STRUCTURAL AND FUNCTIONAL STUDIES ON *DROSOPHILA* ADH

Lluís Ribas i de Pouplana

Thesis presented for the degree of
Doctor of Philosophy.
University of Edinburgh, 1992.

Contents: Page number

"Pixo a l'abisme,

al fons la mar blava,

allí el cap de Bagur,

aqui el cap de la fava"

J.M. de Segarra

"Potser ja seré lluny, però el camí
no voldria oblidar-lo perquè sempre
sigui benigne i fàcil el retorn.
Tancant els ulls veuré de nou la casa
i l'heura i el xiprer quan, a sol post,
fa de bon seure a sota l'olivera ..."

Miquel Martí i Pol

"All science is either physics or stamp-collecting."

Rutherford

Als de casa.

Foreword .

Although the final decision of coming to Scotland was mine there is a group of people who are also to blame, and to be thanked, for suggesting and supporting such move. My family, Montserrat and Professor Roser Gonzàlez-Duarte helped to provide both the opportunity and the necessary encouragement that it took to leave Catalonia.

A part of my learning process since 1989 has been in the field of biochemical research. Dr. Linda Gilmore's supervision has always been enthusiastic and encouraging. I specially appreciate the independence I was given to pursue my research, it did not make life easier but it made it much more interesting.

Eventually Scotland started to become 'home', as friends slowly appeared. I first shared the feeling of being an alien with João, then Myrtle appeared with her wordily sapience. Chickpea not only increased my list of funnily-named friends, but took a fair deal of headaches out of my acquaintance. Malcolm's sense of humour kept my brain active and alert, if only we had had a common language... . Undeservedly Fiona was left with the horrible task of sharing most of these years with me, with her love and maturity she has made all the difference, and is responsible for many magic times.

The 'lab 327' people and the crystallography group of the department provided the environment for many discussions, celebrations, questions, answers, fights and all sort of events. I owe them a big deal.

Dr. N.C. Price was another notable scientific stimulus, who must also be thanked for the invaluable help with all the denaturation data and the use of the Stirling CD facility. This facility is run by Sharon Kelly, who, although theologically speaking very distant, made me actually look forward to my visits to Stirling University.

I must also thank Dr. Steve Chapman, for his help to analyze the computer model, and Dr. Joan Fibla for the gift of the anti-ADH antibodies and the technical assistance with the immunodetection.

Drs. Atrian, Albalat, Fibla, Marfany, Gonzalez-Duarte, Ribas, Collet, and Vilageliu, of the genetics department of the University of Barcelona, offered indispensable technical help at many stages of this thesis, they also pushed me forwards with innumerable E-mail messages of support.

The economical support was provided by a British Council/La Caixa fellowship, a faculty of Medicine scholarship and, when everything run out, by my family.

Three friends,Andrea, Cristina and Eugeni, have been in Scotland all the time, with a fantastic reassuring obstinacy. It seems to be impossible to get rid of them, and is just as well.

Finally, I must thank my family. Without them I simply would not have managed. They always offered their help before anyone, and it was always the best help too. This work is for them. I am sure they will understand it better than anyone else.

Abbreviations.

1. Miscellaneous.

A- adenine

Adh- alcohol dehydrogenase gene

ADH- alcohol dehydrogenase enzyme

ATP- adenosine-5'-triphosphate

bp- base pairs

C- cytosine

C- - carboxy-

CD- circular dichroism

cDNA- complementary deoxyribonucleic acid

D. - *Drosophila*

d- - deoxy-

DNA- deoxyribonucleic acid

ds- double stranded

DTT- dithiothreitol

E. coli - *Escherichia coli*

EDTA- ethylendiaminetetraacetic acid

FPLC- fast pressure liquid chromatography

G- guanine

GdnHCl- guanidinium hydrochloride

Iso- isopropanol

kb- kilobase

Kcat- catalytic constant

Km- apparent Michaelis constant

LDH- lactate dehydrogenase

Mr- relative molecular mass

NAD- nicotinamide adenine dinucleotide

NADH- nicotinamide adenine dinucleotide reduced

OD- optical density

PAGE- polyacrylamide gel electrophoresis

RNA- ribonucleic acid

rpm- revolutions per minute

SDH- 3α–20β hydroxysteroid dehydrogenase

SDS- sodium dodecyl sulphate

SDW- sterile distilled water

ss- single stranded

T- thymine

Tris- tris (hydroxymethyl) amino methane

Vmax- apparent maximum velocity

wt- wild type

Yeast- *Saccharomyces cerevisiae*

2. Amino acids :

A	Ala	Alanine
C	Cys	Cysteine
D	Asp	Aspartic acid
E	Glu	Glutamic acid
F	Phe	Phenylalanine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Methionine

N	Asn	Asparagine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
T	Thr	Threonine
V	Val	Valine
W	Trp	tryptophan
Y	Tyr	Tyrosine

Abstract.

Structural and functional studies of *Drosophila* alcohol dehydrogenase.

The enzyme alcohol dehydrogenase of the fruit-fly *Drosophila* (DADH) catalyzes the same reaction as the mammalian alcohol dehydrogenases, transforming alcohols into aldehydes through the reduction of nicotinamide-adenine-dinucleotide. Despite this identical chemical behaviour the enzyme's structure belongs to a different class of proteins, called short-chain dehydrogenases, which have a totally different three-dimensional architecture to the mammalian alcohol dehydrogenases. The study of the structure-fuction relationships of DADH is of interest because, while we still lack a crystal structure for the protein, large amounts of biochemical,evolutionary and genetical data have accumulated which require structural information on the enzyme for its proper interpretation.

The aim of this project was two-fold: to set up a suitable system for the undertaking of protein engineering studies on DADH and to start such studies by producing and analyzing a first set of site-directed mutants. The first part of the project involved: a) creating suitable genetic vectors for the introduction of mutations, their sequencing and the expression of the mutated enzymes in yeast; b) the development of a purification method to obtain pure enzyme solutions from the expressing yeast culture; c) the development of biochemical and biophysical assays for the evaluation of the mutation's effects and, d) the construction of a three-dimensional model for the enzyme that could offer structural explanations to such effects.

The second part consisted of the actual introduction of five different mutations in the enzyme's sequence, and their further evaluation using the system set up in the first place.

Declaration :

The work presented in this thesis is the result of my own research unless otherwise stated, and it has not been presented for any other academic degree elsewhere.

Lluís Ribas

Edinburgh, Scotland. 28th of September of 1992.

1. Introduction

1.1 Historical review of *Drosophila* adh research.

The *Drosophila* alcohol dehydrogenase gene/enzyme system has been the object of scientific analysis for some twenty-eight years. During the same period the appearance of genetic engineering techniques and the massive increase in protein structure information has allowed researchers to gain a more molecular insight into biological problems. This has not only altered the nature of our scientific methodology, but to a large extent it has also turned scientific interests towards questions that are better suited to our new technologies.

The research in the adh system is a very good example of this scientific migration towards more molecular approaches, and, at the same time, it constitutes a rather peculiar case, having remained a good model system for many of the newly appearing research disciplines.

In the 1960's the *Adh* locus of *Drosophila melanogaster* started to be used as a genetic marker for population studies (Johnson & Denniston, 1964). Today the adh system is used in research in ecology, enzymology, molecular evolution, molecular genetics and protein structure (Chambers, 1988).

Highlighting the main breakthroughs of all this scientific effort is, of course, a matter of opinion. However, of special relevance to this thesis must be :

(a) The determination of the sequence of the alcohol dehydrogenase enzyme from *D. melanogaster* (Thatcher, 1980), (b) the cloning and sequencing of the *Adh* gene of *D. melanogaster* (Benyajati et al., 1980), (c) the discovery of the family of the short-chain dehydrogenases into which the *Drosophila* alcohol dehydrogenases

(DADH) were soon included (Jörnvall et al. 1981), (d) the detailed analysis of the kinetic behaviour of the enzyme (see Winberg et al.,1982a, 1982b, 1983, 1985, 1986, 1988a, 1988b, 1989) and, more recently, the determination of the three-dimensional structure of a member of the short-chain dehydrogenase family, 3α,20β-hydroxysteroid dehydrogenase (Ghosh et al., 1991).

The determination of the DADH sequence allowed biochemists to start to analyze the structure-function relationships of the enzyme, and, through comparison between the behaviour of various natural and laboratory-induced isoenzymes, start to identify the residues with possible roles in the enzyme's mechanism (Thatcher & Retzios, 1980). At the same time the first computer analysis of the sequence, in the form of secondary structure predictions, were published (Thatcher & Sawyer, 1980; Benyajati et al., 1980).

The cloning of the D. melanogaster *Adh* gene (Benyajati et al., 1980; Goldberg, 1980) confirmed the protein sequence, and set the ground for further extensive research on the molecular regulation of the gene expression.

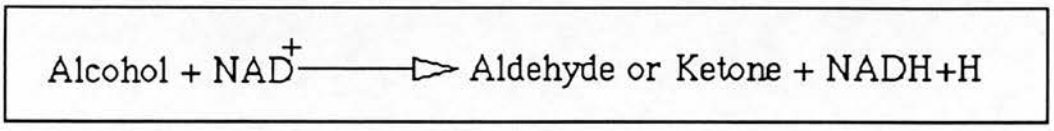
Most of our knowledge about the kinetics and reaction mechanisms of DADH come from the work of Winberg and collaborators. Their detailed enzymological studies of the enzyme have produced a complete picture of the reaction mechanism, specificity, and order.

Recently, more detailed structural studies have started to be undertaken. This has been due to the appearance of the techniques of site-directed mutagenesis (Kunkel, 1985, Fritz and Eckstein, 1985) and to the availability of the cloned *adh* gene, which has allowed the construction of expression systems in microbial organisms (Atrian et al., 1990, Chen et al., 1990).

These structural studies, however, lack the vital information that only a three-dimensional structure for the molecule can provide. The publication of the crystallographic structure of a member of the short-chain dehydrogenases family represents a very important step forward (Ghosh et al., 1991). Based on this structure techniques of protein modelling can be used to obtained a realistic picture of the three-dimensional organization of DADH.

1.2 The classification of dehydrogenases

The group of enzymes known as alcohol dehydrogenases (E.C. 1.1.1.1), by definition, catalyse the reaction shown below :



But, despite their identical catalytic reaction, alcohol dehydrogenases do not constitute a uniform family of enzymes with generally applicable characteristics. Today they are divided into three main groups, which, according to their monomer sizes, are named: short-chain, medium-chain and long-chain dehydrogenases with monomers of about 250, 380, or 750 amino acids respectively.

The best characterized of the three groups is the medium-chain one, to which the mammalian alcohol dehydrogenases belong. These enzymes form soluble dimers or tetramers, require zinc to carry out their reaction, and prefer primary alcohols to other substrates. They are present in a variety of organisms, from mammals and yeast to plants, being generally involved in the catabolic metabolism of ethanol, often in the form of several isoenzymes displaying tissue and substrate specificity (Brändén et al, 1975). Some members of this family, notably a lens crystallin, lack the metal cofactor or, indeed, any dehydrogenase activity. ξ–Crystallin seems to have evolved towards structural roles but it has kept significant sequence similarity to the rest of enzymes of the group (Borrás et al.1989).

High resolution three-dimensional structures are available for the human and horse liver enzymes, and their catalytic mechanism is very well understood (Eklund & Brändén, 1987). The structure of each monomer has two domains, a substrate-binding domain in the N-terminal half, and a nucleotide-binding domain in their C-terminal half.

Only parts of the nucleotide binding domain show some sequence similarity to the short-chain dehydrogenases, and, as will be described later, their reaction mechanisms are very different. There are few reasons to believe that medium-chain dehydrogenases and short-chain dehydrogenases may have evolved from a common ancestor, and they are frequently pointed to as a classical example of convergent evolution (Brändén & Tooze, 1991).

The long-chain dehydrogenases form a recently defined group of enzymes, and are, unquestionably the least known of the three groups. These enzymes are found in ethanol fermentative bacteria, and show sequence similarity to methanol dehydrogenases from methylotrophic bacteria. They require pyrroloquinoline quinone as a prosthetic group, and form membrane-bound polyenzymatic complexes with respiratory functions (Inoue et al., 1989).

The third group of dehydrogenases, the short-chain ones, were defined as a class of their own for the first time in 1981 (Jörnvall et al., 1981). The first two known members of this family were the *Klebsiella* ribitol dehydrogenase and DADH. A glucose dehydrogenase was soon added to the list of members of the group (Jörnvall et al., 1984), and new related enzymes have been discovered, notably among hormone dehydrogenase enzymes. The list of members of the family as of 1991 can be seen in table 1.

The sequence identity between the different members of the family is generally low, mainly ranging about 25%, and they show a great disparity of substrates. Their sequence similarities, however, are uniform throughout the sequence, their sizes are similar, they all present the same patterns of hydrophilicity and similar secondary structure prediction patterns, none of them uses metal cofactors, and the proposed catalytic residues are conserved throughout the family (Persson et al., 1991).

Enzyme(DH=dehydrogenase)	Species	Reference
Alcohol DH	*D.melanogaster* (I)	Thatcher, 1980
Glucose DH	*B. megaterium* (B)	Jany et al., 1984
Ribitol DH	*E. aerogenes* (B)	Irie et al., 1987
Sorbitol-6P-DH	*E. coli* (B)	Yamada & Saier, 1987
Steroid DH	*Eubacterium sp.*	White et al., 1988
11β-Hydroxysteroid DH	Rat (M)	Agarwal et al., 1989
17β-Hydroxysteroid DH	Human (M)	Peltoketo et al., 1988
20β-Hydroxysteroid DH	*S. hydrogenans* (B)	Marekov, Krook & Jörnvall, 1990
cis-Benzene glycol DH	*P. putida* (B)	Dothie et al., 1985
Dihydrodiol DH	*P.pseudoalcaligenes* (B)	Furukawa, et al.,1987
2,3Dihydroxybenzoate DH	*E. coli* (B)	Liu,et al.1989
1,2Dihydroxybenzoate DH	*P.pseudoalcaligenes* (B)	Neidle et al., 1989
FixR protein	*B. japonicim* (B)	Thöny et al., 1897
NodG protein	*R. meliloti* (B)	Debellé & Sharma, 1986
Acetoacetyl-CoA reductase	*A. eutrophus* (A)	Peoples & Sinskey, 1989
Acetoacetyl-CoA reductase	*Z. ramigera* (B)	Peoples & Sinskey, 1989
Putative oxoacyl reductase	*S. coelicolor* (B)	Hallam,et al. 1988
Adipocyte P27 protein	Mouse (M)	Navre & Ringold, 1988
15-Hydroxyprostaglandin DH	Human (M)	Krook, et al.1990
3β-Hydroxysteroid DH	*P. testosteroni* (B)	Yin et al., 1991

Table 1. List of short-chain dehydrogenases. Adapted from Persson et al., 1991. (I)= insect. (B)= bacteria. (M)= mammal.

1.3 The structure of short-chain dehydrogenases

1.3.1 Structural similarities between members of the family

The disparity of enzymes listed in table 1 immediately suggests a large amount of evolutionary divergence between the different members of the short-chain dehydrogenase family. They are, nonetheless, assumed to have a common structure which, if this common fold assumption is correct, has now been revealed with the solving of the three-dimensional structure of the enzyme 3α,20β–hydroxysteroid dehydrogenase (Ghosh et al., 1991).

Short-chain dehydrogenases form dimers or tetramers of about 250 amino acids per monomer and do not contain metal ions. The members of the family show varying amounts of sequence similarity to each other. In most of the cases sequence identities are about 25 %, but they range between 14% to 54% (Persson et al., 1991). The necessary amount of sequence identity required to be able to assume common folding between polypeptides of more than 80 amino acids has been calculated at about 24.8 % residues (Sander & Schneider, 1991). This value is very close to the one generally observed between short-chain dehydrogenases.

Other common aspects to all short-chain dehydrogenases support the view of a shared structure. The majority of the enzymes of the family catalyse red-ox reactions with the participation of NAD^+, and most of them share a common sequence near the N-terminus of the form GxGxxG (G being glycine and x any amino acid).

ADHs from all *Drosophila* species have the sequence GxGxxG with the exception of *D. lebanonensis* which has the sequence AxGxxG (A being alanine) (Villaroya et al., 1989).

The sequence GxGxxG was initially pointed at as a universal feature of nucleotide binding folds (Argos, 1990). It was also suggested that the sequence GxGxxG was universal for all NAD^+ specific dehydrogenases while those using $NADP^+$ had the sequence GxGxxA (Scrutton et al., 1990). Scrutton and collaborators (1990) proposed that the substitution of the last residue of this sequence to alanine or glycine should be enough to shift the cofactor specificity of any nucleotide binding fold. Their proposal was strongly supported by their experiments on glutathione reductase (Scrutton et al., 1990).

This view, however, has been challenged by the work of Lilley and colleagues (1991), who have discovered the opposite specificity relationship between the last base of the sequence and the cofactors NAD^+ and $NADP^+$ in the enzyme glutamate dehydrogenase of *Clostridium symbiosum*. This enzyme is highly selective for NAD^+ despite having the sequence GxGxxA in its nucleotide binding fold.

Their results, and the sequences of the short-chain dehydrogenases, indicate that the structure and specificity of the poly-glycine loops of nucleotide binding folds do not only depend on the exact sequence and distribution of the glycine residues.

Proteins	residues	sequence
Glyceraldehyde phosphate DH	1-15	SKIGID**G**F**G**RI**G**RLV
Lactate DH	21-35	NKITVV**G**V**G**AV**G**MAC
Horse liver alcohol DH	193-207	STCAVF**G**L**G**GV**G**SVI
Drosophila alcohol DH	8-22	NVIFVA**G**L**G**GI**G**LDT
Glutamate DH	233-247	KTVALA**G**S**G**NV**A**WGA

Table 2. Sequence alignment of the poly-glycine loop of nucleotide-binding domains of some NAD^+-specific dehydrogenases (DH).

The poly-glycine motif was identified as soon as the first sequences of DADH were obtained, and the presence of a nucleotide binding domain at the N-terminus of the enzyme was hypothesized on the basis of its presence and from secondary structure predictions (Thatcher & Sawyer, 1980; Benyajati et al., 1980). At the time a two domain structure for the enzyme was proposed, clearly influenced by the known structure of horse liver alcohol dehydrogenase (Eklund et al., 1976). The presence of an N-terminal nucleotide binding domain and of a C-terminal substrate binding domain was hypothesized. This view has been proven to be wrong with the solving of the $3\alpha,20\beta$-hydroxysteroid dehydrogenase structure.

Another part of sequence shared among all short-chain dehydrogenases is found between positions 150 and 160 (DADH residue numbers will always be used unless otherwise stated otherwise). Tyrosine 152 and lysine 156 are conserved in all the sequences, and have been pointed to as the possible amino acids involved in the proton transfer reaction (Krook et al., 1990). In fact, site-directed mutagenesis of the conserved tyrosine to phenylalanine or alanine render DADH, or 15-hydroxyprostaglandin dehydrogenase respectively, completely inactive (Albalat et al., 1992; Ensor & Tai, 1991).

A common fold for all the members of the family finds further confirmation in the similarity of sequence-based analysis, like hydrophobicity plots and secondary structure predictions (Persson et al., 1991).

Several secondary structure predictions for DADH have been published (Thatcher & Sawyer, 1980; Benyajati et al., 1980; Ribas de Pouplana et al., 1991). They show values of secondary structure typical of a soluble, globular protein, and seem to detect an alternation of helices and strands in the N-terminal part of the enzyme.

Circular dichroism analysis of DADH (Ribas de Pouplana et al., 1991; see results section) gave a value of helical structure of about 28%, and a value of β-strand of about 40%, very similar to the secondary structure contents of horse liver alcohol dehydrogenase (Eklund et al., 1976).

1.3.2 The structure of 3α,20β hydroxysteroid dehydrogenase.(SDH)

The solving of the structure of SDH (Ghosh et al, 1991) is one of the main steps towards an understanding of the family of the short-chain dehydrogenases.

Contrary to the predictions, the tetrameric SDH is composed of four single-domain monomers. Each monomer constitutes a classical α–β alternating mononucleotide binding fold, somehow resembling a cubical structure. Such a cube is composed of three layers of secondary structure : an internal, parallel β–sheet, packed on both sides by α–helices. SDH has both the cofactor and the substrate binding sites in the superior side of the cube, at the C-terminal side of each β-strand (figure 1-I).

Each monomer has two main interaction sites with other monomers. One contact involves the pairing of the C-terminal β–strand of the β–sheet of one monomer to the C-terminal β–strand of another monomer.

These two adjacent β–strands form an antiparallel and unconnected hairpin, practically creating a 14-stranded β–sheet that expands along both monomers. The active sites of each monomer are placed at the opposite faces of the dimer.

If, for the sake of clarity we take the described dimer as a unit, the tetramer is then formed by the lateral assembly of two such units, through the interaction of helices E and F (and their preceding turns) of each monomer with the same helices and turns of an equivalent

monomer from the second unit.(figure 1-II). Of course the real assembly order of the monomers is not necessarily the one used for the description of the quaternary structure of the enzyme.

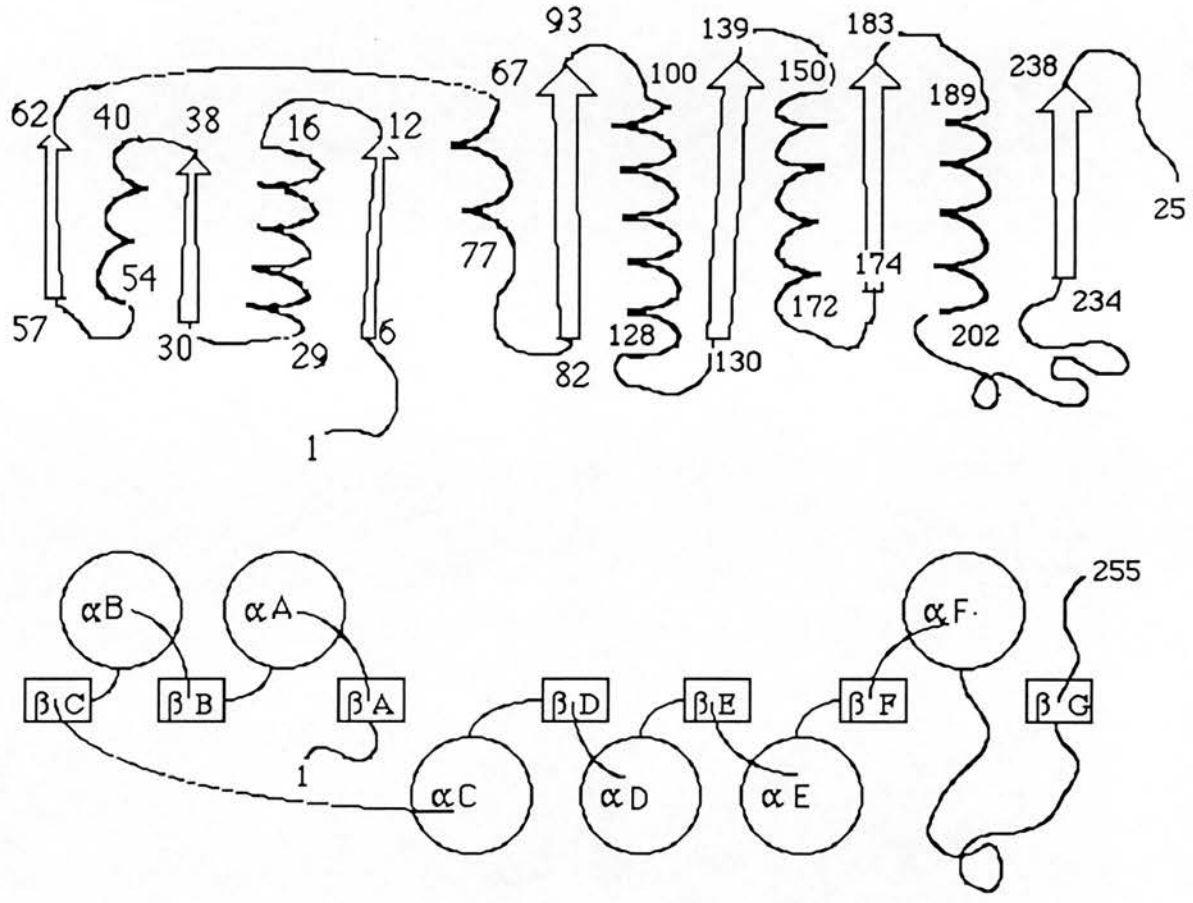


Figure 1-I. Schematic representation of the SDH structure. The top view is perpendicular to the B-sheet. The bottom view looks down the B-sheet from the C-terminal end of the strands.

The putative binding site for the steroid substrate of SDH is placed by Ghosh and coworkers just above strand F, closely in contact with its C-terminal loop (λF'). This binding site would have some structural contributions from the loop λG, from the adjacent subunit. However, this positioning is still hypothetical, since no steroid or substrate-analogue was present in the crystals used for data collection.

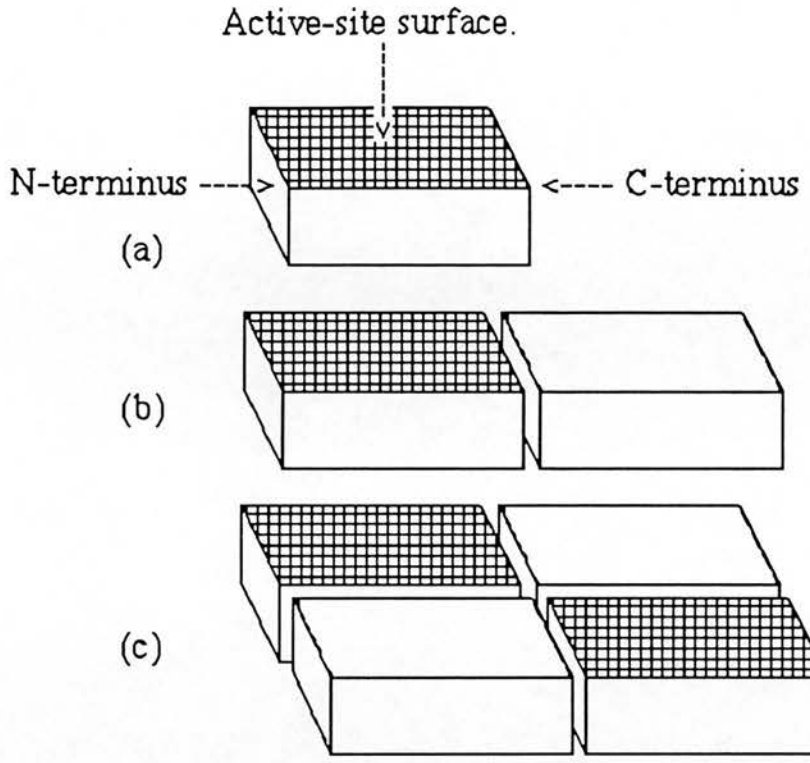


Figure 1-II. Schematic representation of the assembly of the SDH tetramer. (a) monomer. (b) Dimer formation through pairing of the C-terminal strands of their B-sheet. (c) Pairing of dimers along helices-D and -E.

Of more relevance to this project is the architecture of the NAD^+ binding site. The crystals of SDH used for the determination of the enzyme structure had been grown in the presence of 4mM NADH. A region of electron density was assigned by Ghosh et al. to the cofactor structure. This solution locates the cofactor molecule in a totally different position from the one previously found in other dehydrogenases.

The NADH molecule in SDH is bound to a cleft generated by N-terminal ends of α–helices B, C and D; C-terminal ends of β-strands A, B and C and midsections of βD and αE.

The nicotinamide end of the coenzyme is buried, in contact with residues Ile11 to Arg16, Ala36 to Leu39, Ala88 to Ile90 and Ile110, its carboxamide group being at hydrogen bonding distance of Asp 37, a residue conserved in most nucleotide binding folds (Eklund & Brändén, 1987). The ribose ring of the cofactor is close to Ile10, while its adenine moiety is sandwiched between Trp67 and Ile117.

The cofactor is in an extended conformation, 2 Å longer than observed in medium-chain dehydrogenases (Eklund et al., 1984).

An approximate comparison of the relative positioning of the NAD^+cofactor in medium-chain dehydrogenases and in the SDH structure is shown in figure 1-III. The cofactor in SDH is located much more externally than in other nucleotide binding folds.

As a consequence, the residues that in other dehydrogenases make contacts with the adenine (exterior) part of the cofactor in SDH are located around the nicotinamide region, close to the supposed reaction centre.

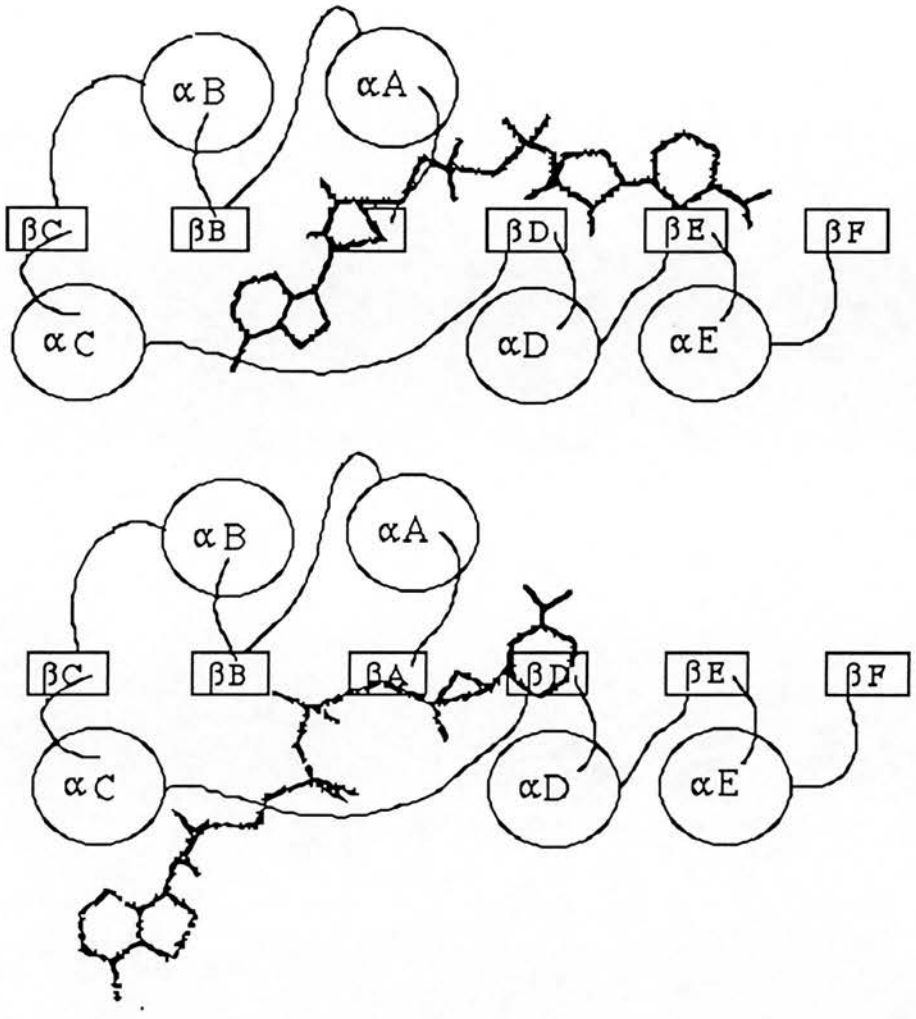


Figure 1-III. Schematic representation of the cofactor positioning in the nucleotide binding fold of horse liver alcohol dehydrogenase (top) and in SDH (bottom).

This result is in apparent contradiction with other published observations from mutagenesis and kinetic experiments (Winberg, 1989; Chen et al., 1991). The poly-glycine loop, at positions 14 to 19, is, as mentioned earlier, a common feature of many nucleotide binding folds. Its role in the structure is meant to be that of allowing a very sharp turn of the polypeptide backbone in that area, thus providing the necessary space for entry of the cofactor into the binding cleft.

The relative positioning of this loop with the cofactor is invariable in all dehydrogenases: always being contiguous to the phosphate chain of the cofactor (Eklund & Brändén, 1987). In SDH the poly-glycine loop is located close to the nicotinamide end, near what would be expected to be the reaction centre of the enzyme.

Another conserved residue in most dehydrogenases (including SDH) is Asp38. Its role in medium-chain dehydrogenases is that of providing a hydrogen bond between the enzyme and the ribose of the adenine moiety of the cofactor. Its mutation in DADH affects the recognition of the cofactor by the enzyme (Chen et al., 1991; this thesis) in a way that suggests that it could play a similar role to that in medium-chain dehydrogenases.

The mutation of Asp38 to Gln changes the enzyme's specificity towards $NADP^+$ (Chen et al., 1991), again suggesting that the residue is in close proximity of the adenine ribose of the coenzyme. In SDH, however, Asp38 is in hydrogen-bonding distance of the carboxamide group of the nicotinamide moiety. It is important to notice that in the SDH structure another negative residue, Glu65, has a relative position with the adenine ribose of the NAD^+ similar to the one of Asp38 in other medium-chain dehydrogenase structures.

Clearly, the positioning of the cofactor in the SDH structure is in contradiction with the function of the mentioned residues in other studied enzymes and with biochemical data obtained for other short-chain dehydrogenases. Although the origin of the identical residues in nucleotide binding folds is still in dispute; whether evolutionary remains of a common enzyme or product of convergent evolution, one would expect their function to be the same, since a common role is the only obvious reason for their appearance in so many different proteins.

A number of possibilities could explain the discrepancies between the positioning of NAD^+ in the SDH structure and the rest of data in the subject.

a) The NAD^+ binding site of SDH has changed in respect to other dehydrogenases to accommodate larger substrates than alcohols.

b) There are more than one possible NAD^+ binding positions in SDH. It is known that the enzyme is capable of recognizing its steroid substrate in two different orientations and catalyses the dehydrogenase reaction with both extremes of the steroid molecule (Ghosh et al., 1991). This suggests a large degree of flexibility in the architecture of the reaction centre. Perhaps such flexibility is in part achieved through differential positioning of the cofactor molecule.

c) Short-chain dehydrogenases have a different binding mechanism for NAD^+ from the rest of dehydrogenases. As mentioned, this is in contradiction with the data on DADH, but this possibility can not be ruled out until structures for other short-chain dehydrogenases are obtained.

d) The cofactor position in the SDH is an artefact of crystallization.

To summarize, the solved structure of SDH probably shows the general fold of all short-chain dehydrogenases. This structure can now be used to model the structures of other members of the family. The binding site for the substrate proposed by Ghosh and colleagues is hypothetical, and in any case, can not be used to make predictions about the substrate binding site of other short-chain dehydrogenases that oxidize very different molecules. The cofactor positioning in the SDH is novel, and in contradiction with biochemical data obtained for DADH. The reasons for this discrepancy are not clear, and can not be analyzed in depth as, unfortunately, Gosh and colleagues have not

released the whole set of published data to the Brookhaven protein data bank for public examination. This unreleased data and the solving of other short-chain dehydrogenases's structures will be required to understand these structural discrepancies.

1.4 The biochemistry of *Drosophila* alcohol dehydrogenase.

1.4.1 The *Adh* gene and its regulation.

The enzyme alcohol dehydrogenase of *Drosophila melanogaster* is encoded by a single gene, located at position 50.1 in the left arm of the second chromosome (Grell et al., 1965). The function of the gene is to produce high levels of ADH (around 1% of total protein) in larvae and adults. The gene was cloned simultaneously in two different laboratories, during the early stages of gene cloning and construction of genetic libraries (Maniatis et al., 1978; Goldberg, 1980). The total length of the gene is 903 base pairs (bp), and it contains two small introns of 65 and 70 bp, respectively.

The first intron is located between codons 32 and 33, the second one between codons 167 and 168 (Benyajati et al., 1980). The *Adh* gene of *D. melanogaster* is unusual in that it has two different promoters, (distal and proximal, with respect to the ATG codon) which regulate the expression of the gene at different stages of development (Benyajati et al., 1983).

In larvae, transcription begins from the proximal promoter, located 70 bp 5' to the ATG. This transcript ends with a short 3' noncoding segment, that includes the polyadenylation signal. In adults, the transcription begins at the distal promoter, located some 700 bp 5' to the proximal promoter, and the transcripts require the splicing of a

large extra intron of 654 bp to produce the mature mRNA (Sofer & Martin, 1987). The final polypeptide is identical in both cases, and the reasons for the different promoters are not clear.

An isolated transcription factor that initiates expression from the distal promoter is implicated in the regulation of, at least, two other *Drosophila* genes (dopadecarboxylase and antennapedia) (England et al., 1990), so it seems likely that the differential use of promoters may be connected to more global changes in gene expression programs during development. The structure of the gene is shown in figure 1-IV.

The research on the molecular regulation of the *Adh* gene expression has unveiled a fascinating and extremely complex mechanism of control. Many questions remain unanswered, and the *Adh* expression system is probably one of the more intensively studied models for eukariotic gene regulation. The description of the system is beyond the scope of the this thesis (for a review see Chambers, 1991).

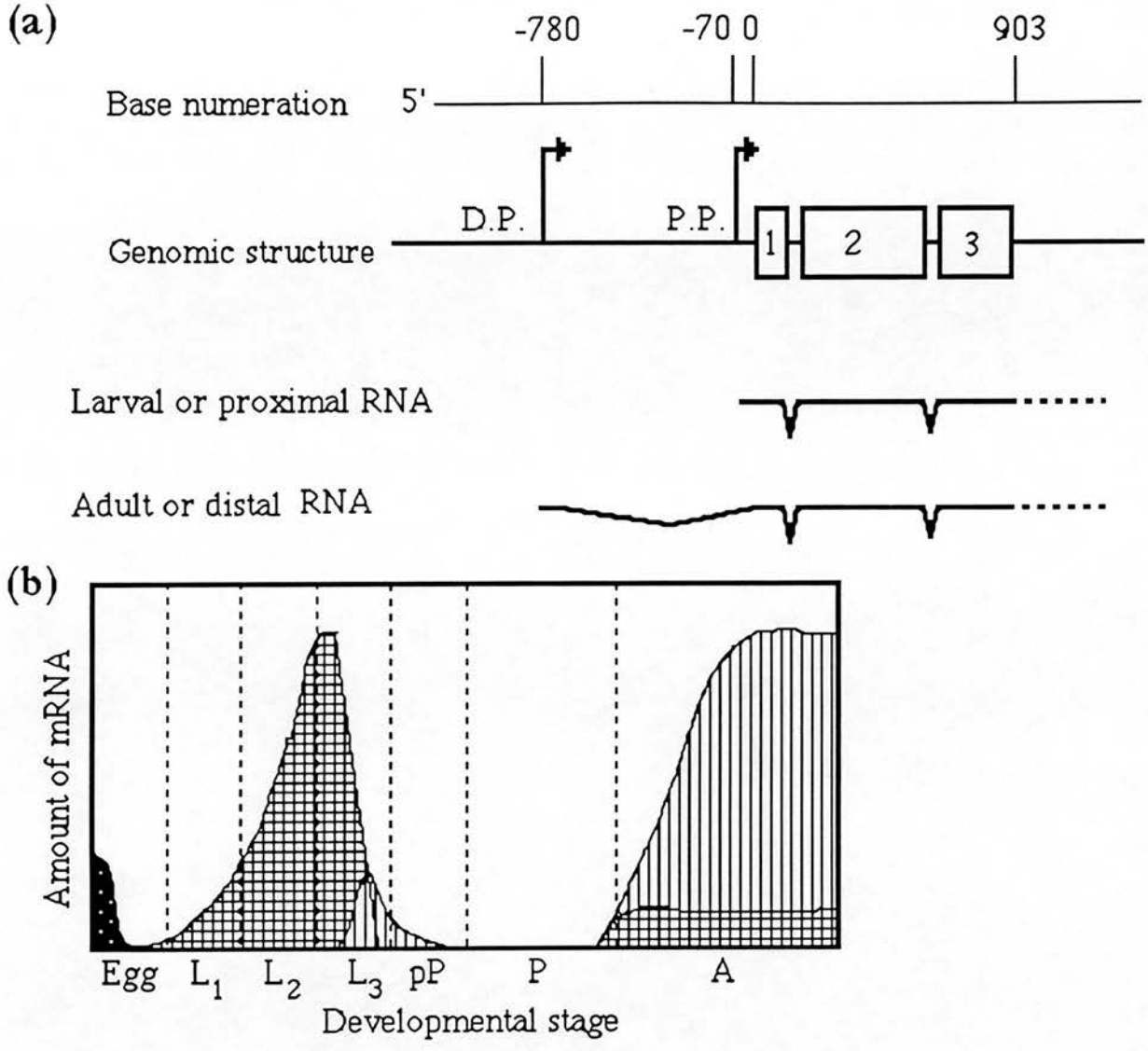


Figure 1-IV. (a) Structure of the genomic region of *Adh* in *D. melanogaster*. The boxes represent the exons of the gene. D.P.= distal promoter, P.P. = proximal promoter. Dashed lines represent the untranslated regions of the mRNA, including poly-A signals. (b) The level of expression of the distal and proximal mRNA's during development. The black area represents maternal mRNA. Hatched areas represent proximal mRNA and dashed ones distal mRNA. (L= different larval stages, pP= prepupae, P=pupae, A=adult)

1.4.2 The *Adh* product, allelic variation.

1.4.2.1 Allelic variation within *D. melanogaster.*

The variation among *Adh* alleles within *D. melanogaster* has been characterized both at the genetic and biochemical levels. The genetic variation has been extensively analyzed by Kreitman and collaborators (Kreitman,1983; McDonald & Kreitman, 1991).

In an attempt to study the mechanisms of genetic evolution, the sequences of eleven *Adh* genes from five different populations of *D. melanogaster* were compared (Kreitman, 1983). Out of 43 genetic polymorphisms only one was a non-silent change, which resulted in the substitution of Lys192 by Thr. This very low ratio of substitutions to silent mutations was interpreted as an indication of a high selective pressure upon mutation. This nucleotide change defines the two main isoenzymes found in natural populations: ADH-F (with Thr192) and ADH-S (with Lys192). Other isoenzymes are known and will be described later on, but their frequency of appearance is very small compared to the one of the F and S isoenzymes.

From the distribution of the polymorphisms Kreitman was able to divide the eleven sequences into two groups. Those groups actually separated the F alleles from the S alleles. Only three different polymorphisms distinguished the two groups, suggesting a recent separation event, but the group of *Adh*-S showed more heterogeneity within itself than the *Adh*-F group. This suggests that *Adh*-S was the ancestor allele. Interestingly a similar two-group classification has come out of the biochemical analysis, which separates the molecules between isoenzymes with Lys and Thr at positions 192.

The biochemical analysis of the melanogaster isoenzymes has been carried out by different groups over the years (Day et al., 1974; Winberg et al., 1982; Chambers et al., 1984). Chambers and collaborators analyzed thirteen different ADH variants from different populations (see table 3). The analysis of the mobility of the isoenzymes in isoelectrofocusing gels gave four different groups of enzymes, namely ultraslow (isoenzyme ADH-US), slow (isoenzymes ADH-S and ADH-Ss), fast (isoenzymes ADH-F, ADH-FChD, ADH-F71K, ADH-Fr, ADH-Fs and ADH-FR.city) and ultrafast (isoenzymes ADH-UF and ADH-D).

Classically, ADH isoenzymes have been classified using this method because it allows a fast determination of the allelic content of many individuals. Hence, for example, ADH-F (for "fast") migrates further in isoelectrofocusing gels than ADH-S (for "slow"), due to the mentioned substitution of a Lys in ADH-S for a Thr in ADH-F, at position 192. However, the method only provides information about the isoelectric point of each variant, and so other biochemical parameters are necessary to compare the functionality of isoenzymes.

Drosophila ADHs are more active with secondary alcohols than with primary ones (Winberg et al., 1982). An activity ratio was defined as the ratio of activity recorded with propan-2-ol to that recorded with ethanol (Gibson et al., 1980). This ratio, plotted against the specific activity with propan-2-ol, was used to compare the thirteen ADH variants.

Isoenzyme	Phenotype compared to ADH-S	Distribution	Changes relative to ADH-S
ADH-US	Ultra-slow mobility.	Africa	not sequenced
ADH-Ss	Heat sensitive	N. America	not sequenced
ADH-S	Standard ADH-S	World-wide	Standard
ADH-F	See text.	World-wide	Lys192 to Thr
ADH-Fr (a)	Heat-stable	N. America	Lys192 to Thr Pro214 to Ser
ADH-FChD	Heat-stable	Australia	Lys192 to Thr Pro214 to Ser
ADH-71K (a)	Heat-stable	Laboratory strain	Lys192 to Thr Pro214 to Ser
ADH-FR.city	Heat-stable	N. America	not sequenced
ADH-F'	Stable enzyme	Africa	Ala 51 to Glu
ADH-UF	Ultra-fast mobility	Spain	Asn8 to Ala Ala45 to Asp

Table 3. Data from Chambers, 1988. (a) Updated in Chambers, 1991. The reference sequence is that of ADH-S (Ashburner, 1985).

The result of the comparison clearly suggested the presence of two different classes of enzymes : an ADH-S-like group, with values of activity ratio between 4.14 and 4.86, and an ADH-F-like group with variants of activity ratio between 6.16 and 7.44. The members of the F group had higher specific activities than the S-like isoenzymes.

In general, those enzymes with Thr192 belonged to the F functional class, although not necessarily to the F-mobility class. Similarly most functionally S-like isoenzymes had Lys192, although their mobilities were not necessarily S-like. This clearly suggested that the residue at position 192 was the determinant of the functional nature of the isoenzyme, while the other changes present in the compared ADHs had little effect on the functionality of the molecule, although they affected the electrophoretic mobility.

The determination of proper kinetic constants with purified enzymes was carried out by Winberg and collaborators (Winberg et al., 1982a, 1985). The isoenzymes ADH-F, ADH-UF and ADH-S and the ADH from *D. lebanonensis* were purified and compared. The main difference between ADH-S and the F-class enzymes is that the former binds NAD^+ more tightly. Winberg and colleagues (1983) proved that the release of the cofactor was the rate-limiting step in the ADH reaction. Hence the higher affinity for NAD^+ showed by ADH-S must be responsible for its lower specific activity, and it suggests a positioning for residue 192 in proximity to the cofactor.

An exception to the two class distribution of isoenzymes is the thermo-stable enzyme ADH-FChD, which belongs to the S-like functional class but has a Thr192, thus having F-like mobility. This enzyme has a specific activity half the way between ADH-F and ADH-S, but an activity ratio very similar to ADH-S. ADH-FChD, in fact, belongs to a whole subgroup of thermo-stable isoenzymes : ADH-FChD, ADH-FR and ADH-71K. These were initially found as isolated polymorphisms in different parts of the world, but the sequencing of the different enzymes has revealed a single change for all of them with respect to the isoenzyme ADH-F: Pro214 to Ser (Chambers, 1991).

These isoenzymes display a marked increase in thermal stability (Chambers, 1984; Hernandez et al., 1986), retaining 75% of the activity after 60 minutes at 44 °C; a temperature at which, after 40 minutes, ADH-F and ADH-S lose all activity (Chambers, 1984). As already mentioned, a comparison of their biochemical properties reveals that these isoenzymes are more similar to ADH-S than to ADH-F, despite the fact of having a threonine at position 192. Clearly the Pro214 to Ser substitution has a double effect: it increases the stability of the enzymes and, at the same time, decreases their specific activity, perhaps increasing the strengths of binding to the cofactor. Interestingly, no ADH-S with a Pro214-Ser mutation has been described. The investigation of this observation is part of the subject of this thesis.

1.4.2.2 Allelic variation of *Adh* /ADH in the genus *Drosophila*

A great deal of research has been dedicated to comparing the characteristics of the *Adh* genes and their products from different species of *Drosophila*, mostly with a view to unravelling the neutral or adaptive nature of their genetic evolution.

The genus *Drosophila* is divided into various radiations, or subgenera, which, in turn, separate into groups and subgroups. The two radiations that have been more intensively studied are *Sophophora* and *Drosophila*, which are thought to have diverged approximately 40-60 million years (Myr) ago (Beverley & Wilson, 1985). Once again, the research on the variation of the *Adh* /ADH system has concentrated on : a) the genetic polymorphisms between species, and b) the biochemical differences in the ADH enzymes.

Many *Drosophila Adh* genes have been cloned and sequenced to date, and every article reporting a new sequence invariably compares it to those previously reported from different species (Kreitman, 1983; Coyne and Kreitman, 1986; Bodmer and Ashburner, 1984; Schaeffer and Aquadro, 1987; Rowan and Dickinson, 1988; Fischer and Maniatis, 1985; Atkinson et al., 1988).

For the sake of clarity I will divide the gene in 5' regulatory sequences, coding region (including introns) and 3' genomic regions. The regulatory sequences at the upstream end of the gene are very different between the members of the subgenera Sophophora and Drosophila. In the former two different promoters exist, as described previously. In the case of some members of the Drosophila radiation, like *D. mulleri* , *D. mojavensis* and *D. hidey*, the system is totally different : the gene has been duplicated and one of the copies has a larval promoter while the other gene has an adult promoter (Fischer and Maniatis, 1985, Atkinson et al., 1988). Other members of this subgenus like *D. affinisdisjuncta*, however, have a two promoters structure, similar to *D. melanogaster* (Rowan & Dickinson, 1988). This variation in promoters organization makes finding an explanation to the different expression mechanisms very difficult, because *D. affinisdisjuncta*, belonging to the subgenus Drosophila, is a much closer species to *D. mulleri* than to *D. melanogaster*, which, confusedly, belongs to the subgenus Sophophora.

In the coding region the general structure of the genes is similar, and the two introns found in the Sophophora group keep the same approximate length and position in all the species analyzed. A striking observation is a different grade of mutability in different parts of the coding region.

This phenomenon was first described by Bodmer and Ashburner, and has been confirmed in other genes. There is a significant difference in substitutive mutation accumulation between the three exons of the gene, with exon three accumulating more mutations than one and two, and number two being the most conserved of the three.

Despite the lack of structural evidence this distribution of mutation rate was speculated to be due to conformational constraints, which were connected to the distribution of functional domains in the enzyme. It was suggested that the cofactor and substrate binding sites would be encoded by exons one and two (Bodmer and Ashburner, 1984; Schaeffer & Aquadro, 1987). More surprising was the finding that the rate of silent mutation was also nonrandomly distributed, showing a similar pattern as the one found for the substitutive mutations. This observation can be explained by unknown mutational constraints in supposed silent sites or by a change of mutational rates across the gene (Schaeffer and Aquadro, 1987).

The overall amino acid variability encoded by the different genes does not exceed 30% even in very distant species like *D. affinisdisjuncta* and *D. melanogaster* (79% identical). This is not surprising when we consider that all the comparisons are made within a single genus.

The downstream region of the gene had been shown by Kreitman (Kreitman,1983) to have a high degree of conservation. During their analysis of this part of the *D. pseudoobscura* gene, Schaeffer and Aquadro (1987) discovered an open reading frame, which had a high sequence identity with the *Adh* gene. This was proposed to be the result of an ancient duplication of the gene. Such gene duplications have been detected at the 3' end other *Adh* genes (Atkinson et al., 1988, Marfany and Gonzàlez-Duarte, 1991).

Their transcription to mRNA has been proven but their translation product, if any, has not yet been identified (Schaeffer & Aquadro, 1988).

The biochemical characterization of the ADH enzymes from different *Drosophila* species has been the subject of intense research, in an attempt to find a correlation between the observed genetic variations and functional differences, either in the enzyme or the metabolism of the whole fly . The different species of *Drosophila* studied (mainly by the group of Gonzàlez-Duarte : Juan & Gonzàlez-Duarte, 1980; Atrian & Gonzàlez-Duarte, 1982; Atrian & Gonzàlez-Duarte, 1985; Winberg et al., 1986; Hernández et al., 1988) occupy very different ecological environments. The main physiological variable concerning the ADH enzyme is , of course, the amount of alcohol present in the media, and most of the studies attempted to find correlations between the activity of the ADH enzyme of each species with its resistance to alcohols, and the amount of them that the flies are likely to encounter in the wild.

Large variations of ADH activity can be found in crude extracts of different flies, and these, generally, correlate well with the kind of environment where the flies are found. In most cases the species more resistant to alcohol ingestion were the ones with higher ADH activity in crude extracts (Atrian & Gonzàlez-Duarte, 1982). However, when the enzymes were purified and their kinetic characteristics obtained they were found to be quite similar to each other in terms of affinity for substrate and cofactor, and in specific activity (Atrian & Gonzàlez-Duarte, 1985).

A variation was found in the amount of enzyme that could be recovered from each species, suggesting that the difference in ethanol tolerance depends on the amount of active enzyme in the fly more than on the specific activity of the enzyme.

This could be due to increased expression or a lower turn-over, which, in turn, could be a function of the enzyme's sequence-induced stability. The possible structural effects of the characterized polymorphisms will be discussed in 1.3.5.

Although the amount of ADH activity per fly, their resistance to ethanol and their presence or absence in alcohol-rich environments correlate quite well, there are some exceptions. *D. simulans*, a sibling species of *D. melanogaster* that has a somewhat smaller amount of ADH activity to *D. melanogaster*, completely avoids the alcoholic environments where *D. melanogaster* thrives. Possibly, this indicates that ecological behaviour is a function of a large variety of factors, not necessarily limited to ADH metabolism.

The kinetic parameters for many different pure *Drosophila* ADH enzymes have been obtained. In general, and despite sequence differences of up to 21 %, as in the case of *D. lebanonensis* and *D. melanogaster*, their kinetic values are similar (Winberg et al., 1986; Atrian & Gonzàlez-Duarte, 1985). Enzymes purified from flies, bacteria (Chen et al., 1990), or yeast (see results section), show the same catalytic uniformity.

The Km for NAD^+ is around 200 μM for most species, Km's for propan-2-ol are in the order of 1 mM and the ones for ethanol about 8 mM. k_{cat} values for isopropanol are around 7/sec, and for ethanol around 4/sec. The variability in catalytic activity is also small, and ADH-S and ADH-F differ from each other by a maximum factor of three. A certain amount of variability in these values can be found in the literature, probably the more reliable measures are the ones by Winberg and collaborators (Winberg et al; 1982a, 1985, 1986, 1988a) because the authors developed and used a method for the quantification of enzyme active-sites in solution.

It is clear that the amount of polymorphism encountered in different *Drosophila* ADH isoenzymes does not result in large kinetic variation. The only reported kinetic value that significantly differs from the rest is the Km for isopropanol of the ADH from *D. affinisdisjuncta* (Green et al., 1989), which is about two orders of magnitude higher than the mean value of the rest of species. This enzyme, expressed and purified from bacteria, has a Km value for NAD^+ of the same order as in other species, and a similar specific activity. Unfortunately, the kinetic constants for ethanol are not reported by the authors.

It seems clear from the studies at the genetic level that there is a strong evolutionary pressure against changes in the genome that result in amino acid variation. At the same time, most of the polymorphisms at the protein level seem to have little effect on the enzyme's catalytic mechanism. Also, Middleton and Kacser have reported that the variations in ADH activity due to genetic polymorphisms do not cause variation in the metabolic flux of alcohol utilization (Middleton & Kacser, 1983). This seems to be in good agreement with Kimura's neutral evolution theory, where most changes would be deleterious, and the majority of the non-deleterious would be neutral (Kimura, 1983).

In contrast to this fact, McDonald & Kreitman (1991), with a simple analysis of the ratios of amino acid substitutions to silent mutations between and within species, conclude that such ratios can not be explained by a neutral theory of evolution. They find many more amino acid replacements compared to silent mutations between species than within single species. If neutral evolution was the main cause for sequence variation those rates would be expected to be equal both between and within related species.

If there is an effective selective pressure behind the sequence variation in ADH it does not seems to be working solely on the catalytic properties of the enzyme. As mentioned before, there may be many possible factors involved in the suitability of a specific gene or protein sequence. If the amount of ADH is important for the metabolism of the fly, as the results described earlier suggest, then the stability of the enzyme and its levels of expression may be more relevant factors than catalytic alcohol dehydrogenase activity. Perhaps the differences found in stability between the ADHs of different species are showing a selection trend (Atrian & Gonzàlez-Duarte, 1985).

The amount of gene expression is related to the codon usage in each species (Grantham et al., 1981), and a difference in third-codon position nucleotide content has been reported between the subgroups *Drosophila* and *Sophophora* (Starmer & Sullivan, 1989). This difference in third-base frequency seems to be specific for the *Adh* gene, and it could mean that the evolutionary pressure concerning *Adh* is applied at the expression level by altering the third-base frequency.

The third-base nucleotide content between the three species studied by McDonald & Kreitman (1991) are similar, as they used members of the *Sophophora* group only. This means that their definition of silent mutations is correct. But codon usage values will have to be taken into consideration when defining the concept of silent genetic change if comparisons between *Drosophila* groups are to be made.

Finally, it has to be mentioned that the metabolic role of *Drosophila* ADH is far from well understood, having even been suggested that the enzyme could be involved in pheromone metabolism (Winberg et al., 1982b) (see 1.4) . Clearly, no explanation for the observed non-neutral evolution can be offered until the role of ADH is clear.

1.5 The enzymology and structure of Drosophila ADH

As mentioned previously, most of the enzymological research on *Drosophila* ADH has been carried out by Winberg and colleagues (see Winberg et al., 1982a, 1982b, 1983, 1984, 1985, 1986, 1988, 1988b, 1989). Their analysis of the catalytic mechanism of the isoenzymes ADH-F, ADH-S and ADH-UF of *D. melanogaster* and the ADH enzyme of *D. lebanonensis* offers a large amount of information concerning the architecture and mechanism of the active site, the nature of the residues involved in catalysis, and the differences between isoenzymes.

1.5.1 The mechanism and kinetics of the cofactor binding and reduction

The nucleotide-binding site of *Drosophila* ADH is supposed to be very similar to the one found in medium-chain dehydrogenases. As seen in section 1.2.2, this is confirmed by the structure of 3α–20β SDH.

It was found that the reaction catalysed by *Drosophila* ADHs follows a compulsory-order mechanism, where the cofactor binds first, followed by the alcohol (Winberg et al., 1985, 1986, 1988a). In the case of secondary alcohols the dissociation of the enzyme-NADH complex (ER) becomes the rate-limiting step of the reaction. This is due to the high binding affinity between the enzyme and NADH, which makes the release of the reduced coenzyme a slower process than the proton transfer. As a consequence, the strength of binding of the enzyme to the coenzyme is directly related to the kinetic efficiency of the reaction, and the isoenzymes that show strongest binding to NAD^+ have the lowest specific activities.

Similar to horse-liver ADH (LADH), the analysis of the pH dependence of the cofactor-enzyme complex formation showed that the enzyme-NAD^+ complex (EO) was inhibited by the protonation of two groups with pKs of 6.5 and around 9, respectively.

The ER complex, on the other hand, was only dependant on the protonation of a group with a pk of around 8. (Winberg & McKinley-McKee, 1988a).

This implies that a group with a pk of 6.5 is responsible for the inhibition of the formation of the EO complex but not of the ER complex, hence this group must interact with the nicotinamide end of the cofactor. This situation is similar in LADH, where residue His51 binds to the oxygens of the nicotinamide ribose and is suggested to repel the oxidized cofactor when positively charged.

The ADHs of *D. melanogaster* have four histidines, two of which can be chemically modified (Thatcher, 1981), but inactivation results from the modification of only one of them (Retzios, 1982), and can be protected by the cofactors NAD^+ and NADH.

Another residue that affects coenzyme binding is Lys or Thr192. As described before, position 192 is the only amino acid difference found between ADH-F and ADH-S. These isoenzymes differ in the strength with which they can bind NAD^+, with ADH-S (Lys192) having a stronger interaction than ADH-F (Thr-192).

If thio-NAD^+ is used, the difference is abolished and the binding in ADH-S takes ADH-F-like values, suggesting that Lys192 interacts with the nicotinamide part of the cofactor, possibly binding to the carboxamide group. Lys192 could be the residue that, in ADH-S, shows a pk of about 9.2 affecting the coenzyme recognition. The pH dependence of this process has not been determined in the Thr192 containing enzymes (Winberg, 1989).

Drosophila ADH contains two cysteines and no disulfide bridges, and these cysteines have been pointed at as possible catalytic residues. Their chemical modification was shown to abolish enzyme activity, while the presence of the coenzyme protected them against chemical attack (Thatcher, 1977, 1981; Chambers, 1984).

However, the substitution of both cysteines to alanines did not affect the enzymes activity (Chen et al., 1990), proving that they are not involved in the reaction mechanism. Nevertheless, they might be in contact with the coenzyme and the substrate, which would explain the observation that both substrate and cofactor protect Cys218 against chemical modification. Interestingly, the binding of substrate and coenzyme seems to make Cys135 accessible to chemical attack (Chen et al., 1990).

The investigation of the pH dependence of the ER and EO dissociation process showed the existence of a group, with a pk of 7.6 whose protonation increases EO dissociation. Protonation of the same residue also unstabilises the binding of alcohols and ethanol competitive inhibitors to the EO complex (Winberg et al., 1988). Since this group does not affect acetaldehyde binding or NADH dissociation in the pH range 6 to 10, it must have a pk higher than 10 when in the ER complex. Winberg and collaborators compared this behaviour to the pk's of the zinc atom in medium-chain enzymes, which are 9 in the free enzyme, 7.6 in the EO complex and 11.2 in the ER complex (Winberg et al., 1988; Petterson, 1987). This residue has not been determined in Drosophila, although Tyr152 seems to be the a probable candidate, in view of its total conservation and because its substitution abolishes all enzyme activity (Krook et al., 1990; Ensor & Tai, 1991; Albalat et al., 1992).

Other residues are known to affect the binding of the coenzyme to the enzyme. As described earlier isoenzymes where the residue Pro214 is changed to Ser show increased thermal stability and a reduction in specific activity that makes them more similar to ADH-S despite the fact that they contain Thr at position 192. This effect is due to an increase in the binding energy to NAD^+, which suggests that this mutation occurs in the proximity of the cofactor binding site or somehow modifies it.

At positions 1 and 82 respectively, all *D. melanogaster* isoenzymes have Ser and Gln, while ADH from *D. simulans* has Ala and Lys. The binding of the *D. simulans* ADH to the coenzyme is stronger than in ADH-S from *D. melanogaster*, and it results in an even smaller catalytic activity (Bodmer & Ashburner, 1984; Winberg, 1989).

This suggests that residue 82 is involved in the binding of the cofactor (since residue number 1 is in an extremely variable region of the enzyme, and is very likely to have a disordered structure). A possible role of this residue is suggested by a computer model for the structure of *Drosophila* ADH (see modelling section of this thesis).

Asp38 is a conserved residue in most nucleotide-binding folds where it hydrogen bonds to one of the hydroxyl groups of the adenine ribose of the cofactor (see 1.2.2). As discussed earlier, in SDH this residue seems to have a different relative positioning in respect of the coenzyme, and a glutamic acid at position 65 seems to be closer to the adenine ribose of the cofactor. In the light of a multiple sequence alignment it has also been suggested that another aspartate residue (Asp64) maybe the conserved aspartate in the nucleotide binding folds of short-chain dehydrogenases (Persson et al., 1991).

However these observations are in contradiction with experimental results that sustain that Asp38 of *D. melanogaster* ADH interacts with the cofactor in a similar way to the seen in medium-chain dehydrogenases (discussed in 1.2.2)(Chen et al., 1991).

Site-directed mutagenesis experiments have demonstrated the anticipated proximity of the poly-glycine sequence G-X-G-X-X-G, at positions 14-19 to the cofactor (Chen et al., 1990; Results section of this thesis). The mutation of Gly14 to alanine reduces the enzymes activity to 30%. Its mutation to valine totally inactivates the enzyme (Chen et al., 1990; Results section of this thesis). This is in perfect agreement with the structural role assumed for this part of the sequence, that is, to provide an empty space to allow the accommodation of the cofactor inside the binding-site. As predicted the bigger the group that is introduced into the sequence the larger the effect on the enzyme's mechanism, presumably because of increasing distortion of the nucleotide recognition. The mutation of Gly19 to Ala, also affects the structure of the nucleotide-binding site (Results section of this thesis).

1.5.2 The recognition of the substrate and kinetics of its oxidation

All *Drosophila* ADHs are more efficient oxidizers of secondary alcohols like propan-2-ol than of primary alcohols like ethanol (Winberg et al., 1982). This specificity indicates that the substrate binding sites of *Drosophila* ADHs are probably very different to the medium-chain structures, which recognize primary alcohols more efficiently.

The substrate specificities, and the activity ratios (see 1.3.1), of *D. melanogaster* ADH-F, ADH-S and ADH-UF and *D. lebanonensis* ADH are very similar (Winberg et al., 1982, 1985). The substitution of the non-transferable hydrogen of ethanol by a methyl group in propan-2-ol increases the k_{cat} value of the enzymes by a factor of 20 (Winberg et al., 1982a; Chambers, 1984; Atrian & González-Duarte, 1985).

Through the comparison of the specific activity of ADH-S and ADH-UF with a large range of primary and secondary alcohols, Winberg and collaborators (Winberg et al., 1982b) concluded that the enzyme binds secondary alcohols through a double hydrophobic interaction, which, in the case of propan-2-ol, affects the two methyl groups of the alcohol. This is supported by the fact that the enzymes are not active with charged or polar alcohols (Winberg et al., 1984, 1986). This is different to the substrate-binding mechanism of horse-liver ADH, where the methyl group of the ethanol is bound to a single hydrophobic funnel in the active site (Eklund & Brändën, 1987).

Winberg (1989) assigns a difference on binding energy of -1.6 to -1.8 kCal/mol between primary and secondary alcohols. Using the calculated values for the free-energy of transfer of a methyl group from water to n-octanol (-0.68 kCal/mol) he proposes that the second methyl group of propan-2-ol binds to a region of the enzyme that is about 2.5 times more hydrophobic than n-octanol.

The specificity of the active site is, however, extremely flexible and the enzymes manage to oxidize alcohols of very different structures and sizes, suggesting a very open conformation of the substrate binding-site. In contrast, the enzymes show a marked preference for the R-enantiomers of the substrates (Winberg et al., 1982a).

The difference in substrate specificity between *Drosophila* ADHs and the medium-chain ADHs makes clear that the structures of the binding-sites are different between the two classes of enzymes, and so it is not possible to extrapolate crystallographic data from the latter to the *Drosophila* enzymes. At the same time the similarities in substrate recognition between the ADHs of different *Drosophila* species do not allow the identification of possible residues involved in the binding site.

1.5.3 The reaction mechanism

The dehydrogenase reaction is started with the binding of the cofactor molecule to the enzyme. The reasons for this order are not clear, but the observation that coenzyme binding may induce conformational changes (Chen et al., 1990) may indicate that structural shifts induced by the cofactor may be necessary for substrate recognition, as it is found in medium-chain ADHs. Once the substrate is bound the reaction proceeds, supposedly with the participation of Tyr152 or lys156. The speed of the reaction is much lower if primary alcohols are being metabolized, and hydride transfer becomes a rate-limiting process (Winberg et al., 1986).

If secondary alcohols are used as substrate, the rate-limiting step becomes the dissociation of the reduced coenzyme from the enzyme. This has been proved by:a) the similar Vm shown for different secondary alcohols, b) the substrate activation with some secondary alcohols, c) the lack of primary kinetic isotope effects when deuterated secondary alcohols were used (Winberg et al., 1984, 1985, 1986).

In the reverse reaction, the maximum velocity for acetaldehyde reduction was found to be the dissociation rate of NAD^+ from the EO complex, while with acetone a ternary complex interconversion seemed to be rate limiting (Winberg & McKinley-McKee, 1988).

This reaction mechanism seems to be general to all the *Drosophila* ADHs tested so far, as they all show similar substrate specificity and consistently similar Vmax values for the secondary alcohols tested (Thatcher, 1977; Chambers, 1984; Juan & González-Duarte, 1980; Atrian & González-Duarte, 1985).

In contrast to this well documented mechanism it must be mentioned that Heinstra and collaborators (Heinstra et al., 1988) have reported that NADH can act as a competitive inhibitor with respect to ethanol as a variable substrate, this observation would suggest a rapid equilibrium random mechanism for the oxidation of primary alcohols. This appears to be an isolated observation.

1.6 The physiological role of Drosophila ADH

The main physiological role of *Drosophila* ADH is the dietary utilization of alcohols. In addition, a function of detoxification when high levels of alcohols are present in the environment is also clear, although the experimental results are better understood if the processes of ethanol utilization and of ethanol resistance are treated as different ones (Chambers, 1991). Other functionalities, like a role in pheromone metabolism, have been proposed (Winberg et al., 1982: Madhavan et al., 1973), but they have not been further investigated.

1.6.1 Ethanol as a nutrient.

From experiments with *Adh* null strains, and population studies of flies starved in ethanol rich environments it is clear that ethanol can be used as a nutritional source (Parsons, 1981; Parsons & Spence, 1981). The metabolic pathway of this utilization is not well understood, although it is known that the ingestion of ethanol produces an accumulation of triacylglycerides in the fat body. Acetate is supposed to be an intermediate in this process (Geer et al., 1985).

ADH activity seems to be the main control point of this pathway, having a control coefficient of nearly one according to ^{13}NMR studies on the ethanol --> lipid flux (Freriksen et al., 1991). This is in contrast with the observation that flies with different *Adh* genotypes and different global ADH activities do not show differences on the flux of the alcohol metabolizing pathway (Middleton & Kacser, 1983). A second enzymatic activity, oxidizing acetaldehyde to acetate, has been shown for *Drosophila* ADH (Brooks et al., 1985), although this activity has too high optimum pH (>10) to be significant *in vivo* (Chambers, 1988).

The main question surrounding the role of ADH in the utilization of ethanol is whether there is an adaptive connexion between the amount of ADH activity that one fly obtains from its ADH isoenzymes and its fitness in the population. The fact that ADH is such an abundant protein in the fly's body (about 1% of total protein), and that *D. melanogaster* and other species are mostly associated with fermenting habitats makes it unavoidable that a connexion between ADH activity and ecological or evolutionary fitness be searched.

When comparisons between different species are made, a correlation between amount of ADH activity and presence in ethanol-rich niches seems to be well stablished (Oakeshott et al., 1982). However, when the comparisons are made within *D.melanogaster*'s different allelic populations results are not so clear, and ADH-F flies, which contain more ADH activity do not always outperform ADH-S flies when grown in ethanol-rich environments (Oakeshott et al., 1984).

This last observation would agree with the previously mentioned results of Middleton and Kacser and suggest that the differences observed between different species may be the result of other, not taken into account, differences between the species compared. The net contribution of the ADH enzyme to the ecological and physiological relationship between *Drosophila* and ethanol is not totally defined yet.

1.6.2 Alcohols as poisons.

Once again, a large amount of contradictory data has accumulated concerning the relationship between ADH activity and the tolerance to ethanol or other alcohols. Experiments involving inbred lines of flies or laboratory strains generally show that ADH-F strains have an advantage over ADH-S strains when grown in ethanol rich media (> 6% ethanol) (Maroni et al., 1982). However, experiments using freshly captured flies fail to demonstrate any increase of the ADH-F alleles when grown in the same conditions (Gibson & Wilks, 1988). At high concentrations of ethanol a toxic build up of acetaldehyde occurs, and so the capacity of resisting high concentrations of ethanol will depend on the capacity to digest ethanol and, also, on the capacity to remove the accumulating acetaldehyde (Anderson & Barnett, 1991).

Probably, a more important alcohol in terms of toxicity is propan-2-ol, whose oxidation product, acetone, is highly toxic to the fly (Geer et al., 1985). When flies are grown in the presence of propan-2-ol the ADH activity is rapidly suppressed. This is due to the covalent attachment of NADH and acetone to the enzyme, creating an inert tertiary complex.

When *Drosophila* ADH is analyzed in isoelectrofocusing gels three forms are observed with different mobilities: Adh-1, Adh-3 and Adh-5, corresponding to dimers with two, one or nil NAD^+ molecules bound, respectively (Johnson & Denniston, 1964). In normal conditions, the Adh-5 population represents about 90% of the total amount of enzyme (Winberg, 1989), but the amount of Adh-3 and Adh-1 can be increased if the enzyme is incubated with a mixture of NAD^+ and acetone due to the formation of this ternary complex (Jacobsen et al., 1972).

The ternary complex is irreversible, because *In vivo* the recovery of ADH activity after incubation with propan-2-ol is totally due to *de novo* synthesis of enzyme (González-Duarte & Atrian, 1986).

Propan-2-ol can be an abundant alcohol in the organic environments where *Drosophila* larvae prefer to grow, and it has been proposed that the inactivation of the ADH activity is in fact a protection mechanism against an excessive build-up of acetone in the larvae (Chambers, 1988). An $NADP^+$-dependent aldo-keto reductase activity has been reported in *D. melanogaster* and *D. hidey* , it is independent of ADH activity and it has been proposed to have a detoxifying role, converting aldehydes and ketones back to alcohols, a role that ADH cannot perform due to the formation of the inert ternary complex (Atrian & Gonzàlez-Duarte, 1985).

1.6.3 Other environmental factors affecting ADH activity and distribution

Adh alleles are not randomly distributed, and the different polymorphisms follow a clinal geographic distribution. The search for the factors that determine such distribution is intense, because these may be the same factors that determine the actual existence of the polymorphisms (Vigue & Johnson, 1973). The frequency of the allele *Adh*-F increases with latitude both north and south of the equator, while *Adh*-S behaves in an opposite way (Oakeshott et al., 1984; Vigue & Johnson, 1983).

This allelic distribution is quite constant in all continents (Chambers, 1988), despite some variations like the very high frequency of ADH-FChD in China (~ 30% of wild populations), which is thought to be due to the fact that China is the area were the mutation first appeared.

Which are the factors that govern this distribution?. Temperature is probably the main candidate. A correlation between altitude and increase in ADH-F frequency has been found in mountains of Mexico and Russia (Grossman et al., 1970; Pipkin et al., 1976). ADH-F, which seems to be more common in cooler locations is reportedly less thermo-stable than ADH-S, an allele which attains maximum frequencies in the tropics.

As mentioned earlier, ADH-F is believed to have appeared from ADH-S (Kreitman, 1983), and, from the clinal data, it would appear to have colonized areas of lower temperatures more efficiently. Perhaps flies with an F genotype tend to occupy niches where S genotypes no longer have the advantage of thermal stability.

Against the clinal observations we must cite the report of Sampsell & Barnette (1985), which shows that ADH-F flies grown at 29 °C have more enzyme activity than if grown at 20 °C. And, at both temperatures, ADH-F flies have more ADH activity than ADH-S flies.

Also, a clinal analysis of the distribution of *Adh* alleles in South Africa failed to show a latitudinal distribution, but found a positive correlation between *Adh*-S distribution and increase in rainfall (Getz & Grant, 1990). The authors try to explain such correlation with a possible increase of amount of alcohols in zones of higher humidity, due to lower evaporation. In such conditions they argue that a more active enzyme like ADH-F would be selected against, because of the larger amount of toxic intermediates that would accumulate in the flies with such genotype.

Once again there seems to be a general trend that correlates *Adh* allelic distribution with environmental factors (latitude, temperature), but this correlation is confronted with contradictory reports that reminds us that the factors involved in metabolic and ecological processes are far to complicated to be explained by a single parameter.

The alcohol dehydrogenase gene/enzyme system, because of its well described polymorphism and distribution, and its central role on the metabolism of *Drosophila*, seemed a good candidate for obtaining simple and general answers to many evolutionary and ecological questions. As usual, this has proved not to be the case.

1.7 Aims and strategy of the project

This project was originally planned as part of a collaborative effort between the Department of Genetics of the University of Barcelona and the Department of Biochemistry of the University of Edinburgh. The overall goal remains the elucidation of the structure of *Drosophila* ADH, the analysis of the structure-function relationships in the enzyme and the study of the effect of the introduction of induced polymorphisms into *Drosophila* populations. The project described here aimed to investigate the role of residues of interest in the protein by site-directed mutagenesis with the use of various biochemical and biophysical characterization techniques.

Specifically, two main aspects of the enzyme structure were to be investigated :

a) The existence of the proposed nucleotide-binding site at the N-terminus of the protein was to be investigated through mutagenesis of the residues Gly14, Gly19 and Asp38.

b) The thermal stability effect of the mutation Pro214 to Ser was to be investigated, first by introducing the mutation into an ADH-S isoenzyme, where it has not been found in natural populations.

To carry out these experiments the *Adh*-S gene was cloned into a plasmid where all the mutagenesis and sequence analysis were performed. The mutated genes where then subcloned again into a yeast expression vector, and the mutated enzymes purified to homogeneity from yeast cultures.

The mutant enzymes were then analyzed with a variety of techniques in order to characterize the effect of the introduced changes.

During the process of this research it became clear that the structure of Drosophila ADH would not be readily available due to the low quality of the crystals obtained from the enzyme. Since a three-dimensional structure is essential for the proper understanding of site-directed mutagenesis experiments, an alternative three-dimensional model of the enzyme was necessary.

The recent publication of the structure of 3α,20β–hydroxysteroid dehydrogenase, and the release of its α-carbon coordinates in the Brookhaven database, provided the necessary material to build such a three-dimensional model for Drosophila ADH, using the new techniques of protein modelling. Such a model offers the unusual possibility of analyzing the structural evolution of an enzyme for which a large number sequences with a relatively small divergence time are available.

2. Experimental Materials and Methods.

2.1 Materials.

2.1.1 Cloning, expression, mutagenesis and sequencing materials.

2.1.1.1 Plasmids.

1. pUC115-3008, from Dr. W. Sofer, Waksman Institute, Rutgers University, USA. Vector pUC118, containing the *Drosophila melanogaster Adh*-S gene,

2. M13mp18, from Bethesda Research Laboratories (UK) Ltd, Cambridge, UK. M13mp18*E. coli* bacteriophage DNA. Used for mutagenesis and sequencing procedures.

3. PVT-U-101, from Dr. S. Atrian, Dept. of Genetics, Barcelona University, Catalonia. Phagemid, capable of replicating in yeast, or in bacteria as a bacteriophage. It contains the expression promoter of the yeast ADH-1 gene, for protein expression in this organism.

2.1.1.2 Strains.

1. *Escherichia coli* TG1. K12, A(lac-pro), SupE, thi, hsdDS/F'traD36, pro A^+B^+, lac I^q, lacZ ΔM15. From Amersham plc., Aylesbury, UK.

2. *Escherichia coli* BW313. dut, ung, thi-1, relA, spoT1/F'lysA. From Dr. T. McNally, Biochemistry Dept., Leeds University.

3. *Sacharomyces cerevisiae* WV36.201. Mat a, ura 3, trp1, ade2, ade4, adh1Δ, adh2Δ, adh3. From Dr. S. Atrian.

2.1.1.3 Enzymes

BamHI, EcoRI, HindIII, PstI, XbaI, SacI and SphI were from Northumbria Biologicals Ltd., Cramlington, UK.

SpeI was from New England Biolabs, Cambridge, USA.

T4 DNA ligase and Rnase A were from Boehringer Mannheim, Lewes, UK.

2.1.1.4 Media, antibiotics and related

Bacto peptone, purified agar and yeast nitrogen base w/o amino acids were from Difco labs, East Molesly, UK.

Tryptone and yeast extract were from Oxoid Ltd., Haverhill, UK.

Ampicillin, Xgal, IPTG and all amino acids and nucleotides were from Sigma Chemical Co., Poole, UK.

2.1.1.5 Oligonucleotides and radiochemicals

All oligonucleotides were synthesized by and purchased from Oswel DNA service, Dept. of Chemistry, University of Edinburgh, except the M13 sequencing oligonucleotides, which were from Amersham.

Deoxyadenosine 5'-α- (^{35}S) thiotriphosphate, triethylammonium salt, was from Amersham.

2.1.1.6 Kits.

All sequencing reactions were performed using the Sequenase™ DNA sequencing kit, provided by Cambridge Biosciences, Cambridge, UK.

2.1.1.7 Hardware.

DNA agarose electrophoresis was performed in a Pharmacia (Pharmacia, Uppsala, Sweden) gel electrophoresis apparatus GNA-100, connected to a Sandos (Sandos, Stockport, UK) power unit.

DNA sequencing was done on a Pharmacia sequencing cast connected to a Pharmacia power unit.

2.1.1.8 Chemicals

All chemicals were supplied by Sigma, BDH (BDH, Poole, UK) or IBI (IBI, New Haven, USA), and were of the best grade commercially available.

2.1.1.9 Miscellaneous

Hyperfilm MP X-ray film was from Amersham. Polaroid 665 film and agarose were from Sigma. Plastic tubes (Falcon, 5 and 50 ml) were from A&J Beveridge (Beveridge, Edinburgh, UK).

2.1.2 Protein purification and analysis materials.

2.1.2.1 Chemicals.

Sepharose G-25, Blue Sepharose, Q-Sepharose FF, S-Sepharose FF, Superose 6 and Superose 12, either unpacked or prepacked, were from Pharmacia.

Ammonium sulphate, dichloroisocumarin, 1,10 phenanthroline, E-64, TEMED, ammonium persulfate, SDS, DTT, acrylamide, N,N'-methylene-bis-acrylamide and Trizma base were from Sigma. Methanol, H_2O_2 and 4-chloro-1-naphtol were from Merck, Schuchardt, Germany.

2.1.2.2 hardware

Protein electrophoresis in SDS-polyacrylamide gels was done in either Bio-Rad miniprotean systems (Bio-Rad, Hemel Hempstead, UK) or in Hoeffer (Hoeffer, San Francisco, USA) SE 280 gel electrophoresis units.

Protein blotting to nitrocellulose membrane was done in a Bio-Rad mini trans-blot electrophoretic transfer cell using nitrocellulose filters (Schleicher & Schuell, Dassel, Germany).

All Chromatographic steps were performed using :

- Gilson minipuls3 peristaltic pump
- Pharmacia-LKB redifrac fraction collector.
- Isco UA-5 UV detector and chart recorder.
- Pharmacia FPLC system LCC-500.

2.1.3 Enzyme characterization materials.

2.1.3.1 Enzymes.

Horse liver alcohol dehydrogenase was from Boehringer Mannheim. yeast alcohol dehydrogenase was from Sigma. Alcohol dehydrogenase from *Drosophila melanogaster* and *Drosophila lebanonensis* were purified from adult flies by Drs Atrian and González-Duarte as published (Ribas de Pouplana et al., 1991).

2.1.3.2 Chemicals.

Guanidinium chloride (Aristar grade), propan-2-ol and NAD^+ (both AnalaR grade) were from BDH.

NADH (grade I), Trizma base and kanamycin were from Sigma.

2.1.3.3 Hardware

Circular dichroism spectra were recorded using a Jasco J-600 spectropolarimeter.

Spectrophotometric assays were performed using a Phillips PU 8720 UV/VIS scanning spectrophotometer.

Temperature-controlled incubations and reactions were carried out using Grant instruments Ltd. (Cambridge, UK) thermostatted water-baths.

EX BIBL. UNIV. EDINBURGENSIS

2.2 Experimental methods.

Although various alternative methods were tried at various points of this project only the ones that were used for the final experiments will be described here. The results obtained with the rest of methods used will be discussed in the experimental results section.

2.2.1 Standard DNA and protein techniques.

The following DNA and protein techniques were performed according to standard protocols, as described in "Guide to Molecular Cloning" 2nd edition (Sambrook et al., 1989) :

Technique	Page n°
-Enzymatic cleavage of double stranded DNA	1.85
-Ligation of DNA fragments	1.68-1.69
-Growth and storage of *E. coli*	A.5
-Preparation and transformation of competent *E. coli*	1.76-1.84
-Purification of plasmids and phage DNA from *E.coli* .	1.25-1.28
-Protein SDS-polyacrylamide gel electrophoresis and Coomassie blue staining.	18.47-8.48, 18.55

2.2.2 Site-directed mutagenesis.

The relevance of the principle of site-directed mutagenesis (SDM) to the bio-molecular sciences is evidenced by the large number of methods that have been designed to perform it. All these methods share a common initial step where a single-stranded DNA (ssDNA) template containing the sequence to mutagenize is annealed to a complementary oligonucleotide whose sequence has the desired modification. This oligonucleotide is then used as a primer for the polymerization of the complementary strand of the template, giving way to a hybrid double-stranded DNA molecule, one strand of which contains the desired mutation. The crucial step is the selection of the mutated strand over the wild type one. This is achieved by different strategies in each method (for a review see Carter, 1986).

The method used in this project is that designed by Kunkel (Kunkel, 1985), which is based on the utilization of the *E. coli* strain BW313, a strain deficient for the *Dut* and *Ung* genes. These genes code for the enzymes dUTPase and uracil glycosidase, which are utilized by the cell to digest uracil-containing DNA and to modify uracil to stop its incorporation into replicating DNA. The lack of these enzymes creates a bacterium that contains high levels of uracil in all its cellular DNA molecules. As a consequence ssDNA isolated from such a strain will be unable to survive in a normal cell, because it will be digested by the uracil glycosidase. If such ssDNA is used as a template for annealing with a mutation-containing oligonucleotide, and a second, mutant strand is obtained by polymerization with normal nucleotides, a hybrid molecule is obtained. One, containing the wild type sequence and a high number of uracil bases, and another containing the desired mutation and normal base content.

If such double-stranded DNA is used to transform a *Ung*$^+$ strain of *E. coli* only the mutation-containing strand should survive since the uracil containing DNA will be digested by uracil glycosidase.

This method provides an economic, efficient and elegant way of introducing specific mutations into a DNA sequence.

Procedure :

M13mp18 ssDNA containing the intronless *Adh*-S sequence (hereinafter '*Adh* gene') was purified from phages isolated from a single phage plaque in *E. coli* BW313 infected plates. This DNA, supposedly containing large quantities of uracil bases was tested for such quality by transforming *E. coli* TG1 cells with it. An infection efficiency of less than ten plaques per plate was taken as an indication of a satisfactory presence of uracil in the ssDNA.

