Three-dimensional Structure of a Highly Thermostable Enzyme, 3-Isopropylmalate Dehydrogenase

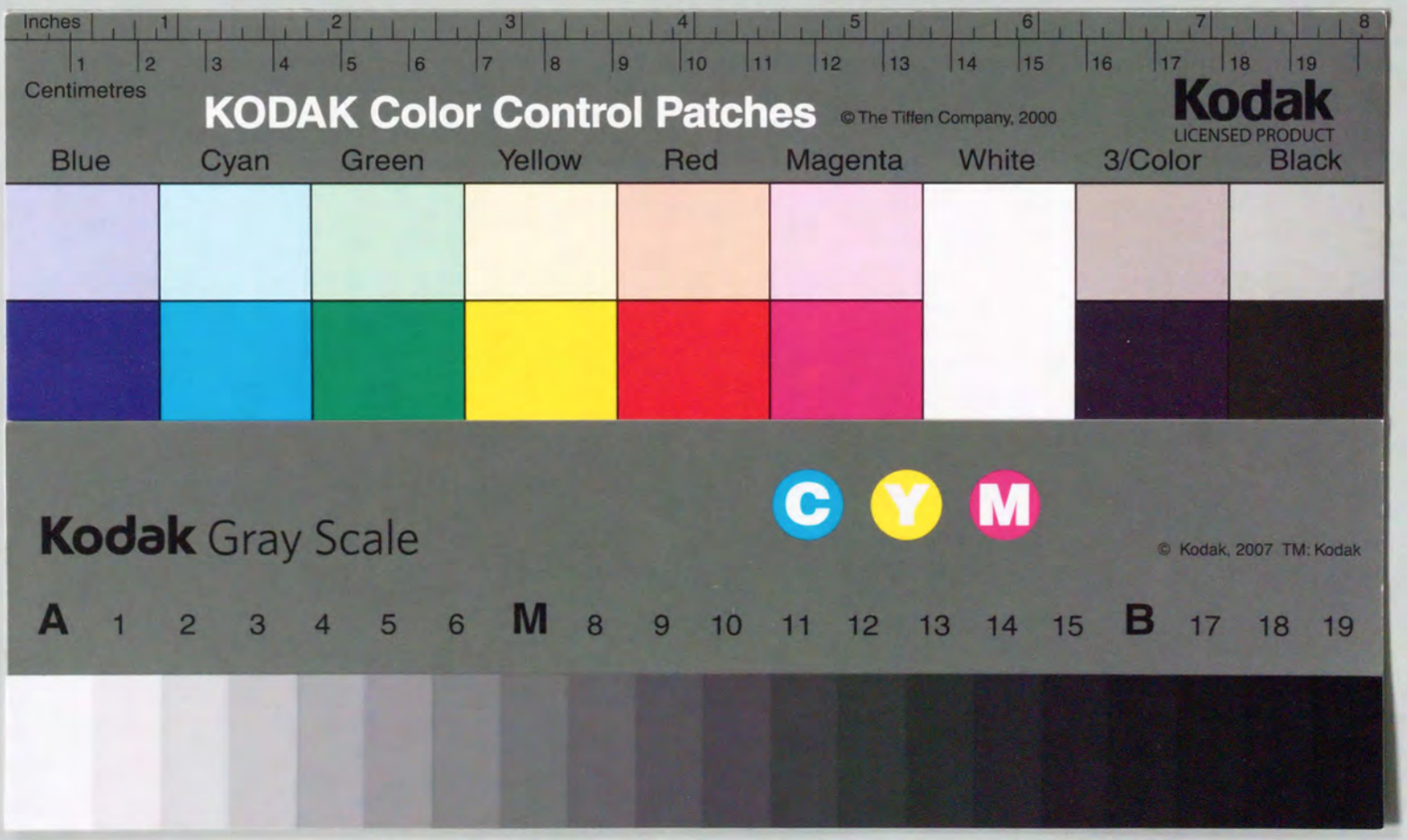
A Doctoral Thesis

by

Katsumi Imada

Submitted to the Faculty
of Science, Osaka University

February, 1992



Three-dimensional Structure of a Highly Thermostable Enzyme, 3-Isopropylmalate Dehydrogenase

A Doctoral Thesis

by

Katsumi Imada

Submitted to the Faculty

of Science, Osaka University

February, 1992

Approval

Februaty, 1992

This thesis is approved
as to style and content
by

勝部幸輝

Member-in-chief

高木俊夫

Member

小高忠男

Member

小林雅通

Member

Akira Nakamura

Member

Acknowledgements

The present work has been performed under the direction of Professor Yukiteru Katsube of Institute for Protein Research, Osaka University.

The author would like to express his sincere gratitude to Professor Y. Katsube for his cardinal guidance, discussion, and encouragements throughout this investigation.

The author is grately obliged to Associate Professor Yoshiki Matsuura and Dr. Mamoru Sato, Institute for Protein Research, Osaka University, for their helpful advice and continuing encouragement.

The author is deeply indebted to Professor Tairo Oshima, Dr. Hiromi Kirino and Mr. Kentaro Miyazaki, Department of Life Science, Tokyo Institute of Technology for their useful information, fruitful discussion and kindly supplying purified enzyme sample.

The author would like to express his gratitude Professor Nobuo Tanaka, Mr. Ko Onodera, Department of Life Science, Tokyo Institute of Technology for their helpful discussion, suggestion and kind supports.

The author also wishes his sincere thanks to Mr. Koichi Numata, Toyota Co. Ltd. for his helpful advice and discussion.

And the author gratefully acknowledges the members of Oshima Laboratory for their collaboration and discussion.

The author sincerely thanks to all the member of Katsube Laboratory for their kind assistances and friendship.

Finally, the author thanks his parents and all presents for their incessant understanding and encouragement.

今田勝巳

Katsumi Imada

February, 1992

Contents

Chapter-1

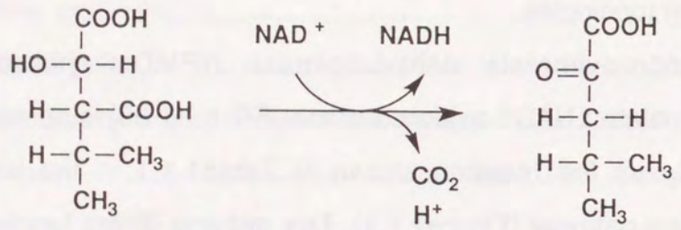
General introduction

Dehydrogenase catalyzes the dehydrogenation of a substrate molecule with NAD as the reducing co-enzyme. The three-dimensional structures of such well-known enzymes as the alcohol and lactate dehydrogenases (LADH and LDH) have been established at high resolution by X-ray diffraction analysis (Rossmann, 1975). They share a common structural feature in their NAD-binding domains for the bindings of NAD moieties and for the folding topologies. The remaining parts of the structures form substrate-binding domains that have different substrate specificities and folding topologies.

3-isopropylmalate dehydrogenase (IPMDH) [threo-d-3-isopropylmalate : NAD^+ oxidoreductase, EC 1.1.1.85] is an enzyme that catalyzes the reaction shown in Table1.1.1 in the leucine biosynthesis pathway (Figure1.1.1). This enzyme differs functionally from the well-known enzymes. IPMDH is bi-functional and catalyzes the decarboxylation of the substrate molecule simultaneously with dehydrogenation; whereas, most of the dehydrogenases characterized by X-ray analysis are mono-functional and have no enzymatic function other than dehydrogenation. X-ray analysis of a bi-functional enzyme, the isocitrate dehydrogenase (ICDH) from *E. coli* (Hurley et al., 1909) has shown that the ICDH is distinct from the well-known enzymes both in its amino acid sequence and in its

Table 1.1.1 Property of Tt-IPMDH

Source	*Thermus thermophilus* HB8
Molecular Weight	73,600 (36,800×2)
Quarternary Structure	identical subunit dimer
pI	4.7
Optimum pH	7.6 at 75°C
Function	



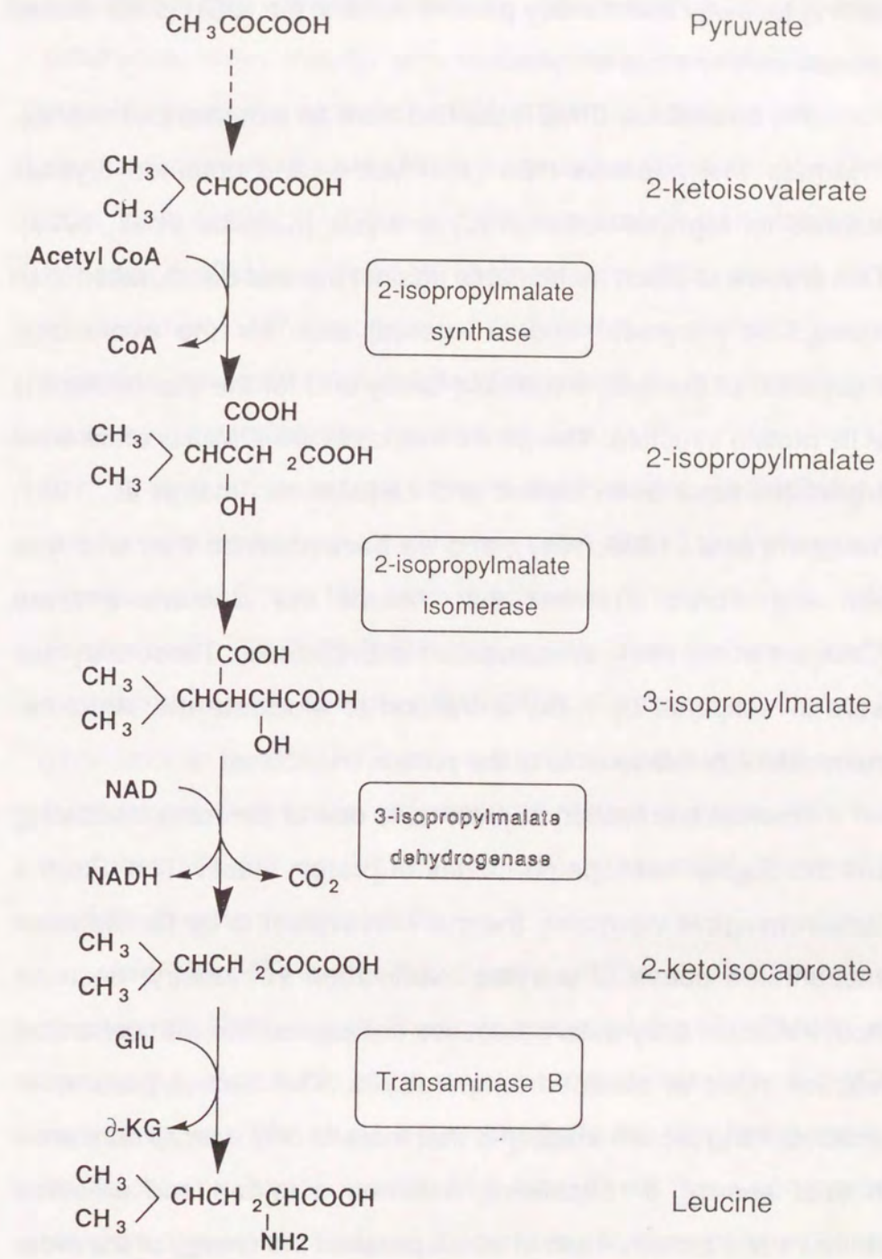


Figure. 1.1.1 The path way of leucine biosynthesis

folding topology and thereby pointed out that the ICDH is not related evolutionally to those enzymes.

We crystallized IPMDH purified from an extreme thermophile, *Thermus thermophilus* HB8 (Tt-IPMDH), and obtained crystals suitable for high resolution X-ray analysis (Katsube et al., 1988). This enzyme is much more stable against thermal denaturation than mesophilic enzymes, and of interest both for the evolutional implication of the dehydrogenase family and for the thermostability of its protein structure. The genes that code the IPMDHs of different organisms have been cloned and sequenced (Imai et al., 1987; Sekiguchi et al., 1986, 1987), and we have obtained their wild-type and engineered enzymes that include the chimeric enzyme (Onodera et al., 1991) overproduced in *E. coli* cells. These enzymes were investigated by X-ray diffraction to elucidate the structure-thermostability relationship of the protein molecules.

Thermal inactivation of proteins is one of the most interesting and thoroughly investigated modes of protein inactivation. From a biotechnological viewpoint, thermal inactivation is by far the most encountered cause of enzyme inactivation in industry, because most industrial enzymatic processes are carried out with enhanced reaction rates at elevated temperatures. The central problem in understanding protein stability is that there is only a small difference in total energy, 5-15kcal/mol, between a folded and unfolded structure of a protein, each of which possess the energy of the order of 10 million kcal/mol (Baldwin et al., 1987; Pace, 1986). Clearly, it is almost an impossible task to derive a structural difference between

the folded and unfolded states by calculations from the energy difference, even though one can calculate the energies of all possible structures of both the folded and the unfolded polypeptide chains. However, when the three-dimensional structures of a protein from thermophiles and the homologous protein from mesophiles become available, and when engineered proteins based on the three-dimensional structures of the wild-type proteins become available, we expect that detailed information on the stability would be obtained by comparing these structures.

Among enzymes from extreme thermophiles, Tt-IPMDH is the first enzyme whose gene coding has been cloned and whose over-expression in *E. coli* has been succeeded (Nagahari et al., 1980; Tanaka et al., 1981). The estimated denaturation temperature of the expressed enzyme is higher than 360K. As described above, the gene coding for IPMDH from mesophiles also has been cloned. Furthermore, their fusion enzymes and mutant enzymes have been produced. Therefore IPMDH will be an excellent material for studying protein thermostability.

We describe here the details of the three-dimensional structure of IPMDH from *Thermus thermophilus* (Tt-IPMDH) and discuss the relationship of the structure to thermostability in IPMDH, expecting that this study may contribute for the furtherance of research of thermal inactivation based on three-dimensional structure of proteins. Chapter 2 describes structure determination of Tt-IPMDH by X-ray crystallographic method. The details of the Tt-IPMDH structure are given in chapter 3. The obtained structure is

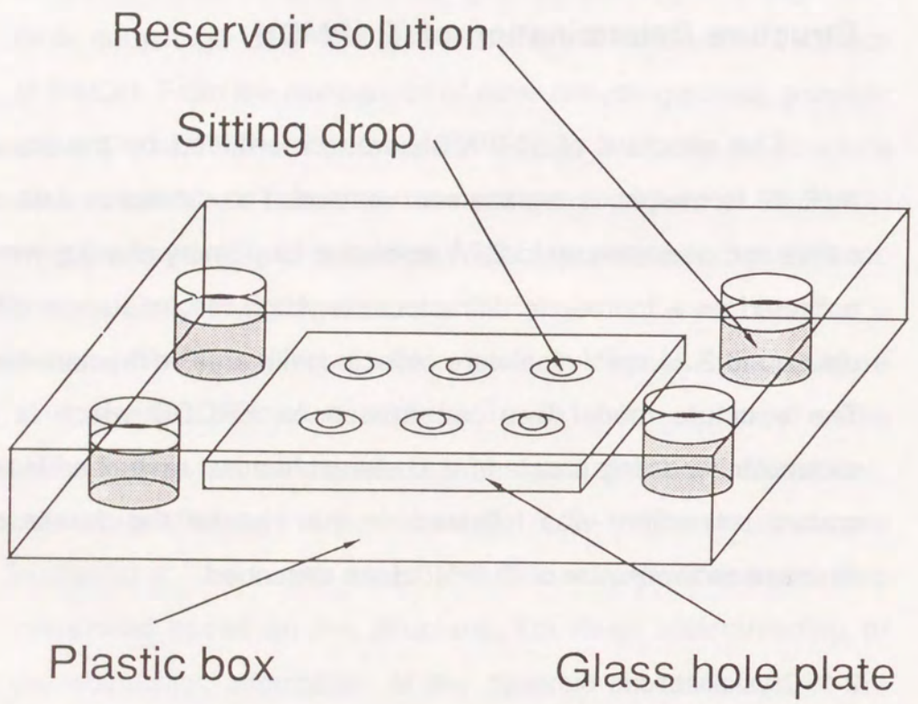
distinct from the other dehydrogenases except for isocitrate dehydrogenase. Chapter 4 deals with comparison of IPMDH with other dehydrogenases and also discusses the molecular evolution of IPMDH. From the comparison of other dehydrogenases, possible active site of IPMDH is described in chapter 5. During the structure determination, we found that the diffraction pattern of Tt-IPMDH crystal was varied with a change in the concentration of precipitant. It is expected that some structural changes occurred. Chapter 6 describes the structures of such crystals. In chapter 7, the structure determination procedure and the resultant structures of some chimeric enzymes are described. With these structural informations, the factors affecting the thermostability of Tt-IPMDH are discussed in chapter 8. The results of some mutational experiments are also interpreted based on the structure. For deep understanding of thermostability, information of the dynamic structure and of the structure under elevated temperatures are required. In chapter 9, we show a possibility of the structural analysis under elevated temperatures and B-factor analysis including the dynamic information.

Chapter-2
Structure Determination of Tt-IPMDH

The structure of Tt-IPMDH was determined by the usual multiple isomorphous replacement method. The diffraction data of native and derivative up to 2.7Å resolution for primary phasing were collected on a four circle diffractometer. High resolution intensity data up to 2.2Å resolution were collected with an IP-diffractometer. The structure model was constructed on FRODO which is a molecular modeling program, and stereochemical restrained least square refinement was followed. In this chapter the details of structure determination of Tt-IPMDH are described.

2-1 Crystallization

IPMDH of *T. thermophilus* HB8 was purified from *E.coli* C600 cells carrying pHB2, which clones the gene coding for the *T. thermophilus* enzyme, according to the procedure described by Yamada et al. (1990). The enzyme was crystallized with the sitting-drop vapor diffusion method from a solution containing 33 mg/ml IPMDH using ammonium sulfate as the precipitant at pH 7.5 (Figure2.1.1.). Hexagonal bipyramid crystals (Figure2.1.2) were grown up to the size of 0.5X0.5X1.0mm^3 in a week at room temperature (Katsube et al., 1988).



Sitting drop	4.0mM	Phosphate buffer
	33mg/ml	IPMDH
	0.20~0.70M	$(NH_4)_2SO_4$
Reservoir solution	4.0mM	Phosphate buffer
	0.70~1.00M	$(NH_4)_2SO_4$

Figure 2.1.1 Crystallization of Tt-IPMDH by the sitting drop vapor diffusion methods

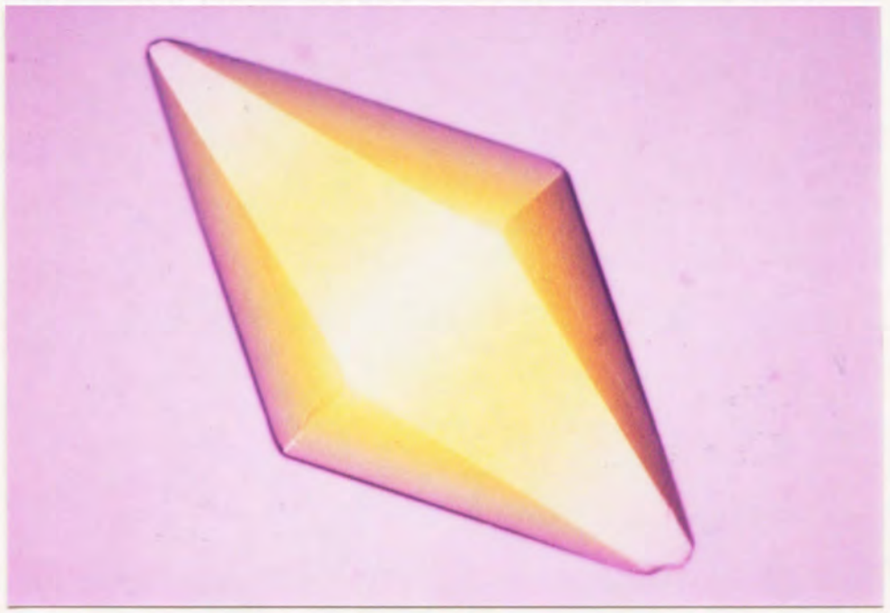


Figure 2.1.2 Crystal of Tt-IPMDH

2-2 Crystallographic parameters

Unit cell parameters were determined from precession photographs (Figure2.2.1), and more refined values were obtained from the measurement using four circle diffractometer. The crystals belong to trigonal system, space group $P3_221$ with unit cell dimensions a=b=78.6Å, c=158.1Å. In order to estimate the number of molecules per crystal asymmetric unit, the density measurement was carried out by a linear gradient method as shown in Figure2.2.2. Sodium nitrate was used as a standard solution for density calibration of the linear gradient solution. The densities of crystal and mother solution were estimated to be 1.18g/cm^3 and 1.09g/cm^3, respectively. Assuming the density of protein, D_P, is 1.37g/cm^3 (Matthews, 1968), the number of protein molecules in an unit cell, Z, can be calculated as follows,

$$Dc\text{-}Ds=\frac{M_W \cdot Z}{N_A \cdot V}(1-D_P \cdot Ds)$$

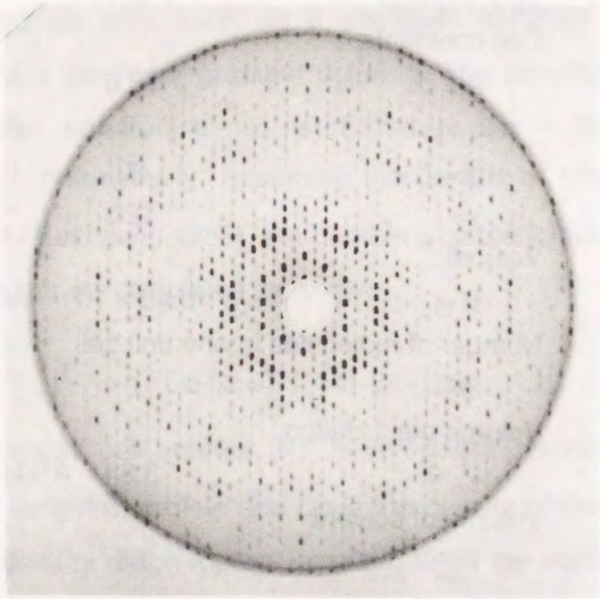
where V is unit cell volume, N_A is Avogadro's constant, M_W is molecular weight of protein and Dc is the density of the crystal and Ds is the density of the mother solution. From the equation, Z was determined to 6 and, therefore, there is one subunit per asymmetric unit. All the crystallographic parameters are given in Table2.2.1

2-3 Heavy-atom derivatives

Heavy-atom derivatives were prepared by a conventional soaking method using K_2PtCl_4, $NaAu(CN)_4$ and $K_3UO_2F_5$ as heavy-atom reagents. Because of the appearance of an insoluble

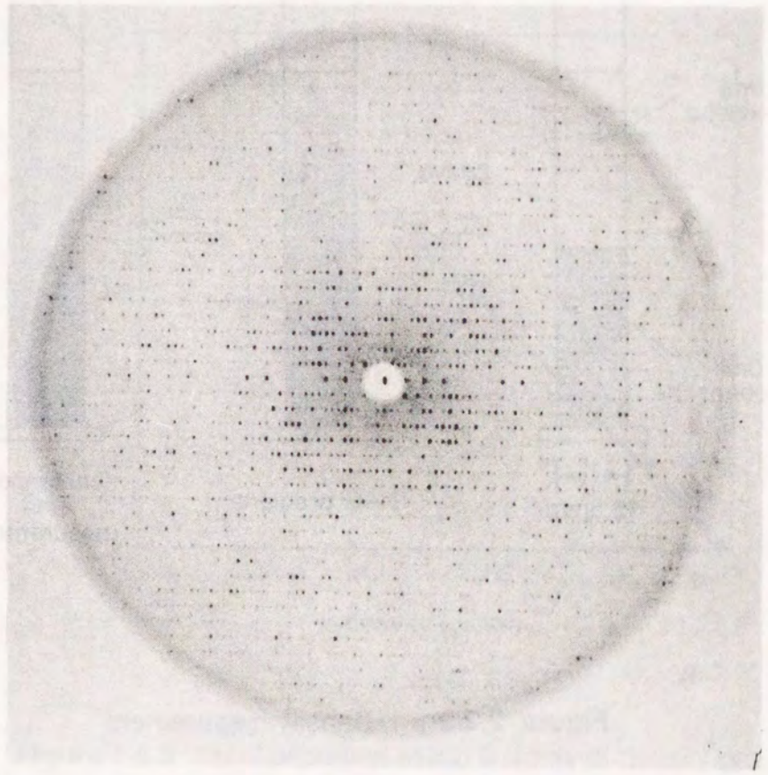
Table 2.2.1 Crystal parameters of IPMDH

Crystal system	Trigonal	
Space group	$P3_221$	
Cell constants		
$a = b =$	78.6	Å
$c =$	158.1	Å
$\alpha = \beta =$	90	°
$\gamma =$	120	°
Volume		
$V =$	8.45×10^5	Å^3
Number of molecules in one unit cell		
$Z =$	6	
Volume per 1 dalton		
$V_m =$	3.84	Å^3/dalton
Volume of solvent		
$V_{solv} =$	68	%
Volume of protein		
$V_{prot} =$	32	%



(h k 0) plane

Figure 2.2.1 Precession photograph of an IPMDH crystal.



(h 0 l) plane

Figure 2.2.1 Continued

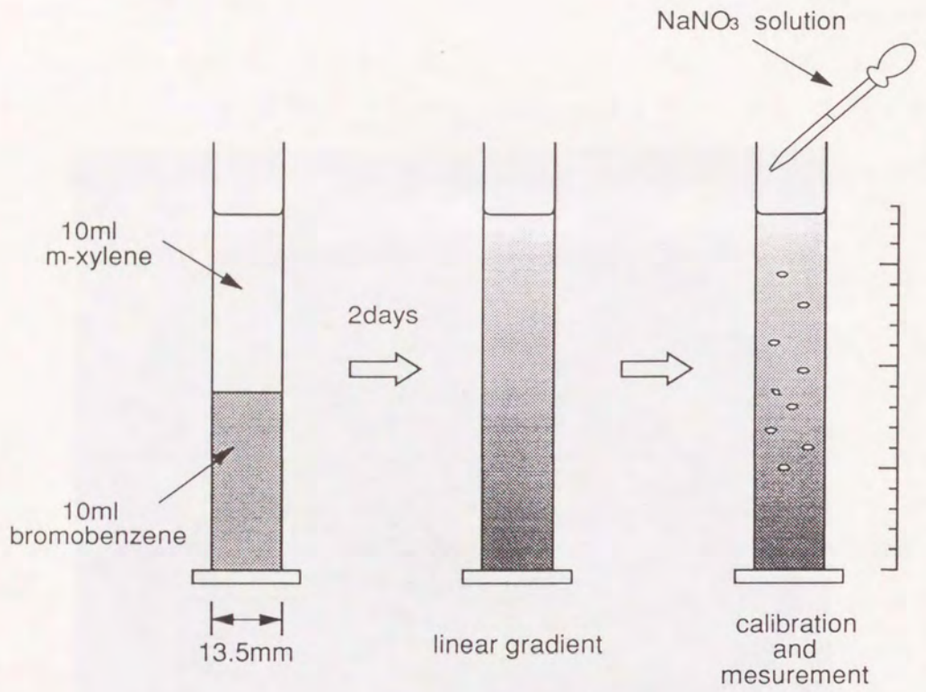


Figure 2.2.2 Density mesurement

Density gradient was formed using bromobenzene and m-xylene, whose density is 1.54 and 0.87g/cm^3, respectively. Bromobenzene was poured into a cylinder, then m-xylene is added calmly. After two days, they are mixed and linear density gradient was formed. Sodium nitrate solution was dropped into the mixture for calibration of the gradient. Crystals were put into the gradient and read their density from their floating position.

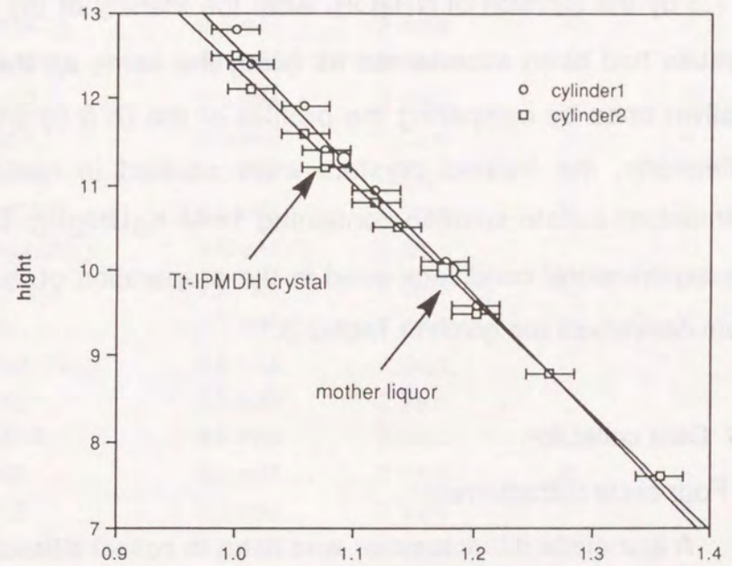


Figure 2.2.3 Calibration of the linearity of density gradients

salt in phosphate buffer, the preparation of uranium derivative was not straightforward. The problem was solved by modifying the native crystals before soaking as follows: the crystals were placed in nonbuffered ammonium sulfate solution and kept for at least 1 day to remove phosphate species. The pH of the solution was adjusted to 7.5 by the addition of NH_4OH. After the identity of the resulting crystals had been ascertained as being the same as the original (native) ones by comparing the profiles of the (h 0 0) and (0 0 l) reflections, the treated crystals were soaked in non-buffered ammonium sulfate solution containing 1mM $K_3UO_2F_5$. Details of the experimental conditions used in the preparation of the heavy-atom derivatives are given in Table2.3.1.

2-4 Data collection

Four circle diffractometer

A four circle diffractometer was used to collect diffraction data to a resolution of 2.7Å from the native and heavy-atom derivative crystals. The anomalous data for K_2PtCl_4 derivative were also collected upto 6Å resolution to determine the absolute configuration. The X-ray source was nickel-filtered copper Kα radiation from a rotating-anode X-ray generator (Rigaku RU200) operated at 40kV, 200mA. Anisotropic absorption was corrected by the method described by North et al. (1968). After correction for absorption and X-ray radiation damage, the full diffraction data sets were obtained by merging and scaling the respective data sets according to the

Table 2.3.1 List of the compounds checked for heavy atom derivatives

Compounds	Concentration	Soaking time	broken?	useable?
K_2PtCl_4†	0.5 mM	3 days	○	○
K_2PtCl_6	saturated	4 days	×	×
$K_2Pt(NO_2)_4$	0.5 mM	6 days	×	×
$K_2Pt(CN)_4$	0.5 mM	6 days	○	×
K_2PdCl_4	0.5 mM	6 days	○	×
Na_2IrCl_6	0.5 mM	1 days	×	×
$NaAuCl_4$	1.0 mM	2 days	×	×
$KAu(CN)_2$	1.0 mM	2 days	×	×
$Na[Au(CN)_4]$†	0.5 mM	3 days	○	○
C_2H_5HgCl	1.0 mM	2 days	○	×
K_2HgI_4	0.5 mM	2 days	×	×
$Hg(CH_3CO_2)_2$	0.5 mM	4 days	○	×
PCMB	0.5 mM	2 days	○	×
PCMBS	0.5 mM	2 days	○	×
PHMB	0.5 mM	2 days	×	×
EMTS	0.5 mM	3 days	×	×
Na_2WO_4	0.5 mM	3 days	○	×
$UO_2(CH_3CO_2)_2$¶	1.0 mM	9 hours	○	○
$K_3UO_2F_5$†	1.0 mM	9 hours	○	○
$Sm(NO_3)_3$	1.0 mM	1 days	×	×
$Sm_2(SO_4)_3$	0.5 mM	1 days	○	×

† The compound used for the MIR analysis.
¶ Radiation damage was too large to collect the diffraction data.

EMTS : Ethylmercury (II) thio salicylate
PCMB : p-Chloromercury (II) Benzoate
PCMBS : p-Chloromercury (II) Benzenesulfonate
PHMB : p-Hydroxymercury (II) Benzoate

method of Hamilton et al. (1965). A summary of the data collection and statistics is given in Table2.4.1.

IP-diffractometer

High resolution intensity data up to 2.2Å resolution were collected from the native crystal with an IP-diffractometer (Rigaku R-AXIS IIc). This system is based on the use of Arndt-Wonacott oscillation method (Arndt & Wonacott, 1977; Rossmann, 1979) for data collection with a reusable film, the Fuji imaging plate, as the two-dimensional X-ray detector (Amemiya et al., 1988). The details about this system is described by Yamamoto (1991) and Sato et al. (1991). The X-ray source was Cu-Kα radiation from a rotating anode X-ray generator (Rigaku RU200) operated at 40kV 100mA, monochromatized by a graphite crystal. Crystals of Tt-IPMDH were mounted with approximately c* axis, which is the shortest axis of inverse lattice, parallel to the spindle axis (oscillation axis). The orientation of a crystal was determined from three still photographs (Higashi, 1989). Exposure time was 40min for each data frame, and it took about only 14 hours for all measurements. Intensities recorded on each frame were integrated and indexed by the data processing package PROCESS, which is based on the oscillation film processing system MOSFLM (Leslie & Wonacott, 1986). Then each data set was combined and scaled by the method of Fox, G. C & Holmes, K. C. (1966). All the processing were implemented on VAX 3100 computer. Summary of the data collection and statistics of intensity data are given in Table2.4.2.

Table 2.4.1

Summary of data collection from a four-circle diffractometer and statistics

	Resolution	No. of crystal used	No. of independent reflections	R-merge§	R-iso¶
Native	2.7Å	6	14,808	0.035	
K_2PtCl_4	2.7Å	11	14,727	0.048	0.271
$K_3UO_2F_5$	2.7Å	13	14,808	0.047	0.179
$NaAu(CN)_4$	2.7Å	8	15,088	0.031	0.173
K_2PtCl_4†	6.0Å	1	+1,594	0.035	
			-1,582		

§ R-merge = $\Sigma\Sigma \mid F_i(h) / G_i - \langle F(h) \rangle \mid / \Sigma\Sigma \langle F(h) \rangle$
where h is the unique reflection index and $F_i(h)$ the structure amplitude of the symmetry equivalent reflections giving a mean value of $\langle F(h) \rangle$

¶ R-iso = $\Sigma \mid F_P - F_{PH} \mid / \Sigma \mid F_P \mid$

† Anomalous data were collected in order to determine the absolute configulation. This data set is not involved in phase determination.

Table 2.4.2

Summary of data collection from an X-ray imaging plate system and statistics

X-ray Source	Cu-Kα
X-ray generator	Rigaku RU200
Focus size	0.3 × 3mm
X-ray power	40kV, 100mA
Monochromatization	graphite plate
IP size	200×200mm
Pixel size	105μm
No. of crystal used	1
φ(spindle) - axis	approx. c axis
Crystal - to - IP distance	90mm
Resolution limit	2.2Å
Oscillation range per frame	1.5Å
No. of frames	21
Total oscillation range	31.5Å
Exposure time	40min / frame
No. of observed reflections	
full	25,263
partial	17,236
total	42,499
No. of independent reflections	24,478
Completeness †	83%
R - merge ‡	
full reflections	3.79%
partial reflections	4.74%
total reflections	4.19%
No. of rejected reflections ¶	
full reflections	0
partial reflections	2

† Considering the blind region.

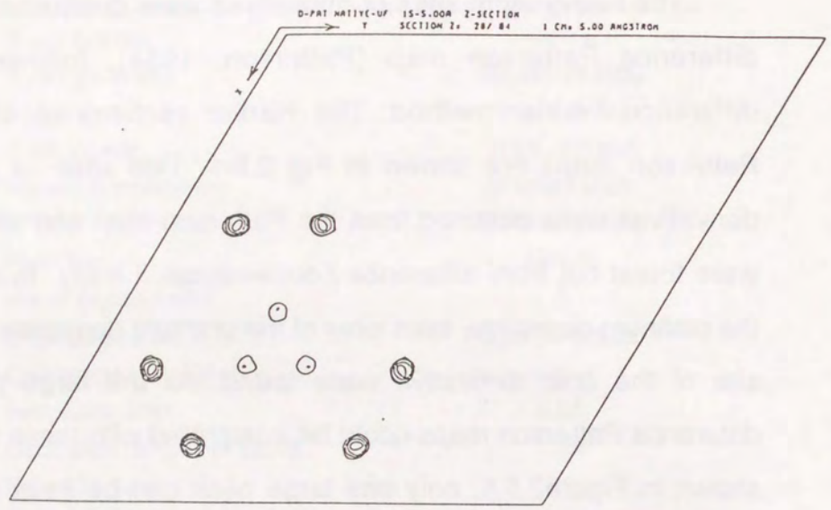
‡ *R*-merge = $\Sigma\Sigma | I_i(h) / G_i - \langle I(h) \rangle | / \Sigma\Sigma \langle I(h) \rangle$
where G_i is the inverse scale factor, $I(h)$ the diffraction intensity of the symmetry equivalent reflections, and $\langle I(h) \rangle$ the mean value of $I(h)$.

¶ Rejection criteria are $C_R = 0.3\ I_{mean} + 0.1 \langle I(h) \rangle$ for the reflections measured more than twice and $C_R = 3\ (0.3\ I_{mean} + 0.1 \langle I(h) \rangle)$ for the reflections with only two observations, where I_{mean} is the mean of all $I(h)$ values. The reflections with the differences $D_i = | \langle I(h) \rangle - I_i(h) / G_i | > C_R$ are rejected (Rossmann *et al.*, 1979).

