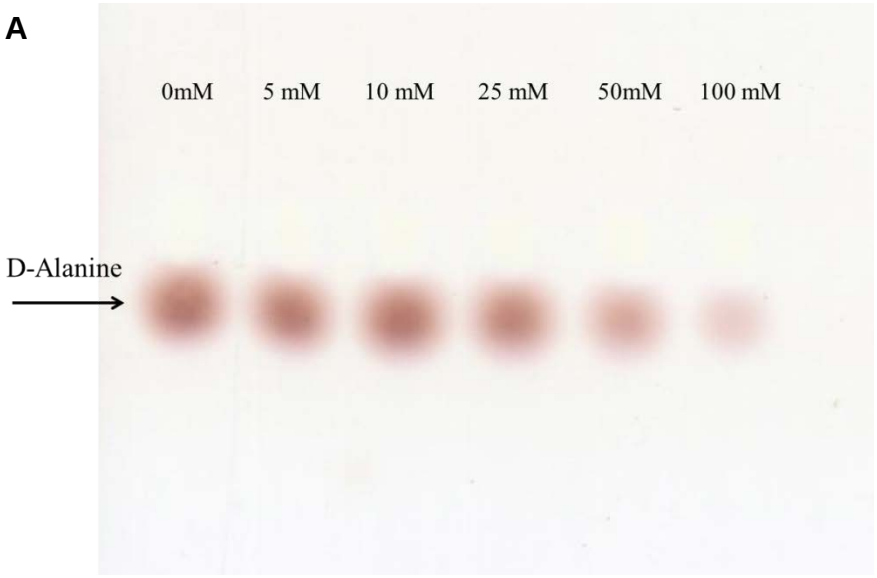


Fig. 4-7 TLC of (*R*)-TA activity assay following inhibited by *S*-MBA at different concentration. The ninhydrin spots represent reaction product, D-Alanine.

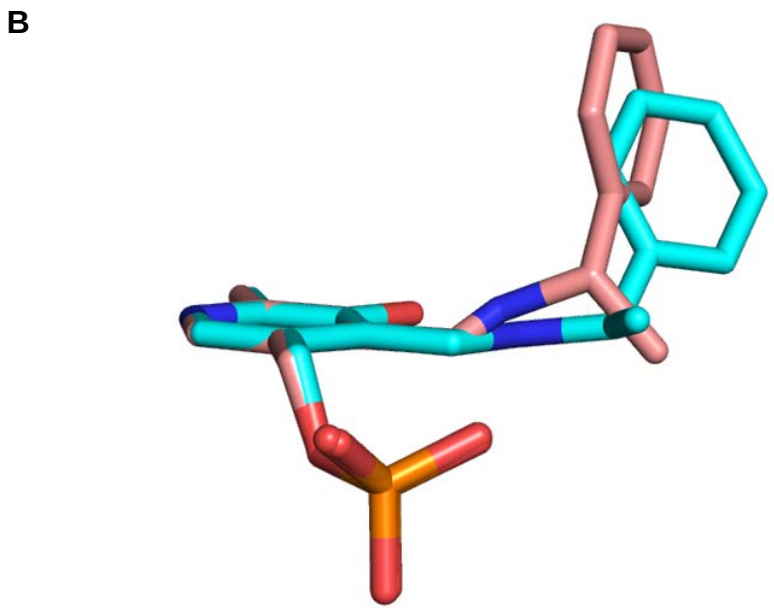


Fig. 4-8 Comparison of external aldimine formed by PLP and *R*-MBA (cyan) and *S*-MBA (salmon), respectively.