Two µg of uracil-rich ssDNA were mixed with 2 µl of KGBx10 buffer and the phosphorylated oligonucleotide containing the desired mutation. The molar ratio between ssDNA template and oligonucleotide was typically 0.2 to 0.5, water was added to a final volume of 16 µl. This mixture was incubated at 65°C for 3 min and allowed to cool down to room temperature, for at least one hour, to achieve optimum annealing. The polymerization reaction was started by adding to the mixture 2µl of a solution containing dATP, dTTP, dCTP and dGTP, at a concentration of 7.5 µM each, 1 µl of Klenow fragment of DNA polymerase at a concentration of 5 units/µl, and 1 µl of DNA ligase at a concentration of 10 units/µl.

The polymerization and ligation reactions were allowed to proceed for 30 min at room temperature. The presence of dsDNA was checked by agarose gel electrophoresis and, if this check was satisfactory, the DNA from the reaction was separated from the enzymes by a phenol extraction followed by ethanol precipitation. The final DNA sample obtained was used to transform *E. coli* TG1 cells. From the bacteriophage plaques obtained three were randomly chosen, and their ssDNA purified and sequenced to confirm the presence of the desired mutation.

2.2.3 DNA subcloning.

Using standard DNA subcloning techniques the *Adh* gene contained in the vector pUC-*Adh*-S was removed from this vector and subcloned into M13mp18 for the mutagenesis experiments.

The *Adh* genes from the M13 clones were then subcloned into PVT-U-101, for further introduction and expression in yeast using convenient restriction sites (see results chapter for a description of the cloning steps and vectors used).

2.2.4 DNA sequencing.

DNA sequencing was used at three different stages of the site-directed mutagenesis process:

First, the *Adh* gene in the plasmid pUC-*Adh*- S, provided by Dr. Sofer, was totally sequenced to confirm its expected sequence. Secondly, the ssDNA from the phage plaques obtained at the last stage of the mutagenesis process were sequenced to confirm the presence of the desired mutations. Finally, once each mutated gene was subcloned

into the expression vector PVT-U-101 and transformed into yeast for expression, the plasmid was recovered and the *Adh* gene fully sequenced, to ensure that the introduced mutation was the only difference between the mutant gene and the original *Adh*-S.

Procedures :

ssDNA sequencing of *Adh* in either M13 or PVT vectors was done following Sanger's dideoxynucleotide method, which was used as supplied in kit form (Sequenase™ DNA sequencing kit, from Cambridge Bioscience). Sequencing gels were run on a Pharmacia gel cast, fixed, dried on 3MM Whatman paper and autoradiographed using Hyperfilm MP.

During the sequencing of the *Adh* gene present in the pUC-ADH vector the following steps were performed in order to obtain ssDNA:

The dsDNA pUC-*Adh* was purified from 100 ml of liquid culture of *E. coli* TG1 cells by CsCl gradient ultracentrifugation. 10µg of the DNA were denatured by adding NaOH and EDTA to a final concentration of 0.2M and 0.2mM, respectively. The solution was then neutralized by addition of ammonium acetate to a final concentration of 0.2M. The DNA was precipitated with ethanol, resuspended in 30µl of ice-cold water and used immediately for standard sequencing.

2.2.5 Growth and transformation of yeast .

The yeast strain *S. cerevisiae* WV.36.201, was used for the expression of the *Adh* gene as published (Atrian et al., 1990). This strain is auxotrophic for tryptophan, adenine and uracil, and it can be grown either in rich non-selective media as YEPD (1% yeast extract, 2% bactopeptone, 2% glucose) or in selective media supplemented with tryptophan, adenine and uracil, known as YD (0.67% yeast nitrogen base, 2% glucose, 1% casein amino acids, 0.002% tryptophan, 0.002% adenine, 0.002% uracil).

The expression plasmid used throughout this project, PVT-U-101 carries s functional *Ura* gene, so WV.36.201 cells carrying the plasmid can be detected by their ability to grow in selective media without uracil.

All the growth curves carried out with the strains expressing different ADH variants were done as follows :

Single colonies were inoculated into small cultures (~5 ml) of selective media and grown overnight at 30 °C in an orbital shaker. After 24 hours, the optical density at 600 nm (OD_{600}) was measured for every culture. Different volumes of each small culture were transferred to 50 ml cultures of either selective or non-selective media, to obtain initial inoculates with approximately equal number of cells.

The growth of each culture was monitored by measuring the OD_{600} of each culture at regular intervals. When necessary the cultures were diluted before measuring their optical density to ensure that OD values stayed on the 0.1 to 1.5 OD units range.

Transformation of yeast :

The method described by Ito and colleagues (Ito et al., 1983) was used. This is based on the treatment of yeast cells in exponential growth state with alkali cations, in our case lithium acetate.

This method has a lower efficiency than that of Beggs (Beggs, 1978), which used enzymatic digestion of the yeast cell wall with lyticase, but it is simpler, faster and, for the purposes of this thesis, the most convenient.

Procedure :

A 100 ml culture of WV.36.201 strain was grown up to a OD_{600} of 0.3 and the cells were harvested with a 2 min centrifugation at 2000 rpm in a bench centrifuge. The collected cells were washed in 5 ml of TE buffer (10 mM Tris-HCl pH 8, 1mM EDTA), centrifuged again and resuspended in 5 ml of 0.3 M lithium acetate.

The cells were incubated in this solution for an hour at room temperature in an orbital shaker, then centrifuged again, resupended once more in 5 ml 0.3M lithium acetate and divided into 0.3 ml aliquots. Between 2 and 5 µg of DNA were added to each aliquot together with 10µg of fragmented genomic DNA (from herring sperm). The aliquots were then left at room temperature for 30 min before 0.7 ml of a 50% solution of polyethylene glycol in TE buffer were added. The mixtures were homogenized by inversions and incubated at 30°C in a water bath for 30 min. The cells were then heat-shocked at 42 °C for 5 min, centrifuged with a 10 second pulse in a bench microfuge and resuspended in YEPD media. After 1 or 2 hours at 30 °C, the cells were again centrifuged, resuspended in 200 µl of water, plated onto plates containing YD without uracil, and left incubating at 30 °C. Colonies appeared in 4-5 days.

2.2.6 Analysis of yeast transformants and DNA isolation from yeast .

The yeast colonies that, after transformation, grew on a uracil deficient selective medium had to be checked for the presence of the expression construct, and this construct had to be isolated for the sequencing of its *Adh* gene.

In practice, an intermediate step was used where the plasmid contained in the transformed yeast was first transferred into *E. coli,* from which it was later purified and analyzed. This technique, known as plasmid rescue, provides the simpler way of checking for the presence of a certain plasmid in yeast , provided such plasmid is capable of replicating in bacteria and has a functional selection marker in this organisms (Hoffman & Winston, 1987).

Procedure :

Yeast colonies that grew in uracil deficient YD were inoculated in 5 ml liquid cultures of the same media and incubated at 30 °C for 24-48 hours. Cells from 1 ml aliquots of culture were then collected by a 5 second centrifugation in a bench microfuge, and resuspended in 200 µl of a solution containing : 2% Triton X-100, 1% SDS, 100 mM NaCl, 10 mM Tris pH 8.0, 1 mM EDTA.

To this solution 200 µl of phenol were added, and, immediately, glass beads were added up to the meniscus of the solution surface. This mixture was vortexed for 30 seconds and the tubes centrifuged in a bench microfuge for 5 min. The objectives of this methods are: a) to break the yeast cells and liberate their possible plasmid content and, b) the separation of the DNA from the cellular protein by a phenol separation.

The aqueous phase obtained after the centrifugation was collected, and its DNA content precipitated in ethanol and resuspended in a small volume of water. This was then used to transform *E. coli* TG1 cells, and the presence of the plasmid was checked by testing the resistance to ampicillin of the transformed bacteria.

The bacterial clones that showed resistance to ampicillin were isolated, and their plasmid DNA isolated by standard methods. This DNA could be used to check the presence of the correct construct by restriction analysis in agarose gels or to confirm the presence of the right gene sequence by DNA sequencing techniques.

2.2.7 The analysis of *Adh* expression in yeast .

The analysis of the expression of any gene can be undertaken at three different levels. First, the content of mRNA being transcribed can be quantified. This analysis gives a good measure of the amount of expression activity at the gene level, and is the method of choice when the effect of regulatory sequences on the amount of gene expression is being studied. However, it does not necessarily give information on the amount of protein actually being produced, because of the effect of the codon content of the mRNA on its level of translation efficiency, and it does not provide any information on the levels of active protein.

The amounts of total protein being synthesized from mRNA can be checked by immunoblotting analysis of the protein content of the cells. This method requires the availability of antibodies against the expressed protein, but it can give very accurate accounts of the amounts of protein being produced. The method will not discriminate between inactive or active forms of the protein, and so it can not be used to

check the protein activity levels in the expressing cells if, for instance, the expressed polypeptide is an enzyme.

The quantification of the amount of protein activity in the expressing cells can be achieved *in vivo*, if the activity of the expressed molecule affects the metabolism of the expressing organism, or *in vitro*, either in crude extracts or after the purification of the expression product, if a suitable activity assay is available.

For the analysis of the expression of wild type and mutant DADHs in yeast three different approaches were used. The amount of total protein produced was analyzed by immunoblotting, the expressing cells were analyzed in terms of their growth behaviour (see 2.2.5) and the amount of DADH activity was measured in cell crude extracts.

2.2.7.1 Immunoblotting analysis of DADH expression in yeast .

A series of mouse monoclonal and polyclonal antibodies raised against *Drosophila* ADH by Dr. J. Fibla (Fibla et al., 1989), of the Genetics Department of the University of Barcelona were used for the immunoblotting analysis.

Procedure :

Yeast cells containing the PVT-U-101 expression vectors with the wild type and mutant *Adh* genes were grown in 50 ml cultures of YD media without uracil. Yeast cells without plasmid, and cells with the PVT-U-101 plasmid without the *Adh* gene were used as controls.

When each culture reached stationary growth the cells were harvested by centrifugation at 5000 rpm, for 15 min at 4°C. The cell pellet obtained was resuspended in an equivalent volume of ice-cold lysis buffer (20 mM Tris HCl pH 8.0, 10 mM $MgCl_2$, 1 mM EDTA, 5% glycerol, 1 mM DTT, 0.3 M ammonium sulphate, 100 µM

dichloroisocumarin, 100 μM 1,10 phenathroline and 20 μM E-64). Cells were lysed by filling the solution with glass beads up to the surface meniscus and vortexing for 6 min in two min intervals, separated by 2 min periods where the extract was cooled in ice. The resulting crude extract was centrifuged at 10000 rpm for 30 min to remove all insoluble material and cellular debris, and the supernatant was kept for further analysis.

The total protein content of the crude extracts was determined using the Bradford method (Bradford, 1976), and equal amounts of total protein from each different strain analyzed were loaded onto a SDS-polyacrylamide gel, for standard electrophoresis separation.

After electrophoresis the gels were transferred to a renaturizing buffer (80% 50 mM Tris HCl pH 7.6, 20% glycerol) as a previous step to the electrotransfer of the protein bands to a nitrocellulose filter.

The bands were transferred to the filter with a Bio-Rad mini trans-blot electrophoretic transfer cell, for 30 min. After that period the filter was removed form the transfer apparatus for immunological detection.

The protein-free areas of the filter were first saturated by incubating the filter in a 10% solution of freeze-dried milk in phosphate buffer saline (PBS, 0.8% NaCl, 0.02% KCl, 0.11% Na_2HPO_4, 0.02% KH_2PO_4), for 2 hours at room temperature, this step is designed to stop non specific binding of the first antibody to the nitrocellulose fibres of the filter.

Excess saturating protein was removed by washing the filter in PBS in three 15-min periods. The filter was then incubated with the anti-ADH polyclonal or monoclonal antibodies at adequate dilutions for 1 hour. The excess antibody was removed by washing the filter in a PBS solution containing 0.5% freeze-dried milk in three 5-min steps.

A second antibody, raised against mouse IgG antibodies and complexed to the enzyme peroxidase, was then added to the filter at a convenient dilution, for a period of 1 hour. The excess antibody was again removed by washing the filter with PBS containing 0.5% freeze-dried milk. The filter was then developed by adding two solutions simultaneously (solution 1: 50 ml of PBS containing 0.035% hydrogen peroxide. Solution 2: 50 ml of a 60% PBS, 40% Methanol, 0.001% 4-chloro-1-naphthol). Bands containing material recognized by the peroxidase-complexed antibody appeared as a result of the oxidation of 4-chloro-1. naphtol, which gives a coloured compound.

2.2.7.2 DADH activity measurements in crude extracts.

Crude extracts were prepared as described in the previous chapter and their protein content also quantified using the method of Bradford (Bradford, 1976), or the method of Layne (Layne, 1957). This last method uses the differential absorbance of proteins and DNA at 280 and 260 nm. The concentration of protein is given by the equation :

$$[\text{protein}] = 1.55x\ (A_{280}) - 0.76x(A_{260})$$

Although this method is simpler and faster than the Bradford analysis it has the disadvantage of not being able to discriminate between protein absorbance at 280 nm or other aromatic compounds that might be present in the crude extract. Once the total protein content of the crude extracts was known, three different volumes of the protein solution (typically 1, 5 and 10 µl) were assayed for enzyme activity. The ADH activity was determined at 25 °C by monitoring the increase in optical absorbance at 340 nm due to the production of NADH. The assays were performed in 1 ml cuvettes, the assay mix contained 1 mM NAD, 10 mM isopropanol and 100 mM Tris-HCl pH 8.7.

The crude extract sample was added first to the cuvette, and 1 ml of reaction mix was pipetted inside the cuvette to rapidly mix the components of the reaction. The reactions were linear for at least two min, and one unit of enzyme was defined as the amount of enzyme that catalyses an increment of 1 unit of absorbance at 340 nm per minute.

The activity of each sample and its protein concentration were used to determine the specific activity of each crude extract, and the total amount of ADH activity recovered from each different strain.

2.2.8 Protein purification methods.

Purification of DADH from either adult flies or third-instar larvae have been reported for years, and a considerable amount of expertise exists. However, the purification of the enzyme from yeast had not been attempted before this project.

For this reason, a series of different chromatographic steps were tested trying to take advantage of the already existing expertise in the purification of the enzyme from flies. However, the final purification procedure from yeast is quite different. Only the techniques that were finally used will be described here (a more detailed account of the results given by the other techniques tested is given in the results section), but it is worth mentioning that one of the most effective purification steps when dealing with flies or larvae, the affinity column Blue-Sepharose (Ribas de Pouplana et al., 1991), was shown to be less effective in the purification of the enzyme from yeast . Another set-back of this technique, of course, was that very large behaviour differences would have been expected when trying to purify enzymes with mutations in the nucleotide binding site.

2.2.8.1 Crude extract preparation and Ammonium sulphate fractionation.

For the purpose of the purification of DADH large cultures with high cell density were required. These were obtained by growing cells from a single colony in a pre-culture of 50 ml of YD minus uracil, for 24 to 48 hours. These cultures were then transferred to 750 ml cultures of YEPD media, were they were allowed to grow at 30 °C, under vigorous shaking, until stationary phase was reached.

Similarly as with small crude extract preparation cells were harvested by 15 min centrifugation at 5000 rpm, resuspended in an equal volume of the lysis buffer and the cells were lysed with glass beads as described in 2.2.7.1 .

The ammonium sulphate concentration of the crude extract was then raised to 35% saturation by slowly adding the necessary amount of the salt while the extract was in ice and in moderate agitation. The salt concentration was allowed to equilibrate for 30 min and then the crude extract was centrifuged at 35000 rpm, for 20 min at 4 °C, to remove all cellular debris and proteins that precipitated at 35% saturated ammonium sulphate or less. In a similar fashion, ammonium sulphate was again slowly added to the supernatant solution to a final 70% ammonium sulphate saturation, and after a 30 min equilibration in ice it was centrifuged again in the same conditions described. The amount of ADH activity left in the supernatant was measured and, if found to be less than 10%, the supernatant was discarded.

2.2.8.2 Sepharose G-25 and Q-Sepharose FF steps.

The pellet obtained at the last step of the ammonium sulphate fractionation was resuspended in a 25 mM Tris HCl pH 8.0 and the solution (typically about 5 ml) loaded onto a Sepharose G-25 column (20x400 mm), with a flow rate of 1 ml/min. Fractions containing ADH activity were pooled (typically a total volume of ~ 12 ml) and passed through a previously equilibrated Q-Sepharose FF column (15x60 mm) with a flow rate of 60 ml/hour. Both chromatographic steps were performed at 4 °C.

Fractions containing DADH activity did not stick to this column and were directly recovered, identified and pooled (typically about 15 ml total volume).

2.2.8.3 Sample concentration and Superose-12 steps.

The volume typically collected from the Q-Sepharose FF step (~15 ml) was too large to be applied to the Superose-12 column available in the Department (25 ml), so the sample had to be reduced. This was achieved by saturating the sample with ammonium sulphate and centrifuging the precipitated protein at 35000 rpm, for 30 min at 4 °C. The resulting pellet was then resuspended in a small volume (~1 ml) of 25 mM Tris HCl pH 8.0 buffer.

The solution was injected onto the FPLC Superose-12 column in steps of 200 µl. The column was run at a flow rate of 0.2 ml/min, at room temperature. Fractions of 500 µl were collected ant assayed for ADH activity. Those with such activity were kept and analyzed for purity in a SDS-polyacrylamide electrophoresis gel stained with Coomassie blue.

2.2.9 Characterization methods for pure DADHs.

Once the different ADH enzymes had been purified they had to be analyzed in order to investigate the effects of the introduced mutations.

For this purpose three main analytical procedures were used : a) the stability of the enzymes and the strength of their binding to the coenzyme were analyzed by guanidine hydrochloride denaturation kinetics monitored by circular dichroism and enzyme activity analysis, this allowed the calculation of the $\Delta\Delta G$ values of the cofactor binding process, b) the thermal stability of the molecules was analyzed by incubating them at different temperatures and monitoring the loss of activity with time, and c) the kinetic parameters Km_{app}, Vmax and k_{cat} were determined both for isopropanol and for NAD.

2.2.9.1 Guanidine hydrochloride denaturation studies.

The study of the denaturation kinetics of enzymes is an effective way of quantifying the difference in free energy between the unfolded and the folded states of the molecules, and obtaining information about their structural stability (for a review see Pace, 1986). The denaturation agents normally used are either temperature or chemical agents (mainly urea and guanidine hydrochloride (GdnHCl)); and the systems used to monitor the denaturation process are mainly optical systems (fluorescence, circular dichroism (CD), UV difference spectroscopy, optical rotatory dispersion), nuclear magnetic resonance, differential calorimetry or biological activity measurements.

In this study the denaturation process of the wild type and mutant DADH's in the presence of GdnHCl was monitored both by CD analysis and by ADH activity measurements. Also, the effect of the binding of NAD^+ on the enzyme's stability was studied by incubating the enzyme with the cofactor prior to denaturation.

The denaturation kinetics of the enzymes monitored by ADH activity measurements were used to determine the free energy values of the folded enzymes in solution.

2.2.9.1.1 Circular dichroism analysis.

Circular dichroism activity (CD) is defined as the differential absorption of left- and right-circularly polarized light by any molecule in solution. This difference in absorption is directly dependent on the structural asymmetry of the molecule, and so a molecule with a random conformation in solution will show little CD activity, since no average asymmetry will be present.

Molecules with high structural asymmetry and rigid conformations will have measurable amounts of CD activity at certain wavelengths, but will lose such property if their structure becomes random. Because CD measurements are, in fact, optical absorbance measures, only the regions of the light spectra where a molecule shows optical absorbance will be useful to determine the CD activity of such molecule.

These characteristics of CD make it a highly informative way of analyzing the conformation of proteins in solution. Protein structures have two main regions of optical absorbance: the polypeptide backbone of proteins in solution absorbs and is optically active in the far UV region, below approximately 240 nm.

The aromatic residues of polypeptides, on the other hand, have high optical absorbance properties in the near UV region, around 280 nm. Far UV CD activity of proteins is, hence, a measure of their backbone asymmetry, which is a direct function of the amount of secondary structure of the molecule. If a protein is subjected to denaturation the fixed conformation of its structure disappears, and so does its CD activity. In this way, CD values can be used to monitor the denaturation of a certain polypeptide in solution.

Alpha-helical structures are much more asymmetrical than extended or random regions and so they are the main contributors to the far UV spectra of proteins (for reviews on CD techniques see Johnson, 1988 and Johnson, 1990).

GdnHCl is an excellent denaturing agent for such studies because it does not absorb far UV radiation and, consequently, does not interfere with the CD measurements (Pace, 1986)

Procedure :

ADH-F from *D. melanogaster*, ADH from *D. lebanonensis*, and ADH's from horse-liver and yeast were used for the CD analysis. All enzymes were diluted in the same buffer, and their concentrations had been determined by amino acid analysis. GdnHCl solutions were prepared from fresh GdnHCl, and their concentrations were checked by refractive-index measurements (Nozaki, 1972). All CD analyses were performed at the Scottish CD facility, Department of Molecular and Biological Sciences, University of Stirling, which is run by S. M. Kelly and Dr. N. C. Price.

Firstly, the necessary incubation time to reach denaturation equilibrium was determined for each enzyme using a low (0.5 M) concentration of GdnHCl. All incubations were carried out at 25 °C, for the amount of time found to be required to reach equilibrium.

A fixed volume of enzyme was added to a microcentrifuge tube, and a fixed volume of appropriately diluted GdnHCl was added to obtain the desired final GdnHCl concentration. This mixture was left at 25 °C for the period required to reach equilibrium and the CD spectrum was then obtained, using cell pathlengths of 0.02 or 0.1 cm, depending on the enzyme concentration. The samples where the cofactor binding effect was being measured were treated equally, but for the addition of NAD^+ to a final 220 μM concentration 5 min prior to the addition of the denaturant. CD spectra were recorded over the range 250-190 nm. Molar ellipticity values (CD activity/mol of enzyme) were calculated by using a value of 108 for the mean residue weight (Thatcher, 1980; Jörnvall, 1970; Jörnvall et al., 1975).

The secondary structure contents were derived by the CONTIN procedure (Provencher & Glöckner, 1981).

2.2.9.1.2 Denaturation monitored by ADH activity measurements.

The monitoring of the biological activity of an enzyme can be a very sensitive way of detecting structural changes in the protein's structure and of determining its denaturation kinetics. As it would be expected this method of characterization is much more sensitive to subtle effects of small structural modifications than methods like CD or fluorescence spectroscopy. However, the results from this kind of analysis can be difficult to interpret because the adding of protein denaturants can seriously alter the solution conditions, hence altering the catalytic constants of the enzymes being measured. Also, as the level of denaturant increases and the proportion of folded enzyme decreases changes in the relative amounts of substrates to enzyme can occur, causing byphasic effects in the denaturation process (Shrake & Ross,

1990). This last problem can be overcome by ensuring that the substrate concentrations are at saturating levels at all times.

Procedure :

The denaturation process and enzyme activity determination for each denaturant concentration point were done in duplicate, and the analyses of the denaturation in the presence or absence of NAD^+ were done simultaneously. The GdnHCl, isopropanol, NAD^+ and buffer stock solutions used were always the same for all the analyses carried out in a single day to minimize variation.

As a previous step to the development of the assay the following controls were performed :

a) The time required for the denaturation process to reach equilibrium at a low concentration of denaturant was determined and used in all analyses. It was found that 15 min at 25 °C was enough to reach such equilibrium at 0.3 M GdnHCl (figure 2.2.9-I).

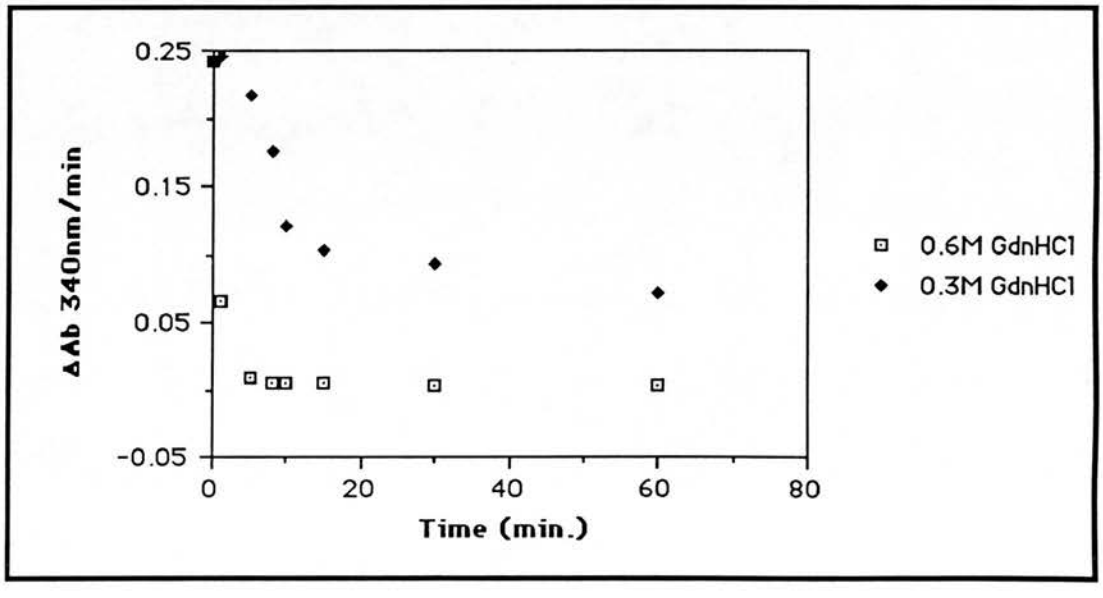


Figure 2.2.9-I. Evolution of ADH activity with time of incubation in GdnHCl.

b) The reversibility of the denaturation process was analyzed by incubating DADH in 0.6M GdnHCl for 15 min and then diluting the sample 5 fold in buffer containing 0M, 0.2M, 0.4M and 0.6M GdnHCl. ADH activity was increasingly recovered with the dilution of the denaturant, and with time, showing the reversibility of the denaturation process (figure 2.2.9-II).

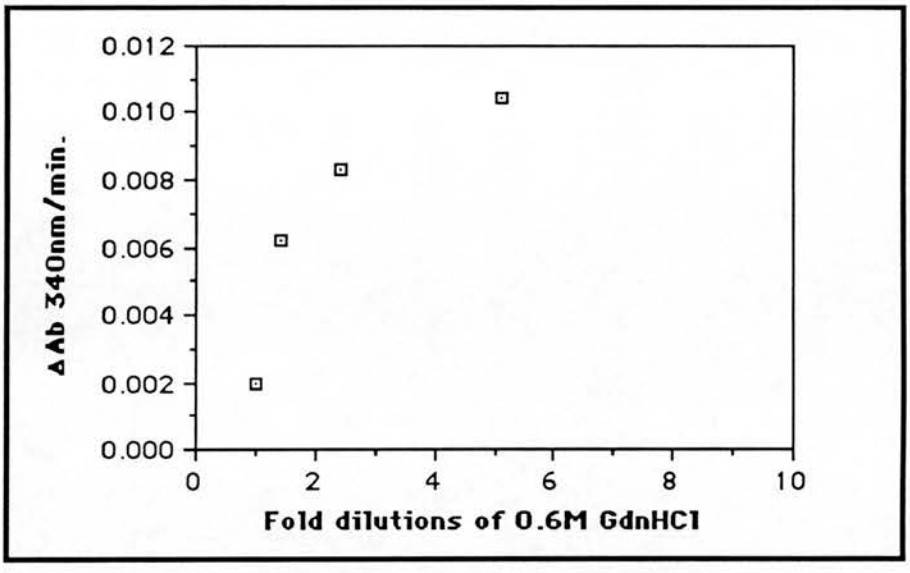


Figure 2.2.9-II. Recovery of ADH activity upon dilution .

---

c) The additive effect of the coenzyme in protecting DADH against denaturation was checked by incubating the enzyme with increasing amounts of cofactor. It was shown that the protection provided by the coenzyme increased with the coenzymes concentration. (figure 2.2.9-III)

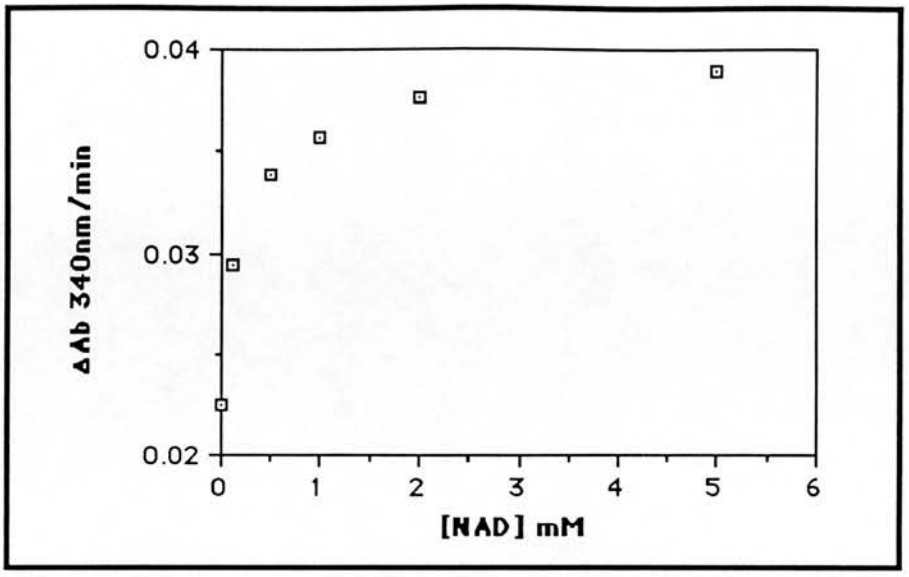


Figure 2.2.9-III. Effect of increasing concentrations of NAD^+ on the denaturation on the loss of ADH activity.

---

d) The specificity effect of NAD^+ was checked by incubating the enzyme with a control molecule, of a similar molecular weight but of a totally different structure. Kanamycin was used for this purpose and it was shown to have no effect on the denaturation process of DADH (figure 2.2.9-IV).

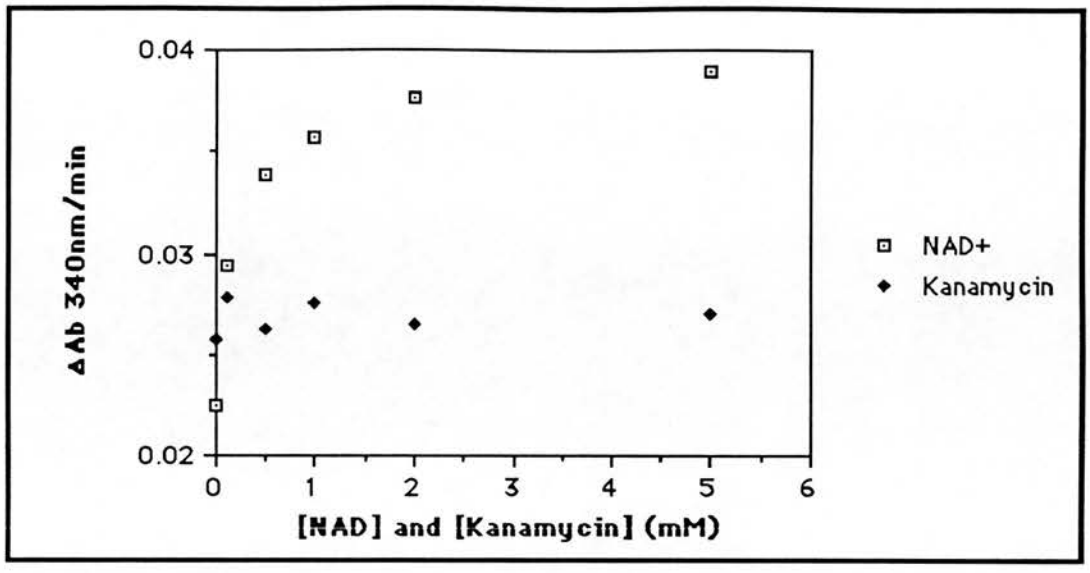


Figure 2.2.9-IV. The effects of addition of NAD^+ and kanamycin on the loss of ADH activity.

Small aliquots of the enzyme (typically 20 µl) were placed at the bottom of four spectrophotometer cuvettes. To two of the cuvettes 10 µl of a 10 mM solution of NAD^+ were added and left to equilibrate at 25 °C for 5 min. After that period a calculated amount of buffer (100 mM Tris HCl pH 8.7) was added to the four tubes, immediately followed by the necessary amount of 2M GdnHCl to reach the desired denaturant concentration in a total volume of 100 µl. The solutions were mixed and left to incubate for 15 min at 25 °C.

During that period the reaction buffer was prepared for the four cuvettes. The reaction buffer contained NAD, isopropanol and GdnHCl in 100 mM Tris HCl pH 8.7.

The concentrations of substrates and denaturant were such as to make the final concentrations in the reaction as follows : NAD: 1 mM, isopropanol: 10 mM, GdnHCl: same concentration as the one the enzyme had been previously incubated in.

2.2.9.1.3 Determination of the conformational free energy of the enzymes.

Provided that the denaturation kinetics of a protein comply with certain conditions (i.e., denaturation equilibrium reached, reversibility of the reaction, denaturation process following a two-step mechanism) the free energy corresponding to the initial state of the protein (ΔG(water)) can be estimated, and hence the differential ΔG induced by specific mutations, or by the effect of the binding of a cofactor or substrate, can be quantified. An example of an ideal denaturation curve following the mentioned conditions is shown in figure 2.2.9-V.

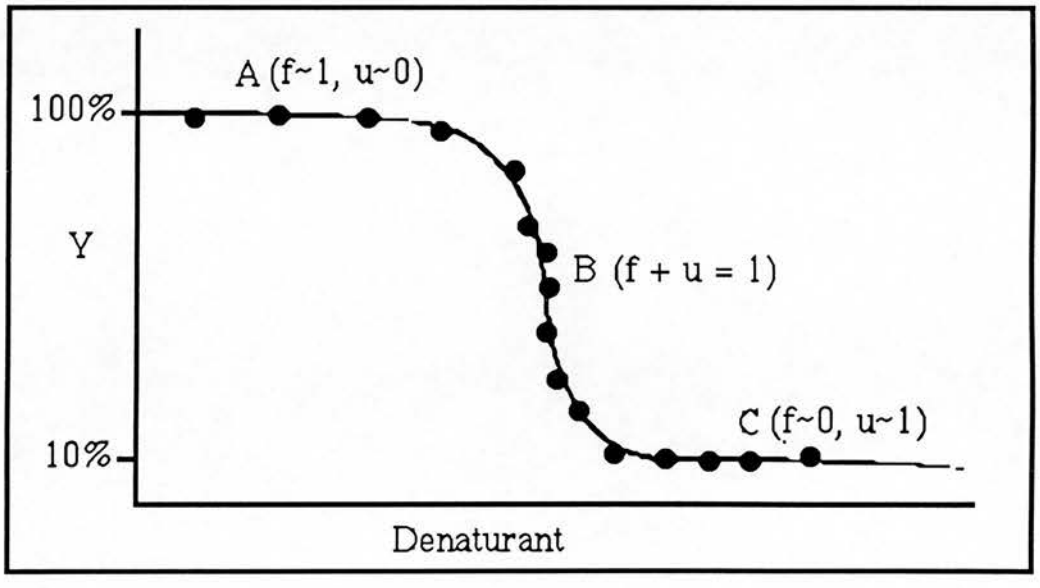


Figure 2.2.9-V. Ideal denaturation curve of a protein upon addition of denaturant.

In figure 2.2.9-V, Y represents the parameter used to measure the unfolding process, A B and C are the three different regions of the curve, f and u represent the proportion of folded and unfolded protein at each region of the curve. In region A, at low concentration of

denaturant, the folded state of the protein predominates. Region B corresponds to the concentrations of denaturant where the folded-unfolded transition occurs, while region C is the zone where the unfolded state of the protein predominates.

If one defines Ff and Fu, as the fraction of unfolded and folded protein present at any given concentration of denaturant then, at any point of the curve

$$Ff + Fu = 1$$

and, clearly, the Y value at any point of the curve will be

$$Y = YfFf + YuFu$$

where Yf and Yu are the characteristic Y values of the folded and unfolded states of the protein.

Combining both equations the value of Fu at any point is

$$Fu = (Yf-Y)/(Yf-Yu)$$

Since the equilibrium constant of unfolding K equals

$$K = Fu/(1-Fu) = Fu/Ff = (Yf-Y)/(Yf-Yu)$$

the Boltzmann equation can then be used to calculate the ΔG at any point of the curve

$$\Delta G = -RT \ln K = -RT \ln (Yf-Y)/(Yf-Yu)$$

where R is the gas constant (1.987 cal/deg/mol) and T is the absolute temperature.

The value of K, however, can only be accurately measured in the transition region B, because of the substantial experimental error that occurs in regions A and C, where the one of the two conformational populations is very small.

For this reason the ΔG values in the transition region B are used to obtain the ΔG(water) value of the protein by extrapolation from the trend of the calculated transition ΔG's, as shown in figure 2.2.9-VI.

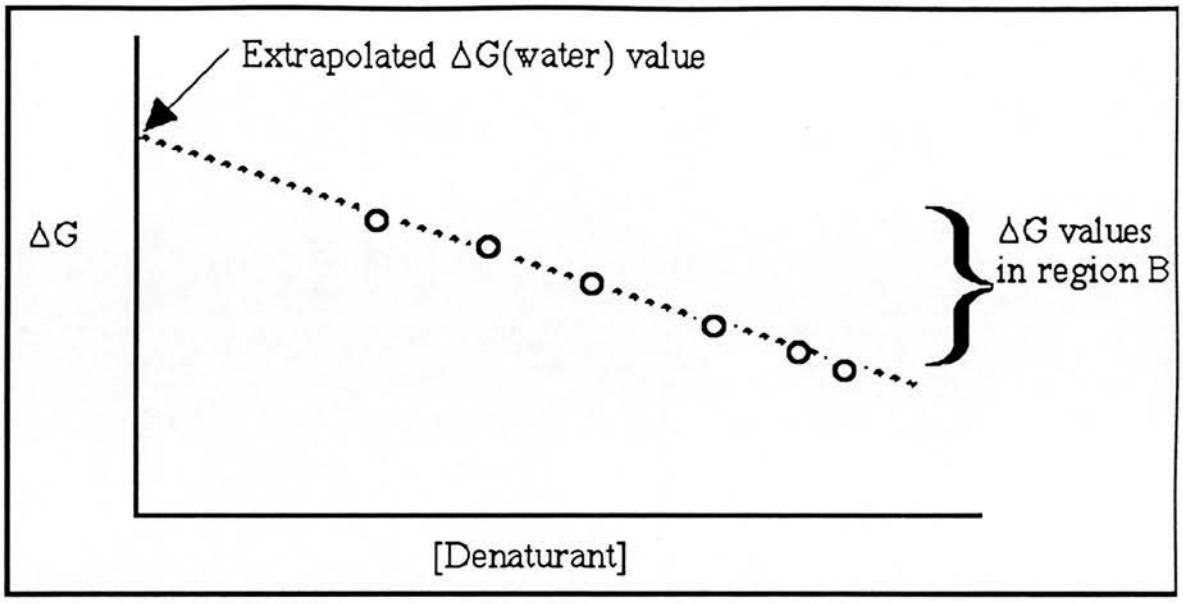


Figure 2.2.9-VI. The extrapolation of the ΔG(water) value of a protein from points of the denaturation curve.

Clearly, if the introduction of a certain mutation in the protein, or the binding of a ligand, affects the thermal stability of the molecule (without affecting the unfolding pathway), the Fu/Ff ratio at any denaturant concentration will change, and different ΔG(water) values will be obtained. The difference between the wild type and mutant (or enzyme ligand) ΔG(water) values will be the ΔΔG effect of the introduction of the mutation (or of the ligand binding), as seen in figure 2.2.9-VII .

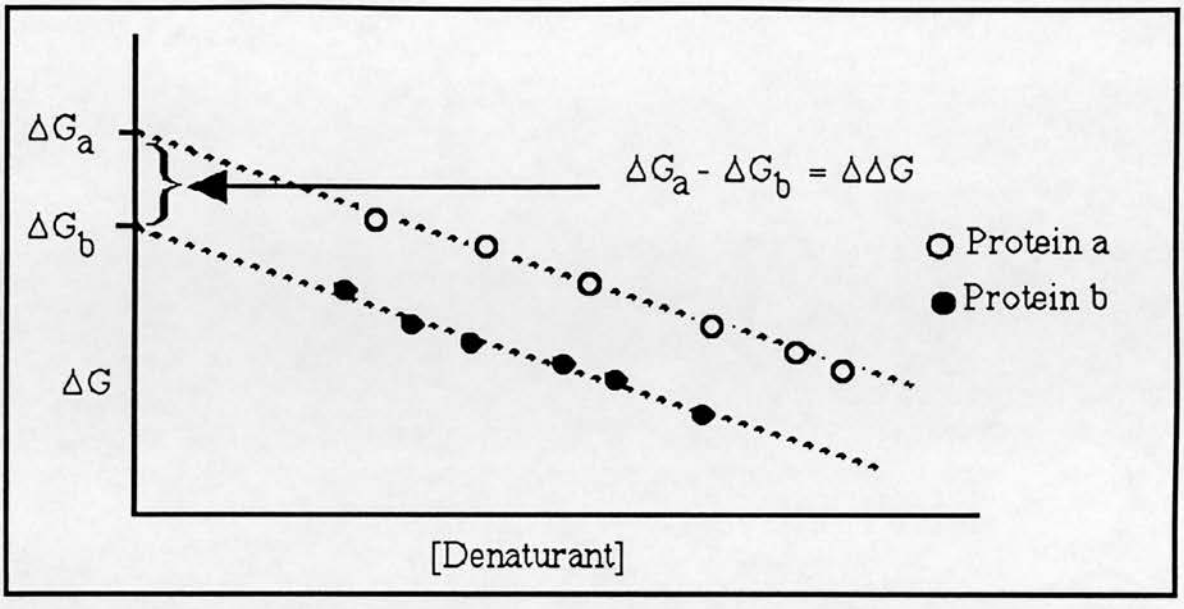


Figure 2.2.9-VII. The calculation of the ΔΔG effect by subtraction of the ΔG(water) extrapolated values.

2.2.9.2 Determination of the thermal stability of DADH's

The molecular effect of chemical denaturants like GdnHCl or urea is not well understood (Creighton, 1988). It is generally believed that their denaturing effect is due to their ability to increase the solubility of hydrophobic groups of proteins, and hence decreasing the hydrophobic energy that contributes to the tight packing of protein cores.

The effect of rising temperature is quite different and much more extensive because a rise in the general energy state of the system decreases the energetic constrains of not only the hydrophobic interactions but also of the electrostatic and structural (backbone rigidity etc.) forces that contribute to the architecture of the protein.

As described in the introduction, the thermal stability of DADH's is one of the candidates to explain the differential distribution of

isoenzymes, for this reason thermal stability analyses of DADH's are common in the literature (eg. Atrian & Gonzàlez-Duarte, 1985).

Since the conformational free energies of DADH's had not been previously analyzed, the thermal stability of each enzyme purified in this project was determined in order to be able to compare the effects of the introduced mutations in DADH to published reports in stability.

Procedure :

Aliquots of each purified enzyme were incubated at three different temperatures (0°C, 25°C and 40°C). Small volumes were removed from the aliquots at fixed time periods and the ADH activity present was immediately analyzed as described previously. All activity assays were done in duplicate.

2.2.9.3 Determination of kinetic constants.

Determination of apparent Km's:

The Km_{app} values for NAD, $NADP^+$ and propan-2-ol were calculated maintaining the concentration of one of the substrates constant in the assay conditions and varying the concentration of the other substrate.

The assay conditions were as described in 2.2.7.2 except for the substrate being analyzed. The Km_{app} for isopropanol was calculated keeping the NAD^+ concentration at 1 mM and using the following concentrations of propan-2-ol : 0, 0.25, 0.5, 1, 4 and 16 mM.

The Km_{app} for NAD^+ was calculated with a constant concentration of propan-2-ol of 10 mM and with the following concentrations of the cofactor : 0, 0.05, 0.2, 0.8, 1.6 and 3.2 mM.

Similarly, the Km_{app} for $NADP^+$ was calculated keeping the propan-2-ol concentration at 10 mM and using the following concentrations of $NADP^+$: 0, 0.1, 0.4, 1.6, 3.2 and 6.4 mM.

All assays were done in triplicate, standard deviations were calculated and an additional control with the higher substrate concentrations and without enzyme was performed.

The Km_{app} values were obtained from a Hanes plot ([substrate] vs. [substrate]/velocity), where Km_{app} values are obtained from the plot equation, and equal -x when y=0.

Determination of Vmax and k_{cat} values.

Vmax values were also determined from the Hanes plot, as they are defined as 1/m, where m is the slope of the equation.

k_{cat} values were calculated as $k_{cat} = Vmax/e_o$, where e_o is the enzyme concentration which was calculated using the published molar absorption coefficients for *D. melanogaster* ADH (Cornish-Bowden & Wharton, 1988; Juan & Gonzàlez-Duarte, 1981). Although an enzymatic assay is available for the determination of the active site concentration of the enzyme (Winberg et al., 1985), it requires large amounts of enzyme at high concentrations, and is not normally used in the literature.

3. Cloning, mutagenesis, sequencing and expression of ADH. Results and discussion.

3.1. Cloning of the *Adh* gene into M13mp18 and PVT-U-100.

The *Adh* gene included in the vector pUC-118 had to be subcloned into M13mp18 to obtain single-stranded DNA for the introduction of site-directed mutations and to confirm the final *Adh* sequence by DNA sequencing.

The restriction map of the plasmid pUC-118 is shown in figure 3.1-I. The plasmid was digested with the enzymes HindIII and SpeI to isolate a 1.3 kb fragment containing the *Adh* coding sequence. Simultaneously, the M13mp18 DNA was digested with the restriction enzymes HindIII and XbaI, which cut once through the polylinker region of the phage's DNA. The digested M13mp18 DNA and the 1.3 kb fragment containing *Adh* were ligated taking advantage of the compatibility of the hanging sequences generated by SpeI and XbaI (see figure 3.1-II for representation of the resulting M13-Adh plasmid).

Once the subcloning conditions were optimized an M13mp18 clone containing the *Adh* insert was obtained, as demonstrated by the restriction analysis in figure 3.1-III. This clone was called M13-Adh.

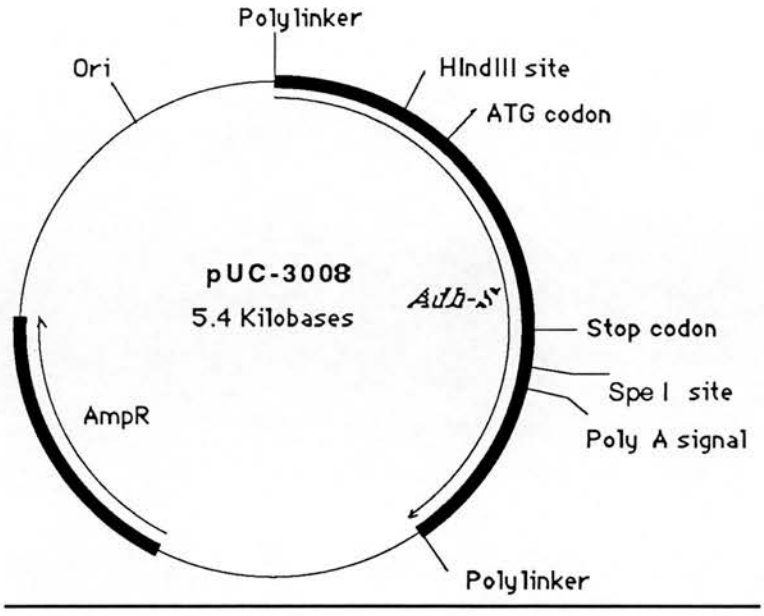


Figure 3.1-I. Map of the plasmid pUC115-3008, provided by Dr. Sofer, containing the *Adh* gene without introns.

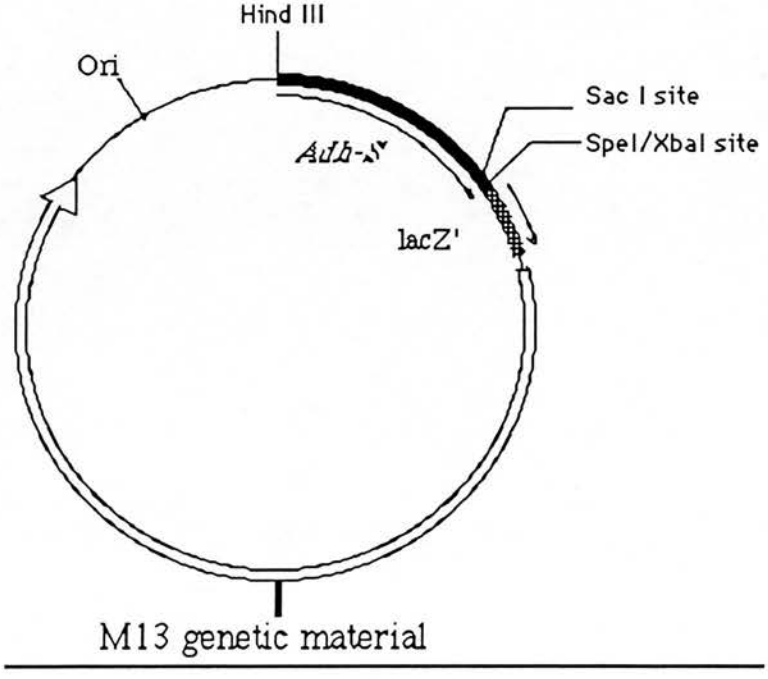


Figure 3.1-II. Map of the plasmid M13-Adh obtained by subcloning the Hind III/Spe I fragment of pUC-3008 containing the *Adh-S* gene into M13mp18 digested with Hind III and Xba I.

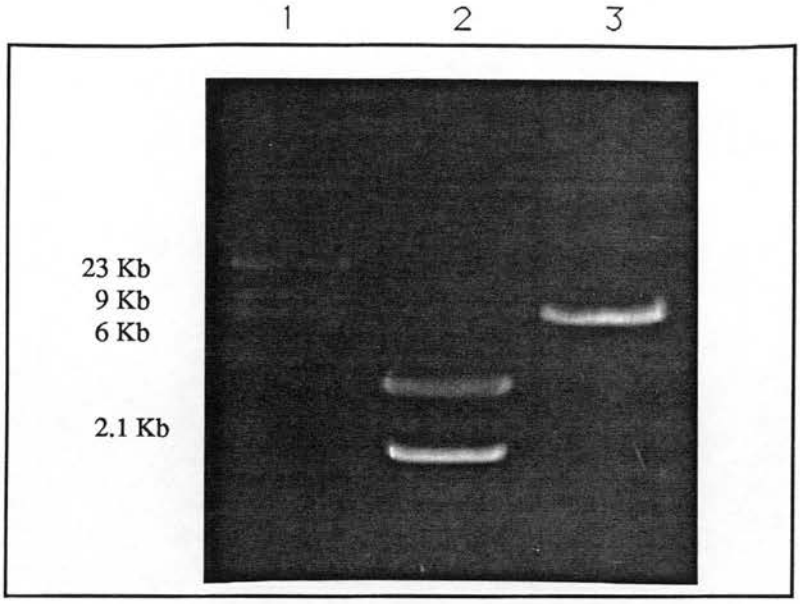


Figure 3.1-IIIa. Restriction analysis of the plasmids pUC-3008 and M13mp18 digested with the enzymes Hind III and Spe I. (1) Molecular weight marker. (2) pUC-3008. (3) M13mp18.

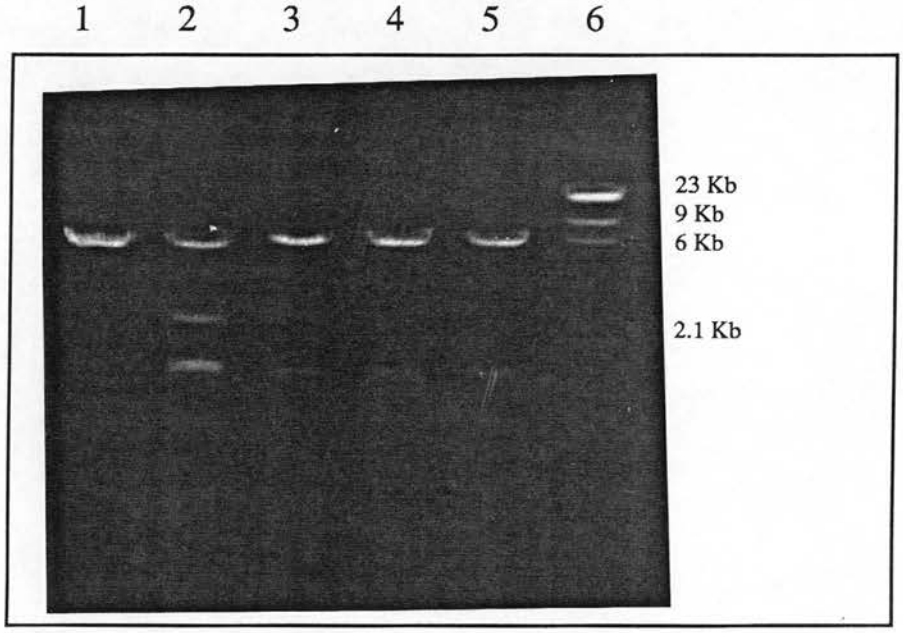


Figure 3.1-IIIb. Restriction analysis of M13mp18 and M13-Adh digested with the enzymes Hind III and Sac I. (1) M13mp18. (2-5) M13-Adh, the band of 1.3 Kb containing the *Adh* gene can be seen at the end of each track. The intermediate band in tracks 2 and 5 corresponds to undigested plasmid in super-coiled form. (6) Molecular weight marker.

3.2. Site-directed mutagenesis of the M13-Adh clone and sequencing of the mutations.

Rationale of the mutations introduced:

Table 3.2-1 summarizes all the mutations that were introduced into the *Adh* gene of the M13-Adh clone, all mutations were introduced using the Kunkel site-directed mutagenesis system (Kunkel, 1985).

Mutation	*Adh-S* sequence and mutations introduced
Gly 14 to Ala	37-GTTGCCGGTCTGGGAGGC-54 **GCT**
Gly 19 to Ala	49-GGAGGCATTGGTCTGGAC-63 **GCT**
Asp 38 to Ala	109-ATCCTCGACCGCATTGAG-126 **GCC**
Asn 136 to Stop	403-ATCTGCAACATTGGATCC-420 **TAA**
Pro 214 to Ser	634-CCCACCCAGCCATCGTTG-651 **TCA**

Table 3.2-1. List of the mutations introduced into the *Adh-S* gene. Below the native *Adh* sequence the codon mutated is indicated and the bases changed are shown in bold characters.

Mutations G14A, G19A, D38A and 136Stop: as described in the introduction, the main structural interpretation that could be made from the sequence of *Drosophila* ADHs was the existence of a nucleotide binding domain at the N-terminal part of the protein. Although such hypothesis is now supported by the evidence from the crystal structure of SDH, this information was only published at the end of this project, and the need still remains for detailed information on the enzyme-coenzyme interactions in *Drososphila* ADH.

For these reasons a series of mutations affecting residues thought to be important for the cofactor recognition were designed. Residues Gly 14 and Gly 19, located in the poly-glycine region of the N-terminal sequence of the enzyme, were mutated to alanines in order to try to disrupt the space that the glycines are meant to provide in the cofactor-binding site. During the development of this project similar or identical experiments have been reported in the literature (Scrutton et al., 1990; Chen et al., 1991; Feeney et al., 1990), confirming the importance of the poly-glycine sequence for cofactor recognition.

Residue Asp 38 in the short-chain dehydrogenases was originally thought to correspond structurally to the Asp 223 of Horse liver alcohol dehydrogenase, and to the conserved aspartic acid of other nucleotide binding domains, where it creates a hydrogen bond with one of the hydroxyl groups of the adenine ribose of the cofactor. To test such possibility we decided to mutate Asp 38 to alanine, in order to destroy any possible hydrogen bond of the side chain with the coenzyme molecule.

The availability of antibodies against the enzyme encouraged the attempt to introduce a stop codon at the end of the supposed N-terminal NAD-binding domain. The purpose of introducing this mutation was to test whether the N-terminal part of *Drosophila* ADH would be capable of folding into a functional binding domain. The presence of a soluble polypeptide expressed from the truncated gene could be tested with antibodies against *Drosophila* ADH, and its dinucleotide binding capacity analyzed by techniques like equilibrium dialysis.

Mutation P214S : The mutation P214S in ADH, as described in the introduction, is found in some wild populations of *D. melanogaster,* always in the ADH-F form of the enzyme. The mutation has a double effect: it increases the thermal stability of the enzyme, and it modifies

its kinetic properties making it ADH-S like, presumably by increasing the binding affinity of the enzyme with the cofactor, hence slowing down the turn-over rate (Chambers, 1984; Chambers, 1988). So far, no ADH-S isoenzyme containing the P214S mutation has been found. The effect of the mutation has hence a double interest: it is intrinsically interesting to study the physical effect of the change that induces an increase in structural stability; it is also interesting to find out what kind of effect the change has in an ADH-S form of the enzyme.

Mutagenesis results :

The frequency of mutation was always 75% or higher (four clones were selected for sequencing after every mutagenesis experiment, and a minimum of 3 were always found to be mutant). The sequences of all the mutants produced are shown in figure 3.2-I. At this point no attempt was made to sequence the whole gene as this was to be done for all mutants, after the yeast transformation steps.

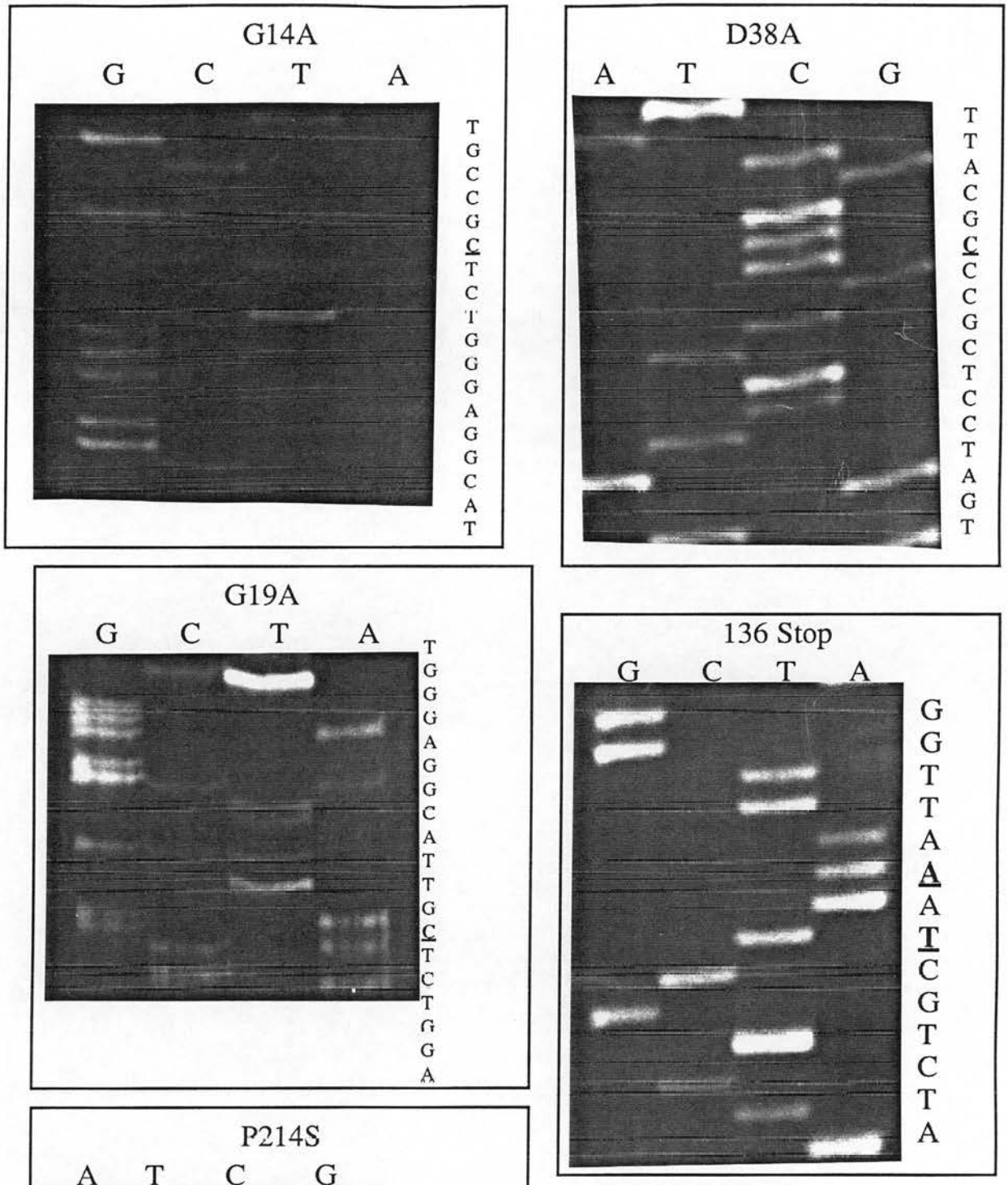


Figure 3.2-I. Autoradiographs of the sequencing gels showing the five mutations introduced into the *Adh* gene. The mutated bases are underlined.

3.3. Cloning of the wild type and mutated M13-Adh clones into PVT-U-100.

The transfer of the wild type and all mutated *Adh* genes, except Adh136Stop, into PVT-U-100 was carried out using the restriction sites HindIII and SacI, both present in the polylinkers of M13mp18 and PVT-U-100 in the right orientation for expression from the ADH1 promoter present in the plasmid (see figure 3.3-I for a restriction map of PVT-U-100).

Adh136Stop was subcloned into PVT-U-100 using the sites HindIII and BamHI, which produce a fragment of 570 bp, containing the initiation codon ATG and the introduced stop codon ATT.

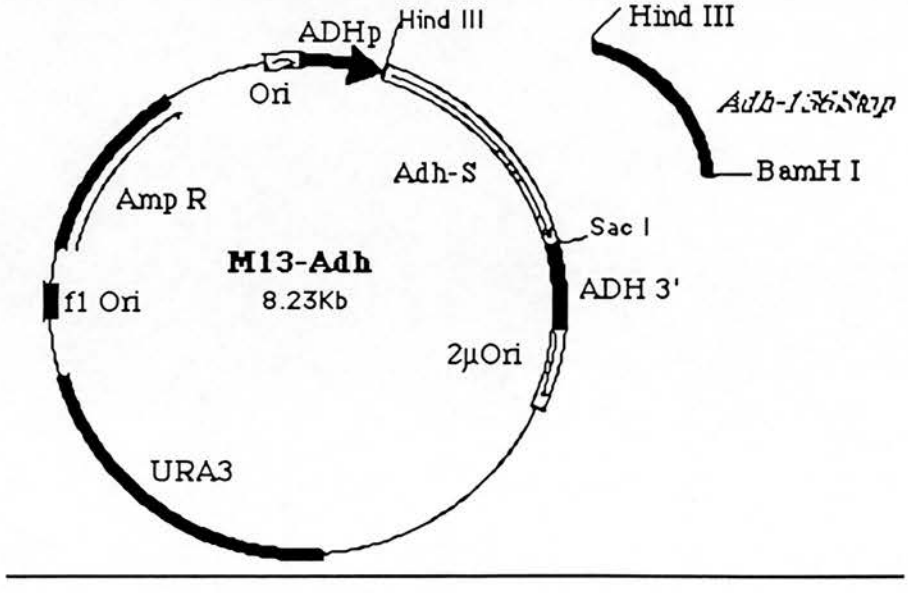


Figure 3.3-I. Map of the PVT-Adh plasmids obtained by subcloning the HindIII/SacI fragment of M13-Adh that contained the *Adh* gene into M13mp18 digested with HindIII and SacI. The mutant Adh-136Stop was digested from M13 with HindIII and BamHI (the latter cutting the gene 30 bases upstream from the introduced stop codon) and subcloned into M13mp18 using the same restriction sites. .

Photographs of the agarose gels containing restriction analysis samples of the PVT-U-100 clones with the wild type and mutagenized *Drosophila Adh* genes are shown in figure 3.3-II. It can be seen that new plasmids, carrying the *Adh* containing fragments, were obtained.

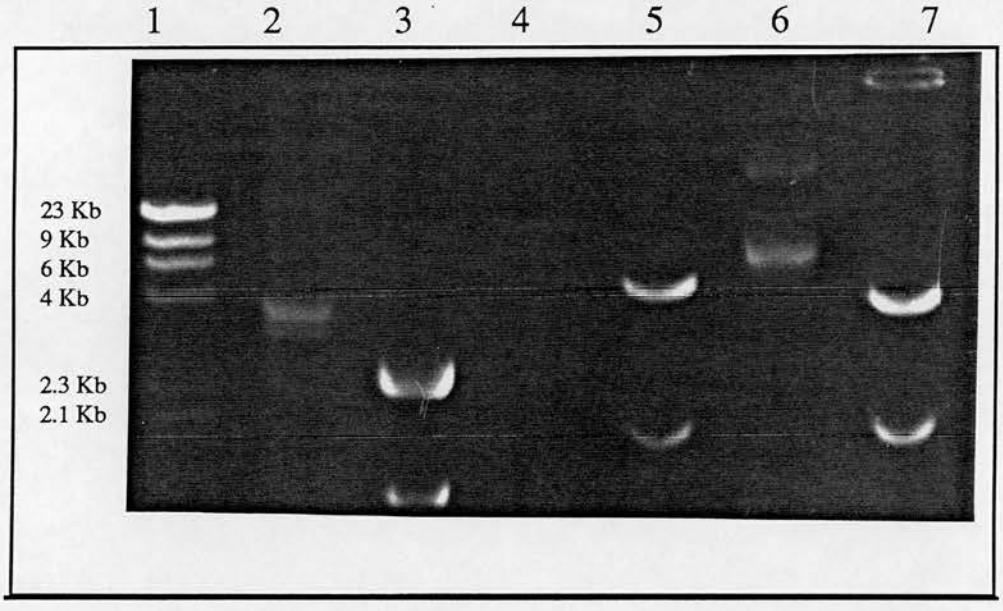


Figure 3.3-II. Restriction analysis of the PVT-Adh . (1) Molecular weight marker. (2) pUc-3008. (3) pUC-3008 digested with HInd III and Spe I as in figure 3.1-III. (4) PVT-Adh. (5) PVT-Adh digested with Hind III and Eco RI, this digestion produces a fragment of 1.8 Kb containing PVT material, and a larger fragment of 6.4 Kb containing the entire coding sequence of *Adh*. (6) PVT-Adh136Stop. (7) PVT-Adh136 Stop digested with Hind III and Eco RI, the large fragment contains the *Adh* -136Stop codifying sequence. As it can be seen the larger band is smaller than that containing the entire *Adh* coding sequence run in track 5.

3.4 Transformation of yeast, rescue and sequencing of the whole *Adh* genes.

The plasmids obtained by subcloning of the various *Adh* mutants into PVT-U-100 were used to transform the yeast strain WV.36.201 as described previously. Once the yeast clones containing the desired plasmids were obtained, the plasmids were rescued, their presence tested by restriction analysis and their *Adh* genes wholly sequenced to confirm that only the desired changes were present.

To our surprise several unreported mutations were found in the sequence. All the new changes were found in all the plasmids sequenced, ruling out possible artifacts of mutagenesis or yeast transformation, and pointing to a common origin for all of them.

To confirm the presence of the changes in the original *Adh* sequence obtained from Dr. Sofer the whole *Adh* gene of the plasmid pUC-118 was sequenced in both orientations, between the nucleotide 129 upstream from the initiation codon to nucleotide 201 downstream from the termination codon. This confirmed that all the new changes detected had their origin in the starting *Adh* coding sequence.

Table 3.4-1 lists all the changes found with respect to the 'consensus' ADH-S sequence as described by Kreitman (Kreitman, 1983). Autoradiographs of the sequencing gels containing the described changes are shown in figure 3.4-I.

	Position of the change	Change from consensus	Transcriptional effect
1	Nucleotide 518	T changes to C	Ile 173 changes to Thr
2	Nucleotide 531	C changes to G	Silent change
3	Nucleotide 540	C changes to T	Silent change
4	Nucleotide 684	G changes to A	Silent change
5	Nucleotides 786 to 795	Deletion of A	Past termination codon
6	Nucleotide 819	Insertion of a T	Past termination codon
7	Nucleotides 849 to 850	Deletion of a T	Past termination codon

Table 3.4-1. List of changes found in the *Adh* sequence used in this thesis relative to the 'consensus' sequence proposed by Kreitman (Kreitman, 1983). The nucleotide numbering starts at the first nucleotide of the initiation codon of the gene.

Discussion :

Of all the changes found in the *Adh* sequence, three of them, at positions 531, 540 and 684, had already been described as polymorphisms by Kreitman, although these were not found together in the same gene.

Of the four remaining changes, two (insertion of T at position 819, and deletion of a T at positions 849-850) have not been reported previously. The remaining two (change of a C by a T at position 518, and deletion of an A between positions 786-795) are located in areas found by Kreitman to be polymorphic between different populations, but the changes that we found are of different nature. The mutation at position 518, that creates a change in the protein sequence was, in fact, found by Kreitman at position 519, where the change is silent.

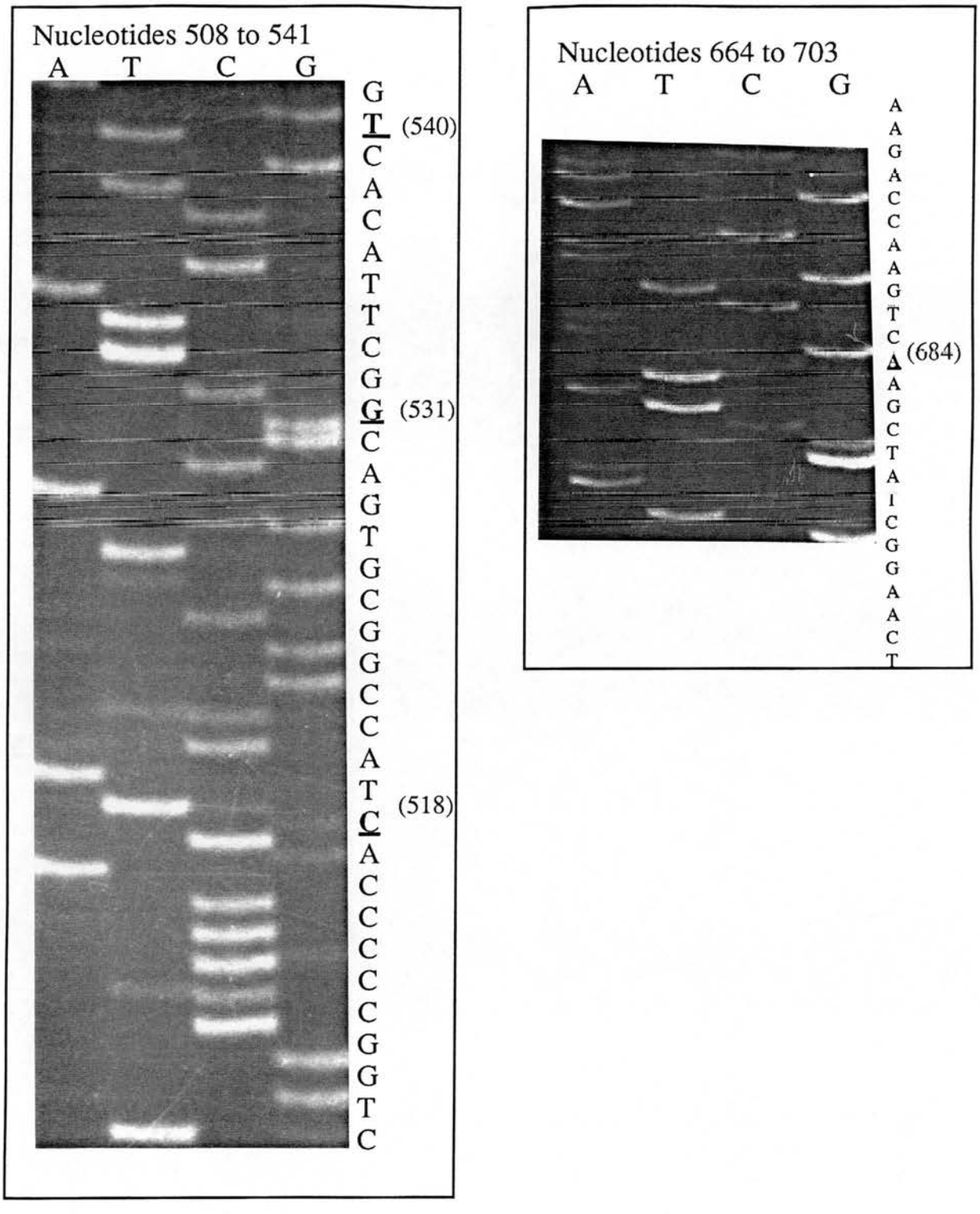


Figure 3.4-I. Autoradiographs showing the new nucleotide changes found inside the coding region of the *Adh* gene provided by Dr. Sofer. The new nucleotides found are underlined and their correspondent number (from the initiation codon) is shown.

The strategy used by Dr. Sofer's group for the construction of the intron-less *Adh* gene is of relevance in this discussion because it involved various genetic manipulations that could have resulted in the introduction of changes in the sequence.

The final gene is in fact a mosaic, constructed by hybridising a DNA primer, containing a fragment of an ADH-S clone ranging from nucleotides 568 to 1003, to an mRNA molecule encoding for ADH-F. The DNA primer did not contain any of the intron sequences of the *Adh* gene and it was extended with reverse transcriptase, using the mRNA molecule as template, and producing a hybrid DNA-RNA molecule. The RNA was then digested leaving a single stranded DNA molecule encoding for *Drosophila* ADH, without the intron sequences.

This DNA molecule originated from an *Adh*-F mRNA from nucleotides 1 to 567, and from an *Adh* -S gene from nucleotides 711 (this numbering includes intron sequences) to its end. Dr. Sofer's group was interested in constructing an intronless *Adh* -S 'gene' for expression studies, and since the only known difference between ADH-S and ADH-F is at amino acid 192 (caused by a change at nucleotide 575) they assumed that the DNA produced would encode for an -S type ADH. They confirmed that by checking the isoelectrofocusing analysis, but the final sequence of this mosaic 'gene' was not obtained (Clarke, 1988). To further complicate matters, the sequence of the mRNA used for the construction of the intronless *Adh* was not checked.

As can be seen from all the exposed and table 3.4-1, the changes numbered 1, 2 and 3, found during this project, must belong to the sequence of the mRNA molecule, while the rest of the changes correspond to the original sequence of*Adh* -S .

However one cannot rule out the possibility that some of these changes may have been introduced during the genetic engineering manipulations used for the construction of the final coding sequence. Strongly against such possibility argues the fact that most of the changes are either previously described polymorphisms or similar to these in both position and nature. The only two novel changes are in the non coding upstream region of the gene.

Of major relevance to this work was the finding of the change at position 518, which determines a non-reported difference in the sequence of the enzyme, substituting the expected isoleucine by a threonine. The amino acid that is changed, Ile 173, is conserved in all the *Drosophila* ADH enzymes sequenced so far. Although a change of an isoleucine for a threonine can be considered as a rather conservative substitution its appearance made it mandatory to fully characterize the new enzyme (which, for simplicity, we will continue to call ADH-S in this thesis), in order to be able to compare the mutations introduced in it with the rest of published data on ADH-S.

3.5 Analysis of growth behaviour of the different yeast strains and of ADH activity in crude extracts.

The yeast clones transformed with the PVT-U-100 plasmids containing the wild type, G14A, G19A, D38A, 136Stop and P214S forms of *Adh* will be named respectively : WVwt, WV14A, WV19A, WV38A, WV136Stop and WV214S. The cells containing the plasmid without the *Adh* gene will be named WV. (Growth curves shown in figure 3.5-I)

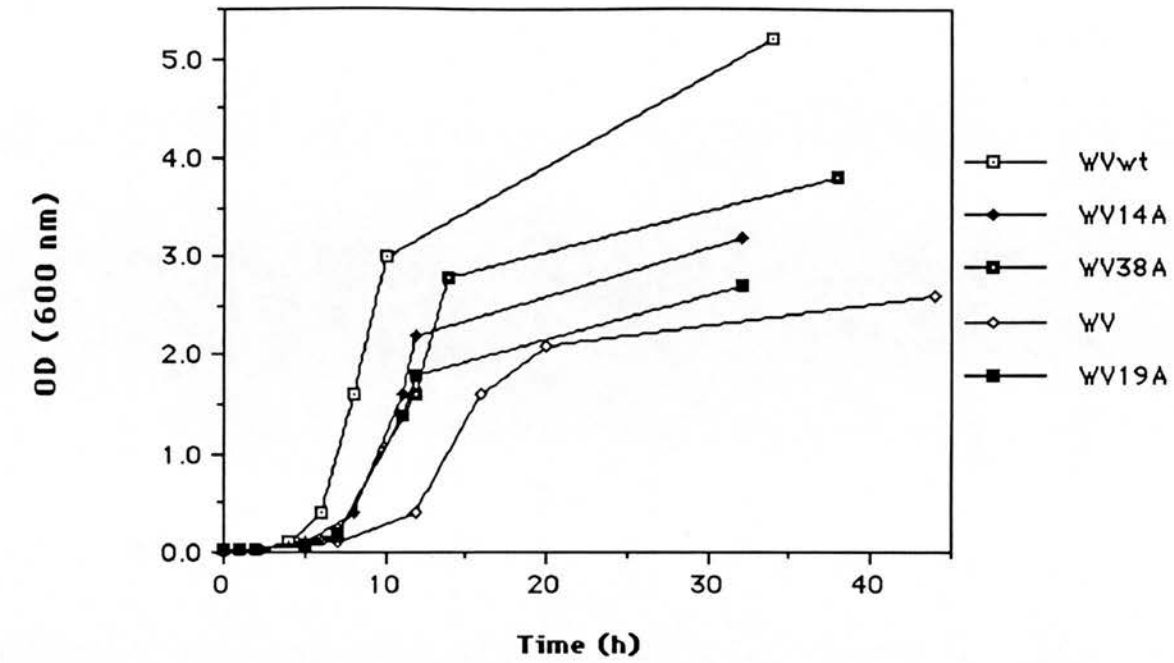


Figure 3.5-Ib

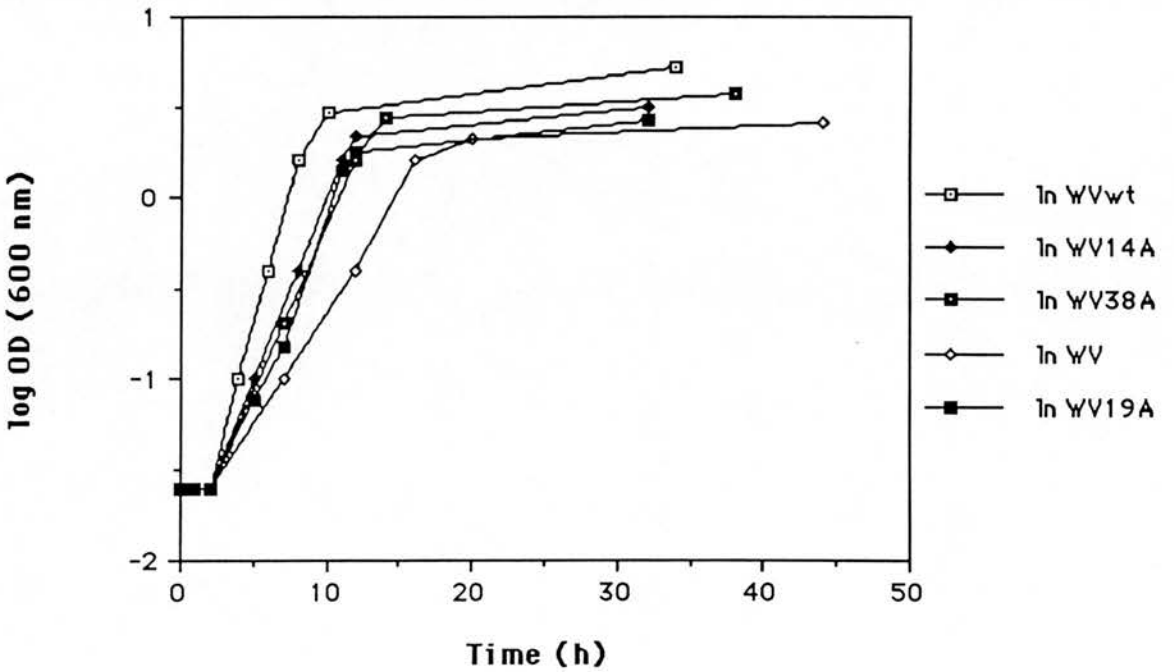


Figure 3.5-I . A) Growth curves for the different strains described. Strains WV136Stop and WV214S showed identical behaviour to strain WV, and are not displayed for clarity. B) Logarithmic transformation of the original growth values.

The growth curve for each of these clones was obtained several times and a consistent difference was found in the growth behaviour of the different cells analyzed (figure 3.5-I). Simultaneously, the amount of ADH specific activity found in crude extracts of the different clones at stationary phase of growth was tested (figure 3.5-II).

Figure 3.5-II

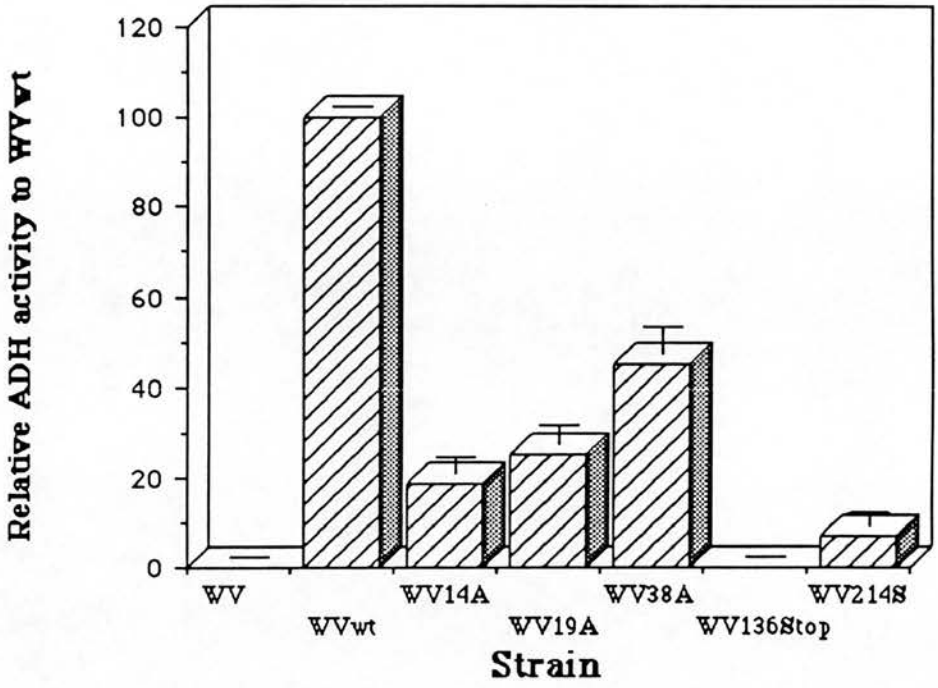


Figure 3.5-II. The specific activity values for the crude extracts of the different strains relative to the activity of the activity of the crude extract of the WVwt strain.

The relationship between the different clones according to maximum cell density attained in liquid culture is as follows :

WVwt > WV38A > WV14A > WV19A > WV214S = WV136Stop = WV

Similarly, the relationships according to ADH specific activity in crude extract are :

WVwt > WV38A > WV19A > WV14A > WV214S > WV136Stop = WV = 0

The main differences between these two comparisons are : a) crude extracts of WV19A cells have higher specific activity than WV14A cells, but they grow to slightly lower cell densities in cell cultures. b) WV214S cells do not grow to higher cell densities than WV136Stop or WV cells, but a certain amount of ADH activity can be detected in the former, while, as expected, neither WV136Stop or WV cells have any detectable ADH activity in their crude extracts.

Discussion :

As clearly shown by the growth curves of the different strains used in this thesis the expression of the *Drosophila Adh* gene has an enhancing effect on the growth rate and on the attainable cell density properties of its expressing yeast cells. Although it would be necessary to check whether such effects are lost when the enzyme is expressed in wild type yeast cells, it seems very probable that the reason for such enhancement may be a complementation of the ADH activity that in the strains used has been eliminated through deletion of the three yeast *Adh* genes.

In wild type yeast growing in glucose-rich media and in aerobic conditions the main ADH activity (carried out by the ADH-1 isoenzyme) is dedicated to the production of ethanol from the acetaldehyde that is obtained from the glycolytic pathway. When ethanol is used as a carbon source the first step of utilization involves its oxidation to acetaldehyde by the effect of ADH-2 (Wills, 1990). In the case of the WV.36.201 cells used in this project, the lack of ADH-1 should, in glucose-rich conditions, lead to an accumulation of acetaldehyde, to a decrease in the available pool of NAD^+ in the cell and to a lack of ethanol as a metabolite once the glucose pool has been depleted. Also, the lack of ADH-2 should lead to a poor utilization of ethanol.