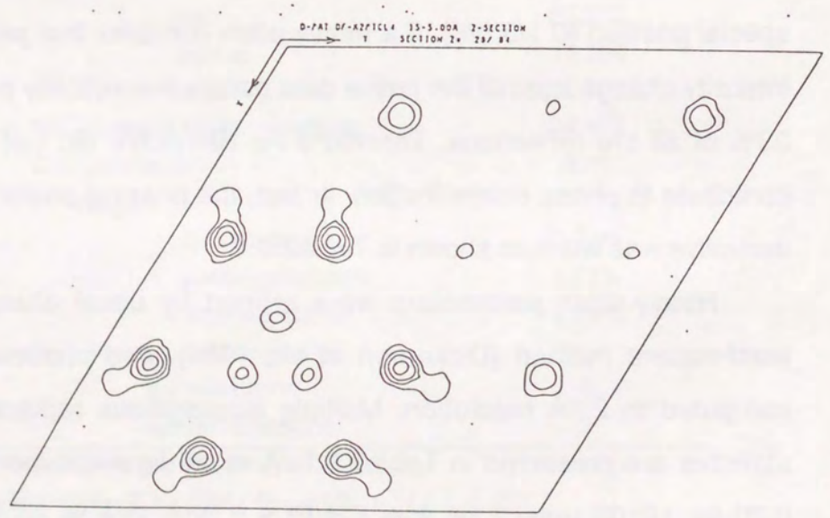
2-5 Phase determination

The heavy-atom sites of derivatives were determined from a difference Patterson map (Patterson, 1934), followed by a difference Fourier method. The Harker sections of difference Patterson maps are shown in Fig 2.5.1. Two sites of platinum derivatives were obtained from the Patterson map and other sites were found out from difference Fourier maps. Finally, five sites of the platinum derivative, fives sites of the uranium derivative and one site of the gold derivative were found. All the large peaks in difference Patterson maps could be interpreted with these sites. As shown in Figure2.5.1, only one large peak can be seen near the three fold axis in the Harker section of the Patterson map of Au derivative. This indicates that the site of Au atom locates near the special position (0 1/3 5/6). If a heavy-atom occupies that position, intensity change against the native data occurs theoretically only for 33% of all the reflections. Therefore Au derivative did not much contribute to phase determination. In fact, the phasing power of Au derivative was week as shown in Table2.5.1.

Heavy-atom parameters were refined by usual alternating least-square method (Dickerson et al., 1961), and phases were computed to 2.7Å resolution. Multiple isomorphous replacement statistics are presented in Table2.5.1. A mean figure of merit was 0.70 for 10103 reflections with $F/\sigma(F) > 3$ from 20Å to 2.7Å. The electron density map to 2.7Å resolution was highly clear, and the overall backbone of IPMDH molecule could be successfully traced.

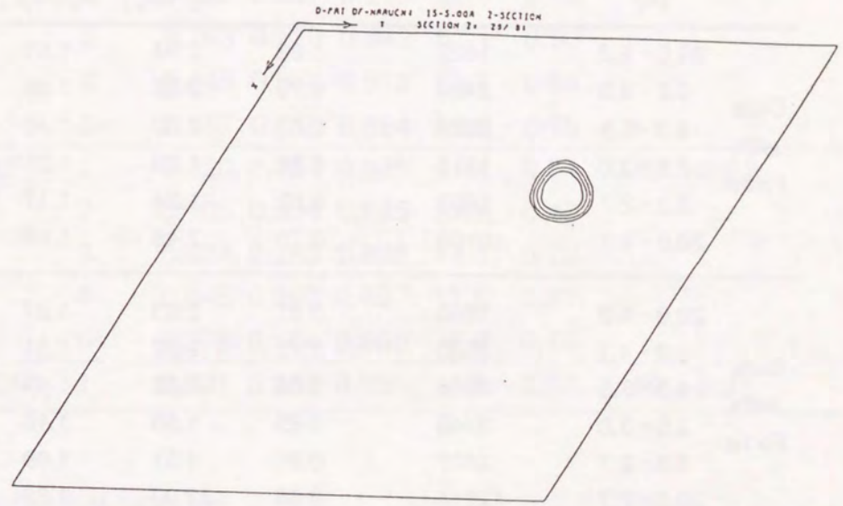


$K_3UO_2F_5$ derivative



K_2PtCl_4 derivative

Figure 2.5.1 The Harker section (W=1/3) of difference Patterson maps.



$NaAu(CN)_4$ derivative

Figure 2.5.1 Continued

Table 2.5.1 Summary of M.I.R statistics

	Resolution (Å)	No. of reflections	<m>†	Phasing power¶		
				K_2PtCl_4	$K_3UO_2F_5$	$NaAu(CN)_4$
Data with F>3σ	20.0 ~ 6.2	1622	0.83	2.64	1.91	2.12
	6.2 ~ 4.3	2436	0.73	2.35	1.56	1.49
	4.3 ~3.5	2624	0.67	2.00	1.45	1.16
	3.5 ~ 3.0	1818	0.66	1.59	1.21	1.04
	3.0 ~ 2.7	1603	0.62	1.34	1.17	1.15
	20.0 ~ 2.7	10103	0.70	2.48	1.69	1.56
Data with F>1σ	20.0 ~ 6.2	1665	0.81	2.63	1.87	2.04
	6.2 ~ 4.3	2560	0.72	2.32	1.51	1.43
	4.3 ~ 3.5	2936	0.66	1.98	1.41	1.07
	3.5 ~ 3.0	2448	0.65	1.55	1.15	1.89
	3.0 ~ 2.7	2407	0.61	1.31	1.09	1.10
	20.0 ~ 2.7	12016	0.68	2.09	1.39	1.23

† <m> : Figure of merit

¶ Phasing power is the mean value of the heavy-atom structure factor amplitude ($f_{r.m.s}$) divided by the residual lack-of-closure error ($E_{r.m.s}$).

phasing power = $(f_{r.m.s} / E_{r.m.s})$ $\quad f_{r.m.s} = (S\, f_H^2 / n)^{1/2}$

$E_{r.m.s} = (S\, (F_{PH} - |F_P + f_H|)^2 / n)^{1/2}$

where f_H is the structure factor amplitude for the heavy-atom, F_P the structure factor amplitude for the native crystal, and F_{PH} the structure factor amplitude for the derivative crystals.

Table 2.5.2 Refined heavy atom parameters

	Site	x	y	z	B‡	G§	R_k¶	R_c†
K_2PtCl_4	1	0.911	0.509	0.969	30.5	1.42	0.12	0.59
	2	0.127	0.597	0.875	36.6	0.99		
	3	0.763	0.138	0.842	57.1	0.90		
	4	0.448	0.884	0.912	55.1	0.84		
	5	0.157	0.653	0.864	62.6	0.76		
$K_3UO_2F_5$	1	0.720	0.154	0.925	11.4	0.54	0.09	0.61
	2	0.105	0.534	0.885	29.9	0.40		
	3	0.924	0.283	0.896	14.5	0.18		
	4	0.645	0.096	0.897	17.6	0.31		
	5	0.276	0.584	0.950	6.4	0.17		
$NaAu(CN)_4$	1	0.021	0.336	0.831	1.5	0.67	0.09	0.61

‡ Overall temperature factor (Å²)

§ Occpancy of a heavy atom site(%)

¶ Krout R : $R_K = \dfrac{\sum_{\mathbf{h}} |F_{PH}(\mathbf{h}) - F_{PH(calc)}(\mathbf{h})|}{\sum_{\mathbf{h}} |F_{PH}(\mathbf{h})|}$

where F_{PH} is the structure factor of the derivative.

† Cullis R : $R_C = \dfrac{\sum_{\mathbf{h}} |(F_{PH}(\mathbf{h}) - F_P(\mathbf{h})) - F_H(\mathbf{h})|}{\sum_{\mathbf{h}} |F_{PH}(\mathbf{h}) - F_P(\mathbf{h})|}$

where F_{PH} and F_P are the respective structure factors for native and derivative crystals, and F_H the structure factor for the heavy-atom.

The anomalous data of K_2PtCl_4 derivative up to 6Å were used for determination of absolute configuration. Bijovoet-difference Fourier maps were calculated to determine the enantiomorph of the crystal. Coefficient used in the Fourier synthesis were

$$|F_{PH}(+)| - |F_{PH}(-)|$$

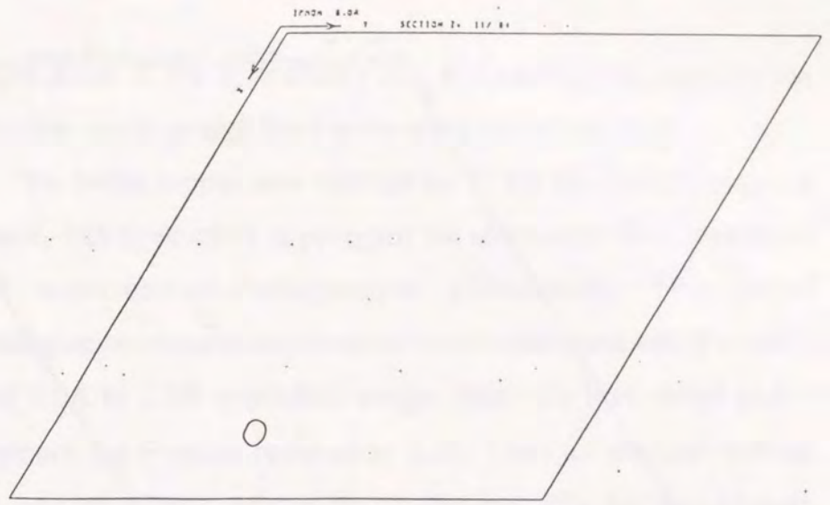
and phase were

$$\phi_{best} - \pi/2$$

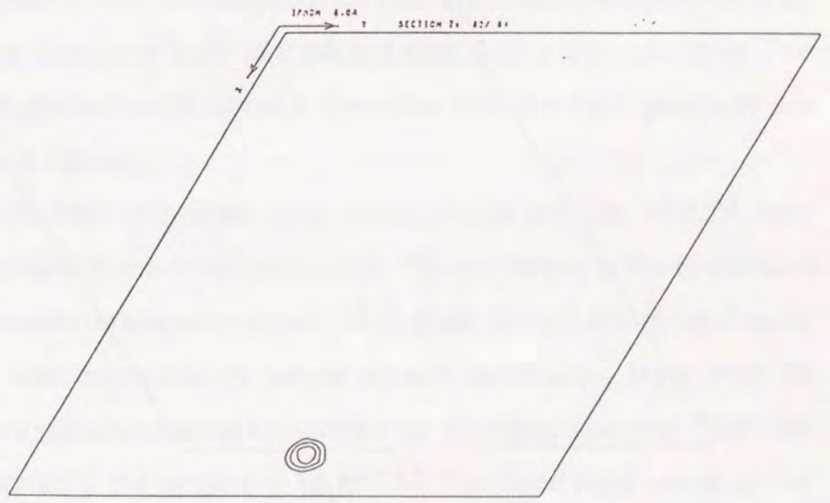
where ϕ_{best} is the best phase (Blow & Crick, 1959) determined from Au-derivative and anomalous data of Pt-derivative. A difference Fourier map between Au-derivative and a native phased with anomalous data of Pt-derivative were also calculated. As shown in Figures2.5.2 and 2.5.3, $P3_121$ gives relatively noisy map, while $P3_221$ gives large peak at the heavy-atom position. Hence, the space group was determined to $P3_221$.

2-6 Model building and refinement

The initial positions of α-carbon were located by manual interpretation on a mini map calculated from the diffractometer data up to 2.7Å resolution and MIR 'best' phases (Brow & Crick, 1959). The MIR electron density map was quite fine. The solvent-protein boundary was clearly appeared and the electron density for many side chains was apparent. A molecular model was constructed using the model building program FRODO (Jones, 1978) implemented on an Evans and Sutherland PS390 graphics system linked to a Micro-VAX II computer. The construction of the model was greatly aided by knowledge of the complete amino acids sequence data of Tt-

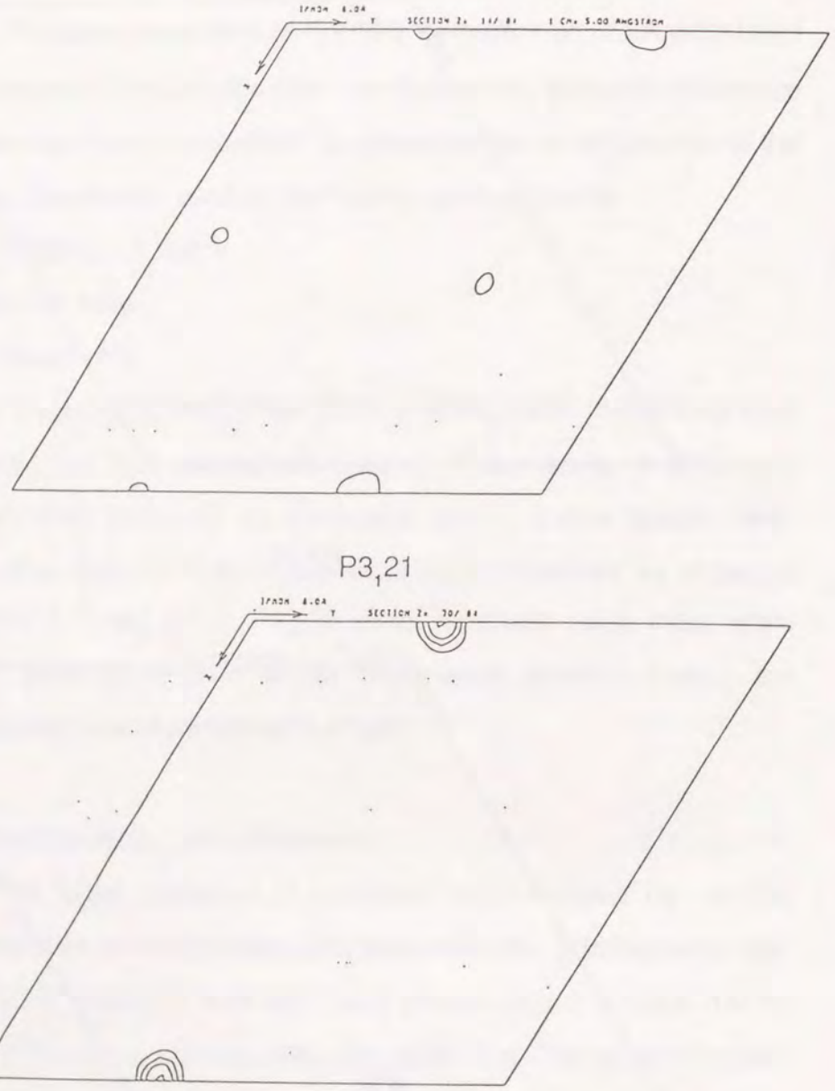


$P3_121$



$P3_221$

Figure 2.5.2 Bijovoet difference Fourier maps
coefficient : $F_{PH}(+) - F_{PH}(-)$
phase : $\phi_{best} - \pi/2$
The peak is corresponding to Pt1 site in Table2.5.2



P3$_1$21

P3$_2$21

NaAu(CN)4 derivative

Figure 2.5.3 Difference fourier map phased with anomalous data of K_2PtCl_4
coefficient : $F_{PH} - F_P$ where F_{PH} is the Au-derivative data in this figure
phase : ϕ_{best} determined from anomolus data of K_2PtCl_4
The peak is corresponding to Au1 site in Table 2.5.2.

IPMDH. Most of the side chains and the carbonyl oxygens of the main chain could be well fitted in the electron density map.

The initial model was refined by PROLSQ (Hendrickson & Konnert, 1980) which is a program for stereochemical restrained least squares crystallographic refinement. The initial crystallographic R-value was 0.42 for 8126 reflections with $F > 3\sigma(F)$ in the 5.0Å to 2.7Å resolution range. After the first round of the refinement the R-value reduced to 0.33. Then an electron density map was calculated based on coefficients 2Fo-Fc and phases ϕ_{calc}, and displayed again on FRODO. The refined model was modified to fit the density map and refined again. After several refinements and manual modifications the R-factor fell down to 0.23 for the data from 5.0Å to 2.7Å collected with a diffractometer. The target sigmas and final r.m.s. deviation from the ideal geometry are given in Table2.6.1.

Further refinement was continued by the use of 2.2Å data collected with the IP-diffractometer. The resolution of the refinement was extended stepwise from 2.7Å to 2.2Å. Then a difference Fourier map was calculated to locate solvent molecules. More than 60 solvent molecules were identified from the difference map. They are positioned in the range of 2.5Å to 3.5Å from fixed polar atoms on the molecular surface of the enzyme or other solvent molecules. These solvent molecules were included in the least-square refinement with unit occupancy. At this stage restraints on individual B-factors were released. Solvent molecules whose B-factors exceeded 100Å^2 were eliminated. ' Omit ' maps (Artymuik & Blake, 1981; Rice, 1981) were

calculated at certain stages in the refinement and were useful for locating the less well ordered regions. The final R-factor is 0.182 for 20,307 reflections with $F > 3\sigma(F)$ in the 5.0Å to 2.2Å resolution range. The dependency of the R-factor on the resolution is given in Table2.6.2. The target sigmas and final r.m.s.deviation of the model are given in Table2.6.1. A part of the final electron density map is shown in Figure2.6.1.

The final model has 2,590 protein atoms and 63 solvent molecules. Two of the solvent molecules seems like SO_4 or PO_4 ions from the shape of electron density. They were treated as SO_4 ion in the refinement. The B-factor for the protein atoms are in the range from 21.0 to 60.5Å^2 and for the solvent atoms are from 21.4 to 86.2Å^2. The average B-factor for all protein atoms took somewhat large value, 33.2Å^2, because of the sharp decrease of diffraction intensities beyond 2.7Å resolution. The variation in B-factor for both main chain and all atoms is shown in Figure2.6.2. Residues whose main chain B-factors exceed 50Å^2 include the residues from 78 to 84 and this region is a surface loop extended to the solvent region.

Figure2.6.3 shows the Ramachandran plot (Ramakrishnan & Ramachandran, 1965) of the main chain dihedral angles. Most of non-glycine residues lie in normal allowed regions. Only three non-glycine residues, Arg176, Asp231 and Ile284, are exceptional. Arg176 and Asp231 are involved in loop regions between α-helix and β-strand, and Ile284 exists in the middle of the long loop region constructed by residues from 272 to 287. There is one cis peptide that is Pro143.

Table 2.6.1 Summary of least-squares parameters and deviations

	Target	r.m.s.deviation	
		diffractometer	imaging plate
Bonding distances (Å)			
1-2 bond	0.020	0.013	0.015
1-3 angle	0.030	0.035	0.036
1-4 planar	0.050	0.048	0.050
Planar groups (Å)	0.020	0.010	0.012
Chiral volumes (Å^3)	0.150	0.161	0.177
Non-bonded contacts (Å)			
Single torsion	0.500	0.259	0.207
Multiple torsion	0.500	0.396	0.241
Possible hydrogen bond	0.500	0.385	0.249
Torsion angles (deg.)			
Planar	3.0	1.9	2.2
Staggered	15.0	27.9	23.7
Orthonormal	20.0	42.1	31.1
Thermal factors (Å^2)			
Main-chain bond	1.000	0.516	0.500
Main-chain angle	1.500	0.919	0.852
Side-chain bond	1.500	0.800	1.052
Side-chain angle	2.000	1.319	1.640

Table 2.6.2 Dependency of the R-factors on resolution

Resolution (Å)	No. of reflections	*R* - factor	*R* - factor (accumulated)
5.00 ~ 4.00	2,283	0.131	0.131
4.00 ~ 3.30	3,529	0.154	0.143
3.30 ~ 2.90	3,482	0.197	0.157
2.90 ~ 2.63	3,383	0.219	0.167
2.63 ~ 2.45	2,785	0.236	0.173
2.45 ~ 2.32	2,414	0.233	0.177
2.32 ~ 2.20	2,431	0.251	0.182
5.00 ~ 2.20	20,307	-	0.182

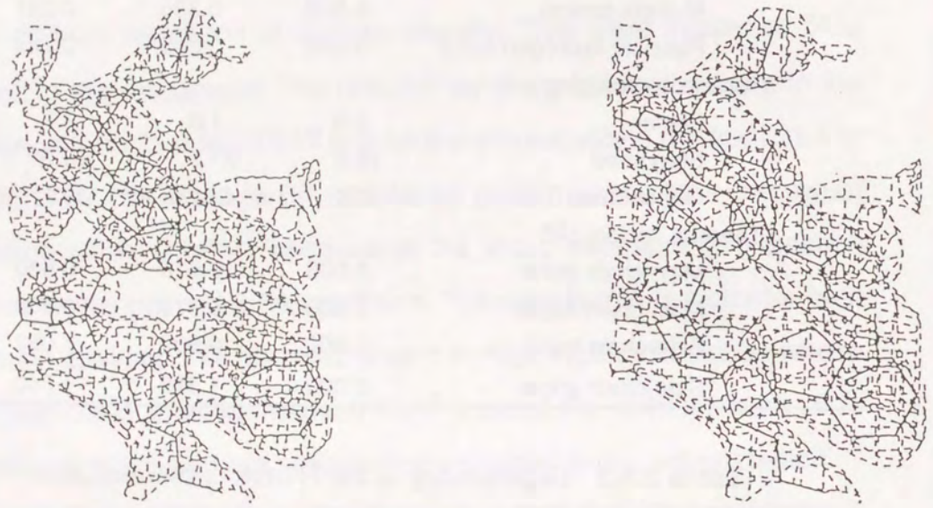


Figure 2.6.1 Selected views of the electron density in the final $2F_o$-F_c map

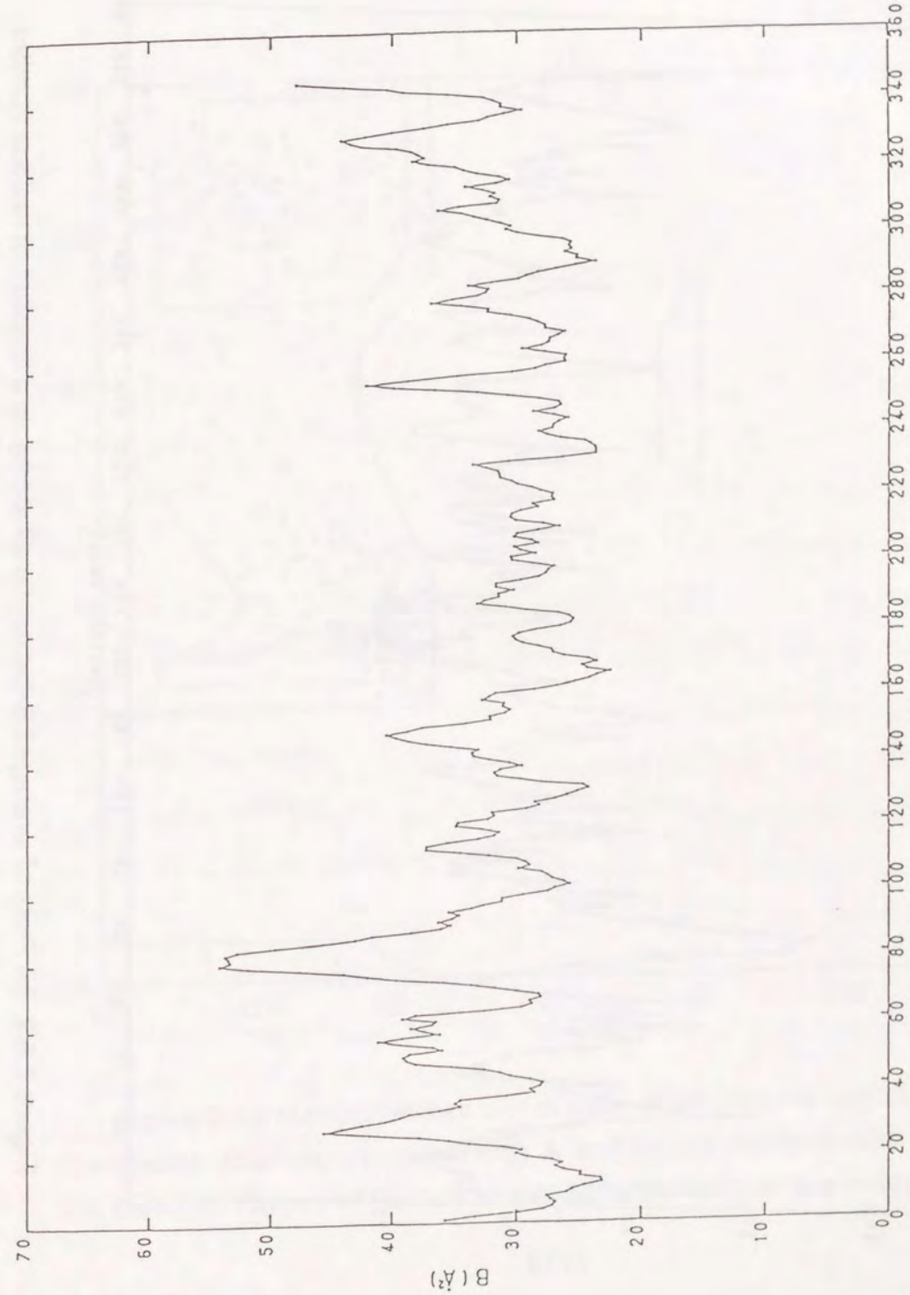


Figure 2.6.2 (a) Plots of averaged B-factors for mainchain atoms as a function of residue number

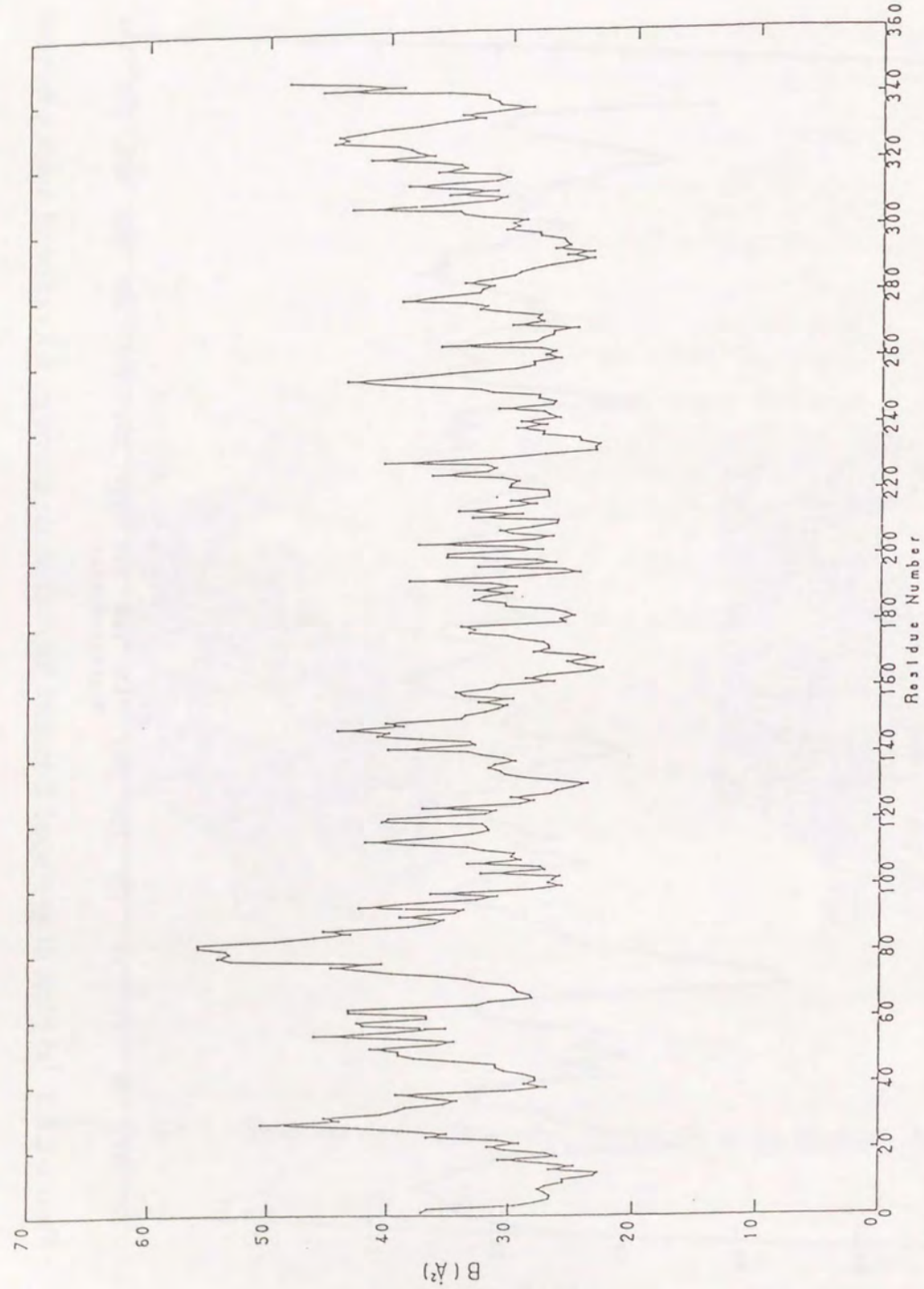


Figure 2.6.2 (b) Plots of averaged B-factors for all atoms as a function of residue number

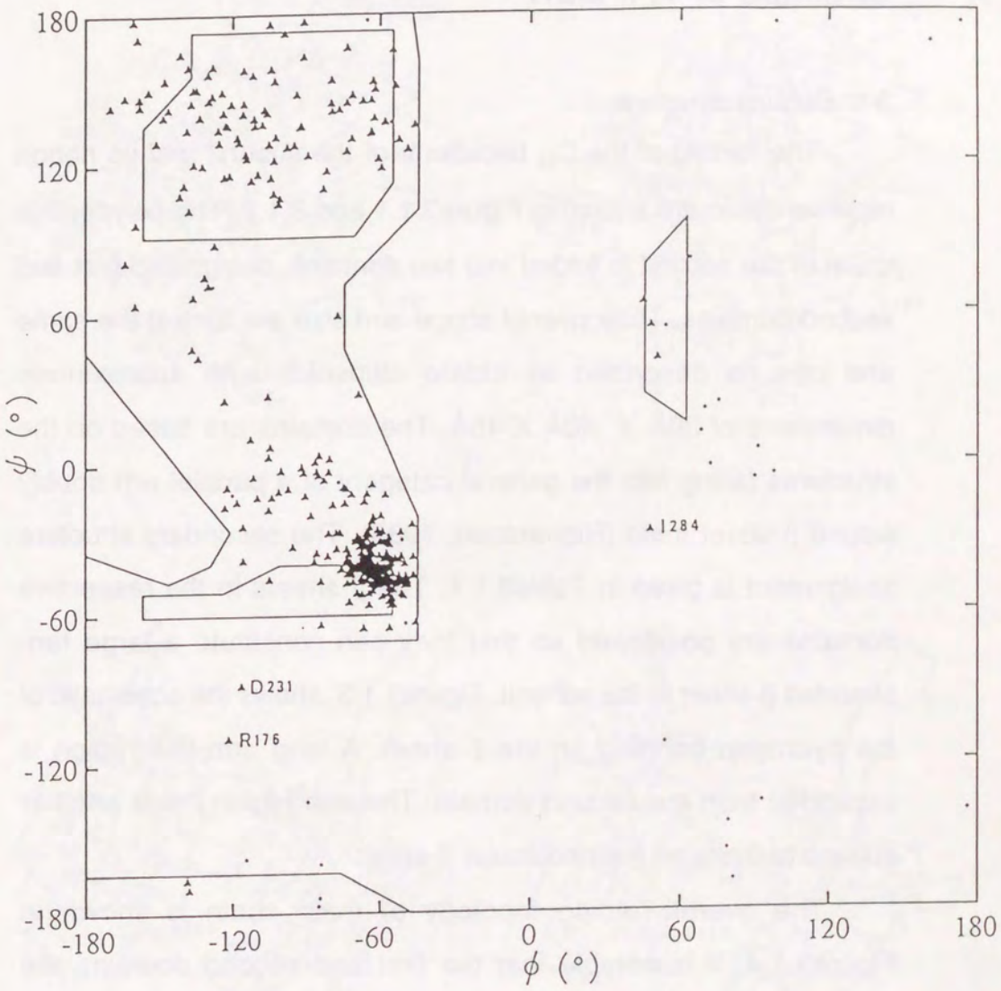


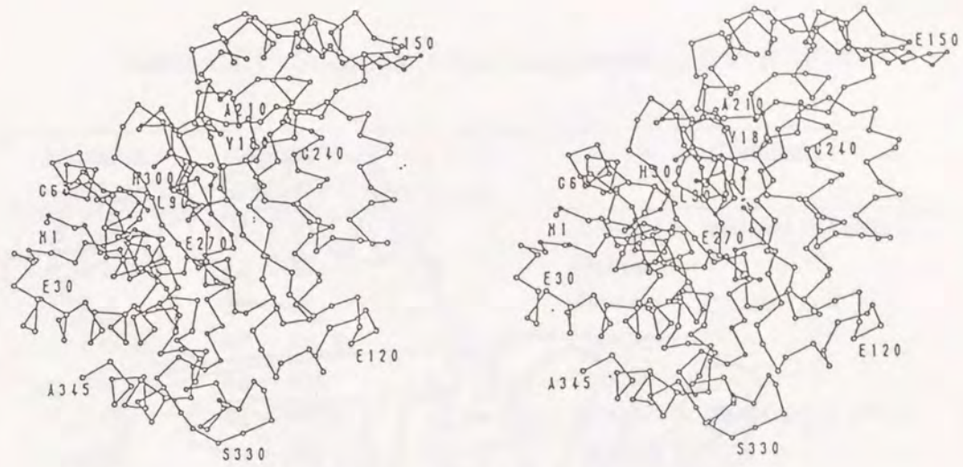
Figure 2.6.3 Ramachandran plot of main-chain dihedral angles. Non-glycine residues are shown with ▲ and glycine residues with ・. The preferred regions of Ramakrishana & Ramachandran are indicated.

Chapter-3
Structure of Tt-IPMDH

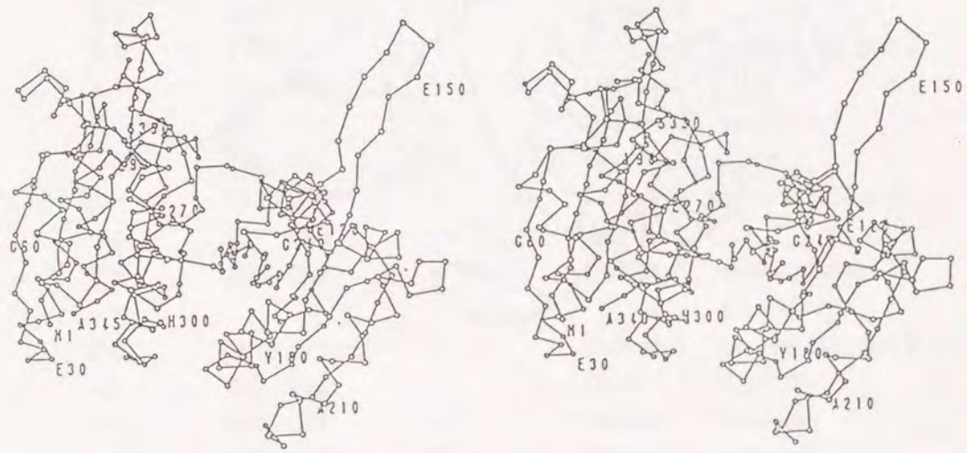
3-1 Subunit structure

The folding of the C_α backbone of the subunit and its ribbon representation are shown in Figure3.1.1 and 3.1.2. The polypeptide chain of the subunit is folded into two domains, designated first and second domains. Their overall shape and size are almost the same and can be described as oblate ellipsoids with approximate dimensions of 30Å X 40Å X 45Å. The domains are based on the structures falling into the general category of a parallel α/β doubly wound β-sheet motif (Richardson, 1985). The secondary structure assignment is given in Table3.1.1. The β-sheets in the respective domains are positioned so that they can constitute a large ten-stranded β-sheet in the subunit. Figure3.1.3. shows the schematic of the hydrogen bonding on the β-sheet. A long arm like region is expanded from the second domain. The arm region holds another subunit to make an intermolecular β-sheet.

The overall folding topology of main chain is shown in Figure3.1.4. It is notable that the first and second domains are topologically identical around their central β-sheets. The topology of both domains of IPMDH are quite distinct from the NAD binding domain of well-known enzymes.



(a)



(b)

Figure 3.1.1 Stereo drawing of the Cα backbone of a subunit of Tt-IPMDH. Amino acid residues and their numbers are given near some Cα positions. (a) same direction of Figure 3.1.2, (b) another view

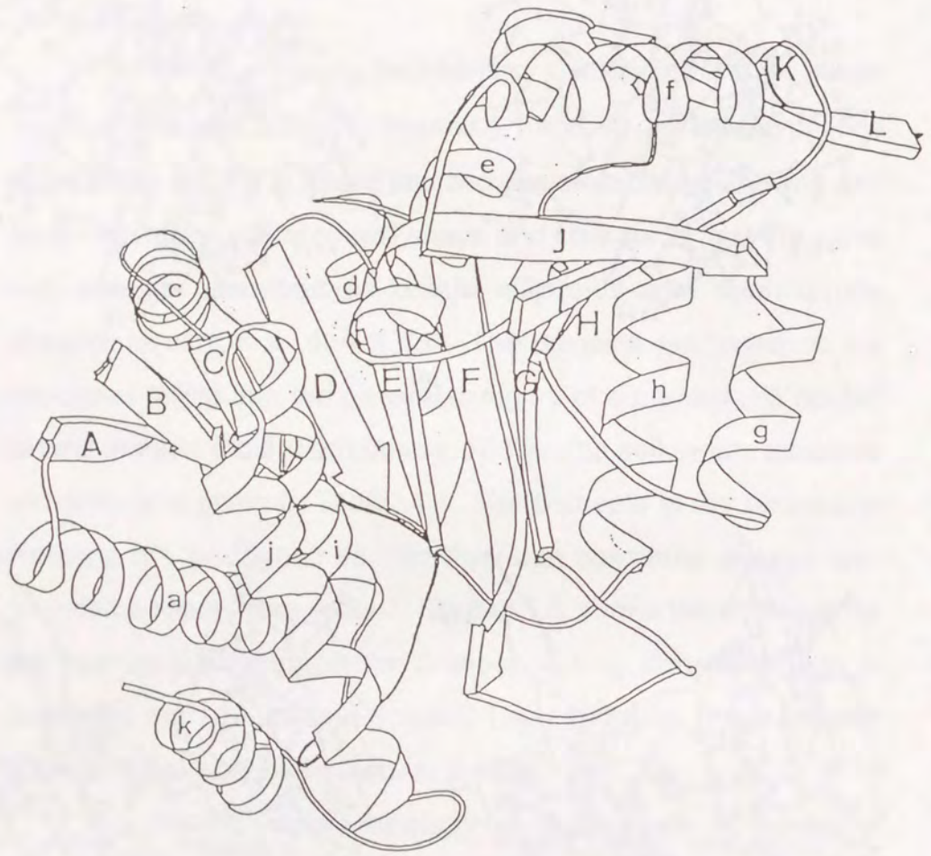


Figure 3.1.2 Ribbon (Priestle, 1988) representation of the polypeptide chain of a subunit of Tt-IPMDH. β-strands and α-helices are labeled with captital and small letters which are same as the letters in Table 3.1.1.