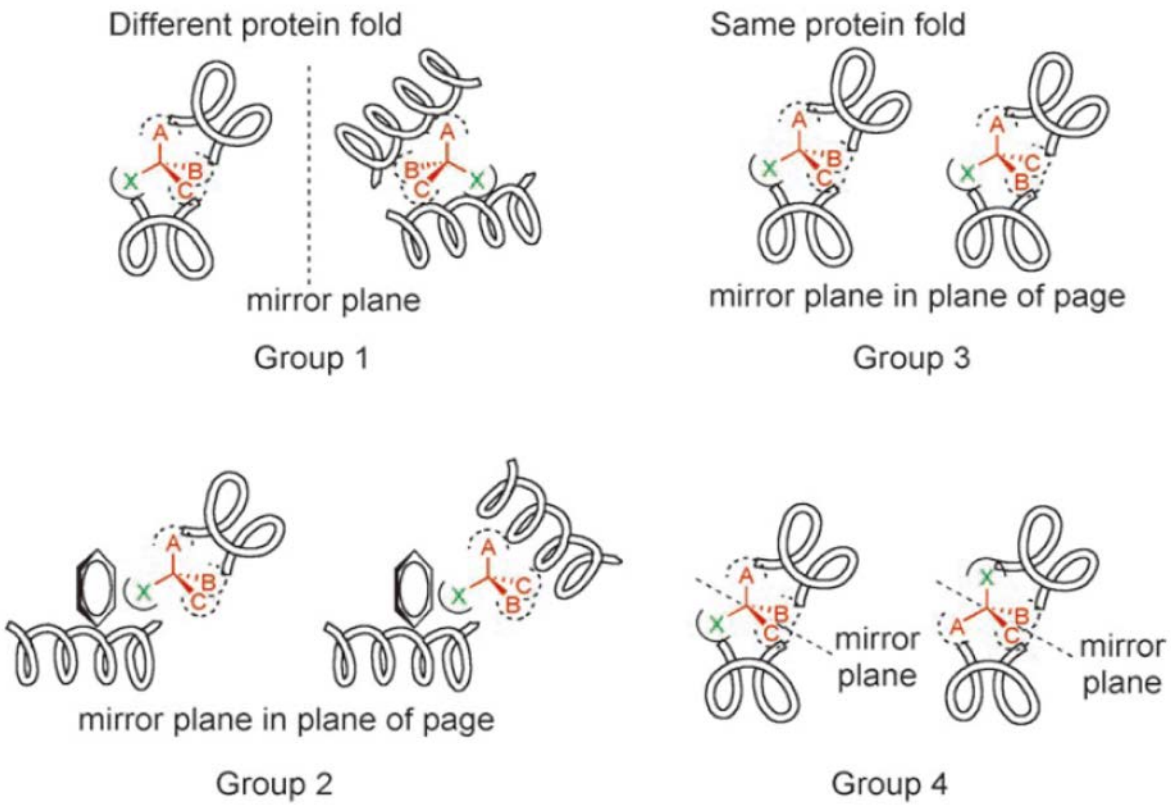


Fig. 4-9 Classification of enantiocomplementary enzyme pairs according to the relationship between the mirror-image active sites. X represents a substituent directly involved in bond making and bond breaking; A, B, and C represent nonreacting substituents. The benzene ring in the schematic illustration for group 2 represents a cofactor,such as flavin, pyridoxal, or nicotinamide. [31]