Drosophila ADH could be partially complementing both alcohol dehydrogenase activities, by shifting the equilibrium towards the production of ethanol as acetaldehyde accumulated and, inversely, catalyzing the oxidation of ethanol once the glucose pool had been depleted. This explanation would agree with the fact that the growth capacity of the cells used in this project generally correlates with the ADH activity detectable in their crude extracts, which, in turn, gives a preliminary indication of the level of activity of the different mutants expressed in the cells.

However, the level of ADH activity found in the crude extracts and the effect of each enzyme in its expressing cells is not only a function of the enzyme's catalytic efficiency. It is a combination of chemical efficiency with expression and stability, and so it will also depend on the level of production of each mutant, in its physical stability, and in its resistance to proteolytic cleavage. Differences in these two last factors may be the causes for the apparent contradiction between the results of growth of WV14A and WV19A cells and their relative ADH activity in crude extracts. WV14A cells grow slightly faster and to slightly higher densities than WV19A cells, but the crude extracts of the latter have higher ADH specific activity. This could be due to a relatively higher instability of the enzyme ADH G14A, that could result in a higher loss of its activity, due to denaturation or proteolysis, during the crude extract preparation process.

Without any analysis of enzyme production or kinetic evidence from purified enzyme solutions it was not possible to know at this stage if all the detected differences are, or are not, due to differences in the level of expression, enzyme activity or stability of each of the ADH variants. These analyses will be described in the rest of this chapter.

3.6 Immunological analysis of ADH expression.

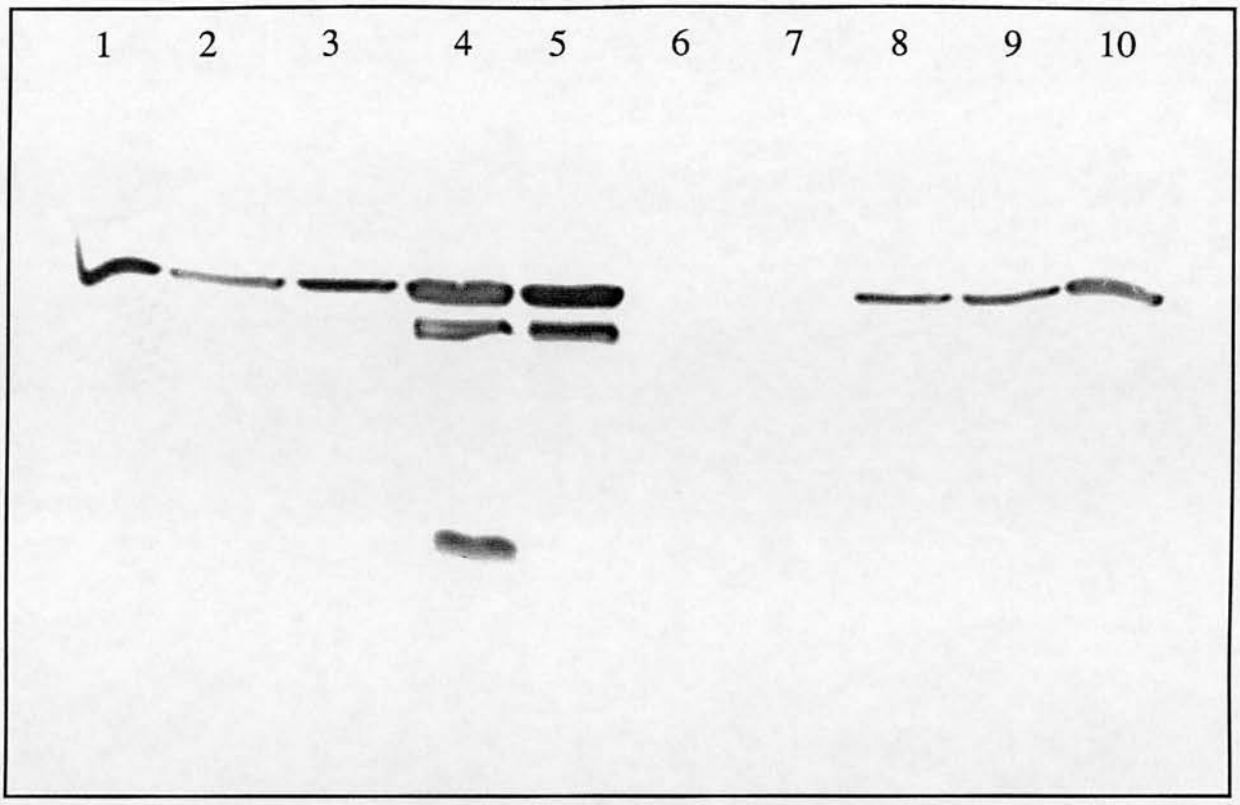
The presence of cross-reacting material with monoclonal and polyclonal antibodies against *Drosophila* ADH was tested in WV, WVwt, WV14A, WV38A and WV136Stop cells. The results for the different clones and antibodies used can be seen in figure 3.6-I.

Because similar amounts of total protein were used from all the cells analyzed the intensity of the recognized bands can be used to assess the relative amounts of ADH in each cell type. All the antibodies used recognized a band of the same molecular weight as *Drosophila* ADH in WVwt, WV14A and WV38A cells, and several bands of lower molecular weight that can be attributed to degradation products because their band patterns are similar to the ones found in crude extracts from *Drosophila* flies (Fibla et al., 1989). No such bands were detected in the WV and WV136Stop cells.

Three of the antibodies used (polyclonal serum and the monoclonal antibodies LLBE8 and 10BB2) cross-reacted with other proteins of higher molecular weight that were present in all the clones tested, including WV (which does not contain an *Adh* gene). The monoclonal antibody MMBB8 did not cross-react with any bands in the WV or WV136Stop cells, hence showing specific reaction with ADH.

As it can be seen in figure 3.6-I, no significant differences in amount of seropositive material can be seen between clones WVwt, WV14A and WV38A. The rest of the clones analyzed (WV and WV136Stop) had no seropositive material when the monoclonal antibody MMBB8 was used, and no band of lower molecular weight than ADH could be detected.

Monoclonal antibody MMBB8



Polyclonal antibody

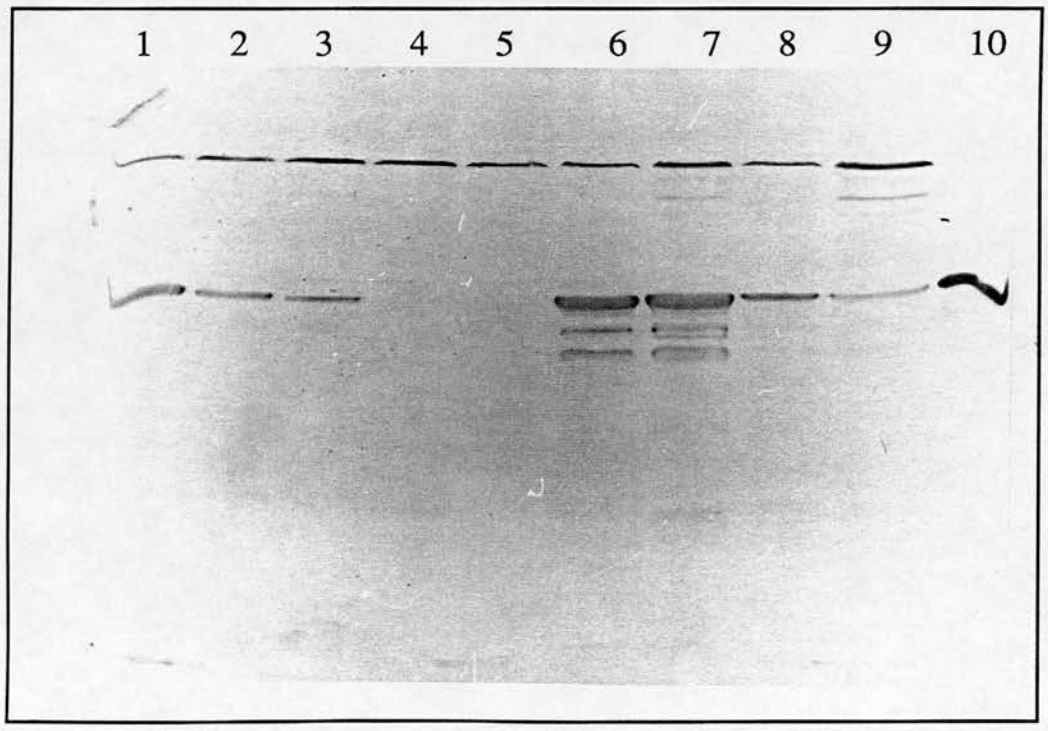


Figure 3.6-I. Immunodetection of ADH in crude extracts of yeast cells. Lanes 1 to 5 show the results with similar amounts of crude extract of the strains WVwt, WVG14A, WVD38A, WV136Stop and WVPVT respectively. Lane 10 is a purified ADH control. Lanes 6-9 are not relevant.

Discussion :

The results obtained seem to indicate that the level of expression of the wild type enzyme is largely equivalent to the expression of the mutants ADH G14A and ADH D38A. That, in turn, points to differences in intrinsic enzyme properties as the cause for the differential growth characteristics seen in the previous section, at least for the strains expressing these three enzymes. The final determination of the nature of these differences required the analysis of pure preparations of the enzymes.

A second important result of the immunological analysis of the different strains was the finding of no cross-reacting material in the WV136Stop cells, specially when polyclonal antibodies were used. Although such result is not conclusive, it seems to suggest that no soluble peptide is being produced by that strain, either because the truncated gene is not expressed, its mRNA its not correctly processed or translated, or because the translated peptide is not stable and either precipitates or is degraded. In any case, no further analyses were done on that mutant. It is unfortunate that the expression levels of the *Adh* genes could only be tested for five of the seven strains used in the project, but, as it will be seen in the purification results, there is no evidence of different levels of expression in the untested strains WVG19A and WVP214S.

4. Purification and characterization of wild type and mutant ADH enzymes. Results and Discussion.

4.1 Purification of *Drosophila* ADH from yeast cells .

General considerations :

In the Materials and Methods section of this thesis I have described only the procedures that form the final purification protocol used to obtain pure ADH from yeast cultures. However many different purification techniques were actually tried before reaching the definite purification scheme. In this section the results of each of the steps finally used for all the enzymes purified will be described, and a mention will be made of the results that were obtained with related techniques that were also tried.

4.1.1 Cell harvesting and lysis .

The procedure for cell harvesting and lysis was essentially a scale-up of the methods used to check for the presence of ADH activity in crude cell extracts, and so the results obtained from this step are essentially the same as those described in section 3.5. The total protein concentrations at this stage of the purification oscillated from 1 to 10 mg/ml depending on the cell density attained by the yeast cultures used (See tables 4.1-1 to 4.1-5).

4.1.2 Ammonium sulphate fractionations and de-salting procedures.

Two ammonium sulphate cuts were carried out on the crude extracts at concentrations of salt of 35% saturation and 70% saturation. That step had four main consequences on the sample : the lysate was homogenized and freed of general debris, the total volume of the samples was reduced by a factor of approximately 4, the total protein concentration was increased and the specific ADH activity of the samples was enhanced.

The specific activity values suggest a purification factor of around three after the salt cut. However, the treatment of ADH with ammonium sulphate enhances the activity of the enzyme, which suggests that such purification factor may partially be an artifact.

The ammonium sulphate that remained in the crude solution after the second fractionation was removed by running the samples through a Sephadex G-25 column as described. This step increased the volume of the sample by a factor of 3 to 5 and allowed recoveries of 80-85% of the total enzyme activity (see results tables 4.1-1 to 4.1-5).

Membrane dialysis was tried as an alternative at this stage of the purification but the method is slower and allows for much poorer recoveries of enzyme activity. Typically 30 to 40 % of the enzyme activity was lost, probably due to enzyme denaturation and proteolysis.

4.1.3 Anion exchange chromatography.

As described in the materials and methods section, it was felt that ionic exchange columns would be more suitable for the purification of ADH than the previously described affinity methods. To determine the

optimal conditions to apply to our system a battery of small columns of anionic and cationic exchange nature was prepared (see table 4.1.4-I).

The most encouraging results were obtained with the ion exchange matrix Q-Sepharose, equilibrated at pH 8.00. Under this conditions a large amount of the protein content of the samples was retained by the column, but almost 90% of the loaded ADH activity was eluted directly out of the matrix, giving rise to a 20 fold purification or more for the ADH wild type enzyme (tables 4.1-1 to 4.1-5, and figure 4.1.4-I).

A set back of this procedure is the loss of the sample concentrating advantages of normal ion-exchange methods, where, after a positive binding of the protein of interest, a salt gradient can be applied to elute the material in a small volume of solvent. In our case the volume of the sample was slightly increased after the Q-Sepharose column.

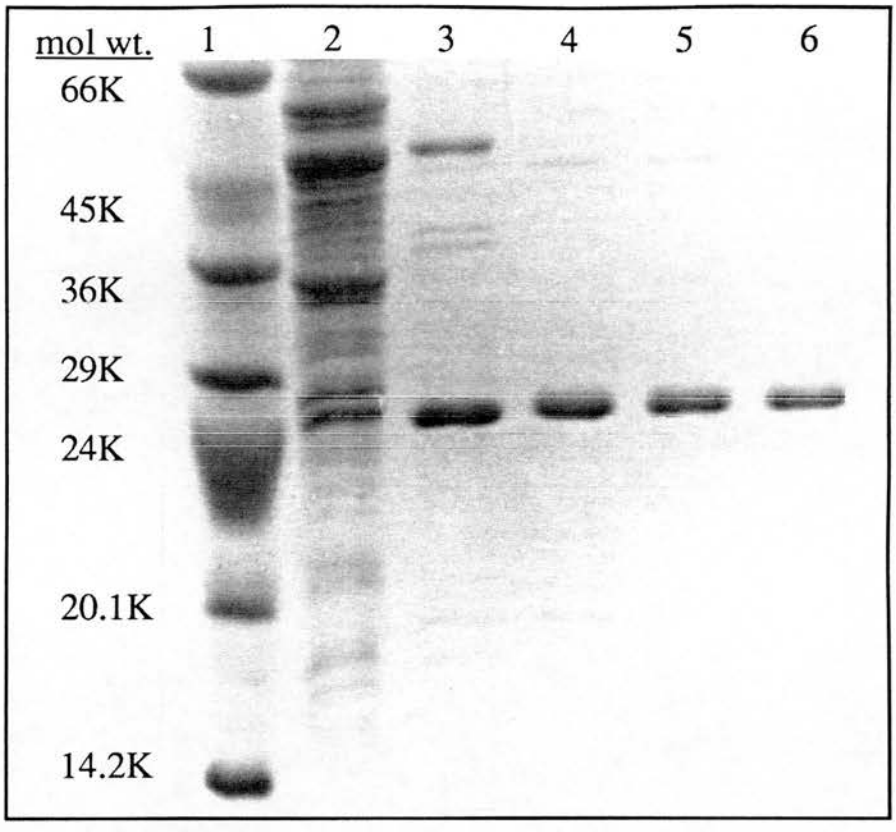


Figure 4.1.4-I. Photograph of an SDS-polyachrilamide gel showing samples at different stages of the purification of ADH-S. (1) Molecular weight markers. (2) Crude extract of WVwt yeast cells. (3) Sample of the fractions with ADH activity pooled after the Q-Sepharose column. (4-6) Consecutive fractions with ADH activity eluted from the Superose-12 column.

4.1.4 Gel filtration chromatography and previous concentrating step.

As can be seen in figure 4.1.4-I, a remarkable purification of the ADH enzyme was achieved after the anion-exchange column, but some contaminants were still present at this stage. The main contaminant band appeared to be of a much higher subunit mass (~55 kDa) than the ADH monomer. A good choice to separate these two molecules was, hence, a gel-filtration technique.

Three different gel-filtration approaches were tried, the Pharmacia Superose12 FPLC column (exclusion mass of $2x10^6$ Da), a combination of a Superose12 column followed by a Superose6 column (exclusion mass of $4x10^7$ Da), and the Pharmacia Superdex FPLC column. The best results were obtained with the first approach. The combination of Superose 12 followed by Superose 6 gave slightly better resolution of peaks but, due to the lower flux that Superose 6 columns can support, the processing times were dramatically increased. That factor, coupled to the lack of refrigeration system in the FPLC system used, gave rise to unacceptably high losses of enzyme activity due to denaturation.

On the other hand, the Superdex column gave poor resolution of peaks compared to the Superose12 column alone. In figure 4.1.5-I a chromatogram of a typical Superose12 separation is shown. That separation yielded samples of pure ADH when analyzed by Coomasie staining of SDS-polyacrylamide gels (figure 4.1-I). One of the mutants produced, D38A, behaved differently (figure 4.1.5-II), and was always eluted with variable, but detectable, amounts of a second protein of an apparently slightly larger monomeric molecular weight.

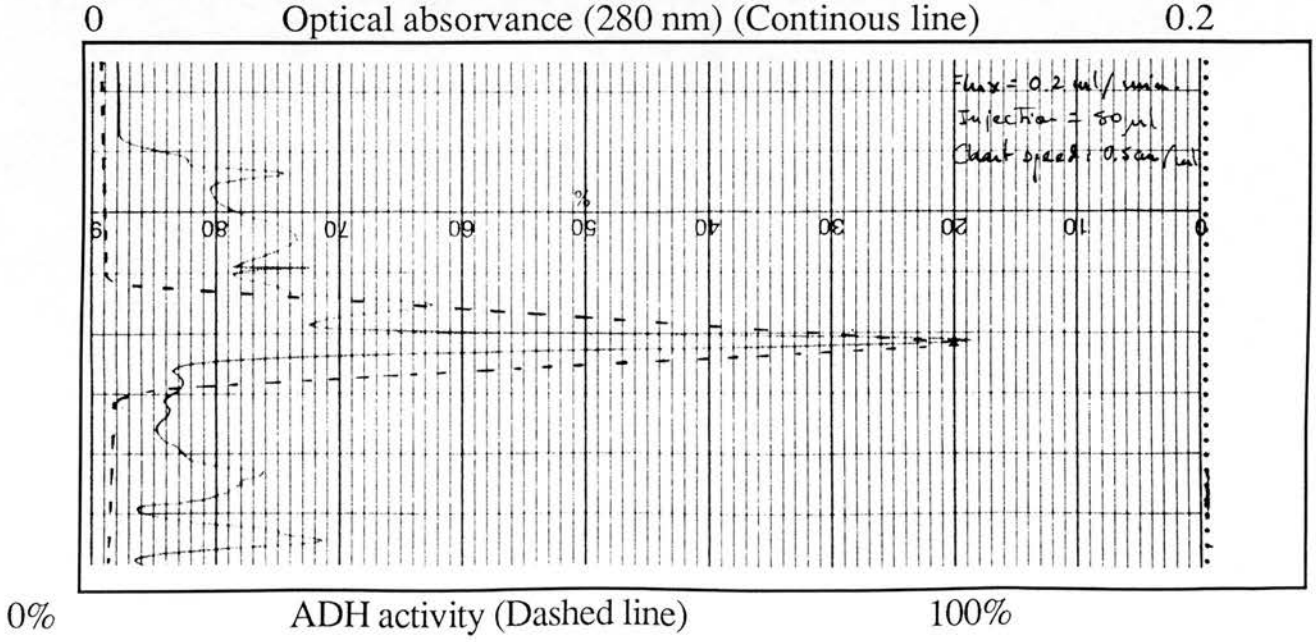


Figure 4.1.5-I. Chromatogram of the Superose 12 results.

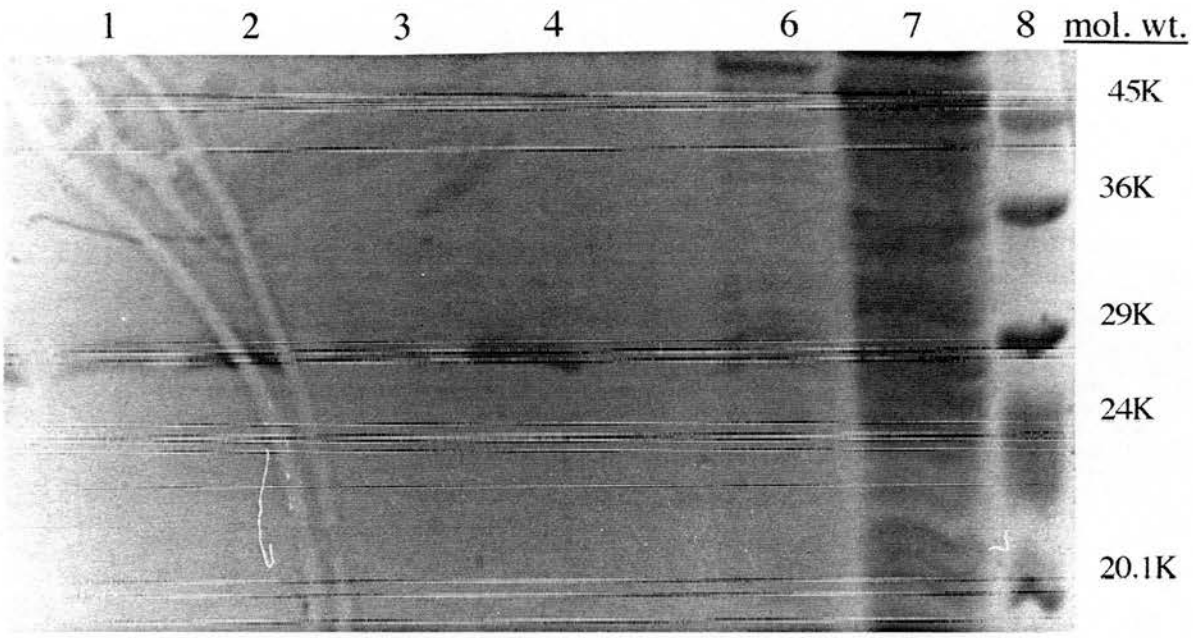


Figure 5.1.5-II. SDS-polyacrylamide gel with samples of various stages of the purification of ADH-D38A. (1-3) Fractions from the Superose-12 column with higher ADH activity.(4) Pure ADH-S control. (5) sample of the fractions with ADH activity pooled after the Q-sepharose column. (6) Crude extract of WVD38A yeast cells. (7) Molecular weight markers.

By reducing the volume of the fractions collected the amount of such contaminant was reduced, but, as a result of its presence the purification of ADH D38A could not be achieved at better than 80% of the total protein. The difference in behaviour of this mutant enzyme is difficult to explain. Although its difference in charge may be responsible for differences in its aggregation properties this enzyme eluted from the Superose column at the same times as those of the rest of the ADHs purified, and so the possibility that D38A could be somehow aggregating with the contaminant protein seems unlikely.

Although this technique allowed for efficient recoveries of purified enzyme it had a major inconvenience in the low capacity of the columns that were available. The volume of the Superose columns was 20 ml, hence limiting the volume of the sample to a maximum of 200 µl if optimal resolution was to be achieved.

As has been described in the previous chapter the peculiar nature of the ionic-exchange method used actually increased the volume of the final sample to typical values of about 15 ml. The processing of these volumes with the Superose 12 columns available would have been extremely slow and inefficient, so the introduction of a concentrating step was decided.

This was achieved by saturating the sample obtained from the Q-Sepharose column with ammonium sulphate, collecting the precipitated protein by centrifugation and resuspending in a smaller volume of buffer (~ 500 µl). This step accomplished the necessity of reducing the total volume of protein sample, and allowed a much more efficient use of the gel-filtration system. However, this step proved to have two main problems which, as can be seen in tables 4.1-1 to 4.1-5 made of it the most inefficient step of the purification procedure.

Firstly, the recoveries of ADH activity from this step were far from ideal, and from 30 to 40 % of the activity recovered from the Q-Sepharose column was normally lost. The explanations for this poor efficiency are probably that : a) the protein concentrations of the samples eluting from the anion-exchange columns were too low to allow a very efficient protein precipitation and, b) the necessity to keep the final volume of the sample to a minimum meant that poor solubilization efficiencies were attained.

Despite the problems described this concentrating step was carried out on all the samples due to its global convenience. There are a number of alternative techniques that should be tried in the future to try to reduce the loss of enzyme activity at this stage of the purification. Techniques of sample concentration by filtration, either using vacuum, tangential flow or centrifugal force could provide better ways to concentrate the samples.

4.1.5 General evaluation of the purification protocol.

To summarize, a rapid, small-scale method to purify *Drosophila* ADH from yeast cultures expressing different forms of the enzyme was developed. The purification grade obtained was to homogeneity in SDS-polyacrylamide gels stained with Coomassie blue, except for the mutant D38A. A second contaminant, that amounted to 20% of the total protein, could not be removed from the preparations of this latter mutant with the method described. This purification method was used to obtain enzyme preparations for the characterization of the effects of the mutations introduced into the wild type *Drosophila* enzyme.

ADH WT	Total volume	Total protein	Enzyme units	Specific activity	%Yield	Purification (times)
C. E.	9 ml	90 mg	93	1u./mg	100	0
A.S. 1	2.4 ml	19.2 mg	72	3.7u./mg	77	3.7
G-25	12 ml	18.5 mg	70	3.8u./mg	75	3.8
Q-Seph.	15 ml	2.6 mg	63.5	24u./mg	68	24
A.S. 2	0.5 ml	2.3 mg	40	17u./mg	43	17
Superose	4.5 ml	1.3 mg	31	24u./mg	33	24

Table 4.1-1. Purification results for ADH-S.

ADH G14A	Total volume	Total protein	Enzyme units	Specific activity	%Yield	Purification (times)
C. E.	15 ml	60 mg	22.5	0.3u./mg	100	0
A.S. 1	4 ml	18 mg	15.8	0.8u./mg	70	3.5
G-25	12 ml	16 mg	14.8	0.9u./mg	66	3.6
Q-Seph.	13 ml	1.9 mg	13.9	7.3u./mg	51	23.5
A.S. 2	0.4 ml	1.3 mg	7.6	5.8u./mg	34	17
Superose	4 ml	0.5 mg	6.3	12.4u./mg	28	22

Table 4.1-2. Purification results for ADH G14A.

ADH G19A	Total volume	Total protein	Enzyme units	Specific activity	%Yield	Purification (times)
C. E.	7 ml	13 mg	10.5	0.8	100	0
A.S. 1	2 ml	3.5 mg	7.6	2.2	76	2.7
G-25	6 ml	3.2 mg	6	1.9	57	2.3
Q-Seph.	8 ml	0.35 mg	4.5	12.9	43	16
A.S. 2	.25 ml	0.22 mg	3.5	15.9	33	20
Superose	2 ml	0.09 mg	2.9	32.6	28	41

Table 4.1-3. Purification results for the enzyme ADH G19A

ADH D38A	Total volume	Total protein	Enzyme units	Specific activity	%Yield	Purification (times)
C. E.	20 ml	85 mg	134	1.6u./mg	100	0
A.S. 1	4 ml	19.7 mg	92	4.6u./mg	69	2.9
G-25	20 ml	19.6 mg	68	3.4u./mg	51	2.2
Q-Seph.	25 ml	4.1 mg	57.5	14u./mg	43	8.7
A.S. 2	0.5 ml	3.4 mg	40	12u./mg	31	7.3
Superose	6 ml	1.6 mg	32	20u./mg	24	12.5

Table 4.1-4. Purification results for the enzyme ADH D38A

ADH P214S	Total volume	Total protein	Enzyme units	Specific activity	%Yield	Purification (times)
C. E.	7 ml	11 mg	5.5	0.5u./mg	100	0
A.S. 1	2 ml	4.4 mg	4.9	1.1u./mg	85	2.2
G-25	6 ml	4.2 mg	4.5	1.1u./mg	82	2.2
Q-Seph.	8 ml	0.8 mg	4.1	5.1u./mg	74	10.2
A.S. 2	0.25 ml	0.6 mg	2.3	3.8u./mg	42	7.6
Superose	2.5 ml	0.25mg	1.9	7.5u./mg	34	15

Table 4.1-5. Purification results for the enzyme ADH P214S.

4.2 Characterization of the wild type and mutant ADH enzymes.

4.2.1 Enzyme kinetics .

Results:

Figures 4.2.1-I to 4.2.1-XIV show the Michaelis-Menten and Hanes plots for each of the enzymes and substrates studied. Table 4.2.1-1 summarizes the Km, k_{cat} and k_{cat}/Km values obtained for all the enzymes analyzed and for all substrates tested.

	NAD			$NADP^+$			Isopropanol		
	Km (mM)	k_{cat} (s^{-1})	k_{cat}/Km	Km (mM)	k_{cat} (s^{-1})	k_{cat}/Km	Km (mM)	k_{cat} (s^{-1})	k_{cat}/Km
ADH wt	0.18	16.6	89.5	1.4	3.9	1.8	0.6	16.7	27.7
G14A	0.36	7.1	19.8	2.4	1.3	0.5	0.64	7.14	11.2
G19A	0.32	17.2	53.7	No detectable activity			0.66	17.2	26.06
D38A	0.45	9	20	4.27	1	0.2	1.99	9	4.5
P214S	0.09	3	33.3	1.9	0.5	0.3	0.47	3	6.66

Table 4.2.1-1 . Kinetic results for all the enzymes and substrates studied. All the rate values measured had standard errors lower than ± 5%, except for two points : 6.4 mM $NADP^+$ for the enzyme ADH D38A (± 24%, figure 4.2.1-X) and 0.8 mM NAD^+ for the enzyme P214S (±9.5%, figure 4.2.1-XII).

Figure 4.2.1-Ia

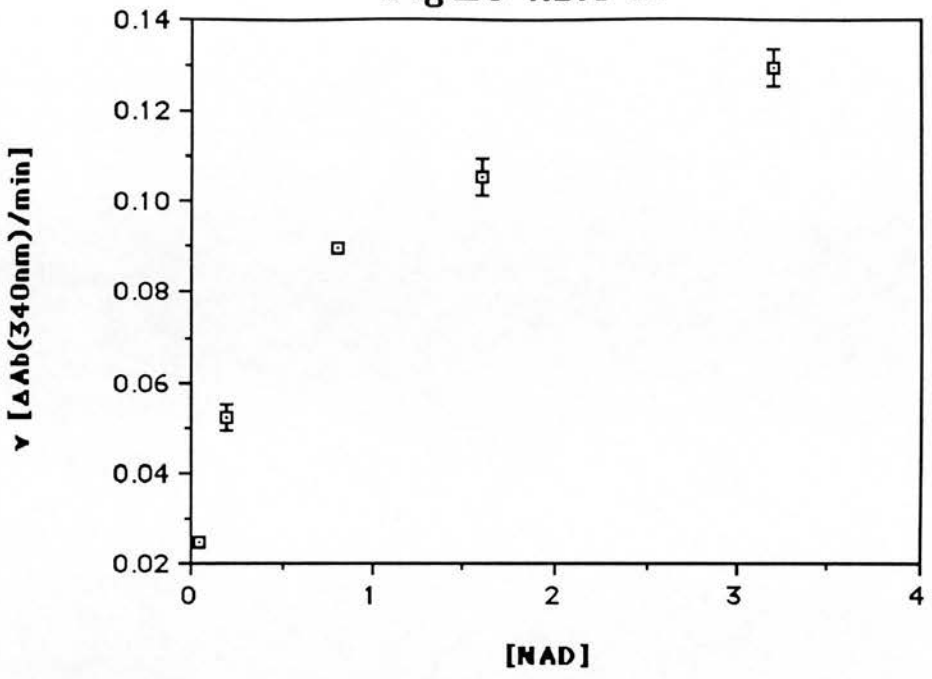


Figure 4.2.1-Ib

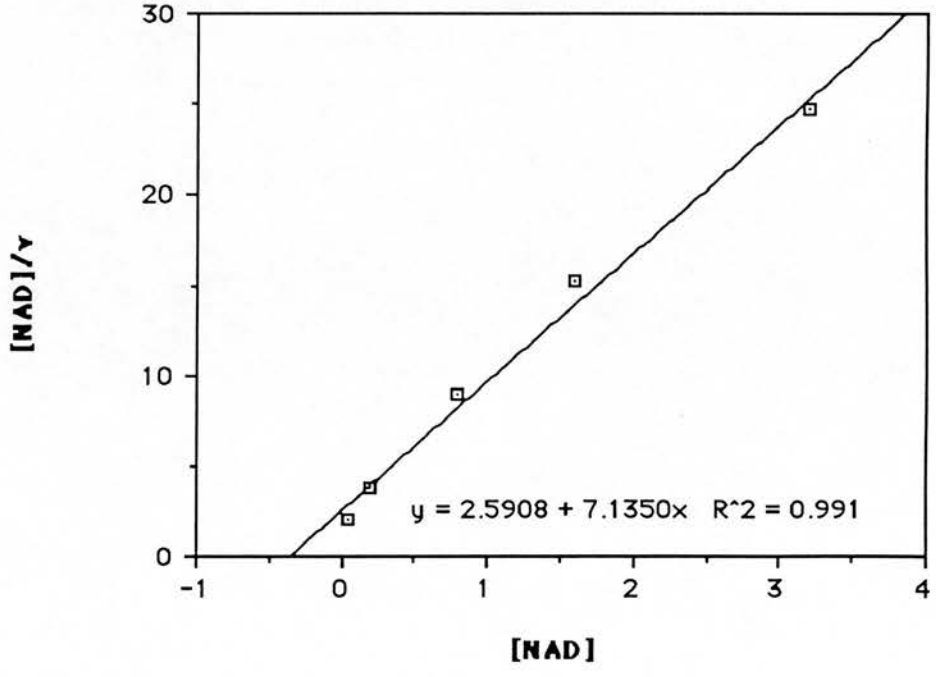


Figure 4.2.1-I . Kinetic plots for ADH wild type and NAD^+, (a) Michaelis-Menten plot. (b) Hanes plot using the mean rate values (simple regression calculated with the program CRICKET GRAPH). The equation defining the regression line and its regression coefficient are shown at the base of the plot.

Figure 4.2.1-IIa

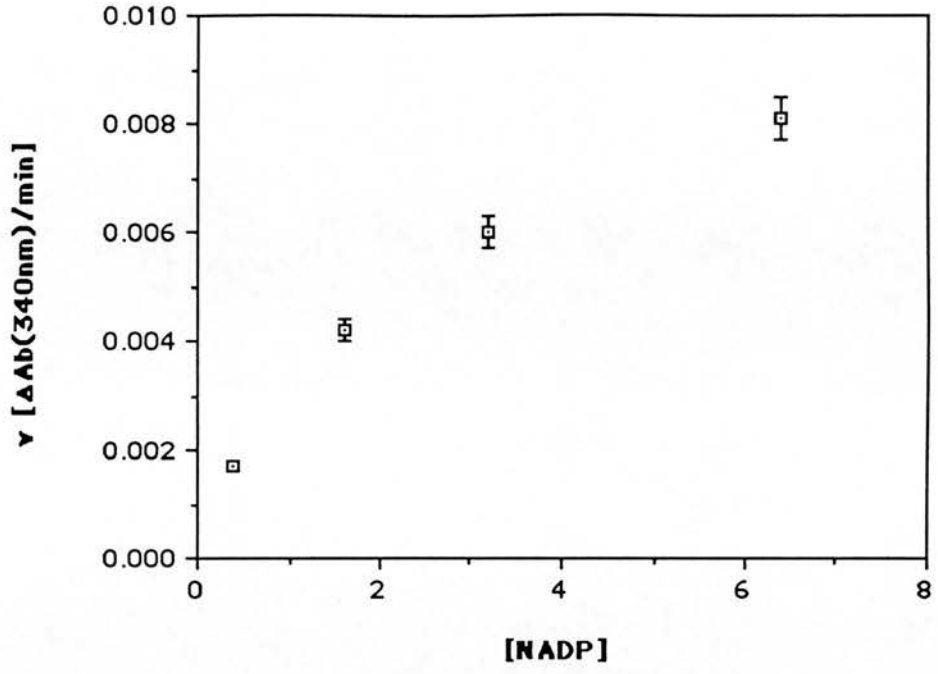


Figure 4.2.1-IIb

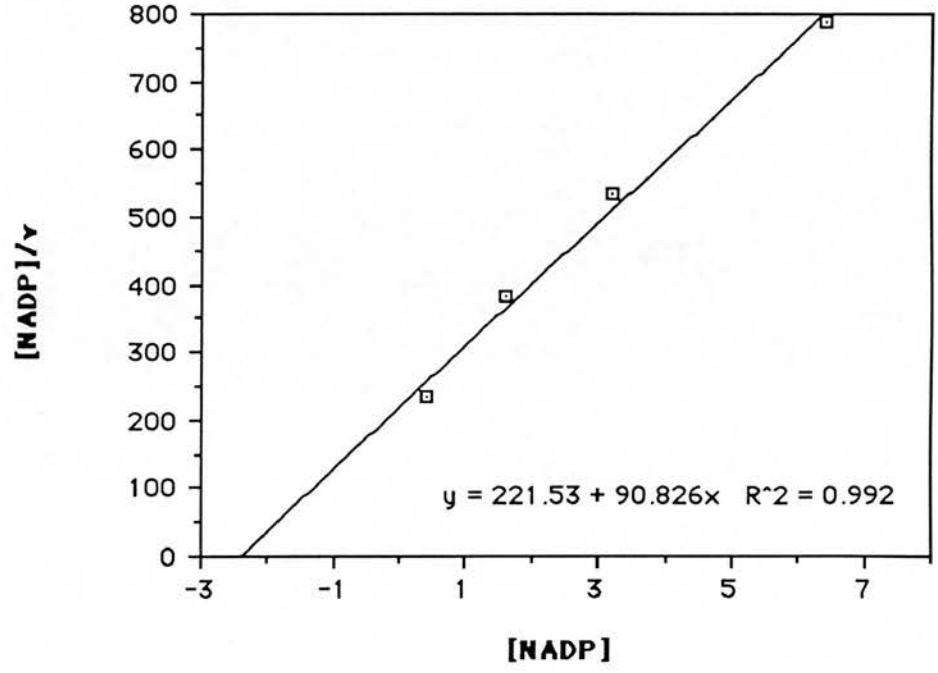


Figure 4.2.1-II. Kinetic plots for ADH wild type and $NADP^+$. (a) Michaelis-Menten plot. (b) Hanes plot using the mean rate values (simple regression calculated with the program CRICKET GRAPH). The equation defining the regression line and its regression coefficient are shown at the base of the plot.

Figure 4.2.1-IIIa

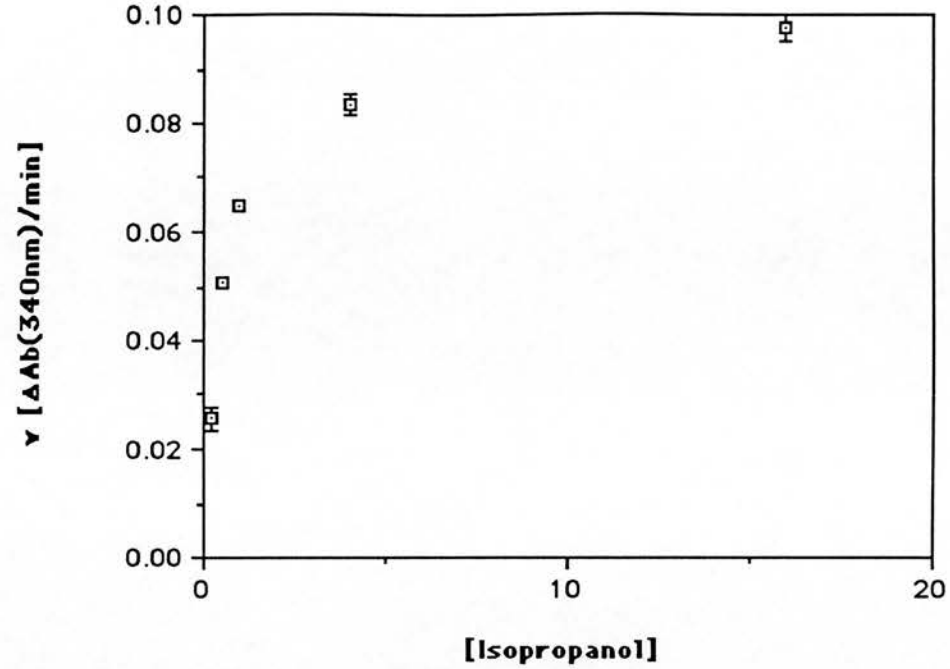


Figure 4.2.1-IIIb

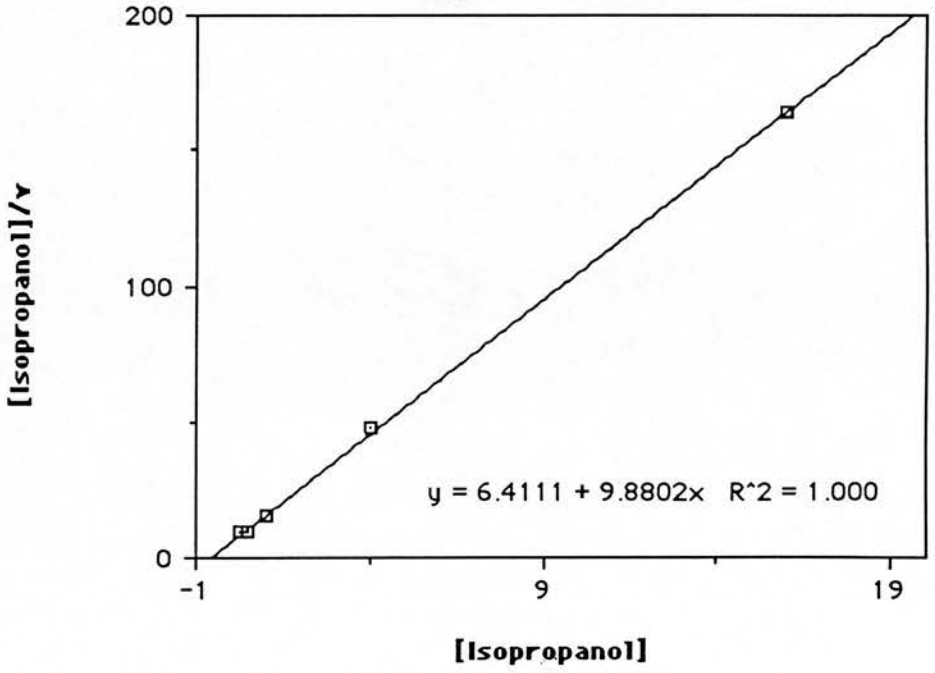


Figure 4.2.1-III . Kinetic plots for ADH wild type and isopropanol. (a) Michaelis-Menten plot. (b) Hanes plot using the mean rate values (simple regression calculated with the program CRICKET GRAPH). The equation defining the regression line and its regression coefficient are shown at the base of the plot.

Figure 4.2.1-IVa

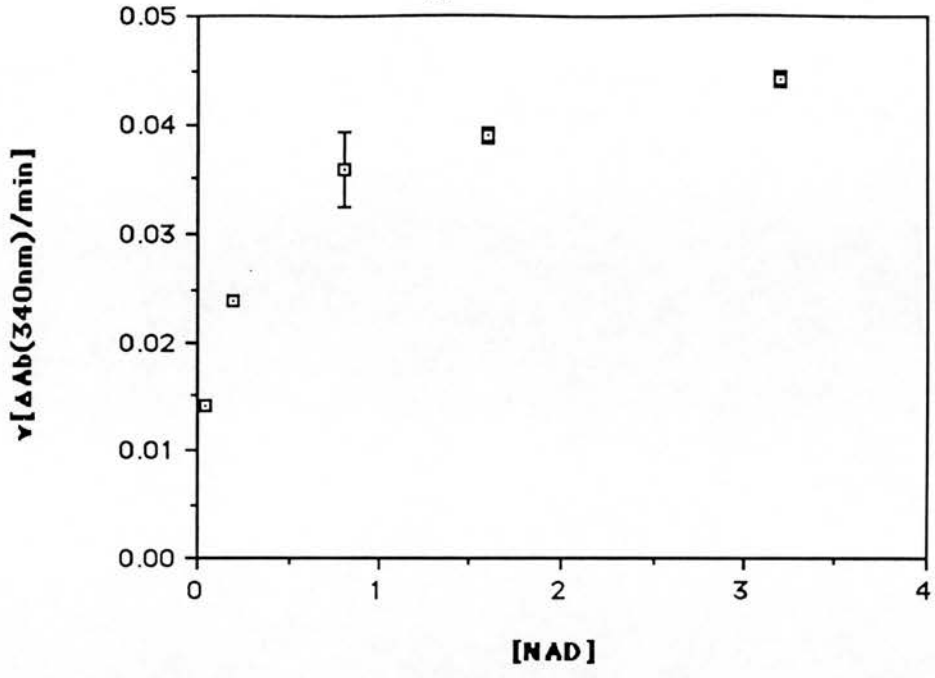


Figure 4.2.1-IVb

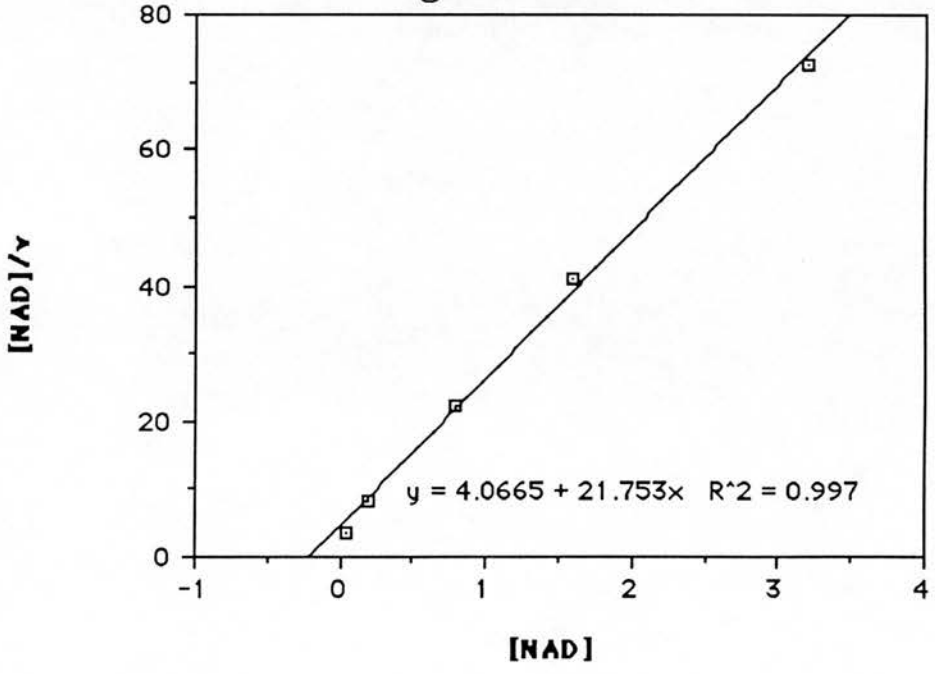


Figure 4.2.1-IV. Kinetic plots for ADH G14A and NAD^+. (a) Michaelis-Menten plot. (b) Hanes plot using the mean rate values (simple regression calculated with the program CRICKET GRAPH). The equation defining the regression line and its regression coefficient are shown at the base of the plot.

Figure 4.2.1-Va

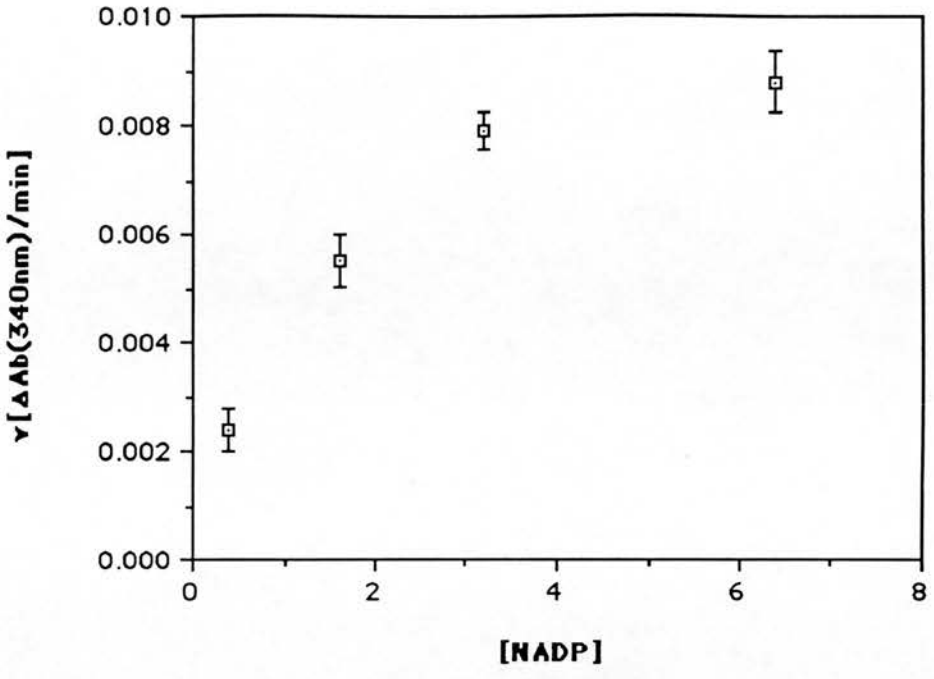


Figure 4.2.1-Vb

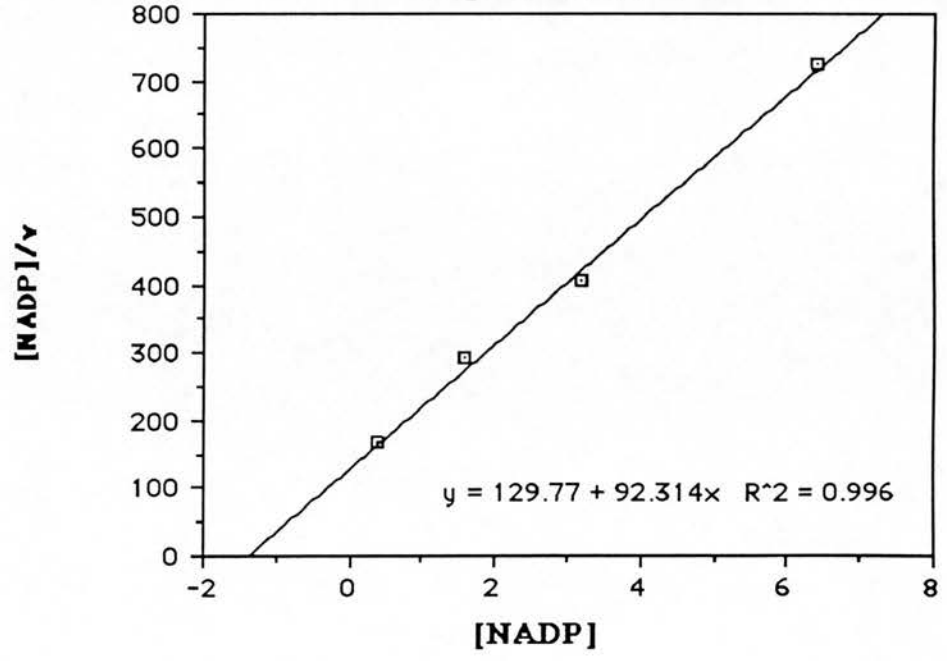


Figure 4.2.1-V. Kinetic plots for ADH G14A and $NADP^+$. (a) Michaelis-Menten plot. (b) Hanes plot using the mean rate values (simple regression calculated with the program CRICKET GRAPH). The equation defining the regression line and its regression coefficient are shown at the base of the plot.

Figure 4.2.1-VIa

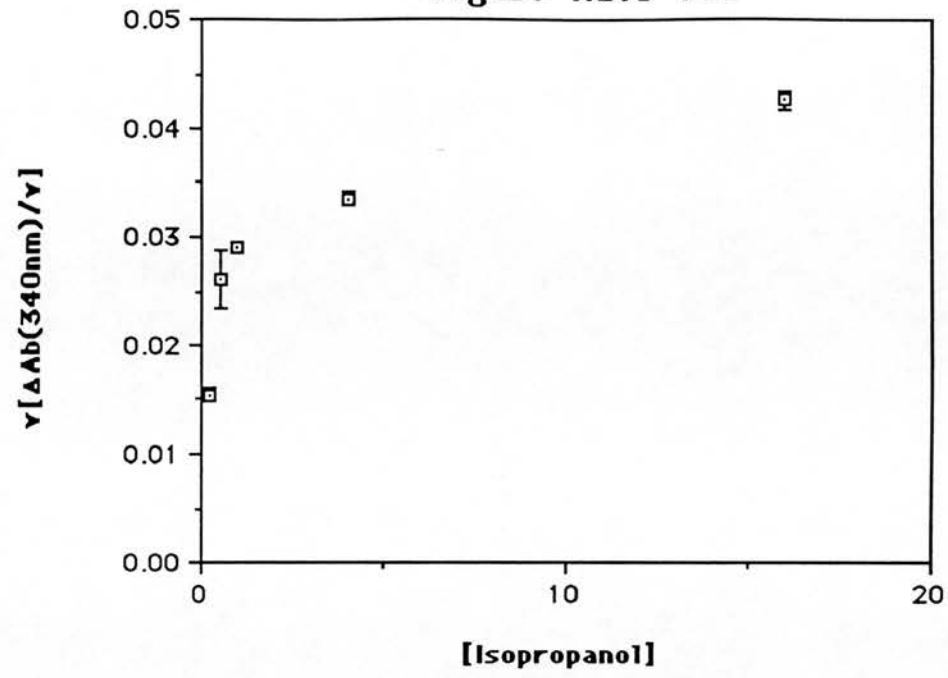


Figure 4.2.1-VIb

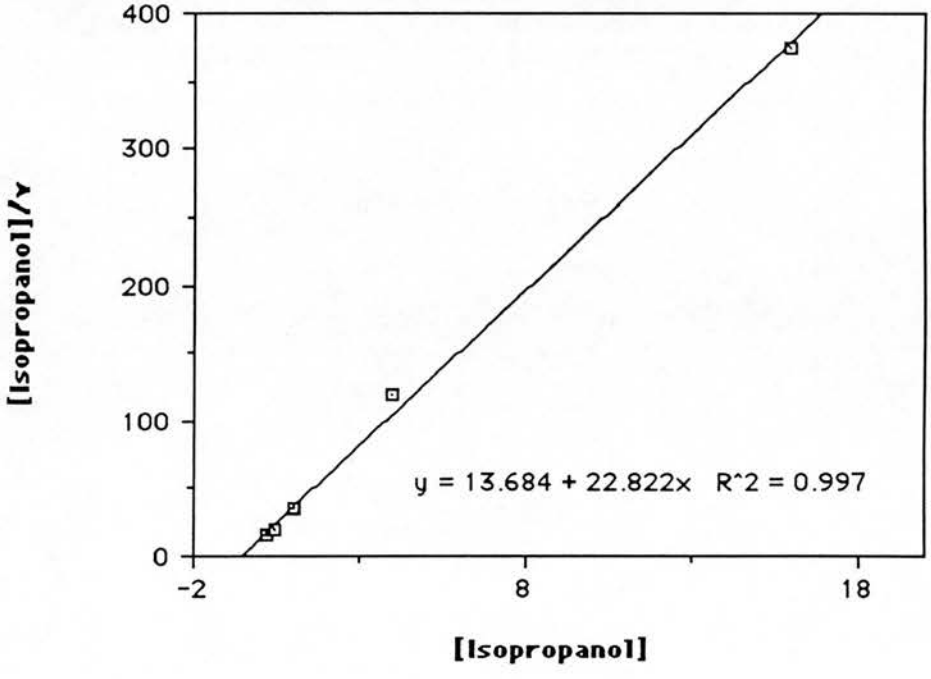


Figure 4.2.1-VI. Kinetic plots for ADH G14A and isopropanol. (a) Michaelis-Menten plot. (b) Hanes plot using the mean rate values (simple regression calculated with the program CRICKET GRAPH). The equation defining the regression line and its regression coefficient are shown at the base of the plot.

Figure 4.2.1-VIIa

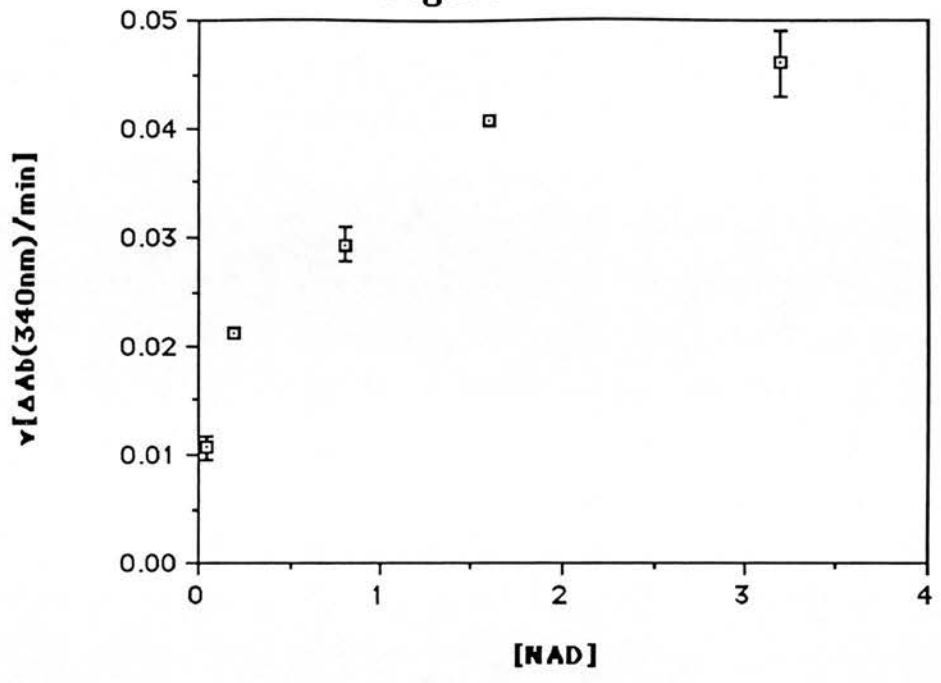


Figure 4.2.1-VIIb

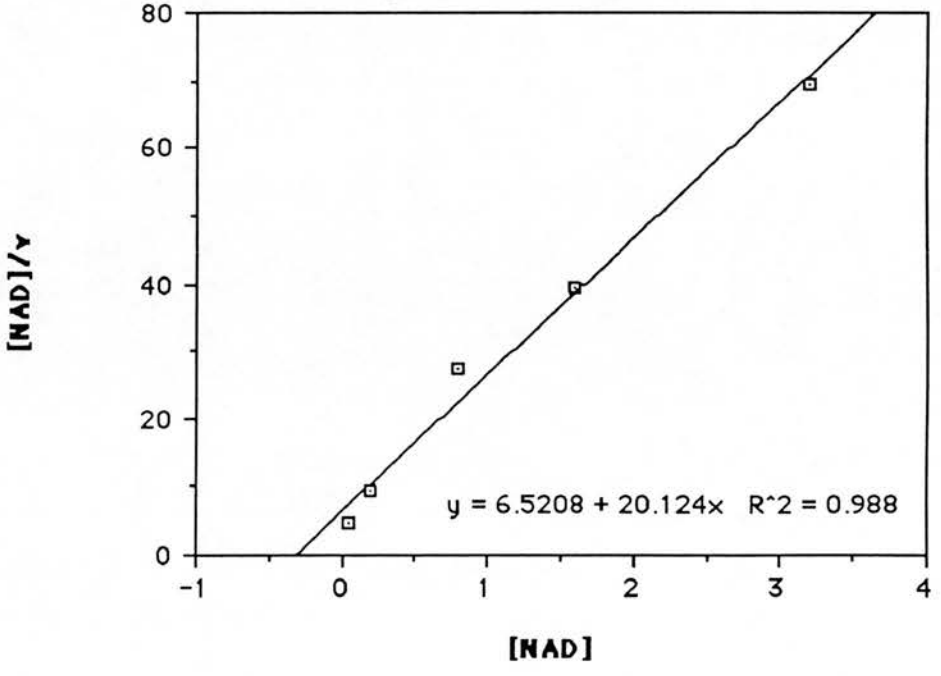


Figure 4.2.1-VI. Kinetic plots for ADH G19A and NAD^+. (a) Michaelis-Menten plot. (b) Hanes plot using the mean rate values (simple regression calculated with the program CRICKET GRAPH). The equation defining the regression line and its regression coefficient are shown at the base of the plot.

Figure 4.2.1-VIIIa

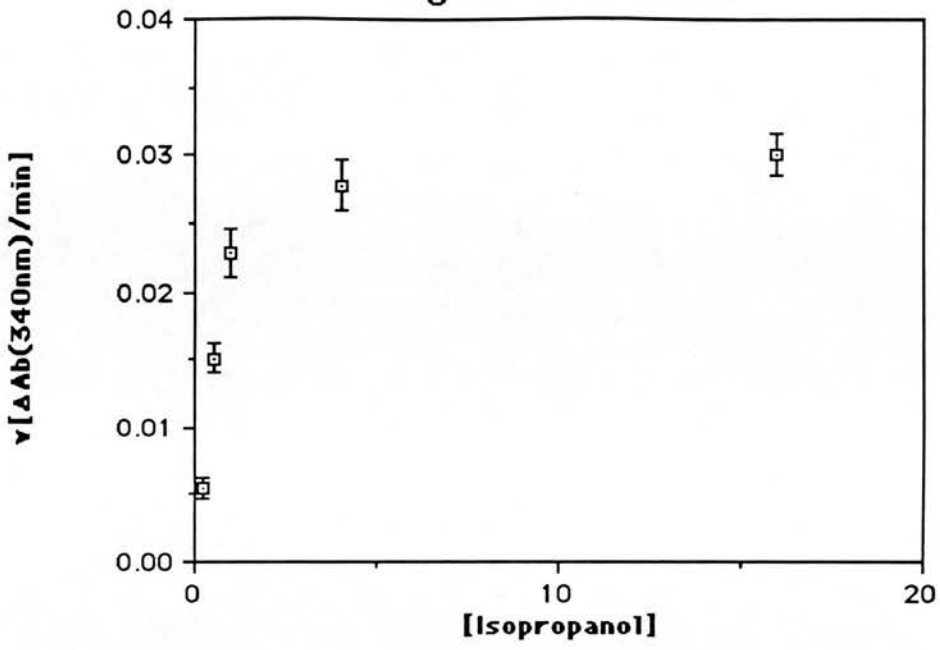


Figure 4.2.1-VIIIb

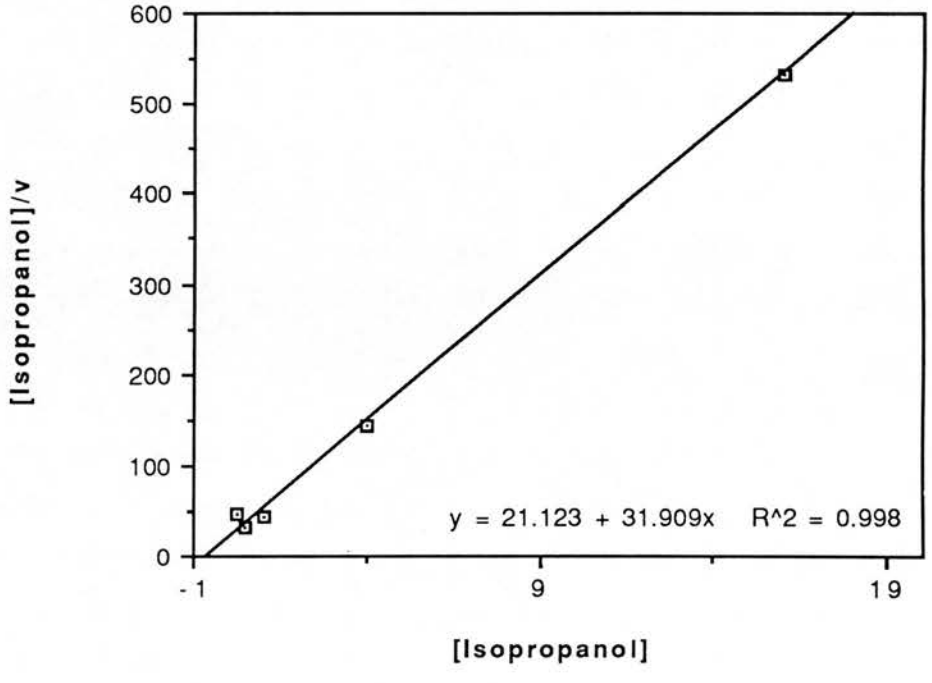


Figure 4.2.1-VIII. Kinetic plots for ADH G19A and isopropanol. (a) Michaelis-Menten plot. (b) Hanes plot using the mean rate values (simple regression calculated with the program CRICKET GRAPH). The equation defining the regression line and its regression coefficient are shown at the base of the plot.

Figure 4.2.1-IXa

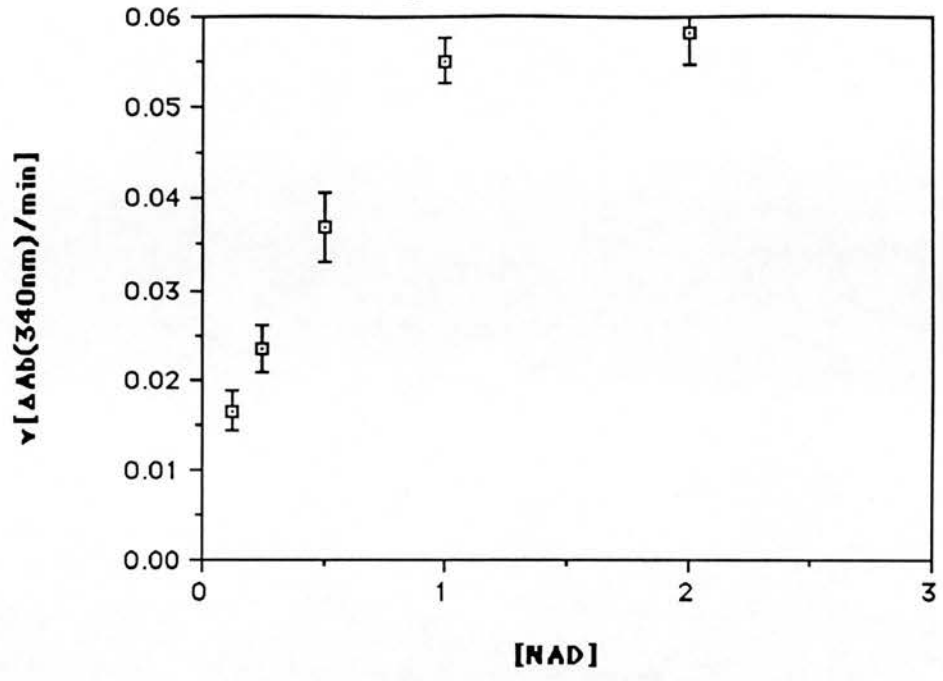


Figure 4.2.1-IXb

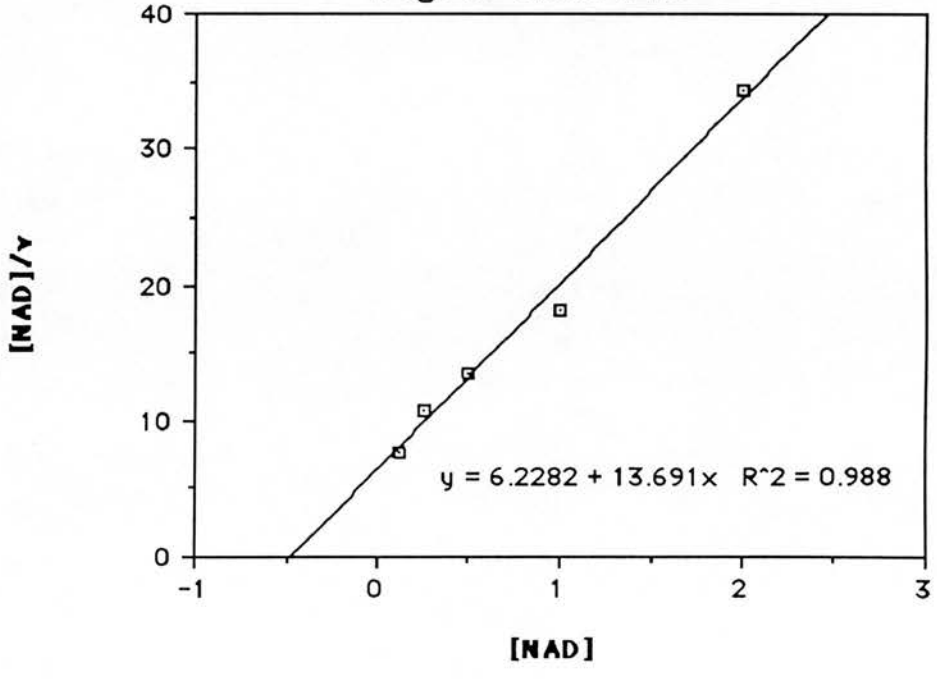


Figure 4.2.1-IX. Kinetic plots for ADH D38A and NAD^+. (a) Michaelis-Menten plot. (b) Hanes plot using the mean rate values (simple regression calculated with the program CRICKET GRAPH). The equation defining the regression line and its regression coefficient are shown at the base of the plot.

Figure 4.2.1-Xa

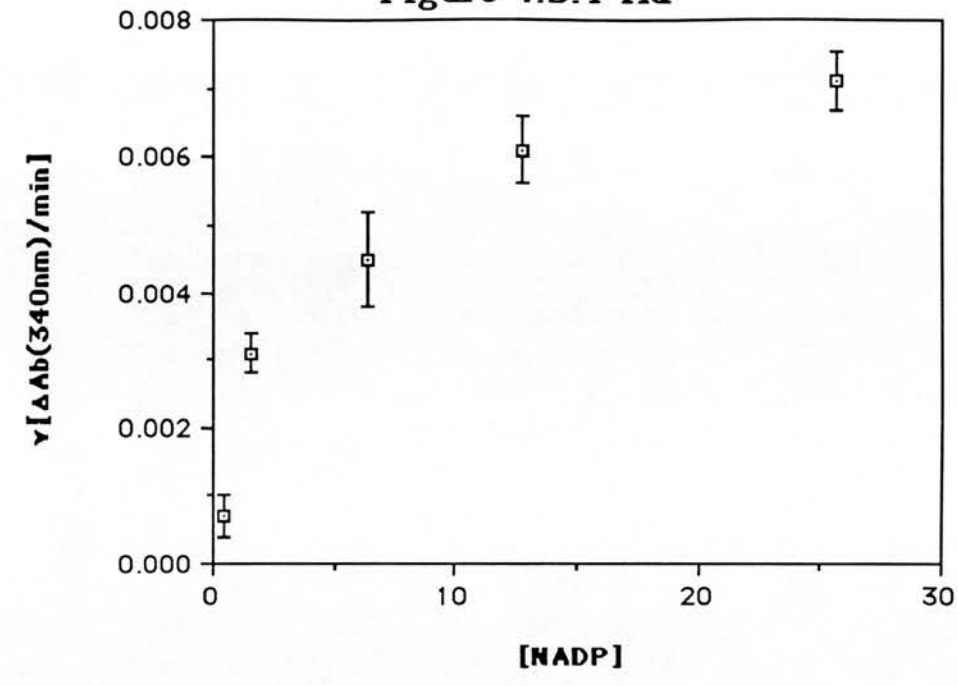


Figure 4.2.1-Xb

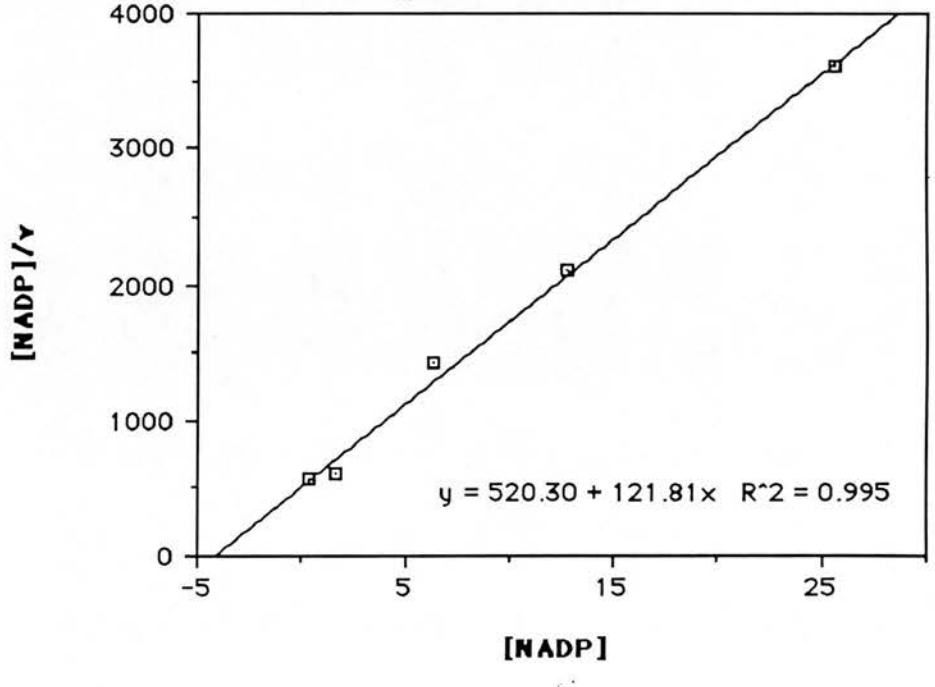


Figure 4.2.1-X. Kinetic plots for ADH D38A and $NADP^+$. (a) Michaelis-Menten plot. (b) Hanes plot using the mean rate values (simple regression calculated with the program CRICKET GRAPH). The equation defining the regression line and its regression coefficient are shown at the base of the plot.

Figure 4.2.1-XIa

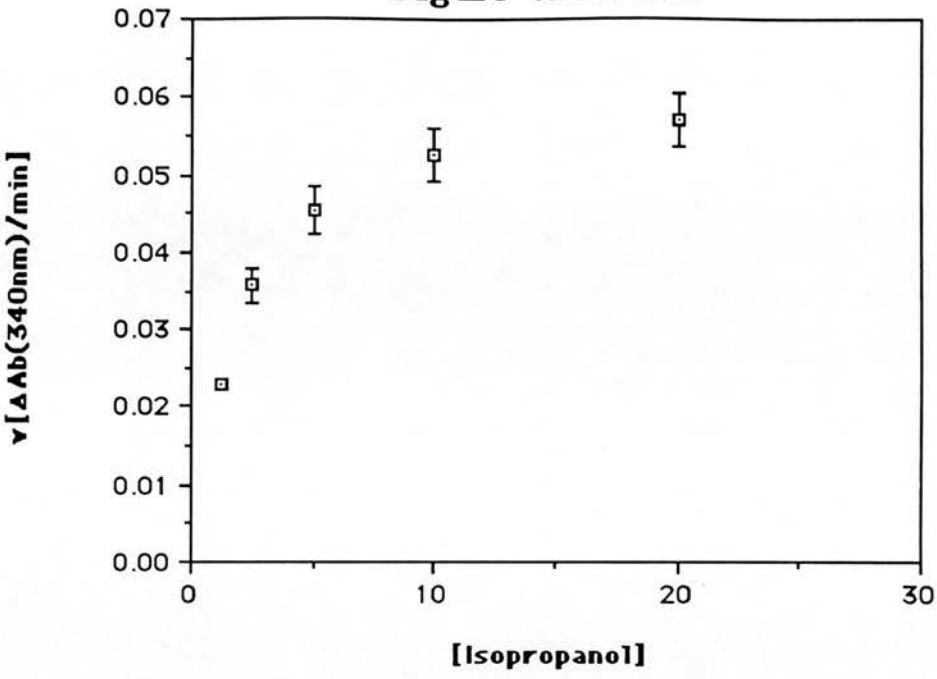


Figure 4.2.1-XIb

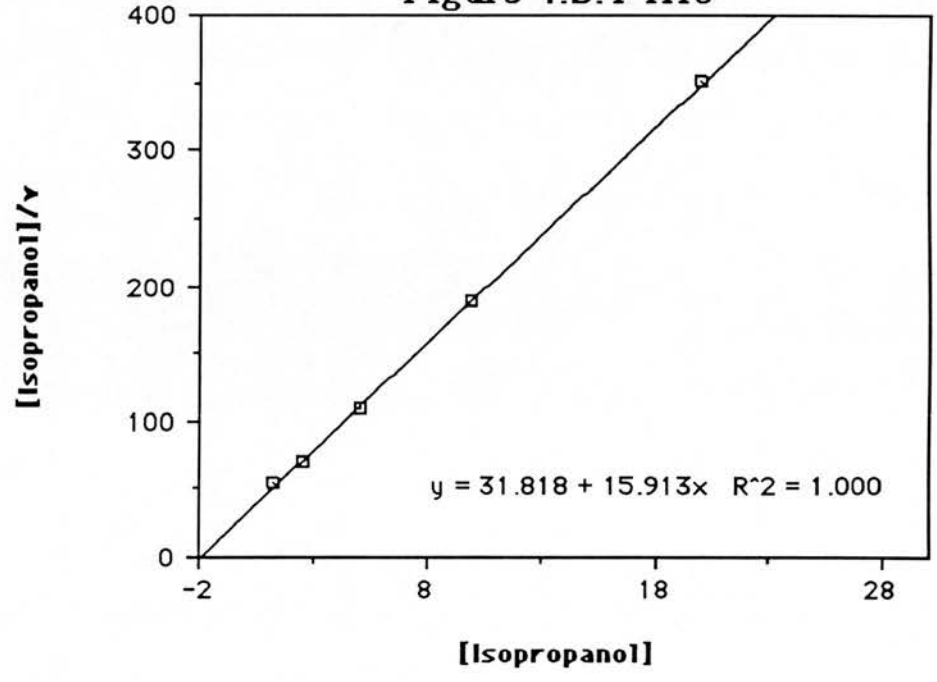


Figure 4.2.1-XI. Kinetic plots for ADH D38A and isopropanol. (a) Michaelis-Menten plot. (b) Hanes plot using the mean rate values (simple regression calculated with the program CRICKET GRAPH). The equation defining the regression line and its regression coefficient are shown at the base of the plot.

Figure 4.2.1-XIIa

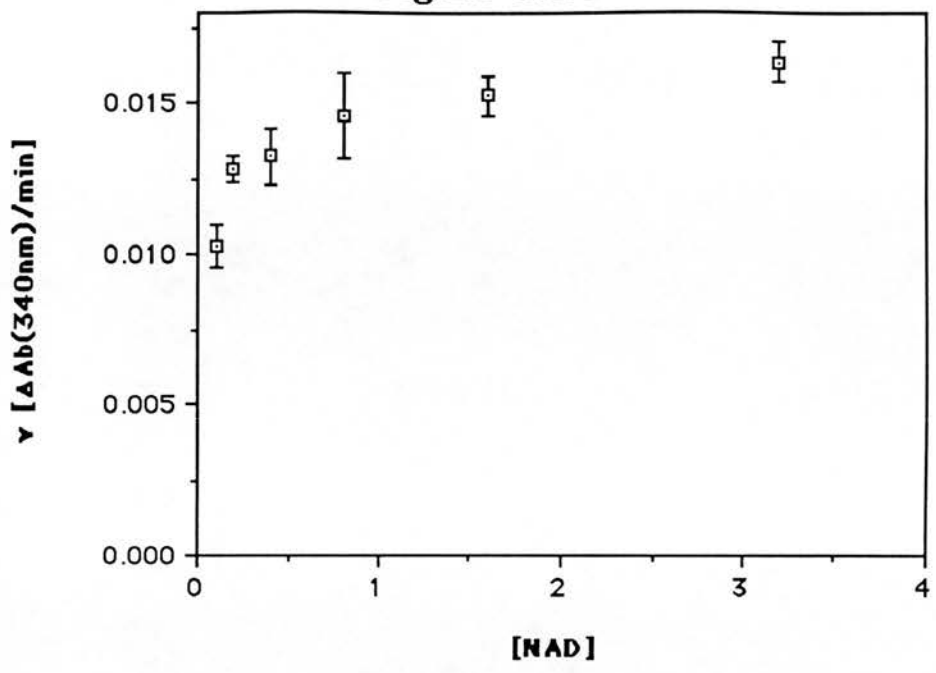


Figure 4.2.1-XIIb

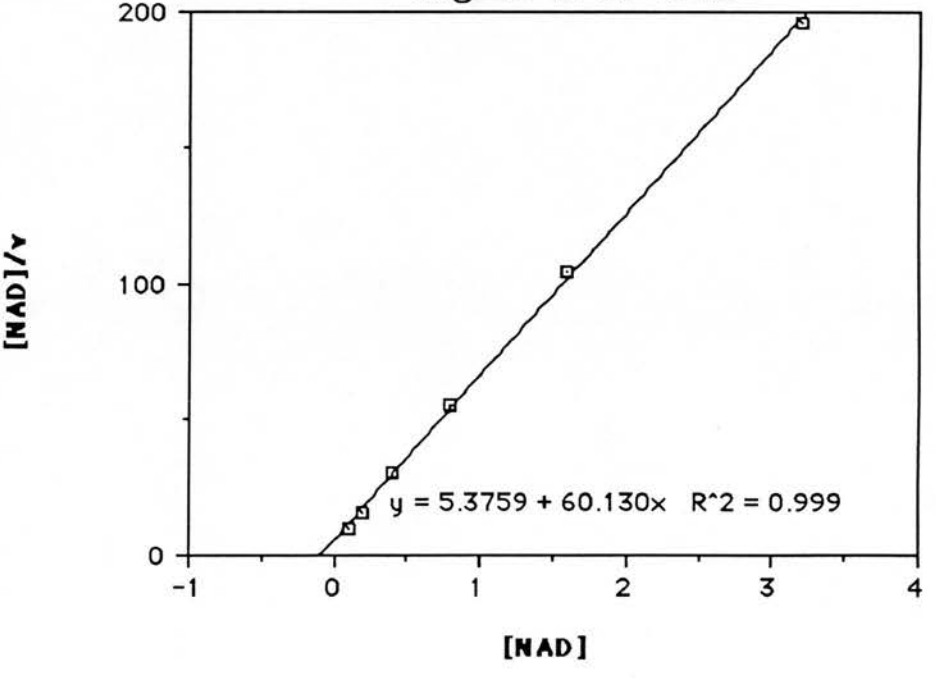


Figure 4.2.1-XII . Kinetic plots for ADH P214S and NAD^+. (a) Michaelis-Menten plot. (b) Hanes plot using the mean rate values (simple regression calculated with the program CRICKET GRAPH). The equation defining the regression line and its regression coefficient are shown at the base of the plot.

Figure 4.2.1-XIIIa

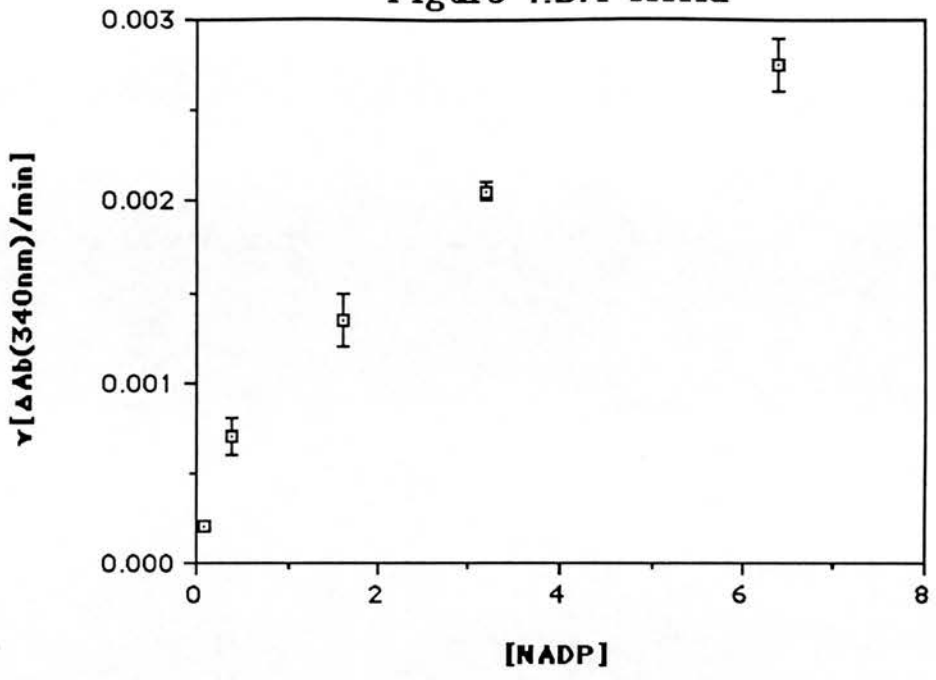


Figure 4.2.1-XIIIb

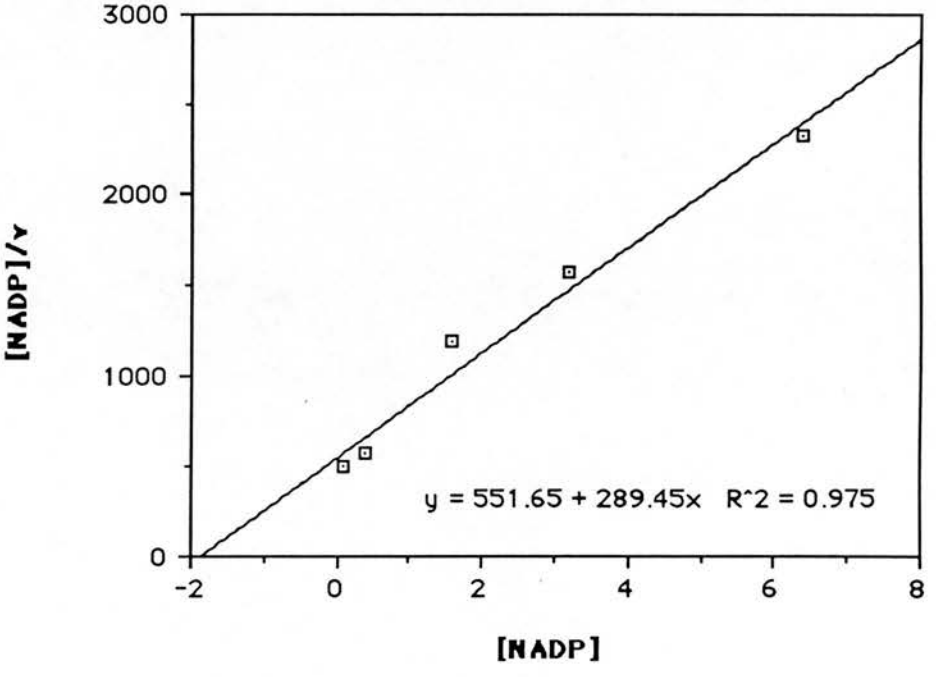


Figure 4.2.1-XIII. Kinetic plots for ADH P214S and $NADP^+$. (a) Michaelis-Menten plot. (b) Hanes plot using the mean rate values (simple regression calculated with the program CRICKET GRAPH). The equation defining the regression line and its regression coefficient are shown at the base of the plot.