Table 3.1.1 Secondary structural elements of Tt-IPMDH

Structural element	Residue numbers	No. of amino acid		Comment
β-B	2 - 6	5	Domain 1	parallel to A and C
α-a	12 - 30	19	Domain 1	
β-A	35 - 40	6	Domain 1	parallel to B
α-b	42 - 48	7	Domain 1	
α-c	56 - 63	8	Domain 1	
β-C	66 - 69	4	Domain 1	parallel to B and D
α-d	86 - 95	10	Domain 1	
β-F	100 - 110	11	Domain 2	antiparallel to E and G
β-G	126 - 133	8	Domain 2	antiparallel to F parallel to H
β-K	144 - 147	4	Arm region	antiparallel to L
β-L	150 - 158	9	Arm region	antiparallel to K and L'
α-e	159 - 175	17	Domain 2	
β-I	179 - 185	7	Domain 2	parallel to H and J
α-f	190 - 204	15	Domain 2	
β-J	208 - 215	8	Domain 2	parallel to I
α-g	216 - 225	10	Domain 2	
β-H	232 - 236	5	Domain 2	parallel to G and I
α-h	237 - 250	14	Domain 2	
β-E	259 - 262	4	Domain 1	antiparallel to F parallel to D
β-D	267 - 271	5	Domain 1	parallel to C and E
α-i	288 - 298	11	Domain 1	α-bundle
α-j	305 - 321	17	Domain 1	α-bundle
α-k	333 - 343	11	Domain 1	α-bundle

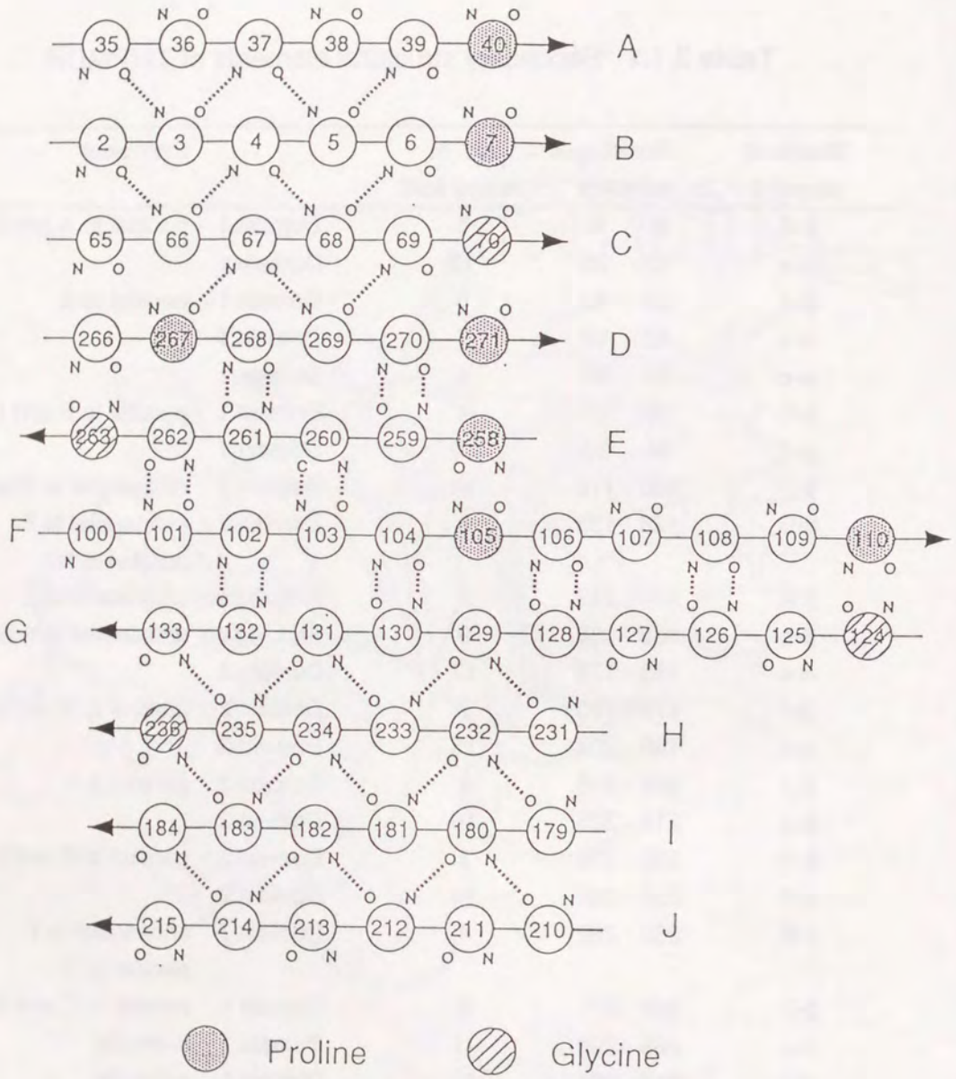


Figure 3.1.3 Hydrogen bonding diagram on the ten-stranded central β-sheet of a subunit of Tt-IPMDH. The letters showing the strands are the same as in Table 3.1.1.

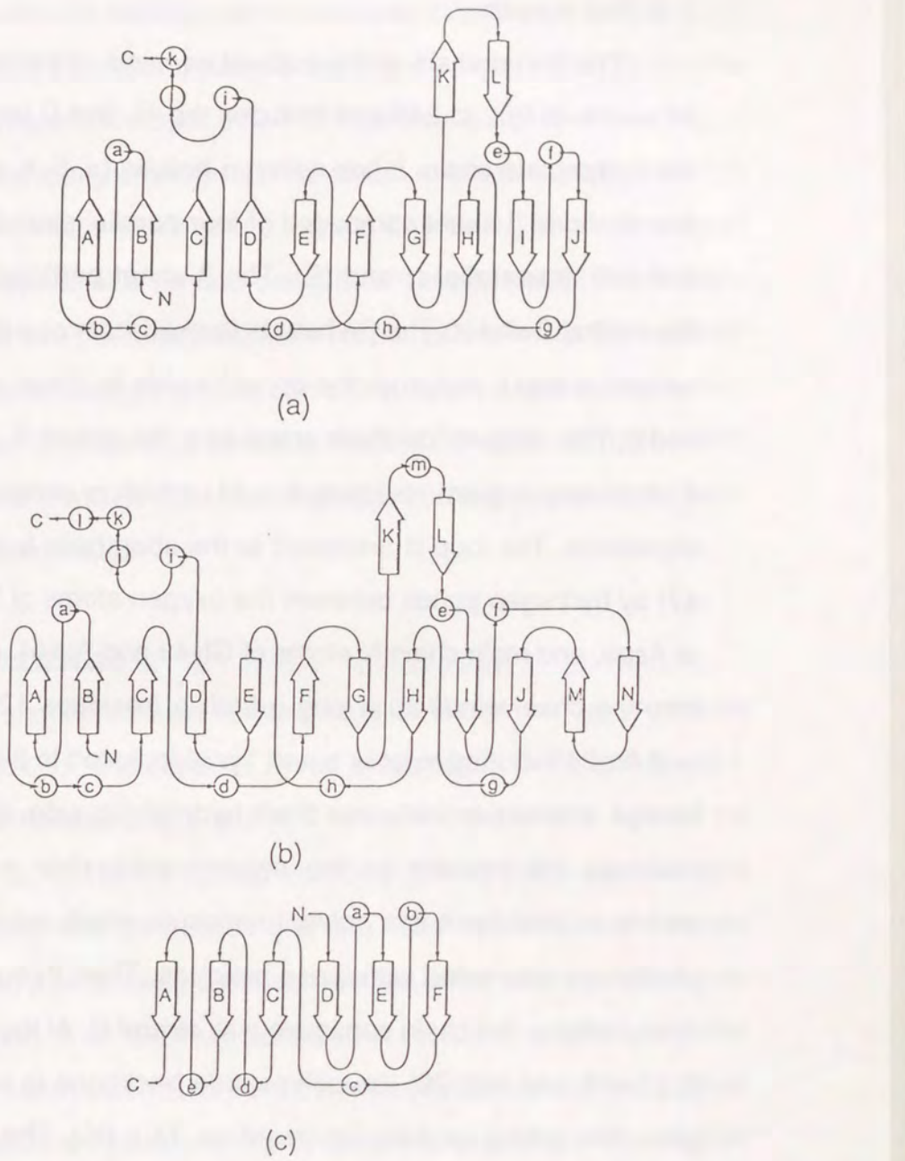


Figure 3.1.4 Folding topologies of the polypeptide chains of (a) a subunit of Tt-IPMDH, (b) a subunit of *E.coli* ICDH and (c) NAD-binding domain of LADH. Spheres and arrows represent α-helices and β-strands.

3-2 First domain

The first domain of the subunit consists of residues from 1 to 99 and from 252 to 345 and includes the N- and C-terminal ends of the polypeptide chain. It has seven α-helices (a, b, c, d, i, j, k) and a five stranded β-sheet composed of four parallel strands (A, B, C, D) and one antiparallel strand (E). The β-sheet participate in forming the central β-sheet. The β-sheet is flanked from one side by two α-helices, a and i, and from the opposite side by three α-helices, b, c and d. The polypeptide chain starts from the strand B. There follows a short loop region (residues 8 ~ 11) which is conserved in many organisms. The loop is anchored to the short helix b (residues 43 ~ 47) by hydrogen bonds between the oxygen atoms of the side chain of Asp9, and main chain N atoms of Gly43 and Ala44. After the short loop the chain winds up a long α-helix a (residues 12 ~ 30). Glu17 and Arg24 included in helix a and Tyr36 included in β-strand A have charge interaction between their hydrophilic side chains. These residues are present on the solvent accessible surface of the molecule, and there is a solvent molecule which may be sulfate or phosphate interacted with these residues. Then through the strand A and helix c, the chain continues with strand C. At the end of strand C (Ser71 and Val172) the polypeptide backbone is fairly bent and goes into a long loop region (residues 71 ~ 86). The side chain of Lys76 which belongs to the loop makes charge interaction with the carboxyl group of Asp47 which is in helix c. After the loop, the polypeptide chain passes the helix d and goes to the second domain. Helix d may be concerned to the substrate binding and the

possible substrate binding site is discussed in section 5-1. The chain returned from the second domain continues with strand E, D and the long loop region (residues 272 ~ 284) which is highly conserved in IPMDH from many organisms. The amino acid residues from Ile285 to the C-terminal end constitute three α-helices, i, j and k., arranged with an antiparallel α up-and-down helix bundle motif (Richardson, 1985). Of these three helices, helix i is comprised of apolar residues and positioned away from the solvent surface. Between the helix j and k, there is an unusual structure Pro323 - Pro324 - Pro325 which is peculiar to IPMDHs from thermophiles. These three residues have all trans conformation.

3-3 Second domain

The second domain is composed of the amino acid residues from 100 to 251, and contains seven β-strands F, G, H, I, J, K and L and four α-helices e, f, g and h. The central β-sheet is formed by strands F, G, H, I and J flanked from one side by helices e and f and from another side by helices g and h. The sequences of strands F,G,H,I and J joined with α-helices are the same as those of strands E,D,C,B and A in the first domain. In fact, the rotation function (Lattman, 1985) calculated from the intensity data of native crystals gave a significant peak at the position corresponding to the relative disposition of the first and second domains.

The second domain starts from the β-strand F. As shown in Figure3.1.3, there is a proline residue (Pro105) in the middle of the strand F. At this position the strand is slightly twisted. There is a

middle loop between the strand F and G. This loop including residues 110 ~ 120 makes both hydrophobic and hydrophilic contact with the other subunit. The Glu120 - Glu121 - Ile122 - Ala123 residues form a one turn α-helical structure. Strand G leads into the arm like region through the short loop which constructed from Leu134 to Gly141. Only one cis-peptide can be found in IPMDH, that is Pro143 which is present at the entrance of the arm like region.

Returned from the arm region, the polypeptide chain forms the helix e. Five basic residues such as arginine and lysine are present in the C terminal end of this helix. The polypeptide chain continues with the strand I (residues 179 ~ 185), helix f (residues 190 ~ 204), and strand J (residues 208 ~ 215). The short loop consists of residues 186 ~ 189 protrude to the cleft constructed by the dimer. Helix g and h makes subunit-subunit contacts with the symmetry related g and h helices of the other subunit.

3-4 Arm region

The amino acid resides from Pro143 to Ser158 form a long arm like polypeptide chain that protrudes from the second domain. The arm-like polypeptide chain forms, in part, an anti-parallel β-sheet consisting of strands K (residues 143 ~ 147) and L (residues 149 ~ 158) and runs over to another subunit to make up intersubunit hydrogen bonds. The arrangement of the β-sheet is shown in Figure3.4.1.

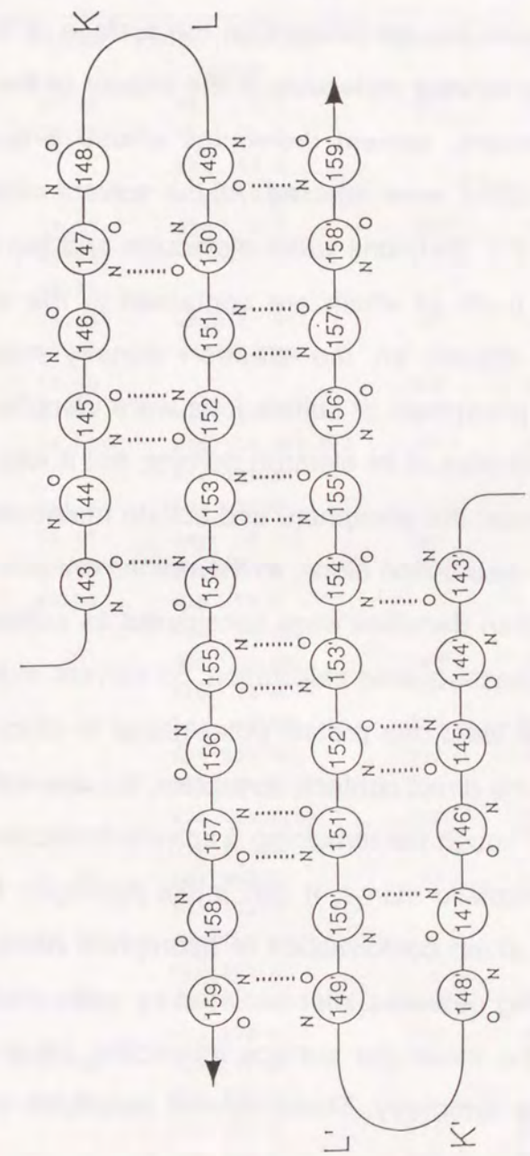


Figure 3.4.1 Hydrogen bonding diagram on the inter-subunit β-sheet consisting of four β-strands K, L, K' and L'.

3-5 Solvent structure

Distinct electron density peaks corresponding to relatively fixed solvent molecules are present on the surface of the enzyme molecule and no solvent molecules in the interior of the molecule. After the refinement, solvent molecules whose B-factors were diverged over 100Å^2 were rejected. All the solvent molecules are listed in Table3.5.1. Sixty-one water molecules and two phosphate or sulfate ions (both of which are contained in the solution for crystallization) appear on the electron density map at 2.2Å resolution. The phosphate or sulfate ions were identified from the tetrahedral distribution of its electron density, but it was difficult to distinguish between the phosphate and sulfate molecules from the electron density distribution alone. In the refinement procedure, the tetrahedral electron densities were considered as sulfate ions and included in the least-squares calculation. 55 solvent molecules are within 3.5Å of at least one protein polar group or atom. 3 solvent molecules have no direct contacts to protein, but are within 3.5Å of other solvents. Though the remaining 5 solvent molecules are apart from other hydrophilic atoms, it can make hydrogen bonds with alternative side chain conformation of hydrophilic residues. Some hydrogen bonding networks intermediated by water molecules are found around the molecular surface interacting other molecules related by lattice symmetry. These solvent molecules are given in Table3.8.1.

It is characteristic that the number of the solvent molecules appearing on the electron density map is few in comparison with the

Table 3.5.1 Fixed solvent molecules and their protein hydrogen bonds

Solvent	B-factor	Protein Hydrogen bonds	Solvent	B-factor	Protein Hydrogen bonds
346	56.36	Val 3 O	363	25.71	Asp 127 O
347	38.20	Phe 39 O			Pro 227 O
		Phe 41 O			Phe 230 O
		Gly 8 N			Asp 127 N
348	38.15	Gly 8 O	364	44.75	Asp 127 Oδ2
349S	69.95				Val 128 O
O1	69.17				Arg 176 Nη2
O2	68.52	Arg167' Nη1	365	25.68	Glu 133 Oε1
O3	68.26	Arg167' Nη2			Glu 161 O
O4	69.40		366	26.98	Glu 155 Oε1,Oε2
350	30.30	Glu 14 O			Ile 238 N
351	28.40	Glu 14 Oε1,Oε2	367	36.01	Arg 156 O
		Ala 285 N			Glu 161 Oε1,Oε2
352	35.58	Lys 21 O			Gly 137 N
353	30.59	Asp 9 Oδ1	368	21.41	Ile 138 O
354	45.55	Pro 56 O			Gly 141 O
355	46.74	Gly 74 O			Asn 153 Oδ1
		Asp 78 Oδ2			Thr 154 N
		Glu 87 Oε1	369	32.18	Glu 163 Oε1
		Gly 74 N			Arg 167 Nη2
356	63.23	Asp 87 Oδ1	370	26.13	Glu 171 Oε2
		Lys 107 Nζ	371	36.60	Glu 171 Oε1
357	36.86	Val 108 O			Glu 171 O
		Glu 113 Oε1	372	46.03	Ala 172 O
		Gly 125 N			Glu 299 Oε2
358	41.52	Val 108 O	373	34.50	His 179 Nε2
		Val 126 O			His 179 O
359	33.82	Leu 112 O			Val 232 O
		Pro 251 O			Asp 231 N
360	26.89	Ile 122 O	374	21.73	Ser 182 Oγ
		Ala 228 N	375	30.78	Val 183 O
361	51.81	Glu 120 N			Thr 235 Oγ1
362	41.32	Glu 121 O			Gly 236 N
		ILe 122 O	376	49.41	Leu 292 O
		Arg 124 O	377	52.51	His 300 Nε2

Table 3.5.1 continued

Solvent	B-factor	Protein Hydrogen bonds	Solvent	B-factor	Protein Hydrogen bonds
378	45.84	Asn 286 O	394	54.99	Ser 244 O
		Ala 289 N	395	49.81	Glu 155 Oε1
379	59.90		396	44.26	Gly 10 N
380	24.17	His 273 Nε2	397	68.36	Trp 77 O
		Asn 286 O			Glu 87 Oε2
		Ala 290 N	398	41.27	
381	47.23		399	52.92	Glu 334 Oε1
382	33.48	Lys 282 Nζ	400	43.36	Ala 338 O
383	24.97	Gly 281 O	401	67.25	
384	34.23	Ile 284 O	402	68.73	Lys 310 Nζ
		Ala 285 O	403	52.26	Lys 310 O
		Ala 331 O	404	45.90	Val 168 O
385	29.73	Glu 334 N	405	55.30	
386	42.69	Lys 185 N	406	45.17	Glu 51 Oε1
387	45.23	Ala 280 O			Phe 53 N
388	63.14	Asp 278 O	407	46.53	Asp 98 Oδδ2
		Ala 280 O			Leu 99 O
389	48.17		408S	86.18	
390	65.22		O1	85.39	
391	83.91		O2	85.52	
392	58.07	Lys 197 Nζ	O3	85.72	
393	44.61	Arg 196 Nη1,Nη2	O4	85.06	

many known proteins with well-refined and well-ordered crystal structures. As there are no other peaks on the electron density map, the water molecules that occupy the large portion of the molecular suface are considered to be similar to those of bulk water (Creighton, 1984).

3-6 Main chain hydrogen bonds

A summary of the hydrogen bonds involving in main chain atoms is given in Table3.6.1~3.6.3. There are 209 main chain to main chain, 46 main chain to side chain and 64 main chain to solvent molecule hydrogen bonds. There are 4 main chain to main chain and 2 main chain to side chain hydrogen bonds between residues in first and second domain. Intersubunit contact region involves 8 main chain to main chain and 3 main chain to side chain hydrogen bonds. These hydrogen bonds listed in Table3.9.2.

3-7 Heavy-atom binding sites

Three kinds of heavy-atom reagents were used to determine the crystal structure. The best of the derivatives used in the initial phase determination was potassium tetrachloroplatinate(II). There are five sites in the platinum derivative. Generally, the site of platinum is close to the sulfur atom of methionine. In the case of IPMDH, there exist methionine residues near all the platinum sites. The major site of the derivative is close to the Sδ atom of Met296 and His300. This methionine is located at the C terminal end of helix i, and His300 is on the following loop. The second platinum site, Pt2,

Table 3.6.1 Hydrogen bonds between main chains

Main Chain (donner)	Main Chain (acceptor)	Comment	Main Chain (donner)	Main Chain (acceptor)	Comment
Val 3 N	Ala 35 O	βB-βA	Ala 64 N	Val 61 O	αc end
Ala 4 N	Ala 66 O	βB-βC	Glu 65 N	Lys 2 O	βC-βB
Val 5 N	Glu 37 O	βB-βA	Val 67 N	Pro 267 O	βC-βD
Leu 6 N	Leu 68 O	βB-βC	Leu 68 N	Ala 4 O	βC-βB
Asp 9 N	Ser 71 O		Leu 69 N	Phe 269 O	βC-βD
Ile 11 N	Ser 275 O		Gly 70 N	Leu 6 O	βC-βB
Gly 12 N	Asp 9 O	αa	Trp 77 N	Gly 74 O	
Val 15 N	Ile 11 O	αa	Asp 78 N	Pro 75 O	
Thr 16 N	Gly 12 O	αa	Ile 84 N	Pro 81 O	
Glu 17 N	Pro 13 O	αa	Gly 89 N	Ser 85 O	
Ala 18 N	Glu 14 O	αa	Leu 90 N	Pro 86 O	αd
Ala 19 N	Val 15 O	αa	Leu 91 N	Glu 87 O	αd
Leu 20 N	Thr 16 O	αa	Ser 92 N	Thr 88 O	αd
Lys 21 N	Glu 17 O	αa	Leu 93 N	Gly 89 O	αd
Val 22 N	Ala 18 O	αa	Arg 94 N	Leu 90 O	αd
Leu 23 N	Ala 19 O	αa	Lys 95 N	Leu 91 O	αd
Arg 24 N	Leu 20 O	αa	Ser 96 N	Ser 92 O	αd
Ala 25 N	Lys 21 O	αa	Gln 97 N	Leu 93 O	αd end
Leu 26 N	Val 22 O	αa	Asp 98 N	Lys 95 O	αd end
Asp 27 N	Leu 23 O	αa	Leu 99 N	Arg 94 O	αd end
Glu 28 N	Arg 24 O	αa	Ala 101 N	Leu 262 O	βF-βE
Ala 29 N	Ala 25 O	αa	Asn 102 N	Arg 132 O	βF-βG
Glu 30 N	Leu 26 O	αa	Leu 103 N	Ala 260 O	βF-βE
Leu 32 N	Asp 27 O	αa end	Arg 104 N	Ile 130 O	βF-βG
Ala 35 N	Met 1 O	βA-βB	Ala 106 N	Val 128 O	βF-βG
Glu 37 N	Val 3 O	βA-βB	Val 108 N	Val 126 O	βF-βG
Phe 39 N	Val 5 O	βA-βB	Leu 112 N	Phe 109 O	
Gly 42 N	Pro 52 O		Leu 115 N	Leu 112 O	
Ile 46 N	Gly 42 O	αb	Ser 116 N	Glu 113 O	
Asp 47 N	Gly 43 O	αb	Ile 122 N	Lys 119 O	
Ala 48 N	Ala 44 O	αb	Ala 123 N	Lys 119 O	
Thr 57 N	Pro 54 O		Arg 124 N	Glu 120 O	
Arg 58 N	Pro 54 O		Val 126 N	Ala 123 O	
Lys 59 N	Glu 55 O		Val 128 N	Ala 106 O	βG-βF
Gly 60 N	Pro 56 O	αc	Leu 129 N	Asp 231 O	βG-βH
Val 61 N	Thr 57 O	αc	Ile 130 N	Arg 104 O	βG-βF
Glu 62 N	Arg 58 O	αc	Val 131 N	Val 233 O	βG-βH
Glu 63 N	Lys 59 O	αc	Arg 132 N	Asn 102 O	βG-βF

Hydrogen bonds between first and second domains are shaded.

Table 3.6.1 continued

Main Chain (donner)	Main Chain (acceptor)	Comment	Main Chain (donner)	Main Chain (acceptor)	Comment
Glu 133 N	Thr 235 O	βG-βH	Glu 201 N	Lys 197 O	αf
Leu 134 N	Phe 100 O	βG-βF	Val 202 N	Thr 198 O	αf
Phe 140 N	Gly 137 O		Gly 203 N	Val 199 O	αf
Gly 141 N	Gly 137 O		Arg 204 N	Glu 201 O	αf
Gly 145 N	Trp 152 O	βK-βL	Gly 205 N	Val 202 O	αf end
Trp 152 N	Gly 145 O	βL-βK	Tyr 206 N	Gly 203 O	αf end
Glu 161 N	Ser 158 O	αe start	Val 209 N	Tyr 206 O	
Val 162 N	Ser 158 O	αe start	Ala 210 N	Lys 178 O	βJ-βI
Glu 163 N	Lys 159 O	αe	Glu 212 N	Val 180 O	βJ-βI
Arg 164 N	Pro 160 O	αe	Gln 214 N	Ser 182 O	βJ-βI
Ala 166 N	Val 162 O	αe	Val 216 N	Asp 184 O	βJ-βI
Arg 167 N	Glu 163 O	αe	Met 219 N	Tyr 215 O	αg
Val 168 N	Arg 164 O	αe	Ala 220 N	Val 216 O	αg
Ala 169 N	Val 165 O	αe	His 222 N	Ala 218 O	αg
Phe 170 N	Ala 166 O	αe	Leu 223 N	Met 219 O	αg
Glu 171 N	Arg 167 O	αe		Ala 220 O	αg
Ala 172 N	Val 168 O	αe	Val 224 N	Met 221 O	αg
Ala 173 N	Ala 169 O	αe	Arg 225 N	Met 221 O	αg
Arg 174 N	Phe 170 O	αe	Ser 226 N	His 222 O	αg end
Lys 175 N	Ala 172 O	αe end	Arg 229 N	Ser 226 O	
Arg 176 N	Ala 173 O	αe end	Phe 230 N	Pro 227 O	
Lys 178 N	Ala 173 O	αe end	Val 232 N	His 179 O	βH-βI
Val 180 N	Ala 210 O	βI-βJ	Val 233 N	Leu 129 O	βH-βG
Val 181 N	Val 232 O	βI-βH	Val 234 N	Val 181 O	βH-βI
Ser 182 N	Glu 212 O	βI-βJ	Thr 235 N	Val 131 O	βH-βG
Val 183 N	Val 234 O	βI-βH	Gly 240 N	Gly 236 O	αh start
Asp 184 N	Gln 214 O	βI-βJ	Asp 241 N	Asn 237 O	αh
Val 188 N	Lys 185 O		Ile 242 N	Ile 238 O	αh
Leu 189 N	Lys 185 O		Leu 243 N	Phe 239 O	αh
Glu 193 N	Leu 189 O	αf start	Ser 244 N	Gly 240 O	αh
Phe 194 N	Glu 190 O	αf	Asp 245 N	Asp 241 O	αh
Trp 195 N	Val 191 O	αf		Ile 242 O	αh
Arg 196 N	Gly 192 O	αf	Leu 246 N	Ile 242 O	αh
Lys 197 N	Glu 193 O	αf	Ala 247 N	Leu 243 O	αh
Thr 198 N	Phe 194 O	αf	Ser 248 N	Ser 244 O	αh
Val 199 N	Trp 195 O	αf	Val 249 N	Leu 246 O	αh
Glu 200 N	Arg 196 O	αf	Leu 250 N	Ala 247 O	αh
	Lys 197 O	αf	Gly 252 N	Val 249 O	αh end

Table 3.6.1 continued

Main Chain (donner)	Main Chain (acceptor)	Comment	Main Chain (donner)	Main Chain (acceptor)	Comment
Gly 255 N	Ser 253 O		Leu 304 N	Leu 298 O	
Leu 256 N	Ser 253 O		Val 305 N	Gly 303 O	
Ala 260 N	Leu 103 O	βE-βF	Ala 308 N	Leu 304 O	αj start
Ser 261 N	Val 268 O	βE-βD	Arg 309 N	Val 305 O	αj
Leu 262 N	Ala 101 O	βE-βF	Lys 310 N	Glu 306 O	αj
Arg 264 N	Asp 98 O		Val 311 N	Leu 307 O	αj
Gly 265 N	Gln 97 O		Glu 312 N	Ala 308 O	αj
Val 268 N	Ser 261 O	βD-βE	Asp 313 N	Arg 309 O	αj
Phe 269 N	Val 67 O	βD-βC	Ala 314 N	Lys 310 O	αj
Glu 270 N	Ser 259 O	βD-βE	Val 315 N	Val 311 O	αj
Val 272 N	Leu 257 O	βD-βE	Ala 316 N	Glu 312 O	αj
Ile 279 N	Ala 276 O		Lys 317 N	Asp 313 O	αj
Ala 280 N	Pro 277 O		Ala 318 N	Ala 314 O	αj
Lys 282 N	Ile 279 O		Leu 319 N	Val 315 O	αj
Ile 291 N	Pro 287 O	αi start	Leu 320 N	Lys 317 O	αj
Leu 292 N	Thr 288 O	αi	Glu 321 N	Lys 317 O	αj
Ser 293 N	Ala 289 O	αi	Thr 322 N	Ala 318 O	αj end
	Ala 290 O	αi	Leu 327 N	Pro 324 O	
Ala 294 N	Ala 290 O	αi	Phe 336 N	Gly 332 O	αk start
Ala 295 N	Ile 291 O	αi	Thr 337 N	Thr 333 O	αk
Met 296 N	Leu 292 O	αi	Ala 338 N	Glu 334 O	αk
Met 297 N	Ser 293 O	αi	Thr 339 N	Ala 335 O	αk
Leu 298 N	Ala 294 O	αi	Val 340 N	Phe 336 O	αk
Glu 299 N	Ala 295 O	αi end	Leu 341 N	Thr 337 O	αk
His 300 N	Met 296 O	αi end	Arg 342 N	Ala 338 O	αk
Ala 301 N	Met 297 O	αi end	His 343 N	Thr 339 O	αk
Phe 302 N	Met 297 O		Leu 344 N	Leu 341 O	αk end
	Leu 298 O		Ala 345 N	Leu 341 O	αk end
Gly 303 N	Glu 296 O				

Hydrogen bonds between first and second domains are shaded.

Table 3.6.2 Hydrogen bonds between mainchain and sidechain

Side Chain (donner)	Main Chain (acceptor)	Comment	Main Chain (donner)	Side Chain (acceptor)	Comment
Thr					
Thr 16 Oγ1	Gly 12 O	αa-αa	Lys 2 N	Glu 65 Oε1	βB-cCjoint
Thr 57 Oγ1	Pro 40 O	βA-αc	Leu 34 N	Asp 27 Oδ1	aAloop-αa
Thr 88 Oγ1	Arg 82 O	Cdloop-αd	Gly 43 N	Asp 9 Oδ2	αb-Baloop
Thr 198 Oγ1	Phe 194 O	αf-αf	Gly 44 N	Asp 9 Oδ1	αb-Baloop
Thr 235 Oγ1	Val 183 O	βI-βH	Gly 73 N	Asp 9 Oδ2	Cdloop-Baloop
Thr 266 Oγ1	Val 61 O	EDloop-αc	Gly 111 N	Glu 113 Oε1	FGloop-FGloop
	Val 64 O	-αc		Glu 113 Oε2	-FGloop
Thr 322 Oγ1	Ala 318 O	αj-αj	Leu 118 N	Ser 116 Oγ	FGloop-FGloop
Thr 333 Oγ1	Gly 283 O	αk-Diloop	Thr 135 N	Glu 133 Oε1	βG-GKloop
Thr 337 Oγ1	Thr 333 O	αk-αk	Gly 136 N	Glu 133 Oε1	GKloop-GKloop
Ser				Glu 133 Oε2	-GKloop
Ser 71 Oγ	Asp 9 O	Cdloop-Baloop	Ile 138 N	Glu 155 Oε2	GKloop-βL
Ser 92 Oγ	Thr 88 O	αd-αd	Ser 158 N	Glu 161 Oε1	βL-αe
Ser 96 Oγ	Leu 93 O	αd-αd	Arg 177 N	Asp 231 Oδ2	eIloop-gHloop
Ser 116 Oγ	Leu 250 O	FGloop-αh	His 179 N	Asp 231 Oδ2	βI-gHloop
Ser 248 Oγ	Asp 245 O	αh-αh	Ala 186 N	Asp 184 Oδ1	βI-βI
Ser 261 Oγ	Ala 260 O	βE-βE	Asn 237 N	Tyr 157 Oη	αh-βL
Ser 275 Oγ	Gly 73 O	Diloop-Cdloop	Thr 266 N	Gln 97 Oε1	EDloop-dFloop
Ser 293 Oγ	Ala 289 O	αi-αi	Gly 281 N	Glu 14 Oε1	Diloop-αa
Arg			Gly 283 N	Glu 14 Oε2	Diloop-αa
Arg 144 Nη1	Tyr 139 O	βK-GKloop			
Arg 176 Nη2	Asp 127 O	eIloop-βG			
Arg 176 Nη1	Asp 231 O	eIloop-gHloop			
Arg 177 Nη2	Asp 127 O	eIloop-βG			
Arg 177 Nη2	Phe 230 O	eIloop-gHloop			
Arg 309 Nη2	Lys 175 O	αj-αe			
Gln Asn					
Gln 97 Nε2	Ser 96 O	dFloop-dFloop			
Asn 102 Nδ2	Ala 260 O	βF-βE			
Asn 237 Nδ2	Leu 134 O	αh-GKloop			
	Gly 136 O	-GKloop			

Hydrogen bonds between first and second domains are shaded.

Table 3.6.3 Hydrogen bonds between mainchain and solvent molecule

Solvent (donner)	Main Chain (acceptor)	Comment
Wat 346 O	Val 3 O	βB
Wat 347 O	Phe 39 O	βA
	Phe 41 O	Ab joint
Wat 348 O	Gly 8 O	Baloop
Wat 350 O	Glu 14 O	αa
Wat 351 O	Ile 279 O	Diloop
Wat 352 O	Lys 21 O	αa
Wat 354 O	Pro 56 O	αc
Wat 355 O	Gly 74 O	Cdloop
Wat 357 O	Val 108 O	βF
Wat 358 O	Val 108 O	βF
	Val 126 O	βG
Wat 359 O	Leu 112 O	FGloop
	Pro 251 O	hEloop
Wat 360 O	Ile 122 O	FGloop
Wat 362 O	Glu 121 O	FGloop
	Ile 122 O	FGloop
	Arg 124 O	FGloop
Wat 363 O	Asp 127 O	βG
	Pro 227 O	gHloop
	Phe 230 O	gHloop
Wat 364 O	Val 128 O	βG
Wat 365 O	Glu 161 O	αe
Wat 367 O	Arg 156 O	βL
Wat 368 O	Ile 138 O	GKloop
	Gly 141 O	GKloop
Wat 371 O	Glu 171 O	αe
Wat 372 O	Ala 172 O	αe
Wat 373 O	His 179 O	βI
	Val 232 O	βH
Wat 375 O	Val 183 O	βI
Wat 376 O	Leu 292 O	αi
Wat 378 O	Asn 286 O	Diloop
Wat 380 O	Asn 286 O	Diloop
Wat 383 O	Gly 281 O	Diloop
Wat 384 O	Ile 284 O	Diloop
	Ala 285 O	Diloop
	Ala 331 O	jkloop
Wat 387 O	Ala 280 O	Diloop
Wat 388 O	Asp 278 O	Diloop
	Ala 280 O	Diloop
Wat 394 O	Ser 244 O	αh
Wat 397 O	Trp 77 O	Cdloop
Wat 400 O	Ala 338 O	αk
Wat 403 O	Lys 310 O	αj
Wat 404 O	Val 168 O	αe
Wat 407 O	Leu 99 O	dFloop
Wat 408 O		

Main Chain (donner)	Solvent (acceptor)	Comment
Gly 8 N	Wat 347 O	Baloop
Gly 10 N	Wat 396 O	Baloop
Phe 53 N	Wat 406 O	bcloop
Gly 74 N	Wat 355 O	Cdloop
Glu 120 N	Wat 361 O	FGloop
Gly 125 N	Wat 357 O	FGloop
Asp 127 N	Wat 363 O	βG
Gly 137 N	Wat 367 O	GKloop
Thr 154 N	Wat 368 O	βL
Lys 185 N	Wat 386 O	βI
Asp 231 N	Wat 373 O	gHloop
Gly 236 N	Wat 375 O	βH
Ile 238 N	Wat 366 O	αh
Ala 285 N	Wat 351 O	Diloop
Ala 289 N	Wat 378 O	αi
Ala 290 N	Wat 380 O	αi
Glu 334 N	Wat 385 O	αk

is close to Met221 and Asp245 which are on the helix g and h, respectively. They are involved in the pocket described in the following section 3-9, and exposed to solvent. The minor site Pt5 is 4.2Å from the Pt2 site. Pt5 is also in contact to Met221 and Arg225. This site has relatively large B-factor. The third site, Pt3, is close to Met146 which belongs to arm region and Lys197' which is on helix f of another subunit. The fourth site, Pt4, is at the N-terminal end of the polypeptide chain. The Sδ atom of Met 1 is also close to the Pt4 site.