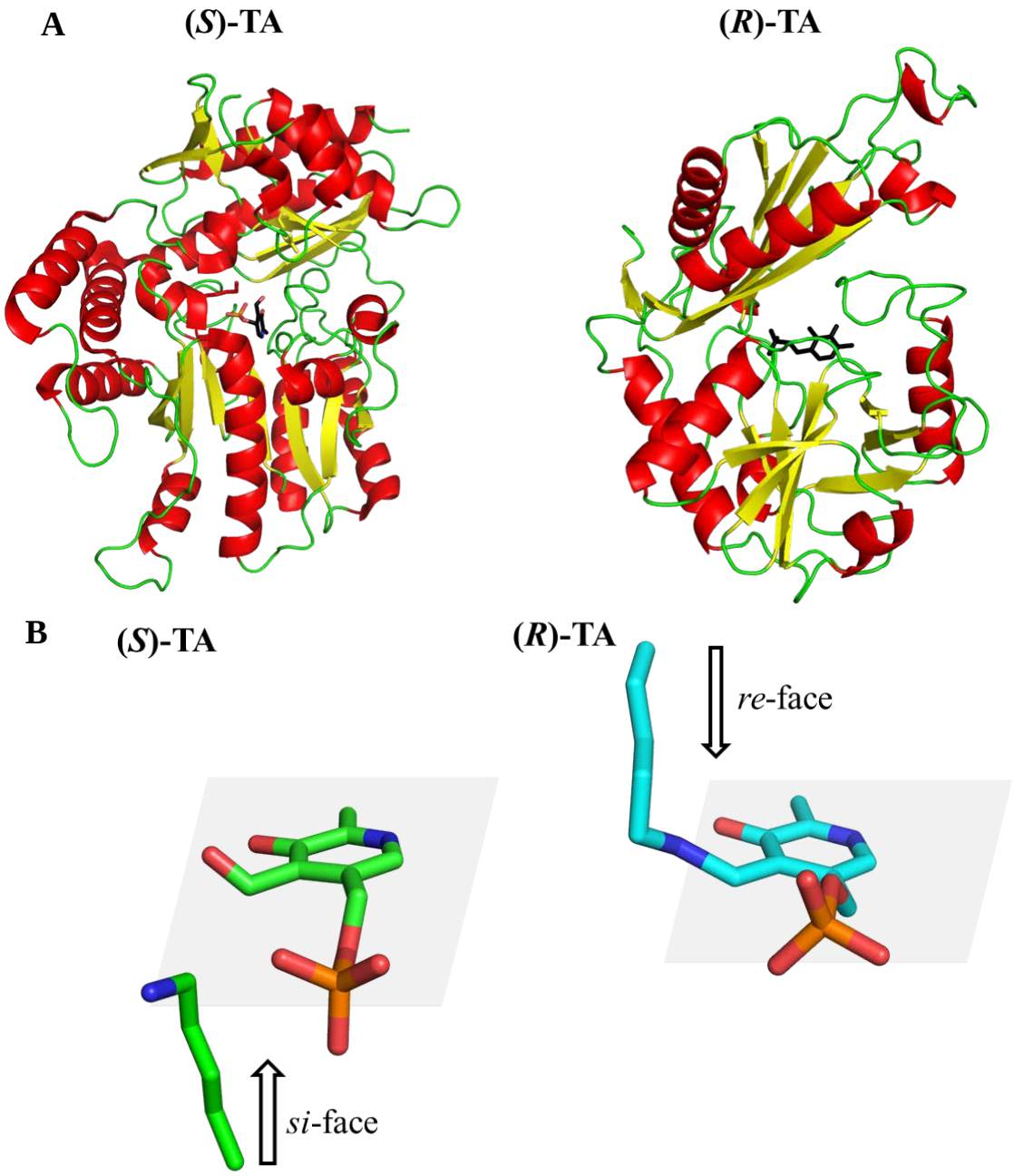


Fig. 4-10 (A) Fold type comparison of (*R*)-TA and an (*S*)-TA from *Chromobacterium violaceum* (PDB entry, 4B9B). The monomers are shown as a ribbon diagram; α-helices are shown in red, β-strands in yellow and loops in green. PLP is shown as stick models. (B) The catalytic lysine residues of the (*S*)-TA and (*R*)-TA are positioned on the *si*-face and *re*-face of cofactor PLP, respectively.