Figure 4.2.1-XIVa

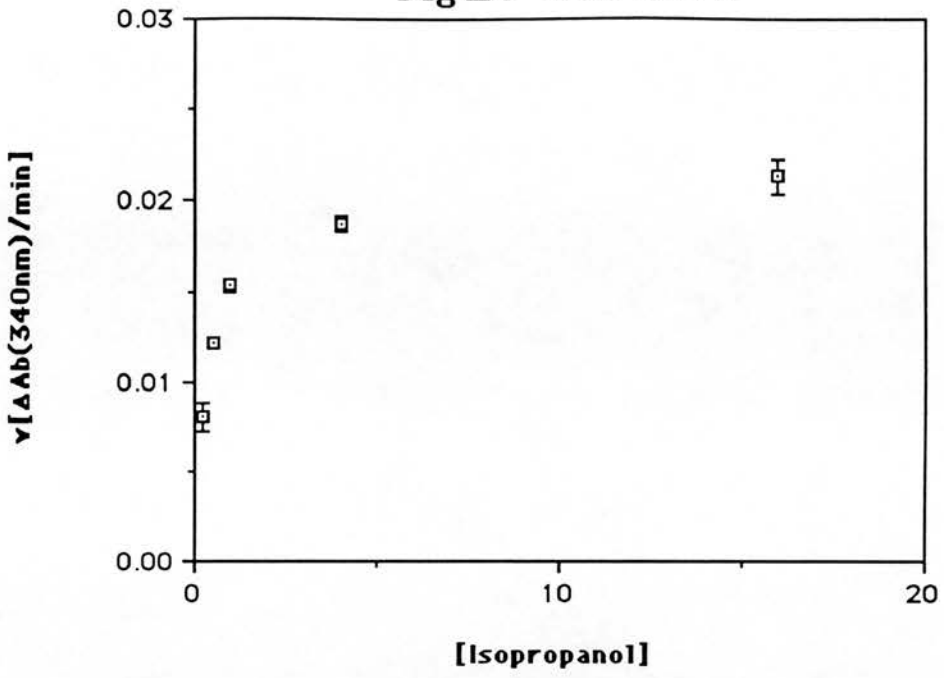


Figure 4.2.1-XIVb

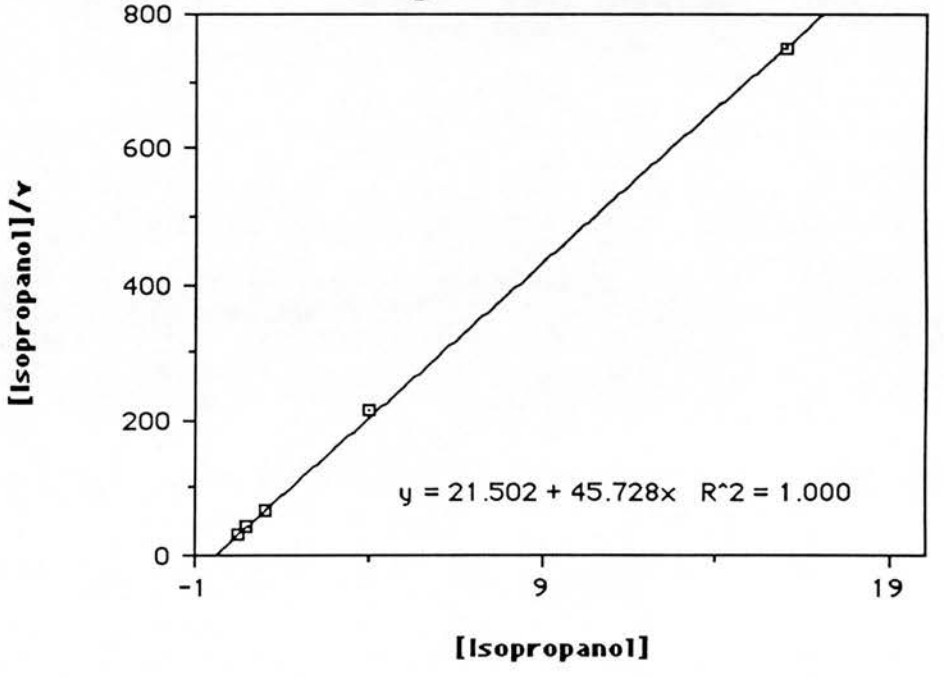


Figure 4.2.1-XIV. Kinetic plots for ADH P214S and isopropanol. (a) Michaelis-Menten plot. (b) Hanes plot using the mean rate values (simple regression calculated with the program CRICKET GRAPH). The equation defining the regression line and its regression coefficient are shown at the base of the plot.

Discussion:

As described in the introduction chapter, Drosophila ADH catalyzes the oxidation of alcohols through a ternary-complex, ordered mechanism, where NAD^+ binds first to the enzyme. The release of the reduced cofactor is the limiting step of the reaction, and it can be seen by studying different isoenzymes that an increase in affinity for the cofactor results in a slower reaction due to a decrease in the dissociation constant of NADH (Chambers, 1988).

This immediately suggests that a mutation that may be capable of decreasing the affinity constant of the enzyme for the cofactor without affecting the reaction rate will result in a faster enzyme. The Michaelis constant (Km) is not a simple measure of affinity, it is negatively dependent on the constant of formation of the enzyme-substrate complex, and positively dependent on the dissociation constant of such a complex plus the constant of its transformation into the complex enzyme-substrate.

Due to this fact, it is expected that mutations that increase the Km values for a substrate in ADH will fall into two categories: those that affect the affinity of the enzyme for the cofactor but do not affect the rate of reaction and those that, while affecting the affinity for NAD, do also decrease its rate of reduction. The first class of mutations would be expected to accelerate the reaction rate of the enzyme, because a lower affinity of the enzyme for the cofactor would allow the release of NADH to occur at higher rate, hence approaching the reaction rate to the rate of proton transfer. The second class of mutants would probably be slower enzymes, because of loss of reaction efficiency.

Clearly, any mutation is likely to have a certain effect on all the rates, and its global effect in the Km will be the result of the sum of the effects in all the constants. The two cases described above are the extreme cases where one constant is affected by the substitution while the others are not. However, from the comparison of the values of the different mutants it should be possible to draw conclusions on the predominant effect of each change.

Enzyme ADH-S:

The kinetic results obtained with the ADH-S enzyme expressed and purified in this project are similar to the various kinetic data published. Perhaps more importantly, they are very similar to the kinetic values published by Chen and colleagues (1990), which were used as reference for their site-directed mutagenesis experiments on the enzyme. This result has two main implications for this project: it makes the kinetic results directly comparable to the ones reported by Chen and colleagues, and it shows that the amino acid change at position 173 caused by the unexpected change found in the gene sequence has no effect on the enzyme's mechanism.

Enzyme ADH G14A:

The G14A mutation has already been analyzed by other investigators (Chen et al., 1990), and they reported similar kinetic effects to the ones found in this work. The enzyme shows a two fold increase in the Km for both NAD^+ and $NADP^+$, while it seems to have no effect in the recognition of isopropanol. These results agree with the expected positioning of G14 in the enzyme, at the nucleotide binding site.

The fact that the k_{cat} is also reduced two fold may indicate that the mutation somehow disrupts the normal positioning of the NAD, impairing the catalytic efficiency of the enzyme. Because the same kind of effect is seen for both NAD^+ and $NADP^+$ it is likely that the residue G14 does not interact with the phospho group of $NADP^+$.

Enzyme ADH G19A:

This mutation also produces a two-fold increase in Km for the cofactor but, interestingly, it does not impair the speed of the reaction. In fact, a small increase in the k_{cat} value is obtained. This increase is probably not significant, but it shows that in this mutant, the loss in cofactor recognition is probably compensated by a faster rate of enzyme turn-over, which leaves the catalytic efficiency of the enzyme untouched.

A very important indication of the spatial positioning of this residue is given by the fact that no detectable activity could be measured with this mutant when $NADP^+$ was used as a cofactor. This seems to indicate that residue 19 is in close proximity to the phospho group of $NADP^+$, and that the introduction of a methyl group totally disrupts its environment. Again the mutation does not seems to have any effect on substrate recognition, supporting a location distant from the alcohol binding site and from the reaction centre.

Enzyme ADH D38A:

Mutations in this amino acid have already been reported by Chen and colleagues (1991), although no substitution to alanine was done in their work. Chen and colleagues worked on the hypothesis that D38 may be the residue that is responsible for hydrogen-bonding the hydroxyl groups of the adenine-ribose of NAD.

As mentioned in the introduction, this hypothesis is based on the general structure of NAD^+ binding sites, although some alignment results and the structure of SDH indicate that such interactions may be between the cofactor and residue 64, another aspartic acid (Marekov et al., 1991; Ghosh et al., 1991).

Chen and colleagues found that mutating D38 to hydrophobic (leucine) or positively charged (arginine) residues resulted in inactivation of the enzyme. But mutating the aspartic side-chain to asparagine, hence removing its charge but maintaining its capacity to form hydrogen bonds, resulted in an enzyme capable of utilizing $NADP^+$ as efficiently as NAD.

The mutation to asparagine increased the Km for NAD^+ by 50%, and also increased the catalytic efficiency (k_{cat}) two fold. This is expected since, as discussed, a mutation that removes binding energy between the cofactor and enzyme without affecting their three-dimensional positioning will produce a faster enzyme. The same mutation, however, reduced the Km for $NADP^+$ by a factor of 60 and increased the k_{cat} by a factor of 10, making the enzyme almost as efficient with $NADP^+$ as with NAD.

In their work the authors argue that the negative charge of aspartic acid must be responsible for the low affinity of ADH for $NADP^+$, due to direct charge repulsion between the side-chain and the phospho group of the cofactor. They propose that such repulsive interaction is eliminated when a glutamine is introduced at position 38, producing an enzyme capable of using both NAD^+ and $NADP^+$.

There can be little doubt than the removal of the negative charge of the aspartic acid allows the enzyme to use $NADP^+$. However that is not necessarily an indication of direct interaction between the cofactor and the enzyme. In fact, Chen et al. do not try to explain how is the extra

phospho group is sterically accommodated in the NAD^+ binding site of the enzyme. This is of major importance, because if residue 38 is located within hydrogen bonding distance of the hydroxyl groups of the adenine ribose a substitution of one of these hydroxyls by a phospho group is likely to create pure steric clashes that would not be avoided if a residue of similar size to aspartate is placed at position 38.

Nothing can be concluded from the inactivation that the authors find when substituting the residue 38 by leucine and arginine, because the structural effects of these mutations were not investigated.

The introduction of an alanine at position 38 done as part of this project introduces at this position a side-chain of smaller size, not capable of creating hydrogen bonds. With respect to NAD^+ the mutation results in a two-fold increase in Km and a two-fold decrease in k_{cat}, indicating that the mutation probably affects both the affinity for the cofactor and the reaction efficiency.

Since the enzyme is active and the mutation abolishes the negative charge of the aspartate 38, if the proposition of Chen and colleagues is true one would expect this mutant to have better kinetic constants for $NADP^+$ than the wild type enzyme. Instead ADH D38A shows an 3-fold increase in Km for $NADP^+$, and a four fold decrease in k_{cat}, suggesting that the effect of the mutation has similar effects in the kinetics of both NAD^+ and $NADP^+$.

The kinetic results for isopropanol of ADH D38A are surprising, since the mutation increases the Km for the alcohol by three-fold. This result could suggests that residue 38 may be positioned close to the active site, or interacting with residues involved in the recognition of the alcohol. This would argue against the proposed hypothesis of residue 38 having a similar role as that in long-chain dehydrogenases, and, by default, would support the works of Marekov and colleagues

(1991) and Ghosh and colleagues (1991). Surprisingly, Chen and colleagues (1991) do not report the kinetic constants for alcohol of their D38N mutant.

The results obtained for our mutant at position 38 could have been affected by the fact that the purity of the enzyme was not higher than 80%, with a second contaminant that could not be removed (see purification results). However, interferences from such contaminant seem unlikely because it has no isopropanol dehydrogenase activity, since crude extracts of WV.36.201 cells show no such activity.

Enzyme ADH P214S:

The P214S substitution is found in nature, but only in ADH-F type isoenzymes. It seems to have two main consequences: it largely increases the thermal stability of the molecule and it changes the kinetic characteristics of the enzyme making it more similar to ADH-S. (Chambers, 1984; Chambers, 1988; Winberg, 1989).

The reasons for the increase in thermal stability are not clear, but the change in the kinetic constants are thought to be due to an increase in the affinity of the enzyme for the cofactor, which, as discussed, does slow down the enzyme velocity (Winberg, 1989).

The same mutation at position 214 has not been reported in an ADH-S enzyme. That fact becomes clear when the kinetic constants of the ADH P214S produced in this project are analyzed. The enzyme has a two-fold reduced Km for the cofactor in respect to the wild type enzyme which, as expected, results in a slower reaction speed, with a k_{cat} reduced by a factor of 5.5, making this enzyme the slowest of all those tested.

The mutation has little effect on the Km for isopropanol, causing a reduction of about 20%. Winberg and colleagues (1982) have proposed that the substrate binds to ADH in two hydrophobic pockets.

The residue at position 214 is clearly affecting the interaction with the cofactor NAD, and, since it is located at he C-terminal part of the enzyme, is also possible that it could be in the environments of the reaction centre. However, from the kinetic results obtained in this project it seems likely that this residue is not in direct contact with the substrate.

The effect of this mutation on the stability of the enzyme required other kinds of analysis to be investigated. They will be described in the next sections of this chapter.

4.2.2 Guanidine-induced denaturation studies of *Drosophila* ADH monitored by CD and enzyme activity .

4.2.2.1 Studies based on enzyme activity measurements.

Figures 4.2.2-I to 4.2.2-VI show the results obtained from monitoring the enzyme activity remaining in enzyme solutions incubated in different concentrations of guanidine chloride. The second part of each figure shows the transformation used to calculate, in each case, the $\Delta G_{(water)}$ (units: cal/mol) of the holoenzymes and of the enzymes saturated with cofactor NAD^+. The results are joined and processed in table 4.2.2.-1.

	$GdnHCl_{50}$		$\Delta G_{(water)}$(Kcal/mol)		$\Delta\Delta G_{(NAD)}$
	$-NAD^+$	$+NAD^+$	$-NAD^+$	$+NAD^+$	
ADH-F	0.4	0.58	3.2	4.5	1.3
ADH-S	0.38	0.74	3.9	7.4	3.5
ADH-G14A	0.26	0.52	2.6	5.9	3.3
ADH-G19A	0.18	0.19	2.8	3.2	0.4
ADH-D38A	0.28	0.31	2.4	2.7	0.3
ADH-P214S	0.64	0.95	7.7	11.4	3.7

Table 4.2.2-1. Joined results of the denaturation studies monitored by enzyme activity. The "$GdnHCl_{50}$" values show the guanidine hydrochloride concentrations at which 50% of the initial activity is lost. $\Delta\Delta G_{(NAD)}$ values are the difference between the extrapolated $\Delta G_{(water)}$ value of each enzyme with and without NAD^+ .

Figure 4.2.2-Ia

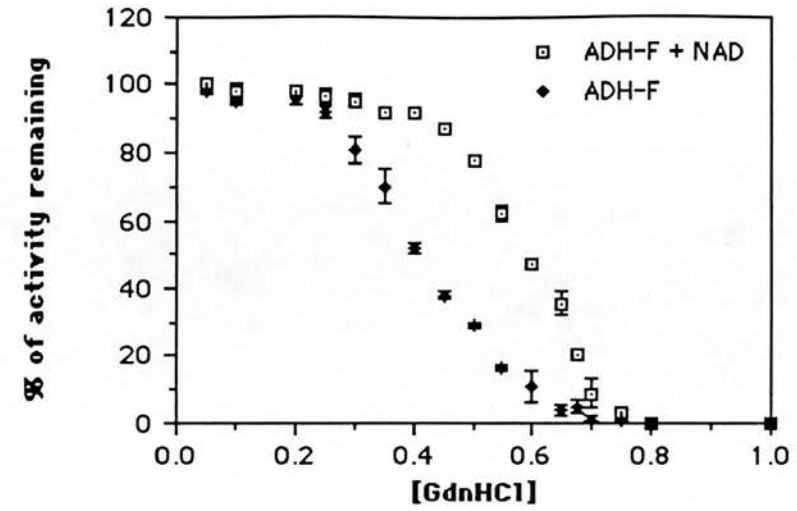


Figure 4.2.2-Ib

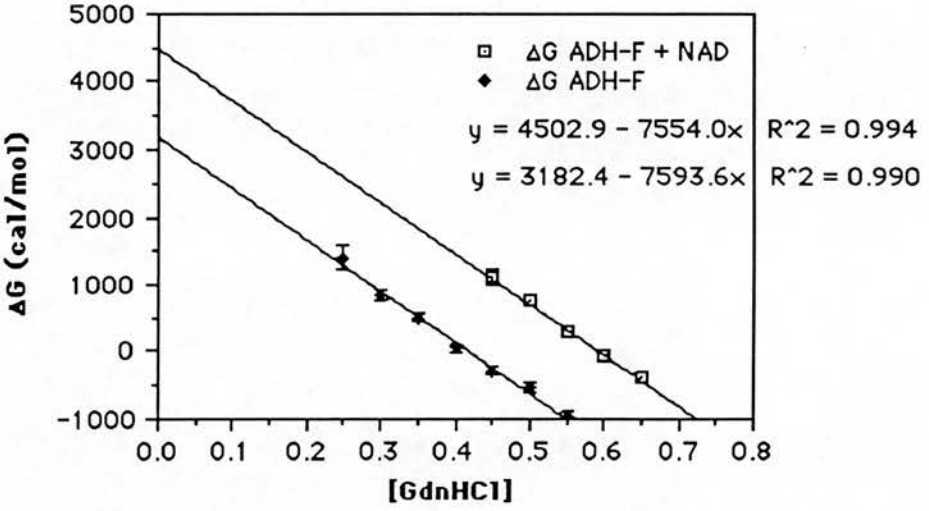


Figure 4.2.2-I . a) Loss of enzyme activity of a solution of ADH-F purified from *D. melanogaster* after incubation with increasing concentrations of guanidine hydrochloride, with and without NAD. b) The transformation of the ratio of inactivated enzyme against activated enzyme into free energy values, and their extrapolation to the $\Delta G_{(water)}$ (cal/mol) value for the enzyme with and without NAD.

Figure 4.2.2-IIa

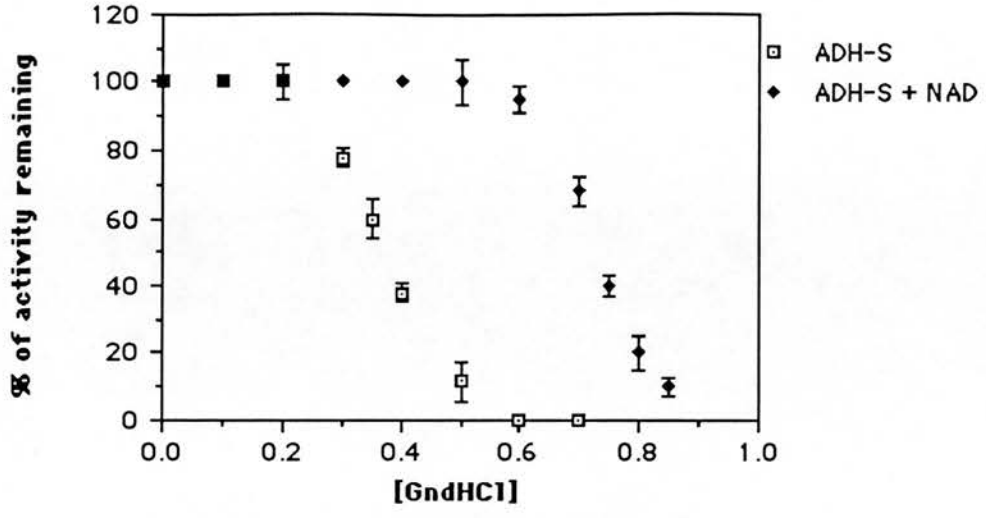


Figure 4.2.2-IIb

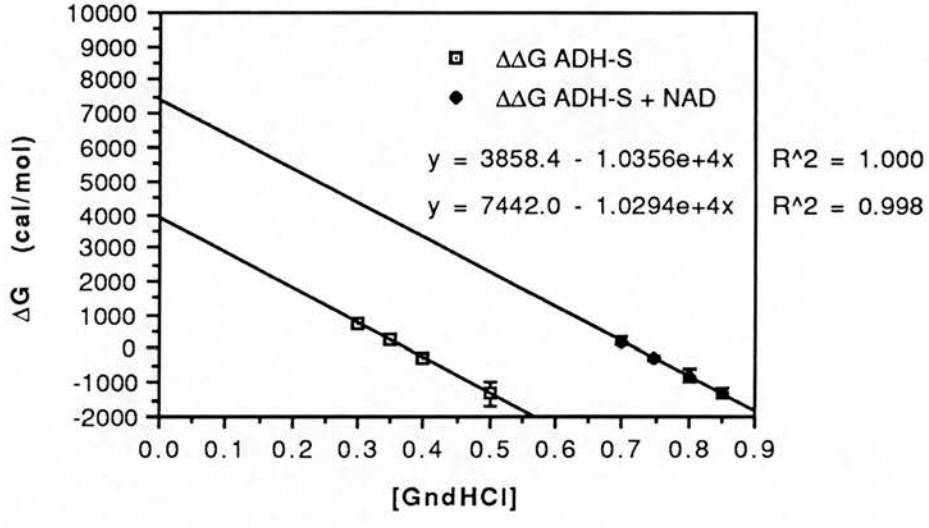


Figure 4.2.2-II . a) Loss of enzyme activity of a solution of ADH-S purified from *S. cerevisiae* after incubation with increasing concentrations of guanidine hydrochloride, with and without NAD. b) The transformation of the ratio of inactivated enzyme against activated enzyme into free energy values, and their extrapolation to the $\Delta G_{(water)}$ (cal/mol) value for the enzyme with and without NAD.

Figure 4.2.2-IIIa

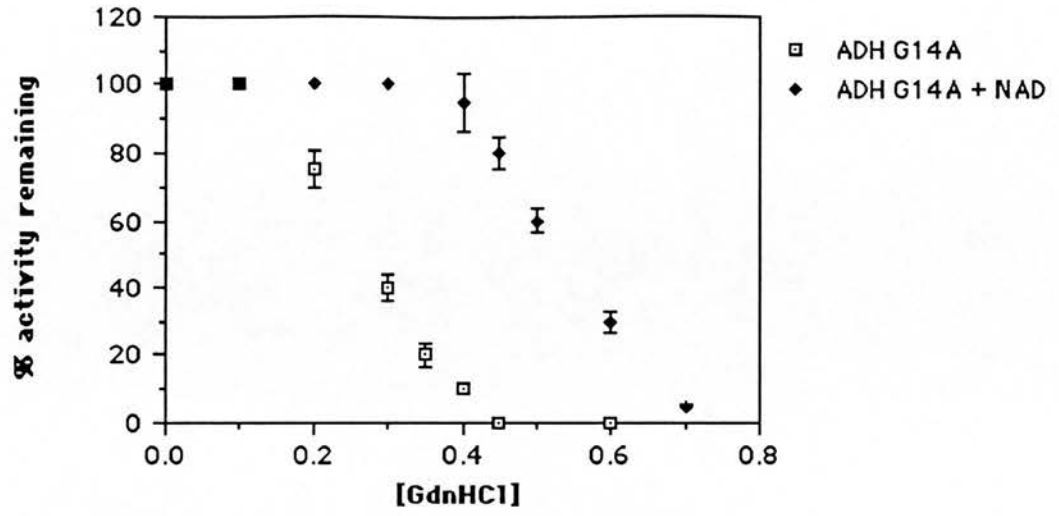


Figure 4.2.2-IIIb

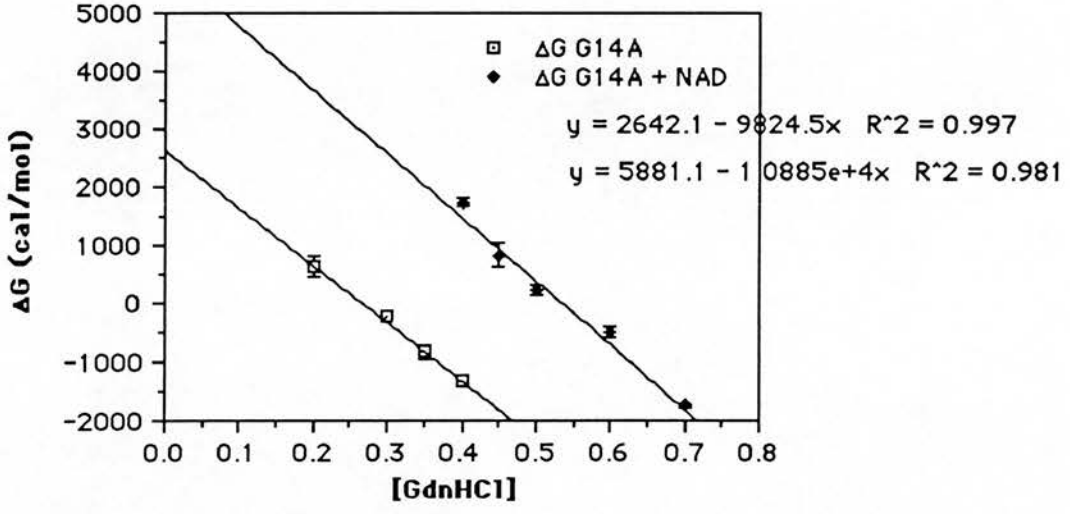


Figure 4.2.2-III . a) Loss of enzyme activity of a solution of ADH G14A purified from *S. cerevisiae* after incubation with increasing concentrations of guanidine hydrochloride, with and without NAD. b) The transformation of the ratio of inactivated enzyme against activated enzyme into free energy values, and their extrapolation to the $\Delta G_{(water)}$(cal/mol) value for the enzyme with and without NAD.

Figure 4.2.2-IVa

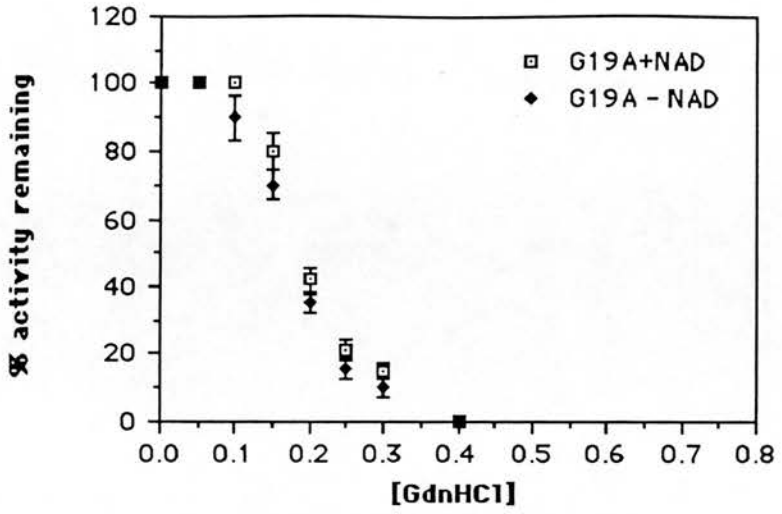


Figure 4.2.2-IVb

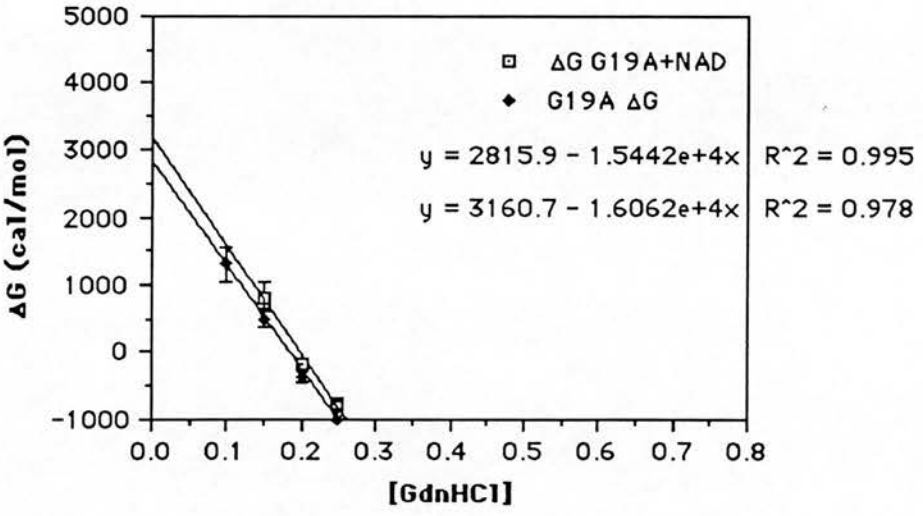


Figure 4.2.2-IV . a) Loss of enzyme activity of a solution of ADH G19A purified from *S. cerevisiae* after incubation with increasing concentrations of guanidine hydrochloride, with and without NAD. b) The transformation of the ratio of inactivated enzyme against activated enzyme into free energy values, and their extrapolation to the $\Delta G_{(water)}$ (cal/mol) value for the enzyme with and without NAD.

Figure 4.2.2-Va

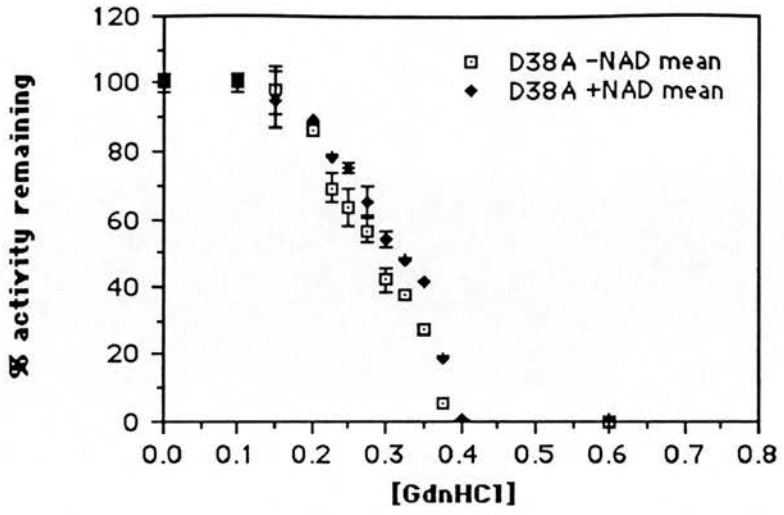


Figure 4.2.2-Vb

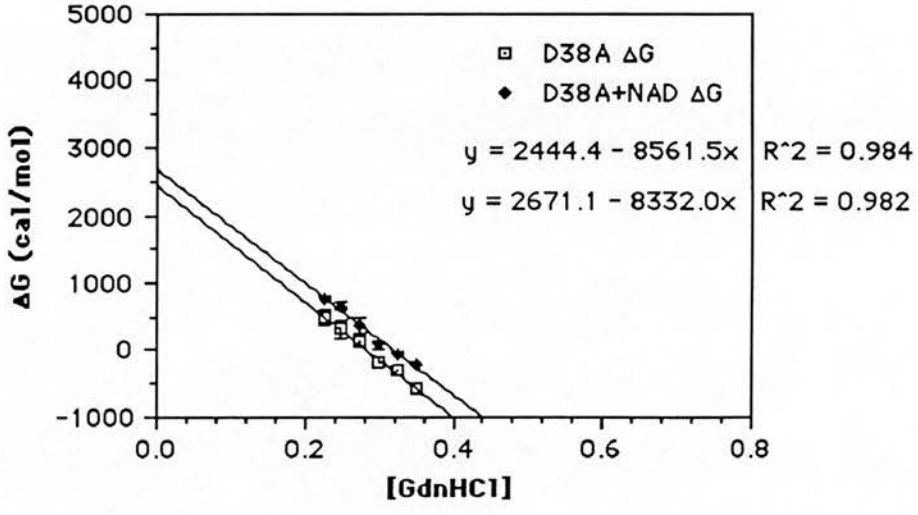


Figure 4.2.2-V . a) Loss of enzyme activity of a solution of ADH D38A purified from *S. cerevisiae* after incubation with increasing concentrations of guanidine hydrochloride, with and without NAD. b) The transformation of the ratio of i

nactivated enzyme against activated enzyme into free energy values, and their extrapolation to the $\Delta G_{(water)}$ (cal/mol) value for the enzyme with and without NAD.

Figure 4.2.2-VIa

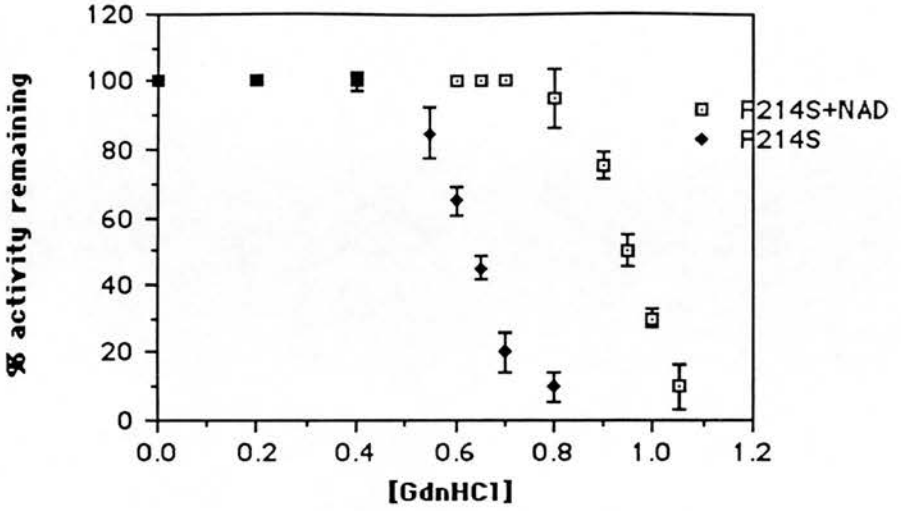


Figure 4.2.2-VIb

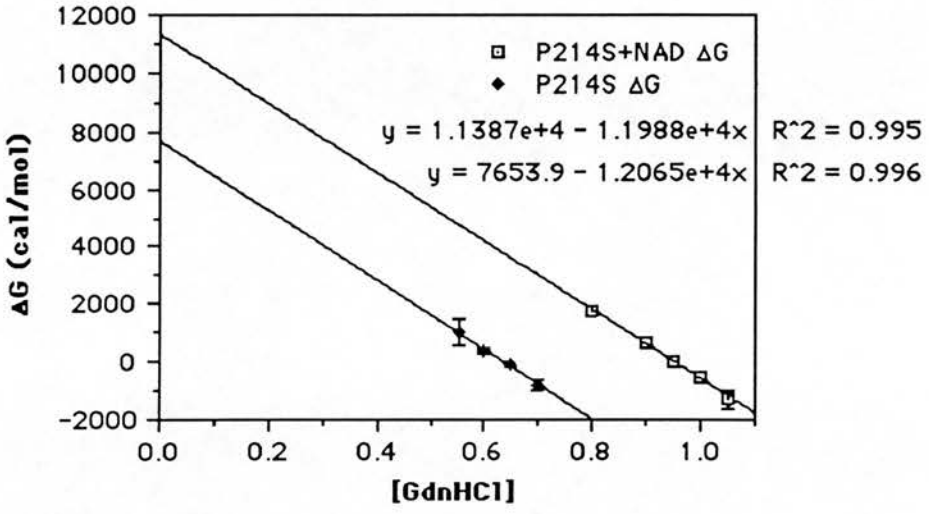


Figure 4.2.2-VI . a) Loss of enzyme activity of a solution of ADH P214S purified from *S. cerevisiae* after incubation with increasing concentrations of guanidine hydrochloride, with and without NAD. b) The transformation of the ratio of inactivated enzyme against activated enzyme into free energy values, and their extrapolation to the $\Delta G_{(water)}$ (cal/mol) value for the enzyme with and without NAD.

Discussion:

The stability of each of the enzymes tested is determined by the $\Delta G_{(water)}$ value extrapolated from the ΔG values of various points in the denaturation curves. As discussed in the Materials and Methods chapter, the binding of the cofactor induces an increase in enzyme stability which can be quantified by determining the $\Delta G_{(water)}$ value for each enzyme when NAD^+ is present in the solution. The difference between the value obtained in the absence and presence of the cofactor, the $\Delta\Delta G_{(NAD)}$value of cofactor binding, is a measure of the energy effect of cofactor binding. Since the energy of binding between the enzyme and the cofactor is a major determinant of the reaction speed it would be expected to find a correlation between the catalytic constants of each enzyme and its $\Delta\Delta G_{(NAD)}$ value of cofactor binding.

ADH-S and ADH-F:

Previous attempts to compare the stability properties of these two natural isoenzymes using crude thermal denaturation studies showed that results were variable and dependant on assay conditions (Chambers, 1984). However, when the stability of the enzymes was investigated by thermal gradient electrophoresis the isoenzyme ADH-S was found to be more stable than ADH-F (Thatcher and Sheikh, 1981). This difference in stability seems to be confirmed by our studies, which show a difference in free energy of about 700 cal/mol in favour of ADH-S.

Confirming the also known fact that ADH-S binds tighter to NAD^+ than ADH-F, a difference between the $\Delta\Delta G_{(NAD)}$ values of about 2.3kCal/mol was found. This value is similar to the calculated energies for a hydrogen bond between tyrosyl-tRNA synthetases and their substrate (Fersht, 1987), which range from 0.5 to 1.8 kCal/mol.

Although the estimated energy of a hydrogen bond depends on a large number of factors and the values estimated in the literature vary largely between 1 and 20 kCal/mol (Williams, 1991), the result obtained in this project can at least be compared to the published values, and seems to suggest that the difference in binding NAD^+ between ADH-F and ADH-S can be due to a single interaction. To know how this relates to the amino acid substitution at position 192 that constitutes the only difference between these two isoenzymes will require three-dimensional information on the contacts between protein and cofactor.

Interestingly, the difference between the apo forms and the holo forms of the two isoenzymes are not equal, the difference between the apo forms is 674.3 cal/mol while the difference between the holo forms is 2939 cal/mol. This difference seems to show that the substitution at position 192 does have a larger effect on the binding of the cofactor than on the intrinsic energy of the enzyme.

Enzymes ADH G14A, G19A and D38A:

The $\Delta G_{(water)}$ values of these three enzymes are all reduced from the original wild type ADH-S by similar amounts (from 1.4 to 1 Kcal/ml). However, they present variation in the slopes of their denaturation curves, which is thought to be an indication of changes in the solvation energy of the denatured state of the polypeptide (Green et al., 1992).

This variation in slope results in an apparent contradiction, where ADH G19A loses 50% of its activity at lower GndHCl concentrations despite having a higher $\Delta G_{(water)}$ value than the other two enzymes.

If this interpretation of the slope of the denaturation curves is correct, ADH G19A is marginally more stable in water than the other two enzymes (a maximum of 371 cal/mol of difference). There is a large variation, however, in the energy that NAD^+ binding transfers to each of these mutants. ADH G14A shows an energy of NAD^+ binding only

slightly smaller than the one of ADH-S. This result is very interesting when compared to the kinetic values obtained for this mutant with NAD. ADH G14A shows a two-fold increase in Km and a two-fold reduction in k_{cat}, suggesting that the mutation primarily impairs the reactivity of the enzyme. This is confirmed by this analysis, showing that the overall energy of binding is not changed, hence the reason for the increase in the Km is probably due to changes in the catalytic constant. A possible structural explanation is that the mutation induces a small repositioning of NAD, maintaining its normal interactions but affecting the architecture of the reaction centre.

ADH G19A and ADH D38A, on the other hand, show a very small increase in the energy of binding of NAD. Their $\Delta\Delta G_{(NAD)}$values of about 0.3 Kcal/mol represent a reduction on binding energy of about 3.3 Kcal/mol in respect to ADH-S. In the case of ADH G19A this result fits very well with the catalytic constants determined for this enzyme, where the rise in Km is not accompanied by a decrease in k_{cat}, and hence one can deduce that the recognition of the cofactor has not been affected, but that the interaction energy between both enzymes has been reduced.

ADH D38A has a similar behaviour to ADH G19A, but these results are difficult to interpret because the mutation could have various effects on the recognition system of the enzyme. Not taking into account trivial structural distortions of three-dimensional structure, there are two major possibilities for the kind of effect observed: the side-chain binds directly with the cofactor, and its loss reduces the interaction energy, or the side-chain is not in contact with the NAD^+ molecule, but its absence increases the entropic energy of the surroundings of the binding surface, hence decreasing the free energy of binding.

As described in the Materials and Methods chapter, all the denaturation curves in the presence of cofactor were carried out by incubating the enzymes with 1 mM NAD. This concentration was used because it is about ten fold the published Km values of the wild type enzyme for NAD, hence ensuring saturation. Since the Km values of the mutants ADH G14A, ADH G19A and ADH D38A are higher than the Km of the wild type enzyme it could be possible that the difference in the protective effect of NAD^+ was due to the fact that the enzymes were not saturated.

However the mutant ADH G14A shows a similar $\Delta\Delta G_{(NAD)}$value to the wild type enzyme despite having a higher Km, suggesting that this enzyme was also saturated. To find out whether ADH G19A and ADH D38A were undersaturated the same analysis should be repeated at higher concentrations of NAD.

Enzyme ADH P214S:

The results obtained with this enzyme confirm that the substitution of the proline at position 214 for a serine has the same stabilizing effect in ADH-S as the one observed in ADH-F isoenzymes. The mutant enzyme has a $\Delta G_{(water)}$ value of 7.7 Kcal/mol, 3.8 Kcal/mol higher than the original ADH-S molecule. From the kinetic studies it is clear that ADH P214S has a higher affinity for the cofactor than ADH-S, but this doesn't seem to be reflected in the difference in $\Delta\Delta G_{(NAD)}$values for both enzymes which are quite similar (only 150 cal/mol higher for ADH P214S).

It may be possible that the stabilizing effect of substrates follows a sigmoidal behaviour, reaching saturation above a certain value of $\Delta G_{(water)}$ for the enzyme. This is however pure speculation as very little information is available on the energy behaviour of protein-ligand complexes.

Another possibility may be that the residue at position 214 does not in fact interact with the cofactor directly, but with parts of the enzyme structure that are involved in the recognition of the NAD^+ molecule. This last possibility would explain why the mutation has a large effect on the $\Delta G_{(water)}$ of the apoenzyme, but does not change considerably the $\Delta\Delta G$ value of NAD^+ binding.

One of the purposes of developing the denaturation assay described in this project was that of trying to correlate changes in the kinetic behaviour of the enzyme with physical parameters directly linked to the molecule's structure and its interaction with the substrates. Although it is possible to correlate the energy of NAD^+ binding of some of the mutants with their kinetic behaviour a general relationship does not exist between the $\Delta\Delta G_{(NAD)}$values obtained and the Km or k_{cat} constants determined.

Clearly, other factors than binding energy are determining the changes in catalytic values; perhaps alterations of the binding pocket that result in positional distortions of the substrates. In any case the determination of the molecular effect of each mutation will require the availability of a high-resolution crystal structure for ADH.

4.2.2.2 Circular dichroism monitoring of the denaturation kinetics.

Results :

Figure 4.2.2-VII displays the results obtained by monitoring the circular dichroism values of solutions of ADH-F purified from *D. melanogaster* larvae which were incubated at increasing concentrations of guanidine hydrochloride.

Figure 4.2.2-VII

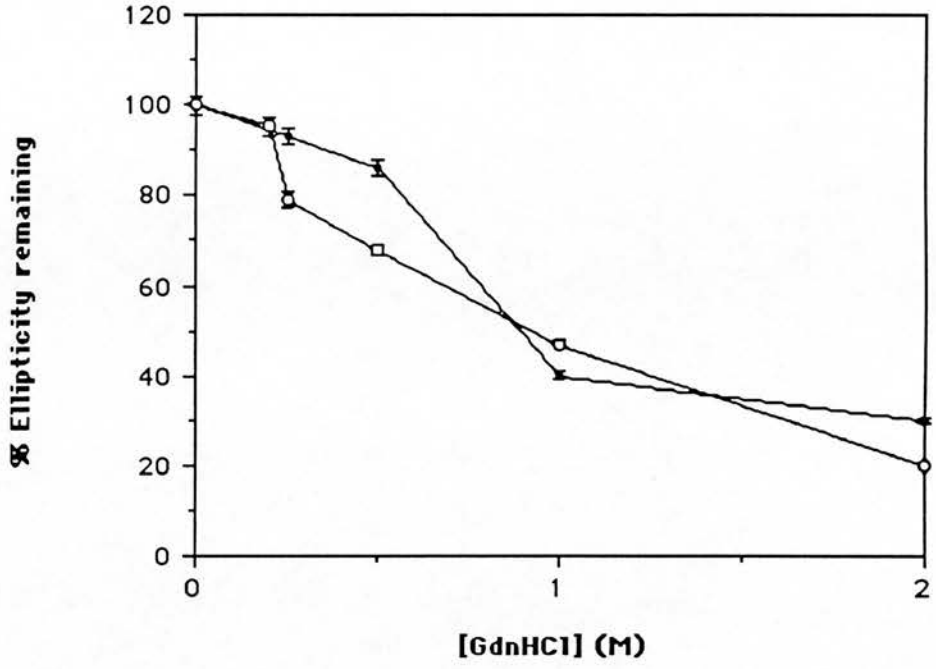


Figure 4.2.2-VII- Denaturation curve of ADH-F monitored with circular dichroism. (•) Enzyme in the presence of NAD. (◦) Enzyme in the absence of NAD.

Discussion:

Since the parameter measured in circular dichroism of far UV is the amount of secondary structure it is expected that the denaturation curve would show a more shallow gradient that the ones obtained when monitoring enzyme activity. However, the CD results still confirm the protective effect of the cofactor on the enzyme, at similar concentrations of denaturant as found in the enzyme activity analyses.

4.2.2.3 Comparative studies with horse liver ADH and yeast ADH.

To be able to compare the denaturation kinetics and the effect of NAD^+ on the unfolding of *Drosophila* ADH to other alcohol dehydrogenases, the same kind of analyses performed on DADH were carried out with solutions of alcohol dehydrogenase from horse liver and from yeast. The results are shown in figures 4.2.2-VIII and -IX.

Discussion:

The monitoring of the denaturation kinetics of two long-chain dehydrogenases was done with the intention of obtaining data on other NAD^+ binding enzymes that could be compared with the results obtained with *Drosophila* ADH. However, both horse liver ADH and yeast ADH proved to behave in a quite different manner in their denaturation process. From our data, it appears that horse liver ADH is more stable to GdnHCl than is the short chain alcohol dehydrogenase from *Drosophila* . Moreover, it appears that only the latter enzyme can be stabilised by addition of NAD. The precise structural reasons for this contrasting behaviour must await a detailed structure of the *Drosophila* enzyme, but presumably are a consequence of different domain and subunit organisations in the two dimeric enzymes.

In the experiments with the tetrameric long-chain enzyme from yeast it was found that NAD^+ offers no protection against inactivation or unfolding by GdnHCl, i.e. a similar result to the horse liver (long-chain) enzyme. In the case of the yeast enzyme inactivation occurs at lower concentrations of GdnHCl, consistent with earlier reports (Brändén et al., 1975) which indicate that the yeast enzyme is less stable than the horse-liver enzyme towards thermal inactivation. The greater instability of the yeast enzyme is thought to be related to a greater ease of dissociation into monomers.

Previous observations on the effect of NAD^+ on the stability of the yeast and horse-liver enzymes have been somewhat contradictory. Thus, Sekuzu et al. (1957) reported that NAD^+ (> 1 mM) protected yeast alcohol dehydrogenase against inactivation by urea, but Wiseman & Williams (1971) found that NAD^+ at these high concentrations had a destabilising effect on the enzyme against thermal inactivation.

In the case of the horse-liver enzyme, there appears to be at least one intermediate in the unfolding process as indicated by the plateau region of figure 4.2.2-VIII (Strambini & Gonelli, 1990). In the plateau region (0.5 M GdnHCl) there is a 30% loss of activity and a 10% loss of secondary structure (as judged by the ellipticity at 225 nm). The nature of this (these) intermediate(s), including its (their) quaternary structure, would require further investigation, possibly by use of nuclear magnetic resonance studies or cross-linking analysis.

Figure 4.2.2-VIII

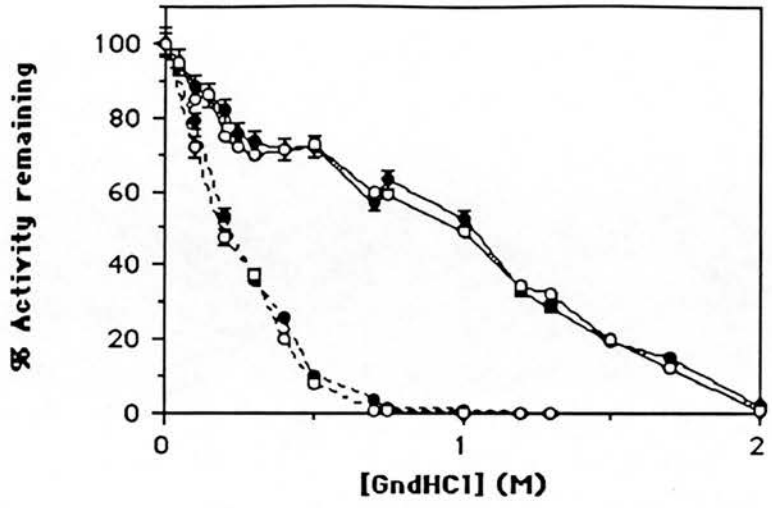


Figure 4.2.2-VIII. Results of the monitoring of the decrease in enzyme activity for horse liver ADH (continuous line) and yeast ADH (dashed line) when incubated with increasing concentrations of guanidine hydrochloride in the presence (o) or absence (•) of NAD.

Figure 4.2.2-IX

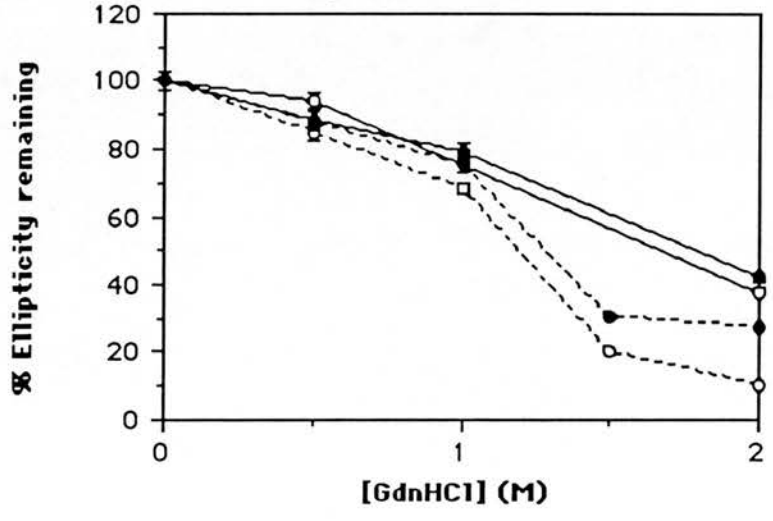


Figure 4.2.2-IX. Results of monitoring the decrease in circular dichroism activity for horse liver ADH (continuous line) and yeast ADH (dashed line) when incubated with increasing concentrations of guanidine hydrochloride in the presence (o) or absence (•) of NAD.

4.2.3 Thermal denaturation studies of *Drosophila* ADH.

As described in the materials and methods chapter, the loss of enzyme activity with time during incubation at different temperatures was analyzed for the enzyme ADH-S and all the mutants produced. Figures 4.2.3-I to 4.2.3-III show the results obtained from the experiments carried out at 4, 25 and 40 °C respectively.

Discussion:

Although thermal denaturation has been used regularly as a way to measure the stability of *Drosophila* ADH from crude extracts and in purified form (Chambers, 1984; Hernandez et al., 1988; Chen et al., 1990, 1991) the results from this kind of analysis are subject to too many different factors (thermal stability, enzyme concentration, purity, proteolytic activity) to be able to be used reliably to determine small differences in stability.

Large modifications, as in the case of the thermo-stable ADH-FChD, can however be consistently observed (Chambers, 1984). The results of the experiments carried out in this project show that, over the time analyzed all the enzymes are stable when kept at 4°C. At 25 °C two groups of enzymes seem to appear, ADH-S and ADH P214S are stable at this temperature, and retain 80 and 95% of their respective activities after 150 minutes of incubation. The ADHs G14A, G19A and D38A are much more labile, retaining 50, 40 and 5% of their respective original activities. Proteolytic cleavage cannot be discarded in any case, especially in the case of ADH D38A, where pure enzyme solutions could not be obtained. However, the results correlate quite well with the $\Delta G_{(water)}$ values obtained for all the enzymes, confirming the general pattern of stability.

At 40 °C all the enzymes but ADH P214S had lost all their activity after 30 minutes of incubation. This is slightly faster than other reported experiments (Chambers, 1984; Chen et al., 1991).

ADH P214S retains 50 % of its activity after 40 minutes at 40 °C, and 20% after 60 minutes. This results confirmed the increased structural stability that this mutation confers to the molecule.

Figure 4.2.3-I

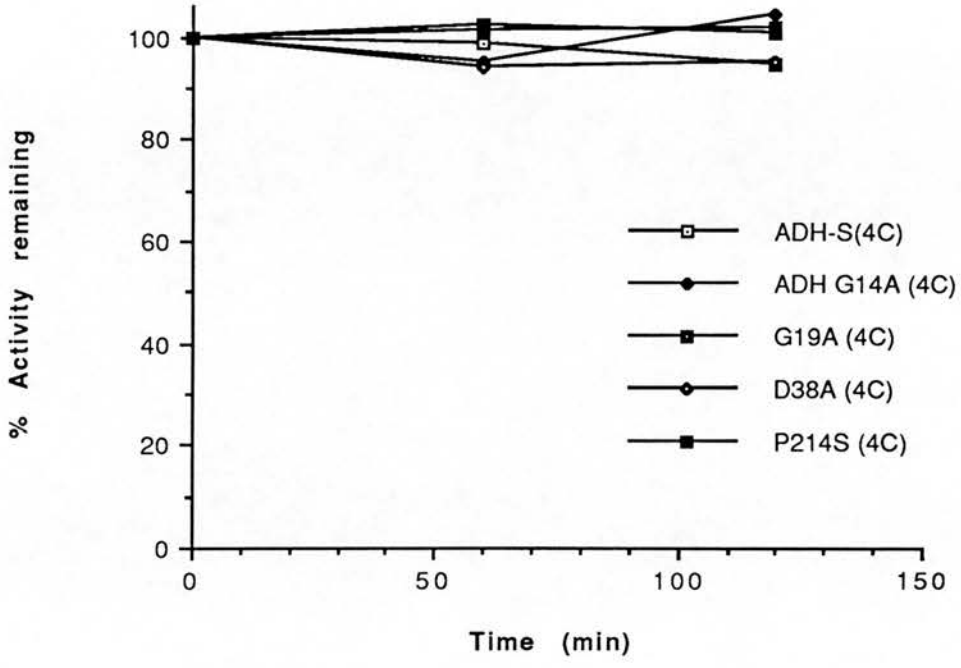


Figure 4.2.3-I. Loss of activity at 4°C with time.

Figure 4.2.3-II

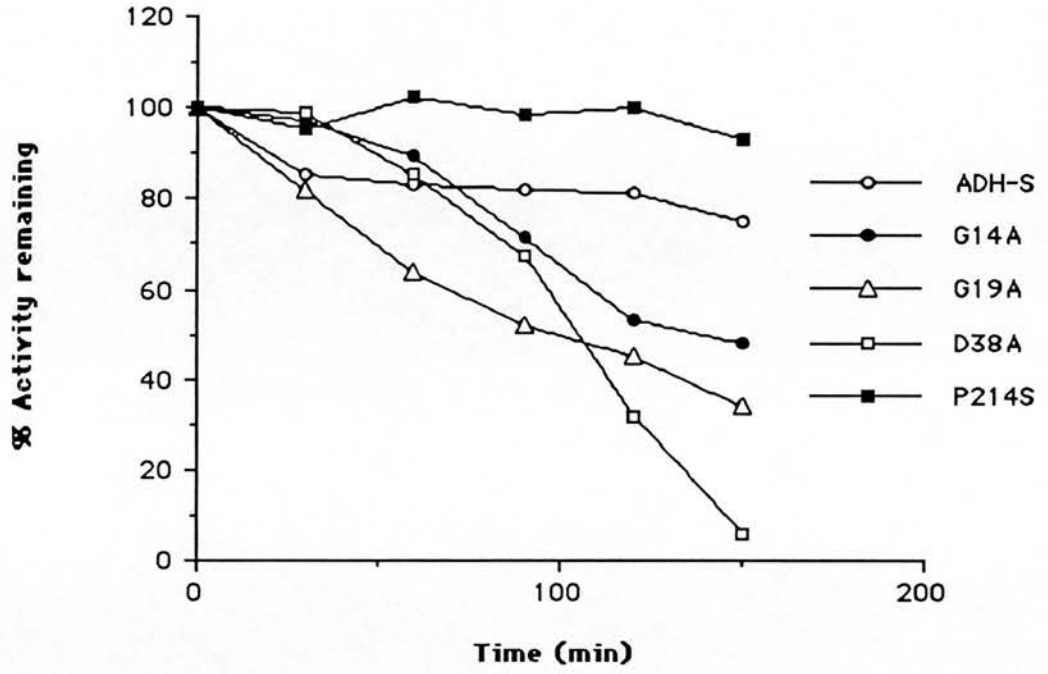


Figure 4.2.3-II. Loss of activity at 25 °C.

Figure 4.2.3-III

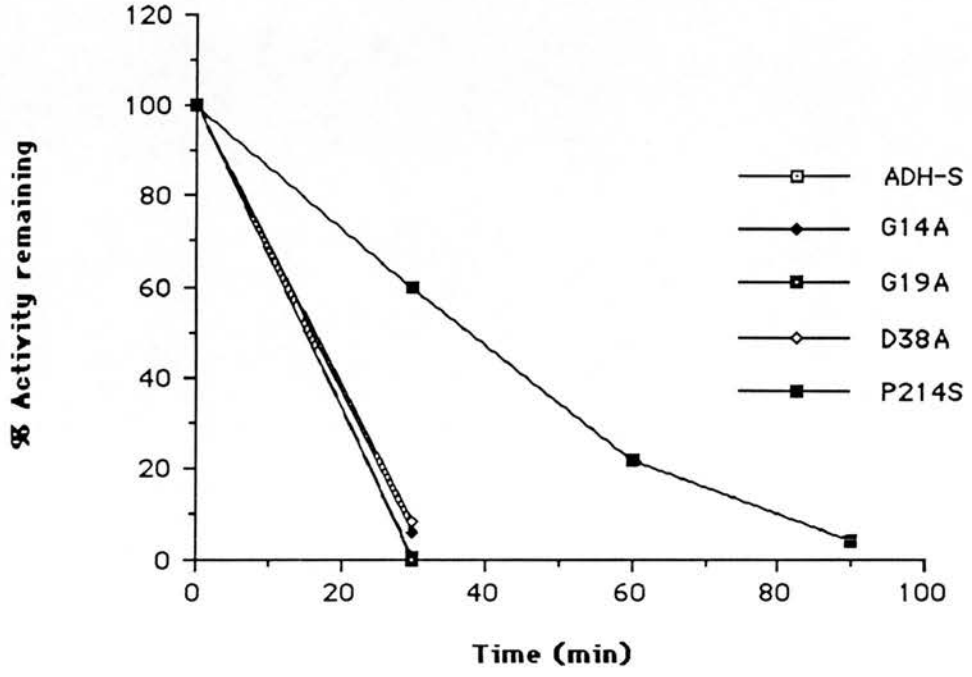


Figure 4.2.3-III. Loss of activity at 40°C.

---

5. Computer modelling of *Drosophila* ADH. Introduction, methods, results and discussion.

5.1 Introduction

Several attempts at crude three-dimensional modelling of *Drosophila* ADH have been published over the years (Thatcher & Retzios, 1980; Thatcher & Sawyer, 1980; Benyajati et al., 1980). These attempts were directed at the N-terminal part of the enzyme, and tried to superimpose the predicted secondary structure of this part of the sequence on the NAD-binding domain of horse liver alcohol dehydrogenase.

Although the basic philosophy of these experiments is the same that is behind the modern techniques of protein modelling, the lack of a proper frame-structure for the model and the crudeness of the comparisons gave very little information about the actual structure of ADH. Today, the need to close the gap between the huge sequence databases and the structural database of the Brookhaven databank has forced researchers to develop a large number of methods that allow the construction of accurate protein models for sequences with sufficient similarity to existing structures.

The definition of 'sufficient similarity' is not a simple one, because we still do not understand what are the determinants of protein structure. The parameter of sequence identity alone is not very adequate unless other information is used, like length of the sequences compared, biological activity, etc.. For sequences of over 80 amino acids it is believed that identities of over 25% indicate the existence of a common fold between both proteins (Sander & Schneider, 1991).

However, some ancient protein families, like the aminoacyl-tRNA synthetases, have common structures despite having sequences far below 20% identical (Schimmel, 1987). It is clear from the existing databases that the variability in protein folds is much smaller than that found in the sequences, and the assumption that protein families that share biological function, monomer length and significant sequence identity will have similar three-dimensional folds is generally accepted (Blundell et al., 1987). As discussed in the introduction the family of the short-chain dehydrogenases complies with these three conditions, and hence the availability of a crystal structure for a member of the family made it possible and logical to try to model the structure of *Drosophila* ADH on it.

The rationale of a modelling experiment starts by identifying the sequence equivalences between the frame-structure and the protein to be modelled. This is achieved through sequence alignment analysis and by comparison of the secondary structure contents of the frame-structure with the secondary structure predictions of the problem protein.

Once the parts of the problem sequence that correspond to the main secondary structure elements of the frame-structure have been determined the process of building the three-dimensional model begins. Firstly, the pieces of sequence of the problem protein can be fit into the secondary elements of the structure taking into account their predicted secondary structure, the distribution of hydrophobicity and charges, the presence of residues normally found at definite positions in secondary structure elements, etc. (Richardson, 1985).

The result of this first exercise provides an initial base, over which the parts of the problem sequence that showed poor relationships to the frame structure have to be added. These sequences will normally correspond to the loops connecting major secondary structure elements, and the way of incorporating them will depend on the amount of three-dimensional information available. In the cases where only one structure is used as reference, loops with little sequence identity to the original ones can be constructed through extensive database search procedures, which will identify regions of the database with better sequence identities and similar length and N- and C-terminal structural positioning.

The result at this stage is a crude approximation to the final model, which includes the total sequence of the problem protein, with large numbers of incorrect angle and bond geometries, large steric clashes, and many local structural incongruities.

The task of improving this first model is done by means of hand refinement, energy minimisation, and molecular dynamics. All three procedures try to correct errors in the structure, hand refinement can be used to remove large steric clashes or improve the environment of determinate side-chains. Energy minimisation is effective in reducing the structural constraints and correcting angles and distances between atoms at a local scale. Finally, molecular dynamics procedures are used to search for a global minimal energy conformation for the whole molecule.

Clearly the whole procedure must be monitored at all stages by comparing the incoming results with biological data on the protein and general data on protein structures, to ensure that no gross mistake is incorporated at any point in the construction of the model.

5.2 Modelling methods.

5.2.1 Sequence alignment procedures.

The direct comparison of DNA and protein sequences was initially used by evolutionary biologists as a way to define pathways of speciation. With the growth of the databases, the appearance of the first three-dimensional structures of proteins and the growth of interest in the molecular structure-function relationships of enzymes, the methods of sequence comparison and alignment have become basic tools for the detection of topological and functional similarities.

Several methods have been devised to carry out pair-wise and multiple sequence alignment analysis (for reviews see Doolittle, 1990 or Argos et al., 1991). They differ in their way of scoring sequence similarities and in the method of regulating the creation of gaps in the sequence alignment. Changes in these parameters can largely affect the sensitivity of the alignment procedure, while the characteristics of the alignment algorithm will determine the computational efficiency of the method.

However, the importance of correctly determining the proper alignment parameters, which is usually a matter of experience, increases with the phylogenetic distance between the compared sequences , and with their difference in length (Argos et al., 1991). In the case reported here, where the length of all the sequences compared is very similar, the determination of the conserved residues is straightforward, and no big differences in the alignment results could be seen between different alignment methods.

One pair-wise alignment algorithm (GAP, UWGCG package, Devereux et al., 1984), and two multiple alignment programs (PILEUP, UWGCG package, Devereux et al., 1984; and CLUSTAL, Higgins & Sharp, 1988) were used for the final model construction. All the sequences used were retrieved from either the Protein Information Resource (PIR) databank (National Biomedical Research Foundation, Maryland, USA), or from the Genebank (EMBL, Heidelberg, Germany).

The alignment program GAP uses the Needleman and Wunsch (1970) algorithm, which gradually searches all the possible alignment positions by locating the sequences in a matrix and calculating the scores of all possible residue pairs using the Dayhoff PAM matrix (Dayhoff *et al.*, 1983). Once this matrix is calculated the best alignment is found by checking the pairs with higher scores, starting at the bottom right (the last residues of each sequence) of the matrix. The penalties for gaps are applied to the the pair scores, allowing breaks to be introduced only if the final score of the pairing is still higher than without the gap.

The multiple alignment algorithms used in this project apply a pair-wise system by which all possible pairs between all the sequences are aligned, the best pair is chosen and the rest are added and aligned successively by order of affinity to the first pair.

The search for the best possible alignment between *Drosophila* ADH and the sequence of SDH (corresponding sequence to the frame structure) was started by carrying out multiple alignment analysis of several *Drosophila* ADHs, with the purpose of defining the conserved parts of the sequence which, in turn, will correspond to the structural core of the structure or to residues with a direct role in the catalytic mechanism.

Since the core of related protein structures tend to be more conserved (Greer, 1981), the conserved sequences between the ADHs analyzed are the parts of the enzyme's sequence that should have clearer equivalents in the sequence of SDH. This expected conservation was then used to analyze the result of the ADH-S alignment with the SDH sequence, to search for significant alignment results.

5.2.2 Secondary structure prediction methods.

The prediction of the secondary structure contents of proteins from their primary sequence has been attempted by many authors for the last 20 years (Lim, 1974; Chou & Fasman, 1978; Garnier et al., 1978; Levin & Garnier, 1988; Gilbert, 1992).

The first, and still most widely used methods, use statistical analysis of the tendency of each amino acid to appear at certain kind of secondary structure type. The sequence, divided in short windows of a few residues, is compared to these calculated values for each residue and a prediction of structure is made on the basis of the score obtained. The problem with this statistical approach is that is largely biassed by the database used for the calculations, and that, considering the sequence of the protein as a linear concatenation of residues, the effect of long-range interactions is not taken into account. These algorithms rarely manage to predict the correct conformation for more than 65% of the total sequence analyzed, and the introduction of three-dimensional information into the predicted algorithms has been attempted recently, with promising results (Garratt et al., 1991; Gilbert, 1992).

In the process of molecular modelling the secondary structure prediction of the problem sequence is a tool that, in conjunction with the sequence alignments, is used to decide what parts of the problem protein will be modelled into the secondary structure elements of the frame structure.

To such purpose we used a protein prediction program (PREDICT, E.E. Eliopoulos, University of Leeds, UK) that contains in itself six different classical secondary prediction algorithms (Garnier et al., 1978; Lim, 1974; Chou & Fasman, 1974; Nagano, 1973; Burgess et al., 1974; Dufton & Hider, 1977). These prediction algorithms differ among each other in the scoring matrices used to analyze the sequence, the way they define secondary structure elements and the length of sequence windows used for the analysis. PREDICT carries out each individual prediction analysis and finds a consensus prediction based on the overall results.

5.2.3 Model construction methods.

The process of building and manipulating the computer model for ADH was done using the modelling package SYBYL version 5.4 (Tripos Associates Inc.), run on an Evans and Sutherland ESV10 workstation. The frame structure of SDH used in this project has only been released to the databank in the form of α-carbon coordinates, so it was necessary to build a proper protein backbone based on this information. This is done using the 'backbone construct' facility of SYBYL, based on a algorithm designed by Claessens and colleagues (1989).

This method uses a database of high resolution crystal structures to search for fragments of 4 residues with total overall length equal to groups of 4 contiguous α-carbons in the structure to be built. The algorithm proceeds linearly, one residue at a time, from the N-terminus of the structure, finding suitable groups of peptides which are then further selected by comparing the distances between all the α-carbons in the fragments to the ones in the protein being constructed. The best fragments are kept and selected again on the basis of an RMS fit between the equivalent α-carbons in the fragment and in the structure. This method produces a protein backbone with an expected mean error of 0.3 to 0.6 Å.

Once the backbone was constructed the sidechains were added. The procedure for the incorporation of sidechains does not rely in database searching methods, as these will be extremely computationally expensive. It relies rather in general conformational values and, as a result, is expected to introduce a significant amount of error in the structure (estimated by SYBYL manufacturers at ~2.45Å).

For the purposes of the construction of a protein model this is not a very large problem. The correctness of the α-carbon coordinates is much more crucial, as the side chains are directly substituted with the problem sequence and steric clashes and high energy conformations between rotamers can be lately removed by molecular dynamics (see section 5.2.4).

Once the sequence of ADH had been fitted to the parts of the structure suggested by the sequence alignments and the secondary structure predictions, the sequences of ADH that had no identity with the SDH sequence had to be incorporated to the model.

These sequences invariably corresponded to connective structures between helices and strands, and were constructed using a loop search algorithm included in SYBYL. The loop construction procedure requires an initial educated guess of which residues may constitute the end and beginning of the particular pair of secondary structure elements that are to be connected.

Once these have been established the program will search the database of crystallographic structures for sequences that have the following characteristics: same length as the sequence that is going to be incorporated to the structure, similar relative three-dimensional positioning of the residues defined as previous and following to the loop sequence and sequence similarity to the sequence of the problem protein. The program will produce the 25 loop structures that best fit these parameters, and these can be then analyzed for coherence with the adjacent parts of the model.

Every small modification made to the SDH structure was followed by a cycle of energy minimisation, with the intention of eliminating structural and steric constraints introduced during the previous stages. This process calculates the energy state of each atom for each of its six degrees of freedom (three translational and three rotational). The value obtained, calculated in arbitrary units, is a function of the amount of bad contacts and angles of each residue. The minimisation process used in this project is a combination of an initial, non-derivative method, which will work atom by atom reducing the energy of each of them below a certain threshold followed by 100 cycles of a first-derivative method known as steepest descent, which adjusts the coordinates of all the atoms contemporarily, using the first derivative of the energy equation for all the degrees of freedom.

During the performance of this kind of energy minimisation the electrostatic energies of the model atoms were not considered, because their inclusion made the calculations far too intensive for the computer power available. It is important to stress that these energy corrections were only applied locally, during the process of construction of the first complete model. These methods are not capable of searching the entire experimental space for a protein structure, and are not adequate for the searching of a global minima state for the model.

5.2.4 Molecular dynamics methods.

As argued in the last section, the methods of energy minimisation used for local modelling operations are not powerful enough to find an energy state of global minimum for the model. This is due to the fact that this methods do not force the energy of the model up at any moment, they just look for possible energy decreasing pathways from the given conformation.

Molecular dynamics methods, in contrast, search the conformational space of the structure analyzed in a much thorough way, by first increasing the internal energy of the atoms of the models by thousands of degrees. This sets the atoms in theoretical motion and, as they are computationally cooled down, allows them to reach much lower energy configurations (Karplus & Petsko, 1990).

The program X-PLOR (Brunger et al., 1987) was used for the molecular dynamics calculations. Typically, the structure was subjected to 200 iterations of energy minimisation to remove large steric clashes before the molecular dynamics procedure. The dynamics analysis was started by increasing the energy of the molecule to 2000 °K, followed

by 34 steps of cooling where the temperature was decreased 50 °K at a time, to reach a final 300 °K.

Although the inclusion of solvent in molecular dynamics calculations is the recommended procedure (Gunsteren & Mark, 1992), water molecules were not included in our calculations because the whole modelling process was carried out in a single monomer. The reason for this is that the modelling of a monomer-monomer interface is far too complex to be carried out with sequences of low identity.

Because the protein surfaces involved in subunit interactions normally have large amounts of hydrophobic residues (Chothia & Lesk, 1987) the inclusion of solvent in the molecular dynamic calculations performed in a single subunit will produce large structural shifts due to the high solvation energy of the hydrophobic residues exposed in the monomer-monomer interaction surface.

5.2.5 Quality assessment methods.

Unlike in protein crystallography, where the quality of a certain crystal structure can be assessed by physical parameters like atomic resolution or temperature factors, there is no experimental way of quantifying the quality of a computer model. This problem of quality assessment is a major one and attempts have been made to provide tools that may allow to examine the likeliness of a model.

Some of them are classical crystallographic parameters, like the Ramachandran analysis of the φ and ϕ angles of the amino acids (Ramachandran & Sasisekharan, 1968), or the rational rules that determine side chains orientations (Blundell et al. 1987); others rely on the comparisons between the three-dimensional analysis of the

environments of the residues of the model with the environments found in crystal structures of the database (Lüthy et al., 1992).

When biological data is available, as in the case of ADH, it can be used to asses the likelihood of the final residues distribution, together with the positions of variable and conserved parts of the protein sequence.

The quality of the model obtained for ADH was studied using all the methods mentioned. The Ramachandran analysis was performed using SYBYL. The program PROFILE (Lüthy et al., 1992), compares the environment of each residue of the model to the normal environments found for residues of the same type in the same kind of secondary structure elements in the database of crystal structures. It produces a plot that indicates the parts of the model where anomalous environments are found, and these have been shown in various examples to correspond to incorrectly constructed models. PROFILE was used to analyze both the SDH structure and the ADH model obtained for it, to check for regions of impaired quality.

The most important quality assessment made involved the contrast of the biologic data described in the introduction chapter with the final model structure. The residues known to have some role in the enzymes mechanism can be checked for positioning compatible with their expected role, and regions of the ADH sequence that are strictly conserved, or largely variable, can be used to look for structural incongruities (no variable sequences should be expected to be found in the core of the structure, for instance).

5.3 Results and discussion.

5.3.1 Sequence alignment and secondary structure predictions. Results and discussion.

The sequence alignment used for the assignment of ADH sequences to the SDH structure is shown in figure 5.3-I. The total identity between the two sequences amounts to 21.5%, a value slightly lower than the found by Sander and Schneider (1991) to be necessary to be able to assume a common fold between two polypeptides. However, as discussed along this thesis, there is good evidence to believe that both enzymes share a common fold.

The similarity between both sequences is not uniformly distributed along the alignment but rather clustered in 11 groups separated by divergent patches of sequence or, in two cases, by gaps.

The distribution of this conserved clusters is identical to the one found when ADH is aligned to 15-prostaglandin-dehydrogenase (figure 5.3-II), another member of the family, suggesting that the conserved clusters may correspond to conserved parts of the structure of these enzymes, possibly the structural core of the fold.

When the three sequences (ADH, SDH and 15-prostaglandin-dehydrogenase) are aligned together the amount of similarity drops, as it would be expected from a triad of enzymes that have evolved in independent directions, but regions of higher similarity still appear at the same positions of the conserved clusters between ADH and SDH (figure 5.3-III). The comparison between the secondary structure contents of SDH with the alignment results (figure 5.3-IV) reveals that most of the conserved clusters do in fact correspond to the strands of SDH.

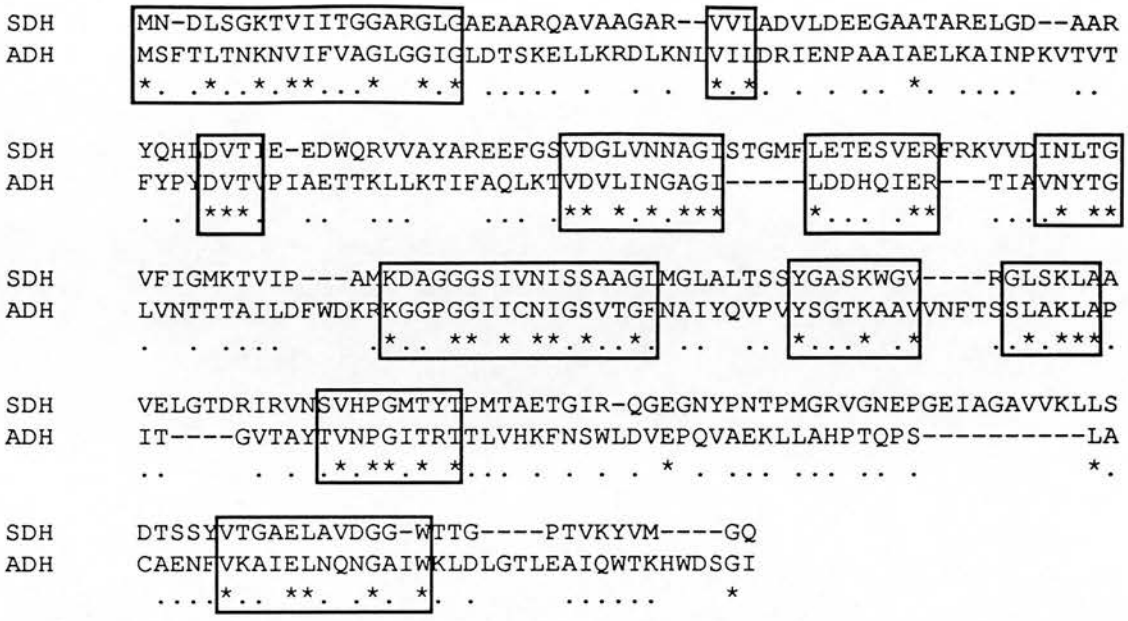


Figure 5.3-I. Alignment results obtained from the sequences of ADH and SDH with the program GAP, using a gap penalty of 10.

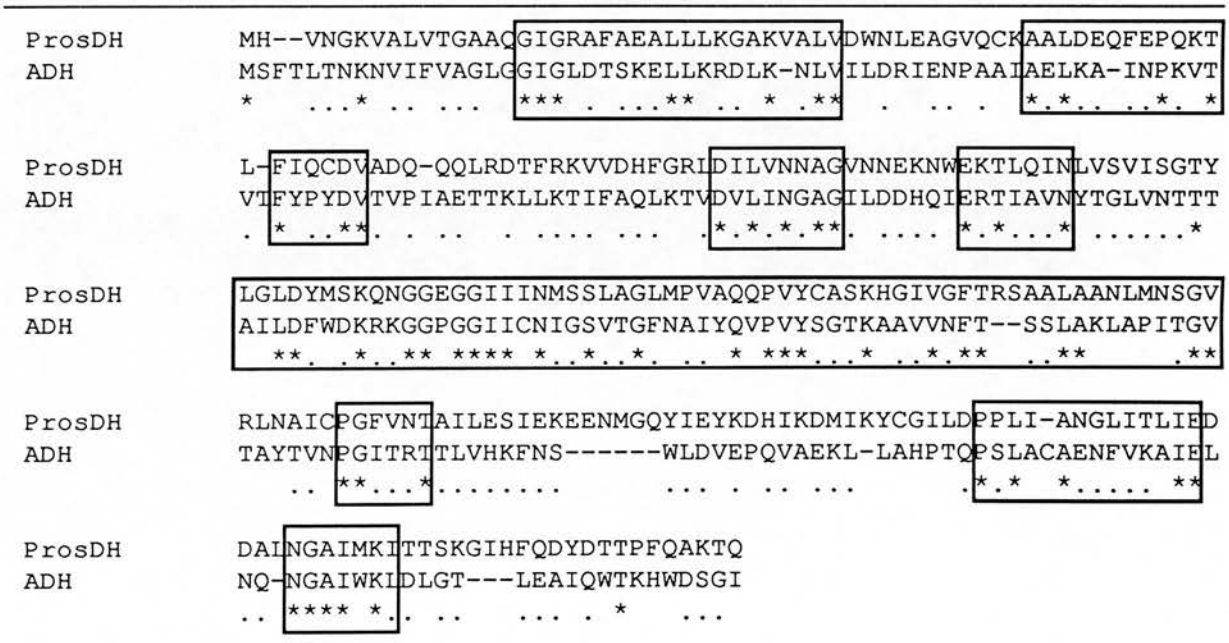


Figure 5.3-II. Alignment result of the comparison of the sequences of ADH and 15-prostaglandin dehydrogenase (ProsDH) with the program GAP, using a gap penalty value of 10.

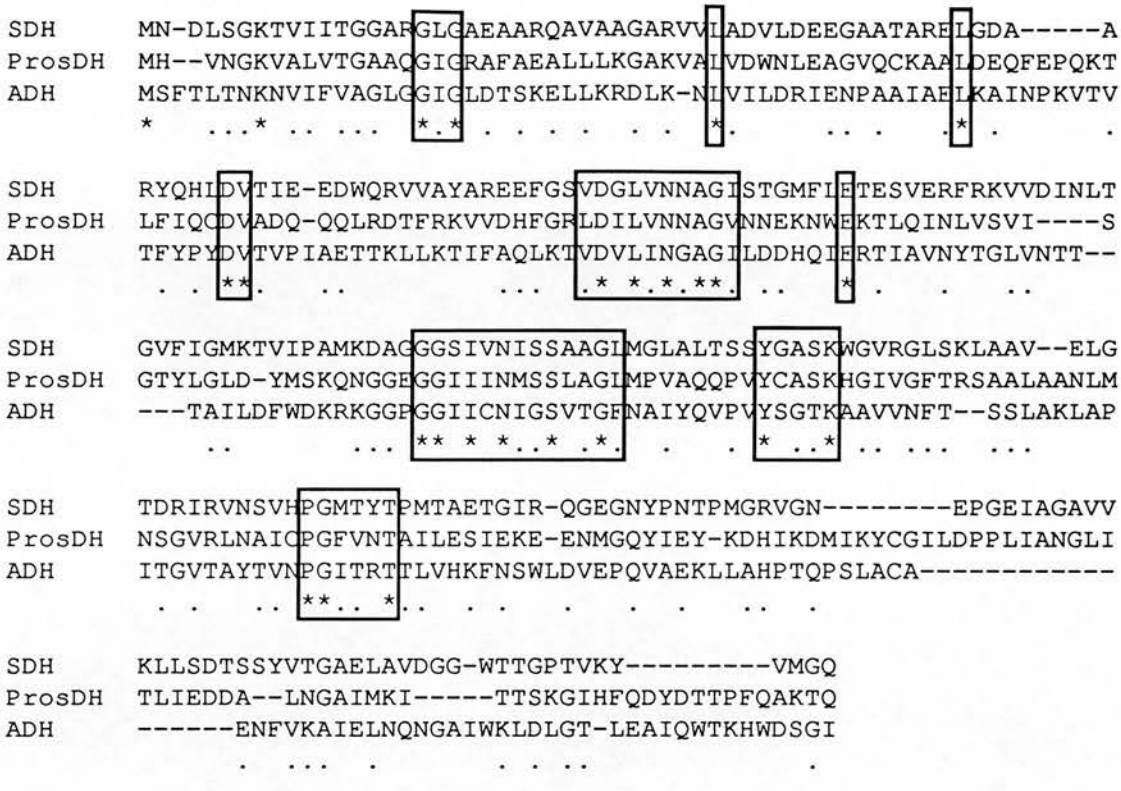


Figure 5.3-III. Alignment results from the comparison of the sequences of ADH, SDH, and 15-prostaglandin dehydrogenase (ProsDH), using the multiple alignment program CLUSTAL, with gap penalty values of 2.