The major site, UF1, of the uranium derivative is close to Glu201, Glu163 and Glu17' (superscript prime denotes the symmetry related molecule costructing dimer) which is in the symmetry related molecule. UF4 is 7Å from UF1 site and close to Arg204, Glu201 and Tyr36'. These two sites are in the lattice contact region, and the interacting molecules are related by crystallographic symmetry not to form dimer. The second site, UF2, is contact to Asp241, Asp245 and Asp217. It is 4.63Å from the Pt2 site. The third site, UF3, is at the position occupied by Oε1 atom of Tyr157 in the native structure. As this site is buried in the molecule, it is of lower occupancy. In the uranium derivative, Tyr157 may be turned around the C_α-C_β bond to make the space for the uranium atom. Glu133 and Glu161 are close to the UF3 site. The fifth site, UF5, is also of lower occupancy and close to Asp78 and Glu87.

The only site of gold derivative, Au1 is in contact to Glu155 and

Table 3.7.1 List of the residues contact with heavy-atoms

derivative	site	contact residue
K_2PtCl_4	Pt1	Met296, His300
	Pt2	Met221, Asp245
	Pt3	Met146, Lys197'
	Pt4	Met 1
	Pt5	Met221, His222, Arg225
$K_3UO_2F_5$	U1	Glu201, Glu163, Glu 17'
	U2	Asp241, Asp245, Asp217
	U3	Glu133, Glu161
	U4	Arg204, Glu201', Tyr36'
	U5	Asp78, Gln87
$NaAu(CN)_4$	Au1	Glu155

Superscript prime denotes the symmetry related molecule.

located close to the two-fold symmetry axis. As described in section2-5, the site is near the special position (0 1/3 5/6)..

3-8 Lattice contacts

The subunit contacts with three other subunits which are related by the crystallographic symmetry or lattice translation. The most extensive lattice contacts are the subunit-subunit interactions in the dimer which is described in section3-9. The other contacts are listed in Table3.8.1. The contacts with neighbor molecules mainly occur between the B-a loop, helix a, D-i loop and helix k in the first domain and helix e, f-J loop and E-D loop in the second domain. Hydrogen bonds mediated by water molecules greatly contribute to the contacts. On the contrary, there are few direct interactions between the molecules. Hydrophobic interaction could not be found in the contacts with neighbor molecules.

3-9. Quarternary structure

One subunit of Tt-IPMDH is in close contact with a second subunit to give rise to an identical dimer in solution. The spatial arrangement of these subunits is depicted in Figure3.9.1. The two subunits are related by a crystallographic two-fold axis and interact in their second domains. Most of the subunit contacts are essentially hydrophobic in nature.

The close contacts are present in four regions. The dominant interactions are made by helix g and helix h, and form a hydrophobic

Table 3.8.1 Hydrogen bonds contributing lattice contacts

protein	intermidiate water	intermidiate water (neigbor)	neighbor protein
Gly 8 O	Wat 348		Gly 205 O
Asp 9 Oδ1	Wat 353		
Gly 10 N	Wat 396		
Glu 14 O	Wat 350	Wat 369	Arg 167 Nη2
Glu 334 Oδ1	Wat 399		Glu 163 Oδ1
	Wat 398		
Lys 21 Nζ			
Ala 280 O	Wat 387	Wat 381	
		Wat 404	Val 168 O
Gly 281 O	Wat 383		Arg 164 O
Lys 282 Nζ			Gly 263 O
	Wat 382		Gly 265 O

Line indicates hydrogen bonding.
The neighbor molecule constructing a dimer is eliminated from the table.

core. Intersubunit hydrogen bonds and van der Waals interactions are given in Table3.9.1 and 3.9.2. The subunit contact region is shown in Figure3.9.2.

The first contact region is around the F-G loop where the side-chain of Pro117 and Leu118 are in close proximity to Pro117' and Leu118' , and the Nζ atom of Lys 119 (Lys119') hydrogen bonds to both of the peptide oxygen atoms of Ser116' (Ser116) and Lys118' (Lys118) (Figure3.9.3).

The second is around the helices g, h, g' and h', where the side-chains of Val216, Ala220 and Val 224 in helix g and Ile238, Phe239, Ile242 and Leu246 in helix h face toward the corresponding residues in helices g' and h' and constitute the hydrophobic core at the center of the dimer. The spatial alignment of the helices are shown in Figure3.9.4. The side chains of Met221 and 221' are also in close proximity in this core.

The third region is around Ile138, Tyr139, Val188', Leu189'. and also Ile138', Tyr139', Val188, Leu189. The side chains of these hydrophobic residues are part of the constituents of the hydrophobic core (Figure3.9.5).

The final contact region is the arm region which protrudes from the second domain. Strands K and L form the intersubunit β-sheet. with strands K' and L' by hydrogen bonding. Minor contacts also occur around the arm region. There exist hydrophobic interactions between Ala151 in the arm and Val191' and Phe194' in helix f. Nη atom of Arg144 and Oε atom of Glu190' form an intersubunit salt bridge (Figure3.9.6).

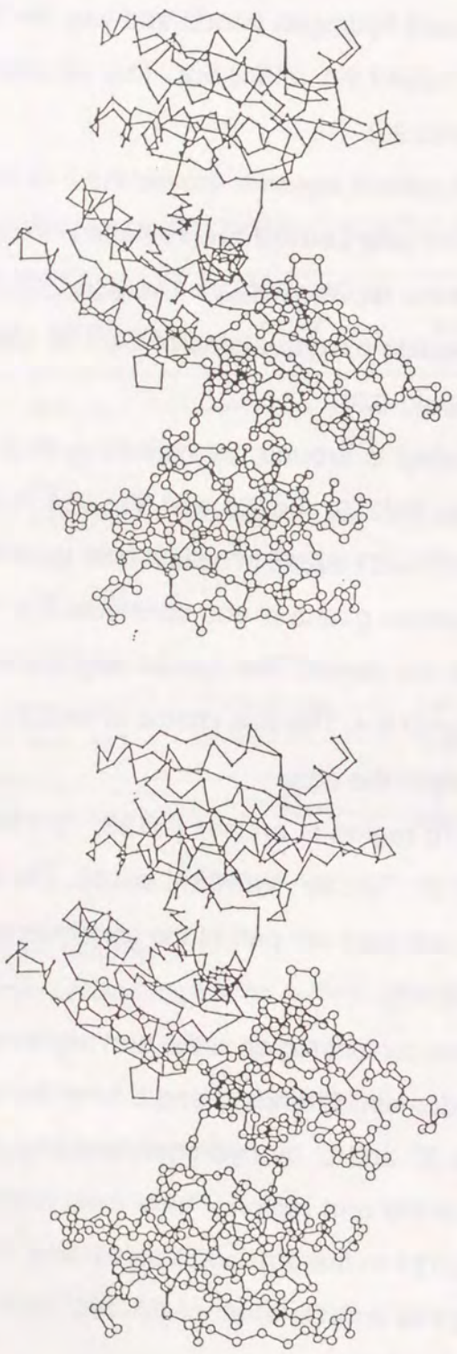


Figure 3.9.1 A stereo drawing of the Cα backbone of Tt-IPMDH dimer. One subunit of the dimer is drawn with large balls and thick sticks and the other with small balls and thin sticks.

Table 3.9.1 Intermolecular hydrophobic contacts at dimer interface

distance	≤4Å	4Å≤ ≤5Å	distance	≤4Å	4Å≤ ≤5Å
Arg 114'		Lys 119	Ala 151'	Val 191	Arg 156
Leu 115'		Lys 119		Phe 194	Glu 190
Ser 116'	Lys 119		Trp 152'	Val 191	Thr 154
Pro 117'	Val 224	Leu 118			Glu 155
		Lys 119			Arg 156
		Ile 122	Asn 153'	Leu 189	Thr 154
		Arg 225			Glu 155
Leu 118'	Lys 119	Pro 117			Glu 190
		Leu 118			Val 191
Lys 119'	Ser 116	Arg 114	Thr 154'		Trp 152
	Leu 118	Leu 115			Asn 153
		Pro 117			Thr 154
		Lys 119	Glu 155'	Ile 138	Trp 152
Ile 122'		Pro 117			Asn 153
Ile 138'	Glu 155	Ile 138	Arg 156'		Glu 150
	Leu 189				Ala 151
Tyr 139'	Val 188	Lys 185			Trp 152
		Leu 189	Tyr 157'	Glu 150	
Arg 144'	Glu 190		Ser 158'		Ala 149
Gly 145'	Glu 190				Glu 150
Met 146'	Glu 190	Phe 194	Lys 159'		Ala 149
	Glu 193		Lys 185'		Tyr 139
Ser 147'		Phe 194			Ile 238
Glu 148'		Phe 194	Val 188'	Tyr 139	
Ala 149'	Phe 194	Ser 158	Leu 189'	Ile 138	Tyr 139
		Lys 159		Asn 153	Ile 238
Glu 150'	Arg 156	Ser 158	Glu 190'	Arg 144	Ala 151
	Tyr 157			Gly 145	Asn 153
	Phe 194			Met 146	

Superscript prime denotes the symmetry related molecule constructiong dimer.

Table 3.9.1 continued

distance	≤4Å	4Å≤ ≤5Å	distance	≤4Å	4Å≤ ≤5Å
Val 191'	Ala 151	Asn 153	Ile 238'	Phe 239	Lys 185
	Trp 152				Leu 189
Glu 193'	Met 146		Phe 239'	Ile 238	Ile 242
Phe 194'	Ala 149	Met 146	Asp 241'		Asp 217
	Glu 150	Ser 147	Ile 242'		Val 216
	Ala 151	Glu 148			Asp 217
Val 216'		Ile 242			Ala 220
Asp 217'		Asp 241			Phe 239
		Ile 242			Ile 242
		Asp 245	Asp 245'	Met 221	Asp 217
Ala 220'		Ile 242	Leu 246'	Val 224	Ala 220
		Leu 246			Met 221
Met 221'	Asp 245	Leu 246			Leu 246
	Val 249		Val 249'	Met 221	
Val 224'	Pro 117	Leu 250		Val 224	
	Val 224			Arg 225	
	Leu 246		Leu 250'		Val 224
	Val 249		Leu 254'		Arg 225
Arg 225'	Val 249	Pro 117			
		Leu 254			

Table 3.9.2 Intermolecular hydrophilic contacts at dimer interface

Donner	Acceptor	Comment
Lys 119 Nζ	Glu 113' O	FGloop-FGloop
	Ser 116' O	-FGloop
	Leu 118' O	-FGloop
Arg 144 Nη2	Glu 190' Oε2	βK-αf
Ala 151 N	Tyr 157' O	βL-βL
Glu 155 N	Asn 153' O	βL-βL

Superscript prime denotes the symmetry related molecule constructiong dimer.

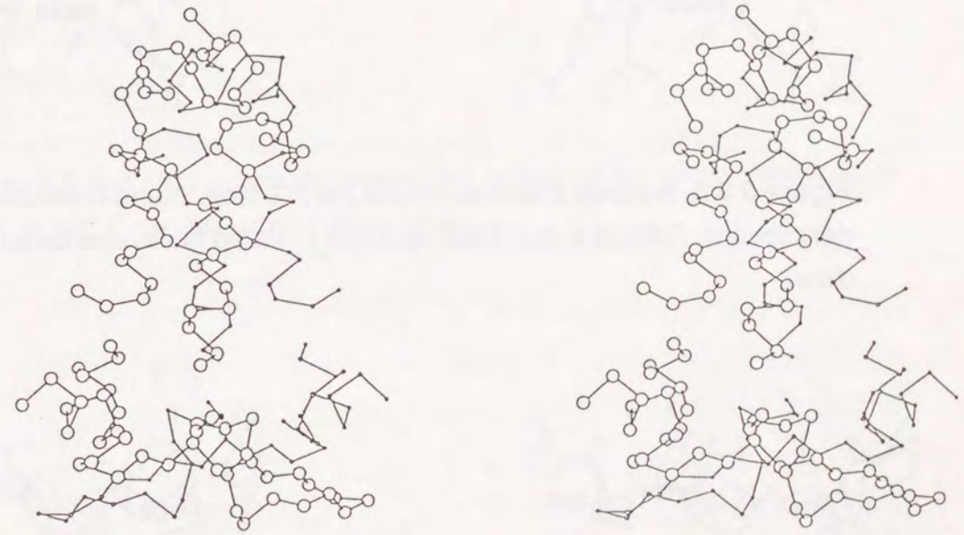


Figure 3.9.2 A stereo drawing of the main chain atoms of subunit contact region. One subunit of the dimer is drawn with large balls and sticks and the other with small balls and sticks.

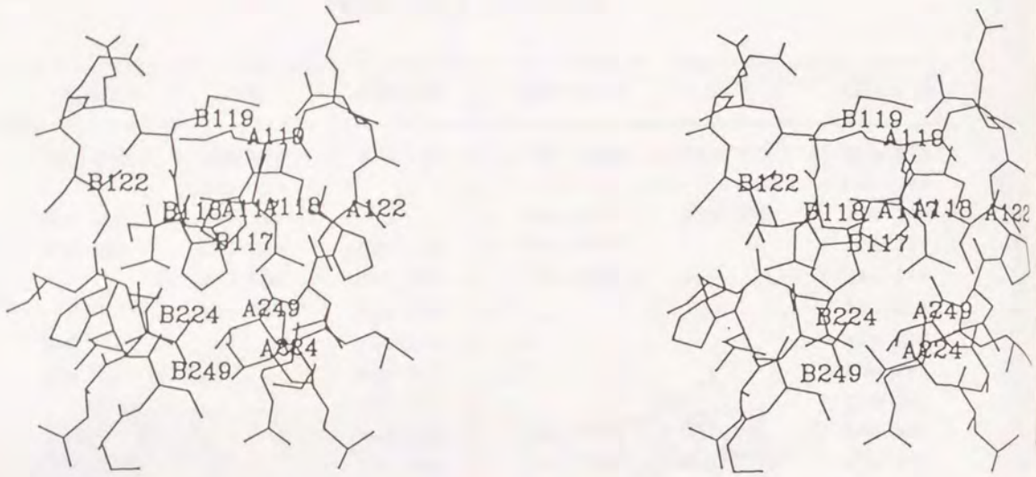


Figure 3.9.3 A stereo drawing around the FG-loop. A and B denote subunit descriptions. Subunit A and B are symmetry related molecules constructing a dimer.

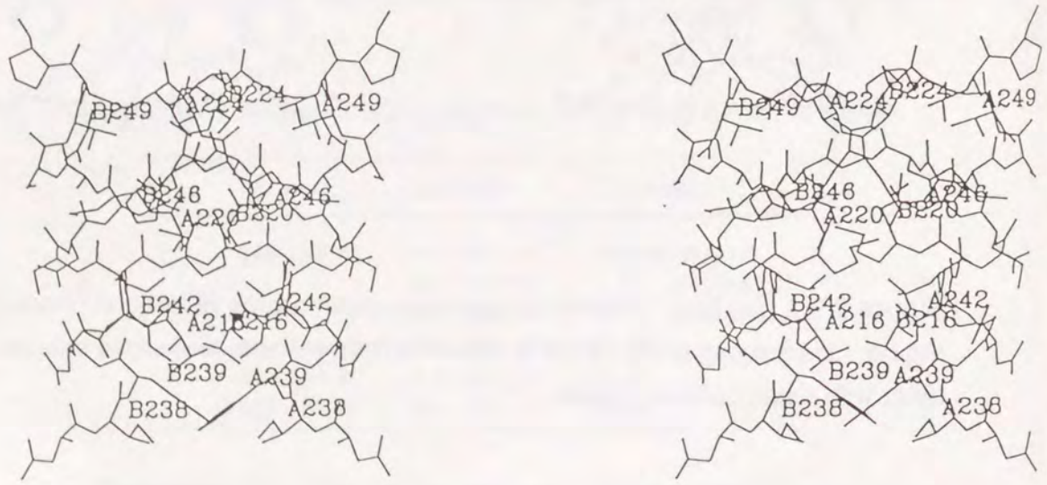


Figure 3.9.4 A stereo drawing of four helices g,h,g',h' constructiong the central hydrophobic core. A and B denote subunit descriptions.

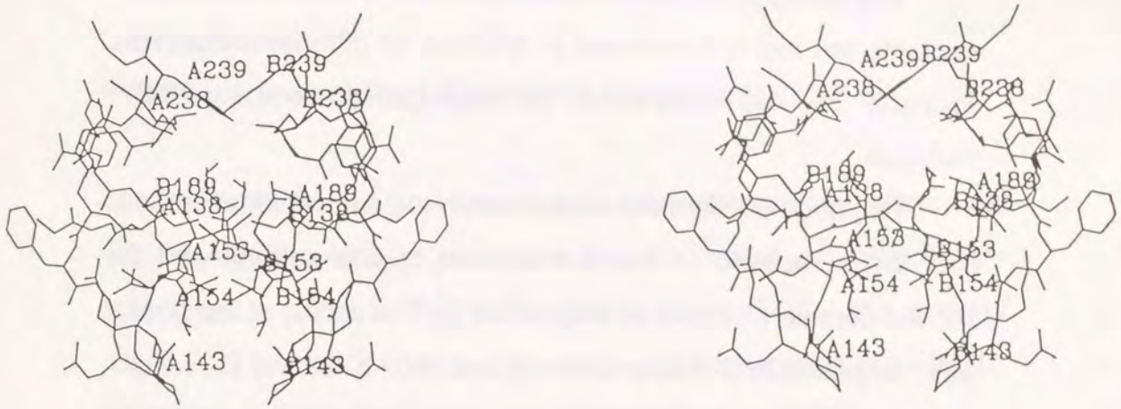


Figure 3.9.5 A stereo drawing around the bottom of the central hydrophobic core. A and B denote subunit descriptions.

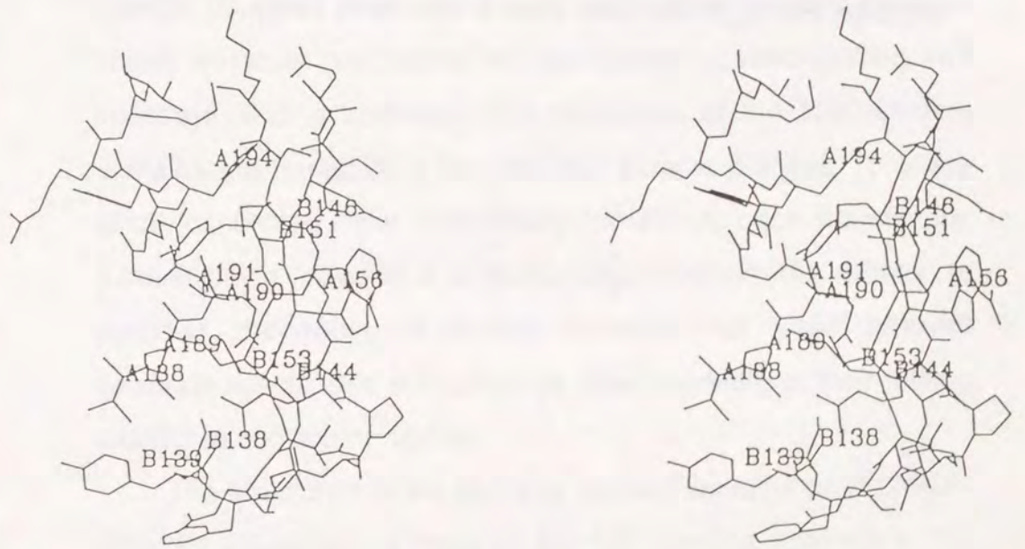


Figure 3.9.6 A stereo drawing of the intersubunit β-sheet and its surrouiding residues. A and B denote subunit descriptions.

The amino acid residues that participate in these close subunit contacts are highly conserved in IPMDHs of different organisms. Probably they are important to form the dimeric form for IPMDH molecule.

Two large pockets are recognized in the dimeric structure, and they are composed of the first domain of one subunit and the second domain of the other (Figure3.9.1). The ceiling of the pocket was composed of C-d loop, D-i loop and helix d. Strand D, E, F, G, helix h, G-K loop and h-E loop make the wall of the pocket. The bottom of the pocket is formed by another subunit. I'-f' loop, F'-G' loop, helix f' and g' are main components. The significance of these findings is that they indicate that the dimeric form is essential for Tt-IPMDH to express its enzymatic activity. This is further supported by the results of a recent X-ray analysis of *E. coli* ICDH-substrate complex (Hurley et al., 1990; Dean & Koshland, 1990).

Chapter-4
Comparison with Other Dehydrogenases

Tt-IPMDH was compared with other dehydrogenases based on three-dimensional structures for elucidating an evolutionary implication of dehydrogenase family. In addition, the primary structure of IPMDHs from other organisms are compared with Tt-IPMDH and their three dimensional structure is predicted.

4-1 Well known dehydrogenases

A comparison was made between the structure of Tt-IPMDH and those of well known dehydrogenases such as LADH (Branden et al., 1973; Åkeson & Jones, 1981), LDH (Adams et al., 1970), GAPDH (Buehner et al., 1974). The polypeptide chains of the well-known enzymes are folded into two domains; NAD-binding and substrate-binding domains. The structures of the NAD-binding domains are based on a six-stranded parallel β-sheet. They are strictly conserved under evolutionary constraints, each strand in the β-sheet being joined in a common sequence with the α-helix. In contrast, the substrate-binding domains that have different substrate specificities and catalyses differ markedly in their folding topologies and conformations.

The structures of the first and second domains of Tt-IPMDH show no similarities to those of the NAD-binding domains in the joined sequences of β-strands with α-helices. The parallel β-strands

A, B, C and D in the first domain are connected in the sequence B→A→C→D, and the β-strands G, H, I and J in the second domain in the sequence of G→I→J→H, whereas the β-strands C, D, E and F in the NAD-binding domains are connected in the common sequence of D→E→F→C (Figure3.1.4). We therefore infer that IPMDH is not related evolutionarily to those well-known enzymes.

4-2 ICDH from *E. coli*

The structure of Tt-IPMDH was compared with isocitrate dehydrogenase (ICDH) (Hurley et al., 1989), which is a bifunctional enzyme catalyzing decarboxylation and dehydrogenation. Tt-IPMDH shows marked similarities to *E. coli* ICDH, both in its amino acid sequence and in its overall folding topology (Figure3.1.4). A minor difference is the presence of strands M, N and helices l, m in ICDH. The β-strands M and N are inserted in the fJ-loop of IPMDH and participate in the formation of the central β-sheet in the second domain. Likewise, helices l and m are located in the jk- and KL-loops, respectively, of IPMDH. A comparison of their amino acid sequences shows that the regions corresponding to these extra secondary structures of ICDH are absent in Tt-IPMDH (Figure4.2.1). These similarities indicate that IPMDH and ICDH are diverged from a common ancestral protein.

Further determination of three dimensional structure of bifunctional enzymes that catalyses both decarboxylation and dehydrogenation will allow for clear evolutionary implication of the enzymes.

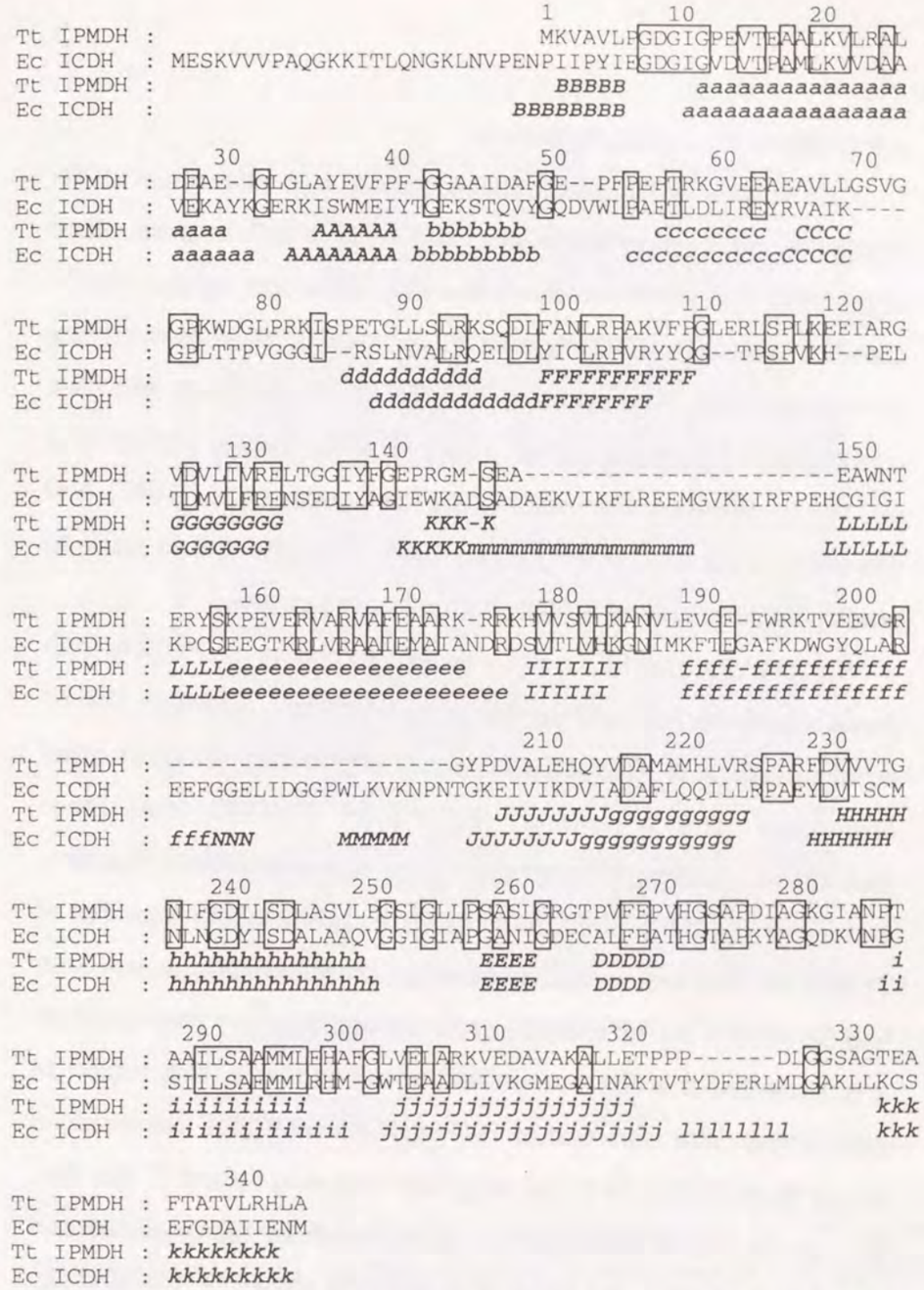


Figure 4.2.1 Alignment of the amino acid sequences of T.thermophilus IPMDH and E.coli ICDH.

Residues that are conserved between these two enzymes are boxed. The secondary structures of the enzymes are shown below the sequences with boldfaced letters. The capital and small letters designate β-strands and α-helices.

4-3 IPMDHs from other organisms

As shown in Figure4.3.1, the amino acid sequence of Tt-IPMDH is highly homologous to those of IPMDHs from other organisms; i.e *Yarrowia lipolytica* (Yl) (Davidow et al., 1987), *Candida utilis* (Cu) (Hamasawa et al., 1987), *Saccharomyces cerevisiae* (Sc) (Andreadis et al., 1984), *E. coli* (Ec), *Bacillus coagulans* (Bc) (Sekiguchi et al., 1986) , *Bacillus subtilis* (Bs) (Imai et al., 1987), *Bacillus caldotenax* (Bt) (Sekiguchi et al., 1987) and *Thermus aquaticus* (Ta) (Kirino, 1991). Therefore it is easy to predict the structures of IPMDH from other organisms.

Except for Yl-IPMDH, the polypeptide chains of IPMDHs from these organisms have almost the same length as Tt-IPMDH. The N-terminal region of Yl-IPMDH is much longer than those of other organisms. The N-terminal regions of IPMDHs from other mesophiles are slightly longer than those of thermophiles. Probably, β-strand B of IPMDH from mesophiles is slightly longer than that of Tt-IPMDH. The amino acid sequences corresponding to a-A loop and β-strand A are different between mesophiles and thermophiles. It is expected that the structures corresponding to this region of mesophiles differ from that of Tt-IPMDH. Yl-IPMDH is five residues longer than Tt-IPMDH in the long loop following strand C. But Bs-, Bc-, Bt- and Ec-IPMDH are the same length as thermophiles and Cu- and Sc-IPMDH are rather two residues short. Yl-, Cu- and Sc-IPMDH have five residual insertions in the short loop following the strand E. Beyond the residue number 320 of Tt-IPMDH, the length

Figure 4.3.1 Alignment of the amino acid sequences of the IPMDHs from *Yarrowia lipolytica* (Yl), *Candida utilis* (Cu), *Saccharomyces cerevisiae* (Sc), *E. coli* (Ec), *Bacillus coagulans* (Bc), *B. subtilis* (Bs), *Bacillus caldotenax* (Bt), *T. thermophilus* (Tt) and *Thermus aquaticus* (Ta). Amino acid residues identical with those of Tt- IPMDH are shaded.

```
          300        310        320        330        340
Yl  FSLNMKPAGDAVEAAVKESVEAGITTADIGGSSSTSEVGDLLPTRSRSCSRRSKSFLRRIDG
    hhh  hhhhhhhhhhhhhhhhh              hhhhhh
    sssss                    ssssss        sssssss            sssss

Cu  LSLDLYEEGVAVETAVKQVLDAGIRTGDLKGTNSTTEVGDAVA---EAVKKILA
    hhhhhhhhhhhhhhhhhhhhhh               hhhhhhh---hhhhhhhh
    ssssss   sssssssssssssssss      sssssssss   ---  sssss

Sc  LSLNLPEEGKAIEDAVKKVLDAGIRTGDLGGSNSTTEVGDAVA---EEVKKILA
    hhhhhhhhhhhhhhhhhhhhhh               hhhhhhh---hhhhhhhh
    sssss          sssssssssssss      ssssss   ---  sssss

Ec  YSLDADDARFAIERAINRALEEGIRTGDLARGVPPLVPMKWAISLPAM
        hhhhhhhhhhhhhhhhhhh           hhhhhhhhhh
    sss                        sssssssssssssssssss

Bc  YSFGLEKEAAALEKAVDDVAQDGYCTGDLQVANGKVVSTIEITDRLIEKLNNSAARTRIFQ
    hhhhhhhhhhhhhhhhhhhhh                hhhhhhhhhhhhhh
    sssss              sssssssssssssssssssssssssssssssss      sssss

Bs  TSFGLEEEAKAVEDAVNKVLASGKRTRDLARSEEFS-STQAIT---EEVKAAIMSANTISNV

Bt  LSFGLTAEAGGRARVWQALALGSGSRLGQRRPHLSTNEMV------EEIKAAVLDYTAIAQI
    hhhhhhhhhhhhhhhhhhhhh                hhhh------hhhhhhhh   hhhhh
    ssssss        sssssss               sssss------sssssssssssssss

Ta  PRFGLVELAKRVEDAVAKALRETPPP-DLGGSAGTQAFTEEVLRH-------L
    hhhhhhhhhhhhhhhhhhhhhhh    -     hhhhhhhhhhhhh-------h
      sssssss                  -     sssssssssss-------s

Tt  HAFGLVELARKVEDAVAKALLETPPP-DLGGSAGTEAFTATVLRH-------LA
        jjjjjjjjjjjjjjjjj      -     kkkkkkkkkkk

Yl  RSKLTRLRVGLPAGSASV

Bt  MTVYA
    hhhh
    sssss
```

Figure 4.3.2 Prediction of secondary structures of IPMDHs C-terminal region from various organisms. 'h' shows the possible helix and 's' the possible strand. Sequences are alligned to Tt-IPMDH and identical residues with Tt-IPMDH are shaded. Residue numbers are corresponding to Tt-IPMDH. The secondary structure of Tt-IPMDH is also shown with bold faced letters in the figure as a reference.

and the sequence of amino acids is quite different among various organisms. But from the prediction of secondary structure (Chou & Fasman, 1978; Osgthor & Robson, 1978), it is expected that they have a helix in their C-terminal region after a loop structure composed of 10~15 residues (Figure4.3.2).

4-4 *T. aquaticus* IPMDH

Thermus aquaticus is also extreme thermophile and the amino acid sequence is previously reported (Kirino, 1991), and the number of amino acid is 344 residues (Figure4.4.1). Thirty-three residues different from that of *T. thermophilus* (10% of all the residues). In these replacement of residues, eight residues are alternation from arginine to lysine or lysine to arginine. Almost all the replaced residues are distributed on the molecular surface. Only two exceptional residues, Val180 and Val181 of Tt-IPMDH are buried in the molecule. Thr88 is varied to serine in Ta-IPMDH. The Oγ atom of Thr88 and the carbonyl oxygen atom of Arg82 make hydrogen bond in Tt-IPMDH. As serine has also Oγ atom, the hydrogen bond may be also present in Ta-IPMDH.

```
                1         10         20         30         40         50
Tt IPMDH :      MKVAVLPGDGIGPEVTEAALKVLRALDEAEGLGLAYEVFPFGGAAIDAFGE
Ta IPMDH :      *R******************R**K****R*****T**T*********GY*
Secondary
structure:       BBBBB      aaaaaaaaaaaaaaaaaaa     AAAAAA bbbbbbb

               50        60         70         80         90        100
Tt IPMDH :      PFPEPTRKGVEEAEAVLLGSVGGPKWDGLPRKISPETGLLSLRKSQDLFAN
Ta IPMDH :      ****V******AE**********T***A********S***A**********
Secondary
structure:          cccccccc  CCCC                 ddddddddddd    FFF

                     110        120        130        140        150
Tt IPMDH :      LRPAKVFPGLERLSPLKEEIARGVDVLIVRELTGGIYFGEPRGMSEAEAWN
Ta IPMDH :      **************************************************
Secondary
structure:      FFFFFFFF               GGGGGGGG           KKKK  LLLL

                    160        170        180        190        200
Tt IPMDH :      TERYSKPEVERVARVAFEAARKRRKHVVSVDKANVLEVGEFWRKTVEEVGR
Ta IPMDH :      *************K**********R*LT**********************PQ
Secondary
structure:      LLLLLeeeeeeeeeeeeeeeee   IIIIIII    fffffffffffffff

                    210        220        230        240        250
Tt IPMDH :      GYPDVALEHQYVDAMAMHLVRSPARFDVVVTGNIFGDILSDLASVLPGSLG
Ta IPMDH :      *******D***********KN*****************************
Secondary
structure:          JJJJJJJJgggggggggg       HHHHHhhhhhhhhhhhhhh

                    260        270        280        290        300
Tt IPMDH :      LLPSASLGRGTPVFEPVHGSAPDIAGKGIANPTAAILSAAMMLEHAFGLVE
Ta IPMDH :      **************************************************
Secondary
structure:        EEEE    DDDDD                 iiiiiiiiii      jj

                    310        320        330        340
Tt IPMDH :      LARKVEDAVAKALLETPPPDLGGSAGTEAFTATVLRHLA
Ta IPMDH :      **KR*********R*************Q***EE*****-
Secondary
structure:      jjjjjjjjjjjjjj           kkkkkkkkkkk
```

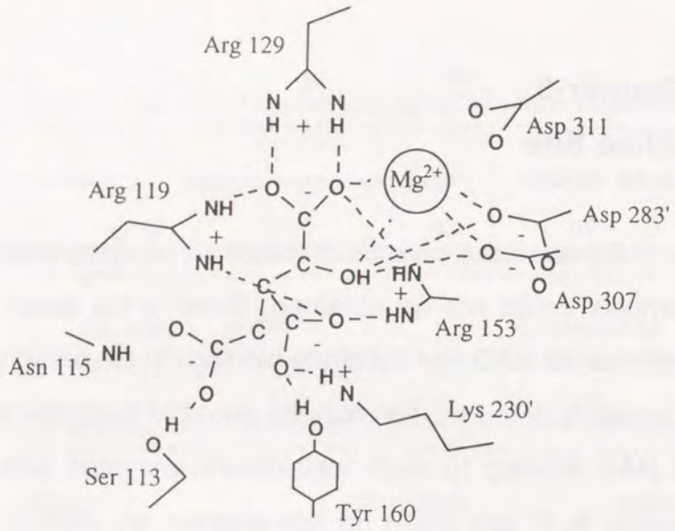
Figure 4.4.1 Comparison of amino acid sequences of Tt- and Ta-IPMDH. Asteriscs denote the same amino acid residues as Tt-IPMDH.