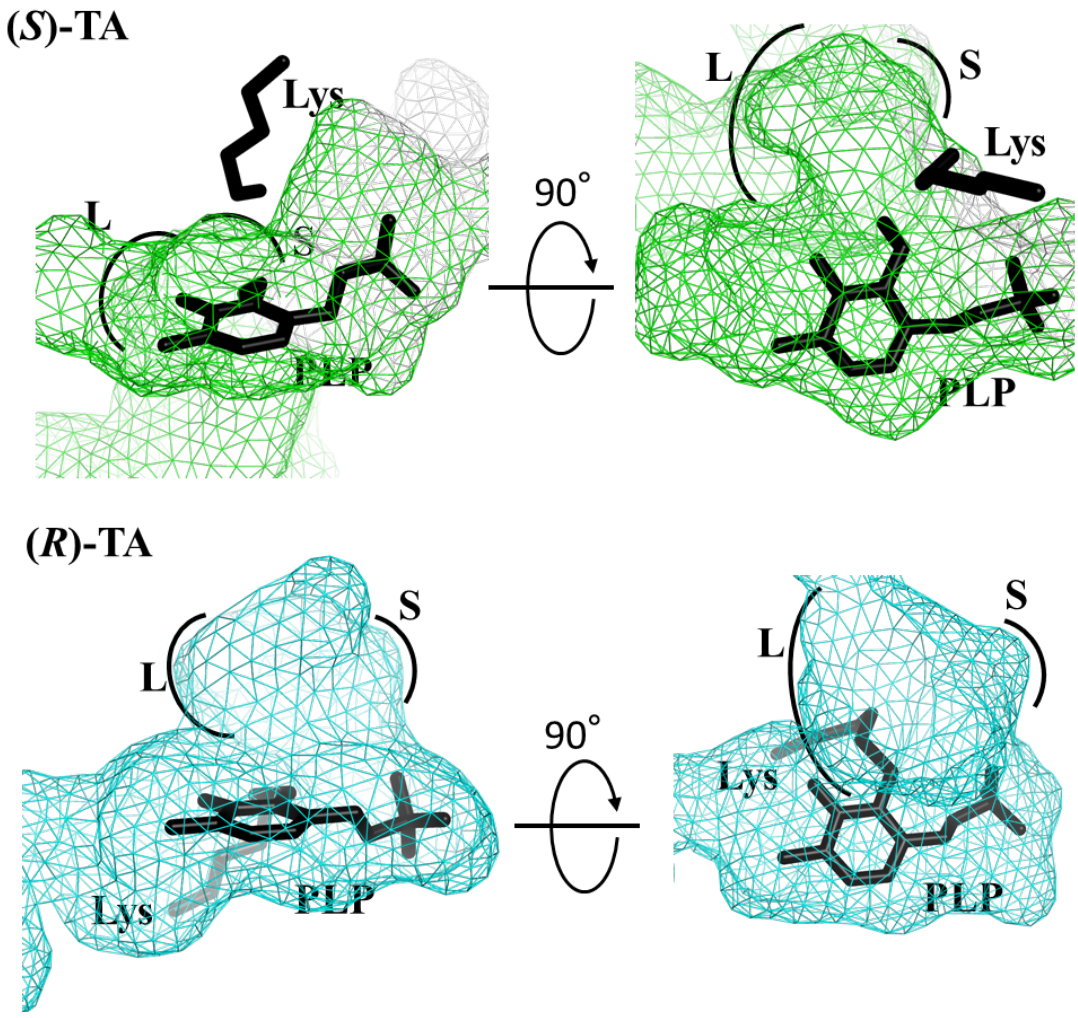


Fig. 4-11 The mesh represents the cavity of the active site of (*S*)-TA (up) and (*R*)-TA (below). The right panels are rotated along the horizontal axis from the left panels. The PLP and the side chain of Lys are shown as stick models embedded in the cavity. Large binding pockets (L) and small binding pockets (S) are labeled.