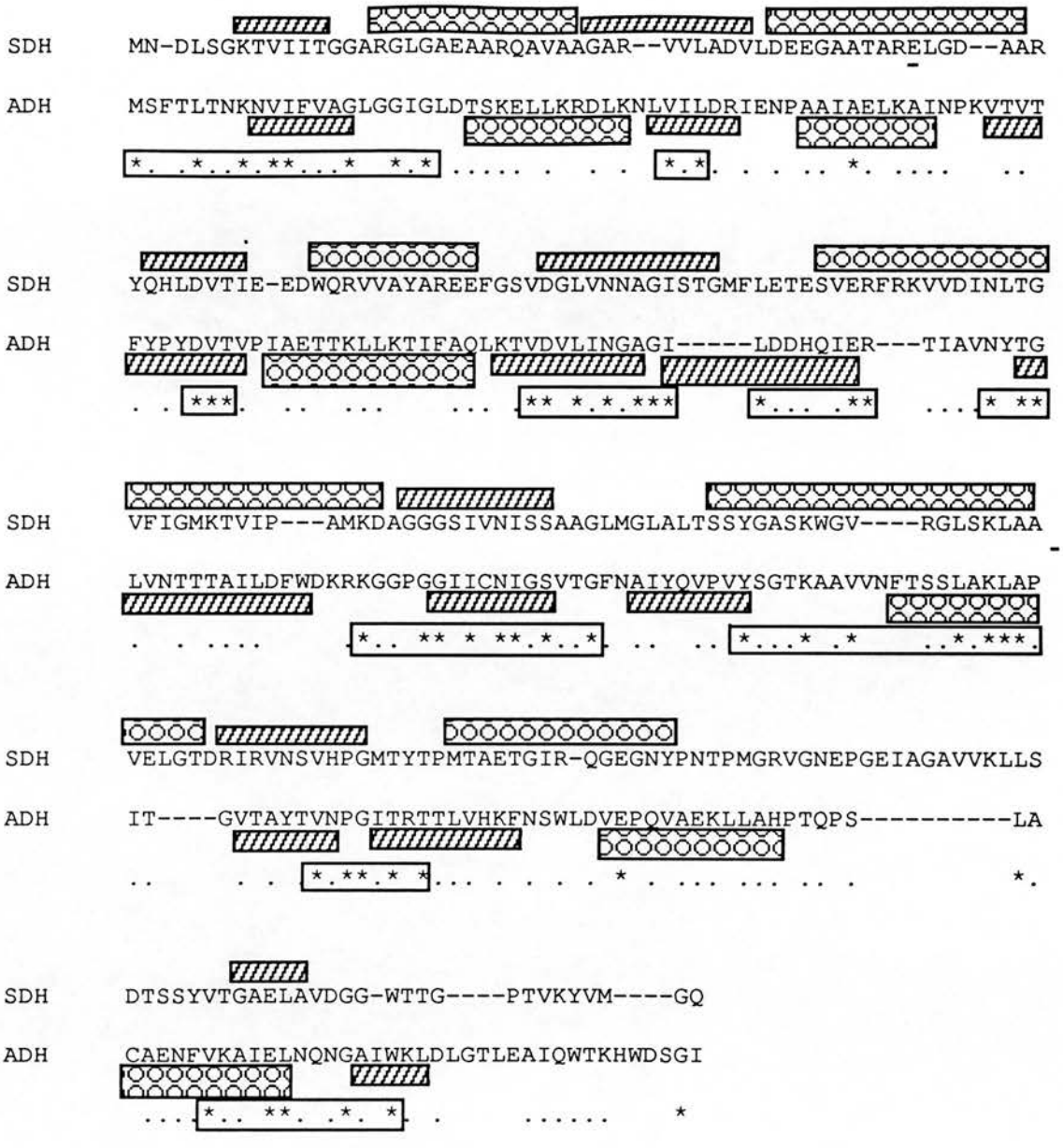


Figure 5.3-IV. Alignment of the sequences of SDH and ADH as in figure 5.3-I. On top of the SDH sequence the helical and extended regions of the structure are indicated with circles and bars respectively. The same code, underneath the ADH sequence, is used to show the results of the secondary structure predictions on ADH, for comparison.

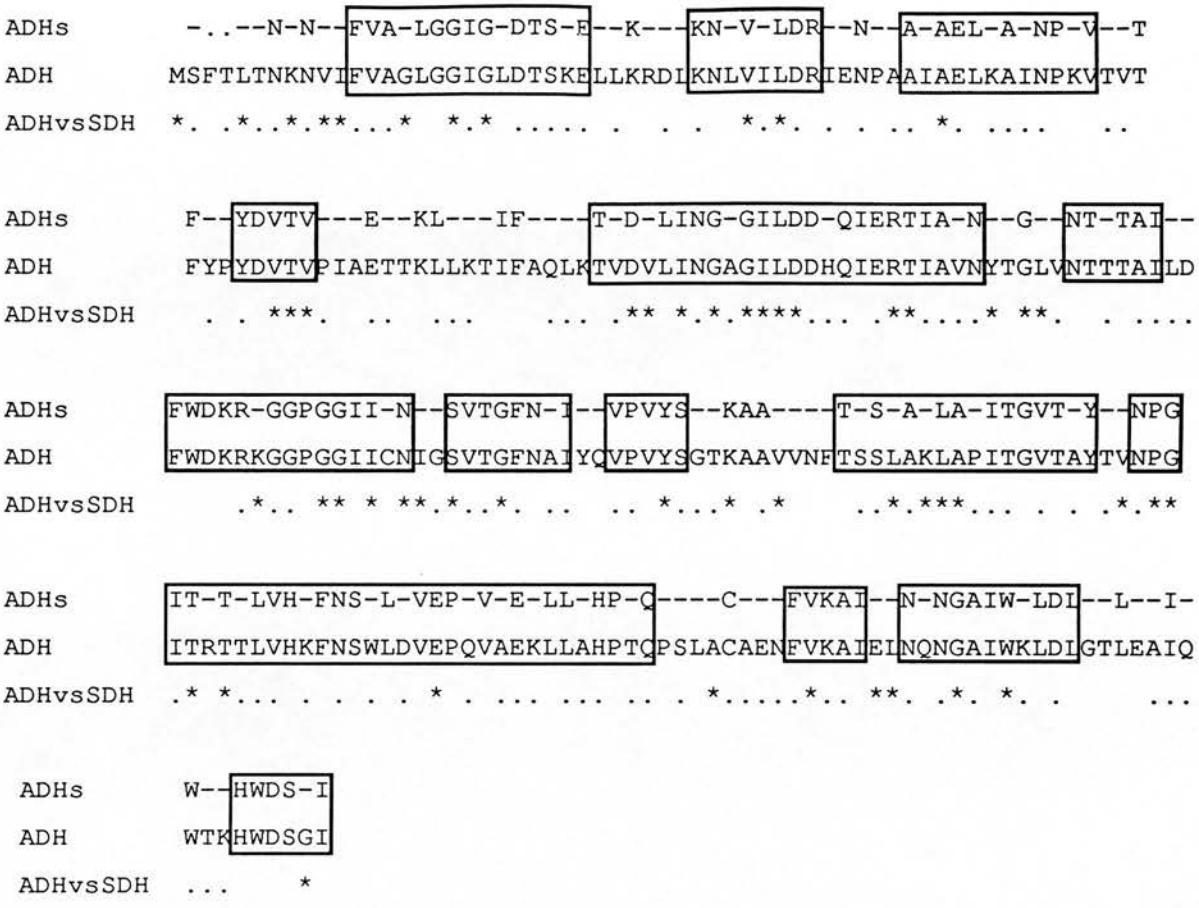


Figure 5.3-V. Alignment of the sequence of ADH from *D. melanogaster* with the conserved residues from the multiple alignment of 25 *Drosophila* ADHs. All the regions of the *Drosophila* sequences that have two or more conserved residues separated by not more than one variable amino acid are boxed. The multiple alignment was done with the program CLUSTAL. The conserved residues found between ADH and SDH are marked at the bottom.

Of the eleven clusters of conserved sequences that can be seen in the alignment between ADH and SDH, seven are at the positions of the seven strands of SDH. The strands of SDH account for only 20% of the total sequence, so it can be considered significant that more than 50% of the conserved residues between SDH and ADH falls into the strand sequences.

It is very interesting to notice that of the six helices of SDH only one shows some sequence similarity to ADH. This conserved helix (helix-E) forms one of the main monomer-monomer binding surfaces in the SDH tetramer, and the same region of ADH that shows identity to this part of SDH constitutes part of the larges similarity cluster between ADH and 15-prostaglandin-dehydrogenase. It seems to indicate that, with the core of the protein, the region involved in this monomer-monomer interaction has also been preserved in the family. Since ADH is a dimer, this similarity is an indication of the possible location of the monomer-monomer binding-site.

Figure 5.3-V shows an alignment of the sequence of *Drosophila melanogaster* ADH (the sequence used for the model) with the conserved sequence from the alignment of the 25 different *Drosophila* ADHs available in the databases. The residues that are conserved between ADH and SDH are also shown, and it can be seen that all the regions of the sequence that have high identity with SDH correspond to well preserved areas among ADHs. This seems to indicate that the variable regions are the same between ADHs and short-chain dehydrogenases in general, a clear suggestion of the existence of a common fold for the whole family.

All the secondary structure predictions of ADH published (Thatcher & Sawyer, 1980; Benyajati et al., 1980; Ribas de Pouplana et al., 1991) show an alternation of strands and at the N-terminal part of the protein that was normally considered to correspond to the N-terminal domain.

In figure 5.3-IV the prediction done with PREDICT is superimposed to the sequence alignment between SDH and ADH, and to the secondary structure contents of SDH. In general, the predictions for ADH match quite closely the structure of SDH for the first 100 residues. The first divergence, appears at the sequences that match the helix-D of SDH. This sequence has little similarity to SDH and is, in fact predicted as extended structure. Helix-D of SDH forms also part of the monomer-monomer binding surface, like helix-E, although contributing less surface to the recognition.

From this position onwards the alignment of the secondary structures of SDH with the predictions for ADH becomes less clear. The stranded regions of SDH have clear equivalents in the predictions for ADH, but a number of shifts are introduced in the predictions for the remaining helical structures. The region with similarity to helix-E of SDH contains a predicted helical structure of shorter length, this discrepancy could be due to the special residue content of this part of the sequence, which is very hydrophobic. The hydrophobicity of that fragment could again be an indication of involvement in the monomer-monomer recognition surface.

A region of similar length to the helix-G of SDH is predicted in ADH with a shift nine residues. helix-G is delimited by the presence of prolines at both its N- and C-terminal, as does the region predicted as helix in ADH. The structure of SDH, between helix-G and the next strand contains a large loop region that was poorly resolved during the crystallographic study of the enzyme (Ghosh et al., 1991). The

sequence of ADH that is aligned to that part of SDH, between residues 215 and 230, has a predicted helical configuration.

5.3.2 Structure modelling. Results and discussion.

Table 5.3.1 shows the fragments of the ADH sequence that, based on the alignment and secondary structure prediction results were selected to construct the main secondary structure elements of the model. As it can be seen, the length of the secondary structure elements was not necessarily conserved. Other parameters, mainly the presence of prolines, alignment gaps or regions of poor sequence similarity, were taken into account for the assignment of the length of a particular element. The residues of the regions that are shown in table 5.3.1 were directly substituted into the SDH structure and were locally refined to remove steric clashes. Once the core structure was created the connecting loops were searched using the method described earlier in this chapter, the list of loops used in the model is shown in table 5.3.2.

The crystal structure of SDH revealed a large loop of 32 residues between helix-G and strand-G (Ghosh et al., 1991). This region of the enzyme is virtually impossible to model because that part of the SDH sequence has no similarity to ADH, and it is far too long to use any standard database search algorithms. Another part of the enzyme that could not be modelled is the c-terminal end, after the strand-G. This region provides very low electron density and has a very large temperature factor in the SDH structure. Again, this region has poor similarity to ADH, in fact, it appears to be a very variable region among all short-chain dehydrogenases. These regions were simply substituted with the sequence of ADH and energy minimized without any further modification. It is important to mention at this point that

the high temperature factor and the poor electron density of the c-terminal tail of SDH suggests that this region may be undergoing large structural shifts. The possible relevance of this fact will become apparent later.

Once the whole polypeptide was constructed and local structural incongruities removed by means of energy minimisation the whole molecule was minimized with molecular dynamics, using X-PLOR. The final fold of the model can be seen in figure 5.3-VI .

Strands	(SDH) (ADH)	Helices	(SDH) (ADH)
Strand-A	7-KTVIIT-12 7-KNVIFV-12	Helix-A	16-ARLGAEAARQAVAA-29 18-IGLDTSKELLKR-29
Strand-B	32-RVVLADV-38 34-LVILDR-39	Helix-B	40-DEEGAATARELGDAA-54 44-AAIAELKAIN-53
Strand-C	58-HLDVTI-63 62-PYDVTV-67	Helix-C	67-WQRVVAYAREE-77 72-TTKLLKTIFAQ-82
Strand-D	81-VDGLVNNAGIS-91 86-VDVLINGAGI-95	Helix-D	100-SVERFRKVVDINLTGVFIGMKTVIP-124 103-RTIAVNYTGLVNTTTAILDFWDKRK-127
Strand-E	132-GSIVNIS-138 132-GIICNIG-138	Helix-E	150-SSYGASKWGVRGLSKLAAVELGT-172 151-VYSGTKAAVVNFTSSLAKLA-170
Strand-F	174-RIRVNSVH-181 175-VTAYTVN-181	Helix-F	189-MTAETGIRQGEGNY-202 189-LVHKFNSWLDVE-200
Strand-G	234-GAELA-238 235-KLDLG-239		

Table 5.3.1. List of the secondary structure elements of the SDH structure and the fragments of the ADH sequence used to substitute them in the model.

The backbone of the ADH model has a mean structural difference of 1.6 Å rms in respect to the backbone of SDH. Obviously the fold shape is generally identical, and the process of modelling has not disrupted the general hydrogen bond pattern of the proteins core, suggesting a compatibility of the fold with the sequence of ADH. Apart from the

low-electron density zones of SDH, the main differences between the original structure and the model are at the regions of the loops between strand-D and helix-E, between strand-E and helix-F and between strand-F and helix-G, which is expected since these regions are the ones that were constructed using the database search methods for loop sequences. Large structural differences were introduced by the dynamics procedure in the large loop between residues 200 and 230, and at the C-terminal tail of the enzyme.

It is difficult to interpret the meaning of the structural features of the ADH model. The sequence alignments and secondary structure prediction analysis clearly suggest that the regions of higher sequence divergence are located at the loop regions of SDH. That sequence differences may not necessarily mean a large structural change in these areas, and the results from the loop searching algorithm offer only the more probable conformation for the sequences searched given the used database. Table 5.3.2 shows the list os loop sequences constructed using the loop search algorithm.

In general, the structural accuracy of a certain model is a direct function of its sequence identity with the original structure (Blundell et al., 1987), as a result the resolution of the ADH model is bound to be low, and extreme care must be taken before linking its structural features to functional properties of the enzyme.

However, the model can be tested with the biochemical data available on the enzyme, and it may offer possible explanations to its catalytic behaviour. This possible explanations can then be used as working hypothesis for experimental tests, a possibility that is not existent at the moment.

Loop sequence	Frame structure	Å RMS	% Identity
14-GLGI-17	1PYP	0.9	50
30-DLKN-33	8ADH	0.6	6
40-IENP-43	1LZT	0.49	6
54-PKVTVTFY-61	5CPA	0.48	11
68-PIAE-71	3TLN	0.74	27
83-LKT-85			
96-LDDHQIE-102	3ADK	0.72	25
128-GGPG-131	3BCL	0.56	30
139-SVTGFNAIYQ VP-150	2HCO	0.67	15
171-PITG-174	1PHH	0.57	46
182-PGITRTT-188			

Table 5.3.2. List of the fragments of the ADH sequence used to construct the loops connecting the secondary elements of the model. The frame structure is the set of coordinates from which the best fitting loop coordinates were selected. Å RMS refers to the spatial difference between the residues at the beginning and end of the loop region in the model and the equivalent regions in the loop of the frame structure. The identity value refers to the sequence identity between the loop added to the model and the equivalent loop of the frame structure.

An indication of the soundness of the model comes from the position on it of the parts of the *Drosophila* sequences that are not conserved among ADHs. The analysis of variability carried out by Villaroya and Juan (1991) was used for that analysis.

The variable regions found by these authors accumulate at the furthest possible position of the NAD binding site, mainly at the loops preceding the helices of the structure (figure 5.3-VII). The regions of the model at the C-terminal ends of helices, which construct the surface of the structure that are meant to construct the active-site of the enzyme, are highly conserved. The two largest helices of the model, D and E, are also mainly conserved, but not so the region of the last strand of the model and its preceding loop. Since the two large helices and the last strand form the two main subunit contact surfaces of the SDH tetramer that difference in conservation found in ADH may be of relevance for the identification of the dimer interface, as will be discussed later.

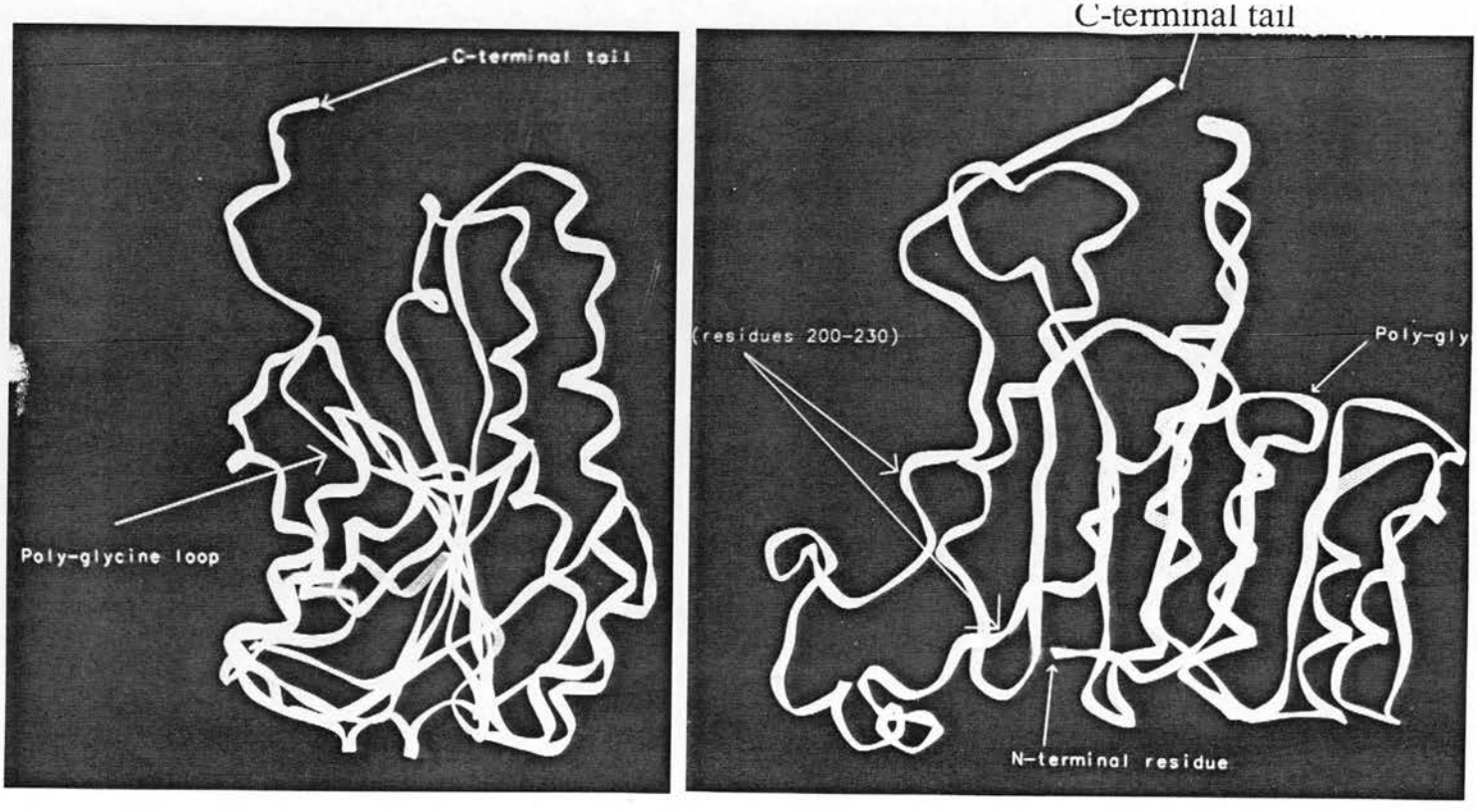


Figure 5.3-VI. Ribbon representation of the general fold of the model.

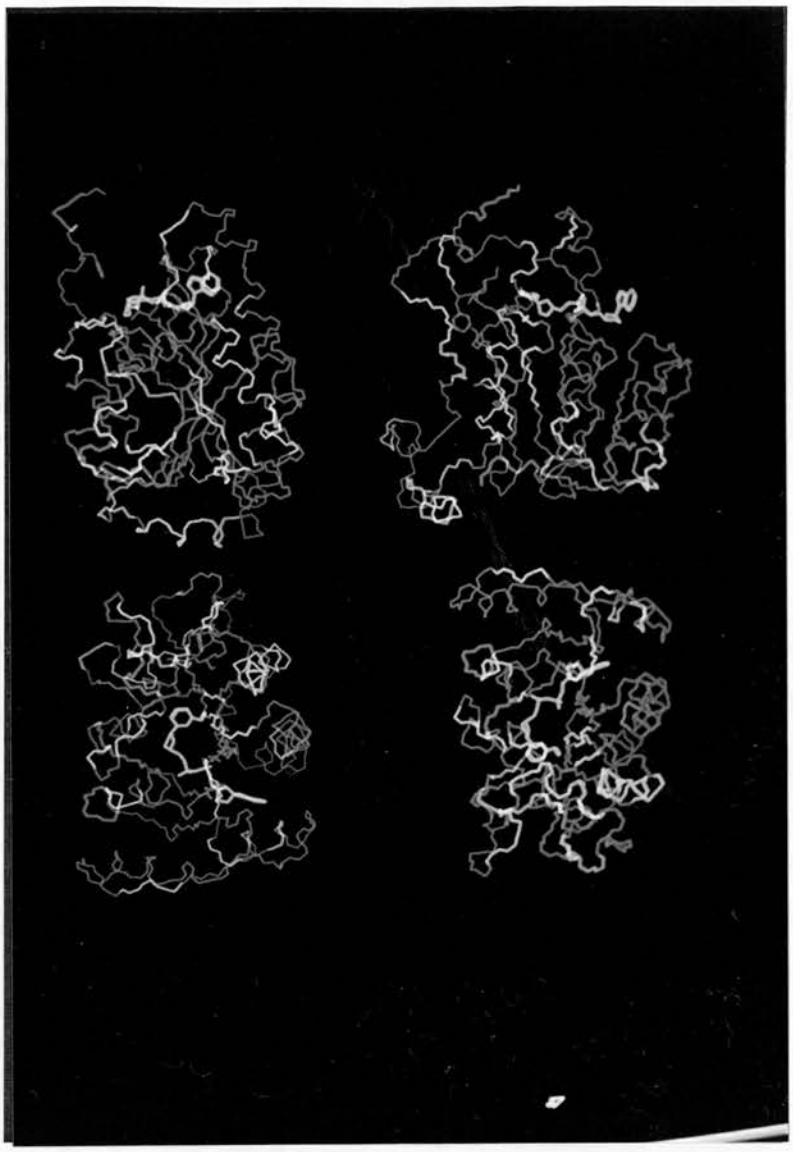


Figure 5.3-VII. Analysis of the distribution of the variable regions of *Drosophila* ADH sequences in the model for ADH. The variability analysis of Villaroya & Juan (1991) was used to colour the regions differentially. The regions were divided in 5 groups according to their variability score.Totally conserved regions are in dark blue (score 0), scores 1-4 are in pale blue, 5-10 in green, 11-15 in orange and highly variable areas (more than 15) are in red. The NAD molecule in its proposed positioning is in bright yellow. The bottom left view is over the NAD molecule with the enzyme below it, the bottom right view is the opposite, the enzyme being above the cofactor.

5.3.2.1 NAD^+ binding site. Results and discussion.

As discussed in the introduction, Ghosh and colleagues found that the position of NAD^+ in the SDH structure was far more external to the positions found in other dehydrogenases. From their cofactor positioning they suggested an steroid binding site that would include a cleft that runs along the surface of the protein. Residues Tyr 152 and Lys 156, conserved in all the members of the family, seemed to have a simple role of recognition of the substrate, being too far away from the cofactor to be able to participate in the catalysis. Ghosh and colleagues proposed that residue Arg 16, which is only present in three other short-chain dehydrogenases (Persson et al., 1991), could be involved in catalysis.

The positioning of the NAD^+ molecule in ADH in a similar position as the one found in SDH leaves the nicotinamide part of the cofactor at least 10 Å away from any residue with catalytic activity. Since the results of the site-directed mutagenesis experiments exposed in this thesis and by others (Chen et al., 1990 and 1991) suggest a very different positioning of NAD, it was attempted to find an alternative location in the constructed model that could satisfy the range of biochemical observations made on ADH.

The effect of the mutation at position G19, which selectively knocks out activity with $NADP^+$, suggested a binding more likely to that found in medium-chain dehydrogenases. Similarly the results obtained comparing the binding strength to NAD^+ of ADH-S and ADH-F suggest that the residue 192 is in the proximity of the cofactor. Also, the results of mutating the residue Tyr 152 (Ensor and Tai, 1991; Albalat et al., 1992) suggest that this residue can be involved in catalysis.

A molecule of NAD, in a similar conformation to the one described by Ghosh and colleagues as bound to SDH, was used to try different positionings that could agree with an structural proximity of all the mentioned residues to the cofactor molecule. A rotation and translation movement of the cofactor, without modifying its conformation, was enough to locate it in a position where the carbamide atom of the nicotinamide ring finds an environment typical of a non-metal dehydrogenase active site (Adams, 1987). Figure 5.3-VIII shows the environments of the NAD^+ molecule in this conformation.

The loop between residues 39 and 46 interacts with the adenine ring and its adjacent ribose. The strand preceding that loop, together with its following helix and the strand formed by residues 92 to 94 form a cleft that could accommodate both the adenine ring and its ribose.

Their extension of burial inside the cleft is difficult to determine, but seems probable that the adenine ribose may pack against the hydrophobic Ile 40, Val 12 and Ala 93, which form the bottom of the cleft, this environment is similar to the one found in other dehydrogenases, like malate dehydrogenase (Grau, 1982).

Unfortunately it is not possible to decide whether the ribose may be hydrogen bonding to residue Asp 38 or to Glu 41, as it would depend in its position inside the cleft. If this positioning of the cofactor is the correct one, the residue Asp 64, proposed by Jörnvall and collaborators (Persson et al., 1991) as the one responsible for hydrogen bonding to the ribose, is too far away to carry out this function.

The ribose of the adenine moiety of the cofactor, and the first atoms of the pyrophosphate are in close contact with the loop from residues 11 to 19. This is the polyglycine loop characteristic of all nucleotide binding domains. The residues 12 to 15 of the loop bend inwards creating a a pocket close to the ribose, capable of accommodating a

phosphate group as the one in $NADP^+$. Figure 5.3-IX shows this loop with the cofactor. Residue Gly 14 is at the bottom of the pocket mentioned, pointing outwards. The incorporation of a methyl group by mutating residue 14 to alanine seems likely to produce steric clashes with the neighbouring loop at about position 40, but the methyl group is well accommodated by the loose packing around residue 14 itself.

In contrast, Gly 19 is in very close packing to the backbone of residues 12 and 13, and the introduction of a methyl group at its position could distort the overall structure of the loop, especially the pocket created by residues 12 to 15. That may be able to explain the inability of the ADH G19A to use $NADP^+$, because its perturbation of the loop may be enough to make impossible the accommodation of the phosphate group of that cofactor.

This may also be the reason for the difference found in $\Delta\Delta G$ of NAD^+ binding between the two enzymes. While ADH G14A has a $\Delta\Delta G$ value of NAD^+ binding similar to the wild type enzyme, in ADH G19A this energy is reduced by 90%.

Maybe this is due to unfavourable interactions caused by the larger perturbation of the mutation G19A compared to G14A. Although both glycines at positions 14 and 19 are conserved among all the short-chain dehydrogenases (Persson et al., 1991), an exception exists among the *Drosophila* ADHs, the enzyme from *D. lebanonensis* having an alanine at position 14.

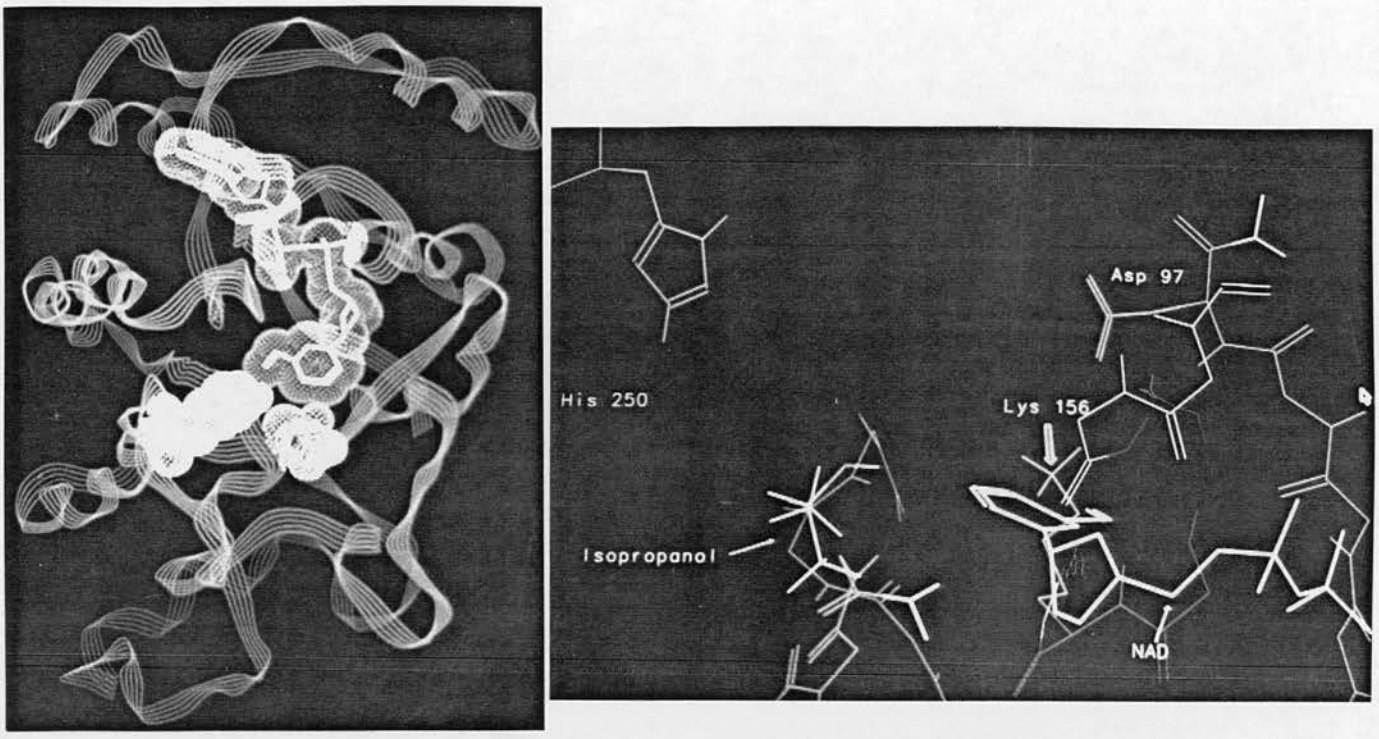


Figure 5.3-VIII. The suggested environment of the NAD molecule in the ADH model. The ribbon diagram alsho shows the Van der Waals surfaces of Lys 152 (yellow) and isopropanol (green).

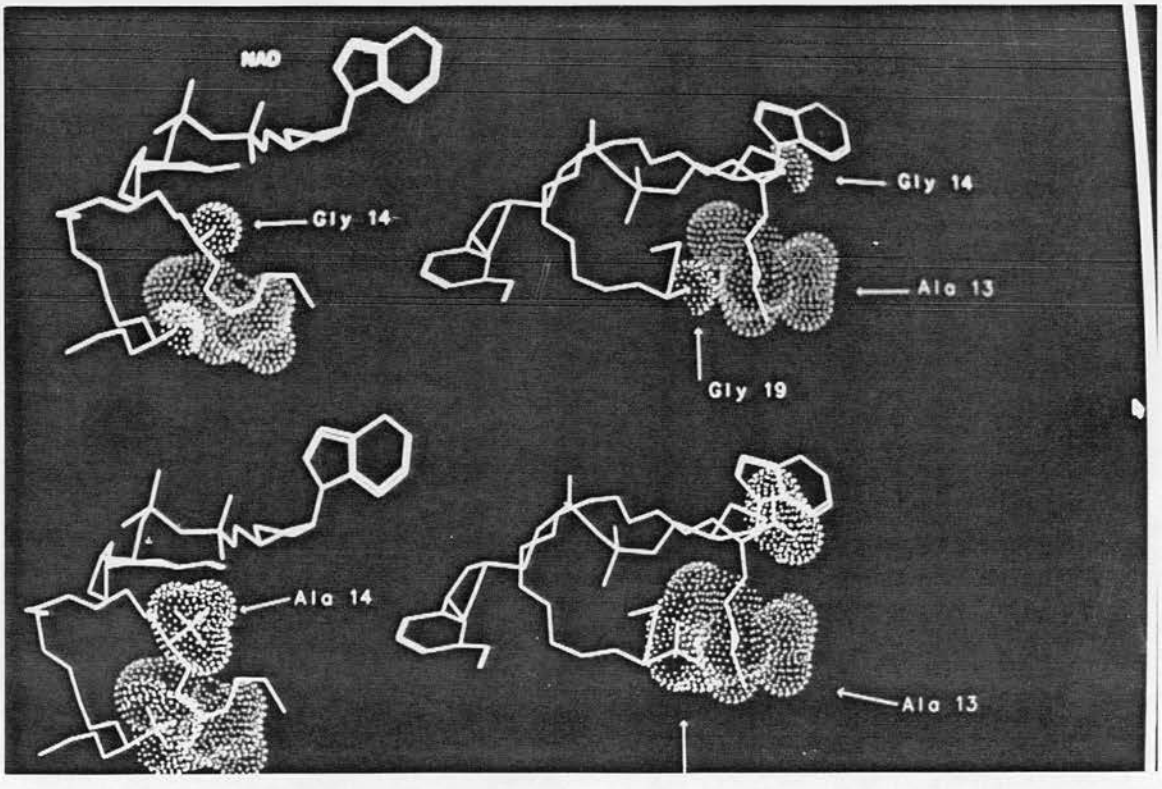


Figure 5.3-IX. The modelled polyglycine loop and the possible effect of the mutations G14A and G19A. the top images show different perspectives of the loop with its original sequence and the Van der Waals surface of A13, G14 and G19. The bottom views are identical except for residues 14 and 19, which are now alanines.

Another part of the model in contact with this adenine moiety of the cofactor is the sequence Gly-Ala-Gly, at positions 92 to 94. The Ala-Gly sequence is conserved in 17 of the 20 sequences of short-chain dehydrogenases aligned by Jörnvall and colleagues (Persson et al., 1991), and the three residues are totally conserved among all the *Drosophila* enzymes. These three residues form part of the strand-D of the model, and they role is probably the same postulated for the Gly-rich loop of NAD^+ binding structures, that is to produce the space to allow the cofactor to enter the binding pocket (Brändén et al., 1975). Such recognition role for these residues has not been proposed so far, and it should be tested by mutating these sidechains to bulkier ones that may affect the binding of the cofactor without modifying the extended conformation of this part of the molecule.

The rest of the pyrophosphate groups and the ribose of the nicotinamide moiety have major interactions with the loop formed by residues 182 to 190, the residue Thr 185 is at 2.5 Å of the hydroxyl groups of the ribose, and could hydrogen bond to them. The structure of this loop is interesting for two reasons. Firstly, the residue 192, responsible for the difference between ADH-F and ADH-S as it changes from threonine to lysine, modifies the strength of recognition of the cofactor. In this model, residue 192 is located at the exterior part of the first turn of the helix that follows the 182-190 turn.

In this position a direct interaction between that residue and the cofactor seems unlikely, and an explanation of the effect of the mutation is not simple. A possible reason could be an interaction between the residue 186, an arginine that points towards the exterior of the loop, and the residue at position 192. But such theory is pure speculation and would require several mutagenesis experiments of the residues in the area to be properly analyzed.

Secondly, the loop at this region is of interest because it has a striking similarity with a region of bacterial lactate dehydrogenases also involved in the catalytic mechanism. The region between residues 100 to 120 of the lactate dehydrogenase enzyme form part of the highly mobile loop that directly participates in catalysis through closure over the active site (Parker & Holbrook, 1977).

The analysis of the sequences of this loops in various lactate dehydrogenases done by Gernstein and Chothia (1991) shows a high degree of sequence conservation between the loops of bacterial dehydrogenases. As can be seen in figure 5.3-X, the region between 182 and 190 of ADH shares a high similarity with that region of lactate dehydrogenase. It also shares a similar three-dimensional function, connecting an internal strand of the nucleotide binding fold to an external helix. In lactate dehydrogenase the loop in question closes over the active site, removing water, and allowing the reaction to proceed (Parker & Holbrook, 1977).

```
B. stearo.  LDH        AGANQKPGETRLDLVDKNIAIFRS
B. megat.   LDH        AGANQAPGETRLDLVEKNVXIFEX
L. casei    LDH        AGAPKQPGETRLDLVNKLNKILKS
                       *** : **********    *:

Drosophila  ADH  176-TAYTVNPGITRTTLVHKFNSWLDV-199
                       ::    ** ** :**:*    ::
```

Figure 5.3-X. Alignment of the sequence of the catalytic loop of various bacterial LDHs (modified from Gernstein and Chothia, 1991) with the sequence of *Drosophila* ADH around residue 192.

The similarity between these two protein regions may be not significative, but it is tempting to speculate that ADH may have a similar loop-closing mechanism. Clearly, all dehydrogenases require structural shifts to exclude water from the active site and allow the proton transfer reaction to occur.

The two domains of horse liver alcohol dehydrogenase close over each other to allow water exclusion (Eklund & Brändén, 1987), but smaller enzymes like lactate dehydrogenase use smaller structural shifts. Perhaps the 182-190 loop, maybe with other structures in the enzyme, undergoes similar structural shifts, closing over the region of the nicotinamide ring of the cofactor, and excluding water. In such situation, the effect of change at position 192 may have a different implication, as it may be affecting the dynamics of the loop.

In the binding position being described, the nicotinamide ring of the cofactor lies over a set of hydrophobic residues, namely Ile 95, Leu 96 and Ile 137. The C-4 of the ring gets in close proximity to residues Lys 156 and Ser 139, also in the vicinity of the nicotinamide ring are residues aspartic 97 and tyrosine 152. Although the conformation of this last sidechain seems to point its OH group in a different direction, the definition of the model is not good enough to provide information about the real conformation of the sidechains. Although the process of molecular dynamics can find the minimal energy conformation for them, a conformation of the tyrosine pointing towards the nicotinamide ring of the cofactor can not be ruled out.

5.3.2.2 A reaction mechanism hypothesis.

As described in the introduction a number of residues are good candidates to participate in the catalytic mechanism of ADH. The chemical modification of one of the four histidines of ADH results in the inactivation of the enzyme (thatcher, 1981; Retzios, 1982). Of the four histidines only three are totally conserved, at positions 191, 210 and 250.

In the model presented here His 210 is not in the proximity of the active site of the enzyme, His 191 is the same helix as Lys 192, and its chemical modification could have a certain effect on the interaction of its preceding loop with the cofactor but the residue does not directly interact with either cofactor or substrate. On the other hand His 250 is part of the C-terminal tail, a region of very little sequence similarity to the SDH sequence, and which, in the latter, and on the basis of its poor electron density, is probably a largely mobile structure.

Using pH dependent kinetics, Winberg and colleagues identified a group in the enzyme which, in the range of pH 6-10, shows different effect on the binding of the alcohol or aldehyde. The binding of acetaldehyde is not affected in that pH range, but that of isopropanol is destabilized by a residue with a pK of 7.6 (Winberg & McKinley-McKee, 1988).

Winberg and colleagues propose the existence of a group that binds the hydroxyl of the alcohol and the oxygen of the acetaldehyde, and which has a pK of 7.6 when bound to the former, and one over 10 when bound to the latter. The four candidates proposed for such behaviour are histidines, tyrosines, lysines and cysteines. Cysteines, however have been shown to play no role on the activity of the enzyme by site-directed mutagenesis (Chen et al., 1990).

The two best known classes of dehydrogenase reactions are epitomized by horse liver alcohol dehydrogenase and by lactate dehydrogenase (Adams, 1987). The former uses zinc in the active site to act as a Lewis acid and catalyze the proton transfer reaction (Eklund and Brändén, 1987), the latter does not have any metals and rely on a histidine group which, in the dehydrogenase reaction, removes the hydrogen of the lactate's hydroxyl (Parker & Holbrook, 1977), the limiting step of that reaction is the dissociation of the reduced cofactor, which depends on the opening of a mobile loop of the enzyme that covers the active site.

Drosophila ADH contains no metal ions, and the limiting step of its reaction is also the dissociation of NADH (Thatcher, 1981; Winberg et al., 1986), the similarity to the lactate dehydrogenase mechanism also extends to clear structural identity between lactate and isopropanol. However, the pH dependence of ADH seems to be more similar to horse liver alcohol dehydrogenase than to lactate dehydrogenase (Winberg and McKinley-McKee, 1988).

The proposed binding position of NAD^+ in the computer model for ADH clearly delimited a possible active site space (figure 5.3-XI). The analysis of this space showed an striking similarity with the three-dimensional arrangement of the lactate dehydrogenase binding site.

In this hypothetical reaction centre lysine 156 and aspartic 97 are aligned in a very similar positioning to that of arginine 109 and aspartic 168. No histidine residues are present in that region of the molecule, but the C-terminal tail of the enzyme is in the vicinity, and histidine 250 could easily reach the active site in the right angle if the C-terminal tail closed over the proposed catalytic centre (figure 5.3-XII).

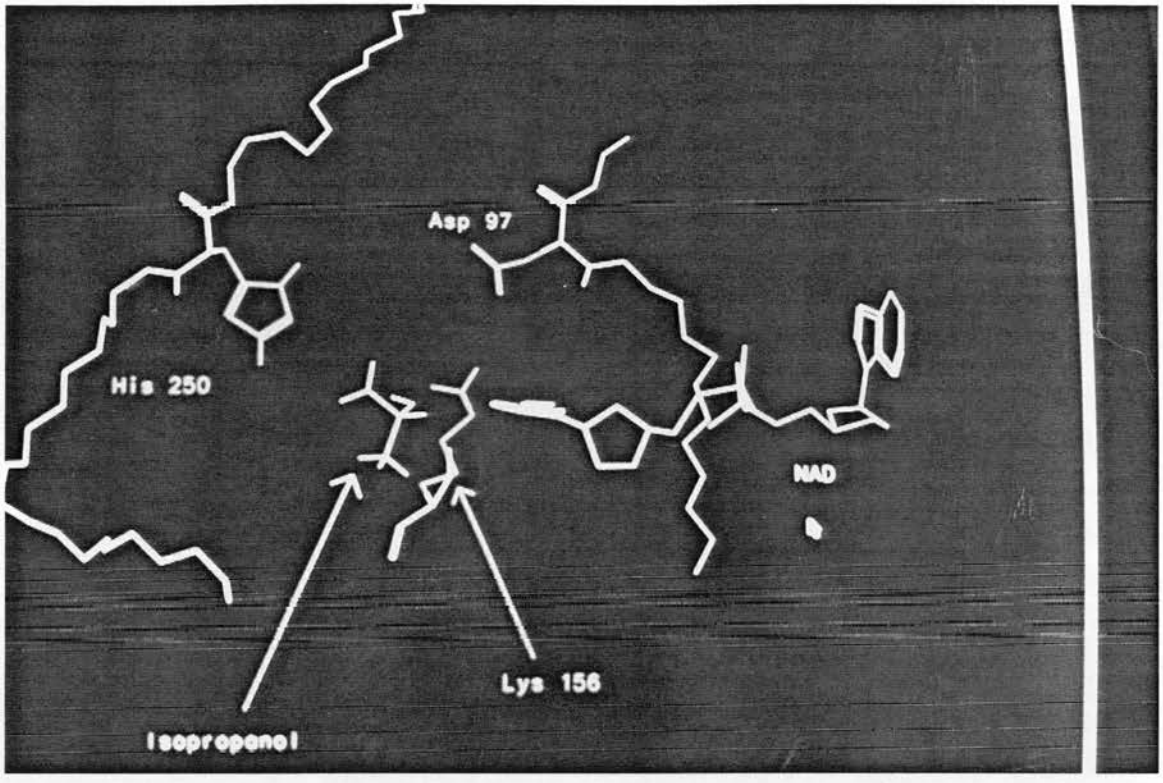


Figure 5.3-XI. Photograph of the space around the carbamide group of the NAD^+ in the model. The isopropanol molecule placed in the active site is shown.

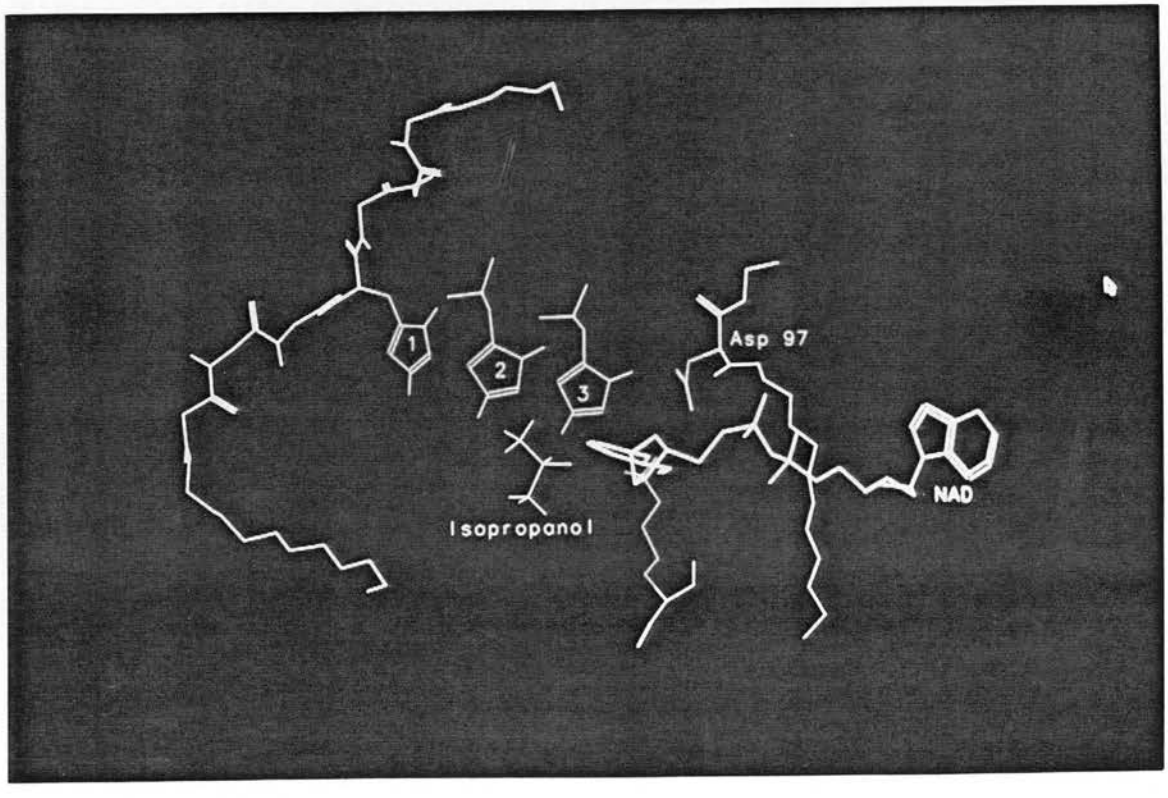


Figure 5.3-XII. Representation of the proposed closing of the C-terminal tail of the enzyme over the active site.

An attempt was made to locate an isopropanol molecule in the right positioning in respect to the C-4 of the nicotinamide ring, and no steric clashes or conformational constrains were found that prevented such orientation. In this positioning one can propose a theory for a reaction mechanism involving residues aspartic 97, lysine 156 and histidine 250, which would be essentially identical to the reaction mechanism of lactate dehydrogenase (Parker & Holbrook, 1977; Clarke et al., 1989).

We propose that the cofactor binds first and somehow modifies the environment of the active-site allowing it to receive the substrate molecule (lysine 156 hydrogen bonds to the oxygen of the substrate's hydroxyl) and triggering the closing of the C-terminal loop over the reaction centre. Histidine 250 acts then as a general base binding to the hydrogen of the alcohol's hydroxyl group, and promotes the hydride transfer to the NAD. The imidazolium ion is stabilized by aspartate 97.

Once the dehydrogenase reaction is complete the loop opens its conformation, perhaps due to electrostatic repulsion with NADH and the products are released. Having lost the stabilizing effect of Asp 97, His 250 loses its proton to the solvent, and the reaction is repeated (figure 5.3-XIII).

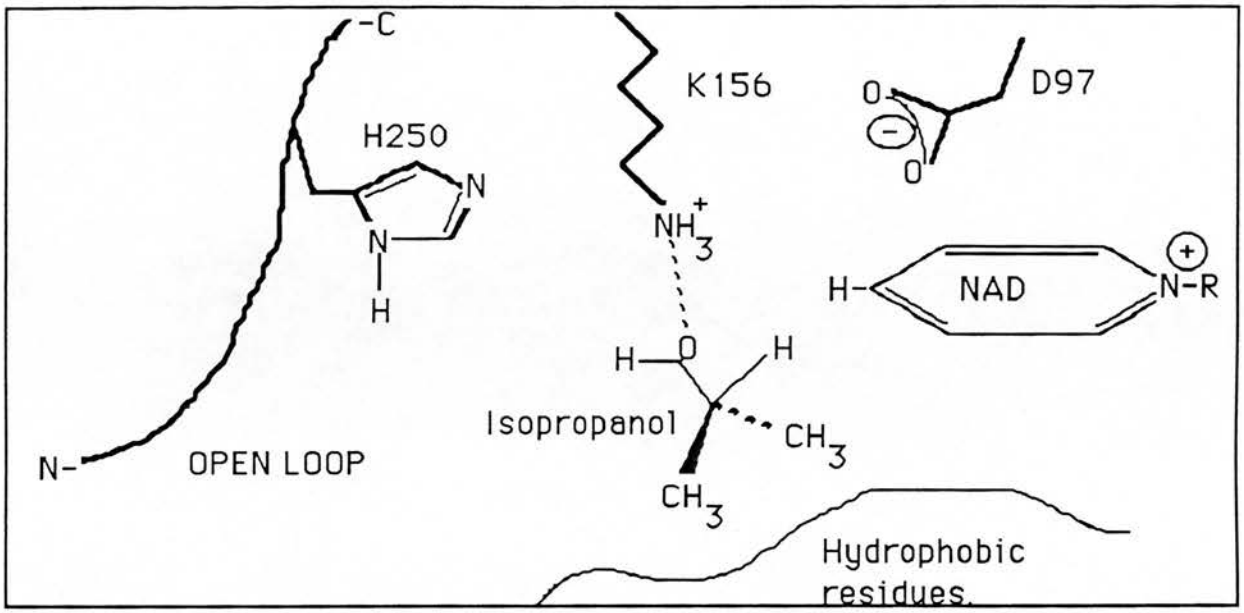


Figure 5.3-XIIIa. First step of the proposed reaction mechanism for *Drosophila* ADH. The NAD^+ molecule enters the active site followed by the alcohol, which is stabilized by hydrophobic interactions and by a hydrogen bond with the residue Lys 156.

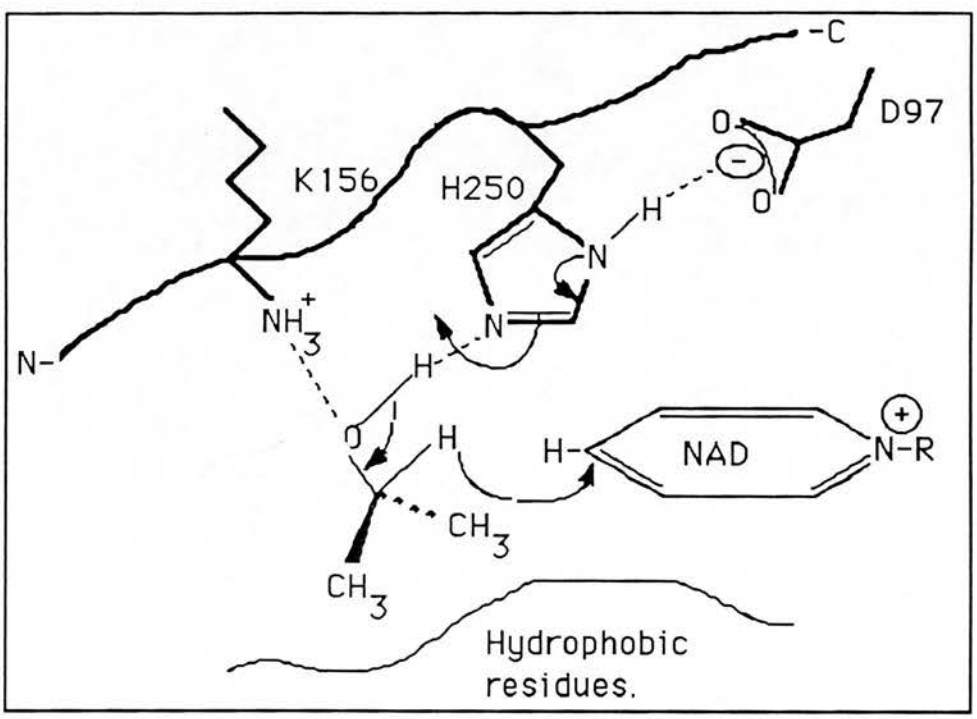


Figure 5.3-XIIIb . The C-terminal loop closes over the active site and His 250 acts as a general base promoting the hydride transfer to NAD^+.

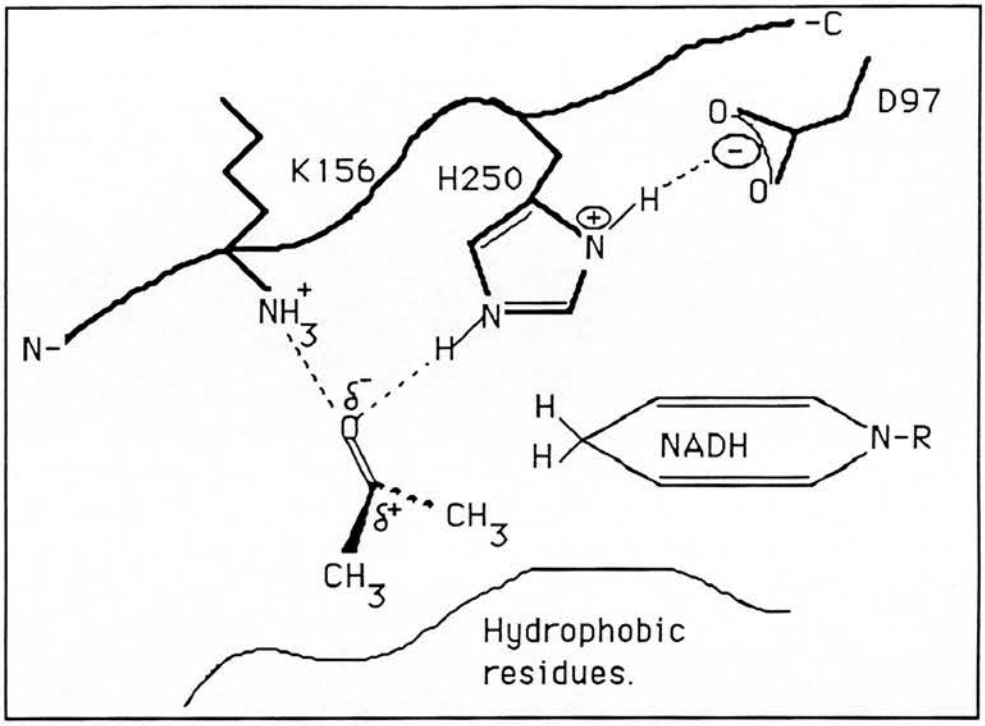


Figure 5.3-XIIIc. The product ketone is formed by transfer of the proton of the hydroxyl to His 250. The imidazolium ion formed is stabilized by Asp 97.

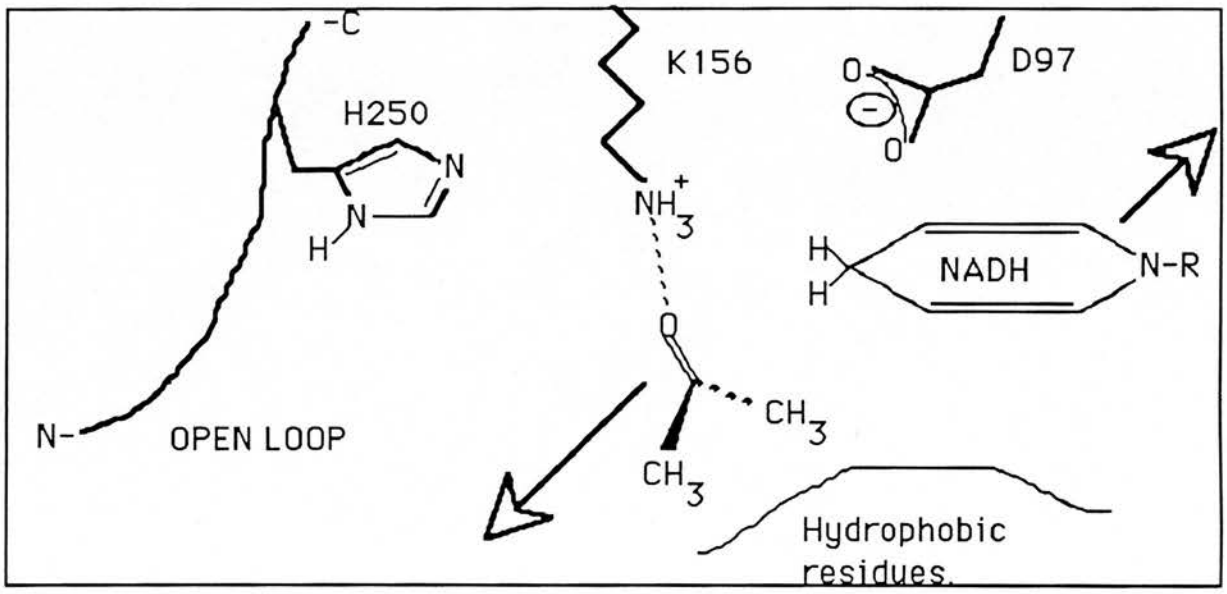


Figure 5.3-XIIId. The C-terminal loop opens and products dissociate from the enzyme. The imidazolium ion looses the stabilizing effect of Asp 97 and deprotonates.

This reaction mechanism offers explanations for the effect of chemical modification of histidines, and for the rate-limiting effect of the release of NADH, as this would be a function of the opening of the C-terminal loop, also, all the residues involved are totally conserved in all the ADHs sequenced so far.

It should be stressed, however, that this reaction mechanism can not apply to all the group of short-chain dehydrogenases, because the residue His250 is not conserved in the family (Persson et al., 1991). Whether this is due to the existence of different kinds of reaction mechanisms in short-chain dehydrogenases, to the fact that *Drosophila* ADH does not use His250 in its reaction mechanism, or to different positionings of catalytic His residues in different members of the family will require further investigation.

The positioning of the isopropanol molecule, which is highly determined by the presence of the cofactor and other sidechains, is such to pack the methyl groups of the substrate against residue 140, a valine also conserved in all ADHs. This is consistent with the observations of Winberg and colleagues, which lead them to believe that each methyl group of the substrate interacts with a hydrophobic group of the enzyme. Tyrosine 152, is part of the active site architecture, and it could be hydrogen bonding to the cofactor. Although its mutation to phenylalanine inactivates the enzyme (Albalat et al., 1992) that could be due to non-specific perturbations of the active site structure.

Holbrook and colleagues (Clarke et al., 1989) have studied the reaction mechanism of lactate dehydrogenase with a series of site-directed mutants whose effects were found to be highly predictable using their reaction mechanism.

The same kind of analysis, helped by the model, should be undertaken to prove or dismiss the hypothetical reaction centre proposed here.

5.3.2.2 Monomer-monomer binding site. Results and discussion.

The three-dimensional structure of SDH shows that the four subunits forming the tetramer assemble through two main interactions. One is formed through the antiparallel pairing of the last strand of the fold of two subunits, forming a twisted sheet of 14 strands that runs along two subunits. This type of interaction is similar to the assembly method of horse liver alcohol dehydrogenase (Eklund And Brändén, 1987), where the two nucleotide-binding folds are joined to form the dimer.

The other main interaction involves the pairing of helices D and E against the same helices of another subunit. Each helix runs along its equivalent helix, creating a two-fold axis of symmetry. The most important of this is the contact between helix E and its preceding loop, with themselves. In this interaction less important contacts exist between the large loop following helix-G of one subunit with the C-terminal tail of the other subunit.

ADH is a dimer, and if, as suspected, shares a common origin with SDH, it is reasonable to expect that its monomer-monomer binding site will be one of the two found in SDH. If this is accepted, then a certain sequence similarity between the common recognition sequences is also expectable, while the region that serves as monomer-monomer binding site only in SDH will probably have evolved differently, as it has different constrains in each of the enzymes.

It is obvious from the alignments shown in figures 5.3-I and 5.3-II, that the region at the C-terminal of both enzymes is highly divergent, while there is a region of high similarity between positions 127 to 170 of the ADH sequence. The last twenty-odd residues of that region correspond to the helix-E of SDH, being one of the helical regions of SDH with highest identity with a sequence of ADH.

The similarity in this region is even higher between ADH and prostaglandin dehydrogenase, making that region of the sequence the most likely subunit interaction area of ADH. The construction of the dimer was attempted by duplicating the model structure for the monomer and assembling the two monomers with the same orientation of their helices E as the one found in SDH (figure 5.3-XIV). Interestingly, the regions modelled as helices D and E in ADH have a large number of hydrophobic residues, that can not be organized in the normal amphiphilic distribution of helices on the surface of proteins. This abundance of hydrophobic residues may be the reason why the secondary structure predictions of this part of the enzyme has a prediction for strands, despite being aligned with parts of SDH that are helical. The abundance of hydrophobic residues is also consistent with a monomer-monomer binding role, as this kind of surfaces are very rich in hydrophobic-hydrophobic interactions (Richardson, 1985).

Once a model dimer of ADH is constructed in this way the main features of the interactions between the two monomers are as follows :

The main helix-helix interaction is that of helix-E, with its N-terminal loop. The loop of one monomer interacts with the C-terminal end of the same helix of the second monomer, through a hydrophobic packing of the residues Val 149 and Pro 150 of the loop with Leu 166 and Leu 169 of the helix. With the exception of *D. affinisdisjuncta* and *D. picta* (where the residue Leu 166 is isoleucine), the four residues are conserved among the 25 ADHs aligned.

A second hydrophobic interaction between the central part of the pair of helices-E seem to occur between residue Phe 162 and Gly 154, residue 162 is conserved in all ADHs, and in prostaglandin dehydrogenase, residue 154 is a glycine in half of the sequences and an alanine in the rest, including prostaglandin dehydrogenase.

At the middle point of this helices interaction, residues Ala 157 and Ala 158 pack against their equivalents in the other helix, both residues are also conserved among all ADHs. Finally, the sidechain of Ser 165 is at hydrogen bonding distance (2.9 Å) of the Gln 148 sidechain of the pairing helix, both residues are also totally conserved, with the exception of *D. mojavensis*, where position 148 is a proline in the enzyme encoded by the second *Adh* gene of this species.

The second helix-helix interaction is the antiparallel pairing of helix-E. In this case a very interesting combination of hydrophobic interactions occurs. Helix-E has three pairs of hydrophobic residues pointing out of the enzyme's core, Leu 112 and Val 113, at the N-terminus of the helix, Ile 105 and Ala 106 at the middle and Ile 119 and Leu 120 at the C-terminus. At the pairing of the helices the first pair of residues faces the last pair of the other helix, while the middle pair packs against its equivalent. The middle pair, Ile 105-Ala 106 is conserved through all the ADHs, but the other two pairs are not.

Half of the sequences of the aligned ADHs have Leu-Val as first pair, while the other half have Thr-Val, and exactly the same half that have Leu-Val as first pair has Ile-Leu as third pair, while the rest have now Ile-Met.

Clearly, if the interaction between the two helices is totally symmetric, the first residue of the first pair will face the second residue of the third pair, and vice-versa. And it is highly interesting to see that the enzymes where the first residue of the first pair is hydrophobic (Leu), have a hydrophobic residue as second amino acid of the third pair, while the enzymes where the first residue of the first pair has hydrogen-bonding capabilities (Thr), have a second residue of the third pair with the same capacity (Met).

This suggests that the evolution of this amino acids may have been coupled, perhaps to ensure optimal arrangement of the sidechains at this interaction. It would be interesting to check whether the two groups of enzymes differ in strength of monomer-monomer association.

Finally, the association of the models suggests the possibility of interactions between the C-terminal tail of one monomer with the large 30-residues loop between helix-G and strand-G, from positions 200 to 230. This suggestion would be too weak to be mentioned if it was not the only simple way to explain the effect of the mutation P214S on the stability of the enzyme.

The residue 214 of the model, located at the mentioned loop, is too far away from the core of the enzyme, or from the active site of its own monomer for the effect of its change to serine to be easily explained. If, however, the loop that contains it is interacting with the other subunit of the dimer, the change in stability due to the mutation may be related to the strength of that interaction. Clearly, further experimentation will be necessary to test this possibility.

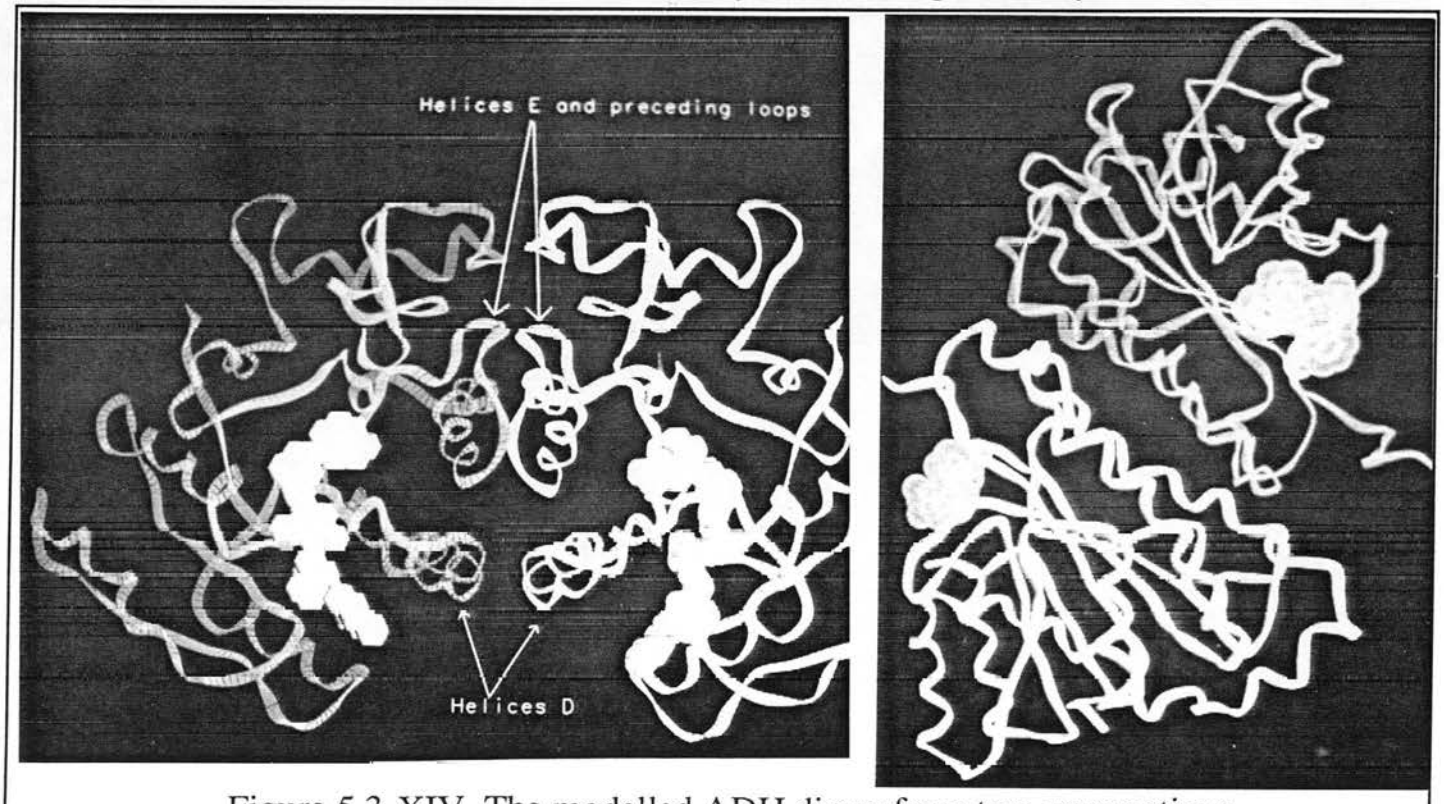


Figure 5.3-XIV. The modelled ADH dimer from two perspectives.
An NAD^+ molecule is added to each monomer at their proposed location.

5.3.2.4 Assessment of the structural correctness of the model.

The compatibility of the model with the existing biochemical data on ADH has been discussed in the previous sections of this chapter. Other methods of assessment that are independent of biological function were used to test for structural incongruities in the proposed model.

The Ramachandran plot of SDH shows six nonglycine residues in non allowed regions of the plot (Ghosh et al., 1991), all belonging to loop areas of poor electron density. The model of *Drosophila* ADH has 21 residues in non allowed regions of the plot. Of these, eleven are located at the large loop preceding strand-G, which could not be modelled. Six are at the C-terminal tail, which was not modelled either, and the 4 remaining ones are distributed in three loop regions. That result indicates that the two non-modelled regions of the structure probably have very different conformations to SDH.

A more sophisticated method of assessing the structural correctness of a protein structure has been devised recently by Lüthy and colleagues (Lüthy et al., 1992). The method, included in the program PROFILE, tests the structural likeliness of a set of coordinates by checking the environment of each residue and comparing it to the average environments found for residues in the same three-dimensional positioning in a database of high-resolution structures.

A set of 20 parameters is used to define each residue, and its environments are defined by three parameters: the area of the residue that is buried, the fraction of the side-chain area that is covered by polar atoms (O and N) and the local secondary structure.

The compatibility of segments of the sequence with their assigned three-dimensional structure can be assessed by plotting the average score of each residue (in a window of defined length) against the sequence number. The method has been proven able to discriminate between correct models and misfolded proteins (Lüthy et al., 1992), and to detect mistaken regions of crystallographic structures. As an average the authors found that correct models scored generally above 0.2 all around the sequence, and never had negative values (figure 5.3-XV).

The scoring plots calculated by the authors for various proteins shows a large variability in the score along their sequences. This is expectable since the method, albeit more sophisticated and thorough than the other assessment techniques, still relies in the statistical occurrence of certain kind of environments and their variability in the database used. Regions that are highly constrained structurally in the database are likely to give extreme scores, because the similar regions of the model will either fit the database largely or present large differences. in contrast variable structural domains may show a more graded score as different levels of similarity may be encountered. Seemingly, protein regions that are poorly represented in the database may give systematical low scores, their low frequency being interpreted as structural errors.

Although the authors do not discuss the statistical bias that the database used may be introducing in the analysis, it is reasonable to expect that regions were the environments of residues are not constant, like subunit interaction regions for instance, will score low values, while regions like B-sheets, where the packing of side-chains follows general and abundant patterns, will be judged with higher accuracy.

The results obtained on the SDH structure constructed from the α-carbon coordinates are shown in figure 5.3-XV. The plot shows large fluctuations in the scores obtained along the sequence.

Since the side-chain environments were constructed artificially it is difficult to use this analysis to asses the quality of the structure, however the side-chains were added in a knowledge-based way, and regions with low scoring may be indicating regions of the enzyme were the structure shows peculiarities. On the other hand the regions with higher scores probably correspond to regions of the molecule were the side-chains have been correctly constructed thanks to a regular backbone fold. 110 residues (43% of the total sequence of SDH) has scores below 0.2, which would be considered a sign of misfolding, although such conclusion can not be reached here, as discussed earlier. Three regions (residues 55-70, 90-105 and 188-192) score below zero with some or all of the window lengths used for the analysis (see figure 5.3-XV, and figure 1-I, page 12, for the structure of SDH). The first of them corresponds to strand C, which is partly exposed. The other two regions correspond to loop regions at the C-terminal ends of strands D and F. The regions scoring above 0.2 cover regions 10-35, 75-85, 115-180 and 200-210. This covers the majority of the sequence that forms the internal B-strands of the monomer (not the last strand, G), helices A, C, part of helix D and helix E. The initial region of the large loop between residues 202-234 scores also above 0.2, but the quality of the model decreases steadily from then on.

The length of the window used to analyze the structures is arbitrary; although the authors of the program suggest a window of 21 residues I could find no significant differences in the overall shape of the plot using windows of 15, 20, 25 or 30 residues. The plots shown in this thesis were produced using windows of 20 residues.

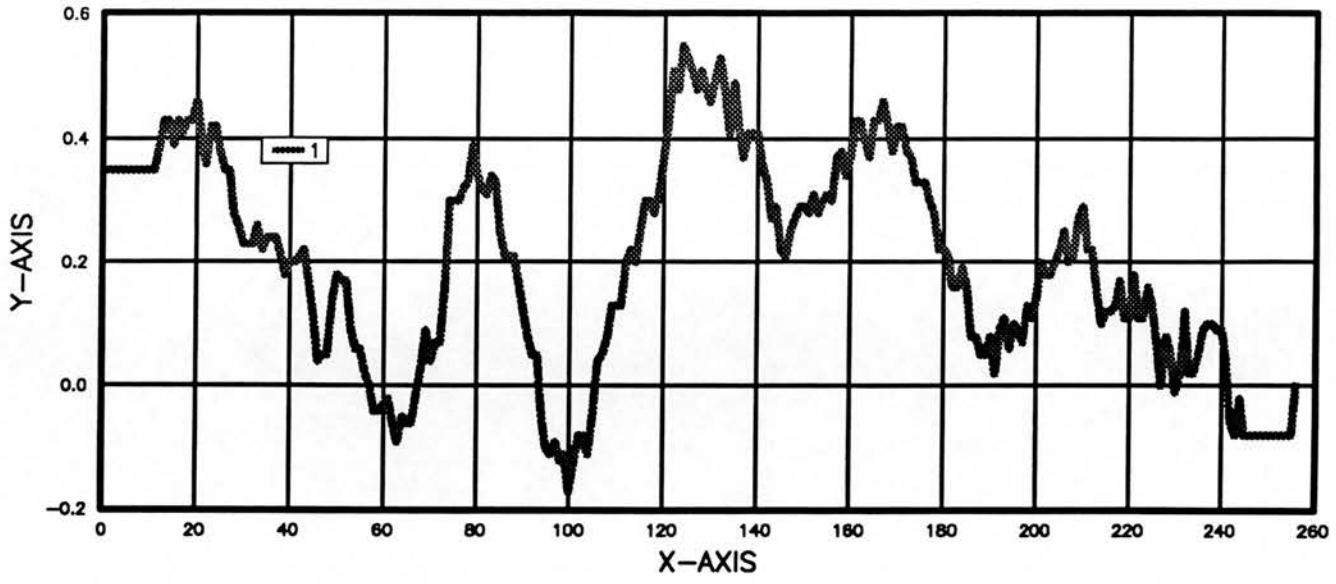


Figure 5.3-XV. PROFILE plot for the SDH structure used as basis for the modelling of ADH. Protein sequences form the X axis and the score at each position the Y axis.

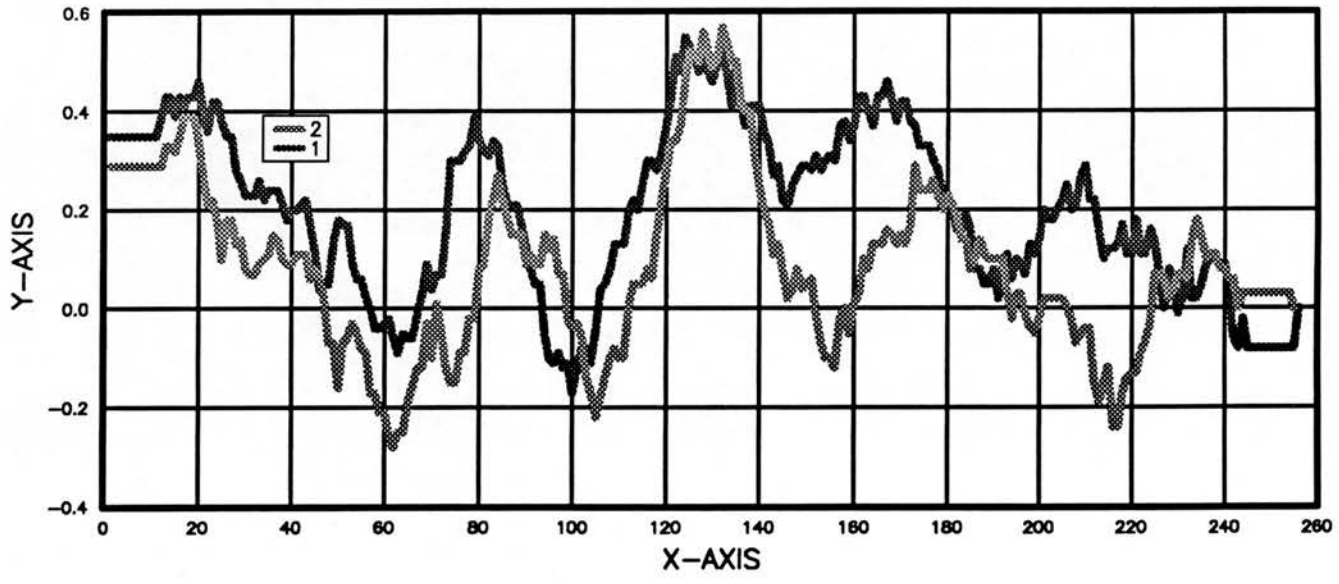


Figure 5.3-XVI. Comparision of the PROFILE plot for SDH (dark line) with the plot for the ADH model (lighter line). Protein sequences form the X axis and the score at each position the Y axis.

The external regions of the structure generally perform worse than the internal ones, and regions in the two main subunit interactions of the tetramer have scores below 0.1 (residues 90-110 which correspond to helix D and its preceding loop, and 230-245 which cover the last strand of the structure), confirming the possibility that regions with peculiar structural constraints may be underscored by the program.

The comparison of the PROFILE plots for ADH and SDH (figure 5.3-XVI) immediately shows that both structures behave in a very similar way along their sequences although the score for the ADH model is consistently lower (by a difference of about 0.1) than that for SDH.

A possible interpretation of this behaviour may be the following : the construction of a model of ADH based on the SDH structure can introduce two main kind of architectural errors: firstly, errors of ineffective packing, that is regions where the general fold of the model is correctly similar to the frame structure but minor changes in the environment of the side chains have not been well accounted for, and produce a general all-along decrease in the quality of the final structure; secondly, gross errors of folding, where the final model is largely incorrect, due to big differences between the frame structure and the real architecture of the modelled enzyme.

The first kind of error may be expected to produce a uniform decrease in the score of the model as calculated by PROFILE, but the general pattern of the plot would stay similar because the sequence of the model is equally suited to its fold as the one of the original structure.

In the second case, however, one may expect large changes in the pattern of the plot, because as sequences that actually create different

kinds of folds are compared, their suitability for the measured structure will be very different.

The region between residues 10 to 85 of SDH and ADH seem to show the first kind of error in the model. The pattern of both plots is largely the same, but the ADH model accumulates a constant amount of penalty which may be assigned to small mistakes in the side-chain packing. The region between residues 90 and 140, is the best performing one in both enzymes, and buy very similar values.

The region 140-190, where the hypothetical active site of ADH has been proposed (see 5.3.2.2) displays again a parallel behaviour between SDH and ADH, but the difference between the score of both models is here of about 0.2 in favour of SDH.

Between residues 190 and 225 a difference in the pattern appears, and the score of ADH continues to fall while SDH improves in the plot. This is what one would expect in the second kind of mistake discussed earlier, and it is very interesting to see this happening in the part of the sequence that covers the large loop between residues 200 and 230, which has already been described as a largely unmodelable region, due to its highly undefined nature. Very significantly, the quality of the ADH score rises again when the sequence of the modelled last strand comes along, between residues 235 and 239, reaching similar values to SDH in this zone.

If the interpretation of the differences between both plots discussed earlier is correct, the differences found in the PROFILE analysis of SDH and ADH suggest the following conclusions :

- The general fold of SDH is compatible with the sequences of ADH that have been assigned to it, although a certain amount of systematical error is introduced mainly in the regions 10-85 and 140-190.

- The region between residues 85-180 of ADH seems correctly folded, implicating that the whole N-terminal region of the model up to residue 180 has been correctly assigned to the sequence of ADH.

- Although the profiles follow the same pattern a large difference of about 0.2 exists between both structures in the region 140-180, this is the region proposed as active site of ADH. The difference may be due to different functionalities of that part of the structure between both enzymes, and/or to mistakes in the final packing of the side-chains of ADH in that area.

The fact that helix-E, the main subunit interaction surface of the proposed ADH dimer, is in this part of the structure may be another reason for the divergence, as the orientation of the residues of this helix is difficult to define in the model.

- The loop region between residues 200-230 is probably wrong in the model, and this region of ADH may be very different to the equivalent area of SDH. That was suggested too by the Ramachandran plots of the ADH model.

Finally, the sequence assigned to the last strand of ADH is probably correctly located, as it reaches a similar score to the one obtained by the SDH structure in the same region. Unfortunately nothing can be said about the C-terminal tail of the enzymes, because the window regions used for the analysis leave the values for the first and last residues of the plot without any meaning.

6. Final conclusions and future work.

1- A suitable system for the mutagenesis and expression of *Drosophila* ADH in yeast has been devised and a fast and simple method for the purification of ADH from yeast cells has also been produced. The method has been used to purify wt ADH, and three of the four mutant enzymes studied, to homogeneity. The method should be modified to improve the purification of the enzyme ADH D38A.

2- The sequence of the chimeric gene used for the expression of ADH was obtained and several unreported mutations were found. One of them encodes for a new amino acid difference in the enzyme, which does not seem to have any effect on the functional characteristics of the protein.

3-A denaturation assay, coupled to enzyme activity determination, for the calculation of the $\Delta G_{(water)}$ values of the enzymes and the $\Delta\Delta G$ effect of cofactor binding was devised. The assay, and circular dichroism measurements were used to study the denaturation of *Drosophila* ADH-F, ADH-S, horse liver ADH and yeast ADH (see paper attached).

4-Four different site-directed mutations were introduced in ADH. The mutant enzymes were totally sequenced, expressed in yeast, purified and analyzed kinetically. Their $\Delta G_{(water)}$ values and the $\Delta\Delta G$ effect of cofactor binding were calculated for all of them. They were also analyzed for their resistance to thermal denaturation.

5- The mutation performed in residue P214 confirms that the thermostable effect that this mutation confers to the isoenzyme ADH-F can be transfered to an ADH-S isoenzyme reproducing the P214S change. The mutation G19A was found to knock out detectable activity with NADP, suggesting a positioning of the cofactor that is in contradiction with that found by Ghosh and colleagues (1991) in the structure of SDH. Similarly the mutations at positions 14 and 38 confirmed the importance of that part of the molecule for the recognition of NAD.

The results obtained complement and confirm other work published during the development of this thesis on the importance of residues G14, and D38 of ADH in the recognition of NAD.

6- Based on the structure of SDH a computer model was constructed to try to find structural explanations for the observed effects of the mutations introduced. The model is largely compatible with the existent biochemical data on the enzyme. The positioning of NAD^+ supported by the results of the G19A mutation suggests a well defined possible active site. A reaction mechanism based on that putative reaction centre is proposed. The mechanism is very similar to that of lactate dehydrogenase, and would involve the closure and opening of the C-terminal tail of the enzyme which carries the proposed catalytic residue H250.

7- A possible monomer-monomer interaction site was identified in ADH, and a hypothetical model of the dimeric structure was constructed.

The project presented here describes the set-up and initial utilization of a general system for the study of an enzyme. There are several ways in which the project may be continued:

1- If large quantities of the enzyme are desired for the production of crystals the expression system should be optimized, possibly by attempting the expression of the enzyme in bacteria.

2- The structural model presented in this thesis does now suggest a larger number of residues that can be studied for their possible role in the mechanisms and structure of the enzyme. This is a significant improvement as it introduces a rational basis to the design of further ADH mutants. A number of mutations can now be introduced to test the hypothesized active-site and the monomer-monomer interaction surface, the recognition mechanism of NAD, the broad substrate specificity displayed by the enzyme, the function of the region around residue 192, and the role of the large loop at the region 200-235.

3- The model itself should be improved as better data for SDH becomes available. An improved model might be used efficiently for the application of molecular replacement techniques to the existing crystallographic data on ADH.