Chapter-5
Active Site

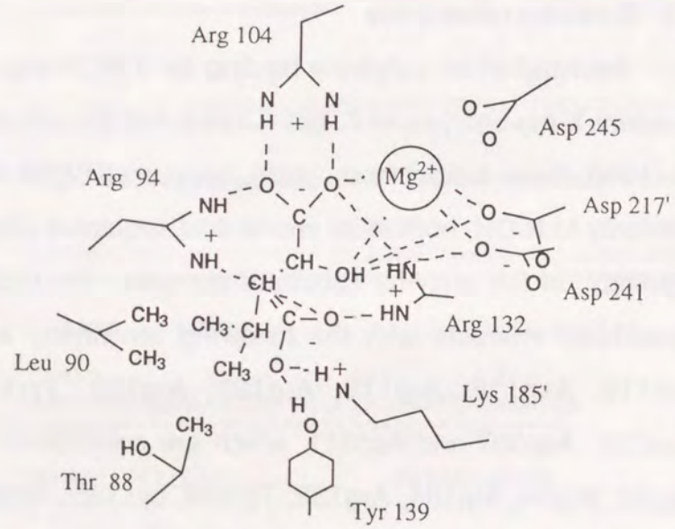
Because the crystals of substrate-enzyme and NAD-enzyme complex could not be obtained, there is no direct experimental evidence for NAD and substrate bindings in the present analysis. But a tentative determination may be possible from the X-ray analyses of NAD binding to such well-known enzymes and of substrate binding to *E. coli* ICDH. In this chapter, we discuss the possible substrate and NAD binding sites.

5-1 Substrate binding site

Information on substrate binding for IPMDH was derived from a recent X-ray analysis of E. coli ICDH-substrate complex (Hurley et al., 1990; Dean & Koshland, 1990), because IPMDH shows marked similarity to ICDH, both in its amino acid sequence and in its folding topology. In the enzyme-substrate complex, the substrate moiety (isocitrate) interacts with the following ten amino acid residues: Ser113, Asn115, Arg119, Arg129, Arg153, Tyr160, Lys230', Asp283', Asp307 and Asp311, which are corresponding to The88, Leu90, Arg94, Arg104, Arg132, Tyr139, Lys185', Asp217', Asp241 and Asp245 in IPMDH in amino acid sequence, respectively. As for the amino acid residues responsible for this substrate interaction, ICDH differs from IPMDH only in Ser113 and Asn115, which are replaced by Thr88 and Leu90 in in IPMDH. Moreover, the amino



I C D H



I P M D H

Figure 5.1.1 Schematic drawing of the substrate binding site of ICDH and possible substrate binding site of IPMDH

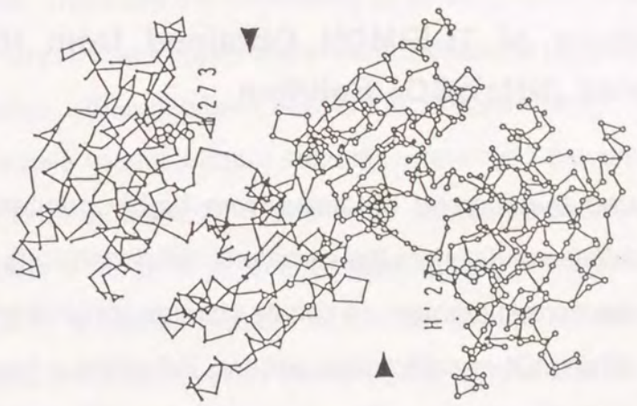
acid residues that correspond to the residues of ICDH for the substrate interaction are conserved among the IPMDHs of different organisms, the only exception being recognized at Thr88 of Tt-IPMDH. It is therefore proposed that The88, Leu90, Arg94, Arg104, Arg132, Tyr139, Lys185', Asp217', Asp241 and Asp245 of Tt-IPMDH are concerned with substrate (3-isopropylmalate) binding, and Leu90 in the first domain specially with the recognition and regulation of the substrate binding.

Further information on the substrate binding site of IPMDH was obtained from kinetic analysis, chemical modification studies and site directed mutagenesis experiments (Miyazaki, 1991). In the presence of isopropylmalate and Mn^{2+}, which is necessary for enzymatic activity, His273 was protected from the chemical modification with diethylpyrocarbonate. His273 is present near the substrate binding site (Miyazaki et al., 1989) and exists in the pocket (Figure5.1.2), indicating that the active site is included in this pocket. Furthermore, the R94Q and R132Q, which were mutants substituted by arginine to glutamine, completely lost their activities, thereby showing the responsibility of these residues for the enzymatic activity. In addition, the R104Q mutant decreased the affinity to isopropylmalate. The affinity of 3-isopropylmalate-1-carboxyamide, which is a substrate analogue, to enzyme was considerably reduced in R104Q mutant. These analyses show Arg104 interacts 1-carboxyl group of 3-isopropylmalate.

From these results, we propose the substrate binding model for IPMDH as shown in Figure5.1.1 and strongly suggest that the dimeric form is essential for IPMDH to display the enzymatic activity.

5-2 Possible NAD binding site

For the well-known enzymes, the basic structures of their NAD-binding domains are composed of a six-stranded parallel β-sheet surrounded by some α-helices (Rossmann, 1975), the structural half comprising a common β–α–β–α–β fold centered on a highly conserved sequence, Gly - X - Gly - X - X - Gly (where X is any amino acid residues) (Scrutton et al., 1990; Wierenga et al., 1985). The dipole moment of one of the α-helices contributes to the binding of the NAD moiety near the C-terminal ends of the β-strands by interaction favorably with the pyrophosphate moiety (Wierenga et al., 1985). Tt-IPMDH also has the β–α–β–α–β folds common to the NAD-binding domains, represented by the B-a-A-b-c-C fold in the first domain and I-f-J-g-H in the second domain. But, as IPMDH has no highly conserved sequences, such as Gly - X - Gly - X - X - Gly, in its amino acid sequence, it is difficult in the present analysis to determine which β–α–β–α–β folds in IPMDH are concerned with NAD binding.



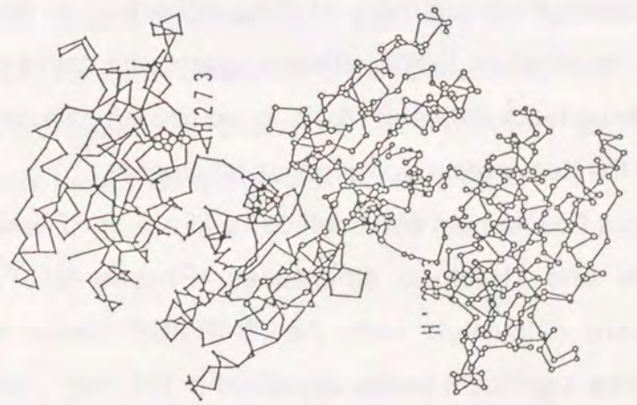


Figure 5.2.1 A stereo drawing of the Cα backbone of Tt-IPMDH dimer with side chains of His273 residue. One subunit of the dimer is drawn with large balls and thick sticks and the other with small balls and thin sticks. His 273 is drawn with the largest ball. Arrows indicate the clefts which are considered as active sites.

Chapter-6

The Structure of Tt-IPMDH Obtained from Highly Concentrated $(NH_4)_2SO_4$ Solution

We found that crystal obtained from highly concentrated ammonium sulfate solution has the same unit-cell dimensions as the crystal from the normal ammonium sulfate solution (original crystal), but give an different X-ray diffraction pattern, indicating a structural difference in both crystals. To elucidate the difference, we determined the three-dimensional structure of the enzyme crystallized from highly concentrated $(NH_4)_2SO_4$ solution (in this thesis, we call it 'S-Cryst') .

6-1 Data collection and reduction

The intensity data was collected up to 2.2Å resolution with an IP-diffractometer (RAXIS II-c). All the data could be collected from only one crystal. The summary of data collection is given in Table6.1.1. The structure factor variance against the native crystal was 12.9% for up to 5Å data and 18.4% for up to 3.0Å data, which is equivalent to the heavy atom isomorphous differences.

Difference Fourier map with coefficients $||F_{S\text{-}Cryst}| - |F_{native}||$ was calculated to find structural differences. Phases for Fourier synthesis were calculated from the Tt-IPMDH native model structure. Three significant peeks appeared in the map ; close to Arg132, Ser182 and Met297 sites. These peeks were considered as

solvent molecules, because there are also present in the native model. Therefore the occupancy of solvent molecules are increased in S-Cryst. The Fourier maps with coefficients $||2F_{S\text{-}Cryst}| - |F_{calc}||$ and $||F_{S\text{-}Cryst}| - |F_{calc}||$, where F_{calc} is the structure factor calculated from the native model without solvents, were also calculated to find the structural differences. There were additional peaks considered as solvent molecules in these maps. Using the modeling program FRODO, additional solvent molecules of thirteen were found on the molecular surface. With these solvents, the structure model was refined by PROLSQ. The crystallographic R-value of the initial model was 30.2%, and after ten cycles it was reduced to 19.6%. When the native model structure with solvent molecules was used for the initial model, the R-value could not be reduced less than 21%. The refinement statistics are given in Table6.2.2.

6-2 Structure description and comparison with Tt-IPMDH

As shown in Table6.2.1, the r.m.s. difference for all protein atoms between the S-Cryst model and the native model is 0.24Å. Hence, the major difference of these two models is the distribution of solvent molecules bound on the molecular surface. The fixed solvent molecules are listed in Table6.2.3.

Table 6.1.1

Summary of data collection of S-Cryst and statistics

X-ray Source	Cu-Kα
X-ray generator	Rigaku RU200
Focus size	0.3 × 3mm
X-ray power	40kV, 100mA
Monochromatization	graphite plate
IP size	200×200mm
Pixel size	105μm
No. of crystal used	1
φ(spindle) - axis	approx. c axis
Crystal - to - IP distance	90mm
Resolution limit	2.1Å
Oscillation range per frame	1.5Å
No. of frames	21
Total oscillation range	31.5°
Exposure time	20min / frame
No. of observed reflections	
full	27,725
partial	19,360
total	47,085
No. of independent reflections	27,415
Completeness †	81%
R - merge ‡	
full reflections	4.20%
partial reflections	4.87%
total reflections	4.48%
No. of rejected reflections ¶	
full reflections	1
partial reflections	0

† Considering the blind region.

‡ R-merge = $\Sigma\Sigma | I_i(h) / G_i - \langle I(h) \rangle | / \Sigma\Sigma \langle I(h) \rangle$
where G_i is the inverse scale factor, $I(h)$ the diffraction intensity of the symmetry equivalent reflections, and $\langle I(h) \rangle$ the mean value of $I(h)$.

¶ Rejection criteria are $C_R = 0.3\ I_{mean} + 0.1 \langle I(h) \rangle$ for the reflections measured more than twice and $C_R = 3\ (0.3\ I_{mean} + 0.1 \langle I(h) \rangle)$ for the reflections with only two observations, where I_{mean} is the mean of all $I(h)$ values. The reflections with the differences $D_i = | \langle I(h) \rangle - I_i(h) / G_i | > C_R$ are rejected (Rossmann *et al.*, 1979).

Table 6.1.2 Summary of least-squares parameters and deviations of the crystal kept in the solution of high precipitant concentration at 2.2Å resolution (imaging plate)

	Target	r.m.s deviation
Bonding distances (Å)		
1-2 bond	0.020	0.014
1-3 angle	0.030	0.035
1-4 planar	0.050	0.047
Planar groups (Å)	0.020	0.011
Chiral volumes (Å³)	0.150	0.175
Non-bonded contacts (Å)		
Single torsion	0.500	0.201
Multiple torsion	0.500	0.243
Possible hydrogen bond	0.500	0.262
Torsion angles (deg.)		
Planar	3.0	2.5
Staggered	15.0	23.0
Orthonormal	20.0	30.6
Thermal factors (Å²)		
Main-chain bond	1.000	0.465
Main-chain angle	1.500	0.803
Side-chain bond	1.500	0.962
Side-chain angle	2.000	1.496

Table 6.1.3 Dependency of the R-factors on resolution

Resolution (Å)	No. of reflections	*R* - factor	*R* - factor (accumulated)
5.00 ~ 4.00	2,281	0.137	0.137
4.00 ~ 3.30	3,518	0.167	0.152
3.30 ~ 2.90	3,533	0.212	0.168
2.90 ~ 2.63	3,402	0.240	0.180
2.63 ~ 2.45	2,879	0.255	0.187
2.45 ~ 2.32	2,518	0.263	0.192
2.32 ~ 2.20	2,520	0.264	0.196
5.00 ~ 2.20	20,651	-	0.196

Table 6.2.1 Difference of the coordinate between native and S-Cryst

	Main chain atom	Side chain atom	All protein atom
mean distance	0.13Å	0.22Å	0.17Å
r.m.s.deviation	0.15Å	0.31Å	0.24Å
<0.5Å	1373	1149	2522
0.5Å≤ <1.0Å	6	49	55
1.0Å≤ <1.5Å	1	7	8
1.5Å≤ <2.0Å	0	2	2
2.0Å≤	0	3	3

All the above data were calculated after fitted the coordinate of S-Cryst with that of native crystal by least square method. Solvent molecules were eliminated from the calculation.

Table 6.2.2 List of atoms whose coordinates are differnt more than 1.5Å between native and S-Cryst structure

residue	atom	distance (Å)	comment
Asp 27	Oδ2	1.91	C-end of α-a
Glu 63	Cγ	2.18	C-end of α-c
	Cδ	1.71	
	Oε1	3.76	
Arg176	Nη1	3.31	C-end of α-e

Table 6.2.3 Fixed solvent molecules and their protein hydrogen bonds

Solvent	B-factor	Protein Hydrogen bonds	Solvent	B-factor	Protein Hydrogen bonds
346	56.34	Val 3 O	363	25.23	Asp 127 O
347	48.81	Phe 39 O			Pro 227 O
		Phe 41 O			Phe 230 O
		Gly 8 N			Asp 127 N
348	44.73	Gly 8 O	364	39.63	Asp 127 Oδ2
349S	72.15				Val 128 O
O1	71.96				Arg 176 Nη2
O2	71.46	Arg167' Nη1	365	29.88	Glu 133 Oε1
O3	71.32	Arg167' Nη2			Glu 161 O
O4	72.18		366	18.91	Glu 155 Oε1,Oε2
350	33.31	Glu 14 O			Ile 238 N
351	27.98	Glu 14 Oε1,Oε2	367	34.47	Arg 156 O
		Ile 279 O			Glu 161 Oε1,Oε2
		Gly 283 N			Gly 137 N
		Ala 285 N	368	32.59	Ile 138 O
352	29.41	Lys 21 O			Gly 141 O
353	28.84	Asp 9 Oδ1			Asn 153 Oδ1
354	59.59	Pro 56 O			Thr 154 N
355	37.89	Gly 74 O	369	34.53	Glu 163 Oε1
		Glu 87 Oε1			Arg 167 Nη2
		Gly 74 N	370	27.71	
356	50.97	Lys 107 Nζ	371	37.86	Glu 171 Oε1
357	36.86	Val 108 O			Glu 171 O
		Glu 113 Oε1	372	53.09	Ala 172 O
		Gly 125 N			Glu 299 Oε2
358	36.84	Val 108 O	373	28.61	His 179 Nε2
		Val 126 O			His 179 O
359	35.86	Leu 112 O			Val 232 O
		Pro 251 O			Asp 231 N
360	28.62	Ile 122 O	374	16.08	Ser 182 Oγ
		Ala 228 N	375	29.17	Thr 235 Oγ1
361	43.99	Glu 120 N	376	55.38	Leu 292 O
362	44.30	Glu 121 O			Glu 312 Oε1
		ILe 122 O	377	55.47	Glu 299 Oε1
		Arg 124 O			

Solvent molecules whose coordinates are different more than 0.5Å between native and S-Cryst structure are shaded. Solvent molecules peculiar to S-Cryst crystal are represented by bold faced letters.

Table 6.2.3 continued

Solvent	B-factor	Protein Hydrogen bonds	Solvent	B-factor	Protein Hydrogen bonds
378	63.84	Asn 286 O	**402**	60.68	Glu 312 Oε2
		Ala 289 N			Arg 176 Nη1
379	62.84		403	49.17	Lys 310 O
380	21.43	His 273 Nε2			His 343 O
		Asn 286 O	404	48.45	Val 168 O
		Ala 290 N	405	62.36	
381	46.25		406	61.42	Glu 51 Oε1
382	32.77	Lys 282 Nζ			Phe 53 N
383	26.40	Gly 281 O	407	36.99	Asp 98 Oδ2
384	37.32	Ile 284 O			Leu 99 O
		Ala 285 O	408S	86.20	
		Ala 331 O	O1	85.58	
385	45.87	Glu 334 N	O2	85.63	
386	44.91	Lys 185 N	O3	85.69	
387	55.71		O4	85.34	
388	65.48	Pro 277 O	**409**	54.07	Glu 65 O
		Ala 280 O	**410**	48.84	Ala 64 O
389	52.38				Thr 266 Oγ1
390	28.95		**411**	64.88	Arg 264 O
391	82.33		**412**	48.33	Glu 14 Oε2
392	53.43	Lys 197 Nζ	**413**	35.21	
393	63.17	Lys 2 Nζ	**414**	44.24	
394	52.36	Ser 244 O	**415**	56.16	
395	43.24	Glu 155 Oε1	**416**	44.71	Asp 127 Oδ1
396	35.06	Gly 10 N	**417**	45.45	Gly 203 O
397	56.04	Trp 77 O			Tyr 206 O
		Asp 78 Oδ2			Val 209 O
		Glu 87 Oε2	**418**	53.63	Asp 241 Oδ2
398	34.51				Asp 241 O
399	46.87				Asp 245 Oδ1
400	45.94	Ala 338 O	**419**	73.01	
401	65.91		**420**	70.18	
			421	30.76	Asp 47 Oδ1

Solvent molecules whose coordinates are different more than 0.5Å between native and S-Cryst structure are shaded. Solvent molecules peculiar to S-Cryst crystal are represented by bold faced letters.

Chapter-7
Structure of Chimeric IPMDH

If certain amino acid residues of Bs-IPMDH were replaced with those of Tt-IPMDH, it would be expected that the enzyme could be a highly thermostable. Only a few residue may be essential for the thermostability. Point mutational technique such as site directed mutagenesis is general approach to determine which residues are essential for the thermostability. But the method requires a lot of substitution experiments from residue to residue. Recent development of genetic manipulation techniques, gene fusion techniques, enables us to create a fusion enzyme, i.e., 'chimeric' enzyme. The gene fusion technique is powerful and efficient, because the fusion enzyme is equal to a multi-mutated enzyme whose characters may be inherited from both parents.

As shown in Figure7.0.1, chimeric IPMDHs were produced by gene fusion technique between *T.thermophilus* (Tt) and *Bacillus subtilis* (Bs) which is a mesophile (Akutsu, 1989; Numata et al, 1990, 1991). These enzymes show various thermostability (7.0.2). The chimeric enzyme, 4M6T, is more thermostable than the enzyme from Bs-IPMDH and less thermostable than that from Tt-IPMDH. Although the chimeric enzyme, 2T2M6T, contains a larger amount of residues from Tt-IPMDH, it is slightly less stable than 4M6T. This indicates that one of essential regions for thermostability exists between the 75th and the 135th amino acid residue of Tt-IPMDH.

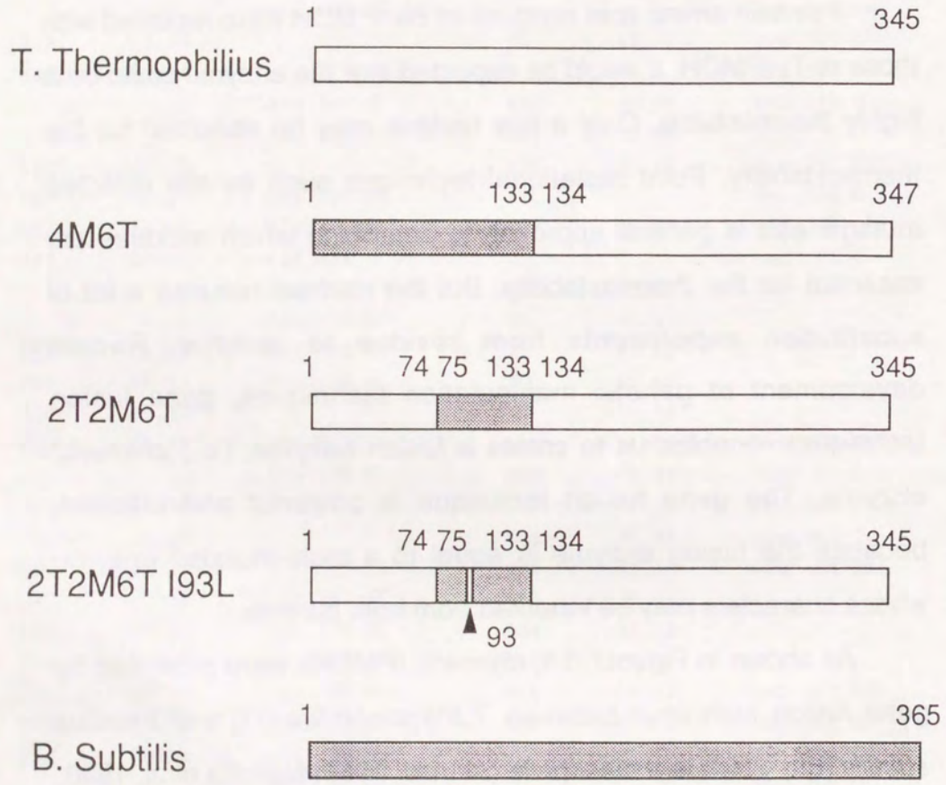


Figure 7.0.1 Construction of chmeric enzymes

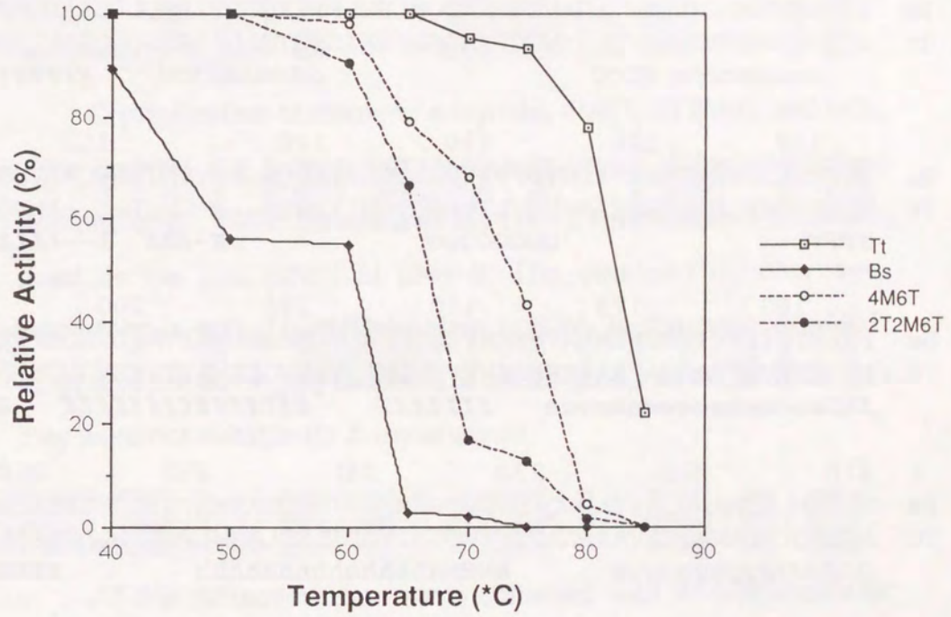


Figure 7.0.2 Remaining activity after the heat treatment for 10min.

```
     1        10        20        30        40        50
Bs   MKKRIALLPGDGIGPEVLESATDVLKSVAERFNHEFEFEYGLIGGAAIDEHHNP
Tt     MKVAVLPGDGIGPEVTEAALKVLRALDEAEGLGLAYEVFPFGGAAIDAFGEP
        BBBBB     aaaaaaaaaaaaaaaaaa     AAAAAA bbbbbbb

            60        70        80        90        100
Bs   LPEETVAACKNAEAILLGAVGGPKWDQNLSELRPEKGLLSIRKQLDLFANLRPV
Tt   FPEPTRKGVEEAEAVLLGSVGGPKWDGLPRKISPETGLLSLRKSQDLFANLRPA
     ccccccccc CCCC                 ddddddddd    FFFFFFF

       110       120       130       140       150
Bs   KVFESLSDRSPLKKEYIDNVDFVIVRELTGGLYFGQPSKRYVNTEGEQEAVDTL
Tt   KVFPGLERLSPLKEEIARGVDVLIVRELTGGIYFGEP--R-GMSEAE--AWNTE
     FFFF               GGGGGGGG            K-KKK  L--LLLLL

        160       170       180       190       200
Bs   FYKRTEIERVIREGFKMAATRKGKVTSVDKANVLESSRLWREVAEDVAQEFPDV
Tt   RYSKPEVERVARVAFEAARKRRKHVVSVDKANVLEVGEFWRKTVEEVGRGYPDV
     LLLeeeeeeeeeeeeeeee   IIIIIII    fffffffffffffff   JJ

     210       220       230       240       250       260
Bs   KLEHMLVDNAAMQLIYAPNQFDVVVTENMFGDILSDEASMLTGSLGMLPSASLS
Tt   ALEHQYVDAMAMHLVRSPARFDVVVTGNIFGDILSDLASVLPGSLGLLPSASLG
     JJJJJJgggggggggg       HHHHHhhhhhhhhhhhhh         EEEE

            270       280       290       300       310
Bs   SSGLHLFEPVHGSAPDIAGKGMANPFAAILSAAMLLRTSFGLEEEAKAVEDAVN
Tt   R-GTPVFEPVHGSAPDIAGKGIANPTAAILSAAMMLEHAFGLVELARKVEDAVA
        DDDDD                 iiiiiiiiii      jjjjjjjjjjj

        320       330       340
Bs   KVLASGKRTRDLARSEEFS-STQAITEEVKAAIMSANTISNV
Tt   KALLETPPP-DLGGSAGTEAFTATVLRHLA
     jjjjj             kkkkkkkkkkk
```

Figure 7.0.3 Alignment of the amino acid sequences of Tt-IPMDH and Bs IPMDH. Secondary structure of Tt-IPMDH is also shown.

Furthermore, the mutant produced by substituting leucine for Ile93 of 2T2M6T, I93L, shows higher thermostability than 4M6T and 2T2M6T. X-ray crystallographic analysis of 4M6T, 2T2M6T and I93L was carried out to interpret these results from the three-dimensional structure.

7-1 Crystallization

Crystallization of chimeric enzymes, 4M6T, 2T2M6T and I93L, were carried out with a hanging drop vapor diffusion method (Sakurai et al., 1991; Onodera et al., 1991). Ammonium sulfate was used as the precipitant at pH6~8. The obtained crystals were isomorphous with Tt-IPMDH native crystal. In the case of 4M6T, other crystals belonging to different space group were obtained, but they were not suitable for X-ray analysis.

7-2 Data collection

All the diffraction data were collected with IP-diffractometer (RAXIS II-c). The intensity data could be collected up to 2.2Å for 4M6T, 2.0Å for 2T2M6T and 1.9Å for I93L. The X-ray source was Cu-Kα radiation from a rotating anode X-ray generator (Rigaku Ru-200) with fine focus mode operated at 40kV 100mA, monochromatized by Ni coated mirrors. The beam was focused on the imaging plate by Franks double-mirror optics (Franks, 1955). Crystals were mounted the c* axis parallel to the spindle axis. The summary of data collection was given in Table7.2.1 ~ 7.2.3. For 2.7Å resolution data, intensity variance against Tt-IPMDH crystal was

Table 7.2.1

Summary of data collection of 4M6T-crystal and statistics

X-ray Source	Cu-Kα
X-ray generator	Rigaku RU200
Focus size	0.3 × 3mm
X-ray power	40kV, 100mA
Monochromatization	Ni & mirror
IP size	200×200mm
Pixel size	105μm
No. of crystal used	1
φ(spindle) - axis	approx. c axis
Crystal - to - IP distance	90mm
Resolution limit	2.2Å
Oscillation range per frame	1.5Å
No. of frames	21
Total oscillation range	31.5°
Exposure time	40min / frame
No. of observed reflections	
full	18,554
partial	8,507
total	25,612
No. of independent reflections	16,420
Completeness †	57.3%
R - merge ‡	
full reflections	5.49%
partial reflections	6.18%
total reflections	5.62%
No. of rejected reflections ¶	
full reflections	0
partial reflections	0

† Considering the blind region.

‡ R-merge = $\Sigma\Sigma \mid I_i(h) / G_i - \langle I(h) \rangle \mid / \Sigma\Sigma \langle I(h) \rangle$
where G_i is the inverse scale factor, $I(h)$ the diffraction intensity of the symmetry equivalent reflections, and $\langle I(h) \rangle$ the mean value of $I(h)$.

¶ Rejection criteria are $C_R = 0.3\ I_{mean} + 0.1 \langle I(h) \rangle$ for the reflections measured more than twice and $C_R = 3\ (0.3\ I_{mean} + 0.1 \langle I(h) \rangle)$ for the reflections with only two observations, where I_{mean} is the mean of all $I(h)$ values. The reflections with the differences $D_i = \mid \langle I(h) \rangle - I_i(h) / G_i \mid > C_R$ are rejected (Rossmann *et al.*, 1979).

Table 7.2.2

Summary of data collection of 2T2M6T-crystal and statistics

X-ray Source	Cu-Kα
X-ray generator	Rigaku RU200
Focus size	0.3 × 3mm
X-ray power	40kV, 100mA
Monochromatization	Ni & mirror
IP size	200×200mm
Pixel size	105μm
No. of crystal used	1
φ(spindle) - axis	approx. c axis
Crystal - to - IP distance	86mm
Resolution limit	2.0Å
Oscillation range per frame	1.5Å
No. of frames	21
Total oscillation range	31.5°
Exposure time	40min / frame
No. of observed reflections	
full	30,349
partial	11,767
total	42,116
No. of independent reflections	24,762
Completeness †	62.3%
R - merge ‡	
full reflections	5.01%
partial reflections	6.69%
total reflections	5.48%
No. of rejected reflections ¶	
full reflections	0
partial reflections	11

† Considering the blind region.

‡ R-merge = $\Sigma\Sigma \mid I_i(h) / G_i - \langle I(h) \rangle \mid / \Sigma\Sigma \langle I(h) \rangle$
where G_i is the inverse scale factor, $I(h)$ the diffraction intensity of the symmetry equivalent reflections, and $\langle I(h) \rangle$ the mean value of $I(h)$.

¶ Rejection criteria are $C_R = 0.3\ I_{mean} + 0.1 \langle I(h) \rangle$ for the reflections measured more than twice and $C_R = 3\ (0.3\ I_{mean} + 0.1 \langle I(h) \rangle)$ for the reflections with only two observations, where I_{mean} is the mean of all $I(h)$ values. The reflections with the differences $D_i = \mid \langle I(h) \rangle - I_i(h) / G_i \mid > C_R$ are rejected (Rossmann *et al.*, 1979).

Table 7.2.3

Summary of data collection of I93L-crystal and statistics

X-ray Source	Cu-Kα
X-ray generator	Rigaku RU200
Focus size	0.3 × 3mm
X-ray power	40kV, 100mA
Monochromatization	Ni & mirror
IP size	200×200mm
Pixel size	105μm
No. of crystal used	1
φ(spindle) - axis	approx. c axis
Crystal - to - IP distance	86mm
Resolution limit	1.9Å
Oscillation range per frame	1.5Å
No. of frames	21
Total oscillation range	31.5°
Exposure time	40min / frame
No. of observed reflections	
full	28,055
partial	10,551
total	38,606
No. of independent reflections	24,155
Completeness †	55.1%
R - merge ‡	
full reflections	4.79%
partial reflections	5.56%
total reflections	5.00%
No. of rejected reflections ¶	
full reflections	0
partial reflections	0

† Considering the blind region.

‡ R-merge = $\Sigma\Sigma \mid I_i(h) / G_i - \langle I(h) \rangle \mid / \Sigma\Sigma \langle I(h) \rangle$
where G_i is the inverse scale factor, $I(h)$ the diffraction intensity of the symmetry equivalent reflections, and $\langle I(h) \rangle$ the mean value of $I(h)$.

¶ Rejection criteria are $C_R = 0.3\ I_{mean} + 0.1 \langle I(h) \rangle$ for the reflections measured more than twice and $C_R = 3\ (0.3\ I_{mean} + 0.1 \langle I(h) \rangle)$ for the reflections with only two observations, where I_{mean} is the mean of all $I(h)$ values. The reflections with the differences $D_i = \mid \langle I(h) \rangle - I_i(h) / G_i \mid > C_R$ are rejected (Rossmann *et al.*, 1979).

26.0%, 25.4% and 23.9% for 4M6T, 2T2M6T and I93L, respectively. The intensity variance of I93L against 2T2M6T was 6.20%, which is significant difference compared with R-merge values.

7-3 Model building and refinement

As crystals of chimeric enzymes were isomorphous with that of Tt-IPMDH, difference Fourier and 'omit' maps were directly calculated. But because of the lack and poorness of high angle data above 3Å resolution, 4M6T gave an indistinct electron density map. Therefore it is impossible to make up the molecular model of 4M6T. The molecular models of 2T2M6T and I93L were constructed on FRODO.

2T2M6T

In order to emphasize the structural difference between 2T2M6T and Tt-IPMDH on the electron density, the 'omit' maps with coefficients $|F_{2T2M6T}| - |F_{calc}|$ and $2|F_{2T2M6T}| - |F_{calc}|$ were calculated by using phases of Tt-IPMDH structure. In the first map, residues from 79 to 100 of Tt-IPMDH were omitted from the calculation of F_{calc}s and phases, and in the second map, residues from 105 to 130 were omitted. Fixed solvent molecules were also eliminated from the calculation of the maps. These two omit maps were fairly clear and the structure model was successfully constructed. The model was refined by PROLSQ and the statistics are given in Table7.3.1. The final R-factor for the model was reduced to 19.6%. The overall B-

Table 7.3.1 Summary of least-squares parameters and deviations of 2T2M6T and I93L chimera at 2.2Å resolution

	Target	r.m.s deviations	
		2T2M6T	I93L
Bonding distances (Å)			
1-2 bond	0.020	0.014	0.016
1-3 angle	0.030	0.035	0.039
1-4 planar	0.050	0.048	0.053
Planar groups (Å)	0.020	0.011	0.012
Chiral volumes (Å^3)	0.150	0.162	0.172
Non-bonded contacts (Å)			
Single torsion	0.500	0.213	0.214
Multiple torsion	0.500	0.275	0.280
Possible hydrogen bond	0.500	0.295	0.270
Torsion angles (deg.)			
Planar	3.0	2.5	2.4
Staggered	15.0	23.3	23.3
Orthonormal	20.0	27.5	27.5
Thermal factors (Å^2)			
Main-chain bond	1.000	0.454	0.495
Main-chain angle	1.500	0.787	0.847
Side-chain bond	1.500	0.869	0.953
Side-chain angle	2.000	1.378	1.498

Table 7.3.2 Dependency of the R-factors on resolution

(a) 2T2M6T

Resolution (Å)	No. of reflections	*R* - factor	*R* - factor (accumulated)
5.00 ~ 4.00	2,225	0.139	0.139
4.00 ~ 3.30	3,370	0.172	0.156
3.30 ~ 2.90	3,151	0.209	0.170
2.90 ~ 2.63	2,812	0.246	0.181
2.63 ~ 2.45	2,052	0.260	0.187
2.45 ~ 2.32	1,576	0.274	0.192
2.32 ~ 2.20	1,426	0.291	0.196
5.00 ~ 2.20	16,612	-	0.196

(b) I93L

Resolution (Å)	No. of reflections	*R* - factor	*R* - factor (accumulated)
5.00 ~ 4.00	2,199	0.133	0.133
4.00 ~ 3.30	3,278	0.169	0.152
3.30 ~ 2.90	2,950	0.209	0.166
2.90 ~ 2.63	2,489	0.248	0.178
2.63 ~ 2.45	1,702	0.263	0.185
2.45 ~ 2.32	1,270	0.285	0.190
2.32 ~ 2.20	1,058	0.324	0.195
5.00 ~ 2.20	14,946	-	0.195

factor derived from Willson plot (Wilson, 1949) was 28Å2 and was assigned to initial individual B-factors.

I93L

After the refinement of 2T2M6T, model building of I93L was started. The omit maps with coefficients $|F_{I93L}| - |F_{calc}|$ and $2|F_{I93L}| - |F_{calc}|$ were calculated. F_{calc}s and phases were derived from the refined model of 2T2M6T without the residues from 92 to 94 and fixed solvent molecules. The electron density of leucine 93 appeared quite clear and some other fixed solvent molecules were also clearly shown. Then the molecular model was refined by PROLSQ. An initial R-factor was 42.7% and after nine cycles it resulted in 19.5% for 2.2Å resolution. The refinement statistics are shown in Table7.3.1.

7-4 Structure description and comparison with Tt-IPMDH

2T2M6T

The final model of 2T2M6T contains 2607 protein atoms and 62 solvent molecules. The model was fitted to the Tt-IPMDH model with least squares method using the mainchain atoms. Table7.4.1 gives the mean and r.m.s. deviation between the 2T2M6T and Tt-IPMDH. In the calculation, the common side chain atoms were included. In Figure7.4.1, the structure of 2T2M6T was overwritten on that of Tt-IPMDH. As shown in Figure7.4.1, the two regions in the mainchain structure of 2T2M6T is different from the Tt-IPMDH structure. One is the region of the residues from 77 to 93 which

Table 7.4.1 differences of the coordinate between native and 2T2M6T

	Main chain atom	Side chain atom¶	All protein atom¶
mean distance	0.29Å	0.43Å	0.35Å
r.m.s.deviation	0.40Å	0.66Å	0.53Å
0.5Å≤ <1.0Å	99	215	314
1.0Å≤ <1.5Å	24	28	52
1.5Å≤ <2.0Å	6	12	18
2.0Å≤	8	24	32

The coordinate of 2T2M6T were fitted to that of Tt-IPMDH by least square fitting program MFIT using main chain atoms.