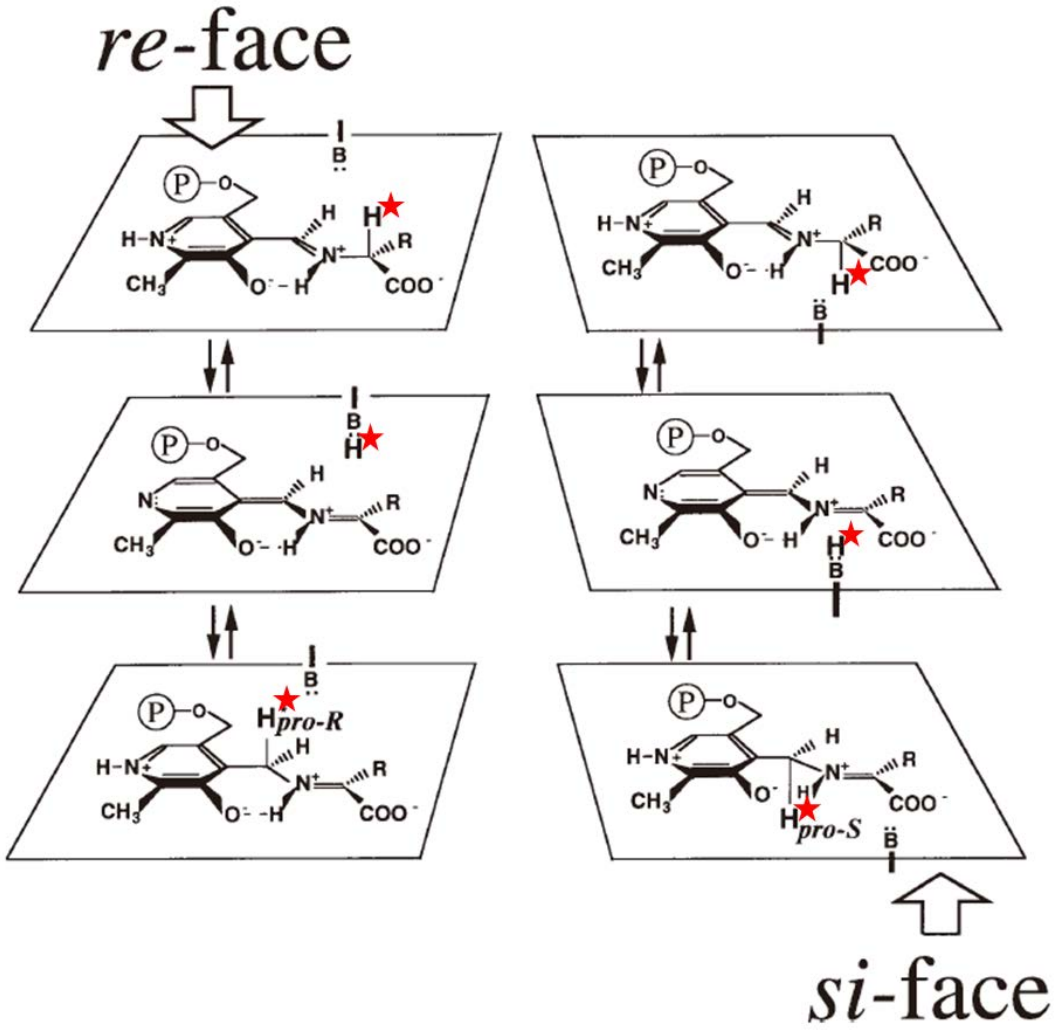


Fig. 4-12 Stereospecificity for the hydrogen transfer in transaminases. B represents the catalytic residue lysine. The hydrogen atoms involved in the hydrogen transfer are labeled with a red star. [33]