7. Bibliography.

Adams, M.J. (1987) in Enzyme mechanisms (Page, M.I. & Williams, A., eds.), pp. 477-505, The Royal Society of Chemistry, Cambridge, UK

Agarwal, A.K., Monder, C., Eckstein, B. & White, P.C. (1989) J. Biol. Chem. **264**, 18939-18943

Albalat, R., Gonzàlez-Duarte, R. & Atrian, S. (1992) FEBS lett. **308**, 235-239

Argos, P. (1990) Methods Enzymol. **182**, 751-776

Argos, P., Vingron, M. & Vogt, G. (1991) Protein Engng. **4**, 375-383

Ashburner, M. (1985) Dros. Inform. Serv. **62**, 9-13

Atkinson, P.W., Mills, L.E., Starmer, W.T. & Sullivan, D.T. (1988) Genetics **120**, 713-723

Atrian, S., Gonzalez-Duarte, R. & Fothergill-Gilmore, L.A. (1990) Gene **93**, 205-212

Atrian, S. & Gonzàlez-Duarte, R. (1982) in Advances in Genetics, Development and Evolution of Drosophila (Lakovaara, S., ed.), pp. 251-261, Plenum Press, New york

Atrian, S. & Gonzàlez-Duarte, R. (1985) Biochem. Genet. **23**, 891-911z

Beggs, J.D. (1978) Nature **275**, 104-109

Benyajati, C., Place, A.R., Powers, D.A. & Sofer, W. (1980) Proc. Natl. Acad. Sci. U. S. A. **78**, 2717-2721

Benyajati, C., Spoerel, N., Haymerle, H. & Ashburner, M. (1983) Cell **33**, 125-133

Beverley, S.M. & Wilson, A.C. (1985) Proc. Natl. Acad. Sci. U. S. A. **82**, 4753-4757

Blundell, T.L., Sibanda, B.L., Sternberg, M.J.E. & Thornton, J.M. (1987) Nature **326**, 347-352

Bodmer, M. & Ashburner, M. (1984) Nature **309**, 425-430

Borrás, T., Persson, B. & Jörnvall, H. (1989) Biochemistry **24**, 6133-6139

Bradford, M.M. (1976) Anal. Biochem. **72**, 248-254

Brändén, C.-I., Jörnvall, H., Eklund, H. & Furugren, B. (1975) in The Enzymes Vol. XI (P.D. Boyer, ed.) Academic Press, New York

Branden, C.-I. & Tooze, J. (1991) in Introduction to Protein Structures p. 145, Garland Publishers.

Brooks, W.M., Moxon, L.N., Field, J., Irving, M.G. & Doddrell, D.M. (1985) Biochem. Biophys. Res. Commun. **128**, 107-112

Brünger, A.T., Kuriyan, J. & Karplus, M. (1987) Science **235**, 458-460

Carter, P. (1986) Biochem. J. **237**, 1-7

Chambers, G.K. (1984) Biochem. Genet. **22**, 529-549

Chambers, G.K. (1988) in Advances in Genetics, Vol. 25 (Caspari, E.W. & Scandalios, J.G., eds.), pp. 39-107, Academic Press, Inc.,

Chambers, G.K. (1991) Comp. Biochem. Phys. **99**, 723-730

Chambers, G.K., Wilks, A.V. & Gibson, J.B. (1981) Aust. J. Biol. Sci. **34**, 625

Chen, Z., Lee, W.R. & Chang, S.H. (1991) Eur. J. Biochem. **202**, 263-267

Chen, Z., Lu, L., Shirley, M., Lee, W.R. & Chang, S.H. (1990) Biochemistry **29**, 1112-1118

Chothia, C. & Lesk, A.M. (1987) J. Mol. Biol. **196**, 901-918

Chou, P.Y. & Fasman, G.D. (1974) Biochemistry **13**, 222-245

Claessens, M., Van Cutsem, E., Lasters, I. & Wodak, S. (1989) Protein Engng. **2**, 335-345

Clark, W. (1988) PhD Thesis, Rutgers University.

Clarke, A.R., Atkinson, T. & Holbrook, J.J. (1989) Trends in Biochemical Sciences **14**, 101-104

Cornish-Bowden, A. & Wharton, R.P. (1988) in Introduction to enzyme kinetics: In Focus (Rickwood, D., ed.) IRL press, Oxford

Coyne, J.A. & Kreitman, M. (1986) Evolution **40**, 673-691

Creighton, T.E.. (1984) in Proteins pp. 150-151, W.H. Freeman, New York

Day, T.H., Hillier, P.C. & Clarke, B (1974) Biochem. Genet. **11**, 141.

Dayhoff, M.O., Barker, W.C. & Hunt, L.T. (1983) Methods Enzymol. **91**, 524-545

Debellé, S. & Sharma, S.B. (1986) Nucleic Acids Res. **14**, 7453-7472

Devereaux, J., Haeberli, P. & Smithies, O.A. (1984) Nuc. Acids Res. **12**, 765-784

Doolitle, R.F.(editor of the volume, the whole book is relevant) (1990) Methods Enzymol. **183**

Dothie, J.M., Giglio, J.R., Moore, C.H., Taylor, S.S. & Hartley, B.S. (1985) Biochem. J. **230**, 569-578

Dufton, M.J. & Hider, R.C. (1977) J. Mol. Biol. **115**, 177-193

Eklund, H. & Brändén, C.-I. (1987) in Biological Macromolecules & Assembly, Vol. 3 (Jurnak, & McPherson, , eds.), pp. 74-142, Wiley,

Eklund, H., Nordström, B., Zeppezauer, E., Söderlund, G., Ohlsson, I., Boiwe, T., Söderberg, T., Tapia, O., Brändén, C.-I. & Åkeson, Å. (1976) J. Mol. Biol. **102**, 27-59

Eklund, H., Samama, J.-P. & Jones, T.A. (1984) Biochemistry **23**, 5982-5996

England, B.P., Heberlein, U. & Tjian, R. (1990) J. Biol. Chem. **265**, 5086-5094

Ensor, C.M. & Tai, H.H. (1991) BBRC **176**, 840-845

Feeney, R., Clarke, A.R. & Holbrook, J.J. (1990) Biochem. Biophys. Res. Commun. **166**, 667-672

Fersht, A. (1987) Trends in Biochemical Sciences **12**, 301-309

Fibla, J., Enjuanes, L. & Gonzàlez-Duarte, R. (1989) BBRC **160**, 638-646

Fischer, J.A. & Maniatis, T. (1985) Nucleic Acids Res. **13**, 6899-6917

Furukawa, K., Arimura, N. & Miyazaki, T. (1987) J. Bacteriol. **169**, 427-429

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

Garrat, R.C., Thornton, J.M. & Taylor, W.R. (1991) FEBS lett. **280**, 141-146

Geer, B.W., Langevin, M.L. & McKechnie, S.W. (1985) Biochem. Genet. **23**, 607-622

Gernstein, M. & Chothia, C. (1991) J. Mol . Biol. **220**, 133-149

Getz, C. & Grant, W.S. (1990) Biochem. Genet. **28**, 503-507

Ghosh, D., Weeks, C.M., Grochulski, P., Duax, W.L., Erman, M., Rimsay, R.L. & Orr, J.C. (1991) Proc. Natl. Acad. Sci. U. S. A. **88**, 10064-10068

Gibson, J.B., Chambers, G.K., Wilks, A.V. & Oakeshott, J.G. (1980) Aust. J. Biol. Sci **33**, 479

Gibson, J.B. & Wilks, A.V. (1988) Heredity **60**, 403-407

Gilbert, R.J. (1992) J. Mol. Graphics **10**, 112-119

Goldberg, D.A. (1980) Proc. Natl. Acad. Sci. U. S. A. **77**, 5794-5798

González-Duarte, R. & Atrian, S. (1986) Heredity **17**, 123-128

Grantham, R.C., Gautier, C., Gouy, M., Jacobzone, M. & Mercier, R. (1981) Nucleic Acids Res. **9**, 43-74

Grau, U.M. (1982) in The Pyridine Nucleotide Coenzymes p. 135 (Everse, J., Anderson, B. & You, K., eds.) Academic Press, New York

Green, M.M., Fang, X., Churchill, P. & Brennan, M.D. (1989) Arch. Biochem. Biophys. **273**, 440-448

Green, M.S., Meeker, M.K. & Shortle, D. (1992) Biochemistry **31**, 5717-5728

Greer, J. (1981) J. Mol. Biol. **153**, 1027-1042

Grell, E.H., Jacobson, K.B. & Murphy, J.B. (1965) Science **149**, 80-82

Grossman, A.I., Koreneva, L.G. & Ulitskaya, L.E. (1970) Sov. Genet. **6**, 211-214

Hallam, S.E., Malpartida, F. & Hopwood, D.A. (1988) Gene **74**, 305-320

Heinstra, P.W.H., Thörig, G.E.W., Scharloo, W., Drenth, W. & Nolte, R.J.M. (1988) Biochim. Biophys. Acta. **967**, 224-228

Hernández, J.J., Vilageliu, Ll. & Gonzàlez-Duarte, R. (1988) Genetica **77**, 15-24

Hernández, M., Padrón, G. & Cabrera, V.M. (1986) Genet. Res. Camb. **47**, 143-146

Higgins, P.G. & Sharp, P.M. (1988) Gene **73**, 237-244

Hovik, R., Winberg, J.O. & McKinley-McKee, J.S. (1984) Insect Biochem. **14**, 345-351

Inoue, T., Sunagawa, M., Mori, A., Chuhei, I., Fukuda, M., Tagaki, M. & Yano, K. (1989) J. Bacteriol. **171**, 3115-3122

Irie, S., Doi, S., Yorifuji, T., Tkagi, M. & Yano, K. (1987) J. Bacteriol. **169**, 5174-5179

Ito, H., Fukuda, I., Murata, K. & Kimura, A. (1983) J. Bacteriol. **153**, 163-168

Jacobsen, K.B., Murphy, J.B., Knopp, J.A. & Ortiz, J.R. (1972) Arch. Biochem. Biophys. **149**, 22

Jany, K.-D., Ulmer, W., Fröscle, M. & Pfleiderer, G. (1984) FEBS lett. **165**, 6-10

Johnson, F.M. & Denniston, C. (1992) Nature **204**, 906-907

Jörnvall, H., Persson, M. & Jeffrey, J. (1981) Proc. Natl. Acad. Sci. U. S. A. **78**, 4226-4230

Jörnvall, H., Von Bahr-Lindström, H., Jany, K.D., Ulmer, W. & Fröschle, M. (1984) FEBS lett. **165**, 190-196

Karplus, M. & Petsko, G.A. (1990) Nature **347**, 631-639

Kimura, M. (1983) The Neutral Theory of Molecular Evolution, Cambridge Univ. Press, London and New York

Kreitman, M. (1983) Nature **304**, 412-417

Krook, M., Marekov, L. & Jörnvall, H. (1990) Biochemistry **29**, 738-743

Kunkel, T.A. (1985) *Proc. Natl. Acad. Sci.* **82**, 488-492

Layne, E. (1957) Methods Enzymol. **3**, 447-454

Liu, J., Duncan, K. & Walsh, C.T. (1989) J. Bacteriol. **171**, 791-798

Levin, J.M. & Garnier, J. (1988) Biochim. Biophys. Acta **955**, 283-295

Lilley, K.S., Baker, P.J., Britton, K.L., Stillman, T.J., Brown, P.E., Moir, A.J.G., Engel, P.C., Rice, D.W., Bell, J.E. & Bell, E. (1991) Biochim. Biophys. Acta **1080**, 191-197

Lim, V.I. (1974) J. Mol. Biol. **88**, 857-872

Lüthy, R., Bowie, J.U. & Eisenberg, D. (1992) Nature **356**, 83-85

Madhavan, K., Conscience-Egli, M., Sieber, F & Ursprung, H. (1973) J. Insect Physiol. **19**, 235-241

Maniatis, T., Hardison, R.C., Lacy, E., Lauer, J., O'Connel, C., Quon, D., Sim, J.K. & Efstratiadis, A. (1978) Cell **15**, 687-701

Marekov, L., Krook, M. & Jörnvall, H. (1990) FEBS lett. **266**, 51-54

Maroni, G., Laurie-Ahlberg, C.C., Adams, D.A. & Wilson, A.N. (1982) Genetics **101**, 431-446

McDonald, J.H. & Kreitman, M. (1991) Nature **351**, 652-654

Middleton, R.J. & Kacser, H. (1983) Genetics **105**, 663-650

Nagano, K. (1973) J. Mol. Biol. **75**, 401-405

Navre, M. & Ringold, G.M. (1988) J. Cell Biol. **107**, 279-286

Needleman, S.B. & Wunsch, C.D. (1970) J. Mol. Biol. **48**, 443-453

Neidle, E.L., Hartnett, C., Ornston, N.L., Bairoch, A., Rekik, M. & Harayama, S. (1989) Swiss Prot Databank

Nozaki, Y. (1972) Methods Enzymol. **26**, 43-50

Oakeshott, J.G., Gibson, J.B., Anderson, P.R., Knibb, W.R., Anderson, D.G. & Chambers, G.K. (1982) Evolution **36**, 86-96

Oakeshott, J.G., Gibson, J.B. & Wilson, S.R. (1984) Heredity **53**, 51-67

Pace, N.C. (1986) Methods Enzymol. (1986) **131**, 266-280

Parker, D.M. & Holbriook, J.J. (1977) in Pyridine Nucleotide Dependent Enzymes p. 485 (Sund, H., ed.) Walter de Gruyter, Berlin

Parsons, P.A. (1981) Aust. J. Zool. **29**, 33-39

Parsons, P.A. & Spence, G.E. (1981) Am. Nat. **117**, 568-571

Peltoketo, H., Isomaa, V., Mäentausta, O. & Vihko, R. (1988) FEBS lett. **239**, 73-77

Peoples, O.R. & Sinskey, A.J. (1989) J. Biol. Chem. **264**, 15293-15297

Persson, B., Krook, M. & Jörnvall, H. (1991) Eur. J. Biochem. **200**, 537-543

Petterson, G. (1987) Crit. Rev. Biochem. **21**, 349

Pipkin, S.B., Franklin-Springer, E., Law, S. & Lubega, S. (1976) J. Hered. **67**, 258-266

Provencher, S.W. & Glöckner, J. (1981) Biochemistry **20**, 33-37

Ramachandran, G.N. & Sasisekharan, V. (1968) Adv. Prot. Chem. **23**, 283-437

Retzios, A. (1982) PhD Thesis, Edinburgh University, Edinburgh

Ribas de Pouplana, Ll., Atrian, S., Gonzàlez-Duarte, R., Fothergill-Gilmore, L.A., Kelly, S.M. & Price, N.C. (1991) Biochem. J. **276**, 433-438

Richardson, J.S. (1985) Methods Enzymol. **115**, 349-358

Rowan, R.G. & Dickinson, W.J. (1988) J. Mol. Evol. **28**, 43-54

Sambrook., J., Fritsch, E.F. & Maniatis, T. (1989) Molecular Cloning Vol. 1-3 Cold Spring Harbor Laboratory Press

Sampsell, B.M. & Barnette, V.C. (1985) Biochem. Genet. **23**, 53-59

Sander, C. & Schneider, R. (1991) Proteins **9**, 69-78

Schaeffer, S.W. & Aquadro, C.F. (1987) Genetics **117**, 61-73

Schimmel, P. (1987) Annu. Rev. Biochem. **56**, 125-158

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

Sekuzu, I., Yamashita, J., Nozaki, M., Hagihara, B., Yonetani, T. & Okunuki, K. (1957) J. Biochem. **44**, 601-614

Shrake, A. & Ross, P.D. (1990) J. Biol. Chem. **265**, 5055-5059

Sofer, W. & Martin, P. (1987) Ann. Rev. Genet. **21**, 203-225

Starmer, W.T. & Sullivan, D.T. (1989) Mol. Biol. Evol. **6**, 546-552

Strambini, G.B. & Gonelli, M. (1990) Biochemistry **29**, 196-203

Thatcher, D.R. (1977) Biochem. J. **163**, 317-323

Thatcher, D.R. (1980) Biochem J. **187**, 875-886

Thatcher, D.R. (1981) Biochem. Soc. Trans. **9**, 299

Thatcher, D.R. & Retzios, A. (1980) Protides Biol. Fluids **28**, 157-160

Thatcher, D.R. & Sawyer, L. (1980) Biochem J. **187**, 884-886

Thatcher, D.R. & Sheikh, R. (1981) Biochem. J. **197**, 111-117

Thöny, B., Fischer, H.-M., Anthamatten, D., Bruderer, D. & Hennecke, H. (1987) Nucleic Acids Res. **15**, 8479-8499

Van Gunsteren, W.F. & Mark, A.E. (1992) Eur. J. Biochem. **204**, 947-961

Vigue, C.L. & Johnson, F.M. (1973) Biochem. Genet. **9**, 213-226

Villaroya, A., Juan, E., Egestad, E. & Jörnvall, H. (1989) Eur. J. Biochem. **180**, 191-197

White, W.B., Franklund, C.V., Coleman, J.P. & Hylemon, P.B. (1988) J. Bacteriol. **170**, 4555-4561

Williams, D.H. (1991) Aldrichimica acta **24**, 71-80

Wills, C. (1990) Critical reviews in Biochem. and Mol. Biol. **25**, 245-277

Winberg, J.-O., Hovik, R. & McKinley-McKee, J.S. (1985) Biochem. Genet. **23**, 205-216

Winberg, J.-O., Hovik, R., McKinley-McKee, J.S., Juan, E. & Gonzalez-Duarte, R. (1986) Biochem J. **235**, 481-490

Winberg, J.-O., Thatcher, D.R. & McKinley-McKee, J.S. (1982) Biochim. Biophys. Acta **704**, 17-25

Winberg, J.-O., Thatcher, D.R. & McKinley-McKee, J.S. (1982) Biochim. Biophys. Acta **704**, 7-16

Winberg, J.-O., Thatcher, D.R. & McKinley-McKee, J.S. (1983) Biochim. Genet. **21**, 63-80

Winberg, J.O. (1989) PhD Thesis, University of Oslo, Oslo

Winberg, J.O. & McKinley-McKee, J.S. (1988) Biochem. J. **251**, 223-227

Winberg, J.O. & McKinley-McKee, J.S. (1988) Biochem. J. **255**, 589-599

Wiseman, A. & Williams, N.J. (1971) Biochim. Biophys. Acta **250**, 1-5

Yamada, M. & Saier, M.H.Jr (1987) J. Biol. Chem. **262**, 5455-5463

Yin, S.-J., Vagelopoulos, N., Lundquist, G. & Jörnvall, H. (1991) Eur. J. Biochem. **197**, 359-365

8. Published Article

Biochem. J. (1991) 276, 433–438 (Printed in Great Britain)

Structural properties of long- and short-chain alcohol dehydrogenases

Contribution of NAD^+ to stability

Lluís RIBAS DE POUPLANA,* Silvia ATRIAN,† Roser GONZÀLEZ-DUARTE,†
Linda A. FOTHERGILL-GILMORE,* Sharon M. KELLY‡ and Nicholas C. PRICE‡
*Department of Biochemistry, University of Edinburgh, George Square, Edinburgh EH8 9XD, Scotland, U.K.,
†Departament de Genetica, Universitat de Barcelona, Av. Diagonal 645, 08071 Barcelona, Spain, and
‡Department of Biological and Molecular Sciences, University of Stirling, Stirling FK9 4LA, Scotland, U.K.

Structural studies were undertaken on long-chain and short-chain alcohol dehydrogenases (from horse liver and *Drosophila* respectively). Far-u.v. c.d. measurements were used to estimate the secondary structure contents of the enzymes. For the horse liver enzyme, the results agree well with the X-ray data; for the *Drosophila* enzyme (for which a crystal structure is not yet available), the results are in good agreement with those obtained by applying a range of structure-prediction procedures to the amino acid sequence of this enzyme. The conformational stabilities of the two enzymes were investigated by studying the unfolding brought about by guanidinium chloride (GdnHCl) by using activity and c.d. measurements. The unfolding of the *Drosophila* enzyme was analysed in terms of a two-state model; the presence of the substrate NAD^+ leads to considerable protection against unfolding. By contrast, the unfolding of the horse liver enzyme shows a plateau effect at intermediate concentrations of GdnHCl, indicating that a two-state model is not appropriate in this case. NAD^+ affords little, if any, protection against unfolding for the horse liver enzyme.

INTRODUCTION

Alcohol dehydrogenases (ADH; EC 1.1.1.1) are enzymes that occur widely across the range of living organisms, where they are important for the detoxification and metabolism of ethanol and other alcohols. They catalyse the oxidation of primary and secondary alcohols to the corresponding aldehydes and ketones, with the concomitant reduction of NAD^+. They can be grouped into two main families (Jörnvall *et al.*, 1981, 1984), which appear to be unrelated at the sequence level. Long-chain ADHs are characterized by a preference for primary alcohols (Dalziel & Dickinson, 1966), by a requirement for Zn^{2+} ions (Dunn & Hutchinson, 1973) and by hydride transfer with 4-*pro*-*R* stereospecificity (You, 1982). In contrast, short-chain ADHs prefer secondary alcohols (Sofer & Ursprung, 1968; Winberg *et al.*, 1982), have no requirement for metal cofactors (Chambers, 1984) and show 4-*pro*-*S* stereospecificity (Benner *et al.*,1985). It is clear from these observations that the catalytic mechanisms of the two types of ADH must be different, although the overall reactions are the same.

Long-chain alcohol and polyol dehydrogenases have approx. 350 amino acid residues per subunit, and examples include the familiar mammalian liver ethanol dehydrogenase, as well as ADHs from birds, plants, yeast and bacteria. Mammalian sorbitol dehydrogenase is also a member of the long-chain ADH family (Jörnvall *et al.*, 1981, 1984). The short-chain ADHs have been found in mammals, insects and bacteria. These enzymes have approx. 250 amino acid residues per subunit, and examples include *Drosophila* ADH (Thatcher, 1980), bacterial ribitol, glutamate and 20β-hydroxysteroid dehydrogenases (Jörnvall *et al.*, 1984; Marekov *et al.*, 1990), and mammalian 17-hydroxysteroid (Peltoketo *et al.*, 1988) and 15-hydroxyprostaglandin (Krook *et al.*, 1990) dehydrogenases.

The three-dimensional structure of the long-chain ADH from horse liver is known in great detail from crystallographic and sequence studies (Eklund *et al.*, 1976; Jörnvall, 1970; reviewed in Eklund & Brändén, 1987). It is reasonable to assume that the other members of the homologous long-chain ADH family have similar topologies. The short-chain family is as yet much less well characterized, although suitable crystals of *Drosophila* ADH have recently been obtained, and diffraction data are currently being collected (E. Gordon & L. Sawyer, personal communication). In the absence of a crystal structure, it is of interest to explore predicted structures derived from sequences and spectroscopic data. A secondary-structure prediction for ADH from *Drosophila melanogaster* (Thatcher & Sawyer, 1980) gives a strong suggestion that the *N*-terminal half of the enzyme folds into the alternating α-helix/β-strand structure characteristic of nucleotide-binding domains [reviewed in Richardson (1981)]. The *C*-terminal half does not appear to have an alternating α/β structure.

Each subunit of the long-chain dehydrogenases (as exemplified by the horse liver enzyme) folds into two domains, with the active site at the interface between the domains [reviewed in Eklund & Brändén (1987)]. One molecule of NAD^+ binds to this interdomain region and makes contacts with both domains. It thus may be expected that the cofactor could stabilize the enzyme to a variety of insults such as heat or denaturing agents. It is the main aim of the present study to compare the structural properties of long-chain and short-chain ADHs (such as those from horse liver and from *Drosophila*), and the contribution which NAD^+ makes to the stability of the enzymes against unfolding by guanidinium chloride (GdnHCl).

EXPERIMENTAL

Materials

ADHs from horse liver and from yeast (*Saccharomyces cerevisiae*) were purchased from Boehringer and Sigma respectively. Before use, the crystalline suspension of liver ADH was dialysed for 18 h against 20 mM-Tris/HCl, pH 8.0. Yeast ADH was

Abbreviations used: GdnHCl, guanidinium chloride; ADH, alcohol dehydrogenase.

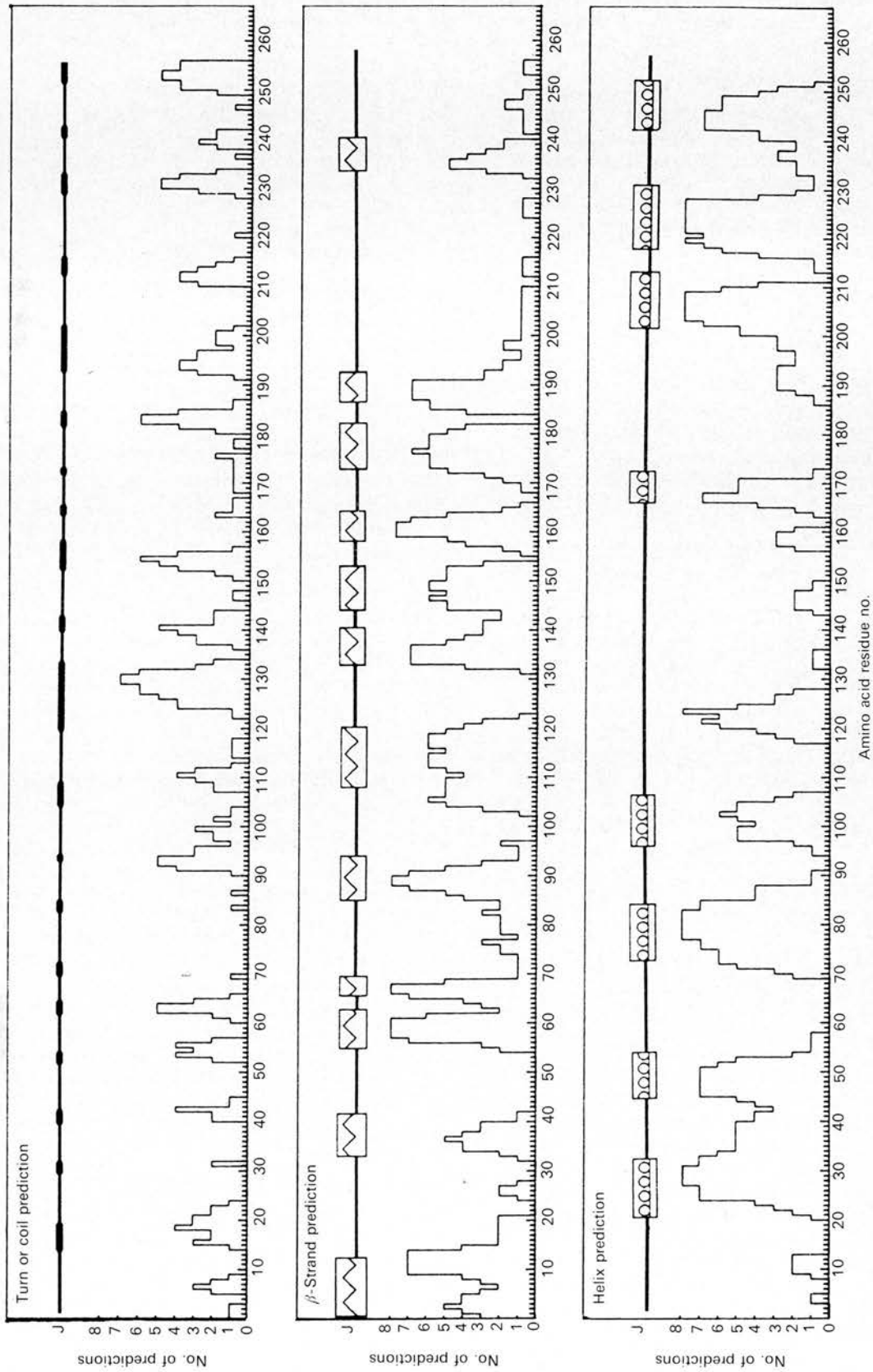


Fig. 2. Secondary structure prediction for *D. melanogaster* ADH

The histograms show the scores of the residues according to the prediction package PREDICT. The interpretation of those scores is shown in lines over the histograms. α-Helix, 29.6 %; *B*-strand, 33.5 %; other structures, 36.9 %. The interpretation of the results of the prediction procedure was undertaken in the light of the known secondary-structural pattern in liver ADH and other dehydrogenases (Eklund & Brändén, 1987).

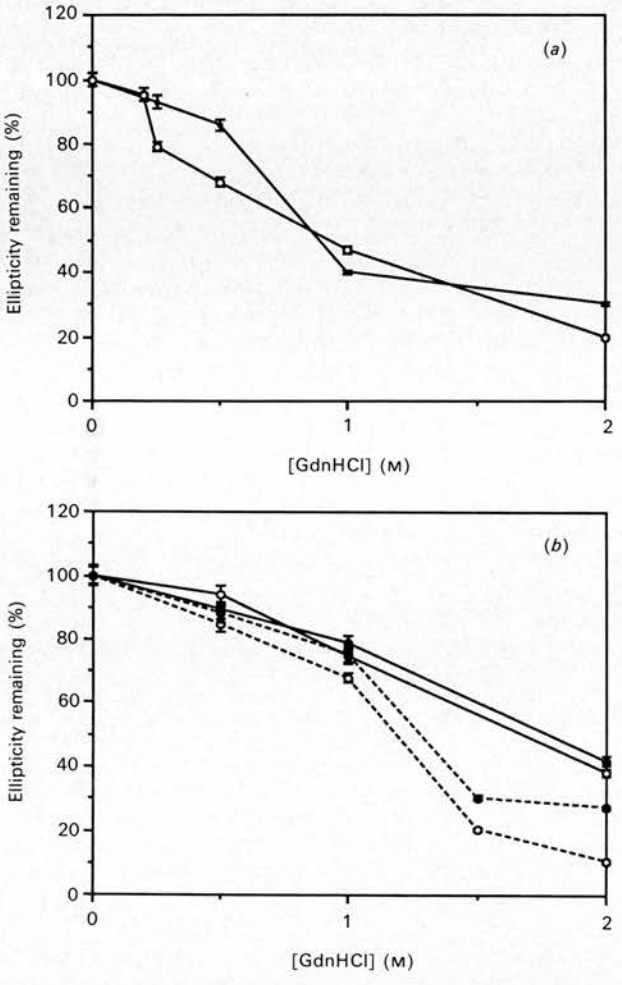


Fig. 4. Structural changes in ADH upon addition of GdnHCl as demonstrated by c.d.

(*a*) *D. melanogaster*; (*b*) horse liver (——) and yeast (----). ○, Without NAD^+; ●, with NAD^+.

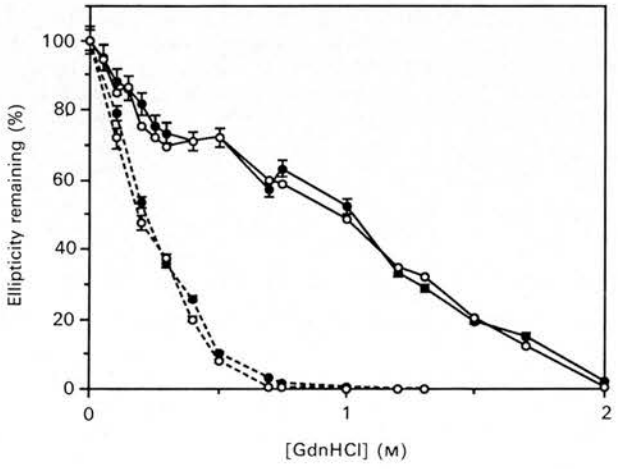


Fig. 5. Inactivation of horse liver (——) and yeast (----) ADH by GdnHCl

Details of the procedures are given in the Experimental section. ●, Enzyme in the presence of NAD^+; ○, enzyme in the absence of NAD^+.

are depicted in Fig. 4(*b*). As with the *Drosophila* enzyme, these changes occur at higher GdnHCl concentrations than do the changes in activity. There is no significant effect of NAD^+ on the structural changes, consistent with the data from the activity studies. Similar protection experiments with yeast ADH also showed that NAD^+ provides little protection to this enzyme (results not shown).

DISCUSSION

In this paper we report on the analysis of the secondary structure of the short-chain ADH from *Drosophila*. The far-u.v. c.d. spectrum provides estimates for α-helix and β-sheet which are in good agreement with results of the structure-prediction methods, and the validity of the c.d. approach has been confirmed with reference to the horse liver enzyme, where a crystal structure is available. Previously Benyajati *et al.* (1981) published an amino acid sequence (derived from cDNA sequence) for the *Drosophila* enzyme. Apart from residues 25 and 251, their sequence was identical with that published previously by Thatcher (1980), which was determined by direct protein sequencing. However, Benyajati *et al.* (1981) reported the result of a structure prediction using the Chou–Fasman method as suggesting that the enzyme contained 50% α-helix, a value considerably larger than that reported by Thatcher & Sawyer (1980) using this predictive method. The cause of this discrepancy is not clear, but it should be noted that the results of our multiple prediction studies are closer to the value obtained by Thatcher & Sawyer (1980). Benyajati *et al.* (1981) also mentioned that analysis of the c.d. spectrum gave a value of 50% α-helix, in accordance with their structure-prediction results. We are unable to explain the discrepancy between their value and our values (Table 1), because they did not provide the primary data on which the analysis was based.

At present, no crystal-structure information is available for the *Drosophila* enzyme, although preliminary results are encouraging (E. Gordon & L. Sawyer, personal communication). We are therefore at present unable to confirm the validity of the c.d. analysis. The analysis suggests that the enzyme contains significant amounts of both α-helix and β-strand structures. The presence of large amounts of these secondary structures would indicate that a typical NAD^+-binding site exists in this enzyme; this proposal would be consistent with other data, including the binding of the enzyme to Cibacron Blue (Juan & Gonzalez-Duarte, 1980).

The conformational stability ($\Delta G_{(H_2O)}$) of the *Drosophila* enzyme has been estimated by a linear extrapolation procedure as shown in Fig. 3(*b*). Although the justification for using this procedure has been the subject of considerable debate [see Pace (1986) and references therein], the value obtained [13.5 kJ/mol (3.20 kcal/mol)] can at least be used as a basis for discussion. It lies at the lower end of the range of values [21–63 kJ/mol (5–15 kcal/mol)] quoted by Pace (1990) for the stabilities of a number of globular proteins. However, it should be noted that in any particular case the exact value can be markedly affected by factors such as temperature, pH and ionic strength, and no attempt was made in the present study to optimize the stability of the enzyme.

The inclusion of NAD^+ (220 μM) serves to increase the conformational stability of the *Drosophila* enzyme by 6.61 kJ/mol (1.57 kcal/mol) (Fig. 3*b*). Although this value may seem low when the number of likely interactions between the enzyme and the substrate is considered, it is of a similar magnitude to other examples of ligand-induced stabilization, e.g. the stabilization of 2.1 kJ/mol (0.5 kcal/mol) afforded to lysozyme by inclusion of the competitive inhibitor $(N\text{-acetylglucosamine})_3$