¶ The common side chain atoms are included. e.g. Cβ atom of 106 (valine in 2T2M6T, alanine in Tt-IPMDH) is included in the calculation.

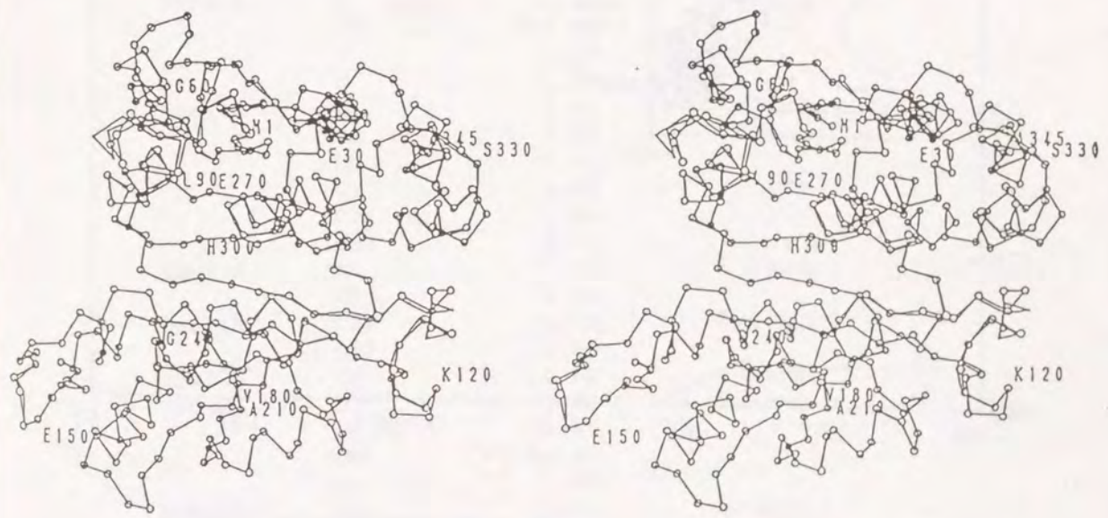


Figure 7.4.1 Stereo drawing of the Cα backbone of a subunit of 2T2M6T overwritten on the that of Tt-IPMDH
Ball and stick model represents 2T2M6T and thin stick model Tt-IPMDH

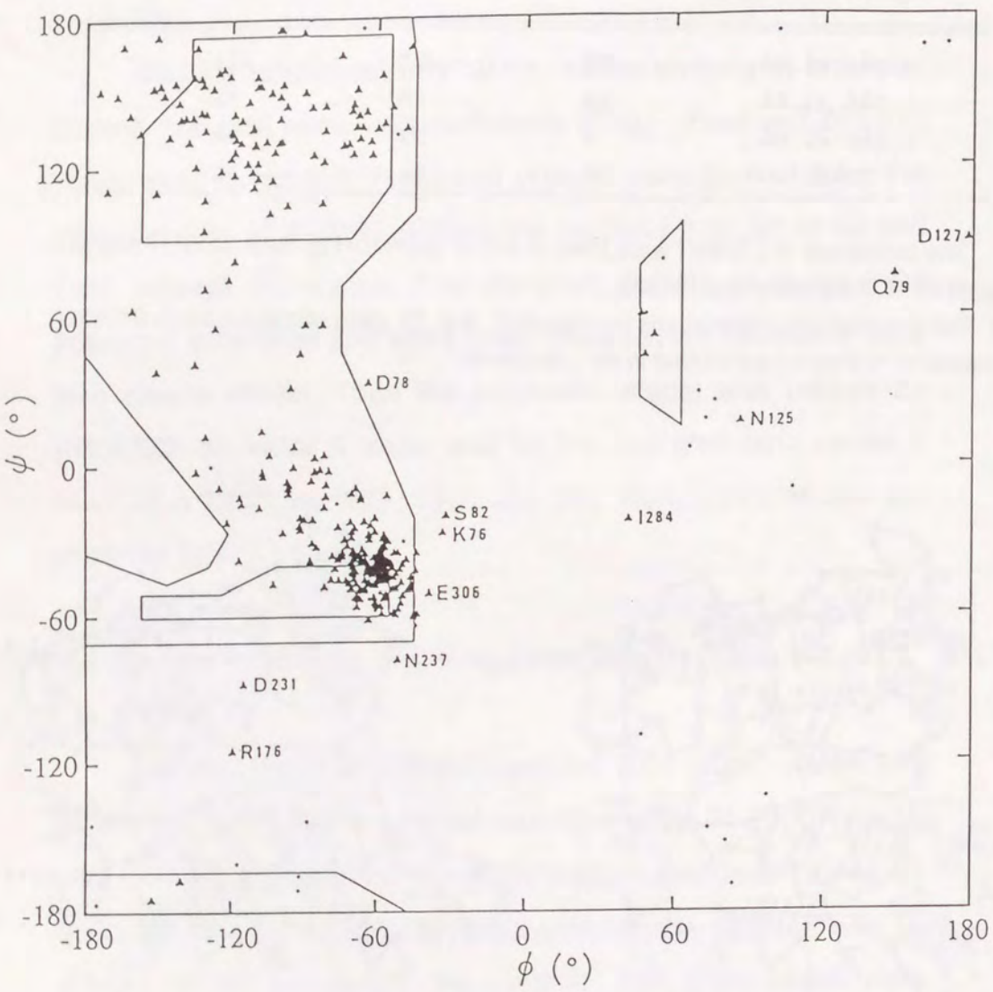


Figure 7.4.2 Ramachandran plot of main-chain dihedral angles of 2T2M6T

Non-glycine residues are shown with ▲ and glycine residues with
The preferred regions of Ramakrishana & Ramachandran are indicated.

Table 7.4.2 List of atoms whose coordinates are differnt more than 1.5Å between native and 2T2M6T structure

residue	atom	distance (Å)	comment
Leu 23	Cδ1	2.64	middle of α-a
	Cδ2	2.47	
Arg 58	Nη1	2.18	middle of α-c
	Nη2	1.86	
Ser 82	C	1.61	C-d loop
	O	2.08	
Glu 83	N	1.54	C-d loop
	Cα	2.05	
	Cβ	2.57	
	Cγ	4.60	
	Cδ	6.31	
Leu 84	Cβ	1.91	C-d loop
Leu 91	Cδ1	2.32	middle of α-d
	Cδ2	2.34	
Arg 94	Cζ	2.12	C-end of α-d
	Nη1	3.57	
	Nη2	3.51	
Glu110	Cα	1.51	C-end of β-F
	Cβ	1.63	
	C	1.63	
	O	3.13	
Ser111	N	2.37	F-G loop
	Cα	2.48	
	C	2.09	
	O	1.65	
Leu112	N	2.36	F-G loop
	Cα	2.02	
	Cβ	2.43	
	Cγ	2.23	
	Cδ1	2.69	
	Cδ2	1.76	
Ser113	N	1.74	F-G loop
	Cβ	1.63	
Leu118	Cδ1	2.79	F-G loop
	Cδ2	1.80	
Lys120	Cγ	1.92	F-G loop
Glu121	Cγ	2.90	F-G loop
	Cδ	3.60	
	Oε1	2.43	
	Oε2	5.83	
Arg176	Nη2	1.90	e-I loop
Leu257	Cγ	1.52	
	Cδ1	2.34	
	Cδ2	1.97	
Ile279	Cγ2	2.08	D-i loop
Lys310	Cε	3.34	middle of α-j
	Nζ	4.21	
Leu320	Cδ1	2.12	C-end of α-j
	Cδ2	1.61	

Table 7.4.3 Fixed solvent molecules and their protein hydrogen bonds of 2T2M6T chimera

Solvent	B-factor	Protein Hydrogen bonds	Solvent	B-factor	Protein Hydrogen bonds
346	65.39	Val 3 O	362	42.34	Glu 161 Oε1
347	38.52	Phe 39 O	363	26.63	Asp 127 O
		Phe 41 O			Pro 227 O
		Gly 8 N			Asp 127 N
348	31.93	Gly 8 O	364	39.37	
349S	80.91		365	31.27	Glu 133 Oε1
O1	80.38				Glu 161 O
O2	80.12	Arg167' Nη1	366	29.99	Glu 155 Oε1,Oε2
O3	79.97	Arg167' Nη2			Ile 238 N
O4	80.26		367	49.95	Arg 156 O
350	35.21	Glu 14 O			Glu 161 Oε1,Oε2
351	28.36	Glu 14 Oε1,Oε2			Gly 137 N
		Ile 279 O	368	33.03	Ile 138 O
		Lys 282 N			Asn 153 Oδ1
		Gly 283 N			Thr 154 N
		Ala 285 N	369	29.14	Glu 163 Oε1
352	41.87	Lys 21 O			Arg 167 Nη2
353	28.62	Asp 9 Oδ1	370	31.20	Glu 171 Oε2
354	43.03	Gly 10 O	371	38.41	Glu 171 Oε1
		Glu 14 Oε1			Glu 171 O
		Glu 14 N	372	38.94	Ala 172 O
355	32.74	Gly 74 O			Glu 299 Oε2
		Asp 78 Oδ2			His 300 Nε2
		Glu 87 Oε1	373	28.34	His 179 Nε2
		Gly 74 N			His 179 O
356	46.88				Val 232 O
357	42.82	Ala 316 O			Asp 231 N
358	65.62	Val 126 O	374	36.24	Ser 182 Oγ
359	42.65	Leu 112 O	375	36.37	Val 183 O
		Pro 251 O			Gly 236 N
360	35.60	Tyr 122 O	376	42.37	Leu 292 O
		Ala 228 N			Glu 312 Oε1
361	51.77	Gly 203 O	377	45.04	Glu 299 Oε1

Table 7.4.3 continued

Solvent	B-factor	Protein Hydrogen bonds	Solvent	B-factor	Protein Hydrogen bonds
378	37.94	Gly 255 O	394	31.47	Pro 258 O
		Asn 286 O	395	44.70	Glu 155 Oε1
		Ala 289 N	396	32.58	Gly 10 N
379	55.51	Pro 105 O	397	58.06	Ser 82 Oγ
380	36.91	His 273 Nε2	398	40.30	
		Asn 286 O	399	43.12	
381	41.18		400	57.14	His 213 O
382	31.79	Lys 282 Nζ			Gln 214 N
383	27.08		401	50.74	
384	54.53	Ile 284 O	402	57.82	Glu 306 O
		Ala 285 O	403	57.17	Lys 310 O
385	51.25	Glu 334 N			His 343 O
386	38.34	Lys 185 N			Lys 310 Nζ
387	52.95		404	29.85	Val 168 O
388	59.65	Pro 277 O	405	55.68	Pro 323 O
		Ala 280 O			Ser 330 N
389	35.91				Ala 331 N
390	37.41		406	39.12	Glu 51 Oε1
391	61.82	Arg 264 Nη1			Phe 53 O
392	69.58	Glu 201 Oε2			Phe 53 N
		Lys 197 Nζ	407	41.36	Leu 134 O
393	48.28	Arg 196 Nη1			Arg 132 Nη2

correspond to C-d loop and the first half of helix d of Tt-IPMDH. Ramachandran plot (Figure7.4.2) shows that Pro75, Lys76, Asp78, Gln79 and Ser82 are deviate from the allowable region. So there are some stress in the structure of this region. The other is the residues from 105 to 128 which correspond to the second half of strand F, F-G loop and the first half of strand G of Tt-IPMDH. F-G loop contributes the dimer interaction. The side chains of Leu118 and Lys119, which play a main role of the dimer interaction, move more than 1.5Å from those of Tt-IPMDH. Therefore the dimer interaction form is slightly changed. Asn125 and Asp127 are outside of the allowable region of Ramachandran plot. This region also has structural stress. The atoms which deviate more than 1.5Å from those of Tt-IPMDH are listed in Table7.4.2. Fixed solvent molecules are also different from Tt-IPMDH and listed in Table7.4.3.

I93L

In spite of the only different residue, I93L mutant is much more thermostable than 2T2M6T. The Ramachandran plots (Figure7.4.2 and 7.4.4) indicate 2T2M6T has much strain than I93L. The structure of I93L was superimposed on that of 2T2M6T with least squares method. In the calculation, all the atoms except for the side chain atoms of the 93 residue were involved. Fig 7.4.3 shows main chains of 2T2M6T and I93L superimposed. The main chain atoms of Asp78, Gln79 and Glu83 deviate more than 0.5Å from each other. But the positional differences of the other main chain atoms are less than 0.5Å. This shows that substitution of the residue 93 does not

Table 7.4.4 differences of the coordinate between 2T2M6T and I93L

	Main chain atom	Side chain atom¶	All protein atom¶
mean distance	0.16Å	0.09Å	0.25Å
r.m.s.deviation	0.48Å	0.12Å	0.68Å
<0.5Å	1372	1205	2579
0.5Å≤ <1.0Å	6	19	25
1.0Å≤ <1.5Å	1	4	5
1.5Å≤ <2.0Å	1	1	2
2.0Å≤	0	26	26

The coordinate of I93L were fitted to that of 2T2M6T by least square fitting program MFIT using all protein atoms except for side chains of the residue 93.
¶ The common side chain atoms of the residue 93 are included.

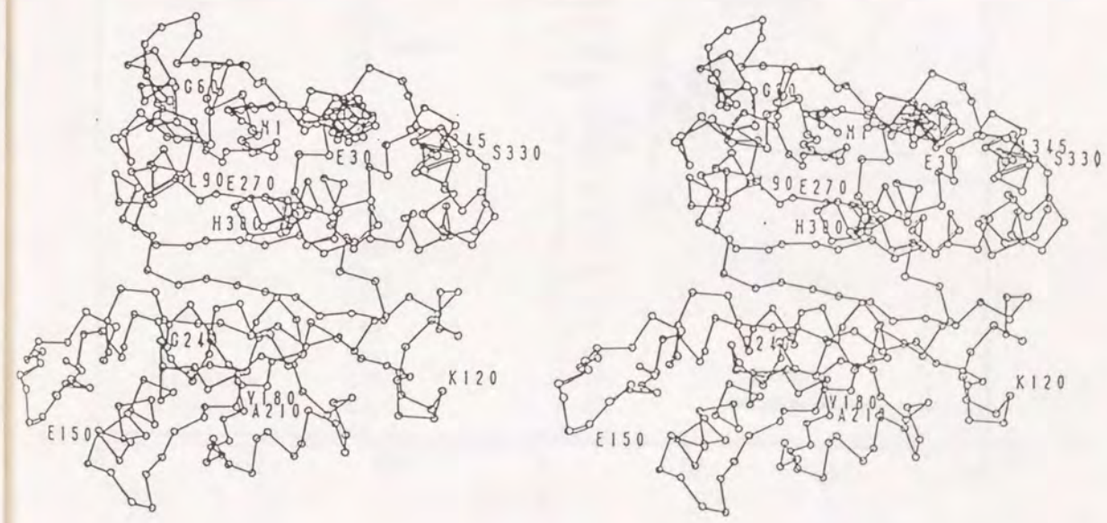


Figure 7.4.3 Stereo drawing of the Cα backbone of a subunit of I93L overwritten on the that of 2T2M6T
Ball and stick model represents I93L and thin stick model 2T2M6T.

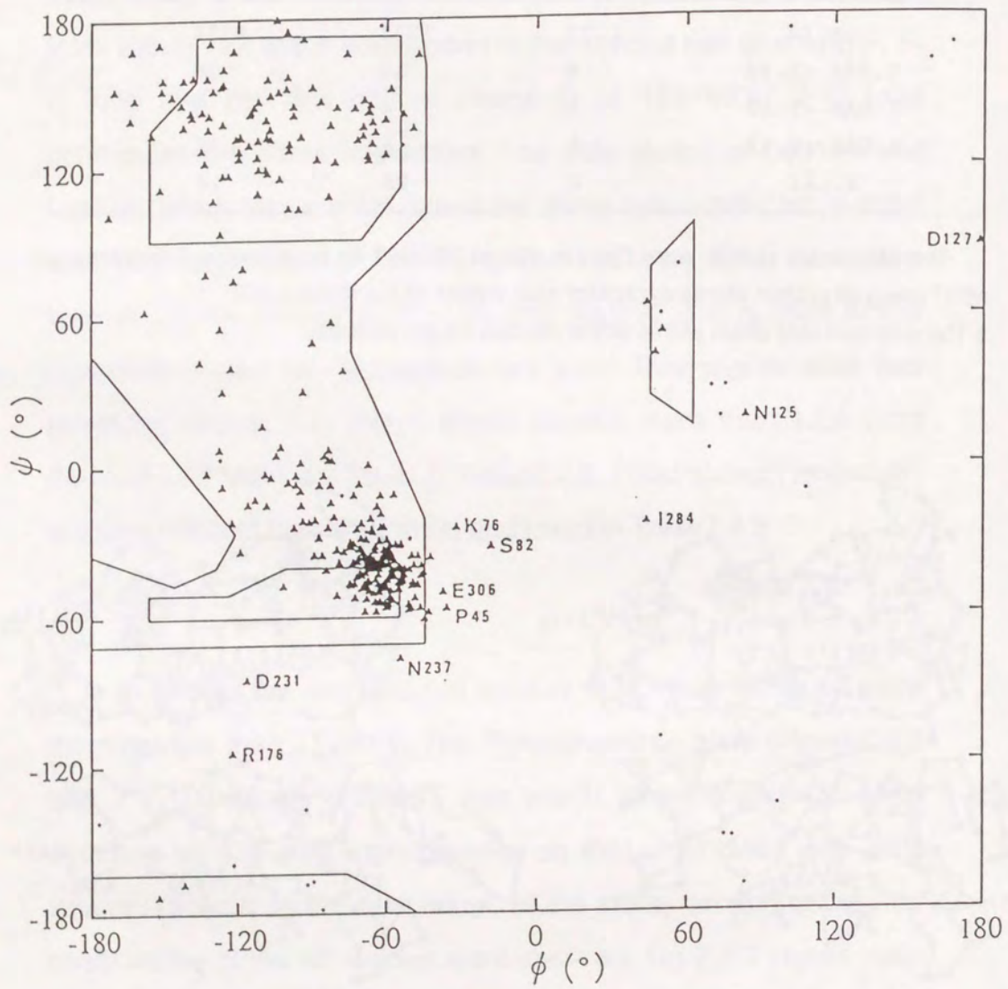


Figure 7.4.4 Ramachandran plot of main-chain dihedral angles of I93L
Non-glycine residues are shown with ▲ and glycine residues with
The preferred regions of Ramakrishana & Ramachandran are indicated.

Table 7.4.5 List of atoms whose coordinates are differnt more than 1.5Å between 2T2M6T and I93L structure

residue	atom	distance (Å)	comment
Glu 55	Cγ	2.70	C-end of α-c
	Cδ	4.34	
	Oε1	6.29	
	Oε2	4.46	
Lys 59	Nζ	2.16	middle of α-c
Gln 79	Cα	1.62	C-d loop
	Cβ	3.25	
	Cγ	4.16	
	Cδ	6.08	
	Oε1	7.42	
	Nε2	6.28	
Glu 83	Cβ	1.52	C-d loop
	Cγ	3.98	
	Cδ	5.26	
	Oε1	6.70	
	Oε2	5.43	
Arg 94	Cγ	2.18	C-end of α-d
	Cδ	2.24	
	Cζ	2.61	
	Nη1	3.65	
	Nη2	4.71	
Glu121	Cγ	2.66	F-G loop
	Cδ	3.81	
	Oε1	2.37	
	Oε2	5.86	
Leu254	Cδ1	3.18	h-E loop
Lys310	Cε	2.91	middle of α-j
	Nζ	3.80	

Table 7.4.6 Fixed solvent molecules and their protein hydrogen bonds of I93L chimera

Solvent	B-factor	Protein Hydrogen bonds	Solvent	B-factor	Protein Hydrogen bonds
346	56.93	Glu 62 O	361	49.40	Gly 203 O
		Ala 64 O			Val 209 O
		Thr 266 Oγ1	362	43.24	Glu 161 Oε1
347	25.21	Phe 39 O	363	37.10	Asp 127 O
		Phe 41 O			Pro 227 O
		Gly 8 N			Ala 228 O
348	33.30	Gly 8 O			Asp 127 N
349S	87.92		364	38.04	
O1	87.25		365	57.55	
O2	87.11	Arg167' Nη1	366	32.11	Glu 155 Oε1,Oε2
O3	86.86	Arg167' Nη2			Ile 238 N
O4	87.21		367	44.37	Glu 17 Oε1
350	43.16	Gly 8 O	368	31.07	Ile 138 O
351	35.68	Glu 14 Oε1,Oε2			Asn 153 Oδ1
		Ile 279 O			Thr 154 N
		Lys 282 N	369	33.36	Glu 163 Oε1,Oε2
		Gly 283 N			Arg 167 Nη2
352	46.32	Lys 21 O	370	31.58	Glu 171 Oε2
353	24.89	Asp 9 Oδ1	371	23.96	Glu 171 Oε1
354	40.48	Gly 10 O			Glu 171 O
		Ile 11 O	372	38.61	Ala 172 O
		Glu 14 Oε1			Glu 299 Oε2
		Glu 14 N			His 300 Nε2
355	29.44	Gly 74 O	373	22.97	His 179 Nε2
		Asp 78 Oδ2			His 179 O
		Glu 87 Oε1			Val 232 O
		Gly 74 N			Asp 231 N
356	31.59		374	31.96	Ser 182 Oγ
357	47.64	Ala 316 O	375	28.63	Val 183 O
		Ser 111 N			Gly 236 N
358	49.21	Val 126 O	376	51.67	Leu 292 O
359	47.35	Phe 109 N			Glu 312 Oε1
360	53.80	Tyr 122 O	377	48.92	
		Ala 228 N			

Table 7.4.6 continued

Solvent	B-factor	Protein Hydrogen bonds	Solvent	B-factor	Protein Hydrogen bonds
378	23.29	Gly 255 O	394	57.84	Asp 78 O
		Asn 286 O	395	43.00	
		Ala 289 N	396	34.93	Gly 10 N
379	51.82	Pro 105 O	397	49.57	Gly 255 O
380	32.24	His 273 Nε2			Asn 286 O
		Asn 286 O			His 273 Nε2
		Ala 290 N	398	44.64	
381	39.88		399	35.10	Glu 334 Oε1
382	38.52		400	60.12	
383	25.27		401	49.03	Met 146 O
384	50.61	Ile 284 O	402	62.67	Glu 306 O
		Ala 285 O	403	62.15	Lys 310 O
385	48.61		404	30.68	
386	28.33	Lys 185 N	405	41.72	Pro 323 O
387	46.83				Ala 331 N
388	34.76	His 300 O	406	76.94	
389	45.60		407	49.67	Asp 127 Oδ2
390	47.64	Arg 390 Nη2			Phe 128 O
391	63.52				Arg 176 Nη2
392	50.63	Glu 201 Oε1	408	23.03	Asp 245 Oδ1,Oδ2
		Lys 197 Nζ	409	40.17	Gly 252 O
393	52.77	Arg 196 Nη1,Nη2			Ser 253 Oγ

largely affect the overall structure at room temperature, probably the effect of the substitution on the structure may appear at high temperatures.

B-factor of chimeric enzymes

The distribution of B-factor as a function of the residue number is shown in Figure7.4.5 and Figure7.4.6. The distribution patterns of 2T2M6T and I93L are very similar to each other, but these differ from Tt-IPMDH. The B-factors of the residues corresponding to F-G loop in 2T2M6T and I93L have considerably high compared to those of Tt-IPMDH.

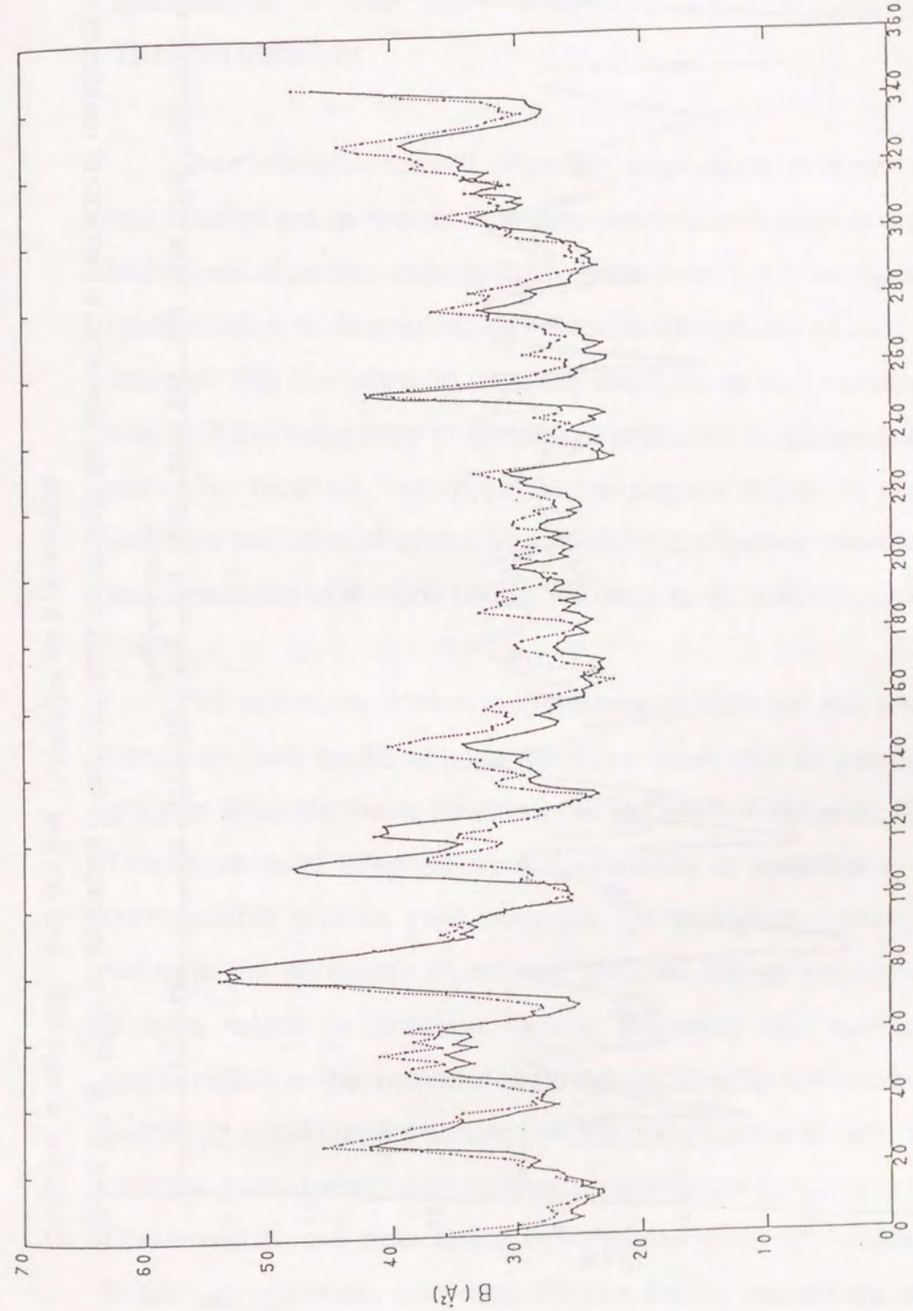


Figure 7.4.5 Plots of averaged B-factors of 2T2M6T for mainchain atoms as a function of residue number Solid line shows 2T2M6T, and dot line Tt-IPMDH as a reference.

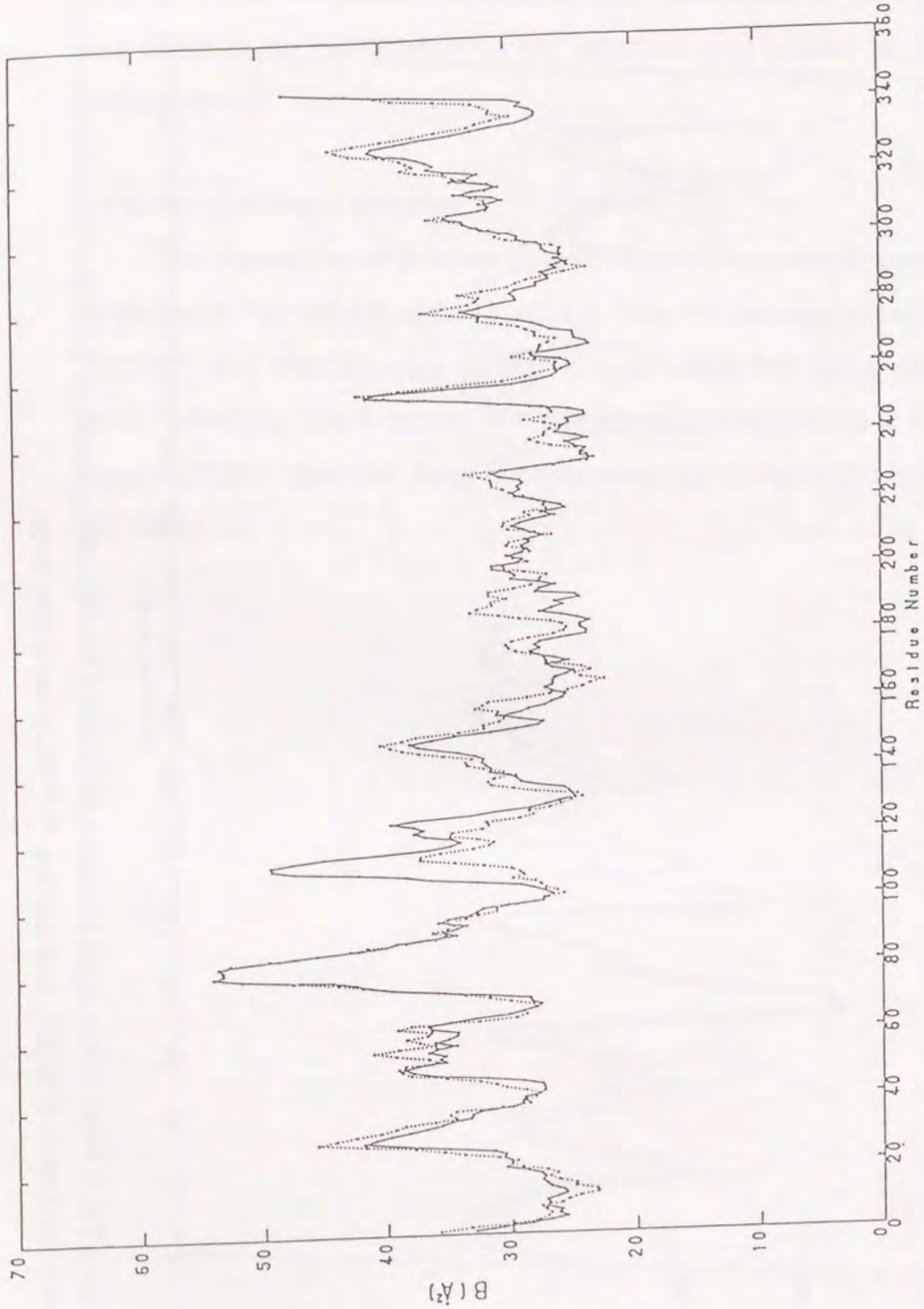


Figure 7.4.6 Plots of averaged B-factorsof I93L for mainchain atoms as a function of residue number
Solid line shows I93L, and dot line Tt-IPMDH as a reference.

Chapter-8
Thermostability

Dramatic development of protein engineering in recent years has enabled one to change individual amino acid residue at will, and techniques of protein engineering may be used to assess the role of each residue in thermostability. Results from these studies have revealed that the following covalent reactions at high temperature may limit thermostability of protein: deamidation of asparagine and glutamine residues, hydrolysis of the peptide bonds at aspartic residues, oxidation of cysteine residues, thiol-disulfide interchange, and destruction of disulfide bonds (Klibanov et al., 1987; Volkin et al., 1987).

The extensive protein engineering studies on the enzyme lysozyme from bacteriophage T4 have been carried out by the group of Brian Matthews (Matthews et al., 1987; Alber et al., 1987). They have used three different approaches to engineer a more thermostable protein than wild-type T4 lysozyme, namely, (1) reducing the difference in entropy between folded and unfolded protein, which in practice means reducing the number of conformation in the unfolded state by introducing novel disulfide bonds, (2) stabilizing the helices by enhancing dipoles of helices, (3) increasing the number of hydrophobic interactions in the inner core. These results are now being used to increase the stability of industrially important enzymes. Proline theory (Matthews et al.,

1987) for increasing protein thermostability also has been proposed. This theory suggests that a protein would be thermostabilized by increasing the frequency of proline occurrence at β-turns (Suzuki et al., 1987) and the total number of hydrophobic residues present in the protein.

Now a days, it has been known that various factors contribute to stabilize a protein structure; disulfide bonds (Pantliano et al., 1987; Perry & Wetzel, 1984,1986), shorter loops, proline residues (Matthews et al., 1987), hydrophilic inter-or-intramolecular interactions such as hydrogen bonds, salt bridges (Perutz, 1978; Alber et al., 1987) and hydrophobic inter-or-intramolecular interactions (Yutani et al., 1977,1987).

As described in section 4-3, the amino acid sequences of IPMDHs are highly conserved among various kinds of organisms, such as mesophiles, moderate thermophiles and extreme thermophiles. From the comparative study of these IPMDHs, some special residues and regions which were conserved only in thermophiles were found. They may contribute to the thermostability of thermophilic IPMDHs through the factors described above. But the primary structure itself can not explain the high thermostability of the enzyme. To understand the thermostability, it is necessary to interpret the meaning of the primary structural feature of Tt-IPMDH from the standpoint of three-dimensional structure. Furthermore the results from mutational analysis and the property of chimeric enzymes must be also analyzed in the term of three-dimensional structure.

In this chapter, we discuss the factors contribute to the thermostability of Tt-IPMDH based on the three dimensional structure.

8-1 Disulfide bonds

There are some examples of stabilizing proteins by formation of a disulfide bond. T4-lysozyme was stabilized by artificial introduction of the disulfide bond (Pantliano et al., 1987; Perry & Wetzel, 1984,1986). As the disulfide bond reduces the entropies of unfolded state, the protein inducting the disulfide bond is generally stabilized. In the case of IPMDH from *T.thermophilus*, it does not contain any cysteine residues. Tt-IPMDH can not stabilize by this way.

On the other hand, IPMDHs from Yl, Cu, Sc and Ec contain more than two cysteine residues. To make sure of possibility of the disulfide bond, these cysteine residues are assigned on the structure from Tt-IPMDH. But it was impossible to find such a cysteine pair making the disulfide bond.

8-2 Shorter loop

From the comparative study, it has been said that loop regions of thermophiles are tend to be shorter than mesophiles, but such a tendency could not be found in IPMDH (Figure4.3.1).

Table 8.3.1 Proline residues in Tt-IPMDH

Residue	conserved?	comment	
7	C	C end of β-B	
13	C	N end of α-a	
40	T	C end of β-B	
52	C	b-c loop	
54	C	b-c loop	
56	N	N end of α-c	
75	C	C-d loop	
81	C	C-d loop	
86	C	N end of α-d	
105	C	middle of β-F	
110	T	C end of β-F	
117	C	F-G loop	
143	C	N end of β-K	cis proline
160	C	N end of α-e	
207	C	N end of β-J	
227	C	g-H loop	
251	C	C end of α-h	
258	C	N end of β-E	
267	C	N end of β-D	
271	C	C end of β-D	
277	C	D-i loop	
287	C	N end of α-i	
323	T	j-k loop	**pro**-pro-pro
324	T	j-k loop	pro-**pro**-pro
325	T	j-k loop	**pro-pro-pro**

C : The proline residue conserved between IPMDHs from mesophiles and thermophiles.

T : The proline residue conserved only between IPMDHs from thermophiles.

N : The proline residue peculiar to Tt-IPMDH.

8-3 Proline residue

It has been suggested that proline residues may contribute to the thermostability of enzymes through their entropy effect (Matthews et al., 1987). As a proline residue is restrained its ϕ angle, the conformation of an unfolded chain is restricted. Hence, the entropy of unfolded state is reduced. There are 25 proline residues in IPMDH from *T.thermophilus*, and six proline residues of these, Pro40, Pro56, Pro110, Pro323, Pro324 and Pro325, are peculiar to *T.thermophilous* (Table8.3.1). Pro40 and Pro110 are conserved in *T.aquaticus* which is also extreme thermophile. Pro40 exists at the end of the β-strand A, and the main chain is bent sharply at this residue. Pro110 is also at the end of the β-strand F. The geometry of residues 323 ~ 325 were described previously (section 3-2). Most of the proline residues in Tt-IPMDH are conserved in mesophiles, and need for keeping its structure, but not for thermostability. But Pro40 and Pro323 ~ Pro325 may contribute to the thermostability of IPMDH. For Pro110, the point mutational study of chimeric IPMDH has shown that it does not contribute to the thermostability.

8-4 Hydrogen bonds and electrostatic interactions

It is well known that hydrophilic interactions, such as hydrogen bonds, play an important role for the thermostability of a protein (Perutz, 1978). The investigation of T4 lysozyme mutants has shown that a hydrogen bond network of Asp159, Thr157 and Thr155 is important for the protein stability (Alber et al., 1987). The investigation of neutral protease, also has shown that the addition of

one hydrogen bond stabilizes the enzyme (Yabuki et al, 1988). We searched hydrogen bonds and electrostatic interactions which may contribute to the thermostability of IPMDH. Table8.4.1 gives the hydrophilic interactions which are conserved only in extreme thermophiles. Tyr36, Arg58, Ser71, Gln97, Thr288 and Thr322 are peculiar to extreme thermophiles, and may contribute to thermostability. As chimera enzymes, 4M6T and 2T2M6T, however, show equivalent thermostability, the hydrophilic interactions by the residues from 1 to 78 give no contribution for the thermostability. From the comparison of Bs, the residues which has a possibility of contribution for thermostability are listed in Table8.4.2.