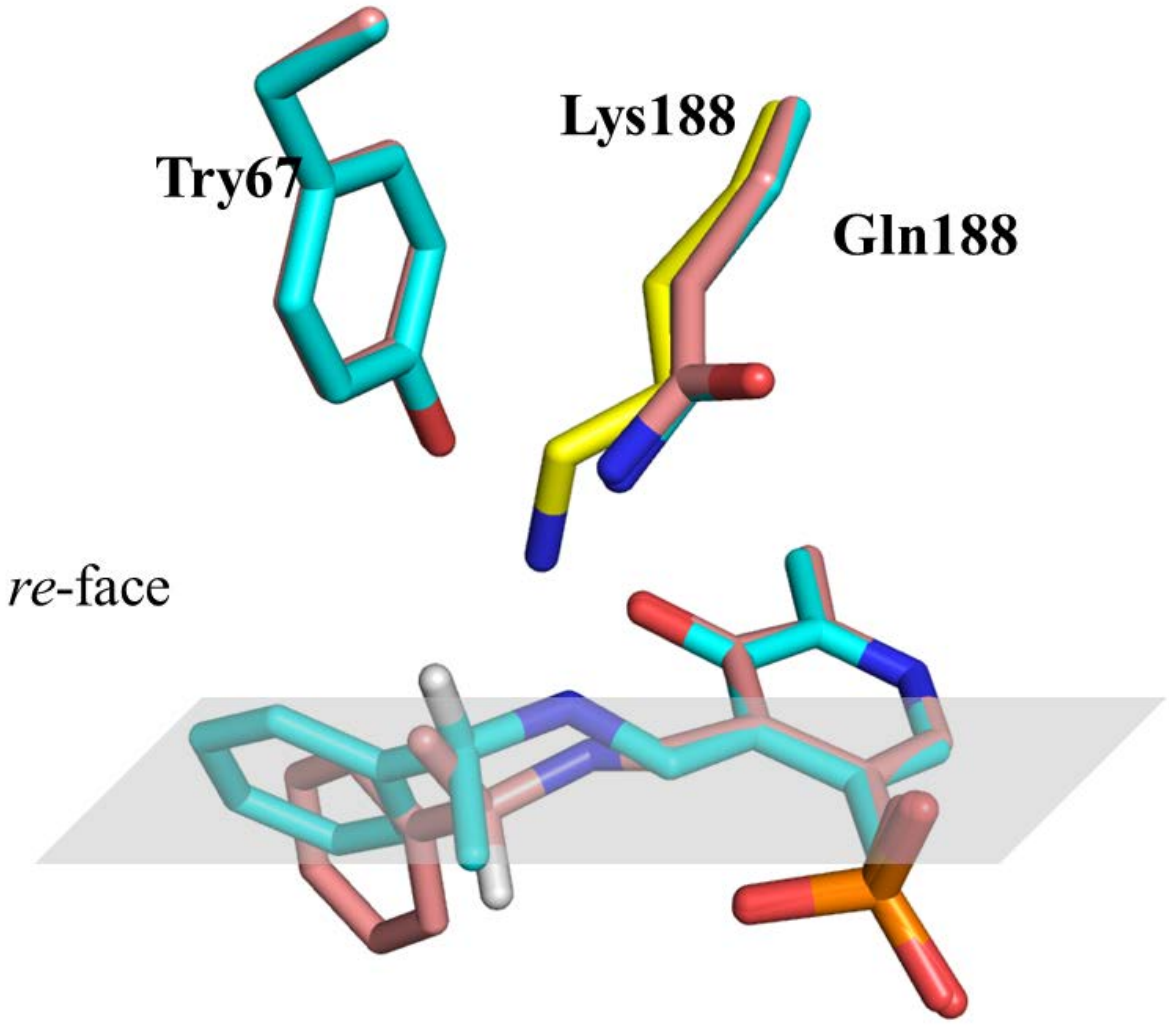


Fig. 4-13 Proposed mechanism of stereoselectivity of (*R*)-TA. The Try67, Gln188 and external aldimine intermediate of (*R*)-TA^{REX} and (*R*)-TA^{SEX} are shown as stick in cyan and salmon, respectively. The hydrogen atom (white stick) at C-2 position of *R*-MBA pointed to the *re*-face while that at C-2 position of *S*-MBA point to the *si*-face. The Lys188 from (*R*)-TA^{WT} is shown as yellow stick.

Conclusions

I purified, crystalized and determined the structure of the (*R*)-TA^{WT} from *Arthrobacter* sp. KNK168. (*R*)-TA^{WT} showed typical structural architecture of the fold-type IV class of PLP-dependent enzymes, which is constituted by two domains and the linker loop connecting these two domains. (*R*)-TA existed as homodimers in crystals and possessed the active site between the two protomers. The active site was formed by the cavity accommodating the cofactor PLP and a glycerol molecule from protein storage buffer.

To clarify the substrate recognition mechanism, I also solved the crystal structures of two ketone substrate complexes ((*R*)-TA^{PYR} and (*R*)-TA^{METH}); the PMP-bound form ((*R*)-TA^{PMP}) and the external aldimine intermediate ((*R*)-TA^{K188Q_REX}) formed by the substrate *R*-MBA and cofactor PLP. The two ketone substrate complexes revealed the binding mode of the ketone substrates and that the Arg138 is a key residue for recognizing ketone substrate of (*R*)-TA. In (*R*)-TA^{PMP}, cofactor PLP and catalytic residue Lys188 showed conformation change comparing those in the structure of (*R*)-TA^{WT}. These conformation changes are probably important for catalyzing the next half-reaction of transamination. In addition, alignment of the ketone substrate complexes and PMP intermediate showed an approximate model of PMP-amino acceptor intermediate. (*R*)-TA^{K188Q_REX} intermediate confirmed a hypothesis previously suggested that the active site of ω-amine transaminase consists of a small binding pocket and a large binding pocket.

To illustrate structural basis of the successful molecular engineering performed before, I solved the crystal structures the G136F mutant in complex with PLP ((*R*)-TA^{G136F}) and the mutant with 28 mutations (Mut28), which was identical to the previously engineered mutant ((*R*)-TA^{Mut28}). (*R*)-TA^{Mut28} showed enlarged binding

pockets. In addition, loops 129-146 in (*R*)-TA^{Mut28} and (*R*)-TA^{G136F} show significantly different conformation in that bulky side chain of 136th residue causes conformational alteration of loop 129-146. Conformation change of this loop results in the drastically modified substrate-binding pocket in volume, shape, and surface property. Docking simulations of Substrate 1 and Substrate 2 using crystal structures in this thesis were performed to explain how their original and improved enzymes recognize these compounds. Three docking models revealed the structure basis of the key mutations for changing substrate specificity of ATA-117.