9. Coordinates of the ADH model in Brookhaven database format

```
HEADER    PROTEIN                                        27-AUG-92
COMPND    ADH.EVOLVING
AUTHOR    Lluis Ribas de Pouplana   PhD project Edinburgh 1989-1992
SEQRES   1     255  SER PHE THR LEU THR ASN LYS ASN VAL ILE PHE VAL ALA
SEQRES   2     255  GLY LEU GLY GLY ILE GLY LEU ASP THR SER LYS GLU LEU
SEQRES   3     255  LEU LYS ARG ASP LEU LYS ASN LEU VAL ILE LEU ASP ARG
SEQRES   4     255  ILE GLU ASN PRO ALA ALA ILE ALA GLU LEU LYS ALA ILE
SEQRES   5     255  ASN PRO LYS VAL THR VAL THR PHE TYR PRO TYR ASP VAL
SEQRES   6     255  THR VAL PRO ILE ALA ILE THR THR LYS LEU LEU LYS THR
SEQRES   7     255  ILE PHE ALA GLN LEU LYS THR VAL ASP VAL LEU ILE ASN
SEQRES   8     255  GLY ALA GLY ILE LEU ASP ASP HIS GLN ILE GLU ARG THR
SEQRES   9     255  ILE ALA VAL ASN TYR THR GLY LEU VAL GLN THR THR THR
SEQRES  10     255  ALA ILE LEU ASP PHE TRP ASP LYS ARG LYS GLY GLY PRO
SEQRES  11     255  GLY GLY ILE ILE CYS ASN ILE GLY SER VAL THR GLY PHE
SEQRES  12     255  ASN ALA ILE TYR GLN VAL PRO VAL TYR SER GLY THR LYS
SEQRES  13     255  ALA ALA VAL VAL ASN PHE THR SER SER LEU ALA LYS LEU
SEQRES  14     255  ALA PRO ILE THR GLY VAL THR ALA TYR THR VAL ASN PRO
SEQRES  15     255  GLY ILE THR ARG THR THR LEU VAL HIS LYS PHE ASN SER
SEQRES  16     255  TRP LEU ASP VAL GLU PRO GLN VAL ALA GLU LYS LEU LEU
SEQRES  17     255  ALA HIS PRO THR GLN PRO SER LEU ALA CYS ALA GLU ASN
SEQRES  18     255  PHE VAL LYS ALA ILE GLU LEU ASN GLN ASN GLY ALA ILE
SEQRES  19     255  TRP LYS LEU ASP LEU GLY THR LEU GLU ALA ILE GLN TRP
SEQRES  20     255  THR LYS HIS TRP ASP SER GLY ILE
ATOM      1  N   SER     1      -6.758  -8.423 -15.621  1.00  0.00
ATOM      2  CA  SER     1      -5.902  -7.396 -15.001  1.00  0.00
ATOM      3  C   SER     1      -7.008  -6.458 -14.317  1.00  0.00
ATOM      4  O   SER     1      -8.192  -6.646 -14.575  1.00  0.00
ATOM      5  CB  SER     1      -5.135  -6.601 -16.063  1.00  0.00
ATOM      6  OG  SER     1      -5.659  -6.824 -17.399  1.00  0.00
ATOM      7  HG  SER     1      -5.141  -6.305 -18.073  1.00  0.00
ATOM      8  H   SER     1      -6.963  -8.333 -16.574  1.00  0.00
ATOM      9  N   PHE     2      -6.725  -5.387 -13.485  1.00  0.00
ATOM     10  CA  PHE     2      -7.654  -4.436 -12.799  1.00  0.00
ATOM     11  C   PHE     2      -8.544  -5.205 -11.847  1.00  0.00
ATOM     12  O   PHE     2      -9.660  -4.737 -11.472  1.00  0.00
ATOM     13  CB  PHE     2      -8.411  -3.800 -13.925  1.00  0.00
ATOM     14  CG  PHE     2      -8.669  -2.296 -13.670  1.00  0.00
ATOM     15  CD1 PHE     2      -8.478  -1.803 -12.383  1.00  0.00
ATOM     16  CD2 PHE     2      -9.063  -1.432 -14.704  1.00  0.00
ATOM     17  CE1 PHE     2      -8.679  -0.457 -12.109  1.00  0.00
ATOM     18  CE2 PHE     2      -9.236  -0.107 -14.324  1.00  0.00
ATOM     19  CZ  PHE     2      -9.065   0.392 -13.077  1.00  0.00
ATOM     20  H   PHE     2      -5.787  -5.251 -13.409  1.00  0.00
ATOM     21  N   THR     3      -8.163  -6.438 -11.458  1.00  0.00
ATOM     22  CA  THR     3      -8.929  -7.322 -10.629  1.00  0.00
ATOM     23  C   THR     3      -8.093  -7.598  -9.363  1.00  0.00
ATOM     24  O   THR     3      -7.143  -8.431  -9.405  1.00  0.00
ATOM     25  CB  THR     3      -9.134  -8.558 -11.542  1.00  0.00
ATOM     26  OG1 THR     3      -9.511  -8.121 -12.864  1.00  0.00
ATOM     27  CG2 THR     3     -10.272  -9.397 -11.021  1.00  0.00
ATOM     28  HG1 THR     3      -9.921  -7.238 -12.660  1.00  0.00
ATOM     29  H   THR     3      -7.245  -6.731 -11.591  1.00  0.00
ATOM     30  N   LEU     4      -8.312  -6.867  -8.305  1.00  0.00
ATOM     31  CA  LEU     4      -7.529  -7.161  -7.127  1.00  0.00
ATOM     32  C   LEU     4      -8.129  -8.412  -6.553  1.00  0.00
ATOM     33  O   LEU     4      -9.139  -8.423  -5.860  1.00  0.00
ATOM     34  CB  LEU     4      -7.603  -6.059  -6.034  1.00  0.00
ATOM     35  CG  LEU     4      -6.553  -6.056  -4.956  1.00  0.00
ATOM     36  CD1 LEU     4      -5.195  -5.800  -5.414  1.00  0.00
ATOM     37  CD2 LEU     4      -6.799  -4.798  -4.165  1.00  0.00
ATOM     38  H   LEU     4      -8.945  -6.137  -8.304  1.00  0.00
ATOM     39  N   THR     5      -7.365  -9.461  -6.936  1.00  0.00
ATOM     40  CA  THR     5      -7.807 -10.824  -6.545  1.00  0.00
ATOM     41  C   THR     5      -7.145 -11.032  -5.190  1.00  0.00
ATOM     42  O   THR     5      -6.370 -11.943  -4.911  1.00  0.00
ATOM     43  CB  THR     5      -7.274 -11.861  -7.536  1.00  0.00
```

```
ATOM     44  OG1 THR     5      -7.833 -11.434  -8.776  1.00  0.00
ATOM     45  CG2 THR     5      -7.690 -13.379  -7.324  1.00  0.00
ATOM     46  HG1 THR     5      -7.153 -10.754  -9.117  1.00  0.00
ATOM     47  H   THR     5      -6.611  -9.341  -7.534  1.00  0.00
ATOM     48  N   ASN     6      -7.344 -10.059  -4.283  1.00  0.00
ATOM     49  CA  ASN     6      -6.574 -10.148  -2.987  1.00  0.00
ATOM     50  C   ASN     6      -7.269  -9.542  -1.754  1.00  0.00
ATOM     51  O   ASN     6      -7.539  -8.318  -1.681  1.00  0.00
ATOM     52  CB  ASN     6      -5.168  -9.454  -3.117  1.00  0.00
ATOM     53  CG  ASN     6      -4.473  -9.545  -4.451  1.00  0.00
ATOM     54  OD1 ASN     6      -4.963  -9.147  -5.530  1.00  0.00
ATOM     55  ND2 ASN     6      -3.360 -10.311  -4.436  1.00  0.00
ATOM     56 HD22ASN     6      -2.918 -10.582  -5.291  1.00  0.00
ATOM     57 HD21ASN     6      -3.118 -10.671  -3.582  1.00  0.00
ATOM     58  H   ASN     6      -7.959  -9.298  -4.406  1.00  0.00
ATOM     59  N   LYS     7      -7.790 -10.306  -0.745  1.00  0.00
ATOM     60  CA  LYS     7      -8.366  -9.721   0.453  1.00  0.00
ATOM     61  C   LYS     7      -7.323  -8.827   1.098  1.00  0.00
ATOM     62  O   LYS     7      -6.086  -9.089   1.044  1.00  0.00
ATOM     63  CB  LYS     7      -8.757 -10.854   1.338  1.00  0.00
ATOM     64  CG  LYS     7      -9.466 -10.322   2.569  1.00  0.00
ATOM     65  CD  LYS     7      -9.405 -11.156   3.819  1.00  0.00
ATOM     66  CE  LYS     7      -9.440 -10.218   4.967  1.00  0.00
ATOM     67  NZ  LYS     7      -9.095 -10.975   6.159  1.00  0.00
ATOM     68  HZ1 LYS     7      -8.884 -10.388   6.975  1.00  0.00
ATOM     69  HZ2 LYS     7      -8.227 -11.540   5.947  1.00  0.00
ATOM     70  HZ3 LYS     7      -9.880 -11.605   6.431  1.00  0.00
ATOM     71  H   LYS     7      -7.836 -11.297  -0.827  1.00  0.00
ATOM     72  N   ASN     8      -7.757  -7.780   1.789  1.00  0.00
ATOM     73  CA  ASN     8      -6.820  -6.723   2.161  1.00  0.00
ATOM     74  C   ASN     8      -7.134  -6.191   3.554  1.00  0.00
ATOM     75  O   ASN     8      -8.266  -6.490   4.005  1.00  0.00
ATOM     76  CB  ASN     8      -6.943  -5.635   1.057  1.00  0.00
ATOM     77  CG  ASN     8      -8.324  -5.125   0.641  1.00  0.00
ATOM     78  OD1 ASN     8      -9.047  -4.487   1.386  1.00  0.00
ATOM     79  ND2 ASN     8      -8.648  -5.331  -0.626  1.00  0.00
ATOM     80 HD22ASN     8      -9.547  -5.031  -0.971  1.00  0.00
ATOM     81 HD21ASN     8      -8.025  -5.930  -1.126  1.00  0.00
ATOM     82  H   ASN     8      -8.645  -7.784   2.048  1.00  0.00
ATOM     83  N   VAL     9      -6.155  -5.565   4.184  1.00  0.00
ATOM     84  CA  VAL     9      -6.462  -4.932   5.527  1.00  0.00
ATOM     85  C   VAL     9      -6.447  -3.358   5.621  1.00  0.00
ATOM     86  O   VAL     9      -7.196  -2.836   6.491  1.00  0.00
ATOM     87  CB  VAL     9      -5.484  -5.379   6.669  1.00  0.00
ATOM     88  CG1 VAL     9      -5.870  -5.050   8.064  1.00  0.00
ATOM     89  CG2 VAL     9      -5.425  -6.817   6.609  1.00  0.00
ATOM     90  H   VAL     9      -5.275  -5.925   4.046  1.00  0.00
ATOM     91  N   ILE    10      -5.699  -2.576   4.812  1.00  0.00
ATOM     92  CA  ILE    10      -5.615  -1.124   4.915  1.00  0.00
ATOM     93  C   ILE    10      -5.423  -0.429   6.329  1.00  0.00
ATOM     94  O   ILE    10      -6.386  -0.199   7.123  1.00  0.00
ATOM     95  CB  ILE    10      -6.837  -0.463   4.233  1.00  0.00
ATOM     96  CG1 ILE    10      -7.282  -1.115   2.953  1.00  0.00
ATOM     97  CG2 ILE    10      -6.429   0.975   3.896  1.00  0.00
ATOM     98  CD1 ILE    10      -8.510  -0.404   2.258  1.00  0.00
ATOM     99  H   ILE    10      -5.317  -3.015   4.001  1.00  0.00
ATOM    100  N   PHE    11      -4.180   0.000   6.552  1.00  0.00
ATOM    101  CA  PHE    11      -3.741   0.649   7.767  1.00  0.00
ATOM    102  C   PHE    11      -4.402   2.012   8.040  1.00  0.00
ATOM    103  O   PHE    11      -5.312   2.103   8.894  1.00  0.00
ATOM    104  CB  PHE    11      -2.158   0.750   7.697  1.00  0.00
ATOM    105  CG  PHE    11      -1.417  -0.305   8.582  1.00  0.00
ATOM    106  CD1 PHE    11      -0.287   0.056   9.180  1.00  0.00
ATOM    107  CD2 PHE    11      -1.860  -1.621   8.615  1.00  0.00
ATOM    108  CE1 PHE    11       0.480  -0.934   9.830  1.00  0.00
ATOM    109  CE2 PHE    11      -1.152  -2.626   9.207  1.00  0.00
```

```
ATOM    110  CZ  PHE    11       0.040  -2.214   9.806  1.00  0.00
ATOM    111  H   PHE    11      -3.483  -0.047   5.826  1.00  0.00
ATOM    112  N   VAL    12      -4.023   3.102   7.404  1.00  0.00
ATOM    113  CA  VAL    12      -4.712   4.382   7.610  1.00  0.00
ATOM    114  C   VAL    12      -6.222   4.473   7.409  1.00  0.00
ATOM    115  O   VAL    12      -6.839   5.287   8.054  1.00  0.00
ATOM    116  CB  VAL    12      -3.965   5.455   6.730  1.00  0.00
ATOM    117  CG1 VAL    12      -4.186   5.098   5.238  1.00  0.00
ATOM    118  CG2 VAL    12      -4.511   6.853   7.008  1.00  0.00
ATOM    119  H   VAL    12      -3.199   3.041   6.913  1.00  0.00
ATOM    120  N   ALA    13      -6.740   3.560   6.649  1.00  0.00
ATOM    121  CA  ALA    13      -8.094   3.276   6.317  1.00  0.00
ATOM    122  C   ALA    13      -9.194   4.348   6.335  1.00  0.00
ATOM    123  O   ALA    13      -9.612   4.862   5.348  1.00  0.00
ATOM    124  CB  ALA    13      -8.605   2.181   7.293  1.00  0.00
ATOM    125  H   ALA    13      -6.056   3.063   6.123  1.00  0.00
ATOM    126  N   GLY    14      -9.515   4.747   7.559  1.00  0.00
ATOM    127  CA  GLY    14     -10.644   5.629   7.814  1.00  0.00
ATOM    128  C   GLY    14     -10.531   7.116   7.515  1.00  0.00
ATOM    129  O   GLY    14     -11.279   7.862   8.131  1.00  0.00
ATOM    130  H   GLY    14      -8.909   4.578   8.295  1.00  0.00
ATOM    131  N   LEU    15      -9.616   7.572   6.692  1.00  0.00
ATOM    132  CA  LEU    15      -9.559   8.988   6.352  1.00  0.00
ATOM    133  C   LEU    15      -9.903   9.049   4.893  1.00  0.00
ATOM    134  O   LEU    15      -9.231   8.402   4.074  1.00  0.00
ATOM    135  CB  LEU    15      -8.172   9.635   6.604  1.00  0.00
ATOM    136  CG  LEU    15      -8.128  11.175   6.487  1.00  0.00
ATOM    137  CD1 LEU    15      -8.619  11.835   7.792  1.00  0.00
ATOM    138  CD2 LEU    15      -6.649  11.664   6.217  1.00  0.00
ATOM    139  H   LEU    15      -8.866   6.965   6.440  1.00  0.00
ATOM    140  N   GLY    16     -10.984   9.721   4.593  1.00  0.00
ATOM    141  CA  GLY    16     -11.531   9.935   3.282  1.00  0.00
ATOM    142  C   GLY    16     -10.642  10.509   2.238  1.00  0.00
ATOM    143  O   GLY    16     -11.096  11.025   1.235  1.00  0.00
ATOM    144  H   GLY    16     -11.490  10.158   5.276  1.00  0.00
ATOM    145  N   GLY    17      -9.291  10.425   2.273  1.00  0.00
ATOM    146  CA  GLY    17      -8.416  10.986   1.266  1.00  0.00
ATOM    147  C   GLY    17      -7.812   9.861   0.477  1.00  0.00
ATOM    148  O   GLY    17      -7.470  10.040  -0.703  1.00  0.00
ATOM    149  H   GLY    17      -8.931  10.026   3.089  1.00  0.00
ATOM    150  N   ILE    18      -7.526   8.704   1.091  1.00  0.00
ATOM    151  CA  ILE    18      -6.861   7.552   0.474  1.00  0.00
ATOM    152  C   ILE    18      -7.489   6.168   0.826  1.00  0.00
ATOM    153  O   ILE    18      -8.046   5.584  -0.074  1.00  0.00
ATOM    154  CB  ILE    18      -5.328   7.324   0.835  1.00  0.00
ATOM    155  CG1 ILE    18      -4.679   8.558   1.563  1.00  0.00
ATOM    156  CG2 ILE    18      -4.612   7.170  -0.523  1.00  0.00
ATOM    157  CD1 ILE    18      -4.614   8.316   3.050  1.00  0.00
ATOM    158  H   ILE    18      -7.763   8.615   2.083  1.00  0.00
ATOM    159  N   GLY    19      -7.537   5.670   2.093  1.00  0.00
ATOM    160  CA  GLY    19      -8.018   4.350   2.443  1.00  0.00
ATOM    161  C   GLY    19      -9.377   4.020   2.057  1.00  0.00
ATOM    162  O   GLY    19      -9.655   3.145   1.164  1.00  0.00
ATOM    163  H   GLY    19      -7.154   6.250   2.770  1.00  0.00
ATOM    164  N   LEU    20     -10.313   4.665   2.707  1.00  0.00
ATOM    165  CA  LEU    20     -11.761   4.493   2.464  1.00  0.00
ATOM    166  C   LEU    20     -12.250   4.598   1.038  1.00  0.00
ATOM    167  O   LEU    20     -12.960   3.700   0.558  1.00  0.00
ATOM    168  CB  LEU    20     -12.589   5.501   3.263  1.00  0.00
ATOM    169  CG  LEU    20     -14.141   5.611   3.399  1.00  0.00
ATOM    170  CD1 LEU    20     -14.729   4.377   4.069  1.00  0.00
ATOM    171  CD2 LEU    20     -14.381   6.867   4.177  1.00  0.00
ATOM    172  H   LEU    20     -10.018   5.286   3.485  1.00  0.00
ATOM    173  N   ASP    21     -11.653   5.513   0.291  1.00  0.00
ATOM    174  CA  ASP    21     -11.782   5.694  -1.192  1.00  0.00
ATOM    175  C   ASP    21     -11.219   4.500  -2.054  1.00  0.00
```

```
ATOM    176  O   ASP    21     -11.817   4.212  -3.128  1.00  0.00
ATOM    177  CB  ASP    21     -11.119   7.052  -1.661  1.00  0.00
ATOM    178  CG  ASP    21     -12.126   8.165  -1.881  1.00  0.00
ATOM    179  OD1 ASP    21     -12.556   8.808  -0.932  1.00  0.00
ATOM    180  OD2 ASP    21     -12.523   8.311  -3.010  1.00  0.00
ATOM    181  H   ASP    21     -11.015   6.079   0.797  1.00  0.00
ATOM    182  N   THR    22     -10.059   3.928  -1.617  1.00  0.00
ATOM    183  CA  THR    22      -9.585   2.815  -2.335  1.00  0.00
ATOM    184  C   THR    22     -10.606   1.659  -2.226  1.00  0.00
ATOM    185  O   THR    22     -10.886   0.939  -3.198  1.00  0.00
ATOM    186  CB  THR    22      -8.244   2.287  -1.805  1.00  0.00
ATOM    187  OG1 THR    22      -7.341   3.418  -1.541  1.00  0.00
ATOM    188  CG2 THR    22      -7.642   1.348  -2.777  1.00  0.00
ATOM    189  HG1 THR    22      -7.661   3.843  -0.816  1.00  0.00
ATOM    190  H   THR    22      -9.720   4.101  -0.717  1.00  0.00
ATOM    191  N   SER    23     -11.182   1.395  -1.033  1.00  0.00
ATOM    192  CA  SER    23     -12.073   0.301  -0.868  1.00  0.00
ATOM    193  C   SER    23     -13.212   0.248  -1.845  1.00  0.00
ATOM    194  O   SER    23     -13.598  -0.822  -2.379  1.00  0.00
ATOM    195  CB  SER    23     -12.719   0.257   0.504  1.00  0.00
ATOM    196  OG  SER    23     -11.696   0.511   1.452  1.00  0.00
ATOM    197  HG  SER    23     -11.280  -0.263   1.813  1.00  0.00
ATOM    198  H   SER    23     -10.852   1.887  -0.258  1.00  0.00
ATOM    199  N   LYS    24     -13.822   1.451  -1.976  1.00  0.00
ATOM    200  CA  LYS    24     -14.942   1.622  -2.900  1.00  0.00
ATOM    201  C   LYS    24     -14.738   0.940  -4.163  1.00  0.00
ATOM    202  O   LYS    24     -15.664   0.211  -4.600  1.00  0.00
ATOM    203  CB  LYS    24     -15.177   3.060  -3.280  1.00  0.00
ATOM    204  CG  LYS    24     -15.638   3.961  -2.161  1.00  0.00
ATOM    205  CD  LYS    24     -15.575   5.376  -2.694  1.00  0.00
ATOM    206  CE  LYS    24     -16.116   6.384  -1.665  1.00  0.00
ATOM    207  NZ  LYS    24     -15.542   6.147  -0.396  1.00  0.00
ATOM    208  HZ1 LYS    24     -14.591   5.789  -0.589  1.00  0.00
ATOM    209  HZ2 LYS    24     -16.134   5.408   0.015  1.00  0.00
ATOM    210  HZ3 LYS    24     -15.407   7.064   0.101  1.00  0.00
ATOM    211  H   LYS    24     -13.568   2.141  -1.363  1.00  0.00
ATOM    212  N   GLU    25     -13.517   1.075  -4.784  1.00  0.00
ATOM    213  CA  GLU    25     -13.171   0.452  -6.053  1.00  0.00
ATOM    214  C   GLU    25     -13.461  -1.055  -6.012  1.00  0.00
ATOM    215  O   GLU    25     -14.094  -1.611  -6.894  1.00  0.00
ATOM    216  CB  GLU    25     -11.725   0.740  -6.387  1.00  0.00
ATOM    217  CG  GLU    25     -11.290   2.190  -6.500  1.00  0.00
ATOM    218  CD  GLU    25      -9.762   2.346  -6.748  1.00  0.00
ATOM    219  OE1 GLU    25      -8.993   2.361  -5.801  1.00  0.00
ATOM    220  OE2 GLU    25      -9.357   2.380  -7.894  1.00  0.00
ATOM    221  H   GLU    25     -12.766   1.623  -4.385  1.00  0.00
ATOM    222  N   LEU    26     -12.935  -1.667  -4.936  1.00  0.00
ATOM    223  CA  LEU    26     -12.809  -3.079  -4.708  1.00  0.00
ATOM    224  C   LEU    26     -14.241  -3.687  -4.410  1.00  0.00
ATOM    225  O   LEU    26     -14.477  -4.879  -4.667  1.00  0.00
ATOM    226  CB  LEU    26     -11.979  -3.459  -3.488  1.00  0.00
ATOM    227  CG  LEU    26     -10.673  -2.770  -3.387  1.00  0.00
ATOM    228  CD1 LEU    26     -10.177  -3.063  -2.038  1.00  0.00
ATOM    229  CD2 LEU    26      -9.698  -3.065  -4.533  1.00  0.00
ATOM    230  H   LEU    26     -12.593  -1.120  -4.210  1.00  0.00
ATOM    231  N   LEU    27     -15.203  -2.953  -3.778  1.00  0.00
ATOM    232  CA  LEU    27     -16.480  -3.537  -3.532  1.00  0.00
ATOM    233  C   LEU    27     -17.277  -3.318  -4.792  1.00  0.00
ATOM    234  O   LEU    27     -18.087  -4.162  -5.083  1.00  0.00
ATOM    235  CB  LEU    27     -17.198  -2.872  -2.416  1.00  0.00
ATOM    236  CG  LEU    27     -16.646  -3.351  -1.080  1.00  0.00
ATOM    237  CD1 LEU    27     -16.289  -2.199  -0.242  1.00  0.00
ATOM    238  CD2 LEU    27     -17.600  -4.361  -0.514  1.00  0.00
ATOM    239  H   LEU    27     -15.019  -2.008  -3.578  1.00  0.00
ATOM    240  N   LYS    28     -17.254  -2.138  -5.466  1.00  0.00
ATOM    241  CA  LYS    28     -18.039  -1.965  -6.633  1.00  0.00
```

```
ATOM    242  C    LYS    28     -17.658  -2.949  -7.793  1.00  0.00
ATOM    243  O    LYS    28     -18.513  -3.512  -8.486  1.00  0.00
ATOM    244  CB   LYS    28     -17.890  -0.545  -7.191  1.00  0.00
ATOM    245  CG   LYS    28     -18.530   0.616  -6.390  1.00  0.00
ATOM    246  CD   LYS    28     -18.094   1.938  -7.013  1.00  0.00
ATOM    247  CE   LYS    28     -18.688   3.025  -6.150  1.00  0.00
ATOM    248  NZ   LYS    28     -18.028   4.306  -6.459  1.00  0.00
ATOM    249  HZ1  LYS    28     -18.207   4.533  -7.441  1.00  0.00
ATOM    250  HZ2  LYS    28     -16.995   4.255  -6.285  1.00  0.00
ATOM    251  HZ3  LYS    28     -18.386   5.058  -5.859  1.00  0.00
ATOM    252  H    LYS    28     -16.825  -1.378  -5.066  1.00  0.00
ATOM    253  N    ARG    29     -16.364  -2.891  -8.043  1.00  0.00
ATOM    254  CA   ARG    29     -15.811  -3.604  -9.225  1.00  0.00
ATOM    255  C    ARG    29     -14.910  -4.610  -8.634  1.00  0.00
ATOM    256  O    ARG    29     -14.498  -4.486  -7.491  1.00  0.00
ATOM    257  CB   ARG    29     -15.054  -2.635 -10.155  1.00  0.00
ATOM    258  CG   ARG    29     -14.906  -3.126 -11.545  1.00  0.00
ATOM    259  CD   ARG    29     -14.593  -1.961 -12.435  1.00  0.00
ATOM    260  NE   ARG    29     -15.660  -0.961 -12.448  1.00  0.00
ATOM    261  CZ   ARG    29     -15.442   0.230 -12.978  1.00  0.00
ATOM    262  NH1  ARG    29     -16.438   1.067 -13.145  1.00  0.00
ATOM    263  NH2  ARG    29     -14.211   0.601 -13.331  1.00  0.00
ATOM    264  HE   ARG    29     -16.482  -1.158 -11.984  1.00  0.00
ATOM    265  HH12ARG    29     -16.405   2.060 -13.444  1.00  0.00
ATOM    266  HH11ARG    29     -17.324   0.644 -12.842  1.00  0.00
ATOM    267  HH22ARG    29     -14.057   1.460 -13.730  1.00  0.00
ATOM    268  HH21ARG    29     -13.359   0.039 -13.051  1.00  0.00
ATOM    269  H    ARG    29     -15.692  -2.496  -7.400  1.00  0.00
ATOM    270  N    ASP    30     -14.409  -5.604  -9.373  1.00  0.00
ATOM    271  CA   ASP    30     -13.451  -6.564  -8.882  1.00  0.00
ATOM    272  C    ASP    30     -13.766  -7.386  -7.577  1.00  0.00
ATOM    273  O    ASP    30     -12.905  -8.181  -7.170  1.00  0.00
ATOM    274  CB   ASP    30     -11.989  -5.845  -8.786  1.00  0.00
ATOM    275  CG   ASP    30     -11.574  -4.853  -7.743  1.00  0.00
ATOM    276  OD1  ASP    30     -11.647  -3.674  -7.958  1.00  0.00
ATOM    277  OD2  ASP    30     -11.279  -5.237  -6.615  1.00  0.00
ATOM    278  H    ASP    30     -14.683  -5.662 -10.280  1.00  0.00
ATOM    279  N    LEU    31     -14.978  -7.172  -7.013  1.00  0.00
ATOM    280  CA   LEU    31     -15.426  -7.619  -5.666  1.00  0.00
ATOM    281  C    LEU    31     -14.957  -9.043  -5.293  1.00  0.00
ATOM    282  O    LEU    31     -15.143 -10.030  -6.040  1.00  0.00
ATOM    283  CB   LEU    31     -16.984  -7.577  -5.536  1.00  0.00
ATOM    284  CG   LEU    31     -17.907  -8.333  -6.625  1.00  0.00
ATOM    285  CD1  LEU    31     -19.112  -9.121  -6.114  1.00  0.00
ATOM    286  CD2  LEU    31     -18.386  -7.231  -7.538  1.00  0.00
ATOM    287  H    LEU    31     -15.533  -6.491  -7.448  1.00  0.00
ATOM    288  N    LYS    32     -14.481  -9.267  -4.087  1.00  0.00
ATOM    289  CA   LYS    32     -13.902 -10.544  -3.789  1.00  0.00
ATOM    290  C    LYS    32     -14.197 -10.838  -2.326  1.00  0.00
ATOM    291  O    LYS    32     -15.195 -10.356  -1.786  1.00  0.00
ATOM    292  CB   LYS    32     -12.406 -10.512  -3.919  1.00  0.00
ATOM    293  CG   LYS    32     -11.925 -11.798  -4.633  1.00  0.00
ATOM    294  CD   LYS    32     -11.113 -11.401  -5.818  1.00  0.00
ATOM    295  CE   LYS    32     -11.584 -12.321  -6.892  1.00  0.00
ATOM    296  NZ   LYS    32     -11.064 -11.886  -8.217  1.00  0.00
ATOM    297  HZ1  LYS    32     -11.344 -10.906  -8.357  1.00  0.00
ATOM    298  HZ2  LYS    32     -11.463 -12.586  -8.878  1.00  0.00
ATOM    299  HZ3  LYS    32     -10.040 -11.882  -8.177  1.00  0.00
ATOM    300  H    LYS    32     -14.305  -8.574  -3.441  1.00  0.00
ATOM    301  N    ASN    33     -13.489 -11.731  -1.637  1.00  0.00
ATOM    302  CA   ASN    33     -13.796 -12.102  -0.233  1.00  0.00
ATOM    303  C    ASN    33     -13.497 -11.014   0.710  1.00  0.00
ATOM    304  O    ASN    33     -12.574 -10.256   0.605  1.00  0.00
ATOM    305  CB   ASN    33     -13.049 -13.373   0.156  1.00  0.00
ATOM    306  CG   ASN    33     -11.571 -13.252   0.548  1.00  0.00
ATOM    307  OD1  ASN    33     -10.731 -13.233  -0.332  1.00  0.00
```

```
ATOM    308  ND2 ASN    33     -11.138 -13.139   1.782  1.00  0.00
ATOM    309  HD22ASN    33     -10.163 -13.128   2.020  1.00  0.00
ATOM    310  HD21ASN    33     -11.782 -13.102   2.494  1.00  0.00
ATOM    311  H   ASN    33     -12.662 -12.094  -1.964  1.00  0.00
ATOM    312  N   LEU    34     -14.404 -11.048   1.657  1.00  0.00
ATOM    313  CA  LEU    34     -14.640 -10.211   2.811  1.00  0.00
ATOM    314  C   LEU    34     -13.416  -9.346   3.247  1.00  0.00
ATOM    315  O   LEU    34     -12.506  -9.752   3.973  1.00  0.00
ATOM    316  CB  LEU    34     -15.036 -11.078   3.971  1.00  0.00
ATOM    317  CG  LEU    34     -16.538 -11.469   4.242  1.00  0.00
ATOM    318  CD1 LEU    34     -16.677 -12.898   4.753  1.00  0.00
ATOM    319  CD2 LEU    34     -17.114 -10.361   5.109  1.00  0.00
ATOM    320  H   LEU    34     -15.164 -11.641   1.548  1.00  0.00
ATOM    321  N   VAL    35     -13.272  -8.134   2.707  1.00  0.00
ATOM    322  CA  VAL    35     -12.220  -7.158   2.915  1.00  0.00
ATOM    323  C   VAL    35     -12.546  -6.397   4.201  1.00  0.00
ATOM    324  O   VAL    35     -13.741  -6.244   4.531  1.00  0.00
ATOM    325  CB  VAL    35     -12.086  -6.118   1.712  1.00  0.00
ATOM    326  CG1 VAL    35     -11.852  -6.974   0.571  1.00  0.00
ATOM    327  CG2 VAL    35     -13.255  -5.275   1.371  1.00  0.00
ATOM    328  H   VAL    35     -13.942  -8.072   2.085  1.00  0.00
ATOM    329  N   ILE    36     -11.477  -5.881   4.879  1.00  0.00
ATOM    330  CA  ILE    36     -11.563  -5.047   6.079  1.00  0.00
ATOM    331  C   ILE    36     -10.696  -3.813   6.163  1.00  0.00
ATOM    332  O   ILE    36      -9.766  -3.786   5.364  1.00  0.00
ATOM    333  CB  ILE    36     -11.260  -5.846   7.409  1.00  0.00
ATOM    334  CG1 ILE    36     -10.139  -6.805   7.211  1.00  0.00
ATOM    335  CG2 ILE    36     -12.604  -6.415   7.958  1.00  0.00
ATOM    336  CD1 ILE    36      -9.959  -7.715   8.382  1.00  0.00
ATOM    337  H   ILE    36     -10.640  -5.964   4.393  1.00  0.00
ATOM    338  N   LEU    37     -10.952  -2.823   7.004  1.00  0.00
ATOM    339  CA  LEU    37     -10.158  -1.584   7.104  1.00  0.00
ATOM    340  C   LEU    37      -9.632  -1.597   8.548  1.00  0.00
ATOM    341  O   LEU    37     -10.405  -1.646   9.491  1.00  0.00
ATOM    342  CB  LEU    37     -11.049  -0.330   6.902  1.00  0.00
ATOM    343  CG  LEU    37     -11.823  -0.071   5.608  1.00  0.00
ATOM    344  CD1 LEU    37     -13.170  -0.699   5.554  1.00  0.00
ATOM    345  CD2 LEU    37     -12.180   1.417   5.612  1.00  0.00
ATOM    346  H   LEU    37     -11.668  -2.936   7.645  1.00  0.00
ATOM    347  N   ASP    38      -8.296  -1.584   8.795  1.00  0.00
ATOM    348  CA  ASP    38      -7.667  -1.667  10.116  1.00  0.00
ATOM    349  C   ASP    38      -7.474  -0.274  10.743  1.00  0.00
ATOM    350  O   ASP    38      -6.413   0.371  10.706  1.00  0.00
ATOM    351  CB  ASP    38      -6.309  -2.413  10.074  1.00  0.00
ATOM    352  CG  ASP    38      -5.605  -2.510  11.427  1.00  0.00
ATOM    353  OD1 ASP    38      -4.597  -1.867  11.572  1.00  0.00
ATOM    354  OD2 ASP    38      -6.022  -3.180  12.349  1.00  0.00
ATOM    355  H   ASP    38      -7.612  -1.429   8.043  1.00  0.00
ATOM    356  N   ARG    39      -8.552   0.201  11.352  1.00  0.00
ATOM    357  CA  ARG    39      -8.559   1.497  12.043  1.00  0.00
ATOM    358  C   ARG    39      -7.384   1.826  13.000  1.00  0.00
ATOM    359  O   ARG    39      -7.636   1.846  14.189  1.00  0.00
ATOM    360  CB  ARG    39      -9.912   1.560  12.781  1.00  0.00
ATOM    361  CG  ARG    39     -10.194   0.366  13.694  1.00  0.00
ATOM    362  CD  ARG    39     -11.258   0.861  14.649  1.00  0.00
ATOM    363  NE  ARG    39     -11.446  -0.145  15.633  1.00  0.00
ATOM    364  CZ  ARG    39     -12.028   0.133  16.804  1.00  0.00
ATOM    365  NH1 ARG    39     -12.127  -0.875  17.672  1.00  0.00
ATOM    366  NH2 ARG    39     -12.485   1.344  17.004  1.00  0.00
ATOM    367  HE  ARG    39     -11.008  -1.026  15.465  1.00  0.00
ATOM    368  HH12ARG    39     -12.752  -0.648  18.479  1.00  0.00
ATOM    369  HH11ARG    39     -11.886  -1.808  17.308  1.00  0.00
ATOM    370  HH22ARG    39     -12.971   1.657  17.835  1.00  0.00
ATOM    371  HH21ARG    39     -12.304   2.027  16.273  1.00  0.00
ATOM    372  H   ARG    39      -9.244  -0.509  11.349  1.00  0.00
ATOM    373  N   ILE    40      -6.194   2.284  12.457  1.00  0.00
```

```
ATOM    374  CA  ILE    40      -4.942   2.659  13.168  1.00  0.00
ATOM    375  C   ILE    40      -5.144   3.461  14.509  1.00  0.00
ATOM    376  O   ILE    40      -4.504   3.230  15.540  1.00  0.00
ATOM    377  CB  ILE    40      -4.100   3.460  12.138  1.00  0.00
ATOM    378  CG1 ILE    40      -2.797   3.947  12.752  1.00  0.00
ATOM    379  CG2 ILE    40      -4.967   4.678  11.733  1.00  0.00
ATOM    380  CD1 ILE    40      -1.734   4.236  11.644  1.00  0.00
ATOM    381  H   ILE    40      -6.046   2.166  11.456  1.00  0.00
ATOM    382  N   GLU    41      -6.094   4.361  14.480  1.00  0.00
ATOM    383  CA  GLU    41      -6.295   5.359  15.557  1.00  0.00
ATOM    384  C   GLU    41      -7.446   4.921  16.564  1.00  0.00
ATOM    385  O   GLU    41      -7.604   5.494  17.664  1.00  0.00
ATOM    386  CB  GLU    41      -6.597   6.716  14.914  1.00  0.00
ATOM    387  CG  GLU    41      -7.790   6.678  13.932  1.00  0.00
ATOM    388  CD  GLU    41      -8.183   7.935  13.142  1.00  0.00
ATOM    389  OE1 GLU    41      -8.649   8.848  13.795  1.00  0.00
ATOM    390  OE2 GLU    41      -8.013   7.945  11.973  1.00  0.00
ATOM    391  H   GLU    41      -6.852   4.169  13.852  1.00  0.00
ATOM    392  N   ASN    42      -8.197   3.871  16.192  1.00  0.00
ATOM    393  CA  ASN    42      -9.203   3.187  17.051  1.00  0.00
ATOM    394  C   ASN    42     -10.419   3.831  17.614  1.00  0.00
ATOM    395  O   ASN    42     -10.951   3.592  18.725  1.00  0.00
ATOM    396  CB  ASN    42      -8.457   2.527  18.284  1.00  0.00
ATOM    397  CG  ASN    42      -7.693   1.260  17.916  1.00  0.00
ATOM    398  OD1 ASN    42      -8.304   0.245  17.626  1.00  0.00
ATOM    399  ND2 ASN    42      -6.407   1.082  18.063  1.00  0.00
ATOM    400 HD22ASN    42      -5.949   0.283  17.651  1.00  0.00
ATOM    401 HD21ASN    42      -6.095   1.813  18.633  1.00  0.00
ATOM    402  H   ASN    42      -7.825   3.428  15.351  1.00  0.00
ATOM    403  N   PRO    43     -11.027   4.736  16.873  1.00  0.00
ATOM    404  CA  PRO    43     -12.235   5.383  17.238  1.00  0.00
ATOM    405  C   PRO    43     -13.429   4.755  16.443  1.00  0.00
ATOM    406  O   PRO    43     -13.845   5.212  15.322  1.00  0.00
ATOM    407  CB  PRO    43     -11.992   6.833  16.942  1.00  0.00
ATOM    408  CG  PRO    43     -10.993   6.824  15.765  1.00  0.00
ATOM    409  CD  PRO    43     -10.583   5.326  15.616  1.00  0.00
ATOM    410  N   ALA    44     -14.108   3.765  17.029  1.00  0.00
ATOM    411  CA  ALA    44     -15.133   2.937  16.393  1.00  0.00
ATOM    412  C   ALA    44     -16.176   3.540  15.456  1.00  0.00
ATOM    413  O   ALA    44     -16.705   2.800  14.667  1.00  0.00
ATOM    414  CB  ALA    44     -15.815   2.170  17.501  1.00  0.00
ATOM    415  H   ALA    44     -14.050   3.850  18.013  1.00  0.00
ATOM    416  N   ALA    45     -16.528   4.836  15.589  1.00  0.00
ATOM    417  CA  ALA    45     -17.504   5.517  14.857  1.00  0.00
ATOM    418  C   ALA    45     -17.312   5.399  13.374  1.00  0.00
ATOM    419  O   ALA    45     -18.295   5.643  12.625  1.00  0.00
ATOM    420  CB  ALA    45     -17.478   7.055  15.161  1.00  0.00
ATOM    421  H   ALA    45     -15.958   5.439  16.113  1.00  0.00
ATOM    422  N   ILE    46     -16.112   5.094  12.830  1.00  0.00
ATOM    423  CA  ILE    46     -15.850   4.868  11.451  1.00  0.00
ATOM    424  C   ILE    46     -16.821   3.808  10.910  1.00  0.00
ATOM    425  O   ILE    46     -17.099   3.742   9.724  1.00  0.00
ATOM    426  CB  ILE    46     -14.342   4.504  11.182  1.00  0.00
ATOM    427  CG1 ILE    46     -13.384   5.476  11.826  1.00  0.00
ATOM    428  CG2 ILE    46     -14.115   4.606   9.730  1.00  0.00
ATOM    429  CD1 ILE    46     -13.736   6.987  11.716  1.00  0.00
ATOM    430  H   ILE    46     -15.241   5.054  13.364  1.00  0.00
ATOM    431  N   ALA    47     -17.448   2.921  11.740  1.00  0.00
ATOM    432  CA  ALA    47     -18.439   1.943  11.300  1.00  0.00
ATOM    433  C   ALA    47     -19.438   2.287  10.171  1.00  0.00
ATOM    434  O   ALA    47     -19.424   1.560   9.173  1.00  0.00
ATOM    435  CB  ALA    47     -19.344   1.502  12.500  1.00  0.00
ATOM    436  H   ALA    47     -17.238   2.998  12.732  1.00  0.00
ATOM    437  N   GLU    48     -20.073   3.461  10.196  1.00  0.00
ATOM    438  CA  GLU    48     -21.057   3.702   9.189  1.00  0.00
ATOM    439  C   GLU    48     -20.450   4.045   7.832  1.00  0.00
```

```
ATOM    440  O   GLU    48     -21.074   3.759   6.806  1.00  0.00
ATOM    441  CB  GLU    48     -21.954   4.735   9.683  1.00  0.00
ATOM    442  CG  GLU    48     -23.048   4.193  10.596  1.00  0.00
ATOM    443  CD  GLU    48     -23.101   5.025  11.917  1.00  0.00
ATOM    444  OE1 GLU    48     -24.129   5.610  12.275  1.00  0.00
ATOM    445  OE2 GLU    48     -22.065   5.065  12.668  1.00  0.00
ATOM    446  H   GLU    48     -19.839   4.072  10.935  1.00  0.00
ATOM    447  N   LEU    49     -19.250   4.630   7.879  1.00  0.00
ATOM    448  CA  LEU    49     -18.558   4.954   6.667  1.00  0.00
ATOM    449  C   LEU    49     -18.317   3.680   5.871  1.00  0.00
ATOM    450  O   LEU    49     -18.626   3.514   4.688  1.00  0.00
ATOM    451  CB  LEU    49     -17.286   5.662   7.128  1.00  0.00
ATOM    452  CG  LEU    49     -17.512   6.865   7.983  1.00  0.00
ATOM    453  CD1 LEU    49     -16.161   7.378   8.354  1.00  0.00
ATOM    454  CD2 LEU    49     -18.322   7.975   7.375  1.00  0.00
ATOM    455  H   LEU    49     -18.749   4.630   8.647  1.00  0.00
ATOM    456  N   LYS    50     -17.808   2.688   6.566  1.00  0.00
ATOM    457  CA  LYS    50     -17.634   1.324   5.994  1.00  0.00
ATOM    458  C   LYS    50     -18.906   0.770   5.431  1.00  0.00
ATOM    459  O   LYS    50     -18.980   0.434   4.260  1.00  0.00
ATOM    460  CB  LYS    50     -17.047   0.392   7.017  1.00  0.00
ATOM    461  CG  LYS    50     -15.699   0.939   7.605  1.00  0.00
ATOM    462  CD  LYS    50     -15.230   0.189   8.882  1.00  0.00
ATOM    463  CE  LYS    50     -13.989   0.762   9.555  1.00  0.00
ATOM    464  NZ  LYS    50     -13.587   0.132  10.761  1.00  0.00
ATOM    465  HZ1 LYS    50     -13.169  -0.823  10.647  1.00  0.00
ATOM    466  HZ2 LYS    50     -14.403   0.068  11.408  1.00  0.00
ATOM    467  HZ3 LYS    50     -12.865   0.685  11.330  1.00  0.00
ATOM    468  H   LYS    50     -17.432   2.797   7.445  1.00  0.00
ATOM    469  N   ALA    51     -19.971   0.785   6.254  1.00  0.00
ATOM    470  CA  ALA    51     -21.284   0.223   5.890  1.00  0.00
ATOM    471  C   ALA    51     -21.998   0.716   4.601  1.00  0.00
ATOM    472  O   ALA    51     -22.785  -0.036   4.004  1.00  0.00
ATOM    473  CB  ALA    51     -22.335   0.386   7.001  1.00  0.00
ATOM    474  H   ALA    51     -19.882   1.292   7.123  1.00  0.00
ATOM    475  N   ILE    52     -21.774   1.973   4.146  1.00  0.00
ATOM    476  CA  ILE    52     -22.328   2.439   2.904  1.00  0.00
ATOM    477  C   ILE    52     -21.871   1.549   1.761  1.00  0.00
ATOM    478  O   ILE    52     -22.708   1.266   0.923  1.00  0.00
ATOM    479  CB  ILE    52     -21.903   3.851   2.653  1.00  0.00
ATOM    480  CG1 ILE    52     -22.427   4.689   3.798  1.00  0.00
ATOM    481  CG2 ILE    52     -22.512   4.349   1.352  1.00  0.00
ATOM    482  CD1 ILE    52     -21.761   6.077   3.968  1.00  0.00
ATOM    483  H   ILE    52     -21.354   2.525   4.732  1.00  0.00
ATOM    484  N   ASN    53     -20.612   1.144   1.679  1.00  0.00
ATOM    485  CA  ASN    53     -20.016   0.317   0.578  1.00  0.00
ATOM    486  C   ASN    53     -20.584  -1.115   0.479  1.00  0.00
ATOM    487  O   ASN    53     -20.934  -1.520  -0.622  1.00  0.00
ATOM    488  CB  ASN    53     -18.509   0.263   0.812  1.00  0.00
ATOM    489  CG  ASN    53     -17.854   1.572   0.703  1.00  0.00
ATOM    490  OD1 ASN    53     -17.362   1.956  -0.366  1.00  0.00
ATOM    491  ND2 ASN    53     -17.660   2.289   1.774  1.00  0.00
ATOM    492 HD22ASN    53     -17.195   3.188   1.727  1.00  0.00
ATOM    493 HD21ASN    53     -17.918   2.012   2.682  1.00  0.00
ATOM    494  H   ASN    53     -20.059   1.328   2.479  1.00  0.00
ATOM    495  N   PRO    54     -20.865  -1.860   1.578  1.00  0.00
ATOM    496  CA  PRO    54     -21.842  -2.875   1.616  1.00  0.00
ATOM    497  C   PRO    54     -23.208  -2.561   1.052  1.00  0.00
ATOM    498  O   PRO    54     -23.631  -3.296   0.137  1.00  0.00
ATOM    499  CB  PRO    54     -21.835  -3.278   3.075  1.00  0.00
ATOM    500  CG  PRO    54     -20.381  -3.260   3.390  1.00  0.00
ATOM    501  CD  PRO    54     -20.069  -1.938   2.774  1.00  0.00
ATOM    502  N   LYS    55     -23.725  -1.451   1.545  1.00  0.00
ATOM    503  CA  LYS    55     -25.127  -1.063   1.221  1.00  0.00
ATOM    504  C   LYS    55     -25.533  -0.826  -0.262  1.00  0.00
ATOM    505  O   LYS    55     -26.562  -1.309  -0.641  1.00  0.00
```

```
ATOM    506  CB  LYS    55     -25.703   0.263   1.879  1.00  0.00
ATOM    507  CG  LYS    55     -26.499  -0.159   3.073  1.00  0.00
ATOM    508  CD  LYS    55     -26.630   1.076   3.974  1.00  0.00
ATOM    509  CE  LYS    55     -27.433   0.884   5.294  1.00  0.00
ATOM    510  NZ  LYS    55     -27.429   1.932   6.265  1.00  0.00
ATOM    511  HZ1 LYS    55     -27.493   2.872   5.781  1.00  0.00
ATOM    512  HZ2 LYS    55     -26.493   2.016   6.697  1.00  0.00
ATOM    513  HZ3 LYS    55     -28.170   1.869   7.057  1.00  0.00
ATOM    514  H   LYS    55     -23.141  -0.851   1.983  1.00  0.00
ATOM    515  N   VAL    56     -24.658  -0.157  -1.036  1.00  0.00
ATOM    516  CA  VAL    56     -24.850   0.075  -2.444  1.00  0.00
ATOM    517  C   VAL    56     -24.197  -1.032  -3.343  1.00  0.00
ATOM    518  O   VAL    56     -23.993  -0.711  -4.489  1.00  0.00
ATOM    519  CB  VAL    56     -24.317   1.473  -2.933  1.00  0.00
ATOM    520  CG1 VAL    56     -25.225   2.539  -2.497  1.00  0.00
ATOM    521  CG2 VAL    56     -22.870   1.684  -2.447  1.00  0.00
ATOM    522  H   VAL    56     -23.881   0.214  -0.584  1.00  0.00
ATOM    523  N   THR    57     -24.030  -2.316  -2.895  1.00  0.00
ATOM    524  CA  THR    57     -23.366  -3.358  -3.712  1.00  0.00
ATOM    525  C   THR    57     -24.046  -4.665  -3.325  1.00  0.00
ATOM    526  O   THR    57     -25.261  -4.863  -3.557  1.00  0.00
ATOM    527  CB  THR    57     -21.772  -3.502  -3.490  1.00  0.00
ATOM    528  OG1 THR    57     -21.209  -2.204  -3.408  1.00  0.00
ATOM    529  CG2 THR    57     -21.108  -4.447  -4.591  1.00  0.00
ATOM    530  HG1 THR    57     -21.446  -1.867  -2.533  1.00  0.00
ATOM    531  H   THR    57     -24.380  -2.520  -2.046  1.00  0.00
ATOM    532  N   VAL    58     -23.473  -5.587  -2.600  1.00  0.00
ATOM    533  CA  VAL    58     -24.124  -6.805  -2.292  1.00  0.00
ATOM    534  C   VAL    58     -24.224  -6.744  -0.789  1.00  0.00
ATOM    535  O   VAL    58     -23.171  -6.852  -0.125  1.00  0.00
ATOM    536  CB  VAL    58     -23.307  -8.036  -2.878  1.00  0.00
ATOM    537  CG1 VAL    58     -23.716  -8.130  -4.306  1.00  0.00
ATOM    538  CG2 VAL    58     -21.766  -7.878  -2.886  1.00  0.00
ATOM    539  H   VAL    58     -22.599  -5.380  -2.176  1.00  0.00
ATOM    540  N   THR    59     -25.410  -6.602  -0.229  1.00  0.00
ATOM    541  CA  THR    59     -25.563  -6.378   1.219  1.00  0.00
ATOM    542  C   THR    59     -25.407  -7.653   2.049  1.00  0.00
ATOM    543  O   THR    59     -26.314  -8.144   2.689  1.00  0.00
ATOM    544  CB  THR    59     -26.951  -5.660   1.359  1.00  0.00
ATOM    545  OG1 THR    59     -27.558  -5.477   0.020  1.00  0.00
ATOM    546  CG2 THR    59     -26.781  -4.325   2.071  1.00  0.00
ATOM    547  HG1 THR    59     -27.168  -4.663  -0.351  1.00  0.00
ATOM    548  H   THR    59     -26.236  -6.613  -0.659  1.00  0.00
ATOM    549  N   PHE    60     -24.221  -8.278   2.025  1.00  0.00
ATOM    550  CA  PHE    60     -24.077  -9.527   2.747  1.00  0.00
ATOM    551  C   PHE    60     -22.810  -9.431   3.611  1.00  0.00
ATOM    552  O   PHE    60     -22.490 -10.369   4.272  1.00  0.00
ATOM    553  CB  PHE    60     -23.950 -10.651   1.819  1.00  0.00
ATOM    554  CG  PHE    60     -25.259 -11.058   1.117  1.00  0.00
ATOM    555  CD1 PHE    60     -25.496 -10.779  -0.251  1.00  0.00
ATOM    556  CD2 PHE    60     -26.206 -11.687   1.880  1.00  0.00
ATOM    557  CE1 PHE    60     -26.780 -11.114  -0.782  1.00  0.00
ATOM    558  CE2 PHE    60     -27.444 -12.018   1.282  1.00  0.00
ATOM    559  CZ  PHE    60     -27.757 -11.750  -0.007  1.00  0.00
ATOM    560  H   PHE    60     -23.409  -7.853   1.641  1.00  0.00
ATOM    561  N   TYR    61     -22.177  -8.272   3.841  1.00  0.00
ATOM    562  CA  TYR    61     -20.918  -8.055   4.564  1.00  0.00
ATOM    563  C   TYR    61     -21.184  -7.683   6.012  1.00  0.00
ATOM    564  O   TYR    61     -21.864  -6.711   6.263  1.00  0.00
ATOM    565  CB  TYR    61     -20.086  -6.991   3.957  1.00  0.00
ATOM    566  CG  TYR    61     -19.557  -7.312   2.573  1.00  0.00
ATOM    567  CD1 TYR    61     -20.392  -7.220   1.497  1.00  0.00
ATOM    568  CD2 TYR    61     -18.244  -7.731   2.461  1.00  0.00
ATOM    569  CE1 TYR    61     -19.882  -7.503   0.247  1.00  0.00
ATOM    570  CE2 TYR    61     -17.729  -8.028   1.226  1.00  0.00
ATOM    571  CZ  TYR    61     -18.569  -7.892   0.152  1.00  0.00
```

```
ATOM    572  OH  TYR    61     -18.130  -8.149  -1.069  1.00  0.00
ATOM    573  HH  TYR    61     -18.951  -8.220  -1.698  1.00  0.00
ATOM    574  H   TYR    61     -22.728  -7.489   3.723  1.00  0.00
ATOM    575  N   PRO    62     -20.872  -8.501   7.044  1.00  0.00
ATOM    576  CA  PRO    62     -20.531  -8.075   8.375  1.00  0.00
ATOM    577  C   PRO    62     -19.040  -7.848   8.520  1.00  0.00
ATOM    578  O   PRO    62     -18.284  -8.689   8.069  1.00  0.00
ATOM    579  CB  PRO    62     -21.120  -9.167   9.311  1.00  0.00
ATOM    580  CG  PRO    62     -21.971 -10.108   8.391  1.00  0.00
ATOM    581  CD  PRO    62     -21.283  -9.887   7.087  1.00  0.00
ATOM    582  N   TYR    63     -18.607  -6.763   9.167  1.00  0.00
ATOM    583  CA  TYR    63     -17.206  -6.557   9.371  1.00  0.00
ATOM    584  C   TYR    63     -16.943  -6.599  10.929  1.00  0.00
ATOM    585  O   TYR    63     -17.786  -6.278  11.783  1.00  0.00
ATOM    586  CB  TYR    63     -16.771  -5.198   8.852  1.00  0.00
ATOM    587  CG  TYR    63     -16.812  -4.876   7.275  1.00  0.00
ATOM    588  CD1 TYR    63     -16.556  -3.553   6.912  1.00  0.00
ATOM    589  CD2 TYR    63     -17.027  -5.853   6.370  1.00  0.00
ATOM    590  CE1 TYR    63     -16.502  -3.269   5.554  1.00  0.00
ATOM    591  CE2 TYR    63     -16.973  -5.613   5.030  1.00  0.00
ATOM    592  CZ  TYR    63     -16.712  -4.293   4.653  1.00  0.00
ATOM    593  OH  TYR    63     -16.746  -4.049   3.301  1.00  0.00
ATOM    594  HH  TYR    63     -16.360  -4.798   2.877  1.00  0.00
ATOM    595  H   TYR    63     -19.285  -6.109   9.440  1.00  0.00
ATOM    596  N   ASP    64     -15.687  -7.010  11.216  1.00  0.00
ATOM    597  CA  ASP    64     -15.250  -7.037  12.566  1.00  0.00
ATOM    598  C   ASP    64     -13.764  -7.011  12.564  1.00  0.00
ATOM    599  O   ASP    64     -13.091  -7.685  11.759  1.00  0.00
ATOM    600  CB  ASP    64     -15.674  -8.296  13.336  1.00  0.00
ATOM    601  CG  ASP    64     -14.948  -8.617  14.664  1.00  0.00
ATOM    602  OD1 ASP    64     -14.695  -9.810  14.965  1.00  0.00
ATOM    603  OD2 ASP    64     -14.787  -7.754  15.538  1.00  0.00
ATOM    604  H   ASP    64     -15.082  -7.287  10.500  1.00  0.00
ATOM    605  N   VAL    65     -13.096  -6.066  13.283  1.00  0.00
ATOM    606  CA  VAL    65     -11.704  -5.866  13.509  1.00  0.00
ATOM    607  C   VAL    65     -11.369  -4.902  14.632  1.00  0.00
ATOM    608  O   VAL    65     -12.108  -3.959  14.851  1.00  0.00
ATOM    609  CB  VAL    65     -10.979  -5.325  12.288  1.00  0.00
ATOM    610  CG1 VAL    65     -10.265  -6.536  11.759  1.00  0.00
ATOM    611  CG2 VAL    65     -11.875  -4.503  11.315  1.00  0.00
ATOM    612  H   VAL    65     -13.696  -5.487  13.848  1.00  0.00
ATOM    613  N   THR    66     -10.262  -5.277  15.304  1.00  0.00
ATOM    614  CA  THR    66      -9.554  -4.580  16.360  1.00  0.00
ATOM    615  C   THR    66     -10.190  -4.386  17.719  1.00  0.00
ATOM    616  O   THR    66     -11.383  -4.335  17.937  1.00  0.00
ATOM    617  CB  THR    66      -9.099  -3.189  15.806  1.00  0.00
ATOM    618  OG1 THR    66      -8.682  -3.305  14.436  1.00  0.00
ATOM    619  CG2 THR    66      -8.057  -2.659  16.692  1.00  0.00
ATOM    620  HG1 THR    66      -8.783  -2.509  13.937  1.00  0.00
ATOM    621  H   THR    66      -9.809  -6.083  15.004  1.00  0.00
ATOM    622  N   VAL    67      -9.291  -4.536  18.733  1.00  0.00
ATOM    623  CA  VAL    67      -9.604  -4.362  20.179  1.00  0.00
ATOM    624  C   VAL    67      -8.844  -3.015  20.444  1.00  0.00
ATOM    625  O   VAL    67      -7.721  -2.965  20.069  1.00  0.00
ATOM    626  CB  VAL    67      -8.958  -5.495  21.054  1.00  0.00
ATOM    627  CG1 VAL    67      -8.847  -5.327  22.546  1.00  0.00
ATOM    628  CG2 VAL    67      -9.962  -6.582  21.059  1.00  0.00
ATOM    629  H   VAL    67      -8.349  -4.656  18.536  1.00  0.00
ATOM    630  N   PRO    68      -9.470  -1.966  20.915  1.00  0.00
ATOM    631  CA  PRO    68      -8.845  -0.592  21.032  1.00  0.00
ATOM    632  C   PRO    68      -7.437  -0.731  21.568  1.00  0.00
ATOM    633  O   PRO    68      -6.597  -0.017  21.013  1.00  0.00
ATOM    634  CB  PRO    68      -9.836   0.136  21.929  1.00  0.00
ATOM    635  CG  PRO    68     -10.539  -1.029  22.636  1.00  0.00
ATOM    636  CD  PRO    68     -10.845  -1.982  21.498  1.00  0.00
ATOM    637  N   ILE    69      -6.999  -1.498  22.613  1.00  0.00
```

```
ATOM    638  CA  ILE    69      -5.591  -1.657  22.953  1.00  0.00
ATOM    639  C   ILE    69      -4.951  -2.630  21.984  1.00  0.00
ATOM    640  O   ILE    69      -4.535  -3.787  22.226  1.00  0.00
ATOM    641  CB  ILE    69      -5.373  -2.100  24.508  1.00  0.00
ATOM    642  CG1 ILE    69      -5.961  -3.518  24.840  1.00  0.00
ATOM    643  CG2 ILE    69      -6.186  -1.213  25.467  1.00  0.00
ATOM    644  CD1 ILE    69      -5.058  -4.701  25.094  1.00  0.00
ATOM    645  H   ILE    69      -7.653  -1.795  23.222  1.00  0.00
ATOM    646  N   ALA    70      -4.729  -2.186  20.757  1.00  0.00
ATOM    647  CA  ALA    70      -4.273  -2.932  19.628  1.00  0.00
ATOM    648  C   ALA    70      -2.892  -3.516  19.707  1.00  0.00
ATOM    649  O   ALA    70      -1.953  -2.808  19.363  1.00  0.00
ATOM    650  CB  ALA    70      -4.450  -2.059  18.322  1.00  0.00
ATOM    651  H   ALA    70      -4.875  -1.245  20.620  1.00  0.00
ATOM    652  N   ILE    71      -2.839  -4.780  20.145  1.00  0.00
ATOM    653  CA  ILE    71      -1.568  -5.518  20.170  1.00  0.00
ATOM    654  C   ILE    71      -1.772  -6.533  19.100  1.00  0.00
ATOM    655  O   ILE    71      -2.786  -7.194  18.990  1.00  0.00
ATOM    656  CB  ILE    71      -1.489  -6.097  21.600  1.00  0.00
ATOM    657  CG1 ILE    71      -1.066  -4.988  22.452  1.00  0.00
ATOM    658  CG2 ILE    71      -0.527  -7.275  21.763  1.00  0.00
ATOM    659  CD1 ILE    71      -1.371  -5.401  23.938  1.00  0.00
ATOM    660  H   ILE    71      -3.684  -5.272  20.358  1.00  0.00
ATOM    661  N   THR    72      -0.736  -6.644  18.295  1.00  0.00
ATOM    662  CA  THR    72      -0.590  -7.675  17.210  1.00  0.00
ATOM    663  C   THR    72      -1.488  -8.959  17.236  1.00  0.00
ATOM    664  O   THR    72      -2.382  -9.081  16.388  1.00  0.00
ATOM    665  CB  THR    72       0.860  -8.144  17.137  1.00  0.00
ATOM    666  OG1 THR    72       1.712  -6.987  17.061  1.00  0.00
ATOM    667  CG2 THR    72       1.056  -9.132  16.051  1.00  0.00
ATOM    668  HG1 THR    72       2.169  -6.960  16.208  1.00  0.00
ATOM    669  H   THR    72      -0.102  -5.961  18.254  1.00  0.00
ATOM    670  N   THR    73      -1.273  -9.763  18.267  1.00  0.00
ATOM    671  CA  THR    73      -2.107 -10.925  18.554  1.00  0.00
ATOM    672  C   THR    73      -3.607 -10.821  18.377  1.00  0.00
ATOM    673  O   THR    73      -4.126 -11.493  17.615  1.00  0.00
ATOM    674  CB  THR    73      -1.800 -11.338  20.011  1.00  0.00
ATOM    675  OG1 THR    73      -0.388 -11.143  20.314  1.00  0.00
ATOM    676  CG2 THR    73      -2.207 -12.808  20.198  1.00  0.00
ATOM    677  HG1 THR    73      -0.067 -11.816  20.921  1.00  0.00
ATOM    678  H   THR    73      -0.560  -9.453  18.929  1.00  0.00
ATOM    679  N   LYS    74      -4.295  -9.784  18.909  1.00  0.00
ATOM    680  CA  LYS    74      -5.756  -9.541  18.750  1.00  0.00
ATOM    681  C   LYS    74      -6.182  -9.255  17.304  1.00  0.00
ATOM    682  O   LYS    74      -7.273  -9.573  16.855  1.00  0.00
ATOM    683  CB  LYS    74      -6.215  -8.322  19.579  1.00  0.00
ATOM    684  CG  LYS    74      -5.857  -8.482  21.105  1.00  0.00
ATOM    685  CD  LYS    74      -6.944  -9.428  21.655  1.00  0.00
ATOM    686  CE  LYS    74      -6.656 -10.143  22.970  1.00  0.00
ATOM    687  NZ  LYS    74      -6.368  -9.074  23.932  1.00  0.00
ATOM    688  HZ1 LYS    74      -7.183  -8.374  24.114  1.00  0.00
ATOM    689  HZ2 LYS    74      -5.478  -8.510  23.588  1.00  0.00
ATOM    690  HZ3 LYS    74      -6.074  -9.656  24.740  1.00  0.00
ATOM    691  H   LYS    74      -3.630  -9.220  19.385  1.00  0.00
ATOM    692  N   LEU    75      -5.241  -8.767  16.604  1.00  0.00
ATOM    693  CA  LEU    75      -5.497  -8.312  15.261  1.00  0.00
ATOM    694  C   LEU    75      -5.361  -9.455  14.322  1.00  0.00
ATOM    695  O   LEU    75      -6.235  -9.792  13.496  1.00  0.00
ATOM    696  CB  LEU    75      -4.505  -7.222  14.947  1.00  0.00
ATOM    697  CG  LEU    75      -4.604  -5.804  15.637  1.00  0.00
ATOM    698  CD1 LEU    75      -3.357  -4.907  15.217  1.00  0.00
ATOM    699  CD2 LEU    75      -5.896  -5.081  15.219  1.00  0.00
ATOM    700  H   LEU    75      -4.408  -8.531  17.059  1.00  0.00
ATOM    701  N   LEU    76      -4.321 -10.223  14.592  1.00  0.00
ATOM    702  CA  LEU    76      -3.868 -11.408  13.781  1.00  0.00
ATOM    703  C   LEU    76      -5.016 -12.411  13.798  1.00  0.00
```

```
ATOM    704  O   LEU    76      -5.388 -13.042  12.784  1.00  0.00
ATOM    705  CB  LEU    76      -2.665 -12.067  14.403  1.00  0.00
ATOM    706  CG  LEU    76      -1.254 -11.561  14.477  1.00  0.00
ATOM    707  CD1 LEU    76      -0.323 -12.722  14.796  1.00  0.00
ATOM    708  CD2 LEU    76      -0.884 -10.936  13.123  1.00  0.00
ATOM    709  H   LEU    76      -3.804  -9.914  15.319  1.00  0.00
ATOM    710  N   LYS    77      -5.542 -12.570  15.003  1.00  0.00
ATOM    711  CA  LYS    77      -6.663 -13.516  15.223  1.00  0.00
ATOM    712  C   LYS    77      -8.000 -13.233  14.470  1.00  0.00
ATOM    713  O   LYS    77      -8.605 -14.200  13.976  1.00  0.00
ATOM    714  CB  LYS    77      -7.041 -13.466  16.702  1.00  0.00
ATOM    715  CG  LYS    77      -5.999 -13.901  17.634  1.00  0.00
ATOM    716  CD  LYS    77      -5.457 -15.328  17.625  1.00  0.00
ATOM    717  CE  LYS    77      -5.734 -15.875  18.938  1.00  0.00
ATOM    718  NZ  LYS    77      -7.175 -15.974  19.171  1.00  0.00
ATOM    719  HZ1 LYS    77      -7.533 -16.668  18.485  1.00  0.00
ATOM    720  HZ2 LYS    77      -7.673 -15.094  19.053  1.00  0.00
ATOM    721  HZ3 LYS    77      -7.308 -16.355  20.176  1.00  0.00
ATOM    722  H   LYS    77      -5.180 -12.045  15.741  1.00  0.00
ATOM    723  N   THR    78      -8.562 -12.046  14.329  1.00  0.00
ATOM    724  CA  THR    78      -9.896 -11.816  13.772  1.00  0.00
ATOM    725  C   THR    78      -9.842 -12.152  12.266  1.00  0.00
ATOM    726  O   THR    78     -10.546 -12.881  11.604  1.00  0.00
ATOM    727  CB  THR    78     -10.298 -10.330  13.974  1.00  0.00
ATOM    728  OG1 THR    78      -9.114  -9.586  13.606  1.00  0.00
ATOM    729  CG2 THR    78     -10.647  -9.978  15.381  1.00  0.00
ATOM    730  HG1 THR    78      -8.663  -9.329  14.420  1.00  0.00
ATOM    731  H   THR    78      -8.067 -11.214  14.430  1.00  0.00
ATOM    732  N   ILE    79      -8.843 -11.434  11.702  1.00  0.00
ATOM    733  CA  ILE    79      -8.534 -11.408  10.278  1.00  0.00
ATOM    734  C   ILE    79      -8.635 -12.833   9.730  1.00  0.00
ATOM    735  O   ILE    79      -9.614 -13.065   9.014  1.00  0.00
ATOM    736  CB  ILE    79      -7.151 -10.787  10.117  1.00  0.00
ATOM    737  CG1 ILE    79      -6.898  -9.428  10.603  1.00  0.00
ATOM    738  CG2 ILE    79      -7.210 -10.529   8.579  1.00  0.00
ATOM    739  CD1 ILE    79      -5.419  -8.986  10.502  1.00  0.00
ATOM    740  H   ILE    79      -8.230 -10.944  12.313  1.00  0.00
ATOM    741  N   PHE    80      -7.714 -13.713  10.002  1.00  0.00
ATOM    742  CA  PHE    80      -7.608 -15.036   9.501  1.00  0.00
ATOM    743  C   PHE    80      -8.713 -15.919  10.007  1.00  0.00
ATOM    744  O   PHE    80      -9.455 -16.460   9.165  1.00  0.00
ATOM    745  CB  PHE    80      -6.273 -15.771   9.879  1.00  0.00
ATOM    746  CG  PHE    80      -5.325 -16.141   8.713  1.00  0.00
ATOM    747  CD1 PHE    80      -5.142 -17.463   8.295  1.00  0.00
ATOM    748  CD2 PHE    80      -4.674 -15.098   8.110  1.00  0.00
ATOM    749  CE1 PHE    80      -4.318 -17.788   7.274  1.00  0.00
ATOM    750  CE2 PHE    80      -3.839 -15.384   7.071  1.00  0.00
ATOM    751  CZ  PHE    80      -3.660 -16.732   6.672  1.00  0.00
ATOM    752  H   PHE    80      -7.064 -13.386  10.613  1.00  0.00
ATOM    753  N   ALA    81      -8.943 -16.114  11.308  1.00  0.00
ATOM    754  CA  ALA    81      -9.891 -17.129  11.839  1.00  0.00
ATOM    755  C   ALA    81     -11.371 -16.962  11.372  1.00  0.00
ATOM    756  O   ALA    81     -12.200 -17.874  11.551  1.00  0.00
ATOM    757  CB  ALA    81     -10.031 -17.145  13.441  1.00  0.00
ATOM    758  H   ALA    81      -8.373 -15.510  11.979  1.00  0.00
ATOM    759  N   GLN    82     -11.702 -15.800  10.756  1.00  0.00
ATOM    760  CA  GLN    82     -13.107 -15.606  10.383  1.00  0.00
ATOM    761  C   GLN    82     -13.365 -15.600   8.971  1.00  0.00
ATOM    762  O   GLN    82     -14.391 -15.360   8.412  1.00  0.00
ATOM    763  CB  GLN    82     -13.772 -14.938  11.519  1.00  0.00
ATOM    764  CG  GLN    82     -15.255 -15.029  11.699  1.00  0.00
ATOM    765  CD  GLN    82     -15.821 -16.459  11.996  1.00  0.00
ATOM    766  OE1 GLN    82     -16.961 -16.535  12.442  1.00  0.00
ATOM    767  NE2 GLN    82     -15.120 -17.624  11.901  1.00  0.00
ATOM    768  HE22GLN    82     -15.556 -18.503  12.097  1.00  0.00
ATOM    769  HE21GLN    82     -14.163 -17.575  11.653  1.00  0.00
```

```
ATOM    770  H   GLN    82     -11.259 -14.811  11.006  1.00  0.00
ATOM    771  N   LEU    83     -12.532 -16.549   8.364  1.00  0.00
ATOM    772  CA  LEU    83     -12.693 -16.960   6.986  1.00  0.00
ATOM    773  C   LEU    83     -12.398 -15.920   5.948  1.00  0.00
ATOM    774  O   LEU    83     -12.810 -16.063   4.775  1.00  0.00
ATOM    775  N   LYS    84     -11.717 -14.877   6.348  1.00  0.00
ATOM    776  CA  LYS    84     -11.265 -13.838   5.431  1.00  0.00
ATOM    777  C   LYS    84      -9.806 -14.208   5.081  1.00  0.00
ATOM    778  O   LYS    84      -9.083 -14.556   6.009  1.00  0.00
ATOM    779  N   THR    85      -9.452 -14.070   3.817  1.00  0.00
ATOM    780  CA  THR    85      -8.043 -14.332   3.466  1.00  0.00
ATOM    781  C   THR    85      -7.505 -13.029   2.876  1.00  0.00
ATOM    782  O   THR    85      -8.173 -12.519   1.974  1.00  0.00
ATOM    783  N   VAL    86      -6.334 -12.609   3.357  1.00  0.00
ATOM    784  CA  VAL    86      -5.823 -11.339   2.763  1.00  0.00
ATOM    785  C   VAL    86      -4.631 -11.646   1.899  1.00  0.00
ATOM    786  O   VAL    86      -3.855 -12.569   2.215  1.00  0.00
ATOM    787  N   ASP    87      -4.366 -11.062   1.099  1.00  0.00
ATOM    788  CA  ASP    87      -3.328 -11.309   0.104  1.00  0.00
ATOM    789  C   ASP    87      -2.488 -10.151  -0.058  1.00  0.00
ATOM    790  O   ASP    87      -2.242  -9.592  -0.496  1.00  0.00
ATOM    791  CB  ASP    87      -4.177 -12.027  -0.931  1.00  0.00
ATOM    792  CG  ASP    87      -3.335 -12.906  -1.905  1.00  0.00
ATOM    793  OD1 ASP    87      -2.565 -13.772  -1.437  1.00  0.00
ATOM    794  OD2 ASP    87      -3.484 -12.810  -3.119  1.00  0.00
ATOM    795  H   ASP    87      -4.768 -10.322   1.408  1.00  0.00
ATOM    796  N   VAL    88      -3.065  -8.993   0.233  1.00  0.00
ATOM    797  CA  VAL    88      -2.361  -7.787  -0.005  1.00  0.00
ATOM    798  C   VAL    88      -2.461  -6.971   1.278  1.00  0.00
ATOM    799  O   VAL    88      -3.564  -6.920   1.917  1.00  0.00
ATOM    800  CB  VAL    88      -2.934  -7.004  -1.306  1.00  0.00
ATOM    801  CG1 VAL    88      -4.343  -6.499  -1.042  1.00  0.00
ATOM    802  CG2 VAL    88      -2.085  -5.751  -1.586  1.00  0.00
ATOM    803  H   VAL    88      -3.878  -8.852   0.653  1.00  0.00
ATOM    804  N   LEU    89      -1.421  -6.364   1.741  1.00  0.00
ATOM    805  CA  LEU    89      -1.399  -5.575   3.028  1.00  0.00
ATOM    806  C   LEU    89      -1.164  -4.150   2.506  1.00  0.00
ATOM    807  O   LEU    89      -0.239  -3.978   1.641  1.00  0.00
ATOM    808  CB  LEU    89      -0.154  -6.012   3.848  1.00  0.00
ATOM    809  CG  LEU    89      -0.313  -6.360   5.356  1.00  0.00
ATOM    810  CD1 LEU    89       1.001  -6.822   5.887  1.00  0.00
ATOM    811  CD2 LEU    89      -0.725  -5.136   6.155  1.00  0.00
ATOM    812  H   LEU    89      -0.642  -6.441   1.243  1.00  0.00
ATOM    813  N   ILE    90      -1.929  -3.188   2.893  1.00  0.00
ATOM    814  CA  ILE    90      -1.628  -1.829   2.446  1.00  0.00
ATOM    815  C   ILE    90      -1.027  -1.252   3.642  1.00  0.00
ATOM    816  O   ILE    90      -1.770  -0.754   4.519  1.00  0.00
ATOM    817  CB  ILE    90      -2.841  -0.856   2.140  1.00  0.00
ATOM    818  CG1 ILE    90      -3.394  -1.331   0.848  1.00  0.00
ATOM    819  CG2 ILE    90      -2.476   0.629   1.889  1.00  0.00
ATOM    820  CD1 ILE    90      -4.189  -2.720   0.766  1.00  0.00
ATOM    821  H   ILE    90      -2.653  -3.322   3.470  1.00  0.00
ATOM    822  N   ASN    91       0.339  -1.269   3.731  1.00  0.00
ATOM    823  CA  ASN    91       1.010  -0.546   4.820  1.00  0.00
ATOM    824  C   ASN    91       0.959   0.921   4.360  1.00  0.00
ATOM    825  O   ASN    91       1.780   1.430   3.579  1.00  0.00
ATOM    826  CB  ASN    91       2.544  -0.919   5.017  1.00  0.00
ATOM    827  CG  ASN    91       2.703  -2.332   5.523  1.00  0.00
ATOM    828  OD1 ASN    91       2.081  -2.748   6.504  1.00  0.00
ATOM    829  ND2 ASN    91       3.527  -3.123   4.904  1.00  0.00
ATOM    830  HD22ASN    91       3.630  -3.979   5.228  1.00  0.00
ATOM    831  HD21ASN    91       4.025  -2.758   4.149  1.00  0.00
ATOM    832  H   ASN    91       0.865  -1.781   3.053  1.00  0.00
ATOM    833  N   GLY    92      -0.068   1.573   4.793  1.00  0.00
ATOM    834  CA  GLY    92      -0.294   2.985   4.555  1.00  0.00
ATOM    835  C   GLY    92      -0.233   3.649   5.870  1.00  0.00
```

```
ATOM    836  O   GLY    92      -1.245   3.614   6.528  1.00  0.00
ATOM    837  H   GLY    92      -0.648   0.983   5.282  1.00  0.00
ATOM    838  N   ALA    93       0.913   4.066   6.442  1.00  0.00
ATOM    839  CA  ALA    93       0.909   4.751   7.681  1.00  0.00
ATOM    840  C   ALA    93       1.455   6.101   7.405  1.00  0.00
ATOM    841  O   ALA    93       2.112   6.346   6.407  1.00  0.00
ATOM    842  CB  ALA    93       1.855   4.176   8.762  1.00  0.00
ATOM    843  H   ALA    93       1.693   4.061   5.903  1.00  0.00
ATOM    844  N   GLY    94       0.971   7.102   8.140  1.00  0.00
ATOM    845  CA  GLY    94       1.411   8.456   8.051  1.00  0.00
ATOM    846  C   GLY    94       0.230   9.408   8.477  1.00  0.00
ATOM    847  O   GLY    94      -0.936   8.995   8.561  1.00  0.00
ATOM    848  H   GLY    94       0.307   6.853   8.835  1.00  0.00
ATOM    849  N   ILE    95       0.653  10.644   8.716  1.00  0.00
ATOM    850  CA  ILE    95      -0.247  11.715   9.266  1.00  0.00
ATOM    851  C   ILE    95       0.267  12.877   8.483  1.00  0.00
ATOM    852  O   ILE    95       0.661  12.727   7.291  1.00  0.00
ATOM    853  CB  ILE    95      -0.013  11.966  10.762  1.00  0.00
ATOM    854  CG1 ILE    95       0.227  10.721  11.575  1.00  0.00
ATOM    855  CG2 ILE    95      -1.227  12.662  11.215  1.00  0.00
ATOM    856  CD1 ILE    95       1.032  11.024  12.846  1.00  0.00
ATOM    857  H   ILE    95       1.528  10.935   8.416  1.00  0.00
ATOM    858  N   LEU    96       0.421  14.125   8.949  1.00  0.00
ATOM    859  CA  LEU    96       0.746  15.146   8.018  1.00  0.00
ATOM    860  C   LEU    96       2.017  15.928   8.224  1.00  0.00
ATOM    861  O   LEU    96       2.361  16.192   9.355  1.00  0.00
ATOM    862  CB  LEU    96      -0.393  16.091   7.995  1.00  0.00
ATOM    863  CG  LEU    96      -1.604  15.467   7.381  1.00  0.00
ATOM    864  CD1 LEU    96      -2.828  16.136   7.982  1.00  0.00
ATOM    865  CD2 LEU    96      -1.430  15.418   5.851  1.00  0.00
ATOM    866  H   LEU    96       0.283  14.328   9.910  1.00  0.00
ATOM    867  N   ASP    97       2.622  16.249   7.058  1.00  0.00
ATOM    868  CA  ASP    97       3.805  17.083   7.072  1.00  0.00
ATOM    869  C   ASP    97       3.511  18.545   7.468  1.00  0.00
ATOM    870  O   ASP    97       3.831  19.532   6.827  1.00  0.00
ATOM    871  CB  ASP    97       4.398  16.975   5.648  1.00  0.00
ATOM    872  CG  ASP    97       5.834  16.509   5.515  1.00  0.00
ATOM    873  OD1 ASP    97       6.117  15.323   5.525  1.00  0.00
ATOM    874  OD2 ASP    97       6.653  17.366   5.354  1.00  0.00
ATOM    875  H   ASP    97       2.275  15.990   6.230  1.00  0.00
ATOM    876  N   ASP    98       2.856  18.786   8.609  1.00  0.00
ATOM    877  CA  ASP    98       2.345  20.070   9.033  1.00  0.00
ATOM    878  C   ASP    98       2.416  19.987  10.531  1.00  0.00
ATOM    879  O   ASP    98       1.480  19.608  11.330  1.00  0.00
ATOM    880  CB  ASP    98       0.947  20.241   8.567  1.00  0.00
ATOM    881  CG  ASP    98       0.646  21.326   7.541  1.00  0.00
ATOM    882  OD1 ASP    98       0.690  20.988   6.379  1.00  0.00
ATOM    883  OD2 ASP    98       0.350  22.433   7.949  1.00  0.00
ATOM    884  H   ASP    98       2.875  18.083   9.250  1.00  0.00
ATOM    885  N   HIS    99       3.653  20.255  10.940  1.00  0.00
ATOM    886  CA  HIS    99       4.024  20.353  12.321  1.00  0.00
ATOM    887  C   HIS    99       5.378  20.922  12.312  1.00  0.00
ATOM    888  O   HIS    99       6.092  20.694  11.387  1.00  0.00
ATOM    889  CB  HIS    99       4.094  18.987  13.006  1.00  0.00
ATOM    890  CG  HIS    99       2.898  18.138  13.362  1.00  0.00
ATOM    891  ND1 HIS    99       1.595  18.434  13.444  1.00  0.00
ATOM    892  CD2 HIS    99       3.014  16.803  13.655  1.00  0.00
ATOM    893  CE1 HIS    99       0.920  17.381  13.761  1.00  0.00
ATOM    894  NE2 HIS    99       1.777  16.421  13.865  1.00  0.00
ATOM    895  HE2 HIS    99       1.631  15.580  14.300  1.00  0.00
ATOM    896  HD1 HIS    99       1.203  19.278  13.217  1.00  0.00
ATOM    897  H   HIS    99       4.420  20.397  10.365  1.00  0.00
ATOM    898  N   GLN   100       5.722  21.684  13.310  1.00  0.00
ATOM    899  CA  GLN   100       7.111  22.160  13.195  1.00  0.00
ATOM    900  C   GLN   100       8.143  21.044  13.427  1.00  0.00
ATOM    901  O   GLN   100       8.100  20.435  14.491  1.00  0.00
```

```
ATOM    902  CB  GLN   100       7.293  23.326  14.133  1.00  0.00
ATOM    903  CG  GLN   100       6.898  23.104  15.570  1.00  0.00
ATOM    904  CD  GLN   100       6.681  24.322  16.435  1.00  0.00
ATOM    905  OE1 GLN   100       7.083  25.423  16.043  1.00  0.00
ATOM    906  NE2 GLN   100       6.057  24.231  17.600  1.00  0.00
ATOM    907  HE22GLN   100       6.219  25.048  18.147  1.00  0.00
ATOM    908  HE21GLN   100       5.600  23.426  17.990  1.00  0.00
ATOM    909  H   GLN   100       5.211  21.872  14.083  1.00  0.00
ATOM    910  N   ILE   101       8.936  20.716  12.378  1.00  0.00
ATOM    911  CA  ILE   101      10.093  19.855  12.313  1.00  0.00
ATOM    912  C   ILE   101      10.231  18.922  13.538  1.00  0.00
ATOM    913  O   ILE   101      10.167  17.715  13.310  1.00  0.00
ATOM    914  CB  ILE   101      11.372  20.651  12.233  1.00  0.00
ATOM    915  CG1 ILE   101      11.349  22.079  11.542  1.00  0.00
ATOM    916  CG2 ILE   101      12.272  19.729  11.436  1.00  0.00
ATOM    917  CD1 ILE   101      11.861  23.169  12.563  1.00  0.00
ATOM    918  H   ILE   101       8.595  21.045  11.536  1.00  0.00
ATOM    919  N   GLU   102      10.520  19.390  14.738  1.00  0.00
ATOM    920  CA  GLU   102      10.604  18.555  15.966  1.00  0.00
ATOM    921  C   GLU   102       9.360  17.743  16.268  1.00  0.00
ATOM    922  O   GLU   102       9.532  16.551  16.427  1.00  0.00
ATOM    923  CB  GLU   102      11.022  19.441  17.169  1.00  0.00
ATOM    924  CG  GLU   102      12.454  19.882  16.916  1.00  0.00
ATOM    925  CD  GLU   102      12.499  21.405  16.840  1.00  0.00
ATOM    926  OE1 GLU   102      12.601  22.112  17.866  1.00  0.00
ATOM    927  OE2 GLU   102      12.475  21.924  15.751  1.00  0.00
ATOM    928  H   GLU   102      10.798  20.343  14.778  1.00  0.00
ATOM    929  N   ARG   103       8.197  18.381  16.243  1.00  0.00
ATOM    930  CA  ARG   103       6.980  17.633  16.383  1.00  0.00
ATOM    931  C   ARG   103       6.952  16.434  15.400  1.00  0.00
ATOM    932  O   ARG   103       6.772  15.253  15.767  1.00  0.00
ATOM    933  CB  ARG   103       5.813  18.586  16.178  1.00  0.00
ATOM    934  CG  ARG   103       5.600  19.330  17.530  1.00  0.00
ATOM    935  CD  ARG   103       4.307  20.148  17.390  1.00  0.00
ATOM    936  NE  ARG   103       4.443  21.193  16.375  1.00  0.00
ATOM    937  CZ  ARG   103       3.415  21.993  16.157  1.00  0.00
ATOM    938  NH1 ARG   103       3.560  23.026  15.319  1.00  0.00
ATOM    939  NH2 ARG   103       2.242  21.918  16.784  1.00  0.00
ATOM    940  HE  ARG   103       5.266  21.280  15.878  1.00  0.00
ATOM    941  HH12ARG   103       2.833  23.727  15.257  1.00  0.00
ATOM    942  HH11ARG   103       4.362  23.058  14.757  1.00  0.00
ATOM    943  HH22ARG   103       1.655  22.696  16.553  1.00  0.00
ATOM    944  HH21ARG   103       2.121  21.214  17.553  1.00  0.00
ATOM    945  H   ARG   103       8.156  19.355  16.108  1.00  0.00
ATOM    946  N   THR   104       7.176  16.743  14.116  1.00  0.00
ATOM    947  CA  THR   104       7.068  15.772  12.992  1.00  0.00
ATOM    948  C   THR   104       8.088  14.601  13.213  1.00  0.00
ATOM    949  O   THR   104       7.749  13.420  13.082  1.00  0.00
ATOM    950  CB  THR   104       7.392  16.523  11.655  1.00  0.00
ATOM    951  OG1 THR   104       6.840  17.843  11.772  1.00  0.00
ATOM    952  CG2 THR   104       6.722  15.964  10.457  1.00  0.00
ATOM    953  HG1 THR   104       7.521  18.489  11.503  1.00  0.00
ATOM    954  H   THR   104       7.627  17.609  13.948  1.00  0.00
ATOM    955  N   ILE   105       9.343  14.914  13.569  1.00  0.00
ATOM    956  CA  ILE   105      10.402  13.936  13.865  1.00  0.00
ATOM    957  C   ILE   105      10.067  13.061  15.089  1.00  0.00
ATOM    958  O   ILE   105      10.219  11.863  15.076  1.00  0.00
ATOM    959  CB  ILE   105      11.727  14.707  14.116  1.00  0.00
ATOM    960  CG1 ILE   105      12.183  15.409  12.877  1.00  0.00
ATOM    961  CG2 ILE   105      12.830  13.743  14.531  1.00  0.00
ATOM    962  CD1 ILE   105      13.375  16.414  12.886  1.00  0.00
ATOM    963  H   ILE   105       9.530  15.860  13.589  1.00  0.00
ATOM    964  N   ALA   106       9.608  13.727  16.141  1.00  0.00
ATOM    965  CA  ALA   106       9.295  12.971  17.346  1.00  0.00
ATOM    966  C   ALA   106       8.177  11.944  17.128  1.00  0.00
ATOM    967  O   ALA   106       8.393  10.800  17.512  1.00  0.00
```

```
ATOM    968  CB  ALA   106       8.896  13.916  18.479  1.00  0.00
ATOM    969  H   ALA   106       9.658  14.685  16.138  1.00  0.00
ATOM    970  N   VAL   107       7.059  12.322  16.537  1.00  0.00
ATOM    971  CA  VAL   107       6.018  11.308  16.373  1.00  0.00
ATOM    972  C   VAL   107       6.301  10.474  15.147  1.00  0.00
ATOM    973  O   VAL   107       5.418   9.743  14.700  1.00  0.00
ATOM    974  CB  VAL   107       4.636  11.944  16.284  1.00  0.00
ATOM    975  CG1 VAL   107       4.255  12.383  17.680  1.00  0.00
ATOM    976  CG2 VAL   107       4.628  13.055  15.223  1.00  0.00
ATOM    977  H   VAL   107       6.990  13.257  16.247  1.00  0.00
ATOM    978  N   ASN   108       7.542  10.432  14.621  1.00  0.00
ATOM    979  CA  ASN   108       7.812   9.526  13.449  1.00  0.00
ATOM    980  C   ASN   108       7.661   8.048  13.777  1.00  0.00
ATOM    981  O   ASN   108       7.648   7.247  12.847  1.00  0.00
ATOM    982  CB  ASN   108       9.164   9.740  12.883  1.00  0.00
ATOM    983  CG  ASN   108       9.182   9.658  11.358  1.00  0.00
ATOM    984  OD1 ASN   108       8.811   8.668  10.797  1.00  0.00
ATOM    985  ND2 ASN   108       9.664  10.643  10.672  1.00  0.00
ATOM    986  HD22ASN   108       9.493  10.635   9.697  1.00  0.00
ATOM    987  HD21ASN   108      10.060  11.420  11.111  1.00  0.00
ATOM    988  H   ASN   108       8.333  10.900  14.981  1.00  0.00
ATOM    989  N   TYR   109       7.423   7.636  15.072  1.00  0.00
ATOM    990  CA  TYR   109       7.198   6.200  15.524  1.00  0.00
ATOM    991  C   TYR   109       6.052   5.621  14.711  1.00  0.00
ATOM    992  O   TYR   109       6.224   4.511  14.147  1.00  0.00
ATOM    993  CB  TYR   109       6.830   6.035  17.034  1.00  0.00
ATOM    994  CG  TYR   109       7.761   6.750  18.004  1.00  0.00
ATOM    995  CD1 TYR   109       9.156   6.712  17.989  1.00  0.00
ATOM    996  CD2 TYR   109       7.119   7.462  18.951  1.00  0.00
ATOM    997  CE1 TYR   109       9.796   7.484  18.950  1.00  0.00
ATOM    998  CE2 TYR   109       7.732   8.211  19.871  1.00  0.00
ATOM    999  CZ  TYR   109       9.087   8.229  19.855  1.00  0.00
ATOM   1000  OH  TYR   109       9.791   9.061  20.692  1.00  0.00
ATOM   1001  HH  TYR   109      10.723   9.062  20.525  1.00  0.00
ATOM   1002  H   TYR   109       7.292   8.398  15.685  1.00  0.00
ATOM   1003  N   THR   110       5.045   6.461  14.462  1.00  0.00
ATOM   1004  CA  THR   110       3.759   6.103  13.782  1.00  0.00
ATOM   1005  C   THR   110       4.064   5.641  12.351  1.00  0.00
ATOM   1006  O   THR   110       3.357   4.787  11.771  1.00  0.00
ATOM   1007  CB  THR   110       2.880   7.350  13.828  1.00  0.00
ATOM   1008  OG1 THR   110       3.218   8.001  15.139  1.00  0.00
ATOM   1009  CG2 THR   110       1.470   7.031  13.785  1.00  0.00
ATOM   1010  HG1 THR   110       3.428   8.858  14.909  1.00  0.00
ATOM   1011  H   THR   110       5.022   7.339  14.974  1.00  0.00
ATOM   1012  N   GLY   111       5.148   6.126  11.798  1.00  0.00
ATOM   1013  CA  GLY   111       5.449   5.635  10.489  1.00  0.00
ATOM   1014  C   GLY   111       6.570   4.605  10.392  1.00  0.00
ATOM   1015  O   GLY   111       6.433   3.438   9.978  1.00  0.00
ATOM   1016  H   GLY   111       5.857   6.641  12.320  1.00  0.00
ATOM   1017  N   LEU   112       7.712   5.144  10.787  1.00  0.00
ATOM   1018  CA  LEU   112       8.992   4.387  10.875  1.00  0.00
ATOM   1019  C   LEU   112       9.017   3.005  11.473  1.00  0.00
ATOM   1020  O   LEU   112       8.955   1.978  10.776  1.00  0.00
ATOM   1021  CB  LEU   112      10.140   5.105  11.694  1.00  0.00
ATOM   1022  CG  LEU   112      10.904   6.325  11.139  1.00  0.00
ATOM   1023  CD1 LEU   112      11.784   6.844  12.215  1.00  0.00
ATOM   1024  CD2 LEU   112      11.627   6.024   9.804  1.00  0.00
ATOM   1025  H   LEU   112       7.761   6.146  10.991  1.00  0.00
ATOM   1026  N   VAL   113       9.057   2.880  12.787  1.00  0.00
ATOM   1027  CA  VAL   113       9.187   1.613  13.429  1.00  0.00
ATOM   1028  C   VAL   113       7.945   0.740  13.313  1.00  0.00
ATOM   1029  O   VAL   113       8.165  -0.417  13.120  1.00  0.00
ATOM   1030  CB  VAL   113       9.591   1.805  14.887  1.00  0.00
ATOM   1031  CG1 VAL   113      11.004   2.314  14.947  1.00  0.00
ATOM   1032  CG2 VAL   113       8.609   2.704  15.605  1.00  0.00
ATOM   1033  H   VAL   113       8.880   3.652  13.320  1.00  0.00
```

```
ATOM   1034  N    GLN   114       6.734   1.303  13.197  1.00  0.00
ATOM   1035  CA   GLN   114       5.507   0.525  13.107  1.00  0.00
ATOM   1036  C    GLN   114       5.448  -0.475  11.896  1.00  0.00
ATOM   1037  O    GLN   114       5.320  -1.664  12.033  1.00  0.00
ATOM   1038  CB   GLN   114       4.332   1.455  13.072  1.00  0.00
ATOM   1039  CG   GLN   114       3.011   0.886  13.622  1.00  0.00
ATOM   1040  CD   GLN   114       1.848   1.070  12.718  1.00  0.00
ATOM   1041  OE1  GLN   114       0.958   0.244  12.631  1.00  0.00
ATOM   1042  NE2  GLN   114       1.778   2.082  11.816  1.00  0.00
ATOM   1043  HE22GLN   114       1.209   1.969  11.110  1.00  0.00
ATOM   1044  HE21GLN   114       2.396   2.825  11.883  1.00  0.00
ATOM   1045  H    GLN   114       6.692   2.308  13.276  1.00  0.00
ATOM   1046  N    THR   115       5.655   0.060  10.720  1.00  0.00
ATOM   1047  CA   THR   115       5.587  -0.613   9.396  1.00  0.00
ATOM   1048  C    THR   115       6.707  -1.705   9.316  1.00  0.00
ATOM   1049  O    THR   115       6.444  -2.790   8.826  1.00  0.00
ATOM   1050  CB   THR   115       5.777   0.470   8.380  1.00  0.00
ATOM   1051  OG1  THR   115       4.685   1.414   8.606  1.00  0.00
ATOM   1052  CG2  THR   115       5.779  -0.092   6.900  1.00  0.00
ATOM   1053  HG1  THR   115       5.112   2.241   8.897  1.00  0.00
ATOM   1054  H    THR   115       5.782   1.060  10.663  1.00  0.00
ATOM   1055  N    THR   116       7.883  -1.459   9.794  1.00  0.00
ATOM   1056  CA   THR   116       8.962  -2.392   9.797  1.00  0.00
ATOM   1057  C    THR   116       8.612  -3.593  10.669  1.00  0.00
ATOM   1058  O    THR   116       8.901  -4.776  10.383  1.00  0.00
ATOM   1059  CB   THR   116      10.169  -1.642  10.338  1.00  0.00
ATOM   1060  OG1  THR   116      10.111  -0.344   9.745  1.00  0.00
ATOM   1061  CG2  THR   116      11.466  -2.373  10.141  1.00  0.00
ATOM   1062  HG1  THR   116       9.820   0.352  10.342  1.00  0.00
ATOM   1063  H    THR   116       8.097  -0.512  10.057  1.00  0.00
ATOM   1064  N    THR   117       7.924  -3.254  11.816  1.00  0.00
ATOM   1065  CA   THR   117       7.482  -4.232  12.777  1.00  0.00
ATOM   1066  C    THR   117       6.400  -5.083  12.050  1.00  0.00
ATOM   1067  O    THR   117       6.413  -6.323  12.212  1.00  0.00
ATOM   1068  CB   THR   117       6.992  -3.496  13.983  1.00  0.00
ATOM   1069  OG1  THR   117       8.145  -3.017  14.617  1.00  0.00
ATOM   1070  CG2  THR   117       6.179  -4.307  14.938  1.00  0.00
ATOM   1071  HG1  THR   117       8.493  -2.298  14.048  1.00  0.00
ATOM   1072  H    THR   117       7.559  -2.357  11.950  1.00  0.00
ATOM   1073  N    ALA   118       5.486  -4.556  11.229  1.00  0.00
ATOM   1074  CA   ALA   118       4.410  -5.367  10.692  1.00  0.00
ATOM   1075  C    ALA   118       4.897  -6.406   9.676  1.00  0.00
ATOM   1076  O    ALA   118       4.355  -7.521   9.504  1.00  0.00
ATOM   1077  CB   ALA   118       3.415  -4.480  10.002  1.00  0.00
ATOM   1078  H    ALA   118       5.645  -3.677  10.930  1.00  0.00
ATOM   1079  N    ILE   119       6.106  -6.258   9.054  1.00  0.00
ATOM   1080  CA   ILE   119       6.617  -7.178   7.993  1.00  0.00
ATOM   1081  C    ILE   119       6.921  -8.565   8.547  1.00  0.00
ATOM   1082  O    ILE   119       6.458  -9.582   8.003  1.00  0.00
ATOM   1083  CB   ILE   119       7.930  -6.635   7.339  1.00  0.00
ATOM   1084  CG1  ILE   119       7.742  -5.229   6.853  1.00  0.00
ATOM   1085  CG2  ILE   119       8.290  -7.509   6.074  1.00  0.00
ATOM   1086  CD1  ILE   119       6.488  -4.845   5.985  1.00  0.00
ATOM   1087  H    ILE   119       6.659  -5.447   9.226  1.00  0.00
ATOM   1088  N    LEU   120       7.385  -8.616   9.806  1.00  0.00
ATOM   1089  CA   LEU   120       7.672  -9.863  10.438  1.00  0.00
ATOM   1090  C    LEU   120       6.355 -10.450  10.911  1.00  0.00
ATOM   1091  O    LEU   120       6.080 -11.617  10.655  1.00  0.00
ATOM   1092  CB   LEU   120       8.704  -9.665  11.619  1.00  0.00
ATOM   1093  CG   LEU   120       9.920  -8.872  11.184  1.00  0.00
ATOM   1094  CD1  LEU   120       9.832  -7.435  11.850  1.00  0.00
ATOM   1095  CD2  LEU   120      11.158  -9.648  11.536  1.00  0.00
ATOM   1096  H    LEU   120       7.299  -7.766  10.297  1.00  0.00
ATOM   1097  N    ASP   121       5.480  -9.667  11.578  1.00  0.00
ATOM   1098  CA   ASP   121       4.352 -10.221  12.260  1.00  0.00
ATOM   1099  C    ASP   121       3.333 -10.903  11.404  1.00  0.00
```

```
ATOM   1100  O   ASP   121       2.579 -11.799  11.801  1.00  0.00
ATOM   1101  CB  ASP   121       3.751  -9.054  13.146  1.00  0.00
ATOM   1102  CG  ASP   121       4.662  -8.546  14.225  1.00  0.00
ATOM   1103  OD1 ASP   121       5.335  -9.279  14.918  1.00  0.00
ATOM   1104  OD2 ASP   121       4.738  -7.333  14.458  1.00  0.00
ATOM   1105  H   ASP   121       5.709  -8.663  11.527  1.00  0.00
ATOM   1106  N   PHE   122       3.382 -10.510  10.142  1.00  0.00
ATOM   1107  CA  PHE   122       2.544 -11.059   9.115  1.00  0.00
ATOM   1108  C   PHE   122       3.228 -12.209   8.306  1.00  0.00
ATOM   1109  O   PHE   122       2.582 -13.203   8.019  1.00  0.00
ATOM   1110  CB  PHE   122       2.133  -9.981   8.140  1.00  0.00
ATOM   1111  CG  PHE   122       0.963  -9.248   8.710  1.00  0.00
ATOM   1112  CD1 PHE   122      -0.292  -9.803   8.672  1.00  0.00
ATOM   1113  CD2 PHE   122       1.047  -8.029   9.272  1.00  0.00
ATOM   1114  CE1 PHE   122      -1.501  -9.208   9.123  1.00  0.00
ATOM   1115  CE2 PHE   122      -0.037  -7.343   9.766  1.00  0.00
ATOM   1116  CZ  PHE   122      -1.328  -7.915   9.688  1.00  0.00
ATOM   1117  H   PHE   122       4.044  -9.846   9.896  1.00  0.00
ATOM   1118  N   TRP   123       4.508 -11.973   7.967  1.00  0.00
ATOM   1119  CA  TRP   123       5.327 -12.913   7.183  1.00  0.00
ATOM   1120  C   TRP   123       5.051 -14.379   7.058  1.00  0.00
ATOM   1121  O   TRP   123       4.837 -14.863   5.934  1.00  0.00
ATOM   1122  CB  TRP   123       6.817 -12.738   7.636  1.00  0.00
ATOM   1123  CG  TRP   123       7.950 -13.411   6.864  1.00  0.00
ATOM   1124  CD1 TRP   123       8.402 -14.661   7.175  1.00  0.00
ATOM   1125  CD2 TRP   123       8.750 -12.885   5.865  1.00  0.00
ATOM   1126  NE1 TRP   123       9.425 -14.931   6.427  1.00  0.00
ATOM   1127  CE2 TRP   123       9.674 -13.917   5.598  1.00  0.00
ATOM   1128  CE3 TRP   123       8.778 -11.726   5.127  1.00  0.00
ATOM   1129  CZ2 TRP   123      10.630 -13.807   4.630  1.00  0.00
ATOM   1130  CZ3 TRP   123       9.718 -11.593   4.145  1.00  0.00
ATOM   1131  CH2 TRP   123      10.665 -12.620   3.879  1.00  0.00
ATOM   1132  HE1 TRP   123       9.923 -15.764   6.504  1.00  0.00
ATOM   1133  H   TRP   123       4.936 -11.066   8.332  1.00  0.00
ATOM   1134  N   ASP   124       4.940 -15.046   8.213  1.00  0.00
ATOM   1135  CA  ASP   124       4.766 -16.467   8.286  1.00  0.00
ATOM   1136  C   ASP   124       3.454 -16.890   7.637  1.00  0.00
ATOM   1137  O   ASP   124       3.430 -17.864   6.850  1.00  0.00
ATOM   1138  CB  ASP   124       4.771 -16.986   9.725  1.00  0.00
ATOM   1139  CG  ASP   124       3.725 -16.396  10.705  1.00  0.00
ATOM   1140  OD1 ASP   124       3.393 -15.250  10.667  1.00  0.00
ATOM   1141  OD2 ASP   124       3.121 -17.104  11.532  1.00  0.00
ATOM   1142  H   ASP   124       5.036 -14.529   9.047  1.00  0.00
ATOM   1143  N   LYS   125       2.386 -16.144   7.767  1.00  0.00
ATOM   1144  CA  LYS   125       1.065 -16.372   7.181  1.00  0.00
ATOM   1145  C   LYS   125       1.000 -16.152   5.690  1.00  0.00
ATOM   1146  O   LYS   125       0.279 -16.836   4.957  1.00  0.00
ATOM   1147  CB  LYS   125       0.082 -15.427   7.839  1.00  0.00
ATOM   1148  CG  LYS   125      -0.138 -15.868   9.328  1.00  0.00
ATOM   1149  CD  LYS   125      -0.914 -15.026  10.272  1.00  0.00
ATOM   1150  CE  LYS   125      -0.359 -15.470  11.646  1.00  0.00
ATOM   1151  NZ  LYS   125      -1.467 -15.470  12.534  1.00  0.00
ATOM   1152  HZ1 LYS   125      -2.179 -16.195  12.113  1.00  0.00
ATOM   1153  HZ2 LYS   125      -1.898 -14.524  12.656  1.00  0.00
ATOM   1154  HZ3 LYS   125      -1.091 -15.744  13.471  1.00  0.00
ATOM   1155  H   LYS   125       2.469 -15.254   8.204  1.00  0.00
ATOM   1156  N   ARG   126       1.845 -15.219   5.299  1.00  0.00
ATOM   1157  CA  ARG   126       1.873 -14.898   3.857  1.00  0.00
ATOM   1158  C   ARG   126       2.586 -15.953   3.049  1.00  0.00
ATOM   1159  O   ARG   126       2.261 -16.078   1.861  1.00  0.00
ATOM   1160  CB  ARG   126       2.563 -13.536   3.708  1.00  0.00
ATOM   1161  CG  ARG   126       2.361 -12.443   4.726  1.00  0.00
ATOM   1162  CD  ARG   126       0.944 -11.862   4.655  1.00  0.00
ATOM   1163  NE  ARG   126       0.842 -10.751   3.751  1.00  0.00
ATOM   1164  CZ  ARG   126      -0.249  -9.967   3.729  1.00  0.00
ATOM   1165  NH1 ARG   126      -0.311  -8.944   2.877  1.00  0.00
```

```
ATOM   1166  NH2 ARG   126      -1.311 -10.334   4.490  1.00  0.00
ATOM   1167  HE  ARG   126       1.695 -10.411   3.243  1.00  0.00
ATOM   1168  HH12ARG   126      -1.113  -8.367   2.915  1.00  0.00
ATOM   1169  HH11ARG   126       0.352  -8.865   2.174  1.00  0.00
ATOM   1170  HH22ARG   126      -2.110  -9.816   4.431  1.00  0.00
ATOM   1171  HH21ARG   126      -1.347 -11.223   5.020  1.00  0.00
ATOM   1172  H   ARG   126       2.482 -14.759   5.889  1.00  0.00
ATOM   1173  N   LYS   127       3.520 -16.661   3.605  1.00  0.00
ATOM   1174  CA  LYS   127       4.318 -17.625   2.924  1.00  0.00
ATOM   1175  C   LYS   127       3.696 -18.983   3.033  1.00  0.00
ATOM   1176  O   LYS   127       3.738 -19.820   2.116  1.00  0.00
ATOM   1177  CB  LYS   127       5.652 -17.599   3.622  1.00  0.00
ATOM   1178  CG  LYS   127       6.783 -18.126   2.807  1.00  0.00
ATOM   1179  CD  LYS   127       8.161 -18.205   3.379  1.00  0.00
ATOM   1180  CE  LYS   127       8.293 -19.203   4.590  1.00  0.00
ATOM   1181  NZ  LYS   127       7.844 -20.573   4.270  1.00  0.00
ATOM   1182  HZ1 LYS   127       6.958 -20.608   3.752  1.00  0.00
ATOM   1183  HZ2 LYS   127       7.762 -21.079   5.155  1.00  0.00
ATOM   1184  HZ3 LYS   127       8.564 -21.026   3.721  1.00  0.00
ATOM   1185  H   LYS   127       3.647 -16.470   4.564  1.00  0.00
ATOM   1186  N   GLY   128       3.251 -19.360   4.260  1.00  0.00
ATOM   1187  CA  GLY   128       2.438 -20.518   4.413  1.00  0.00
ATOM   1188  C   GLY   128       1.154 -20.406   3.588  1.00  0.00
ATOM   1189  O   GLY   128       0.470 -21.432   3.375  1.00  0.00
ATOM   1190  H   GLY   128       3.610 -18.928   5.052  1.00  0.00
ATOM   1191  N   GLY   129       0.638 -19.262   3.227  1.00  0.00
ATOM   1192  CA  GLY   129      -0.505 -19.057   2.321  1.00  0.00
ATOM   1193  C   GLY   129       0.020 -18.366   1.020  1.00  0.00
ATOM   1194  O   GLY   129      -0.309 -17.184   0.888  1.00  0.00
ATOM   1195  H   GLY   129       0.957 -18.456   3.670  1.00  0.00
ATOM   1196  N   PRO   130       0.772 -19.048   0.156  1.00  0.00
ATOM   1197  CA  PRO   130       1.850 -18.568  -0.715  1.00  0.00
ATOM   1198  C   PRO   130       1.547 -17.399  -1.699  1.00  0.00
ATOM   1199  O   PRO   130       1.483 -17.472  -2.901  1.00  0.00
ATOM   1200  CB  PRO   130       2.310 -19.717  -1.491  1.00  0.00