8-5 Hydrophobic interactions

We picked up common residues conserved in thermophiles but not in mesophiles. By putting these residues on the structure, some residues which may contribute to thermostabilization by hydrophobic interactions were found. Those are Val191, Phe194, Val224, Ile238, Leu246 and Val249. All of them contribute to the subunit-subunit interactions. Particularly, Leu246, which lies at the center of hydrophobic core made by four α-helices and plays an important role on intersubunit hydrophobic interactions, is replaced by glutamic acid in mesophiles. Figure8.5.1 shows the replacement of Leu246 to glutamic acid. This replacement seems to be unfavorable for the dimer interaction. In fact, the multi-mutant replaced Leu246 and Val249 by Glu and Met became unstable and

Table 8.4.1 Hydrophilic interactions conserved in extreme thermophiles, Tt- and Ta-, not in Bs-IPMDH

Side Chain (donner)	Main Chain (acceptor)	Comment
Thr		
Thr 16 Oγ1	Gly 12 O	αa-αa
Thr 198 Oγ1	Phe 194 O	αf-αf
Thr 266 Oγ1	Val 61 O	EDloop-αc
	Val 64 O	-αc
Thr 322 Oγ1	Ala 318 O	αj-αj
Thr 333 Oγ1	Gly 283 O	αk-Diloop
Ser		
Ser 71 Oγ	Asp 9 O	Cdloop-Baloop
Ser 96 Oγ	Leu 93 O	αd-αd
Arg		
Arg 177 Nη2	Asp 127 O	eIloop-βG
Arg 177 Nη2	Phe 230 O	eIloop-gHloop
Gln Asn		
Gln 97 Nε2	Ser 96 O	dFloop-dFloop

Side Chain (donner)	Main Chain (acceptor)	Comment
Leu 34 N	Asp 27 Oδ1	aAloop-αa
Gly 111 N	Glu 113 Oε1	FGloop-FGloop
	Glu 113 Oε2	-FGloop
Ile 138 N	Glu 155 Oε2	GKloop-βL
Thr 266 N	Gln 97 Oε1	EDloop-dFloop

Side Chain (donner)	Side Chain (acceptor)	Comment
Arg 58 Nη1	Glu 55 Oε1	αc-bcloop
	Glu 55 Oε2	
Gln 97 Nε2	Ser 96 Oγ	dFloop-dFloop
Ser 158 Oγ	Glu 161 Oε1	bL-αe
Lys 159 Nζ	Glu 163 Oε2	αe-αe
	Glu 201 Oε1	-αf
Arg 167 Nη1	Tyr 206 Oη	αe-αf
Arg 174 Nη1	Asp 208 Oδ1	αe-βJ
	Asp 208 Oδ2	
Arg 174 Nη2	Glu 171 Oε1	αe-αe
Arg 177 Nη1	Asp 231 Oδ1	eIloop-gHloop
Asn 286 Nδ2	Thr 288 Oγ1	Diloop-αi

Table 8.4.2 Possible hydrophilic interactions contributing to thermostability

Donner	Acceptor	Comment
Thr 88 Oγ1	Arg 82 O	Cdloop-Cdloop
Thr 198 Oγ1	Phe 194 O	αf-αf
Thr 266 Oγ1	Val 61 O	EDloop-αc
	Val 64 O	-αc
Thr 322 Oγ1	Ala 318 O	αj-αj
Thr 333 Oγ1	Gly 283 O	αk-Diloop
Ser 96 Oγ	Leu 93 O	αd-αd
Gln 97 Nε2	Ser 96 O	dFloop-dFloop
Gly 111 N	Glu 113 Oε1	FGloop-FGloop
	Glu 113 Oε2	-FGloop
Ile 138 N	Glu 155 Oε2	GKloop-βL
Thr 266 N	Gln 97 Oε1	EDloop-dFloop

Donner	Acceptor	Comment
Gln 97 Nε2	Ser 96 Oγ	dFloop-dFloop
Ser 158 Oγ	Glu 161 Oε1	bL-αe
Arg 167 Nη1	Tyr 206 Oη	αe-αf
Arg 174 Nη1	Asp 208 Oδ1	αe-βJ
	Asp 208 Oδ2	
Arg 174 Nη2	Glu 171 Oε1	αe-αe
Arg 178 Nζ	Asp 208 Oδ2	eIloop-gHloop
Gln 214 Nε2	Glu 212 Oε2	βJ-βJ
Arg 225 Nη1	Met 221 Sδ	αg-αg
Asn 286 Nδ2	Thr 288 Oγ1	Diloop-αi

easy to dissociate into the monomers by the presence of urea (Kirino, 1991; Kirino et al., 1991). Ile238 and Val224 in helix g are replaced by Met and Ile in mesophiles. Val191 and Phe194 interacting with the arm region of another subunit are altered by serine and leucine in mesophiles. As the size of these residues is different between mesophiles and thermophiles, intersubunit interactions of thermophiles may differ from those of mesophiles.

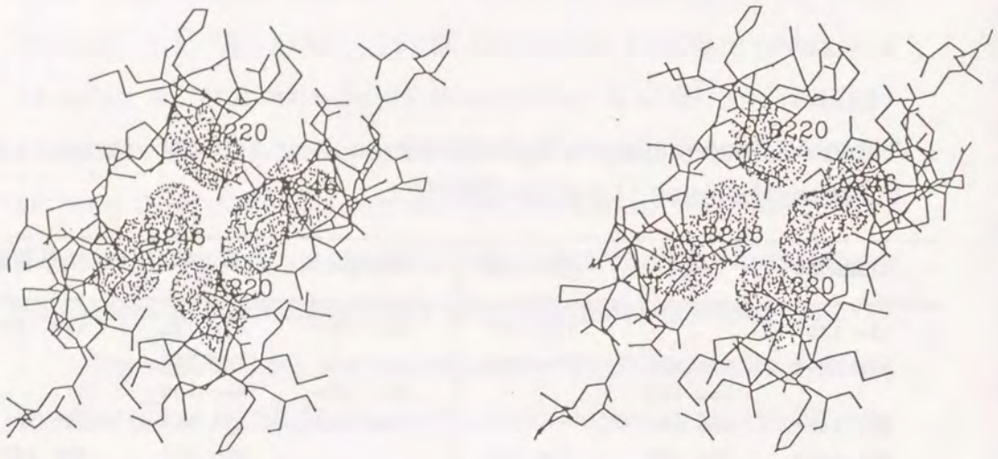
The results from the thermostability of mutants also indicate importance of hydrophobic interactions. Argos et al. (1979) and Matthews et al.(1987) proposed that the replacement of glycine in a helix to alanine stabilizes an enzyme. This strategy was applied to IPMDH. Three Gly residues, Gly89, Gly192 and Gly240 were selected and substituted to Ala. G89A mutant was thermally stabilized more than Tt-IPMDH, but the remaining mutants, G192A and G240A, were destabilized (Kirino,1991). On the basis of the structure, these results may be interpretated as follows;

There is space around Gly89 in the molecule of Tt-IPMDH. On the other hand, additional C_β atom in G89A may be filling the space (Figure8.5.2). As the result, hydrophobic interactions are increased and the mutant is stabilized. However, there is no enough space around the other glycine residues to add C_β atom. The distance between the C_β atom of Ala192 and the carbonyl carbon atom of Lys185 is 2.36Å. The C_γ atom of Arg132 is within 2.4Å from the C_β atom of Ala240. Therefore the introduction of alanine in the 192 and 240 sites causes unfavorable contact with neighbor atoms and conformational distortion of helix f and h (Figure8.5.3 and 8.5.4).

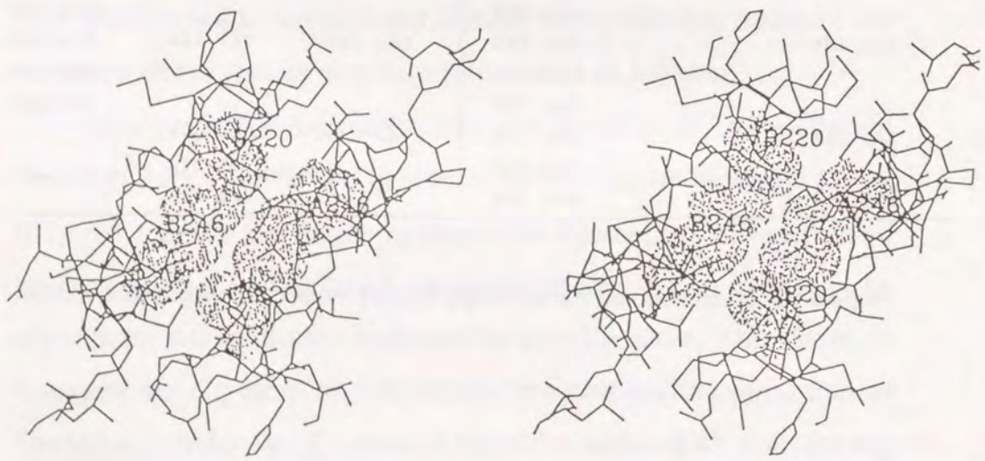
Table 8.5.1 Intermolecular hydrophobic contacts at dimer interface not conserved between Tt and Bs-IPMDH

distance	≤4Å	4Å≤ ≤5Å	distance	≤4Å	4Å≤ ≤5Å
Ile 122'		Pro 117	Val 191'	Ala 151	Asn 153
Ile 138'	Glu 155	Ile 138		Trp 152	
	Leu 189		Glu 193'	Met 146	
Gly 145'	Glu 190		Phe 194'	Ala 149	Met 146
Met 146'	Glu 190	Phe 194		Glu 150	Ser 147
	Glu 193			Ala 151	Glu 148
Ser 147'		Phe 194	Val 224'	Pro 117	Leu 250
Ala 149'	Phe 194	Ser 158		Val 224	
		Lys 159		Leu 246	
Trp 152'	Val 191	Thr 154		Val 249	
		Glu 155	Arg 225'	Val 249	Pro 117
		Arg 156			Leu 254
Glu 155'	Ile 138	Trp 152	Ile 238'	Phe 239	Lys 185
		Asn 153			Leu 189
Arg 156'		Glu 150	Leu 246'	Val 224	Ala 220
		Ala 151			Met 221
		Trp 152			Leu 246
Ser 158'		Ala 149	Val 249'	Met 221	
		Glu 150		Val 224	
Lys 159'		Ala 149		Arg 225	

Superscript prime denotes the symmetry related molecule constructiong dimer.

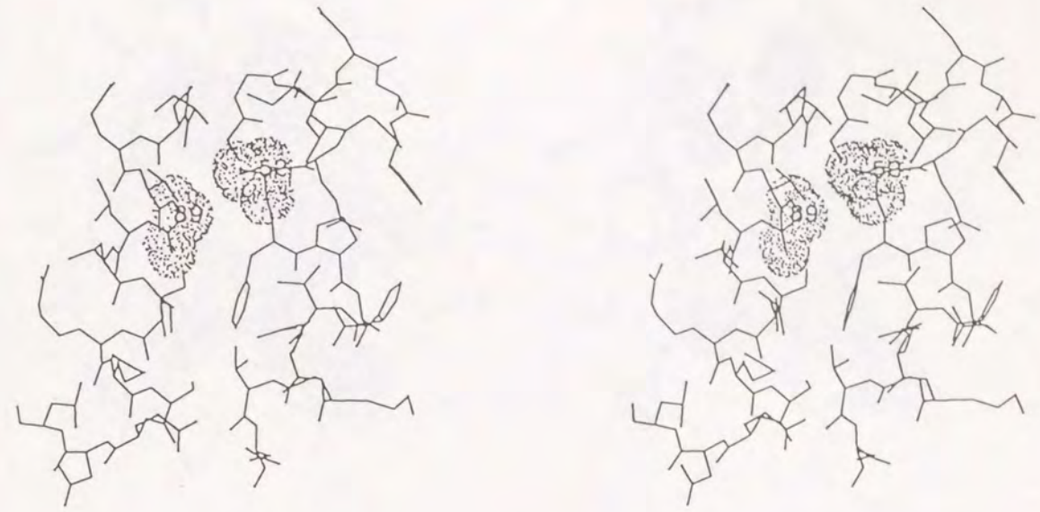


(a)

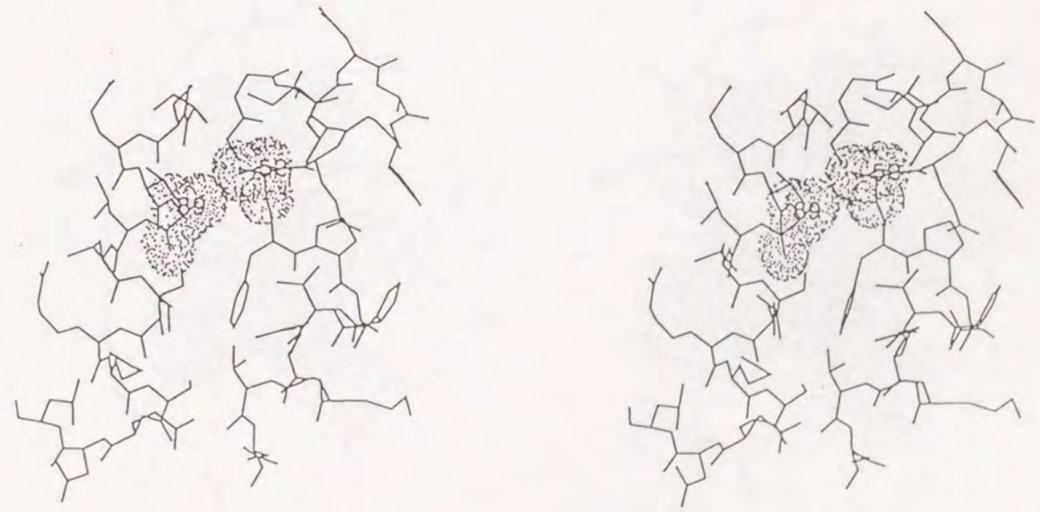


(b)

Figure 8.5.1 Stereo drawing around residue 246
(a) before substituion of L246E (b) after substitution of L246E

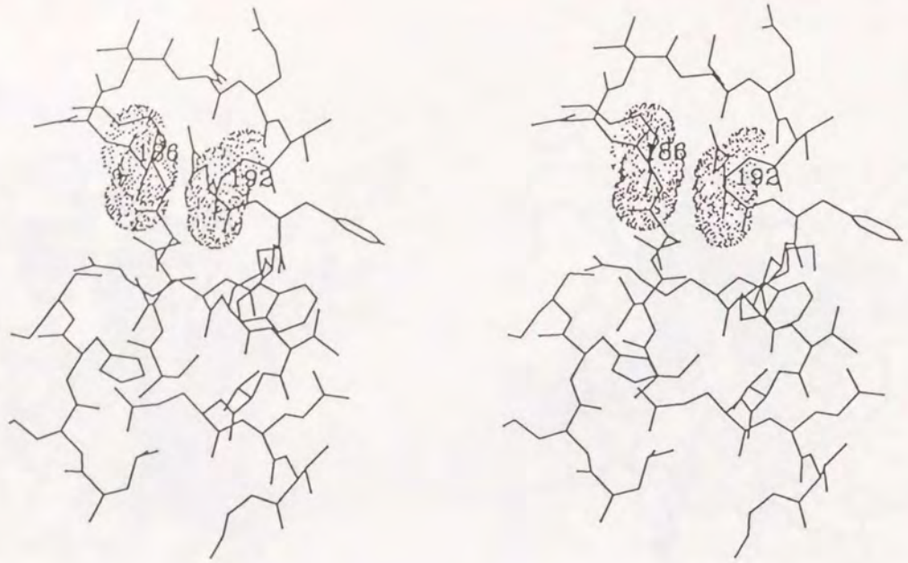


(a)

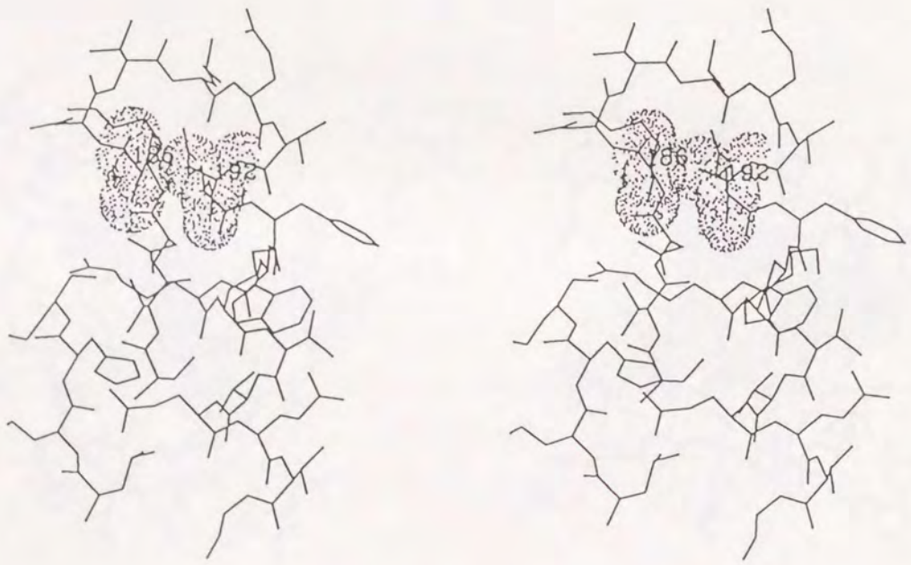


(b)

Figure 8.5.2 Stereo drawing around residue 89
(a) before substituion of G89A (b) after substitution of G89A

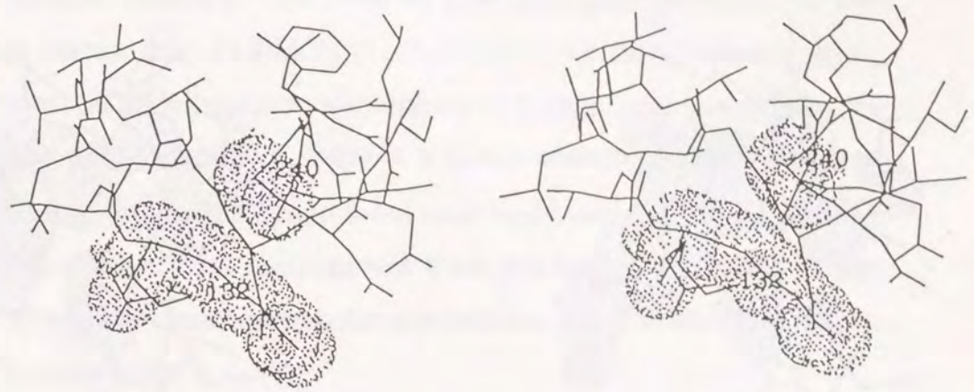


(a)

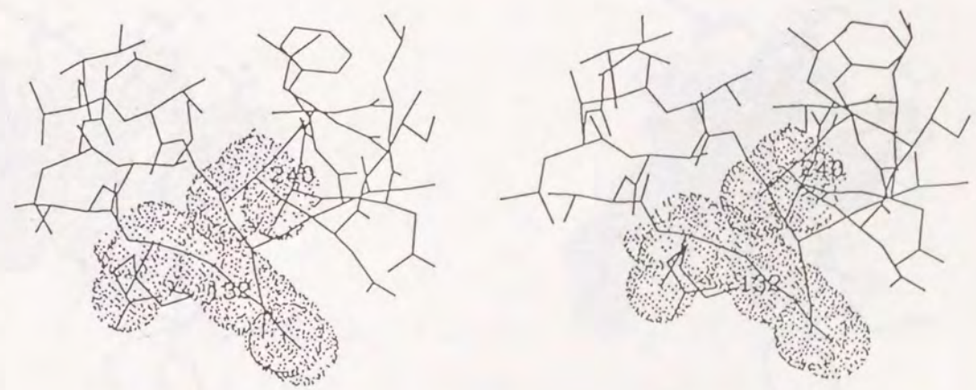


(b)

Figure 8.5.3 Stereo drawing around residue 192
(a) before substituion of G192A (b) after substitution of G192A

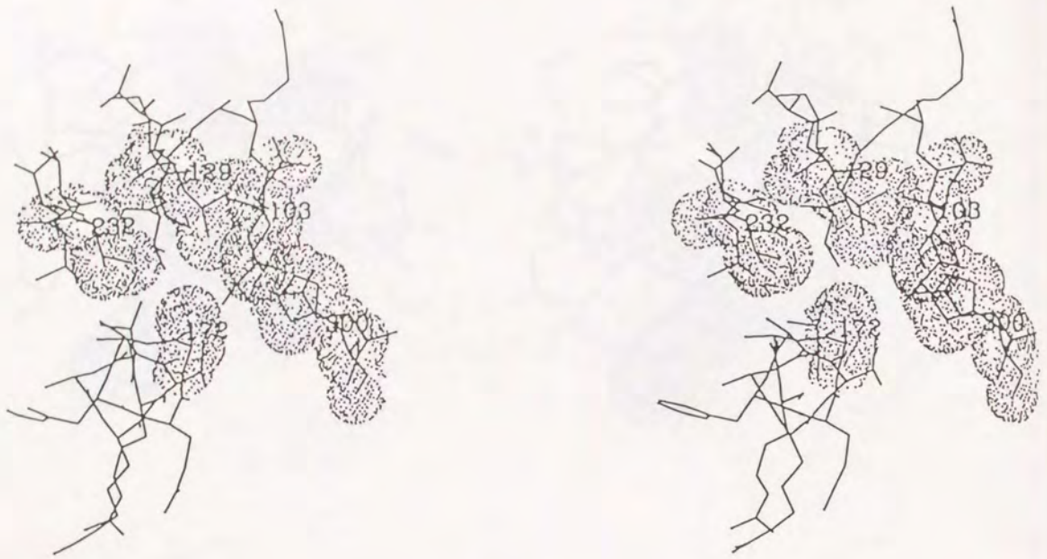


(a)

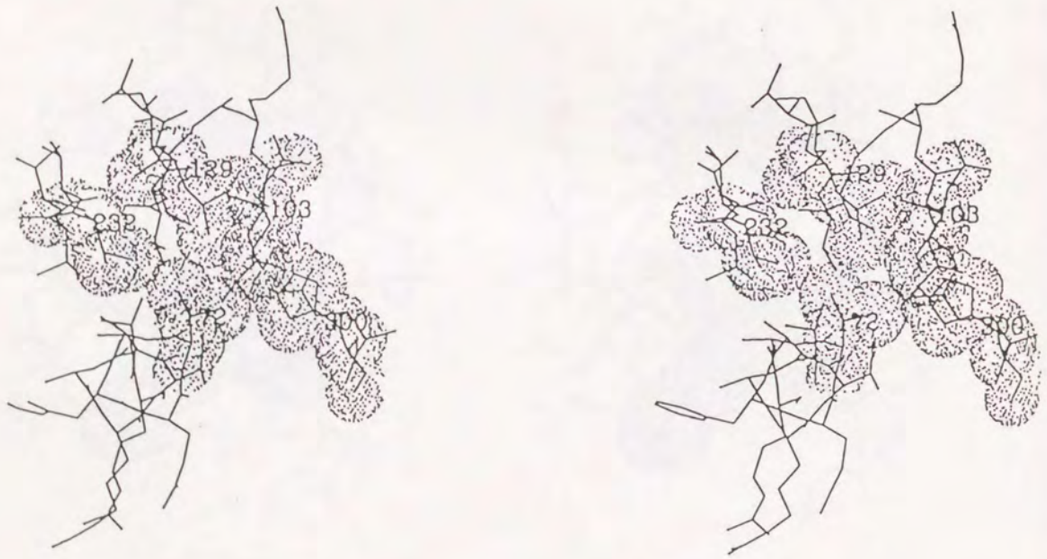


(b)

Figure 8.5.4 Stereo drawing around residue 240
(a) before substituion of G240A (b) after substitution of G240A



(a)



(b)

Figure 8.5.5 Stereo drawing around residue 172
(a) before substituion of A172V (b) after substitution of A172V

Another mutant, A172V obtained from suppressor mutation method, is stabler than Tt-IPMDH (Tamakoshi et al., 1990; Kotsuka et al., 1991). This substitution also enhances hydrophobic interactions. As shown in Figure8.5.5, there is a space enough for the C_γ atom of valine. Around the space there exist hydrophobic residues such as Leu129 etc. The replacement from Ala to Val at the 172nd site should increase hydrophobic interactions and therefore the mutant is stabilized.

Chapter-9
Structure and B-factor Analysis

In the previous chapter, we discussed the thermostability based on the statical structures of Tt- and chimeric IPMDH. But for deep understanding of thermostability, we need techniques which give dynamic structural information. In this chapter, we describe how to determine the structure under elevated temperatures and how to use B-factors to obtain dynamical information of each atom.

These approaches should be useful for understanding not only the stability of enzymes but also their enzymatic functions. From this view point, the structure of Tt-IPMDH was determined at 45°C, 39°C, 20°C and 10°C. The reason why these conditions are chosen is as follows; at room temperature, Tt-IPMDH has the same activity as Bs-IPMDH, and the activity increases as the temperature increases; the highest activity shows at about 75°C. In Tt-IPMDH, there is an turning-point in Arhenius plot around 37~39°C, and therefore the activity becomes considerably high under elevated temperatures. Therefore it is expected that the structural change, including B-factors of the individual atoms, may occur around 38°C. Hence, the measurements were carried out above the temperature, 45°C, at the temperature, 39°C, below the temperature, 20°C, and extra below the temperature, 10°C.

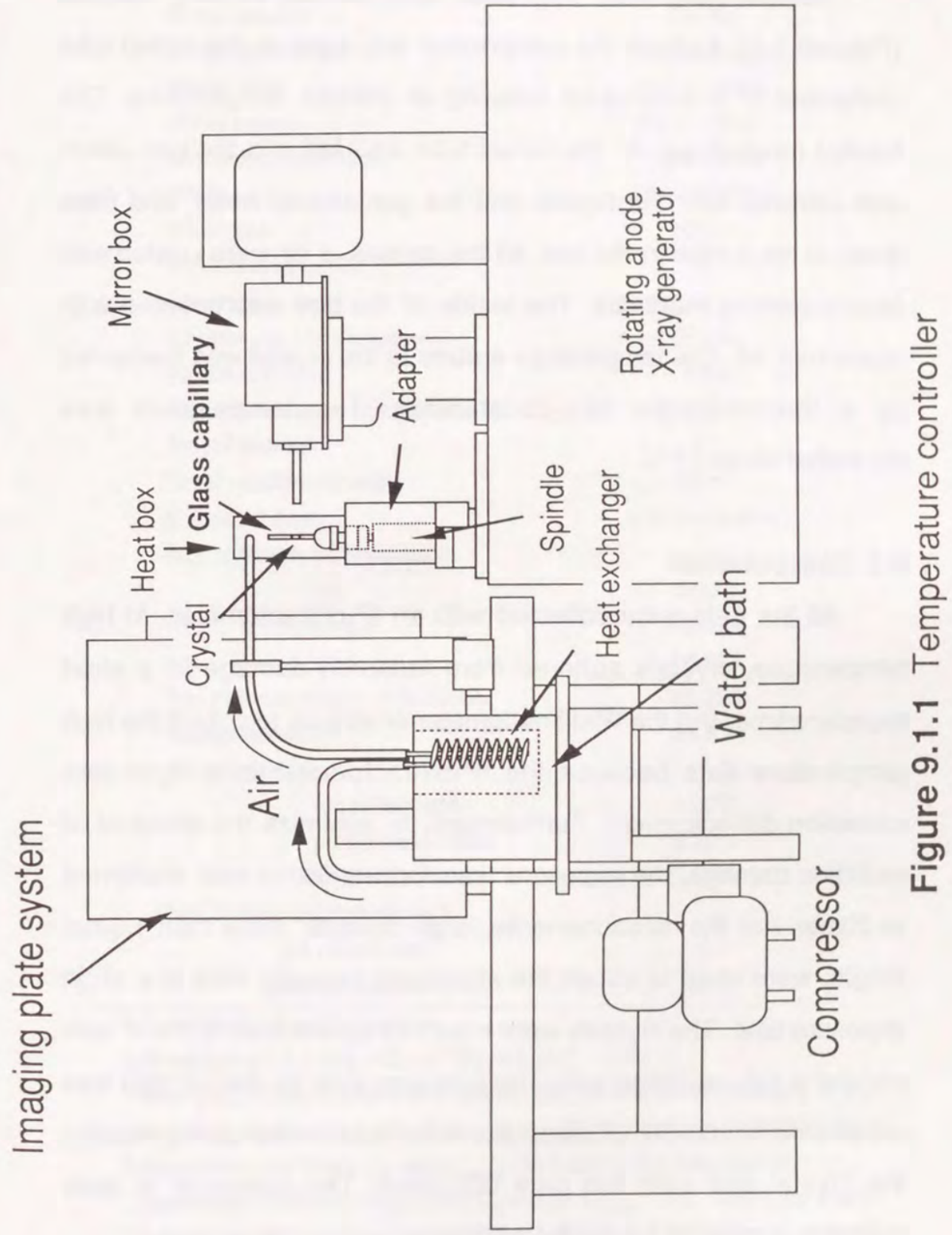


Figure 9.1.1 Tempe

9-1 Temperature control

Temperature was controlled with the air blowing method (Figure9.1.1). Air from the compressor was send to the coiled tube immersed in a water bath keeping at definite temperature. The heated (cooled) air by the coiled tube was led into the box which was covered over the crystal and the goniometer head, and blew down to the crystal in the box. All the passes of air were coated with heat insulating materials. The inside of the box was covered with aluminium foil. The temperature around at the crystal was measured by a thermo-couple (Cu-constantan). The temperature was controlled within ±1°C.

9-2 Data collection

All the data were collected with an IP-diffractometer. At high temperature, crystals suffered from radiation damage in a short exposure time. But the IP-diffractometer enable us to collect the high temperature data because the IP-diffractometer is a rapid-data collection diffractometer. Furthermore, to minimize the effect of of radiation damage, the exposure time for one frame was shortened to 20min. For the measurements, large crystals, more than 1.5mm length, were used to obtain the significant intensity data in a short exposure time. The crystals were mounted approximately the c* axis parallel to the oscillation axis. The exposure time for the full data was within nine hours. Still photographs were taken twice during heating the crystal and after the data collection. The summary of data collection is given in Table9.2.1 ~9.2.3.

Table 9.2.1

Summary of data collection of 45DEG-crystal and statistics

X-ray Source	Cu-Kα
X-ray generator	Rigaku RU200
Focus size	0.3 × 3mm
X-ray power	40kV, 100mA
Monochromatization	Ni & mirror
IP size	200×200mm
Pixel size	105μm
No. of crystal used	1
ϕ(spindle) - axis	approx. c axis
Crystal - to - IP distance	90mm
Resolution limit	2.5Å
Oscillation range per frame	1.5Å
No. of frames	17
Total oscillation range	25.5°
Exposure time	20min / frame
No. of observed reflections	
full	5,855
partial	10,959
total	16,814
No. of independent reflections	11,356
Completeness †	55.8%
R - merge ‡	
full reflections	4.99%
partial reflections	5.59%
total reflections	5.39%
No. of rejected reflections ¶	
full reflections	0
partial reflections	2

† Considering the blind region.

‡ R-merge $= \Sigma\Sigma \mid I_i(h) / G_i - \langle I(h) \rangle \mid / \Sigma\Sigma \langle I(h) \rangle$
where G_i is the inverse scale factor, $I(h)$ the diffraction intensity of the symmetry equivalent reflections, and $\langle I(h) \rangle$ the mean value of $I(h)$.

¶ Rejection criteria are $C_R = 0.3\, I_{mean} + 0.1 \langle I(h) \rangle$ for the reflections measured more than twice and $C_R = 3\,(0.3\, I_{mean} + 0.1 \langle I(h) \rangle)$ for the reflections with only two observations, where I_{mean} is the mean of all $I(h)$ values. The reflections with the differences $D_i = \mid \langle I(h) \rangle - I_i(h) / G_i \mid > C_R$ are rejected (Rossmann *et al.*, 1979).

Table 9.2.2

Summary of data collection of 38DEG-crystal and statistics

X-ray Source	Cu-Kα
X-ray generator	Rigaku RU200
Focus size	0.3 × 3mm
X-ray power	40kV, 100mA
Monochromatization	Ni & mirror
IP size	200×200mm
Pixel size	105μm
No. of crystal used	1
φ(spindle) - axis	approx. c axis
Crystal - to - IP distance	90mm
Resolution limit	2.5Å
Oscillation range per frame	1.5Å
No. of frames	21
Total oscillation range	31.5°
Exposure time	20min / frame
No. of observed reflections	
full	4,876
partial	12,098
total	16,974
No. of independent reflections	11,682
Completeness †	58.2%
R - merge ‡	
full reflections	6.75%
partial reflections	8.33%
total reflections	7.88%
No. of rejected reflections ¶	
full reflections	9
partial reflections	18

† Considering the blind region.

‡ R-merge = $\Sigma\Sigma \mid I_i(h) / G_i - \langle I(h) \rangle \mid / \Sigma\Sigma \langle I(h) \rangle$
where G_i is the inverse scale factor, $I(h)$ the diffraction intensity of the symmetry equivalent reflections, and $\langle I(h) \rangle$ the mean value of $I(h)$.

¶ Rejection criteria are $C_R = 0.3\ I_{mean} + 0.1 \langle I(h) \rangle$ for the reflections measured more than twice and $C_R = 3\ (0.3\ I_{mean} + 0.1 \langle I(h) \rangle)$ for the reflections with only two observations, where I_{mean} is the mean of all $I(h)$ values. The reflections with the differences $D_i = \mid \langle I(h) \rangle - I_i(h) / G_i \mid > C_R$ are rejected (Rossmann *et al.*, 1979).

Table 9.2.3

Summary of data collection of10DEG-crystal and statistics

X-ray Source	Cu-Kα
X-ray generator	Rigaku RU200
Focus size	0.3 × 3mm
X-ray power	40kV, 100mA
Monochromatization	Ni & mirror
IP size	200×200mm
Pixel size	105μm
No. of crystal used	1
φ(spindle) - axis	approx. c axis
Crystal - to - IP distance	90mm
Resolution limit	2.1Å
Oscillation range per frame	1.5Å
No. of frames	21
Total oscillation range	31.5°
Exposure time	20min / frame
No. of observed reflections	
full	27,450
partial	9,784
total	37,234
No. of independent reflections	22,003
Completeness †	64.7%
R - merge ‡	
full reflections	6.78%
partial reflections	6.69%
total reflections	6.75%
No. of rejected reflections ¶	
full reflections	135
partial reflections	4

† Considering the blind region.

‡ R-merge = $\Sigma\Sigma \mid I_i(h) / G_i - \langle I(h) \rangle \mid / \Sigma\Sigma \langle I(h) \rangle$
where G_i is the inverse scale factor, $I(h)$ the diffraction intensity of the symmetry equivalent reflections, and $\langle I(h) \rangle$ the mean value of $I(h)$.

¶ Rejection criteria are $C_R = 0.3\ I_{mean} + 0.1 \langle I(h) \rangle$ for the reflections measured more than twice and $C_R = 3\ (0.3\ I_{mean} + 0.1 \langle I(h) \rangle)$ for the reflections with only two observations, where I_{mean} is the mean of all $I(h)$ values. The reflections with the differences $D_i = \mid \langle I(h) \rangle - I_i(h) / G_i \mid > C_R$ are rejected (Rossmann *et al.*, 1979).

9-3 Data processing

The crystal heated at 45°C (45DEG)

At first all the data were merged and scaled, The Rmerge value for scaling was more than 10%. The first four frames which were in poor agreement with the other data were eliminated. The final Rmerge for the frames was 5.39% for 17 frames

The crystal heated at 39°C (39DEG)

Although the first frame could not be indexed, the remaining frames were successfully indexed and scaled. The final Rmerge value for scaling was 7.88% for 21 frames.

The crystal heated at 10°C (10DEG)

The Rmerge value for this data was 6.75%. Rmerge showed the normal value. The eighth and ninth frames failed in indexing. Most of the rejected reflections were detected on last seven frames. It is probably due to the misalignment of crystal and wetness on the capillary with dew.

These data differed significantly from the native data which were collected at 20°C (20DEG).

The average change of the structure factors due to heating the crystal from 20°C to 45°C was 25.7% for the reflections up to 2.7Å resolution. This variation is equal to the difference between

Table 9.3.1 Summary of data processing

	detector	resolution Å	completeness	R-merge§	R-iso¶ Native	R-iso¶ SO_4
Native (20°C)	IP*	2.2	83%	4.19%	—	18.4%#
SO_4‡ (20°C)	IP	2.1	81%	4.48%	18.4%#	—
10DEG	IP	2.2	65%	6.75%	7.7%†	19.8%$
39DEG	IP	2.5	58%	7.88%	8.7%†	16.6%$
45DEG	IP	2.5	56%	5.39%	25.3%$	11.5%$
Pt-derivative	diffractometer	2.7		4.80%	27.1%$	
Au-derivative	diffractometer	2.7		3.10%	17.3%$	
U-derivative	diffractometer	2.7		4.70%	17.9%$	

‡ see chapter-6

* imaging plate

§ R-merge = $\Sigma\Sigma \mid F_i(h) / G_i - < F(h) > \mid / \Sigma\Sigma < F(h) >$
where h is the unique reflection index and $F_i(h)$ the structure amplitude of the symmetry equivalent reflections giving a mean value of $< F(h) >$

¶ R-iso = $\Sigma \mid F_P - F_{PH} \mid / \Sigma \mid F_P \mid$

\# compared up to 3.0Å resolution

$ compared up to 2.7Å resolution

† compared up to 2.5Å resolution

isomorphous derivatives (Table9.3.1). 38DEG and 10DEG showed the variation of 8.7% and 7.7%, respectively. These are almost equal to the Rmerge and it is hence expected that the structures of these two temperatures are almost same as that of 20DEG. The summary of data processing is shown in Table9.3.1.