Finally, I resolved a structure of the K188Q mutant of (*R*)-TA complexed with a non-reactive amino donor substrate (*S*-MBA). The results of UV-visible spectroscopy also indicated that *S*-MBA could bind to the K188Q mutant. The results of TLC implied that *S*-MBA is an inhibitor of (*R*)-TA. Based on these results, I discussed the stereoselectivity of (*R*)-TA. The large- and small-binding pockets of (*R*)-TA, to some extent, could determine the stereoselectivity of (*R*)-TA at substrate binding step. In addition, the orientation of the hydrogen at C-2 position of MBA is a key factor to determine the chiral discrimination of (*R*)-TA, and the reaction of *S*-MBA which can bind to (*R*)-TA would stop at external aldimine state. Thus, the stereoselectivity of ω-TA occurs at the substrate binding step or external aldmine step if the substrate could bind to the active site.

References

1. G. L. Plosker, D. P. Figgitt, Repaglinide. *Pharmacoeconomics* **22**, 389-411 (2004)

2. J. Eron *et al*., The klean study of fosamprenavir-ritonavir versus lopinavir-ritonavir, each in combination with abacavir-lamivudine, for initial treatment of HIV infection over 48 weeks: a randomised non-inferiority trial. *Lancet* **368**, 476-482 (2006)

3. C. Chen, Physicochemical, pharmacological and pharmacokinetic properties of the zwitterionic antihistamines cetirizine and Levocetirizine. *Curr. Med. Chem.* **15**, 2173-2191 (2008)

4. G. A. Block *et al*., Cinacalcet for secondary hyperparathyroidism in patients receiving hemodialysis. *N. Engl. J. Med*. **350**, 1516-1525 (2004).

5. M. Höhne, U. T. Bornscheuer, Biocatalytic routes to optically active amines. *ChemCatChem* **1**, 42-51 (2009).

6. J. S. Shin, B. G. Kim, Asymmetric synthesis of chiral amines with ω-transaminase. *Biotechnol. Bioeng.* **65**, 206-211 (1999).

7. P. P. Taylor, D. P. Pantaleone, R. F. Senkpeil, I. G. Fotheringham, Novel biosynthetic approaches to the production of unnatural amino acids using transaminases. *Trends Biotechno.* **16**, 412-418 (1998).

8. J. S. Shin, B. G. Kim, Exploring the active site of amine:pyruvate aminotransferase on the basis of the substrate structure−reactivity relationship: how the enzyme controls substrate specificity and stereoselectivity. *J. Organic. Chem.* **67**, 2848-2853 (2002).

9. D. Koszelewski, K. Tauber, K. Faber, W. Kroutil, ω-Transaminases for the synthesis of non-racemic α-chiral primary amines. *Trends Biotechnol.* **28**, 324-332 (2010).

10. M. S. Humble *et al*., Crystal structures of the *Chromobacterium violaceum* ω-transaminase reveal major structural rearrangements upon binding of coenzyme PLP. *FEBS. J.* **279**, 779-792 (2012)

11. C. Rausch, A. Lerchner, A. Schiefner, A. Skerra, Crystal structure of the ω-aminotransferase from Paracoccus denitrificans and its phylogenetic relationship with other class III amino-transferases that have biotechnological potential. *Proteins* **81**, 774-787 (2013).

12. C. Sayer, M. N. Isupov, A. Westlake, J. A. Littlechild, Structural studies of Pseudomonas and Chromobacterium ω-aminotransferases provide insights into their differing substrate specificity. *Acta Crystallogr. D* **69**, 564-576 (2013).

13. A. Iwasaki, Y. Yamada, Y. Ikenaka, J. Hasegawa, Microbial synthesis of (*R*)- and (*S*)-3,4-dimethoxyamphetamines through stereoselective transamination. *Biotechnol. Lett*. **25**, 1843-1846 (2003).

14. A. Iwasaki, Y. Yamada, N. Kizaki, Y. Ikenaka, J. Hasegawa, Microbial synthesis of chiral amines by (*R*)-specific transamination with *Arthrobacter sp.* KNK168. *Appl Microbiol Biotechnol* **69**, 499-505 (2006)

15. A. Iwasaki, K. Matsumoto, J. Hasegawa, Y. Yasohara, A novel transaminase, (*R*)-amine:pyruvate aminotransferase, from *Arthrobacter sp.* KNK168 (FERM BP-5228): purification, characterization, and gene cloning. *Appl Microbiol Biotechnol* **93**, 1563-1573 (2012).