ATOM   1201  CG  PRO   130       1.712 -20.873  -0.803  1.00  0.00
ATOM   1202  CD  PRO   130       0.361 -20.340  -0.291  1.00  0.00
ATOM   1203  N   GLY   131       1.276 -16.310  -1.047  1.00  0.00
ATOM   1204  CA  GLY   131       0.981 -15.037  -1.732  1.00  0.00
ATOM   1205  C   GLY   131       1.541 -13.987  -0.795  1.00  0.00
ATOM   1206  O   GLY   131       2.763 -13.961  -0.478  1.00  0.00
ATOM   1207  H   GLY   131       1.476 -16.295  -0.099  1.00  0.00
ATOM   1208  N   GLY   132       0.543 -13.329  -0.242  1.00  0.00
ATOM   1209  CA  GLY   132       0.705 -12.228   0.673  1.00  0.00
ATOM   1210  C   GLY   132       1.573 -11.076   0.205  1.00  0.00
ATOM   1211  O   GLY   132       2.680 -10.848   0.611  1.00  0.00
ATOM   1212  H   GLY   132      -0.267 -13.816  -0.390  1.00  0.00
ATOM   1213  N   ILE   133       0.995 -10.360  -0.736  1.00  0.00
ATOM   1214  CA  ILE   133       1.631  -9.210  -1.354  1.00  0.00
ATOM   1215  C   ILE   133       1.676  -8.125  -0.285  1.00  0.00
ATOM   1216  O   ILE   133       0.704  -7.594   0.229  1.00  0.00
ATOM   1217  CB  ILE   133       0.780  -8.742  -2.615  1.00  0.00
ATOM   1218  CG1 ILE   133       0.778  -9.674  -3.789  1.00  0.00
ATOM   1219  CG2 ILE   133       1.381  -7.441  -3.208  1.00  0.00
ATOM   1220  CD1 ILE   133      -0.124 -10.905  -3.719  1.00  0.00
ATOM   1221  H   ILE   133       0.022 -10.631  -0.891  1.00  0.00
ATOM   1222  N   ILE   134       2.813  -7.691   0.185  1.00  0.00
ATOM   1223  CA  ILE   134       2.932  -6.558   1.116  1.00  0.00
ATOM   1224  C   ILE   134       3.167  -5.237   0.332  1.00  0.00
ATOM   1225  O   ILE   134       4.192  -5.106  -0.287  1.00  0.00
ATOM   1226  CB  ILE   134       4.067  -6.691   2.186  1.00  0.00
ATOM   1227  CG1 ILE   134       3.933  -7.874   3.060  1.00  0.00
ATOM   1228  CG2 ILE   134       4.058  -5.395   2.977  1.00  0.00
ATOM   1229  CD1 ILE   134       4.942  -8.957   2.640  1.00  0.00
ATOM   1230  H   ILE   134       3.707  -8.013  -0.078  1.00  0.00
ATOM   1231  N   CYS   135       2.227  -4.291   0.377  1.00  0.00
```

```
ATOM   1232  CA  CYS   135       2.308  -3.151  -0.474  1.00  0.00
ATOM   1233  C   CYS   135       2.425  -1.861   0.356  1.00  0.00
ATOM   1234  O   CYS   135       1.568  -1.671   1.246  1.00  0.00
ATOM   1235  CB  CYS   135       1.093  -2.982  -1.371  1.00  0.00
ATOM   1236  SG  CYS   135       1.112  -3.940  -2.930  1.00  0.00
ATOM   1237  H   CYS   135       1.700  -4.245   1.118  1.00  0.00
ATOM   1238  N   ASN   136       3.432  -1.026   0.059  1.00  0.00
ATOM   1239  CA  ASN   136       3.765   0.174   0.858  1.00  0.00
ATOM   1240  C   ASN   136       3.510   1.369   0.024  1.00  0.00
ATOM   1241  O   ASN   136       3.864   1.511  -1.148  1.00  0.00
ATOM   1242  CB  ASN   136       5.180   0.209   1.268  1.00  0.00
ATOM   1243  CG  ASN   136       5.687  -1.070   1.908  1.00  0.00
ATOM   1244  OD1 ASN   136       5.441  -1.494   3.069  1.00  0.00
ATOM   1245  ND2 ASN   136       6.407  -1.794   1.089  1.00  0.00
ATOM   1246  HD22ASN   136       6.867  -2.575   1.395  1.00  0.00
ATOM   1247  HD21ASN   136       6.400  -1.485   0.140  1.00  0.00
ATOM   1248  H   ASN   136       3.899  -1.311  -0.690  1.00  0.00
ATOM   1249  N   ILE   137       3.038   2.481   0.606  1.00  0.00
ATOM   1250  CA  ILE   137       2.880   3.759  -0.120  1.00  0.00
ATOM   1251  C   ILE   137       4.216   4.516  -0.186  1.00  0.00
ATOM   1252  O   ILE   137       5.117   4.036   0.559  1.00  0.00
ATOM   1253  CB  ILE   137       1.703   4.542   0.659  1.00  0.00
ATOM   1254  CG1 ILE   137       1.130   5.674  -0.147  1.00  0.00
ATOM   1255  CG2 ILE   137       2.300   4.997   1.941  1.00  0.00
ATOM   1256  CD1 ILE   137       0.334   5.137  -1.331  1.00  0.00
ATOM   1257  H   ILE   137       2.864   2.417   1.582  1.00  0.00
ATOM   1258  N   GLY   138       4.499   5.538  -0.947  1.00  0.00
ATOM   1259  CA  GLY   138       5.749   6.217  -1.077  1.00  0.00
ATOM   1260  C   GLY   138       5.255   7.650  -1.355  1.00  0.00
ATOM   1261  O   GLY   138       4.030   7.929  -1.237  1.00  0.00
ATOM   1262  H   GLY   138       3.673   5.965  -1.384  1.00  0.00
ATOM   1263  N   SER   139       6.121   8.617  -1.479  1.00  0.00
ATOM   1264  CA  SER   139       5.736   9.927  -1.873  1.00  0.00
ATOM   1265  C   SER   139       6.970  10.721  -2.333  1.00  0.00
ATOM   1266  O   SER   139       8.046  10.632  -1.642  1.00  0.00
ATOM   1267  CB  SER   139       5.079  10.759  -0.752  1.00  0.00
ATOM   1268  OG  SER   139       4.258  10.236   0.262  1.00  0.00
ATOM   1269  HG  SER   139       3.590   9.698  -0.217  1.00  0.00
ATOM   1270  H   SER   139       7.068   8.448  -1.233  1.00  0.00
ATOM   1271  N   VAL   140       6.861  11.479  -3.451  1.00  0.00
ATOM   1272  CA  VAL   140       7.996  12.203  -4.050  1.00  0.00
ATOM   1273  C   VAL   140       9.124  12.730  -3.182  1.00  0.00
ATOM   1274  O   VAL   140      10.271  12.602  -3.576  1.00  0.00
ATOM   1275  CB  VAL   140       7.498  13.412  -4.877  1.00  0.00
ATOM   1276  CG1 VAL   140       6.757  12.944  -6.143  1.00  0.00
ATOM   1277  CG2 VAL   140       6.568  14.311  -4.037  1.00  0.00
ATOM   1278  H   VAL   140       5.983  11.499  -3.933  1.00  0.00
ATOM   1279  N   THR   141       8.796  13.342  -2.026  1.00  0.00
ATOM   1280  CA  THR   141       9.746  14.021  -1.125  1.00  0.00
ATOM   1281  C   THR   141      11.032  13.148  -0.901  1.00  0.00
ATOM   1282  O   THR   141      12.099  13.697  -1.050  1.00  0.00
ATOM   1283  CB  THR   141       9.077  14.194   0.243  1.00  0.00
ATOM   1284  OG1 THR   141       7.690  14.211   0.030  1.00  0.00
ATOM   1285  CG2 THR   141       9.531  15.457   0.993  1.00  0.00
ATOM   1286  HG1 THR   141       7.286  14.126   0.885  1.00  0.00
ATOM   1287  H   THR   141       7.861  13.375  -1.763  1.00  0.00
ATOM   1288  N   GLY   142      10.954  11.830  -0.737  1.00  0.00
ATOM   1289  CA  GLY   142      12.072  10.894  -0.515  1.00  0.00
ATOM   1290  C   GLY   142      12.978  10.647  -1.717  1.00  0.00
ATOM   1291  O   GLY   142      14.068  10.085  -1.553  1.00  0.00
ATOM   1292  H   GLY   142      10.050  11.432  -0.936  1.00  0.00
ATOM   1293  N   PHE   143      12.553  11.154  -2.834  1.00  0.00
ATOM   1294  CA  PHE   143      13.268  10.916  -4.099  1.00  0.00
ATOM   1295  C   PHE   143      14.097  12.130  -4.486  1.00  0.00
ATOM   1296  O   PHE   143      15.268  12.030  -4.862  1.00  0.00
ATOM   1297  CB  PHE   143      12.287  10.604  -5.203  1.00  0.00
```

```
ATOM   1298  CG  PHE   143      11.669   9.211  -5.246  1.00  0.00
ATOM   1299  CD1 PHE   143      12.137   8.377  -6.218  1.00  0.00
ATOM   1300  CD2 PHE   143      10.745   8.791  -4.343  1.00  0.00
ATOM   1301  CE1 PHE   143      11.669   7.110  -6.331  1.00  0.00
ATOM   1302  CE2 PHE   143      10.288   7.507  -4.488  1.00  0.00
ATOM   1303  CZ  PHE   143      10.730   6.634  -5.472  1.00  0.00
ATOM   1304  H   PHE   143      11.773  11.697  -2.921  1.00  0.00
ATOM   1305  N   ASN   144      13.439  13.319  -4.440  1.00  0.00
ATOM   1306  CA  ASN   144      14.101  14.549  -4.834  1.00  0.00
ATOM   1307  C   ASN   144      15.271  14.864  -3.944  1.00  0.00
ATOM   1308  O   ASN   144      15.159  14.821  -2.729  1.00  0.00
ATOM   1309  CB  ASN   144      13.164  15.771  -4.934  1.00  0.00
ATOM   1310  CG  ASN   144      13.782  16.863  -5.766  1.00  0.00
ATOM   1311  OD1 ASN   144      14.363  17.774  -5.200  1.00  0.00
ATOM   1312  ND2 ASN   144      13.667  16.797  -7.108  1.00  0.00
ATOM   1313  HD22ASN   144      14.066  17.517  -7.677  1.00  0.00
ATOM   1314  HD21ASN   144      13.179  16.041  -7.546  1.00  0.00
ATOM   1315  H   ASN   144      12.479  13.410  -4.183  1.00  0.00
ATOM   1316  N   ALA   145      16.402  15.174  -4.599  1.00  0.00
ATOM   1317  CA  ALA   145      17.621  15.484  -3.862  1.00  0.00
ATOM   1318  C   ALA   145      17.643  16.766  -3.066  1.00  0.00
ATOM   1319  O   ALA   145      18.086  16.700  -1.912  1.00  0.00
ATOM   1320  CB  ALA   145      18.773  15.509  -4.851  1.00  0.00
ATOM   1321  H   ALA   145      16.370  15.177  -5.554  1.00  0.00
ATOM   1322  N   ILE   146      17.155  17.879  -3.651  1.00  0.00
ATOM   1323  CA  ILE   146      17.037  19.118  -2.888  1.00  0.00
ATOM   1324  C   ILE   146      15.753  19.159  -2.084  1.00  0.00
ATOM   1325  O   ILE   146      15.715  19.839  -1.070  1.00  0.00
ATOM   1326  CB  ILE   146      17.116  20.321  -3.852  1.00  0.00
ATOM   1327  CG1 ILE   146      18.366  20.191  -4.746  1.00  0.00
ATOM   1328  CG2 ILE   146      17.209  21.636  -3.059  1.00  0.00
ATOM   1329  CD1 ILE   146      18.167  21.001  -6.039  1.00  0.00
ATOM   1330  H   ILE   146      16.824  17.853  -4.597  1.00  0.00
ATOM   1331  N   TYR   147      14.687  18.435  -2.507  1.00  0.00
ATOM   1332  CA  TYR   147      13.434  18.405  -1.750  1.00  0.00
ATOM   1333  C   TYR   147      13.421  17.353  -0.657  1.00  0.00
ATOM   1334  O   TYR   147      12.409  17.168  -0.001  1.00  0.00
ATOM   1335  CB  TYR   147      12.315  18.179  -2.796  1.00  0.00
ATOM   1336  CG  TYR   147      10.904  18.015  -2.242  1.00  0.00
ATOM   1337  CD1 TYR   147      10.123  16.950  -2.694  1.00  0.00
ATOM   1338  CD2 TYR   147      10.385  18.918  -1.310  1.00  0.00
ATOM   1339  CE1 TYR   147       8.785  16.847  -2.300  1.00  0.00
ATOM   1340  CE2 TYR   147       9.051  18.805  -0.908  1.00  0.00
ATOM   1341  CZ  TYR   147       8.237  17.792  -1.431  1.00  0.00
ATOM   1342  OH  TYR   147       6.885  17.716  -1.095  1.00  0.00
ATOM   1343  HH  TYR   147       6.596  18.439  -0.548  1.00  0.00
ATOM   1344  H   TYR   147      14.740  17.865  -3.327  1.00  0.00
ATOM   1345  N   GLN   148      14.555  16.652  -0.443  1.00  0.00
ATOM   1346  CA  GLN   148      14.676  15.516   0.496  1.00  0.00
ATOM   1347  C   GLN   148      15.022  16.013   1.890  1.00  0.00
ATOM   1348  O   GLN   148      15.510  15.282   2.697  1.00  0.00
ATOM   1349  CB  GLN   148      15.861  14.645  -0.043  1.00  0.00
ATOM   1350  CG  GLN   148      15.839  13.189   0.436  1.00  0.00
ATOM   1351  CD  GLN   148      16.823  12.323  -0.364  1.00  0.00
ATOM   1352  OE1 GLN   148      17.985  12.168   0.007  1.00  0.00
ATOM   1353  NE2 GLN   148      16.423  11.657  -1.509  1.00  0.00
ATOM   1354  HE22GLN   148      17.071  11.341  -2.141  1.00  0.00
ATOM   1355  HE21GLN   148      15.457  11.608  -1.588  1.00  0.00
ATOM   1356  H   GLN   148      15.408  16.723  -1.042  1.00  0.00
ATOM   1357  N   VAL   149      14.630  17.193   2.333  1.00  0.00
ATOM   1358  CA  VAL   149      15.070  17.747   3.641  1.00  0.00
ATOM   1359  C   VAL   149      14.271  17.032   4.811  1.00  0.00
ATOM   1360  O   VAL   149      13.079  16.806   4.699  1.00  0.00
ATOM   1361  CB  VAL   149      14.887  19.252   3.635  1.00  0.00
ATOM   1362  CG1 VAL   149      15.724  19.863   2.503  1.00  0.00
ATOM   1363  CG2 VAL   149      13.432  19.620   3.446  1.00  0.00
```

```
ATOM   1364  H   VAL   149      14.081  17.746   1.712  1.00  0.00
ATOM   1365  N   PRO   150      14.875  16.785   5.993  1.00  0.00
ATOM   1366  CA  PRO   150      14.670  15.556   6.737  1.00  0.00
ATOM   1367  C   PRO   150      13.164  15.248   7.022  1.00  0.00
ATOM   1368  O   PRO   150      12.669  14.212   6.765  1.00  0.00
ATOM   1369  CB  PRO   150      15.487  15.752   8.012  1.00  0.00
ATOM   1370  CG  PRO   150      15.375  17.295   8.147  1.00  0.00
ATOM   1371  CD  PRO   150      15.598  17.814   6.765  1.00  0.00
ATOM   1372  N   VAL   151      12.513  16.231   7.703  1.00  0.00
ATOM   1373  CA  VAL   151      11.133  16.307   8.231  1.00  0.00
ATOM   1374  C   VAL   151      10.414  15.045   8.445  1.00  0.00
ATOM   1375  O   VAL   151      10.246  14.703   9.653  1.00  0.00
ATOM   1376  CB  VAL   151      10.121  17.213   7.411  1.00  0.00
ATOM   1377  CG1 VAL   151      10.643  18.630   7.669  1.00  0.00
ATOM   1378  CG2 VAL   151      10.183  17.086   5.916  1.00  0.00
ATOM   1379  H   VAL   151      13.050  17.060   7.766  1.00  0.00
ATOM   1380  N   TYR   152       9.986  14.381   7.396  1.00  0.00
ATOM   1381  CA  TYR   152       9.343  13.090   7.457  1.00  0.00
ATOM   1382  C   TYR   152      10.007  12.274   6.299  1.00  0.00
ATOM   1383  O   TYR   152       9.932  11.052   6.298  1.00  0.00
ATOM   1384  CB  TYR   152       7.795  13.232   7.115  1.00  0.00
ATOM   1385  CG  TYR   152       6.898  12.297   7.846  1.00  0.00
ATOM   1386  CD1 TYR   152       6.907  12.364   9.258  1.00  0.00
ATOM   1387  CD2 TYR   152       6.065  11.446   7.178  1.00  0.00
ATOM   1388  CE1 TYR   152       5.972  11.550   9.967  1.00  0.00
ATOM   1389  CE2 TYR   152       5.235  10.626   7.812  1.00  0.00
ATOM   1390  CZ  TYR   152       5.201  10.692   9.173  1.00  0.00
ATOM   1391  OH  TYR   152       4.344   9.824   9.818  1.00  0.00
ATOM   1392  HH  TYR   152       4.726   9.771  10.676  1.00  0.00
ATOM   1393  H   TYR   152      10.147  14.876   6.523  1.00  0.00
ATOM   1394  N   SER   153      10.729  12.883   5.391  1.00  0.00
ATOM   1395  CA  SER   153      11.346  12.280   4.171  1.00  0.00
ATOM   1396  C   SER   153      12.000  10.940   4.283  1.00  0.00
ATOM   1397  O   SER   153      11.839  10.075   3.393  1.00  0.00
ATOM   1398  CB  SER   153      12.227  13.300   3.708  1.00  0.00
ATOM   1399  OG  SER   153      11.417  14.475   3.680  1.00  0.00
ATOM   1400  HG  SER   153      12.014  15.223   3.540  1.00  0.00
ATOM   1401  H   SER   153      10.904  13.816   5.565  1.00  0.00
ATOM   1402  N   GLY   154      12.691  10.710   5.376  1.00  0.00
ATOM   1403  CA  GLY   154      13.393   9.398   5.525  1.00  0.00
ATOM   1404  C   GLY   154      12.360   8.223   5.617  1.00  0.00
ATOM   1405  O   GLY   154      12.728   7.104   5.238  1.00  0.00
ATOM   1406  H   GLY   154      12.700  11.401   6.088  1.00  0.00
ATOM   1407  N   THR   155      11.159   8.458   6.155  1.00  0.00
ATOM   1408  CA  THR   155      10.100   7.455   6.239  1.00  0.00
ATOM   1409  C   THR   155       9.712   6.972   4.773  1.00  0.00
ATOM   1410  O   THR   155       9.603   5.752   4.544  1.00  0.00
ATOM   1411  CB  THR   155       8.808   8.057   6.899  1.00  0.00
ATOM   1412  OG1 THR   155       9.206   8.775   8.049  1.00  0.00
ATOM   1413  CG2 THR   155       7.782   6.945   7.133  1.00  0.00
ATOM   1414  HG1 THR   155       8.946   8.228   8.816  1.00  0.00
ATOM   1415  H   THR   155      11.025   9.295   6.531  1.00  0.00
ATOM   1416  N   LYS   156       9.630   7.955   3.906  1.00  0.00
ATOM   1417  CA  LYS   156       9.332   7.644   2.510  1.00  0.00
ATOM   1418  C   LYS   156      10.484   7.016   1.768  1.00  0.00
ATOM   1419  O   LYS   156      10.377   6.216   0.905  1.00  0.00
ATOM   1420  CB  LYS   156       8.841   8.957   1.826  1.00  0.00
ATOM   1421  CG  LYS   156       7.471   9.361   2.356  1.00  0.00
ATOM   1422  CD  LYS   156       7.427  10.865   2.071  1.00  0.00
ATOM   1423  CE  LYS   156       6.451  11.602   2.960  1.00  0.00
ATOM   1424  NZ  LYS   156       6.130  12.891   2.479  1.00  0.00
ATOM   1425  HZ1 LYS   156       5.950  13.436   3.311  1.00  0.00
ATOM   1426  HZ2 LYS   156       5.299  12.754   1.904  1.00  0.00
ATOM   1427  HZ3 LYS   156       6.998  13.270   1.948  1.00  0.00
ATOM   1428  H   LYS   156       9.619   8.914   4.210  1.00  0.00
ATOM   1429  N   ALA   157      11.683   7.361   2.223  1.00  0.00
```

```
ATOM   1430  CA  ALA   157      12.820   6.716   1.651  1.00  0.00
ATOM   1431  C   ALA   157      12.883   5.218   2.012  1.00  0.00
ATOM   1432  O   ALA   157      13.231   4.356   1.217  1.00  0.00
ATOM   1433  CB  ALA   157      14.017   7.394   2.066  1.00  0.00
ATOM   1434  H   ALA   157      11.718   8.225   2.810  1.00  0.00
ATOM   1435  N   ALA   158      12.602   4.907   3.264  1.00  0.00
ATOM   1436  CA  ALA   158      12.561   3.568   3.801  1.00  0.00
ATOM   1437  C   ALA   158      11.655   2.739   2.999  1.00  0.00
ATOM   1438  O   ALA   158      12.099   1.732   2.448  1.00  0.00
ATOM   1439  CB  ALA   158      12.007   3.724   5.205  1.00  0.00
ATOM   1440  H   ALA   158      12.378   5.589   3.885  1.00  0.00
ATOM   1441  N   VAL   159      10.428   3.277   2.846  1.00  0.00
ATOM   1442  CA  VAL   159       9.454   2.430   2.208  1.00  0.00
ATOM   1443  C   VAL   159       9.738   2.048   0.769  1.00  0.00
ATOM   1444  O   VAL   159       9.115   1.100   0.274  1.00  0.00
ATOM   1445  CB  VAL   159       7.905   2.998   2.271  1.00  0.00
ATOM   1446  CG1 VAL   159       7.402   3.158   3.700  1.00  0.00
ATOM   1447  CG2 VAL   159       7.848   4.342   1.612  1.00  0.00
ATOM   1448  H   VAL   159      10.181   4.242   3.031  1.00  0.00
ATOM   1449  N   VAL   160      10.741   2.717   0.127  1.00  0.00
ATOM   1450  CA  VAL   160      11.151   2.356  -1.241  1.00  0.00
ATOM   1451  C   VAL   160      12.132   1.176  -1.132  1.00  0.00
ATOM   1452  O   VAL   160      11.831   0.172  -1.695  1.00  0.00
ATOM   1453  CB  VAL   160      11.852   3.440  -2.059  1.00  0.00
ATOM   1454  CG1 VAL   160      11.778   3.191  -3.543  1.00  0.00
ATOM   1455  CG2 VAL   160      11.319   4.766  -1.715  1.00  0.00
ATOM   1456  H   VAL   160      11.308   3.366   0.627  1.00  0.00
ATOM   1457  N   ASN   161      13.148   1.454  -0.278  1.00  0.00
ATOM   1458  CA  ASN   161      14.310   0.598  -0.269  1.00  0.00
ATOM   1459  C   ASN   161      14.076  -0.796   0.320  1.00  0.00
ATOM   1460  O   ASN   161      14.757  -1.797  -0.053  1.00  0.00
ATOM   1461  CB  ASN   161      15.533   1.277   0.475  1.00  0.00
ATOM   1462  CG  ASN   161      16.104   2.386  -0.346  1.00  0.00
ATOM   1463  OD1 ASN   161      16.978   2.307  -1.204  1.00  0.00
ATOM   1464  ND2 ASN   161      15.723   3.602  -0.067  1.00  0.00
ATOM   1465  HD22ASN   161      16.170   4.135  -0.693  1.00  0.00
ATOM   1466  HD21ASN   161      15.101   3.792   0.697  1.00  0.00
ATOM   1467  H   ASN   161      13.005   2.047   0.484  1.00  0.00
ATOM   1468  N   PHE   162      13.331  -0.900   1.415  1.00  0.00
ATOM   1469  CA  PHE   162      13.048  -2.245   1.915  1.00  0.00
ATOM   1470  C   PHE   162      12.640  -3.325   0.993  1.00  0.00
ATOM   1471  O   PHE   162      12.961  -4.505   1.156  1.00  0.00
ATOM   1472  CB  PHE   162      12.096  -2.060   3.050  1.00  0.00
ATOM   1473  CG  PHE   162      12.880  -2.342   4.313  1.00  0.00
ATOM   1474  CD1 PHE   162      13.302  -3.631   4.613  1.00  0.00
ATOM   1475  CD2 PHE   162      13.262  -1.289   5.178  1.00  0.00
ATOM   1476  CE1 PHE   162      14.116  -3.847   5.754  1.00  0.00
ATOM   1477  CE2 PHE   162      14.027  -1.539   6.264  1.00  0.00
ATOM   1478  CZ  PHE   162      14.478  -2.793   6.578  1.00  0.00
ATOM   1479  H   PHE   162      12.939  -0.126   1.923  1.00  0.00
ATOM   1480  N   THR   163      11.890  -2.934  -0.055  1.00  0.00
ATOM   1481  CA  THR   163      11.346  -3.960  -0.961  1.00  0.00
ATOM   1482  C   THR   163      12.537  -4.637  -1.632  1.00  0.00
ATOM   1483  O   THR   163      12.724  -5.816  -1.436  1.00  0.00
ATOM   1484  CB  THR   163      10.388  -3.202  -1.954  1.00  0.00
ATOM   1485  OG1 THR   163       9.442  -2.656  -1.075  1.00  0.00
ATOM   1486  CG2 THR   163       9.773  -3.982  -3.118  1.00  0.00
ATOM   1487  HG1 THR   163       9.048  -1.863  -1.444  1.00  0.00
ATOM   1488  H   THR   163      11.505  -2.016  -0.067  1.00  0.00
ATOM   1489  N   SER   164      13.328  -3.897  -2.464  1.00  0.00
ATOM   1490  CA  SER   164      14.586  -4.366  -3.073  1.00  0.00
ATOM   1491  C   SER   164      15.495  -5.036  -1.987  1.00  0.00
ATOM   1492  O   SER   164      16.108  -6.014  -2.285  1.00  0.00
ATOM   1493  CB  SER   164      15.270  -3.066  -3.724  1.00  0.00
ATOM   1494  OG  SER   164      14.400  -2.404  -4.563  1.00  0.00
ATOM   1495  HG  SER   164      14.353  -2.905  -5.409  1.00  0.00
```

```
ATOM   1496  H   SER   164      13.131  -2.941  -2.592  1.00  0.00
ATOM   1497  N   SER   165      15.402  -4.552  -0.717  1.00  0.00
ATOM   1498  CA  SER   165      16.092  -5.207   0.330  1.00  0.00
ATOM   1499  C   SER   165      15.586  -6.593   0.782  1.00  0.00
ATOM   1500  O   SER   165      16.336  -7.542   0.908  1.00  0.00
ATOM   1501  CB  SER   165      16.154  -4.285   1.506  1.00  0.00
ATOM   1502  OG  SER   165      16.859  -3.013   1.322  1.00  0.00
ATOM   1503  HG  SER   165      16.397  -2.497   0.670  1.00  0.00
ATOM   1504  H   SER   165      15.009  -3.650  -0.564  1.00  0.00
ATOM   1505  N   LEU   166      14.275  -6.828   0.803  1.00  0.00
ATOM   1506  CA  LEU   166      13.717  -8.099   1.247  1.00  0.00
ATOM   1507  C   LEU   166      13.733  -9.106   0.186  1.00  0.00
ATOM   1508  O   LEU   166      14.217 -10.229   0.260  1.00  0.00
ATOM   1509  CB  LEU   166      12.243  -7.823   1.796  1.00  0.00
ATOM   1510  CG  LEU   166      12.060  -7.075   3.114  1.00  0.00
ATOM   1511  CD1 LEU   166      10.593  -6.894   3.499  1.00  0.00
ATOM   1512  CD2 LEU   166      12.717  -7.894   4.219  1.00  0.00
ATOM   1513  H   LEU   166      13.647  -6.095   0.639  1.00  0.00
ATOM   1514  N   ALA   167      13.409  -8.630  -1.022  1.00  0.00
ATOM   1515  CA  ALA   167      13.179  -9.432  -2.232  1.00  0.00
ATOM   1516  C   ALA   167      14.235 -10.357  -2.725  1.00  0.00
ATOM   1517  O   ALA   167      13.972 -11.294  -3.475  1.00  0.00
ATOM   1518  CB  ALA   167      12.829  -8.408  -3.341  1.00  0.00
ATOM   1519  H   ALA   167      13.154  -7.647  -1.100  1.00  0.00
ATOM   1520  N   LYS   168      15.495 -10.101  -2.290  1.00  0.00
ATOM   1521  CA  LYS   168      16.639 -11.045  -2.510  1.00  0.00
ATOM   1522  C   LYS   168      16.385 -12.455  -1.965  1.00  0.00
ATOM   1523  O   LYS   168      16.416 -13.405  -2.741  1.00  0.00
ATOM   1524  CB  LYS   168      17.973 -10.568  -1.930  1.00  0.00
ATOM   1525  CG  LYS   168      18.461  -9.461  -2.792  1.00  0.00
ATOM   1526  CD  LYS   168      19.835  -9.039  -2.291  1.00  0.00
ATOM   1527  CE  LYS   168      20.281  -7.730  -2.851  1.00  0.00
ATOM   1528  NZ  LYS   168      21.690  -7.581  -2.652  1.00  0.00
ATOM   1529  HZ1 LYS   168      22.107  -8.226  -3.371  1.00  0.00
ATOM   1530  HZ2 LYS   168      21.993  -7.954  -1.704  1.00  0.00
ATOM   1531  HZ3 LYS   168      22.046  -6.633  -2.791  1.00  0.00
ATOM   1532  H   LYS   168      15.670  -9.205  -1.851  1.00  0.00
ATOM   1533  N   LEU   169      16.271 -12.612  -0.678  1.00  0.00
ATOM   1534  CA  LEU   169      16.024 -13.898  -0.221  1.00  0.00
ATOM   1535  C   LEU   169      14.538 -14.115  -0.153  1.00  0.00
ATOM   1536  O   LEU   169      14.092 -14.492   0.916  1.00  0.00
ATOM   1537  CB  LEU   169      16.780 -13.968   1.086  1.00  0.00
ATOM   1538  CG  LEU   169      18.251 -14.324   1.132  1.00  0.00
ATOM   1539  CD1 LEU   169      18.867 -13.916   2.515  1.00  0.00
ATOM   1540  CD2 LEU   169      18.372 -15.855   0.927  1.00  0.00
ATOM   1541  H   LEU   169      16.165 -11.853  -0.074  1.00  0.00
ATOM   1542  N   ALA   170      13.768 -13.913  -1.234  1.00  0.00
ATOM   1543  CA  ALA   170      12.324 -14.024  -1.167  1.00  0.00
ATOM   1544  C   ALA   170      11.832 -15.185  -2.001  1.00  0.00
ATOM   1545  O   ALA   170      12.140 -15.321  -3.188  1.00  0.00
ATOM   1546  CB  ALA   170      11.691 -12.779  -1.654  1.00  0.00
ATOM   1547  H   ALA   170      14.137 -13.630  -2.126  1.00  0.00
ATOM   1548  N   PRO   171      11.011 -16.109  -1.423  1.00  0.00
ATOM   1549  CA  PRO   171      10.376 -17.237  -2.082  1.00  0.00
ATOM   1550  C   PRO   171       9.398 -16.740  -3.202  1.00  0.00
ATOM   1551  O   PRO   171       8.220 -16.750  -3.050  1.00  0.00
ATOM   1552  CB  PRO   171       9.624 -17.934  -0.992  1.00  0.00
ATOM   1553  CG  PRO   171      10.485 -17.703   0.182  1.00  0.00
ATOM   1554  CD  PRO   171      10.861 -16.221  -0.001  1.00  0.00
ATOM   1555  N   ILE   172       9.852 -16.263  -4.359  1.00  0.00
ATOM   1556  CA  ILE   172       9.015 -15.762  -5.462  1.00  0.00
ATOM   1557  C   ILE   172       8.089 -16.726  -6.156  1.00  0.00
ATOM   1558  O   ILE   172       7.810 -16.731  -7.368  1.00  0.00
ATOM   1559  CB  ILE   172       9.881 -15.100  -6.538  1.00  0.00
ATOM   1560  CG1 ILE   172      11.080 -15.928  -6.946  1.00  0.00
ATOM   1561  CG2 ILE   172      10.302 -13.734  -6.016  1.00  0.00
```

```
ATOM   1562  CD1 ILE   172      11.700 -15.559  -8.281  1.00  0.00
ATOM   1563  H   ILE   172      10.845 -16.121  -4.385  1.00  0.00
ATOM   1564  N   THR   173       7.453 -17.590  -5.415  1.00  0.00
ATOM   1565  CA  THR   173       6.359 -18.420  -5.791  1.00  0.00
ATOM   1566  C   THR   173       5.135 -17.562  -5.805  1.00  0.00
ATOM   1567  O   THR   173       4.270 -17.634  -6.682  1.00  0.00
ATOM   1568  CB  THR   173       6.292 -19.600  -4.779  1.00  0.00
ATOM   1569  OG1 THR   173       7.665 -20.138  -4.627  1.00  0.00
ATOM   1570  CG2 THR   173       5.444 -20.737  -5.339  1.00  0.00
ATOM   1571  HG1 THR   173       7.687 -20.720  -3.913  1.00  0.00
ATOM   1572  H   THR   173       7.829 -17.870  -4.576  1.00  0.00
ATOM   1573  N   GLY   174       4.958 -16.681  -4.875  1.00  0.00
ATOM   1574  CA  GLY   174       3.768 -15.821  -4.792  1.00  0.00
ATOM   1575  C   GLY   174       4.057 -14.609  -3.829  1.00  0.00
ATOM   1576  O   GLY   174       3.677 -13.454  -4.046  1.00  0.00
ATOM   1577  H   GLY   174       5.783 -16.498  -4.374  1.00  0.00
ATOM   1578  N   VAL   175       4.906 -14.908  -2.922  1.00  0.00
ATOM   1579  CA  VAL   175       5.464 -13.946  -1.963  1.00  0.00
ATOM   1580  C   VAL   175       6.097 -12.802  -2.827  1.00  0.00
ATOM   1581  O   VAL   175       6.984 -13.140  -3.697  1.00  0.00
ATOM   1582  CB  VAL   175       6.614 -14.493  -1.078  1.00  0.00
ATOM   1583  CG1 VAL   175       7.072 -13.455  -0.056  1.00  0.00
ATOM   1584  CG2 VAL   175       6.165 -15.886  -0.405  1.00  0.00
ATOM   1585  H   VAL   175       5.137 -15.855  -2.839  1.00  0.00
ATOM   1586  N   THR   176       5.603 -11.584  -2.614  1.00  0.00
ATOM   1587  CA  THR   176       5.907 -10.395  -3.337  1.00  0.00
ATOM   1588  C   THR   176       5.594  -9.145  -2.484  1.00  0.00
ATOM   1589  O   THR   176       4.764  -9.195  -1.654  1.00  0.00
ATOM   1590  CB  THR   176       5.009 -10.360  -4.632  1.00  0.00
ATOM   1591  OG1 THR   176       5.416 -11.414  -5.441  1.00  0.00
ATOM   1592  CG2 THR   176       5.085  -9.069  -5.428  1.00  0.00
ATOM   1593  HG1 THR   176       4.918 -12.173  -5.059  1.00  0.00
ATOM   1594  H   THR   176       4.928 -11.583  -1.872  1.00  0.00
ATOM   1595  N   ALA   177       6.218  -7.940  -2.838  1.00  0.00
ATOM   1596  CA  ALA   177       6.006  -6.686  -2.125  1.00  0.00
ATOM   1597  C   ALA   177       6.044  -5.536  -3.164  1.00  0.00
ATOM   1598  O   ALA   177       6.606  -5.832  -4.182  1.00  0.00
ATOM   1599  CB  ALA   177       7.072  -6.366  -1.061  1.00  0.00
ATOM   1600  H   ALA   177       6.754  -7.908  -3.708  1.00  0.00
ATOM   1601  N   TYR   178       5.528  -4.357  -2.999  1.00  0.00
ATOM   1602  CA  TYR   178       5.578  -3.277  -3.988  1.00  0.00
ATOM   1603  C   TYR   178       5.360  -1.994  -3.248  1.00  0.00
ATOM   1604  O   TYR   178       4.784  -1.969  -2.113  1.00  0.00
ATOM   1605  CB  TYR   178       4.534  -3.305  -5.029  1.00  0.00
ATOM   1606  CG  TYR   178       5.252  -3.343  -6.340  1.00  0.00
ATOM   1607  CD1 TYR   178       5.897  -2.230  -6.803  1.00  0.00
ATOM   1608  CD2 TYR   178       5.186  -4.494  -7.104  1.00  0.00
ATOM   1609  CE1 TYR   178       6.491  -2.135  -8.044  1.00  0.00
ATOM   1610  CE2 TYR   178       5.819  -4.387  -8.346  1.00  0.00
ATOM   1611  CZ  TYR   178       6.443  -3.261  -8.832  1.00  0.00
ATOM   1612  OH  TYR   178       7.079  -3.305 -10.040  1.00  0.00
ATOM   1613  HH  TYR   178       7.628  -2.505 -10.220  1.00  0.00
ATOM   1614  H   TYR   178       5.126  -4.162  -2.118  1.00  0.00
ATOM   1615  N   THR   179       5.935  -0.952  -3.875  1.00  0.00
ATOM   1616  CA  THR   179       5.656   0.436  -3.479  1.00  0.00
ATOM   1617  C   THR   179       4.925   1.240  -4.581  1.00  0.00
ATOM   1618  O   THR   179       5.332   1.197  -5.740  1.00  0.00
ATOM   1619  CB  THR   179       7.076   1.016  -3.071  1.00  0.00
ATOM   1620  OG1 THR   179       7.807   0.004  -2.379  1.00  0.00
ATOM   1621  CG2 THR   179       6.941   2.251  -2.203  1.00  0.00
ATOM   1622  HG1 THR   179       8.077   0.374  -1.460  1.00  0.00
ATOM   1623  H   THR   179       6.513  -1.096  -4.632  1.00  0.00
ATOM   1624  N   VAL   180       3.864   1.953  -4.190  1.00  0.00
ATOM   1625  CA  VAL   180       3.117   2.779  -5.198  1.00  0.00
ATOM   1626  C   VAL   180       3.712   4.179  -4.826  1.00  0.00
ATOM   1627  O   VAL   180       3.674   4.511  -3.601  1.00  0.00
```

```
ATOM   1628  CB  VAL   180       1.560   2.663  -4.827  1.00  0.00
ATOM   1629  CG1 VAL   180       0.749   3.025  -6.026  1.00  0.00
ATOM   1630  CG2 VAL   180       1.143   1.259  -4.366  1.00  0.00
ATOM   1631  H   VAL   180       3.798   2.167  -3.288  1.00  0.00
ATOM   1632  N   ASN   181       4.258   5.027  -5.710  1.00  0.00
ATOM   1633  CA  ASN   181       4.808   6.365  -5.258  1.00  0.00
ATOM   1634  C   ASN   181       3.769   7.477  -5.322  1.00  0.00
ATOM   1635  O   ASN   181       3.225   7.872  -4.285  1.00  0.00
ATOM   1636  CB  ASN   181       6.049   6.827  -6.054  1.00  0.00
ATOM   1637  CG  ASN   181       6.599   8.215  -5.584  1.00  0.00
ATOM   1638  OD1 ASN   181       7.199   8.391  -4.559  1.00  0.00
ATOM   1639  ND2 ASN   181       6.398   9.172  -6.462  1.00  0.00
ATOM   1640  HD22ASN   181       6.452  10.080  -6.065  1.00  0.00
ATOM   1641  HD21ASN   181       6.067   9.041  -7.409  1.00  0.00
ATOM   1642  H   ASN   181       4.107   4.808  -6.655  1.00  0.00
ATOM   1643  N   PRO   182       3.470   8.120  -6.479  1.00  0.00
ATOM   1644  CA  PRO   182       3.311   9.616  -6.619  1.00  0.00
ATOM   1645  C   PRO   182       3.104  10.586  -5.425  1.00  0.00
ATOM   1646  O   PRO   182       4.113  11.143  -4.934  1.00  0.00
ATOM   1647  CB  PRO   182       2.246   9.744  -7.695  1.00  0.00
ATOM   1648  CG  PRO   182       2.744   8.626  -8.656  1.00  0.00
ATOM   1649  CD  PRO   182       3.064   7.503  -7.712  1.00  0.00
ATOM   1650  N   GLY   183       1.896  10.643  -4.894  1.00  0.00
ATOM   1651  CA  GLY   183       1.661  11.491  -3.801  1.00  0.00
ATOM   1652  C   GLY   183       0.637  12.463  -4.331  1.00  0.00
ATOM   1653  O   GLY   183      -0.112  12.212  -5.255  1.00  0.00
ATOM   1654  H   GLY   183       1.085  10.196  -5.244  1.00  0.00
ATOM   1655  N   ILE   184       0.542  13.628  -3.649  1.00  0.00
ATOM   1656  CA  ILE   184      -0.455  14.636  -4.006  1.00  0.00
ATOM   1657  C   ILE   184      -1.884  14.132  -3.999  1.00  0.00
ATOM   1658  O   ILE   184      -2.507  14.082  -5.050  1.00  0.00
ATOM   1659  CB  ILE   184      -0.053  15.354  -5.314  1.00  0.00
ATOM   1660  CG1 ILE   184       1.414  15.825  -5.243  1.00  0.00
ATOM   1661  CG2 ILE   184      -0.970  16.567  -5.570  1.00  0.00
ATOM   1662  CD1 ILE   184       1.629  16.772  -4.045  1.00  0.00
ATOM   1663  H   ILE   184       1.147  13.833  -2.877  1.00  0.00
ATOM   1664  N   THR   185      -2.428  13.745  -2.825  1.00  0.00
ATOM   1665  CA  THR   185      -3.739  13.119  -2.822  1.00  0.00
ATOM   1666  C   THR   185      -4.933  14.068  -2.905  1.00  0.00
ATOM   1667  O   THR   185      -5.285  14.883  -2.029  1.00  0.00
ATOM   1668  CB  THR   185      -3.870  12.253  -1.538  1.00  0.00
ATOM   1669  OG1 THR   185      -2.589  11.706  -1.255  1.00  0.00
ATOM   1670  CG2 THR   185      -4.751  11.097  -1.812  1.00  0.00
ATOM   1671  HG1 THR   185      -2.748  11.087  -0.550  1.00  0.00
ATOM   1672  H   THR   185      -1.899  13.743  -1.994  1.00  0.00
ATOM   1673  N   ARG   186      -5.490  14.009  -4.116  1.00  0.00
ATOM   1674  CA  ARG   186      -6.692  14.767  -4.391  1.00  0.00
ATOM   1675  C   ARG   186      -7.842  13.986  -3.747  1.00  0.00
ATOM   1676  O   ARG   186      -8.188  12.910  -4.244  1.00  0.00
ATOM   1677  CB  ARG   186      -6.860  14.860  -5.865  1.00  0.00
ATOM   1678  CG  ARG   186      -6.240  16.034  -6.605  1.00  0.00
ATOM   1679  CD  ARG   186      -4.716  15.929  -6.617  1.00  0.00
ATOM   1680  NE  ARG   186      -3.969  17.159  -6.687  1.00  0.00
ATOM   1681  CZ  ARG   186      -3.617  17.764  -7.847  1.00  0.00
ATOM   1682  NH1 ARG   186      -2.957  18.970  -7.795  1.00  0.00
ATOM   1683  NH2 ARG   186      -3.712  17.168  -9.039  1.00  0.00
ATOM   1684  HE  ARG   186      -3.595  17.586  -5.886  1.00  0.00
ATOM   1685  HH12ARG   186      -2.584  19.292  -8.664  1.00  0.00
ATOM   1686  HH11ARG   186      -2.767  19.460  -6.880  1.00  0.00
ATOM   1687  HH22ARG   186      -3.503  17.726  -9.845  1.00  0.00
ATOM   1688  HH21ARG   186      -3.936  16.169  -9.143  1.00  0.00
ATOM   1689  H   ARG   186      -5.028  13.328  -4.726  1.00  0.00
ATOM   1690  N   THR   187      -8.453  14.535  -2.710  1.00  0.00
ATOM   1691  CA  THR   187      -9.637  13.988  -2.069  1.00  0.00
ATOM   1692  C   THR   187     -10.788  13.697  -3.102  1.00  0.00
ATOM   1693  O   THR   187     -11.029  14.507  -4.009  1.00  0.00
```

```
ATOM   1694  CB  THR   187      -9.995  15.076  -1.002  1.00  0.00
ATOM   1695  OG1 THR   187      -8.893  15.378  -0.176  1.00  0.00
ATOM   1696  CG2 THR   187     -11.145  14.589  -0.145  1.00  0.00
ATOM   1697  HG1 THR   187      -8.570  14.608   0.105  1.00  0.00
ATOM   1698  H   THR   187      -8.168  15.394  -2.335  1.00  0.00
ATOM   1699  N   THR   188     -11.304  12.472  -2.917  1.00  0.00
ATOM   1700  CA  THR   188     -12.422  11.918  -3.597  1.00  0.00
ATOM   1701  C   THR   188     -11.962  11.454  -4.985  1.00  0.00
ATOM   1702  O   THR   188     -12.283  10.344  -5.340  1.00  0.00
ATOM   1703  CB  THR   188     -13.639  12.913  -3.752  1.00  0.00
ATOM   1704  OG1 THR   188     -13.675  13.621  -2.492  1.00  0.00
ATOM   1705  CG2 THR   188     -15.016  12.259  -3.977  1.00  0.00
ATOM   1706  HG1 THR   188     -14.421  14.193  -2.474  1.00  0.00
ATOM   1707  H   THR   188     -11.054  11.969  -2.134  1.00  0.00
ATOM   1708  N   LEU   189     -11.228  12.270  -5.756  1.00  0.00
ATOM   1709  CA  LEU   189     -10.859  12.034  -7.113  1.00  0.00
ATOM   1710  C   LEU   189      -9.613  11.143  -7.244  1.00  0.00
ATOM   1711  O   LEU   189      -9.256  10.689  -8.373  1.00  0.00
ATOM   1712  CB  LEU   189     -10.684  13.479  -7.664  1.00  0.00
ATOM   1713  CG  LEU   189     -11.771  14.239  -8.398  1.00  0.00
ATOM   1714  CD1 LEU   189     -13.145  14.048  -7.878  1.00  0.00
ATOM   1715  CD2 LEU   189     -11.280  15.705  -8.359  1.00  0.00
ATOM   1716  H   LEU   189     -10.903  13.088  -5.343  1.00  0.00
ATOM   1717  N   VAL   190      -8.914  10.866  -6.182  1.00  0.00
ATOM   1718  CA  VAL   190      -7.744   9.998  -6.169  1.00  0.00
ATOM   1719  C   VAL   190      -8.089   8.663  -6.757  1.00  0.00
ATOM   1720  O   VAL   190      -7.122   7.969  -7.018  1.00  0.00
ATOM   1721  CB  VAL   190      -7.261   9.989  -4.662  1.00  0.00
ATOM   1722  CG1 VAL   190      -8.423   9.352  -3.817  1.00  0.00
ATOM   1723  CG2 VAL   190      -5.976   9.183  -4.492  1.00  0.00
ATOM   1724  H   VAL   190      -9.204  11.292  -5.357  1.00  0.00
ATOM   1725  N   HIS   191      -9.305   8.222  -7.093  1.00  0.00
ATOM   1726  CA  HIS   191      -9.526   7.012  -7.848  1.00  0.00
ATOM   1727  C   HIS   191      -8.610   6.854  -9.038  1.00  0.00
ATOM   1728  O   HIS   191      -7.999   5.798  -9.145  1.00  0.00
ATOM   1729  CB  HIS   191     -10.980   6.960  -8.287  1.00  0.00
ATOM   1730  CG  HIS   191     -11.977   6.731  -7.146  1.00  0.00
ATOM   1731  ND1 HIS   191     -12.004   5.684  -6.337  1.00  0.00
ATOM   1732  CD2 HIS   191     -12.933   7.587  -6.754  1.00  0.00
ATOM   1733  CE1 HIS   191     -12.976   5.900  -5.404  1.00  0.00
ATOM   1734  NE2 HIS   191     -13.540   7.082  -5.659  1.00  0.00
ATOM   1735  HE2 HIS   191     -14.310   7.533  -5.269  1.00  0.00
ATOM   1736  HD1 HIS   191     -11.350   4.937  -6.416  1.00  0.00
ATOM   1737  H   HIS   191     -10.053   8.757  -6.783  1.00  0.00
ATOM   1738  N   LYS   192      -8.285   7.990  -9.696  1.00  0.00
ATOM   1739  CA  LYS   192      -7.350   8.123 -10.790  1.00  0.00
ATOM   1740  C   LYS   192      -5.996   7.415 -10.456  1.00  0.00
ATOM   1741  O   LYS   192      -5.709   6.402 -11.087  1.00  0.00
ATOM   1742  CB  LYS   192      -7.209   9.589 -11.100  1.00  0.00
ATOM   1743  CG  LYS   192      -8.492   9.999 -11.844  1.00  0.00
ATOM   1744  CD  LYS   192      -8.685  11.537 -11.903  1.00  0.00
ATOM   1745  CE  LYS   192      -7.580  12.425 -12.540  1.00  0.00
ATOM   1746  NZ  LYS   192      -7.915  13.837 -12.577  1.00  0.00
ATOM   1747  HZ1 LYS   192      -8.912  13.925 -12.889  1.00  0.00
ATOM   1748  HZ2 LYS   192      -7.198  14.279 -13.190  1.00  0.00
ATOM   1749  HZ3 LYS   192      -7.784  14.248 -11.614  1.00  0.00
ATOM   1750  H   LYS   192      -8.665   8.789  -9.263  1.00  0.00
ATOM   1751  N   PHE   193      -5.392   7.732  -9.294  1.00  0.00
ATOM   1752  CA  PHE   193      -4.078   7.185  -9.053  1.00  0.00
ATOM   1753  C   PHE   193      -4.194   5.835  -8.347  1.00  0.00
ATOM   1754  O   PHE   193      -3.428   4.865  -8.591  1.00  0.00
ATOM   1755  CB  PHE   193      -3.183   8.038  -8.191  1.00  0.00
ATOM   1756  CG  PHE   193      -3.062   9.562  -8.467  1.00  0.00
ATOM   1757  CD1 PHE   193      -2.725  10.084  -9.676  1.00  0.00
ATOM   1758  CD2 PHE   193      -3.248  10.446  -7.419  1.00  0.00
ATOM   1759  CE1 PHE   193      -2.560  11.446  -9.815  1.00  0.00
```

```
ATOM   1760  CE2 PHE   193      -3.085  11.827  -7.566  1.00  0.00
ATOM   1761  CZ  PHE   193      -2.724  12.345  -8.783  1.00  0.00
ATOM   1762  H   PHE   193      -5.958   8.167  -8.688  1.00  0.00
ATOM   1763  N   ASN   194      -5.202   5.708  -7.541  1.00  0.00
ATOM   1764  CA  ASN   194      -5.401   4.508  -6.705  1.00  0.00
ATOM   1765  C   ASN   194      -5.861   3.270  -7.528  1.00  0.00
ATOM   1766  O   ASN   194      -5.787   2.114  -7.119  1.00  0.00
ATOM   1767  CB  ASN   194      -6.427   4.873  -5.573  1.00  0.00
ATOM   1768  CG  ASN   194      -5.835   5.546  -4.390  1.00  0.00
ATOM   1769  OD1 ASN   194      -4.777   6.184  -4.414  1.00  0.00
ATOM   1770  ND2 ASN   194      -6.374   5.275  -3.258  1.00  0.00
ATOM   1771  HD22ASN   194      -5.924   5.860  -2.578  1.00  0.00
ATOM   1772  HD21ASN   194      -7.147   4.675  -3.157  1.00  0.00
ATOM   1773  H   ASN   194      -5.787   6.490  -7.436  1.00  0.00
ATOM   1774  N   SER   195      -6.248   3.572  -8.789  1.00  0.00
ATOM   1775  CA  SER   195      -6.462   2.462  -9.697  1.00  0.00
ATOM   1776  C   SER   195      -5.175   1.689  -9.811  1.00  0.00
ATOM   1777  O   SER   195      -5.273   0.484  -9.981  1.00  0.00
ATOM   1778  CB  SER   195      -6.867   3.058 -11.032  1.00  0.00
ATOM   1779  OG  SER   195      -8.164   3.573 -10.982  1.00  0.00
ATOM   1780  HG  SER   195      -7.958   4.496 -10.637  1.00  0.00
ATOM   1781  H   SER   195      -6.299   4.516  -9.054  1.00  0.00
ATOM   1782  N   TRP   196      -3.926   2.182  -9.657  1.00  0.00
ATOM   1783  CA  TRP   196      -2.692   1.371  -9.719  1.00  0.00
ATOM   1784  C   TRP   196      -2.672   0.280  -8.701  1.00  0.00
ATOM   1785  O   TRP   196      -2.084  -0.794  -9.019  1.00  0.00
ATOM   1786  CB  TRP   196      -1.572   2.356  -9.629  1.00  0.00
ATOM   1787  CG  TRP   196      -0.650   2.637 -10.855  1.00  0.00
ATOM   1788  CD1 TRP   196       0.681   3.036 -10.629  1.00  0.00
ATOM   1789  CD2 TRP   196      -1.048   2.563 -12.206  1.00  0.00
ATOM   1790  NE1 TRP   196       1.171   3.191 -11.799  1.00  0.00
ATOM   1791  CE2 TRP   196       0.193   2.963 -12.707  1.00  0.00
ATOM   1792  CE3 TRP   196      -2.092   2.264 -13.092  1.00  0.00
ATOM   1793  CZ2 TRP   196       0.415   3.057 -14.039  1.00  0.00
ATOM   1794  CZ3 TRP   196      -1.846   2.377 -14.400  1.00  0.00
ATOM   1795  CH2 TRP   196      -0.572   2.797 -14.896  1.00  0.00
ATOM   1796  HE1 TRP   196       2.062   3.492 -11.943  1.00  0.00
ATOM   1797  H   TRP   196      -3.785   3.123  -9.433  1.00  0.00
ATOM   1798  N   LEU   197      -3.361   0.381  -7.507  1.00  0.00
ATOM   1799  CA  LEU   197      -3.384  -0.718  -6.561  1.00  0.00
ATOM   1800  C   LEU   197      -4.185  -1.940  -7.109  1.00  0.00
ATOM   1801  O   LEU   197      -3.893  -3.128  -6.841  1.00  0.00
ATOM   1802  CB  LEU   197      -4.099  -0.278  -5.353  1.00  0.00
ATOM   1803  CG  LEU   197      -3.248   0.282  -4.144  1.00  0.00
ATOM   1804  CD1 LEU   197      -4.009   1.422  -3.541  1.00  0.00
ATOM   1805  CD2 LEU   197      -3.084  -0.769  -3.041  1.00  0.00
ATOM   1806  H   LEU   197      -3.852   1.164  -7.221  1.00  0.00
ATOM   1807  N   ASP   198      -5.177  -1.676  -7.891  1.00  0.00
ATOM   1808  CA  ASP   198      -5.977  -2.771  -8.470  1.00  0.00
ATOM   1809  C   ASP   198      -5.215  -3.582  -9.561  1.00  0.00
ATOM   1810  O   ASP   198      -5.738  -4.590 -10.053  1.00  0.00
ATOM   1811  CB  ASP   198      -7.254  -2.141  -8.987  1.00  0.00
ATOM   1812  CG  ASP   198      -8.056  -1.424  -7.893  1.00  0.00
ATOM   1813  OD1 ASP   198      -7.684  -0.359  -7.409  1.00  0.00
ATOM   1814  OD2 ASP   198      -9.070  -1.975  -7.451  1.00  0.00
ATOM   1815  H   ASP   198      -5.354  -0.767  -8.185  1.00  0.00
ATOM   1816  N   VAL   199      -4.010  -3.153  -9.947  1.00  0.00
ATOM   1817  CA  VAL   199      -3.245  -3.862 -10.918  1.00  0.00
ATOM   1818  C   VAL   199      -1.957  -4.428 -10.319  1.00  0.00
ATOM   1819  O   VAL   199      -1.148  -5.106 -10.989  1.00  0.00
ATOM   1820  CB  VAL   199      -2.786  -2.998 -12.177  1.00  0.00
ATOM   1821  CG1 VAL   199      -3.696  -3.493 -13.307  1.00  0.00
ATOM   1822  CG2 VAL   199      -2.856  -1.476 -11.985  1.00  0.00
ATOM   1823  H   VAL   199      -3.589  -2.490  -9.333  1.00  0.00
ATOM   1824  N   GLU   200      -1.708  -4.247  -9.020  1.00  0.00
ATOM   1825  CA  GLU   200      -0.434  -4.671  -8.532  1.00  0.00
```

```
ATOM   1826  C   GLU   200      -0.498  -6.171  -8.180  1.00  0.00
ATOM   1827  O   GLU   200      -1.557  -6.767  -7.898  1.00  0.00
ATOM   1828  CB  GLU   200      -0.077  -3.936  -7.234  1.00  0.00
ATOM   1829  CG  GLU   200       0.386  -2.467  -7.518  1.00  0.00
ATOM   1830  CD  GLU   200       0.693  -1.674  -6.295  1.00  0.00
ATOM   1831  OE1 GLU   200       1.762  -1.185  -6.157  1.00  0.00
ATOM   1832  OE2 GLU   200      -0.233  -1.655  -5.501  1.00  0.00
ATOM   1833  H   GLU   200      -2.427  -3.917  -8.398  1.00  0.00
ATOM   1834  N   PRO   201       0.615  -6.935  -8.389  1.00  0.00
ATOM   1835  CA  PRO   201       1.425  -6.862  -9.609  1.00  0.00
ATOM   1836  C   PRO   201       0.930  -7.952 -10.573  1.00  0.00
ATOM   1837  O   PRO   201       1.691  -8.726 -11.047  1.00  0.00
ATOM   1838  CB  PRO   201       2.797  -7.032  -9.086  1.00  0.00
ATOM   1839  CG  PRO   201       2.719  -6.945  -7.598  1.00  0.00
ATOM   1840  CD  PRO   201       1.381  -7.609  -7.395  1.00  0.00
ATOM   1841  N   GLN   202      -0.352  -7.935 -10.966  1.00  0.00
ATOM   1842  CA  GLN   202      -0.944  -8.937 -11.765  1.00  0.00
ATOM   1843  C   GLN   202      -0.542  -8.757 -13.208  1.00  0.00
ATOM   1844  O   GLN   202      -0.453  -9.714 -13.970  1.00  0.00
ATOM   1845  CB  GLN   202      -2.492  -9.056 -11.864  1.00  0.00
ATOM   1846  CG  GLN   202      -3.414  -8.897 -10.675  1.00  0.00
ATOM   1847  CD  GLN   202      -4.169  -7.637 -10.724  1.00  0.00
ATOM   1848  OE1 GLN   202      -4.803  -7.246 -11.737  1.00  0.00
ATOM   1849  NE2 GLN   202      -4.039  -6.843  -9.685  1.00  0.00
ATOM   1850  HE22GLN   202      -4.583  -6.042  -9.724  1.00  0.00
ATOM   1851  HE21GLN   202      -3.481  -7.148  -8.923  1.00  0.00
ATOM   1852  H   GLN   202      -0.931  -7.129 -10.773  1.00  0.00
ATOM   1853  N   VAL   203      -0.350  -7.492 -13.605  1.00  0.00
ATOM   1854  CA  VAL   203       0.103  -6.999 -14.921  1.00  0.00
ATOM   1855  C   VAL   203      -0.738  -7.536 -16.078  1.00  0.00
ATOM   1856  O   VAL   203      -1.678  -8.313 -15.960  1.00  0.00
ATOM   1857  CB  VAL   203       1.690  -7.266 -15.270  1.00  0.00
ATOM   1858  CG1 VAL   203       2.497  -6.600 -14.136  1.00  0.00
ATOM   1859  CG2 VAL   203       2.050  -8.673 -15.456  1.00  0.00
ATOM   1860  H   VAL   203      -0.307  -6.766 -12.915  1.00  0.00
ATOM   1861  N   ALA   204      -0.550  -6.942 -17.204  1.00  0.00
ATOM   1862  CA  ALA   204      -1.387  -7.187 -18.385  1.00  0.00
ATOM   1863  C   ALA   204      -1.025  -8.518 -18.998  1.00  0.00
ATOM   1864  O   ALA   204      -0.001  -9.238 -18.693  1.00  0.00
ATOM   1865  CB  ALA   204      -1.128  -6.048 -19.416  1.00  0.00
ATOM   1866  H   ALA   204       0.219  -6.318 -17.250  1.00  0.00
ATOM   1867  N   GLU   205      -2.009  -9.070 -19.729  1.00  0.00
ATOM   1868  CA  GLU   205      -1.799 -10.287 -20.528  1.00  0.00
ATOM   1869  C   GLU   205      -0.792 -10.021 -21.646  1.00  0.00
ATOM   1870  O   GLU   205       0.096 -10.828 -22.018  1.00  0.00
ATOM   1871  CB  GLU   205      -3.108 -10.736 -21.100  1.00  0.00
ATOM   1872  CG  GLU   205      -3.958 -11.413 -20.074  1.00  0.00
ATOM   1873  CD  GLU   205      -3.762 -12.897 -20.015  1.00  0.00
ATOM   1874  OE1 GLU   205      -4.493 -13.606 -20.712  1.00  0.00
ATOM   1875  OE2 GLU   205      -2.879 -13.349 -19.212  1.00  0.00
ATOM   1876  H   GLU   205      -2.889  -8.779 -19.603  1.00  0.00
ATOM   1877  N   LYS   206      -0.822  -8.753 -22.114  1.00  0.00
ATOM   1878  CA  LYS   206       0.089  -8.445 -23.204  1.00  0.00
ATOM   1879  C   LYS   206       1.351  -7.941 -22.461  1.00  0.00
ATOM   1880  O   LYS   206       1.479  -6.823 -22.019  1.00  0.00
ATOM   1881  CB  LYS   206      -0.346  -7.253 -23.987  1.00  0.00
ATOM   1882  CG  LYS   206      -1.270  -7.737 -24.995  1.00  0.00
ATOM   1883  CD  LYS   206      -1.553  -6.709 -26.079  1.00  0.00
ATOM   1884  CE  LYS   206      -2.548  -7.422 -26.863  1.00  0.00
ATOM   1885  NZ  LYS   206      -3.218  -6.457 -27.725  1.00  0.00
ATOM   1886  HZ1 LYS   206      -2.756  -6.174 -28.589  1.00  0.00
ATOM   1887  HZ2 LYS   206      -3.362  -5.581 -27.160  1.00  0.00
ATOM   1888  HZ3 LYS   206      -4.137  -6.824 -27.979  1.00  0.00
ATOM   1889  H   LYS   206      -1.479  -8.126 -21.765  1.00  0.00
ATOM   1890  N   LEU   207       2.364  -8.829 -22.359  1.00  0.00
ATOM   1891  CA  LEU   207       3.701  -8.880 -21.751  1.00  0.00
```

```
ATOM   1892  C   LEU   207       3.898 -10.389 -21.838  1.00  0.00
ATOM   1893  O   LEU   207       3.769 -10.853 -23.006  1.00  0.00
ATOM   1894  CB  LEU   207       3.667  -8.430 -20.284  1.00  0.00
ATOM   1895  CG  LEU   207       3.620  -6.922 -19.938  1.00  0.00
ATOM   1896  CD1 LEU   207       2.615  -6.762 -18.794  1.00  0.00
ATOM   1897  CD2 LEU   207       5.016  -6.332 -19.674  1.00  0.00
ATOM   1898  H   LEU   207       2.148  -9.663 -22.899  1.00  0.00
ATOM   1899  N   LEU   208       4.195 -11.206 -20.846  1.00  0.00
ATOM   1900  CA  LEU   208       4.442 -12.675 -21.068  1.00  0.00
ATOM   1901  C   LEU   208       3.301 -13.451 -20.307  1.00  0.00
ATOM   1902  O   LEU   208       3.445 -14.545 -19.786  1.00  0.00
ATOM   1903  CB  LEU   208       5.904 -12.979 -20.512  1.00  0.00
ATOM   1904  CG  LEU   208       6.690 -14.206 -20.846  1.00  0.00
ATOM   1905  CD1 LEU   208       6.580 -14.688 -22.348  1.00  0.00
ATOM   1906  CD2 LEU   208       8.060 -13.899 -20.408  1.00  0.00
ATOM   1907  H   LEU   208       4.315 -10.878 -19.974  1.00  0.00
ATOM   1908  N   ALA   209       2.178 -12.721 -20.325  1.00  0.00
ATOM   1909  CA  ALA   209       0.874 -13.150 -19.764  1.00  0.00
ATOM   1910  C   ALA   209       0.960 -13.881 -18.445  1.00  0.00
ATOM   1911  O   ALA   209       0.813 -15.103 -18.340  1.00  0.00
ATOM   1912  CB  ALA   209       0.113 -14.066 -20.868  1.00  0.00
ATOM   1913  H   ALA   209       2.082 -11.914 -20.812  1.00  0.00
ATOM   1914  N   HIS   210       1.315 -13.053 -17.517  1.00  0.00
ATOM   1915  CA  HIS   210       1.565 -13.307 -16.107  1.00  0.00
ATOM   1916  C   HIS   210       2.136 -14.675 -15.689  1.00  0.00
ATOM   1917  O   HIS   210       3.380 -14.668 -15.581  1.00  0.00
ATOM   1918  CB  HIS   210       0.347 -13.142 -15.187  1.00  0.00
ATOM   1919  CG  HIS   210      -0.945 -12.493 -15.635  1.00  0.00
ATOM   1920  ND1 HIS   210      -1.178 -11.382 -16.356  1.00  0.00
ATOM   1921  CD2 HIS   210      -2.175 -12.978 -15.220  1.00  0.00
ATOM   1922  CE1 HIS   210      -2.486 -11.179 -16.397  1.00  0.00
ATOM   1923  NE2 HIS   210      -3.061 -12.127 -15.723  1.00  0.00
ATOM   1924  HE2 HIS   210      -3.984 -11.981 -15.385  1.00  0.00
ATOM   1925  HD1 HIS   210      -0.525 -10.637 -16.523  1.00  0.00
ATOM   1926  H   HIS   210       1.550 -12.192 -17.763  1.00  0.00
ATOM   1927  N   PRO   211       1.513 -15.815 -15.312  1.00  0.00
ATOM   1928  CA  PRO   211       2.197 -16.963 -14.864  1.00  0.00
ATOM   1929  C   PRO   211       3.532 -17.376 -15.570  1.00  0.00
ATOM   1930  O   PRO   211       4.494 -17.794 -14.963  1.00  0.00
ATOM   1931  CB  PRO   211       1.206 -18.096 -14.890  1.00  0.00
ATOM   1932  CG  PRO   211      -0.096 -17.359 -14.946  1.00  0.00
ATOM   1933  CD  PRO   211       0.255 -16.215 -15.810  1.00  0.00
ATOM   1934  N   THR   212       3.512 -17.276 -16.879  1.00  0.00
ATOM   1935  CA  THR   212       4.596 -17.650 -17.780  1.00  0.00
ATOM   1936  C   THR   212       5.930 -17.046 -17.376  1.00  0.00
ATOM   1937  O   THR   212       7.031 -17.599 -17.610  1.00  0.00
ATOM   1938  CB  THR   212       4.263 -17.149 -19.225  1.00  0.00
ATOM   1939  OG1 THR   212       2.813 -17.274 -19.391  1.00  0.00
ATOM   1940  CG2 THR   212       5.018 -18.005 -20.231  1.00  0.00
ATOM   1941  HG1 THR   212       2.428 -16.400 -19.173  1.00  0.00
ATOM   1942  H   THR   212       2.755 -16.837 -17.335  1.00  0.00
ATOM   1943  N   GLN   213       5.995 -15.865 -16.638  1.00  0.00
ATOM   1944  CA  GLN   213       7.193 -15.245 -16.150  1.00  0.00
ATOM   1945  C   GLN   213       7.642 -15.808 -14.819  1.00  0.00
ATOM   1946  O   GLN   213       8.857 -16.168 -14.839  1.00  0.00
ATOM   1947  CB  GLN   213       7.061 -13.694 -15.939  1.00  0.00
ATOM   1948  CG  GLN   213       6.783 -12.937 -17.186  1.00  0.00
ATOM   1949  CD  GLN   213       5.700 -11.807 -17.113  1.00  0.00
ATOM   1950  OE1 GLN   213       4.507 -11.916 -16.734  1.00  0.00
ATOM   1951  NE2 GLN   213       6.173 -10.600 -17.417  1.00  0.00
ATOM   1952 HE22GLN   213       5.471  -9.889 -17.252  1.00  0.00
ATOM   1953 HE21GLN   213       7.103 -10.499 -17.681  1.00  0.00
ATOM   1954  H   GLN   213       5.182 -15.383 -16.412  1.00  0.00
ATOM   1955  N   PRO   214       6.850 -15.993 -13.758  1.00  0.00
ATOM   1956  CA  PRO   214       7.163 -16.957 -12.677  1.00  0.00
ATOM   1957  C   PRO   214       7.923 -18.240 -13.106  1.00  0.00
```

```
ATOM   1958  O   PRO   214       9.057 -18.525 -12.796  1.00  0.00
ATOM   1959  CB  PRO   214       5.805 -17.158 -12.005  1.00  0.00
ATOM   1960  CG  PRO   214       5.180 -15.841 -12.120  1.00  0.00
ATOM   1961  CD  PRO   214       5.565 -15.367 -13.494  1.00  0.00
ATOM   1962  N   SER   215       7.173 -18.979 -13.941  1.00  0.00
ATOM   1963  CA  SER   215       7.665 -20.270 -14.417  1.00  0.00
ATOM   1964  C   SER   215       9.061 -20.307 -14.976  1.00  0.00
ATOM   1965  O   SER   215       9.948 -21.062 -14.609  1.00  0.00
ATOM   1966  CB  SER   215       6.673 -20.829 -15.458  1.00  0.00
ATOM   1967  OG  SER   215       5.316 -20.628 -15.253  1.00  0.00
ATOM   1968  HG  SER   215       5.261 -19.967 -14.563  1.00  0.00
ATOM   1969  H   SER   215       6.253 -18.693 -14.178  1.00  0.00
ATOM   1970  N   LEU   216       9.385 -19.410 -15.968  1.00  0.00
ATOM   1971  CA  LEU   216      10.677 -19.537 -16.595  1.00  0.00
ATOM   1972  C   LEU   216      11.825 -18.824 -15.997  1.00  0.00
ATOM   1973  O   LEU   216      12.991 -19.321 -15.895  1.00  0.00
ATOM   1974  CB  LEU   216      10.545 -19.158 -18.053  1.00  0.00
ATOM   1975  CG  LEU   216       9.794 -20.201 -18.889  1.00  0.00
ATOM   1976  CD1 LEU   216       9.365 -19.478 -20.075  1.00  0.00
ATOM   1977  CD2 LEU   216      10.672 -21.398 -19.304  1.00  0.00
ATOM   1978  H   LEU   216       8.783 -18.731 -16.336  1.00  0.00
ATOM   1979  N   ALA   217      11.493 -17.675 -15.324  1.00  0.00
ATOM   1980  CA  ALA   217      12.523 -16.801 -14.692  1.00  0.00
ATOM   1981  C   ALA   217      13.558 -17.432 -13.779  1.00  0.00
ATOM   1982  O   ALA   217      14.739 -17.542 -14.185  1.00  0.00
ATOM   1983  CB  ALA   217      11.942 -15.687 -13.813  1.00  0.00
ATOM   1984  H   ALA   217      10.598 -17.322 -15.215  1.00  0.00
ATOM   1985  N   CYS   218      13.184 -17.825 -12.536  1.00  0.00
ATOM   1986  CA  CYS   218      14.120 -18.246 -11.487  1.00  0.00
ATOM   1987  C   CYS   218      15.193 -19.226 -11.907  1.00  0.00
ATOM   1988  O   CYS   218      16.386 -19.003 -11.571  1.00  0.00
ATOM   1989  CB  CYS   218      13.225 -18.826 -10.295  1.00  0.00
ATOM   1990  SG  CYS   218      13.805 -18.295  -8.641  1.00  0.00
ATOM   1991  H   CYS   218      12.256 -17.658 -12.280  1.00  0.00
ATOM   1992  N   ALA   219      14.958 -20.320 -12.628  1.00  0.00
ATOM   1993  CA  ALA   219      16.051 -21.247 -13.017  1.00  0.00
ATOM   1994  C   ALA   219      16.821 -20.879 -14.250  1.00  0.00
ATOM   1995  O   ALA   219      18.083 -20.932 -14.322  1.00  0.00
ATOM   1996  CB  ALA   219      15.538 -22.640 -13.158  1.00  0.00
ATOM   1997  H   ALA   219      14.056 -20.337 -12.969  1.00  0.00
ATOM   1998  N   GLU   220      16.137 -20.447 -15.317  1.00  0.00
ATOM   1999  CA  GLU   220      16.785 -20.100 -16.550  1.00  0.00
ATOM   2000  C   GLU   220      17.538 -18.696 -16.441  1.00  0.00
ATOM   2001  O   GLU   220      18.535 -18.489 -17.123  1.00  0.00
ATOM   2002  CB  GLU   220      15.676 -20.114 -17.617  1.00  0.00
ATOM   2003  CG  GLU   220      16.189 -20.326 -19.035  1.00  0.00
ATOM   2004  CD  GLU   220      15.115 -20.621 -20.074  1.00  0.00
ATOM   2005  OE1 GLU   220      14.330 -19.751 -20.321  1.00  0.00
ATOM   2006  OE2 GLU   220      15.145 -21.730 -20.701  1.00  0.00
ATOM   2007  H   GLU   220      15.183 -20.244 -15.227  1.00  0.00
ATOM   2008  N   ASN   221      17.030 -17.764 -15.651  1.00  0.00
ATOM   2009  CA  ASN   221      17.620 -16.486 -15.499  1.00  0.00
ATOM   2010  C   ASN   221      17.856 -16.395 -13.998  1.00  0.00
ATOM   2011  O   ASN   221      17.150 -15.767 -13.185  1.00  0.00
ATOM   2012  CB  ASN   221      16.691 -15.413 -16.027  1.00  0.00
ATOM   2013  CG  ASN   221      17.401 -14.095 -16.088  1.00  0.00
ATOM   2014  OD1 ASN   221      18.110 -13.747 -15.139  1.00  0.00
ATOM   2015  ND2 ASN   221      17.146 -13.379 -17.167  1.00  0.00
ATOM   2016  HD22ASN   221      17.540 -12.520 -17.365  1.00  0.00
ATOM   2017  HD21ASN   221      16.450 -13.789 -17.757  1.00  0.00
ATOM   2018  H   ASN   221      16.277 -18.055 -15.046  1.00  0.00
ATOM   2019  N   PHE   222      18.950 -17.081 -13.576  1.00  0.00
ATOM   2020  CA  PHE   222      19.303 -17.100 -12.168  1.00  0.00
ATOM   2021  C   PHE   222      20.223 -15.896 -11.976  1.00  0.00
ATOM   2022  O   PHE   222      21.411 -16.029 -11.643  1.00  0.00
ATOM   2023  CB  PHE   222      20.052 -18.391 -11.742  1.00  0.00
```

```
ATOM   2024  CG  PHE   222      19.986 -18.832 -10.276  1.00  0.00
ATOM   2025  CD1 PHE   222      20.213 -20.138 -10.007  1.00  0.00
ATOM   2026  CD2 PHE   222      19.718 -17.996  -9.220  1.00  0.00
ATOM   2027  CE1 PHE   222      20.209 -20.667  -8.727  1.00  0.00
ATOM   2028  CE2 PHE   222      19.729 -18.485  -7.927  1.00  0.00
ATOM   2029  CZ  PHE   222      19.967 -19.814  -7.677  1.00  0.00
ATOM   2030  H   PHE   222      19.624 -17.502 -14.133  1.00  0.00
ATOM   2031  N   VAL   223      19.688 -14.718 -12.222  1.00  0.00
ATOM   2032  CA  VAL   223      20.324 -13.448 -11.990  1.00  0.00
ATOM   2033  C   VAL   223      19.206 -12.404 -11.884  1.00  0.00
ATOM   2034  O   VAL   223      19.374 -11.234 -11.431  1.00  0.00
ATOM   2035  CB  VAL   223      21.281 -12.977 -13.157  1.00  0.00
ATOM   2036  CG1 VAL   223      22.675 -13.578 -12.935  1.00  0.00
ATOM   2037  CG2 VAL   223      20.922 -13.501 -14.479  1.00  0.00
ATOM   2038  H   VAL   223      18.813 -14.743 -12.611  1.00  0.00
ATOM   2039  N   LYS   224      17.987 -12.847 -12.237  1.00  0.00
ATOM   2040  CA  LYS   224      16.700 -12.179 -12.070  1.00  0.00
ATOM   2041  C   LYS   224      16.763 -10.736 -12.561  1.00  0.00
ATOM   2042  O   LYS   224      17.172 -10.665 -13.703  1.00  0.00
ATOM   2043  CB  LYS   224      16.327 -12.323 -10.623  1.00  0.00
ATOM   2044  CG  LYS   224      15.871 -13.767 -10.512  1.00  0.00
ATOM   2045  CD  LYS   224      15.825 -14.284  -9.067  1.00  0.00
ATOM   2046  CE  LYS   224      17.157 -14.854  -8.811  1.00  0.00
ATOM   2047  NZ  LYS   224      17.244 -15.878  -7.802  1.00  0.00
ATOM   2048  HZ1 LYS   224      16.533 -16.576  -8.083  1.00  0.00
ATOM   2049  HZ2 LYS   224      16.996 -15.430  -6.868  1.00  0.00
ATOM   2050  HZ3 LYS   224      18.166 -16.281  -7.772  1.00  0.00
ATOM   2051  H   LYS   224      17.978 -13.601 -12.839  1.00  0.00
ATOM   2052  N   ALA   225      16.641  -9.669 -11.694  1.00  0.00
ATOM   2053  CA  ALA   225      16.852  -8.238 -12.001  1.00  0.00
ATOM   2054  C   ALA   225      17.970  -7.929 -12.972  1.00  0.00
ATOM   2055  O   ALA   225      17.656  -7.231 -13.917  1.00  0.00
ATOM   2056  CB  ALA   225      17.201  -7.451 -10.764  1.00  0.00
ATOM   2057  H   ALA   225      16.418  -9.916 -10.808  1.00  0.00
ATOM   2058  N   ILE   226      19.192  -8.480 -12.873  1.00  0.00
ATOM   2059  CA  ILE   226      20.323  -8.175 -13.713  1.00  0.00
ATOM   2060  C   ILE   226      19.980  -8.416 -15.155  1.00  0.00
ATOM   2061  O   ILE   226      20.305  -7.557 -15.985  1.00  0.00
ATOM   2062  CB  ILE   226      21.644  -9.041 -13.279  1.00  0.00
ATOM   2063  CG1 ILE   226      21.992  -8.632 -11.823  1.00  0.00
ATOM   2064  CG2 ILE   226      22.783  -8.883 -14.355  1.00  0.00
ATOM   2065  CD1 ILE   226      22.374  -7.135 -11.834  1.00  0.00
ATOM   2066  H   ILE   226      19.258  -9.157 -12.105  1.00  0.00
ATOM   2067  N   GLU   227      19.402  -9.575 -15.525  1.00  0.00
ATOM   2068  CA  GLU   227      19.179  -9.815 -16.927  1.00  0.00
ATOM   2069  C   GLU   227      17.762  -9.696 -17.439  1.00  0.00
ATOM   2070  O   GLU   227      17.552  -9.042 -18.493  1.00  0.00
ATOM   2071  CB  GLU   227      19.783 -11.195 -17.237  1.00  0.00
ATOM   2072  CG  GLU   227      19.801 -11.517 -18.742  1.00  0.00
ATOM   2073  CD  GLU   227      20.131 -12.916 -19.163  1.00  0.00
ATOM   2074  OE1 GLU   227      19.295 -13.532 -19.811  1.00  0.00
ATOM   2075  OE2 GLU   227      21.150 -13.422 -18.776  1.00  0.00
ATOM   2076  H   GLU   227      19.127 -10.207 -14.868  1.00  0.00
ATOM   2077  N   LEU   228      16.761 -10.183 -16.761  1.00  0.00
ATOM   2078  CA  LEU   228      15.312 -10.138 -17.195  1.00  0.00
ATOM   2079  C   LEU   228      15.014  -8.624 -17.472  1.00  0.00
ATOM   2080  O   LEU   228      14.200  -8.394 -18.366  1.00  0.00
ATOM   2081  CB  LEU   228      14.342 -10.617 -16.092  1.00  0.00
ATOM   2082  CG  LEU   228      14.310 -12.066 -15.661  1.00  0.00
ATOM   2083  CD1 LEU   228      13.451 -12.074 -14.385  1.00  0.00
ATOM   2084  CD2 LEU   228      13.700 -12.954 -16.731  1.00  0.00
ATOM   2085  H   LEU   228      16.941 -10.526 -15.866  1.00  0.00
ATOM   2086  N   ASN   229      15.640  -7.793 -16.697  1.00  0.00
ATOM   2087  CA  ASN   229      15.531  -6.330 -16.666  1.00  0.00
ATOM   2088  C   ASN   229      14.092  -5.745 -16.992  1.00  0.00
ATOM   2089  O   ASN   229      13.503  -5.201 -16.058  1.00  0.00
```

```
ATOM   2090  CB  ASN   229      16.488  -5.734 -17.595  1.00  0.00
ATOM   2091  CG  ASN   229      16.465  -4.208 -17.428  1.00  0.00
ATOM   2092  OD1 ASN   229      15.733  -3.525 -18.116  1.00  0.00
ATOM   2093  ND2 ASN   229      17.217  -3.551 -16.549  1.00  0.00
ATOM   2094 HD22ASN   229      16.988  -2.593 -16.555  1.00  0.00
ATOM   2095 HD21ASN   229      17.890  -4.042 -16.024  1.00  0.00
ATOM   2096  H   ASN   229      16.365  -8.171 -16.147  1.00  0.00
ATOM   2097  N   GLN   230      13.589  -5.821 -18.253  1.00  0.00
ATOM   2098  CA  GLN   230      12.250  -5.388 -18.542  1.00  0.00
ATOM   2099  C   GLN   230      11.189  -6.461 -18.244  1.00  0.00
ATOM   2100  O   GLN   230      10.111  -6.167 -17.666  1.00  0.00
ATOM   2101  CB  GLN   230      12.015  -4.974 -19.979  1.00  0.00
ATOM   2102  CG  GLN   230      12.541  -3.608 -20.465  1.00  0.00
ATOM   2103  CD  GLN   230      11.971  -3.186 -21.812  1.00  0.00
ATOM   2104  OE1 GLN   230      11.595  -2.028 -22.008  1.00  0.00
ATOM   2105  NE2 GLN   230      11.903  -3.930 -22.952  1.00  0.00
ATOM   2106 HE22GLN   230      11.639  -3.363 -23.635  1.00  0.00
ATOM   2107 HE21GLN   230      12.204  -4.835 -23.100  1.00  0.00
ATOM   2108  H   GLN   230      14.156  -6.357 -18.903  1.00  0.00
ATOM   2109  N   ASN   231      11.526  -7.646 -18.665  1.00  0.00
ATOM   2110  CA  ASN   231      10.654  -8.800 -18.530  1.00  0.00
ATOM   2111  C   ASN   231      10.580  -9.460 -17.157  1.00  0.00
ATOM   2112  O   ASN   231      10.931 -10.612 -16.805  1.00  0.00
ATOM   2113  CB  ASN   231      11.030  -9.876 -19.527  1.00  0.00
ATOM   2114  CG  ASN   231       9.928 -10.218 -20.521  1.00  0.00
ATOM   2115  OD1 ASN   231       8.729 -10.406 -20.215  1.00  0.00
ATOM   2116  ND2 ASN   231      10.261 -10.046 -21.810  1.00  0.00
ATOM   2117 HD22ASN   231       9.500 -10.058 -22.504  1.00  0.00
ATOM   2118 HD21ASN   231      11.192  -9.777 -21.984  1.00  0.00
ATOM   2119  H   ASN   231      12.454  -7.849 -18.886  1.00  0.00
ATOM   2120  N   GLY   232      10.060  -8.628 -16.278  1.00  0.00
ATOM   2121  CA  GLY   232       9.867  -9.000 -14.919  1.00  0.00
ATOM   2122  C   GLY   232       8.913  -7.955 -14.395  1.00  0.00
ATOM   2123  O   GLY   232       7.900  -7.552 -14.998  1.00  0.00
ATOM   2124  H   GLY   232       9.612  -7.768 -16.521  1.00  0.00
ATOM   2125  N   ALA   233       9.316  -7.507 -13.214  1.00  0.00
ATOM   2126  CA  ALA   233       8.580  -6.539 -12.366  1.00  0.00
ATOM   2127  C   ALA   233       9.302  -6.405 -11.074  1.00  0.00
ATOM   2128  O   ALA   233      10.370  -7.028 -10.947  1.00  0.00
ATOM   2129  CB  ALA   233       7.233  -7.099 -12.042  1.00  0.00
ATOM   2130  H   ALA   233      10.224  -7.799 -12.875  1.00  0.00
ATOM   2131  N   ILE   234       8.852  -5.607 -10.138  1.00  0.00
ATOM   2132  CA  ILE   234       9.347  -5.526  -8.803  1.00  0.00
ATOM   2133  C   ILE   234      10.826  -5.198  -8.755  1.00  0.00
ATOM   2134  O   ILE   234      11.069  -3.962  -8.860  1.00  0.00
ATOM   2135  CB  ILE   234       9.176  -6.803  -8.057  1.00  0.00
ATOM   2136  CG1 ILE   234       7.784  -7.504  -8.287  1.00  0.00
ATOM   2137  CG2 ILE   234       9.143  -6.388  -6.595  1.00  0.00
ATOM   2138  CD1 ILE   234       7.768  -9.029  -8.017  1.00  0.00
ATOM   2139  H   ILE   234       8.285  -4.920 -10.390  1.00  0.00
ATOM   2140  N   TRP   235      11.886  -6.010  -8.519  1.00  0.00
ATOM   2141  CA  TRP   235      13.311  -5.640  -8.651  1.00  0.00
ATOM   2142  C   TRP   235      13.748  -4.387  -8.007  1.00  0.00
ATOM   2143  O   TRP   235      13.940  -4.316  -6.822  1.00  0.00
ATOM   2144  CB  TRP   235      13.569  -5.772 -10.197  1.00  0.00
ATOM   2145  CG  TRP   235      13.459  -7.123 -10.746  1.00  0.00
ATOM   2146  CD1 TRP   235      13.226  -7.185 -12.088  1.00  0.00
ATOM   2147  CD2 TRP   235      13.394  -8.427 -10.138  1.00  0.00
ATOM   2148  NE1 TRP   235      12.994  -8.488 -12.297  1.00  0.00
ATOM   2149  CE2 TRP   235      13.084  -9.220 -11.193  1.00  0.00
ATOM   2150  CE3 TRP   235      13.524  -9.148  -8.948  1.00  0.00
ATOM   2151  CZ2 TRP   235      12.911 -10.575 -11.114  1.00  0.00
ATOM   2152  CZ3 TRP   235      13.366 -10.517  -8.791  1.00  0.00
ATOM   2153  CH2 TRP   235      13.050 -11.232  -9.879  1.00  0.00
ATOM   2154  HE1 TRP   235      12.915  -8.836 -13.185  1.00  0.00
ATOM   2155  H   TRP   235      11.693  -6.949  -8.367  1.00  0.00
```

```
ATOM   2156  N   LYS   236      13.897  -3.384  -8.768  1.00  0.00
ATOM   2157  CA  LYS   236      14.306  -2.065  -8.392  1.00  0.00
ATOM   2158  C   LYS   236      13.399  -1.070  -9.138  1.00  0.00
ATOM   2159  O   LYS   236      13.854   0.081  -9.320  1.00  0.00
ATOM   2160  CB  LYS   236      15.787  -1.798  -8.774  1.00  0.00
ATOM   2161  CG  LYS   236      16.741  -2.106  -7.698  1.00  0.00
ATOM   2162  CD  LYS   236      18.058  -2.081  -8.352  1.00  0.00
ATOM   2163  CE  LYS   236      18.704  -0.688  -8.556  1.00  0.00
ATOM   2164  NZ  LYS   236      18.290  -0.147  -9.892  1.00  0.00
ATOM   2165  HZ1 LYS   236      17.218   0.033  -9.903  1.00  0.00
ATOM   2166  HZ2 LYS   236      18.546  -0.812 -10.624  1.00  0.00
ATOM   2167  HZ3 LYS   236      18.826   0.721 -10.104  1.00  0.00
ATOM   2168  H   LYS   236      13.837  -3.528  -9.731  1.00  0.00
ATOM   2169  N   LEU   237      12.157  -1.515  -9.273  1.00  0.00
ATOM   2170  CA  LEU   237      11.171  -0.763 -10.039  1.00  0.00
ATOM   2171  C   LEU   237      10.049  -0.438  -9.053  1.00  0.00
ATOM   2172  O   LEU   237       9.457  -1.313  -8.431  1.00  0.00
ATOM   2173  CB  LEU   237      10.552  -1.624 -11.062  1.00  0.00
ATOM   2174  CG  LEU   237      11.588  -2.384 -12.010  1.00  0.00
ATOM   2175  CD1 LEU   237      11.038  -3.515 -12.834  1.00  0.00
ATOM   2176  CD2 LEU   237      12.026  -1.286 -13.034  1.00  0.00
ATOM   2177  H   LEU   237      11.801  -2.188  -8.691  1.00  0.00
ATOM   2178  N   ASP   238       9.790   0.848  -8.864  1.00  0.00
ATOM   2179  CA  ASP   238       8.736   1.229  -8.005  1.00  0.00
ATOM   2180  C   ASP   238       7.715   1.769  -8.955  1.00  0.00
ATOM   2181  O   ASP   238       8.074   2.219 -10.037  1.00  0.00
ATOM   2182  CB  ASP   238       9.163   2.387  -6.973  1.00  0.00
ATOM   2183  CG  ASP   238       9.824   3.625  -7.638  1.00  0.00
ATOM   2184  OD1 ASP   238       9.158   4.633  -7.666  1.00  0.00
ATOM   2185  OD2 ASP   238      10.954   3.522  -8.139  1.00  0.00
ATOM   2186  H   ASP   238      10.346   1.490  -9.318  1.00  0.00
ATOM   2187  N   LEU   239       6.420   1.769  -8.571  1.00  0.00
ATOM   2188  CA  LEU   239       5.463   2.432  -9.460  1.00  0.00
ATOM   2189  C   LEU   239       5.393   3.948  -9.180  1.00  0.00
ATOM   2190  O   LEU   239       4.703   4.515  -8.311  1.00  0.00
ATOM   2191  CB  LEU   239       4.069   1.886  -9.259  1.00  0.00
ATOM   2192  CG  LEU   239       3.864   0.372  -9.579  1.00  0.00
ATOM   2193  CD1 LEU   239       2.500  -0.061  -9.039  1.00  0.00
ATOM   2194  CD2 LEU   239       4.097   0.146 -11.109  1.00  0.00
ATOM   2195  H   LEU   239       6.081   1.186  -7.885  1.00  0.00
ATOM   2196  N   GLY   240       6.181   4.718  -9.954  1.00  0.00
ATOM   2197  CA  GLY   240       6.071   6.183  -9.875  1.00  0.00
ATOM   2198  C   GLY   240       6.077   6.762 -11.266  1.00  0.00
ATOM   2199  O   GLY   240       7.127   6.643 -11.889  1.00  0.00
ATOM   2200  H   GLY   240       6.695   4.293 -10.721  1.00  0.00
ATOM   2201  N   THR   241       4.857   7.142 -11.727  1.00  0.00
ATOM   2202  CA  THR   241       4.706   7.929 -12.878  1.00  0.00
ATOM   2203  C   THR   241       4.887   9.391 -12.500  1.00  0.00
ATOM   2204  O   THR   241       5.758  10.083 -13.043  1.00  0.00
ATOM   2205  CB  THR   241       3.344   7.758 -13.480  1.00  0.00
ATOM   2206  OG1 THR   241       2.455   7.681 -12.320  1.00  0.00
ATOM   2207  CG2 THR   241       3.115   6.651 -14.383  1.00  0.00
ATOM   2208  HG1 THR   241       1.517   7.654 -12.592  1.00  0.00
ATOM   2209  H   THR   241       4.076   6.971 -11.184  1.00  0.00
ATOM   2210  N   LEU   242       4.020   9.936 -11.654  1.00  0.00
ATOM   2211  CA  LEU   242       4.288  11.273 -11.217  1.00  0.00
ATOM   2212  C   LEU   242       5.271  11.103 -10.061  1.00  0.00
ATOM   2213  O   LEU   242       5.052  10.570  -8.970  1.00  0.00
ATOM   2214  CB  LEU   242       3.010  11.851 -10.714  1.00  0.00
ATOM   2215  CG  LEU   242       2.492  13.313 -11.036  1.00  0.00
ATOM   2216  CD1 LEU   242       1.882  13.995  -9.825  1.00  0.00
ATOM   2217  CD2 LEU   242       3.697  14.156 -11.483  1.00  0.00
ATOM   2218  H   LEU   242       3.314   9.393 -11.127  1.00  0.00
ATOM   2219  N   GLU   243       6.474  11.640 -10.361  1.00  0.00
ATOM   2220  CA  GLU   243       7.524  11.709  -9.353  1.00  0.00
ATOM   2221  C   GLU   243       8.380  12.934  -9.571  1.00  0.00
```

```
ATOM   2222  O   GLU   243       8.354  13.821  -8.732  1.00  0.00
ATOM   2223  CB  GLU   243       8.340  10.408  -9.204  1.00  0.00
ATOM   2224  CG  GLU   243       8.505   9.677 -10.552  1.00  0.00
ATOM   2225  CD  GLU   243       9.951   9.716 -10.971  1.00  0.00
ATOM   2226  OE1 GLU   243      10.308  10.499 -11.837  1.00  0.00
ATOM   2227  OE2 GLU   243      10.758   8.963 -10.458  1.00  0.00
ATOM   2228  H   GLU   243       6.660  12.027 -11.267  1.00  0.00
ATOM   2229  N   ALA   244       9.116  13.006 -10.702  1.00  0.00
ATOM   2230  CA  ALA   244       9.876  14.210 -11.002  1.00  0.00
ATOM   2231  C   ALA   244       9.858  14.417 -12.494  1.00  0.00
ATOM   2232  O   ALA   244      10.648  13.806 -13.198  1.00  0.00
ATOM   2233  CB  ALA   244      11.317  14.033 -10.489  1.00  0.00
ATOM   2234  H   ALA   244       9.154  12.249 -11.358  1.00  0.00
ATOM   2235  N   ILE   245       8.942  15.270 -12.998  1.00  0.00
ATOM   2236  CA  ILE   245       8.784  15.352 -14.441  1.00  0.00
ATOM   2237  C   ILE   245       9.144  16.695 -15.027  1.00  0.00
ATOM   2238  O   ILE   245       8.259  17.413 -15.464  1.00  0.00
ATOM   2239  CB  ILE   245       7.380  14.860 -14.852  1.00  0.00
ATOM   2240  CG1 ILE   245       7.135  13.458 -14.256  1.00  0.00
ATOM   2241  CG2 ILE   245       7.276  14.790 -16.389  1.00  0.00
ATOM   2242  CD1 ILE   245       5.669  13.037 -14.472  1.00  0.00
ATOM   2243  H   ILE   245       8.323  15.799 -12.414  1.00  0.00
ATOM   2244  N   GLN   246      10.454  17.020 -15.077  1.00  0.00
ATOM   2245  CA  GLN   246      10.907  18.171 -15.845  1.00  0.00
ATOM   2246  C   GLN   246      10.460  19.544 -15.412  1.00  0.00
ATOM   2247  O   GLN   246      11.289  20.289 -14.915  1.00  0.00
ATOM   2248  CB  GLN   246      10.655  17.971 -17.354  1.00  0.00
ATOM   2249  CG  GLN   246      11.754  17.077 -17.962  1.00  0.00
ATOM   2250  CD  GLN   246      11.359  15.624 -18.013  1.00  0.00
ATOM   2251  OE1 GLN   246      10.376  15.237 -17.401  1.00  0.00
ATOM   2252  NE2 GLN   246      12.127  14.799 -18.756  1.00  0.00
ATOM   2253  HE22GLN   246      11.901  13.825 -18.818  1.00  0.00
ATOM   2254  HE21GLN   246      12.925  15.148 -19.250  1.00  0.00
ATOM   2255  H   GLN   246      11.158  16.426 -14.685  1.00  0.00
ATOM   2256  N   TRP   247       9.182  19.917 -15.630  1.00  0.00
ATOM   2257  CA  TRP   247       8.796  21.311 -15.453  1.00  0.00
ATOM   2258  C   TRP   247       8.098  21.616 -14.155  1.00  0.00
ATOM   2259  O   TRP   247       7.769  20.707 -13.408  1.00  0.00
ATOM   2260  CB  TRP   247       7.785  21.683 -16.558  1.00  0.00
ATOM   2261  CG  TRP   247       8.437  22.398 -17.702  1.00  0.00
ATOM   2262  CD1 TRP   247       9.007  21.838 -18.780  1.00  0.00
ATOM   2263  CD2 TRP   247       8.548  23.894 -17.845  1.00  0.00
ATOM   2264  NE1 TRP   247       9.439  22.793 -19.563  1.00  0.00
ATOM   2265  CE2 TRP   247       9.191  24.019 -19.062  1.00  0.00
ATOM   2266  CE3 TRP   247       8.163  24.990 -17.074  1.00  0.00
ATOM   2267  CZ2 TRP   247       9.493  25.264 -19.611  1.00  0.00
ATOM   2268  CZ3 TRP   247       8.462  26.246 -17.617  1.00  0.00
ATOM   2269  CH2 TRP   247       9.108  26.379 -18.855  1.00  0.00
ATOM   2270  HE1 TRP   247       9.911  22.624 -20.464  1.00  0.00
ATOM   2271  H   TRP   247       8.501  19.272 -15.982  1.00  0.00
ATOM   2272  N   THR   248       7.849  22.925 -13.917  1.00  0.00
ATOM   2273  CA  THR   248       6.981  23.338 -12.823  1.00  0.00
ATOM   2274  C   THR   248       7.673  23.877 -11.601  1.00  0.00
ATOM   2275  O   THR   248       7.675  25.086 -11.429  1.00  0.00
ATOM   2276  CB  THR   248       5.674  22.537 -12.674  1.00  0.00
ATOM   2277  OG1 THR   248       5.106  22.407 -13.984  1.00  0.00
ATOM   2278  CG2 THR   248       4.707  23.299 -11.748  1.00  0.00
ATOM   2279  HG1 THR   248       4.315  21.879 -13.968  1.00  0.00
ATOM   2280  H   THR   248       8.189  23.649 -14.518  1.00  0.00
ATOM   2281  N   LYS   249       8.286  23.032 -10.747  1.00  0.00
ATOM   2282  CA  LYS   249       9.073  23.618  -9.673  1.00  0.00
ATOM   2283  C   LYS   249       8.417  23.949  -8.353  1.00  0.00
ATOM   2284  O   LYS   249       9.160  24.260  -7.435  1.00  0.00
ATOM   2285  CB  LYS   249      10.526  23.109  -9.594  1.00  0.00
ATOM   2286  CG  LYS   249      11.420  24.041 -10.432  1.00  0.00
ATOM   2287  CD  LYS   249      11.048  23.946 -11.925  1.00  0.00
```

```
ATOM   2288  CE  LYS   249      12.064  24.759 -12.754  1.00  0.00
ATOM   2289  NZ  LYS   249      11.724  26.191 -12.811  1.00  0.00
ATOM   2290  HZ1 LYS   249      12.552  26.726 -13.253  1.00  0.00
ATOM   2291  HZ2 LYS   249      11.547  26.603 -11.828  1.00  0.00
ATOM   2292  HZ3 LYS   249      10.851  26.342 -13.428  1.00  0.00
ATOM   2293  H   LYS   249       8.273  22.035 -10.856  1.00  0.00
ATOM   2294  N   HIS   250       7.068  23.904  -8.230  1.00  0.00
ATOM   2295  CA  HIS   250       6.429  24.252  -6.966  1.00  0.00
ATOM   2296  C   HIS   250       6.357  25.736  -6.718  1.00  0.00
ATOM   2297  O   HIS   250       7.389  26.388  -6.765  1.00  0.00
ATOM   2298  CB  HIS   250       7.065  23.538  -5.756  1.00  0.00
ATOM   2299  CG  HIS   250       6.562  22.127  -5.704  1.00  0.00
ATOM   2300  ND1 HIS   250       5.604  21.741  -4.895  1.00  0.00
ATOM   2301  CD2 HIS   250       7.019  21.108  -6.461  1.00  0.00
ATOM   2302  CE1 HIS   250       5.380  20.480  -5.070  1.00  0.00
ATOM   2303  NE2 HIS   250       6.162  20.067  -5.959  1.00  0.00
ATOM   2304  HE2 HIS   250       6.183  19.115  -6.291  1.00  0.00
ATOM   2305  HD1 HIS   250       5.103  22.346  -4.227  1.00  0.00
ATOM   2306  H   HIS   250       6.466  23.636  -8.983  1.00  0.00
ATOM   2307  N   TRP   251       5.147  26.277  -6.446  1.00  0.00
ATOM   2308  CA  TRP   251       5.054  27.702  -6.154  1.00  0.00
ATOM   2309  C   TRP   251       5.453  27.993  -4.730  1.00  0.00
ATOM   2310  O   TRP   251       6.595  28.379  -4.543  1.00  0.00
ATOM   2311  CB  TRP   251       3.741  28.424  -6.531  1.00  0.00
ATOM   2312  CG  TRP   251       3.054  27.845  -7.730  1.00  0.00
ATOM   2313  CD1 TRP   251       1.900  27.160  -7.732  1.00  0.00
ATOM   2314  CD2 TRP   251       3.532  27.948  -9.155  1.00  0.00
ATOM   2315  NE1 TRP   251       1.608  26.844  -8.967  1.00  0.00
ATOM   2316  CE2 TRP   251       2.531  27.287  -9.842  1.00  0.00
ATOM   2317  CE3 TRP   251       4.639  28.517  -9.786  1.00  0.00
ATOM   2318  CZ2 TRP   251       2.557  27.152 -11.229  1.00  0.00
ATOM   2319  CZ3 TRP   251       4.672  28.388 -11.180  1.00  0.00
ATOM   2320  CH2 TRP   251       3.658  27.720 -11.883  1.00  0.00
ATOM   2321  HE1 TRP   251       0.758  26.326  -9.239  1.00  0.00
ATOM   2322  H   TRP   251       4.316  25.719  -6.403  1.00  0.00
ATOM   2323  N   ASP   252       4.570  27.823  -3.717  1.00  0.00
ATOM   2324  CA  ASP   252       3.226  27.349  -3.997  1.00  0.00
ATOM   2325  C   ASP   252       2.150  28.326  -3.590  1.00  0.00
ATOM   2326  O   ASP   252       1.151  27.912  -3.024  1.00  0.00
ATOM   2327  CB  ASP   252       2.970  25.897  -3.541  1.00  0.00
ATOM   2328  CG  ASP   252       3.594  25.613  -2.205  1.00  0.00
ATOM   2329  OD1 ASP   252       3.344  26.322  -1.246  1.00  0.00
ATOM   2330  OD2 ASP   252       4.363  24.673  -2.087  1.00  0.00
ATOM   2331  H   ASP   252       4.817  28.024  -2.766  1.00  0.00
ATOM   2332  N   SER   253       2.340  29.632  -3.890  1.00  0.00
ATOM   2333  CA  SER   253       1.286  30.606  -3.647  1.00  0.00
ATOM   2334  C   SER   253       0.793  30.657  -2.222  1.00  0.00
ATOM   2335  O   SER   253      -0.405  30.777  -2.013  1.00  0.00
ATOM   2336  CB  SER   253       0.131  30.416  -4.654  1.00  0.00
ATOM   2337  OG  SER   253       0.625  30.147  -5.972  1.00  0.00
ATOM   2338  HG  SER   253       1.077  30.908  -6.320  1.00  0.00
ATOM   2339  H   SER   253       3.171  29.968  -4.334  1.00  0.00
ATOM   2340  N   GLY   254       1.692  30.569  -1.219  1.00  0.00
ATOM   2341  CA  GLY   254       1.209  30.613   0.153  1.00  0.00
ATOM   2342  C   GLY   254       2.297  30.892   1.156  1.00  0.00
ATOM   2343  O   GLY   254       3.247  31.586   0.833  1.00  0.00
ATOM   2344  H   GLY   254       2.673  30.463  -1.391  1.00  0.00
ATOM   2345  N   ILE   255       2.163  30.345   2.382  1.00  0.00
ATOM   2346  CA  ILE   255       3.208  30.561   3.369  1.00  0.00
ATOM   2347  C   ILE   255       4.003  29.297   3.540  1.00  0.00
ATOM   2348  CB  ILE   255       2.687  31.046   4.739  1.00  0.00
ATOM   2349  CG1 ILE   255       1.559  32.084   4.579  1.00  0.00
ATOM   2350  CG2 ILE   255       3.852  31.690   5.514  1.00  0.00
ATOM   2351  CD1 ILE   255       0.194  31.369   4.588  1.00  0.00
ATOM   2352  O1  ILE   255       5.214  29.319   3.410  1.00  0.00
ATOM   2353  O2  ILE   255       3.434  28.249   3.798  1.00  0.00
```

```
ATOM   2354  H    ILE   255       1.380  29.769   2.624  1.00  0.00
TER    2355       ILE   255
HETATM 2356  P    NDP     1      -3.035  11.304   7.405  1.00  0.00
HETATM 2357  O1   NDP     1      -2.521  12.425   8.240  1.00  0.00
HETATM 2358  O2   NDP     1      -2.093  10.680   6.435  1.00  0.00
HETATM 2359  O5   NDP     1      -3.743  10.109   8.202  1.00  0.00
HETATM 2360  C5   NDP     1      -4.711  10.331   9.231  1.00  0.00
HETATM 2361  C4   NDP     1      -4.823   9.018  10.028  1.00  0.00
HETATM 2362  O4   NDP     1      -3.821   9.204  10.859  1.00  0.00
HETATM 2363  C1   NDP     1      -4.086   8.637  12.203  1.00  0.00
HETATM 2364  N9   NDP     1      -3.924   9.721  13.183  1.00  0.00
HETATM 2365  C8   NDP     1      -4.036  11.065  13.084  1.00  0.00
HETATM 2366  N7   NDP     1      -3.627  11.824  14.009  1.00  0.00
HETATM 2367  C5   NDP     1      -3.419  10.855  15.046  1.00  0.00
HETATM 2368  C6   NDP     1      -2.819  11.029  16.493  1.00  0.00
HETATM 2369  N1   NDP     1      -2.559   9.888  16.967  1.00  0.00
HETATM 2370  C2   NDP     1      -2.723   8.698  16.253  1.00  0.00
HETATM 2371  N3   NDP     1      -3.298   8.378  15.066  1.00  0.00
HETATM 2372  C4   NDP     1      -3.445   9.633  14.433  1.00  0.00
HETATM 2373  C2   NDP     1      -5.511   8.211  12.025  1.00  0.00
HETATM 2374  C3   NDP     1      -6.101   9.232  10.950  1.00  0.00
HETATM 2375  O3   NDP     1      -4.290  11.743   6.515  1.00  0.00
HETATM 2376  P    NDP     1      -4.908  13.137   6.027  1.00  0.00
HETATM 2377  O1   NDP     1      -4.696  14.277   6.962  1.00  0.00
HETATM 2378  O2   NDP     1      -6.118  12.742   5.253  1.00  0.00
HETATM 2379  O5   NDP     1      -3.908  13.295   4.788  1.00  0.00
HETATM 2380  C5   NDP     1      -3.520  12.357   3.437  1.00  0.00
HETATM 2381  C4   NDP     1      -3.660  13.252   2.192  1.00  0.00
HETATM 2382  O4   NDP     1      -2.381  14.014   2.120  1.00  0.00
HETATM 2383  C1   NDP     1      -1.639  13.659   0.971  1.00  0.00
HETATM 2384  N1   NDP     1      -0.323  13.798   1.096  1.00  0.00
HETATM 2385  C2   NDP     1       0.496  13.367   2.047  1.00  0.00
HETATM 2386  C3   NDP     1       1.764  13.178   1.939  1.00  0.00
HETATM 2387  NC7  NDP     1       2.203  12.727   3.042  1.00  0.00
HETATM 2388  O7   NDP     1       3.407  12.674   2.849  1.00  0.00
HETATM 2389  C4   NDP     1       2.433  13.984   0.873  1.00  0.00
HETATM 2390  C5   NDP     1       1.588  14.640  -0.084  1.00  0.00
HETATM 2391  C6   NDP     1       0.200  14.566  -0.022  1.00  0.00
HETATM 2392  C2   NDP     1      -2.154  12.318   0.655  1.00  0.00
HETATM 2393  C3   NDP     1      -3.668  12.406   0.989  1.00  0.00
HETATM 2394  C2   C       2       4.840  14.545  -1.744  1.00  0.00
HETATM 2395  O1   C       2       5.434  14.219  -0.485  1.00  0.00
HETATM 2396  H5   C       2       6.371  14.370  -0.528  1.00  0.00
HETATM 2397  C3   C       2       5.485  15.827  -2.301  1.00  0.00
HETATM 2398  H9   C       2       6.378  15.564  -2.887  1.00  0.00
HETATM 2399  H10  C       2       5.775  16.483  -1.466  1.00  0.00
HETATM 2400  H11  C       2       4.763  16.350  -2.945  1.00  0.00
HETATM 2401  C4   C       2       5.034  13.393  -2.748  1.00  0.00
HETATM 2402  H12  C       2       4.440  13.591  -3.652  1.00  0.00
HETATM 2403  H13  C       2       4.705  12.449  -2.291  1.00  0.00
HETATM 2404  H14  C       2       6.098  13.319  -3.017  1.00  0.00
HETATM 2405  H8   C       2       3.769  14.657  -1.517  1.00  0.00
CONECT  894  892  893  895
CONECT  895  894
CONECT 1734 1732 1733 1735
CONECT 1735 1734
CONECT 1923 1921 1922 1924
CONECT 1924 1923
CONECT 2303 2301 2302 2304
CONECT 2304 2303
CONECT 2347 2346 2352 2353
CONECT 2352 2347
CONECT 2353 2347
CONECT 2356 2357 2358 2359 2375
CONECT 2357 2356
CONECT 2358 2356
```

```
CONECT 2359 2356 2360
CONECT 2360 2359 2361
CONECT 2361 2360 2362 2374
CONECT 2362 2361 2363
CONECT 2363 2362 2364 2373
CONECT 2364 2363 2365 2372
CONECT 2365 2364 2366
CONECT 2366 2365 2367
CONECT 2367 2366 2368 2372
CONECT 2368 2367 2369
CONECT 2369 2368 2370
CONECT 2370 2369 2371
CONECT 2371 2370 2372
CONECT 2372 2364 2367 2371
CONECT 2373 2363 2374
CONECT 2374 2361 2373
CONECT 2375 2356 2376
CONECT 2376 2375 2377 2378 2379
CONECT 2377 2376
CONECT 2378 2376
CONECT 2379 2376 2380
CONECT 2380 2379 2381
CONECT 2381 2380 2382 2393
CONECT 2382 2381 2383
CONECT 2383 2382 2384 2392
CONECT 2384 2383 2385 2391
CONECT 2385 2384 2386
CONECT 2386 2385 2387 2389
CONECT 2387 2386 2388
CONECT 2388 2387
CONECT 2389 2386 2390
CONECT 2390 2389 2391
CONECT 2391 2384 2390
CONECT 2392 2383 2393
CONECT 2393 2381 2392
CONECT 2394 2395 2397 2401 2405
CONECT 2395 2394 2396
CONECT 2396 2395
CONECT 2397 2394 2398 2399 2400
CONECT 2398 2397
CONECT 2399 2397
CONECT 2400 2397
CONECT 2401 2394 2402 2403 2404
CONECT 2402 2401
CONECT 2403 2401
CONECT 2404 2401
CONECT 2405 2394
MASTER        0    0    0    0    0    0    0    0 2404    1   61   20
END
```