9-4 Refinement

Although the diffraction intensities were collected more than 2.5Å resolution, completeness of the data was about 60%. Therefore Fourier and difference Fourier maps which were calculated based on the 20DEG model, were not so clear that modify the structural difference from 20DEG and solvent molecules were not clearly defined on the electron density maps. Since the number of diffractions, more than ten thousand independent reflections, was enough for the structure refinement, 20DEG model including solvent molecules was used as the initial model for the refinement. The overall B-factor derived from Willson plot was assigned to each atom as the starting individual B-factors.

45DEG

The overall B-factor derived from Willson plot was 38Å2. The R-factor was reduced to 19.5%. The target sigmas and final r.m.s. deviations are given in Table9.4.1.

Table 9.4.1 Summary of least-squares parameters and deviations

	Target	r.m.s deviations		
		45DEG‡	39DEG‡	10DEG¶
Bonding distances (Å)				
1-2 bond	0.020	0.013	0.011	0.013
1-3 angle	0.030	0.035	0.029	0.031
1-4 planar	0.050	0.045	0.041	0.044
Planar groups (Å)	0.020	0.010	0.009	0.010
Chiral volumes (Å³)	0.150	0.166	0.148	0.161
Non-bonded contacts (Å)				
Single torsion	0.500	0.214	0.201	0.200
Multiple torsion	0.500	0.268	0.244	0.239
Possible hydrogen bond	0.500	0.290	0.273	0.244
Torsion angles (deg.)				
Planar	3.0	2.0	1.8	2.0
Staggered	15.0	23.6	23.5	23.2
Orthonormal	20.0	31.1	30.9	30.6
Thermal factors (Å²)				
Main-chain bond	1.000	0.372	0.299	0.334
Main-chain angle	1.500	0.650	0.521	0.569
Side-chain bond	1.500	0.700	0.581	0.675
Side-chain angle	2.000	1.120	0.941	1.041

‡ 2.5Å resolution
¶ 2.2Å resolution

Table 9.4.2 Dependency of the R-factors on resolution

(a) 45DEG

Resolution (Å)	No. of reflections	*R* - factor	*R* - factor (accumulated)
5.00 ~ 4.00	1,576	0.137	0.137
4.00 ~ 3.55	1,369	0.170	0.150
3.55 ~ 3.30	1,030	0.207	0.162
3.30 ~ 3.05	1,305	0.234	0.174
3.05 ~ 2.88	943	0.249	0.181
2.88 ~ 2.70	1,074	0.261	0.188
2.70 ~ 2.50	1,092	0.295	0.195
5.00 ~ 2.20	16,612	-	0.195

(b) 38DEG

Resolution (Å)	No. of reflections	*R* - factor	*R* - factor (accumulated)
5.00 ~ 4.00	1,775	0.123	0.123
4.00 ~ 3.40	2,057	0.152	0.137
3.40 ~ 3.00	1,987	0.188	0.150
3.00 ~ 2.80	1,031	0.224	0.157
2.80 ~ 2.70	497	0.243	0.160
2.70 ~ 2.60	438	0.249	0.163
2.60 ~ 2.50	332	0.287	0.166
5.00 ~ 2.20	8,117	-	0.166

(c) 10DEG

Resolution (Å)	No. of reflections	*R* - factor	*R* - factor (accumulated)
5.00 ~ 4.00	2,135	0.141	0.141
4.00 ~ 3.30	3,146	0.162	0.152
3.30 ~ 2.90	2,987	0.203	0.166
2.90 ~ 2.63	2,600	0.228	0.176
2.63 ~ 2.45	1,892	0.244	0.181
2.45 ~ 2.32	1,396	0.258	0.186
2.32 ~ 2.20	1,223	0.280	0.190
5.00 ~ 2.20	15,379	-	0.190

38DEG

The overall B-factor derived from Willson plot was 35Å^2. The R-factor was reduced to 16.6%. The target sigmas and final r.m.s. deviations are given in Table9.4.1.

10DEG

The overall B-factor derived from Willson plot was 27Å^2. The R-factor was reduced to 18.9%. The target sigmas and final r.m.s. deviations are given in Table9.4.1.

9-5 Structure description in various temperatures

After the refinement, each model was compared with 20DEG model. The mean and r.m.s. deviations are given in Table9.5.1. The number of atoms deviationg more than 1.0Å from those of 20DEG is shown in Table9.6.2. For all atoms in 38DEG and 10DEG the r.m.s deviations are less than 0.15Å, excepting for an atom in 38DEG. On the other hand, the r.m.s deviation of 45DEG resulted in slightly larger value. But after the least-squared fitting with 20DEG model, the r.m.s deviation reduced to 0.29Å. Anyway, the structures at 45°C, 38°C and 10°C are almost the same with 20DEG.

9-6 B-factor

Although the atom positions did not change significantly by heating, the distribution of B-factors in the structure were varied according to the temperature variation. Figure9.6.1 shows the averaged B-factors for 10DEG, 20DEG, 38DEG and 45DEG, as a

Table 9.5.1 differences of the coordinate between 20DEG and 45DEG

	Main chain atom	Side chain atom	All protein atom
mean distance	0.17Å	0.31Å	0.24Å
r.m.s.deviation	0.19Å	0.37Å	0.29Å
0.5Å≤ <1.0Å	0	174	174
1.0Å≤ <1.5Å	0	16	16
1.5Å≤ <2.0Å	0	0	0
2.0Å≤	0	0	0

The coordinate of 45DEG were fitted to that of Tt-IPMDH by least square fitting program MFIT using main chain atoms.

Table 9.5.2 List of atoms whose coordinates are different more than 1.0Å between native and 45DEG structure

residue	atom	distance (Å)	comment
Met 1	Cγ	1.06	
Lys 2	Nζ	1.33	β-B
Leu 32	Cδ1	1.28	
	Cδ2	1.49	
Lys 59	Cγ	1.14	middle of α-c
	Nζ	1.16	
Leu 68	Cδ1	1.32	β-C
Lys 83	Nζ	1.21	
Ile 84	Cδ1	1.09	
Leu112	Cδ1	1.31	
Arg164	Nη2	1.08	middle of α-e
Lys310	Cε	1.24	middle of α-j
Arg342	Nη2	1.02	C end of α-k
Leu344	Cδ1	1.39	
	Cδ2	1.12	

Table 9.5.3 differences of the coordinate between 20DEG and 39DEG

	Main chain atom	Side chain atom	All protein atom
mean distance	0.07Å	0.15Å	0.11Å
r.m.s.deviation	0.09Å	0.18Å	0.14Å
0.5Å≤ <1.0Å	0	13	13
1.0Å≤ <1.5Å	0	1¶	1¶
1.5Å≤ <2.0Å	0	0	0
2.0Å≤	0	0	0

The coordinate of 39DEG was fitted to that of Tt-IPMDH by a least square fitting program MFIT using main chain atoms.

¶ Ser330 O; The difference is 1.03Å.

Table 9.5.4 Differences of the coordinate between 20DEG and 10DEG

	Main chain atom	Side chain atom	All protein atom
mean distance	0.05Å	0.09Å	0.07Å
r.m.s.deviation	0.05Å	0.11Å	0.08Å
0.5Å≤ <1.0Å	0	0	0
1.0Å≤ <1.5Å	0	0	0
1.5Å≤ <2.0Å	0	0	0
2.0Å≤	0	0	0

The coordinate of 10DEG were fitted to that of Tt-IPMDH by least square fitting program MFIT using main chain atoms.

function of the residue number. Each point in Figure9.6.1 shows the averaged B-factor over the main chain atoms of the residue. According to the rise of temperature, B-factors of individual residues are getting higher. As the increment of B-factor for temperature is different with residues, distribution of B-factors in the structure is different with temperature. 10DEG and 20DEG have essentially the same distribution. But B-factors of 45DEG shows interesting behavior in some regions. 38DEG shows the distribution which is close to that of 20DEG. We remarked the regions of which B-factors were lower than their surroundings at low temperature but were high at 45DEG. In other words, we looked for the regions which became flexible at high temperature. For example, the B-factors of residues from 231 to 237 are much lower than their surrounding residues in 10DEG and 20DEG. But the B-factors of these residues at 45°C are higher than those of residues from 238 to 243 which are neighbor of the region of residues from 231 to 237. These regions are shown by arrows in Figure9.6.2.

Figure9.6.3 indicates the distribution of these regions on the three-dimensional structure. This suggests that partial molecular motion arisen by temperature increment localizes intrinsic part in the structure. Heating will trigger not only the motion of the atoms on the surface of the molecules, but also the motion of the atoms inside of the molecule. F1 region shown in Figure9.6.3 is constructed from A,B,C and D in Figure9.6.2 and contains Thr88 and Leu90 which may be important for molecular recognition. E, F, G,H and I in Figure9.6.2 make S1 region in the second domain. As shown in

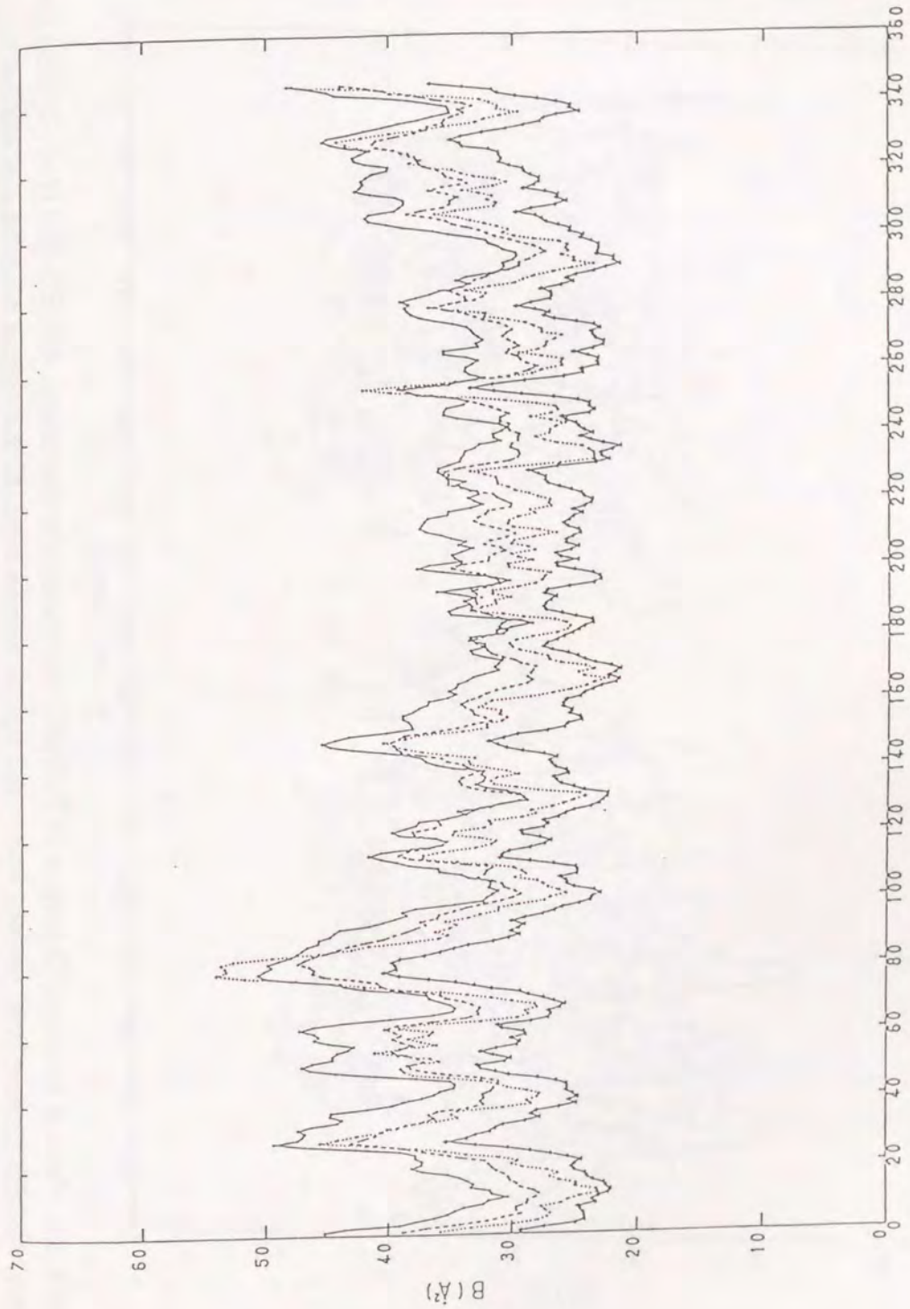


Figure 9.6.1 Plots of averaged B-factors of Tt-IPMDH for various temperature. 45DEG (solid line) ,39DEG (broken line), 20DEG (dot line) and 10DEG (solid line with ▲) are showed in the figure.

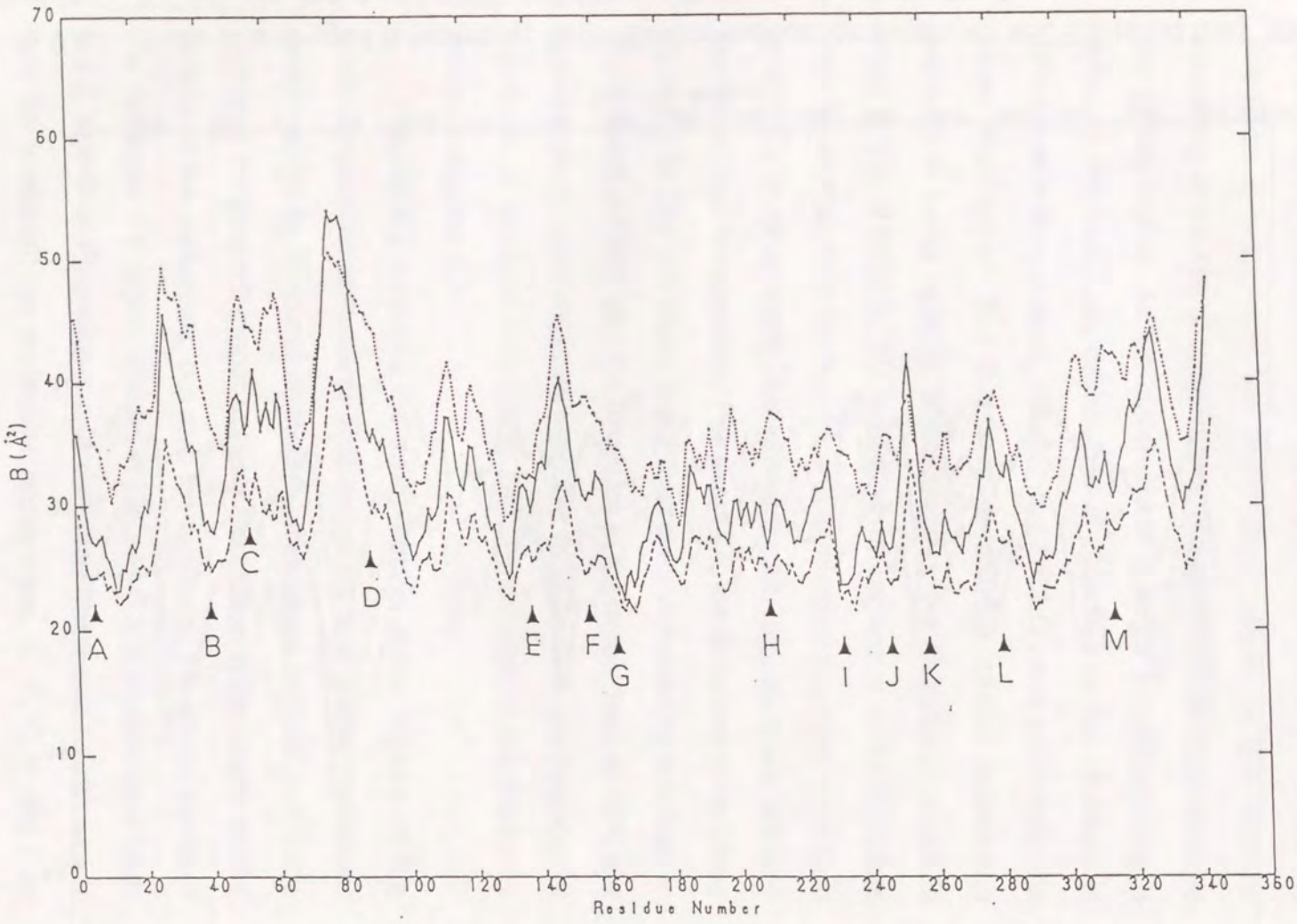


Figure 9.6.2 Plots of averaged B-factors of Tt-IPMDH for various temperature. 45DEG (dot line) , 20DEG (solid line) and 10DEG (broken line) are showed in the figure. Arrows indicate the region described in section 9-7.

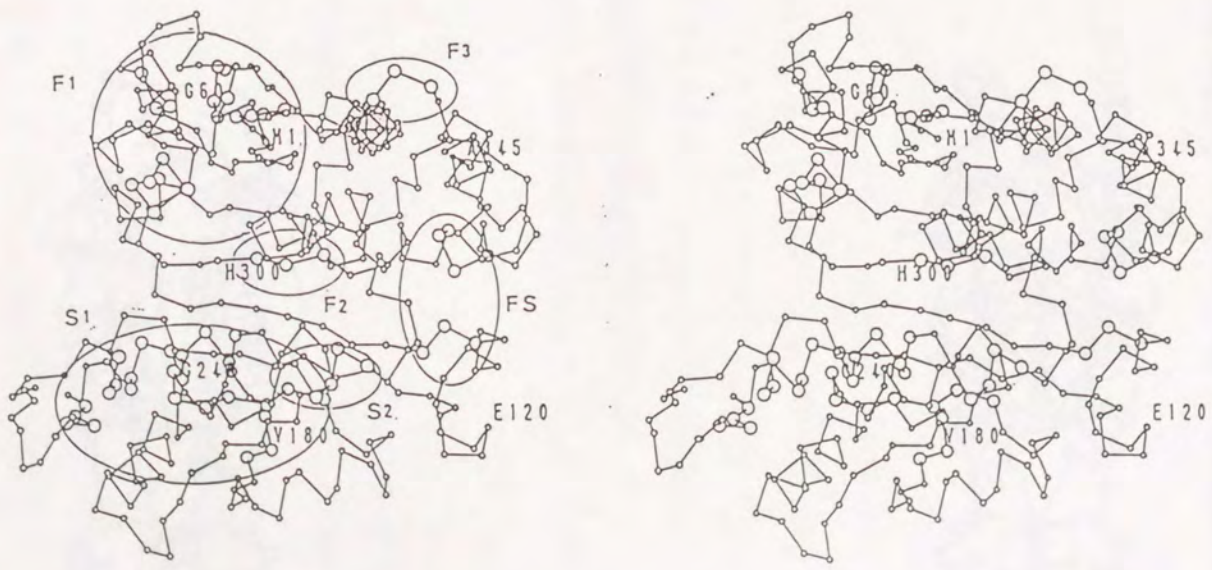
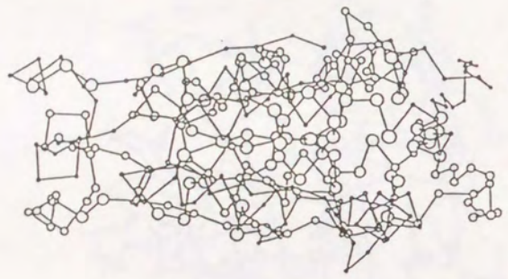


Figure 9.6.3 A stereo view of the distribution of the region whose B-factors are lower than their surrounding residues at low temperature but high at 45DEG on three dimensional structure. Large balls represent the Cα atom in such region.



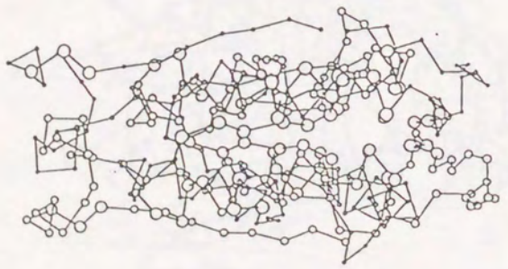


Figure 9.6.4 A stereo view of the distribution of the region where the B-factors are lower than their surrounding residues at low temperature but high at 45DEG on three dimensional structure of Tt-IPMDH second domain. Large balls represent the Cα atom in such region. One subunit is shown with middle balls and sticks and the other subunit with small balls and sticks. The large balls were gathered in the middle of the second domain.

Figure9.6.4, S1 region of which helix e and β-strand H are main components, slices the second domain. S2 region which is a part of helix h involves Asp245 which may be responsible for the substrate binding.

Chapter-10
Summary and Conclusions

The three-dimensional structure of Tt-IPMDH was determined at a resolution of 2.2Å by the X-ray diffraction method using a multiple isomorphous replacement method. Three isomorphous derivatives, K_2PtCl_4, $NaAu(CN)_4$ and $K_3UO_2F_5$, were successfully prepared by a conventional soaking method and subjected to primary phasing. An initial model was constructed by use of the data collected on four circle diffractometer until 2.7Å resolution. High resolution data upto 2.2Å resolution were collected on an ip-diffractometer, and used for stereochemically restrained least-square refinement. The final crystallographic R-factor was reduced to 0.182 for 20307 reflections.

The polypeptide chain of a subunit of Tt-IPMDH is folded into two domains, designated as the first and second domains, with similar conformations and folding topologies. The domains are basically composed of four parallel and one anti-parallel β-strands surrounded by some α-helices. Inter-domain hydrogen bonding occurs between the anti-parallel strands of the respective domains to form a large ten-stranded β-sheet in the subunit. In addition, there are two sub-structures in the subunit; one at the C-terminal region in the first domain, and the other at the armlike region that protrudes from the second domain.

The subunit of Tt-IPMDH is in close contact with another subunit to give rise to an isologous dimer with a crystallographic 2-fold symmetry. Hydrophobic interaction and hydrogen bonding contribute to subunit contact. The dimer has two large pockets that are made up from the first domain of one subunit and the second domain of the other. The pockets may include the amino acid residues responsible for substrate binding and catalysis.

The folding topologies of the first and second domains differ from those of the NAD-binding domains of well-known enzymes such as LADH and LDH. This indicates that IPMDH is not related evolutionally to those well-known enzymes. In contrast, Tt-IPMDH shows a marked similarity to *E. coli* ICDH both in its amino acid sequence and in its folding topology. It is therefore suggested that IPMDH and ICDH have diverged from a common ancestral protein.

The three-dimensional structure of Tt-IPMDH crystallized from highly concentrated ammonium sulfate solution was also determined at a resolution of 2.2Å by the X-ray diffraction method, because the crystal was significantly different from that of original crystal (R_F=18.4% for 3.0Å resolution data). The resultant structure showed that the difference was caused by the difference in the distribution of solvent molecules that were fixed on the protein surface, and not by the conformational difference.

In order to provide an experimental basis for elucidating the structure-thermostability relationship, we also determined the three-dimensional structures of chimeric enzymes of Tt-IPMDH. Chimeric enzymes, 4M6T, 2T2M6T, and its mutant I93L were crystallized isomorphously to Tt-IPMDH by hanging drop vapor diffusion method. Diffraction data were collected for their native crystals with the ip-diffractometer and phased on the basis of the refined Tt-IPMDH structure. The 2.2Å electron density maps of 2T2M6T and I93L were clear enough to build up their atomic structure models. But, the electron density map of 4M6T did not allow us to trace the polypeptide chain, unambiguously due to the poor diffraction ability of the crystal beyond 3.0Å resolution. The structures of 2T2M6T and I93L were refined at 2.2Å resolution by PROLSQ. The final R-factors are 19.6% for 2T2M6T and 19.5% for I93L. Their structures are slightly different in the regions of C-d loop and F-G loop, but show no significant structure differences responsible for their altered thermostabilities.

A comparison of the amino acid sequences of thermophilic and mesophilic IPMDH made on the basis of the three-dimensional structure of Tt-IPMDH proposes that the hydrophobic interaction is a major factor affecting the thermostability of the enzyme. From this point of view, L246E, G89A, G192A, G240A and A172V mutants of Tt-IPMDH were prepared and subjected to their activity measurements. As a result, their thermal stabilities were well correlated with the increase and decrease of hydrophobic

interaction, indicating the significance of the interaction for thermostabilization of protein structure.

A new attempt was made to obtain further information for structure-thermostability relationship. The statical structural studies and the structural studies under room temperature are insufficient to understand the thermostability. To obtain new information on the structural basis of Tt-IPMDH, the three-dimensional structures of Tt-IPMDH at 45°C, 39°C, 20°C and 10°C were also determined by the X-ray diffraction method. All the diffraction data were rapidly collected by the ip-diffractometer equipped with a laboratory-made thermocontroler. The three-dimensional structures determined at 45°C, 39°C and 10°C were almost the same as that determined at 20°C. But the B-factor distribution of amino acid residues showed significant difference between the structures at 39°C and 45°C. Furthermore, the regions showing the different distribution in B-factor were localized in the three-dimensional structure of Tt-IPMDH. This analysis was the first attempt for protein crystallographic investigation and quite useful to elucidate directly the relationship between structure and thermostability. The analyses are now in progress also for various mutant and chimeric enzymes of Tt-IPMDH. The comparative study of the distribution among these enzymes will promise a best understanding of the basic structure affecting the thermostability of Tt-IPMDH.

We noted the regions whose B-factors were lower than their surrounding residues at low temperature but were high at 45°C. Projecting the regions on the three dimensional structure, we found that they were gathered in some place. This means that partial molecular motion arisen by temperature increment localizes in the molecule. The motion was seen not only on the molecular surface but also within the molecule.

References

Adams, M. J., Ford, G. C., Koekoek, R., Lentz Jr., P. J., McPherson Jr., A., Schevitz, R. W., & Wonacott, A. J., (1970). *Nature*, **227**, 1098-1103

Akutsu, N. (1989) A Master thesis, Tokyo Institute of Technology, Yokohama, Japan

Alber, T., Dao-pin, S., Willson, K., Wozniak, J. A., Cook, S. P. & Matthews, B. W., (1987). *Nature* **330**, 41-46

Amemiya, Y. & Miyajima, J., (1988). *Nature*, **336**, 89-90

Andreadis, A., Hsu, Y., Hermodson, M., Kohlhaw, G. & Schmmel, P., (1984). *J. Biol. Chem.*, **259**, 8059-8062

Argos, P., Rossmann, M. G., Gram, U. M., Zuber, H., Frank, G. & Tratschin, J. D., (1979). *Biochemistry*, **18**, 5698-5703

Arndt, U. W. & Wonacott, A. J., (1977). *The Rotation Method in Crystallography*, North-Holland, Amsterdam

Artymuik, P. J. & Blake, C. C., (1981). J. Mol. Biol., 152, 737-762

Åkeson, Å. & Jones, T. A., (1981). *J. Mol. Biol.*, **146**, 561-587

Baldwin, R. L. & Eisenberg, D., (1987). *Protein engineering*, pp.127-148, Alan R. Liss, Inc., New York

Blow, D. M. & Crick, F. H. C., (1959). *Acta Crystallogr.*, **12**, 794

Branden, C. I., Eklund, H., Nordstrom, B., Boiwe, T., Soderlund, G., Zeppezauer, E., Ohlsson, I. & Åkeson, Å., (1973). *Proc. Natl. Acad. Sci., U.S.A.* **70**, 2439-2442

Buehner, M. , Ford, G. C., Moras, D., Olsen, K.W. & Rossmann, M. G., (1974). *J. Mol. Biol*, **90**, 25-49

Chou, P. Y. & Fasman, G. D., (1978). *Ann. Rev. Biochem.*, **47**, 25-76

Creighton, T. E., (1984). *Protein*, pp.251-252, W. H. Freeman and Company, New York

Davidow, L. S., Kaczmarek, F. S., Dezeeuw, J. R., Conlon, W. S., Lauth, M. R., Pereiva, D. A. & Franke, A. E., (1987). *Curr. Genet.*, **11**, 377-383

Dean, A. M. & Koshland, D. E. Jr., (1990). *Science*, **249**, 1044-1046

Dickerson, R. E., Kendrew, J. C. & Strandberg, B. E., (1961). *Acta Crystallogr.*, **14**, 1188-1195

Fox, G. C. & Holmes, K. C., (1966). *Acta Crystallogr.*, **20**, 886-891

Franks, A., (1955). *Proc. Phys. Soc, London Sect.*, **B 68**, 1054

Hamasawa, K., Kobayashi, Y., Harada, S., Yoka, K., Yamasaki, M. & Tamura, G., (1987). *J. General Microbiol.*, **133**, 1089-1097

Hamilton, W. C., Rollett, J. S. & Sparks, R. A., (1965). *Acta Crystallogr.*, **18**, 129-130

Hendrickson, W. A. & Konnert, J. H., (1980). In *Computing in Crystallography* (Diamond, R., Ramaseshan, S. & Venkatesan, K., eds), pp.13.01-13.23, Indian Institute of Science, Bangalore

Higashi, T., (1989). *J. Appl. Cryst.*, **22**, 9-18

Hurley, J. H., Thorsness, P. E., Ramalingam, V., Helmers, N. H., Koshland, D. E. Jr. & Stroud, R. M., (1989). *Proc. Nat. Acad Sci., U.S.A.* **86**, 8635-8639

Hurley, J. H., Dean, A. M., Sohl, J. L., Koshland, D. E. Jr. & Stroud, R. M., (1990). *Science*, **249**, 1012-1016

Imai, R., Sekiguchi, T., Nosoh, Y. & Tsuda, K. (1987) *Nucleic Acids Res.*, **15**, 4988

Jones, T. A., (1978). *J. appl. Crystallogr.*, **11**, 268-279

Katsube, Y., Tanaka, N., Takenaka, A., Yamada, T. & Oshima, T., (1988). *J. Biochem. (Tokyo)*, **104**, 679-680

Kirino, Y. (1991). *A Doctorial Thesis*, Tokyo Institute of Technology, Yokohama, Japan

Kirino, H., Wakagi, T., Imada, K., Katsube, Y., Tanaka, N., Arisaka, F. & Oshima, T., (1991). *Annual Meeting of the Japanese Biochemical Society,* 2401, Tokyo, Japan

Klibanov, A. M. & Ahern, T. J., (1987). *Protein engineering*, pp.127-148, Alan R. Liss, Inc., New York

Kotsuka, T., Tamakoshi, M., Yamagishi, A., & Oshima, T., (1991). *Annual Meeting of the Japanese Biochemical Society,* 2400, Tokyo, Japan

Lattman, E., (1985). *Methods Enzymol.*, **115**, 56-65

Leslie, A. G. W., Brick, P. & Wonacott, A. T., (1986) *Daresbury Lab. Inf. Quart. Prot. Crystallogr.* **18**, 33-39

Matthews, B. W., (1968). *J. Mol. Biol.*, **33**, 491-497

Matthews, B. W., Nicholson, H. & Becktel, W. J., (1987). *Proc. Natl. Sci. U.S.A*, **84**, 6663-6667

Miyazaki, K., Wakagi, T., Yoshida, K. & Oshima, T., (1989). *Annual Meeting of the Japanese Biochemical Society,* 3p-Ae03, Kyoto, Japan

Miyazaki, K. (1991) A Master thesis, Tokyo Institute of Technology, Yokohama, Japan

Nagahari, K., Koshikawa, T. & Sakaguchi, K., (1980). *Gene*, **10**, 137-145

North, A. C. T., Phillips, D. C. & Mathews, F. S., (1968). *Acta Crystallogr.*, **24**, 351-359

Numata, K., Muro, K., Kirino, H., Tanaka, N., Katsube, Y., Umeyama, Y. & Oshima, T., (1990). *Annual Meeting of the Japanese Biochemical Society,* 0058, Osaka, Japan

Numata, K., Hayashi, Y., Tanaka, N., Katsube, Y. & Oshima, T., (1991). *Annual Meeting of the Japanese Biochemical Society,* 2397, Tokyo, Japan

Ohlsson, I., Nordstrom, B., & Branden, C. I, (1974). *J. Mol. Biol.,* **89**, 339-354

Onodera, K., Moriyama, H., Takenaka, A., Tnaka, N., Akutsu, N., Muro, M., Oshima, T., Imada. K., Sato, M. & Katsube, Y., (1991). *J. Biochem. (Tokyo)*, **109**, 1-2

Osguthor, D. J. & Robson, B., (1978). *J. Mol. Biol.* **120**, 97-120

Pace, C. N., (1986). *Meth. Enzym.*, **131**, 266-279

Pantoliano, M. W., Ladner, R. C., Bryan, P. N., Rollence, M. L., Wood, J. F. & Poulos, T. L., (1987). *Biochemistry* **26**, 2077-2082

Patterson, A. L., (1934). *Phys. Rev.*, **46**, 372

Perry, L. J. & Wetzel, R., (1984). *Science* **226**, 555-557

Perry, L. J. & Wetzel, R., (1986). *Biochemistry* **25**, 733-739

Perutz, M. F., (1978). *Science* **201**, 1187-1191

Priestle, J. P., (1988). *J. Appl. Crystallogr.,* **21**, 572-576

Rice, D. W., (1981) *Acta Crystallogr.*, **A37**, 491-500

Ramakrishnan, C. & Ramachandran, G. N., (1965). *Biophys. J.*, **5**, 909-933

Richardson, J. S., (1985). *Methods Enzymol.*, **115**, 341-358

Rossmann, M. G., (1975). In *Structure and Conformation of Nucleic Acids and Protein-Neucleic Acid Interactions* (Sundaralingam, M. & Rao, T., eds), pp. 353-374, University Park Press, Baltimore

Rossmann, M. G., (1979). *J. Appl Cryst.*, **12**, 225-238

Sakurai, M., Onodera, K., Moriyama, H., Tanaka, N., Numata, K. & Oshima, T., (1991). *Annual Meeting of the Japanese Biochemical Society,* 2395, Tokyo, Japan

Sato, M., Yamamoto, M., Imada, K., Katsube, Y., Tanaka, N. & Higashi, T., (1991). *J. Appl. Cryst.*, **23**, (in press)

Scrutton, N. S., Berry, A. & Perham, R. N., (1990). *Nature*, **343**, 38-43

Sekiguchi, T., Ortega-Cesena, J., Nosoh, Y., Ohashi, S., Tsuda, K. & Kanaya, S., (1986). *Biochem. Biophys. Acta*, **867**, 36-44

Sekiguchi, T., Suda, M., Ishii, T., Nosoh, Y. & Tsuda, K., (1987). *Nucleic Acids Res.* **15**, 853

Suzuki, Y., Oishi, K., Nakano, H. & Nagayama, T., (1987), *Appl. Microbiol. Biotechnol.*, **26**, 546-551

Tamakoshi, M., Aizawa, H., Yamagishi, A., & Oshima, T., (1990). *Annual Meeting of the Japanese Biochemical Society,* 0059, Osaka, Japan

Tanaka, T., Kawano, N. & Oshima, T., (1981). *J. Biochem. (Tokyo)*, **89**, 677-682

Volkin, D. B. & Klibanov, A. M., (1987). *J. Biol. Chm.*, **262**, 2945-2950

Wierenga, R. K., De maeyer, M. C. H. & Hol, W. G. J., (1985). *Biochemistry*, **24**, 1346-1357

Wilson, A. J. C., (1949). *Acta Crystallogr.*, **2**, 318

Yabuki, T., Harada, S., Kai, Y., Kasai, N., Imanakla, T., Takagi, M., Kusunoki, M. & Yoshiki, M., (1988). *Annual Meeting of the Japanese Chemical Society (Japan),* pp.19

Yamada, T., Akutsu, N., Miyazaki, K., Kakinuma, K., Yoshida, M. & Oshima, T., (1990). *J. Biochem.* (Tokyo), **108**, 449-456

Yamamoto, M., (1991). A Doctorial Thesis, Osaka University, Osaka, Japan

Yutani, K., Ogasawara, K., Sugino, Y. & Matsushiro, A., (1977). *Nature* **267**, 274-275

Yutani, K., Ogasawara, K., Tsujita, T., Kanemoto, K., Matsumoto, M., Tanaka, S., Miyashita, T., Matsushiro, A., Sugino, Y. & Miles, E. W. (1987). *J. Biol. Chem.* **262**, 13429-13433

Yutani, K., Ogasawara, K., Tsujita, T. & Sugino, Y., (1987). *Proc. Natl. Acad. Sci., U.S.A.* **84**, 4441-4444

List of Publications

(1) Three-dimensional Structure of a Highly Thermostable Enzyme, 3-Isopropylmalate Dehydrogenase of *Thermus thermophilus* at 2.2Å Resolution
Imada, K., Sato, M., Tanaka, N., Katsube, Y., Matsuura, Y. & Oshima, T.
(1991). *J. Mol. Biol.* **222**, 725-738

(2) Crystallization and Preliminary X-ray Studies of a *Bacillus subtilis* and *Thermus thermophilus* HB8 Chimeric 3-sopropylmalate Dehydrogenase
Onodera, K., Moriyama, H., Takenaka, A., Tnaka, N., Akutsu, N., Muro, M., Oshima, T., Imada. K., Sato, M. & Katsube, Y., (1991). *J. Biochem. (Tokyo)*, **109**, 1-2

(3) The Crystal Structure Analysis of 3-Isopropylmalate Dehydrogenase at 2.7Å Resolution
Imada, K., Sato, M., Matsuura, Y., Katsube, Y., Tanaka, N. & Oshima, T.
(1990) *Acta Crystallogr.*, **A46**, Supplement C-130

Other related paper

Development of a High Speed Data Collection System with an Imaging Plate in Consideration of Large Unit Cell Crystals
Sato, M., Yamamoto, M., Imada, K., Katsube, Y., Tanaka, N. & Higashi, T., (1991). *J. Appl. Cryst.*, **23**, (in press)