16. C. K. Savile *et al*., Biocatalytic asymmetric synthesis of chiral amines from ketones applied to sitagliptin manufacture. *Science* **329**, 305-309 (2010)

17. M. Höhne, S. Schätzle, H. Jochens, K. Robins, U. T. Bornscheuer, Rational assignment of key motifs for function guides in silico enzyme identification. *Nat Chem Biol* **6**, 807-813 (2010)

18. W. Kabsch, XDS. *Acta Crystallogr. D* **66**, 125-132 (2010).

19. A. Vagin, A. Teplyakov, Molecular replacement with MOLREP. *Acta Crystallogr. D*

66, 22-25 (2010).

20. M. D. Winn *et al*., Overview of the CCP4 suite and current developments. *Acta Crystallogr. D* **67**, 235-242 (2011).

21. K. Cowtan, The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. D* **62**, 1002-1011 (2006).

22. S. X. Cohen *et al*., Towards complete validated models in the next generation of ARP/wARP. *Acta Crystallogr. D* **60**, 2222-2229 (2004).

23. G. N. Murshudov *et al*., REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D* **67**, 355-367 (2011).

24. P. D. Adams *et al*., PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213-221 (2010).

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

26. T. Clausen, R. Huber, L. Prade, M. C. Wahl, A. Messerschmidt, Crystal structure of Escherichia coli cystathionine [gamma]-synthase at 1.5 Å resolution. *EMBO J*. **17**, 6827-6838 (1998).

27. M. S. Malik, E. S. Park, J. S. Shin, Features and technical applications of ω-transaminases. *Appl. Microbiol. Biot.* **94**, 1163-1171 (2012).

28. F. S. Munsberg *et al*., Revealing the structural basis of promiscuous amine transaminase cctivity. *ChemCatChem* **5**, 154-157 (2013).

29. S. Schätzle *et al*., Enzymatic asymmetric synthesis of enantiomerically pure aliphatic, aromatic and arylaliphatic amines with (*R*)-selective amine transaminases. *Adv. Synth. Catal.* **353**, 2439-2445 (2011).

30. J. S. Shin, B. G. Kim, Substrate inhibition mode of ω-transaminase from

Vibrio fluvialis JS17 is dependent on the chirality of substrate. *Biotechnol. Bioeng*. **77**, 832-837 (2002)

31 P. F. Mugford, U. G. Wagner, Y. Jiang, K. Faber, R. J. Kazlauskas, Enantiocomplementary enzymes: classification, molecular basis for their enantiopreference, and prospects for mirror-image biotransformations. *Angew. Chem. Int. Ed.* **47**, 8782-8793 (2008)

32. E. S. Park, J. Y. Dong, J. S. Shin, Active site model of (*R*)-selective 5-transaminase and its application to the production of D-amino acids. *Appl. Microbiol. Biotechnol.* **98**, 651-660 (2014)

33. K. Soda, T. Yoshimura, N. Esaki, Stereospecificity for the hydrogen transfer of pyridoxal enzyme reactions. *The Chemical Record*. **1**, 373-384 (2001)

Acknowledgments

I would really like to thank my supervisor Prof. Tanokura Masaru. I offer my sincere appreciation and gratitude to his patient advice, warm help and revision of the papers. I received many helpful comments and useful advice to successfully finish this work. I do believe that the experience I learned from Prof. Tanokura Masaru will help me a lot for my future research.

I am grateful to Dr. Koji Nagata, Associate Professor in the Tanokura's laboratory at the University of Tokyo and Dr. Takuya Miyakawa, Project Assistant Professor in Tanokura's laboratory at the University of Tokyo for their comments and advice. I also would like to express my sincere thanks to Dr. Jun Ohtsuka, Project Assistant Professor in Tanokura's laboratory at the University of Tokyo, for his precious advice and suggestions on my research; for explaining various techniques and even teaching the method to conduct research to me; for his patient to discuss with me whenever I met problems.

I am also grateful to Dr. Masahiko Okai, Ms. Tomoko Mase for teaching me a lot of basics at structural biology, which are indispensable for this work, when I was a beginner. I am also deeply indebted to all the members in Tanokura's laboratory. It is a wonderful time for me with these members in my graduate study.

I am also most grateful to the collaborators, Dr. Noriyuki Ito and Dr. Yoshihiko Yasohara, at Frontier Biochemical and Medical Research Laboratories of Kaneka Corporation for sharing me with their experiences on determination of enzymatic activities and providing the data and plasmids. I also thank the beamline staff at the Photon Factory for their kind help with data collection.

Finally, I owe my deepest gratitude to my family and the financial support of MEXT of Japan that made it possible to complete my doctoral thesis.